frontiers | Research Topics

Functional analysis of species-specific noncoding RNAs in plants

Edited by
Yuepeng Song, Mahmoud Yaish, Deqiang Zhang,
Yun Zheng, Lianfeng Gu and Byeong-ha Lee

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-83251-331-6
DOI 10.3389/978-2-83251-331-6

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

Functional analysis of species-specific noncoding RNAs in plants

Topic editors

Yuepeng Song — Beijing Forestry University, China
Mahmoud Yaish — Sultan Qaboos University, Oman
Deqiang Zhang — Beijing Forestry University, China
Yun Zheng — Yunnan Agricultural University, China
Lianfeng Gu — Fujian Agriculture and Forestry University, China
Byeong-ha Lee — Sogang University, Republic of Korea

Citation

Song, Y., Yaish, M., Zhang, D., Zheng, Y., Gu, L., Lee, B.-h., eds. (2023). *Functional analysis of species-specific noncoding RNAs in plants*.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-331-6

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 04 January 2023
DOI 10.3389/fgene.2022.1105433

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Hongliang Zhu,
China Agricultural University, China

*CORRESPONDENCE
Yuepeng Song,
yuepengsong@bjfu.edu.cn

SPECIALTY SECTION
This article was submitted to RNA, a section of the journal Frontiers in Genetics

RECEIVED 22 November 2022
ACCEPTED 12 December 2022
PUBLISHED 04 January 2023

CITATION
Wang M, Xiao Y, Su N and Song Y (2023), Editorial: Functional analysis of species-specific non-coding RNAs in plants. *Front. Genet.* 13:1105433. doi: 10.3389/fgene.2022.1105433

Editorial: Functional analysis of species-specific non-coding RNAs in plants

Menglei Wang[1,2], Yue Xiao[1,2], Nan Su[1,2] and Yuepeng Song[1,2]*

[1]National Engineering Research Center of Tree Breeding and Ecological Restoration, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China, [2]Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China

KEYWORDS

miRNA, non-coding RNA, developmental plasticity, abiotic stress, biotic stress

Editorial on the Research Topic
Functional analysis of species-specific noncoding RNAs in plants

Introduction

In addition to protein-coding RNAs, eukaryotes have different types of non-coding RNAs (ncRNAs) that are involved in the regulation of complex molecular and cellular processes. The sequences and functions of these ncRNAs are species-specific. Research on ncRNAs in plants has flourished during the past decade, due to advances in high-throughput sequencing (HTS) technologies and pioneering studies that have revealed the high frequency of ncRNAs in plants. NcRNAs can be divided into small RNAs, medium ncRNAs, and lncRNAs, among which small RNAs, particularly microRNAs (miRNA) and lncRNAs, play important roles in plant growth and development, critical phase transition, developmental plasticity, and response to biotic and abiotic stresses. Studying ncRNAs and their functions in a wider range of species will help elucidate the evolutionary pathway of ncRNA-mediated regulation mechanisms in plants and provide a theoretical basis for the study of genes that are constantly updated in response to changes in the natural environment. In this Research Topic, we present the latest advances in research of ncRNA regulation in plants, including their regulatory roles in plant developmental plasticity and abiotic and biotic stress responses. To date, ncRNAs have been studied in 179 species (Guo et al., 2022), among which this Research Topic covers 12, three-quarters of which are newly discussed, including the rare wood species *Dalbergia odorifera* and the low-temperature-tolerant *Ammopiptanthus nanus*. The findings reported in these studies provide valuable information toward the further identification of ncRNA functions and targets in a wider range of species.

TABLE 1 Conserved and species-specific miRNA of new species in this Research Topic.

Species	Conserved and species-specific miRNA	Target genes	Function
Betula luminifera	miR156/miR164/miR166	*SPL13/NAC100/HHP1*	Low nitrogen stress
Dalbergia odorifera	miR156	*SPL6, SPL12*	Xylem differentiation
Moso bamboo	miR166	*PeHOX10, PeHOX32*	Vascular tissue differentiation
Vaccinium corymbosum	miR166	*HD-ZIP III*	Freezing, cold, heat, salt stress
Sonneratia apetala	miR396	*UBP*	Salt stress
Ammopiptanthus nanus	miR4415	*L-AO*	Cold acclimation

The full names of the abbreviations in the table are as follows, *SPL: Squamosa promoter-binding protein-like, HHP1: Heptahelical protein 1, HD-ZIP III: Homeobox-leucine zipper proteins, UBP: Ubiquitin-specific protease, L-AO: L-ascorbate oxidase gene.*

Known miRNA regulatory network conserved in new species

MiRNA is an endogenous small ncRNA with important regulatory functions in eukaryotes (Reinhart et al., 2000; Bartel, 2004). Plant miRNA mainly recognizes and degrades target genes or represses the translation of target genes by recognizing complementary sequences at the post-transcriptional level, thereby participating in the regulation of plant growth and developmental plasticity (Song et al., 2019). Conserved miRNA emerged in the early stages of plant evolution, compared to other miRNA families, and they have high expression abundance and function conservation features. MiRNAs typically target multiple members of a gene family due to their mechanism of action, and conserved miRNA tends to target conserved genes with similar functions (Bartel, 2004; Song et al., 2019).

In a study of newly discovered ncRNAs in species such as *Dalbergia odorifera* and *Ammopiptanthus nanus*, 302 conserved miRNAs and 282 novel miRNAs belonging to 79 conserved miRNA families were found, among which the expression patterns and functions of the target genes of conserved miRNAs were consistent with those reported in model plants. These findings suggest that functional conserved miRNAs that emerged in the early evolution of species are functionally conserved in different plant groups, thus determining their critical roles in plant morphological differentiation, developmental plasticity, and environmental stress responses (Table 1).

Species-specific ncRNAs in plant developmental plasticity

LncRNAs are RNAs that are >200 bp in length and lack open reading frames (ORFs) or protein-coding capacity. They have higher tissue specificity, lower expression, and lower sequence conservation among species compared to mRNA in plants (Bardou et al., 2014). Unlike miRNA, lncRNA affects the transcription efficiency of neighboring genes while also forming scaffolds and decoys, or encoding small peptides. Recently, with the continuous development of sequencing technology, lncRNAs have been identified in increasing numbers of plant species in various growth stages, indicating their essential roles in modulating diverse biological regulatory processes in plants (Bartel, 2004; Chekanova, 2015; Chekanova, 2021).

LncRNAs participate in poplar lignin biosynthesis with the involvement of transcription factors and miRNAs. Zhang et al. (2022) reported that differentially expressed woody plant lncRNAs and target genes in two poplar genotypes were directly coexpressed with MYB and VND transcription factors and structural genes in the lignin and flavonoid pathways. Numerous auxin- and gibberellin-related lncRNA-mRNA coexpression networks have also been identified; these may regulate secondary xylem during the formation of tension wood. Together, these results suggest that lncRNAs are widely involved in lignin and flavonoid metabolism in poplars, complementing recent findings on new ncRNA members and their regulatory pathways, and providing a theoretical basis for exploring the function of ncRNAs in the developmental plasticity of plants.

Species-specific miRNAs and lncRNAs in abiotic stress responses in plants

In addition to conserved miRNAs, which have mainly analogical functions among plant species, some species also have non-conserved, species-specific miRNAs. Species-specific miRNAs typically have unique functional roles, and their existence provides clues that can be used to study the functions of corresponding *MIR* genes to understand the evolutionary positions of these genes (Song et al., 2019; Yu et al., 2019).

The leguminous plant *Ammopiptanthus nanus* has excellent tolerance to low temperature, and has therefore been used to study the molecular mechanisms of plant

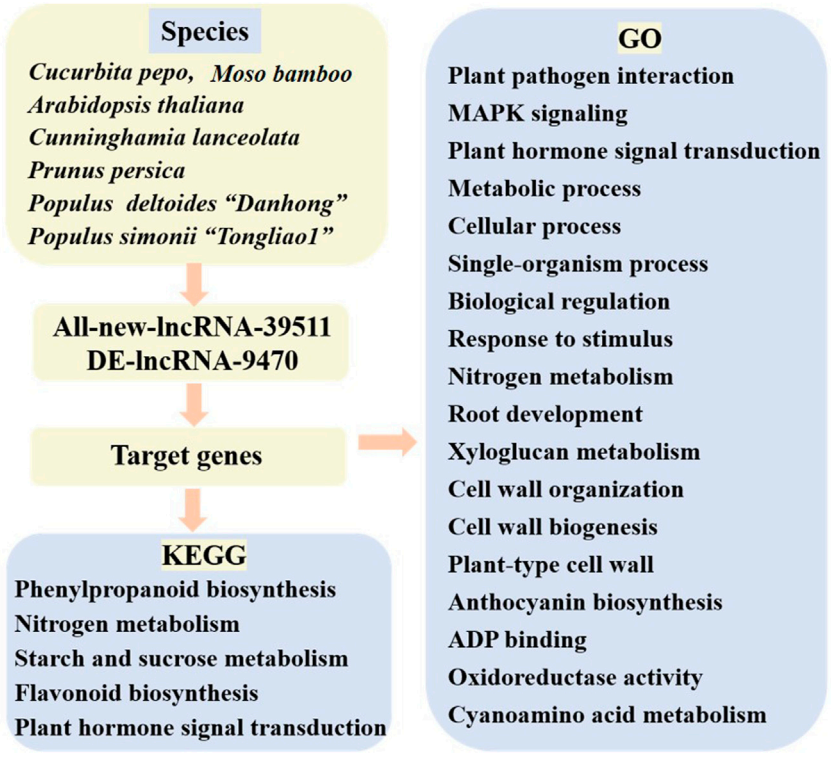


FIGURE 1
Novel lncRNA and their regulatory function of target genes in this issue.

responses to low-temperature stress. The legume-specific miRNA miR4415 is involved in cold acclimation in *A. nanus* by targeting an L-ascorbate oxidase gene that regulates apoplast redox status. This research has provided a basis for investigating cold acclimation regulation by miRNA.

In this Research Topic, we present ncRNA studies on *Cucurbita pepo*, *Moso bamboo*, *Arabidopsis thaliana*, *Cunninghamia lanceolata*, *Prunus persica*, *Populus deltoides* "Danhong", and *Populus simonii* "Tongliao1" that have reported 39,511 lncRNAs, among which most are novel. Together, these studies detected 9,470 differentially expressed lncRNAs showing targeted regulation of genes involved in processes related to powdery mildew, nitrogen metabolism, compression stress, ultraviolet B-induced flesh anthocyanin biosynthesis, wood formation, as well as phenylalanine molecular pathways and other biological processes (Figure 1).

In addition to developmental plasticity in plants, lncRNAs are widely involved in abiotic and biotic stress responses. The stress responses of lncRNAs differ from those of miRNAs. LncRNAs differentially expressed under drought, cold, salt, heat, and abscisic acid stress have been well described in model plants such as *Arabidopsis thaliana* and rice (Bartel, 2004; Zhang et al., 2022).

Light is a critical factor influencing anthocyanin biosynthesis in plants depending on light intensity, duration, and other light qualities. A recent genome-wide study based on transcriptomes of the flesh and juice of peach (*Prunus persica*) fruits at different developmental stages to identify lncRNAs involved in fruit ripening identified differentially expressed lncRNAs including XLOC_011933, XLOC_001865, and XLOC_042291. These are involved in ultraviolet B-induced anthocyanin biosynthesis and positively regulate *UVR8* and *COP10*, which participate in anthocyanin biosynthesis in peach fruits (Zhang et al., 2022).

LncRNAs are also involved in compression stress. In *Cunninghamia lanceolata*, lncRNA transcript_31,838 and transcript_29,184 are significantly correlated with *CESA.2*, *CESA2.2*, *CESA4*, *MANS.1*, *MYB21*, *MYB128*, *PAL1*, and *CCoAOMT1*, and may target these genes through transregulation in response to compression stress, thereby regulating the production of compression wood.

Interactions between miRNAs and lncRNAs in response to stress in plants

LncRNAs also act as precursors to small RNAs and interact with miRNAs in a process known as target mimicry to regulate

miRNA activity and abundance. This mechanism has also been linked to plant responses to abiotic stress. In nitric oxide-treated *rhd6* and wild-type *A. thaliana*, *MIR5658* and *MIR171* precursors are highly upregulated, and together with the novel lncRNAs MSTRG 15935, 15936, and 17,591, interact with differentially expressed protein coding genes involved in hormone signaling, cell wall development, and root hair formation. This makes them candidate genes for cell wall regulation and root hair phenotype recovery under nitric oxide treatment.

RNA sequencing analyses of the expression profiles of mRNAs, miRNAs, and lncRNAs in bamboo roots under different nitrogen treatment levels demonstrated a nitrogen metabolism regulatory network, which included 17 nitrogen metabolic pathway genes, 4 miRNAs targeting three *NPF* genes, and 10 lncRNAs targeting *NPFs* and *GDHs* through 15 transcription factors, indicating an ncRNA nitrogen metabolism regulation mechanism in *Moso bamboo*. Collectively, these findings demonstrate that lncRNA interactions with miRNA are essential in plant responses to abiotic stress.

In response to pathogen attack, plant cells trigger downstream molecular signaling networks. Pathogens such as powdery mildew, stripe rust, and rice blast often reduce the production of economic crops such as wheat and rice. Previous studies have reported the roles of ncRNAs in plant responses to such biological stresses. In one study, overexpression of the lncRNA ALEX1 in rice led to jasmonic acid pathway activation and resistance to bacterial blight (Yu et al., 2020). In another study on *Vitis vinifera*, 71 lncRNAs responsive to powdery mildew and 83 responsive to downy mildew were identified based on transcriptome sequencing responses to these obligate biotrophic fungal phytopathogens (Bhatia et al., 2021). Tian et al. inoculated powdery mildew into *Cucurbita pepo* leaves, and identified 242 differentially expressed lncRNAs. Genome-wide profile analysis predicted interactions between these lncRNAs and miRNAs as well as target genes associated with plant-pathogen interactions, MAPK signaling, and plant hormone signal transduction pathways. These findings suggest that *C. pepo* lncRNAs responsive to powdery mildew may participate in the pathogen response by regulating the expression of genes related to plant-pathogen interactions.

RNA interference (RNAi) technology, a form of post-transcriptional gene silencing induced by double-stranded RNA, has recently been applied to control plant pathogens and pests. RNAi depends on the recognition of target genes by silencing RNA (siRNA) (Fire et al., 1998). RNAi-induced gene silencing is an effective tool for enhancing pest and disease resistance in plants and genetic improvement in crops. Carrying out RNAi using a nano-carrier delivery system has been applied in biopesticide research; for example, multivariate nano-biotics have been successfully prepared to overcome the short duration and poor efficacy of plant pesticides using a nanoparticle-mediated delivery system that simultaneously loaded double-stranded RNA and plant-derived pesticides (Shen et al., 2022). This technique provides a new strategy for the development of biopesticides and opens a new chapter for the integration of ncRNA functional studies of the biological regulation of plant responses to pathogen attack and insect resistance (Li et al., 2022).

Prospects for future ncRNA research

NcRNAs are key regulatory factors in plant growth and development and play an indispensable role in maintaining the balanced expression of functional genes. The continual discovery of new ncRNAs in different species has accelerated progress in ncRNA characterization and functional analysis. However, ncRNAs have been studied in only 1% of plant species, with a focus on model and economically important species, and many species with evolutionary significance remain to be studied. The study of ncRNAs in new species will contribute to the comprehensive analysis of important nodes in species evolution.

The rapid development of ncRNA research in recent years has been made possibly by increasingly updated HTS technology. In the future, single-cell RNA sequencing, which is far superior to the current technology in terms of sequencing throughput, and spatial RNA sequencing, which spatially resolves RNA activity while comprehensively analyzing RNA transcription, will provide new opportunities to discover new ncRNAs and further analyze the biological functions of those that are already known.

In addition to their involvement in the development and stress responses of plants, new research directions to analyze ncRNA functions such as long-distance transport, plant-microbial interactions, and nanomaterial delivery are worthy of attention and further research. Nanoparticles can be applied to carry gene-editing elements through the cell wall for genetic transformation in plants; thus, it is also worth exploring whether ncRNA functions can be studied by combining nanoparticle-mediated delivery systems with gene-editing systems.

Author contributions

YS conceptualized the review, conducted research, and supervised the other contributors. MW prepared the original manuscript draft. YS, NS, and YX reviewed and edited the manuscript. YX and YS visualized the results. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Forestry and Grassland Science and Technology Innovation Youth Talent Project (No. 2020132606), the National Natural Science Foundation of China Project (No. 31770707), and the 111 Project (No. B20050).

Acknowledgments

We are grateful for the miRNA data provided in the miRbase Database (https://www.mirbase.org/) and the Plant microRNA Encyclopedia (http://www.pmiren.com).

Conflict of interest

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

Publisher's note

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

References

Bardou, F., Ariel, F., Simpson, C. G., Romero-Barrios, N., Laporte, P., Balzergue, S., et al. (2014). Long noncoding RNA modulates alternative splicing regulators in Arabidopsis. *Dev. Cell* 30, 166–176. doi:10.1016/j.devcel.2014.06.017

Bartel, D. P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297. doi:10.1016/s0092-8674(04)00045-5

Bhatia, G., Upadhyay, S. K., Upadhyay, A., and Singh, K. (2021). Investigation of long non-coding RNAs as regulatory players of grapevine response to powdery and downy mildew infection. *BMC Plant Biol.* 21, 265. doi:10.1186/s12870-021-03059-6

Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. *Curr. Opin. Plant Biol.* 27, 207–216. doi:10.1016/j.pbi.2015.08.003

Chekanova, J. A. (2021). Plant long non-coding RNAs in the regulation of transcription. *Essays Biochem.* 65, 751–760. doi:10.1042/EBC20200090

Fire, A., Xu, S., Montgomery, K., Kostas, A., Driver, E., and Mello, C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811. doi:10.1038/35888

Guo, Z., Kuang, Z., Zhao, Y., Deng, Y., He, H., Wan, M., et al. (2022). PmiREN2.0: From data annotation to functional exploration of plant microRNAs. *Nucleic Acids Res.* 50, D1475–D1482. doi:10.1093/nar/gkab811

Li, M., Ma, Z., Peng, M., Li, L., Yin, M., Yan, S., et al. (2022). A gene and drug co-delivery application helps to solve the short life disadvantage of RNA drug. *Nano Today* 43, 101452. doi:10.1016/j.nantod.2022.101452

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906. doi:10.1038/35002607

Song, X., Li, Y., Cao, X., and Qi, Y. (2019). MicroRNAs and their regulatory roles in plant–environment interactions. *Annu. Rev. Plant Biol.* 70, 489–525. doi:10.1146/annurev-arplant-050718-100334

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. F., Feng, Y. Z., He, H., Lian, J. P., Yang, Y. W., et al. (2020). Transcriptional landscape of pathogen-responsive lncRNAs in rice unveils the role of ALEX1 in jasmonate pathway and disease resistance. *Plant Biotech. J.* 18, 679–690. doi:10.1111/pbi.13234

Zhang, Y., Zhou, Y., Zhu, W., Liu, J., and Cheng, F. (2022). Non-coding RNAs fine-tune the balance between plant growth and abiotic stress tolerance. *Front. Plant Sci.* 13, 965745. doi:10.3389/fpls.2022.965745

frontiers in Genetics

ORIGINAL RESEARCH
published: 15 November 2021
doi: 10.3389/fgene.2021.762678

Genome-Wide Identification of Long Non-Coding RNAs and Their Potential Functions in Poplar Growth and Phenylalanine Biosynthesis

Lei Zhang[1], *Xiaolan Ge*[1], *Jiujun Du*[1], *Xingqi Cheng*[1], *Xiaopeng Peng*[1]* *and Jianjun Hu*[1,2]*

[1]State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China, [2]Collaborative Innovation Center of Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China

Poplar is an important bioenergy tree species. lncRNAs play important roles in various biological regulatory processes, and their expression pattern is more tissue-specific than mRNAs. In this study, P. deltoides "Danhong" (Pd) and P. simonii "Tongliao1" (Ps) with different growth rates and wood quality were used as experimental materials, and the transcriptomes of their shoot apical meristem, xylem, and phloem were sequenced. Furthermore, high-throughput RNA sequencing analysis revealed that the expression patterns of genes and lncRNAs are different between the two genotypes. 6,355 lncRNAs were identified. Based on target prediction, lncRNAs and target genes were involved in ADP binding, oxidoreductase activity, phenylpropanoid biosynthesis, and cyanoamino acid metabolism. The DElncRNAs in two poplars were co-expressed with transcription factors and structural genes of lignin and flavonoid pathways. In addition, we found the potential target lncRNAs of miRNA. This result provides basic evidence for a better understanding of the regulatory role of lncRNAs in regulating phenylalanine molecular pathways and wood formation.

Keywords: lncRNA–mRNA, poplar, phenylalanine biosynthesis, xylem, hormone

OPEN ACCESS

Edited by:
Deqiang Zhang,
Beijing Forestry University, China

Reviewed by:
Wanwen Zeng,
Nankai University, China
Tao Ma,
Sichuan University, China

***Correspondence:**
Xiaopeng Peng
xp@caf.ac.cn
Jianjun Hu
hujj@caf.ac.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *22 August 2021*
Accepted: *11 October 2021*
Published: *15 November 2021*

Citation:
Zhang L, Ge X, Du J, Cheng X, Peng X and Hu J (2021) Genome-Wide Identification of Long Non-Coding RNAs and Their Potential Functions in Poplar Growth and Phenylalanine Biosynthesis. Front. Genet. 12:762678. doi: 10.3389/fgene.2021.762678

INTRODUCTION

Plants are unique in their ability to continuously produce new organs throughout their life cycles. The process of continuous organogenesis depends on the activity of pluripotent cells (Xue et al., 2020). In trees, this mainly refers to the shoot apical meristem (SAM) affecting high growth and the vascular cambium affecting radial growth (Elo et al., 2009). The SAM generates leaves, stems, and floral organs throughout the lifespan of higher plants (Ha et al., 2010; Xue et al., 2020). The cambium differentiates into xylem and phloem, determined cell types, and cell layers in the secondary xylem (Ye and Zhong, 2015). These complex processes are easily regulated by plant hormones, transcription factors (TFs), miRNAs, and lncRNAs (J. Zhang G. et al., 2018; Xue et al., 2020).

lncRNAs are non-coding transcripts longer than 200 nucleotides (nts), including intergenic, intronic, sense, and antisense types (Ma et al., 2013). Compared with protein-coding genes (PCgenes), most lncRNAs are less conserved between species, lower expression levels, and stronger tissue-specific expression patterns (Liu et al., 2012; Zhou et al., 2017; Xu et al., 2018). lncRNAs can regulate genes expression at transcriptional, posttranscriptional, and epigenetic levels and play an important role in genomic imprinting, chromatin remodeling, transcriptional activation,

transcriptional interference, and cell cycle (Wang and Chekanova, 2017; Sun et al., 2018). With the continuous development of resequencing technology, lncRNAs of more and more species have been identified. They are widely involved in embryo development, seed formation, flower development, secondary growth of wood, and abiotic stress response (Zhou et al., 2017; Severing et al., 2018; Xu et al., 2018; Jiang et al., 2019; Wu et al., 2019). For example, lncRNAs play a potential regulatory role in endosperm and embryo development of castor bean (Xu et al., 2018). *COOLAIR* and *COLDAIR* play an important role in regulating vernalization in *Arabidopsis* (Heo and Sung, 2011). *FLINC* lncRNA participates in ambient temperature–mediated flowering time of *Arabidopsis* (Severing et al., 2018). lncRNAs influence the formation of tension wood by regulating *ARF*s in *Catalpa bungei* (Xiao et al., 2020). lncRNAs are widely involved in the secondary growth, GA response, heat tolerance, low nitrogen stress, salt stress, and other life processes of poplar (Chen et al., 2016; Tian et al., 2016; Ci et al., 2019; Ma et al., 2019; Song et al., 2020).

Populus is often used as a short rotation coppice (SRC) and bioenergy tree species all over the world because of its fast growth and reduced inhibitory extract from wood fermentation during bioenergy conversion (Guerra et al., 2013; Zhang et al., 2020b). So, the growth rate determines the economic benefit and the output of biomass energy. There are significant differences in the growth rate of *P. deltoides* "Danhong" and *P. simonii* "Tongliao1." *P. deltoides* "Danhong" is a southern poplar characterized by fast growth and insect resistance (Zhang et al., 2008). *P. simonii* is a native tree species in northern China; although the growth rate is slow, it is resistant to cold and drought (Wei et al., 2013).

In order to identify the regulation mechanism of growth and wood property differences and provide theoretical basis for breeding new germplasm with fast growing ability, we selected *P. deltoides* "Danhong" and *P. simonii* "Tongliao1" as experimental materials and identified the important lncRNAs that may be involved in growth regulation by sequencing. In this study, the sequencing of lncRNA libraries was constructed from the SAM, phloem, and developing xylem of *P. deltoides* "Danhong" and *P. simonii* "Tongliao1." *P. tricorcarpa* was used as the reference genome for the identification of lncRNA. We identified a total of 6,355 lncRNAs, of which 2,454 were sense_overlapping lncRNAs, 2,004 were lincRNAs, and 1,897 were antisense lncRNAs. The functional prediction of lncRNAs and their expressions as involved in wood development were examined. We investigated putative functional lncRNA candidates by differential expression analysis and co-expression network construction during SAM and xylem development. The important miRNA–lncRNA pairs in phenylalanine biosynthesis and hormone transduction were identified.

MATERIALS AND METHODS

Plant Materials

One-year-old *P. deltoides* "Danhong" (Pd) and *P. simonii* "Tongliao1" (Ps) were cultivated in the experimental field of the Chinese Academy of Forestry, Beijing, China (116.256°E, 40.007°N). We collected shoot apical meristem (SAM, Pd_S and Ps_S) and scraped phloem (inside of the bark, Pd_P and Ps_P) and developing xylem (newly formed xylem cells about 2–3 mm, Pd_X and Ps_X) from Pd and Ps, respectively, at diameter breast height (DBH) during the fast-growing period (July 20, 2019). Each tissue had three biological replicates. The samples (2 genotypes × 3 tissues × 3 biological replicates) used for RNA extraction were frozen immediately in liquid nitrogen and stored at −80°C. Shoot tips and cuneiform blocks (phloem, cambium, and xylem) at DBH for histologic analysis were fixed in a mixture of formalin, glacial acetic acid, and 70% ethanol in the ratio 5:5:90 vol.; FAA under vacuum for at least 24 h.

Histologic Analysis

Stem pieces were embedded with Spurr resin as described by Zhang et al. (2020d). A cross section of 4 μm thick was obtained from the stem by Leica M205FA, while the SAM sections of 40 μm were obtained using a rotary microtome (Leica VT1200S, Wetzlar, Germany). The sections were stained using 0.05% toluidine blue O (TBO) and were examined with a microscope (Zeiss). The number and diameter of vessel cells in the same area (1,260 × 980 μm) were counted by ImageJ (version 1.8.0).

Wood Property Determination

In order to understand the difference of wood properties between Pd and Ps, we measured the plant height and ground diameter and collected the stems to measure the wood properties including basic density, fiber length, fiber width, microfibril angle, cellulose, holocellulose, and lignin content of 1-year-old trees in December 2019. The basic density was determined by using the drainage method. A 10-cm-high wood segment was cut from the base of the trunk without bark and pith. It was softened by heating in 30% nitric acid and a small amount of potassium chlorate and converted into wood pulp by forced oscillation. The length and width of the fiber were measured 50 times using Shyygx Measure 2.0. Wood flour (40–60 mesh) from a 5-cm basal stem segment was used to determine the chemical composition. The content of holocellulose and lignin was calculated according to Chinese standards GB/T 2677.10-995 and GB/T 2677.8-1994, respectively. To evaluate the content of cellulose, the specimens were extracted with a mixed solvent of nitric acid and ethanol (v/v = 1/1) (Zhan et al., 2015). Three replicates were performed for each variety.

Total RNA Isolation, Library Construction, and Illumina Transcriptome Sequencing

Total RNA was isolated from the 18 samples (SAM, phloem, and developing xylem) using the RNAprep Pure Plant Plus Kit (TIANGEN, China). An index of the reference genome (*P. trichocarpa* v3.0) was built using HISAT2 (Kim et al., 2015). StringTie was used to calculate FPKMs of both lncRNAs and coding genes in each sample (Pertea et al., 2015). Sequencing data are available in NCBI SRA database (SRA number: SRP2343030 to SRR13961247).

lncRNA Identification

We used four filtration steps to identify lncRNAs from the transcriptome assembly: 1) Transcripts with an exon number ≥2 and length ≥200 bp were selected. 2) CuffCompare software was used to screen out transcripts that overlap with the database annotation exon field. 3) Evaluation of Coding Potential Calculator (CPC) (Kong et al., 2007), Coding Potential Assessment Tool (CPAT) (Wang et al., 2013), and Coding-Non-Coding Index (CNCI) (Sun et al., 2013) was carried out to screen whether there is coding potential. 4) They were referred to the the Hugo Gene Nomenclature Committee (HGNC) to name the novel_lncRNA of this analysis (Wright, 2014).

Prediction of the Target Gene

Two methods were used to predict lncRNA target genes. *Cis* target genes were predicted according to the location relationship between lncRNA and mRNA, and the screening range was within 10 kb (Jia et al., 2010). Co-expression–related target genes were predicted according to the expression correlation between lncRNA and mRNA, and the screening condition is that the correlation coefficient is greater than 0.95 and p-value< 1.68E-09. The mRNA–lncRNA regulatory network was further modeled and visualized using Cytoscape 3.8 (Otasek et al., 2019).

In order to identify lncRNAs that may be used as precursors of miRNAs, we compared the published miRNAs of *P. trichocarpa* in miRBase (http://www.mirbase.org/search.shtml) with lncRNAs. lncRNAs as targets of miRNAs were predicted by Novomagic, a free online platform for data analysis (https://magic.novogene.com).

Differential Expression Genes and Functional Analysis

To identify the differential expression of lncRNAs and mRNAs between Pd and Ps, we performed the read count of pair-wise comparisons (Pd_S vs Ps_S, Pd_P vs Ps_P, Pd_X vs Ps_X, Pd_S vs Pd_P, Pd_S vs Pd_X, Pd_P vs Pd_X, Ps_S vs Ps_P, Ps_S vs Ps_X, Ps_P vs Ps_X) by DESeq R package with a q-value < 0.05 (Love et al., 2014). Finally, those putative *cis*- and coexpression-targets of lncRNAs were analysed using Gene Ontology (GO) analysis (Ashburner et al., 2000; Conesa et al., 2005), and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment of DE genes was performed based on a corrected p-value < 0.05.

Quantitative Real-Time (qRT)-PCR and Correlation Analysis of Expression Trends

We selected three DELs and three DEGs from the results of the transcriptional analysis and confirmed them through qRT-PCR. *PtrActin* and *PtrUBQ* were used as internal reference genes (Wang et al., 2020a) (**Supplementary Table 1**). Real-time PCR was conducted on a LightCycler 480 (Roche, Basel, Switzerland) using the SYBR Premix Ex Taq™ Kit (Takara, Dalian, China). The relative expression levels of the genes were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001), and the data are presented as mean ± SD from three independent biological replicates.

RESULTS

Differences in Growth and Wood Properties Between Pd and Ps

Wood is the secondary xylem of trees, mainly composed of cellulose, hemicellulose, and lignin. All xylem cell types first undergo secondary cell wall (SCW) thickening and programmed cell death (Zhong and Ye, 2015; Zhang et al., 2020d). The plant height and ground diameter of annual Pd were significantly higher than those of Ps (**Figures 1A,B**). The SAM is an important regulatory site of plant height growth, and phloem and developing xylem are important parts of plant radial growth. The slice results showed that the SAM was surrounded by young leaves, and the SAM of Pd was conical and convex, while the Ps was flat, which suggested that the growth point of Ps might not be obvious enough, which caused the high growth to be slower (**Figure 1C**). The radial section of the stem showed that the phloem of Pd had wider phloem fibers, and the cambium was more obvious (**Figure 1D**). The average diameter of the vessel cells in the xylem of Pd was 72.5 μm, which was significantly larger than that of Ps (**Figures 1D,E**). We further determined the quality of wood. The basic density of basic Pd was less than Ps (**Figure 1F**), and there was no significant difference in the microfibril angle (**Figure 1G**). Compared with Ps, the fiber of Pd is short and thick (**Figures 1H,I**). Further material property determination found that the content of the three major elements of Pd is higher (**Figures 1J–L**).

Identification of lncRNAs From SAM, Phloem, and Developing Xylem RNA-Seq Datasets

As an important fast-growing tree, it is very important to understand the molecular pathways of growth and development of poplar. After trimming adapters and removing low-quality and contaminated reads, in total, 246.24 Gb clean data were obtained from 18 libraries, with an average Q30 of 93.00% (**Supplementary Table 2**). Finally, we identified 6,355 lncRNAs, with protections of 2,454 sense_overlapping lncRNAs, 2,004 of lincRNAs, and 1,897 of antisense lncRNAs (**Table S3**).

In order to analysis the characteristics of these lncRNAs, we evaluated the distribution of chromosome location, transcript length, exon number, and expression level of lncRNAs. In general, lincRNA, sense_overlapping, and antisense lncRNAs were evenly distributed on 19 chromosomes, although they had different emphases (**Figure 2A**). The average length of lncRNAs was 990 bp, and about 63.4% contained two exons (**Figures 2B,C**). Antisense lncRNAs ranged in length from 201 to 9,830 bp, and the average was 940 bp. lincRNAs ranged between 201 and 4,734 bp (average = 783 bp), and average length of sense_overlapping lncRNAs was 1,197 bp. The GC content of antisense lncRNA was 41.36%, which was significantly higher than that of lincRNAs and sense_overlapping lncRNAs (**Figure 2D**). For expression levels, the lncRNA expression levels

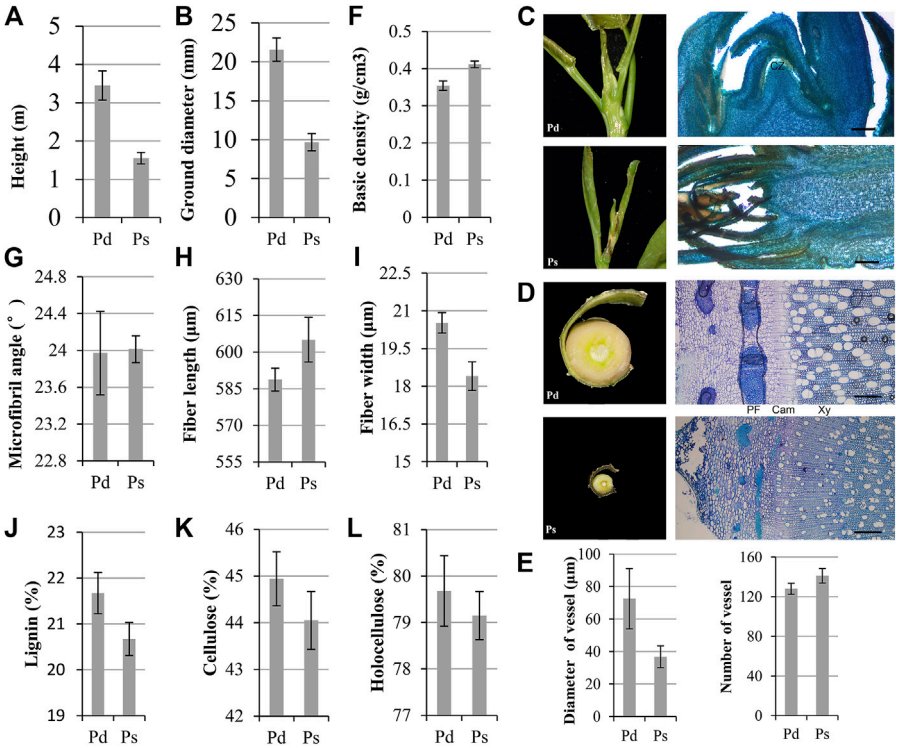


FIGURE 1 | Phenotypic traits of *P. deltoides* "Danhong" (Pd) and *P. simonii* "Tongliao1" (Ps). Annual plant height **(A)** and ground diameter **(B)** of Pd and Ps. **(C)** Shoot apical meristem (SAM) image and section of Pd and Ps. Scale bars = 500 μm. **(D)** Cross sections of phloem–xylem region from 1-year-old trees. Scale bars = 200 μm. **(E)** Vessel size and number of vessel in the same area of Pd and Ps. CZ, central zone; PF, phloem fiber; Cam, cambium; Xy, xylem. Basic density **(F)**, microfibril angle **(G)**, fiber length **(H)**, fiber width **(I)**, lignin **(J)**, cellulose **(K)**, and holocellulose **(L)** content of Pd and Ps. Mean ± SD from three biological replicates.

were different and showed fewer average counts (FPKM = 4.44) than the coding transcripts (FPKM = 18.96) (**Figures 2E,F**).

A principal component analysis (PCA) plot of the whole dataset revealed a sequential order of the different samples. The results showed that the SAM, xylem, and phloem of the two species were clustered into three groups, and the similarity of lncRNAs in tissues was greater than that between genotypes (**Figure 3A**).

Differentially Expressed Anaylsis Between Pd and Ps

In order to further analyze whether these genes were differentially expressed between the two genotypes and different tissues, nine comparative combinations were carried out. Finally, 3,572 differentially expressed (DE) lncRNAs and 27,582 DEmRNAs were obtained. Among them, the DElncRNAs of Pd_S vs Ps_S were the highest in number, including 1957 lncRNAs (**Figures 3B,C**; **Supplementary Figure 1**). DEmRNAs participated in molecular functions such as "ADP binding" and "catalytic activity" (**Supplementary Figure 1C**). We also compared the DElncRNAs in three tissue difference genes between the two genotypes, and there were 980 DElncRNAs in the three comparison combinations (Pd_ X vs Ps_ X, Pd_ P vs Ps_ P and Pd_ S vs Ps_ S) and 322, 333, and 420 specifically expressed lncRNAs in the SAM, xylem, and phloem, respectively (**Figure 3D**).

Enrichment Analysis of lncRNAs With a Potential Regulatory Function

Since lncRNAs play important roles in regulating gene expression, identification and analysis of their target genes may help us explore their potential functions. We calculated and predicted 13,932 co-localization pairs consisting of 3,413 lncRNAs and 10,627 RNAs and identified 72,038 co-expression pairs consisting of 1,975 lncRNAs and 11,709 RNAs (**Tables S4, S5**). To further analyze the function of these lncRNAs, we performed GO and KEGG analyses on their target genes. The colocation target genes of DElncRNAs were mainly enriched in 60 GO terms such as "ADP binding" and "nucleoside binding" (**Supplementary Figure 2**). Some target genes were enriched in the photosynthesis pathway, including 39 lncRNAs and 68 mRNAs. For example, TCONS_00135489 showed the same trend as its target, and the expression level of related genes was high in the SAM (**Supplementary Figure 3**). And 3,489 genes were found by co-location and co-expression of lncRNAs. They participated in immune response, cell death, purine nucleotide binding, and ATP binding progress, and the KEGG enrichment analysis shows that they were enriched in cytochrome P450, chalcone synthase, and phenylalanine ammonia-lyase

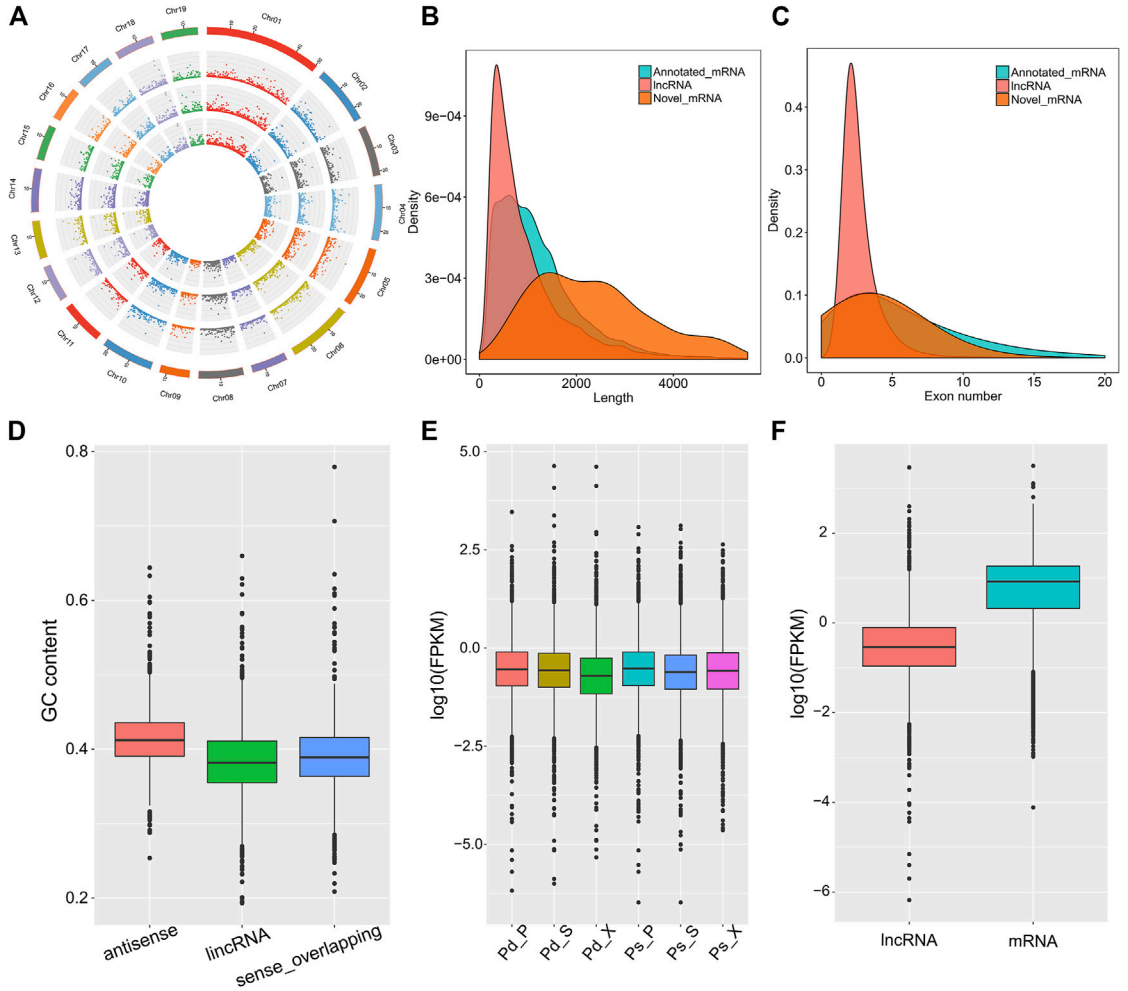


FIGURE 2 | Characterization of cassava lncRNAs in Pd and Ps. **(A)** Distribution of lncRNAs along each chromosome for the SAM, phloem, and xylem in Pd from inside to outside. Distributions of length density **(B)** and exon numbers **(C)** in mRNAs, novel mRNAs, and novel lncRNAs. **(D)** GC content of lncRNAs. **(E)** Fragments per kilobase per million read (FPKM) values of lncRNAs in different samples. **(F)** Comparison of the expression levels of lncRNA and mRNA.

which were important parts of phenylpropanoid biosynthesis (**Figure S4**).

The co-expression genes of Pd_ X vs Ps_ X, Pd_ P vs Ps_ P, and Pd_ S vs Ps_ S were mainly related to "ADP binding", "heme binding", and other biological functions (**Figure 4A**). And they were significantly enriched in the "phenylpropanoid biosynthesis" and "cyanoamino acid metabolism" pathway (**Figure 4B**). These possible target genes provide new insights into the role of lncRNAs in poplar development.

Regulation of lncRNAs and Transcription Factors in Phenylpropanoid Biosynthesis

Based on the predicted GO conditions of growth differential lncRNAs and the pathways associated with target genes, we speculated that lncRNAs might play an important role in phenylalanine biosynthesis in poplar. Phenylpropanoids are a group of plant secondary metabolites derived from phenylalanine, which has a variety of structural and signal molecular functions (Costa et al., 2003). It is the starting compound for biosynthesis of lignin, flavonoids, anthocyanins, etc., and a core mediator of crosstalk between development- and defense-related pathways (Alessandra et al., 2010). Further analysis co-expression network of these lncRNAs, and structure genes of lignins, and flavonoid biosynthesis found that TCONS_00128372, lincRNA, was located in Chr12 and interacted with MYB46, secondary wall–associated NAC domain2 (*SND2*), cinnamate-4-hydroxylase (*C4H*), caffeoyl-CoA 3-O-methyltransferase (*CcoAMT*), and laccase (*LAC*). And sense_overlapping lncRNA TCONS_00079190 co-expressed with *MYB83*, *MYB46*, NAC secondary wall thickening promoting factor1 (*NST1*), and *LAC* (**Figure 5A**). Similarly, the results of the study by Quan et al. (2019) and Zhou et al. (2017) showed that lncRNAs and miRNAs regulated lignin biosynthesis by regulating TFs such as *bHLHs* and *bZIPs* or directly coexpressing with structural genes in the process of lignin

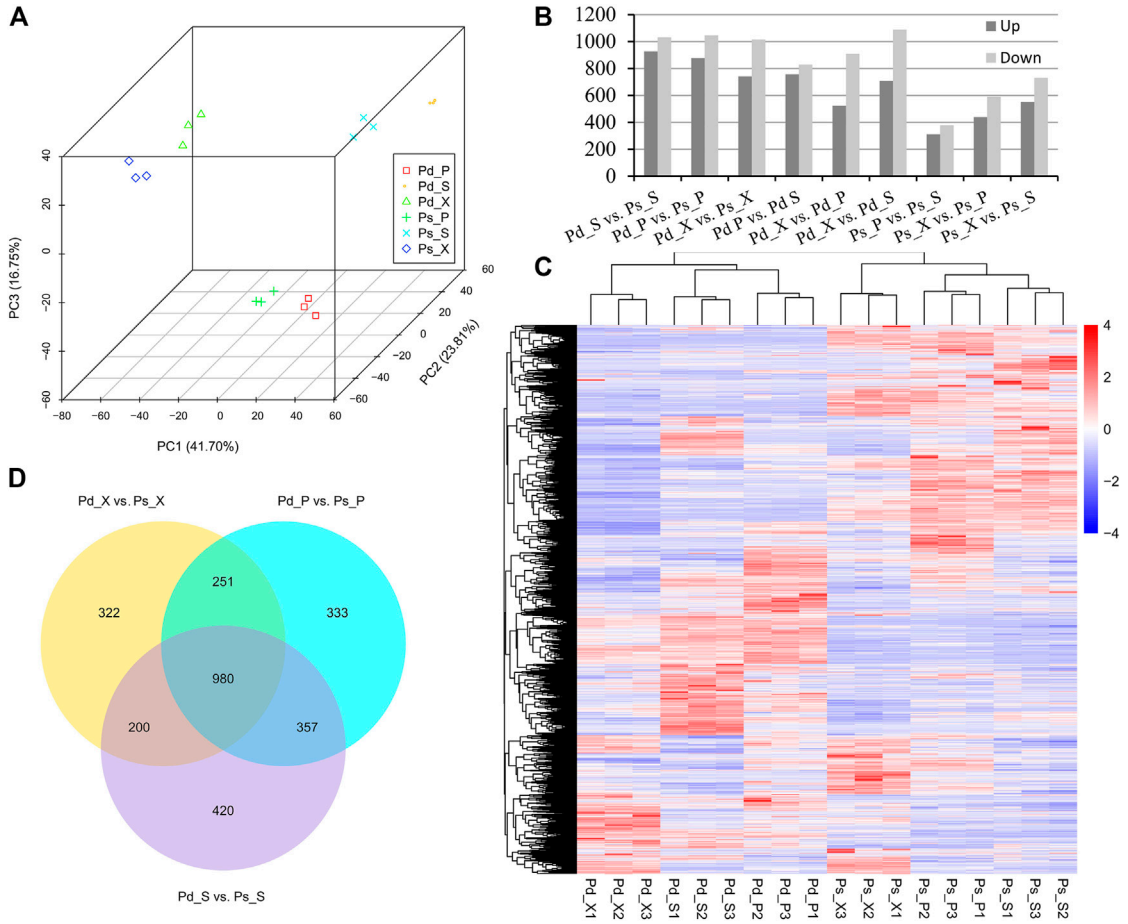


FIGURE 3 | Differentially expressed (DE) lncRNAs identified in pair-wise comparison of 18 samples. **(A)** Principal component analysis (PCA) 3D of the expressed genes showing sample separation. **(B)** Number of DElncRNAs in different comparisons. **(C)** Heat map of DElncRNAs in the two genotypes. **(D)** Venn diagram showing DElncRNAs of Pd_X_vs_Ps_X, Pd_P_vs_Ps_P, and Pd_S_vs_Ps_S.

biosynthesis. And TCONS_00174042, TCONS_00101258, and TCONS_00136338 regulated the structural genes of flavonoid biosynthesis. TCONS_00174042 and TCONS_00101258 co-expressed with *MYB3*, chalcone and stilbene synthase (*CHS*), leucoanthocyanidin reductase (*LAR*), leucoanthocyanidin dioxygenase (*LDOX*), and naringenin 3- dioxygenase (*F3H*). In addition, TCONS_00136338 can coexpress with *LDOX*, *CHS*, *UFGT*, *LAR*, and dihydroflavonol 4-reductase (*DFR*) (**Figure 5B**). Lignin-related lncRNAs and structural genes were highly expressed in xylem of Pd and Ps, while flavonoid-related genes were highly expressed in the SAM (**Figure S5**).

lncRNAs Involved in the Plant Hormone Biosynthesis Signal Transduction Pathway

Hormones such as auxin, cytokinin, and gibberellin play important roles in primary and secondary growth (Nieminen et al., 2008; Xu et al., 2019). In order to further analyze the relationship between lncRNAs and plant hormones, we constructed a co-expression network including the structural genes of auxin, cytokinin, and gibberellin biosynthesis and their possible regulatory lncRNAs (**Figure 6A**). For example, TCONS_00134627 could be coexpressed with *GA2OX8* (Potri.011G134000), *GASA10* (Potri.009G092600), and *SAUR94* (Potri.009G127300) (**Figure 6A**). The expression levels of DEGs and DElncRNAs related to auxin, cytokinin, and gibberellin biosynthesis were shown in the heatmap (**Figure 6B**). The differential genes related to auxins and gibberellins were mainly concentrated in the high expression of Pd_S and Pd_X, while only a few genes are highly expressed in the phloem. And the genes related to cytokinins are highly expressed in Ps_P (**Figure 6B**).

MicroRNAs Involved in the Hormone and Phenylpropanoid Pathway

To further analyze whether lncRNAs were used as ceRNAs to absorb miRNAs, inhibited the effect of miRNAs, and promoted mRNA expression, we conducted sequential analysis. MicroRNAs, a major class of small RNAs with 20–24 nucleotides, create various aspects of plant development and stress responses through posttranscriptionally regulated gene expression (Yu et al., 2019).

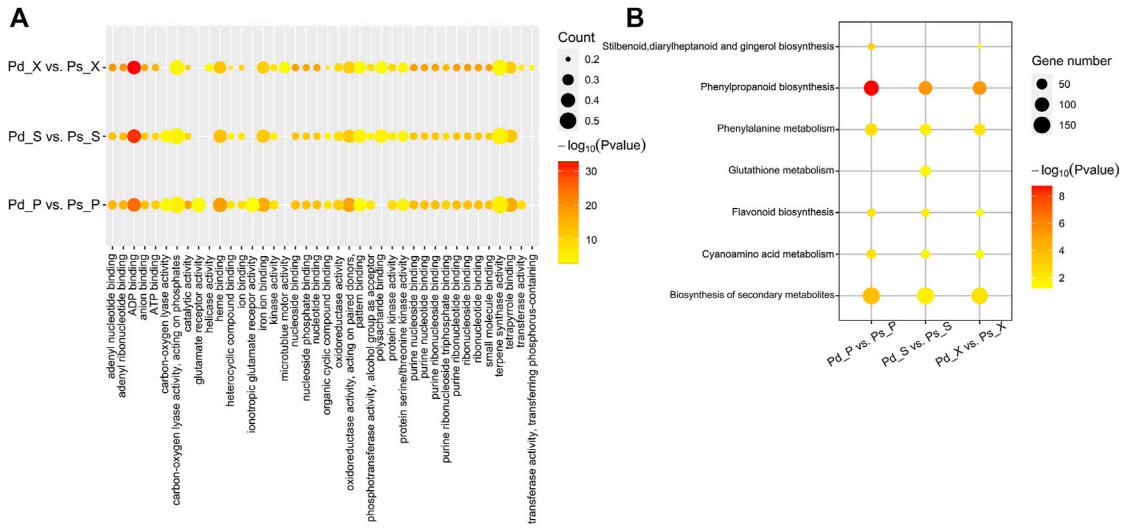


FIGURE 4 | Co-expression mRNA of DElncRANs, GO, and KEGG enrichment. **(A)** GO enrichment analysis of differentially expressed lncRNA-target genes. **(B)** KEGG enrichment analysis of differentially expressed lncRNA-target genes. Node color represents –log10 corrected *p*-value.

A total of 658 miRNA–lncRNA pairs consisting of 188 miRNAs and 200 lncRNAs were identified, including 19 plant hormone–related pairs and 28 phenylalanine-related pairs (**Supplementary Table 6**). TCONS_00066905, hormone-related lncRNA, was predicted to be a target mimic of miR396a and miR396b. In addition, phenylalanine-related TCONS_00023606 and TCONS_00093325 were the target mimics of flavonoid-related regulatory genes miR156h and miR828a (**Figure 7A**). miR396-*GRF* was an important regulatory module of plant growth and development. We found that 12 *GRFs* were differentially expressed and highly expressed in the SAM of Pd, which may be an important reason for the rapid growth of Pd (**Figure 7B**).

Validation of lncRNA and Gene Expression by qRT-PCR

In order to verify the accuracy of transcriptome data, differential genes related to xylem development and hormone signal transduction were selected for qRT-PCR verification. The results showed that the expression trends of three lncRNAs and three mRNAs in six tissues were consistent with those of qRT-PCR, which proved the reliability of transcriptome data (**Figure S6**).

DISCUSSION

Poplar is an important bioenergy tree in the world. Its growth rate and wood quality determine its economic value. With the development of RNA-seq technology, genome-wide mapping has been proved to be a powerful tool for studying primary and secondary growth of poplar. Pd and Ps have great differences in plant height, ground diameter, and other growth characters, as well as wood qualities, such as basic density and lignin content (**Figure 1**). In this study, lncRNAs of woody plants was comprehensively analyzed to study their growth and development and the regulation of wood quality. We identified 6,355 lncRNAs, including 3,572 DElncRNAs and 27, 582 DEGs. lncRNAs in Pd and Ps show similar characteristics with other species (Ci et al., 2019). They are characterized by high tissue specificity and short length. The length of lncRNAs was about 991 bp, and the *cis*- and *trans*- effects were recognized. In this study, we identified a large number of trans-regulatory networks, mainly acting on the phenylalanine pathway. These differentially expressed genes may be an important reason for the differences in growth rate and wood quality.

The SCW provides rigidity and strength for plants to support their body weight and ensure water and nutrient transport (Oda and Fukuda, 2012; Wang and Dixon, 2012). It is mainly composed of lignin, cellulose, and hemicellulose, and their biosynthesis was highly related to the transformation and production of biofuels and biological products (Mohnen, 2008; Hussey et al., 2019). It is regulated by microRNA, *MYB*, *NAC*, and *WRKY* in the SCW complex regulatory network (Zhang et al., 2018b). As an important support to the SCW, lignin determines the conversion efficiency of poplar as biomass energy. In previous studies, lncRNAs participate in the lignin biosynthesis of poplar with TFs and miRNAs (Quan et al., 2019; Zhang et al., 2020c). Also, in our research, we found that 75 lncRNAs, including TCONS_00079190, TCONS_00128372, and TCONS_00007135, can be directly coexpressed with *MYB*s, *VND*s, and lignin structural genes (**Figure 5A**). Flavonoids and lignin are the metabolic pathways of phenylalanine metabolism, and there are some common structural genes. At the same time, flavonoids are important compounds for plants to respond to biological and abiotic stresses. lncRNAs are involved in regulating the anthocyanin biosynthetic pathways in strawberry, buckthorn, and apple (Zhang et al., 2018a; Lin et al., 2018; Yang et al., 2019). We identified a large number of lncRNAs coexpressed with structural genes of the flavonoid pathway, such as *CHS* and *DFR* (**Figure 5B**). The insertion of five *cis*genes encoding gibberellin metabolism or signal proteins affects plant growth (Han

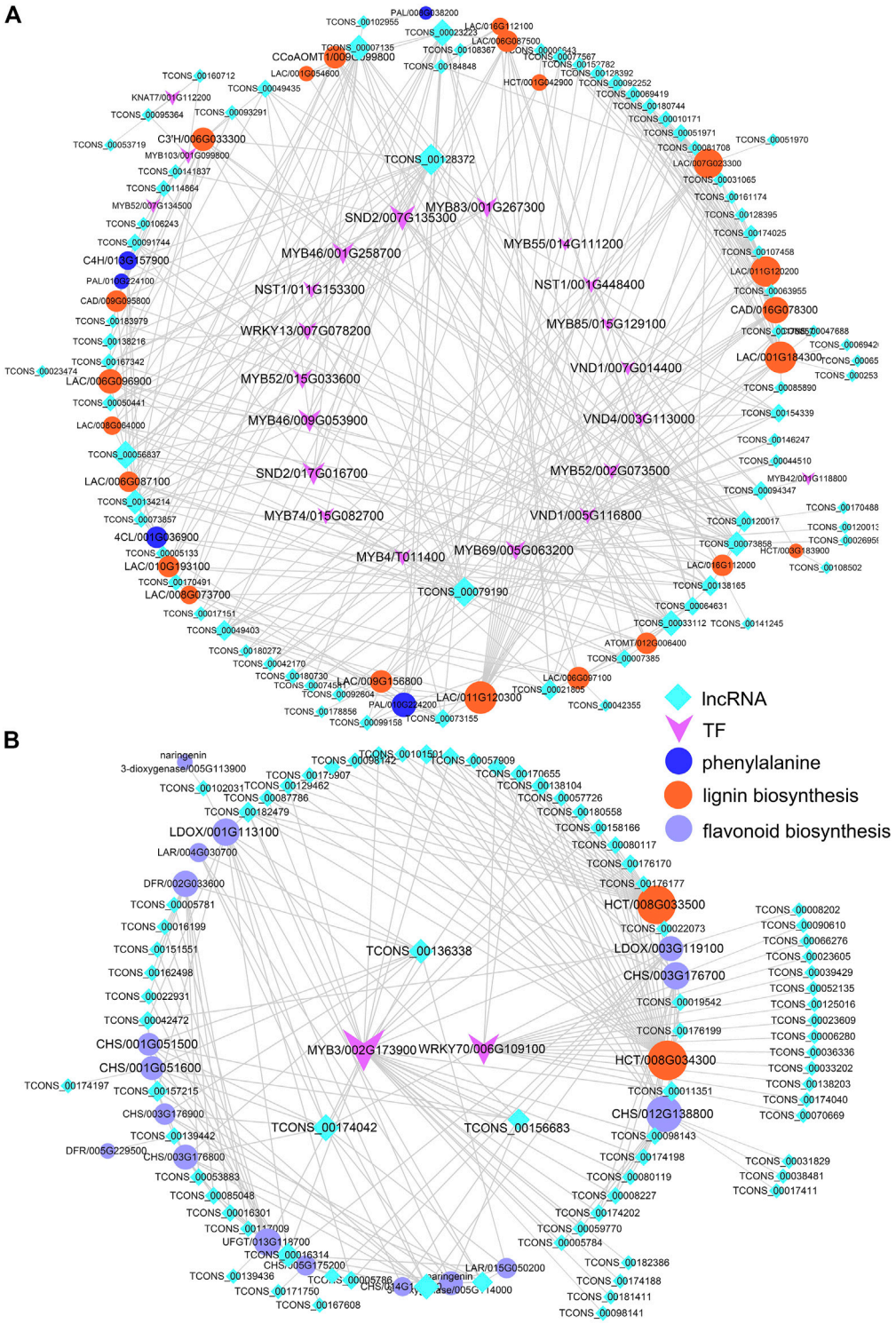


FIGURE 5 | Interaction network diagrams. Interaction networks of differentially expressed lncRNAs, DEmRNAs, and transcription factors (TFs) of phenylalanine biosynthesis, including lignin **(A)** and flavonoid **(B)** biosynthesis.

et al., 2011). The auxin-mediated Aux/IAA-ARF-HB signal cascade regulates the development of the secondary xylem of poplar (Xu et al., 2019). And it was found that a large number of auxin- and gibberellin-related lncRNA–mRNA coexpression networks were identified (**Figure 6**), which was similar to endogenous hormone regulation in the secondary xylem and during tension wood

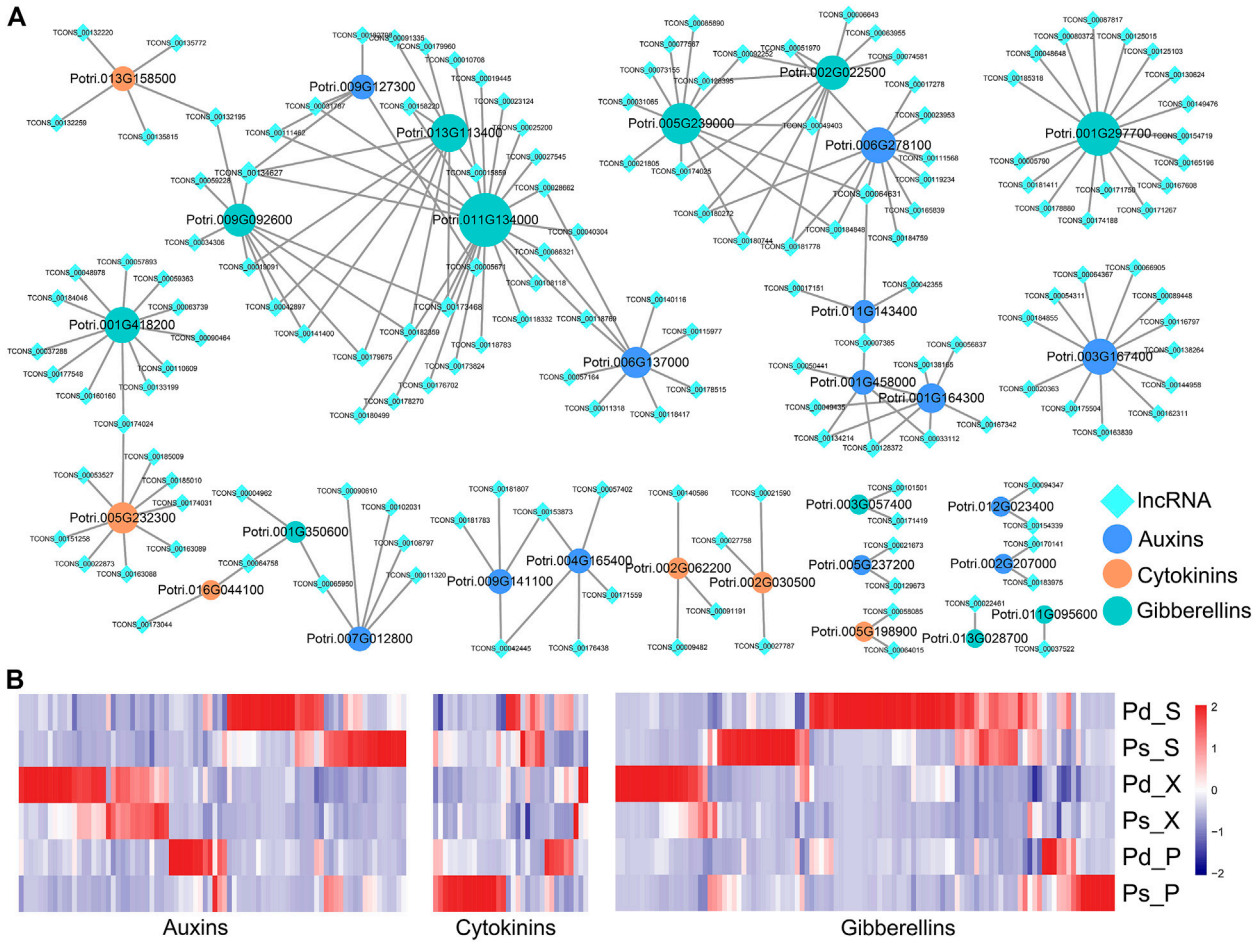


FIGURE 6 | lncRNAs related to plant hormone biosynthesis and signal transduction. **(A)** Co-expression network of DElncRNAs and protein coding genes (PCgenes) involved in auxin, cytokinin, and gibberellin biosynthesis and signal transduction. **(B)** Expression of selected lncRNAs and their predicted co-expression target PCgenes involved in plant hormone biosynthesis signal transduction.

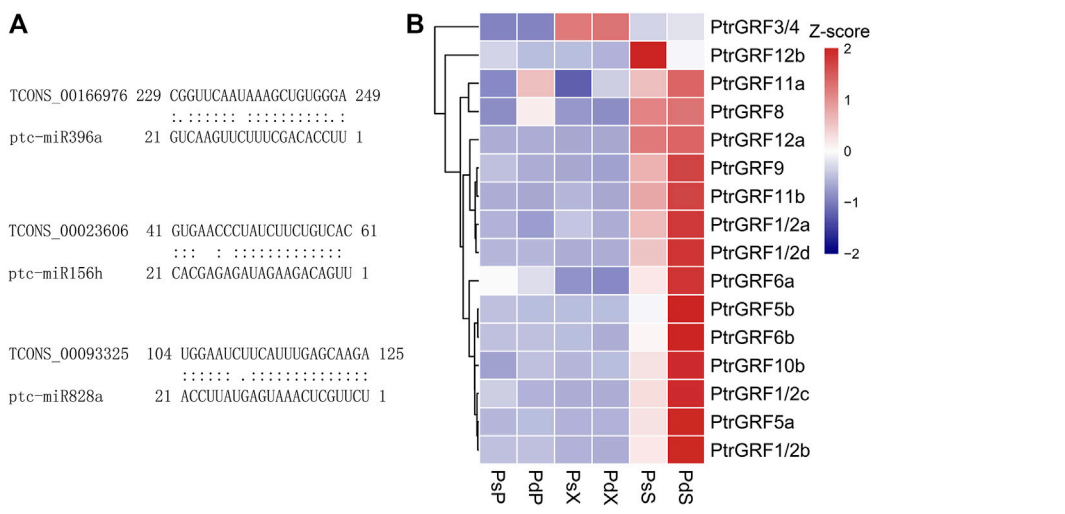


FIGURE 7 | Expression analysis of related lncRNAs as potential targets or target mimics of miRNAs. **(A)** Predicted binding alignment of lncRNAs with ptc-miR396a, ptc-miR156h, and ptc-miR828a. **(B)** Expression patterns of differentially expressed *GRFs and* the target genes of miR396 in the two species.

formation in *Catalpa bungei* (Xiao et al., 2020). Therefore, lncRNAs were widely involved in lignin and flavonoid metabolism and plant growth of Pd and Ps, and they affect their differences.

Posttranscriptional regulation is an important process affecting gene expression and plant development (Li and Zhang, 2016; Yu et al., 2019). miR156 and miR828 are involved in the biosynthesis of flavonoids and anthocyanins by regulating *MYB*s in many species (Tirumalai et al., 2019; B. Zhang et al., 2020a; Wang et al., 2020b). We identified TCONS_00023606 and TCONS_00093325 as the target genes of miRNA156 and miR828, respectively, so they may indirectly participate in the transcriptional regulation of flavonoids through this pathway (**Figure 7**; **Supplementary Table 3**). miR396, *GRF*s, and GRF-INTERACTING FACTORS (*GIF*s) have been proven to control the growth of multiple tissues and organs of multiple species (Debernardi et al., 2014; Liebsch and Palatnik, 2020). *GRF*s are important regulators of the SAM, which are the starting sites of leaf and stem development (Wang Y. et al., 2020). The high expression of *GRFs* in Pd_S may be an important reason for the rapid growth of Pd. lncRNAs–miRNAs–TFs–mRNAs play an important role in regulating the growth of poplars. Therefore, the differences in the growth rate and wood quality of the two poplars may be caused by the joint regulation of these factors, which requires our follow-up for further functional verification.

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

JH conceived and designed the project. LZ, XG, JD, XC, and XP participated in the experiments and data analysis. LZ drafted the manuscript. LZ and JH modified the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Key Research and Development Program of China (2021YFD2200201), the National Natural Science Foundation (32071797), and the National Key Program on Transgenic Research (2018ZX08020002).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.762678/full#supplementary-material

Supplementary Figure 1 | Differentially expressed mRNAs (DEGs) in pair-wise comparison of 18 samples. Number of different genes **(A)** and heat maps **(B)** of different combinations were compared. **(C)** Gene Ontology (GO) enrichment of DEGs in different comparisons. It is detailed enrichment of the biological process (BP). Node color represents –log10 corrected *p*-value.

Supplementary Figure 2 | GO enrichment of lncRNA colocation target genes in nine comparative combinations.

Supplementary Figure 3 | Expression profile of lncRNAs and mRNAs related to photosynthesis in different tissues.

Supplementary Figure 4 | GO **(A)** and KEGG **(B)** analyses of co-location and co-expression mRNAs. BP, biological process; MF, molecular function; CC, cellular component.

Supplementary Figure 5 | Expression patterns of lncRNAs **(A)** and mRNAs **(B)** related to lignin and flavonoid biosynthesis.

Supplementary Figure 6 | lncRNA and mRNA results from RNA-seq verified by qRT-PCR. Mean ± SD from three biological replicates.

Supplementary Table 1 | Primer pairs for qRT- PCR.

Supplementary Table 2 | RNA-seq filter data.

Supplementary Table 3 | Basic information of lncRNAs was identified in *P. deltoides* "Danhong" and *P. simonii* "Tongliao1."

Supplementary Table 4 | Co-location target mRNAs of lncRNAs.

Supplementary Table 5 | Co-expression mRNAs of lncRNAs.

Supplementary Table 6 | Phenylpropanoid biosynthesis–related lncRNAs targeted by miRNAs.

REFERENCES

Alessandra, L., Luca, P., and Adriano, M. (2010). Differential Gene Expression in Kernels and Silks of maize Lines with Contrasting Levels of Ear Rot Resistance after Fusarium Verticillioides Infection. *J. Plant Physiol.* 167, 1398–1406. doi:10.1016/j.jplph.2010.05.015

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* 25, 25–29. doi:10.1038/75556

Chen, M., Wang, C., Bao, H., Chen, H., and Wang, Y. (2016). Genome-wide Identification and Characterization of Novel lncRNAs in *Populus* under Nitrogen Deficiency. *Mol. Genet. Genomics* 291, 1663–1680. doi:10.1007/s00438-016-1210-3

Ci, D., Tian, M., Song, Y., Du, Q., Quan, M., Xuan, A., et al. (2019). Indole-3-acetic Acid Has Long-Term Effects on Long Non-coding RNA Gene Methylation and Growth in *Populus Tomentosa*. *Mol. Genet. Genomics* 294, 1511–1525. doi:10.1007/s00438-019-01593-5

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a Universal Tool for Annotation, Visualization and Analysis in Functional Genomics Research. *Bioinformatics* 21, 3674–3676. doi:10.1093/bioinformatics/bti610

Costa, M. A., Collins, R. E., Anterola, A. M., Cochrane, F. C., Davin, L. B., and Lewis, N. G. (2003). An In Silico Assessment of Gene Function and Organization of the Phenylpropanoid Pathway Metabolic Networks in *Arabidopsis thaliana* and Limitations Thereof. *Phytochemistry* 64, 1097–1112. doi:10.1016/S0031-9422(03)00517-X

Debernardi, J. M., Mecchia, M. A., Vercruyssen, L., Smaczniak, C., Kaufmann, K., Inze, D., et al. (2014). Post-transcriptional Control ofGRFtranscription Factors by microRNA miR396 and GIF Co-activator Affects Leaf Size and Longevity. *Plant J.* 79, 413–426. doi:10.1111/tpj.12567

Elo, A., Immanen, J., Nieminen, K., and Helariutta, Y. (2009). Stem Cell Function during Plant Vascular Development. *Semin. Cel Dev. Biol.* 20, 1097–1106. doi:10.1016/j.semcdb.2009.09.009

Guerra, F. P., Wegrzyn, J. L., Sykes, R., Davis, M. F., Stanton, B. J., and Neale, D. B. (2013). Association Genetics of Chemical wood Properties in Black poplar (Populus Nigra). *New Phytol.* 197, 162–176. doi:10.1111/nph.12003

Ha, C. M., Jun, J. H., and Fletcher, J. C. (2010). Shoot Apical Meristem Form and Function. *Curr. Top. Dev. Biol.* 91, 103–140. doi:10.1016/S0070-2153(10)91004-1

Han, K. M., Dharmawardhana, P., Arias, R. S., Ma, C., Busov, V., and Strauss, S. H. (2011). Gibberellin-associated Cisgenes Modify Growth, Stature and wood Properties in *Populus*. *Plant Biotechnol. J.* 9, 162–178. doi:10.1111/j.1467-7652.2010.00537.x

Heo, J. B., and Sung, S. (2011). Vernalization-mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science* 331, 76–79. doi:10.1126/science.1197349

Hussey, S. G., Grima-Pettenati, J., Myburg, A. A., Mizrachi, E., Brady, S. M., Yoshikuni, Y., et al. (2019). A Standardized Synthetic Eucalyptus Transcription Factor and Promoter Panel for Re-engineering Secondary Cell Wall Regulation in Biomass and Bioenergy Crops. *ACS Synth. Biol.* 8, 463–465. doi:10.1021/acssynbio.8b00440

Jia, H., Osak, M., Bogu, G. K., Stanton, L. W., Johnson, R., and Lipovich, L. (2010). Genome-wide Computational Identification and Manual Annotation of Human Long Noncoding RNA Genes. *RNA* 16, 1478–1487. doi:10.1261/rna.1951310

Jiang, H., Jia, Z., Liu, S., Zhao, B., Li, W., Jin, B., et al. (2019). Identification and Characterization of Long Non-coding RNAs Involved in Embryo Development of *Ginkgo Biloba*. *Plant Signaling Behav.* 14, 1674606. doi:10.1080/15592324.2019.1674606

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317

Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., et al. (2007). CPC: Assess the Protein-Coding Potential of Transcripts Using Sequence Features and Support Vector Machine. *Nucleic Acids Res.* 35, W345–W349. doi:10.1093/nar/gkm391

Li, C., and Zhang, B. (2016). MicroRNAs in Control of Plant Development. *J. Cel. Physiol* 231, 303–313. doi:10.1002/jcp.25125

Liebsch, D., and Palatnik, J. F. (2020). MicroRNA miR396, GRF Transcription Factors and GIF Co-regulators: a Conserved Plant Growth Regulatory Module with Potential for Breeding and Biotechnology. *Curr. Opin. Plant Biol.* 53, 31–42. doi:10.1016/j.pbi.2019.09.008

Lin, Y., Jiang, L., Chen, Q., Li, Y., Zhang, Y., Luo, Y., et al. (2018). Comparative Transcriptome Profiling Analysis of Red- and White-Fleshed Strawberry (Fragaria×ananassa) Provides New Insight into the Regulation of the Anthocyanin Pathway. *Plant Cel Physiol* 59, 1844–1859. doi:10.1093/pcp/pcy098

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., et al. (2012). Genome-wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in. *Plant Cell* 24, 4333–4345. doi:10.1105/tpc.112.102855

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* 25, 402–408. doi:10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8

Ma, J., Bai, X., Luo, W., Feng, Y., Shao, X., Bai, Q., et al. (2019). Genome-wide Identification of Long Noncoding RNAs and Their Responses to Salt Stress in Two Closely Related Poplars. *Front. Genet.* 10, 777. doi:10.3389/fgene.2019.00777

Ma, L., Bajic, V. B., and Zhang, Z. (2013). On the Classification of Long Non-coding RNAs. *RNA Biol.* 10, 924–933. doi:10.4161/rna.24604

Mohnen, D. (2008). Pectin Structure and Biosynthesis. *Curr. Opin. Plant Biol.* 11, 266–277. doi:10.1016/j.pbi.2008.03.006

Nieminen, K., Immanen, J., Laxell, M., Kauppinen, L., Tarkowski, P., Dolezal, K., et al. (2008). Cytokinin Signaling Regulates Cambial Development in poplar. *Proc. Natl. Acad. Sci.* 105, 20032–20037. doi:10.1073/pnas.0805617106

Oda, Y., and Fukuda, H. (2012). Secondary Cell wall Patterning during Xylem Differentiation. *Curr. Opin. Plant Biol.* 15, 38–44. doi:10.1016/j.pbi.2011.10.005

Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., and Demchak, B. (2019). Cytoscape Automation: Empowering Workflow-Based Network Analysis. *Genome Biol.* 20, 185. doi:10.1186/s13059-019-1758-4

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

Quan, M., Du, Q., Xiao, L., Lu, W., Wang, L., Xie, J., et al. (2019). Genetic Architecture Underlying the Lignin Biosynthesis Pathway Involves Noncoding RNAs and Transcription Factors for Growth and wood Properties in Populus. *Plant Biotechnol. J.* 17, 302–315. doi:10.1111/pbi.12978

Severing, E., Faino, L., Jamge, S., Busscher, M., Kuijer-Zhang, Y., Bellinazzo, F., et al. (2018). *Arabidopsis thaliana* Ambient Temperature Responsive lncRNAs. *BMC Plant Biol.* 18, 145. doi:10.1186/s12870-018-1362-x

Song, Y., Chen, P., Liu, P., Bu, C., and Zhang, D. (2020). High-temperature-responsive poplar lncRNAs Modulate Target Gene Expression via RNA Interference and Act as RNA Scaffolds to Enhance Heat Tolerance. *Ijms* 21, 6808. doi:10.3390/ijms21186808

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

Sun, X., Zheng, H., and Sui, N. (2018). Regulation Mechanism of Long Non-coding RNA in Plant Response to Stress. *Biochem. Biophysical Res. Commun.* 503, 402–407. doi:10.1016/j.bbrc.2018.07.072

Tian, J., Song, Y., Du, Q., Yang, X., Ci, D., Chen, J., et al. (2016). Population Genomic Analysis of Gibberellin-Responsive Long Non-coding RNAs inPopulus. *Exbotj* 67, 2467–2482. doi:10.1093/jxb/erw057

Tirumalai, V., Swetha, C., Nair, A., Pandit, A., and Shivaprasad, P. V. (2019). miR828 and miR858 Regulate VvMYB114 to Promote Anthocyanin and Flavonol Accumulation in Grapes. *J. Exp. Bot.* 70, 4775–4792. doi:10.1093/jxb/erz264

Wang, H.-L. V., and Chekanova, J. A. (2017). Long Noncoding RNAs in Plants. *Adv. Exp. Med. Biol.* 1008, 133–154. doi:10.1007/978-981-10-5203-3_5

Wang, H.-Z., and Dixon, R. A. (2012). On-off Switches for Secondary Cell wall Biosynthesis. *Mol. Plant.* 5, 297–303. doi:10.1093/mp/ssr098

Wang, J., Zhou, H., Zhao, Y., Sun, P., Tang, F., Song, X., et al. (2020a). Characterization of poplar Growth-Regulating Factors and Analysis of Their Function in Leaf Size Control. *BMC Plant Biol.* 20, 509. doi:10.1186/s12870-020-02699-4

Wang, L., Park, H. J., Dasari, S., Wang, S., Kocher, J.-P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool Using an Alignment-free Logistic Regression Model. *Nucleic Acids Res.* 41, e74. doi:10.1093/nar/gkt006

Wang, Y., Liu, W., Wang, X., Yang, R., Wu, Z., Wang, H., et al. (2020b). MiR156 Regulates Anthocyanin Biosynthesis through SPL Targets and Other microRNAs in poplar. *Hortic. Res.* 7, 118. doi:10.1038/s41438-020-00341-w

Wei, Z., Du, Q., Zhang, J., Li, B., and Zhang, D. (2013). Genetic Diversity and Population Structure in Chinese Indigenous poplar (*Populus Simonii*) Populations Using Microsatellite Markers. *Plant Mol. Biol. Rep.* 31, 620–632. doi:10.1007/s11105-012-0527-2

Wright, M. W. (2014). A Short Guide to Long Non-coding RNA Gene Nomenclature. *Hum. Genomics* 8, 7. doi:10.1186/1479-7364-8-7

Wu, X., Shi, T., Iqbal, S., Zhang, Y., Liu, L., and Gao, Z. (2019). Genome-wide Discovery and Characterization of Flower Development Related Long Non-coding RNAs in *Prunus Mume*. *BMC Plant Biol.* 19, 64. doi:10.1186/s12870-019-1672-7

Xiao, Y., Yi, F., Ling, J., Yang, G., Lu, N., Jia, Z., et al. (2020). Genome-wide Analysis of lncRNA and mRNA Expression and Endogenous Hormone Regulation during Tension wood Formation in *Catalpa Bungei*. *BMC Genomics* 21, 609. doi:10.1186/s12864-020-07044-5

Xu, C., Shen, Y., He, F., Fu, X., Yu, H., Lu, W., et al. (2019). Auxin-mediated Aux/IAA - ARF - HB Signaling cascade Regulates Secondary Xylem Development in Populus. *New Phytol.* 222, 752–767. doi:10.1111/nph.15658

Xu, W., Yang, T., Wang, B., Han, B., Zhou, H., Wang, Y., et al. (2018). Differential Expression Networks and Inheritance Patterns of Long Non-coding RNAs in castor Bean Seeds. *Plant J.* 95, 324–340. doi:10.1111/tpj.13953

Xue, Z., Liu, L., and Zhang, C. (2020). Regulation of Shoot Apical Meristem and Axillary Meristem Development in Plants. *Ijms* 21, 2917. doi:10.3390/ijms21082917

Yang, T., Ma, H., Zhang, J., Wu, T., Song, T., Tian, J., et al. (2019). Systematic Identification of Long Noncoding RNA S Expressed during Light-induced Anthocyanin Accumulation in Apple Fruit. *Plant J.* 100, 572–590. doi:10.1111/tpj.14470

Ye, Z.-H., and Zhong, R. (2015). Molecular Control of wood Formation in Trees. *Exbotj* 66, 4119–4131. doi:10.1093/jxb/erv081

Yu, Y., Zhang, Y., Chen, X., and Chen, Y. (2019). Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. *Annu. Rev. Cel Dev. Biol.* 35, 407–431. doi:10.1146/annurev-cellbio-100818-125218

Zhan, T., Lu, J., Zhou, X., and Lu, X. (2015). Representative Volume Element (RVE) and the Prediction of Mechanical Properties of Diffuse Porous Hardwood. *Wood Sci. Technol.* 49, 147–157. doi:10.1007/s00226-014-0687-3

Zhang, B., Yang, H.-J., Yang, Y.-Z., Zhu, Z.-Z., Li, Y.-N., Qu, D., et al. (2020a). mdm-miR828 Participates in the Feedback Loop to Regulate Anthocyanin Accumulation in Apple Peel. *Front. Plant Sci.* 11, 608109. doi:10.3389/fpls.2020.608109

Zhang, C., Li, S., Zhao, Z., Hu, J., and Han, Y. (2008). A New poplar Variety *Populus deltoides* 'Danhong'. *Scientia Silvae Sinicae* 44, 169. doi:10.11707/j.1001-7488.20080127

Zhang, G., Chen, D., Zhang, T., Duan, A., Zhang, J., and He, C. (2018a). Transcriptomic and Functional Analyses Unveil the Role of Long Non-coding RNAs in Anthocyanin Biosynthesis during Sea Buckthorn Fruit Ripening. *DNA Res.* 25, 465–476. doi:10.1093/dnares/dsy017

Zhang, J., Song, X., Zhang, L., Jia, H., Peng, X., Zhao, Z., et al. (2020b). Agronomic Performance of 27 Populus Clones Evaluated after Two 3-year Coppice Rotations in Henan, China. *GCB Bioenergy* 12, 168–181. doi:10.1111/gcbb.12662

Zhang, J., Tuskan, G. A., Tschaplinski, T. J., Muchero, W., and Chen, J.-G. (2020c). Transcriptional and post-transcriptional Regulation of Lignin Biosynthesis Pathway Genes in *Populus*. *Front. Plant Sci.* 11, 652. doi:10.3389/fpls.2020.00652

Zhang, J., Xie, M., Tuskan, G. A., Muchero, W., and Chen, J.-G. (2018b). Recent Advances in the Transcriptional Regulation of Secondary Cell wall Biosynthesis in the Woody Plants. *Front. Plant Sci.* 9, 1535. doi:10.3389/fpls.2018.01535

Zhang, L., Liu, B., Zhang, J., and Hu, J. (2020d). Insights of Molecular Mechanism of Xylem Development in Five Black poplar Cultivars. *Front. Plant Sci.* 11, 620. doi:10.3389/fpls.2020.00620

Zhong, R., and Ye, Z.-H. (2015). Secondary Cell walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant Cel Physiol* 56, 195–214. doi:10.1093/pcp/pcu140

Zhou, D., Du, Q., Chen, J., Wang, Q., and Zhang, D. (2017). Identification and Allelic Dissection Uncover Roles of lncRNAs in Secondary Growth of *Populus Tomentosa*. *DNA Res.* 24, 473–486. doi:10.1093/dnares/dsx018

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 Genetics

ORIGINAL RESEARCH
published: 03 March 2022
doi: 10.3389/fgene.2022.843269

Comprehensive Transcriptome Analysis of Stem-Differentiating Xylem Upon Compression Stress in Cunninghamia Lanceolata

Zekun Zhang[1], Huiyuan Wang[1], Ji Wu[1], Yandong Jin[1], Shengwu Xiao[1], Tao Li[1], Xuqinq Liu[1], Hangxiao Zhang[2], Zeyu Zhang[2], Jun Su[2], Jingzao Liu[3], Xiaoyan Wang[4], Yubang Gao[2], Xiangqing Ma[1]* and Lianfeng Gu[2]**

[1]College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, China, [2]College of Forestry, Basic Forestry and Proteomics Research Center, Fujian Agriculture and Forestry University, Fuzhou, China, [3]Taining State-owned Forest Farm, Taining, China, [4]Sanming Forestry Bureau, Sanming, China

OPEN ACCESS

Edited by:
Haifeng Wang,
Guangxi University, China

Reviewed by:
Tao Ma,
Sichuan University, China
Xuan Ma,
Tianjin Normal University, China

***Correspondence:**
Xiangqing Ma
lxymxq@126.com
Yubang Gao
yubanggaofafu@gmail.com
Lianfeng Gu
lfgu@fafu.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *25 December 2021*
Accepted: *11 January 2022*
Published: *03 March 2022*

Citation:
Zhang Z, Wang H, Wu J, Jin Y, Xiao S, Li T, Liu X, Zhang H, Zhang Z, Su J, Liu J, Wang X, Gao Y, Ma X and Gu L (2022) Comprehensive Transcriptome Analysis of Stem-Differentiating Xylem Upon Compression Stress in Cunninghamia Lanceolata. Front. Genet. 13:843269. doi: 10.3389/fgene.2022.843269

Compression wood (CW) in gymnosperm brings great difficulties to wood industry using wood as raw materials since CW presents special wood structure and have different physical and chemical properties from those of normal wood (NW). Chinese fir (*Cunninghamia lanceolata*) is widely distributed in China. However, global transcriptome profiling of coding and long non-coding RNA in response to compression stress has not been reported in the gymnosperm species. In this study, we revealed that CW in Chinese fir exhibited distinct morphology and cytology properties compared with those of NW, including high lignin content, thick and round tracheid cells. Furthermore, we combined both PacBio long-read SMRT sequencing (Iso-Seq) and Illumina short-read RNA-Seq to reveal the transcriptome in stem-differentiating xylem (SDX) under different time points (2, 26, and 74 h) upon compression stress in NW, CW, and OW (opposite wood), respectively. Iso-Seq was successfully assembled into 41,253 *de-novo* full-length transcriptome reference (average length 2,245 bp). Moreover, there were striking differences in expression upon compression stress, which were involved 13 and 7 key enzyme genes in the lignin and cellulose synthesis, respectively. Especially, we revealed 11 secondary growth-related transcription factors show differential expression under compression stress, which was further validated by qRT-PCR. Finally, the correlation between 6,533 differentially expressed coding genes and 372 differentially expressed long non-coding RNAs (lncRNAs) indicates that these lncRNAs may affect cell wall biogenesis and xyloglucan metabolism. In conclusion, our results provided comprehensive cytology properties and full-length transcriptome profiling of wood species upon compression stress. Especially we explored candidate genes, including both coding and long non-coding genes, and provided a theoretical basis for further research on the formation mechanism of CW in gymnosperm Chinese fir.

Keywords: compression wood, Chinese fir (Cunninghamia lanceolata (C.L.), PacBio Iso-seq, LncRNA - long noncoding RNA, lignin biosynthesis, cellulose biosynthesis

INTRODUCTION

Wood plays a very important role in human society. The wooden logs can be processed into plates and sticks, which are used in the manufacture of various furniture and buildings (Ramage et al., 2017). At the same time, the high cellulose content of wood makes it an essential source of raw materials for papermaking and biofuels production (Ragauskas et al., 2006). In natural state, trees develop a special woody tissue due to gravity stimulation or mechanical stress such as strong wind, heavy snow, unstable soil, *etc*. (Du and Yamamoto, 2007). The cross-section of the trunk shows a kind of eccentricity, and one side of the concentric circle is wider (Ruelle, 2014). This special wood structure is called reaction wood (RW) (Dadswell and Wardrop, 1949), which can correct the growth of slanted branches and stems (Sinnott, 1952; Wilson and Archer, 1977).

The RW in angiosperms is called Tension Wood (TW) (Dadswell and Wardrop, 1949), which is formed in the area of the trunk under tension stress (the upper side of the inclined stem) (Scurfield, 1973). Although TW usually has vessels with smaller pore diameters, it has higher cellulose content and lower hemicellulose and lignin content (Dadswell and Wardrop, 1955). TW fiber forms an inner gelatinous cell wall layer (G-layer) which is composed of highly crystalline cellulose (Furuya et al., 1970). However, not all angiosperms have the G-layer (Okuyama et al., 1994; Clair et al., 2006; Sultana et al., 2010). Plant cell walls are composed of the middle lamella, primary cell wall, and secondary cell wall (Dey and Brinson, 1984). The secondary cell wall is usually divided into 3 layers due to different microfibril angle (MFA) (Bidlack, 1992). The S2 layer of TW cell wall has a very low MFA (Donaldson, 2008).

In model species poplars, researchers have carried out a lot of studies on the RW formation through transcriptome sequencing (Yeh et al., 2007; Wang et al., 2013; Martin et al., 2014; Chen et al., 2015a; Chen et al., 2015b; Kerwin, 2021; Ren et al., 2021; Yu et al., 2021). C2H2 transcription factor (TF) (PtaZFP2) negatively regulates the molecular response of poplar and is critical in the process of plants adapting to mechanical stimuli (Martin et al., 2014). In addition to C2H2, 2 TFs (HSFB3-1 and MYB092) in the stems of *Populus tomentosa* were significantly induced in TW and have five common directly targeted genes (Liu et al., 2021). A significantly induced TF (PtrLBD39) from curved poplar trees was revealed from stem xylem based on transcriptome sequencing. Transcriptomic analysis of CRISPR-based *PtrLBD39/22* double mutant revealed that PtrLBD39 indirectly regulates 10 cell wall component genes in TW responsive (Yu et al., 2021). Potential regulatory role during the forming of RW in mature xylem tissues of *Populus tomentosa* is revealed by constructing transcriptome-level regulatory relationship between lncRNAs, miRNAs, and mRNAs in NW, TW, and OW (opposite wood), respectively (Qiu et al., 2015).

The RW in gymnosperms is formed on the underside of the inclined stem (Scurfield, 1973). This special wood structure is called Compression Wood (CW) (Dadswell and Wardrop, 1949). The cell wall of CW tracheid cells is thicker and round in shape (Yoshizawa and Idei, 1987). The content of lignin in CW is increased, and the MFA in the S2 layer is higher (Westing, 1965; Timell, 1982; Donaldson and Singh, 2013). Chinese fir (*Cunninghamia lanceolata*) is an important artificial timber forest species in China, and its area and stock volume account for the highest proportion of artificial forests (Duan et al., 2016; Chen et al., 2020). Thus, Chinese fir forests occupy an extremely important position in the sustainable development of China's forestry.

With the development of sequencing technology, the sRNA library of Chinese fir from leaves, stems, seeds, seedlings, and calli was sequenced using Illumina Next Generation Sequencing (NGS) to identify a broad set of sRNAs, including miRNAs, rasiRNAs, and tasiRNAs (Wan et al., 2012). Small RNA sequencing of vascular cambium reveals the expression changes of miRNA156 and miRNA172 from dormancy to active growth (Qiu et al., 2015). In addition to miRNA, mRNA degradation played a key post-transcriptional regulation during the release of primary dormancy of Chinese fir seed (Cao et al., 2016). Currently, the research of Chinese fir non-coding RNA mainly focuses on the small RNA, and the research on the long non-coding RNA has not been carried out yet.

RW brings great difficulties to the standardized production of the wood industry due to unevenly distributed density of wood material (Ormarssonand et al., 2000; Wimmer and Johansson, 2014). The goal of these industries is to obtain high-purity cellulose (Santos et al., 2013). However, the high lignin content of the CW in gymnosperms has brought great trouble to remove it in pulp processing and biofuel production (Obst, 1990; Himmel et al., 2007). Thus, the studies on the regulation of CW formation at the molecular level are critical in Chinese fir. The process of Chinese fir producing CW provides an excellent model for us to study the secondary growth of Chinese fir (Pilate et al., 2004; Foston et al., 2011). However, the studies of Chinese fir based on sequencing technology have not been widely carried out due to the lack of genomic reference sequence so far. In this study, we combined both long-read SMRT sequencing (Iso-Seq) and short-read RNA-Seq to reveal the transcriptome in stem-differentiating xylem (SDX) of Chinese fir under different time points (2, 26, and 74 h) upon early compression stress in NW, CW, and OW, respectively. Firstly, we revealed the cytology characteristics changes in response to compression stress in Chinese fir. Especially, we revealed transcriptome profile for lignin, cellulose biosynthesis-related genes, and TFs of SDX in response to compression stress. Finally, we found that lncRNAs have potential regulatory functions in this process. Both the transcript profiles and the regulatory network of lncRNA in this study provided many candidate genes for further revealing the formation of CW in Chinese fir. Our research not only provides a foundation for further research on the molecular mechanisms for formation of CW, but also provides new clues to improve wood quality in angiosperms by genetic manipulation.

MATERIALS AND METHODS

Plant Material

Chinese fir was grown in the greenhouse of Fujian Agriculture and Forestry University. The growth condition was maintained at ~25°C with 16 h light/8 h dark photoperiod. We selected Chinese fir with relatively uniform height (1.2–1.3 m) as

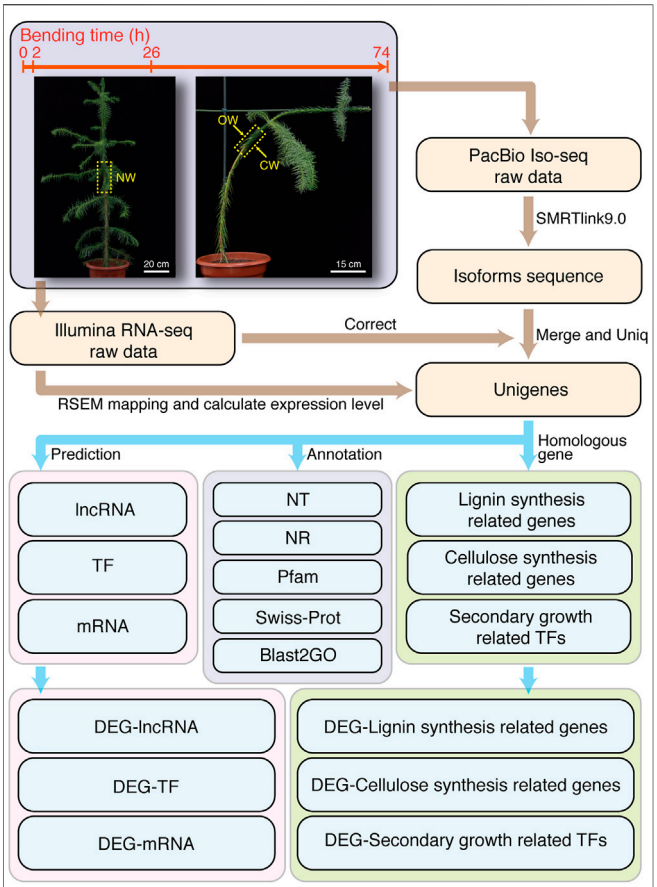


FIGURE 1 | Flow chart of experimental design under artificial bending and Bioinformatic workflow. The bending treatment time point and the material collection time point are 0, 2, 26, and 74 h, respectively.

experimental material. The NW group without bending treatment was used as control (**Figure 1**). We also performed artificial bending treatment following previous method (Paux et al., 2005) to induce the generation of CW in Chinese fir. The angle between the 0.5–0.7 m height of the trunk and the vertical direction is 45° (**Figure 1**). At 10:00 AM on 21 October 2020, SDX tissue samples without bending treatment were collected as NW. The bending treatment of all groups started at 8:00 AM on 22 October 2020. The collection time for 2, 26, and 74 h treatment groups is 10:00 AM on October 22, 23, and 25, respectively (**Figure 1**).

After the bending treatment, the SDX samples were collected immediately. About 20 cm of the trunk at the bent part of the plant was cut. Then we peeled off the bark and scraped the SDX on the upper (the side far from the ground, marked as OW) and lower sides (the side near the ground, marked as CW) of the bent stem. The materials of NW were collected at the equivalent position. All the samples were immersed in liquid nitrogen, and stored at −80°C until RNA isolation.

Anatomy Observation of *Cunninghamia lanceolata*

We collected wood stems from Chinese fir plants that grow straight and lean, respectively. The bark was peeled off, and then the stems were cut into 0.8–1 cm. One-half number of the small fresh stems were cut into 100-micron slices using Leica rotary microtome. The 100-micron slices were dehydrated using alcohol and tert-butanol. Then the water-removed slices were used for scanning electron microscope observation. At the same time, the other small stems were dehydrated and transparentized using alcohol and tert-butanol, and then embedded using paraffin wax. The tissue wax block was cut into 12-micron paraffin slices using Leica rotary microtome. The paraffin slices were dewaxed by xylene and stained using safranin O and fast green. Then the slices were used for upright microscope observation.

Ribonucleic Acid Extraction and Quality Evaluation

The SDX tissue materials were grinded and crushed using the motorized tissue grinders. We extracted the total RNA by using the RNAprep Pure Plant Kit (Polysaccharides & Polyphenolics-rich) (Tiangen, Beijing, China) according to manual. Residual DNA was removed using RNase-free dnase I (Tiangen, Beijing, China). The quantity and purity of RNA samples were evaluated using the NanoDrop 2000 spectrophotometer to meet the requirements of library construction ($1.8 < OD260/280 < 2.2$). The RNA integrity and precise concentration were detected (using Agilent 2,100 RNA 6000 Nano kit) in Agilent 2,100 Bioanalyzer System. RNA samples with RNA Integrity Number (RIN) ≥ 7.5, RNA concentration ≥300 ng/μL, and total RNA amount ≥2 μg were used in the downstream experiment. In total, 14 RNAs from 7 experimental groups (NW, 2hCW, 2hOW, 26hCW, 26hOW, 74hCW, and 74hOW) with two biological repeats were extracted for RNA sequencing and qPCR validation.

Construction of PacBio Iso-Seq Libraries and Sequencing

In total, 14 PacBio Iso-Seq libraries were constructed using the SMARTer PCR cDNA Synthesis Kit (Clontech, #634925). Total RNA samples were reverse-transcribed into first-strand cDNA. The distribution characteristics of amplified products were confirmed by electrophoresis and Agilent 2,100. The cycle number of PCR amplification of cDNA was determined by the pre-amplification to obtain a sufficient amount of cDNA. The purified PCR amplicons were used to construct the SMRT sequencing library using the Template Prep Kit (PacBio, #100–259–100). The sequencing library was purified with PB beads and quantified using Qubit, for which quality control was performed using Agilent 2,100. In brief, 14 libraries with 1–10 kb insert sizes were constructed and sequenced by PacBio Sequel II.

Construction of Illumina Libraries and Sequencing

In this study, 14 libraries were also constructed for second-generation transcriptome sequencing. SDX mRNA was

enriched by Oligo (dT)25 magnetic beads. In first-strand synthesis reaction buffer, with elevated temperature, the rRNA-depleted RNA fragments are cut into smaller fragments using divalent cations. By using random primer and reverse transcriptase, the first-strand cDNA was synthesized. By using buffer, dNTPs (containing dUTP instead of dTTP), RNaseH, and DNA polymerasel, the double-strand cDNA was synthesized. Then, double-strand cDNA was purified. Next, the purified double-stranded cDNA was performed terminal repair, poly-A tail addition, sequencing adapter connection, and the fragment size selection. The sequencing library was obtained by PCR enrichment. The sequencing library was initially quantified using Qubit, and the effective concentration was accurately quantified by qPCR. In total, 14 strand-specific cDNA libraries were constructed for Paired-End RNA Sequencing with reads length of 150 nt.

Full-Length Reference Transcriptome

We used SMRT Link 9.0 command-line tool isoseq3 program (version 3.3.0) to process subreads (https://github.com/PacificBiosciences/IsoSeq). First, we used the ccs module (using the polish parameter) of the isoseq3 program to generate the circular consensus sequencing (CCS read). Next, we used the lima module of the isoseq3 program to remove the primers of the CCS sequence. We used the refine module to trim poly-A tails and remove concatemer sequences. After this step, we obtained full-length non-concatemer reads (FLNC reads). Finally, we used the cluster module to cluster the FLNC sequences and obtain the redundant isoforms sequence. After the combination of the FASTA files from above 14 processed Iso-Seq libraries, CD-HIT-EST software (Huang et al., 2010) was used to remove redundant sequences using following parameter: -c 0.99 -M 0 -T 0 -G 0 -aL 0.90 -AL 100 -aS 0.99 -AS 30. To improve the quality of isoforms sequences (removed redundant), we used Illumina reads to correct isoforms sequences by LoRDEC software (Salmela and Rivals, 2014) using following parameter: -t 5 -b 200 -e 0.4 -T 40 -k 21 -s 3. Finally, we again used CD-HIT-EST software (-c 0.95 -M 0 -T 0) to cluster isoforms sequences after error correction for redundancy removal. In brief, we obtained non-redundant, non-chimeric, full-length unigenes, which makes subsequent analyses more accurate. The flowchart for the bioinformatics analysis is shown in **Figure 1**.

Functional Annotation for Unigenes

For exhaustive annotation of above full-length unigenes, we used BLAST software (version 2.5.0) (Altschul et al., 1997) to search against NR (NCBI Non-redundant Protein Sequence) (Wheeler et al., 2007), NT (Nucleotide Sequence database), and Swiss-Prot (Apweiler et al., 2004), respectively. The NT, NR, and Swiss-Prot annotations correspond to unigenes using blastn, blastx, and blastp, with E-value $<1*10^{-6}$. Moreover, the unigenes were also annotated based on the Pfam (protein family) database (Finn et al., 2014) using HMMSCAN module of HMMER software (version 3.3.2) (Potter et al., 2018). Finally, Gene Ontology (GO) (Ashburner et al., 2000) terms were assigned to unigenes by Blast2GO software (Conesa and Gotz, 2008).

Identification of Long Non-Coding Ribonucleic Acid and Transcription Factor

ORF of unigenes with a base number >200 nt was predicted using getorf module from EMBOSS software (version6.3.1) (Rice et al., 2000). The unigenes whose ORF amino acids were less than 120 were extracted. We used CNCI (Sun et al., 2013), CPC2 (Kang et al., 2017), and PLEK (Li et al., 2014) to identify lncRNA. The sequences predicted by the three software were intersected to obtain high confidence lncRNAs.

For TF identification, we removed lncRNAs from unigenes firstly. Then the CDS region of unigenes was identified and translated into a protein sequence using ESTScan 3.0, which was provided by the Plant Transcription Factor database (PlantTFDB 5.0) from PlantRegMap (Tian et al., 2020). The TF families were identified by mapping the protein sequence to PlantTFDB 5.0.

Quantification of Unigene Upon Bending Treatment

Clean data from 14 Illumina RNA-seq libraries were mapped to full-length unigenes using the RSEM software (--bowtie2, version 1.3.0) (Li and Dewey, 2011). We obtained genes read count using rsem-calculate-expression module of RSEM software from the mapping results. The Spearman's correlation coefficient (SCC) between replicates in each experimental group was calculated as a measure of repeatability. Then, the differentially expressed unigenes among different treatment groups were calculated using DESeq2 (version 1.26.0) (Love et al., 2014) using p-value < 0.05 and fold change>=2.0 as cutoff. According to the classification of lncRNA and TF, the differentially expressed lncRNAs and differentially expressed TFs were obtained.

Co-Expression Analysis of Differentially Expressed Long Non-Coding Ribonucleic Acid and Coding Gene

We further classified differentially expressed unigenes into lncRNAs and coding genes. Pearson correlation coefficient (PCC) between lncRNA and coding gene was calculated and selected to construct the gene co-expression network by Cytoscape software (version 3.8.0) (Kohl et al., 2011) using PCC >0.9 as cutoff.

Gene Ontology Enrichment Analysis

Using the clusterProfiler package (version 4.0.2) (Wu et al., 2021) with p-value < 0.05 as cutoff, we performed GO functions enrichment analysis on differentially expressed coding genes from the 6 treatment groups and from the co-expression network in 20 GO terms, respectively.

Quantitative Real Time-Polymerase Chain Reaction Validation and Expression Analysis

We synthetized the first-strand cDNA using Hifair III SuperMix plus, and 1 μg of total RNA was used. The cDNA was used as

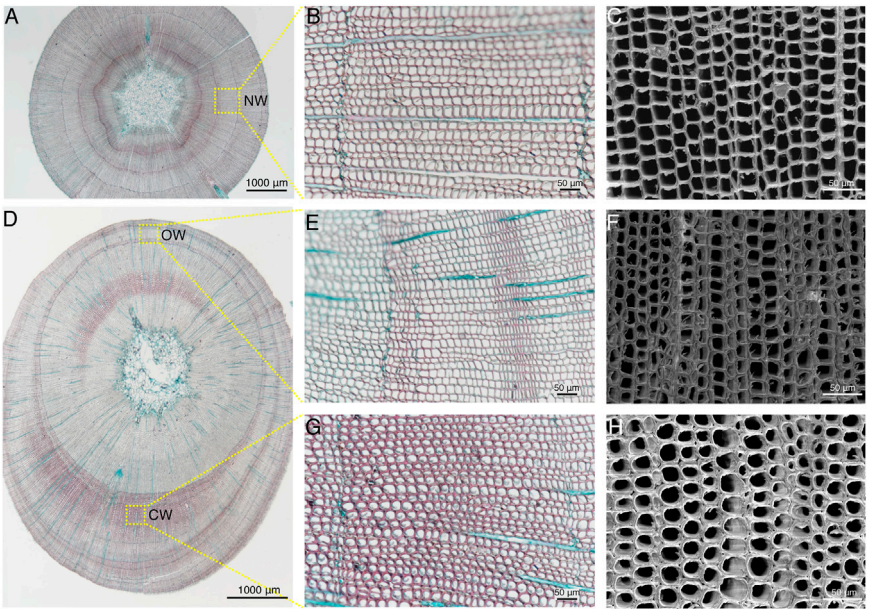


FIGURE 2 | Cytological structure of compression wood (CW) formation in *C. lanceolata*. **(A,D)** Cross sections of straight stem and leaning stem, which were double-stained with Safranin O and Fast Green FCF, respectively. Scale bars = 1,000 μm. **(B,E,G)** Sections of NW, OW, and CW double-stained with Safranin O and Fast Green FCF, respectively. Scale bars = 50 μm. **(C,F,H)** Scanning electron microscopy images of the xylem cross-sections of NW, OW, and CW, respectively. Scale bars = 50 μm.

template, Hieff qPCR SYBR Green Master Mix was used for quantitative reverse transcription-polymerase chain reaction on Mx3005P QPCR system (Stratagene, Santa Clara, CA). Primers for the selected 12 genes (genes related to the lignin and cellulose synthesis pathway, TFs related to the regulation of secondary growth, and predicted lncRNAs) were designed using NCBI Primer-BLAST tool. The primers are listed in **Supplementary Table S1**. We chose the EF1α (elongation factor1-alpha) gene as an internal reference for normalization. This gene is stably expressed in Chinese fir across different tissues and organs. The final 20 μL reaction contained: primers with final concentration of 0.2 μM, 2 μL template cDNA(1:20 dilution), and 1X Power SYBR Green dye. PCR amplification conditions: 95°C for 5 min, 40 cycles of 95°C for 10 s, 60°C for 30 s. After amplification, a thermal denaturation cycle was performed to determine the dissociation curve for verification of amplification specificity: 95°C for 15 s, 60°C for 15 s, and 95°C for 15 s. All qRT-PCR reactions were repeated 4 times.

RESULTS

Cytological Observation of Normal Wood, Compression Wood, and Opposite Wood in Chinese Fir

In order to observe cell structure of the NW, OW, and CW in xylem of Chinese fir, the paraffin sections from straight and leaning trunks were observed by upright microscope, and the fresh slices after dehydrate were observed by scanning electron microscopy (**Figure 2**). The width of the entire xylem of the NW is relatively uniform (**Figure 2A**). The tracheids in the NW were quadrilateral or polygonal and arranged tightly according to the paraffin section under the X200 field of view (**Figure 2B**), which was further confirmed under X500 scanning electron microscope (**Figure 2C**). Compared with the straight Chinese fir, the leaning Chinese fir showed obvious eccentric growth, and the xylem of the CW was wider than that of the OW (**Figure 2D**). Safranin and fast green staining revealed that the color of CW is redder than OW. The lignin content of the xylem in CW is higher than that of OW, which was consistent with the results of earlier studies in Pinus and Chinese fir (Nanayakkara et al., 2009; Li et al., 2021). The tracheid morphology in the OW was consistent with the tracheid morphology in NW (**Figure 2E**), which was further confirmed under the X500 scanning electron microscope (**Figure 2F**). The tracheids in the CW were obviously elliptical or round, with intercellular spaces, and the tracheid cell wall also showed a certain degree of thickening (**Figure 2G**), which were consistent with the results of earlier studies in European yew, red cedar, and Norway spruce (Burgert et al., 2004). The special tracheid morphology in the CW was further confirmed under X500 scanning electron microscope (**Figure 2H**).

Construction of Full-Length Transcriptome Using PacBio Iso-Seq

The full-length transcriptome libraries of 14 SDX samples from Chinese fir were sequenced using PacBio Sequel II platform. A total of 382.78 G subreads bases were generated, which included a total of 236,935,007 reads of insert (ROI). The

statistics of PacBio raw data are shown in **Supplementary Table S2**. The average number of ROI generated by each library was 16,923,929 and the average read quality was >99.9%. After using the SMRTlink9.0 command-line tool to process subreads from the ZMW, we obtained a total of 4,136,486 circular consensus sequencing (CCS) reads, and the average number of CCS reads generated by each library was 295,463. The CCS sequence was processed with 5′, 3′ sequencing primer removal and chimeric sequence removal. Finally, a total of 3,682,735 FLNC sequences were obtained. The average FLNC reads number in each library was 263,052. The FLNC reads were clustered by using the cluster module of isoseq3 software resulted in a total of 415,423 high-quality isoform sequences.

Generation of Non-Redundant Full-Length Unigene

We merged 415,423 high-quality isoform sequences from 14 PacBio full-length transcriptome libraries to perform the first de-redundancy for removing redundant sequences with sequence similarity ≥99% using CD-HIT-EST software. In total, there are 139,318 high-quality sequences remaining. Subsequently, the filtered high-quality sequences were further corrected using LoRDEC software and the short-read sequence generated by Illumina RNA-Seq. The corrected sequence was again removed redundant sequences by using the CD-HIT-EST software with a similarity ≥95%. In total, 41,253 high-quality sequences with an average length of 2,245 bp were obtained as full-length unigenes for consequent analysis.

Integrative Annotation of 41,253 Full-Length Unigenes

The 41,253 unigenes were searched for NT, NR, and Swiss-Prot databases using BLAST for functional annotation. There were 25,453, 34,823, and 27,207 genes which obtained functional annotations from NT, NR, and Swiss-Prot, respectively. Moreover, 32,134 and 26,767 unigenes were annotated using HMMSCAN and BLAST2GO, respectively. The statistics of the full-length unigenes annotations were listed in **Supplementary Table S3**. GO classifications of the full-length unigenes were shown in **Supplementary Table S4**.

Identification of Long Non-Coding Ribonucleic Acid and Transcription Factor

In total, there were 41,225 unigenes remaining after removing the unigenes sequences whose base number was less than 200. Subsequently, the ORF of the remaining sequences was predicted using EMBOSS-getorf to exclude the sequences whose ORF is greater than 120. There were 5,527 remaining unigenes for lncRNA identification. In total, there were 4,949, 5,363, and 4,049 lncRNAs based on the result from CNCI, CPC2, and PLEK, respectively. Finally, a total of 3,602 unigenes were annotated as lncRNAs using the common intersection from three software (**Supplementary Table S5**).

After removing 3,602 lncRNAs from 41,253 unigenes, we annotated TF using 37,651 unigenes according to PlantTFDB database. A total of 1,185 unigenes were annotated as TFs. The most abundant group of TF family was the bHLH family(93), followed by ERF(75), bZIP(74), Trihelix(74), C2H2(73), and C3H(73) family, respectively (**Supplementary Table S6**).

Profiling Differentially Expressed Coding Genes

NGS technology has a greater sequencing depth than third-generation sequencing. Thus, we used the Illumina RNA-seq libraries to quantitatively analyze the expression of the full-length unigenes. The 14 SDX strand-specific mRNA transcriptome libraries were sequenced and generated a total of 324,215,605 paired-end reads. On average, each library produced 23,158,258 reads of 150 nt. The statistics of Illumina RNA-Seq data are shown in **Supplementary Table S7**. The SCC showed good repeatability and reliability between replicates in each experimental group (**Supplementary Figure S1**).

According to the alignment of Illumina RNA-Seq, we obtained all the expression levels of unigenes in 14 samples using RSEM software (**Supplementary Table S8**). The differentially expressed unigenes between the control group (NW) and the six treatment group (2 h_CW, 2 h_OW, 26 h_CW, 26 h_OW, 74 h_CW, 74 h_OW) were identified using DESeq2 R package (p-value < 0.05, $\log_2$FoldChange> 1). From six pairwise comparisons, we obtained the differentially expressed coding genes (DECGs), which included 838 DECGs in all comparisons (**Figure 3A**). Compared with the NW group, the number of the DECGs detected in the 2 h_CW, 2 h_OW, 26 h_CW, 26 h_OW, 74 h_CW, and 74 h_OW was 4,829, 4,470, 3,211, 3,555, 4,219, 4,447, respectively (**Supplementary Table S9**). Overall, the number of DECGs in the 2 h treatment group was the largest, followed by the 74 h treatment group, and the least number of DECGs was in the 26 h treatment group. The number of DECGs unique to the six treatment groups was 770, 499, 209, 251, 285, 390, respectively (**Figure 3A**). The number of unique DECGs in the 2 h treatment group was the largest, and the up-regulated DECGs accounted for about 53% of the total DECGs (**Figure 3B**). The number of up-regulated DECGs in the 26 h treatment group and 74 h treatment group was significantly reduced, accounting for 24 and 17% of the total DECGs, respectively. For down-regulated DECGs, the number of down-regulated DECGs in the 26 and 74 h treatment groups was higher than that in the 2 h treatment group, and the 74 h treatment group had the largest number of down-regulated DECGs (**Figure 3B**). In addition, the number of DECGs obtained by CW compared with OW gradually increased in a time-dependend manner upon compression stress (**Supplementary Figure S2**, **Supplementary Table S10**).

Among the DECGs, we found a total of 371 differentially expressed TFs (DETFs) (**Supplementary Table S11**). The percentage of DETFs in DECGs among the six treatment groups was 4.47, 4.70, 2.96, 3.15, 2.75, and 2.74%, respectively (**Figure 3C**). The proportion of DETFs in CW and OW samples gradually decreases as time increases upon bending treatment

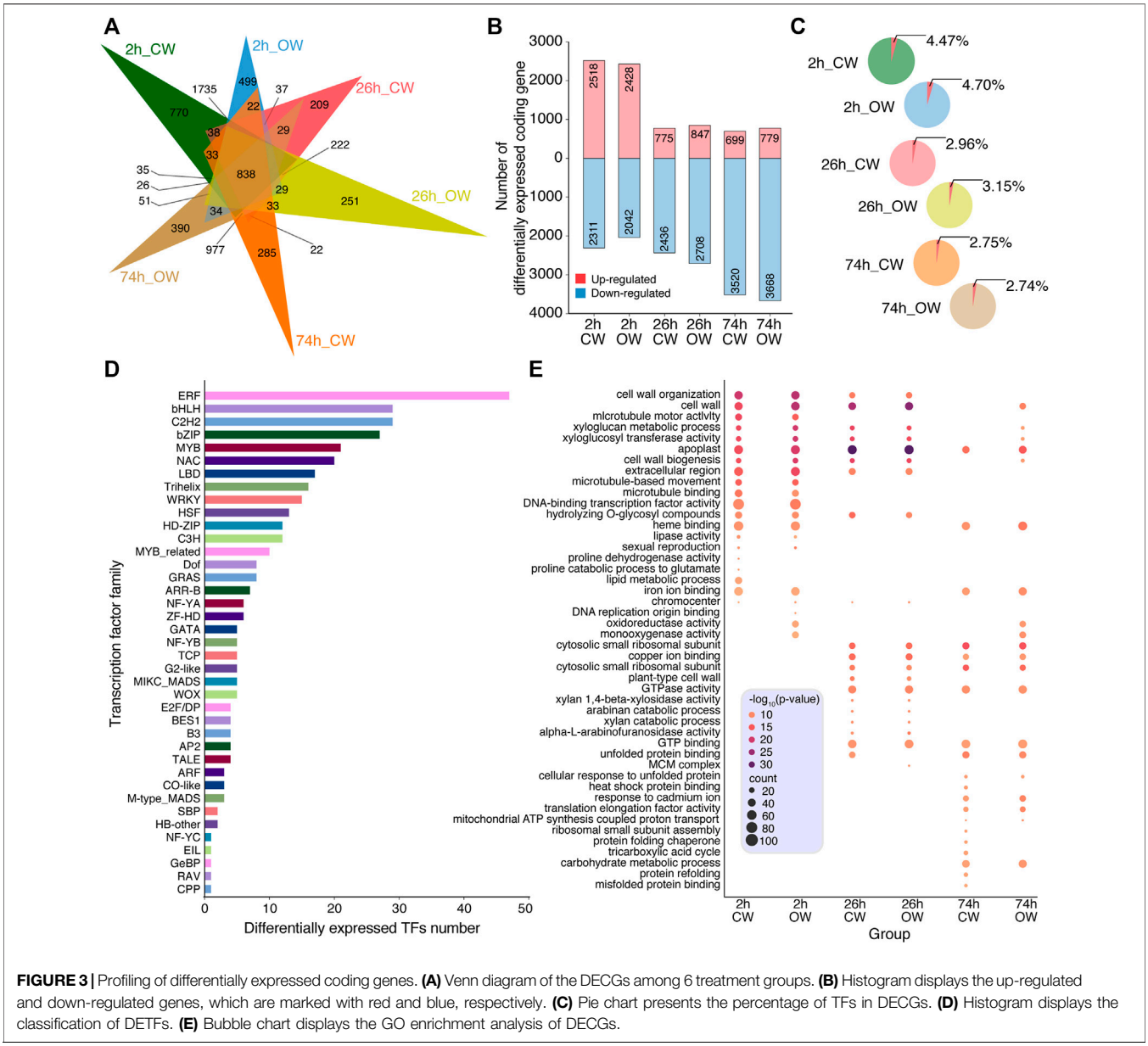


FIGURE 3 | Profiling of differentially expressed coding genes. **(A)** Venn diagram of the DECGs among 6 treatment groups. **(B)** Histogram displays the up-regulated and down-regulated genes, which are marked with red and blue, respectively. **(C)** Pie chart presents the percentage of TFs in DECGs. **(D)** Histogram displays the classification of DETFs. **(E)** Bubble chart displays the GO enrichment analysis of DECGs.

(**Figure 3C**). The 2 h treatment group had the highest proportion of DETFs. Among the DETFs, the number of ERF, bHLH, C2H2, bZIP, MYB, and NAC families was 47, 29, 29, 27, 21, and 20, respectively (**Figure 3D**).

GO enrichment analysis was performed on the DECGs (p-value < 0.05). The top 20 GO terms with the most significant p-value were retained (**Supplementary Table S12**). The DECGs of 2 h_CW and 2 h_OW were enriched in GO terms, including cell wall organization, cell wall, xyloglucan metabolic process, xyloglucosyl transferase activity, cell wall biogenesis, *etc.* (**Figure 3E**). In addition to the above terms, the 26 h_CW and 26 h_OW treatment groups were enriched in plant-type cell wall, xylan 1,4-β-xylosidase activity, arabinose catabolic process, xylan catabolic process, *etc.* (**Figure 3E**). For the 74 h treatment group, only the 74 h_OW treatment group had DECGs enriched in cell wall, xyloglucan metabolic process, cell wall biogenesis, xyloglucosyl transferase activity, *etc.* (**Figure 3E**). Xyloglucan is a component of plant hemicellulose (Fry, 1989). Arabinose is often combined with other monosaccharides, and exists in the form of heteropolysaccharides in hemicellulose, pectic acid, the heartwood of coniferous trees, and in some glycosides (Maki-Arvela et al., 2011). According to the GO enrichment analysis of the DECGs in the six treatment groups, we revealed that the DECGs related to the SDX of Chinese fir are involved in the process of plant cell wall production and plant hemicellulose catabolism after the artificial bending treatment. This might be one of the reasons that Chinese fir produces reaction wood tissue to resist bending stress and restore straight growth.

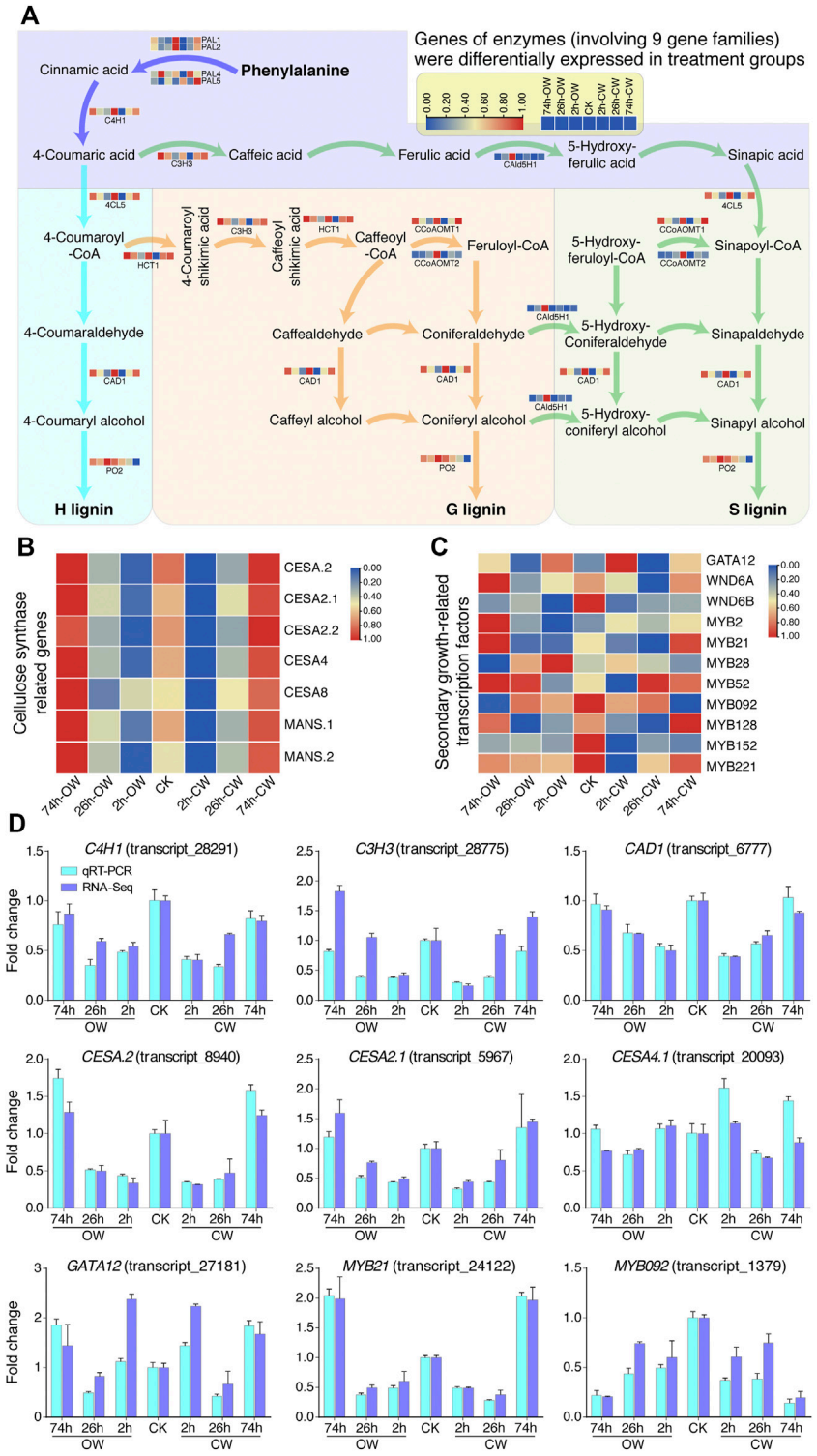


FIGURE 4 | Expression patterns of DEGs related to cell wall synthesis. **(A)** Differential gene expression in lignin synthesis pathway. **(B)** Differential gene expression in cellulose synthesis pathway. **(C)** Differential TF expression in plant secondary growth. In **(A,B)**, and **(C)**, the gene expression was varied from high (red) to low (blue) value. **(D)** Histogram displays qRT-PCR verification of genes involved in cell wall synthesis and transcriptional regulation.

Differential Expression of Genes Related to Cell Wall Synthesis and Transcriptional Regulation

Chinese fir generates CW in response to stress, and the lignin and cellulose content of its xylem have changed. We analyzed the expression of the genes involved in the lignin and cellulose synthesis pathways and the TFs related to the plant secondary growth based on homologous genes from *Populus trichocarpa*. According to RNA-Seq result, we revealed the differential expression levels of lignin, cellulose-related genes, and secondary growth-related TFs (**Figure 4**).

Previous studies have found that the phenylalanine biosynthesizes three kinds of lignin-subunits (H- subunit, G-subunit, S-subunit) through a metabolic grid consisting of 11 enzyme families and 24 metabolites (Chen et al., 2015a; Wang et al., 2018; Liu et al., 2021). In gymnosperms, the G lignin monomer accounts for most of its lignin content. In this study, we found that genes of 13 enzymes (involving 9 gene families) were differentially expressed in treatment groups (**Figure 4A**, **Supplementary Table S13**). Differentially expressed genes involved in lignin synthesis included four phenylalanine ammonia-lyases genes (*PAL1*, *PAL2*, *PAL4*, *PAL5*), one cinnamic acid 4-hydroxylase gene (*C4H1*), one 4-coumaric acid 3-hydroxylase gene (*C3H3*), one 4-coumaric acid: CoA ligase gene (*4CL5*), one cinnamyl alcohol dehydrogenase gene (*CAD1*), one hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase gene (*HCT1*), two caffeoyl-CoA O -methyltransferases genes (*CCoAOMT1*, *CCoAOMT2*), one coniferaldehyde 5-hydroxylase gene (*CAld5H1*), and one peroxidase gene (*PO2*), respectively. Most of the genes involved in the lignin monomer synthesis pathway were generally down-regulated first, and then restored to normal expression levels or up-regulated (**Figure 4A**).

For genes related to the cellulose synthesis pathway (Chen et al., 2015a; Liu et al., 2021), we found 5 cellulose synthase genes (*CESA.2*, *CESA2.1*, *CESA2.2*, *CESA4*, and *CESA8*) and two glycosyltransferase family mannan synthase genes (*MANS.1* and *MANS.2*) are differentially expressed upon bending treatment (**Figure 4B**, **Supplementary Table S14**). Except for the down-regulation of *CESA8* in 26 h_OW, the other genes showed an obvious down-regulation at 2 h stage first, and then recovered to a certain extent at 26 h groups. Finally, their expression levels were up-regulated at 74 h groups.

Two key TFs PtrHSFB3-1 and PtrMYB092 regulate cell wall biosynthesis during the formation of TW in poplar (Liu et al., 2021). PtrMYB092 can directly activate 11 monolignol genes. In the study of the poplar wood formation process, PtrGATA12, as a nuclear localization transcription activator, coordinately regulates the biosynthetic pathway of secondary cell wall components (Ren et al., 2021). In our study, we also detected 11 TFs genes (**Figure 4C**, **Supplementary Table S15**), including *GATA12*, *WND6A*, *WND6B*, *MYB2*, *MYB21*, *MYB092*, *etc.*, which are differentially expressed at different treatment groups. The qRT-PCR analysis of these genes confirmed the expression changes detected by RNA-Seq (**Figure 4D**). The expression analysis of genes related to secondary growth provides valuable clues for explicating the molecular mechanism of CW formation.

Correlation Between Differentially Expressed Long Non-Coding Ribonucleic Acids and Differentially Expressed Coding Genes

Among 9,785 differentially expressed unigenes during the artificial bending process, of which 434 are predicted lncRNAs, and the remaining 9,351 are DECGs (including 371 TFs). The differentially expressed lncRNAs (DElncRNAs) in CW and OW samples at 2, 26, and 74 h treatment groups were shown in **Figure 5A** (**Supplementary Table S16**). The expression changes of these lncRNAs were confirmed by qRT-PCR (**Figure 5B**). And the results of qRT-PCR showed that the expression trends of 3 lncRNAs were consistent with those obtained from RNA-Seq, indicating the accuracy of the RNA-Seq data.

According to the DECGs and DElncRNAs expression, we calculated the PCC (**Supplementary Table S17**), and the correlation between the expression of DElncRNAs and DECGs was shown through the network diagram (**Figure 5C**). We revealed a significant correlation ($r > 0.9$) between 372 DElncRNAs and 6,533 DECGs (6,263 mRNA, 270 TF). GO enrichment analysis was performed on these 6,533 coding genes (p-value < 0.05, keep the 20 GO terms with the most significant p-value, **Figure 5D**, **Supplementary Table S18**). Coding genes presenting correlation with lncRNA were enriched in GO terms such as cell wall, xyloglucan metabolism, cell wall organization, cell wall biogenesis, xyloglucosyl transferase activity, plant-type cell wall, *etc.* (**Figure 5D**). This may imply that lncRNA is involved in regulating the process of compression stress response in Chinese fir.

DISCUSSION

In recent years, many studies have used RW as a model to discover the key genes in the process of wood formation. For example, 9,684 differentially expressed genes were found among TW, OW, and NW transcriptome in *Betula platyphylla* (Wang et al., 2014). In *Populus tomentosa*, 2,417 genes were differentially transcribed between TW and NW (Chen et al., 2015a). In *Betula luminifera*, global transcriptome profiling analysis revealed 13,273 differentially regulated unigenes during the early stage of TW formation (Cai et al., 2018). In addition, 384 up-regulated and 410 down-regulated genes were identified in *Catalpa bungei*, which provides a new basis for explaining the formation of TW (Xiao et al., 2020b). However, most of the studies were carried out in angiosperm, and only a few research were conducted on gymnosperms. Especially the formation of CW in Chinese fir remains unknown.

For non-model organisms, the emergence of transcriptome sequencing technology makes it possible to reveal the special gene expression characteristics without reference genome information (Collins et al., 2008; Ekblom and Galindo, 2011). However, subsequent analysis based on second-generation sequencing technology will encounter great difficulties, especially in the process of assembling the transcripts (Grabherr et al., 2011; Pertea et al., 2015). Moreover, different transcript isoforms

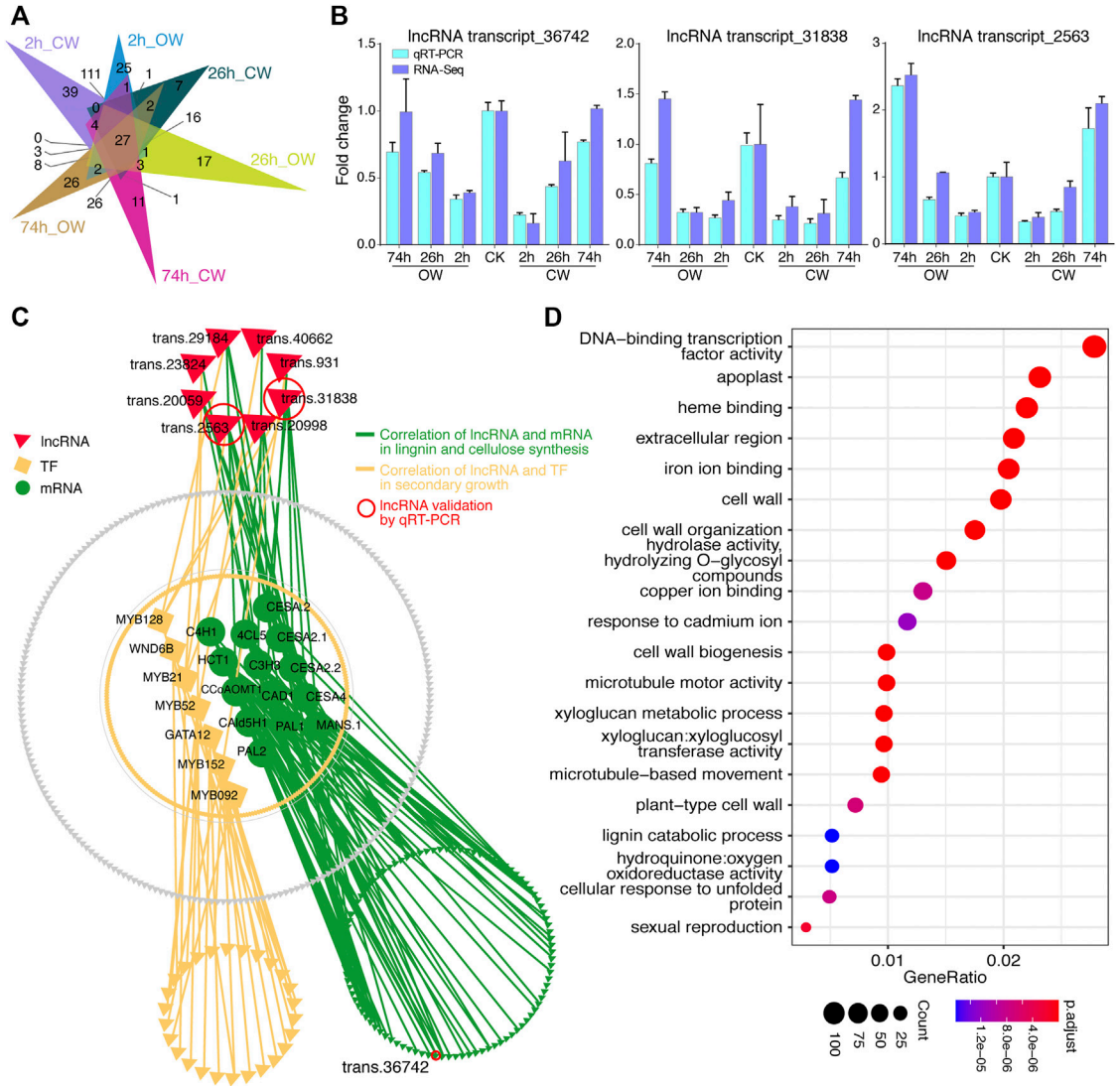


FIGURE 5 | Differential expression analysis of lncRNA. **(A)** Venn diagram shows DElncRNAs of 6 treatment groups. **(B)** Histogram displays qRT-PCR verification of lncRNAs. **(C)** Correlation network diagram between DElncRNAs and DECGs. **(D)** Bubble chart presented GO enrichment analysis of DECGs, which were correlated with lncRNA.

often show high sequence similarity (Martin and Wang, 2011). In this study, we used PacBio Iso-Seq technology to obtain full-length transcriptome of SDX from Chinese fir under artificial bending as reference unigene set. A total of 415,423 full-length sequences from 14 PacBio Ise-Seq libraries were filtered and retained to 41,253, which was further corrected by the short reads from the same batch of RNAs. This makes the full-length unigenes unprecedentedly complete and accurate, and represents the transcriptome characteristics of the Chinese fir under compression stress. The precise and complete unigenes also enable us to obtain more accurate annotation. The high precision full-length unigenes can provide great advantage for the research of Chinese fir.

In our study, we obtained superior transcription maps of Chinese fir NW, CW, and OW SDX under different bending times of 2, 26, and 74 h, respectively. The average length of unigenes we obtained is greater than the 497 bp obtained by previous study (Qiu et al., 2013). In this study, a total of 9,351 DECGs were identified during the process of artificial bending, which indicates that the gene expression of the SDX has changed considerably during this process of CW formation in Chinese fir. In addition, with the continuation of the bending time, the number of DECGs decreased and then increased, which is consistent with previous study in wild-type poplars (Pomies et al., 2017). In another study, the number of differentially expressed genes (DEGs) gradually increased in the 6 h, 48 h, and 7 d groups in *Betula luminifera* (Cai et al., 2018). Therefore, we speculate that the number of DEGs is highest in the early stage (0.5 h, 2 h) under artificial bending, and with the continuation of the bending time, it showed a decreasing trend and then increased.

Moreover, the up-regulated genes in angiosperms accounted for more or nearly half of the DEGs in the initial stage of artificial bending, such as 0.5, 2, or 6 h (Pomies et al., 2017; Cai et al., 2018). While at 24, 48, 72 h, 7 d, 90 d, the DEGs accounted for a larger proportion of down-regulated genes (Chen et al., 2015a; Pomies et al., 2017; Cai et al., 2018; Xiao et al., 2020a; Xiao et al., 2020b). Our study showed the same trend, in the 2 h group, up-regulated genes accounted for 53%, while in the 26 h, 74 h groups, down-regulated genes accounted for a larger proportion (**Figure 2B**). This indicates that the up-regulated genes are dominant in the early phase, while down-regulated genes become increasingly important in the late phase.

In addition, we found that as the bending time increases, the proportion of TF in the DECGs gradually decreases, and the ERF TF family accounts for the largest number of DETFs, which are consistent with previous study in wild-type poplars (Pomies et al., 2017). In addition, study on the secondary growth of poplar has proved that the TF PtrGATA12 directly or indirectly regulates some TFs (WND6A, WND6B, MYB152, and MYB21) to affect the formation of secondary cell walls (Ren et al., 2021). In this study, we explored the continuous expression of TFs related to secondary growth. The *GATA12* expression showed a trend of U-shaped curve (**Figure 3D**). Study on the formation of poplar TW has shown that TF PtrMYB092 can transactivate 11 monoglycol synthesis pathway genes (Liu et al., 2021). In our study, a downward trend was observed in the expression of *MYB092* with the bending time increases, which might affect the expression of monoglycol synthesis pathway genes.

Previous studies have found extensive expression changes in lignin and cellulose synthesis pathway genes during the formation of RW (Paux et al., 2005; Wang et al., 2014; Chen et al., 2015a; Cai et al., 2018; Xiao et al., 2020a; Liu et al., 2021). In *Pinus pinaster*, the cellulose synthase-like A (*CSLA*) gene and the lignin biosynthesis genes, such as *PAL*, *C3H*, *4CL*, *CAD*, *HCT*, *CCoAOMT,* were up-regulated in CW (Villalobos et al., 2012). In *Buxus microphylla* var. *japonica*, the *PAL* gene is up-regulated in RW (Hiraide et al., 2016). In this study, we observed complex changes in lignin synthesis pathway genes, including *PAL4*, *PAL5*, *CAld5H1*, *CCoAOMT2*, and *PO2*, respectively. We speculated that the five genes with special expression trends in the lignin synthesis pathway might be critical in the formation of CW. Other DECGs involved in the synthesis of lignin and cellulose showed a similar trend of decreasing first and then increasing, which seems to be a prevalent response under artificial bending in Chinese fir.

In addition to coding RNA, non-coding RNA also plays an important role in plant transcriptional regulation (Liu et al., 2015; Liaquat et al., 2021; Zhang et al., 2021). Previous study has shown that transcriptome sequencing on the mature xylem tissues of NW, TW, and OW reveals 1,377 possible lncRNAs, of which 776 are differentially expressed in different tissue samples of poplar (Chen et al., 2015b). They report 389 differentially expressed lncRNAs that may target 1,151 genes through *trans*-regulation (Chen et al., 2015b). In our research, we found that there is a significant correlation between the expression levels of 372 DElncRNAs and 6,533 DECGs ($r > 0.9$) (**Figure 5C**). The lncRNA (transcript_31,838) and lncRNA (transcript_29,184) were significantly correlated with *CESA.2*, *CESA2.2*, *CESA4*, *MANS.1*, *MYB21*, *MYB128*, *PAL1*, and *CCoAOMT1*, respectively. Two lncRNAs may target these genes through *trans*-regulation in response to compression stress in Chinese fir. However, further transformation experiment should be carried out to validate the potential regulation in the future.

At present, we only investigated the expression level since Chinese fir genome was unavaiable. Once the genomic is available, our third-generation full-length transcriptome from 14 libraries can be used to investigate the post-transcriptional regulatory mechanism, including alternative splicing (AS) and alternative polyadenylation (APA) (Wang et al., 2017; Zhang et al., 2018; Yin et al., 2019) during the formation of Chinese fir CW in response to compression stress. This will further improve our understanding of the CW formation in Chinese fir. In addition, we expect the knowledge about the target sites of TFs and lncRNAs to obtain a more accurate regulatory network of Chinese fir under artificial bending by using ChIP-seq or DAP-seq when the genome of Chinese fir becomes available.

CONCLUSION

In summary, we provide a full-length reference transcriptome of Chinese fir. Combined with Illumina NGS RNA-Seq, we obtained transcriptomes of the NW, CW, and OW SDX from Chinese fir (*Cunninghamia lanceolata*) under different bent times (2, 26, 74 h). By using 382.78 G subreads bases, we obtained 41,253 high-quality, non-redundant full-length unigenes with average length of 2,245 bp. Moreover, we revealed 20 full-length unigenes related to lignin and cellulose synthesis pathways, and 11 TFs related to secondary growth, which were differentially expressed at different treatment groups, indicating that these genes play a critical role in the process of producing CW in Chinese fir. At the same time, 3,602 possible lncRNAs were predicted, of which 434 lncRNAs were differentially expressed in different treatment groups. 372 DElncRNAs have a significant correlation with 6,533 DECGs. In conclusion, our research provides the first full-length SDX transcription data as an isoform-level reference for Chinese fir. All these findings not only provide a theoretical basis for exploring the regulation mechanism of the production of CW in Chinese fir, but also offer candidate functional genes resources for the production of CW.

DATA AVAILABILITY STATEMENT

The names of the accession number(s) can be found below: PRJNA777739 and PRJNA777106. The datasets presented in this study can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

LG and XM conceived and designed the research. ZkZ performed bioinformatics and experiments. YG, HW, and ZyZ provided assistance for bioinformatics. JW, YJ, SX, TL, JS, JL, and XW

provided assistance for experiments. XL and HZ provided assistance for generating the figure. LG and XM managed the project. ZkZ and LG 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 (31570674), the Innovation Fund for Science & Technology Project from Fujian Agriculture and Forestry University (CXZX2020093A), National Key Research and Development Program of China (2016YFD0600106), and Fujian Forest Seedling Technology Project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.843269/full#supplementary-material

REFERENCES

Altschul, S., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a New Generation of Protein Database Search Programs. *Nucleic Acids Res.* 25 (17), 3389–3402. doi:10.1093/nar/25.17.3389

Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., et al. (2004). UniProt: the Universal Protein Knowledgebase. *Nucleic Acids Res.* 32, 115D–119D. doi:10.1093/nar/gkh131

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556

Bidlack, J. (1992). Molecular Structure and Component Integration of Secondary Cell walls in Plants. *Proc. Okla. Acad. Sci.* 72, 51–56. doi:10.4141/cjps92-097

Burgert, I., Fruhmann, K., Keckes, J., Fratzl, P., and Stanzl-Tschegg, S. (2004). Structure?function Relationships of Four Compression wood Types: Micromechanical Properties at the Tissue and Fibre Level. *Trees* 18 (4), 480–485. doi:10.1007/s00468-004-0334-y

Cai, M., Huang, H., Ni, F., Tong, Z., Lin, E., and Zhu, M. (2018). RNA-seq Analysis of Differential Gene Expression in Betula Luminifera Xylem during the Early Stages of Tension wood Formation. *PeerJ* 6, e5427. doi:10.7717/peerj.5427

Cao, D., Xu, H., Zhao, Y., Deng, X., Liu, Y., Soppe, W. J. J., et al. (2016). Transcriptome and Degradome Sequencing Reveals Dormancy Mechanisms of Cunninghamia Lanceolata Seeds. *Plant Physiol.* 172 (4), 2347–2362. doi:10.1104/pp.16.00384

Chen, H., He, Z., Hong, W., and Liu, J. (2020). An Assessment of Stumpage Price and the Price Index of Chinese Fir Timber Forests in Southern China Using a Hedonic Price Model. *Forests* 11 (4), 436. doi:10.3390/f11040436

Chen, J., Chen, B., and Zhang, D. (2015a). Transcript Profiling of Populus Tomentosa Genes in normal, Tension, and Opposite wood by RNA-Seq. *Bmc Genomics* 16, 164. doi:10.1186/s12864-015-1390-y

Chen, J., Quan, M., and Zhang, D. (2015b). Genome-wide Identification of Novel Long Non-coding RNAs in Populus Tomentosa Tension wood, Opposite wood and normal wood Xylem by RNA-Seq. *Planta* 241 (1), 125–143. doi:10.1007/s00425-014-2168-1

Clair, B., Ruelle, J., Beauchêne, J., Prévost, M. F., and Fournier, M. (2006). Tension Wood and Opposite Wood in 21 Tropical Rain Forest Species. *Iawa J.* 27 (3), 329–338. doi:10.1163/22941932-90000158

Collins, L. J., Biggs, P. J., Voelckel, C., and Joly, S. (2008). An Approach to Transcriptome Analysis of Non-model Organisms Using Short-Read Sequences. *Genome Inform.* 21, 3–14. doi:10.1142/9781848163324_0001

Conesa, A., and Götz, S. (2008). Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. *Int. J. Plant Genomics* 2008, 1–12. doi:10.1155/2008/619832

Dadswell, H. E., and Wardrop, A. B. (1955). The Structure and Properties of Tension Wood. *Holzforschung* 9 (4), 97–104. doi:10.1515/hfsg.1955.9.4.97

Dadswell, H. E., and Wardrop, A. B. (1949). What Is Reaction wood? *Aust. For.* 13 (1), 22–33. doi:10.1080/00049158.1949.10675761

Dey, P. M., and Brinson, K. (1984). Plant Cell-walls. *Adv. Carbohydr. Chem. Biochem.* 42, 265–382. doi:10.1016/S0065-2318(08)60127-4

Donaldson, L. A., and Singh, A. P. (2013). Formation and Structure of Compression wood. *Cell aspects wood formation* 20, 225–256. doi:10.1007/978-3-642-36491-4_9

Donaldson, L. (2008). Microfibril Angle: Measurement, Variation and Relationships - A Review. *Iawa J.* 29 (4), 345–386. doi:10.1163/22941932-90000192

Du, S., and Yamamoto, F. (2007). An Overview of the Biology of Reaction wood Formation. *J. Integr. Plant Biol.* 49 (2), 131–143. doi:10.1111/j.1744-7909.2007.00427.x

Duan, H., Cao, S., Zheng, H., Hu, D., Lin, J., Lin, H., et al. (2016). Variation in the Growth Traits and Wood Properties of Chinese Fir from Six Provinces of Southern China. *Forests* 7 (9), 192. doi:10.3390/f7090192

Ekblom, R., and Galindo, J. (2011). Applications of Next Generation Sequencing in Molecular Ecology of Non-model Organisms. *Heredity* 107 (1), 1–15. doi:10.1038/hdy.2010.152

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the Protein Families Database. *Nucl. Acids Res.* 42, D222–D230. doi:10.1093/nar/gkt1223

Foston, M., Hubbell, C. A., Samuel, R., Jung, S., Fan, H., Ding, S.-Y., et al. (2011). Chemical, Ultrastructural and Supramolecular Analysis of Tension wood in Populus Tremula X alba as a Model Substrate for Reduced Recalcitrance. *Energy Environ. Sci.* 4 (12), 4962–4971. doi:10.1039/c1ee02073k

Fry, S. C. (1989). The Structure and Functions of Xyloglucan. *J. Exp. Bot.* 40 (1), 1–11. doi:10.1093/jxb/40.1.1

Furuya, N., Takahashi, S., and Miyazaki, M. (1970). The Chemical Composition of the Gelatinous Layer from the Tension wood of Populus Euroamericana. *J. Jpn. Wood Res. Soc.* 16 (1), 26–30.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. *Nat. Biotechnol.* 29 (7), 644–652. doi:10.1038/nbt.1883

Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., et al. (2007). Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* 315 (5813), 804–807. doi:10.1126/science.1137016

Hiraide, H., Yoshida, M., Sato, S., and Yamamoto, H. (2016). Common Mechanism of Lignification of Compression Wood in Conifers and Buxus. *Ajps* 07 (7), 1151–1162. doi:10.4236/ajps.2016.77110

Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: a Web Server for Clustering and Comparing Biological Sequences. *Bioinformatics* 26 (5), 680–682. doi:10.1093/bioinformatics/btq003

Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L., et al. (2017). CPC2: a Fast and Accurate Coding Potential Calculator Based on Sequence Intrinsic Features. *Nucleic Acids Res.* 45 (W1), W12–W16. doi:10.1093/nar/gkx428

Kerwin, R. E. (2021). Under Pressure: Transcriptional Regulation of Tension wood in Populus trichocarpa (California poplar). *Plant Physiol.* 186 (1), 212–214. doi:10.1093/plphys/kiab096

Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: Software for Visualization and Analysis of Biological Networks. *Methods Mol. Biol.* 696, 291–303. doi:10.1007/978-1-60761-987-1_18

Li, A., Zhang, J., and Zhou, Z. (2014). PLEK: a Tool for Predicting Long Non-coding RNAs and Messenger RNAs Based on an Improved K-Mer Scheme. *Bmc Bioinformatics* 15, 311. doi:10.1186/1471-2105-15-311

Li, B., and Dewey, C. N. (2011). RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *Bmc Bioinformatics* 12, 323. doi:10.1186/1471-2105-12-323

Li, Z., Zhan, T., Eder, M., Jiang, J., Lyu, J., and Cao, J. (2021). Comparative Studies on wood Structure and Microtensile Properties between Compression and Opposite wood Fibers of Chinese Fir Plantation. *J. Wood Sci.* 67 (1). doi:10.1186/s10086-021-01945-z

Liaquat, F., Munis, M. F. H., Arif, S., Haroon, U., Shi, J., Saqib, S., et al. (2021). PacBio Single-Molecule Long-Read Sequencing Reveals Genes Tolerating Manganese Stress in Schima Superba Saplings. *Front. Genet.* 12, 635043. doi:10.3389/fgene.2021.635043

Liu, B., Liu, J., Yu, J., Wang, Z., Sun, Y., Li, S., et al. (2021). Transcriptional Reprogramming of Xylem Cell wall Biosynthesis in Tension wood. *Plant Physiol.* 186 (1), 250–269. doi:10.1093/plphys/kiab038

Liu, J., Wang, H., and Chua, N.-H. (2015). Long Noncoding RNA Transcriptome of Plants. *Plant Biotechnol. J.* 13 (3), 319–328. doi:10.1111/pbi.12336

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 (12), 550. doi:10.1186/s13059-014-0550-8

Mäki-Arvela, P., Salmi, T., Holmbom, B., Willför, S., and Murzin, D. Y. (2011). Synthesis of Sugars by Hydrolysis of Hemicelluloses- A Review. *Chem. Rev.* 111 (9), 5638–5666. doi:10.1021/cr2000042

Martin, J. A., and Wang, Z. (2011). Next-generation Transcriptome Assembly. *Nat. Rev. Genet.* 12 (10), 671–682. doi:10.1038/nrg3068

Martin, L., Decourteix, M., Badel, E., Huguet, S., Moulia, B., Julien, J. L., et al. (2014). The Zinc finger Protein P Ta ZFP 2 Negatively Controls Stem Growth and Gene Expression Responsiveness to External Mechanical Loads in poplar. *New Phytol.* 203 (1), 168–181. doi:10.1111/nph.12781

Nanayakkara, B., Manley-Harris, M., Suckling, I. D., and Donaldson, L. A. (2009). Quantitative Chemical Indicators to Assess the Gradation of Compression wood. *Holzforschung* 63 (4), 431–439. doi:10.1515/Hf.2009.062

Obst, J. R. (1990). Lignins: Structure and Distribution in wood and Pulp. *MRS Proc.* 197 (1), 11–20. doi:10.1557/PROC-197-11

Okuyama, T., Yamamoto, H., Yoshida, M., Hattori, Y., and Archer, R. (1994). Growth Stresses in Tension wood: Role of Microfibrils and Lignification. *Ann. For. Sci.* 51 (3), 291–300. doi:10.1051/forest:19940308

Ormarssonand, S., Petersson, H., and Dahlblom, O. (2000). Numerical and Experimental Studies on Influence of Compression wood Timber Distortion. *Drying Techn.* 18 (8), 1897–1919. doi:10.1080/07373930008917817

Paux, E., Carocha, V., Marques, C., Mendes de Sousa, A., Borralho, N., Sivadon, P., et al. (2005). Transcript Profiling of Eucalyptus Xylem Genes during Tension wood Formation. *New Phytol.* 167 (1), 89–100. doi:10.1111/j.1469-8137.2005.01396.x

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 (3), 290–295. doi:10.1038/nbt.3122

Pilate, G., Déjardin, A., Laurans, F., and Leplé, J. C. (2004). Tension wood as a Model for Functional Genomics of wood Formation. *New Phytol.* 164 (1), 63–72. doi:10.1111/j.1469-8137.2004.01176.x

Pomiès, L., Decourteix, M., Franchel, J., Moulia, B., and Leblanc-Fournier, N. (2017). Poplar Stem Transcriptome Is Massively Remodelled in Response to Single or Repeated Mechanical Stimuli. *Bmc Genomics* 18, 300. doi:10.1186/s12864-017-3670-1

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER Web Server: 2018 Update. *Nucleic Acids Res.* 46 (W1), W200–W204. doi:10.1093/nar/gky448

Qiu, Z., Li, X., Zhao, Y., Zhang, M., Wan, Y., Cao, D., et al. (2015). Genome-wide Analysis Reveals Dynamic Changes in Expression of microRNAs during Vascular Cambium Development in Chinese Fir, Cunninghamia Lanceolata. *J. Exp. Bot.* 66 (11), 3041–3054. doi:10.1093/jxb/erv103

Qiu, Z., Wan, L., Chen, T., Wan, Y., He, X., Lu, S., et al. (2013). The Regulation of Cambial Activity in Chinese Fir (Cunninghamia Lanceolata) Involves Extensive Transcriptome Remodeling. *New Phytol.* 199 (3), 708–719. doi:10.1111/nph.12301

Ragauskas, A. J., Nagy, M., Kim, D. H., Eckert, C. A., Hallett, J. P., and Liotta, C. L. (2006). From wood to Fuels: Integrating Biofuels and Pulp Production. *Ind. Biotechnol.* 2 (1), 55–65. doi:10.1089/ind.2006.2.55

Ramage, M. H., Burridge, H., Busse-Wicher, M., Fereday, G., Reynolds, T., Shah, D. U., et al. (2017). The wood from the Trees: The Use of Timber in Construction. *Renew. Sustain. Energ. Rev.* 68, 333–359. doi:10.1016/j.rser.2016.09.107

Ren, M., Zhang, Y., Liu, C., Liu, Y., Tian, S., Cheng, H., et al. (2021). Characterization of a High Hierarchical Regulator, PtrGATA12, Functioning in Differentially Regulating Secondary Wall Component Biosynthesis in Populus trichocarpa. *Front. Plant Sci.* 12, 657787. doi:10.3389/fpls.2021.657787

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16 (6), 276–277. doi:10.1016/s0168-9525(00)02024-2

Ruelle, J. (2014). Morphology, Anatomy and Ultrastructure of Reaction wood. *The Biol. React. wood*, 13–35. doi:10.1007/978-3-642-10814-3_2

Salmela, L., and Rivals, E. (2014). LoRDEC: Accurate and Efficient Long Read Error Correction. *Bioinformatics* 30 (24), 3506–3514. doi:10.1093/bioinformatics/btu538

Santos, R. B., Hart, P. W., Jameel, H., and Chang, H. M. (2013). Wood Based Lignin Reactions Important to the Biorefinery and Pulp and Paper Industries. *Bioresources* 8 (1), 1456–1477. doi:10.15376/biores.8.1.1456-1477

Scurfield, G. (1973). Reaction Wood: Its Structure and Function. *Science* 179 (4074), 647–655. doi:10.1126/science.179.4074.647

Sinnott, E. W. (1952). Reaction wood and the Regulation of Tree Form. *Am. J. Bot.* 39 (1), 69–78. doi:10.2307/243809610.1002/j.1537-2197.1952.tb13047.x

Sultana, R. S., Ishiguri, F., Yokota, S., Iizuka, K., Hiraiwa, T., and Yoshizawa, N. (2010). Wood Anatomy of Nine Japanese Hardwood Species Forming Reaction Wood without Gelatinous Fibers. *Iawa J.* 31(2), 191–202. Doi doi:10.1163/22941932-90000016

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 (17), e166. doi:10.1093/nar/gkt646

Tian, F., Yang, D.-C., Meng, Y.-Q., Jin, J., and Gao, G. (2020). PlantRegMap: Charting Functional Regulatory Maps in Plants. *Nucleic Acids Res.* 48 (D1), D1104–D1113. doi:10.1093/nar/gkz1020

Timell, T. E. (1982). Recent Progress in the Chemistry and Topochemistry of Compression wood. *Wood Sci.Technol.* 16 (2), 83–122. doi:10.1007/BF00351097

Villalobos, D. P., Díaz-Moreno, S. M., Said, E.-S. S., Cañas, R. A., Osuna, D., Van Kerckhoven, S. H. E., et al. (2012). Reprogramming of Gene Expression during Compression wood Formation in pine: Coordinated Modulation of S-Adenosylmethionine, Lignin and Lignan Related Genes. *BMC Plant Biol.* 12, 100. doi:10.1186/1471-2229-12-100

Wan, L.-C., Wang, F., Guo, X., Lu, S., Qiu, Z., Zhao, Y., et al. (2012). Identification and Characterization of Small Non-coding RNAs from Chinese Fir by High Throughput Sequencing. *BMC Plant Biol.* 12, 146. doi:10.1186/1471-2229-12-146

Wang, C., Zhang, N., Gao, C., Cui, Z., Sun, D., Yang, C., et al. (2014). Comprehensive Transcriptome Analysis of Developing Xylem Responding to Artificial Bending and Gravitational Stimuli in Betula Platyphylla. *PLoS One* 9 (2), e87566. doi:10.1371/journal.pone.0087566

Wang, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C., Tunlaya-Anukit, S., et al. (2018). Improving wood Properties for wood Utilization through Multi-Omics Integration in Lignin Biosynthesis. *Nat. Commun.* 9, 1579. doi:10.1038/s41467-018-03863-z

Wang, T., Wang, H., Cai, D., Gao, Y., Zhang, H., Wang, Y., et al. (2017). Comprehensive Profiling of Rhizome-Associated Alternative Splicing and Alternative Polyadenylation in Moso Bamboo (Phyllostachys Edulis). *Plant J.* 91 (4), 684–699. doi:10.1111/tpj.13597

Wang, Z., Zhang, H., Yang, J., Chen, Y., Xu, X., Mao, X., et al. (2013). Phylogenetic, Expression, and Bioinformatic Analysis of theABC1Gene Family inPopulus Trichocarpa. *Scientific World J.* 2013, 1–11. doi:10.1155/2013/785070

Westing, A. H. (1965). Formation and Function of Compression wood in Gymnosperms. *Bot. Rev.* 31 (3), 381–480. doi:10.1007/BF02859131

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., et al. (2007). Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 35, D5–D12. doi:10.1093/nar/gkl1031

Wilson, B. F., and Archer, R. R. (1977). Reaction wood: Induction and Mechanical Action. *Annu. Rev. Plant Physiol.* 28 (1), 23–43. doi:10.1146/annurev.pp.28.060177.000323

Wimmer, R., and Johansson, M. (2014). Effects of Reaction wood on the Performance of wood and wood-based Products. *Biol. React. wood*, 225–248. doi:10.1007/978-3-642-10814-3_8

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A Universal Enrichment Tool for Interpreting Omics Data. *The Innovation* 2 (3), 100141. doi:10.1016/j.xinn.2021.100141

Xiao, Y., Yi, F., Ling, J., Wang, Z., Zhao, K., Lu, N., et al. (2020a). Transcriptomics and Proteomics Reveal the Cellulose and Pectin Metabolic Processes in the Tension Wood (Non-G-layer) of Catalpa Bungei. *Ijms* 21 (5), 1686. doi:10.3390/ijms21051686

Xiao, Y., Yi, F., Ling, J., Yang, G., Lu, N., Jia, Z., et al. (2020b). Genome-wide Analysis of lncRNA and mRNA Expression and Endogenous Hormone Regulation during Tension wood Formation in Catalpa Bungei. *BMC Genomics* 21 (1), 609. doi:10.1186/s12864-020-07044-5

Yeh, T. F., Wang, J., Shi, R., Sun, Y. F., and Chang, V. L. (2007). A Novel O-Methyl Transferase-like Gene with a Drastic Ectopic Expression in Response to Tension Wood Formation in Populus Trichocarpa. *Cellulose Chem. Techn.* 41 (9-10), 521–528.

Yin, Z., Zhang, F., Smith, J., Kuo, R., and Hou, Z.-C. (2019). Full-length Transcriptome Sequencing from Multiple Tissues of Duck, Anas platyrhynchos. *Sci. Data* 6, 275. doi:10.1038/s41597-019-0293-1

Yoshizawa, N., and Idei, T. (1987). Some Structural and Evolutionary Aspects of Compression wood Tracheids. *Wood Fiber Sci.* 19 (4), 343–352.

Yu, J., Zhou, C., Li, D., Li, S., Jimmy Lin, Y.-C., Wang, J. P., et al. (2022). A PtrLBD39-Mediated Transcriptional Network Regulates Tension Wood Formation in Populus trichocarpa. *Plant Commun.* 3, 100250. doi:10.1016/j.xplc.2021.100250

Zhang, B., Liu, J., Wang, X., and Wei, Z. (2018). Full-length RNA Sequencing Reveals Unique Transcriptome Composition in Bermudagrass. *Plant Physiol. Biochem.* 132, 95–103. doi:10.1016/j.plaphy.2018.08.039

Zhang, L., Ge, X., Du, J., Cheng, X., Peng, X., and Hu, J. (2021). Genome-Wide Identification of Long Non-coding RNAs and Their Potential Functions in Poplar Growth and Phenylalanine Biosynthesis. *Front. Genet.* 12, 762678. doi:10.3389/fgene.2021.762678

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 04 April 2022
doi: 10.3389/fgene.2022.870446

Small RNA Sequencing Revealed that miR4415, a Legume-Specific miRNA, was Involved in the Cold Acclimation of *Ammopiptanthus nanus* by Targeting an L-Ascorbate Oxidase Gene and Regulating the Redox State of Apoplast

Ming Zhu[1,2,3], *Xue Wang*[4], *Yanqiu Zhou*[1,2,3], *Jinhua Tan*[1,2,3], *Yijun Zhou*[1,2,3]* and *Fei Gao*[1,2,3]*

[1]Key Laboratory of Ecology and Environment in Minority Areas, National Ethnic Affairs Commission, Minzu University of China, Beijing, China, [2]Key Laboratory of Mass Spectrometry Imaging and Metabolomics, National Ethnic Affairs Commission, Minzu University of China, Beijing, China, [3]College of Life and Environmental Sciences, Minzu University of China, Beijing, China, [4]Beijing Center for Disease Prevention and Control, Beijing, China

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Jiaxing Tian,
Beijing Academy of Agricultural and Forestry Sciences, China
Cunfu Lu,
Beijing Forestry University, China

***Correspondence:**
Yijun Zhou
zhouyijun@muc.edu.cn
Fei Gao
gaofei@muc.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: 06 February 2022
Accepted: 07 March 2022
Published: 04 April 2022

Citation:
Zhu M, Wang X, Zhou Y, Tan J, Zhou Y and Gao F (2022) Small RNA Sequencing Revealed that miR4415, a Legume-Specific miRNA, was Involved in the Cold Acclimation of Ammopiptanthus nanus by Targeting an L-Ascorbate Oxidase Gene and Regulating the Redox State of Apoplast. Front. Genet. 13:870446. doi: 10.3389/fgene.2022.870446

MicroRNAs (miRNAs) are small endogenous single-stranded RNAs that regulate plant growth, development, and environmental stress response posttranscriptionally. *Ammopiptanthus nanus*, a rare evergreen broad-leaved shrub in the temperate area of Central Asia, can tolerate freezing stress as low as –30 degrees centigrade in winter, and miRNA might be involved in the cold acclimation which enables *A. nanus* to obtain tolerance to freezing stress. Systematic identification and functional analysis of the miRNAs involved in the cold acclimation in *A. nanus* may promote understanding of the miRNA-mediated gene regulation network underlying cold acclimation. Here, based on small RNA and degradome sequencing, 256 miRNAs and 1,808 miRNA-target pairs were identified in *A. nanus*. A total of 39 cold-responsive miRNAs were identified, of which 29 were upregulated and ten were downregulated. These cold-responsive miRNAs may participate in the cold acclimation by regulating redox homeostasis (miR398, miR4415, and miR408), calcium signaling (miR5225 and miR5211), growth and development (miR159 and miR390), and small RNA–mediated gene silencing (miR168 and miR1515). We found that miR4415, a legume-specific miRNA, is involved in the cold acclimation of *A. nanus* by targeting an L-ascorbate oxidase gene and then regulating the redox state of the apoplast. Our study provides important data for understanding the regulatory role of miRNA in the cold acclimation of *A. nanus*.

Keywords: miRNA, miR4415, cold acclimation, *Ammopiptanthus nanus*, L-ascorbate oxidase

INTRODUCTION

Low temperature is a common unfavorable factor that negatively affects plant growth, development, and productivity and limits the geographical distribution of plant species (Khan et al., 2017; Ding et al., 2020). According to the temperature range, low-temperature stress can be divided into cold stress (0–10°C, chilling temperature) and freezing stress (<0°C, freezing temperature). Cold stress can directly affect the activity of various biological macromolecules, including various enzymes and bio-membranes, resulting in metabolic disorders, photosynthesis inhibition, and increased levels of reactive oxygen species (ROS) (Liu et al., 2018). Under freezing stress, ice crystals are first formed in the intercellular space or apoplast of plants. The ice crystals gradually increase, leading to severe cell dehydration and plasma membrane damage and rupture, further causing intracellular protein denaturation, precipitation of various biological molecules, disorder of cell metabolism, large accumulation of ROS, and even cell death (Bredow and Walker, 2017). Plants in temperate regions can gradually develop tolerance to freezing stress by exposure to low nonfreezing temperatures in autumn, a process known as cold acclimation. During the long evolutionary process, plants have evolved fine-tuned molecular regulatory mechanisms at transcriptional, posttranscriptional, translational, and posttranslational levels to cope with low-temperature conditions (Krutzfeldt and Stoffel, 2006; Liu et al., 2019). Recent studies have shown that in addition to signal transduction components such as mitogen-activated protein kinase (MAPK) pathway and transcription factors such as dehydration responsive element-binding (DREB) that play key roles in stress signal transduction and gene expression regulation and noncoding RNAs, such as microRNA (miRNA), also play essential regulatory roles in cold acclimation (Yu et al., 2019).

miRNAs are a class of endogenous single-stranded non-coding small RNAs (sRNAs) of about 18–24 nt. miRNAs play vital regulatory roles in plant growth and development and respond to environmental factors by negatively regulating protein-coding genes at the posttranscriptional level, thus regulating various biological processes (Song et al., 2019). It has been shown that miRNAs are involved throughout the development of reproductive organs, leaves, and stems and respond to high salinity, drought, and low-temperature stress (Ma et al., 2019). The application of high-throughput sequencing techniques to investigate the effects of cold stress on the expression of the conserved and nonconserved miRNAs in plants has been reported in *Arabidopsis thaliana* (Sunkar and Zhu, 2004), *Oryza sativa* (Lv et al., 2010), *Zea mays* (Yang et al., 2011), *Glycine max* (Xu et al., 2016), and *Triticum aestivum* (Wang et al., 2014). Several miRNAs were found to be involved in plant response to cold stress, including miR156, miR159, miR167, miR168, miR169, miR172, miR319, miR393, miR396, miR398, and miR408 (Megha et al., 2018), and further functional studies revealed the regulatory roles of cold-responsive miRNAs in cold stress response in plants. For example, it was demonstrated that miR156 enhanced cold stress tolerance in rice by targeting the expression of transcription factor genes such as *SQUAMOSA promoter binding protein-like* (*SPL*) (Zhou and Tang, 2019). Furthermore, miR528 decreased the expression of transcription factor *MYB30* by regulating an *F-box domain–containing protein gene* (*Os06g06050*), leading to enhanced cold stress tolerance in rice (Tang and Thompson, 2019).

Ammopiptanthus nanus is a desert shrub belonging to Ammopiptanthus, Leguminosae. This plant species is mainly distributed in the junction of the western Tianshan Mountains and Kunlun Mountains in Xinjiang, Wuqia county of Xinjiang Uygur Autonomous Region, and Uzbekistan. *A. nanus* originates in the Paleotropical region and is a tertiary relic that survived through Mediterranean retreat and climate aridity. *A. nanus* is of considerable scientific value in investigating the changes of paleogeography, paleoclimate, and paleoflora in Central Asia. In addition, *A. nanus* plays a vital role in the maintenance of the fragile ecosystems in arid areas in Central Asia. It can also be used as an appropriate plant species for windbreak, sand fixation, and greening in arid areas. *A. nanus* has evolved high tolerance to low temperature, drought, and other adverse conditions in temperate deserts. It is noteworthy that *A. nanus* is a rare, evergreen, broad-leaved woody plant in the desert area of Central Asia. *A. nanus* can tolerate low-temperature environments as low as –30°C. Thus, *A. nanus* was proposed to be a crucial material for analyzing the mechanism of low-temperature tolerance and screening candidate genes used for genetic engineering for improving low-temperature tolerance of economic plants (Gao et al., 2016; Cao et al., 2019; Cao et al., 2020; Zhang et al., 2021). However, system identification of miRNAs in *A. nanus* has not been reported, and how miRNAs respond to cold acclimation in *A. nanus* is not clear till now.

In the current study, we conducted genome-wide identification of the conserved and species-specific miRNAs in *A. nanus* based on deep sequencing and performed the expression profiling of miRNAs and their targets under cold acclimation. Several conserved and species-specific miRNAs and their targets were identified, and some miRNAs were demonstrated to play roles in cold acclimation of *A. nanus via* negatively regulating their targets. We further analyzed the biological functions of a lineage-specific miRNA, miR4415, in cold acclimation. Our results provided important data for understanding the miRNA-mediated gene regulation network underlying the cold acclimation in *A. nanus*.

MATERIALS AND METHODS

Plant Materials and Cold Acclimation Treatment

Seeds of *A. nanus* were collected from Wuqia county (75°1′E, 39°43′N), Xinjiang Autonomous Region, China. After surface sterilization with 75% (v/v) ethanol and soaking in water for 48 h at room temperature, the seeds were planted in commercial pots containing a 3:1 (v/v) mixture of perlite and vermiculite. Seedlings were grown in a plant incubator with a photoperiod of 16 h, at 22–28°C and relative humidity of 35%. The seedlings were irrigated every 4 days with a solution of half-strength

Hoagland. Eight weeks after germination, seedlings with similar heights were randomly divided into two groups. The first group of seedlings were kept in the plant incubator at 22–28°C for 7 days, and half of the plants were used as the control group (CK), and the other half were exposed to freezing stress treatment (–30°C) for 1 day and used as the freezing-stressed control group (SCK). The other group of seedlings were exposed to cold treatment (5°C/4°C, light/dark) for 7 days for cold acclimation, and half of the plants were used as the cold-acclimated group (CA), and the other half were exposed to freezing stress treatment (–30°C) for 1 day and used as the freezing-stressed cold-acclimated group (SCA). Leaf samples were harvested from each group and used for biochemical parameter measurement. Leaf samples from CK and CA were used for small RNA sequencing, and leaf samples from CK were used for degradome sequencing.

Biochemical Parameter Measurement

Malondialdehyde (MDA) and relative electrolyte leakage (REL) were measured using the methods described by Yasar and Yang, respectively (Yang et al., 1996; Yasar et al., 2008). The activities of malate dehydrogenase (MDH) and ascorbate oxidase (AO) were measured using kits manufactured by Solarbio Science & Technology Co., Ltd. (Beijing, China). The content of ascorbate (ASA) and dehydroascorbate (DHA) was measured using kits manufactured by Solarbio Science & Technology Co., Ltd. (Beijing, China). Five independent plant samples were used (n = 5) for each biochemical parameter measurement.

Small RNA and Degradome Sequencing

RNA samples were prepared using the TRIzol reagent. The concentration and purity of total RNA samples were assayed using Bioanalyzer 2100 and the RNA 6000 Nano LabChip Kit (Agilent, CA, United States) with a RIN value of >7.0. For small-RNA sequencing, approximately 1 μg of total RNA was used to construct a small-RNA library according to the protocol of the TruSeq Small RNA Sample Prep Kits (Illumina, San Diego, United States). Single-end sequencing (36 bp) was conducted on an Illumina Hiseq2500 at the LC Sciences (Hangzhou, China) by following the manufacturer's instructions. Three independent biological replicates were set for each group, and a total of six high-throughput small RNA sequencing data were surveyed. The raw data of small RNA sequencing have been submitted to the NCBI SRA database with accession numbers SRR10906487, SRR10906486, SRR10906485, SRR10906484, SRR10906483, and SRR10906482.

For degradome sequencing, 150 ng of the poly (A) RNA sample was used as input RNA and annealed with biotinylated random primers. After RNA fragments were captured using biotinylated random primers, a 5′ adapter was ligated to the RNAs with monophosphate at the 5′ end. Next, reverse transcription and PCR were performed. Finally, the resulting library was single-end sequenced using an Illumina Hiseq2500 platform at the LC Sciences (Hangzhou, China) by following the vendor's recommended protocol. The raw data of degradome sequencing have been submitted to the NCBI SRA database under the accession number SRR17711209.

Identification of the Conserved and Species-Specific miRNAs

miRNAs were predicted using the ACGT101-miRprogram (version 4.2, LC Sciences), and the predicted miRNAs were then checked manually to remove false miRNAs. The genome sequence of *A. nanus* was downloaded from GigaBase (Gao et al., 2018). The conserved miRNAs and species-specific miRNAs were determined by aligning to the mature sequences of miRNAs from all green plants in the miRBase database (http://www.mirbase.org) (Kozomara et al., 2019). Species-specific miRNAs are miRNAs that are not aligned to the miRNAs deposited in miRBase (mismatches >4 nt) and meet the criteria of high-confidence miRNAs according to the latest miRNA identification protocol (Kozomara and Griffiths-Jones, 2014).

miRNA Target Prediction and Degradome Analysis

The targets of miRNAs were predicted by using psRNATarget with the settings (maximum expectation: 4.0; length for complementary scoring: 20 bp; target accessibility–allowed maximum energy to unpair the target site: 25.0; flanking length around target site for target accessibility analysis: 17 bp in upstream and 13 bp in downstream; range of central mismatch leading to translational inhibition: 9–11 nt) (Dai et al., 2018). CleaveLand 3.0 was used for analyzing sequencing data with default parameters (Addo-Quaye et al., 2009). CPC2.0 was used to evaluate the coding potential of miRNA targets (Kang et al., 2017).

Differential Expression Analysis of miRNAs

The read counts from CK and CA were normalized to transcripts per million reads (TPM) (Dewey, 2010), and the differentially expressed miRNAs were determined by comparing the conserved or species-specific miRNA expression between two groups. Genes with a fold change ≥1.5 or ≤0.67 and a p-value <0.05 were identified as differentially expressed miRNAs. The heatmap was generated by using TBtools (Chen et al., 2020).

qRT-PCR Analysis of miRNAs and Their Targets

RNA samples were prepared from the leaves using TRIzol reagent. qRT-PCR analysis of miRNAs and their targets were performed using the methods described previously (Abla et al., 2019). All primers were designed based on the sequences of selected miRNAs and their targets (**Supplementary Table S1**). The internal control genes are U6 snRNA (for miRNA) and 18S rRNA (for target gene). Three technical replicates were performed for each analyzed gene, and the expression levels of miRNAs and their target genes were calculated using the $2^{-\Delta\Delta Ct}$ method (Varkonyi-Gasic et al., 2007).

Promoter Analysis

Plant promoter prediction software PlantCARE (http://intra.psb.ugent.be: 8080/PlantCARE) was used to predict the *cis*-

acting elements involved in stress response or hormone response, and 1000 bp upstream sequences of the stem-loop precursors of each miRNA were used for promoter prediction.

Identification, Multiple Sequence Alignment, and Phylogenetic Analysis of miR4415

To identify miR4415 and their precursors, mature sequences of know miR4415 were downloaded from miRBase and used to search for homologous sequences using the BLAST program, allowing four mismatches. Then, 400-bp flanking sequence was retrieved from reference sequences and folded using the online tool mfold to evaluate their secondary structure. Only those genomic loci with a stem-loop structure with an MFE value less than 30 kcal/mol and having the typical secondary structure for miRNA precursors were retained for subsequent analysis. All plant genomes in Phytozome v12 (https://genome.jgi.doe.gov/portal/PhytozomeV12) were downloaded and used for identifying miR4415. Multiple sequence alignment was performed using Clustal and visualized with Weblogo. The phylogenetic tree of the precursor sequences of miR4415 was built by MEGA X using the neighbor-joining (NJ) method.

Dual-Luciferase Reporter Assay

Dual-luciferase reporter assay was performed according to the method for studying miRNA function (Clement et al., 2015). The precursor sequence of *A. nanus* miR4415 was ligated to the pGreen II 62 SK vector, and the coding sequence of L-AO was ligated to the pGreen II 0800-LUC vector. All primers were designed based on the sequences of Pre-miR4415, L-AO, pGreen II 62 SK, and pGreen II 0800-LUC vector (**Supplementary Table S1**). Extraction of *A. thaliana* protoplasts was performed as described by Yoo et al. (2007). The luciferase signals were detected using a Bio-Lite TM Luciferase Assay system kit manufactured from Vazyme Biotech Co., Ltd. (Nanjing, China). Five independent repeats were used for this assay.

Subcellular Localization Analysis

Subcellular localization analysis of L-AO protein was performed as described in the literature (Yin et al., 2021). The coding sequence of the *L-AO* gene was ligated to pCAMBIA1303, and *L-AO* was transiently expressed in tobacco leaves for 48–72 h and imaged by using a confocal laser-scanning microscope (Leica TCS Sp8).

Extraction of the Leaf Apoplast Fluid of *Ammopiptanthus nanus*

The leaf apoplast fluid of *A. nanus* was obtained using the infiltration–centrifugation methods (O'Leary et al., 2014). Water was used as the buffer, a 60-ml syringe was used for leaf infiltration, and the resulting leaf apoplast fluid was recovered by centrifugation for 10 min at 3,000 g.

RESULTS

Cold Acclimation Enhances the Tolerance to Freezing Stress in *Ammopiptanthus nanus*

To evaluate the effect of cold acclimation on the tolerance to freezing stress in *A. nanus*, cold-acclimated and untreated *A. nanus* seedlings were subjected to freezing stress treatment, and MDA and REL, two biochemical indicators reflecting the degree of cell damage, were measured. The seedlings in SCK exhibited significant wilting, while the seedlings in SCA showed only slight wilting (**Figure 1A**). The MDA and REL values in SCA were significantly lower than those of SCK (**Figures 1B,C**), indicating that cold acclimation alleviated the damage to seedlings of *A. nanus* caused by the freezing stress treatment. In other words, high freezing tolerance can be obtained by cold acclimation in *A. nanus* seedlings.

Cold Acclimation Resulted in a Significant Alteration in the Small RNA Population in *Ammopiptanthus nanus*

To evaluate the influence of cold acclimation on the sRNA population and identify miRNAs involved in cold acclimation in *A. nanus*, six sRNA libraries were constructed by using leaf samples isolated from CK and CA seedlings. CK and CA samples from *A. nanus* were sequenced by high-throughput sequencing technology. Approximately, 18 M raw sequence reads were obtained for each library. After removing adapters, low-quality sequences and sequences smaller than 18 nt, 12.28–16.10 M clean reads with sizes ranging from 18 to 25 nt were obtained for these sRNA libraries (**Table 1**).

The size distribution pattern of the three sRNA biological replicates in each group (CK or CA) was similar (**Figure 2**). In group CK, the top four abundant small RNA size categories were 24, 21, 22, and 20 nt groups, and such a size distribution pattern was similar to that of the typical plant mature sRNAs reported in several plant species, for example, *A. thaliana*, *Medicago truncatula* (Szittya et al., 2008), *O. sativa* (Morin et al., 2008), and *Citrus trifoliate* (Song et al., 2010). While there was a significant difference in the sRNA size distribution pattern between the CK and CA groups, the top two abundant sRNA size categories were 21 nt and 24 nt. These data suggested that cold acclimation resulted in a significant alteration in the sRNA population in *A. nanus* leaves.

Identification of the Conserved and the Species-Specific miRNAs in *Ammopiptanthus nanus*

To identify miRNAs, all sRNAs of the six libraries were mapped to *A. nanus* genome sequences, and the flanking sequences were folded to determine whether a qualified hairpin could be formed. Consequently, 188 conserved mature miRNA sequences and 107 corresponding stem-loop precursors were identified (**Supplementary Table S2**). All the stem-loop precursors

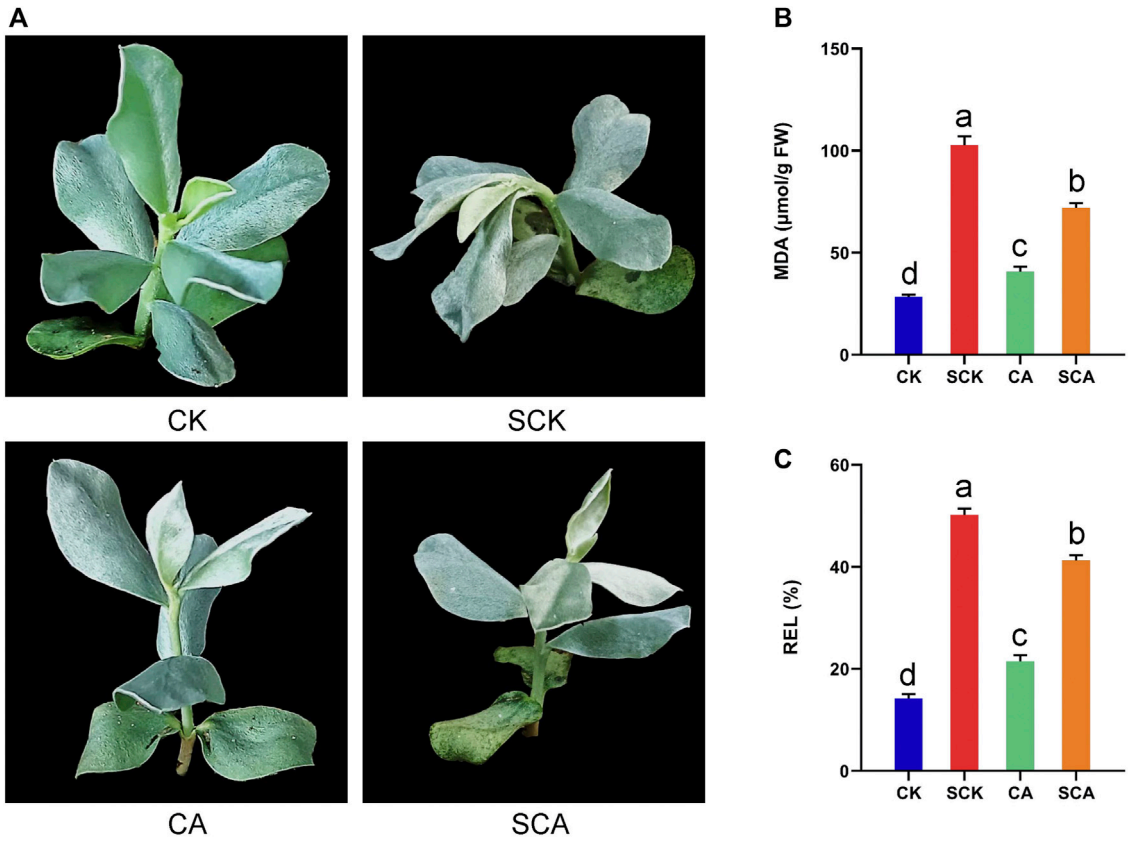


FIGURE 1 | Cold acclimation enhanced the freezing stress tolerance of *A. nanus*. **(A)** Phenotype of *A. nanus* seedling in CK, SCK, CA, and SCA. The change of MDA **(B)** and REL **(C)** in the leaves of *A. nanus* before (CK and CA) and after freezing treatment (SCK and SCA). Values are expressed as means ± SD. Duncan's method was used for multiple comparison analysis of variance, $n = 3$.

TABLE 1 | Statistics of the clean reads generated by high-throughput sequencing.

Library	Number of raw reads	Number of high-quality reads	Reads without 3′ adapter	Reads without insert fragment	5′ adapter contaminants	Reads with polyA	Reads smaller than 18nt	Number of clean reads
miCK_1	176,958,37 (100%)	173,638,56 (98.12%)	94,823 (0.54%)	158,912 (0.91%)	71,408 (0.41%)	728 (0.01%)	4,759,128 (27.40%)	12,278,857 (70.71%)
miCK_2	18,724,118 (100%)	18,316,541 (97.82%)	70,928 (0.38%)	147,893 (0.80%)	52,832 (0.28%)	1340 (0.01%)	3,316,221 (18.10%)	14,727,327 (80.40%)
miCK_3	19,377,347 (100%)	19,001,586 (98.06%)	61,387 (0.32%)	88,535 (0.46%)	59,142 (0.31%)	1097 (0.01%)	2,690,773 (14.16%)	16,100,652 (84.73%)
miCA_1	18,647,262 (100%)	18,277,927 (98.01%)	108,694 (0.59%)	128,250 (0.70%)	53,975 (0.29%)	918 (0.01%)	3,387,959 (18.53%)	14,598,131 (79.86%)
miCA_2	19,485,083 (100%)	19,126,024 (98.15%)	66,229 (0.34%)	109,935 (0.57%)	82,378 (0.43%)	576 (0.01%)	5,235,308 (27.37%)	13,631,598 (71.27%)
miCA_3	18,046,481 (100%)	17,705,703 (98.11%)	66,387 (0.37%)	130,810 (0.73%)	75,991 (0.42%)	504 (0.01%)	4,288,459 (24.22%)	13,143,552 (74.23%)

showed similarity to the miRNA precursors of green plants in the miRBase (**Supplementary Table S2**), indicating that these conserved miRNAs are generated from conserved stem-loop precursors. Most of the conserved miRNAs are supported by the simultaneous presence of the hairpin 5p and 3p arms, but no miR* sequences have been found for 80 stem-loop precursors. Among the 80 precursors, 35 with only 5p arm miRNAs and 45 with 3p arm miRNAs were found.

By aligning the stem-loop sequences obtained above with the miRBase database, 41 conserved miRNA families were identified

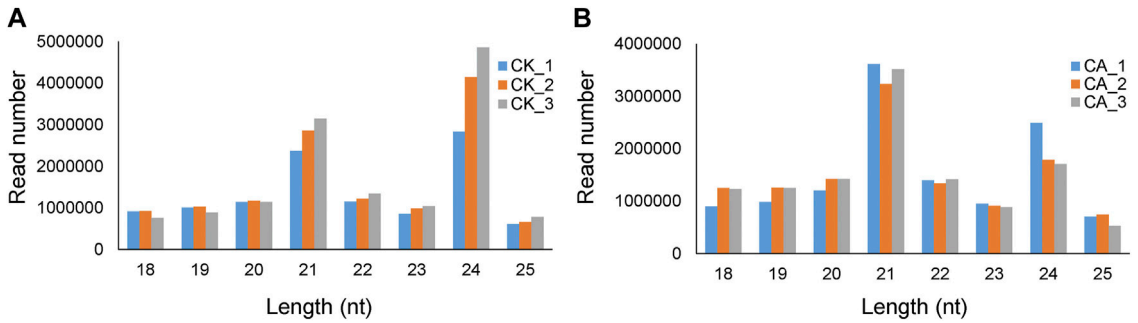


FIGURE 2 | Length distribution of sRNA sequences in the six sRNA libraries prepared from leaves of the *A. nanus* seedlings cultured in untreated conditions (CK) **(A)** and under cold acclimation (CA) **(B)**.

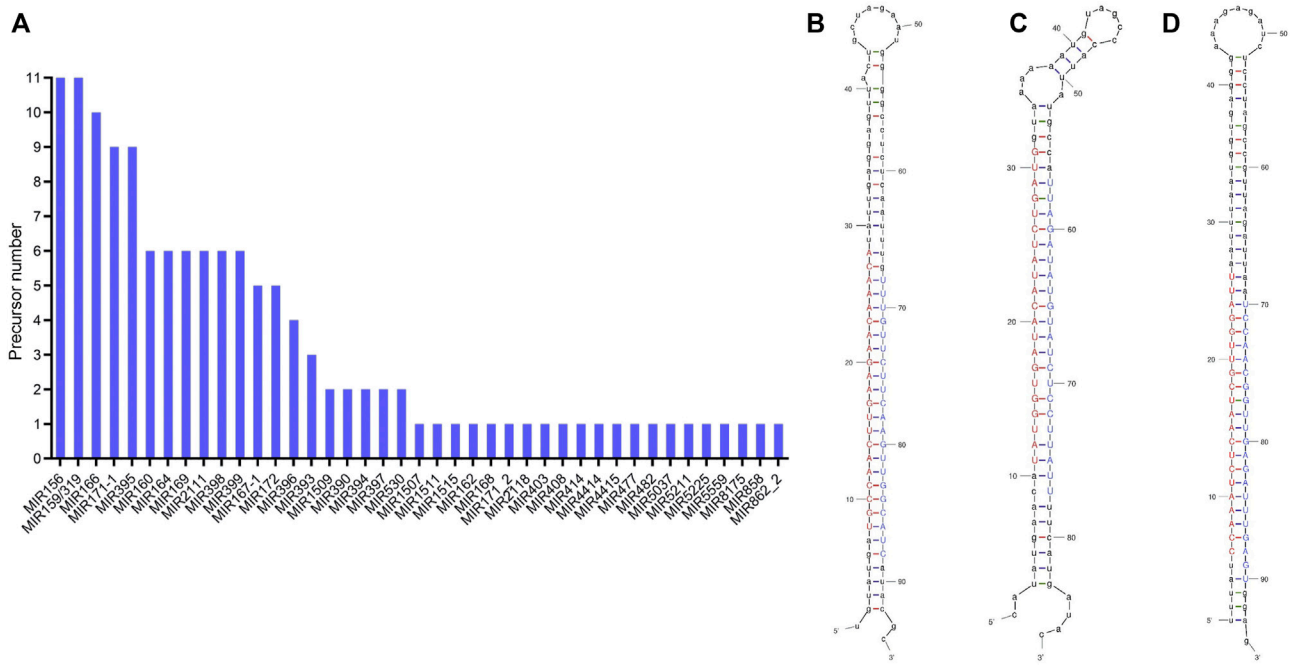


FIGURE 3 | Statistics on the number of precursors of the miRNA families **(A)** and the secondary structure of the miRNA precursors of ana-miRN4 **(B)**, ana-miRN9 **(C)**, and ana-miRN33 **(D)**. Mature sequences of miRNAs are shown in uppercase, with red representing 5p and blue representing 3p. The structure diagrams were generated by using the UNAFold Web Server.

(**Figure 3A**). MIR159 and MIR156, with up to 11 precursor loci, are the largest miRNA families, followed by MIR166 (10), MIR395 (9), and MIR171-1 (9). Twenty miRNA families, such as MIR162, MIR394, MIR477, MIR862-2, MIR1507, MIR1515, and MIR5559, had only one predicted precursor.

In the current study, all predicted miRNAs that have no homologs (with less than four mismatches) in miRBase were classified as species-specific miRNAs. According to the criteria for distinguishing the high- and low-confidence miRNA annotations (Kozomara and Griffiths-Jones, 2014), a total of 68 nonconserved miRNAs and 39 corresponding stem-loop precursors were identified with high confidence (**Figures 3B–D**; **Table 2**; **Supplementary Tables S3**).

Target Identification of miRNAs from *Ammopiptanthus nanus* Using Bioinformatics Analysis and Degradome Sequencing

To dissect the function of the predicted miRNAs in *A. nanus*, psRNAtarget, a miRNA target predicting tool, was used to identify miRNA target genes, and a total of 15,005 miRNA target pairs were predicted. Among them, 159 conserved miRNAs formed 10,159 miRNA-target pairs (**Supplementary Table S4**), and 64 nonconserved miRNAs formed 4,846 miRNA-target pairs (**Supplementary Table S5**).

TABLE 2 | Species-specific miRNAs predicted from *A. nanus*.

ID	miRNA name	Sequence	Length (nt)	Total reads	ΔG (kcal/mol)
N1	ana-miRN1-5p	AAGCTCTTGCTAGGTTGATTG	21	12	−70.20
	ana-miRN1-3p	ATTGCCTACAGTTAGATCCTTA	22	281	
N2	ana-miRN2-5p	AGGATATTGCTGGGTTGATTG	21	57	−76.80
	ana-miRN2-3p	ATTACCTATAGTTTTATCCT	20	3	
N3	ana-miRN3-5p	TACTAATTTGGTCCTCAAGAA	21	11	−66.80
	ana-miRN3-3p	TCTTGAGGATCAAATTTGTATT	22	34	
N4	ana-miRN4-5p	TGCCAACTTGAAGAACAAACA	21	150	−56.60
	ana-miRN4-3p	TTTGTTCTTCAAGTTGGCATC	21	40,865	
N5	ana-miRN5-5p	TGTTTGGGAGAAGTGGGAAAG	21	243	−55.80
	ana-miRN5-3p	TTTTCCGCTTCTTTCAAACAG	21	1791	
N6	ana-miRN6-5p	TCATTTAATCCCTATACGTAG	21	25	−38.80
	ana-miRN6-3p	TAACGGGAGGGACCAAATGGG	21	101	
N7	ana-miRN7-5p	TATGGTGATACATATCTGATG	21	709	−36.10
	ana-miRN7-3p	TTAGATATGTATCTCCTTATT	21	772	
N8	ana-miRN8-5p	TCCGTGGCCAAGATGACGAAGAA	23	1437	−40.70
	ana-miRN8-3p	CTCTTCATCTTCCTCGGATTC	21	43	
N9	ana-miRN9-5p	TTTAGGATAGGCTTTGATACC	21	630	−54.30
	ana-miRN9-3p	TATCAAAGCCTATCCGAGATC	21	16	
N10	ana-miRN10-5p	ATTCTTCATTTAGTCTCTATA	21	334	−40.70
	ana-miRN10-3p	TAGGGACTAAATGAAGAATCT	21	10	
N11	ana-miRN11-5p	CATTAACGAAATTCCAAAAAC	21	69	−93.40
	ana-miRN11-3p	TTTTGGAATTTCGTTAATATC	21	393	
N12	ana-miRN12-5p	TCCGCCGAGACTCGTCTGCA	20	172	−81.90
	ana-miRN12-3p	CAGACGGGTCATGGCAGACGA	21	21	
N13	ana-miRN13-5p	TTTCCTGTCACATTCTCTGTC	21	49	−93.70
	ana-miRN13-3p	CAGAGAATGTGACAGGAAAGTG	22	12	
N14	ana-miRN14-5p	TTGTACGATTTTGGTCCCTCA	21	12	−85.00
	ana-miRN14-3p	TGAGGGACCAAAATTACACAAT	22	43	
N15	ana-miRN15-5p	TGGGATTACAGGGGTTCTCTT	21	10	−82.30
	ana-miRN15-3p	AGAGAACCCCTGTAATCCCAG	21	36	
N16	ana-miRN16-5p	TTACGGGACATTCTTATGTGGC	22	14	−43.40
	ana-miRN16-3p	CACATAGGAATGACACGTAAGCT	23	20	
N17	ana-miRN17-5p	TTTTGTTTTTGGTCCCTGTCA	21	18	−54.00
	ana-miRN17-3p	CAGGGACCAAAAACAAAATTTT	22	10	
N18	ana-miRN18-5p	CCCACCCACACAGTTGACCTAA	22	1057	−37.10
	ana-miRN18-3p	AGGTCAAATTTGTGGGGGGTG	21	67	
N19	ana-miRN19-5p	GAGGGACTAAATGAAGAATTT	21	105	−123.00
	ana-miRN19-3p	ATTCTTCATTTAGTCTCTATA	21	347	
N20	ana-miRN20-5p	TAAGGACTAAATGAAGAATTT	21	32	−63.90
	ana-miRN20-3p	ATTCTTCATTTAGTCTCTATA	21	347	
N21	ana-miRN21-5p	CATTAACGAAATTCCAAAAAC	21	67	−78.70
	ana-miRN21-3p	TTTTGGAATTTCGTTAATATC	21	365	
N22	ana-miRN22-5p	CGTCCTGGTGAAACGCGCCACT	22	307	−70.10
	ana-miRN22-3p	TGGCGCGTTCCACCAGGATGC	21	11	
N23	ana-miRN23-5p	TCCACTGAACGTAATTAACCA	21	227	−57.70
	ana-miRN23-3p	ATTAATTACGTTCAGTGGATG	21	13	
N24	ana-miRN24-5p	TTGATCAACTCCTCCACCGTGA	22	25	−34.90
	ana-miRN24-3p	CACGGTGAGAGTTGGACTTTGC	22	200	
N25	ana-miRN25-5p	TCCACTGAACGTAATTAACCA	21	210	−90.90
	ana-miRN25-3p	ATTAATTACGTTCAGTGGATG	21	15	
N26	ana-miRN26-5p	TACCAATTTGATCCTCAAGAAT	22	10	−68.90
	ana-miRN26-3p	TTTGAGGATCAAATTGGTATT	21	194	
N27	ana-miRN27-5p	TGGCGCGACGCAGTGGAAGGC	21	72	−114.50
	ana-miRN27-3p	TCCACTGCGTCGCGCCACGTG	21	67	
N28	ana-miRN28-5p	TGGCGCGACGCAGTGGAAGGC	21	43	−83.10
	ana-miRN28-3p	TCCACTGCGTCGCGCCACGTG	21	67	
N29	ana-miRN29-5p	TCACTCATCGTTGGATCAATC	21	75	−32.50
	ana-miRN29-3p	TTGATCTAACGGTGCGTGAAT	21	18	
N30	ana-miRN30-5p	TAGCACATCATGTCCCAATCA	21	12	−47.80
	ana-miRN30-3p	ATTGGGATATAATGTGATACAT	22	55	
N31	ana-miRN31-5p	CATTAACCACGATTTTGAACG	21	37	−91.40
	ana-miRN31-3p	TTCAAAATCGTGGTTAATGAA	21	27	
N32	ana-miRN32-5p	TCCGGATCCTCTAACTTTAGG	21	49	−55.70
	ana-miRN32-3p	TAAAGTTAGAGGATCCAAATC	21	20	

(Continued on following page)

TABLE 2 | (*Continued*) Species-specific miRNAs predicted from *A. nanus*.

ID	miRNA name	Sequence	Length (nt)	Total reads	ΔG (kcal/mol)
N33	ana-miRN33-5p	CCAAATCTCAATCGTTGGATT	21	39	−58.90
	ana-miRN33-3p	TCCAACGGTTGAGATTTGAGT	21	34	
N34	ana-miRN34-5p	CATTAACCACGATTTTGAACG	21	37	−72.40
	ana-miRN34-3p	TTCAAAATCGTGGTTAATGAA	21	21	
N35	ana-miRN35-5p	TATCAATTTGGTCCTCAAGAA	21	32	−56.07
	ana-miRN35-3p	TCTTGAGGATCAAATTTGTATT	22	34	
N36	ana-miRN36-5p	TTTGGATCCTCTAACTTTAGG	21	15	−56.58
	ana-miRN36-3p	TAAAGTTAGAGGATCCGGATC	21	26	
N37	ana-miRN37-5p	TTTGGATCCTCTAACTTTAGG	21	15	−50.50
	ana-miRN37-3p	TAAAGTTAGAGGATCCGGATC	21	26	
N38	ana-miRN38-5p	TATCAATTTGGTCCTCAAGAA	21	32	−50.10
	ana-miRN38-3p	TTGAGGACCAAATTAATATTTT	22	13	
N39	ana-miRN39-5p	CGGAAAATTGTTGCAGTTAAGC	22	41	−83.10
	ana-miRN39-3p	TTAACTGCAACAATTTGTCCAT	22	2655	

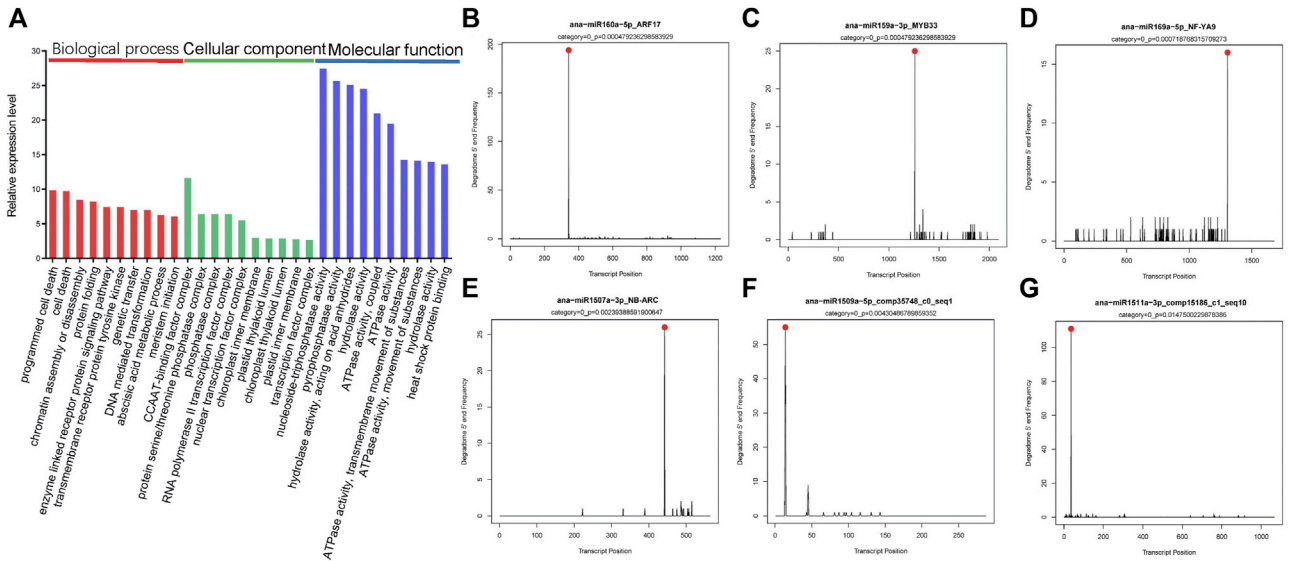


FIGURE 4 | GO analysis and T-plots of miRNA targets in *A. nanus*. GO analysis of the miRNA targets **(A)**. Red columns represent BP, green columns represent CC, and blue columns represent MF. T-plots of miR160-*ARF* **(B)**, miR159-*MYB* **(C)**, miR169-*NF-YA* **(D)**, miR1507-*NB-ARC* **(E)**, miR1509-*comp35748_c0_seq1* **(F)**, and miR1511-*comp15186_c1_seq10* **(G)** pairs identified by degradome sequencing. Cleavage sites on target genes are indicated by red dots.

Through Gene Ontology (GO) functional classification analysis, the top 10 of the biological process (BP), cellular component (CC), and molecular function (MF) categories were obtained (**Figure 4A**). Among these groups, the term regulation of programmed cell death (GO: 0012501), cell death (GO: 0008219), CCAAT-binding factor complex (GO: 0016602), protein serine/threonine phosphatase complex (GO: 0008287), nucleoside triphosphatase activity (GO: 0017111), and pyrophosphatase activity (GO: 0016462) was the top two abundant terms in BP, CC, and MF categories. These results indicated miRNAs might be involved in pathways such as programmed cell death, chromatin assembly or disassembly, signal transduction, enzyme activity, transcription, and secondary metabolism in *A. nanus* by negatively regulating their target genes.

To experimentally identify the target genes of the predicted miRNAs in *A. nanus* in batch, degradome sequencing was conducted. High-throughput sequencing produced 11,048,980 reads representing the 5' ends of uncapped, poly-adenylated RNAs. The total number of signatures matching the transcriptome sequences was 10,151,291 (91.88%), and the number of distinct sequences matching *A. nanus* transcriptome sequences was 2,440,623. CleaveLand pipeline was used to identify the cleaved miRNA targets in the *A. nanus* transcriptome. A total of 1,808 nonredundant putative splicing sites with *p*-values ≤1 were identified confidently

TABLE 3 | Noncanonical targets of the predicted miRNAs identified by degradome sequencing in *A. nanus*.

miRNA	Non-canonical target	Annotation/coding potential
ana-miR1507a-3p	comp13515_c0_seq1	Pre-rRNA-processing protein TSR2, conserved region
ana-miR1507a-3p	comp15069_c7_seq8	Basic-leucine zipper (bZIP) transcription factor family protein
ana-miR1507a-3p	comp19088_c0_seq1	TIM23-2 | translocase inner membrane subunit 23–2
ana-miR1509a-5p	comp15920_c0_seq10	hydrolases; protein serine/threonine phosphatases (PP)
ana-miR1509a-5p	comp15920_c0_seq1	hydrolases; protein serine/threonine phosphatases
ana-miR1511a-3p	comp15499_c1_seq10	Bromo-adjacent homology (BAH) domain-containing protein
ana-miR1511a-3p	comp15499_c1_seq2	Bromo-adjacent homology (BAH) domain-containing protein
ana-miR1511a-3p	comp15499_c1_seq9	Bromo-adjacent homology (BAH) domain-containing protein
ana-miR1511a-3p	comp15745_c1_seq1	GAUT6 | galacturonosyltransferase 6
ana-miR1511a-3p	comp19214_c0_seq1	ARM repeat superfamily protein
ana-miR1511a-3p	comp6070_c1_seq1	SWIB complex BAF60b domain-containing protein
ana-miR156d-3p	comp13662_c0_seq4	GTE8 | global transcription factor group E8
ana-miR156i-3p	comp15773_c2_seq28	DNA glycosylase superfamily protein (DGS)
ana-miR156j-5p	comp11270_c0_seq3	BPM2 | BTB-POZ and MATH domain 2
ana-miR156j-5p	comp16323_c0_seq1	*Glycine* cleavage T-protein family
ana-miR159h-3p	comp4558_c0_seq1	Mitochondrial transcription termination factor family protein
ana-miR159i-3p	comp14691_c0_seq1	Ypt/Rab-GAP domain of gyp1p superfamily protein (Gyp1p)
ana-miR171-1b-5p	comp15661_c0_seq18	Phox-associated domain, Phox-like, Sorting nexin, C- terminal (PX)
ana-miR171-1f-3p	comp14326_c1_seq1	Nucleotide/sugar transporter family protein
ana-miR396c-5p	comp11385_c0_seq1	LACS1 | AMP-dependent synthetase and ligase family protein
ana-miR396c-5p	comp19381_c0_seq1	BSD domain-containing protein
ana-miR1509a-5p	comp15155_c0_seq14	Noncoding
ana-miR1509a-5p	comp35748_c0_seq1	Noncoding
ana-miR1509b-5p	comp14487_c0_seq2	Noncoding
ana-miR1509b-5p	comp15155_c0_seq6	Noncoding
ana-miR1511a-3p	comp15186_c1_seq10	coding
ana-miR1511a-3p	comp15876_c0_seq40	Noncoding
ana-miR1511a-3p	comp390_c0_seq1	Noncoding
ana-miR156d-5p	comp12590_c1_seq4	coding
ana-miR156h-5p	comp12590_c1_seq2	coding
ana-miR164a-3p	comp14847_c0_seq4	Noncoding
ana-miR164c-3p	comp21396_c0_seq1	Noncoding
ana-miR396b-5p	comp10567_c1_seq1	Noncoding
ana-miR396b-5p	comp10567_c1_seq2	Noncoding

(categories ≤ 4) (**Supplementary Table S6**). Among these predicted targets, 81 were classified in category 0, and they were the most credible targets. Of the 81 targets, 47 were known targets that were reported in other plant species previously, and 34 were noncanonical targets that were identified for the first time. Most of the 34 targets (21/34) were protein-coding genes, including *comp13662_c0_seq4* and *comp15069_c7_seq8*, which encode a global transcription factor group E8 and a basic-leucine zipper (bZIP) transcription factor family protein, respectively. Of the other 13 noncanonical targets, ten were predicted to be noncoding transcripts by CPC2.0, a coding potential calculator (Kang et al., 2017). (**Figures 4B–G**; **Table 3**).

Identification of the Cold Stress-Responsive miRNAs in *Ammopiptanthus nanus*

We analyzed miRNA expression by comparing the abundance of each miRNA in the group CK and CA base on the high-throughput miRNA sequencing (miRNA-seq) (**Figure 5A**), and 39 cold-responsive miRNAs were identified, including 29 conserved miRNAs and ten nonconserved. Twenty-one conserved miRNAs were upregulated under cold acclimation, including ana-miR156a-5p, ana-miR159a-3p, ana-miR162a-3p, ana-miR166a-3p, ana-miR168a-5p, ana-miR390a-5p, ana-miR395a-3p, ana-miR399a-3p, ana-miR1507a-3p, ana-miR1515a-5p, ana-miR2118a-3p, ana-miR5211a-5p, and ana-miR5225a-3p, and eight conserved miRNAs were downregulated, that is, ana-miR1507a-5p, ana-miR156g-3p, ana-miR160a-3p, ana-miR171-1g-3p, ana-miR396b-3p, ana-miR398b-3p, ana-miR408a-5p, and ana-miR4415a-3p. Among the differentially expressed nonconserved miRNAs, eight were upregulated under cold acclimation, that is, ana-miRN8-5p, ana-miRN9-5p, ana-miRN18-5p, ana-miRN29-5p, ana-miRN6-3p, ana-miRN8-3p, ana-miRN15-3p, and ana-miRN39-3p and two were downregulated, that is, ana-miRN7-5p and ana-miRN19-3p.

qRT-PCR experiments were conducted to validate the results of high-throughput sequencing. Twelve conserved miRNAs, including ana-miR159a-3p, ana-miR168a-5p, ana-miR390a-5p, ana-miR399a-3p, ana-miR1507a-3p, ana-miR1515a-5p, ana-miR2118a-3p, ana-miR5211a-5p, ana-miR5225a-3p, ana-miR398b-3p, ana-miR408a-3p, and ana-miR4415a-3p, and five nonconserved miRNAs, that is, ana-

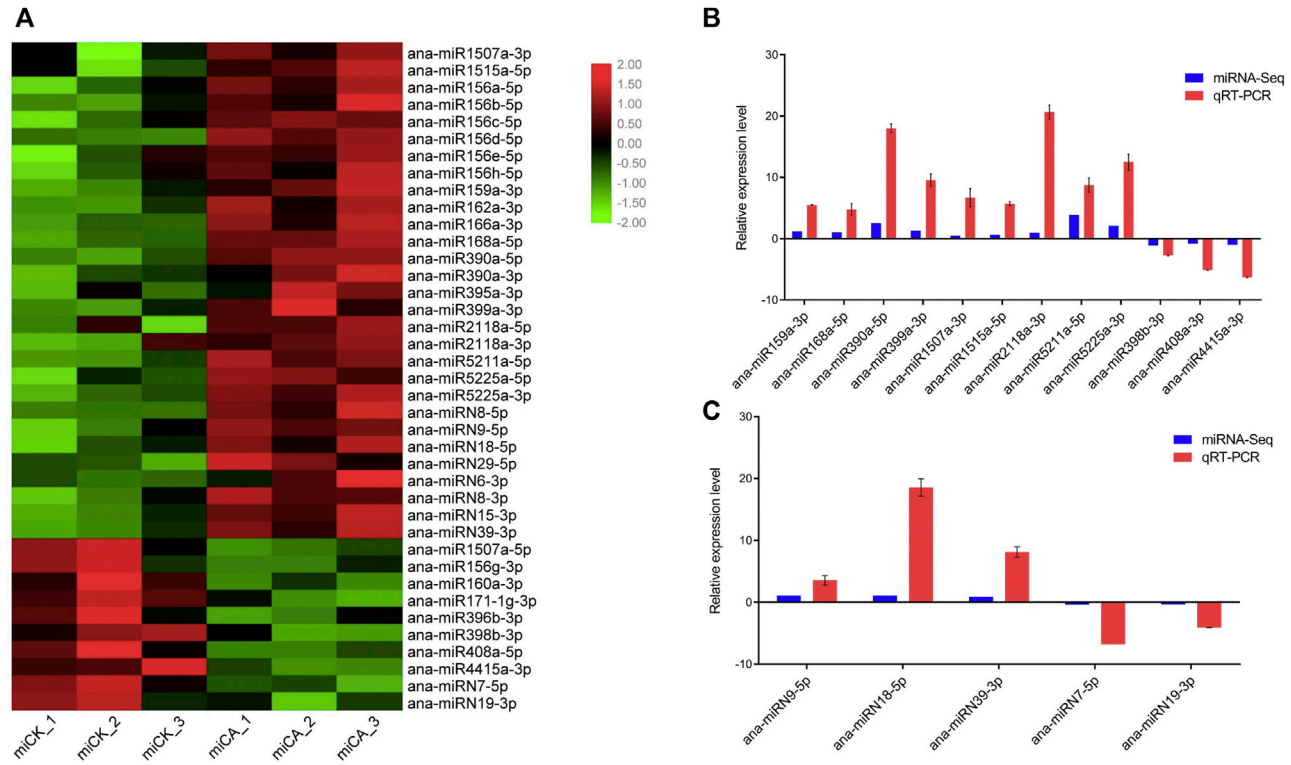


FIGURE 5 | Expression of cold-responsive miRNAs in *A. nanus*. **(A)** Expression of cold-responsive conserved miRNAs and non-conserved miRNAs. **(B)** qRT-PCR validation of the cold-responsive conserved miRNAs. **(C)** qRT-PCR validation of the cold-responsive nonconserved miRNAs. The *Y*-axis shows the relative expression levels of conserved miRNAs and non-conserved miRNAs, with U6 as an internal reference, averaged over three technical replicates. Values are expressed as means ± SD (n = 3).

Element	Function	ana-miR159a-3p	ana-miR168a-5p	ana-miR399a-3p	ana-miR5225a-3p	ana-miR5211a-5p	ana-miR398b-3p	ana-miR408a-3p	ana-miR4415a-3p	ana-miR390a-5p	ana-miR1507a-3p	ana-miR1515a-5p	ana-miR2118a-3p
MBS	Drought		1	2	1		1		1			1	
LTR	Low temperature			1		1				1	2		
TC-rich repeat	Defense and stress		1	1						1	1		
ARE	Stress	5	3	4		3		5	1	2	2	4	2
TGA element	Auxin				1			1		1	2	1	
AuxRR core	Auxin								1				
GARE motif	GA						1		1				
P Box	GA									1	2		
ABRE	ABA		8				1			1		2	3
ERE	Ethylene	3			1	1	1	2	2				3
CGTCA motif	MeJA	1		1	2	2				2	2		
TGACG motif	MeJA	1		1	2	2				2	2		
TCA element	SA	1		1						1	2	1	
W-box	Defense				1					1	1		1

FIGURE 6 | Abiotic stress–responsive *cis*-elements predicted from the promoters of cold-responsive miRNA genes.

miRN9-5p, ana-miRN18-5p, and ana-miRN39-3p, ana-miRN7-5p, and ana-miRN19-3p were selected for qRT-PCR analysis (**Figures 5B,C**). The change patterns of *A. nanus* miRNAs revealed by qRT-PCR analysis were consistent with those calculated based on the miRNA-seq results.

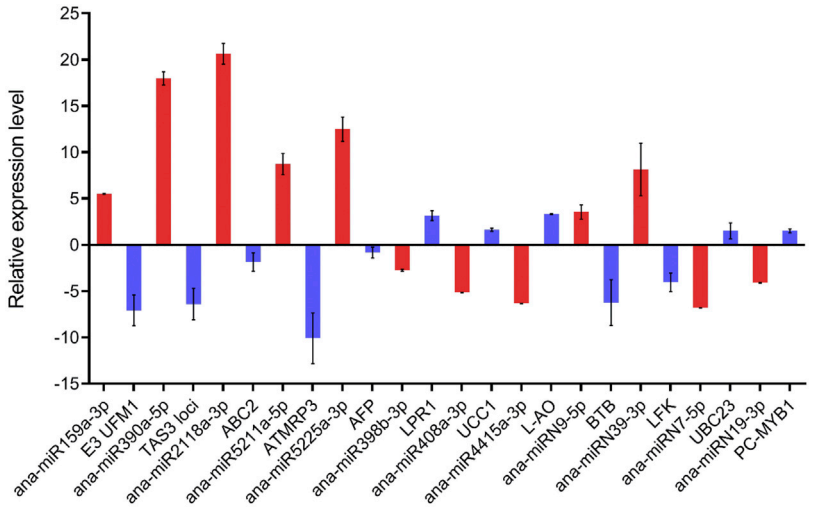


FIGURE 7 | qRT-PCR analysis of the targets of the 12 selected cold-responsive miRNAs. Red columns represent miRNAs and blue columns represent target genes. The *Y*-axis shows the relative expression with U6 or actin as the internal reference, and data are averaged over three technical replicates. Values are expressed as means ± SD.

Cis-Elements Involved in Abiotic Stress Responses in the Promoters of the Cold-Responsive miRNA Genes

Cis-acting elements have been reported to be associated with plant responses to abiotic stresses, including cold stress (Zhang et al., 2005). To identify the *cis*-acting elements associated with the expression patterns of the cold-responsive miRNAs, we analyzed the 1000-bp upstream promoter sequence of the 12 cold-responsive miRNAs, which represented 12 cold-responsive miRNA families with higher abundance, by using a plant promoter prediction software, PlantCARE (http://intra.psb.ugent.be: 8080/PlantCARE For) (Lescot et al., 2002), and consequently, several stress response- and hormone response–related *cis*-acting elements, including the low-temperature responsive elements (LTRs), were found.

Among the various *cis*-acting elements predicted from the promoters of the 12 cold-responsive miRNAs (**Figure 6**), the AU-rich element (ARE) was the most frequently detected *cis*-acting element, which was found in promoters of ten miRNAs, followed by ethylene response factor (ERE), appeared in promoters of seven miRNA genes. ARE has been shown to respond to hypoxic, low-temperature, and dehydration stress (Jang, 2016). Ethylene is an important gaseous phytohormone, playing a crucial role in plant growth and development and the response to biotic or abiotic stresses, and genes with ERE located in the promoter may be involved in ethylene response in plants (Rajesh et al., 2017). We also found several other *cis*-acting elements, including MBS, LTR, TC-rich repeat, TGA element, AuxRR core, GARE motif, P Box, ABRE, CGTCA motif, TGACG motif, TCA element, and W-box. ABREs were found in the promoter regions of most ABA-responsive genes and were reported to respond to abiotic stress in *A. thaliana* (Nagatoshi et al., 2018; Chang et al., 2019), and MBS is the MYB-binding site involved in drought inducibility (Lv et al., 2020).

Expression Analysis of the Targets of the Cold-Responsive miRNAs in *Ammopiptanthus nanus*

To investigate whether the expression of the predicted target negatively correlated with the expression of the corresponding miRNAs, we selected 12 predicted targets and analyzed their expression in cold-acclimated seedlings by qRT-PCR (**Figure 7**). After cold acclimation, seven miRNAs, including ana-miR159a-3p, ana-miR390a-5p, ana-miR2118a-3p, ana-miR5211a-5p, ana-miR5225a-3p, ana-miRN9-5p, and ana-miRN39-3p were upregulated, and the corresponding targets *E3 UFM1-protein ligase-like protein* (*E3 UFM1*), *TAS3*, *ABC2 homolog 13* (*ABC2*), *ATMRP3*, *Ankyrin repeat family protein* (*AFP*), *BTB-POZ and MATH domain 4* (*BTB*), and *LFK L-fructokinase* (*LFK*) were downregulated. The expression levels of ana-miR398b-3p, ana-miR408a-3p, ana-miR4415a-3p, ana-miRN7-5p, and ana-miRN19-3p were downregulated, and the corresponding targets *LPR1 Cupredoxin superfamily protein* (*LPR1*), *UCC1 uclacyanin 1* (*UCC1*), *L-AO*, *UBC23 ubiquitin-conjugating enzyme 23* (*UBC23*), *PC-MYB1 homeodomain-like protein* (*PC-MYB1*), *auxin response factor 3* (*ARF3*), *TIR1-NBS-LRR* (*TIR1*), *TIR2-NBS-LRR* (*TIR2*), *cold shock domain protein 1* (*CSDP1*), and *DICER-LIKE2* (*DCL2*) were upregulated. The expression of these 12 selected targets displayed the opposite pattern to that of the corresponding miRNAs after cold acclimation.

miR4415 is a Legume-Specific miRNA

To understand the roles of miRNAs in cold acclimation, miR4415 was selected for further functional study. In the miRBase database (miRBase 21), there are two mature sequences of miR4415, namely, gma-miR4415a and gma-miR4415b. The mature sequences of miR4415 and the corresponding precursor sequences from *A. mongolicus* and *Astragalus membranaceus*

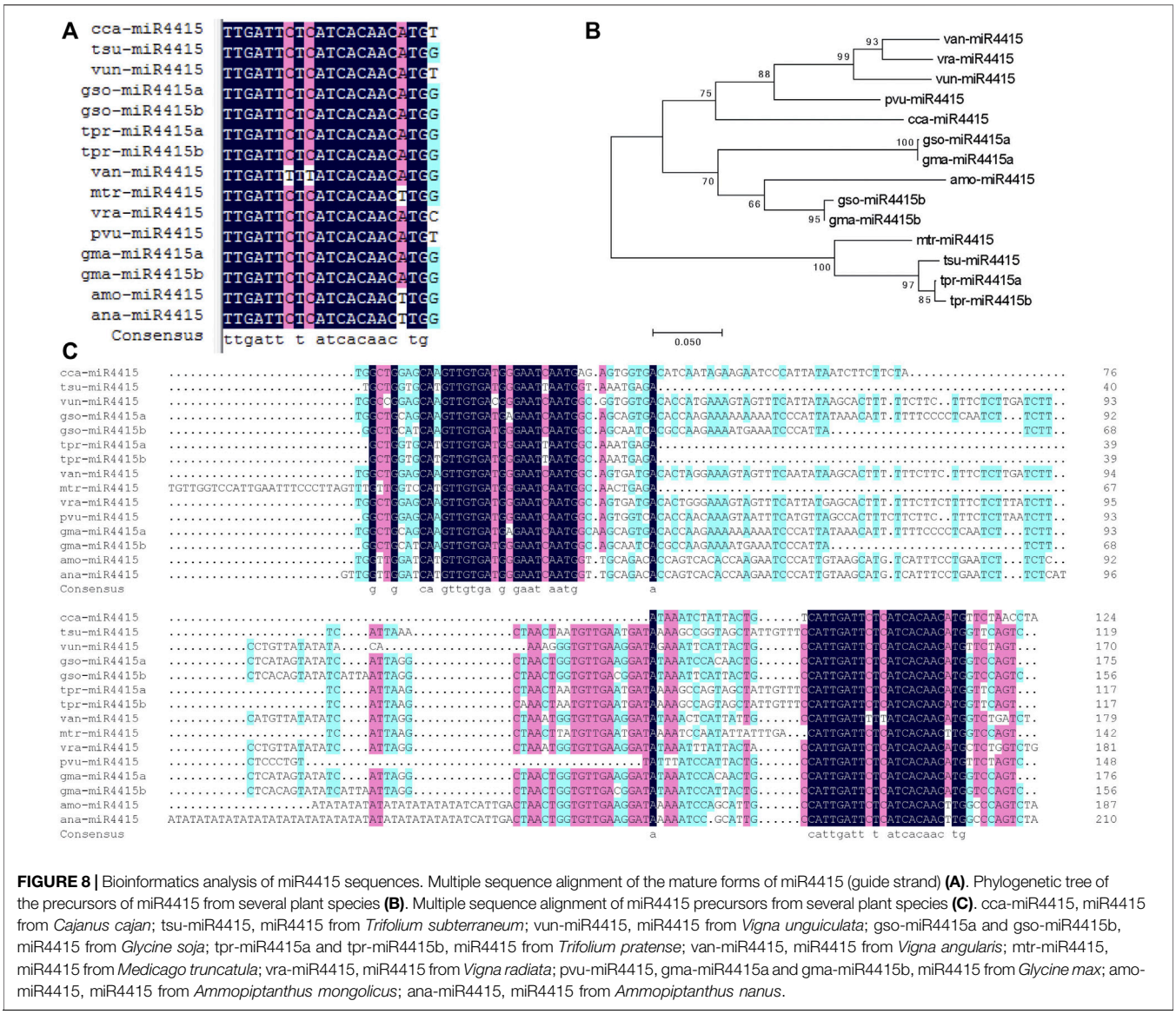


FIGURE 8 | Bioinformatics analysis of miR4415 sequences. Multiple sequence alignment of the mature forms of miR4415 (guide strand) **(A)**. Phylogenetic tree of the precursors of miR4415 from several plant species **(B)**. Multiple sequence alignment of miR4415 precursors from several plant species **(C)**. cca-miR4415, miR4415 from *Cajanus cajan*; tsu-miR4415, miR4415 from *Trifolium subterraneum*; vun-miR4415, miR4415 from *Vigna unguiculata*; gso-miR4415a and gso-miR4415b, miR4415 from *Glycine soja*; tpr-miR4415a and tpr-miR4415b, miR4415 from *Trifolium pratense*; van-miR4415, miR4415 from *Vigna angularis*; mtr-miR4415, miR4415 from *Medicago truncatula*; vra-miR4415, miR4415 from *Vigna radiata*; pvu-miR4415, gma-miR4415a and gma-miR4415b, miR4415 from *Glycine max*; amo-miR4415, miR4415 from *Ammopiptanthus mongolicus*; ana-miR4415, miR4415 from *Ammopiptanthus nanus*.

have been identified by small RNA sequencing in previous studies (Gao et al., 2018; Abla et al., 2019). miR4415 is absent in both *A. thaliana* and rice, and thus we speculated that miR4415 might be a lineage-specific miRNA. We used the BLAST program to compare the mature sequences of the abovementioned miR4415 with the genomic sequences of all plant species in Phytozome v12 to detect the mature sequences of miR4415 in other plant species. Consequently, 11 additional mature sequences of miR4415 (guide strand) that were not included in miRBase 21 were identified from nine plant species, that is, *Glycine soja*, *Trifolium pratense*, *Cicer arietinum*, *Cajanus cajan*, *Vigna unguiculata*, *Vigna angularis*, *Medicago truncatula*, *Trifolium subterraneum*, and *Vigna radiata*. It is worth noting that all miR4415 (guide strand) are located at 3p of the stem-loop precursors.

Multiple sequence alignment analysis indicated that the mature forms of miR4415 from different species differed significantly at nucleotide 21, and minor differences were also found at nucleotide 7, 9, and 18 (**Figure 8A**). Multiple sequence alignment analysis of the precursor sequences revealed that in addition to the strong conservation of miR4415-3p and miR4415-5p, some nucleotides in other sites of the precursor also showed a high degree of conservation, such as the two nucleotides before the 5′ end of miR4415-3p (**Figure 8C**). These conserved nucleotides may play important roles in the formation of hairpin structures.

Phylogenetic analysis using the precursors of the 12 miR4415 indicated that all miR4415 were clustered into three clades. The first clade contained *van-miR4415, vra-miR4415, vun-miR4415, pvu-miR4415,* and *cca-miR4415*, the second contained *gso-miR4415a, gma-miR4415a, gso-miR4415b, gma-miR4415,* and *amo-miR4415*, while *mtr-miR4415, tsu-miR4415, tpr-miR4415a,* and *tpr-miR4415b*

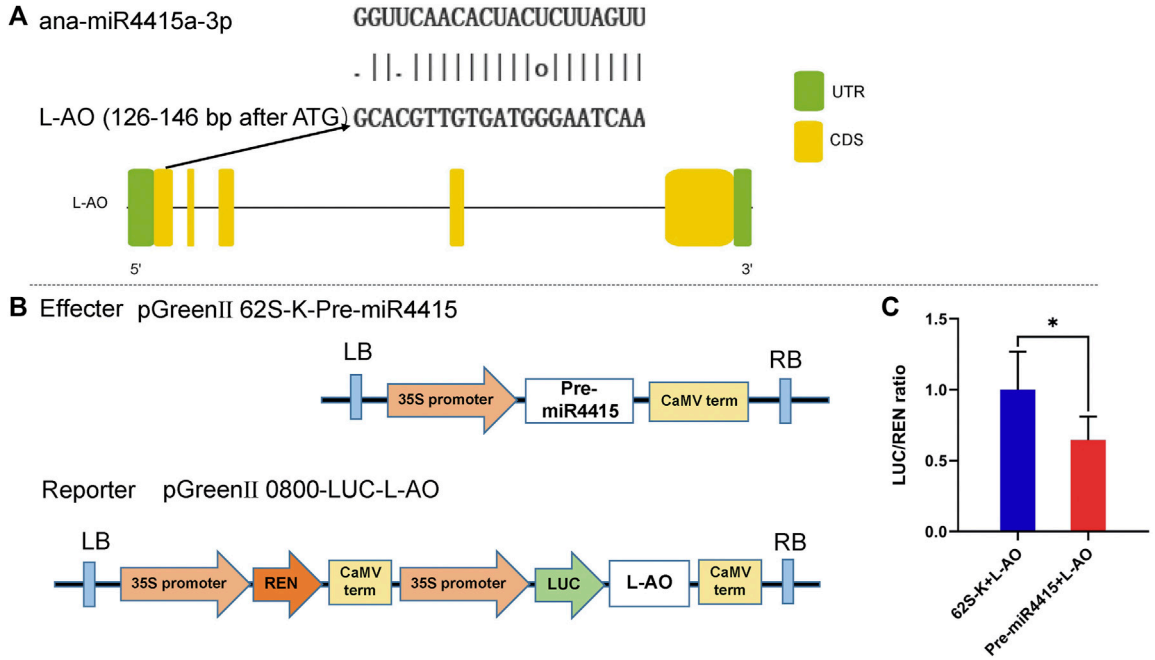


FIGURE 9 | *L-AO* gene was targeted by ana-miR4415a-3p. Binding sites of ana-miR4415a-3p on *L-AO* as predicted by psRNAtarget **(A)**. Dual-luciferase reporter vector construction for dual-luciferase reporter assay **(B)**. Ratio of LUC/REN in the 62 SK + L-AO and Pre-miR4415 + L-AO cells **(C)**. Values are expressed as means ± SD (n = 5). Student's t-test, $^*p < 0.05$.

were clustered into the third branch. *Gma-miR4415a* and *gso-miR4415a* in soybean and wild soybean were clustered into one clade, while *gma-miR4415b* and *gso-miR4415b* were clustered into another clade (**Figure 8B**). Thus, *miR4415a* and *miR4415b* were already present before the species differentiation in soybean and wild soybean.

miR4415 Targets an L-Ascorbate Oxidase

An *L-AO* gene (*EVM0030483.1*) was predicted to be the target of ana-miR4415a-3p by using the online software psRNAtarget (**Figure 9A**). Dual-luciferase reporter assay, a reliable method to detect the targeting relationship between miRNAs and their possible targets, was conducted to validate the prediction experimentally. Compared with 62 SK + L-AO, the Luciferase/Renilla (LUC/REN) ratio was significantly decreased in the Pre-miR4415 + L-AO cells (**Figure 9B**), indicating that miR4415 negatively regulated the expression of L-AO (**Figure 9C**). The opposite change pattern between ana-miR4415a-3p and *L-AO* (*EVM0030483.1*) in cold-acclimated *A. nanus* revealed by qRT-PCR analysis also supported the targeting relationship of miR4415 and *L-AO* (*EVM0030483.1*) (**Figure 7**).

Ammopiptanthus nanus L-Ascorbate Oxidase is Located in the Apoplast

The L-AO (EVM0030483.1) protein was predicted to be located in the apoplast bioinformatically. To determine the subcellular localization of *A. nanus* L-AO (EVM0030483.1) experimentally, the *L-AO* (*EVM0030483.1*) gene was fused with the *GFP* gene and transiently expressed in *Nicotiana tabacum*. The GFP signal was mainly observed in the intercellular space, validating the apoplastic localization of L-AO protein (EVM0030483.1) (**Figure 10A**).

L-AO Regulated the Contents of ASA and DHA in the Leaf Apoplast of *Ammopiptanthus nanus* Under Cold Acclimation

L-AO catalyzes the oxidation of ascorbic acid, which may affect the contents of ASA and DHA in the leaf apoplast of *A. nanus*. To assess the effect of upregulated L-AO on the concentration of ASA and DHA in the apoplast under cold acclimation, we extracted the apoplast fluid and measured the AO activity and ASA and DHA levels in the apoplast fluid of *A. nanus* before and after cold acclimation. Leaf extracts of *A. nanus* have MDH activity (4,755 u/g FW), while the MDH activity of the leaf apoplast liquid of *A. nanus* was not detected, indicating that the apoplast liquid we prepared had no cytoplasmic contamination and can be used for measurement of AO activity and the levels of ASA and DHA. The AO activity assay indicated that the AO activity in the apoplast of *A. nanus* was elevated after cold acclimation (**Figure 10B**). ASA and DHA measurements indicated a significant decrease in ASA (**Figure 10C**), a marked increase in DHA (**Figure 10D**), and a remarkable decrease in ASA/DHA ratio (**Figure 10E**) in the leaf apoplast liquid of *A. nanus* after cold acclimation.

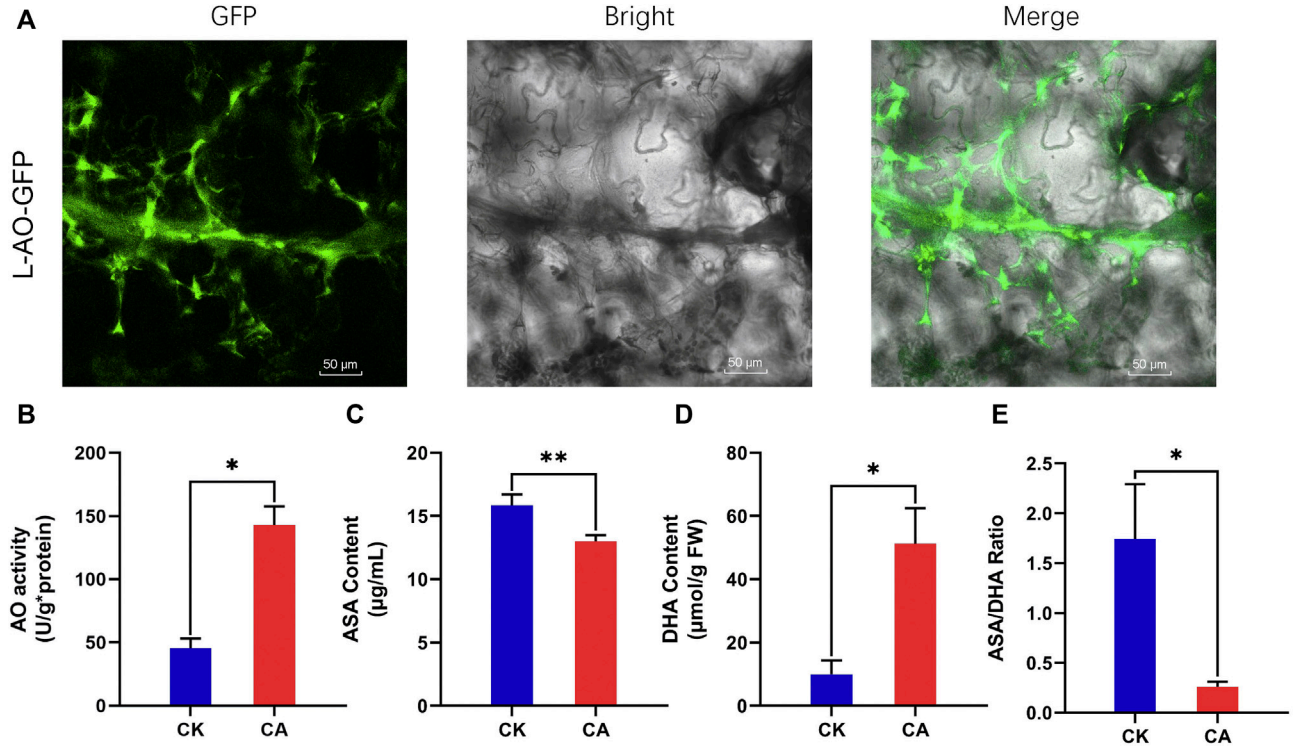


FIGURE 10 | Subcellular localization of L-AO-GFP and analysis of apoplastic redox homeostasis in leaves of *A. nanus* under cold acclimation. **(A)** Subcellular localization of L-AO-GFP. Bars = 50 μm **(B)** AO activity. **(C)** Content of ASA. **(D)** Content of DHA. **(E)** ASA/DHA ratio. Values are expressed as means ± SD. Student's t-test, $^{*}p < 0.05$.

DISCUSSION

A. nanus is a rare temperate evergreen broad-leaved plant, and it is of great significance to study its tolerance mechanism to freezing stress. As a class of endogenous regulatory RNA molecules, miRNAs not only participate in the regulation of plant growth and development but also in the response to environmental stresses (Song et al., 2019). Our study demonstrated that cold acclimation can improve the tolerance of *A. nanus* to the freezing stress, and miRNAs might play regulatory roles in cold acclimation in *A. nanus*. Analyzing the expression changes of miRNAs and their targets under cold acclimation is helpful in understanding the gene expression regulation network associated with the cold acclimation process. Our study systematically identified the miRNA and the corresponding targets that might be involved in cold acclimation and provided important data for understanding the molecular regulation mechanism of cold acclimation of *A. nanus*.

The Conserved miRNAs Predicted in *Ammopiptanthus nanus*

The completion of genome sequencing of *A. nanus* makes it possible to systematically identify the miRNA in *A. nanus*. Most of the identified miRNAs in *A. nanus* were conserved miRNAs, and the precursor sequences of these miRNAs were distributed on different assembled genome contigs. Of them, several conserved miRNA loci were clustered in genome contigs (**Supplementary Table S2**). For example, three precursors of miR2111, that is, *ana-MIR2111a*, *ana-MIR2111b*, and *ana-MIR2111c* were clustered in Contig00573, and the other three miR2111 precursors, that is, *ana-MIR2111d*, *ana-MIR2111e*, and *ana-MIR2111f* were clustered in Contig00782. There are nine precursors in the MIR395 family, and seven of them were clustered in Contig00419, including *ana-MIR395a*, *ana-MIR395b*, *ana-MIR395c*, *ana-MIR395d*, *ana-MIR395e*, *ana-MIR395f*, and *ana-MIR395g*, and the other two precursors, that is, *ana-MIR395h* and *ana-MIR395i* were distributed closely in Contig00478. Some of these clustered miRNA loci are probably generated *via* gene duplication.

Based on their conservation among plant species, the 41 conserved miRNA families in *A. nanus* can be classified into four groups. Ten miRNA families, including MIR156, MIR159, MIR160, MIR166, MIR167_1, MIR 171, MIR390, MIR395, MIR396, and MIR408, showed a high level of conservation among multiple plant species and are presented in Bryophyta, Lycopodiophyta, Coniferophyta, and Magnoliophyta, but not in Chlorophyta. This group contains the most conserved miRNAs in the plant kingdom. Group 2 comprised 12 miRNA families, that is, MIR162, MIR164, MIR168, MIR169, MIR172, MIR393, MIR394, MIR397, MIR398, MIR399, MIR2111, and MIR2118.

These miRNAs are mainly identified in Magnoliophyta and they represent classical higher plant miRNAs. miRNAs in Group 3 (MIR403, MIR858, MIR1509, MIR1511, and MIR5225) are present only in eudicotyledons; hence, they might originate before the emergence of eudicotyledons. The remaining miRNA families, including MIR1507, MIR1515, MIR4415, MIR5037, and MIR862_2, were identified only in legume plants such as *G. max* and *M. truncatula* and might be legume-specific miRNAs.

The Species-Specific miRNAs Predicted in *Ammopiptanthus nanus*

Besides the 68 nonconserved miRNA, we also identified approximately 264 miRNA candidates in *A. nanus*. Although their putative precursors can be folded into qualified hairpins (Axtell and Meyers, 2018), the reads mapped to 5′ arm or 3′ arm of the hairpins were not found or the abundance of 5p or 3p in sRNA libraries is less than ten; thus, these sRNA cannot be annotated as miRNA with high confidence. Maybe increasing the sequencing depth will result in identifying more novel miRNAs.

The sequencing frequencies of the 68 identified species-specific miRNAs varied greatly (**Supplementary Table S3**). Although the expression levels of most of the species-specific miRNAs were very low, some of them were expressed in a considerable amount. More than one-third of the species-specific miRNAs (24/64, 35.29%) were sequenced more than 100 times, and five species-specific miRNAs (ana-miRN8-5p, ana-miRN18-5p, ana-miRN4-3p, ana-miRN39-3p, and ana-miRN5-3p) have read numbers of more than 1,000 (**Supplementary Table S3**). Among all the identified species-specific miRNAs, ana-miRN4-3p was the most abundant, with a read number of more than 40,000 (**Supplementary Table S3**). Some of the species-specific miRNAs with high expression levels might have developed under environmental pressure, and thus might be involved in the adaptive evolution of *A. nanus*.

The Noncanonical Targets Predicted by Degradome Sequencing Analysis

Degradome sequencing has been proven to be an efficient method to identify plant miRNA target genes (Ma X. et al., 2015). In *A. nanus*, many miRNA-target pairs identified by degradome sequencing have been reported in other plant species, including *SQUAMOSA promoter binding protein-like (SPL)* genes as the targets of miR156, the *MYB* genes as the targets of miR159, *ARFs* as the targets of miR160, miR167, miR393, *endoribonuclease Dicer-like 1 (DCL1)* as the target of miR162, *nuclear factor Y genes (NF-Y)* as the target of miR169, *Integrase-type DNA-binding superfamily protein (AP2)* as the target of miR172, *ATP sulfurylase 1 (APS1)* as the target of miR395, *Laccase/Diphenol oxidase family protein (LAC)* as the target of miR397, *phosphate 2* as the target of miR399, *Argonaute family protein (AGO2)* as the target of miR403, *NB-ARC domain-containing disease resistance protein* as the targets of miR1507 and miR2118, and *autoinhibited Ca^{2+}-ATPase* (*ACA*) as the target of miR5225. These results demonstrated the effectiveness of our degradome sequencing analysis.

Moreover, our degradome sequencing analysis also identified some noncanonical target genes of *A. nanus* miRNAs, which were not reported in other plant species. Among the 34 highly reliable target genes, which were classified in category 0, 21 had homologs in *A. thaliana* protein, and 13 cannot be annotated by the *A. thaliana* protein dataset, which were likely to be noncoding RNAs. Most of the noncanonical targets predicted in our study have no homology with the targets of the conserved miRNAs; thus, further investigations need to be conducted to understand the biological roles of these novel miRNA-target regulatory motifs.

The miRNAs Were Involved in the Cold Acclimation in *Ammopiptanthus nanus*

A total of 39 cold-responsive miRNAs were identified in *A. nanus*, including 29 conserved miRNAs and ten nonconserved. The 29 conserved miRNAs belonged to 20 miRNA families such as MIR159, MIR168, MIR399, MIR5211, MIR5225, MIR398, MIR408, and MIR4415. Twelve conserved miRNAs that represent most of these cold-responsive conserved miRNAs and five nonconserved miRNAs with higher abundance were further validated by qRT-PCR. These cold-responsive miRNAs might participate in cold acclimation by regulating multiple biological processes, including ROS homeostasis (miR398, miR408, and miR4415), development (miR159 and miR390), defense (miR2118 and miR1507), calcium signaling (miR5225 and miR5211), and small RNA (sRNA)–mediated gene silencing (miR168 and miR1515) (**Figure 11**).

Calcium is a ubiquitous signal and an essential macronutrient in plants, and calcium signaling plays a vital role in environmental stress signal transduction. Calcium ATPases catalyze the hydrolysis of ATP coupled with the translocation of calcium from the cytosol into the endoplasmic reticulum or out of the cell and are essential components of stress signaling. In *A. nanus*, miR5225 and miR5211 were predicted to target genes encoding calcium-transporting ATPases, and both were upregulated after cold acclimation, indicating that these miRNAs might be involved in cold acclimation by regulating the calcium signaling pathway.

Three miRNAs, that is, miR398, miR408, and miR4415 might be involved in cold acclimation by regulating their targets, *copper/zinc superoxide dismutase gene (CSD)*, *ceruloplasmin/laccase*, and *L-AO*, respectively. It is intriguing that all three targets encode multicopper oxidases (MCO). Multicopper oxidases are a class of enzymes that catalyze the oxidation of various substrates *via* the reduction of O_2 to H_2O without releasing activated oxygen species (Komori and Higuchi, 2015). Of them, ascorbate oxidase (AO, EC 1.10.3.3) is localized in the cell wall and widespread in the plant kingdom (Sanmartin et al., 2007). L-AO oxidizes AsA to DHA by monodehydroascorbate (MDHA) with the participation of O^{2-} and H_2O_2 while reducing oxygen to water, thereby regulating the redox state of the apoplast.

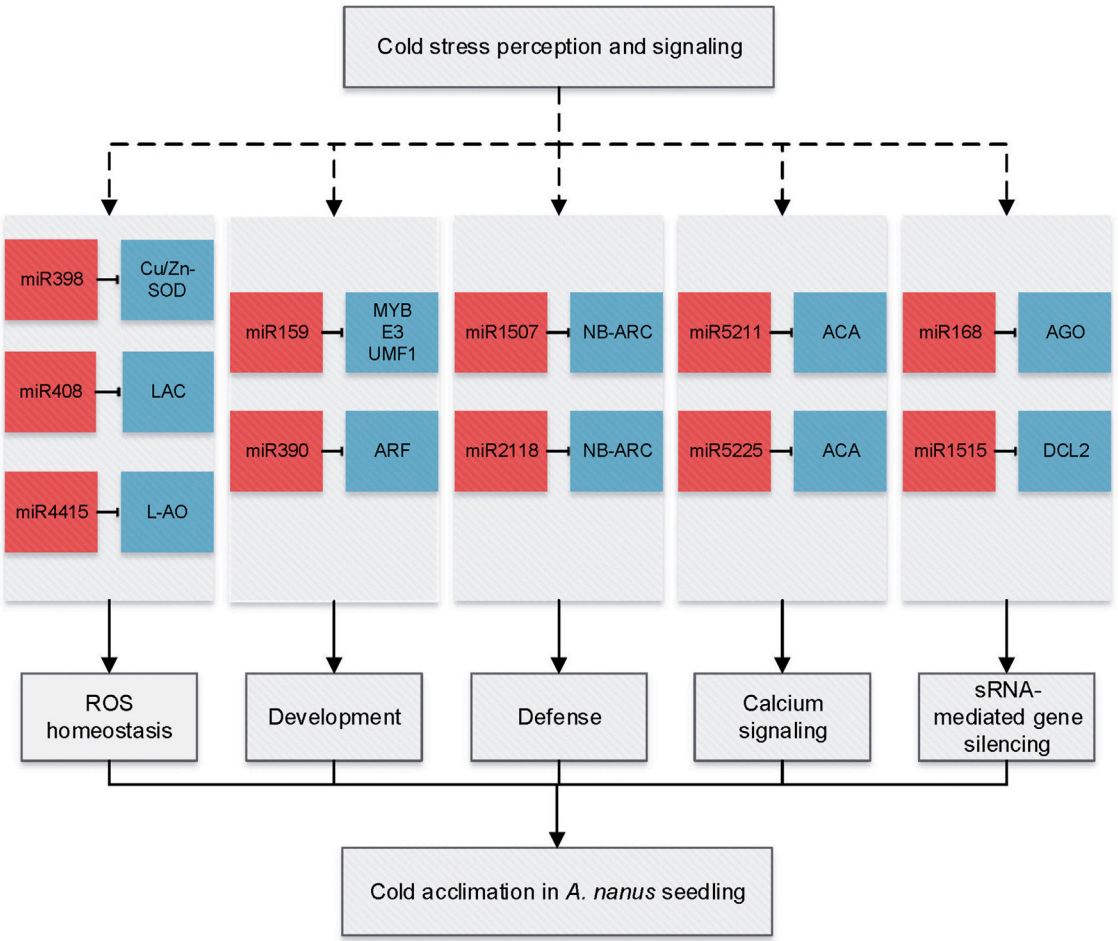


FIGURE 11 | miRNA-mediated gene regulation network in response to cold acclimation in *A. nanus* leaves. miRNAs are indicated by red squares and targets by blue squares.

Laccases are classified as benzene diol oxygen reductases (EC 1.10.3.2) and are involved in lignin degradation and detoxification of lignin-derived product wound repair, disease resistance, stress resistance, and pigment synthesis in plants (Ranocha et al., 1999). Ceruloplasmin plays a vital role in copper metabolism and is also involved in scavenging anion radicals in plants (Bielli and Calabrese, 2002). miR408 was found to be downregulated in our study, which is consistent with previous reports in wheat and grape (Wang et al., 2014) but different from the observations in *A. thaliana* (Liu et al., 2008). Overexpression of miR408 promoted drought tolerance in chickpea (Hajyzadeh et al., 2015) and cold and oxidative stresses tolerance in *A. thaliana* (Ma C. et al., 2015).

CSDs are copper- and zinc-dependent enzymes that catalyze the conversion of superoxide radicals to H_2O_2, which can subsequently be reduced to H_2O, and CSDs are proposed to play essential roles in scavenging the excessive ROS generated in plants under stressful conditions. Several reports have shown that miR398 plays an important role in plant antioxidant stress by regulating the expression of its target gene CSD (Jones Rhoades and Bartel, 2004; Sunkar and Zhu, 2004), although the expression pattern of miR398 was distinctly different. It has also been found that miR398 is upregulated in wheat but downregulated in rice when subjected to heat stress (Sailaja et al., 2014). Our data showed that miR398, miR408, and miR4415 were downregulated and their targets were upregulated in cold-acclimated *A. nanus* seedlings, and these miRNAs might contribute to cold acclimation by regulating ROS homeostasis in *A. nanus*.

miR159 has been reported to be induced by various abiotic or biotic stresses (Millar et al., 2019) and transient or mild upregulation of miR159 was observed in *A. thaliana* under cold stress (Liu et al., 2008). In *A. nanus*, we found that ana-miR159a-3p was upregulated under cold acclimation. The targets of ana-miR159a-3p in *A. nanus* encoded an MYB65 and an E3 UFM1-protein ligase–like protein (E3 UFM1). MYB TFs were involved in the regulation of various biological processes in plant growth and development in several plant species such as *A. thaliana* (Dubos et al., 2010; Chen et al., 2013) and rice (Yang et al., 2012). E3 ligase modulates plant abiotic stress responses through posttranslational modifications (Sun and Chen, 2004;

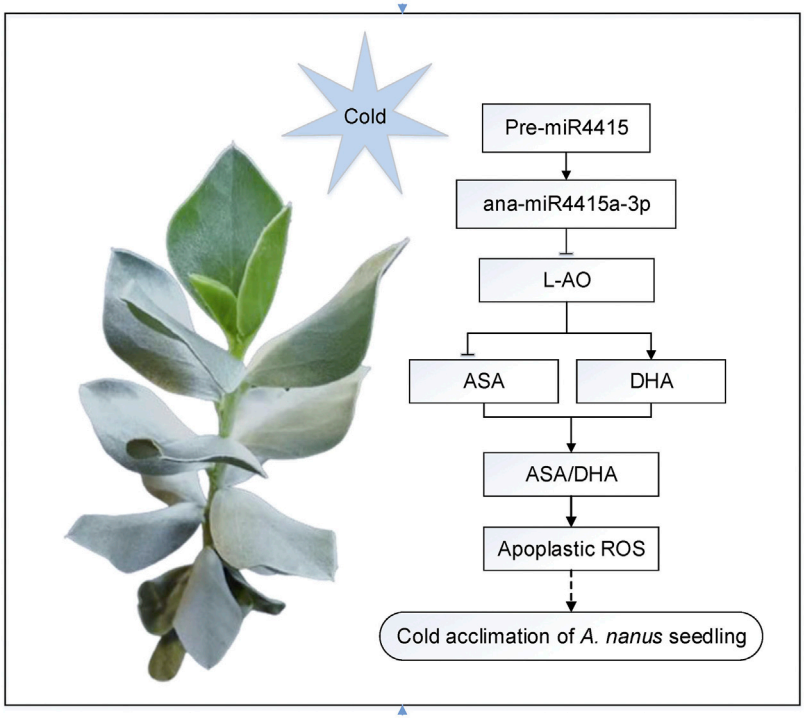


FIGURE 12 | Proposed model of the regulatory role of miR4415 in cold acclimation in *A. nanus*. Under cold acclimation, expression of the miR4415 precursor was inhibited, resulting in downregulation of the ana-miR4415a-3p and upregulation of *L-AO*. Upregulated L-AO catalyzes the oxidation of ASA to form DHA, lowering the AsA/DHA ratio, thereby changing the redox status of the apoplast. Change in apoplastic redox state regulates the tolerance to freezing stress in *A. nanus* seedlings.

Lyzenga and Stone, 2012). Ana-miR159a-3p may be involved in cold acclimation of *A. nanus* by negatively regulating *ATMYB65* and *E3 UFM1*.

MiR4415 is involved in Cold Acclimation in *Ammopiptanthus nanus* by Regulating the Apoplastic Redox Status by Targeting an L-AO

Most miRNAs are conserved, to some extent, among different species, but not all miRNAs are highly conserved. Some miRNAs exist only in a specific evolutionary branch, and these miRNAs are called lineage-specific miRNAs. The lineage-specific miRNAs were evolved/developed in a specific evolutionary branch, hence contributing to developmental novelties during evolution (Zhang et al., 2008), and might be closely associated with some unique regulatory mechanism of the plants in this evolutionary branch. Our results demonstrated that miR4415 is a legume-specific miRNA family, which has only been reported in legumes such as *G. max* (Lei et al., 2019), *G. soja* (Zeng et al., 2012), *Cajanus cajan* (Nithin et al., 2017), *Phaseolus vulgaris* (Formey et al., 2015), *A. membranaceus* (Abla et al., 2019), and *A. mongolicus* (Gao et al., 2018).

miR4415 has been reported to be induced by aluminum treatment and drought stress but downregulated by biotic stress in *G. soja* (Zeng et al., 2012). miR4415 was observed to be downregulated by cold stress treatment in *A. nanus* and *A. membranaceus* (Abla et al., 2019) and by osmotic stress in *A. mongolicus* (Gao et al., 2018). These results indicated its involvement in the abiotic and biotic stress response.

AO was a key enzyme regulating the ROS status of the apoplast and an important regulatory enzyme for plant growth, development, environmental stress perception, and stress signal transduction (Foyer et al., 2020). AO can catalyze the oxidation of ASA to unstable MDHA and then form dehydroascorbic acid through nonenzymatic disproportionation. The redox status of the apoplast is largely determined by AO (Karpinska et al., 2018), and the change of ASA/DHA ratio caused by environmental stress may regulate gene expression and protein abundance through the change of redox state of the apoplast.

In *A. nanus*, expression of the miR4415 precursor was inhibited after cold acclimation, leading to downregulation of the ana-miR4415a-3p and upregulation of the *L-AO* gene. An elevated level of L-AO promotes the oxidation of ASA to form DHA and lowers the AsA/DHA ratio, resulting in alteration of the redox state of the apoplast. Change in the apoplastic redox state regulates the tolerance to freezing stress in *A. nanus* seedlings (**Figure 12**). A similar regulatory pathway has been reported in rice, in which miR528 regulated the apoplastic ROS state by targeting the *L-AO* gene, leading to enhanced tolerance to salt stress (Wang et al., 2021) and viral infection (Wu et al., 2017).

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/, SRR10906482https://www.ncbi.nlm.nih.gov/, SRR10906483https://www.ncbi.nlm.nih.gov/, SRR10906484https://www.ncbi.nlm.nih.gov/, SRR10906485https://www.ncbi.nlm.nih.gov/, SRR10906486https://www.ncbi.nlm.nih.gov/, SRR10906487https://www.ncbi.nlm.nih.gov/, SRR17711209.

AUTHOR CONTRIBUTIONS

YiZ and FG contributed to the conception and design of the study. MZ performed the statistical analysis. MZ, JT, XW, YaZ, and FG performed the investigation. MZ and FG wrote the first draft of the manuscript. FG and YiZ wrote sections of the manuscript. All the authors contributed to manuscript revision and read and approved the submitted version.

FUNDING

This study was funded by the National Natural Science Foundation of China (Grant Number 31770363 and 31670335), the Ministry of Education of China under "Double First-Class" projects (Number Yldxxk201819), Key Laboratory of Ecology and Environment in Minority Areas (Minzu University of China), the National Ethnic Affairs Commission (No. KLEEMA202106), and Beijing Advanced Discipline for Mass Spectrometry Imaging and Metabolomics (No. 104-01900403).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.870446/full#supplementary-material

Supplementary Table S1 | Primers used for qRT-PCR analysis.

Supplementary Table S2 | Conserved miRNAs identified from *A. nanus*.

Supplementary Table S3 | Species-specific miRNAs identified in *A. nanus*.

Supplementary Table S4 | Targets of the conserved miRNAs predicted in *A. nanus* by psRNATarget.

Supplementary Table S5 | Targets of the species-specific miRNAs predicted in *A. nanus* by psRNATarget.

Supplementary Table S6 | Targets of the miRNAs identified in *A. nanus* by degradome analysis.

REFERENCES

Abla, M., Sun, H., Li, Z., Wei, C., Gao, F., Zhou, Y., et al. (2019). Identification of miRNAs and Their Response to Cold Stress in *Astragalus Membranaceus*. *Biomolecules* 9, 182. doi:10.3390/biom9050182

Addo-Quaye, C., Miller, W., and Axtell, M. J. (2009). CleaveLand: A Pipeline for Using Degradome Data to Find Cleaved Small RNA Targets. *Bioinformatics* 25, 130–131. doi:10.1093/bioinformatics/btn604

Axtell, M. J., and Meyers, B. C. (2018). Revisiting Criteria for Plant microRNA Annotation in the Era of Big Data. *Plant Cell* 30, 272–284. doi:10.1105/tpc.17.00851

Bielli, P., and Calabrese, L. (2002). Structure to Function Relationships in Ceruloplasmin: A 'Moonlighting' Protein. *Cell Mol. Life Sci. (Cmls)* 59, 1413–1427. doi:10.1007/s00018-002-8519-2

Bredow, M., and Walker, V. K. (2017). Ice-Binding Proteins in Plants. *Front. Plant Sci.* 8, 2153. doi:10.3389/fpls.2017.02153

Cao, S., Wang, Y., Li, X., Gao, F., Feng, J., and Zhou, Y. (2020). Characterization of the AP2/ERF Transcription Factor Family and Expression Profiling of DREB Subfamily under Cold and Osmotic Stresses in *Ammopiptanthus Nanus*. *Plants* 9, 455. doi:10.3390/plants9040455

Cao, S., Wang, Y., Li, Z., Shi, W., Gao, F., Zhou, Y., et al. (2019). Genome-Wide Identification and Expression Analyses of the Chitinases under Cold and Osmotic Stress in *Ammopiptanthus Nanus*. *Genes* 10, 472. doi:10.3390/genes10060472

Chang, H.-C., Tsai, M.-C., Wu, S.-S., and Chang, I.-F. (2019). Regulation of ABI5 Expression by ABF3 during Salt Stress Responses in *Arabidopsis Thaliana*. *Bot. Stud.* 60, 16. doi:10.1186/s40529-019-0264-z

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* 13, 1194–1202. doi:10.1016/j.molp.2020.06.009

Chen, Y., Chen, Z., Kang, J., Kang, D., Gu, H., and Qin, G. (2013). AtMYB14 Regulates Cold Tolerance in Arabidopsis. *Plant Mol. Biol. Rep.* 31, 87–97. doi:10.1007/s11105-012-0481-z

Clément, T., Salone, V., and Rederstorff, M. (2015). Dual Luciferase Gene Reporter Assays to Study miRNA Function. *Methods Mol. Biol.* 1296, 187–198. doi:10.1007/978-1-4939-2547-6_17

Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: A Plant Small RNA Target Analysis Server (2017 Release). *Nucleic Acids Res.* 46, W49–W54. doi:10.1093/nar/gky316

Dewey, C. N. (2010). RNA-Seq Gene Expression Estimation with Read Mapping Uncertainty. *Bioinformatics* 26, 493–500. doi:10.1093/bioinformatics/btp692

Ding, Y., Shi, Y., and Yang, S. (2020). Molecular Regulation of Plant Responses to Environmental Temperatures. *Mol. Plant* 13, 544–564. doi:10.1016/j.molp.2020.02.004

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB Transcription Factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi:10.1016/j.tplants.2010.06.005

Formey, D., Iñiguez, L. P., Peláez, P., Li, Y.-F., Sunkar, R., Sánchez, F., et al. (2015). Genome-Wide Identification of the Phaseolus Vulgaris sRNAome Using Small RNA and Degradome Sequencing. *BMC Genomics* 16, 423. doi:10.1186/s12864-015-1639-5

Foyer, C. H., Kyndt, T., and Hancock, R. D. (2020). Vitamin C in Plants: Novel Concepts, New Perspectives, and Outstanding Issues. *Antioxid. Redox Signaling* 32, 463–485. doi:10.1089/ars.2019.7819

Gao, F., Wang, X., Li, X., Xu, M., Li, H., Abla, M., et al. (2018). Long-Read Sequencing and De Novo Genome Assembly of *Ammopiptanthus Nanus*, a Desert Shrub. *Gigascience* 7, 074. doi:10.1093/gigascience/giy074

Gao, F., Wang, N., Li, H., Liu, J., Fu, C., Xiao, Z., et al. (2016). Identification of Drought-Responsive microRNAs and Their Targets in *Ammopiptanthus Mongolicus* by Using High-Throughput Sequencing. *Sci. Rep.* 6, 34601. doi:10.1038/srep34601

Hajyzadeh, M., Turktas, M., Khawar, K. M., and Unver, T. (2015). miR408 Overexpression Causes Increased Drought Tolerance in Chickpea. *Gene* 555, 186–193. doi:10.1016/j.gene.2014.11.002

Jang, J.-C. (2016). Arginine-Rich Motif-Tandem CCCH Zinc finger Proteins in Plant Stress Responses and Post-Transcriptional Regulation of Gene Expression. *Plant Sci.* 252, 118–124. doi:10.1016/j.plantsci.2016.06.014

Jones-Rhoades, M. W., and Bartel, D. P. (2004). Computational Identification of Plant MicroRNAs and Their Targets, Including a Stress-Induced miRNA. *Mol. Cel* 14, 787–799. doi:10.1016/j.molcel.2004.05.027

Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L., et al. (2017). CPC2: A Fast and Accurate Coding Potential Calculator Based on Sequence Intrinsic Features. *Nucleic Acids Res.* 45, W12–W16. doi:10.1093/nar/gkx428

Karpinska, B., Zhang, K., Rasool, B., Pastok, D., Morris, J., Verrall, S. R., et al. (2018). The Redox State of the Apoplast Influences the Acclimation of Photosynthesis and Leaf Metabolism to Changing Irradiance. *Plant Cel Environ* 41, 1083–1097. doi:10.1111/pce.12960

Khan, T. A., Fariduddin, Q., and Yusuf, M. (2017). Low-temperature Stress: Is Phytohormones Application a Remedy? *Environ. Sci. Pollut. Res.* 24, 21574–21590. doi:10.1007/s11356-017-9948-7

Komori, H., and Higuchi, Y. (2015). Structural Insights into the O^2 Reduction Mechanism of Multicopper Oxidase. *J. Biochem.* 158, 293–298. doi:10.1093/jb/mvv079

Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: From microRNA Sequences to Function. *Nucleic Acids Res.* 47, D155–D162. doi:10.1093/nar/gky1141

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: Annotating High Confidence microRNAs Using Deep Sequencing Data. *Nucl. Acids Res.* 42, D68–D73. doi:10.1093/nar/gkt1181

Krützfeldt, J., and Stoffel, M. (2006). MicroRNAs: A New Class of Regulatory Genes Affecting Metabolism. *Cel Metab.* 4, 9–12. doi:10.1016/j.cmet.2006.05.009

Lei, P., Han, B., Wang, Y., Zhu, X., Xuan, Y., Liu, X., et al. (2019). Identification of microRNAs that Respond to Soybean Cyst Nematode Infection in Early Stages in Resistant and Susceptible Soybean Cultivars. *Int. J. Mol. Sci.* 20, 5634. doi:10.3390/ijms20225634

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., et al. (2002). PlantCARE, a Database of Plant Cis-Acting Regulatory Elements and a portal to Tools for In Silico Analysis of Promoter Sequences. *Nucleic Acids Res.* 30, 325–327. doi:10.1093/nar/30.1.325

Liu, C. T., Wang, W., Mao, B. G., and Chu, C. C. (2018). Cold Stress Tolerance in Rice: Physiological Changes, Molecular Mechanism, and Future Prospects. *Yi Chuan* 40, 171–185. doi:10.16288/j.yczz.18-007

Liu, H.-H., Tian, X., Li, Y.-J., Wu, C.-A., and Zheng, C.-C. (2008). Microarray-Based Analysis of Stress-Regulated microRNAs in *Arabidopsis Thaliana*. *Rna* 14, 836–843. doi:10.1261/rna.895308

Liu, Y., Dang, P., Liu, L., and He, C. (2019). Cold Acclimation by the CBF-COR Pathway in a Changing Climate: Lessons from *Arabidopsis thaliana*. *Plant Cel Rep* 38, 511–519. doi:10.1007/s00299-019-02376-3

Lv, D.-K., Bai, X., Li, Y., Ding, X.-D., Ge, Y., Cai, H., et al. (2010). Profiling of Cold-Stress-Responsive miRNAs in rice by Microarrays. *Gene* 459, 39–47. doi:10.1016/j.gene.2010.03.011

Lv, L.-M., Zuo, D.-Y., Wang, X.-F., Cheng, H.-L., Zhang, Y.-P., Wang, Q.-L., et al. (2020). Genome-Wide Identification of the Expansin Gene Family Reveals that Expansin Genes Are Involved in Fibre Cell Growth in Cotton. *BMC Plant Biol.* 20, 223. doi:10.1186/s12870-020-02362-y

Lyzenga, W. J., and Stone, S. L. (2012). Abiotic Stress Tolerance Mediated by Protein Ubiquitination. *J. Exp. Bot.* 63, 599–616. doi:10.1093/jxb/err310

Ma, C., Burd, S., and Lers, A. (2015a). miR408is Involved in Abiotic Stress Responses in Arabidopsis. *Plant J.* 84, 169–187. doi:10.1111/tpj.12999

Ma, F., Liu, Z., Huang, J., Li, Y., Kang, Y., Liu, X., et al. (2019). High-Throughput Sequencing Reveals microRNAs in Response to Heat Stress in the Head Kidney of Rainbow trout (*Oncorhynchus Mykiss*). *Funct. Integr. Genomics* 19, 775–786. doi:10.1007/s10142-019-00682-3

Ma, X., Tang, Z., Qin, J., and Meng, Y. (2015b). The Use of High-Throughput Sequencing Methods for Plant microRNA Research. *RNA Biol.* 12, 709–719. doi:10.1080/15476286.2015.1053686

Megha, S., Basu, U., and Kav, N. N. V. (2018). Regulation of Low Temperature Stress in Plants by microRNAs. *Plant Cel Environ* 41, 1–15. doi:10.1111/pce.12956

Millar, A. A., Lohe, A., and Wong, G. (2019). Biology and Function of miR159 in Plants. *Plants* 8, 255. doi:10.3390/plants8080255

Morin, R. D., Aksay, G., Dolgosheina, E., Ebhardt, H. A., Magrini, V., Mardis, E. R., et al. (2008). Comparative Analysis of the Small RNA Transcriptomes of *Pinus Contorta* and *Oryza Sativa*. *Genome Res.* 18, 571–584. doi:10.1101/gr.6897308

Nagatoshi, Y., Fujita, M., and Fujita, Y. (2018). Casein Kinase 2 α and β Subunits Inversely Modulate ABA Signal Output in Arabidopsis Protoplasts. *Planta* 248, 571–578. doi:10.1007/s00425-018-2919-5

Nithin, C., Thomas, A., Basak, J., and Bahadur, R. P. (2017). Genome-Wide Identification of miRNAs and lncRNAs in *Cajanus Cajan*. *BMC Genomics* 18, 878. doi:10.1186/s12864-017-4232-2

O'leary, B. M., Rico, A., Mccraw, S., Fones, H. N., and Preston, G. M. (2014). The Infiltration-Centrifugation Technique for Extraction of Apoplastic Fluid from Plant Leaves Using Phaseolus Vulgaris as an Example. *J. Vis. Exp.*, 52113. doi:10.3791/52113

Rajesh, K. S., Bharath, B. R., Rao, C. V., Bhat, K. I., Bhat, K. S. C., and Bhat, P. (2017). Neutralization of Naja naja Venom Induced Lethality, Edema and Myonecrosis by Ethanolic Root Extract of *Coix Lacryma-Jobi*. *Toxicol. Rep.* 4, 637–645. doi:10.1016/j.toxrep.2017.11.005

Ranocha, P., Mcdougall, G., Hawkins, S., Sterjiades, R., Borderies, G., Stewart, D., et al. (1999). Biochemical Characterization, Molecular Cloning and Expression of Laccases - a Divergent Gene Family - in Poplar. *Eur. J. Biochem.* 259, 485–495. doi:10.1046/j.1432-1327.1999.00061.x

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

Sanmartin, M., Pateraki, I., Chatzopoulou, F., and Kanellis, A. K. (2007). Differential Expression of the Ascorbate Oxidase Multigene Family during Fruit Development and in Response to Stress. *Planta* 225, 873–885. doi:10.1007/s00425-006-0399-5

Song, C., Wang, C., Zhang, C., Korir, N. K., Yu, H., Ma, Z., et al. (2010). Deep Sequencing Discovery of Novel and Conserved microRNAs in Trifoliate orange (*Citrus Trifoliata*). *BMC Genomics* 11, 431. doi:10.1186/1471-2164-11-431

Song, X., Li, Y., Cao, X., and Qi, Y. (2019). MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. *Annu. Rev. Plant Biol.* 70, 489–525. doi:10.1146/annurev-arplant-050718-100334

Sun, L., and Chen, Z. J. (2004). The Novel Functions of Ubiquitination in Signaling. *Curr. Opin. Cel Biol.* 16, 119–126. doi:10.1016/j.ceb.2004.02.005

Sunkar, R., and Zhu, J.-K. (2004). Novel and Stress-Regulated MicroRNAs and Other Small RNAs from Arabidopsis[W]. *Plant Cell* 16, 2001–2019. doi:10.1105/tpc.104.022830

Szittya, G., Moxon, S., Santos, D. M., Jing, R., Fevereiro, M. P., Moulton, V., et al. (2008). High-Throughput Sequencing of Medicago Truncatula Short RNAs Identifies Eight New miRNA Families. *BMC Genomics* 9, 593. doi:10.1186/1471-2164-9-593

Tang, W., and Thompson, W. A. (2019). OsmiR528 Enhances Cold Stress Tolerance by Repressing Expression of Stress Response-Related Transcription Factor Genes in Plant Cells. *Curr. Genomics* 20, 100–114. doi:10.2174/1389202920666190129145439

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, B., Sun, Y.-F., Song, N., Wei, J.-P., Wang, X.-J., Feng, H., et al. (2014). MicroRNAs Involving in Cold, Wounding and Salt Stresses in *Triticum Aestivum* L. *Plant Physiol. Biochem.* 80, 90–96. doi:10.1016/j.plaphy.2014.03.020

Wang, M., Guo, W., Li, J., Pan, X., Pan, L., Zhao, J., et al. (2021). The miR528-AO Module Confers Enhanced Salt Tolerance in rice by Modulating the Ascorbic Acid and Abscisic Acid Metabolism and ROS Scavenging. *J. Agric. Food Chem.* 69, 8634–8648. doi:10.1021/acs.jafc.1c01096

Wu, J., Yang, R., Yang, Z., Yao, S., Zhao, S., Wang, Y., et al. (2017). ROS Accumulation and Antiviral Defence Control by microRNA528 in rice. *Nat. Plants* 3, 16203. doi:10.1038/nplants.2016.203

Xu, S., Ha, C. H., Wang, W., Xu, X., Yin, M., Jin, F. Q., et al. (2016). PECAM1 Regulates Flow-Mediated Gab1 Tyrosine Phosphorylation and Signaling. *Cell Signal.* 28, 117–124. doi:10.1016/j.cellsig.2015.12.007

Yang, A., Dai, X., and Zhang, W.-H. (2012). A R2R3-Type MYB Gene, OsMYB2, Is Involved in Salt, Cold, and Dehydration Tolerance in rice. *J. Exp. Bot.* 63, 2541–2556. doi:10.1093/jxb/err431

Yang, G., Rhodes, D., and Joly, R. (1996). Effects of High Temperature on Membrane Stability and Chlorophyll Fluorescence in Glycinebetaine-Deficient and Glycinebetaine-Containing maize Lines. *Funct. Plant Biol.* 23, 437. doi:10.1071/pp9960437

Yang, G., Zou, H., Wu, Y., Liu, H., and Yuan, Y. (2011). Identification and Characterisation of Candidate Genes Involved in Chilling Responses in Maize (*Zea mays* L.). *Plant Cel Tiss Organ. Cult* 106, 127–141. doi:10.1007/s11240-010-9900-8

Yasar, F., Ellialtioglu, S., and Yildiz, K. (2008). Effect of Salt Stress on Antioxidant Defense Systems, Lipid Peroxidation, and Chlorophyll Content in green Bean. *Russ. J. Plant Physiol.* 55, 782–786. doi:10.1134/s1021443708060071

Yin, Z., Wang, N., Duan, W., Pi, L., Shen, D., and Dou, D. (2021). Phytophthora Capsici CBM1-Containing Protein CBP3 Is an Apoplastic Effector with Plant Immunity-Inducing Activity. *Mol. Plant Pathol.* 22, 1358–1369. doi:10.1111/mpp.13116

Yoo, S.-D., Cho, Y.-H., and Sheen, J. (2007). Arabidopsis Mesophyll Protoplasts: A Versatile Cell System for Transient Gene Expression Analysis. *Nat. Protoc.* 2, 1565–1572. doi:10.1038/nprot.2007.199

Yu, Y., Zhang, Y., Chen, X., and Chen, Y. (2019). Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. *Annu. Rev. Cel Dev. Biol.* 35, 407–431. doi:10.1146/annurev-cellbio-100818-125218

Zeng, Q.-Y., Yang, C.-Y., Ma, Q.-B., Li, X.-P., Dong, W.-W., and Nian, H. (2012). Identification of Wild Soybean miRNAs and Their Target Genes Responsive to Aluminum Stress. *BMC Plant Biol.* 12, 182. doi:10.1186/1471-2229-12-182

Zhang, R., Wang, Y.-Q., and Su, B. (2008). Molecular Evolution of a Primate-Specific microRNA Family. *Mol. Biol. Evol.* 25, 1493–1502. doi:10.1093/molbev/msn094

Zhang, W., Ruan, J., Ho, T.-H. D., You, Y., Yu, T., and Quatrano, R. S. (2005). Cis-Regulatory Element Based Targeted Gene Finding: Genome-Wide Identification of Abscisic Acid- and Abiotic Stress-Responsive Genes in *Arabidopsis Thaliana*. *Bioinformatics* 21, 3074–3081. doi:10.1093/bioinformatics/bti490

Zhang, Y., Cao, Y., Zheng, H., Feng, W., Qu, J., Fu, F., et al. (2021). Ectopic Expression of Antifreeze Protein Gene from Ammopiptanthus Nanus Confers Chilling Tolerance in maize. *Crop J.* 9, 924–933. doi:10.1016/j.cj.2020.08.011

Zhou, M., and Tang, W. (2019). MicroRNA156 Amplifies Transcription Factor-Associated Cold Stress Tolerance in Plant Cells. *Mol. Genet. Genomics* 294, 379–393. doi:10.1007/s00438-018-1516-4

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fgene.2022.883422

Integrated mRNA and Small RNA Sequencing Reveals microRNAs Associated With Xylem Development in *Dalbergia odorifera*

Wenxiu Zhao[1,2], *Xiangxu Meng*[1,2], *Jiahong Xu*[1], *Zijia Liu*[2], *Yangyang Hu*[1], *Bingyu Li*[1], *Jinhui Chen*[1,2]* *and Bing Cao*[2]*

[1]*Key Laboratory of Genetics and Germplasm Innovation of Tropical Special Forest Trees and Ornamental Plants, Ministry of Education/Engineering Research Center of Rare and Precious Tree Species in Hainan Province, School of Forestry, Hainan University, Haikou, China,* [2]*Sanya Nanfan Research Institute of Hainan University, Hainan Yazhou Bay Seed Laboratory, Sanya, China*

Dalbergia odorifera is a rare and precious rosewood specie, whose wood is a very high-quality material for valuable furniture and carving crafts. However, limited information is available about the process of wood formation in *D. odorifera*. To determine genes that might be closely associated with the xylem differentiation process, we analyzed the differentially expressed genes (DEGs) and microRNAs (miRNAs) from specific xylem tissues of *D. odorifera* by RNA sequencing (RNA-seq) and small RNA sequencing (small RNA-seq). In total, we obtained 134,221,955 clean reads from RNA-seq and 90,940,761 clean reads from small RNA-seq. By comparing the transition zone (Dotz) and sapwood (Dosw) samples, a total of 395 DEGs were identified. Further analysis revealed that DEGs encoded for WRKY transcription factors (eight genes), lignin synthesis (*PER47*, *COMT*, *CCR2*), cell wall composition (*UXS2*), gibberellin synthesis (*KAO2*, *GA20OX1*), jasmonic acid synthesis (*OPR2, CYP74A*), and synthesis of flavonoids (*PAL2*) and terpenoids (*CYP71A1*). Subsequently, a preliminary analysis by small RNA-seq showed that the expressions of 14 miRNAs (such as miR168a-5p, miR167f-5p, miR167h-5p, miR167e, miR390a, miR156g, novel_52, and novel_9) were significantly different between Dotz and Dosw. Further analysis revealed that the target genes of these differentially expressed miRNAs were enriched in the GO terms “amino acid binding,” “cellulase activity,” and “DNA beta-glucosyltransferase activity”. Further, KEGG pathway annotation showed significant enrichment in “fatty acid elongation” and “biosynthesis of unsaturated fatty acids”. These processes might be participating in the xylem differentiation of *D. odorifera*. Next, expression correlation analysis showed that nine differentially expressed miRNAs were significantly negatively associated with 21 target genes, which encoded for proteins such as *pyrH*, *SPL6*, *SPL12*, *GCS1*, and *ARF8*. Overall, this is the first study on miRNAs and their potential functions in the xylem development of *D. odorifera*, which provides a stepping stone for a detailed functional investigation of *D. odorifera* miRNAs.

Keywords: *Dalbergia odorifera*, microRNA (miRNA), xylem differentiation, RNA sequencing, species-specific noncoding RNAs

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Liu-Qiang Wang,
Chinese Academy of Forestry, China
Deng Shu Rong,
Beijing Forestry University, China

***Correspondence:**
Jinhui Chen
jinhuichen@hainanu.edu.cn
Bing Cao
caobing@hainanu.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *25 February 2022*
Accepted: *23 March 2022*
Published: *25 April 2022*

Citation:
Zhao W, Meng X, Xu J, Liu Z, Hu Y, Li B, Chen J and Cao B (2022) Integrated mRNA and Small RNA Sequencing Reveals microRNAs Associated With Xylem Development in Dalbergia odorifera. Front. Genet. 13:883422. doi: 10.3389/fgene.2022.883422

INTRODUCTION

Dalbergia odorifera is a semi-deciduous tree that belongs to the Leguminosae family (Vatanparast et al., 2013; Wariss et al., 2017), with the characteristics of easy germination, good tolerance to drought and barren soil, high disease resistance, and high adaptability (Liu et al., 2017; Hong et al., 2020). The heartwood of *D. odorifera* is exquisite in color and pattern, hard in texture, and has a unique aroma. Therefore, it has always been considered as a very high-quality material for making precious furniture, musical instruments, decorations, and handicrafts (Meng et al., 2019). Due to its extremely high application value, *D. odorifera* was over-exploited, the wild resources were extensively destroyed. The natural forests became increasingly scarce and were almost extinct (Liu et al., 2019). The slow speed of heartwood formation further prevents the existing resources of *D. odorifera* from meeting the demand of domestic and foreign wood markets (Cui et al., 2020). Therefore, it is urgent to cultivate *D. odorifera* trees with good wood quality.

Wood formation involves the biosynthesis and deposition of lignin and cellulose on the cell wall, which comprises a complex network of multiple genes (Mizrachi and Myburg, 2016; Wang Y.-H. et al., 2018). Further, the heartwood formation also involves the generation of special metabolites during its transformation from sapwood (Ye and Zhong, 2015; Celedon and Bohlmann, 2018). The main components of heartwood in *D. odorifera* are essential oils and flavonoids (The, 2017; Zhao et al., 2020). Essential oils are secondary metabolites having aromatic odor and volatile properties, and include terpenoids, aromatic compounds, aliphatic compounds, and compounds containing sulfur and nitrogen (Yang et al., 2022). The biosynthesis and accumulation of lignin, cellulose, essential oils, and flavonoids are important factors that influence xylem differentiation and wood formation in *D. odorifera*. Still, the identity of genes and other regulators involved in the process of xylem differentiation of *D. odorifera* under natural growth remain poorly known.

Wood refers to all tissues within the vascular cambium in the hard stems of perineal plants. Usually, in the cross-section of the wood, we can observe the heartwood is the dark and hard textured central area, while the sapwood is the light and soft textured outer area (Islam et al., 2012; Mizrachi and Myburg, 2016; Chen et al., 2018). As the trees grow, sapwood gradually transforms into the heartwood. This process usually takes place in a narrow transition zone, which comprises living cells that consume reserves such as starch (Bergström, 2003). In general, the transition zone is one to two rings wide and is adjacent to the heartwood (Celedon and Bohlmann, 2018). The xylem tissue of the transition zone is often regarded as an important material for studying wood formation. For example, in *Taiwania cryptomerioides*, the molecular mechanism of the autolysis of the cellular components of ray parenchyma cells in the transition zone, during the heartwood formation, was elucidated by transcriptome sequencing of sapwood and transition zone, in combination with the microscopy and high-performance liquid chromatography technology (Yeh et al., 2020). In *Pinus sylvestris*, transcriptomic sequencing revealed that stilbene and resinic acid were synthesized in transition zone and sapwood, respectively (Lim et al., 2016). These studies suggest that sapwood and transition zone are ideal plant parts for studying the dynamics of gene expression and molecular regulatory mechanisms of wood formation (Yeh et al., 2020).

MicroRNAs (miRNAs) now form an important component of machinery that regulate gene expression. MiRNAs are a class of small, single-strand non-coding RNAs that are approximately 21 nucleotides in length (Rogers and Chen, 2013; Yu et al., 2017). They exert a negative regulation on the expression of target genes through sequence-based complementarity, resulting in mRNA cleavage or translation repression (Chen et al., 2015; Song et al., 2019; Wang M et al., 2021). Studies have shown that miRNAs play important roles in plant growth and development, including organ morphogenesis, hormone secretion, and stress response (Aukerman and Sakai, 2003; Sunkar et al., 2012; Dong et al., 2022). For example, in *Osmanthus fragrans*, miR858 affects flavonoid content in flower tissues by negatively regulating the *MYB1* gene (Shi et al., 2021). Further, some miRNAs, such as miR397a (Lu et al., 2013), miR875 (Zhao et al., 2015), miR257 (Chen et al., 2016), and miR475b (Xiao et al., 2017), have been identified to participate in different processes of wood formation. However, the expression profile and functions of miRNAs in *D. odorifera* have not been reported. A detailed understanding of the complex molecular mechanism of *D. odorifera* xylem differentiation has also been lacking.

Here, we have investigated the expression profile of miRNAs and their potential role in regulating xylem differentiation in *D. odorifera*. We performed small RNA sequencing (small RNA-seq) and mRNA sequencing (mRNA-seq) analyses from three biological replicates of two tissue types (sapwood and transition zone) of *D. odorifera*. To the best of our knowledge, this is the first study to identify miRNAs of *D. odorifera* and describe their potential role in xylem differentiation. This study will broaden our understanding of the complex molecular mechanism of *D. odorifera* xylem differentiation. The results of sequencing analysis provide abundant candidate miRNAs and mRNAs, which are important for the innovation of *D. odorifera* germplasm resources and the cultivation of good wood quality varieties.

MATERIALS AND METHODS

Plant Materials

The three well-developed *D. odorifera* trees (about the age of 7 years) were grown at an artificial nursery of Hainan Province in China (19°38′56″N, 110°14′29″E). These three individual trees, which have formed heartwood, were selected as biological replicates without any treatment. These three trees were used for all the experiments. Samples of sapwood (Dosw1, Dosw2, Dosw3) and transition zone (Dotz1, Dotz2, Dotz3) were taken from xylem tissues near the cambium and near the heartwood of each tree, respectively, by following the previously described protocols (Yeh et al., 2020). All the samples were isolated from the trees with the help of a sharp chisel after removing the bark, and the tissues were immediately frozen in liquid nitrogen and stored at −80°C until RNA isolation.

Library Construction and Sequencing

Total RNA from Dotz and Dosw groups of samples was extracted by the CTAB method. Agarose gel electrophoresis (2% agarose gel) was used to evaluate RNA degradation and contamination, and a micro-spectrophotometer was used to assess the purity of RNA (260/280 ratio: 1.9–2.2, 260/230 ratio: ≥ 2.0). RNA integrity was evaluated with the help of RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2,100 system (Agilent Technologies, CA, United States), and samples with an RNA integrity number (RIN) greater than 7.0 were used for further experiments.

Magnetic beads with oligo (dT) were used to enrich mRNA by binding to its polyadenylated tail through A-T complementary pairing. A fragmentation buffer was used for fragmenting the mRNA. cDNA was synthesized using random hexamer primers, buffer, dNTPs, and DNA polymerase I. Double-stranded cDNA was purified by AMPure XP beads, and subjected to end repair, the addition of the poly-A tail, ligation of the sequencing linker, and fragment size selection. Finally, the cDNA libraries were subjected to PCR enrichment and sequenced with the Illumina HiSeq 2,500. After the original data was filtered, the redundancy was removed to obtain clean reads. The HISAT2 software was used to compare the clean reads of RNA-Seq to the reference genome sequence of *D. odorifera*.

The NEBNext® Multiplex small RNA library prep set was used to generate small RNA-seq libraries for Illumina® (NEB, Ipswich, MA, United States.) by following the manufacturer's instructions. Then, the TruSeq SR Cluster Kit v3-cBot-HS (Illumina) was used to generate a cluster on cBot cluster generation system. Finally, all small RNA-seq libraries were sequenced on an Illumina HiSeq 2,500 platform.

Analysis of miRNA and mRNA Expression Profiles

After sequencing, clean reads were obtained by removing the reads containing poly-N, ploy A/T/G/C, with 5′ adapter contaminants, without 3′adapter or the insert tag, and the low-quality reads. In addition, Q20, Q30, and GC-content of the raw reads also were calculated. At the end, all the downstream analyses were performed on sequences ranging of 18–30 bp in length. Bowtie was used to locate/align the small RNAs to the reference genome of *D. odorifera* and analyze the distribution of small RNAs over the reference genome. The reads that mapped onto the reference genome of *D. odorifera* were compared with sequences in miRBase (20.0) to obtain the known miRNA; whereas novel miRNAs were predicted with the help of miREvo (Wen et al., 2012) and mirDeep2 (Friedlander et al., 2012). The expression level of miRNA was estimated by the TPM (transcript per million) value. The DESeq R package (3.0.3) was used to perform differential expression analysis of Dotz and Dosw with the $|\log_2 (\text{fold change})| \geq 1$ and p-value < 0.05 as the threshold.

Clean reads of RNA-seq were obtained by removing low-quality sequencing fragments from the raw reads. Then, the read count of each gene was obtained by mapping the clean reads to the reference genome of *D. odorifera* with the help of HISAT2. The expression levels of mRNAs were estimated by the calculating FPKM (fragments per kilobase of exon model per million) values. A $|\log_2 (\text{fold change})| \geq 1$ and q-value < 0.05 were used as criteria to screen differentially expressed genes (DEGs) between Dotz and Dosw samples.

Prediction of the Potential Target Genes of miRNA

The psRobot is focused on plant small RNA analysis, which has been widely used for target gene predictions (Wu et al., 2012; Zhao et al., 2020; Liao et al., 2021). Therefore, we used its online version (http://omicslab.genetics.ac.cn/psRobot/target_prediction_1.php/) to predict the potential target genes of miRNAs, using the default parameters (Wu et al., 2012). Then, correlation analysis of accumulation of target genes and miRNAs was performed. When the expression levels of target genes had a strong negative correlation of at least −0.8 (p < 0.05) with miRNAs, they were selected for further analysis. All DEGs and the target genes of miRNAs were mapped to individual terms in the GO database (http://www.geneontology.org/) and the number of genes per term was calculated. Then, the GOseq software was used for GO enrichment analysis of DEGs. Analysis of gene regulatory pathways in the KEGG pathway database (http://www.genome.jp/kegg/pathway.html/) was performed with the help of KOBAS (3.0) software. Finally, Cytoscape 3.9.0 was used to construct a co-expression network graph.

Validation of miRNA and Gene Expression by RT-qPCR

cDNAs were synthesized by reverse transcription of total RNA from six *D. odorifera* samples (Dosw1, Dosw2, Dosw3, Dotz1, Dotz2, and Dotz3). Gene-specific primers for the targets were designed with the help of Primer Premier v5 software (**Supplementary Table S1**). Four important DEGs (*WRKY22*, *AP2*, *FPP7*, and *PAL2*), five miRNA target genes (*SPL12*, *ARF8*, *FH20*, *GCS1*, and *MMT1*) of miRNA-mRNA correlation network, and six differentially expressed miRNAs (miR156g, miR167e, miR168a-5p, novel_9, novel_15, and novel_52) were chosen to verify expression levels. For the genes, RT-qPCR analysis was conducted with TB Green® Premix Ex Taq™ (Tli RNaseH Plus; Takara, Beijing, China) following the manufacturer's recommendations. The amplification was performed on a preheated (94°C) thermal cycler, and samples were incubated at 94°C for 2 min, followed by 40 cycles of 95 °C for 5 s and 60°C for 30 s. The *actin* gene served as an internal control for normalization (Meng et al., 2019). We used the miRNA RT-qPCR Detection Kit (Aidlab, Beijing, China) for RT-qPCR analysis of miRNAs, following the manufacturer's recommendations. PCR amplification was performed at a preheated (94°C) thermal cycler and samples were incubated at 94 °C for 2 min, followed by 40 cycles of 94°C for 15 s and 60°C for 40 s. The *U6* gene served as an internal control for normalization (Meng et al., 2022). The $2^{-\triangle\triangle Ct}$ method was used to calculate the expression levels of the miRNAs and genes against the internal controls (Schmittgen and Livak, 2008). Three

TABLE 1 | Summary of mRNA sequencing datasets.

Sample	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC (%)	Mapping rate (%)
Dotz1	21,748,426	20,567,345	96.42	90.26	45.48	93.25
Dotz2	21,250,794	20,373,919	96.51	90.47	45.19	95.51
Dotz3	27,412,017	25,875,484	97.17	92.16	45.41	94.21
Dosw1	22,942,939	21,492,446	96.40	90.41	46.17	95.51
Dosw2	24,101,319	22,895,259	96.80	91.18	45.46	95.37
Dosw3	23,998,691	23,017,502	96.80	91.16	45.70	95.42

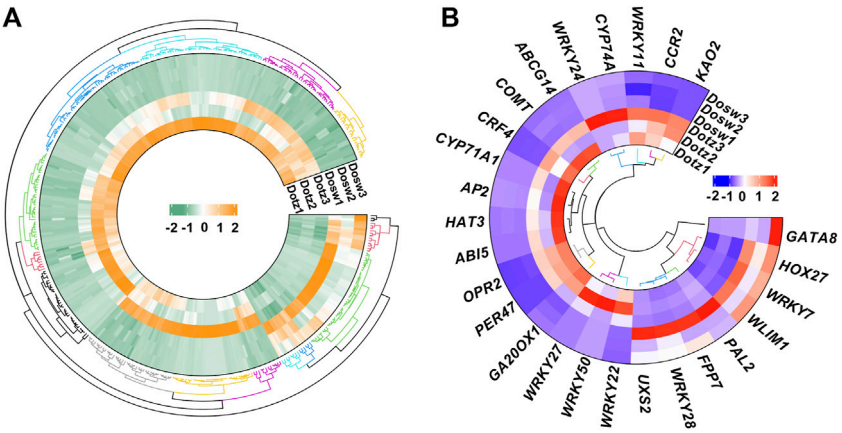


FIGURE 1 | Expression profiles of mRNAs in *D. odorifera*. Hierarchical clustering of all differentially expressed mRNAs **(A)**. Hierarchical clustering of expression of important differentially expressed genes **(B)**.

technical replicates per sample were analyzed to ensure reproducibility and reliability.

RESULTS

mRNA Expression Profile of Xylem Differentiation in *D. odorifera*

To understand the molecular mechanism of xylem differentiation, six cDNA libraries from the Dotz and Dosw of *D. odorifera* were sequenced. RNA-seq generated 20,567,345 (Dotz1), 20,373,919 (Dotz2), 25,875,484 (Dotz3), 21,492,446 (Dosw1), 22,895,259 (Dosw2) and 23,017,502 (Dosw3) clean reads (**Table 1**). Of these, 19,643,871 (Dotz1), 19,430,607 (Dotz2), 24,690,387 (Dotz3), 20,041,706 (Dosw1), 21,867,262 (Dosw2) and 21,684,789 (Dosw3) clean reads, respectively, were mapped to the *D. odorifera* genome with mapping ratios of 93.25–95.51% (**Table 1**). These results show that RNA-seq captured a significant portion of the genes in the genome of *D. ororifera*. For differential analysis of gene expression between the Dotz and Dosw groups, the FPKM values were used to normalize the reads from RNA-seq, and a cutoff of $|\log_2 (\text{fold change})| \geq 1$ and q-value <0.05 was used. A total of 395 mRNAs were differentially accumulated, of which 72 were up- and 323 were down-regulated, respectively, in Dotz compared to Dosw (**Figure 1A**; **Supplementary Table S2**). All 395 differentially expressed mRNAs (**Supplementary Table S2**) obtained from Dotz and Dosw were used for subsequent analysis.

Functional Annotation of the DEGs

To annotate and reveal the function of DEGs in different tissues of xylem, we used GO classification. Up-regulated mRNAs were associated with 11 molecular functions, 49 biological processes, and four cellular components (p-value < 0.05), including "anion transport," "anion transmembrane transporter activity," and "inorganic anion transport" associated with ion transport (**Figure 2A**). Down-regulated mRNAs were associated with 41 molecular functions, 54 biological processes, and three cellular components (p-value < 0.05), including "inorganic anion transmembrane transporter activity", "transmembrane transport", "sulfate transport" and "sulfate transmembrane transporter activity" terms associated with anion transport (**Figure 2B**). The DEGs were further referenced through the KEGG database. Eight KEGG pathways (p-value < 0.05) were significantly enriched, including "carotenoid biosynthesis," "diterpenoid biosynthesis," and "alpha-linolenic acid metabolism" (**Figure 3**).

Functional annotation of DEGs revealed the processes, such as the transcriptional regulation, lignin synthesis, flavonoids, and terpenoids synthesis, which could play important roles in the xylem differentiation of *D. odorifera*. We found that 11 DEGs might encode for transcription factors, of which eight belong to

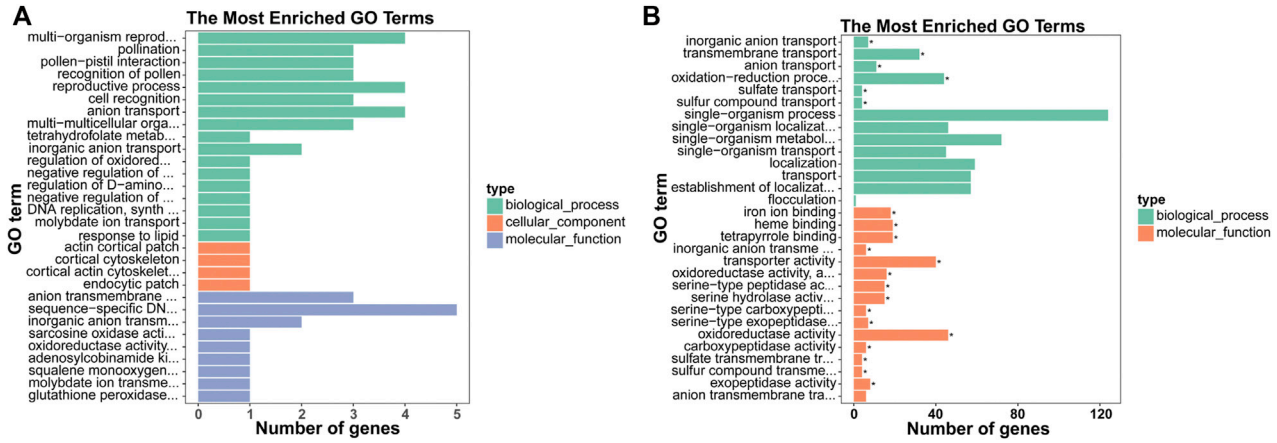


FIGURE 2 | GO analysis of the biological functions of mRNAs. The x-axis represents the number of genes, and the y-axis represents the GO term. GO enrichment analysis of the up- **(A)** and down-regulated **(B)** mRNAs are presented.

the WRKY transcription factor family (**Figure 1B**). Four genes (*PER47* (Peroxidase 47), *COMT* (Caffeic acid 3-O-methyltransferase), *CCR2* (Cinnamoyl-CoA reductase 2), and *UXS2* (UDP-glucuronic acid decarboxylase 2)) are related to lignin synthesis and cell wall composition, two genes (*KAO2* (Ent-kaurenoic acid oxidase 2) and *GA20OX1* (Gibberellin 20 oxidase 1)) are part of gibberellin synthesis, two genes (*OPR2* (12-oxophytodienoate reductase 2) and *CYP74A* (Allene oxide synthase)) participate in jasmonic acid synthesis, and two genes (*PAL2* (Phenylalanine ammonia-lyase 2) and *CYP71A1* (Cytochrome P450 71A1)) are related to flavonoids and terpenoids synthesis (**Figure 1B**).

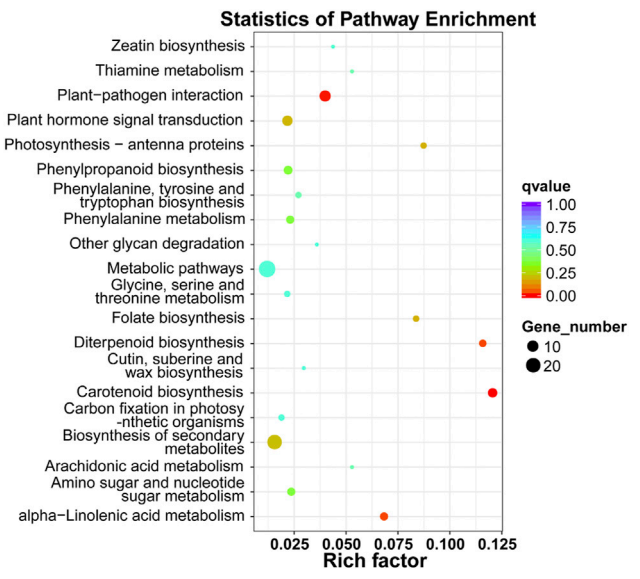


FIGURE 3 | KEGG pathway enrichment analysis of differentially expressed mRNAs. The x-axis represents the rich factor, whereas the y-axis represents the KEGG pathway term. The size of the dots represents the number of genes, and the color of the dots represents the qvalue.

Sequencing of Small RNAs

Small RNA-seq was performed to unveil the possible role of miRNAs in regulating gene expression during xylem differentiation in *D. odorifera*. 15,809,909 (Dotz1), 15,412,631 (Dotz2), 15,469,161 (Dotz3), 15,422,891 (Dosw1), 15,394,233 (Dosw2) and 15,755,203 (Dosw3) raw reads were obtained from the small RNA libraries generated from the Dotz and Dosw groups. After removing connectors and low-quality reads, a total of 45,496,409 (Dotz) and 45,444,352 (Dosw) clean reads were obtained (**Table 2**). The length of most clean reads ranged between 21 and 24 nucleotides (**Supplementary Figure S1**). In comparison by Bowtie software, 90.87–95.30% of small RNA reads matched with the genome sequence of *D. odorifera* (**Table 2**).

Identification of Known and Novel miRNA

The 18–30 nt clean reads that mapped to the reference genome were compared to sequences in miRBase to identify conserved/known miRNAs. Subsequently, novel miRNAs were predicted using miREvo and miRDeep2 tools on the basis of structural characteristics of miRNA. We identified 38 known miRNAs in Dotz and 36 in Dosw samples, respectively. Together, these accounted for a total of 40 known miRNAs that belonged to 22 miRNA-families (**Figure 4A**; **Supplementary Table S3**; **Supplementary Figure S2**). Further analysis of the distribution of the first base of these known miRNAs showed that "U" was the most dominant base (**Supplementary Table S4**; **Supplementary Figure S3**). Also, a total of 123 novel miRNAs were identified, where 122 were present in Dotz and all 123 in Dosw (**Figure 4A**; **Supplementary Table S5**; **Supplementary Figure S4**). The statistical analysis of the distribution of the first base of these novel miRNAs also found that "U" was the most dominant base (**Supplementary Table S6**; **Supplementary Figure S5**).

TABLE 2 | Summary of small RNA sequencing datasets.

Sample	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC (%)	Mapping rate (%)
Dotz1	15,809,909	15,359,940	97.42	92.25	50.44	94.25
Dotz2	15,412,631	15,004,955	97.17	91.47	50.13	90.87
Dotz3	15,469,161	15,131,514	97.43	92.22	50.08	93.12
Dosw1	15,422,891	15,024,534	97.33	91.99	50.81	95.30
Dosw2	15,394,233	15,027,909	97.35	91.90	50.16	93.03
Dosw3	15,755,203	15,391,909	97.36	91.88	49.18	95.20

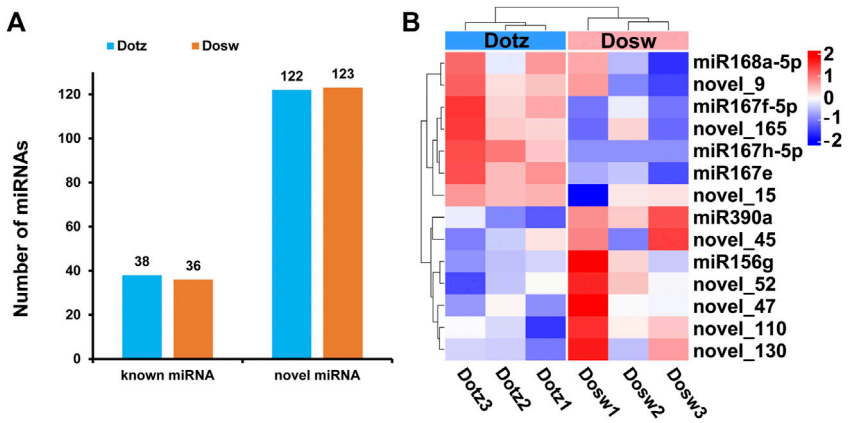


FIGURE 4 | The expression profile of miRNAs in xylem of *D. odorifera*. **(A)** The number of known and novel miRNAs. The x-axis represents the type of miRNAs, and the y-axis represents the number of miRNAs. **(B)** Hierarchical clustering of all differentially expressed miRNAs.

Differentially Expressed miRNAs and Their Targets

To understand the miRNA-driven mechanism of regulation of xylem differentiation, miRNAs that accumulated differentially between the two tissue types were identified by the TPM method. Compared with Dosw, four known (miR168a-5p, miR167f-5p, miR167h-5p, and miR167e) and three novel miRNAs (novel_15, novel_9, and novel_165) were up-regulated in Dotz, whereas two known (miR390a, miR156g) and five novel miRNAs (novel_45, novel_52, novel_110, novel_47, and novel_130) were down-regulated (**Figure 4B**). Among these differentially expressed miRNAs, the most up-regulated miRNAs were miR167e (52.21 fold), miR167h-5p (20.53 fold), and miR167f-5p (16.33 fold). The top down-regulated miRNAs were miR390a (52.75 fold), novel_45 (21.44 fold), and novel_52 (17.17 fold) (**Supplementary Table S7**).

To understand the potential function of differentially expressed miRNAs, 1,056 candidate target genes were predicted with the help of psRobot. Then, the target genes with a strong negative correlation of at least −0.8 ($p < 0.05$) with miRNAs were selected for further analysis (**Supplementary Table S8**). We, thus obtained 21 putative target genes for nine differentially expressed miRNAs. Subsequently, GO enrichment analysis indicated that these 21 target genes are associated with 42 GO terms (p-value < 0.05) from biological processes, molecular functions, and cellular components. Few examples of such processes included "amino acid binding," "cellulase activity," and "DNA beta-glucosyltransferase activity" (**Supplementary Figure S6A**). Further, KEGG enrichment indicated that the target genes were enriched in five pathways (p-value < 0.05), including "protein processing in endoplasmic reticulum," "fatty acid elongation," and "biosynthesis of unsaturated fatty acids" (**Supplementary Figure S6B**).

A Key miRNA-mRNA Regulatory Network During Xylem Differentiation of *D. odorifera*

To explore the relationship between miRNAs and mRNAs during xylem differentiation of *D. odorifera*, a regulatory network diagram of the nine differential expression miRNAs and their 21 target genes was constructed (**Figure 5**; **Supplementary Table S8**). It was evident that one miRNA could target 1-7 mRNAs, among which miR156g and novel_52 could negatively regulate seven and six mRNAs, respectively, while miR167h-5p and miR167f-5p could simultaneously regulate one mRNA (*pyrH*, Uridylate kinase). Similarly, miR156g might target the transcripts of Squamosa promoter-binding-like protein 6 (*SPL6*), which is involved in programmed cell death (Wang Q.-L. et al., 2018), as well as the *SPL12*. A novel miRNA, novel_130, might target mRNA of mannosyl-oligosaccharide glucosidase (*GCS1*), involved in cellulose synthesis. The up-regulated miRNA

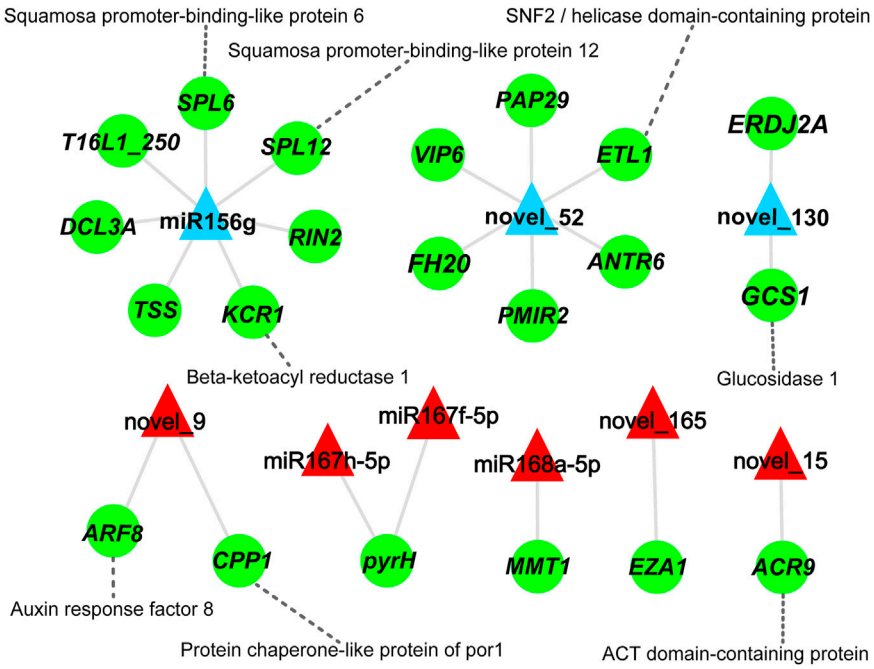


FIGURE 5 | miRNA-mRNA correlation network. Green circles represent the genes, red and blue triangles represent up- and down-regulated miRNAs, respectively. The gray solid line represents the targeted regulatory relationship between miRNAs and genes, and the gray dotted line shows the protein encoded by the gene.

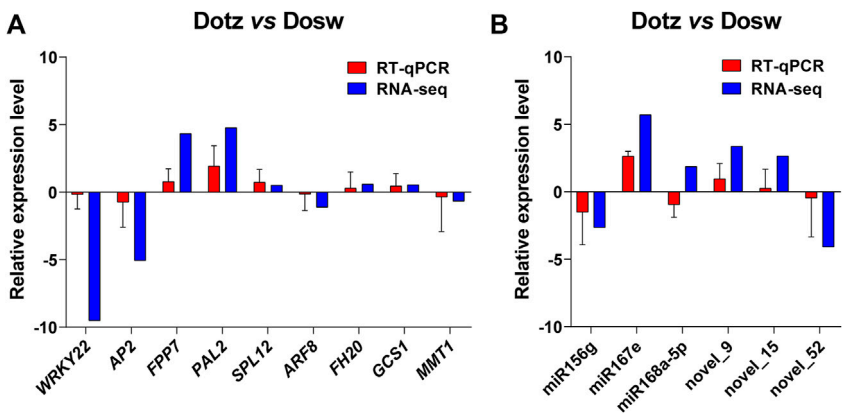


FIGURE 6 | RT-qPCR of the expression levels of miRNAs and genes in Dosw and Dotz from the xylem in *D. odorifera*. The accumulation levels of mRNAs of the genes **(A)**, and miRNAs **(B)** in RT-qPCR and RNA-seq analyses are presented. The *actin* was used as an internal control for genes, and the *U6* was used as the internal control for miRNAs.

novel_9 was predicted to target auxin response factor 8 (*ARF8*) that is involved in auxin dynamics balance (Tian et al., 2004) (**Figure 5**).

Validation of Gene Expression Levels by RT-qPCR

We determined the relative transcript levels of miRNAs and genes identified from DEGs and the interaction network by RT-qPCR, and observed that only miR168a-5p was different with RNA-seq. The expression patterns of all nine genes and other miRNAs estimated by RT-qPCR and RNA-seq followed a similar trend in Dotz and Dosw samples. Although the fold-change (FC) values calculated by sequencing did not exactly match the expression values obtained by RT-qPCR, the expression profiles were basically consistent for all tested miRNAs and genes (**Figure 6**). These analyses validate the reliability of the gene expression values generated from sequencing results.

DISCUSSION

D. odorifera is an important, rare, and precious rosewood. However, only a limited number of studies (Cui et al., 2019) are available that describe the molecular mechanism of wood formation of *D. odorifera*, which is comprised of multiple and complex pathways. RNA-seq technology now provides a convenient tool for a better understanding of the mechanisms of plant growth and development at the molecular level (Yeh et al., 2020). In this study, we explored the potential mechanism of xylem differentiation by investigating the accumulation of miRNAs and mRNAs in different xylem tissues. Through the construction of a regulatory network between differentially accumulating miRNAs and their target genes, we obtained valuable miRNA-mRNA pairs involved in xylem differentiation and identified their interaction and potential role. These data provide a deeper understanding of the molecular mechanism of wood formation in *D. odorifera*.

mRNA Sequencing Analysis

By analyzing the patterns of mRNA accumulations in two tissue types related to wood formation in *D. odorifera*, we identified 395 differentially expressed mRNAs between Dotz and Dosw (**Figure 1A**). Further, the enriched GO terms associated with the DEGs were largely expected. Anion transport and inorganic anion transport were highly enriched, indicating the function of these genes in wood formation. KEGG analysis showed that DEGs were commonly enriched in some pathways, including "carotenoid biosynthesis," "diterpenoid biosynthesis," and "alpha-linolenic acid metabolism". This result is also consistent with our expectations as these three pathways might play an important role in the differentiation of xylem, especially in the biosynthesis of terpenoids and plant hormones (Biesgen and Weiler, 1999; Regnault et al., 2014).

We have identified genes that are possibly involved xylem differentiation process. Eleven genes that may encode transcription factors were identified, eight of which belong to the WRKY family. Our findings on differential expression of many genes encoding transcription factors are supported by previous investigations on other trees, such as rubber tree (Meng et al., 2021) and *Populus tomentosa* (Chen et al., 2015). In *P. tomentosa*, some genes of WRKY family were also differentially accumulated amoung the xylem tissues of tension wood, opposite wood, and normal wood (Chen et al., 2015). Further, earlier studies have shown that WRKY transcription factors were widely involved in plant responses to biotic, abiotic, and hormonal stresses, and regulate the biosynthesis of phenols, terpenes, and alkaloids (Schluttenhofer and Yuan, 2015). For example, *MdWRKY11* promotes the expression of *F3H*, *FLS*, *DFR*, *ANS,* and *10UFGT* in *Malus domestica* callus, and then increases the accumulation of flavonoids and anthocyanins (Wang N et al., 2018). Further, we also found some genes (such as *PER47*, *COMT*, *CCR2*, *UXS2*, *KAO2*, *GA20OX1*, *OPR2*, *CYP74A*, *PAL2*, and *CYP71A1*) related to lignin synthesis and cell wall composition, gibberellin synthesis, jasmonic acid synthesis, and flavonoids and terpenoids synthesis. Similarly, in the study of Sun et al. (2020), seven PAL family genes involved in the phenylpropane pathway were also significantly differentially expressed between the discolored wood by pruning and the normal wood in *D. odorifera*. In addition, *COMT* silenced transgenic poplar trees have significantly reduced 17% lignin levels (Jouanin et al., 2000). The trunk dry mass of a three-month-old *PdGA20ox1* overexpressing transgenic poplar was four times higher than that of untransformed control poplar; the contents of xylose and glucose were also significantly increased in these transgenic plants (Park et al., 2015).

MiRNA Sequencing Analysis

MiRNAs are an important class of non-coding factors that regulate genes expression at the post-transcriptional level in nearly every aspect of plant development such as germination, growth, tissue differentiation, and flowering (Das et al., 2019; Glazińska et al., 2019; Qiu et al., 2019). The results of small RNA-seq demonstrate that the xylem of *D. odorifera* contains a large and diverse small RNA population, a finding similar to previous studies in rubber tree (Meng et al., 2022) and Chinese fir (Wan et al., 2012).

A total of 40 known miRNAs belonging to 22 miRNA families and 123 novel miRNAs were identified in these small RNA libraries, among which 14 miRNAs were differentially expressed (**Figure 4**). Similarly, the miR156, miR159, miR166, miR319, miR396, miR398, and miR408 families have also been identified in the xylem of rubber tree (Meng et al., 2022). The miR396a and miR156g are also differentially expressed in the primary stem, transition stem, and secondary stem of *Populus trichocarpa* (Wang R et al., 2021). Then, we found many miRNA families in *D. odorifera*, for example, miR156 (Wang et al., 2011), miR166 (Chen et al., 2018), and miR397 (Lu et al., 2013), are known to play a role in wood formation in other species. Subsequently, functional analysis of the target genes of the differentially expressed miRNAs revealed that these genes were commonly enriched in a few pathways including, "protein processing in endoplasmic reticulum," "fatty acid elongation" and "biosynthesis of unsaturated fatty acids" These results further suggest that these differentially expressed miRNAs possibly regulate xylem differentiation of *D. odorifera*.

Integration Analysis of Differentially Expressed miRNAs and Target Genes

We predicted 1,267 target-miRNA pairs between the 14 differentially expressed miRNAs and 1,056 candidate genes with the help of psRobot in Dotz and Dosw. Among these miRNA target related gene pairs, 22 showed a significantly negative correlation. Some studies have found that miR156 regulates plant growth and development, morphogenesis, anthocyanin accumulation, gibberellin synthesis, and stress response by inhibiting the expression of SPL transcription factors in *A. thaliana* at the post-transcriptional level (Schwarz et al., 2008; Jung et al., 2011). In our study, miR156g appears to

negatively regulate seven target genes, including *SPL6* and *SPL12*. Novel miRNAs were equally differentially accumulated and significantly negatively correlated to their target genes, thus indicating an important regulatory role in xylem differentiation. In our study, the novel_130 is predicted to target the expression of up-regulated *GCS1* in Dotz. In *A. thaliana*, α-Glucosidase I (encoded by *GCS1*) is required for cellulose biosynthesis and morphogenesis (Gillmor et al., 2002). These results indicate that novel_130 may affect cellulose synthesis in the Dotz of *D. odorifera* by regulating *GCS1*.

CONCLUSION

In summary, mRNA and small RNA profiles were first revealed in the process of xylem differentiation of *D. odorifera*. A total of 395 differentially expressed mRNAs were identified, many of which are involved in diterpenoid biosynthesis and alpha-linolenic acid metabolism, and controlled synthesis of terpenoids. Further, eight genes encoding the WRKY transcription factors, and some genes related to lignin synthesis, cell wall composition, gibberellin synthesis, jasmonic acid synthesis, flavonoids, and terpenoids synthesis (such as *PER47*, *COMT*, *CCR2*, *UXS2*, *KAO2*, *GA20OX1*, *OPR2*, *CYP74A*, *PAL2*, and *CYP71A1*), were also identified. Subsequently, 14 differentially expressed miRNAs between Dotz and Dosw were found, and nine of these were significantly negatively correlated to the expression of 21 target genes. This evidence provides valuable information for further functional characterization of the miRNAs and their targets in the xylem differentiation of *D. odorifera*.

DATA AVAILABILITY STATEMENT

The small RNA and transcriptome data for *D. odorifera* reported in this paper has been deposited at the Genome Sequence Archive (Wang et al., 2017) in BIG Data Center (BIG Data Center Members, 2019), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession numbers CRA006116 and CRA006117, and are publicly available at https://bigd.big.ac.cn/gsa.

AUTHOR CONTRIBUTIONS

JC and BC designed the research; WZ, XM, JX, ZL, YH and BL performed the research. All authors analyzed and interpreted the data; JC and WZ wrote the paper; All authors commented on the manuscript.

FUNDING

This work was supported by the 2019 Hainan Provincial Basic and Applied Research (Natural Science), High-level Talents Program/Hainan Provincial Natural Science Foundation of China (No. 2019RC160), the National Natural Science Foundation of China (31960321), and the Scientific Research Fund Project of Hainan University (KYQD (ZR)1830).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.883422/full#supplementary-material

REFERENCES

Aukerman, M. J., and Sakai, H. (2003). Regulation of Flowering Time and floral Organ Identity by a microRNA and its APETALA2-like Target Genes. *Plant Cell* 15, 2730–2741. doi:10.1105/tpc.016238

Bergstrom, B. (2003). Chemical and Structural Changes during Heartwood Formation in *Pinus Sylvestris*. *Forestry* 76, 45–53. doi:10.1093/forestry/76.1.45

Biesgen, C., and Weiler, E. W. (1999). Structure and Regulation of OPR1 and OPR2 , Two Closely Related Genes Encoding 12-oxophytodienoic Acid-10,11-Reductases from Arabidopsis thaliana. *Planta* 208, 155–165. doi:10.1007/s004250050545

BIG Data Center Members (2019). Database Resources of the BIG Data Center in 2019. *Nucleic Acids Res.* 47, D8–D14. doi:10.1093/nar/gky993

Celedon, J. M., and Bohlmann, J. (2018). An Extended Model of Heartwood Secondary Metabolism Informed by Functional Genomics. *Tree Physiol.* 38, 311–319. doi:10.1093/treephys/tpx070

Chen, B., Du, Q., Chen, J., Yang, X., Tian, J., Li, B., et al. (2016). Dissection of Allelic Interactions Among Pto-miR257 and its Targets and Their Effects on Growth and wood Properties in *Populus*. *Heredity* 117, 73–83. doi:10.1038/hdy.2016.26

Chen, J., Chen, B., and Zhang, D. (2015). Transcript Profiling of *Populus Tomentosa* Genes in normal, Tension, and Opposite wood by RNA-Seq. *BMC Genomics* 16, 1–16. doi:10.1186/s12864-015-1390-y

Chen, S.-Y., Yen, P.-L., Chang, T.-C., Chang, S.-T., Huang, S.-K., and Yeh, T.-F. (2018). Distribution of Living ray Parenchyma Cells and Major Bioactive Compounds during the Heartwood Formation of *Taiwania Cryptomerioides* Hayata. *J. Wood Chem. Technol.* 38, 84–95. doi:10.1080/02773813.2017.1372478

Cui, Z., Li, X., Xu, D., and Yang, Z. (2020). Changes in Non-structural Carbohydrates, wood Properties and Essential Oil during Chemically-Induced Heartwood Formation in *Dalbergia Odorifera*. *Front. Plant Sci.* 11, 1–13. doi:10.3389/fpls.2020.01161

Cui, Z., Yang, Z., and Xu, D. (2019). Synergistic Roles of Biphasic Ethylene and Hydrogen Peroxide in Wound-Induced Vessel Occlusions and Essential Oil Accumulation in *Dalbergia Odorifera*. *Front. Plant Sci.* 10, 1–10. doi:10.3389/fpls.2019.00250

Das, A., Nigam, D., Junaid, A., Tribhuvan, K. U., Kumar, K., Durgesh, K., et al. (2019). Expressivity of the Key Genes Associated with Seed and Pod Development Is Highly Regulated via lncRNAs and miRNAs in Pigeonpea. *Sci. Rep.* 9, 18191. doi:10.1038/s41598-019-54340-6

Dong, Q., Hu, B., and Zhang, C. (2022). MicroRNAs and Their Roles in Plant Development. *Front. Plant Sci.* 13, 824240. doi:10.3389/fpls.2022.824240

Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). MiRDeep2 Accurately Identifies Known and Hundreds of Novel microRNA Genes in Seven Animal Clades. *Nucleic Acids Res.* 40, 37–52. doi:10.1093/nar/gkr688

Gillmor, C. S., Poindexter, P., Lorieau, J., Palcic, M. M., and Somerville, C. (2002). α-Glucosidase I Is Required for Cellulose Biosynthesis and Morphogenesis in Arabidopsis. *J. Cel Biol.* 156, 1003–1013. doi:10.1083/jcb.200111093

Glazińska, P., Kulasek, M., Glinkowski, W., Wojciechowski, W., and Kosiński, J. (2019). Integrated Analysis of Small RNA, Transcriptome and Degradome Sequencing Provides New Insights into floral Development and Abscission in Yellow Lupine (*Lupinus Luteus* L.). *Ijms.* 20, 5122. do:doi:10.3390/ijms20205122

Hong, Z., Li, J., Liu, X., Lian, J., Zhang, N., Yang, Z., et al. (2020). The Chromosome-Level Draft Genome of *Dalbergia Odorifera*. *Gigascience* 9, 1–8. doi:10.1093/gigascience/giaa084

Islam, M. A., Begum, S., Nakaba, S., and Funada, R. (2012). Distribution and Pattern of Availability of Storage Starch and Cell Death of ray Parenchyma Cells of a conifer Tree (*Larix Kaempferi*). *Res. J. Recent Sci.* 1, 28–37.

Jouanin, L., Goujon, T., de Nadaï, V., Martin, M.-T., Mila, I., Vallet, fnm., et al. (2000). Lignification in Transgenic Poplars with Extremely Reduced Caffeic Acid O-Methyltransferase Activity. *Plant Physiol.* 123, 1363–1374. doi:10.1104/pp.123.4.1363

Jung, J.-H., Seo, P. J., Kang, S. K., and Park, C.-M. (2011). MiR172 Signals Are Incorporated into the miR156 Signaling Pathway at the *SPL3/4/5* Genes in *Arabidopsis* Developmental Transitions. *Plant Mol. Biol.* 76, 35–45. doi:10.1007/s11103-011-9759-z

Liao, Y., Zhang, Q., Cui, R., Xu, X., Zhu, F., Cheng, Q., et al. (2021). High-throughput Sequencing Reveals the Regulatory Networks of Transcriptome and Small RNAs during the Defense against *Marssonina Brunnea* in Poplar. *Front. Plant Sci.* 12, 719549. doi:10.3389/fpls.2021.719549

Lim, K.-J., Paasela, T., Harju, A., Venäläinen, M., Paulin, L., Auvinen, P., et al. (2016). Developmental Changes in Scots pine Transcriptome during Heartwood Formation. *Plant Physiol.* 172, 1403–1417. doi:10.1104/pp.16.01082

Liu, F., Hong, Z., Xu, D., Jia, H., Zhang, N., Liu, X., et al. (2019). Genetic Diversity of the Endangered *Dalbergia Odorifera* Revealed by SSR Markers. *Forests* 10, 225. doi:10.3390/f10030225

Liu, X., Xu, D., Yang, Z., and Zhang, N. (2017). Geographic Variations in Seed Germination of *Dalbergia Odorifera* T. Chen in Response to Temperature. *Ind. Crops Prod.* 102, 45–50. doi:10.1016/j.indcrop.2017.03.027

Lu, S., Li, Q., Wei, H., Chang, M.-J., Tunlaya-Anukit, S., Kim, H., et al. (2013). Ptr-miR397a Is a Negative Regulator of Laccase Genes Affecting Lignin Content in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10848–10853. doi:10.1073/pnas.1308936110

Meng, H., Yang, Y., Gao, Z.-H., and Wei, J.-H. (2019). Selection and Validation of Reference Genes for Gene Expression Studies by RT-PCR in *Dalbergia Odorifera*. *Sci. Rep.* 9, 1–10. doi:10.1038/s41598-019-39088-3

Meng, X., Kong, L., Zhang, Y., Wu, M., Wang, Y., Li, J., et al. (2022). Gene Expression Analysis Revealed Hbr-miR396b as a Key Piece Participating in Reaction wood Formation of *Hevea Brasiliensis* (Rubber Tree). *Ind. Crops Prod.* 177, 114460. doi:10.1016/j.indcrop.2021.114460

Meng, X., Wang, Y., Li, J., Jiao, N., Zhang, X., Zhang, Y., et al. (2021). RNA Sequencing Reveals Phenylpropanoid Biosynthesis Genes and Transcription Factors for *Hevea Brasiliensis* Reaction wood Formation. *Front. Genet.* 12, 763841. doi:10.3389/fgene.2021.763841

Mizrachi, E., and Myburg, A. A. (2016). Systems Genetics of wood Formation. *Curr. Opin. Plant Biol.* 30, 94–100. doi:10.1016/j.pbi.2016.02.007

Park, E.-J., Kim, H.-T., Choi, Y.-I., Lee, C., Nguyen, V. P., Jeon, H.-W., et al. (2015). Overexpression of gibberellin *20-oxidase 1* from *Pinus densiflora* results in Enhanced wood Formation with Gelatinous Fiber Development in a Transgenic Hybrid poplar. *Tree Physiol.* 35, tpv099. doi:10.1093/treephys/tpv099

Qiu, L., Chen, R., Fan, Y., Huang, X., Luo, H., Xiong, F., et al. (2019). Integrated mRNA and Small RNA Sequencing Reveals microRNA Regulatory Network Associated with Internode Elongation in Sugarcane (*Saccharum Officinarum* L.). *BMC Genomics* 20, 817. doi:10.1186/s12864-019-6201-4

Regnault, T., Davière, J.-M., Heintz, D., Lange, T., and Achard, P. (2014). The Gibberellin Biosynthetic genes *AtKAO1* and *AtKAO2* have Overlapping Roles throughout *Arabidopsis* Development. *Plant J.* 80, 462–474. doi:10.1111/tpj.12648

Rogers, K., and Chen, X. (2013). Biogenesis, Turnover, and Mode of Action of Plant microRNAs. *The Plant Cell* 25, 2383–2399. doi:10.1105/tpc.113.113159

Schluttenhofer, C., and Yuan, L. (2015). Regulation of Specialized Metabolism by WRKY Transcription Factors. *Plant Physiol.* 167, 295–306. doi:10.1104/pp.114.251769

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing Real-Time PCR Data by the Comparative CT Method. *Nat. Protoc.* 3, 1101–1108. doi:10.1038/nprot.2008.73

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

Shi, Y., Xia, H., Cheng, X., and Zhang, L. (2021). Genome-wide miRNA Analysis and Integrated Network for Flavonoid Biosynthesis in *Osmanthus Fragrans*. *BMC Genomics* 22, 1–11. doi:10.1186/s12864-021-07439-y

Song, X., Li, Y., Cao, X., and Qi, Y. (2019). MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. *Annu. Rev. Plant Biol.* 70, 489–525. doi:10.1146/annurev-arplant-050718-100334

Sun, Y., Gao, M., Kang, S., Yang, C., Meng, H., Yang, Y., et al. (2020). Molecular Mechanism Underlying Mechanical Wounding-Induced Flavonoid Accumulation in *Dalbergia Odorifera* T. Chen, an Endangered Tree that Produces Chinese Rosewood. *Genes* 11, 478. doi:10.3390/genes11050478

Sunkar, R., Li, Y.-F., and Jagadeeswaran, G. (2012). Functions of microRNAs in Plant Stress Responses. *Trends Plant Sci.* 17, 196–203. doi:10.1016/j.tplants.2012.01.010

The, S. N. (2017). A Review on the Medicinal PlantDalbergia odoriferaSpecies: Phytochemistry and Biological Activity. *Evid Based. Complement. Altern. Med.* 2017, 1–27. doi:10.1155/2017/7142370

Tian, C. E., Muto, H., Higuchi, K., Matamura, T., Tatematsu, K., Koshiba, T., et al. (2004). Disruption and Overexpression of *auxin Response Factor 8* gene of *Arabidopsis* affect Hypocotyl Elongation and Root Growth Habit, Indicating its Possible Involvement in Auxin Homeostasis in Light Condition. *Plant J.* 40, 333–343. doi:10.1111/j.1365-313X.2004.02220.x

Vatanparast, M., Klitgård, B. B., Adema, F. A. C. B., Pennington, R. T., Yahara, T., and Kajita, T. (2013). First Molecular Phylogeny of the Pantropical Genus *Dalbergia*: Implications for Infrageneric Circumscription and Biogeography. *South Afr. J. Bot.* 89, 143–149. doi:10.1016/j.sajb.2013.07.001

Wan, L.-C., Wang, F., Guo, X., Lu, S., Qiu, Z., Zhao, Y., et al. (2012). Identification and Characterization of Small Non-coding RNAs from Chinese Fir by High Throughput Sequencing. *BMC Plant Biol.* 12, 1–15. doi:10.1186/1471-2229-12-146

Wang, J.-W., Park, M. Y., Wang, L.-J., Koo, Y., Chen, X.-Y., Weigel, D., et al. (2011). MiRNA Control of Vegetative Phase Change in Trees. *Plos Genet.* 7, e1002012. doi:10.1371/journal.pgen.1002012

Wang, M., Yang, C., Wei, K., Zhao, M., Shen, L., Ji, J., et al. (2021). Temporal Expression Study of miRNAs in the crown Tissues of winter Wheat Grown under Natural Growth Conditions. *BMC Genomics* 22, 1–15. doi:10.1186/s12864-021-08048-5

Wang, N., Liu, W., Zhang, T., Jiang, S., Xu, H., Wang, Y., et al. (2018). Transcriptomic Analysis of Red-Fleshed Apples Reveals the Novel Role of *MdWRKY11* in Flavonoid and Anthocyanin Biosynthesis. *J. Agric. Food Chem.* 66, 7076–7086. doi:10.1021/acs.jafc.8b01273

Wang, R., Reng, M., Tian, S., Liu, C., Cheng, H., Liu, Y., et al. (2021). Transcriptome-wide Identification and Characterization of microRNAs in Diverse Phases of wood Formation in *Populus trichocarpa*. *G3 Genes/Genomes/Genetics* 11, jkab195. doi:10.1093/g3journal/jkab195

Wang, Y.-H., Wu, X.-J., Sun, S., Xing, G.-M., Wang, G.-L., Que, F., et al. (2018). DcC4H and DcPER Are Important in Dynamic Changes of Lignin Content in Carrot Roots under Elevated Carbon Dioxide Stress. *J. Agric. Food Chem.* 66, 8209–8220. doi:10.1021/acs.jafc.8b02068

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: Genome Sequence Archive *. *Genomics, Proteomics & Bioinformatics* 15, 14–18. doi:10.1016/j.gpb.2017.01.001

WangQ.-L., ., Sun, A.-Z., Chen, S.-T., Chen, L.-S., and Guo, F.-Q. (2018). *SPL6* Represses Signalling Outputs of ER Stress in Control of Panicle Cell Death in rice. *Nat. Plants* 4, 280–288. doi:10.1038/s41477-018-0131-z

Wariss, H. M., Yi, T.-S., Wang, H., and Zhang, R. (2017). Characterization of the Complete Chloroplast Genome of *Dalbergia Odorifera* (Leguminosae), a Rare and Critically Endangered Legume Endemic to China. *Conservation Genet. Resour.* 10, 527–530. doi:10.1007/s12686-017-0866-2

Wen, M., Shen, Y., Shi, S., and Tang, T. (2012). miREvo: an Integrative microRNA Evolutionary Analysis Platform for Next-Generation Sequencing Experiments. *BMC Bioinformatics* 13, 140. doi:10.1186/1471-2105-13-140

Wu, H.-J., Ma, Y.-K., Chen, T., Wang, M., and Wang, X.-J. (2012). PsRobot: a Web-Based Plant Small RNA Meta-Analysis Toolbox. *Nucleic Acids Res.* 40, W22–W28. doi:10.1093/nar/gks554

Xiao, L., Quan, M., Du, Q., Chen, J., Xie, J., and Zhang, D. (2017). Allelic Interactions Among Pto-miR475b and its Four Target Genes Potentially Affect Growth and wood Properties in *Populus*. *Front. Plant Sci.* 8, 1–14. doi:10.3389/fpls.2017.01055

Yang, K., Han, H., Li, Y., Ye, J., and Xu, F. (2022). Significance of miRNA in Enhancement of Flavonoid Biosynthesis. *Plant Biol. J.* 24, 217–226. doi:10.1111/plb.13361

Ye, Z.-H., and Zhong, R. (2015). Molecular Control of wood Formation in Trees. *Exbotj* 66, 4119–4131. doi:10.1093/jxb/erv081

Yeh, T. F., Chu, J. H., Liu, L. Y., and Chen, S. Y. (2020). Differential Gene Profiling of the Heartwood Formation Process in *Taiwania Cryptomerioides* Hayata Xylem Tissues. *Ijms* 21, 960. doi:10.3390/ijms21030960

Yu, Y., Jia, T., and Chen, X. (2017). The'how' and 'where' of Plant Micro RNA S. *New Phytol.* 216, 1002–1017. doi:10.1111/nph.14834

Zhao, X., Wang, C., Meng, H., Yu, Z., Yang, M., and Wei, J. (2020). *Dalbergia Odorifera*: A Review of its Traditional Uses, Phytochemistry, Pharmacology, and Quality Control. *J. Ethnopharmacology* 248, 112328. doi:10.1016/j.jep.2019.112328

Zhao, Y., Lin, S., Qiu, Z., Cao, D., Wen, J., Deng, X., et al. (2015). MicroRNA857 Is Involved in the Regulation of Secondary Growth of Vascular Tissues in *Arabidopsis*. *Plant Physiol.* 169, 2539–2552. doi:10.1104/pp.15.01011

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 16 May 2022
doi: 10.3389/fgene.2022.919856

Conservation and Diversity of miR166 Family Members From Highbush Blueberry (*Vaccinium corymbosum*) and Their Potential Functions in Abiotic Stress

Yuening Li[1†], *Xianglong Wang*[1†], *Qingxun Guo*[1], *Xinsheng Zhang*[1], *Lianxia Zhou*[1], *Yang Zhang*[2] *and Chunyu Zhang*[1*]

[1]College of Plant Science, Jilin University, Changchun, China, [2]Helong Forestry Co., Ltd, Changbai Mountain Forest Industry Group, Yanji, China

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Chunlong Li,
Huazhong Agricultural University, China
Dong Meng,
Cornell University, United States

***Correspondence:**
Chunyu Zhang
chunyuzhang1979@126.com

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

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: 14 April 2022
Accepted: 22 April 2022
Published: 16 May 2022

Citation:
Li Y, Wang X, Guo Q, Zhang X, Zhou L, Zhang Y and Zhang C (2022) Conservation and Diversity of miR166 Family Members From Highbush Blueberry (Vaccinium corymbosum) and Their Potential Functions in Abiotic Stress. Front. Genet. 13:919856. doi: 10.3389/fgene.2022.919856

MicroRNA166 (miR166) is highly conserved and has diverse functions across plant species. The highbush blueberry (*Vaccinium corymbosum*) genome is thought to harbor 10 miRNA166 loci (*Vco-miR166*), but the extent of their evolutionary conservation or functional diversification remains unknown. In this study, we identified six additional *Vco-miR166* loci based on conserved features of the miR166 family. Phylogenetic analyses showed that mature Vco-miR166s and their precursor cluster in several clades are evolutionary conserved with diverse species. The *cis*-regulatory elements in the *Vco-miR166* promoters indicated functions related to different phytohormones and defense responses. We also identified putative targets of vco-miR166s, which targeted the same gene families, suggesting the functional conservation and diversification of *Vco-miR166* family members. Furthermore, we examined the accumulation patterns of six mature Vco-miR166s in response to abiotic stresses by stem-loop reverse RT-qPCR, which revealed their upregulation under freezing, cold, and heat stress, while they were downregulated by drought compared to control growth conditions. However, *Vco-miR166* members showed different expression patterns when exposed to salt stress. These results showed that conserved Vco-miR166 family members display functional diversification but also coordinately influence plant responses to abiotic stress.

Keywords: miR166, highbush blueberry, evolutionary conservation, functional diversification, abiotic stress, target genes

INTRODUCTION

MicroRNAs (miRNAs) are small [20–24 nucleotides (nt)] non-coding regulatory RNAs derived from longer precursor (pre-miRNA) molecules with stem-loop structures. The pre-miRNA is processed by DICER LIKE1 (DCL1) into the mature miRNA, which then interacts with its target transcripts. Mature miRNAs mediate almost all plant cellular and metabolic processes by regulating gene expression at the posttranscriptional level (Thakur et al., 2011; Fang and Wang, 2021).

In plants, miR166 is a highly conserved miRNA implicated in a wide range of cellular and physiological processes by regulating its cognate target genes encoding members from the homeodomain leucine zipper (HD-ZIP III) family of transcription factors (Miyashima et al., 2013; Singh et al., 2014; Yadav et al., 2021). Recent evidence showed that the miR166 family plays crucial roles in response to abiotic stress. In mulberry (*Morus multicaulis*), miR166f targets the transcripts encoding an HD-ZIP III transcription factor whose expression is induced by drought stress and is a positive regulator of drought stress tolerance (Li R. et al., 2018). In Arabidopsis (*Arabidopsis thaliana*), multiple miR166 members accumulate when seedlings are exposed to low temperature (Zhou et al., 2008). In potato (*Solanum tuberosum*), the levels of miR166 family members were downregulated by 4°C treatment but were upregulated by salt stress (Kitazumi et al., 2015; Ou et al., 2015). In barley (*Hordeum vulgare*), miR166 was induced in leaves upon dehydration stress (Kantar et al., 2010). In wild emmer wheat (*Triticum turgidum*), miR166 was downregulated by drought stress (Kantar et al., 2011). In soybean (*Glycine max*), miR166 members play an important role in response to cold, drought, and salinity stress (Li et al., 2017).

Blueberries (*Vaccinium* spp.) are an economically important small fruit crop. The main distribution of commercial blueberries from northern to southern China comprises lowbush blueberry (*V. angustifolium*), highbush blueberry (*V. corymbosum*), and rabbiteye blueberry (*V. ashei*). The complement of miRNAs has been identified in blueberries using high-throughput sequencing. For example, 412 conserved miRNAs belonging to 29 families were identified in fruits from rabbiteye blueberry (Yue et al., 2017); similarly, 127 known and 101 novel miRNAs were identified from flowers, young fruits, and ripe fruits of rabbiteye blueberry (Li G. et al., 2018). In highbush blueberry, 84 known miRNAs and 16 novel miRNAs were identified from fruits, with 10 miRNAs from miR166 family (Hou et al., 2017). However, the function of miR166s in highbush blueberry fruits and the genomic sequence of *miR166* loci are unknown. In particular, the response of the highbush blueberry *miR166*s to abiotic stress has not been reported.

As mature miRNAs are highly conserved within the same family, sequence prediction allows the identification of multiple members of a given miRNA family, including miR166 (Li et al., 2017). The sequence and assembly of the highbush blueberry genome were recently completed, opening the doors to a systematic identification of miRNAs and their precursors in this species (Colle et al., 2019). In this study, we predicted the sequences of miRNA166 family members, their precursors, and their encoding genes by mining the highbush blueberry genome sequence. We demonstrated the evolutionary conservation and functional diversification of this family based on a sequence analysis of mature Vco-miR166s, *Vco-miR166* loci, and *Vco-miR166* promoters and predicted vco-miR166 target genes. We also examined the accumulation patterns of six Vco-miR166s in response to abiotic stresses. The results help elucidate the functions of the miR166 family in highbush blueberry.

MATERIALS AND METHODS

Plant Materials and Stress Treatments

One-year-old plants from the highbush blueberry cultivar 'Bluecrop', grown in a growth chamber, were used for salt, drought, freezing, cold, and heat treatments. Plants were also maintained in tissue culture on modified woody plant medium with Murashige and Skoog (MS) vitamins, 1.0 mg/L *trans*-zeatin, and 3% (w/v) sucrose. Plants were subcultured every 28 d and were used for drought stress. All plants used for stress treatments were grown at 25°C under a 16-h-light/8-h-dark photoperiod. Plants were exposed to −5°C for 0, 0.5, 1, 2, 4, or 6 h for freezing treatment; 4°C for 0, 2, 4, 6, 9, 12, for 24 h as cold treatment; and 40°C for 0, 0.5, 1, 1.5, 2, and 3 h as heat treatment. Plants were watered with a 200 mM NaCl solution for 0, 1, 2, 6, 9, 12, or 24 h as salt treatment. For drought treatment, plants in tissue culture were transferred to medium (overlaid with 20% (w/v) PEG-6000 for 1 day, and then PEG-6000 was removed) for 0, 1, 2, 6, 9, 12, and 24 h. For all treatments, the 0-h treatment served as the control. Leaves were collected and immediately frozen in liquid nitrogen for microRNA extraction.

Identification of Novel MiR166s in Highbush Blueberry

The precursor sequences of miR166 from sweet orange (*Citrus sinensis*), grape (*Vitis vinifera*), and apple (*Malus domestica*) were downloaded from the miRbase database (https://www.mirbase.org) and used as a query against the *V. corymbosum* cv. Draper v1.0 genome scaffolds database (https://www.vaccinium.org/). Homologs were screened as previously described (Li et al., 2017; Hou et al., 2017). The secondary structures and the negative minimal folding free energy (MFE) of each miRNA were assessed and computed using the RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) (Mathews et al., 2004; Gruber et al., 2008). The minimal folding free energy index (MFEI) was calculated with the following equation: MFEI = [(MFE/length of the RNA sequence) *100]/(G + C)%. The candidates were then confirmed by comparison with miR166s from other species using the NCBI database (https://www.ncbi.nlm.nih.gov). The mature miR166s of highbush blueberry were predicted by searching against the miRbase database using the precursor sequences of miR166 of highbush blueberry.

Phylogenetic Analysis and Sequences Alignment

Phylogenetic trees were separately constructed using the UPGMA method with the MEGA X program from miR166 precursors and mature miR166s from highbush blueberry (Vco-miR166s), grape (Vvi-miR166s), sweet orange (Csi-miR166s), and apple (Mdm-miR166s) (Kumar et al., 2018). Sequence alignments were performed using DNAMAN software, version 8.192.

TABLE 1 | Characteristics of three newly identified miR166 precursors in highbush blueberry.

Locus ID	Length (nt)	NM[a]	MFE[b] kcal/mol	MFEI[c]	Sequence region on Draper genome	Scaffold_ID on Draper genome	Sequence comparison with similar species in NCBI			
							Accession	Species	Name	Identity (%)
Vco-miR166g	107	3	−45.30	0.85	9,638,758–9,638,864	VaccDscaff1	KT004773	*Camellia sinensis*	*Csi-miR166g*	90
Vco-miR166h	114	3	−51.70	1.08	37,329,235–37,329,348	VaccDscaff19	NR_127759	*Vitis vinifera*	*Vvi-miR166g*	90
Vco-miR166i	186	4	−67.9	0.89	28,809,023–28,809,208	VaccDscaff33	NR_107,986	*Solanum lycopersicum*	*Sly-miR166b*	63

[a]*NM, number of mismatches between predicted 3′ and 5′ arms of the miRNA.*
[b]*Minimal folding free energy.*
[c]*Minimal folding free energy index.*

Promoter Analysis of *Vco-miR166*s

The upstream sequences of *Vco-miR166*s were retrieved from the highbush blueberry genome (https://www.vaccinium.org/). The transcription start sites (TSSs) and TATA-box of *Vco-miR166*s were predicted by TSSP software in Softberry (http://linux1.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter). About 1,600 bp of sequence upstream of the TSS was analyzed for the presence of known *cis*-acting regulatory sequences held in the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Prediction and Functional Analysis of Vco-miR166 Targets

The target sequences of highbush blueberry Vco-miR166s were predicted with the psRNATarget program (http://plantgrn.noble.org/psRNATarget/) against American cranberry (*Vaccinium macrocarpon*) transcripts (Polashock et al., 2014; Dai et al., 2018). The obtained target sequences were then used to query target genes in the *V. corymbosum* cv. Draper v1.0 genome transcript database (Colle et al., 2019). The functions of Vco-miR166s target genes were predicted by BLAST against the Swiss-Prot database (https://ww.uniprot.org). Genetic similarity and genetic distance were calculated by DICE and Nei's similarity coefficient, respectively. The phylogenetic tree was constructed using the UPGMA method with NTSYSpc-2.11 F software.

Reverse-Transcription Quantitative PCR (RT-qPCR)

The microRNAs were extracted using an EASYspin Plant microRNA rapid extraction kit (Aidlab, Beijing, China) from control and abiotic stress samples. Reverse transcription and qPCR were performed with the Mir-X miRNA qRT-PCR TB Green Kit (Takara Inc., Dalian, China) using specific stem-loop RT and forward RT-qPCR primers. qPCR was conducted using the ABI StepOnePlus Real-Time PCR System using *U6* as the reference. Primer sequences are listed in **Supplementary Table S1**.

RESULTS

Identification and Characterization of Vco-miR166s in Highbush Blueberry

Based on the conserved features shared by pre-miR166s sequences from sweet orange (*Citrus sinensis*), grape (*Vitis vinifera*), and apple (*Malus domestica*), we performed a BLAST search of the highbush blueberry genome. We identified three new pre-miR166s, designated here Vco-miR166g, h, and i, following a previous report (Hou et al., 2017) (**Table 1**). The length of these three new pre-miRNAs ranged from 107 to 186 nt, with three to four predicted mismatches between the 5′ and 3′ arms of the miRNA (**Supplementary Figure S1**). The minimal folding free energy (MFE) of these new mature Vco-miRNAs was lower (from −67.9 to −43.30) and their MFE index (MFEI) (from 0.85 to 1.08) was higher than those of transfer RNAs (tRNAs), ribosomal RNAs (rRNA), and messengers RNAs (mRNAs). These results suggest that these predicted precursors may be *bona fide* precursors of highbush blueberry miRNAs.

BLAST analysis determined that Vco-miR166g, Vco-miR166h, and Vco-miR166i precursors share 90%, 90%, and 63% identity with Csi-miR166g, Vvi-MIR166g, and Sly-miR166 b precursors, respectively. We thus concluded that these precursors likely belong to the miR166 family of miRNAs (**Table 1**). At the same time, these three precursors from highbush blueberry showed conserved features in the 3′ arm and differ by several bases in the 5′ arm with other plant species (**Supplementary Figure S2**).

From the three pre-miR166s above, we predicted six putative mature miR166 sequences (**Table 2**), located along the 3′ or 5′ arm of the secondary stem-loop structure of each pre-miR166 sequence by searching against the miRbase database. We designated these candidate miRNAs as Vco-miR166g-3p, Vco-miR166g-5p, Vco-miR166h-3p, Vco-miR166h-5p, Vco-miR166i-3p, and Vco-miR166i-5p. We noticed that the sequence of Vco-miR166i-3p was identical to that of Vco-miR166b-3p, while Vco-miR166i-5p was identical to Vco-miR166d-5p. In addition, BLAST analysis showed that Vco-miR166g-3p, Vco-miR166h-3p, and Vco-miR166i-3p are identical in sequence to known mature miR166s from other

TABLE 2 | Characteristics of six newly identified mature miR166s in highbush blueberry.

ID	Sequence	Length (nt)	Homologs	Species	NM[a]
Vco-miR166g-3p	UCUCGGACCAGG CUUCAUUCC	21	Gma-miR166h-3p, Csi-miR166 b,d,g-3p	*Glycine max*, Citrus sinensis	0
Vco-miR166g-5p	GGGAAUGCUGUC UGGUUCGAG	21	Csi-miR166c-5p	Citrus sinensis	1
Vco-miR166h-3p	AUUUCGGACCAG GCUUCAUUCC	22	Lja-miR166-3p	Lotus japonicus	0
Vco-miR166h-5p	GGAAUGUUGUCU GGUUCGAGA	21	Csi-miR166c-5p, Zma-miR166g-5p, Bdi-miR166d-5p, Ata-miR166e-5p, Vca-miR166a-5p	Citrus sinensis, *Brachypodium distachyon*, Aegilops tauschii, *Vriesea carinata*	1
Vco-miR166i-3p	UCGGACCAGGCU UCAUUCCCC	21	Bdi-miR166 b,c,d,i-3p, Gma-miR166c,i-3p, Stu-miR166a,c,d,h-3p, Aly-miR166h-3p, Ata-miR166a,b,d,e-3p, Vca-miR166a,b,c-3p, Eun-miR166-3p, Fve-miR166d-3p, Cas-miR166c,f-3p	*Brachypodium distachyon*, *Glycine max*, *Solanum tuberosum*, *Arabidopsis lyrata*, Aegilops tauschii, *Vriesea carinata*, *Eugenia uniflora*, *Fragaria vesca*, *Camelina sativa*	0
Vco-miR166i-5p	GGGAUGUUGUCU GGCUCGAUG	21	Cly-miR166c-5p, Zma-miR166c-5p, Csi-miR166a,e,f-5p, Aly-miR166a,c,d-5p, Gma-miR166a,c-5p, Osa-miR166d-5p, Mtr-miR166g-5p, Bdi-miR166e-5p, Stu-miR166a-5p, Vca-miR166b-5p, Eun-miR166-5p	Solanum lycopersicum, Zea mays, Citrus sinensis, *Arabidopsis lyrata*, *Glycine max*, Medicago truncatula, *Brachypodium distachyon*, *Solanum tuberosum*, Vriesea carinata, Eugenia uniflora	2

[a]*Number of mismatches between predicted Vco-miR166s, and their homologs.*

species, with the number of mismatches ranging between one (for Vco-miR166g-5p and Vco-miR166h-5p) and two (for Vco-miR166i-5p). These results thus indicated that these new Vco-miR166 candidates do in fact belong to the miR166 family and that the 3′ arm sequences are more conserved than the 5′ arm.

Conservation and Diversification of the Vco-miR166 Family

We also aligned the sequences of pre-miR166s from highbush blueberry and other major fruit trees (sweet orange, grape, and apple) and produced the corresponding phylogenetic tree (**Figures 1**, **2**). We observed the highest degree of conservation along the arms of the secondary stem-loop structure, especially on the 3′ arm and in the core region of the mature miRNA (GGA CCAGGCTTCaTTCC), with more variation flanking the mature miRNA (**Figure 1**). The nine Vco-miR166s roughly clustered into clades I and II but largely did not cluster together within each clade (**Figure 2A**). These results supported the evolutionary conservation of miR166 family members across various plant species.

We repeated the phylogenetic analysis separately on the 3′ and 5′ arms of nine mature highbush blueberry miR166 sequences along with related sequences from sweet orange (**Figures 2B,C**). For the analysis of the 3′ arm, Vco-miR166i-3p and Vco-miR166b-3p were identical to Csi-miR166a, c, d, and e-3p, while Vco-miR166g-3p was identical to Csi-miR166b, g, and f-3p in clade I. Eight Vco-miR166s and seven Csi-miR166s clustered into clade I. For the 5′ arm of miR166 sequences, eight Vco-miR166s and eight Csi-miR166s formed four clades, with only clade IV not including a Vco-miR166 member. These results supported the notion that the miR166 family is highly conserved in diverse species. Notably, Vco-miR166e-3p appeared to be distant from other Vco-miR166s based on its position in the phylogenetic tree, indicating the evolutionary diversification of Vco-miR166 members.

Putative *cis*-Regulatory Elements of *Vco-miR166* Promoters

To explore the potential functions of Vco-miR166s, we analyzed the putative *cis*-regulatory elements contained within 1,600 bp of the *Vco-miR166* promoter sequence upstream of the TSS (+1 bp) (**Supplementary Table S2**). We identified core promoter elements (TATA-box), common *cis*-acting elements, enhancer regions (CAAT-box), as well as light-, phytohormone-, and defense-related elements in all *Vco-miR166* promoter regions. Several *Vco-miR166* promoters also contained elements related to the cell cycle and meristem expression, seed development, synthesis, and metabolism (**Supplementary Table S3**). **Table 3** summarizes the phytohormone- and defense-related elements in *Vco-miR166* promoters. For plant hormone–related elements, 77.8% (7/9) of *Vco-miR166* promoters harbored gibberellin- and abscisic acid–responsive elements, with 66.7% (6/9) of *Vco-miR166* promoters also presenting elements involved in methyl jasmonate and auxin responsiveness. By contrast, only *Vco-miR166a* and *Vco-miR166h* promoters contained elements associated with salicylic acid responsiveness. For defense-related elements, we detected anaerobic induction elements in all *Vco-miR166* promoter regions, while drought inducibility, low-temperature, and defense- and stress-responsive elements were present in four, three, and two of the nine *Vco-miR166* promoters, respectively. These results indicated that *Vco-miR166* loci may be involved in responses to different phytohormones and defense, but also reflected their potential functional diversification.

Prediction of Vco-miR166 Target Genes

To obtain a better understanding of Vco-miR166 functions, we predicted their targets with the psRNATarget program against the transcript database for American cranberry (*V. macrocarpon*) transcripts, another *Vaccinium* species, because there is no highbush blueberry database in the psRNATarget program (**Supplementary Table S4**). We then used the American

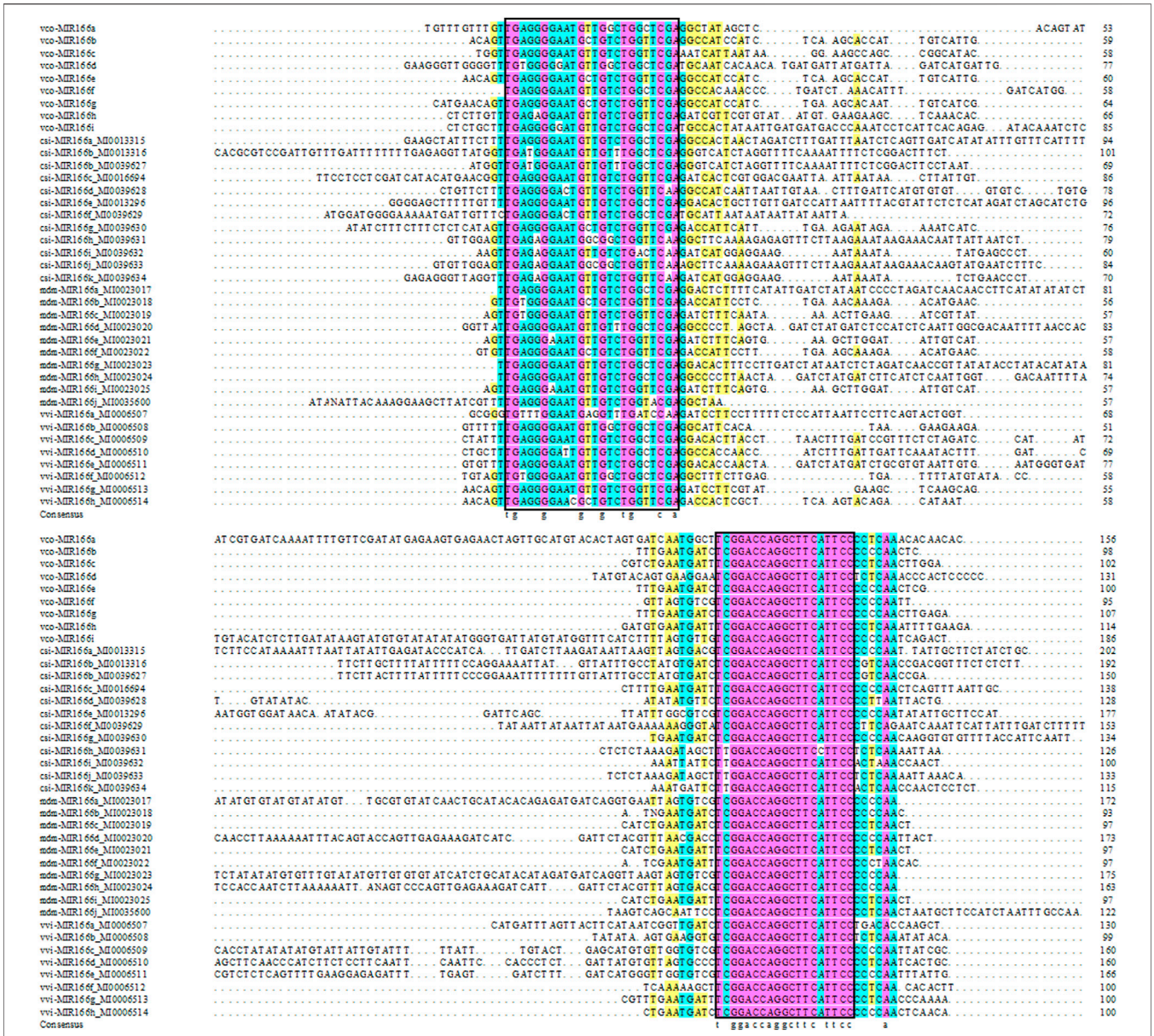


FIGURE 1 | Sequence alignment of miR166 precursors. Vco, *Vaccinium corymbosum*; Csi, *Citrus sinensis*; Vvi, *Vitis vinifera*; Mdm, *Malus domestic*. Pink, 100% conservation; green, ≥75% conservation; yellow, ≥50% conservation. The core region of the mature miRNA is indicated in the black box.

cranberry targets to search for related sequences in the highbush blueberry (*V. corymbosum* cv. Draper v1.0) transcripts (**Supplementary Table S5**). All Vco-miR166s were predicted to target the transcripts of a gene encoding a pentatricopeptide repeat–containing protein (**Supplementary Table S6**). In addition, all Vco-miR166s except Vco-miR166h-3p were predicted to target the transcripts encoding a leucine-rich repeat (LRR) receptor-like serine/threonine-protein kinase. Many Vco-miR166s-3p were predicted to regulate genes encoding homeobox-leucine zipper proteins (HD-ZIP III), while many Vco-miR166s-5p were predicted to target the transcripts for genes encoding phosphatidylinositol/phosphatidylcholine transfer proteins and patellin. These results supported the functional conservation of Vco-miR166 family members. Several Vco-miR166s involved in abiotic stresses, including heat, stress response, and cold, were predicted to target transcripts for cold-regulated proteins, while other Vco-miR166s appeared to target the transcripts encoding many protein families. For example, Vco-miR166a,b,i-3p targeted ethylene-responsive and basic helix-loop-helix (bHLH) transcription factors, Vco-miR166c-3p targeted WD repeat–containing protein, Vco-miR166d-3p targeted bHLH transcription factors, and Vco-miR166e,g-3p targeted WRKY transcription factors. These results illustrated the functional diversification of Vco-miR166 family members in highbush blueberry.

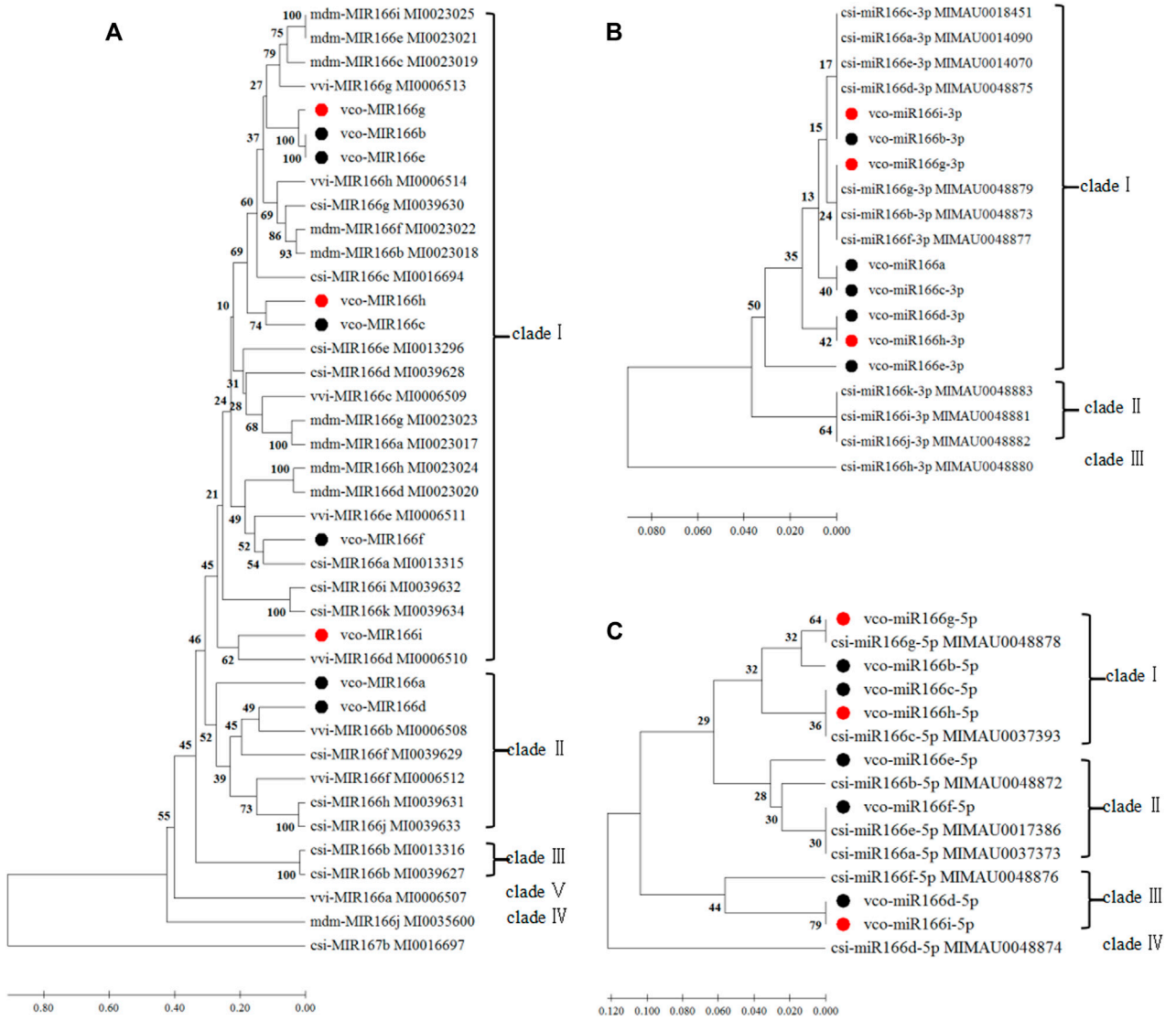


FIGURE 2 | Phylogenetic analyses of *miR166* loci and their predicted 3′ and 5′ arms. Sequences were aligned in MEGA X, and the phylogenetic tree was constructed with the UPGMA method **(A)** Phylogenetic analysis of Vco-miR166 precursors and related sequences from Csi, *Citrus sinensis*; Vvi, *Vitis vinifera* ; and Mdm, *Malus domestica*. Csi-MIR167 b served as the outgroup **(B,C)** Phylogenetic analysis of mature Vco-miR166s and Csi-miR166s for the 3′ arm **(B)** and 5′ arm **(C)**. Black and red dots indicate mature miR166s or their precursors of highbush blueberry from a previous study (Hou et al., 2017) and this study, respectively. The numbers at the nodes indicate the percentage of support from 1,000 bootstrap replicates. The scale indicates the average number of amino acid substitutions per site.

To explore the diversity of Vco-miR166 functions, we calculated the genetic similarity and genetic distance according to the presence or absence of functions of their target genes in **Supplementary Table S6**. The genetic similarity was the highest (0.88) and the genetic distance was the closest (0.12) between Vco-miR166a-3p and Vco-miR166b/i-3p, whereas Vco-miR166b/i-3p and Vco-miR166b-5p exhibited the lowest genetic similarity (0.05) and the farthest genetic distance (3.03) (**Table 4**). The genetic similarity between Vco-miR166h-3p and Vco-miR166i-3p was also low at only 0.12. In fact, a phylogenetic tree according to the genetic similarity between miR166 targets showed that Vco-miR166s are divided into two clades, with Vco-miR166s derived from the same arm clustering together (**Supplementary Figure S3**). These results indicated that the functions of Vco-miR166s are diverse and that Vco-miR166s from the same arm exert similar regulatory functions.

Accumulation of Vco-miR166s Under Abiotic Stress

To explore how Vco-miR166s are related to abiotic stress, we selected six Vco-miR166s to examine their abundance in response to abiotic stresses (**Figure 3**). The six Vco-miR166s had similar accumulation patterns under freezing stress, as they all rapidly increased in abundance, with Vco-miR166h-3p and

TABLE 3 | Putative *cis*-regulatory elements in upstream sequences of blueberry *Vco-miR166*s.

Vco-miR166	Phytohormone-related elements					Defense-related elements			
	Gibberellin responsive	Methyl jasmonate responsiveness	Salicylic acid responsiveness	Auxin responsiveness	Abscisic acid responsiveness	Defense and stress responsiveness	Drought inducibility	Low-temperature responsiveness	Anaerobic induction
Vco-miR166a	GARE-motif, P-box	--	TCA-element	--	--	--	MBS	LTR	ARE
Vco-miR166b	P-box	TGACG-motif, CGTCA-motif	--	TGA-box, TGA-element	ABRE	--	--	--	ARE
Vco-miR166c	--	--	--	AuxRR-core	ABRE	--	--	--	ARE
Vco-miR166d	P-box	--	--	TGA-box	ABRE	TC-rich	MBS	--	ARE
Vco-miR166e	P-box	TGACG-motif, CGTCA-motif	--	TGA-box, TGA-element	ABRE	--	--	--	ARE
Vco-miR166f	GARE-motif	TGACG-motif, CGTCA-motif	--	--	ABRE	--	MBS	--	ARE
Vco-miR166g	GARE-motif	TGACG-motif, CGTCA-motif	--	TGA-element, AuxRR-core	ABRE	--	--	LTR	ARE
Vco-miR166h	--	TGACG-motif, CGTCA-motif	TCA-element	AuxRR-core	ABRE	--	MBS	LTR	ARE
Vco-miR166i	GARE-motif	TGACG-motif, CGTCA-motif	--	--	--	TC-rich	--	--	ARE

TC-rich: cis-acting element involved in defense and stress responsiveness; MBS: MYB, binding site involved in drought-inducibility; LTR: cis-acting element involved in low-temperature responsiveness; ARE: cis-acting regulatory element essential for the anaerobic induction.

Vco-miR166i-3p increasing in abundance by 18.6- and 28.6-fold upon 2 h of freezing stress compared to the control at 0 h. The abundance of the six Vco-miR166 transcripts sharply declined and showed no accumulation after 6 h of cold exposure in 1-year-old blueberry plants, which died after 6 h of –5°C treatment (**Figure 3A**). These results suggested that these six Vco-miR166s, especially Vco-miR166h-3p and Vco-miR166i-3p, may be involved in freezing response and as negative regulators.

In blueberry plants exposed to cold stress, the abundance of Vco-miR166a, Vco-miR166e-3p, and Vco-miR166i-3p reached their maximum levels after 9 h of freezing stress, corresponding to an increase of 4.1-, 10.2-, and 3.9-fold compared to the control at 0 h, respectively, followed by a sharp decrease to control levels for Vco-miR166a and Vco-miR166e-3p or to lower levels than control for Vcc-miR166i-3p. The abundance of Vco-miR166d-5p, Vco-miR166f-5p, and Vco-miR166h-3p was upregulated and reached its highest levels at 24, 4, and 12 h during freezing stress, increasing 12.8-, 15.4-, and 8.5-fold compared to the control samples, respectively (**Figure 3B**). Thus, the Vco-miR166 family appeared to be involved in cold stress responses.

When blueberry plants were exposed to heat stress, Vco-miR166a precursor abundance was upregulated and increased by 38.1-fold at 2 h compared to the control samples at 0 h, followed by a sharp decrease at 3 h. Vco-miR166d-5p and Vco-miR166f-5p showed similar accumulation patterns, with their levels reaching their peak at 2 h and their lowest levels at 1.5 and 3 h. The accumulation of Vco-miR166e-3p, Vco-miR166h-3p, and Vco-miR166i-3p was also upregulated and reached its highest levels at 0.5 h for Vco-miR166h-3p, at 1 and 1.5 h for Vco-miR166e-3p, and at 0.5 and 2 h for Vco-miR166i-3p. All six miRNAs showed a rapid decrease in abundance at 3 h (**Figure 3C**). Therefore, Vco-miR166 family members might respond during early stages of heat stress.

Under drought stress, six Vco-miR166 precursors showed similar trends in their abundance, with a sharp decrease and reaching their lowest levels at 12 and 72 h and a modest increase in their levels at 24 and 48 h (**Figure 3D**). Vco-miR166 members might thus promote the accumulation of their target transcripts and play a positive role in stress adaptation.

In the context of salt stress, Vco-miR166a transcript levels were upregulated after 1 h and then rapidly decreased, with the lowest abundance after 12 h. We observed similar patterns for Vco-miR166d-5p, whose transcript levels gradually increased by 6.9- and 6.0-fold at 6 and 9 h compared to the 0 h control and then rapidly increased until 24 h. Vco-miR166h-3p also increased in abundance before decreasing, but increased rapidly at 24 h. On the contrary, the transcript levels of Vco-miR166e-3p, Vco-miR166f-5p, and Vco-miR166i-3p were downregulated compared to the control under salt stress, especially Vco-miR166e-3p, whose transcript levels rapidly decreased at 1 h and then remained low after salt stress (**Figure 3E**). These results suggested that these Vco-miR166 members might respond to salt stress at an early stage.

TABLE 4 | Genetic distance and genetic similarity between Vco-miR166s from highbush blueberry according to the functions of target genes.

	Vco-miR166a	Vco-miR166b/i-3p	Vco-miR166c-3p	Vco-miR166d-3p	Vco-miR166e-3p	Vco-miR166g-3p	Vco-miR166h-3p	Vco-miR166b-5p	Vco-miR166c-5p	Vco-miR166d/i-5	Vco-miR166e-5p	Vco-miR166f-5p	Vco-miR166g-5p	Vco-miR166h-5p
Vco-miR166a		0.12	1.27	0.31	1.23	1.66	1.28	2.91	2.28	2.56	2.85	2.63	2.83	2.57
Vco-miR166b/i-3	0.88		1.28	0.24	1.15	1.67	1.30	3.03	2.40	2.68	2.97	2.76	2.95	2.70
Vco-miR166c-3p	0.28	0.27		1.02	1.61	1.92	1.66	2.48	2.25	2.94	2.41	2.60	2.40	2.03
Vco-miR166d-3p	0.73	0.78	0.36		1.19	1.62	1.04	2.86	2.01	2.52	2.80	2.59	2.79	2.24
Vco-miR166e-3p	0.29	0.31	0.20	0.31		0.59	1.81	2.15	1.65	1.99	2.09	1.72	2.36	2.22
Vco-miR166g-3p	0.18	0.19	0.14	0.19	0.53		2.12	1.37	1.61	1.96	1.71	1.49	2.16	1.62
Vco-miR166h-3p	0.28	0.27	0.19	0.35	0.16	0.12		2.90	2.26	2.55	2.83	2.33	2.82	2.56
Vco-miR166b-5p	0.05	0.05	0.08	0.06	0.11	0.25	0.05		1.57	1.57	0.82	0.80	0.74	1.01
Vco-miR166c-5p	0.10	0.09	0.10	0.13	0.19	0.20	0.10	0.21		1.53	1.91	1.33	1.90	1.13
Vco-miR166d/i-5	0.08	0.07	0.05	0.08	0.13	0.14	0.08	0.21	0.22		1.51	0.96	2.08	1.44
Vco-miR166e-5p	0.06	0.05	0.09	0.06	0.12	0.18	0.06	0.44	0.15	0.22		0.89	1.49	1.52
Vco-miR166f-5p	0.07	0.06	0.07	0.07	0.17	0.22	0.09	0.45	0.26	0.38	0.40		1.28	0.97
Vco-miR166g-5p	0.06	0.05	0.09	0.06	0.09	0.11	0.06	0.47	0.15	0.12	0.23	0.27		1.22
Vco-miR166h-5p	0.08	0.07	0.13	0.10	0.11	0.20	0.08	0.36	0.32	0.24	0.22	0.38	0.29	

Genetic distances are indicated in the upper right triangle; genetic similarity is shown in the lower left triangle.

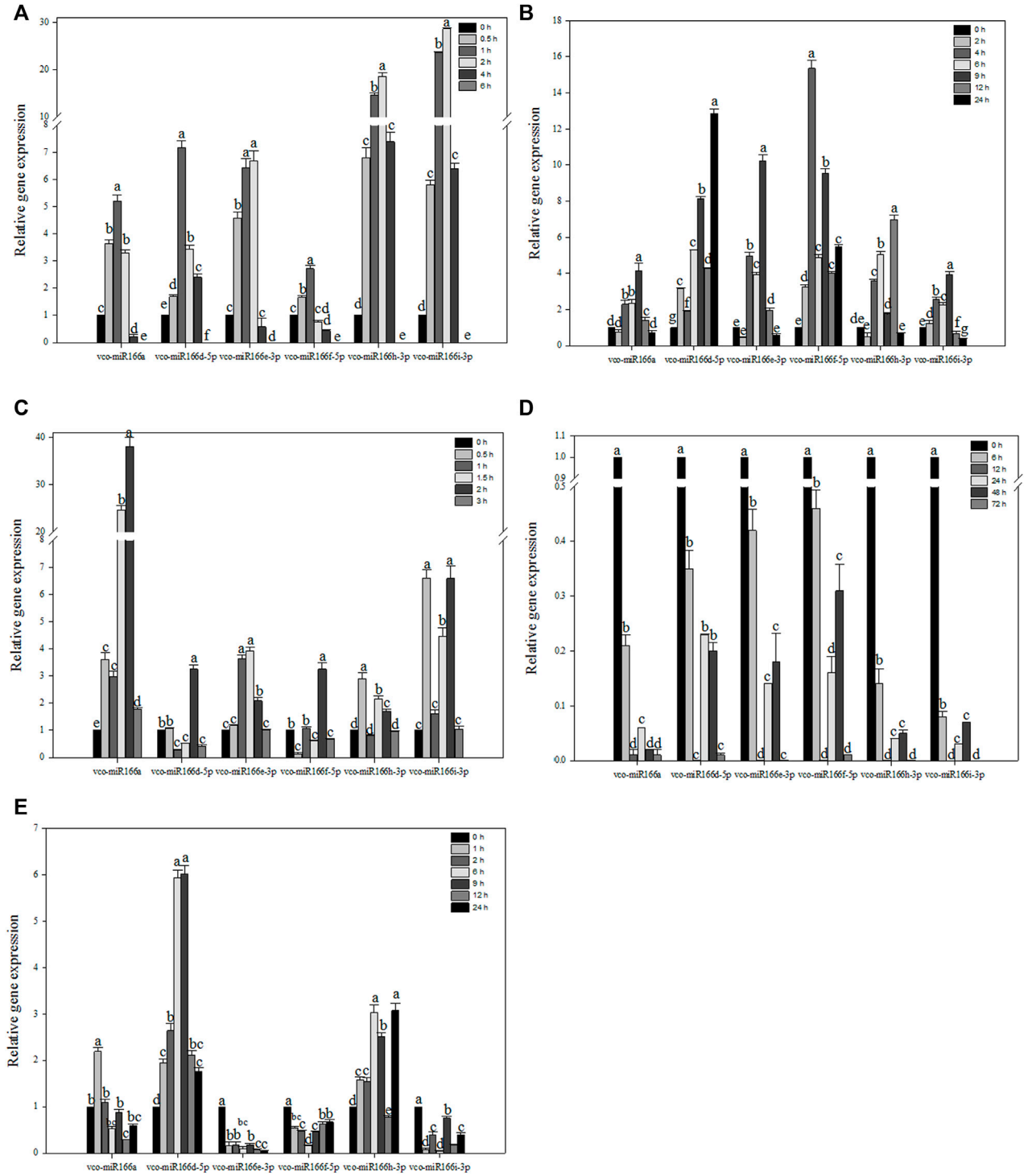


FIGURE 3 | RT-qPCR analysis of Vco-miR166 abundance under various stress conditions. Highbush blueberry plants were exposed to freezing **(A)**, cold **(B)**, heat **(C)**, drought **(D)**, and salt **(E)** stress. Data represent the means ± standard deviation (SD) of three independent biological replicates, each with three technical replicates. Different letters indicate significant differences ($p < 0.05$) between samples, as determined by Tukey's test.

DISCUSSION

miRNAs play various roles in regulating almost all processes of plant growth and in metabolism. Several conserved miRNAs were reported in highbush blueberry based on an analysis of expressed sequence tags (ESTs) (Li et al., 2014), leading to the identification of 10 miR166 family members using high-throughput sequencing of small RNAs and reference sequences from the American cranberry genome, and ESTs and transcriptome data from blueberries during fruit maturation (Hou et al., 2017). The completion of the highbush blueberry genome (Gupta et al., 2015) allowed us to identify three new *Vco-pre-miR166* members and determine their mature miR166s based on the *V. corymbosum* cv. Draper v1.0 genome. We determined here that the newly identified pre-miR166s exhibit low MFE and high MFEI values and form stable secondary hairpin structures (**Table 1** and **Supplementary Figure S1**). In addition, we showed that the mature Vco-miR166 sequences are highly identical to other miR166 family members from other species (**Table 2**). We thus concluded that the three newly identified *Vco-miR166* loci belong to the miR166 family.

The conserved nature of *miR166* family members has been demonstrated in many plant species (Li et al., 2017; Barik et al., 2014). In this study, we showed that the newly identified mature Vco-miR166s derived from the 3′ and 5′ arms have an identical sequence and differ by one or two bases with mature miR166s from other plant species (**Table 2**). Phylogenetic analysis indicated that some Vco-miR166s are identical to mature Csi-miR166s, while the precursor sequences of miR166s from blueberry, sweet orange, grape, and apple also showed a high degree of conservation along the 3′ arm (**Figure 1**). Thus, miR166 family members are highly conserved across diverse plant species, especially along the 3′ arm (Barik et al., 2014). Mature miR166 members are known to regulate the abundance of transcripts encoding HD-ZIP III transcription factors in many species (Boualem et al., 2008; Li R. et al., 2018; Yadav et al., 2021). In this study, we predicted that all Vco-miR166s derived from the 3′ arm target HD-ZIP III transcripts, supporting the functional conservation of miR166 family members in various species (**Supplementary Table S6**).

The multiple *miR166* family members likely diverged from a common ancestor, as reflected by the evolutionary diversification of the miR166 precursor sequences (Barik et al., 2014; Li et al., 2017). In addition, an analysis of the *Vco-miR166* promoters and the putative Vco-miR166 target genes indicated that Vco-miR166 exert diverse functions (Li et al., 2017). Indeed, we detected different defense-related elements and anaerobic induction elements in the promoters of *Vco-miR166* loci (**Table 3**). The prediction of Vco-miR166 target genes showed that most Vco-miR166s target different genes (**Table 4**). These results suggest the functional diversification of Vco-miR166s, in agreement with similar results obtained in other plant species (Barik et al., 2014; Li et al., 2017).

Temperature is an important environmental factor that affects plant growth, yield, and geographical distribution. Plants are constantly subjected to freezing, cold, and heat stress from their surrounding natural environment, with miRNAs playing an active role in temperature stress responses (Sun et al., 2019; Fang and Wang, 2021). In rice (*Oryza sativa*), Osa-miR166 m and Osa-miR166 k are involved in cold stress (Devi et al., 2013). Similarly, Ath-miR166 abundance is upregulated in Arabidopsis seedlings exposed to cold stress (Zhou et al., 2008), while Stu-miR166 (from tomato) and miR166a-e [from almond (*Prunus dulcis*)] abundance decreased under cold stress (Ou et al., 2015; Karimi et al., 2016). In tomato, upregulation of Sly-miR166 and the concomitant downregulation of its target gene encoding an HD-ZIP III showed that Sly-miR166 plays a critical role in regulating cold stress responses in this species (Valiollahi et al., 2014). Under chilling conditions, miR166e-3p abundance in field mustard (*Brassica rapa*) was upregulated, while miR166a levels were downregulated; in almond, the abundance of miR166a-e decreased upon exposure to cold (Zeng et al., 2018; Karimi et al., 2016). In the context of heat exposure, miR166 abundance decreased in radish (*Raphanus sativus*) but increased in rice (Wang et al., 2015; Li et al., 2014a). However, miR166a/b/c-3p abundance increased while that of miR166c/f-5p decreased in *Betula luminifera* under heat stress (Pan et al., 2017). The abundance of almost all Vco-miR166s characterized here transiently increased before dropping under temperature stress. In particular, the levels of Vco-miR166s rapidly increased within 1 or 2 h under freezing stress (**Figure 3A**). At the same time, the promoters of several *Vco-miR166*s (*Vco-miR166a*, *Vco-miR166g*, and *Vco-miR166h*) harbored a low-temperature response element; notably, Vco-miR166s also targeted transcripts encoding HD-Zip III transcription factors, which are involved in temperature stress (Valiollahi et al., 2014). The functions of predicted target genes showed that Vco-miR166f-5p and other Vco-miR166s may target heat and cold shock proteins and directly regulate temperature stress responses (**Supplementary Table S6**). Our results indicated that Vco-miR166s may be early positive regulators of the response to temperature stress.

Drought is another major abiotic stress that limits plant growth and development. Many miRNAs regulate drought responses in plants (Li et al., 2013; Budak et al., 2015; Pagliarani et al., 2017). Indeed, miR166 expression was shown to be downregulated in barley and in emmer wheat, while being upregulated in rubber tree (*Hevea brasiliensis*) in response to drought (Kantar et al., 2010; Kuruvilla et al., 2019; Kantar et al., 2011). In rice, manipulating the levels of miR166 demonstrated its contribution to drought resistance (Zhang et al., 2018). In white mulberry (*Morus alba multicaulis*), miR166f might function as a positive regulator of drought stress (Li R. et al., 2018). In this study, we showed that the abundance of all Vco-miR166s decreases under drought stress. Thus, the Vco-miR166 family might act as a negative regulator of drought stress.

Salt stress is a major abiotic stress that affects plant growth and development. Many miRNAs, including miR166 family members, play important roles in response to salt stress in plants (Si et al., 2014; Sun et al., 2019; Fang and Wang, 2021). The abundance of miR166a-5p and miR166j-3p was upregulated in soybean roots, while that of miR166a-5p was downregulated in barley roots (Wang et al., 2020;

Kuang et al., 2021). However, miR166 family members were upregulated in roots and downregulated in leaves of sweet potato (*Ipomoea batatas*) (Yang et al., 2020), while Pgu-miR166t was also downregulated in the leaves of common guava (*Psidium guajava*) (Sharma et al., 2020). In our study, the levels of Vco-miR166e-3p, Vco-miR166f-5p, and Vco-miR166i-3p were downregulated in highbush blueberry leaves, while those of Vco-miR166d-5p and Vco-miR166h-3p were upregulated upon exposure to salt stress relative to controls. However, Vco-miR166a levels rose at 1 h and dropped at 12 h following exposure to salt stress (**Figure 3E**). These results indicated that different members of the Vco-miR166 family exhibit distinct expression patterns and form a complex regulatory network under salt stress.

Here, we identified three new *miR166* loci in the highbush blueberry genome, which increased the size of the highbush blueberry miR166 family. The resulting 16 *Vco-miR166* members of highbush blueberry exhibited evolutionary conservation and functional diversification based on sequence analysis, promoter dissection, and target gene predictions. The analysis of their upstream sequences and stem-loop RT-qPCR showed that Vco-miR166 family members are involved in various abiotic stresses, possibly by positively and/or negatively regulating target genes.

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

CZ conceived and designed the project. YL, XW, QG, XZ, LZ, and YZ participated in the experiments and data analysis. YL and XW drafted the manuscript. CZ modified the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Natural Science Foundation of Jilin Province, China (20210101006JC).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.919856/full#supplementary-material

REFERENCES

Barik, S., SarkarDas, S., Singh, A., Gautam, V., Kumar, P., Majee, M., et al. (2014). Phylogenetic Analysis Reveals Conservation and Diversification of Micro RNA166 Genes Among Diverse Plant Speciesfication of Micro RNA166 Genes Among Diverse Plant Species. *Genomics* 103, 114–121. doi:10.1016/j.ygeno.2013.11.004

Boualem, A., Laporte, P., Jovanovic, M., Laffont, C., Plet, J., Combier, J.-P., et al. (2008). MicroRNA166 Controls Root and Nodule Development in Medicago Truncatula. *Plant J.* 54, 876–887. doi:10.1111/j.1365-313x.2008.03448.x

Budak, H., Kantar, M., Bulut, R., and Akpinar, B. A. (2015). Stress Responsive miRNAs and isomiRs in Cereals. *Plant Sci.* 235, 1–13. doi:10.1016/j.plantsci.2015.02.008

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, 1–15. doi:10.1093/gigascience/giz012

Dai, X., Zhuang, Z., and Zhao, P. X. (2018). PsRNATarget: a Plant Small RNA Target Analysis Server (2017 Release). *Nucleic Acids Res.* 46, W49–W54. doi:10.1093/nar/gky316

Devi, S. J. S. R., Madhav, M. S., Kumar, G. R., Goel, A. K., Umakanth, B., Jahnavi, B., et al. (2013). Identification of Abiotic Stress miRNA Transcription Factor Binding Motifs (TFBMs) in Rice. *Gene* 531, 15–22. doi:10.1016/j.gene.2013.08.060

Fang, L., and Wang, Y. (2021). MicroRNAs in Woody Plants. *Front. Plant Sci.* 12, 686831. doi:10.3389/fpls.2021.686831

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., and Hofacker, I. L. (2008). The Vienna RNA Websuite. *Nucleic Acids Res.* 36, W70–W74. doi:10.1093/nar/gkn188

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. *GigaSci* 4, 5. doi:10.1186/s13742-015-0046-9

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. *Ijms* 18, 2767. doi:10.3390/ijms18122767

Kantar, M., Unver, T., and Budak, H. (2010). Regulation of Barley miRNAs upon Dehydration Stress Correlated with Target Gene Expression. *Funct. Integr. Genomics* 10, 493–507. doi:10.1007/s10142-010-0181-4

Kantar, M., Lucas, S. J., and Budak, H. (2011). MiRNA Expression Patterns of *Triticum Dicoccoides* in Response to Shock Drought Stress. *Planta* 233, 471–484. doi:10.1007/s00425-010-1309-4

Karimi, M., Ghazanfari, F., Fadaei, A., Ahmadi, L., Shiran, B., Rabei, M., et al. (2016). The Small-RNA Profiles of Almond (*Prunus Ducis* Mill.) Reproductive Tissues in Response to Cold Stress. *Plos One* 11, e0156519. doi:10.1371/journal.pone.0156519

Kitazumi, A., KawaharaOnda, Y. T. S., Onda, T. S., De Koeyer, D., and de los Reyes, B. G. (2015). Implications of miR166 and miR159 Induction to the Basal Response Mechanisms of an Andigena Potato (*Solanum tuberosum* Subsp. Andigena) to Salinity Stress, Predicted from Network Models in Arabidopsis. *Genome* 58, 13–24. doi:10.1139/gen-2015-0011

Kuang, L., Yu, J., Shen, Q., Fu, L., and Wu, L. (2021). Identification of microRNAs Responding to Aluminium, Cadmium and Salt Stresses in Barley Roots. *Plants* 10, 2754. doi:10.3390/plants10122754

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi:10.1093/molbev/msy096

Kuruvilla, L., Sathik, M., Luke, L. P., and Thomas, M. (2019). Identification and Validation of Drought-Responsive microRNAs from Hevea Brasiliensis. *Acta Physiol. Plant.* 41, 14. doi:10.1007/s11738-018-2803-8

Li, J.-s., Fu, F.-l., An, M., Zhou, S.-f., She, Y.-h., and Li, W.-c. (2013). Differential Expression of microRNAs in Response to Drought Stress in Maize. *J. Integr. Agric.* 12, 1414–1422. doi:10.1016/S2095-3119(13)60311-1

Li, X., Xie, X., Li, J., Cui, Y., Hou, Y., Zhai, L., et al. (2017). Conservation and Diversification of the miR166 Family in Soybean and Potential Roles of Newly Identified miR166s. *BMC Plant Biol.* 17, 32. doi:10.1186/s12870-017-0983-9

Li, G., Wang, Y., Lou, X., Li, H., and Zhang, C. (2018a). Identification of Blueberry miRNAs and Their Targets Based on High-Throughput Sequencing and Degradome Analyses. *Ijms* 19, 983. doi:10.3390/ijms19040983

Li, R., Fan, T., Wang, T., Dominic, K., Hu, F., Liu, L., et al. (2018b). Characterization and Functional Analysis of miR166f in Drought Stress Tolerance in Mulberry (*Morus Multicaulis*). *Mol. Breed.* 38, 132. doi:10.1007/s11032-018-0886-y

Li, X., Hou, Y., Zhang, L., Zhang, W., Quan, C., Cui, Y., et al. (2014a). Computation Identification of Conserved microRNAs and Their Targets from Expression Sequence Tags of Blueberry (*Vaccinium Corybosum*). *Plant Signal. Behav.* 9, e29462. doi:10.4161/psb.29462

Li, J., Wu, L.-Q., Zheng, W.-Y., Wang, R.-F., and Yang, L.-X. (2014b). Genome-wide Identification of microRNAs Responsive to High Temperature in Rice (*Oryza Sativa*) by High-Throughput Deep Sequencing. *J. Agro Crop Sci.* 201, 379–388. doi:10.1111/jac.12114

Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004). Incorporating Chemical Modification Constraints into a Dynamic Programming Algorithm for Prediction of RNA Secondary Structure. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7287–7292. doi:10.1073/PANS.040179910110.1073/pnas.0401799101

Miyashima, S., Honda, M., Hashimoto, K., Tatematsu, K., Hashimoto, T., Sato-Nara, K., et al. (2013). A Comprehensive Expression Analysis of the Arabidopsis MICRORNA165/6 Gene Family during Embryogenesis Reveals a Conserved Role in Meristem Specification and a Non-cell-autonomous Function. *Plant Cell Physiol.* 54, 375–384. doi:10.1093/pcp/pcs188

Ou, Y., Liu, X., Xie, C., Zhang, H., Lin, Y., Li, M., et al. (2015). Genome-Wide Identification of microRNAs and Their Targets in Cold-Stored Potato Tubers by Deep Sequencing and Degradome Analysis. *Plant Mol. Biol. Rep.* 33, 584–597. doi:10.1007/s11105-014-0771-8

Pagliarani, C., Vitali, M., Ferrero, M., Vitulo, N., Incarbone, M., Lovisolo, C., et al. (2017). The Accumulation of miRNAs Differentially Modulated by Drought Stress Is Affected by Grafting in Grapevine. *Plant Physiol.* 173, 2180–2195. doi:10.1104/pp.16.01119

Pan, Y., Niu, M., Liang, J., Lin, E., Tong, Z., and Zhang, J. (2017). Idenfification of Heat-Responsive miRNAs to Reverl the miRNA-Mediated Regulatory Network of Heat Stress Response in *Betula Luminifera*. *Trees* 31, 1635–1652. doi:10.1007/s00468-017-1575-x

Polashock, J., Zelzion, E., Fajardo, D., Zalapa, J., Georgi, L., Bhattacharya, D., et al. (2014). The American Cranberry: First Insights into the Whole Genome of a Species Adapted to Bog Habitat. *BMC Plant Biol.* 14, 165. doi:10.1186/1471-2229-14-165

Sharma, A., Ruiz-Manriquez, L. M., Serrano-Cano, F. I., Reyes-Pérez, P. R., Tovar Alfaro, C. K., Barrón Andrade, Y. E., et al. (2020). Idenfification of microRNA and Their Expression in Leaf Tissues of Guava (*Psidium Guajava* L.) under Salinity Stress. *Agronomy* 10, 1920. doi:10.3390/agronomy10121920

Si, J., Zhou, T., Bo, W., Xu, F., and Wu, R. (2014). Genome-Wide Analysis of Salt-Responsive and Novel microRNAs in *Populus Euphratica* by Deep Sequencing. *BMC Genet.* 15, S6. doi:10.1186/1471-2156-15-S1-S6

Singh, A., Singh, S., Panigrahi, K. C. S., Reski, R., and Sarkar, A. K. (2014). Balanced Activity of microRNA166/165 and its Target Transcripts from the Class III Homeodomain-Leucine Zipper Family Regulates Root Growth in *Arabidopsis thaliana*. *Plant Cell Rep.* 33, 945–953. doi:10.1007/s00299-014-1573-z

Sun, X., Lin, L., and Sui, N. (2019). Regulation Mechanism of microRNA in Plant Response to Abiotic Stress and Breeding. *Mol. Biol. Rep.* 46, 1447–1457. doi:10.1007/s11033-018-4511-2

Thakur, V., Wanchana, S., Xu, M., Bruskiewich, R., Quick, W. P., Mosig, A., et al. (2011). Characterization of Statistical Features for Plant microRNA Prediction. *BMC Genomics* 12, 108. doi:10.1186/1471-2164-12-108

Valiollahi, E., Farsi, M., and Kakhki, A. M. (2014). Sly-miR166 and Sly-miR319 Are Components of the Cold Stress Response in Solanum lycopersicum. *Plant Biotechnol. Rep.* 8, 349–356. doi:10.1007/s11816-014-0326-3

Wang, R., Xu, L., Zhu, X., Zhai, L., Wang, Y., Yu, R., et al. (2015). Transcriptome-Wide Characterization of Novel and Heat-Stress-Responsive microRNAs in Radish (Raphanus Sativus L.) Using Next-Generation Sequencing. *Plant Mol. Biol. Rep.* 33, 867–880. doi:10.1007/s11105-014-0786-1

Wang, Q., Yang, Y., Lu, G., Sun, X., Feng, Y., Yan, S., et al. (2020). Genome-wide Identification of microRNAs and Phased siRNAs in Soybean Roots under Long-Term Salt Stress. *Genes Genom.* 42, 1239–1249. doi:10.1007/s13258-020-00990-0

Yadav, A., Kumar, S., Verma, R., Lata, C., Sanyal, I., and Rai, S. P. (2021). MicroRNA 166: an Evolutionarily Conserved Stress Biomarker in Land Plants Targeting HD-ZIP Family. *Physiol. Mol. Biol. Plants* 27, 2471–2485. doi:10.1007/s12298-021-01096-x

Yang, Z., Zhu, P., Kang, H., Liu, L., Cao, Q., Sun, J., et al. (2020). High-throughput Deep Sequencing Reveals the Important Role that microRNAs Play in the Salt Response in Sweet Potato (*Ipomoea Batatas* L.). *BMC Genomics* 21, 164. doi:10.1186/s12864-020-6567-3

Yue, J., Lu, X., Zhang, H., Ge, J., Gao, X., and Liu, Y. (2017). Identification of Conserved and Novel microRNAs in Blueberry. *Front. Plant Sci.* 8, 1155. doi:10.3389/fpls.2017.01155

Zeng, X., Xu, Y., Jiang, J., Zhang, F., Ma, L., Wu, D., et al. (2018). Identification of Cold Stress Responsive microRNAs in Two Winter Turnip Rape (*Brassica Rapa* L.) by High Throughput Sequencing. *BMC Plant Biol.* 18, 52. doi:10.1186/s12870-018-1242-4

Zhang, J., Zhang, H., Srivastava, A. K., Pan, Y., Bai, J., Fang, J., et al. (2018). Knockdown of Rice microRNA166 Confers Drought Resistance by Causing Leaf Rolling and Altering Stem Xylem Development. *Plant Physiol.* 176, 2082–2094. doi:10.1104/pp.17.01432

Zhou, X., Wang, G., Sutoh, K., Zhu, J., and Zhang, W. (2008). Identification of Cold-Inducible microRNAs in Plants by Transcriptome Analysis. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.* 1779, 780–788. doi:10.1016/j.bbagrm.2008.04.005

Conflict of Interest: YZ was employed by Helong Forestry Co., Ltd.

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 16 May 2022
doi: 10.3389/fgene.2022.854346

An Integrated Regulatory Network of mRNAs, microRNAs, and lncRNAs Involved in Nitrogen Metabolism of Moso Bamboo

Tingting Yuan[1,2], *Chenglei Zhu*[1,2], *Guangzhu Li*[1,2], *Yan Liu*[1,2], *Kebin Yang*[1,2], *Zhen Li*[1,2], *Xinzhang Song*[3]* and *Zhimin Gao*[1,2]*

[1]Key Laboratory of National Forestry and Grassland Administration/Beijing for Bamboo and Rattan Science and Technology, Beijing, China, [2]International Center for Bamboo and Rattan, Institute of Gene Science and Industrialization for Bamboo and Rattan Resources, Beijing, China, [3]State Key Laboratory of Subtropical Silviculture, Zhejiang A and F University, Hangzhou, China

OPEN ACCESS

Edited by:
Deqiang Zhang,
Beijing Forestry University, China

Reviewed by:
Su Chen,
Northeast Forestry University, China
Zhongxiong Lai,
Fujian Agriculture and Forestry University, China

***Correspondence:**
Xinzhang Song
songxinzhang@gmail.com
Zhimin Gao
gaozhimin@icbr.ac.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: 13 January 2022
Accepted: 14 April 2022
Published: 16 May 2022

Citation:
Yuan T, Zhu C, Li G, Liu Y, Yang K, Li Z, Song X and Gao Z (2022) An Integrated Regulatory Network of mRNAs, microRNAs, and lncRNAs Involved in Nitrogen Metabolism of Moso Bamboo. Front. Genet. 13:854346. doi: 10.3389/fgene.2022.854346

Nitrogen is a key macronutrient essential for plant growth and development, and its availability has a strong influence on biological processes. Nitrogen fertilizer has been widely applied in bamboo forests in recent decades; however, the mechanism of nitrogen metabolism in bamboo is not fully elucidated. Here, we characterized the morphological, physiological, and transcriptome changes of moso bamboo in response to different schemes for nitrogen addition to illuminate the regulation mechanism of nitrogen metabolism. The appropriate addition of nitrogen improved the chlorophyll content and Pn (net photosynthetic rate) of leaves, the nitrogen and ammonium contents of the seedling roots, the biomass of the whole seedling, the number of lateral roots, and the activity of enzymes involved in nitrogen metabolism in the roots. Based on the whole transcriptome data of the roots, a total of 8,632 differentially expressed mRNAs (DEGs) were identified under different nitrogen additions, such as 52 nitrate transporter genes, 6 nitrate reductase genes, 2 nitrite reductase genes, 2 glutamine synthase genes, 2 glutamate synthase genes (GOGAT), 3 glutamate dehydrogenase genes, and 431 TFs belonging to 23 families. Meanwhile, 123 differentially expressed miRNAs (DEMs) and 396 differentially expressed lncRNAs (DELs) were characterized as nitrogen responsive, respectively. Furthermore, 94 DEM-DEG pairs and 23 DEL-DEG pairs involved in nitrogen metabolism were identified. Finally, a predicted regulatory network of nitrogen metabolism was initially constructed, which included 17 nitrogen metabolic pathway genes, 15 TFs, 4 miRNAs, and 10 lncRNAs by conjoint analysis of DEGs, DEMs, and DELs and their regulatory relationships, which was supported by RNA-seq data and qPCR results. The lncRNA-miRNA-mRNA network provides new insights into the regulation mechanism of nitrogen metabolism in bamboo, which facilitates further genetic improvement for bamboo to adapt to the fluctuating nitrogen environment.

Keywords: moso bamboo, nitrogen metabolism, transcriptome, microRNA, long noncoding RNA

INTRODUCTION

Nitrogen is a key macronutrient for plants and has a strong influence on crop development and productivity. There are different nitrogen forms available for plants, such as nitrate, ammonium, and a small amount of amino acid (Patterson et al., 2010). Nitrate is the major source of nitrogen due to its highly mobile and readily available features in the soil (Jin et al., 2015). Nitrate transporters (NRTs) are responsible for the uptake, transport of nitrate, and intracellular redistribution in plants, namely, NRT1, NRT2, and NRT3. Plants have evolved two different types of NRTs: high-affinity transport system (HATS) and low-affinity transport system (LATS), which are further divided into component and induction transport systems, respectively. After uptake by NRTs, a part of the nitrate is stored or assimilated in the roots while the rest is transported to the shoots (Xu et al., 2012). Several proteins or enzymes related to nitrogen metabolism play important roles in plant nitrogen assimilation, such as nitrate reductase (NR), nitrite reductase (NiR), glutamine synthase (GS), glutamate synthase (GOGAT), and glutamate dehydrogenase (GDH) (Gupta et al., 2019). Significant progress has been made in understanding nitrogen assimilation in plants (Wang et al., 2020; Li et al., 2020; Sanagi et al., 2021). For instance, the overexpression of nitrate transporter and nitrogen assimilation enzyme genes enhanced the ability of nitrogen uptake, with increased nitrate and ammonium contents, total nitrogen content, dry biomass, and yield in transgenic plants (Feng et al., 2011; Fang et al., 2012; Xia et al., 2015; Chen et al., 2016).

TFs have previously been identified to play important roles in nitrogen metabolism (Xu et al., 2016; Liu et al., 2017; Zhang et al., 2021). Many TFs have been reported to participate in the nitrogen response of *Arabidopsis thaliana*, and TCP20 plays a key role in the systemic signaling pathway that directs nitrate foraging by *Arabidopsis* roots (Brooks et al., 2019). The overexpression of *OsNLP1* promotes growth, nitrogen use efficiency, and grain yield in rice (Alfatih et al., 2020). In addition to protein-coding RNAs, emerging evidence has revealed that noncoding RNAs (ncRNAs) also play essential roles in nitrogen metabolism (Liu et al., 2019). Recent studies show that ncRNAs, such as microRNAs (miRNAs) and long noncoding RNAs (lncRNAs), function in many processes related to agricultural traits. Several large-scale investigations and functional studies of ncRNAs in plants suggested that they had significant regulatory effects on the physiological responses through regulating their targeted genes (Fischer et al., 2013), which led to improvements in some important agricultural traits such as productivity, male sterility, nutrient homeostasis, and floral organogenesis (Zhang et al., 2013; Fukuda et al., 2020). It has also been found that the changes in the nitrogen supply status alter the expressions of multiple miRNAs and lncRNAs in several plant species (Nguyen et al., 2015). Many miRNAs in plants were found to be involved in nitrogen metabolism (Lin et al., 2013; Yu et al., 2018), reprogramming root development (Vidal et al., 2010), and homeostasis of other nutrients (Paul et al., 2015). Although species-wide studies on plant miRNAs have been performed, few studies examining the involvement of lncRNAs in regulating nitrogen metabolism in plants have been reported (Chen et al., 2016; Fu et al., 2019; Zhai et al., 2020).

Moso bamboo (*Phyllostachys edulis*) is one of the representatives of woody bamboo, with 4.68 million hectares of forest area, accounting for about 73% of the total bamboo forest area in China (Li and Feng, 2019). Moso bamboo has a shorter growth cycle, strong regeneration ability, excellent mechanical strength, and high elasticity, which have made it a promising substitute for wood (Song et al., 2011). Moreover, bamboo shoot has become one of the indispensable foods to benefit from due to its low sugar and fat, and being rich in fiber characteristics (Hou et al., 2022). In addition, moso bamboo forests have high ecological value due to their strong carbon sequestration ability (Song et al., 2020). In production, a large amount of fertilizers, the majority of which are nitrogen fertilizers, are applied to obtain more bamboo shoots and culms, as well as higher ecological benefits. Recently, the molecular mechanism of nitrogen metabolism in moso bamboo has attracted much attention. A total of 13 *PeAMT*s, 27 *PeNPF*s, and 10 PeNLPs involved in nitrogen metabolism have been identified in moso bamboo (Li et al., 2021; Yuan et al., 2021a; Yuan, et al., 2021b). However, the internal regulation mechanism of nitrogen metabolism is still unclear. Based on the morphological and physiological changes of moso bamboo seedlings under different nitrate additions, we conducted multiple RNA-Seq analyses to find the TF and enzyme genes, as well as the lncRNAs and miRNAs involved in nitrogen metabolism. Finally, a predicted regulatory network of lncRNA-miRNA-mRNA was initially constructed. Our study provided new insights into the regulatory mechanism of nitrogen metabolism in bamboo.

MATERIALS AND METHODS

Plant Material, Growth Conditions, and Nitrogen Treatment

We germinated and grew bamboo seeds in the substrate (peat: vermiculite = 7:3) in a greenhouse with a 16-h light (30°C)/8-h dark (28°C) photoperiod, 300 μmol m^{-2} s^{-1} photon density, and 60% humidity for 60 days. Next, 108 seedlings with a similar height (15 cm) and growth performance were selected and assigned to three groups, with 36 plants in each group. The seeds of these seedlings were removed, the roots were washed with deionized water, and then the plants were placed in a modified Kimura B solution (Wu et al., 2021) containing one of the following NO_3^- concentrations: 0 mM (N0), 6 mM (N6), 12 mM (N12), 18 mM (N18), and 30 mM (N30). The nutrient solution for culture was renewed every 3 days. The treatments were maintained for 2 weeks until distinct morphological differences were observed among the treatments. The bamboo roots were harvested and part of them wrapped in tinfoil and immediately frozen in liquid nitrogen, and then stored at −80°C for further analysis.

Determination of Physiological and Biochemical Indexes of Seedlings

To get the extracts, 0.2 g leaves were ground into powder and extracted with ethanol until the tissue turned white in darkness. The contents of chlorophyll (chlorophyll a and chlorophyll b) in the extracts were determined by a spectrophotometer (Ultrospec 3300 pro, Biochrom Ltd., Cambridge, England) at 470, 649, and 665 nm. Li-6400XT was used to measure the net photosynthetic rate (Pn) of the seedlings under different nitrogen additions. The red and blue lights of the cooperative determination system were used to determine the Pn from 9:00 to 11:00 a.m. The light intensity in the leaf chamber was 1,600 μmol m^{-2} s^{-1}, and the CO_2 concentration was 400 μmol m^{-2} s^{-1}. The leaves were inserted into the leaf chamber, maintaining a normal growth situation as much as possible. The Pn of the seedlings was recorded when the data were stable. The biomass of the whole plant (dry weight) was determined.

The total nitrogen content was determined according to the Kjeldahl method (Ferrari, 1960). The content of nitrate, ammonium, NR (EC 1.7.99.4), NiR (EC 1.7.2.1), GS (EC 6.3.1.2), GOGAT (EC 1.4.7.1), and GDH (EC 1.4.1.2) activities were determined according to the instructions of the corresponding kits (G0440W, G0410W, G0402W, G0408W, G0401W, G0403W, and G0405W) produced by Suzhou Grace Biotechnology Co., Ltd. (Suzhou, China). Briefly, the samples that were weighed and extracted from the corresponding kit were added, followed by homogenization. Centrifugation was carried out, and the supernatant was taken for testing. The OD values were determined at 219, 570, 530, 540, 540, 340, and 450 nm. Then, the content of nitrate and ammonium and activity of the enzymes were calculated using the corresponding calculation formulas (Luo et al., 2013). All the measurements in this part were completed with three biological replicates.

RNA Isolation, Library Construction, and RNA Sequencing

Total RNA was isolated using the TRIzol method for RNA-Seq. The integrity, purity, and concentration of the purified RNA were assessed with the Agilent 2100 Bio analyzer (Agilent, United States). Only high-quality RNA samples were used to construct the sequencing library. For the mRNA and lncRNA sequencing, 5.0 μg of total RNA was used to prepare rRNA (ribosomal RNA) removed strand-specific library using a TruSeq Stranded Total RNA Library Prep with the Ribo-Zero Plant Kit (Illumina, San Diego, CA, United States) according to the manufacturer's instructions. There were three biological replicates per treatment, and a total of nine libraries were prepared. For small RNA sequencing, nine libraries were constructed with 3.0 μg of total RNA and the Truseq Small RNA Sample Prep Kit (Illumina, San Diego, CA, United States).

The denatured libraries were subjected to high-throughput parallel sequencing of both ends of the library using an Illumina HiSeq X™ Ten System sequencing platform. The quality of the raw data was evaluated using FastQC (version 0.10.1) with default settings. The clean data were separated using cutadapt (version 1.9), and the quality threshold was set to Q30, which removed the sequencing adapters and the primer sequence from the raw data to filter out low-quality data. phyllostachys_edulis.gigadb.HIC was used as the reference genome for sequence alignment and subsequent analysis (Zhao et al., 2018). The transcript level was quantified using Cufflinks (version 2.2.1), and the length of the transcript in the sample was normalized to Fragments Per Kilobase of exon model per Million mapped fragment (FPKM) values (Pertea et al., 2016).

Identification, Functional Annotation, and Regulatory Relationships of Differentially Expressed mRNAs

To identify the DEGs between N0, N6, and N18 treatments, the expression level of each transcript was calculated according to the FPKM method. The DEGs were filtered with the criteria of a fold change ($|FC| \geq 2$ and false discovery rate (FDR) < 0.05 by DESeq2_EBSeq. The putative functions of the DEGs were determined according to nonredundant protein (NR) database. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were carried out using the OmicShare online platform (http://www.omicshare.com/tools/) [Accessed July 2021]. For the coexpression network analysis, the Weighted Gene Coexpression Network Analysis (WGCNA) package and the BambooNET database (http://bioinformatics.cau.edu.cn/bamboo/) [Accessed July 2021] (Ma et al., 2018) were used. The binding elements in the promoter were predicted by PlantRegMap (Tian et al., 2019), and the regulatory relationships between TFs, lncRNAs, miRNAs and nitrogen metabolism pathway genes were visually displayed using Cytoscape 3.7.2 (the same below).

Identification of Differentially Expressed lncRNAs, Differentially Expressed miRNAs, and Their Targeted Genes

The prediction of candidate lncRNAs was performed by basic screening and potential coding ability screening. Briefly, the transcript sequences were assembled, annotated, and filtered based on their coding potential and length, in which the transcripts with potential coding ability were removed. Consequently, the remaining genes were regarded as candidate lncRNAs. To further investigate the potential roles of the lncRNAs, the expression level of each lncRNA was calculated according to the FPKM method. Differentially expressed lncRNAs (DELs) were extracted with an absolute value of $|FC| \geq 2$ and FDR < 0.05 by DESeq2_EBSeq. The potential *cis*- and *trans*-targeted mRNAs of DELs were predicted according to their positions on the chromosome and expression correlation, respectively. For *cis* regulation, the lncRNAs were located within 10 kb upstream or downstream of their adjacent mRNA. For *trans* regulation, the expression level of the lncRNAs was opposite to that of their targeted mRNA, and the Pearson correlation between

their expressions was ≥0.9 which was taken as the criteria for their targeted relationship.

The raw data were first quality controlled using the FASTX toolkit software (version 0.0.13, http://hannonlab.cshl.edu/fastx_toolkit/) to obtain clean small RNA reads by filtering out low-quality bases, sequencing adapters, reads shorter than 18 nt, and reads longer than 32 nt. The assembled unique sequences with clean reads were then BLAST searched against the Rfam database (version 12.1, http://rfam.sanger.ac.uk/) [Accessed January 2021] to remove non-miRNA sequences. The remaining reads were used to predict known miRNAs through a BLAST search of the miRbase (version 21.0, http://www.mirbase.org/) [Accessed March 2021], and novel miRNAs through analysis of the hairpin structure of the miRNA precursor with MIREAP (version 0.2) software. The expression level of each miRNA was calculated according to the transcripts per million (TPM) methods. DEMs were extracted with an absolute value of $|FC| \geq 2$ and FDR < 0.05 by DESeq2_EBSeq. The targeted prediction of DEMs was performed with psRobot (version 1.01) (Zhang et al., 2014).

Validation by Quantitative Polymerase Chain Reaction

To maintain the relative gene expression of the DEGs, DEMs, and DELs, qPCR was conducted as described by Yang et al. (2021). Briefly, cDNA was synthesized from 1.0 μg of total RNA with the HiScript® II Q RT SuperMix (Vazyme, Cat#R223, China) for qPCR of mRNAs, the lnRcute lncRNA First-Strand cDNA Synthesis Kit (TIANGEN, Cat#KR202, China) for qPCR of lncRNAs, and the miRNA First-Strand cDNA Synthesis Kit (Vazyme, Cat#MR101, China) for qPCR of miRNAs. A qPCR assay was performed on a qTOWER 2.2 system (Analytik Jena, Germany) using the LightCycler 480 SYBR Green 1 Master kit (Roche, 04887352001), the lnRcute lncRNA qPCR Detection Kit (TIANGEN, Cat#FP402), and the miRNA Universal SYBR® qPCR Master Mix (Vazyme, Cat#MQ101) following the manufacturers' instructions. The qPCR of each gene was carried out with particularly primers (**Supplementary Table S5**), and each experiment used three biological replicates. The $2^{-\Delta\Delta CT}$ method was used to normalize and determine the RNA level relative to an internal reference gene *PeTIP41* (Fan et al., 2013) or *U6* snRNA (Ding et al., 2011).

Statistical Analysis

For experimental variables, one-way analysis of variance (ANOVA) was used with nitrogen treatment as a factor, and differences between the means were considered significant when $p < 0.05$ and extremely significant when $p < 0.01$. The Ct values obtained from the qPCR were normalized and the relative fold changes in transcripts were calculated using the relative expression software tool. The heat map representing the expressions of genes was computed using the $\log_2$ (FPKM) by TBtools.

RESULTS

Morphological and Physiological Changes of Moso Bamboo in Response to Nitrogen Availability

The leaves of seedlings treated with N0 showed most severe chlorosis, while those treated with N18 and N30 showed similar normal color (**Supplementary Figure S1A**). The growth of roots was obviously promoted by nitrogen addition, except that N30 significantly inhibited it (**Supplementary Figure S1B**). Further measurement indicated that the chlorophyll content in the seedling leaves increased significantly under nitrogen addition when compared with that under N0, and it gradually increased with the increase of nitrate addition until it reached its maximum under N18, and then it slightly declined under N30 (**Figure 1A**). Moreover, Pn increased under all nitrogen additions with the maximum under N12, and there was a drop under N18 and N30 (**Figure 1B**), but it was still higher than under N0, which was similar to the changing trend of chlorophyll content. These results indicate that nitrogen addition within a certain concentration could increase the chlorophyll content and Pn of bamboo seedlings. Moreover, the total biomass of the seedlings further supported this inference; the increased biomass under nitrogen additions was not always increasing (**Figure 1C**). The number of lateral roots increased under nitrogen additions, which means nitrogen had a significant promoting effect on lateral roots (**Figure 1D**). The number of lateral roots were the most under N18 and decreased under N30 treatment. Therefore, seedlings under nitrogen additions of N0, N6, and N18 were selected for further studies.

Furthermore, the contents of total nitrogen, ammonium, and nitrate of the selected seedlings were also measured, and it showed that the contents of total nitrogen, i.e., ammonium under N6 and N18 was significantly higher than that under N0 (**Supplementary Figures S2A, B**), which indicated that nitrogen additions promoted nitrate uptake and assimilation in bamboo. However, the nitrate content in the roots under N6 and N18 was similar to that under N0 (**Supplementary Figure S2C**), suggesting that most of the nitrate absorbed by the roots may have been transported to other tissues. Besides, the activities of enzymes involved in nitrogen metabolism were measured, among which the activities of NR, NiR, and GOGAT were significantly increased and those of GS and GDH were significantly inhibited by nitrogen additions (**Supplementary Figures S2D-H**), indicating that nitrogen assimilation in the roots was induced by nitrogen additions in general.

To reveal the transcriptional regulation mechanisms that underlie the morphological and physiological changes of moso bamboo in response to nitrogen availability, genome-wide transcriptional analyses of RNA-seq data generated from the roots of seedlings under N0, N6, and N18 were conducted. A total of 191 million raw reads were obtained, with 22.6–42.2 million reads of each library. After sequence trimming, the number of clean reads per library still ranged from 21.9 to 41.0 million.

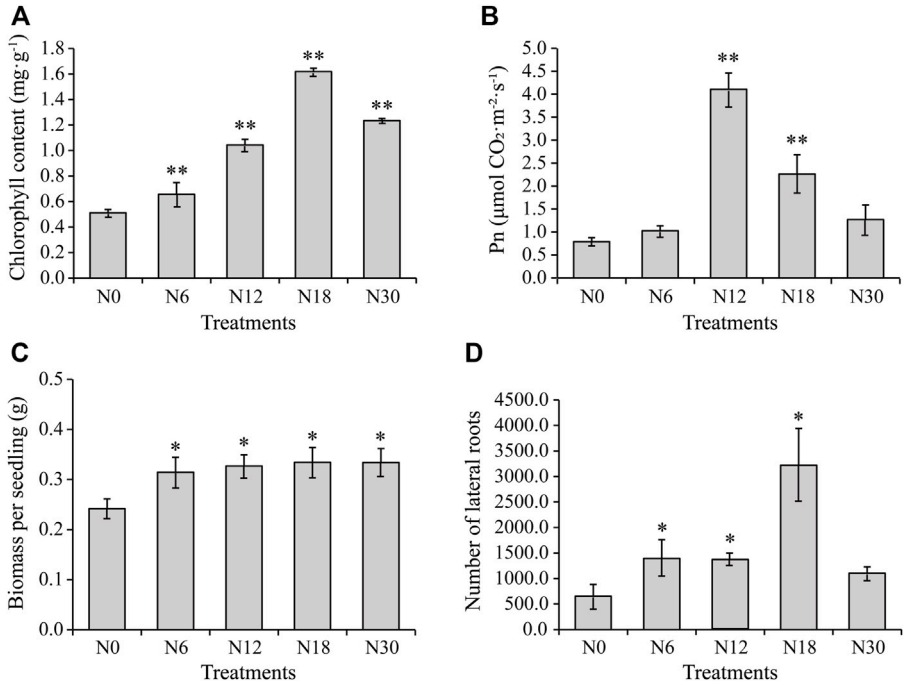


FIGURE 1 | Physiological characteristics of moso bamboo seedlings under different nitrogen additions: **(A)** chlorophyll content of leaves, **(B)** Pn of leaves, **(C)** biomass per seedling, and **(D)** number of lateral roots from seedlings under different nitrogen additions. Data are means ± SD of three replicates ($n = 3$). Asterisks indicate the significant difference between N0 and other nitrogen additions (*means $p < 0.05$, ** means $p < 0.01$, the same below). RNA sequencing elucidated the DEGs in roots under different nitrogen additions.

About 17.7–36.1 million clean reads per library were mapped to the genome of moso bamboo, and the mapping ratio ranged from 71.2 to 87.9%. Finally, 49,292 genes were expressed in at least one sample out of the 59,481 genes detected totally.

A total of 8,632 genes were identified as DEGs, with 6,321 DEGs in N0 *vs*. N6, 7,884 DEGs in N0 *vs*. N18, and 356 DEGs in N6 *vs*. N18 (**Figures 2A, B**; **Supplementary Table S1**). Obviously, the number of DEGs in N0 *vs*. N18 and N0 *vs*. N6 was extremely far more than that in N6 *vs*. N18, indicating that nitrogen addition caused a series of changes in a wide range of biological processes involving a large number of genes in bamboo. Notably, the DEGs in N0 *vs*. N18 were also more than those in N0 *vs*. N6, which reminds us that the higher the nitrogen concentration, the greater the influence on bamboo. However, the concentration of nitrogen under N18 was three times of that under N6, while the number of DEGs in N0 *vs*. N18 was only 1,563 more than that in N0 *vs*. N6, indicating that the number of DEGs may not be totally attributed to the concentration of nitrate, but part of it was caused by the nitrogen present. These results further supported the nitrogen-free, nitrogen addition as well as nitrogen concentration caused by the different numbers of DEGs in bamboo.

Functional Annotation and Classification of the Differentially Expressed mRNAs

To determine the biological functions, the DEGs were subjected to GO and KEGG pathway analyses. Out of 8,632 DEGs, 7,287 were annotated and divided into 53 major GO terms, namely, biological process (20 GO terms), cellular component (17 GO terms), and molecular function (16 GO terms). Of the biological processes, the terms with a high DEG number were metabolic process (GO:0008152), cellular process (GO:0009987), single-organism process (GO:0044699), biological regulation (GO:0065007), and response to stimulus (GO:0050896). Within the cellular component category, the most overrepresented terms were cell (GO:0005623), cell part (GO:0044464), membrane (GO:0016020), membrane part (GO:0044425), and organelle (GO:0043226). As for the molecular function, the most enriched terms were catalytic activity (GO:0003824), binding (GO:0005488), transporter activity (GO:0005215), nucleic acid–binding transcription factor activity (GO:0001071), and enzyme regulator activity (GO:0030234). Among the GO categories, the least frequent GO terms were cell killing (GO:0001906), extracellular matrix part (GO:0044420), and metallochaperone activity (GO:0016530) (**Figure 2C**).

KEGG enrichment analysis showed that the top 20 enriched pathways of the DEGs from N0 *vs*. N6, N0 *vs*. N18, and N6 *vs*. N18 were different from each other. Notably, the enriched pathways of N0 *vs*. N6 and N0 *vs*. N18 were similar, which were greatly different from those of N6 *vs*. N18. In comparisons of N0 *vs*. N6 and N0 *vs*. N18, the most enriched pathway was phenylpropanoid biosynthesis (ko00940), and nitrogen metabolism (ko00910) was also significantly enriched in both comparisons. As for the comparison of N6 vs. N18, starch and sucrose metabolism (ko00500) was the top enriched pathway, and

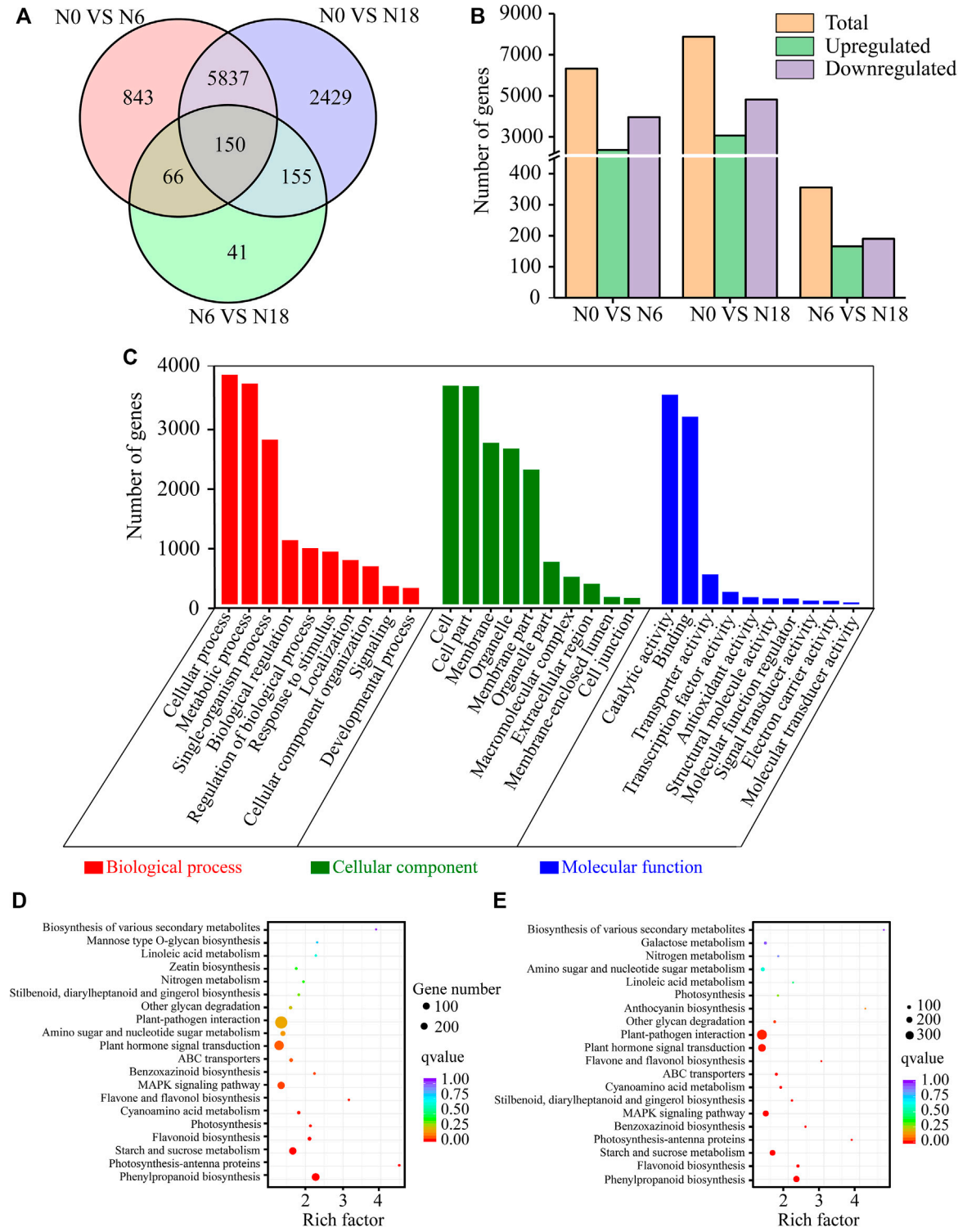


FIGURE 2 | DEGs statistics and functional annotation: **(A)** Venn diagrams of DEGs; **(B)** number of upregulated and downregulated genes; **(C)** GO classification of DEGs; **(D)** KEGG enrichment of DEGs from N0 *vs*. N6; **(E)** KEGG enrichment of DEGs from N0 *vs*. N18.

phenylpropanoid biosynthesis (ko00940) was ranked last, while nitrogen metabolism (ko00910) was not enriched in the top 20 pathways (**Supplementary Figure S3**). These results indicated that similar changes of genes involved in phenylpropanoid biosynthesis and nitrogen metabolism occurred in bamboo under nitrogen addition. However, there were differences

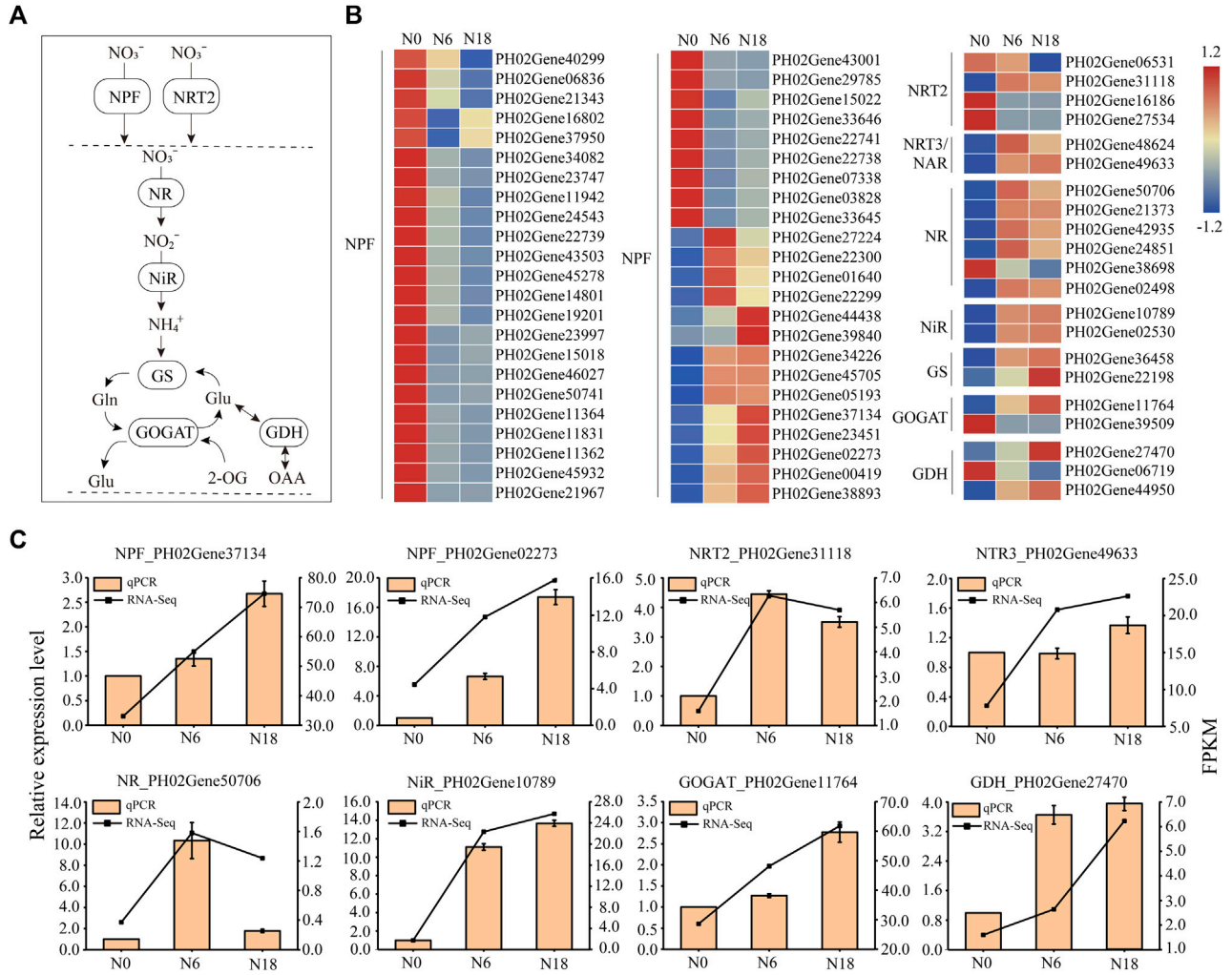
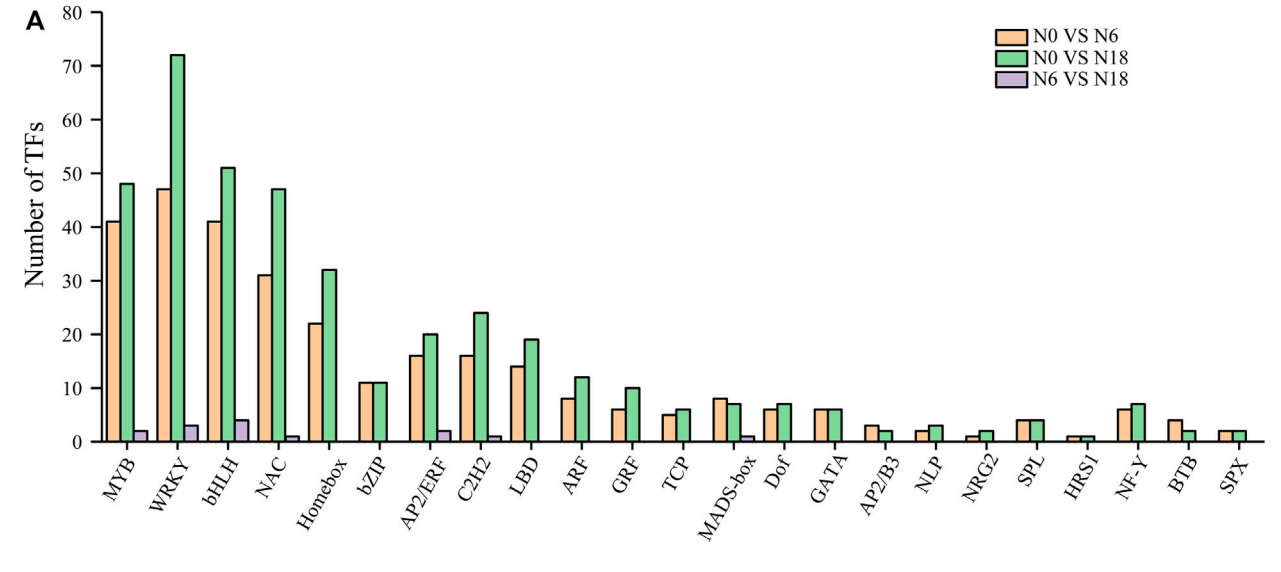


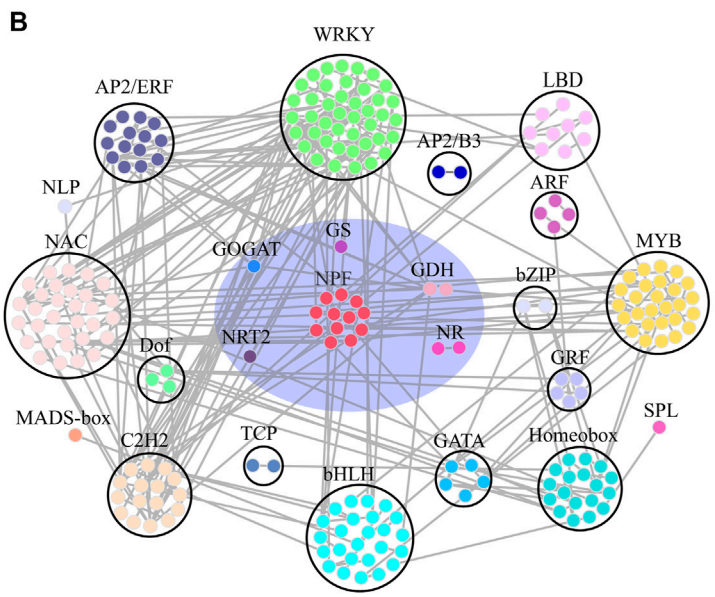
FIGURE 3 | Expression profiles of genes involved in nitrogen metabolism of bamboo roots: **(A)** outline of nitrate uptake and assimilation; **(B)** heat map visualization of DEG expression profiles; **(C)** expression analysis of eight DEGs belonging to the NPF, NRT2, NRT3, NR, NiR, GOGAT, and GDH families. Expression profiles of TFs that participated in nitrogen metabolism.

between N0 *vs*. N6 and N0 *vs*. N18, such as starch and sucrose metabolism were strengthened in N0 *vs*. N18, while phenylpropanoid biosynthesis and nitrogen metabolism were weakened in N0 *vs*. N6, indicating that nitrogen addition might change the flow direction of nitrogen and carbon, resulting in the change of metabolic pathways in bamboo (**Figures 2D, E**).

Expression Patterns of Transporter and Enzyme Genes Involved in Nitrogen Metabolism

The expression patterns of transporter and enzyme genes involved in nitrate uptake, transport, and assimilation were analyzed. Totally, 67 DEGs of 8 families, namely, NPF, NRT2, NRT3/NAR2, NR, NiR, GS, GOGAT, and GDH, were detected (**Figures 3A, B**). Among these families, NPF contained the most DEGs, in which the expressions of 32 members were inhibited and 10 members were induced by nitrogen additions, while 4 members were first induced under N6 and inhibited under N18. The diversity of expression patterns indicated that *PeNPF*s function widely in nitrogen uptake and transport, which is consistent with those NPF members in *Arabidopsis* and rice (Hu et al., 2014; Wang et al., 2018). In addition, three members of *PeNRT2* were induced under N0 and inhibited under N6 and N18, while another member of *PeNRT2* was induced under N6 and N18, which was similar to those members of *NRT3/NAR2*. Most DEGs involved in nitrogen assimilation were induced by nitrogen additions, and their expressions increased with the nitrogen concentration, such as the members of the NiR and GS families. Three DEGs belonging to different nitrogen assimilation families were induced under N0 and suppressed under N6 and N18. These results were further validated by eight randomly selected DEGs





C

TF	PCC	Targeted gene	Binding site
MYB_PH02Gene36545	0.894	NPF_PH02Gene02273	
MYB_PH02Gene01663	0.985	NPF_PH02Gene45278	
MYB_PH02Gene47557	-0.909	NPF_PH02Gene45278	
NAC_PH02Gene44802	0.956	NPF_PH02Gene07338	
NAC_PH02Gene01369	1.000	NPF_PH02Gene45932	
C2H2_PH02Gene34850	0.968	NPF_PH02Gene11942	
LBD_PH02Gene29604	1.000	NPF_PH02Gene22739	
WRKY_PH02Gene06275	-0.964	NRT2_PH02Gene27534	
NAC_PH02Gene03449	0.997	GOGAT_PH02Gene22198	
WRKY_PH02Gene42860	0.997	GDH_PH02Gene27470	
WRKY_PH02Gene03691	0.999	GDH_PH02Gene27470	
WRKY_PH02Gene49262	0.997	GDH_PH02Gene27470	
C2H2_PH02Gene04972	0.987	GDH_PH02Gene27470	
C2H2_PH02Gene06082	0.998	GDH_PH02Gene27470	
MYB_PH02Gene15547	0.984	GDH_PH02Gene06719	

FIGURE 4 | TFs identified from DEGs and their targeted genes involved in nitrogen metabolism: **(A)** statistics of TFs; **(B)** coexpression network of TFs and nitrogen metabolism genes based on WGCNA and BambooNET; **(C)** PCC of TFs and their targeted genes involved in nitrogen metabolism. miRNA sequencing uncovered the DEMs in roots under different nitrogen additions.

using qPCR, which showed a consistent trend with the transcriptome (**Figure 3C**).

TFs play important roles in regulating plant growth and adapting to the environment. According to previous reports, more than 20 TF families are involved in nitrogen-mediated biological processes (Gaudinier et al., 2018; Varala et al., 2018). In this study, 431 TFs belonging to 23 families were identified from the DEGs, and the top 10 families with the most members were WRKY (74);, basic helix–loop–helix (bHLH) (60); MYB (51); NAM, ATAF, and CUC (NAC) (51); ERF/AP2 (24); homeobox (33); C2H2 (24); LBD (20); bZIP (13); and ARF (12). Besides, other families with a few members of DEGs were also concerned due to their significant differences in comparisons of N0 *vs.* N6, N0 *vs.* N18, and N6 *vs.* N18, such as

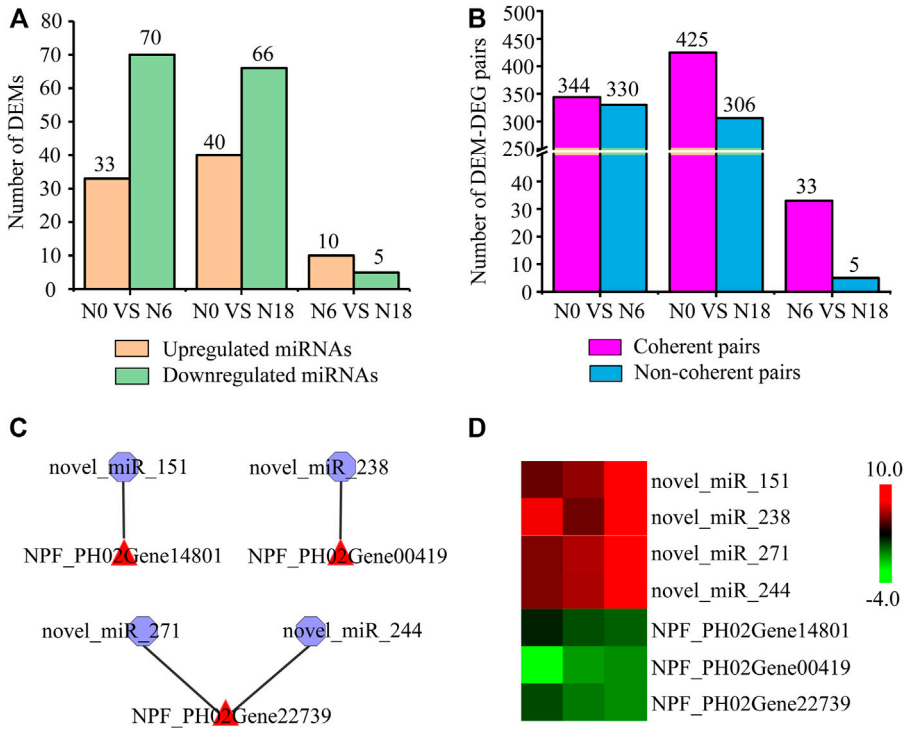


FIGURE 5 | Analysis of DEMs and their targeted genes: **(A)** statistics of DEMs; **(B)** statistics of DEM-DEG pairs with similar or opposite expression patterns; **(C)** DEM-DEG pairs involved in nitrogen metabolism; **(D)** expression analysis of DEM-DEG pairs using transcriptome data. Heat map colors varied from green to red, representing low to high gene expression.

BTB (4), NLP (3), AP2/B3 (3), NRG2 (2), SPX (2), and HRS1 (1) (**Figure 4A**).

To find the transcriptional regulatory relationship, the co-expressed which were differently expressed TFs, and the transporter and enzyme genes involved in nitrogen metabolism were identified. In total, 246 gene pairs from 18 TF families and 6 transporter and enzyme families were identified (**Figure 4B**), among which 34 pairs of TF-nitrogen metabolism genes were found with a Pearson correlation coefficient (|PCC|) > 0.8 and the $p < 0.05$. Furthermore, we analyzed the binding elements in the promoters of the nitrogen metabolic pathway genes in the 34 pairs. Finally, 15 pairs were identified as regulatory pairs, containing 10 nitrogen metabolic pathway genes and 15 TFs. The conserved regulatory elements targeted by TFs in the promoters of the nitrogen metabolic pathway genes are provided in **Figure 4C**.

A total of 43,326,868 raw reads were generated from 9 small RNA root libraries. After removing the reads of low-quality contaminated adapter sequences, the reads with a base length between 18 and 30 nt were further analyzed as typical miRNAs. A total of 383 miRNAs were identified, namely, 62 known miRNAs and 321 newly predicted ones. To explore the expression level changes and potential regulatory roles of these miRNAs, we calculated the expression levels of all the miRNAs and analyzed the DEMs by pairwise comparisons. Among the DEMs identified in moso bamboo, 33, 40, and 10 were upregulated, while 70, 66, and 5 were downregulated in three comparisons (**Figure 5A**; **Supplementary Table S2**). The interaction of miRNA-mRNA can be either coherent with the opposite expression tendency or noncoherent with a similar expression tendency (Garg et al., 2019). Furthermore, 344, 425, and 33 coherent pairs, and 330, 306, and 5 noncoherent pairs were identified in N0 *vs.* N6, N0 *vs.* N18, and N0 *vs.* N18, respectively (**Figure 5B**).

Moreover, among these different expression patterns of DEM-DEG coherent pairs, we focused on those miRNAs which were targeted genes involved in nitrogen metabolism. Totally, 94 DEM-DEG pairs (involved in nitrogen metabolism) were identified, among which 4 were DEM-nitrogen metabolic pathway gene pairs and 90 were DEM-TF pairs. The above four pairs comprised three *NPF* genes (PH02Gene00419, PH02Gene14801, and PH02Gene22739), which were targeted by *novel_miR_238*, *novel_miR_151*, *novel_miR_244*, and *novel_miR_271*, respectively (**Figure 5C**), and the opposite tendency of transcriptome expression patterns further supported the regulatory relationship between them (**Figure 5D**). The 90 pairs included 22 differentially expressed TFs from GRF, ARF, C2H2, NAC, AP2/ERF, bZIP, homeobox, MADS-box, NF-Y, SPL, and WRKY, which were targeted by 29 DEMs (**Supplementary Figure S4**), such as *GRF*_PH02Gene03385 was targeted by *novel_miR_70*, *novel_miR_262*, *novel_miR_7*, *novel_miR_106*, *novel_miR_56*, *novel_miR_247*, *novel_miR_250*, and *novel_miR_162*, while *C2H2*_PH02Gene43070 and *MADS*-

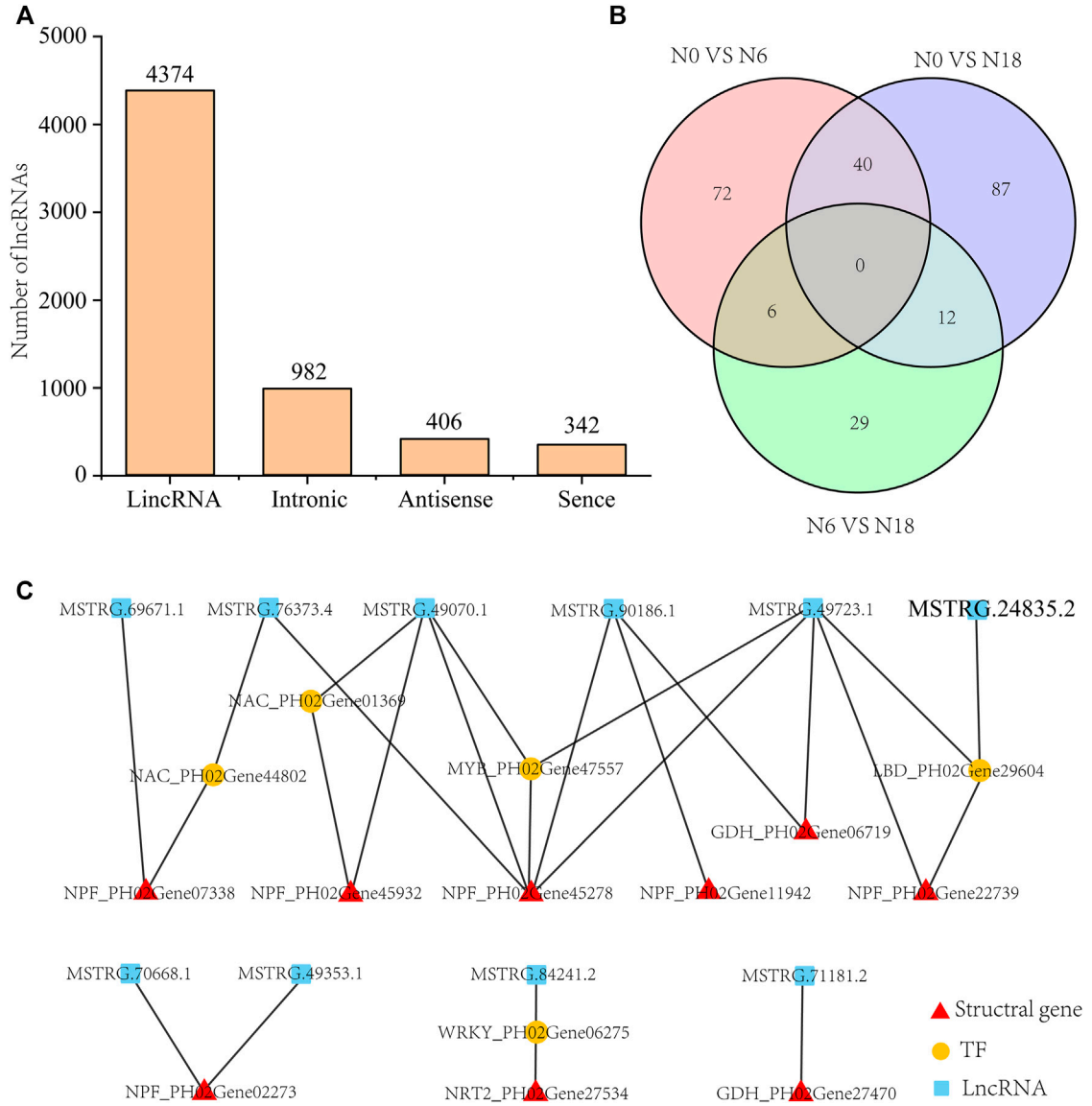


FIGURE 6 | Analysis of DELs and their targeted genes: **(A)** category and number statistics of lncRNAs; **(B)** Venn diagram of DELs from N0 *vs*. N6, N0 *vs*. N18, and N6 *vs*. N18; **(C)** DEL-DEG pairs involved in nitrogen metabolism.

*box*_PH02Gene50032 were targeted by *novel_miR_223* and *novel_miR_333*, respectively.

lncRNAs Sequencing Analysis Uncovered the Differentially Expressed lncRNAs in Roots Under Different Nitrogen Additions

Altogether, 6,104 novel lncRNAs were identified from the 9 cDNA libraries, among which the numbers of lincRNAs, intronic, antisense, and sense were 4,374 (71.7%), 982 (16.1%), 406 (6.7%), and 342 (5.6%), respectively (**Figure 6A**). In the predicted lncRNAs, 396 DELs were identified (**Supplementary Table S3**), namely, 118 DELs (43 upregulated and 75 downregulated) of N0 *vs*. N6, 139 DELs (62 upregulated and 77 downregulated) of N0 *vs*. N18, and 47 DELs (26 upregulated and 21 downregulated) of N6 *vs*. N18 (**Figure 6B**). To further investigate the potential roles of DELs, we first identified the DELs that regulated their targeted genes in *cis* or *trans*. Totally, 237,017 pairs, 247,214 pairs, and 9,315 pairs of DEL-targeted genes were identified, among which 127,314 pairs, 146,008 pairs, and 938 pairs were identified as DEL-DEGs in N0 *vs*. N6, N0 *vs*. N18, and N6 *vs*. N18, respectively.

LncRNAs participated in nitrogen metabolism by directly targeting the nitrogen metabolism pathway genes or indirectly by binding to the TFs that targeted them. Twenty DEL-DEG pairs were identified, which included DEGs of transporters, enzyme genes, and TFs (**Figure 6C**). Eight lncRNAs directly targeted nitrogen metabolic pathway genes, i.e., one-to-one, many-to-one,

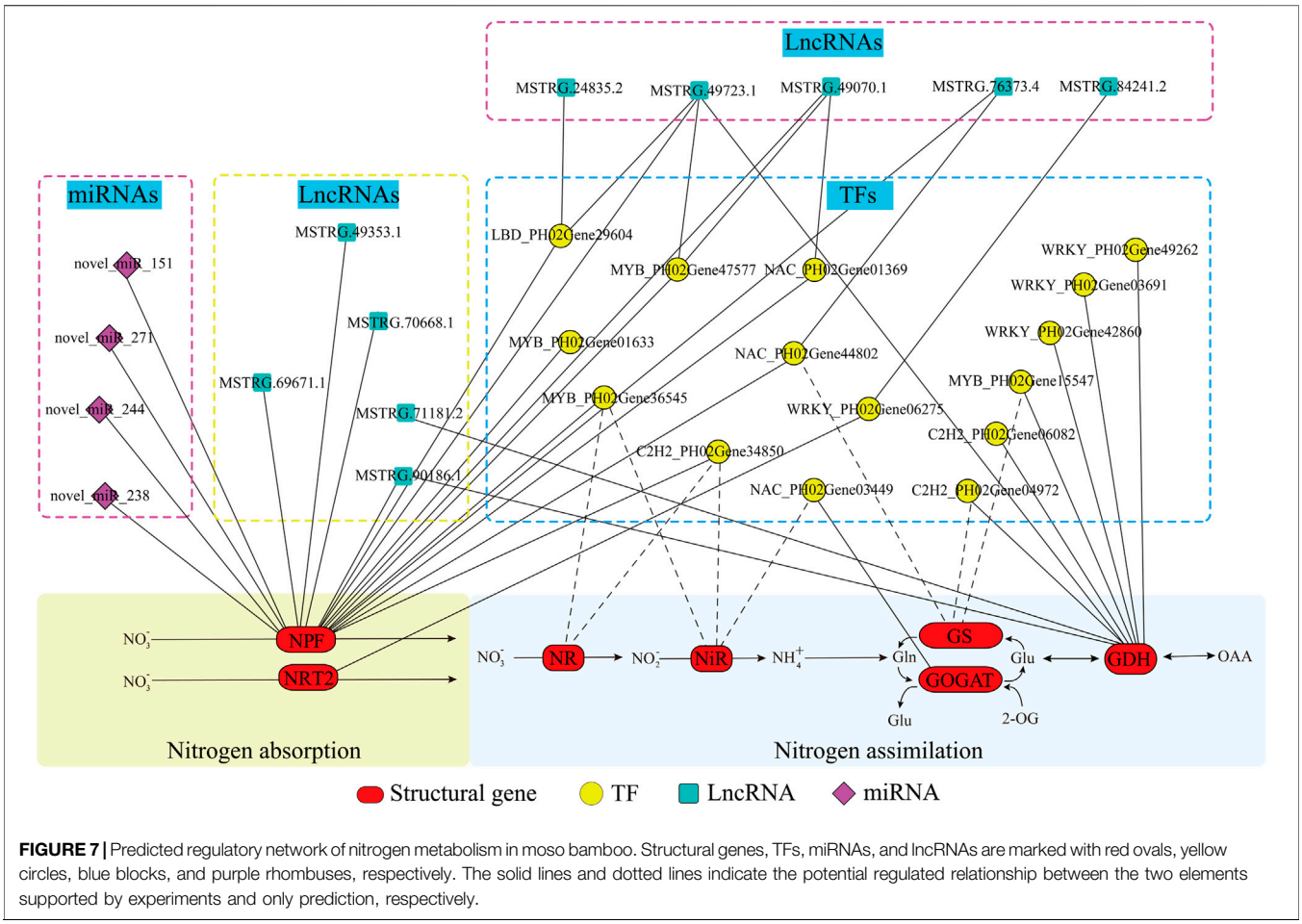


FIGURE 7 | Predicted regulatory network of nitrogen metabolism in moso bamboo. Structural genes, TFs, miRNAs, and lncRNAs are marked with red ovals, yellow circles, blue blocks, and purple rhombuses, respectively. The solid lines and dotted lines indicate the potential regulated relationship between the two elements supported by experiments and only prediction, respectively.

and one-to-many, such as *MSTRG.71181.2* targeted *GDH*_PH02Gene27470, and *MSTRG.90186.1* targeted *NPF*_PH02Gene45278, *NPF*_PH02Gene11942, and *GDH*_PH02Gene06719, while *MSTRG.76373.4*, *MSTRG.49070.1*, and *MSTRG.49723.1* targeted the same *NPF*_PH02Gene45278. Four lncRNAs regulated nitrogen metabolic pathway genes by binding to their targeted TFs, such as NAC_PH02Gene44802, NAC_PH02Gene01369, MYB_PH02Gene47557, LBD_PH02Gene29604, and WRKY_PH02Gene06275. Finally, we identified 10 lncRNAs those which formed 20 gene pairs with the nitrogen metabolic pathway genes or TFs, participating in the response to different nitrogen additions (**Figure 6C**).

Integration and Validation of the Regulatory Network of Nitrogen Metabolism in Moso Bamboo

To further explore the potential regulation mechanism of nitrogen metabolism in moso bamboo, an integrated analysis of lncRNA-miRNA-mRNA was performed. Integrating the above results, a regulatory network was constructed, i.e., 17 nitrogen metabolic pathway genes, 15 TFs, 4 miRNAs, and 10 lncRNAs (**Figure 7**; **Supplementary Table S4**). To validate the genetic elements in the regulatory network, the qPCR method was used. The results showed that the TFs and their targeted genes had similar expression patterns, such as *NAC*_PH02Gene03449 and *GOGAT*_PH02Gene22198 were both downregulated under N6 and upregulated under N18 (**Figure 8A**), which was consistent with the transcriptome data. The same situation was found in *WRKY*_PH02Gene42860 and *GDH*_PH02Gene27470, *C2H2*_PH02Gene34850, and *NPF*_PH02Gene11942 (**Figures 8B, C**). The similar expression patterns further suggested that the enzyme genes were positively regulated by the TFs, which was supported by the binding sites found in the promoter sequences of the enzyme genes (**Figure 4C**).

Novel_miR_151 and *novel_miR_271* were upregulated, while their targeted genes (*NPF*_PH02Gene14801 and *NPF*_PH02Gene22739) were downregulated under N6 and N18 when compared with under N0 (**Figures 8D, E**). In addition, *novel_miR_238* was downregulated and its targeted gene (*NPF*_PH02Gene00419) was upregulated under N6 when compared with under N0, but they were all upregulated under N18 when compared with under N0 (**Figure 8F**). The qPCR of miRNA-mRNA pairs further validated the results of the high-throughput sequencing. *MSTRG.90186.1* and *MSTRG.24835.2* were upregulated, while their targeted genes (*NPF*_PH02Gene45278, *NPF*_PH02Gene11942, and

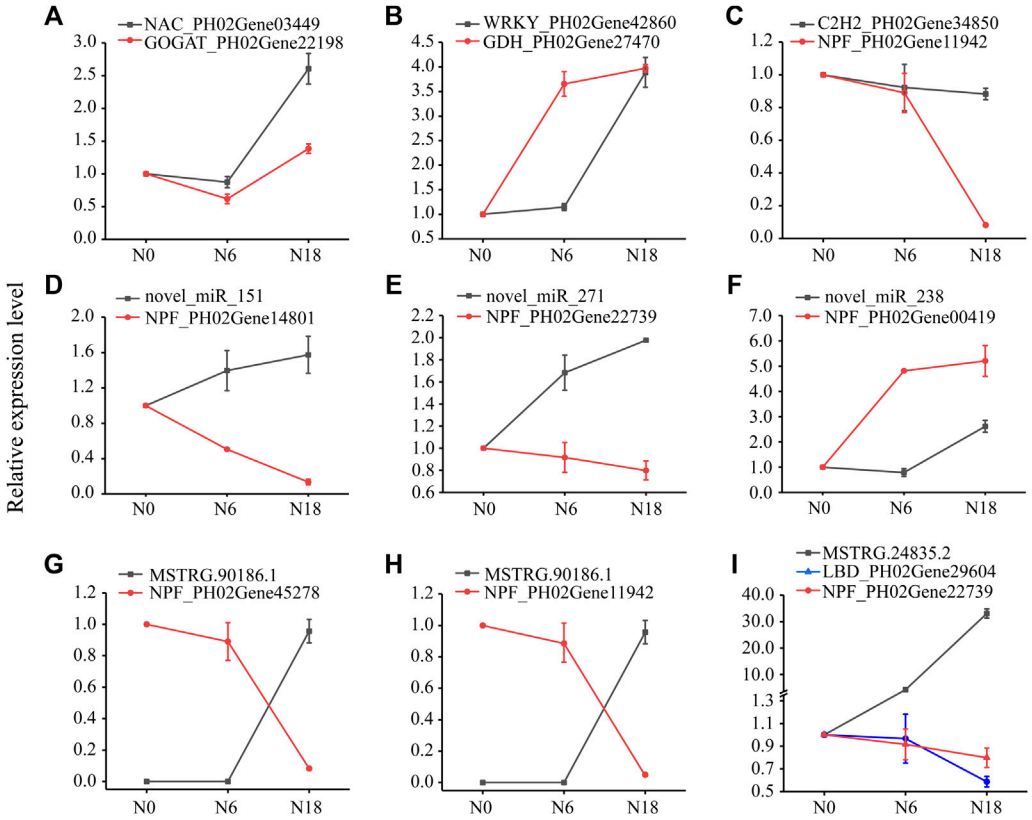


FIGURE 8 | Expression analysis of genes, miRNAs, and lncRNAs using qPCR. Data are means ± SD of three replicates (n = 3): **(A)** NAC-GOGAT pair; **(B)** WRKY-GDH pair; **(C)** WRKY-GDH pair; **(D–F)** miRNA-NPF pairs; **(G–H)** lncRNA-NPF pairs; and **(I)** lncRNA-LBD-NPF pairs.

*LBD*_PH02Gene29604) were downregulated under N6 and N18 when compared with under N0 (**Figures 8G–I**). Meanwhile, *LBD*_PH02Gene29604 and *NPF*_PH02Gene22739 had similar expression patterns, and putative LBD binding sites were found in the promoter of *NPF*_PH02Gene22739 (**Figure 4C**), indicating that *MSTRG.24835.2* indirectly regulates NPF genes by targeting TF. These results provided evidence for elucidating the relationship between the genetic elements in the regulatory network of nitrogen metabolism (**Figure 7**).

DISCUSSION

Nitrogen Availability Affected the Phenotype, Physiology, and Biochemistry of Moso Bamboo

Nitrogen has a wide range of effects on plants: nitrogen deficiency stunts plant growth and development, reduces photosynthesis and leaf area, promotes plant senescence, and ultimately decreases plant productivity (Mu et al., 2016). Nitrogen supply has a great effect on biomass production due to its close relationship with leaf area and the longevity of green leaves (Zhong et al., 2019). In this study, nitrogen additions improved chlorophyll content and Pn of moso bamboo seedling leaves, which led to a biomass increase for the whole plants (**Figure 1**). Similar conclusions were reported in poplar, wheat (*Triticum aestivum*) (Imran et al., 2019), and peanut (*Arachis hypogaea*) (Li et al., 2021). Actually, in our study, the seedlings under N18 treatment had the most chlorophyll accumulation, while the seedlings under N12 had the highest Pn among the treatments (**Figure 1**). This indicates that a moderate nitrogen level is beneficial to the growth of bamboo seedlings, which is similar to the results for spinach (*Spinacia oleracea*) (Lin et al., 2014). Meanwhile, N6 significantly promoted the biomass of seedlings, but the promoting effect was weakened with increasing nitrogen concentrations when compared with that under N6 (**Figure 1**), which are consistent with the results of a poplar study (Luo et al., 2013). Therefore, we hypothesized that there is a threshold for nitrogen uptake and assimilation in moso bamboo.

Besides, there are multiple enzymes involved in nitrogen metabolism, in which NR and NiR achieve the two-step reduction. NR is the rate-limiting enzyme, and its activity directly affects the level of nitrogen metabolism in plants (Huarancca Reyes et al., 2018). The activities of NR and NiR in bamboo roots were promoted by nitrogen additions, but they were weakened when the seedlings were exposed to a higher

nitrogen condition (**Supplementary Figure S2**), which agreed with the NR activity in rice exposed to different nitrogen conditions (Gao et al., 2019). The activity of GS was the highest under N0, while GOGAT activity significantly increased after nitrogen additions (**Supplementary Figure S2**). This may be explained by the feedback regulation between the GS/GOGAT pathway and the posttranscriptional modulation of GS (Ferreira et al., 2019). The activity of GDH was also inhibited by nitrogen additions (**Supplementary Figure S2**), and the reason may be that GDH acted as a link between carbon and nitrogen metabolism, which was influenced by multiple factors (The et al., 2020). In consequence, nitrogen metabolism of bamboo was greatly affected by the fluctuating nitrogen conditions. Further experiments are needed to explain the variations of the enzyme activities in the nitrogen metabolism of moso bamboo.

Nitrogen Additions Affected the Expression of Genes Involved in Nitrogen Metabolism

Nitrogen is absorbed and assimilated by nitrogen metabolic pathway genes, and the most immediate effect caused by nitrogen additions is the transcript change of the genes involved in this pathway (Vidal et al., 2020). According to previous studies, genes encoding NRT1 and NRT2 were rapidly induced by nitrate within hours or days in *Arabidopsis* (Wang et al., 2000). Similarly, the expression of several *NPF* genes was induced in the presence of nitrate, such as *OsNPF7.2* (Hu et al., 2016) and *OsNRT1.1* (Hu et al., 2015). In this study, 10 members of the NPF family were induced by nitrogen additions (**Figure 3**), 32 members were suppressed by nitrogen absence, and 4 members were inhibited by both nitrogen absence and addition (**Figure 3**). This may be related to the functional diversity of NPF family members. Some NPF members transform to low- or high-affinity transporters by phosphorylation to deal with the fluctuant nitrogen environment, such as AtNRT1.1 in *Arabidopsis* and MtNRT1.3 in *Medicago* (*Medicago truncatula*) (Pellizzaro et al., 2014; Wang et al., 2020). It has been shown that *NRT2.1* is induced upon initial nitrate supply and repressed by nitrogen metabolites or high nitrate provision (Muños et al., 2004). In addition, NRT2 could not independently complete the transport of nitrate, and they needed the assistance of NRT3 (Jacquot et al., 2020). In this study, three members of NRT2 were repressed, and one was induced by nitrogen additions (**Figure 3**). Interestingly, the expression pattern of two *NRT3* was similar to that of *NRT2*_PH02Gene31118 but was opposite to that of the other *NRT2* (**Figure 3**). It is speculated that the transport pattern of NRT2 with the partner of NRT3 also existed in moso bamboo, and NRT2 may have dual-affinity transport activity.

NRs play a vital role in nitrogen acquisition, and their mRNAs can be rapidly accumulated in response to nitrate in *Arabidopsis* (Wang et al., 2000). Besides, genes encoding NiR are also induced over a similar range of nitrate concentrations in maize (Trevisan et al., 2011). In bamboo, five NR and two NiR genes were upregulated by nitrogen additions (**Figure 3**), and the activities of NR and NiR were enhanced (**Supplementary Figure S2**), illustrating that the reduction of NR and NiR as well as nitrogen metabolism were strengthened after nitrate addition. With nitrate addition, genes encoding specific isoforms of GS and GOGAT involved in ammonium assimilation were upregulated or downregulated due to their multiple forms in different compartments of the cell (Gaufichon et al., 2016). In *Arabidopsis*, *GLN1;2* coding GS1 was found to be preferentially expressed in the roots and induced by nitrate reduction, while *GLN1;2* and *GLN1;3* did not show marked induction due to a low affinity for ammonium (Konishi et al., 2017). *OsGS1;1* and *OsGS1;2* were expressed in all organs and showed a reciprocal response to ammonium supply in rice roots (Tabuchi et al., 2007). In bamboo, the transcripts of two GS genes were found to be more abundant and had similar expression patterns with most nitrogen metabolic pathway genes after nitrate addition, indicating that they were nitrate induced (**Figure 3**). One GOGAT gene was induced and another was inhibited by nitrate addition. Meanwhile two GDH genes were induced and one was inhibited by nitrate addition (**Figure 3**), indicating the expression diversity of different members in the same family. These results are in accordance with the previous report of GDH genes in tobacco (*Nicotiana tabacum*) (Tercé-Laforgue et al., 2015). In conclusion, nitrogen additions affected the expression of most genes of the nitrogen metabolic pathway, either induced or inhibited, which helped bamboo cope with the fluctuant nitrogen.

TFs and Noncoding RNAs Participated in Nitrogen Metabolism by Regulating Transporter and Enzyme Genes

To date, more than 40 TFs of several families have been identified to be involved in nitrate transport, nitrate reduction, and nitrate assimilation through regulating the expression of genes in nitrogen metabolism (Maeda et al., 2018). Previous studies have shown that *MYB61* was induced by nitrogen and showed a similar expression pattern with nitrogen homeostasis genes (*NRT1.1* and *NIA1*), and MYB bound the promoter of *NRT*, *NiR*, and *GS* (Imamura et al., 2009; Gao et al., 2020). AtNLP6/7 was induced by a nitrate signal, and then bound the promoter of *NRT2.1* and *NIR1* in the presence of nitrate (Liu et al., 2018). LBD37-39 involved in nitrogen metabolism by signaling nitrogen availability, leading to repression of anthocyanin metabolic pathway genes and feedback repression of *NIA* or *NRT2* (Rubin et al., 2009). In this study, 431 TFs belonging to 23 families were identified to participate in nitrogen metabolism, such as MYB, NAC, LBD, WRKY, and C2H2 (**Figure 4**). LBD was predicted to regulate *NPF*, and MYB putatively regulated *NPF*s and *GDH*, due to the predicted binding site found in the promoters as well as the positive correlation of their expression levels (**Figure 4**). In other words, many TFs were involved in bamboo nitrogen metabolism and might play important roles.

Recent research has focused on miRNAs and lncRNAs regulating the expression of specific genes related to nitrogen metabolism (Vidal et al., 2010; Liu et al., 2019). In *Arabidopsis*, *miR827* was involved in the translational repression of NLA,

which directed the ubiquitination of *NRT1.7* and regulated the remobilization of nitrate (Liu et al., 2017). The transcript level assays showed that *miR169o* oppositely regulated *NRT2* under different nitrogen conditions in rice (Yu et al., 2018). In addition, *miR166*, *miR169*, *miR408*, and *miR528* displayed a crucial step in integrating nitrate signals into developmental changes in maize roots upon nitrate shortage (Trevisan et al., 2012). In this study, 383 miRNAs were identified to participate in nitrogen metabolism of moso bamboo, among which four miRNAs targeted three *NPF*s, which was supported by the expression profiles based on the transcriptome data (**Figure 5**). Besides, miRNAs are also involved in the nitrogen metabolism by targeting TFs, which in turn targeted nitrogen metabolic pathway genes. NFYA was targeted by *miR169* while it targeted *NRT1.1* and *NRT2.1* in response to nitrogen changes (Zhao et al., 2011). Also, *miR164* was upregulated under nitrogen starvation and resulted in downregulating the expression of *NAC* (Guo et al., 2005). In this study, 29 DEMs were identified to regulate TFs, such as *osa_miR535_3p* targeted *NF-Y*, 7 DEMs targeted *NAC*, and *novel_miR260* targeted *SPL* (**Supplementary Figure S3**), indicating they play important roles in nitrogen metabolism of moso bamboo.

Although available reports indicated that lncRNAs served as essential regulators of various plant biological processes, studies on lncRNAs involved in nitrogen metabolism are limited. lncRNA *T5120* played an important role in nitrate uptake and nitrate assimilation by regulating the expression of *NTR1.1* in *Arabidopsis*, which resulted in improved crop biomass (Liu et al., 2019). Besides, 2,588 lncRNAs and 388 lncRNAs were identified as participating in nitrogen metabolism in rice and poplar, respectively (Chen et al., 2016; Shin et al., 2018). In this study, 10 lncRNAs were identified, which were speculated to participate in nitrogen metabolism by targeting *NPF*s and *GDH*s directly or indirectly through TFs (**Figure 6**). In total, several miRNAs and lncRNAs involved in nitrogen metabolism of moso bamboo were identified. Further experiments were still needed to verify their functions.

The process of nitrogen metabolism involves complex gene regulation. Based on a series of comprehensive analyses, a multilevel regulatory network, the lncRNA-miRNA-mRNA module of nitrogen metabolism, was successfully established and initially validated in bamboo (**Figures 7**, **8**). The regulatory network helps illustrate the mechanism of nitrogen metabolism in moso bamboo. By manipulating a single TF or lncRNA, targeting multiple genes in the nitrogen metabolism pathway could be realized at the same time, which has a greater effect than overexpression of a single structural gene. The regulatory network provides a more efficient way of manipulating gene expression to improve the nitrogen metabolism of bamboo under different nitrogen supplies.

CONCLUSION

Appropriate nitrogen addition improved nitrogen metabolism and promoted growth of roots, and as a result, Pn and biomass of the seedlings increased. Based on the combination of multiple RNA-seq analyses of the expression profiles of mRNAs, miRNAs, and lncRNAs in bamboo roots under different nitrogen additions, a regulatory network of nitrogen metabolism was constructed, such as 17 nitrogen metabolic pathway genes, 15 TFs, 4 miRNAs, and 10 lncRNAs. The lncRNA-miRNA-mRNA network reveals the regulation mechanism of nitrogen metabolism in moso bamboo and provides candidate gene resources for improving the ability of bamboo to adapt to a fluctuating nitrogen environment.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here https://www.ncbi.nlm.nih.gov/ PRJNA797724 AND PRJNA797734.

AUTHOR CONTRIBUTIONS

ZG and XS conceived and designed the study. TY and CZ performed the experiments. TY, GL, and YL carried out the data analysis. TY prepared the manuscript. KY and ZL performed a critical review of the intellectual content.

FUNDING

This research was funded by the National Natural Science Foundation of China, grant numbers 31930075 and 31971736 and the National Key Research and Development Program of China, grant number 2021YFD2200502.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.854346/full#supplementary-material

REFERENCES

Alfatih, A., Wu, J., Zhang, Z.-S., Xia, J.-Q., Jan, S. U., Yu, L.-H., et al. (2020). Rice NIN-LIKE PROTEIN 1 Rapidly Responds to Nitrogen Deficiency and Improves Yield and Nitrogen Use Efficiency. *J. Exp. Bot.* 71, 6032–6042. doi:10.1093/jxb/eraa292

Brooks, M. D., Cirrone, J., Pasquino, A. V., Alvarez, J. M., Swift, J., Mittal, S., et al. (2019). Network Walking Charts Transcriptional Dynamics of Nitrogen Signaling by Integrating Validated and Predicted Genome-wide Interactions. *Nat. Commun.* 10, 1569. doi:10.1038/s41467-019-09522-1

Chen, J., Zhang, Y., Tan, Y., Zhang, M., Zhu, L., Xu, G., et al. (2016a). Agronomic Nitrogen-Use Efficiency of rice Can Be Increased by drivingOsNRT2.1expression with theOsNAR2.1promoter. *Plant Biotechnol. J.* 14, 1705–1715. doi:10.1111/pbi.12531

Chen, M., Wang, C., Bao, H., Chen, H., and Wang, Y. (2016b). Genome-wide Identification and Characterization of Novel lncRNAs in *Populus* under

Nitrogen Deficiency. *Mol. Genet. Genomics* 291, 1663–1680. doi:10.1007/s00438-016-1210-3

Ding, Y., Chen, Z., and Zhu, C. (2011). Microarray-based Analysis of Cadmium-Responsive microRNAs in rice (*Oryza Sativa*). *J. Exp. Bot.* 62, 3563–3573. doi:10.1093/jxb/err046

Fan, C., Ma, J., Guo, Q., Li, X., Wang, H., and Lu, M. (2013). Selection of Reference Genes for Quantitative Real-Time PCR in Bamboo (*Phyllostachys Edulis*). *PLoS One* 8, e56573. doi:10.1371/journal.pone.0056573

Fang, Z., Xia, K., Yang, X., Grotemeyer, M. S., Meier, S., Rentsch, D., et al. (2012). Altered Expression of thePTR/NRT1homologueOsPTR9affects Nitrogen Utilization Efficiency, Growth and Grain Yield in rice. *Plant Biotechnol. J.* 11, 446–458. doi:10.1111/pbi.12031

Feng, H., Yan, M., Fan, X., Li, B., Shen, Q., Miller, A. J., et al. (2011). Spatial Expression and Regulation of rice High-Affinity Nitrate Transporters by Nitrogen and Carbon Status. *J. Exp. Bot.* 62, 2319–2332. doi:10.1093/jxb/erq403

Ferrari, A. (1960). Nitrogen Determination by a Continuous Digestion and Analysis System. *Ann. N. Y Acad. Sci.* 87, 792–800. doi:10.1111/j.1749-663210.1111/j.1749-6632.1960.tb23236.x

Ferreira, S., Moreira, E., Amorim, I., Santos, C., and Melo, P. (2019). *Arabidopsis thaliana* Mutants Devoid of Chloroplast Glutamine Synthetase (GS2) Have Non-lethal Phenotype under Photorespiratory Conditions. *Plant Physiol. Biochem.* 144, 365–374. doi:10.1016/j.plaphy.2019.10.009

Fischer, J. J., Beatty, P. H., Good, A. G., and Muench, D. G. (2013). Manipulation of microRNA Expression to Improve Nitrogen Use Efficiency. *Plant Sci.* 210, 70–81. doi:10.1016/j.plantsci.2013.05.009

Fu, X.-Z., Zhang, X.-Y., Qiu, J.-Y., Zhou, X., Yuan, M., He, Y.-Z., et al. (2019). Whole-transcriptome RNA Sequencing Reveals the Global Molecular Responses and ceRNA Regulatory Network of mRNAs, lncRNAs, miRNAs and circRNAs in Response to Copper Toxicity in Ziyang Xiangcheng (*Citrus Junos* Sieb. Ex Tanaka). *BMC Plant Biol.* 19, 509. doi:10.1186/s12870-019-2087-1

Fukuda, M., Fujiwara, T., and Nishida, S. (2020). Roles of Non-coding RNAs in Response to Nitrogen Availability in Plants. *Ijms* 21, 8508. doi:10.3390/ijms21228508

Gao, Y., Xu, Z., Zhang, L., Li, S., Wang, S., Yang, H., et al. (2020). MYB61 Is Regulated by GRF4 and Promotes Nitrogen Utilization and Biomass Production in rice. *Nat. Commun.* 11, 5219. doi:10.1038/s41467-020-19019-x

Gao, Z., Wang, Y., Chen, G., Zhang, A., Yang, S., Shang, L., et al. (2019). The Indica Nitrate Reductase Gene *OsNR2* Allele Enhances rice Yield Potential and Nitrogen Use Efficiency. *Nat. Commun.* 10, 5207. doi:10.1038/s41467-019-13110-8

Garg, V., Khan, A. W., Kudapa, H., Kale, S. M., Chitikineni, A., Qiwei, S., et al. (2019). Integrated Transcriptome, Small RNA and Degradome Sequencing Approaches Provide Insights into Ascochyta Blight Resistance in Chickpea. *Plant Biotechnol. J.* 17, 914–931. doi:10.1111/pbi.13026

Gaudinier, A., Rodriguez-Medina, J., Zhang, L., Olson, A., Liseron-Monfils, C., Bågman, A.-M., et al. (2018). Transcriptional Regulation of Nitrogen-Associated Metabolism and Growth. *Nature* 563, 259–264. doi:10.1038/s41586-018-0656-3

Gaufichon, L., Rothstein, S. J., and Suzuki, A. (2016). Asparagine Metabolic Pathways in Arabidopsis. *Plant Cel Physiol* 57, 675–689. doi:10.1093/pcp/pcv184

Guo, H.-S., Xie, Q., Fei, J.-F., and Chua, N.-H. (2005). MicroRNA Directs mRNA Cleavage of the Transcription Factor NAC1 to Downregulate Auxin Signals for Arabidopsis Lateral Root Development. *Plant Cell* 17, 1376–1386. doi:10.1105/tpc.105.030841

Gupta, S., Akhatar, J., Kaur, P., Sharma, A., Sharma, P., Mittal, M., et al. (2019). Genetic Analyses of Nitrogen Assimilation Enzymes in *Brassica Juncea* (L.) Czern & Coss. *Mol. Biol. Rep.* 46, 4235–4244. doi:10.1007/s11033-019-04878-5

Hou, D., Lu, H., Zhao, Z., Pei, J., Yang, H., Wu, A., et al. (2022). Integrative Transcriptomic and Metabolomic Data Provide Insights into Gene Networks Associated with Lignification in Postharvest Lei Bamboo Shoots under Low Temperature. *Food Chem.* 368, 130822. doi:10.1016/j.foodchem.2021.130822

Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., et al. (2015). Variation in *NRT1.1B* Contributes to Nitrate-Use Divergence between rice Subspecies. *Nat. Genet.* 47, 834–838. doi:10.1038/ng.3337

Hu, R., Qiu, D., Chen, Y., Miller, A. J., Fan, X., Pan, X., et al. (2016). Knock-down of a Tonoplast Localized Low-Affinity Nitrate Transporter OsNPF7.2 Affects rice Growth under High Nitrate Supply. *Front. Plant Sci.* 7, 1529. doi:10.3389/fpls.2016.01529

Hu, Y., FernÃ¡ndez, V., and Ma, L. (2014). Nitrate Transporters in Leaves and Their Potential Roles in Foliar Uptake of Nitrogen Dioxideâ€. *Front. Plant Sci.* 5, 360. doi:10.3389/fpls.2014.00360

Huarancca Reyes, T., Scartazza, A., Pompeiano, A., Ciurli, A., Lu, Y., Guglielminetti, L., et al. (2018). Nitrate Reductase Modulation in Response to Changes in C/N Balance and Nitrogen Source in Arabidopsis. *Plant Cel Physiol* 59, 1248–1254. doi:10.1093/pcp/pcy065

Imamura, S., Kanesaki, Y., Ohnuma, M., Inouye, T., Sekine, Y., Fujiwara, T., et al. (2009). R2R3-type MYB Transcription Factor, CmMYB1, Is a central Nitrogen Assimilation Regulator in *Cyanidioschyzon Merolae*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12548–12553. doi:10.1073/pnas.0902790106

Imran, M., Hu, C., Hussain, S., Rana, M. S., Riaz, M., Afzal, J., et al. (2019). Molybdenum-induced Effects on Photosynthetic Efficacy of winter Wheat (*Triticum aestivum* L.) under Different Nitrogen Sources Are Associated with Nitrogen Assimilation. *Plant Physiol. Biochem.* 141, 154–163. doi:10.1016/j.plaphy.2019.05.024

Jacquot, A., Chaput, V., Mauries, A., Li, Z., Tillard, P., Fizames, C., et al. (2020). NRT2.1 C-terminus Phosphorylation Prevents Root High Affinity Nitrate Uptake Activity in *Arabidopsis thaliana*. *New Phytol.* 228, 1038–1054. doi:10.1111/nph.16710

Jin, Z., Zhu, Y., Li, X., Dong, Y., and An, Z. (2015). Soil N Retention and Nitrate Leaching in Three Types of Dunes in the Mu Us Desert of China. *Sci. Rep.* 5, 14222. doi:10.1038/srep14222

Konishi, N., Ishiyama, K., Beier, M. P., Inoue, E., Kanno, K., Yamaya, T., et al. (2017). Contributions of Two Cytosolic Glutamine Synthetase Isozymes to Ammonium Assimilation in Arabidopsis Roots. *Exbotj* 68, erw454–625. doi:10.1093/jxb/erw454

Li, C., Qi, W., Liang, Z., Yang, X., Ma, Z., and Song, R. (2020). A SnRK1-*Zm*RFWD3-Opaque2 Signaling axis Regulates Diurnal Nitrogen Accumulation in maize Seeds. *Plant Cell* 32, 2823–2841. doi:10.1105/tpc.20.00352

Li, L., Li, Q., Davis, K. E., Patterson, C., Oo, S., Liu, W., et al. (2021b). Response of Root Growth and Development to Nitrogen and Potassium Deficiency as Well as microRNA-Mediated Mechanism in Peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* 12, 695234. doi:10.3389/fpls.2021.695234

Li, Y., and Feng, P. (2019). Bamboo Resources in China Based on the Ninth National forest Inventory Data. *World Bamboo and Rattan* 17, 45. doi:10.12168/sjzttx.2019.06.010

Li, Z., Yuan, T., Zhu, C., Yang, K., Song, X., and Gao, Z. (2021a). Analysis of Molecular Characteristics and Gene Expression Pattern of Ammonium Nitrogen Transporter in Moso Bamboo. *Scientia Silvae Sinicae* 57, 70–79. doi:10.11707/j.1001-7488.20210708

Lin, W.-Y., Huang, T.-K., and Chiou, T.-J. (2013). Nitrogen Limitation Adaptation, a Target of microRNA827, Mediates Degradation of Plasma Membrane-Localized Phosphate Transporters to Maintain Phosphate Homeostasis in Arabidopsis. *Plant Cell* 25, 4061–4074. doi:10.1105/tpc.113.116012

Lin, X. Y., Liu, X. X., Zhang, Y. P., Zhou, Y. Q., Hu, Y., Chen, Q. H., et al. (2014). Short-term Alteration of Nitrogen Supply Prior to Harvest Affects Quality in Hydroponic-Cultivated Spinach (Spinacia Oleracea). *J. Sci. Food Agric.* 94, 1020–1025. doi:10.1002/jsfa.6368

Liu, F., Xu, Y., Chang, K., Li, S., Liu, Z., Qi, S., et al. (2019). The Long Noncoding RNA *T5120* Regulates Nitrate Response and Assimilation in Arabidopsis. *New Phytol.* 224, 117–131. doi:10.1111/nph.16038

Liu, K.-h., Niu, Y., Konishi, M., Wu, Y., Du, H., Sun Chung, H., et al. (2017a). Discovery of Nitrate-CPK-NLP Signalling in central Nutrient-Growth Networks. *Nature* 545, 311–316. doi:10.1038/nature22077

Liu, M., Chang, W., Fan, Y., Sun, W., Qu, C., Zhang, K., et al. (2018). Genome-wide Identification and Characterization of *NODULE-INCEPTION-like Protein* (*NLP*) Family Genes in *Brassica Napus*. *Int. J. Mol. Sci.* 19, 2270. doi:10.3390/ijms19082270

Liu, W., Sun, Q., Wang, K., Du, Q., and Li, W. X. (2017b). Nitrogen Limitation Adaptation (NLA) Is Involved in Source-to-sink Remobilization of Nitrate

by Mediating the Degradation of NRT 1.7 in Arabidopsis. *New Phytol.* 214, 734–744. doi:10.1111/nph.14396

Luo, J., Li, H., Liu, T., Polle, A., Peng, C., and Luo, Z.-B. (2013). Nitrogen Metabolism of Two Contrasting poplar Species during Acclimation to Limiting Nitrogen Availability. *J. Exp. Bot.* 64 (14), 4207–4224. doi:10.1093/jxb/ert234

Ma, X., Zhao, H., Xu, W., You, Q., Yan, H., Gao, Z., et al. (2018). Co-expression Gene Network Analysis and Functional Module Identification in Bamboo Growth and Development. *Front. Genet.* 9, 574. doi:10.3389/fgene.2018.00574

Maeda, Y., Konishi, M., Kiba, T., Sakuraba, Y., Sawaki, N., Kurai, T., et al. (2018). A NIGT1-Centred Transcriptional cascade Regulates Nitrate Signalling and Incorporates Phosphorus Starvation Signals in Arabidopsis. *Nat. Commun.* 9, 1376. doi:10.1038/s41467-018-03832-6

Mu, X., Chen, Q., Chen, F., Yuan, L., and Mi, G. (2016). Within-leaf Nitrogen Allocation in Adaptation to Low Nitrogen Supply in maize during Grain-Filling Stage. *Front. Plant Sci.* 7, 699. doi:10.3389/fpls.2016.00699

Muños, S., Cazettes, C., Fizames, C., Gaymard, F., Tillard, P., Lepetit, M., et al. (2004). Transcript Profiling in the Chl1-5 Mutant of Arabidopsis Reveals a Role of the Nitrate Transporter NRT1.1 in the Regulation of Another Nitrate Transporter, NRT2.1[W]. *Plant Cell* 16, 2433–2447. doi:10.1105/tpc.104.024380

Nguyen, G. N., Rothstein, S. J., Spangenberg, G., and Kant, S. (2015). Role of microRNAs Involved in Plant Response to Nitrogen and Phosphorous Limiting Conditions. *Front. Plant Sci.* 6, 629. doi:10.3389/fpls.2015.00629

Patterson, K., Cakmak, T., Cooper, A., Lager, I., Rasmusson, A. G., and Escobar, M. A. (2010). Distinct Signalling Pathways and Transcriptome Response Signatures Differentiate Ammonium- and Nitrate-Supplied Plants. *Plant Cel Environ* 33, 1486–1501. doi:10.1111/j.1365-3040.2010.02158

Paul, S., Datta, S. K., and Datta, K. (2015). MiRNA Regulation of Nutrient Homeostasis in Plants. *Front. Plant Sci.* 06, 232. doi:10.3389/fpls.2015.00232

Pellizzaro, A., Clochard, T., Cukier, C., Bourdin, C., Juchaux, M., Montrichard, F., et al. (2014). The Nitrate Transporter MtNPF6.8 (MtNRT1.3) Transports Abscisic Acid and Mediates Nitrate Regulation of Primary Root Growth in *Medicago Truncatula*. *Plant Physiol.* 166, 2152–2165. doi:10.1104/pp.114.250811

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level Expression Analysis of RNA-Seq Experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11 (9), 1650–1667. doi:10.1038/nprot.2016.095

Rubin, G., Tohge, T., Matsuda, F., Saito, K., and Scheible, W.-R. (2009). Members of theLBDFamily of Transcription Factors Repress Anthocyanin Synthesis and Affect Additional Nitrogen Responses inArabidopsis. *Plant Cell* 21, 3567–3584. doi:10.1105/tpc.109.067041

Sanagi, M., Aoyama, S., Kubo, A., Lu, Y., Sato, Y., Ito, S., et al. (2021). Low Nitrogen Conditions Accelerate Flowering by Modulating the Phosphorylation State of Flowering bHLH 4 in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2022942118. doi:10.1073/pnas.2022942118

Shin, S.-Y., Jeong, J. S., Lim, J. Y., Kim, T., Park, J. H., Kim, J.-K., et al. (2018). Transcriptomic Analyses of rice (*Oryza Sativa*) Genes and Non-coding RNAs under Nitrogen Starvation Using Multiple Omics Technologies. *BMC Genomics* 19, 532. doi:10.1186/s12864-018-4897-1

Song, X., Peng, C., Ciais, P., Li, Q., Xiang, W., Xiao, W., et al. (2020). Nitrogen Addition Increased CO 2 Uptake More Than Non-CO 2 Greenhouse Gases Emissions in a Moso Bamboo forest. *Sci. Adv.* 6, eaaw5790. doi:10.1126/sciadv.aaw5790

Song, X., Zhou, G., Jiang, H., Yu, S., Fu, J., Li, W., et al. (2011). Carbon Sequestration by Chinese Bamboo Forests and Their Ecological Benefits: Assessment of Potential, Problems, and Future Challenges. *Environ. Rev.* 19, 418–428. doi:10.1139/a11-015

Tabuchi, M., Abiko, T., and Yamaya, T. (2007). Assimilation of Ammonium Ions and Reutilization of Nitrogen in rice (*Oryza Sativa* L.). *J. Exp. Bot.* 58, 2319–2327. doi:10.1093/jxb/erm016

Tercé-Laforgue, T., Clément, G., Marchi, L., Restivo, F. M., Lea, P. J., and Hirel, B. (2015). Resolving the Role of Plant NAD-Glutamate Dehydrogenase: III. Overexpressing Individually or Simultaneously the Two Enzyme Subunits under Salt Stress Induces Changes in the Leaf Metabolic Profile and Increases Plant Biomass Production. *Plant Cel Physiol* 56, 1918–1929. doi:10.1093/pcp/pcv114

The, S. V., Snyder, R., and Tegeder, M. (2020). Targeting Nitrogen Metabolism and Transport Processes to Improve Plant Nitrogen Use Efficiency. *Front. Plant Sci.* 11, 628366. doi:10.3389/fpls.2020.628366

Tian, F., Yang, D.-C., Meng, Y.-Q., Jin, J., and Gao, G. (2019). PlantRegMap: Charting Functional Regulatory Maps in Plants. *Nucleic Acids Res.* 48, D1104–D1113. doi:10.1093/nar/gkz1020

Trevisan, S., Begheldo, M., Nonis, A., and Quaggiotti, S. (2012). The miRNA-Mediated post-transcriptional Regulation of maize Response to Nitrate. *Plant Signaling Behav.* 7, 822–826. doi:10.4161/psb.20462

Trevisan, S., Manoli, A., Begheldo, M., Nonis, A., Enna, M., Vaccaro, S., et al. (2011). Transcriptome Analysis Reveals Coordinated Spatiotemporal Regulation of Hemoglobin and Nitrate Reductase in Response to Nitrate in maize Roots. *New Phytol.* 192, 338–352. doi:10.1111/j.1469-8137.2011.03822.x

Varala, K., Marshall-Colón, A., Cirrone, J., Brooks, M. D., Pasquino, A. V., Léran, S., et al. (2018). Temporal Transcriptional Logic of Dynamic Regulatory Networks Underlying Nitrogen Signaling and Use in Plants. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6494–6499. doi:10.1073/pnas.1721487115

Vidal, E. A., Araus, V., Lu, C., Parry, G., Green, P. J., Coruzzi, G. M., et al. (2010). Nitrate-responsive miR393/AFB3 Regulatory Module Controls Root System Architecture in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4477–4482. doi:10.1073/pnas.0909571107

Vidal, E. A., Araus, V., Riveras, E., Brooks, M. D., Krouk, G., Ruffel, S., et al. (2020). Nitrate in 2020: Thirty Years from Transport to Signaling Networks. *Plant Cell* 32, 2094–2119. doi:10.1105/tpc.19.00748

Wang, R., Guegler, K., LaBrie, S. T., and Crawford, N. M. (2000). Genomic Analysis of a Nutrient Response in Arabidopsis Reveals Diverse Expression Patterns and Novel Metabolic and Potential Regulatory Genes Induced by Nitrate. *Plant Cell* 12, 1491–1509. doi:10.1105/tpc.12.8.1491

Wang, R., Qian, J., Fang, Z., and Tang, J. (2020a). Transcriptomic and Physiological Analyses of rice Seedlings under Different Nitrogen Supplies Provide Insight into the Regulation Involved in Axillary Bud Outgrowth. *BMC Plant Biol.* 20, 197. doi:10.1186/s12870-020-02409-0

Wang, W., Hu, B., Li, A., and Chu, C. (2020b). NRT1.1s in Plants: Functions beyond Nitrate Transport. *J. Exp. Bot.* 71, 4373–4379. doi:10.1093/jxb/erz554

Wang, W., Hu, B., Yuan, D., Liu, Y., Che, R., Hu, Y., et al. (2018). Expression of the Nitrate Transporter Gene *OsNRT1.1a/OsNPF6.3* Confers High Yield and Early Maturation in rice. *Plant Cell* 30, 638–651. doi:10.1105/tpc.17.00809

Wu, J., Zhang, Z. S., Xia, J. Q., Alfatih, A., Song, Y., Huang, Y. J., et al. (2021). Rice NIN-LIKE PROTEIN 4 Plays a Pivotal Role in Nitrogen Use Efficiency. *Plant Biotechnol. J.* 19, 448–461. doi:10.1111/pbi.13475

Xia, X., Fan, X., Wei, J., Feng, H., Qu, H., Xie, D., et al. (2015). Rice Nitrate Transporter OsNPF2.4 Functions in Low-Affinity Acquisition and Long-Distance Transport. *J. Exp. Bot.* 66, 317–331. doi:10.1093/jxb/eru425

Xu, G., Fan, X., and Miller, A. J. (2012). Plant Nitrogen Assimilation and Use Efficiency. *Annu. Rev. Plant Biol.* 63, 153–182. doi:10.1146/annurev-arplant-042811-105532

Xu, N., Wang, R., Zhao, L., Zhang, C., Li, Z., Lei, Z., et al. (2016). The Arabidopsis NRG2 Protein Mediates Nitrate Signaling and Interacts with and Regulates Key Nitrate Regulators. *Plant Cell* 28, 485–504. doi:10.1105/tpc.15.00567

Yang, K., Li, L., Lou, Y., Zhu, C., Li, X., and Gao, Z. (2021). A Regulatory Network Driving Shoot Lignification in Rapidly Growing Bamboo. *Plant Physiol.* 187, 900–916. doi:10.1093/plphys/kiab289

Yu, C., Chen, Y., Cao, Y., Chen, H., Wang, J., Bi, Y.-M., et al. (2018). Overexpression of *miR169o*, an Overlapping microRNA in Response to Both Nitrogen Limitation and Bacterial Infection, Promotes Nitrogen Use Efficiency and Susceptibility to Bacterial Blight in rice. *Plant Cel Physiol* 59, 1234–1247. doi:10.1093/pcp/pcy060

Yuan, T., Zhu, C., Li, Z., Song, X., and Gao, Z. (2021b). Identification of NLP Transcription Factors of *Phyllostachys Edulis* and Their Expression Patterns in Response to Nitrogen. *For. Res.* 34, 40–49. doi:10.13275/j.cnki.lykxyj.2021.005.005

Yuan, T., Zhu, C., Yang, K., Song, X., and Gao, Z. (2021a). Identification of Nitrate Transporter Gene Family *PeNPF*s and Their Expression Analysis in *Phyllostachys Edulis*. *For. Res.* 34, 39–48. doi:10.13275/j.cnki.lykxyj.2021.03.001

Zhai, R., Ye, S., Zhu, G., Lu, Y., Ye, J., Yu, F., et al. (2020). Identification and Integrated Analysis of Glyphosate Stress-Responsive microRNAs, lncRNAs, and mRNAs in rice Using Genome-wide High-Throughput Sequencing. *BMC Genomics* 21, 238. doi:10.1186/s12864-020-6637-6

Zhang, T.-T., Kang, H., Fu, L.-L., Sun, W.-J., Gao, W.-S., You, C.-X., et al. (2021). NIN-like Protein 7 Promotes Nitrate-Mediated Lateral Root Development by Activating Transcription of Tryptophan Aminotransferase Related 2. *Plant Sci.* 303, 110771. doi:10.1016/j.plantsci.2020.110771

Zhang, Y.-C., Yu, Y., Wang, C.-Y., Li, Z.-Y., Liu, Q., Xu, J., et al. (2013). Overexpression of microRNA *OsmiR397* Improves rice Yield by Increasing Grain Size and Promoting Panicle Branching. *Nat. Biotechnol.* 31, 848–852. doi:10.1038/nbt.2646

Zhang, Z., Jiang, L., Wang, J., Gu, P., and Chen, M. (2014). MTide: an Integrated Tool for the Identification of miRNA-Target Interaction in Plants. *Bioinformatics* 31, 290–291. btu633. doi:10.1093/bioinformatics/btu633

Zhao, H., Gao, Z., Wang, L., Wang, J., Wang, S., Fei, B., et al. (2018). Chromosome-level Reference Genome and Alternative Splicing Atlas of Moso Bamboo (*Phyllostachys Edulis*). *GigaScience* 7, 1–12. doi:10.1093/gigascience/giy115

Zhao, M., Ding, H., Zhu, J. K., Zhang, F., and Li, W. X. (2011). Involvement of miR169 in the Nitrogen-starvation Responses in Arabidopsis. *New Phytol.* 190, 906–915. doi:10.1111/j.1469-8137.2011.03647.x

Zhong, C., Jian, S.-F., Huang, J., Jin, Q.-Y., and Cao, X.-C. (2019). Trade-off of Within-Leaf Nitrogen Allocation between Photosynthetic Nitrogen-Use Efficiency and Water Deficit Stress Acclimation in rice (Oryza Sativa L.). *Plant Physiol. Biochem.* 135, 41–50. doi:10.1016/j.plaphy.2018.11.021

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 30 June 2022
doi: 10.3389/fgene.2022.930113

In silico Identification of miRNAs and Their Targets in Cluster Bean for Their Role in Development and Physiological Responses

*Vrantika Chaudhary, Sumit Jangra and Neelam R. Yadav**

Department of Molecular Biology, Biotechnology and Bioinformatics, CCS Haryana Agricultural University, Hisar, India

Cluster bean popularly known as "guar" is a drought-tolerant, annual legume that has recently emerged as an economically important crop, owing to its high protein and gum content. The guar gum has wide range of applications in food, pharma, and mining industries. India is the leading exporter of various cluster bean-based products all across the globe. Non-coding RNAs (miRNAs) are involved in regulating the expression of the target genes leading to variations in the associated pathways or final protein concentrations. The understanding of miRNAs and their associated targets in cluster bean is yet to be used to its full potential. In the present study, cluster bean EST (Expressed Sequence Tags) database was exploited to identify the miRNA and their predicted targets associated with metabolic and biological processes especially response to diverse biotic and abiotic stimuli using *in silico* approach. Computational analysis based on cluster bean ESTs led to the identification of 57 miRNAs along with their targets. To the best of our knowledge, this is the first report on identification of miRNAs and their targets using ESTs in cluster bean. The miRNA related to gum metabolism was also identified. Most abundant miRNA families predicted in our study were miR156, miR172, and miR2606. The length of most of the mature miRNAs was found to be 21nt long and the range of minimal folding energy (MFE) was 5.8–177.3 (–kcal/mol) with an average value of 25.4 (–kcal/mol). The identification of cluster bean miRNAs and their targets is predicted to hasten the miRNA discovery, resulting in better knowledge of the role of miRNAs in cluster bean development, physiology, and stress responses.

Keywords: micro RNAs (miRNAs), cluster bean (*Cyamopsis tetragonoloba* L. Taub.), miRNA identification, miRNA targets, non coding RNAs, galactomannan

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Ruili Li,
Beijing Forestry University, China
Gursharn Singh Randhawa,
Sardar Bhagwan Singh Post Graduate Institute of Biomedical Science and Research, India

****Correspondence:***
Neelam R. Yadav
nryadav58@gmail.com

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *27 April 2022*
Accepted: *08 June 2022*
Published: *30 June 2022*

Citation:
Chaudhary V, Jangra S and Yadav NR (2022) In silico Identification of miRNAs and Their Targets in Cluster Bean for Their Role in Development and Physiological Responses. Front. Genet. 13:930113. doi: 10.3389/fgene.2022.930113

1 INTRODUCTION

Cluster bean (*Cyamopsis tetragonoloba* L. Taub.) popularly known as guar is a drought-tolerant, large-seeded legume, mainly adapted to semi-arid parts of the world with minimal requirements. It is commonly used as a source of food and feed in the Indian subcontinent and is a good source of nutrition, as it is a rich source of dietary fiber, folic acid, and Vitamin C (Rajaprakasam et al., 2021). India accounts for 80% of the global cluster bean production making it second-largest cash crop in India (Sahu et al., 2020). In the past two decades, the demand for cluster bean production has increased and cultivation has also been initiated in non-traditional areas like the United States, Australia, Brazil, Italy, Spain, Morocco and Germany, most recently, Russia due to the discovery of a

valuable polysaccharide-galactomannan (guar gum), in the seed endosperm (Teplyakova et al., 2019; Grigoreva et al., 2021a). The gel-like structure formed by hydrated gum has a wide utility in various industries like paper, food and agriculture, cosmetics, explosives, petrochemicals, and water purification where it is used as thickener, stiffener, stabilizer, and strengthening agent (Mudgil et al., 2014; Sharma et al., 2018; Saya et al., 2021). The demand for guar gum has increased by four times in crude oil industry since 2006 (Bhatt et al., 2016). The medicinal properties of cluster bean viz. antihyperlipidemic, antihyperglycemic, and antihypercholesterolemic have made it an important constituent of pharma industry. Recently, it has been used in nanotechnology-based drug delivery and as binder in electrochemical industry (George et al., 2019; Verma and Sharma, 2021; Kaur and Santra, 2022). The global demands for cluster bean in various industries have increased several folds in the past decade making it a crop of high economic importance.

Despite such high economic importance this crop has not been explored and characterized to its full extent at genomic level. Though, recently some studies have been reported (Kaila et al., 2017; Rawal et al., 2017; Tanwar et al., 2017; Al-Qurainy et al., 2019; Chaudhury et al., 2019; Sahu et al., 2020; Rajaprakasam et al., 2021) and the availability of the whole genome sequence which identified 34,680 protein-coding genes from 550.31 Mbp of genome will further help in accelerating the fundamental and applied research in this crop (Gaikwad et al., 2020 Unpublished). When it comes to the non-coding genome, especially microRNA the research is still at its initial stages and yet to be explored to its full potential. miRNAs are single-stranded RNAs with a length of 21–25 nucleotides (nt) that are formed from stem-loop precursors and are highly conserved (Carthew and Sontheimer, 2009; O'Brien et al., 2018). These small non-coding RNAs (miRNAs) are involved in regulation of growth, development, and metabolic pathways. In Leguminosae family, a wide range of miRNAs has been reported from nine species (Kozomara and Griffiths-Jones, 2014). miRNAs and their precursors in response to various physiological stresses have also been reported in important cereals like rice, wheat, maize, sorghum, and barley (Budak and Akpinar, 2015; Djami-Tchatchou et al., 2017; Singroha et al., 2021). All these studies showed that miRNAs play an important role in regulation of physiological responses, developmental and metabolic pathways, and evolution in plants.

When genome sequence is unavailable, EST (expressed sequence tags) and GSS (genome survey sequences) data can be utilized for identification of miRNA and their targets. *In silico* miRNA prediction and characterization are developing as faster and more efficient methods than laboratory-based cloning methods, as they are cheap and quick (Gupta et al., 2017). Such *in silico* methods based on EST and GSS have been used to identify miRNAs and their target in various crop plants including cotton, soybean, lentil, pepper, potato, sweet potato, tomato, tobacco, and horse gram (Qiu et al., 2007; Zhang et al., 2008; Kim et al., 2011; Budheswar et al., 2013; Dehury et al., 2013; Vivek, 2018; Yasin et al., 2020). However, no miRNAs for the cluster bean have been reported in miRBase, and these have yet to be properly studied. Though, few recent studies report the identification of various miRNA families in cluster bean (Tyagi et al., 2018; Sahu et al., 2018). In the present study, we used EST-based identification and characterization of conserved miRNAs belonging to different families and their probable target genes. Predicted miRNAs were also functionally annotated using network analysis and gene ontology. Understanding the role of miRNAs in the control of various biological processes necessitates the study of interactions between individual miRNAs and their target genes (Bansal et al., 2017). Further, the discovery of cluster bean miRNAs and their targets is predicted to hasten miRNA discovery, resulting in a better knowledge of the role of miRNAs in cluster bean development, physiology, and stress responses.

2 MATERIALS AND METHODS

2.1 Sequence Collection and Software Information

A total of 16,503 EST sequences of cluster bean were downloaded from the GenBank nucleotide database (http://www.ncbi.nlm.nih.gov/nucest) available at NCBI. To identify conserved miRNAs genes and their targets, C-mii (version 1.11) programme was used (Numnark et al., 2012). To determine putative miRNA sequences, 4,256 non-redundant mature plant miRNA sequences of viridiplantae (green plants) were also downloaded from miRBASE. BLASTN with an e-value cut-off equal to 10 was performed against the EST sequence of cluster bean and mature miRNA sequences from miRBASE. Next, to remove protein-coding sequences BLASTX with e value b = $1e^{-5}$ was done against UniprotKB/Swissprot (release 2010-12) and UniProtKB/TrEMBL (release 2011-01) protein databases. The primary and precursor miRNAs structures were predicted using RNA database Rfam 10. UNAFold was run at a maximum base pair distance of 3,000, maximum bulge/ interior loop size 30, and single read run of 37°C temperature.

2.2 Prediction of the Secondary Structure of Cluster Bean miRNAs

miRNA sequences that matched at least 18 nt and had a 3 nt mismatch with all known plant mature miRNAs were chosen for further study. The top BLAST hits were collected, and the BLASTX tool was used to execute the anticipated precursor sequences. Non-coding sequences were preserved, while protein-coding precursor sequences were removed. Rfam (10) was used to differentiate miRNA from other RNA families (rRNA, snRNA, and tRNA). UNAFold was used to anticipate secondary structure using parameters such as maximum base pair distance (3,000), maximum bulge/interior loop size (30), and single tread run temperature (37°C).

Using a homology search, the following criteria were utilized to find miRNAs: 1) The predicted mature miRNAs should be 19–25 nucleotides in length, 2) Maximum of four mismatches against the reference miRNA were allowed, 3) The mature miRNA should be localized within a stem-loop structure with one arm, 4) No more than five mismatches were allowed between

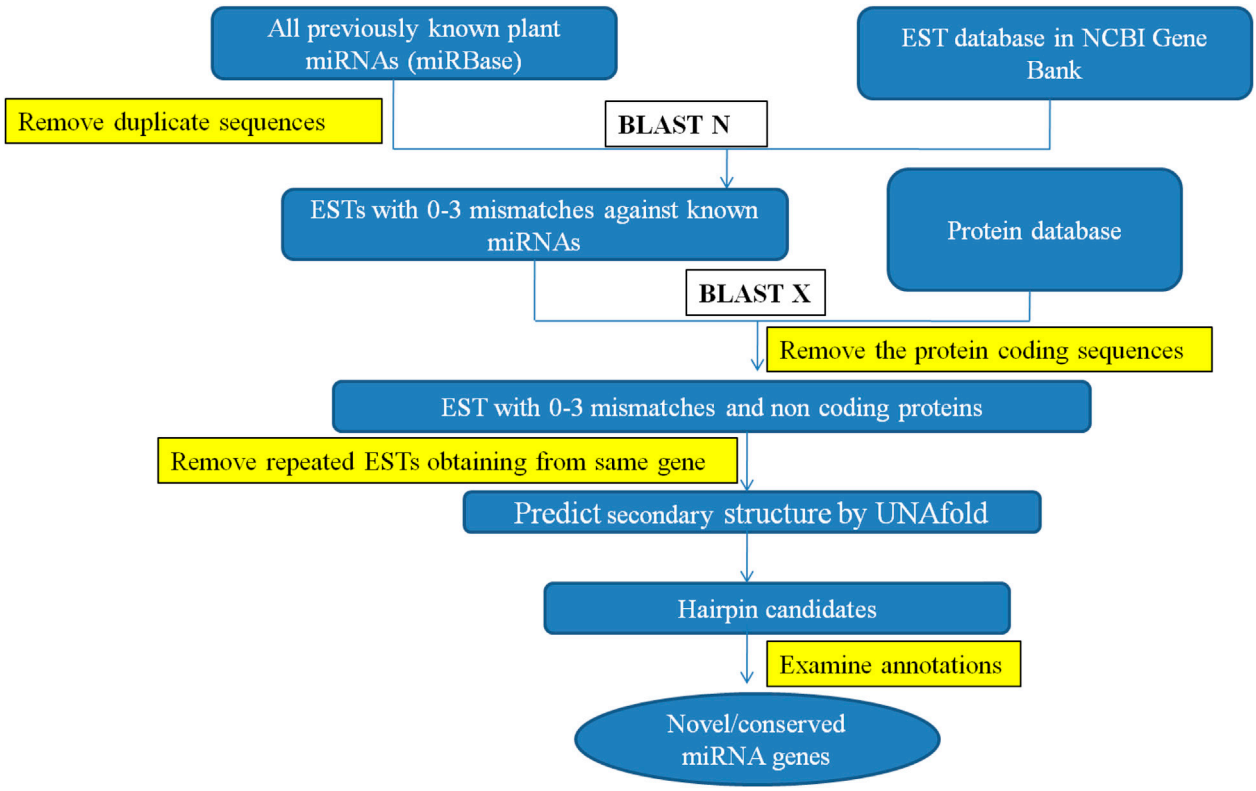


FIGURE 1 | Workflow for identification of miRNA and their targets in cluster bean.

miRNA sequences and the guide miRNA sequence, 5) The A + U content should be high, and the secondary structure's minimal folding free energy (MFE) and MFE index (MFEI) values should be significantly negative.

The MFEI was calculated using the following equation-

MFEI = [(MFE/length of the RNA sequence)∗100]/(G + C)%

2.3 Prediction of miRNA Targets

C-mii and the Plant Target Prediction Tool, both available on the UEA sRNA toolkit (srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi), were used to estimate the likely target locations of detected miRNAs. miRNA-targeted mRNAs have perfectly or almost perfectly complementary sites with miRNAs and miRNAs suppress gene expression by binding to these targeted mRNAs at their complementary sites for direct mRNA cleavage or protein translation suppression (Bartel, 2004; Chen et al., 2004). This suggests that using homology search to predict miRNA targets in plants is an effective method. Target identification was carried out by comparing identified miRNA candidates to the same transcript used for miRNA identification. To predict miRNA-target genes, the following parameters were used: 1) No more than four mismatches were allowed between projected mRNAs and target genes, 2) No mismatches were allowed for the 10^{th} and 11^{th} positions of the complementary site since they are regarded as cleavage sites, and 3) Maximum of four GU pairs were allowed in the complementary alignment. All the steps involved in identification of miRNA and their targets in cluster bean are represented in **Figure 1**.

2.4 Phylogenetic Analysis of Identified miRNAs

The precursor sequences of the identified and well-known cluster bean miRNAs were aligned and phylogenetically analyzed to examine their evolutionary relationships (www.clustal.org/). The evolutionary distances were calculated using the neighbor-joining (NJ) method after 1,000 bootstrapped iterations.

2.5 Analysis of Gene Ontology for Predicted Targets

GO analysis of the identified target transcripts and pathway analysis was performed using the target annotation module in C-mii, which calculates and visualizes the distribution of selected targets, to better understand the role of cluster bean miRNAs and their regulating targets.

2.6 The miRNA-Target Gene Regulatory Network

To study co-regulated targets by miRNA families and to select miRNA-targets based on MFE value, a biological network was created between identified miRNAs and their targets.

GENEMANIA (http://genemania.org/) was used to visualize the biological network of miRNA and its target.

2.7 Identification of Transcription Factors From Expressed Sequence Tags of Cluster Bean

Homology based search against the plant transcription factor database (PlnTFDB) was performed to identify genes encoding transcription factors in the EST of cluster bean (Pérez-Rodríguez, 2010).

3 RESULTS

3.1 Identification of miRNA Using Expressed Sequence Tags Sequences

To identify and characterize the conserved miRNA in cluster bean, a thorough EST-based approach using C-mii (version1.11, Numnark et al., 2012) software was employed as genomic sequence information is limited. The software includes four steps such as sequence loading, homolog search, primary miRNA folding, and precursor miRNA folding. A total of 16,503 EST sequences of cluster bean were downloaded from the public database available at (http://www.ncbi.nlm.nih.gov/nucest) and compared to the non-redundant viridiplantae mature plant miRNA sequences (4,256) retrieved from miRBase (**Supplementary Figure S1**). Sequences longer than 3000 nt or containing characters other than A, T, C, G, U, and N were excluded. The homolog search module scans different plant mature miRNAs from miRBase with 16,503 EST sequences with default parameter setting. Using the plus strand, only 361 miRNA families and 670 members were identified. These sequences were selected as input for the primary miRNA folding module. Next, the primary miRNA folding was done with default parameter settings except for the BLAST e-value ≤ 1E-5 which resulted in successfully folding 3,068 sequences. These sequences were again selected for the next step i.e., the precursor miRNA folding. A different number of sequences remained at each step (**Supplementary Table S1**). The precursor miRNA folding at default parameter settings except for the BLASTX e-value ≤ 1E-5 resulted in 74 miRNA candidates, belonging to 57 miRNA families predicted from cluster bean EST sequences available at NCBI. For reducing the duplicity and improving the accuracy of predicted miRNAs, the false-positive results only plus strand and high negative values of MFEI (b = −0.6) were considered, respectively. The predicted 57 miRNA families showing b = −0.6 MFEI value are listed in **Supplementary Table S1**. All potential miRNAs of cluster bean were located at the 5′ end of the precursor miRNA. Several plant species such as *Arabidopsis thaliana*, *Oryza sativa*, *Glycine max*, *Medicago truncatula*, *Zea mays*, *Vitis vinifera*, *Saccharum officinarum*, *Picea abies*, *Arabidopsis lyrata*, *Selaginella moellendorffii*, *Triticum aestivum*, *Chlamydomonas reinhardtii*, *Pinus taeda* were found to have miRNA family homologs in them. Out of these, *Oryza sativa* (28%), *Arabidopsis thaliana* (23%), and *Glycine max* (12%) had major miRNA family homologs followed by *Medicago truncatula* (9%; **Figure 2**).

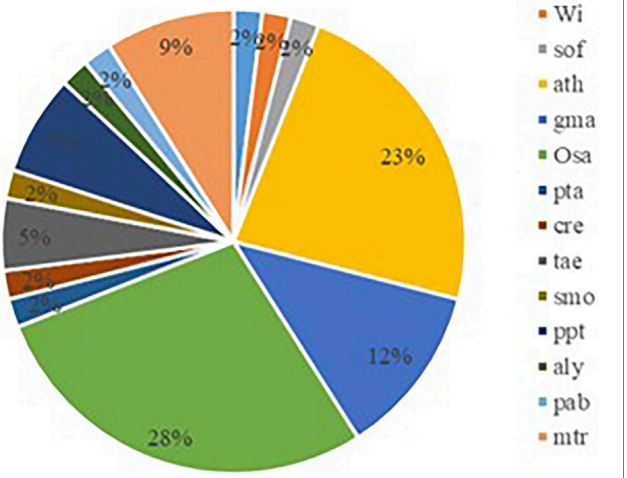


FIGURE 2 | Identified miRNA families of cluster bean which have homologues in other plant species (ath- *Arabidopsis thaliana*, osa-*Oryza sativa*, gma-*Glycine max*, mtr-*Medicago truncatula*, zma-*Zea mays*, vvi-*Vitis vinifera*, sof-*Saccharum officinarum*, pab-*Picea abies*, aly-*Arabidopsis lyrata*, smo-*Selaginella moellendorffii*, tae-*Triticum aesativum*, cre-*Chlamydomonas reinhardtii*, pta- *Pinus taeda*, ppt-*Physcomitrella patens*).

miRNA families which were predicted from EST database were matched against cDNA library of seed development stage and out of 57 EST IDs in which different miRNA families were present, only 19 EST IDs were found to be involved in various functions (**Supplementary Table S2**).

3.2 Characteristics of Newly Identified Cluster Bean miRNA

To validate newly identified miRNAs, various calculated parameters were analyzed for miRNA family member, precursor length of miRNA, A + U content, G + C content, their ratios, the minimal folding free energy (MFE), and minimal folding free energy index (MFEI) for each miRNA precursor.

3.2.1 miRNA Family Members

Most of the miRNA families had one member but some of the miRNAs families were observed to have more than one number. These families were miR156 (9), miR 2606 (7), and miR172 (**Supplementary Figure S2**).

3.2.2 Precursor miRNA Length

The length of precursor miRNAs varied from 35 to 539 nt in length with an average value of 100 nt. miR1132 was found to have the lowest miRNA pre-miRNA length (35 nt) and miR1439 had highest miRNA length (539 nt; **Supplementary Figure S3**).

3.2.3 Mature miRNA Length

The length of most of the mature miRNAs was found to be 21 nt. Based on miRNA length, the predicted miRNA was divided into

TABLE 1 | Summarized parameters of predicted miRNA families in cluster bean.

Predicted miRNA family	AU	GC	Nucleotide content	Pre-miRNA sequence length	MFE	MFEI
miR1044	70.13	29.87	A(39)U(69)G(26)C(20)N(0)	154	−30.3	−0.65
miR1109	69.93	30.07	A(39)U(54)G(23)C(17)N(0)	133	−27.8	−0.69
miR1132	62.86	37.14	A(10)U(12)G(8)C(5)N(0)	35	−12.8	−0.98
miR1134	66.08	33.92	A(20)U(17)G(9)C(10)N(0)	56	−12.5	−0.65
miR1167	58.7	41.3	A(14)U(13)G(11)C(8)N(0)	46	−16.2	−0.85
miR1313	64.9	35.1	A(26)U(35)G(17)C(16)N(0)	94	−22.9	−0.69
miR1439	80.07	19.93	A(240)U(233)G(50)C(68)N(1)	592	−75.3	−0.63
miR1527	63.05	36.95	A(18)U(11)G(7)C(10)N(0)	46	−11	−0.64
miR1533	83.12	16.88	A(35)U(29)G(8)C(5)N(0)	77	−9.9	−0.761
miR1535	42.07	57.93	A(22)U(39)G(41)C(43)N(0)	145	−51.6	−0.61
miR156	56.96	43.04	A(48)U(38)G(31)C(34)N(0)	151	−38.4	−0.59
miR168	47.92	52.08	A(12)U(11)G(13)C(12)N(0)	48	−16.6	−0.66
miR169	56.9	43.1	A(17)U(16)G(13)C(12)N(0)	58	−15.9	−0.63
miR172	62.5	37.5	A(19)U(16)G(9)C(12)N(0)	56	−14.7	−0.7
miR1852	56.1	43.9	A(19)U(26)G(22)C(14)N(1)	82	−21.9	−0.6
miR1857	57.78	42.22	A(9)U(17)G(13)C(6)N(0)	45	−11.8	−0.62
miR2082	52.64	47.36	A(12)U(18)G(19)C(8)N(0)	57	−18.66	−0.66
miR2098	65.52	34.48	A(19)U(38)G(20)C(10)N(0)	87	−21.1	−0.7
miR2105	63.81	36.19	A(30)U(37)G(22)C(16)N(0)	105	−27.7	−0.72
miR2275	47.7	52.3	A(16)U(15)G(16)C(18)N(0)	65	−21.3	−0.62
miR2606	57.5	42.5	A(10)U(13)G(9)C(8)N(0)	40	−10.2	−0.6
miR2628	41.09	58.91	A(21)U(22)G(21)C(9)N(0)	73	−18.6	−0.62
miR2634	74.29	25.71	A(10)U(16)G(5)C(4)N(0)	35	−12.3	−1.36
miR2866	64.95	35.05	A(19)U(43(G(17)C(17)N(1)	97	−21.1	−0.62
miR2919	63.24	36.76	A(21)U(22)G(16)C(9)N(0)	68	−22.2	−0.888
miR393	53.85	46.15	A(11)U(17)G(13)C(11)N(0)	52	−16.2	−0.67
miR396	61.23	38.77	A(10)U(20)G(10)C(9)N(0)	49	−16.5	−0.86
miR397	54.17	45.83	A(8)U(18)G(10)C(12)N(0)	48	−14	−0.63
miR3979	60	40	A(11)U(16)G(8)C(10)N(0)	45	−11.2	−0.62
miR399	57.15	42.85	A(14)U(10)G(10)C(8)N(0)	42	−13.3	−0.73
miR414	49.11	50.89	A(34)U(21)G(29)C(28)N(0)	112	−33.4	−0.6
miR437	58.74	41.26	A(18)U(19)G(16)C(10)N(0)	63	−21.4	−0.82
miR4413	68.43	31.57	A(21)U(31)G(17)C(7)N(0)	76	−14.6	−0.6
miR444	63.83	36.17	A(8)U(22)G(10)C(7)N(0)	47	−10.2	−0.59
miR4995	42.86	57.14	A(13)U(14)G(18)C(18)N(0)	63	−23.7	−0.65
miR5015	54.24	45.76	A(9)U(23)G(15)C(12)N(0)	59	−16.9	−0.625
miR5021	50	50	A(12)U(11)G(14)C(9)N(0)	46	−16.2	−0.7
miR5079	76.93	23.07	A(25)U(35)G(8)C(10)N(0)	78	−11	−0.61
miR5265	67.22	32.78	A(24)U(17)G(11)C(9)N(0)	61	−13.8	−0.69
miR5267	68.75	31.25	A(42)U(46)G(21)C(19)N(0)	128	−37.7	−0.94
miR5338	61.62	38.38	A(90)U(85)G(64)C(45)N(0)	284	−69	−0.63
miR537	71.43	28.57	A(12)U(18)G(10)C(2)N(0)	42	−8.3	−0.69
miR5489	52.31	47.69	A(18)U(16)G(18)C(13)N(0)	65	−20.7	−0.66
miR5542	67.2	32.8	A(36)U(48)G(30)C(11)N(0)	125	−25.2	−0.6
miR5565	39.55	60.45	A(35)U(46)G(33)C(20)N(0)	134	−39.6	−0.74
miR5641	70.28	29.72	A(22)U(30)G(14)C(8)N(0)	74	−21.3	−0.96
miR5658	81.82	18.18	A(14)U(22)G(7)C(1)N(0)	44	−5.8	−0.72
miR5662	51.84	48.16	A(108)U(75)G(91)C(80)N(1)	355	−103.2	−0.6
miR773	53.66	46.34	A(6)U(16)G(12)C(7)N(0)	41	−16.2	−0.85
miR779	57.86	42.14	A(29)U(41)G(33)C(18)N(0)	121	−31.1	−0.609
miR781	68.19	31.81	A(9)U(21)G(9)C(5)N(0)	44	−13.1	−0.93
miR837	54.89	45.11	A(29)U(44)G(30)C(30)N(0)	133	−38.4	−0.63
miR838	49.39	50.61	A(138)U(144)G(150)C(139)N(0)	571	−177.3	−0.613
miR863	63.64	36.36	A(13)U(15)G(9)C(7)N(0)	44	−10.8	−0.67
miR865	76.93	23.07	A(22)U(18)G(4)C(8)N(0)	52	−8.7	−0.72
miR867	81.64	18.36	A(45)U(35)G(9)C(9)N(0)	98	−17.6	−0.9
miR902	68.86	31.14	A(20)U(22)G(11)C(8)N(0)	61	−12.7	−0.66

three groups. These groups had 19, 20, or 21 nucleotide length in mature miRNA except in two groups. First group was 19 nt long, while second group was 20 nt long, and third group was 21 nt long (**Supplementary Table S3**).

3.2.4 Minimal Folding Free Energy

The minimal folding free energy (MFE) has been considered one of the significant features for identification of miRNA. It is an important parameter for determining the secondary structure of

pre-miRNA. Highly negative MFE indicates thermodynamically stable secondary structure of the corresponding sequences. In the present study, the range of MFE (ΔG kcal/mol) was 5.8–177.3 (−kcal/mol) with an average value of 25.4 (−kcal/mol) (**Supplementary Figure S4**).

3.2.5 Minimal Folding Free Energy Index

The index that differentiates pre-miRNA from other coding and non-coding RNAs and RNA fragments is MFEI. The MFEI values observed in the present study ranged from 0.59 to 1.36 −kcal/mol. The mean MFEI values were found to be 0.70 −kcal/mol (**Supplementary Figure S5**). The MFEI values of tRNAs, rRNAs, and mRNA are 0.64, 0.59, and 0.62–0.66 −kcal/mol, respectively which are significantly lower than the cluster bean miRNAs reported in this study, signifying that the identified cluster bean miRNAs are true miRNAs.

3.2.6 AU Content

The analysis also showed that AU contents in all identified miRNA precursors ranged from 41.09 to 83.12% with an average of 61.18% (**Supplementary Figure S6**).

3.2.7 GC Content

GC content was also analyzed for different miRNA families predicted in cluster beans. The GC content ranged from 18.36 to 60.45% with an average of 38.81% (**Supplementary Figure S7**).

3.2.8 Nucleotide Content

The nucleotide content was not uniform throughout the predicted miRNA families (**Supplementary Figure S8**). The A/U and C/G ratio was also calculated as given in **Supplementary Table S4**. The highest A/U was observed for miR1527 (1.6) and the lowest was for miRNA 444 (0.3) whereas the highest C/G ratio was for miR1527 (1.4) and the lowest for miR537 (0.2). All parameters of different characteristics of predicted miRNA families are summarized in **Table 1**.

3.3 Sequence Alignment and Phylogenetic Analysis of New miRNA

Primary and mature plant miRNAs are highly conserved among distantly related plant species. This high level of conservation between taxa was used as an ancillary criterion for miRNA annotation. Comparison of the precursor sequences of the predicted miRNAs with each other showed that most members could have a high degree of sequence similarity with others. Three different clusters were formed which I cluster was the largest and was divided into two sub-clusters. Sub-cluster I consists of miR5542, miR399, miR169 miR393, miR437, miR2606, miR444, miR1857, miR1313, miR5565, miR1109, miR1044, miR5641, miR1533, miR5267, miR1527, miR2666, miR5265, and miR867. Sub-cluster II consists of miR2919, miR397, miR4413, miR865, miR2634, miR773, miR5338, miR2275, miR168, miR4995, miR2093, miR5489, miR5079, miR3979, miR902, miR172, miR414, miR2082, miR838, miR2105, miR5015, miR781, miR396, and miR837and

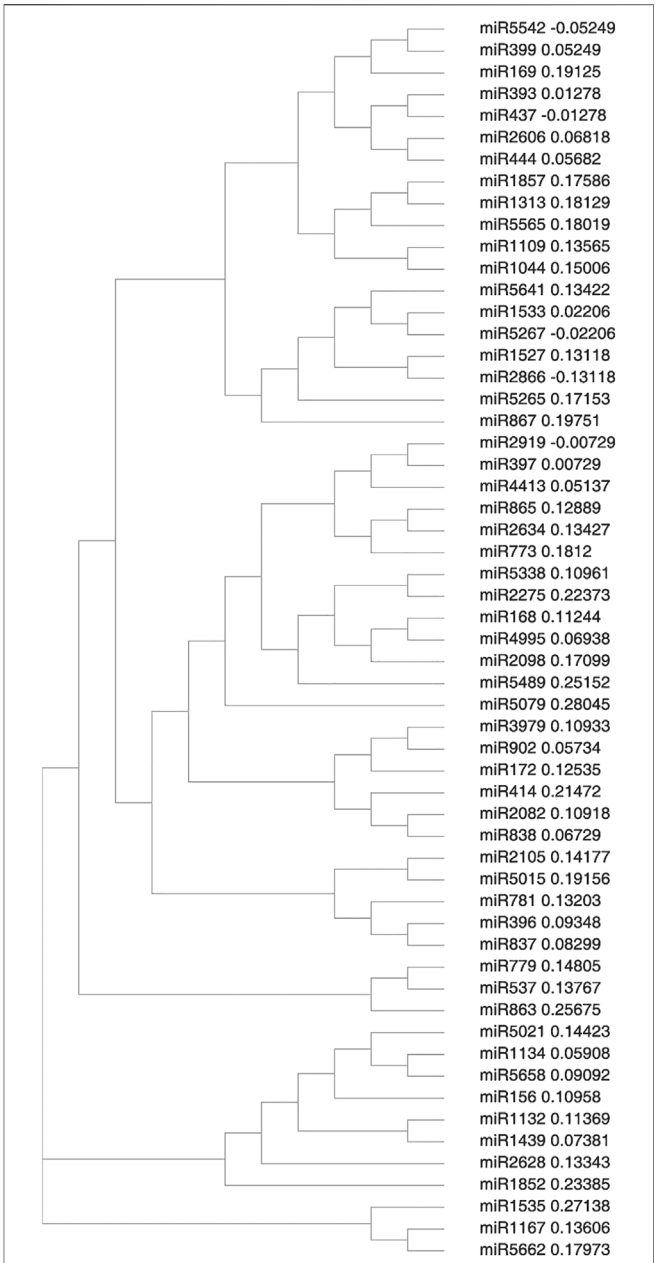


FIGURE 3 | Phylogenetic relationship among predicted miRNA families in cluster bean.

comprised of largest number of miRNA families as in **Figure 3**. Cluster II consists of miR5021, miR1134, miR5658, miR156, miR1132, miR1439, miR2628 and miR1852. Cluster III, which is the smallest of all comprises miR1535, miR1167, and miR5662.

3.4 Target Identification

The target identification procedure using a computational method is simple and efficient, yielding possible targets that aid in the subsequent study. C-mii's target prediction menu consists of four submenus: sequence loading, target scanning,

TABLE 2 | Major potential target genes for predicted miRNAs in cluster bean.

miRNA	Targeted protein	Targeted EST ID	Description
miR156	SPL4, SPL13, SBP2, SBP1, SPL5	EG982824.1 GLL094_E10_036	Squamosa promoter-binding-like protein
	UBC	EG979243.1 GLL062_C07_030	Ubiquitin-conjugating enzyme E2
miR172	2AAG	EG985097.1 GLE041_H10_033	Serine/threonine-protein phosphatase
	ARF	EG990107.1 GLE087_E01_004	ADP-ribosylation factor
miR393	RK20	EG987334.1 GLE062_E02_004	50S ribosomal protein
miR3979	SPZX	EG989678.1 GLE083_C05_022	Serpin-ZX
miR397	G3PA	EG975619.1 GLL029_A09_040	Glyceraldehyde-3-phosphate dehydrogenase A
miR414	AGO1_	EG985306.1 GLE044_A02_008	Protein argonaute
	ALFIN	EG986332.1 GLE052_H02_001	PHD finger protein Alfin1, PHD finger protein ALFIN-LIKE 7
	C3H2	EG982280.1 GLL08_C06_022	Zinc finger CCCH domain-containing
	CD48C	EG983335.1 GLE022_G09_034	Cell division control protein 48 homolog C
	CESA6, CESA5, CESA2, CESA9, CESA9	EG983155.1 GLE021_B02_007	Cellulose synthase A catalytic subunit 5, Cellulose synthase A catalytic subunit 2, Probable cellulose synthase A catalytic subunit 9, Cellulose synthase A catalytic subunit 9
	TGT2	EG986117.1 GLE050_G11_042	Trihelix transcription factor GT-2, Trihelix transcription factor GTL1, Trihelix transcription factor GTL2
miR437	CCMC	EG985481.1 GLE045_F11_043	Putative cytochrome c biosynthesis ccmC-like mitochondrial protein
miR773	P2C	EG988418.1 GLE072_A02_008	Protein phosphatase 2C and cyclic nucleotide-binding/kinase domain-containing protein
miR837	HMDH1	EG988124.1 GLE06_B01_007	3-hydroxy-3-methylglutaryl-coenzyme A reductase 1
	Q8OMT	EG976284.1 GLL034_E07_028	8-hydroxyquercetin 8-O-methyltransferase, Isoflavone-7-O-methyltransferase 9
miR838	NAC	EG989253.1 GLE07_B05_023	NAC domain-containing protein 18, NAC domain-containing protein 68, NAC domain-containing protein 67, NAC domain-containing protein 29, NAC domain-containing protein 48
	HMA4	EG990669.1 GLE091_H04_009	Putative cadmium/zinc-transporting ATPase
miR865	RL	EG988689.1 GLE074_F05_019	60S ribosomal protein
miR902	MNS2	EG989919.1 GLE085_F05_019	Mannosyl-oligosaccharide 1, Mannosyl-oligosaccharide 1
miR1109	AGO4	EG986669.1 GLE056_A10_040	Protein argonaute
miR1132	CML35	EG985964.1 GLE04_C12_046	Probable calcium-binding protein
	RANA1	EG985231.1 GLE043_C03_014	GTP-binding nuclear protein Ran-A1
miR1134	GAST1	EG989127.1 GLE078_H10_033	Gibberellin-regulated protein
miR1439	CML	EG985964.1 GLE04_C12_046	Probable calcium-binding protein
miR1527	CAF	EG981106.1 GLL079_F10_035	Probable CCR4-associated factor 1 homolog 7
miR1533	GALE	EG984760.1 GLE038_G01_002	UDP-glucose 4-epimerase GEPI42, UDP-glucose 4-epimerase, UDP-glucose 4-epimerase 1, UDP-glucose 4-epimerase 2, UDP-glucose 4-epimerase GEPI48
	MT	EG983379.1 GLE011_B12_047	Metallothionein like protein
	JKD_ MGP,NUC	EG990549.1 GLE090_G04_010	Zinc finger protein JACKDAW, Zinc finger protein MAGPIE, Zinc finger protein NUTCRACKER
miR1857	GALE	GO542112.1 Mdfrjg3507l06.g1 Apple_E	UDP-glucose 4-epimerase GEPI48, UDP-glucose 4-epimerase 2, UDP-glucose 4-epimerase 3, UDP-glucose 4-epimerase GEPI42, UDP-glucose 4-epimerase 1
miR2098	RS193	EG981936.1 GLL086_H09_033	40S ribosomal protein
miR2105	REV, HOX9, HOX9, HOX10, HOX10	EG990607.1 GLE091_C08_030	Homeobox-leucine zipper protein
miR2275	HSP81 82 83 90-1	EG986829.1 GLE057_F03_011	Heat shock protein 81-1, Heat shock protein 81-3, Heat shock protein 81-2, Heat shock protein 90-1
miR2628	DXS	EG980801.1 GLL076_E12_044	Probable 1-deoxy-D-xylulose-5-phosphate synthase
miR2606	TSJT1	EG986569.1 GLE055_B02_007	Stem-specific protein TSJT1
miR2634	AG	EG990361.1 GLE089_H09_033	Floral homeotic protein AGAMOUS
	GAOX2	EG987827.1 GLE067_C01_006	Gibberellin 20 oxidase 2,
miR4995	RLA	EG986778.1 GLE057_B04_015	60S acidic ribosomal protein
miR5021	2AAB_	EG986207.1 GLE051_F09_035	Serine/threonine-protein phosphatase 2A
	3MAT 5MAT1,AGCT	EG983947.1 GLE030_E09_036	Malonyl-coenzyme A:anthocyanin 3-O-glucoside-6″-O-malonyltransferase
	AGL11, AGL5, AGL1_,MAD21	EG975055.1 GLL011_C02_006	Agamous-like MADS-box protein, MADS-box transcription factor 21
	ARI7,ARI8,ARI6_	EG987656.1 GLE065_E09_036	Probable E3 ubiquitin-protein ligase
	ASP	EG984370.1 GLE035_A02_008	Aspartic proteinase Asp1
	BGLS	EG983200.1 GLE021_E07_028	Non-cyanogenic beta-glucosidase,Cyanogenic beta-glucosidase
miR5079	EXP13	EG985350.1 GLE044_D07_029	Expansin-A13
miR5565	ADLO1	EG986858.1 GLE057_H05_017	Protein ARABIDILLO 1
miR5641	CAX3	EG984929.1 GLE040_C12_046	Vacuolar cation/proton exchanger
miR5658	AGL11	EG975055.1 GLL011_C02_006	Agamous-like MADS-box protein AGL11
	CSLG2	EG976517.1 GLL036_G04_010	Cellulose synthase-like protein G2, Cellulose synthase-like protein G1, Cellulose synthase-like protein G3
	GASA4	EG987384.1 GLE062_H11_041	Gibberellin-regulated protein 4

(Continued on following page)

TABLE 2 | (*Continued*) Major potential target genes for predicted miRNAs in cluster bean.

miRNA	Targeted protein	Targeted EST ID	Description
	GLYT3	EG975650.1 GLL029_D02_005	Probable glycosyltransferase
	HSFA9, HSF30, HFA2C,HFA6B	EG985970.1 GLE04_D05_021	Heat stress transcription factor A-9, Heat stress transcription factor A-2e, Heat shock factor protein HSF30, Heat stress transcription factor A-2c, Heat stress transcription factor A-6b
miR168	AGO1	EG985306.1 GLE044_A02_008	Protein argonaute 1
	DHSA	EG985277.1 GLE013_C07_030	Succinate dehydrogenase [ubiquinone] flavoprotein subunit
miR169	AL2B7	EG983762.1 GLE029_F12_043	Aldehyde dehydrogenase family 2 member B7
miR399	CYSK	EG985107.1 GLE042_A07_032	Cysteine synthase
miR2082	TBB	EG984572.1 GLE036_H07_025	Tubulin beta-6 chain, Tubulin beta-3 chain, Tubulin beta-7 chain, Tubulin beta-4 chain
miR5662	PIP	EG976742.1 GLL013_A03_016	Probable aquaporin PIP1-2, AquaporinPIP1-2, Aquaporin PIP1-3/PIP1-4, Aquaporin PIP1-3, Aquaporin PIP1-1
miR1535	C3H11,C3H21	EG979135.1 GLL061_C02_006	Zinc finger CCCH domain-containing protein 11, Zinc finger CCCH domain-containing protein 21

TABLE 3 | Common targets of predicted miRNA families in cluster bean

miRNA	Common target	Traits associated	Function
miR172, miR5021	PP2A	Biotic and abiotic stress	Serine/threonine protein phosphatase
miR5021, miR838	FRI	Iron content	Ferritin
miR156, miR1533, miR5021	UBC	Pattern triggered immunity	Ubiquitin-conjugating enzyme
miR837,miR838, miR5021	TRL	Role in plant tolerance of oxidative stress	Thioredoxin
miR5566, miR156	TBG	Plant growth and morphogenesis	Tubulin
miR5021 ,miR5658	PATL	Cell formation	Patellin
miR1533, miR838	P2C	Stress and developmental process	Probable protein phosphatase
miR5021, miR5658, miR396	MADS	Flower development	Floral homeotic protein

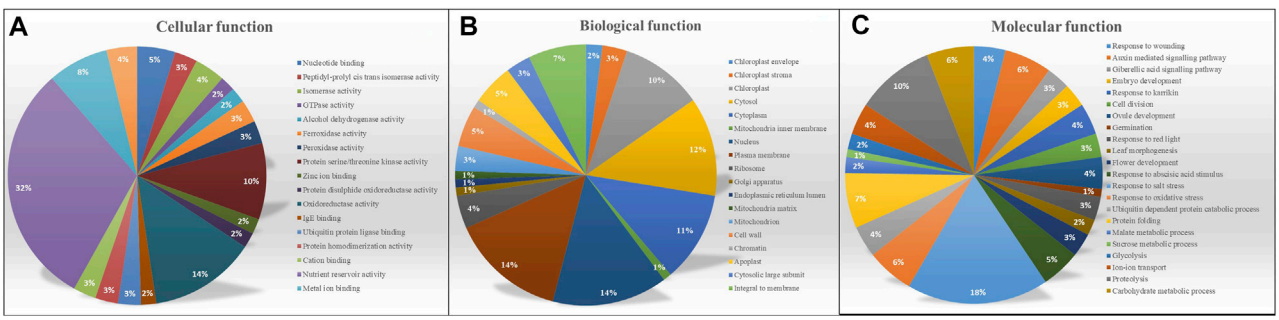


FIGURE 4 | GO annotation for output targets and their distribution in three categories **(A)** Molecular function **(B)** Biological process **(C)** Cellular component.

miRNA-target folding, and target annotation. Different miRNA targets are represented in **Table 2**. The acceptable length of an input sequence for target identification was extended to 20,000 nt. In the present study, 2,554 sequences previously reported as miRNA-specific targets from 47 different plant species were uploaded and selected for target scanning. Using the plus strand, 6,164 sequences were identified as miRNA binding sites for 361 miRNA families and 670 members. All of them were selected as input for miRNA target folding at 37°C. About 3,068 sequences were successfully folded and these sequences were used as input for target annotation using Gene Ontology (GO). The target annotation module predicts the function for potential targets. Target identification results are provided in **Table 2** with their targeted EST ID. A total of 57 predicted miRNAs from 37 miRNA families were found to regulate 623 target transcripts. miRNAs can potentially regulate many distinct genes, implying that these genes belong to several gene families engaged in various biological, cellular, and molecular processes (Dehury et al., 2013). However, several miRNAs with unclear roles were discovered (miR1104, miR1167, miR1313, miR1852, miR2866, miR2919, miR3522, miR4413, miR5267, miR5338, miR537, miR5489, miR779, miR781, miR867, miR863, miR444, miR5265, and miR5542).

miRNA414, miR5658, and miR5021 have the highest number of targets. Some miRNAs such as miR156, miR172, miR414, miR1533, and miR5021 had more than one target while some miRNA families have common targets as represented in **Table 3**.

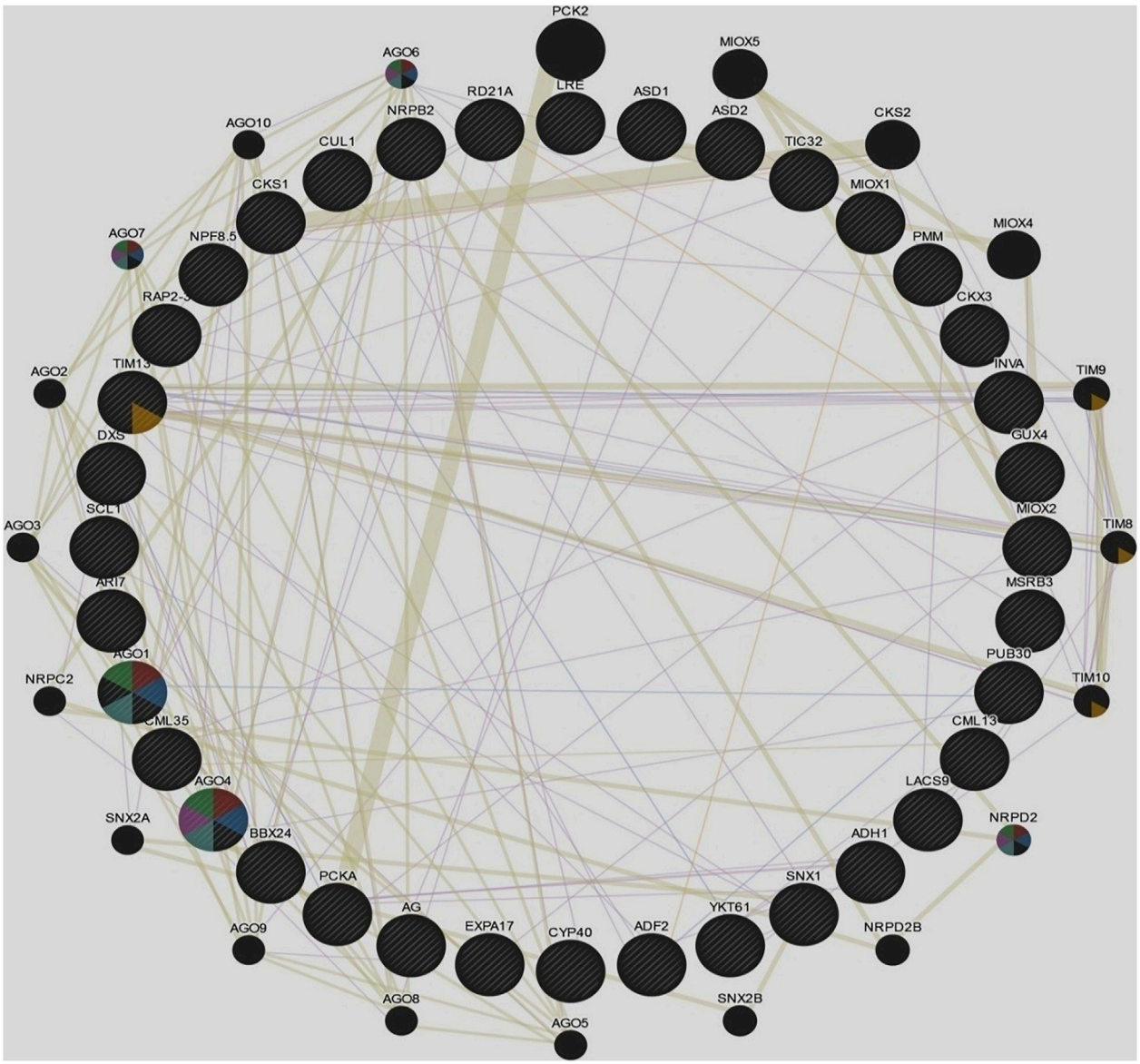


FIGURE 5 | Gene regulatory networks of cluster bean target genes in GENEMANIA against *Arabidopsis thaliana*.

To gain a better understanding of the functional characteristics of these miRNA targets, the GO annotations of all projected targets were examined. All targets were divided into three GO categories based on the ontological meanings of the GO words, as shown in **Figures 4A–C** molecular function, biological process, and cellular component. Predicted targets in the molecular function category were related to nutrition reservoir activity (32%), followed by nucleotide binding and protein serine/threonine kinase activity. The bulk of the targets in the biological process was potentially implicated in cellular activities and metabolic processes involving salt stress (18%) and proteolysis (10%). Other genes were found to be involved in a variety of signaling pathways, including the gibberellic acid signaling pathway, auxin-mediated signaling, and abscisic acid stimulation. Target genes in the cellular component category were assigned to the subcategories cytoplasm (24%) chloroplast (15%) nucleus (14%) mitochondria and membrane. Major GO IDs were selected based on their function and classified into biotic and abiotic, and carbohydrate metabolism-related categories as represented in **Supplementary Table S5**.

The Gene Ontology provides the logical structure of the biological functions and their relationships to one another, manifested as a directed acyclic graph.

1) GO: 0009615 is involved in any process that results from a stimulus of the virus.This GO ID was found to be

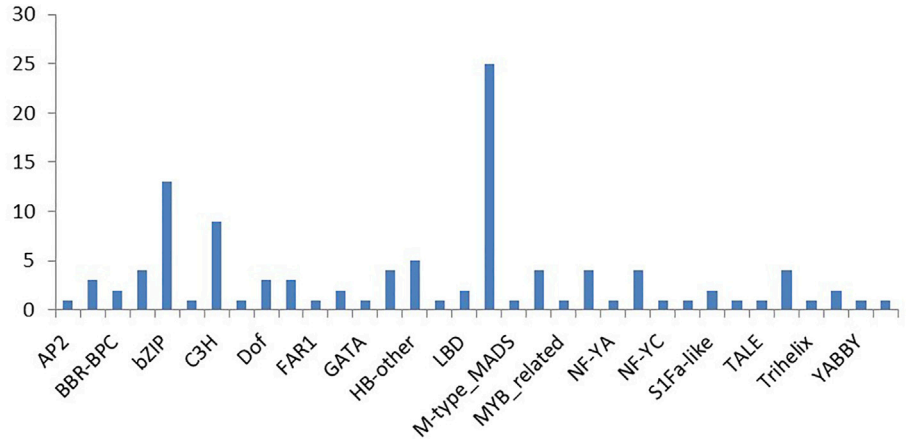


FIGURE 6 | Major transcription factor classes identified from cluster bean ESTs.

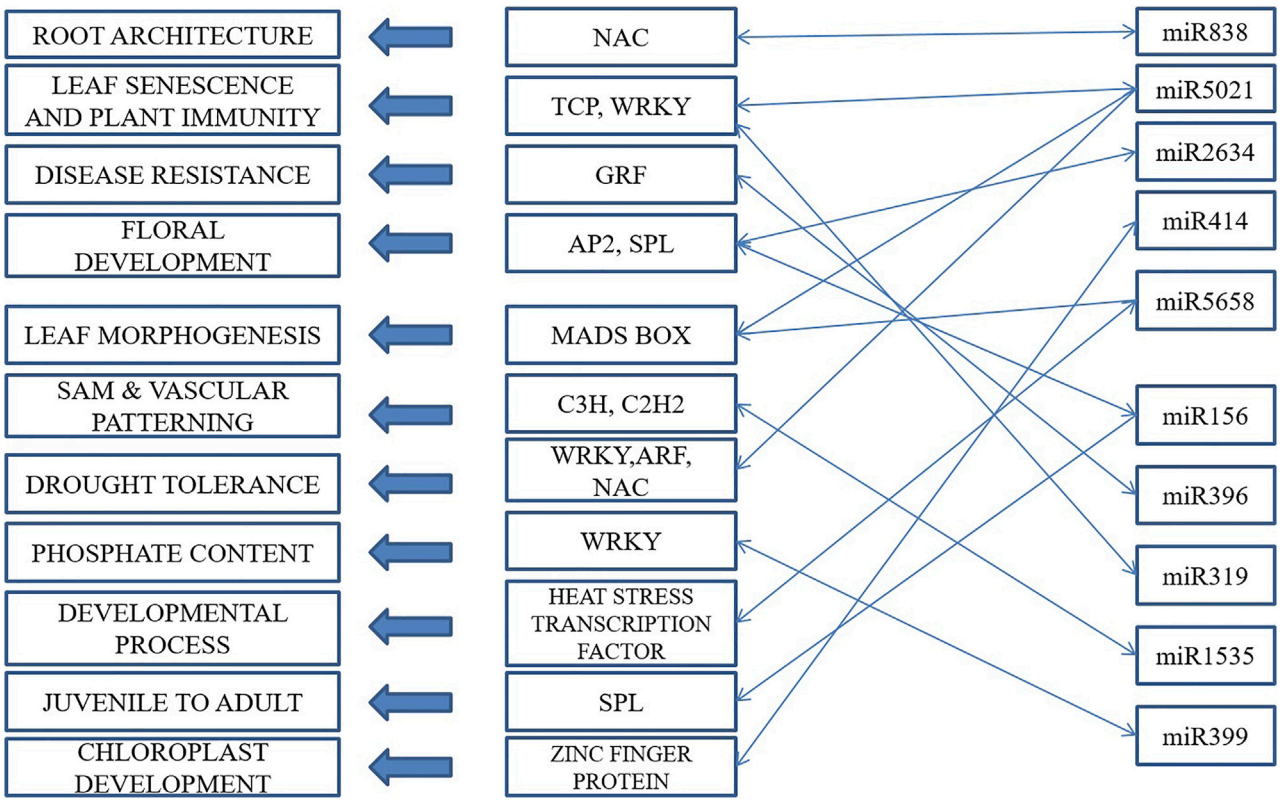


FIGURE 7 | A proposed model of interactions among transcription factors and miRNAs for gene regulation in cluster bean.

associated with other GO IDs such as GO:0008150, GO: 0050896, GO:0009605, GO:0009606, GO:0043207, GO: 0051704, GO:51707.

2) GO: 0009414 relates to a change in the activity of cells as a result of water deprivation. This GO ID was found to be interconnected with GO:0008150, GO:0050896, GO:0042221, GO: 0001101, GO:0010035, GO:0006950, GO:0009415, GO: 1901700, and GO:00009628.
3) GO:0005985 relates to all the chemical reactions and pathways involving sucrose, the disaccharide fructofuranosyl-glucopyranoside. This GO ID was associated with GO: 0008150, GO:0005984, GO:0044262, GO:0009311, GO:

0005975, GO:0044238, GO:0044237, GO:0009987, GO: 0008152, GO:0071704, and GO:008152.

3.5 Gene Regulatory Networks of Cluster Bean Target Genes in GENEMANIA

Target genes of predicted miRNA families of cluster bean were used to make a regulatory network in GENEMANIA and to predict the function of these genes (**Figure 5**). Following genes were found and a regulatory network was predicted against *Arabidopsis thaliana*. The genes were CML 13, MIOX1, SCL1, RD21A, AGO1, GUX4, INVA, TIM13, ADH1, LACS9, CYP40, MIOX2, AGO4, CKS1, AR17, CML35, PMM, ASD1, ADF2, PUB30, CUL1, MSRB3, DXS, AG, TIC32, LRE, PCKA, SNX1, ASD2, CKX3, YKT61, STO, EBP, EXP17 and RPB2. The network represents the involvement of various additional genes and their function (**Supplementary Table S6**) using STRING.

Various genes comprised of different networks such as co-expression (55.01%), shared protein domain (35.32%), predicted (6.35%), and co-localization (3.32%). The given network indicates the involvement of various genes functioning in mainly RNA interference, post-transcriptional gene silencing by RNA, the establishment of protein localization to mitochondrion, and production of siRNA involved in RNA interference.

3.6 Identification of Transcription Factors and Depicting Regulatory Network Model Between miRNA and Transcription Factor

Many of the predicted targets were annotated to be transcription factors. To find some correlation between miRNA and transcription factors, the plant transcription factor database was used to identify transcription factors from the input sequences. EST sequences were used as input sequences and ESTScan 3.0 was employed to identify CDS regions of input nucleic acid sequences which translated them to protein sequences. By checking "Best hit in *Arabidopsis thaliana*," the transcription factors such as B3 family protein, basic pentacysteine 6 and 4, bHLH family protein, basic leucine-zipper 70, TGACG motif-binding factor 6, C3H family protein, MIKC_MADS family protein, nuclear factor Y, and TCP family protein etc. were identified in EST database (**Supplementary Table S7**).

This led to the identification of many transcription factors, the bulk of which belonged to the MIKC MADS family protein basic leucine-zipper and the C3H family protein (**Figure 6**). A model is well represented in **Figure 7**, showing the interaction between miRNA and transcription factor genes, as well as their impact on many developmental processes and metabolic activities in plants. miRNAs and TFs predicted from the EST database enable combinatorial gene regulation with varied roles that can be used to enhance crops. The combination of miRNA and their targets, which are mostly transcription factors, resulted in the construction of a model suggesting a clear-cut link that leads to the regulation of many developmental and metabolic processes. MiR156, for example, interacts with AP2 and SPL to govern floral growth as well as the juvenile-to-adult transition. Similarly, miR5021 and miR5658 target MADS BOX, which is known to play a crucial role in leaf and flower growth. Furthermore, miR5658 targets heat stress transcription factors and is involved in drought stress and developmental processes such as growth and reproduction. miR838 targets NAC, which is essential for root growth.

4 DISCUSSION

In recent times numerous studies carried out by various researchers have demonstrated that plant miRNAs are involved in various biological and developmental processes and physiological responses. Gaining insights into miRNAs and their targets help us to understand the array of biological events underlying miRNA-mediated regulation and will further elucidate the functional importance of miRNAs (Pani and Mahapatra, 2013; Chaudhary, 2019). A substantial number of miRNAs are conserved among the plant species and it has been observed plant miRNAs and their target sites bind to each other with perfect or near to perfect complementarity (Wang et al., 2004; Schwab et al., 2005; Patanun et al., 2013). This provides an opportunity for identification of miRNAs and their targets in different crop plants by comparing/aligning with the known miRNA sequences. This approach has been widely used to identify miRNA and their targets in various crop plants (Sunkar and Jagadeeswaran, 2008; Dehury et al., 2013; Gasparis et al., 2017; Vivek, 2018). Here, we report the identification of miRNAs and their potential targets in cluster bean by mining EST database.

Cluster bean is an emerging economically important legume with wide range of industrial applications (Rajaprakasam et al., 2021). Despite, high industrial importance of the guar gum this crop is considered as orphan as little progress has been made and availability of genetic and genomic resources is very limited in this crop (Ravelombola et al., 2021). Recently, advanced molecular and genomic approaches have been utilized in this orphan crop to combat the yield losses and elucidate the function of galactomannan related genes (Al-Qurainy et al., 2019; Kumar et al., 2020; Rajaprakasam et al., 2021; Grigoreva et al., 2021a; Grigoreva et al., 2021b; Acharya et al., 2022). Attempts have also been made to sequence cluster bean genome (Grigoreva et al., 2019); however, this first sequencing attempt has a gap of 50% (Grigoreva et al., 2019).

Recent studies are now gradually overcoming the limitations, however, still limited information is available on miRNA-mediated post-transcriptional regulation of metabolic pathways in cluster bean. In few reports, deep sequencing has been applied to identify miRNAs in cluster bean (Sahu et al., 2018; Tyagi et al., 2018). Deep sequencing is the most commonly employed technique for miRNA identification in plant species whose whole genome sequence is not known (Peláez et al., 2012; Tyagi et al., 2018; Tiwari et al., 2021). However, this technique generates huge amount of data and needs expertise for data analysis. Moreover, next-generation sequencing (NGS) is costly and not possible for routine applications (Gasparis et al., 2017).

TABLE 4 | List of predicted miRNA families in cluster bean.

Input sequence Id	Precurson start/ Stop	Known mature miRNA sequence	Predicted miRNA sequence	Predicted miRNA family	Homolog miRNA
EG990325.1 GLE089_F01_003	554–644	5′:UGACAGAAGAAAGAGAGC AC:3′	5′:UGAGAGAAGAAAGAGAGG AA:3′	miR156	ath-miR156h
EG980529.1 GLL073_G12_042	382–429	5′:UCGCUUGGGCAGAUCGGG AC:3′	5′:UCGCUUGGUCAGAUCUGC GC:3′	miR168	sof-miR168b
EG984942.1 GLE040_D12_045	201–258	5′:GAGCCAAGGAUGACUUGC CGU:3′	5′:GAGCCAAGGGUGACUUAA UGU:3′	miR169	vvi-miR169l
EG985097.1 GLE041_H10_033	324–379	5′:AGAAUCUUGAUGAUGCUG CAU:3′	5′:AGAAUCUUGAUCAUCCUG UGU:3′	miR172	gma-miR172b
EG987334.1 GLE062_E02_004	369–420	5′:UCCAAAGGGAUCGCAUUG AUC:3′	5′:UCCAAAGGGAUCGCAUGU CAC:3′	miR393	gma-miR393
EG978583.1 GLL055_H01_001	205–253	5′:UUCCACAGCUUUCUUGAA CUU:3′	5′:UUUCACAGCUUUCUUGAU UGU:3′	miR396	osa-miR396c
EG975619.1 GLL029_A09_040	413–460	5′:UCAUUGAGUGCAGCGUUG ACG:3′	5′:UCUCUGAAUGCAGCGUUG ACU:3′	miR397	pab-miR397
EG990966.1 GLE094_G01_002	207–248	5′:UGCCAAAGGAGAAUUGCC C:3′	5′:AUCCAAAGGUGAAUUGCU C:3′	miR399	tae-miR399
EG985485.1 GLE045_G03_010	317–428	5′:UCAUCCUCAUCAUCAUCG UCC:3′	5′:UUAUCAUCAUCCUCAUCA UCC:3′	miR414	osa-miR414
EG978745.1 GLL057_D04_013	137–199	5′:AAAGUUAGAGAAGUUUGA CUU:3′	5′:AUAGCUAGAGAAGUUUGA AAU:3′	miR437	osa-miR437
EG976745.1 GLL038_H10_033	311–357	5′:UGCAGUUGCUGCCUCAAG CUU:3′	5′:UUCAUUUGAAGCCUCAAG CUU:3′	miR444	osa-miR444a.2
EG978435.1 GLL054_D09_037	223–300	5′:UUGAGGUGUUUCUACAGG CUA:3′	5′:UUGAGGAGUUUCUACAUU CAA:3′	miR537	ppt-miR537a
EG989794.1 GLE084_D07_029	632–672	5′:UUUGCUUCCAGCUUUUGU CUC:3′	5′:UUUGCUUCCAGCUUUGGU UUG:3′	miR773	ath-miR773
EG985501.1 GLE045_H05_017	388–541	5′:UUGUAGUGCAUAUUUGUU UU:3′	5′:UUGUUAUGCAUGUUGGUU UU:3′	miR1044	ppt-miR1044
EG989165.1 GLE017_B06_023	380–512	5′:UAGUGGGAGAUUUUGUGU AAC:3′	5′:UAUUGGGAAAUUUUGUGU CGC:3′	miR1109	smo-miR1109
EG974827.1 GLL020_C10_038	335–369	5′:CAUUAUGGAACGGAAGGA G:3′	5′:CAUUAUGGAAAGGAAGGG A:3′	miR1132	tae-miR1132
EG986974.1 GLE05_A06_024	36–91	5′:CAACAACAACAAGAAGAAGAA GAU:3′	5′:GAACAAGAAGAAGAAGAAGAA GAA:3′	miR1134	tae-miR1134
EG975778.1 GLL02_E10_036	44–89	5′:GGGGUGUGAUGAUUUGAA AC:3′	5′:UGGGUGUGAGGAUUUGAG AC:3′	miR1167	cre-miR1167
EG989769.1 GLE084_B06_023	158–251	5′:UACCACUGAAAUUAUUGU UCG:3′	5′:UUCAAUUGAAAUUAUUGG UCG:3′	miR1313	pta-miR1313
EG989141.1 GLE079_A10_040	161–752	5′:UUUUGGAACGGAGUGAGU AUU:3′	5′:UUAUGAAACGGAGUGAGU AAU:3′	miR1439	osa-miR1439
EG981106.1 GLL079_F10_035	51–96	5′:UAACUCAACCUUACAAAA CC:3′	5′:GAACUCAAGCUAACAAAA CC:3′	miR1527	gma-miR1527
EG975750.1 GLL02_C09_038	453–529	5′:AUAAUAAAAAUAAUAAUGA:3′	5′:UUAAUAAAAAUUAUAAAAA:3′	miR1533	gma-miR1533
EG981151.1 GLL010_F06_019	47–191	5′:CUUGUUUGUGGUGAUGUC U:3′	5′:CUGGUCUGUGGUGCUGUC C:3′	miR1535	gma-miR1535
EG983098.1 GLL01_H03_009	96–177	5′:AUAUGGAUUCAGAAUGCA GGU:3′	5′:AUAAGGAAACAGAAUGCA GGA:3′	miR1852	osa-miR1852
EG975952.1 GLL031_C04_014	116–160	5′:UGGUUUUUUUGGAGCAUG AGG:3′	5′:GGAAUUUUUUGGAGCAGG AGG:3′	miR1857	osa-miR1857
EG990914.1 GLE094_C01_006	644–700	5′:UGUGUGUUCCGCUUCUUC UUU:3′	5′:AGCGUCUUCUGCUUCUUC UUU:3′	miR2082	ppt-miR2082
EG983237.1 GLE021_H04_009	192–278	5′:CGGUUUGUCAAGCGGAGU GC:3′	5′:UGCUUUGUCAAGGGGAGU GU:3′	miR2098	osa-miR2098
EG978398.1 GLL054_A10_040	393–497	5′:UUGUGAUGUGAAUGAUUC AU:3′	5′:GGGUGAUGUGAUUGAUUC CU:3′	miR2105	osa-miR2105
EG987207.1 GLE061_C06_022	253–317	5′:AGGAUUAGAGGGACUUGA ACC:3′	5′:ACGCUUCGAGGGACUUGA AUC:3′	miR2275	zma-miR2275c
EG986569.1 GLE055_B02_007	573–612	5′:UACAAUUCCUUAGGUGCU UUU:3′	5′:UUCAGCUCCUUAGGUGCA UUU:3′	miR2606	mtr-miR2606a
EG980814.1 GLL076_F12_043	210–282	5′:CAUGAAAGAAUGAUGAGU AA:3′	5′:CAGGUGAGAAUGAUGAGG AA:3′	miR2628	mtr-miR2628

(Continued on following page)

TABLE 4 | (*Continued*) List of predicted miRNA families in cluster bean.

Input sequence Id	Precurson start/ Stop	Known mature miRNA sequence	Predicted miRNA sequence	Predicted miRNA family	Homolog miRNA
EG983630.1 GLE027_H07_025	570–604	5′:UUUAUUCUCAGUUUGUUG CUC:3′	5′:UUUAUUCUCAGUUUUUUG AUC:3′	miR2634	mtr-miR2634
EG982808.1 GLL094_D08_029	586–682	5′:UCUAGUUUGUGUUCAGCA UC:3′	5′:UCUAAUUUGUGUUCACUU UC:3′	miR2866	osa-miR2866
EG988391.1 GLE071_F12_043	404–471	5′:AAGGGGGGGGGGGGGAAAG A:3′	5′:AGGGGGAGGGGGGGAAAU U:3′	miR2919	osa-miR2919
EG988087.1 GLE016_B04_015	180–224	5′:UCUCUCUCUCCCUUGAAG GC:3′	5′:UCACUCUCUCCCUUGCAG UU:3′	miR3979	osa-miR3979
EG976112.1 GLL032_G08_026	113–188	5′:UAAGAGAAUUGUAAGUCA CU:3′	5′:UCACUGAAUUGUAAGUUA CU:3′	miR4413	gma-miR4413b
EG986778.1 GLE057_B04_015	648–710	5′:AGGCAGUGGCUUGGUUAA GGG:3′	5′:AAGCAGUGGCUUGGUCAA GGC:3′	miR4995	gma-miR4995
EG986237.1 GLE051_H12_041	18–76	5′:UCUGUUGUUGUUGGUGUU AUG:3′	5′:UCUGUUGUUGUUGUUGUU GUU:3′	miR5015	ath-miR5015b
EG985015.1 GLE041_B08_031	5–586	5′:UGAGAAGAAGAAGAAGAA AA:3′	5′:GGAGAAGAAGAGAAAGGA AA:3′	miR5021	ath-miR5021
EG990900.1 GLE094_B02_007	644–721	5′:UUUGGAUCUGUUAUUUUG GUAU:3′	5′:UUUCUUUCUGUUAUUUUG GAAU:3′	miR5079	osa-miR5079
EG982261.1 GLL08_B02_007	154–214	5′:AAGUGAUGUUGGAAUGGU UA:3′	5′:UAGUGAAGUUGGAAUAAU UA:3′	miR5265	mtr-miR5265
EG981105.1 GLL079_F09_035	141–268	5′:AGGCAUUUGCUAGAAUAC ACCCAC:3′	5′:AAGGAUUUGCUAAUAUAC ACCCAC:3′	miR5267	mtr-miR5267a
EG987890.1 GLE067_G12_042	22–305	5′:UGAAGCUUCAGUUGGUUG UAU:3′	5′:AGAAGCUUCAGUUGGUUU UGA:3′	miR5338	osa-miR5338
EG989431.1 GLE017_D06_021	177–241	5′:CAGGUGUUCUCGAUGGCU UCC:3′	5′:CUAGUGAUUUCGAUGGCU UCC:3′	miR5489	osa-miR5489
EG977779.1 GLL049_B04_015	307–431	5′:UUUGAGAAGGUAUCAUGA GAU:3′	5′:UAUGAGAAUGUAUUAUGA GAU:3′	miR5542	osa-miR5542
EG986048.1 GLE050_B05_023	40–173	5′:UUGUUUGGAUGUUGUCGG A:3′	5′:UUGUUUGGAUGCUGAUGG U:3′	miR5565	sbi-miR5565e
EG983888.1 GLE02_H11_041	113–186	5′:UGGAAGAAGAUGAUAGAA UUA:3′	5′:UGGAAGAAGAUGUGAAAA UUA:3′	miR5641	ath-miR5641
EG985672.1 GLE047_E04_012	293–618	5′:AUGAUGAUGAUGAUGAUG AAA:3′	5′:AUGAUGAUGAAGAAGAAG AAG:3′	miR5658	ath-miR5658
EG984029.1 GLE031_E11_044	110–464	5′:AGAGGUGACCAUUGGAGA UG:3′	5′:AGGGGUGACCGUUGGAGA CU:3′	miR5662	ath-miR5662

This limitation can be overcome by using *in silico* approaches that have emerged as fast and cheap alternative and uses databases like EST and GSS (Vivek, 2018).

Similar strategy was applied in this study and EST database was used to mine miRNAs using computational approach. A large number of miRNA families were identified in cluster bean (**Table 4**) belonging to drought, cold, biotic stress tolerance, and salt-specific stress conditions. As previously reported in other species, the majority of the cluster bean precursor miRNA sequences varied in length from 35 nt to 592 nt with an average value of 100 (Wang et al., 2012; Singh et al., 2016). The length of the mature miRNAs was found to be 20–22 nt long. Similar length of mature miRNAs was reported in earlier miRNA identification studies in soybean, and switchgrass (Xie et al., 2010; Li et al., 2015). The nucleotide composition showed that at the first position at 5′ end of mature miRNA uracil was pre-dominant, similar to observations made by Zhang et al. (2006). It was observed that the average AU content (61.18%) was higher than the GC content (38.18). The higher AU content depicts a comparatively less stable pre-miRNA secondary structure that is easily recognized by RISC complex and converted into mature miRNA. The minimal folding free energy (MFE) is an important determining factor for the stability of the secondary structure. The lower the values of MFE, the higher the thermodynamic stability of the secondary structure (Prabu and Mandal, 2010). In the present study, the MFE value was found to be 25.4 (−kcal/mol). The MFE reported in this study was significantly lower than MFE value of earlier reported miRNAs in cluster bean, where the most stable miRNA, cte_miR1134 has MFE of −72.6 kcal/ mol (Sahu et al., 2018). Similarly, the MFE of miRNA families reported by Tyagi et al., 2018 in different tissue-specific miRNAs of cluster bean was quite higher. The MFE values of miRNAs reported in other legumes like peanut (−50.01 kcal/mol; Chi et al., 2011), French bean (-35.0 to -51.2 kcal/mol; Peláez et al., 2012), chickpea (−50.1419 kcal/mol; Kohli et al., 2014), and lentil (-44.07 kcal/mol; Hosseini et al., 2021). Nowadays, to differentiate miRNA from other RNAs an MFEI value of higher than 0.85 is a potential criterion (Wang et al., 2012). The cluster bean pre-miRNAs identified in the present

investigation showed MFEI values ranging from 0.59 to 1.39 (−kcal/mol) with an average value of 0.70 (−kcal/mol). The MFEI values reported in this study were similar to that of wheat (0.79–1.85−kcal/mol; Gasparis et al., 2017) and lentil (−1.02 kcal/mol; Hosseini et al., 2021).

Cluster bean miRNAs were found to be involved in plant development, defense mechanisms, and signal transduction. The projected miRNAs' functional significance can be easily appreciated by obtaining more information about their target genes. Serine/threonine-protein kinase and ubiquitin carrier protein were anticipated to be the targets of miR172, miR5021, miR156, and miR1533, respectively. Previous reports indicate that ubiquitin carrier proteins plays a significant role in the regulation of various physiological processes such as the recycling of aberrant proteins, metabolic regulation, cell cycle control, and transcription factor (TF) activation (Dreher and Callis, 2007). Serine/threonine-protein kinase plays an important function in signal transduction pathways that contributes to plant defense under both biotic and abiotic stresses (País et al., 2009). It has been reported that miRNAs target TFs, signal transduction factors and metabolic transporters (You et al., 2020).

Cluster bean network analysis of projected miRNAs and their target genes revealed information about the co-regulation of numerous target genes. TFs and miRNAs are both thought to be important regulators of transcriptional events and ultimately, the plant metabolic process (Jangra et al., 2018). TFs regulate gene transcription in the promoter area of the gene, whereas miRNAs regulate gene post transcription in the 3′-untranslated region of the gene, and their respective targets form interconnected functional networks that are critical for the execution of any metabolic activity. Annotation of miRNA-targeted unigenes revealed enrichment of a total 39 TFs such as bHLH (250), Myb (150), Myb related (126), ERF (130), bZip and C2H2 (107), WRKY (100) and NAC (99) (Tyagi et al., 2018). A total of 35 novel TFs were identified which belonged to MIKC_MADS family protein, basic leucine-zipper, and C3H family protein, NAC and WRKY classes. Plant MADS-box genes have been identified as regulators of floral organ identity to control various developmental processes such as the determination of meristem identity of vegetative, inflorescence and floral meristems, root growth, ovule and female gametophyte development, flowering time, fruit ripening and dehiscence (Masiero et al., 2011; Castelán-Muñoz et al., 2019). Certain miRNA-targeted TFs such as NAC and ZF-TFs, implicate their involvement in cellulose synthase-like proteins which is a key factor controlling galactomannan synthesis. The presence of such TFs among the identified miRNA further proves their role in galactomannan pathway. The present study demonstrates the information of miR 5658 family targeting CSL G1, G2, G3 proteins in cluster bean. In both the primary and secondary walls of Arabidopsis, CSL genes are responsible for the majority of glucomannan production (Goubet et al., 2009; Gigli-Bisceglia et al., 2020). In plants, miRNAs function as a negative regulator at the post-transcriptional gene level, either facilitating the cleavage of target transcripts or suppressing their translation, resulting in significant changes in metabolic pathway activity. These are known to regulate cell signaling, oxidative stress, abiotic and biotic stress response, and the development of different tissues, including leaf, stem, anther, root, and flower. Cleavage of the target transcript appears to be the most important and common method of gene control by miRNAs in plants (Catalanotto et al., 2016). The *in silico* identification of predicted miRNAs and their target genes from the EST database suggests a regulatory role in developmental stages. The EST database used in the present study is derived from early and late seed development-specific cDNA libraries. It is the time when galactomannan is synthesized and deposited in seed as a food reserve. Since TFs control gene expression, understanding their function will involve detailed genome-wide functional analysis. Nonetheless, the study provides preliminary information on miRNAs and TFs interplay in cluster bean. *In silico* data analysis gives information regarding miRNA identification and their interaction with TFs and target prediction which can be further supported by wet lab. experiment. We have identified miRNA- SSRs and observed DNA polymorphism in cluster bean germplasm (96 genotypes including commercial varieties and wild species). An amplicon of 450 bp and 120 bp specific to miR169 was observed in cluster bean varieties HG 2-20 and RGC 471, respectively. HG 2-20 is tolerant against bacterial leaf blight, alternaria blight, and root rot, while RGC 471 is tolerant to bacterial blight. The miR169 has been found to be associated with biotic stress tolerance (Song et al., 2018; Šečić et al., 2021). The appearance of amplicon specific to miR169 in these varieties clearly demonstrated their role in disease resistance in cluster bean (Unpublished data).

The miRNAs identified in this study are involved in regulating several genes and further gene expression studies are needed which will aid in the discovery of a new dimension of the miRNA regulatory network during plant growth and development in cluster bean.

5 CONCLUSION

Using ESTs from the NCBI database, 57 putative miRNA families were identified in cluster bean, with MFEI ranging from 0.59 to 1.36 (kcal/mol). A total of 623 target genes were predicted for 57 possible miRNA families, and their expression was suppressed either by cleavage or translational repression. Many of them were implicated in various metabolic processes. Furthermore, miRNA target prediction found that most of these genes are classified as transcription factors. A functional study of the targets revealed that they were engaged in the process of plant development. Most of the anticipated targets were shown to be regulated by several miRNAs, according to network analysis. This study provides an important view of conserved and novel miRNAs, their precursors and targets, and associated TFs. This will help the researchers working on cluster bean miRNAs in unraveling the complex gene regulatory molecular networks. The miRNs identified through *in silico* approach need further experimental validation.

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

NY conceived and designed the research. VC carried out the experiments and collected the data. VC, SJ, and NY analyzed the data. VC and SJ wrote the manuscript and NY edited and improved the manuscript. All the authors have read and approved the final manuscript.

FUNDING

The research was funded by the Director of Research, CCS Haryana Agricultural University, Hisar through Research Grant No. C(a) BMB-6-Plan (Agri.).

ACKNOWLEDGMENTS

The authors thank the Director of Research and Head of the Department of Molecular Biology, Biotechnology and Bioinformatics, CCS Haryana Agricultural University, Hisar for providing necessary support and facilities for this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.930113/full#supplementary-material

REFERENCES

Acharya, B. R., Sandhu, D., Dueñas, C., Ferreira, J. F. S., and Grover, K. K. (2022). Deciphering Molecular Mechanisms Involved in Salinity Tolerance in Guar (*Cyamopsis tetragonoloba* (L.) Taub.) Using Transcriptome Analyses. *Plants* 11 (3), 291. doi:10.3390/plants11030291

Al-Qurainy, F., Alshameri, A., Gaafar, A. R., Khan, S., Nadeem, M., Alameri, A. A., et al. (2019). Comprehensive Stress-Based De Novo Transcriptome Assembly and Annotation of Guar (*Cyamopsis tetragonoloba* (L.) Taub.): An Important Industrial and Forage Crop. *Int. J. Genomics.* 2019, 7295859. doi:10.3390/plants1103029110.1155/2019/7295859

Bansal, A., Singh, T. R., and Chauhan, R. S. (2017). A Novel miRNA Analysis Framework to Analyze Differential Biological Networks. *Sci. Rep.* 7 (1), 14604. doi:10.1038/s41598-017-14973-x

Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* 116 (2), 281–297. doi:10.1016/S0092-8674(04)00045-5

Bhatt, R. K., Sharma, R., Mahla, H. R., and Kalia, R. K. (2016). *Guar Production, Utilization and Marketing: Current Scenario and Future Prospects*. Jodhpur: Central Arid Zone Research Institute, 1–40.

Budak, H., and Akpinar, B. A. (2015). Plant miRNAs: Biogenesis, Organization and Origins. *Funct. Integr. Genomics.* 15 (5), 523–531. doi:10.1007/s10142-015-0451-2

Budheswar, D., Debashis, P., Sahu, J., Sahu, M., Sarma, K., Barooah, M., et al. (2013). *In Silico* Identification and Characterization of Conserved miRNAs and Their Target Genes in Sweet Potato (*Ipomoea batatas* L.) Expressed Sequence Tags (ESTs). *Plant Signal. Behav.* 8 (12), e26543. doi:10.4161/psb.26543

Carthew, R. W., and Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136 (4), 642–655. doi:10.1016/j.cell.2009.01.035

Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., et al. (2019). MADS-box Genes Are Key Components of Genetic Regulatory Networks Involved in Abiotic Stress and Plastic Developmental Responses in Plants. *Front. Plant Sci.* 10, 853. doi:10.3389/fpls.2019.00853

Catalanotto, C., Cogoni, C., and Zardo, G. (2016). MicroRNA in Control of Gene Expression: an Overview of Nuclear Functions. *Int. J. Mol. Sci.* 17 (10), 1712. doi:10.3390/ijms17101712

Chaudhary, V. (2019). *Development of Micro-RNA and EST-SSR Markers for Diversity Analysis in Cluster Bean [Cyamopsis tetragonoloba (L.) Taub*. Ph.D. Thesis. Hisar: CCS Haryana Agricultural University.

Chaudhury, A., Kaila, T., and Gaikwad, K. (2019). Elucidation of Galactomannan Biosynthesis Pathway Genes through Transcriptome Sequencing of Seeds Collected at Different Developmental Stages of Commercially Important Indian Varieties of Cluster Bean (*Cyamopsis Tetragonoloba* L.). *Sci. Rep.* 9 (1), 11539. doi:10.1038/s41598-019-48072-w

Chen, C.-Z., Li, L., Lodish, H. F., and Bartel, D. P. (2004). MicroRNAs Modulate Hematopoietic Lineage Differentiation. *Science* 303 (5654), 83–86. doi:10.1126/science.1091903

Chi, X., Yang, Q., Chen, X., Wang, J., Pan, L., Chen, M., et al. (2011). Identification and Characterization of microRNAs from Peanut (*Arachis hypogaea* L.) by High-Throughput Sequencing. *PloS One* 6 (11), e27530. doi:10.1371/journal.pone.0027530

Dehury, B., Panda, D., Sahu, J., Sahu, M., Sarma, K., Barooah, M., et al. (2013). In Silico identification and Characterization of Conserved miRNAs and Their Target Genes in Sweet Potato (*Ipomoea batatas* L.) Expressed Sequence Tags (ESTs). *Plant Signal. Behav.* 8 (12), e26543. doi:10.4161/psb.26543

Djami-Tchatchou, A. T., Sanan-Mishra, N., Ntushelo, K., and Dubery, I. A. (2017). Functional Roles of microRNAs in Agronomically Important Plants-Potential as Targets for Crop Improvement and Protection. *Front. Plant Sci.* 8, 378. doi:10.3389/fpls.2017.00378

Dreher, K., and Callis, J. (2007). Ubiquitin, Hormones and Biotic Stress in Plants. *Ann. Bot.* 99 (5), 787–822. doi:10.1093/aob/mcl255

Gaikwad, K., Ramakrishna, G., Srivastava, H., Saxena, S., Kaila, T., Tyagi, A., et al. (2020). Chromosome Scale Reference Genome of Cluster Bean (*Cyamopsis tetragonoloba* (L.) Taub.). Unpublished. doi:10.1101/2020.05.16.098434

Gasparis, S., Yanushevska, Y., and Nadolska-Orczyk, A. (2017). Bioinformatic Identification and Expression Analysis of New microRNAs from Wheat (*Triticum aestivum* L.). *Acta. Physiol. Plant.* 39, 2360. doi:10.1007/s11738-017-2530-6

George, A., Shah, P. A., and Shrivastav, P. S. (2019). Guar Gum: Versatile Natural Polymer for Drug Delivery Applications. *Eur. Polym. J.* 112, 722–735. doi:10.1016/j.eurpolymj.2018.10.042

Gigli-Bisceglia, N., Engelsdorf, T., and Hamann, T. (2020). Plant Cell Wall Integrity Maintenance in Model Plants and Crop Species-Relevant Cell Wall Components and Underlying Guiding Principles. *Cell. Mol. Life Sci.* 77, 2049–2077. doi:10.1007/s00018-019-03388-8

Goubet, F., Barton, C. J., Mortimer, J. C., Yu, X., Zhang, Z., Miles, G. P., et al. (2009). Cell Wall Glucomannan in Arabidopsis is Synthesised by CSLA Glycosyltransferases, and Influences the Progression of Embryogenesis. *Plant J.* 60 (3), 527–538. doi:10.1111/j.1365-313X.2009.03977.x

Grigoreva, E., Ulianich, P., Ben, C., Gentzbittel, L., and Potokina, E. (2019). First Insights into the Guar (*Cyamopsis tetragonoloba* (L.) Taub.) Genome of the 'Vavilovskij 130' Accession, Using Second and Third-Generation Sequencing Technologies. *Russ. J. Genet.* 55 (11), 1406–1416. doi:10.1134/S102279541911005X

Grigoreva, E., Tkachenko, A., Arkhimandritova, S., Beatovic, A., Ulianich, P., Volkov, V., et al. (2021a). Identification of Key Metabolic Pathways and Biomarkers Underlying Flowering Time of Guar (*Cyamopsis tertragonoloba* (L.) Taub.) via Integrated Transcriptome-Metabolome Analysis. *Genes* 12 (7), 952. doi:10.3390/genes12070952

Grigoreva, E., Barbitoff, Y., Changalidi, A., Karzhaev, D., Volkov, V., Shadrina, V., et al. (2021b). Development of SNP Set for the Marker-Assisted Selection of Guar (*Cyamopsis tetragonoloba* (L.) Taub.) Based on a Custom Reference Genome Assembly. *Plants* 10 (10), 2063. doi:10.3390/plants10102063

Gupta, O. P., Karkute, S. G., Banerjee, S., Meena, N. L., and Dahuja, A. (2017). Contemporary Understanding of miRNA-Based Regulation of Secondary Metabolites Biosynthesis in Plants. *Front. Plant Sci.* 8, 374. doi:10.3389/fpls.2017.00374

Hosseini, S. Z., Ismaili, A., Nazarian Firouzabadi, F., Fallahi, H., and Rezaeinejad, A. (2021). Identification of Heat Stress-Responsive microRNAs in Lentil (*Lens culinaris*) and Their Target Genes. *Iran. J. Biol.* 34 (1), 143–154.

Jangra, S., Chaudhary, V., and Yadav, N. R. (2018). "Transcription Factors and microRNA Interplay: A New Strategy for Crop Improvement," in *Transcriptional and Post-transcriptional Regulation*. Editor K. Ghedira (London: Intech Open). doi:10.5772/intechopen.75942

Kaila, T., Chaduvla, P., Rawal, H., Saxena, S., Tyagi, A., Mithra, S., et al. (2017). Chloroplast Genome Sequence of Clusterbean (*Cyamopsis tetragonoloba* L.): Genome Structure and Comparative Analysis. *Genes* 8 (9), 212. doi:10.3390/genes8090212

Kaur, S., and Santra, S. (2022). Application of Guar Gum and its Derivatives as Green Binder/Separator for Advanced Lithium-Ion Batteries. *ChemistryOpen* 11 (2), e202100209. doi:10.1002/open.202100209

Kim, H.-J., Baek, K.-H., Lee, B.-W., Choi, D., and Hur, C.-G. (2011). *In Silico* identification and Characterization of microRNAs and Their Putative Target Genes in Solanaceae Plants. *Genome* 54 (2), 91–98. doi:10.1139/G10-104

Kohli, D., Joshi, G., Deokar, A. A., Bhardwaj, A. R., Agarwal, M., Katiyar-Agarwal, S., et al. (2014). Identification and Characterization of Wilt and Salt Stress-Responsive microRNAs in Chickpea through High-Throughput Sequencing. *PloS one* 9 (10), e108851. doi:10.1371/journal.pone.0108851

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: Annotating High Confidence microRNAs Using Deep Sequencing Data. *Nucl. Acids Res.* 42 (D1), D68–D73. doi:10.1093/nar/gkt1181

Kumar, S., Palve, A. S., Patel, S. K., Selvanayagam, S., Sharma, R., and Rathore, A. (2020). Development of Genomic Microsatellite Markers in Cluster Bean Using Next-Generation DNA Sequencing and Their Utility in Diversity Analysis. *Curr. Plant Biol.* 21, 100134. doi:10.1016/j.cpb.2019.100134

Li, W., Wang, P., Li, Y., Zhang, K., Ding, F., Nie, T., et al. (2015). Identification of MicroRNAs in Response to Different Day Lengths in Soybean Using High-Throughput Sequencing and qRT-PCR. *PLoS ONE* 10 (7), e0132621. doi:10.1371/journal.pone.0132621

Masiero, S., Colombo, L., Grini, P. E., Schnittger, A., and Kater, M. M. (2011). The Emerging Importance of Type I MADS Box Transcription Factors for Plant Reproduction. *Plant Cell* 23 (3), 865–872. doi:10.1105/tpc.110.081737

Mudgil, D., Barak, S., and Khatkar, B. S. (2014). Guar Gum: Processing, Properties and Food Applications-A Review. *J. Food Sci. Technol.* 51 (3), 409–418. doi:10.1007/s13197-011-0522-x

Numnark, S., Mhuantong, W., Ingsriswang, S., and Wichadakul, D. (2012). C-mii: A Tool for Plant miRNA and Target Identification. *BMC Genomics* 13 (Suppl. 7), S16. doi:10.1186/1471-2164-13-S7-S16

O'Brien, J., Hayder, H., Zayed, Y., and Peng, C. (2018). Overview of microRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front. Endocrinol.* 9, 402. doi:10.3389/fendo.2018.00402

País, S. M., Téllez-Iñón, M. T., and Capiati, D. A. (2009). Serine/threonine Protein Phosphatases Type 2A and Their Roles in Stress Signaling. *Plant Signal. Behav.* 4 (11), 1013–1015. doi:10.4161/psb.4.11.9783

Pani, A., and Mahapatra, R. K. (2013). Computational Identification of microRNAs and Their Targets in *Catharanthus roseus* Expressed Sequence Tags. *Genomics Data* 1, 2–6. doi:10.1016/j.gdata.2013.06.001

Patanun, O., Lertpanyasampatha, M., Sojikul, P., Viboonjun, U., and Narangajavana, J. (2013). Computational Identification of microRNAs and Their Targets in Cassava (*Manihot esculenta* Crantz.). *Mol. Biotechnol.* 53 (3), 257–269. doi:10.1007/s12033-012-9521-z

Pérez-Rodríguez, P., Riaño-Pachón, D. M., Corrêa, L. G. 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 (Suppl. l_1), D822–D827. doi:10.1093/nar/gkp805

Peláez, P., Trejo, M. S., Iñiguez, L. P., Estrada-Navarrete, G., Covarrubias, A. A., Reyes, J. L., et al. (2012). Identification and Characterization of microRNAs in *Phaseolus vulgaris* by High-Throughput Sequencing. *BMC Genomics* 13, 83. doi:10.1186/1471-2164-13-83

Prabu, G. R., and Mandal, A. K. (2010). Computational Identification of miRNAs and Their Target Genes from Expressed Sequence Tags of Tea (*Camellia sinensis*). *Genomics Proteomics Bioinforma.* 8, 113–121. doi:10.1016/S1672-0229(10)60012-5

Qiu, C. X., Xie, F. L., Zhu, Y. Y., Guo, K., Huang, S. Q., Nie, L., et al. (2007). Computational Identification of microRNAs and Their Targets in *Gossypium hirsutum* Expressed Sequence Tags. *Gene* 395 (1-2), 49–61. doi:10.1016/j.gene.2007.01.034

Rajaprakasam, S., Rahman, H., Karunagaran, S., Bapu, K., J.R., G., Kulandivelu, G., et al. (2021). Comparative Transcriptome and Metabolome Profiling in the Maturing Seeds of Contrasting Cluster Bean (*Cyamopsis tetragonoloba* L. Taub) Cultivars Identified Key Molecular Variations Leading to Increased Gum Accumulation. *Gene* 791, 145727. doi:10.1016/j.gene.2021.145727

Ravelombola, W., Manley, A., Adams, C., Trostle, C., Ale, S., Shi, A., et al. (2021). Genetic and Genomic Resources in Guar: A Review. *Euphytica* 217 (11), 119. doi:10.1007/s10681-021-02929-2

Rawal, H. C., Kumar, S., Mithra S V, A., Solanke, A. U., Nigam, D., Saxena, S., et al. (2017). High Quality Unigenes and Microsatellite Markers from Tissue Specific Transcriptome and Development of a Database in Clusterbean (*Cyamopsis tetragonoloba*, L. Taub). *Genes (Basel)* 8 (11), 313. doi:10.3390/genes8110313

Sahu, S., Rao, A. R., Pandey, J., Gaikwad, K., Ghoshal, S., and Mohapatra, T. (2018). Genome-wide Identification and Characterization of lncRNAs and miRNAs in Cluster Bean (*Cyamopsis tetragonoloba*). *Gene* 667 (667), 112–121. doi:10.1016/j.gene.2018.05.027

Sahu, S., Sahu, T. K., Ghosal, S., Gaikwad, K., and Rao, A. R. (2020). Computational Analysis of SNPs and INDELs in Cluster Bean Cultivars Involved in Multiple Trait Expression. *Indian J. Genet. Plant Breed.* 80 (2), 179–185. doi:10.31742/IJGPB.80.2.8

Saya, L., Malik, V., Singh, A., Singh, S., Gambhir, G., Singh, W. R., et al. (2021). Guar Gum Based Nanocomposites: Role in Water Purification through Efficient Removal of Dyes and Metal Ions. *Carbohydr. Polym.* 261, 117851. doi:10.1016/j.carbpol.2021.117851

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

Šečić, E., Kogel, K.-H., and Ladera-Carmona, M. J. (2021). Biotic Stress-Associated microRNA Families in Plants. *J. Plant Physiol.* 263, 153451. doi:10.1016/j.jplph.2021.153451

Sharma, G., Sharma, S., Kumar, A., Al-Muhtaseb, A. a. H., Naushad, M., Ghfar, A. A., et al. (2018). Guar Gum and its Composites as Potential Materials for Diverse Applications: A Review. *Carbohydr. Polym.* 199, 534–545. doi:10.1016/j.carbpol.2018.07.053

Singh, N., Srivastava, S., Shasany, A. K., and Sharma, A. (2016). Identification of miRNAs and Their Targets Involved in the Secondary Metabolic Pathways of Mentha spp. *Comput. Biol. Chem.* 64, 154–162. doi:10.1016/j.compbiolchem.2016.06.004

Singroha, G., Sharma, P., and Sunkur, R. (2021). Current Status of microRNA -mediated Regulation of Drought Stress Responses in Cereals. *Physiol. Plant.* 172 (3), 1808–1821. doi:10.1111/ppl.13451

Song, S., Xu, Y., Huang, D., Ashraf, M. A., Li, J., Hu, W., et al. (2018). Identification and Characterization of miRNA169 Family Members in Banana (*Musa acuminata* L.) that Respond to *fusarium Oxysporum F.* Sp. Cubense Infection in Banana Cultivars. *Peer J.* 6, e6209. doi:10.7717/peerj.6209

Sunkar, R., and Jagadeeswaran, G. (2008). *In Silico* identification of Conserved microRNAs in Large Number of Diverse Plant Species. *BMC Plant Biol.* 8, 37. doi:10.1186/1471-2229-8-37

Tanwar, U. K., Pruthi, V., and Randhawa, G. S. (2017). RNA-seq of Guar (*Cyamopsis tetragonoloba*, L. Taub.) Leaves: De Novo Transcriptome Assembly, Functional Annotation and Development of Genomic Resources. *Front. Plant Sci.* 8, 91. doi:10.3389/fpls.2017.00091

Teplyakova, S. B., Volkov, V. A., Dzyubenko, E. A., and Potokina, E. K. (2019). Variability of the Photoperiod Response in Guar (*Cyamopsis tetragonoloba* (L.) Taub.) Genotypes of Different Geographic Origin. *Vavilov J. Genet. Breed* 23, 730–737. doi:10.18699/VJ19.547

Tiwari, M., Singh, B., Yadav, M., Pandey, V., and Bhatia, S. (2021). High Throughput Identification of miRNAs Reveal Novel Interacting Targets Regulating Chickpea-Rhizobia Symbiosis. *Environ. Exp. Bot.* 186, 104469. doi:10.1016/j.envexpbot.2021.104469

Tyagi, A., Nigam, D., Mithra, S. V. A., Solanke, A. U., Singh, N. K., Sharma, T. R., et al. (2018). Genome-Wide Discovery of Tissue-Specific miRNAs in Clusterbean (*Cyamopsis tetragonoloba*) Indicates Their Association With

Galactomannan Biosynthesis. *Plant Biotechnol. J.* 16 (6), 1241–1257. doi:10.1111/pbi.12866

Verma, D., and Sharma, S. K. (2021). Recent Advances in Guar Gum Based Drug Delivery Systems and Their Administrative Routes. *Int. J. Biol. Macromol.* 181, 653–671. doi:10.1016/j.ijbiomac.2021.03.087

Vivek, A. T. (2018). In Silico Identification and Characterization of microRNAs Based on EST and GSS in Orphan Legume Crop, *Lens culinaris* Medik (Lentil). *Agri Gene* 8, 45–56. doi:10.1016/j.aggene.2018.05.003

Wang, X.-J., Reyes, J. L., Chua, N.-H., and Gaasterland, T. (2004). Prediction and Identification of *Arabidopsis thaliana* microRNAs and Their mRNA Targets. *Genome Biol.* 5 (9), R65. doi:10.1186/gb-2004-5-9-r65

Wang, J., Yang, X., Xu, H., Chi, X., Zhang, M., and Hou, X. (2012). Identification and Characterization of microRNAs and Their Target Genes in *Brassica oleracea*. *Gene* 505 (2), 300–308. doi:10.1016/j.gene.2012.06.002

Xie, F., Frazier, T. P., and Zhang, B. (2010). Identification and Characterization of microRNAs and Their Targets in the Bioenergy Plant Switchgrass (*Panicum virgatum*). *Planta* 232, 417–434. doi:10.1007/s00425-010-1182-1

Yasin, J. K., Mishra, B. K., Pillai, M. A., Verma, N., Wani, S. H., Elansary, H. O., et al. (2020). Genome Wide *In-Silico* miRNA and Target Network Prediction from Stress Responsive Horsegram (*Macrotyloma uniflorum*) Accessions. *Sci. Rep.* 10, 17203. doi:10.1038/s41598-020-73140-x

You, G., Zu, B., Wang, B., Fu, Q., and Li, F. (2020). Identification of miRNA-mRNA-TFs Regulatory Network and Crucial Pathways Involved in Tetralogy of Fallot. *Front. Genet.* 11, 552. doi:10.3389/fgene.2020.00552

Zhang, B., Pan, X., Cobb, G. P., and Anderson, T. A. (2006). Plant microRNA: A Small Regulatory Molecule with Big Impact. *Dev. Biol.* 289 (1), 3–16. doi:10.1016/j.ydbio.2005.10.036

Zhang, B., Pan, X., and Stellwag, E. J. (2008). Identification of Soybean microRNAs and Their Targets. *Planta* 229 (1), 161–182. doi:10.1007/s00425-008-0818-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.

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 01 July 2022
doi: 10.3389/fgene.2022.933022

Genome-Wide Identification of Powdery Mildew Responsive Long Non-Coding RNAs in *Cucurbita pepo*

Jiaxing Tian[1,2], *Guoyu Zhang*[1,2], *Fan Zhang*[1,2], *Jian Ma*[1,2], *Changlong Wen*[1,2]* and *Haizhen Li*[1,2]*

[1]Beijing Vegetable Research Center (BVRC), Beijing Academy of Agriculture and Forestry Sciences (BAAFS), Beijing, China, [2]Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), Ministry of Agriculture, Beijing, China

Cucurbita pepo L. is an essential economic vegetable crop worldwide, and its production is severely affected by powdery mildew (PM). However, our understanding of the molecular mechanism of PM resistance in *C. pepo* is very limited. Long non-coding RNAs (lncRNAs) play an important role in regulating plant responses to biotic stress. Here, we systematically identified 2,363 reliably expressed lncRNAs from the leaves of PM-susceptible (PS) and PM-resistant (PR) *C. pepo*. The *C. pepo* lncRNAs are shorter in length and expressed at a lower level than the protein-coding transcripts. Among the 2,363 lncRNAs, a total of 113 and 146 PM-responsive lncRNAs were identified in PS and PR, respectively. Six PM-responsive lncRNAs were predicted as potential precursors of microRNAs (miRNAs). In addition, 58 PM-responsive lncRNAs were predicted as targets of miRNAs and one PM-responsive lncRNA was predicted as an endogenous target mimic (eTM). Furthermore, a total of 5,200 potential cis target genes and 5,625 potential trans target genes were predicted for PM-responsive lncRNAs. Functional enrichment analysis showed that these potential target genes are involved in different biological processes, such as the plant-pathogen interaction pathway, MAPK signaling pathway, and plant hormone signal transduction pathway. Taken together, this study provides a comprehensive view of *C. pepo* lncRNAs and explores the putative functions of PM-responsive lncRNAs, thus laying the foundation for further study of the regulatory mechanisms of lncRNAs responding to PM.

Keywords: ***Cucurbita pepo*, long non-coding RNA, powdery mildew, target gene, microRNA**

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Qing Yang,
Beijing Forestry University, China
Jin Zhang,
Zhejiang Agriculture and Forestry University, China

***Correspondence:**
Changlong Wen
wenchanglong@nercv.org
Haizhen Li
lihaizhen@nercv.org

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *30 April 2022*
Accepted: *23 May 2022*
Published: *01 July 2022*

Citation:
Tian J, Zhang G, Zhang F, Ma J, Wen C and Li H (2022) Genome-Wide Identification of Powdery Mildew Responsive Long Non-Coding RNAs in Cucurbita pepo. Front. Genet. 13:933022. doi: 10.3389/fgene.2022.933022

INTRODUCTION

Previous transcriptome analyses of many different organisms showed that most transcribed RNAs do not have the ability to encode proteins. Based on transcript lengths, these non-coding RNAs (ncRNAs) can be mainly divided into small non-coding RNAs (sRNAs) and long non-coding RNAs (lncRNAs). LncRNAs refer to ncRNAs that are larger than 200 bp in length but have no apparent coding potential (Kapranov et al., 2007). Depending on the genomic origins relative to the genomic protein-encoding genes, lncRNAs are classified into different categories: intergenic, intronic, antisense, and sense (Ponting et al., 2009). Characteristic analysis showed that the expression level of lncRNAs was lower than that of mRNAs. In addition, there is often low evolutionary conservation and tissue-specific expression type for lncRNAs (Liu et al., 2015). Previous analyses have indicated that lncRNAs can interact with other molecules, functioning at various levels such as

the transcriptional level and translational regulation (Wang and Chekanova, 2017; Lucero et al., 2021). They can regulate protein modification, chromatin remodeling, RNA metabolism, and protein modification *in vivo* through cis- or trans-activation. With the advantage of sequencing technologies, an increasing number of lncRNAs have been recognized in plants, including *Arabidopsis thaliana* (Liu et al., 2012), maize (Li et al., 2014), *Solanum lycopersicum* (Zhu et al., 2015), *Populus tomentosa* (Tian et al., 2016) and *Cucumis melo* (Gao et al., 2020). Emerging studies have indicated the important regulatory role of plant lncRNAs in different biological processes, including photomorphogenesis (Wang et al., 2014), reproduction (Wang et al., 2017), flowering (Wu et al., 2019), development of fruit (Tang et al., 2021), and stress responses (Zhu et al., 2014; Qin et al., 2017; Zhang et al., 2018; Jiang et al., 2019).

Recent research has confirmed that plant lncRNAs play vital roles in stress responses, including pathogen infection. So far, an increasing number of lncRNAs associated with pathogen infection have been found in plants, for instance, in wheat against powdery mildew (PM) (Xin et al., 2011; Zhang H. et al., 2016), in *Arabidopsis* against *Pseudomonas syringe* (Seo et al., 2017), in tomato against *Phytophthora infestans* (Cui et al., 2017; Cui et al., 2019; Jiang et al., 2019; Cui et al., 2020; Hou et al., 2020), in melon against PM (Gao et al., 2020), and in Chinese cabbage against downy mildew (Zhang B. et al., 2021). For example, it was reported that *Arabidopsis* lncRNA ELENA1 plays an important role in enhancing resistance against *Pseudomonas syringe* (Seo et al., 2017). In addition to model plants, pathogen-responsive lncRNAs have also been found in crop plants. Recently, it was reported that the cotton lncRNA (GhlncLOX3) was reported to affect the resistance to *Verticillium dahlia* by influencing the expression of *GhLOX3* related to jasmonic acid biosynthesis (Wang et al., 2021). In tomatoes, lncRNA16397 enhanced the resistance to *Phytophthora infestans* by interacting with glutaredoxin genes (Cui et al., 2017). In addition, tomato lncRNA23468 and lncRNA39026 play important roles in response to *Phytophthora infestans* infection by functioning as endogenous target mimics (eTMs) of miRNAs (Jiang et al., 2019; Hou et al., 2020). More recently, a candidate lncRNA related to downy mildew resistance was identified in Chinese cabbage (Zhang B. et al., 2021). Among the pathogens interacting with plants, there has been considerable research on PM, and several PM-related lncRNAs have been identified. In wheat, PM-induced lncRNAs were identified and characterized using a comparative expression profile analysis of PM-susceptible and PM-resistant wheat (Xin et al., 2011). In melon, several lncRNAs significantly differentially expressed after PM infection were also identified (Gao et al., 2020; Zhou et al., 2020). Moreover, 71 lncRNAs responsive to PM were identified in *Vitis vinifera* and their potential functions related to defense response were further examined (Bhatia et al., 2021). These results suggest that lncRNAs play an essential regulatory role in plant-pathogen interactions.

Cucurbita pepo L. (zucchini, squash) is an essential economic vegetable crop worldwide (Hafez et al., 2018). However, *C. pepo* is very susceptible to the cucurbit PM fungus. Thus far, PM has become one of the most serious diseases affecting *C. pepo* yields in the world. PM is a common and widely distributed fungal disease that affects different kinds of plants (Kusch and Panstruga, 2017; Chen et al., 2020). In cucurbit crops, powdery mildew disease is mainly caused by *Phytophthora xanthii* (Perez-Garcia et al., 2009). It can reduce the fruit yield and quality of *C. pepo* by affecting its photosynthesis, growth, and fruit development (Barickman et al., 2017). The use of chemical and biological fungicides has become one of the main measures for agricultural PM control. However, this strategy increases the resistance of PM fungi and may negatively affect the environment and human health (Zhang S. et al., 2021). Thus, it is of great significance to study the molecular mechanism of PM resistance in the molecular breeding of resistant varieties. The mechanisms of plant defense against PM have been previously reported (Ellinger et al., 2013; Hu et al., 2019). Additionally, several PM resistance-related genes have been identified in wheat (Qie et al., 2019), cucumber (Liu et al., 2017; Xu et al., 2017; Xu et al., 2019), and melon (Romero et al., 2008). In cucumber, several *MLO-like* genes have been identified as susceptibility genes to PM (Schouten et al., 2014; Berg et al., 2015). In addition, *CmMLO2* has been identified in muskmelon as a gene related to PM pathogenesis (Cheng et al., 2013). In pumpkin (*Cucurbita moschata* Duch.), *WRKY21*, *MLO3*, and *SGT1* were identified as candidate genes conferring PM-resistance (Guo et al., 2018). Moreover, the overexpression of *CmSGT1* (Guo et al., 2019) and *CmbHLH87* (Guo et al., 2020) can increase PM resistance in tobacco. Although these studies increase our understanding of the mechanism of plant resistance to PM, the molecular mechanisms underlying the PM resistance in *C. pepo* remain largely unknown. Because PM is one of the factors that severely affect *C. pepo* production, identification of PM-responsive lncRNAs in *C. pepo* is necessary.

In our study, we first conducted systematic identification of lncRNAs in *C. pepo* leaves and then described comprehensive PM-responding lncRNA profiles. In total, we identified 2,363 lncRNAs, including 242 lncRNAs differentially expressed after PM inoculation. To explore the potential function of these PM-responsive lncRNAs, we further predicted interactions between lncRNAs and miRNAs, as well as interactions between lncRNAs and genes. Functional enrichment analysis found that the potential target genes of PM-responsive lncRNAs were significantly associated with various biological processes, such as the plant–pathogen interaction pathway, MAPK signaling pathway, and plant hormone signal transduction pathway. Thus, we selected some candidate lncRNAs with potentially important roles related to PM resistance in *C. pepo*. Our results provide a rich resource for the exploration of the functional roles of *C. pepo* lncRNAs in PM resistance and lay the foundation for the molecular breeding of varieties resistant to PM.

MATERIALS AND METHODS

Plant Materials and Powdery Mildew Inoculation

We used *C. pepo* inbred lines "PS" and "PR," which are susceptible and resistant to PM, respectively, as the host plant.

The seeds of "PS" and "PR" were provided by the National Engineering Research Center for Vegetables (NERCV), Beijing, China. The sterilized seeds were germinated on moistened filter paper in a growth chamber (28°C in darkness). The germinated seeds were then planted in 9-cm-deep plastic pots containing sterilized soil, with one seed per pot. They were grown in a growth chamber with a photoperiod of 16/8 h (day/night) at 28°C/20°C (day/night). PM fungus (*P. xanthii*) was collected from *C. pepo* naturally infected with PM and grown at the NERCV experimental farm. The powdery mildew fungus is purified by inoculating susceptible varieties. The old conidia on the highly infected leaf surfaces were removed 24 h before inoculation so that new conidia could be produced. Then a spore suspension was prepared by soaking heavily infected leaves in tap water containing 0.01% Tween-20. PM inoculation was conducted with *C. pepo* seedlings with three true leaves by spraying with a freshly prepared spore suspension (10^6spores/ml). The inoculated seedlings were placed in a growth chamber as described above. Leaves of both PS and PR plants were harvested at 0 (control), 12, 24, and 48 h post-inoculation (hpi) with three biological replicates for each inbred line per time point. All the harvested leaves were frozen in liquid nitrogen immediately after collection and stored at −80°C until used.

Physiological Characteristics Measurement

Six physiological characteristics were measured using leaf samples collected from PS and PR seedlings at 0, 12, 24, and 48 hpi. Total protein levels, malondialdehyde (MDA) content, the activities of superoxide dismutase (SOD), catalase (CAT), peroxidase (POD), and phenylalanine ammonia-lyase (PAL) were measured using commercial assay kits from Nanjing Jiancheng Bioengineering Institute (Nanjing, China). Detailed procedures were carried out in accordance with the manufacturer's instructions. All the physiological characteristics above were determined using a Tecan Infinite M1000 microplate reader.

RNA Extraction, Library Construction, and Sequencing

High-throughput RNA-seq was conducted using leaf samples collected from the PS and PR seedlings at 0 and 24 hpi. Total RNAs were extracted from all samples using TRIzol reagent (Invitrogen, CA, United States) according to the manufacturer's instructions. After the removal of ribosomal RNA, the generation of sequencing libraries was conducted with a NEBNextR UltraTM Directional RNA Library Prep Kit for IlluminaR (New England Biolabs, United States). Finally, 12 strand-specific libraries were constructed to generate 150 bp paired-end reads on the Illumina HiSeq 4000 platform.

Identification of lncRNAs

By removing reads containing adapters, ploy-N, and low quality, we obtained clean data. The clean reads were mapped to the *C. pepo* reference genome (version 4.1) using HISAT2 (Kim et al., 2015). The genome of *C. pepo* was downloaded from CuGenDB (http://cucurbitgenomics.org/). Then the transcriptome was assembled using StringTie software (version 1.3.1) (Pertea et al., 2016) based on reads with no more than two mismatches. All transcripts were first aligned to housekeeping ncRNA databases to exclude tRNAs, snRNAs, and snoRNAs. Then, the gffcompare program was used for the annotation of assembled transcripts, and the remaining unannotated transcripts were used to identify putative lncRNAs. The criteria used to screen lncRNA candidates are: length >200 bp and exons ≥2.Finally, the protein-coding potential was evaluated using the Coding Potential Calculator (CPC), Coding-Non-Coding-Index (CNCI), Pfam Scan (Pfam), and Coding Potential Assessment Tool (CPAT). Only transcripts with no protein-coding potential in all the four databases were identified as candidate lncRNAs. The types of lncRNAs were classified based on their genomic locations (Roberts et al., 2011).

Characteristic Analysis of lncRNAs

To identify known lncRNAs, lncRNA sequences were subjected to BLAST searches against other plant lncRNAs in GREENC and CANTATAdb using BLASTN (coverage >80% and E-value < 1e-10). We further calculated the repetitive element and GC content of identified lncRNAs using Repeat Masker and EMBOSS explorer's geecee tool, respectively.

Differential Expression Analysis of lncRNAs

To identify lncRNAs related to PM resistance in *C. pepo*, the differential expression was analyzed between 24 hpi and 0 hpi for both PS and PR plants using the DESeq R package (Love et al., 2014). The fold-change (FC) was calculated, representing the change in lncRNA expression between 24 hpi and 0 hpi. The criteria for identifying differentially expressed lncRNA (DEL) is: log2FC ≥ 1 or ≤ −1 with p-value < 0.05.

Predicting Cis-Regulatory Elements of PM-Responsive lncRNAs

In order to find cis-acting motifs, the 1000 bp sequences upstream of the DELs were acquired and submitted to the "Search for Care" tool of the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Prediction of Interactions Between lncRNAs and miRNAs

BLASTN was used to compare all lncRNA sequences with miRNA precursor sequences (identity >90% and E-value < 1.0E-5), so as to predict the possibility of lncRNA as a potential miRNA precursor. No mismatch was allowed in the mature sequence region of miRNA. The lncRNAs potentially targeted by miRNAs were predicted using psRNATarget (https://www.zhaolab.org/psRNATarget/) (Dai and Zhao 2011), with an expectation ≤5. The eTMs for miRNAs were predicted by combining RNAhybrid (https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid/) (Rehmsmeier et al., 2004) with the rules established by Hou et al. (2020). For degradome analysis, an RNA mixture from 12 leaf samples was pooled for degradome library construction and then sequenced on an Illumina HiSeq 2500 instrument (LC Sciences, Hangzhou, China).

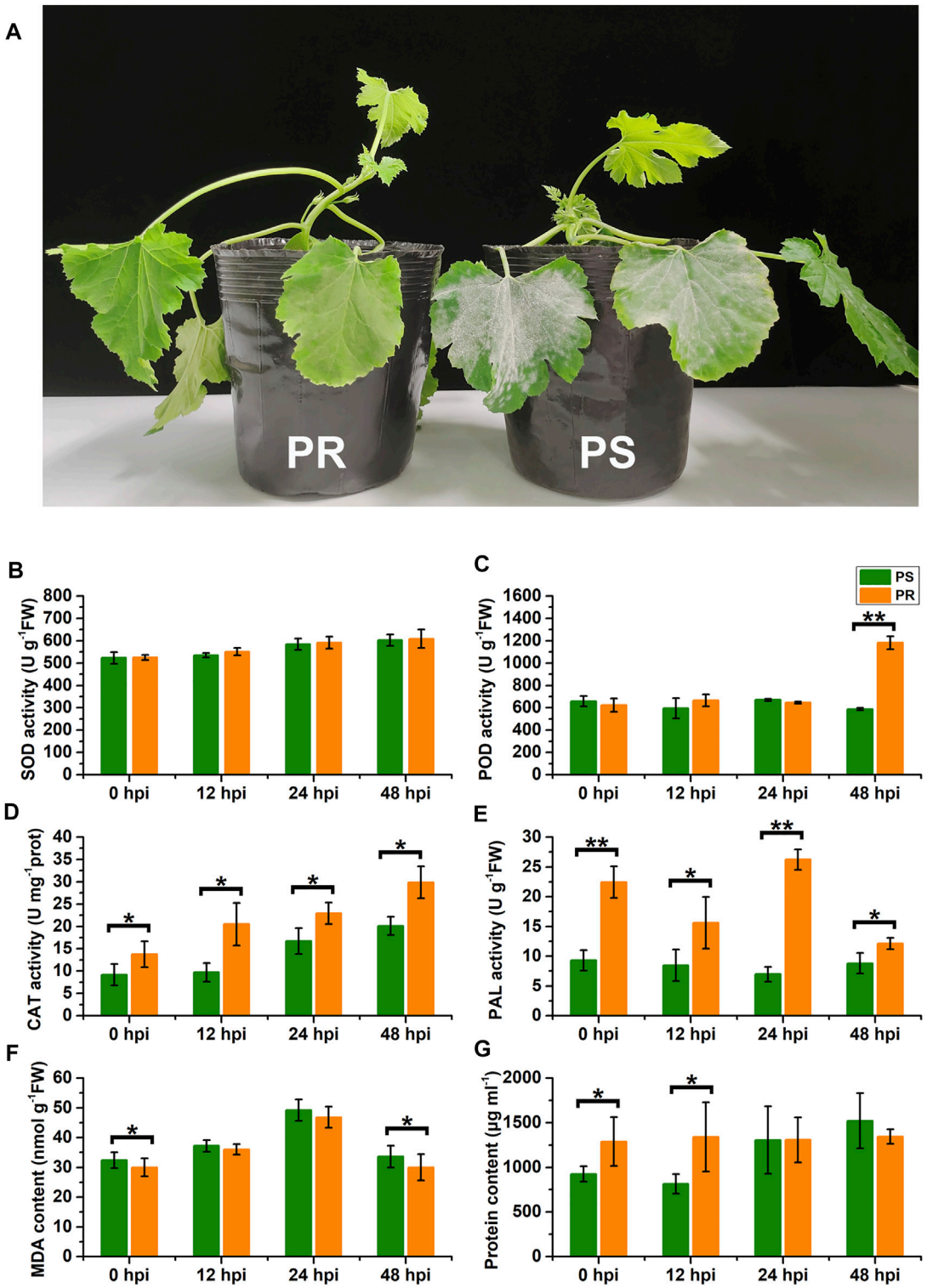


FIGURE 1 | Effects of powdery mildew inoculation on physiological characteristics. **(A)** Different phenotypes of susceptible and resistant *C. pepo* at 7 days after PM inoculation. **(B–G)** Changes in superoxide dismutase (SOD) activity, peroxidase (POD) activity, catalase (CAT) activity, phenylalanine ammonia-lyase (PAL) activity, malondialdehyde (MDA) content, and total protein concentration within leaves of *C. pepo* 0, 12, 24 and 48 h post PM inoculation. PS represents PM-susceptible *C. pepo* and PR represents PM-resistant *C. pepo*. Error bars indicate standard deviations among three biological replicates ($n = 3$). Asterisks indicate significant differences between PS and PR ("*"represents $p < 0.05$; "**" represents $p < 0.01$).

Prediction and Functional Analysis of Target Genes

We predicted potential target genes for PM-responsive lncRNAs based on their regulation modes. Prediction of cis target genes was done according to the positional relationship between lncRNAs and genes. The adjacent genes within 100 kb upstream or downstream of lncRNAs were regarded as potential cis target genes (Jia et al., 2010). Trans target gene prediction was conducted based on expression correlation analysis of lncRNA and mRNA. Trans target genes were genes with a correlation absolute value greater than 0.9 and *p*-value less than 0.01. Gene ontology (GO) enrichment analysis of the target genes was implemented using the topGO R packages. For pathway enrichment analysis, KOBAS (Xie et al., 2011) software was used.

Validation of PM-Responsive lncRNAs by qRT-PCR

Eight PM-responsive lncRNAs were selected and validated using quantitative real-time PCR (qRT-PCR). Total RNA was extracted from leaf samples using Trizol reagent (Thermo Fisher Scientific, United States). Then lnRcute lncRNA First-Strand cDNA Synthesis Kit (Tiangen, China) was used for reverse transcription. The qRT-PCR reactions were conducted with lnRcute lncRNA qPCR Detection Kit (Tiangen, China) on an ABI7500 system (Thermo Fisher Scientific, Waltham, MA, United States). All qRT-PCR amplifications were carried out in triplicate. The *C. pepo actin* gene (Cp4.1LG04g06840) was used as an internal reference gene. The $2^{-\Delta\Delta}$Ct method (Livak and Schmittgen, 2001) was used for data analysis. Primers used for qRT-PCR are listed in **Supplementary Table S1**.

RESULTS

Powdery Mildew Inoculation Affects the Physiological Characteristics of *C. pepo*

To explore the influence of PM disease on the physiology of *C. pepo*, we measured the activities of SOD, POD, CAT, PAL, the MDA content, and the total protein concentration in the leaves of susceptible (PS) and resistant (PR) *C. pepo* harvested at 0, 12, 24, and 48 hpi. The phenotypes of PS and PR plants were significantly different 7 days after PM inoculation (**Figure 1A**). As shown in **Figures 1B–G**, PM inoculation affected all six physiological characteristics from 0 to 48 hpi, and some physiological characteristics significantly differed between PS and PR. For example, the POD activity of PS plants did not show a significant change after PM inoculation, but the POD activity of PR plants significantly increased at 48 hpi (**Figure 1C**). In addition, the CAT activity gradually increased after powdery mildew inoculation, with significant differences between PS and PR plants at 0, 12, 24, and 48 hpi, respectively. Altogether, these differences at the physiological level suggest that the transcriptome of *C. pepo* may have significantly changed after PM inoculation. Moreover, the difference between PS and PR plants in physiological characteristics suggests their different responses to PM inoculation.

Genome-wide Identification and Characterization of lncRNAs in *C. pepo*

To systematically identify lncRNAs in *C. pepo*, we performed RNA sequencing using leaves of PS and PR plants at 0 and 24 hpi. Detailed information regarding RNA-Seq data is provided in **Supplementary Table S2**. As shown in **Supplementary Table S2**, approximately 93.58–95.42% of clean reads were uniquely mapped to the *C. pepo* genome. In total, 2,363 candidate lncRNAs were predicted (**Figure 2A**), including 1,158 lincRNAs, 355 antisense lncRNAs, 462 intronic lncRNAs, and 388 sense lncRNAs (**Figure 2B**). The distribution analysis showed that these 2,363 lncRNAs were evenly distributed on all the 20 *C. pepo* chromosomes (**Figure 2A**). The details of the 2,363 lncRNAs are provided in **Supplementary Table S3**.

The characteristics of *C. pepo* lncRNAs were further analyzed. BLAST analysis of *C. pepo* lncRNAs against the plant non-coding RNAs databases suggested that the overwhelming majority (99.83%) of *C. pepo* lncRNAs could be species-specific. These 2,363 lncRNAs were distributed in the *C. pepo* genome at a density of 9.73 lncRNAs per Mb. The mean GC content of the lncRNAs predicted in this study was 39.47%, and more than half of the lncRNAs (75.92%) did not contain repeated sequences. To further investigate the features of lncRNAs, the length and expression level of *C. pepo* lncRNAs were compared with mRNAs. In general, all the identified lncRNAs showed shorter lengths and lower expression levels as compared to mRNAs in *C. pepo* leaves (**Figures 2C,D**). For example, the mean length of *C. pepo* lncRNAs is 766 nt, which was shorter than the mRNAs of *C. pepo* (median length of 2,255 nt) (**Figure 2C**).

Identification of Differentially Expressed lncRNAs After Powdery Mildew Inoculation

To identify lncRNAs responsive to PM infection, differential expression analysis was conducted for PS and PR plants. The number of DELs in PS at 24 hpi was 113, with 52 (46%) up-regulated and 61 (54%) down-regulated (**Figure 3A**, **Supplementary Table S4**). For PR, 146 DELs were identified at 24 hpi, with 46 (32%) up-regulated and 100 (68%) down-regulated (**Figure 3A**; **Supplementary Table S5**). Compared with PS, more DELs were identified in PR. In addition, there were more down-regulated DELs than up-regulated DELs.

In total, we identified 242 DELs after PM infection ($\log_2$FC > 1 or < −1, and $p < 0.05$), including 39 sense lncRNAs, 33 antisense lncRNAs, 118 lincRNAs, and 52 intronic lncRNAs. Among the 242 DELs, 96 and 129 lncRNAs were specifically differentially expressed in PS and PR plants, respectively (**Figure 3B**). Additionally, 17 DEL were differentially expressed both in PS and PR plants at 24 hpi (**Figure 3B**), with four of them regulated in different directions in PS and PR plants (**Figure 3C**). For example, MSTRG.24598.1 was significantly up-regulated at 24 hpi in PS plants but was significantly down-regulated in PR plants.

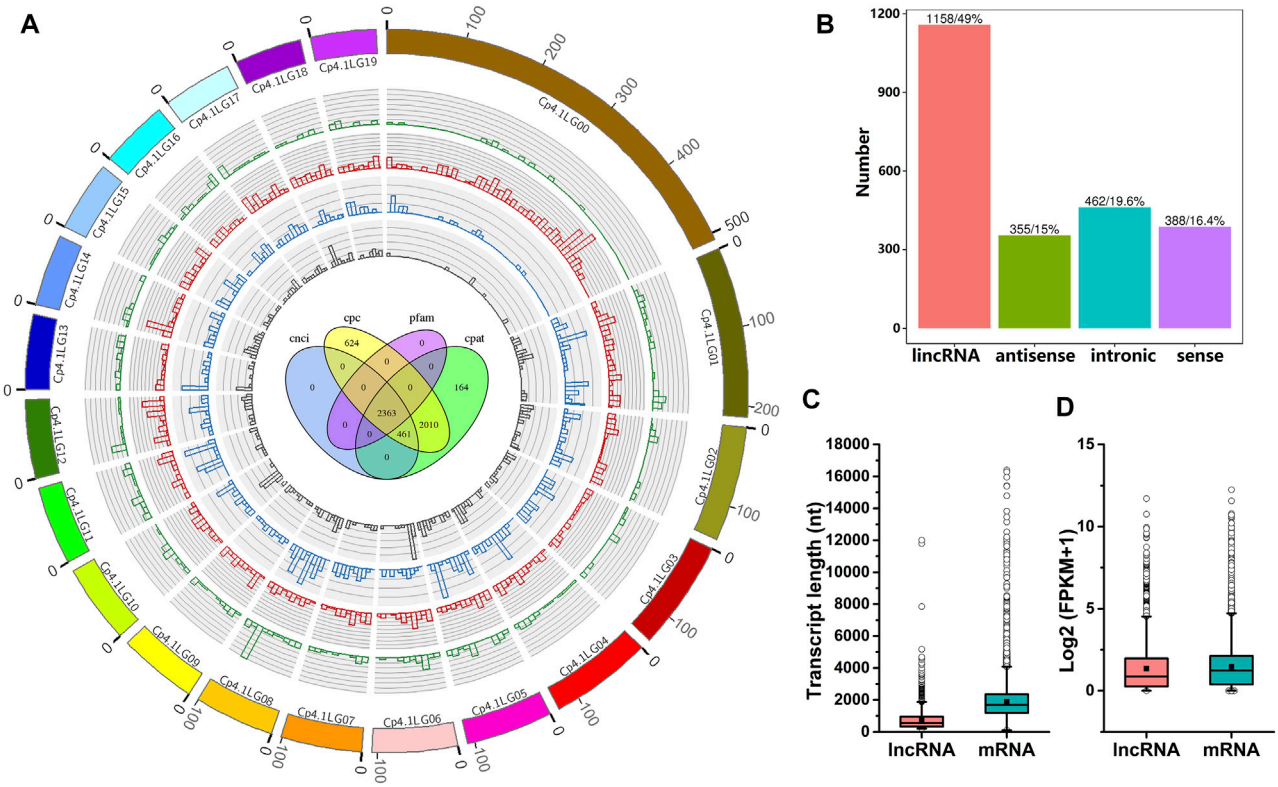


FIGURE 2 | Characteristics of *C. pepo* lncRNAs. **(A)** Distribution of lncRNAs along *C. pepo* chromosomes. The scale unit of chromosomes is 10^5 bp. Green represents sense lncRNAs, red represents lincRNAs, blue represents intronic lncRNAs and grey represents antisense lncRNAs. Venn diagram in the center shows the number of lncRNAs predicted by evaluating protein-coding potential in the Coding Potential Calculator (CPC), Coding-Non-Coding-Index (CNCI), Pfam Scan (Pfam), and Coding Potential Assessment Tool (CPAT). **(B)** Classification of *C. pepo* lncRNAs according to their genomic positions. **(C,D)** Comparisons between lncRNA and mRNA. lncRNAs are shorter **(C)** and expressed at lower levels **(D)** than protein-coding transcripts.

Cis-Elements in the Promoters of PM-Responsive lncRNA Genes

Cis-acting elements are related to responses to stress in plants. To identify cis-acting elements of the PM-responsive lncRNAs, we analyzed the 1000-bp upstream sequence of the 242 PM-responsive lncRNAs using PlantCARE. Some cis-acting elements associated with defense and hormone responsiveness were found in the promoter sequence of 205 PM-responsive lncRNAs (84.71%) (**Figure 4A**; **Supplementary Table S6**). As shown in **Figure 4B**, cis-acting elements associated with defense and stress responsiveness were identified in 50 PM-responsive lncRNAs. Among all the cis-acting elements predicted in this study (**Supplementary Table S6**), ABRE (abscisic acid-responsive) was the most numerous cis-acting element.

Interactions Between lncRNAs and miRNAs

Because lncRNA-miRNA interactions might play a non-negligible role in resistance to disease in plants, all the lncRNAs identified in this study were used to predict potential interactions with miRNAs. In total, 34 lncRNAs were identified as potential precursors of 10 miRNAs (**Supplementary Table S7**), including miR166, which may play an important role in regulating plant responses to pathogens. Among the 34 lncRNAs that are potential precursors of miRNAs, six lncRNAs were PM-responsive. For example, MSTRG.12356.1, which was up-regulated in PR plants after PM inoculation, was predicted to be the precursor of cpe-miR166 (**Figure 5A**). A total of 312 lncRNAs were identified as potential targets of 20 miRNAs belonging to 13 families (**Supplementary Table S8**). Among them, eight lncRNA-miRNA interactions were verified by degradome sequencing (**Supplementary Table S9**). For example, MSTRG.28493.1 was targeted by cpe-miR396b. The degradome T-plot of MSTRG.28493.1 showed a single clear peak at the degradation site (**Figure 5B**). In the present study, it was found that three lncRNAs, including one PM-responsive lncRNA (MSTRG.28570.4), were potential target mimics of two miRNAs (**Figure 5C**). Our results indicated that lncRNAs may participate in the response to PM by interacting with miRNAs.

Target Prediction of PM-Responsive lncRNAs

To analyze the potential functions of PM-responsive lncRNAs, the potential cis and trans target genes of these lncRNAs were predicted. As a result, we identified 5,200 cis target genes for 241 PM-responsive lncRNAs and 5,625 potential trans target

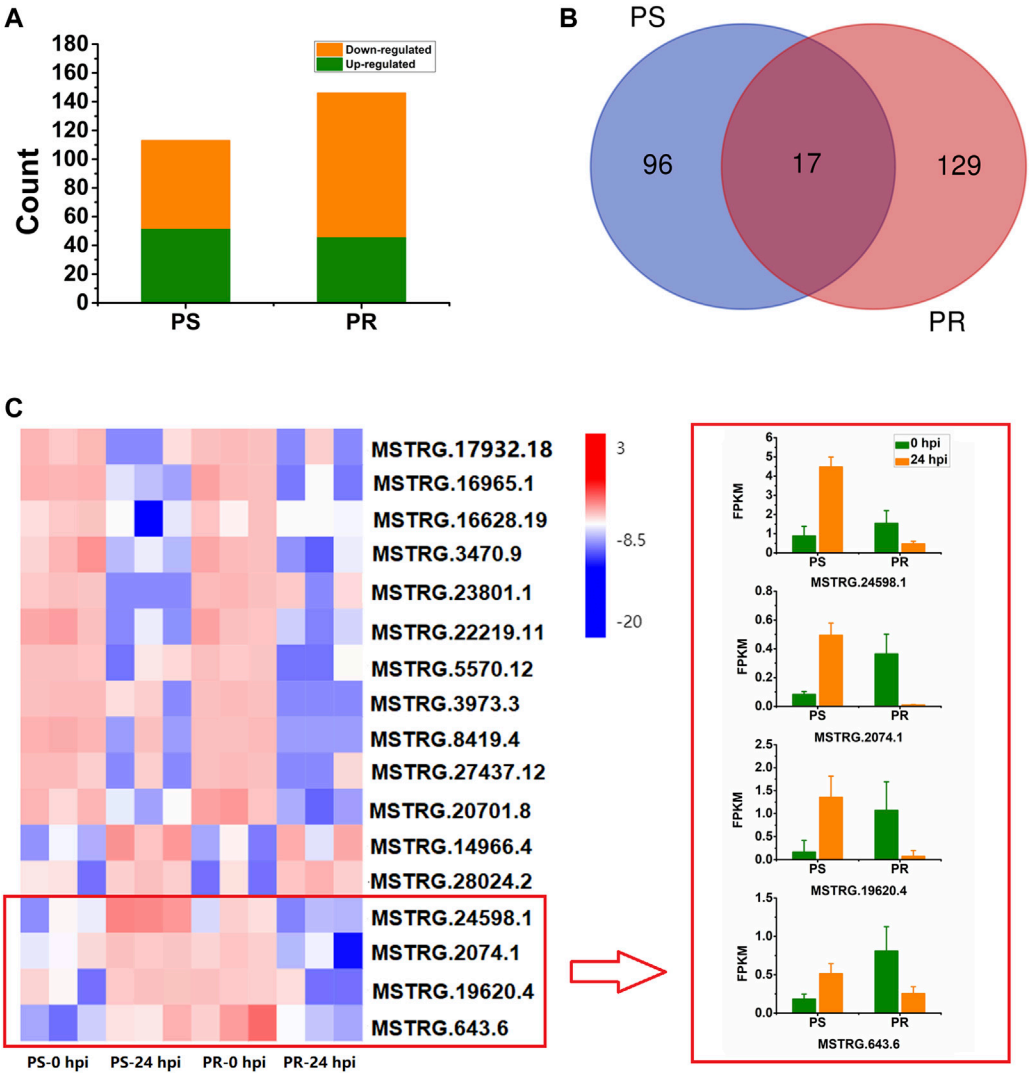


FIGURE 3 | Statistical analysis and expression of PM-responsive lncRNAs in PM-susceptible and PM-resistant *C. pepo*. **(A)** Number of down- and up-regulated lncRNAs at 24 hpi in PS and PR. **(B)** Number of shared and specific PM-responsive lncRNAs in PS and PR. **(C)** Expression of the 17 PM-responsive lncRNAs shared by PS and PR. Four of them showed inconsistent expression patterns between PS and PR. PS represents PM-susceptible *C. pepo* and PR represents PM-resistant *C. pepo*. Error bars indicate standard deviations among three biological replicates (n = 3).

genes for 227 PM-responsive lncRNAs (**Supplementary Table S10**). In addition, 30 potential target genes were predicted as both cis and trans targets of 21 PM-responsive lncRNAs. In total, 16,758 potential lncRNA-target pairs were identified. The results suggest that these PM-responsive lncRNAs in *C. pepo* may play important roles in responding to PM by interacting with their cis or trans target genes.

To explore the expression patterns of PM-responsive lncRNAs and their corresponding target genes, the trends of expression after PM inoculation were analyzed and compared (**Figure 6A**). The lncRNAs and target genes showed not only the same expression tendency but also the opposite expression tendency. The diverse expression relationships between PM-responsive lncRNAs and target genes indicate that the regulatory mechanism of lncRNAs is complex. After PM inoculation, the expression trend of most differentially expressed target genes was the same as that of the corresponding lncRNA. Among all the lncRNA-target pairs, the percentage showing the same expression trend after PM inoculation in PS and PR plants was 93.97 and 82.83%, respectively. For cis target genes, there was little difference in the number of target genes with the same or opposite expression trend as lncRNAs. However, for trans target genes, there were far more target genes with the same expression trend as lncRNAs than those with the opposite expression trend. The qRT-PCR was conducted to confirm the relationship between the expression levels of MSTRG.24598.1

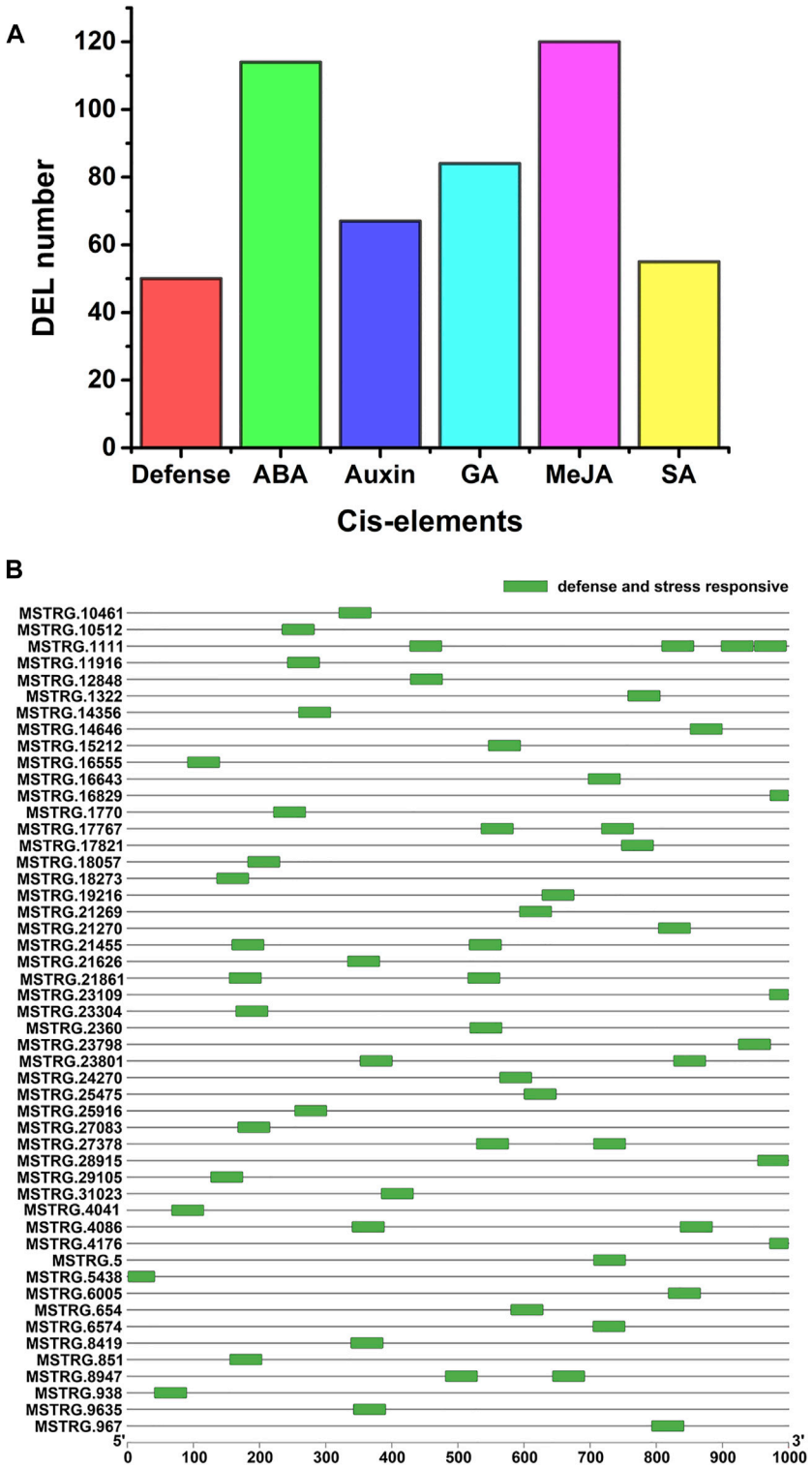


FIGURE 4 | Cis-elements in the promoters of the PM-responsive lncRNA genes. **(A)** Different types of cis-acting elements were identified. Defense represents defense and stress-responsive cis-element, ABA represents abscisic acid-responsive cis-element, Auxin represents auxin-responsive cis-element, GA represents gibberellin responsive cis-element, MeJA represents Methyl Jasmonate responsive cis-element and SA represents salicylic acid-responsive cis-element. **(B)** Distribution of cis-acting elements involved in defense and stress responses in PM-responsive lncRNAs.

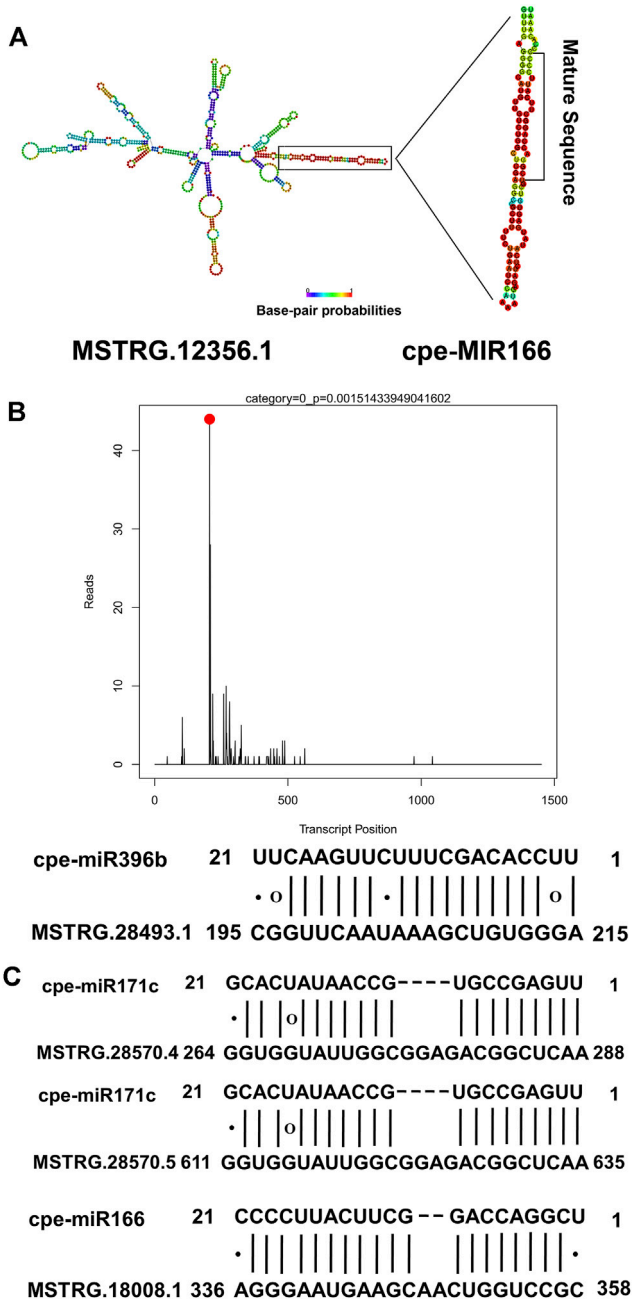


FIGURE 5 | Interactions between lncRNAs and miRNAs. **(A)** MSTRG.12356.1 was predicted to be a potential precursor of cpe-miR166.**(B)** MSTRG.28493.1 was predicted as a potential target of cpe-miR396b. The degradome T-plot of MSTRG.28493.1 showed a single clear peak at the degradation site. **(C)** Three lncRNAs, including one PM-responsive lncRNA (MSTRG.28570.4) were predicted as target mimics of two miRNAs.

and its three putative target genes (**Figure 6B**). As shown in **Figure 6B**, MSTRG.24598.1 and its target genes were up-regulated after PM inoculation in PS, but down-regulated after PM inoculation in PR. The results indicated that MSTRG.24598.1 and its targets shared the same expression trend.

Functional Characterization of PM-Responsive lncRNAs

We next investigated the potential functions of PM-responsive lncRNAs via annotation of their potential target genes (**Supplementary Table S10**). Then GO and pathway enrichment analyses were conducted using the potential target genes. GO enrichment identified 196 terms and 129 terms in PS (**Supplementary Table S11**) and PR plants (**Supplementary Table S12**), respectively. Notably, some GO terms are related to response to stress, such as L-phenylalanine catabolic process, response to stress, and response to jasmonic acid. Moreover, enrichment of the "transcription factor activity, sequence-specific DNA binding" and "transcription factor TFIID complex" categories suggested that PM-responsive lncRNAs may respond to PM by interacting with transcription factors.

We further analyzed the potential target genes of PM-responsive lncRNAs using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The top 20 pathways are listed and shown in **Figure 6C**. The KEGG results showed that the target genes were mainly enriched in different metabolic pathways in PS and PR plants (**Figure 6C**). Among the top 20 enriched pathways, 10 pathways, including circadian rhythm-plant and plant-pathogen interaction, were significantly enriched in PS and PR plants after PM inoculation (**Figure 6C**). These ten pathways may play important regulatory roles in the PM response of *C. pepo*. Interestingly, the most significantly enriched pathway was circadian rhythm-plant in both PS and PR. In addition, we identified some pathways that are only enriched in PS or PR plants (**Figure 6C**). The pathway analysis indicated that M-responsive lncRNAs may respond to PM infection by regulating a wide range of biological processes.

Candidate Regulation Network Involved in PM Response

The above functional analysis suggests that plant-pathogen interaction is one of the most important pathways in PS and PR plants. Additionally, previous studies have shown that MAPK signaling and hormone signal transduction are involved in the plant–pathogen interactions. Therefore, using PM-responsive lncRNAs shared by PS and PR, a candidate regulatory network for PM response was constructed, including the plant–pathogen interaction, MAPK signaling, and the plant hormone signal transduction pathway (**Figure 7**; **Supplementary Table S13**). This candidate regulatory network contains 131 interactions between lncRNAs and potential target genes, including 20 cis-regulated pairs, 110 trans-regulated pairs, and one both-regulated pair. In this regulatory network, only seven lncRNA-target gene pairs have a one-to-one regulatory relationship, which suggests that the regulatory relationship between lncRNA and target genes is complex. A total of 41 and 48 lncRNA-target gene pairs were involved in the MAPK signaling and plant hormone signal transduction pathways, respectively. In addition, 42 lncRNA-target gene pairs (11 lncRNAs and 27 target genes) were involved in the plant-pathogen interaction pathway. In particular, there are potential interactions for MSTRG.24598.1 with nine trans target genes,

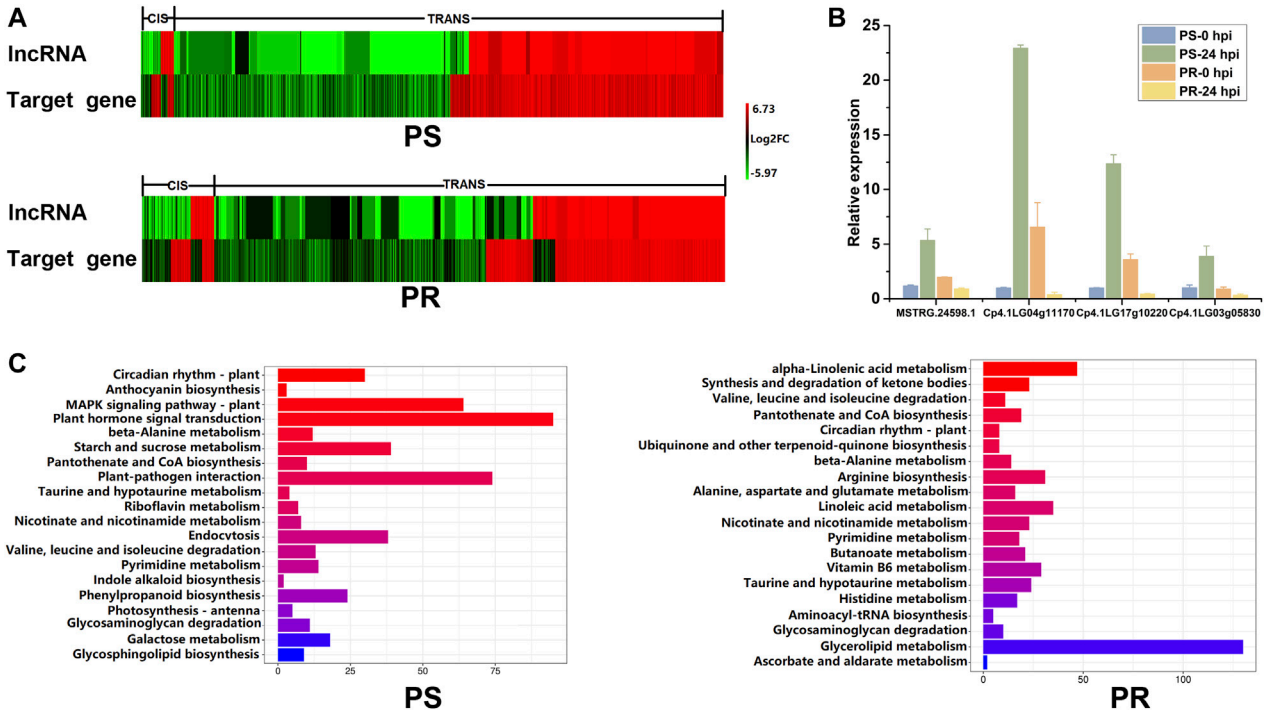


FIGURE 6 | Expression and function analysis of potential target genes of PM-responsive lncRNAs. **(A)** Expression relationship between PM-responsive lncRNAs and their potential target genes. Heat map representing the log_2FC of expression levels of PM-responsive lncRNAs and their putative target genes at 24 h post PM inoculation. **(B)** Confirmation of the expression relationship of MSTRG.24598.1 and its three putative target genes using qRT-PCR. Error bars indicate standard deviations among three biological replicates (n = 3). **(C)** Pathway enrichment analysis for the potential target genes of PM-responsive lncRNAs in PS and PR, respectively. The top 20 pathways were listed and shown. PS represents PM-susceptible *C. pepo* and PR represents PM-resistant *C. pepo*.

including six *calcium-dependent protein kinase* genes and three *WRKY* transcription factors.

Validation of PM-Responsive lncRNAs

To confirm the expression pattern of PM-responsive lncRNAs, we selected eight DELs, including representatives of the four categories of lncRNAs, and conducted validation using qRT-PCR. As shown in **Figure 8**, the qRT-PCR results and the RNA-seq data showed a similar expression pattern for PS and PR plants, despite some differences in their expression levels. For example, the RNA-seq and qRT-PCR results showed that the expression of MSTRG.20701.8 was significantly down-regulated at 24 hpi in PS and PR plants after PM inoculation (**Figure 8A**). In addition, the qRT-PCR results also validated the inconsistent change patterns of MSTRG.24598.1 in PS and PR plants, consistent with the RNA-seq results (**Figure 8B**). These results suggest the reliability of the changing trend of lncRNA expression calculated from the RNA-seq data.

DISCUSSION

Characterization of *C. pepo* lncRNAs

Transcriptome analysis has led to the discovery of non-coding transcripts that were once considered transcriptional noise. Identification of lncRNAs assists in the investigation of the gene regulatory pathways in plants. The existence of lncRNA has been reported in other plants (Liu et al., 2012; Zhang et al., 2014; Tian et al., 2016; Deng et al., 2018), but systematic identification and characterization of lncRNAs in *Cucurbita* still remains to be conducted. The recent availability of *C. pepo* genome sequences has laid a foundation for the genome-wide identification and functional analysis of *C. pepo* lncRNAs (Montero-Pau et al., 2018).

Characteristic analysis showed that the lncRNAs in *C. pepo* shared similar characteristics with the lncRNAs of other species. Previous studies have also revealed that plant lncRNAs are relatively shorter in length and have lower expression as compared to mRNAs. For example, characteristic analysis of rice lncRNAs showed that their length is far below that of mRNAs and they have relatively lower expression levels (Yu et al., 2020). In melon, it was also found that lncRNAs were shorter in length compared to protein-coding transcripts (Gao et al., 2020). Our results showed that the average length of *C. pepo* lncRNAs was only two-fifths of that of *C. pepo* mRNAs, and their expression level was much lower than that of mRNAs in *C. pepo* (**Figure 2C**), which is consistent with previous studies. Moreover, it has been reported that the conservation of plant lncRNAs is extremely low. For example, the majority of lncRNAs in *Capsicum annuum* were found to be novel,

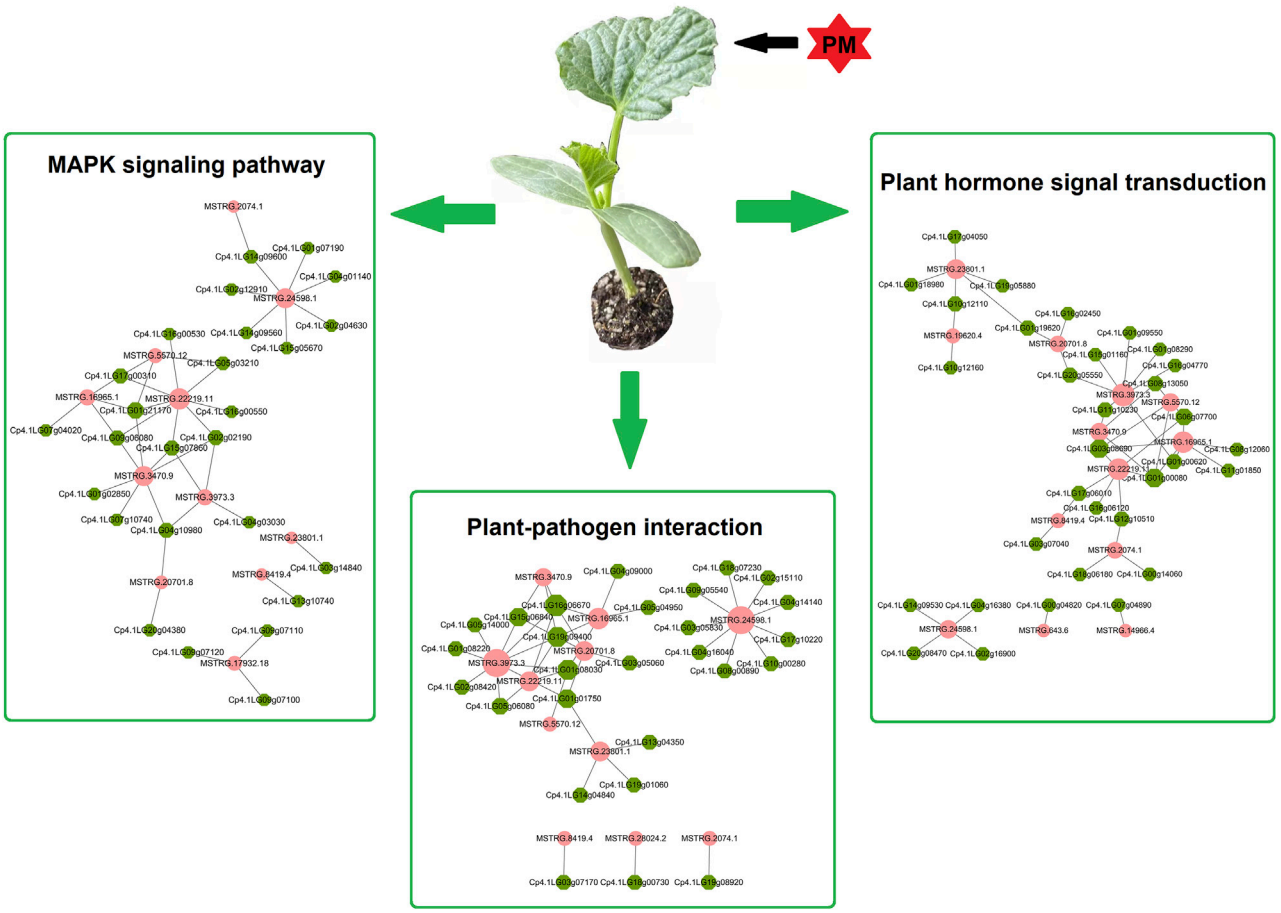


FIGURE 7 | Potential regulation network of lncRNAs and target genes involved in PM response. Potential target genes of PM-responsive lncRNAs are involved in the plant–pathogen interaction, MAPK signalling, and plant hormone signal transduction pathways. The pink circles represent lncRNAs, while the green hexagons represent potential target genes.

because these lncRNAs did not show any similarity with other lncRNAs in public databases (Baruah et al., 2021). BLAST analysis found that the vast majority (99.83%) of lncRNAs identified in this study could be species-specific. Low conservation between species has also been found in other plants (Xin et al., 2011; Liu et al., 2012; Zhang et al., 2014; Li et al., 2014), such as *Arabidopsis*, rice, maize, and wheat. The low conservation of *C. pepo* lncRNAs suggests that they may undergo rapid evolution in a manner similar to that of other plant lncRNAs. Additionally, previous studies in human, animal, and several plant species have reported that lncRNAs can interact with genes and miRNAs (Seo et al., 2017; Cui et al., 2019; Lucero et al., 2021; Tang et al., 2021). In the current study, it was predicted that *C. pepo* lncRNAs would interact with mRNAs and miRNAs. In particular, eight lnRNA-miRNA interactions were verified by degradome sequencing (**Supplementary Table S9**). These results indicate the existence of crosstalk between miRNAs, lncRNAs, and mRNAs in *C. pepo*. Total lncRNAs of *C. pepo* distributed throughout 20 chromosomes (**Figure 2A**) suggest that lncRNA is the important functional component of the *C. pepo* genome. Altogether, our results provide a comprehensive picture of lncRNAs in *C. pepo*.

Potential Function Roles for PM-Responsive lncRNAs in *C. pepo*

So far, many studies have shown that lncRNAs can participate in plant resistance to a variety of diseases (Zhang et al., 2013; Zhang H. et al., 2016; Zhang et al., 2018; Jiang et al., 2020). PM is a severe disease of C. *pepo*, and can cause considerable loss in the yield of C. *pepo*. However, the regulatory mechanism of lncRNA in the response of C. *pepo* to PM remains to be explored. In this study, a large number of significantly differentially expressed lncRNAs were identified after PM infection, suggesting that these PM-responsive lncRNAs may assume an important role in the response of C. *pepo* to PM. Previous studies have proved that lncRNAs can play corresponding functions by interacting with target genes or miRNAs (Kung et al., 2013; Wu et al., 2013; Tang et al., 2021). Thus, we can explore the functions of candidate lncRNAs by analyzing the function of their corresponding target genes or miRNAs.

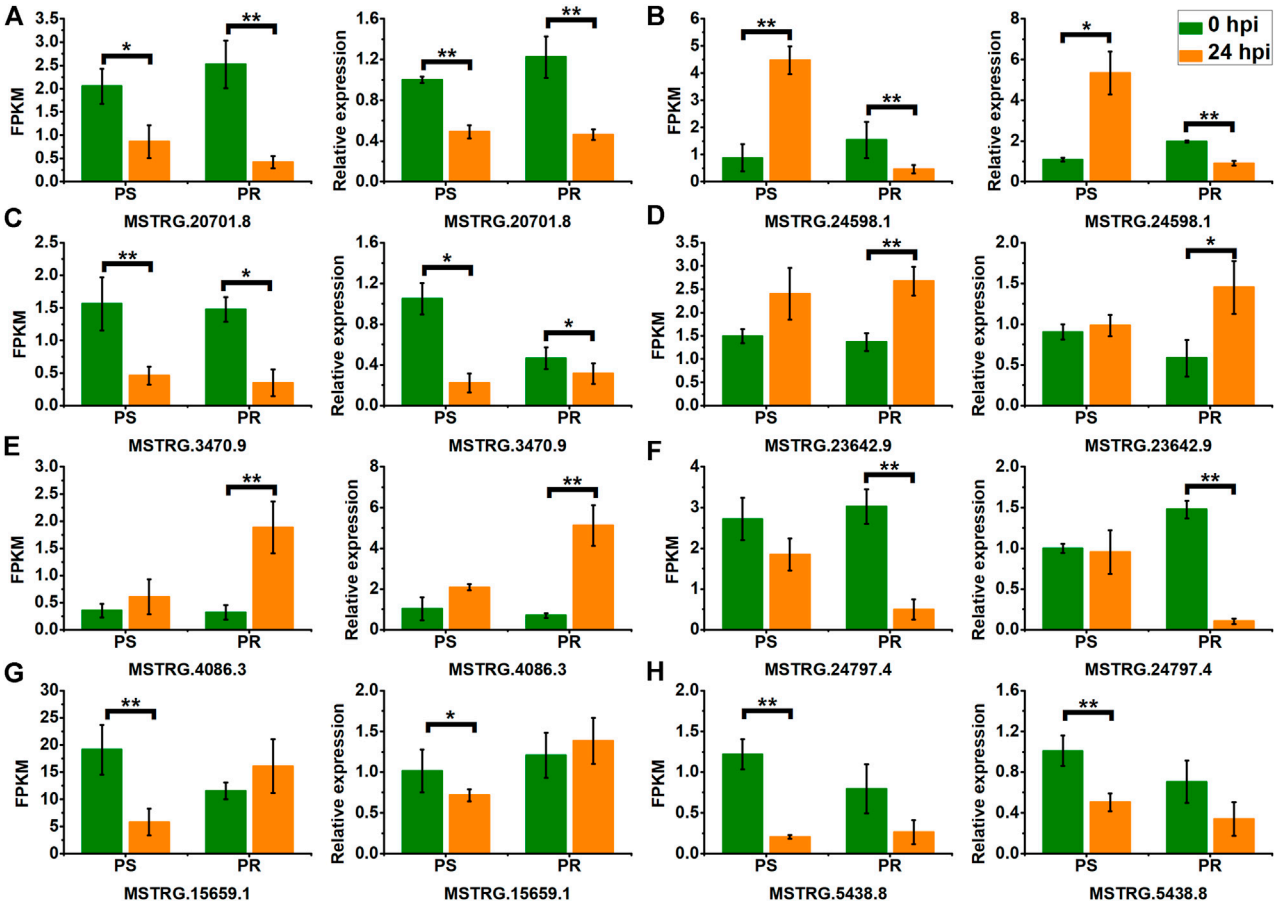


FIGURE 8 | Validation of PM-responsive lncRNAs **(A–H)** using qRT-PCR. PS represents PM-susceptible *C. pepo* and PR represents PM-resistant *C. pepo*. Error bars indicate standard deviations among three biological replicates ($n = 3$). Asterisks indicate significant differences between PS and PR ("*"represents $p < 0.05$; "**" represents $p < 0.01$).

For the functional characterization of potential target genes, plant-pathogen interaction pathway highlights were obtained through pathway analysis (**Figure 6C**). Previous transcriptome analysis in apple and wheat indicated that the plant-pathogen interaction pathway was involved in the response to PM (Tian et al., 2019). Here, our results showed that the plant-pathogen interaction pathway was enriched in PS and PR, indicating that PM-responsive lncRNAs in *C. pepo* may participate in response to PM via regulating the expression of genes related to plant-pathogen interaction (**Figure 7**). In particular, it was predicted that MSTRG.24598.1 interacts with nine trans target genes related to the plant-pathogen interaction pathway. These target genes included six *calcium-dependent protein kinase* genes and three *WRKY* transcription factors, which play an important regulatory role in responding to pathogen invasion (Lecourieux et al., 2006; Li et al., 2019; Pandey and Somssich, 2009; Wei et al., 2018; Chen et al., 2021). Moreover, several target genes of PM-responsive lncRNAs are involved in plant hormone signal transduction. For example, MSTRG.24598.1 was predicted to interact with four target genes (one cis-regulated and three trans-regulated) involved in the plant hormone signal transduction pathway. Previous studies in cucumber suggested that plant hormone signal transduction is involved in the complex regulatory network for PM resistance (Zhang P. et al., 2021; Zheng et al., 2021). These PM-responsive lncRNAs associated with plant-pathogen interaction and plant hormone signal transduction pathways are important candidates for further functional analysis of the regulatory roles of lncRNAs in the *C. pepo* response to PM.

We further investigated the potential interactions between lncRNAs and miRNAs to explore the regulatory roles of PM-responsive lncRNAs in *C. pepo*. A previous study has reported that miR166 was involved in the response to *Verticillium dahlia* in cotton (Zhang T. et al., 2016). In addition, miR166 also acts as an important regulatory factor in the response to *Tomato leaf* curl *New Delhi virus* in tomatoes (Prasad et al., 2022). Here, it was predicted that MSTRG.12356.1 which was significantly up-regulated after PM inoculation was the precursor of cpe-miR166 (**Figure 5A**). This indicated that MSTRG.12356.1 may participate in the response to PM by generating cpe-miR166. Three PM-responsive lncRNAs predicted to be targets of miRNAs were verified by degradome sequencing

(**Supplementary Table S9**). Among these, MSTRG.11047.1 and MSTRG.11047.13 were potential targets of cpe-miR396a. In *Arabidopsis*, it was proven that miR396 is involved in immune responses against fungal pathogens (Soto-Suárez et al., 2017). In addition, miR396 in rice negatively regulates the resistance to rice blast disease (Chandran et al., 2019). Therefore, it indicated that MSTRG.11047.1 and MSTRG.11047.13 might respond to PM by interacting with miRNAs related to disease resistance.

Comparison of PM-Responsive lncRNAs Between Susceptible and Resistant *C. pepo*

Considering that the effects caused by PM disease on *C. pepo* vary substantially between germplasms, a highly susceptible *C. pepo* inbred line (PS) and a highly resistant *C. pepo* inbred line (PR) to PM were selected and compared in our study. The severity of PM disease in PS and PR plants was significantly different after PM inoculation (**Figure 1A**), suggesting there was a significant difference in the PM resistance of PS and PR plants. Moreover, the physiological characteristics differed significantly between PS and PR plants after PM infection (**Figure 1B**). Altogether, the differences between PS and PR plants in phenotype and physiological characteristics suggested their different transcriptional responses to PM.

As we expected, additional PM-responsive lncRNAs were found in PR plants when compared to PS. Among the PM-responsive lncRNAs identified in PS and PR plants, only a small fraction (17 DELs) was shared by PS and PR plants (**Figure 3B**), which is consistent with our previous speculation that there may be significant differences in the transcriptional response between PS and PR. Interestingly, four of the 17 DELs showed inconsistent expression patterns in PS and PR plants (**Figure 3C**), suggesting that they may play an important role in regulating the response of *C. pepo* to PM. All four DELs were significantly up-regulated at 24 hpi in PS but significantly down-regulated at PR. In particular, it was predicted that MSTRG.24598.1 would interact with 241 potential target genes, including hormone-responsive, calcium-binding protein, and transcription factor genes. Transcription factors are important regulators in various plant biological processes, such as stress response (Jiang et al., 2017; Baillo et al., 2019). To date, increasing studies have reported the importance of *WRKY* family transcription factors in plant immunity (Pandey and Somssich, 2009), including response to PM infection (Jiao et al., 2018; Chen et al., 2021). For example, the overexpression of the transcription factor *FvWRKY42* of strawberries can enhance the resistance to PM in *Arabidopsis* (Wei et al., 2018). In the current study, the similar expression patterns of MSTRG.24598.1 and Cp4.1LG17g10220 (*CpWRKY53*) after PM infection (**Figure 6B**) were verified by qRT-PCR results, suggesting that MSTRG.24598.1 might participate the response to PM infection by affecting the expression of *CpWRKY53*. In addition, another trans target gene (Cp4.1LG03g05830) of MSTRG.24598.1 encoding a calcium-binding protein also showed inconsistent expression patterns in PS and PR plants (**Figure 6B**). In plants, calcium-binding protein participated in responses against pathogens, including fungi (Aldon et al., 2018; Li et al., 2019). Here, our results indicated that MSTRG.24598.1 may act as an important regulator in the PM resistance of *C. pepo*.

Several lncRNAs exhibit the same change trend in PS and PR, but their expression levels were different between PS and PR. For example, one PM-responsive lncRNA (MSTRG.20701.8) was down-regulated in PS and PR plants (**Figure 8A**). The fold changes in PS and PR plants were 0.42 and 0.17, respectively. Through target prediction analysis, MSTRG.20701.8 was co-expressed with a *Mildew Resistance Locus O* (*MLO*) gene (Cp4.1LG19g08380, trans-regulated). Consistently, the expression of Cp4.1LG19g08380 was down-regulated in PS and PR plants after PM infection, with a fold change of 0.37 and 0.14, respectively. As a result, the expression level of MSTRG.20701.8 and Cp4.1LG19g08380 in PR plants was significantly lower than that in PS plants after PM infection. It has been demonstrated that induced loss-of-function in specific *MLO* genes was associated with resistance against PM (Zheng et al., 2016; Tapia et al., 2021). For example, tomatoes can obtain high resistance against PM by inducing loss of function of *SlMLO1* (Zheng et al., 2016). These results indicated that *MLO* act as a negative regulator in the response to PM in plants. Thus, MSTRG.20701.8 might increase the PM resistance in PR plants by interacting with the *MLO* gene and decreasing its expression level.

We also identified a large number of PR-specific DELs, which might also play a critical role in resisting PM. Besides PM-responsive lncRNAs shared by PS and PR, a further functional study will be focused on these PR-specific DELs that may contribute to the resistance phenotype of PR. Altogether, our findings will contribute to further understanding of the regulatory mechanisms of lncRNAs in the response of *C. pepo* to PM. These results indicate that there might be critical regulatory functions for lncRNAs in PM resistance by positively and negatively interacting with their corresponding target genes. Subsequently, further molecular studies are required to explore the regulatory mechanisms of these candidate PM-responsive lncRNAs in resisting PM infection.

DATA AVAILABILITY STATEMENT

The raw sequence data have been uploaded to the Genome Sequence Archive (GSA) in the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences (https://bigd.big.ac.cn/), under accession number CRA006731.

AUTHOR CONTRIBUTIONS

JT, CW, and HL planned and designed the research. JT, GZ, FZ, and JM performed experiments. JT analyzed the data. JT, CW, and HL wrote the paper. All authors read and approved the manuscript.

FUNDING

This work was supported by the Project of the National Natural Science Foundation of China (No. 31701940), Beijing Municipal

Excellent Talents Training Program (No. 2018000020060G187), and Innovation and Development Program of Beijing Vegetable Research Center (KYCX202001-02).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.933022/full#supplementary-material

Supplementary Table 1 | Primers used for real-time quantitative PCR.

Supplementary Table 2 | Summary of C. pepo RNA-seq data.

Supplementary Table 3 | Characteristics of all lncRNA identified in this study.

Supplementary Table 4 | PM-responsive lncRNAs in PS.

Supplementary Table 5 | PM-responsive lncRNAs in PR.

Supplementary Table 6 | Cis-elements in the promoters of the PM-responsive lncRNAs.

Supplementary Table 7 | lncRNAs corresponding to miRNA precursors.

Supplementary Table 8 | The lncRNAs predicted as targets of miRNAs.

Supplementary Table 9 | Targets of miRNAs verified by degradome sequencing.

Supplementary Table 10 | The potential target genes of PM-responsive lncRNAs.

Supplementary Table 11 | Enriched GO terms in PS.

Supplementary Table 12 | Enriched GO terms in PR.

Supplementary Table 13 | Candidate regulation network involved in PM response.

REFERENCES

Aldon, D., Mbengue, M., Mazars, C., and Galaud, J.-P. (2018). Calcium Signalling in Plant Biotic Interactions. *Int. J. Mol. Sci.* 19 (3), 665. doi:10.3390/ijms19030665

Baillo, E. H., Kimotho, R. N., Zhang, Z., and Xu, P. (2019). Transcription Factors Associated with Abiotic and Biotic Stress Tolerance and Their Potential for Crops Improvement. *Genes* 10 (10), 771. doi:10.3390/genes10100771

Barickman, T. C., Horgan, T. E., and Wilson, J. C. (2017). Efficacy of Fungicide Applications and Powdery Mildew Resistance in Three Pumpkin Cultivars. *Crop Prot.* 101, 90–94. doi:10.1016/j.cropro.2017.07.025

Baruah, P. M., Krishnatreya, D. B., Bordoloi, K. S., Gill, S. S., and Agarwala, N. (2021). Genome Wide Identification and Characterization of Abiotic Stress Responsive lncRNAs in *Capsicum Annuum*. *Plant Physiol. Biochem.* 162, 221–236. doi:10.1016/j.plaphy.2021.02.031

Berg, J. A., Appiano, M., Santillán Martínez, M., Hermans, F. W., Vriezen, W. H., Visser, R. G., et al. (2015). A Transposable Element Insertion in the Susceptibility Gene *CsaMLO8* Results in Hypocotyl Resistance to Powdery Mildew in Cucumber. *BMC Plant Biol.* 15 (1), 1–17. doi:10.1186/s12870-015-0635-x

Bhatia, G., Upadhyay, S. K., Upadhyay, A., and Singh, K. (2021). Investigation of Long Non-coding RNAs as Regulatory Players of Grapevine Response to Powdery and Downy Mildew Infection. *BMC Plant Biol.* 21 (1), 1–16. doi:10.1186/s12870-021-03059-6

Chandran, V., Wang, H., Gao, F., Cao, X.-L., Chen, Y.-P., Li, G.-B., et al. (2019). miR396-*OsGRF*s Module Balances Growth and Rice Blast Disease-Resistance. *Front. Plant Sci.* 9, 1999. doi:10.3389/fpls.2018.01999

Chen, Q., Yu, G., Wang, X., Meng, X., and Lv, C. (2020). Genetics and Resistance Mechanism of the Cucumber (*Cucumis Sativus* L.) against Powdery Mildew. *J. Plant Growth Regul.* 40, 147–153. doi:10.1007/s00344-020-10075-7

Chen, Y., Jing, X., Wang, S., Wang, J., Zhang, S., and Shi, Q. (2021). Genome-wide Analysis of WRKY Transcription Factor Family in Melon (*Cucumis Melo* L.) and Their Response to Powdery Mildew. *Plant Mol. Biol. Rep.* 39, 686–699. doi:10.1007/s11105-020-01271-6

Cheng, H., Kong, W., Hou, D., Lv, J., and Tao, X. (2013). Isolation, Characterization, and Expression Analysis of *CmMLO2* in Muskmelon. *Mol. Biol. Rep.* 40 (3), 2609–2615. doi:10.1007/s11033-012-2347-8

Cui, J., Luan, Y., Jiang, N., Bao, H., and Meng, J. (2017). Comparative Transcriptome Analysis between Resistant and Susceptible Tomato Allows the Identification of lncRNA16397 Conferring Resistance to *Phytophthora Infestans* by Co-Expressing Glutaredoxin. *Plant J.* 89 (3), 577–589. doi:10.1111/tpj.13408

Cui, J., Jiang, N., Meng, J., Yang, G., Liu, W., Zhou, X., et al. (2019). LncRNA33732-respiratory Burst Oxidase Module Associated with WRKY1 in Tomato-*Phytophthora Infestans* Interactions. *Plant J.* 97 (5), 933–946. doi:10.1111/tpj.14173

Cui, J., Jiang, N., Hou, X., Wu, S., Zhang, Q., Meng, J., et al. (2020). Genome-Wide Identification of lncRNAs and Analysis of ceRNA Networks during Tomato Resistance to *Phytophthora Infestans*. *Phytopathology* 110, 456–464. doi:10.1094/phyto-04-19-0137-r

Dai, X., and Zhao, P. X. (2011). psRNATarget: a Plant Small RNA Target Analysis Server. *Nucleic Acids Res.* 39, W155–W159. doi:10.1093/nar/gkr319

Deng, F., Zhang, X., Wang, W., Yuan, R., and Shen, F. (2018). Identification of *Gossypium Hirsutu*m Long Non-Coding RNAs (lncRNAs) under Salt Stress. *BMC Plant Biol.* 18 (1), 23–14. doi:10.1186/s12870-018-1238-0

Ellinger, D., Naumann, M., Falter, C., Zwikowics, C., Jamrow, T., Manisseri, C., et al. (2013). Elevated Early Callose Deposition Results in Complete Penetration Resistance to Powdery Mildew in *Arabidopsis*. *Plant Physiol.* 161, 1433–1444. doi:10.1104/pp.112.211011

Gao, C., Sun, J., Dong, Y., Wang, C., Xiao, S., Mo, L., et al. (2020). Comparative Transcriptome Analysis Uncovers Regulatory Roles of Long Non-coding RNAs Involved in Resistance to Powdery Mildew in Melon. *BMC Genomics* 21 (1), 1–16. doi:10.1186/s12864-020-6546-8

Guo, W.-L., Chen, B.-H., Chen, X.-J., Guo, Y.-Y., Yang, H.-L., Li, X.-Z., et al. (2018). Transcriptome Profiling of Pumpkin (*Cucurbita Moschata* Duch.) Leaves Infected with Powdery Mildew. *PLoS One* 13 (1), e0190175. doi:10.1371/journal.pone.0190175

Guo, W.-L., Chen, B.-H., Guo, Y.-Y., Yang, H.-L., Mu, J.-Y., Wang, Y.-L., et al. (2019). Improved Powdery Mildew Resistance of Transgenic *Nicotiana Benthamiana* Overexpressing the *Cucurbita Moschata CmSGT1* Gene. *Front. Plant Sci.* 10, 955. doi:10.3389/fpls.2019.00955

Guo, W.-L., Chen, B.-H., Guo, Y.-Y., Chen, X.-J., Li, Q.-F., Yang, H.-L., et al. (2020). Expression of Pumpkin *CmbHLH87* Gene Improves Powdery Mildew Resistance in Tobacco. *Front. Plant Sci.* 11, 163. doi:10.3389/fpls.2020.00163

Hafez, Y. M., El-Nagar, A. S., Elzaawely, A. A., Kamel, S., and Maswada, H. F. (2018). Biological Control of *Podosphaera Xanthii* the Causal Agent of Squash Powdery Mildew Disease by Upregulation of Defense-Related Enzymes. *Egypt J. Biol. Pest Control* 28, 57. doi:10.1186/s41938-018-0058-8

Hou, X., Cui, J., Liu, W., Jiang, N., Zhou, X., Qi, H., et al. (2020). LncRNA39026 Enhances Tomato Resistance to *Phytophthora Infestans* by Decoying miR168a and Inducing PR Gene Expression. *Phytopathology* 110 (4), 873–880. doi:10.1094/PHYTO-12-19-0445-R

Hu, Y., Gao, Y.-R., Yang, L.-S., Wang, W., Wang, Y.-J., and Wen, Y.-Q. (2019). The Cytological Basis of Powdery Mildew Resistance in Wild Chinese Vitis Species. *Plant Physiol. Biochem.* 144, 244–253. doi:10.1016/j.plaphy.2019.09.049

Jia, H., Osak, M., Bogu, G. K., Stanton, L. W., Johnson, R., and Lipovich, L. (2010). Genome-wide Computational Identification and Manual Annotation of Human Long Noncoding RNA Genes. *RNA* 16 (8), 1478–1487. doi:10.1261/rna.1951310

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY Transcription Factors in Plant Responses to Stresses. *J. Integr. Plant Biol.* 59 (2), 86–101. doi:10.1111/jipb.12513

Jiang, N., Cui, J., Shi, Y., Yang, G., Zhou, X., Hou, X., et al. (2019). Tomato lncRNA23468 Functions as a Competing Endogenous RNA to Modulate *NBS-LRR* Genes by Decoying miR482b in the Tomato-*Phytophthora Infestans* Interaction. *Hortic. Res.* 6, 28. doi:10.1038/s41438-018-0096-0

Jiang, N., Cui, J., Hou, X., Yang, G., Xiao, Y., Han, L., et al. (2020). Sl-lncRNA15492 Interacts with Sl-miR482a and Affects *Solanum lycopersicum* Immunity against *Phytophthora Infestans*. *Plant J.* 103, 1561–1574. doi:10.1111/tpj.14847

Jiao, Z., Sun, J., Wang, C., Dong, Y., Xiao, S., Gao, X., et al. (2018). Genome-wide Characterization, Evolutionary Analysis of *WRKY* Genes in Cucurbitaceae Species and Assessment of its Roles in Resisting to Powdery Mildew Disease. *PLoS One* 13 (12), e0199851. doi:10.1371/journal.pone.0199851

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., et al. (2007). RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science* 316, 1484–1488. doi:10.1126/science.1138341

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* 12 (4), 357–360. doi:10.1038/nmeth.3317

Kung, J. T. Y., Colognori, D., and Lee, J. T. (2013). Long Noncoding RNAs: Past, Present, and Future. *Genetics* 193 (3), 651–669. doi:10.1534/genetics.112.146704

Kusch, S., and Panstruga, R. (2017). Mlo-based Resistance: an Apparently Universal"weapon" to Defeat Powdery Mildew Disease. *Mpmi* 30, 179–189. doi:10.1094/MPMI-12-16-0255-CR

Lecourieux, D., Ranjeva, R., and Pugin, A. (2006). Calcium in Plant Defence-Signalling Pathways. *New Phytol.* 171 (2), 249–269. doi:10.1111/j.1469-8137.2006.01777.x

Li, L., Eichten, S. R., Shimizu, R., Petsch, K., Yeh, C.-T., Wu, W., et al. (2014). Genome-wide Discovery and Characterization of Maize Long Non-Coding RNAs. *Genome Biol.* 15 (2), R40. doi:10.1186/gb-2014-15-2-r40

Li, G., Zhou, J., Jia, H., Gao, Z., Fan, M., Luo, Y., et al. (2019). Mutation of a Histidine-Rich Calcium-Binding-Protein Gene in Wheat Confers Resistance to Fusarium Head Blight. *Nat. Genet.* 51 (7), 1106–1112. doi:10.1038/s41588-019-0426-7

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., et al. (2012). Genome-wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in *Arabidopsis*. *Plant Cell* 24, 4333–4345. doi:10.1105/tpc.112.102855

Liu, J., Wang, H., and Chua, N.-H. (2015). Long Noncoding RNA Transcriptome of Plants. *Plant Biotechnol. J.* 13 (3), 319–328. doi:10.1111/pbi.12336

Liu, P. N., Miao, H., Lu, H. W., Cui, J. Y., Tian, G. L., Wehner, T. C., et al. (2017). Molecular Mapping and Candidate Gene Analysis for Resistance to Powdery Mildew in *Cucumis Sativus* Stem. *Genet. Mol. Res.* 16 (3), gmr16039680. doi:10.4238/gmr16039680

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 (4), 402–408. doi:10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Lucero, L., Ferrero, L., Fonouni-Farde, C., and Ariel, F. (2021). Functional Classification of Plant Long Noncoding RNAs: a Transcript Is Known by the Company it Keeps. *New Phytol.* 229 (3), 1251–1260. doi:10.1111/nph.16903

Montero-Pau, J., Blanca, J., Bombarely, A., Ziarsolo, P., Esteras, C., Martí-Gómez, C., et al. (2018). De Novo assembly of the Zucchini Genome Reveals a Whole-genome Duplication Associated with the Origin of the *Cucurbita* Genus. *Plant Biotechnol. J.* 16 (6), 1161–1171. doi:10.1111/pbi.12860

Pandey, S. P., and Somssich, I. E. (2009). The Role of WRKY Transcription Factors in Plant Immunity. *Plant Physiol.* 150 (4), 1648–1655. doi:10.1104/pp.109.138990

Pérez-garcía, A., Romero, D., Fernández-ortuño, D., López-Ruiz, F., De Vicente, A., and Torés, J. A. (2009). The Powdery Mildew Fungus *Podosphaera Fusca* (Synonym *Podosphaera Xanthii*), a Constant Threat to Cucurbits. *Mol. Plant Pathol.* 10, 153–160. doi:10.1111/j.1364-3703.2008.00527.x

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level Expression Analysis of RNA-Seq Experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11 (9), 1650–1667. doi:10.1038/nprot.2016.095

Ponting, C. P., Oliver, P. L., and Reik, W. (2009). Evolution and Functions of Long Noncoding RNAs. *Cell* 136, 629–641. doi:10.1016/j.cell.2009.02.006

Prasad, A., Sharma, N., Chirom, O., and Prasad, M. (2022). The Sly-miR166-*SlyHB* Module Acts as a Susceptibility Factor during ToLCNDV Infection. *Theor. Appl. Genet.* 135 (1), 233–242. doi:10.1007/s00122-021-03962-4

Qie, Y., Sheng, Y., Xu, H., Jin, Y., Ma, F., Li, L., et al. (2019). Identification of a New Powdery Mildew Resistance Gene *pmDHT* at or Closely Linked to the *Pm5* Locus in the Chinese Wheat Landrace Dahongtou. *Plant Dis.* 103, 2645–2651. doi:10.1094/PDIS-02-19-0401-Rea

Qin, T., Zhao, H., Cui, P., Albesher, N., and Xiong, L. (2017). A Nucleus-Localized Long Non-coding RNA Enhances Drought and Salt Stress Tolerance. *Plant Physiol.* 175 (3), 1321–1336. doi:10.1104/pp.17.00574

Rehmsmeier, M., Steffen, P., Höchsmann, M., and Giegerich, R. (2004). Fast and Effective Prediction of microRNA/target Duplexes. *RNA* 10 (10), 1507–1517. doi:10.1261/rna.5248604

Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. (2011). Identification of Novel Transcripts in Annotated Genomes Using RNA-Seq. *Bioinformatics* 27 (17), 2325–2329. doi:10.1093/bioinformatics/btr355

Romero, D., Eugenia Rivera, M., Cazorla, F. M., Codina, J. C., Fernández-Ortuño, D., Torés, J. A., et al. (2008). Comparative Histochemical Analyses of Oxidative Burst and Cell Wall Reinforcement in Compatible and Incompatible Melon Powdery Mildew (*Podosphaera Fusca*) Interactions. *J. Plant Physiol.* 165, 1895–1905. doi:10.1016/j.jplph.2008.04.020

Schouten, H. J., Krauskopf, J., Visser, R. G. F., and Bai, Y. (2014). Identification of Candidate Genes Required for Susceptibility to Powdery or Downy Mildew in Cucumber. *Euphytica* 200 (3), 475–486. doi:10.1007/s10681-014-1216-z

Seo, J. S., Sun, H.-X., Park, B. S., Huang, C.-H., Yeh, S.-D., Jung, C., et al. (2017). *ELF18-INDUCED LONG-NONCODING RNA* Associates with Mediator to Enhance Expression of Innate Immune Response Genes in *Arabidopsis*. *Plant Cell* 29 (5), 1024–1038. doi:10.1105/tpc.16.00886

Soto-Suárez, M., Baldrich, P., Weigel, D., Rubio-Somoza, I., and San Segundo, B. (2017). The Arabidopsis miR396 Mediates Pathogen-Associated Molecular Pattern-Triggered Immune Responses against Fungal Pathogens. *Sci. Rep.* 7 (1), 1–14. doi:10.1038/srep44898

Tang, Y., Qu, Z., Lei, J., He, R., Adelson, D. L., Zhu, Y., et al. (2021). The Long Noncoding RNA FRILAIR Regulates Strawberry Fruit Ripening by Functioning as a Noncanonical Target Mimic. *PLoS Genet.* 17 (3), e1009461. doi:10.1371/journal.pgen.1009461

Tapia, R. R., Barbey, C. R., Chandra, S., Folta, K. M., Whitaker, V. M., and Lee, S. (2021). Evolution of the MLO Gene Families in Octoploid Strawberry (*Fragaria× Ananassa*) and Progenitor Diploid Species Identified Potential Genes for Strawberry Powdery Mildew Resistance. *Hortic. Res.* 8, 153. doi:10.1038/s41438-021-00587-y

Tian, J., Song, Y., Du, Q., Yang, X., Ci, D., Chen, J., et al. (2016). Population Genomic Analysis of Gibberellin-Responsive Long Non-coding RNAs in *Populus*. *J. Exp. Bot.* 67 (8), 2467–2482. doi:10.1093/jxb/erw057

Tian, X., Zhang, L., Feng, S., Zhao, Z., Wang, X., and Gao, H. (2019). Transcriptome Analysis of Apple Leaves in Response to Powdery Mildew (*Podosphaera Leucotricha*) Infection. *Int. J. Mol. Sci.* 20 (9), 2326. doi:10.3390/ijms20092326

Wang, H.-L. V., and Chekanova, J. A. (2017). Long Noncoding RNAs in Plants. *Adv. Exp. Med. Biol.* 1008, 133–154. doi:10.1007/978-981-10-5203-3_5

Wang, H., Chung, P. J., Liu, J., Jang, I.-C., Kean, M. J., Xu, J., et al. (2014). Genome-wide Identification of Long Noncoding Natural Antisense Transcripts and Their Responses to Light in *Arabidopsis*. *Genome Res.* 24 (3), 444–453. doi:10.1101/gr.165555.113

Wang, M., Wu, H.-J., Fang, J., Chu, C., and Wang, X.-J. (2017). A Long Noncoding RNA Involved in Rice Reproductive Development by Negatively Regulating Osa-miR160. *Sci. Bull.* 62 (7), 470–475. doi:10.1016/j.scib.2017.03.013

Wang, G., Wang, X., Zhang, Y., Yang, J., Li, Z., Wu, L., et al. (2021). Dynamic Characteristics and Functional Analysis Provide New Insights into Long Non-coding RNA Responsive to *Verticillium dahliae* Infection in *Gossypium Hirsutum*. *BMC Plant Biol.* 21 (1), 1–13. doi:10.1186/s12870-021-02835-8

Wei, W., Cui, M.-Y., Hu, Y., Gao, K., Xie, Y.-G., Jiang, Y., et al. (2018). Ectopic Expression of *FvWRKY42*, a WRKY Transcription Factor from the Diploid Woodland Strawberry (*Fragaria Vesca*), Enhances Resistance to Powdery Mildew, Improves Osmotic Stress Resistance, and Increases Abscisic Acid Sensitivity in *Arabidopsis*. *Plant Sci.* 275, 60–74. doi:10.1016/j.plantsci.2018.07.010

Wu, H.-J., Wang, Z.-M., Wang, M., and Wang, X.-J. (2013). Widespread Long Noncoding RNAs as Endogenous Target Mimics for microRNAs in Plants. *Plant Physiol.* 161 (4), 1875–1884. doi:10.1104/pp.113.215962

Wu, X., Shi, T., Iqbal, S., Zhang, Y., Liu, L., and Gao, Z. (2019). Genome-wide Discovery and Characterization of Flower Development Related Long Non-coding RNAs in *Prunus Mume*. *BMC Plant Biol.* 19, 64. doi:10.1186/s12870-019-1672-7

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a Web Server for Annotation and Identification of Enriched Pathways and Diseases. *Nucleic Acids Res.* 39 (Suppl. l_2), W316–W322. doi:10.1093/nar/gkr483

Xin, M., Wang, Y., Yao, Y., Song, N., Hu, Z., Qin, D., et al. (2011). Identification and Characterization of Wheat Long Nonprotein Coding RNAs Responsive to Powdery Mildew Infection and Heat Stress by Using Microarray Analysis and SBS Sequencing. *BMC Plant Biol.* 11 (1), 1–13. doi:10.1186/1471-2229-11-61

Xu, Q., Xu, X., Shi, Y., Qi, X., and Chen, X. (2017). Elucidation of the Molecular Responses of a Cucumber Segment Substitution Line Carrying Pm5.1 and its Recurrent Parent Triggered by Powdery Mildew by Comparative Transcriptome Profiling. *BMC Genomics* 18, 21. doi:10.1186/s12864-016-3438-z

Xu, X., Liu, X., Yan, Y., Wang, W., Gebretsadik, K., Qi, X., et al. (2019). Comparative Proteomic Analysis of Cucumber Powdery Mildew Resistance between a Single-Segment Substitution Line and its Recurrent Parent. *Hortic. Res.* 6, 115. doi:10.1038/s41438-019-0198-3

Yu, Y., Zhou, Y. F., Feng, Y. Z., He, H., Lian, J. P., Yang, Y. W., et al. (2020). Transcriptional Landscape of Pathogen-Responsive lncRNAs in Rice Unveils the Role of ALEX1 in Jasmonate Pathway and Disease Resistance. *Plant Biotechnol. J.* 18 (3), 679–690. doi:10.1111/pbi.13234

Zhang, B., Su, T., Li, P., Xin, X., Cao, Y., Wang, W., et al. (2021). Identification of Long Noncoding RNAs Involved in Resistance to Downy Mildew in Chinese Cabbage. *Hortic. Res.* 8, 44. doi:10.1038/s41438-021-00479-1

Zhang, H., Chen, X., Wang, C., Xu, Z., Wang, Y., Liu, X., et al. (2013). Long Non-coding Genes Implicated in Response to Stripe Rust Pathogen Stress in Wheat (*Triticum aestivum* L.). *Mol. Biol. Rep.* 40, 6245–6253. doi:10.1007/s11033-013-2736-7

Zhang, Y.-C., Liao, J.-Y., Li, Z.-Y., Yu, Y., Zhang, J.-P., Li, Q.-F., et al. (2014). Genome-wide Screening and Functional Analysis Identify a Large Number of Long Noncoding RNAs Involved in the Sexual Reproduction of Rice. *Genome Biol.* 15, 512. doi:10.1186/s13059-014-0512-1

Zhang, L., Wang, M., Li, N., Wang, H., Qiu, P., Pei, L., et al. (2018). Long Noncoding RNA S Involve in Resistance to *Verticillium dahliae*, a Fungal Disease in Cotton. *Plant Biotechnol. J.* 16 (6), 1172–1185. doi:10.1111/pbi.12861

Zhang, H., Hu, W., Hao, J., Lv, S., Wang, C., Tong, W., et al. (2016). Genomewide Identification and Functional Prediction of Novel and Fungi-Responsive lincRNAs in *Triticum aestivum*. *BMC Genomics* 17, 238. doi:10.1186/s12864-016-2570-0

Zhang, P., Zhu, Y., and Zhou, S. (2021). Comparative Analysis of Powdery Mildew Resistant and Susceptible Cultivated Cucumber (*Cucumis Sativus* L.) Varieties to Reveal the Metabolic Responses to *Sphaerotheca Fuliginea* Infection. *BMC Plant Biol.* 21 (1), 1–13. doi:10.1186/s12870-020-02797-3

Zhang, S., Liu, J., Xu, B., and Zhou, J. (2021). Differential Responses of *Cucurbita Pepo* to *Podosphaera Xanthii* Reveal the Mechanism of Powdery Mildew Disease Resistance in Pumpkin. *Front. Plant Sci.* 12, 306. doi:10.3389/fpls.2021.633221

Zhang, T., Zhao, Y.-L., Zhao, J.-H., Wang, S., Jin, Y., Chen, Z.-Q., et al. (2016). Cotton Plants Export microRNAs to Inhibit Virulence Gene Expression in a Fungal Pathogen. *Nat. Plants* 2 (10), 1–6. doi:10.1038/nplants.2016.153

Zheng, Z., Appiano, M., Pavan, S., Bracuto, V., Ricciardi, L., Visser, R. G. F., et al. (2016). Genome-wide Study of the Tomato *SlMLO* Gene Family and its Functional Characterization in Response to the Powdery Mildew Fungus *Oidium Neolycopersici*. *Front. Plant Sci.* 7, 380. doi:10.3389/fpls.2016.00380/full

Zheng, L., Zhang, M., Zhuo, Z., Wang, Y., Gao, X., Li, Y., et al. (2021). Transcriptome Profiling Analysis Reveals Distinct Resistance Response of Cucumber Leaves Infected With Powdery Mildew. *Plant Biol.* 23 (2), 327–340. doi:10.1111/plb.13213

Zhou, X., Cui, J., Cui, H., Jiang, N., Hou, X., Liu, S., et al. (2020). Identification of lncRNAs and Their Regulatory Relationships with Target Genes and Corresponding miRNAs in Melon Response to Powdery Mildew Fungi. *Gene* 735, 144403. doi:10.1016/j.gene.2020.144403

Zhu, Q. H., Stephen, S., Taylor, J., Helliwell, C. A., and Wang, M. B. (2014). Long Noncoding RNAs Responsive to *Fusarium oxysporum* Infection in *Arabidopsis thaliana*. *New Phytol.* 201(2), 574–584. doi:10.1111/nph.12537

Zhu, B., Yang, Y., Li, R., Fu, D., Wen, L., Luo, Y., et al. (2015). RNA Sequencing and Functional Analysis Implicate the Regulatory Role of Long Non-coding RNAs in Tomato Fruit Ripening. *J. Exp. Bot.* 66, 4483–4495. doi:10.1093/jxb/erv203

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 11 July 2022
doi: 10.3389/fgene.2022.932832

Integrated Full-Length Transcriptome and MicroRNA Sequencing Approaches Provide Insights Into Salt Tolerance in Mangrove (*Sonneratia apetala* Buch.-Ham.)

Beibei Chen[1], *Zeyi Ding*[1], *Xiang Zhou*[1], *Yue Wang*[2], *Fei Huang*[1], *Jiaxin Sun*[1], *Jinhui Chen*[2]* *and Weidong Han*[1]*

[1]College of Coastal Agricultural Science, Guangdong Ocean University, Zhanjiang, China, [2]Hainan Yazhou Bay Seed Laboratory, Sanya Nanfan Research Institute of Hainan University, Sanya, China

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Hou-Ling Wang,
Beijing Forestry University, China
Yuanyuan Zhao,
Beijing Forestry University, China

***Correspondence:**
Jinhui Chen
jinhuichen@hainanu.edu.cn
Weidong Han
hanwd@gdou.edu.cn

Specialty section:
This article was submitted to RNA,
a section of the journal Frontiers in Genetics

Received: 30 April 2022
Accepted: 08 June 2022
Published: 11 July 2022

Citation:
Chen B, Ding Z, Zhou X, Wang Y, Huang F, Sun J, Chen J and Han W (2022) Integrated Full-Length Transcriptome and MicroRNA Sequencing Approaches Provide Insights Into Salt Tolerance in Mangrove (Sonneratia apetala Buch.-Ham.).
Front. Genet. 13:932832.
doi: 10.3389/fgene.2022.932832

MicroRNAs (miRNAs) are small RNA molecules that serve as key players in plant stress responses. Although stress-regulated miRNAs have been explored in various plants, they are not well studied in mangroves. Herein, we combined PacBio isoform sequencing (Iso-Seq) with BGISEQ short-read RNA-seq to probe the role of miRNAs in the salt stress response of the mangrove plant, *Sonneratia apetala* Buch.-Ham. A total of 1,702,463 circular consensus sequencing reads were generated that produced 295,501 nonredundant full-length transcripts from the leaves of a 1-year-old *S. apetala*. After sequencing nine small RNA libraries constructed from control and 1- and 28-day 300 mM NaCl treatments, we identified 143 miRNAs (114 known and 29 novel) from a total of >261 million short reads. With the criteria of $|\log_2 FC| \geq 1$ and q-value < 0.05, 42 and 70 miRNAs were differentially accumulated after 1- and 28-day salt treatments, respectively. These differential accumulated miRNAs potentially targeted salt-responsive genes encoding transcription factors, ion homeostasis, osmotic protection, and detoxificant-related proteins, reminiscent of their responsibility for salinity adaptation in *S. apetala*. Particularly, 62 miRNAs were *Sonneratia* specific under salt stress, of which 34 were co-expressed with their 131 predicted targets, thus producing 140 miRNA–target interactions. Of these, 82 miRNA-target pairs exhibited negative correlations. Eighteen miRNA targets were categorized for the 'environmental information processing' during KEGG analysis and were related to plant hormone signal transduction (ko04075), MAPK signaling pathway–plant (ko04016), and ABC transporters (ko02010). These results underscored miRNAs as possible contributors to mangrove success in severe environments and offer insights into an miRNA-mediated regulatory mechanism of salt response in *S. apetala*.

Keywords: *Sonneratia* specific, third-generation sequencing, microRNA, target gene, salt stress

INTRODUCTION

Soil salinity, one of the major environmental threats, negatively affects plant growth and development. Worldwide, approximately 20% of irrigated soils and 50% of arable land are exposed to salt stress (Syed et al., 2021). High salinity also leads to secondary stresses such as oxidative stress and nutritional imbalance that result in cellular damage, growth inhibition, and yield decrease. Generally, genetically engineering plants to increase salt tolerance would be a promising approach. Thus, deciphering the molecular mechanism and physiological processes that are involved in response to salt stress would certainly facilitate the identification of candidates for genetic engineering of stress-resistant plants, and could help in coping with stress challenges.

Mangroves are a group of halophytic woody plants, growing in tropical and sub-tropical estuaries and intertidal zones, which play a vital role in water purification, coastal protection and the maintenance of biodiversity (Wang and Gu, 2021). As extremophiles, mangroves have evolutionarily adapted to tolerate high temperatures, flooding, anoxia, and high salt conditions in typically nutrient-poor environments. With continued exposure to highly salinized soil and periodic seawater erosion, salt-tolerance mechanism of mangroves especially differs from that of terrestrial freshwater plants. So, there has been a growing interest in exploring the mechanism of salt stress adaptations of mangroves. Generally, mangroves exhibit several physiological strategies for handling salt, which include salt excretion (e.g., *Sonneratia*, *Aegicaras*, *Avicennia*), salt balance modulation (e.g., *Xylocarpus*, *Bruguiera*, *Rhizophora*), and hyperexclusion (e.g., *Heritiera*) (Dassanayake et al., 2009). Many studies have been devoted to uncovering genetic regulators that modulate mangroves' adaptations to salt environments. Previous research (Srikanth et al., 2016) indicated that fructose-1,6-bisphosphate (FBP) osmotin and aldolase were related to salt resistance in *B. gymnorrhiza* roots. A *vacuolar* Na^+/H^+ *antiporter* served as a crucial actor in cellular salinity adjustments of *R. apiculata* (Saddhe and Kumar, 2019). The mRNA expressions of catalase, Cu–Zn SOD, and ferritin genes have been studied in response to saline stress and their role in oxidative stress response has been confirmed in *A. marina* (Dahibhate et al., 2020).

MicroRNAs (miRNAs) comprise a class of endogenous small non-coding RNA molecules that are approximately 21 nt in length, and are recognized as important regulators of the transcriptional and post-transcriptional expression of genes, mediated via mRNA degradation, transcription inhibition or the repression of mRNA translation (Tyagi et al., 2019). Besides their roles in modulating diverse essential physiological, biochemical and molecular processes, many miRNAs were discovered to respond to various abiotic stimuli in plants, including drought (Li et al., 2021), salinity (Hu et al., 2018; He et al., 2021; Salgado et al., 2021), cold (Hu et al., 2018) and oxidative stress (Li et al., 2020), as well biotic stresses (Salamon et al., 2021). To address salinity-induced stress, multiple gene expression mechanisms have evolved in plants, which relate to a wide range of biological processes, including transcription, signal transduction, energy metabolism, membrane trafficking, protein biosynthesis, and photosynthesis (Pan et al., 2018). Understanding of miRNA-guided biological regulations in plants against salinity could provide the necessary basis for unraveling the complex molecular and genetic mechanism underlying salt-stress resistance. A growing body of evidences has demonstrated that miRNA-mediated gene regulation is indispensable for plants response to salt stress. For instance, miR172c regulates root plastic development in soybean response to salt by targeting *NNC1* (Gupta et al., 2021); the miR398-*CSD* module protects cell membrane structure via detoxifying superoxide radicals against salt stress in tomato (Gao et al., 2022), miR414/miR408/miR164e could modulate gene recombination replication, thereby repairing to resist saline environment (Yang et al., 2020). Despite their critical roles of in-plant tress resistance, miRNAs have been studied in only a few woody plant species, such as poplar (Si et al., 2014), paulownia (Zhao et al., 2018), and willow trees (Zhou J. et al., 2012). In particular, little research has been performed on mangrove miRNAs. Currently, the miRNA repertoires have only been reported for few mangrove species, including *Avicennia marina* (Forsk.) Vierh. (Khraiwesh et al., 2013), *Bruguiera gymnorhiza* (L.) Lam. (Wen et al., 2016; Dasgupta et al., 2017), *Kandelia candel* (Linn.) Druce (Wen et al., 2016), *Rhizophora apiculata* Bl. (Singh et al., 2020) and *Sonneratia alba* Sm. (Wang et al., 2021a). Limited attempt has been made to dissect the molecular basis of salt stress adaptation that is guided by miRNAs (Khraiwesh et al., 2013; Wen et al., 2016).

To detect stress-related miRNAs, it is necessary to compare expression patterns of miRNAs in plants under normal and stress-treated growth conditions. With the implementation of next-generation sequencing, it became easier and cost-efficient to carry out genome-wide mining and identification of stress-responsive miRNAs. However, high throughput sequencing produces short reads and thereby acquires incomplete assembly of sequences. On the other hand, the third-generation long-read sequencing platforms offer better completeness in full-length cDNA sequencing and yield more accurate isoform-level transcripts. Indeed, third-generation PacBio sequencing technology (Iso-Seq) has been successfully harnessed in multiple species and has revealed greater accuracy in exploring genome information (Pearman et al., 2020; Sharma et al., 2021). The combination of second- and third-generation platforms has become an effective approach for critical gene identification and function mining, especially for those with no sequenced genomes available (Mei et al., 2021; Meng et al., 2021; Wang et al., 2021b). In this study, we jointly employed short-read RNA-seq and PacBio Iso-Seq to produce a high-confidence full-length transcriptome dataset of 1-year-old *S. apetala* individuals and further used them to identify salt-responsive miRNAs through constructing small RNA libraries from the leaves of *S. apetala* seedlings treated with 300 mM NaCl for 0 days (control), 1, and 28 days. To study the genetic mechanism of the miRNA-mediated gene modulation in salt stress adaptations, the potential targets of salt-related miRNAs were predicted and further analyzed. Most importantly, *Sonneratia*-specific miRNAs were screened, with the co-expressed miRNA-target regulatory

interactions were investigated using transcriptome data, as well as Gene Ontology (GO) and KEGG analyses. In particular, we focused on negatively correlated *Sonneratia*-specific miRNA-target pairs and the underlying mechanism in salinity response. This study provided systematically characterized salt-related miRNAs of *S. apetala* and the expanded features of putative targets reveal the miRNA inferred regulatory networks responding to excessive saline stress. This would be helpful for further investigation of molecule functions during the salt response in mangroves.

MATERIALS AND METHODS

Plant Materials and Salt Treatments

The seeds of *S. apetala* were collected from adult plants growing along a mangrove coastal belt in Techeng Island (21°09′~21°10′ N, 110°25′~110°27′ E), Guangdong, China and were sown on artificial soil in seedbeds. After approximately 70 days, seedlings with 10–16 cm in height were transplanted into polythene bags and watered every 2 days. For the salt stress treatment, 1-year-old uniformly developed seedlings were selected and salinity stress was induced by adding the salt to the Hoagland's solution. Plants grown on sandy soil were watered with 300 mmol/L NaCl, and the controls with only water. A time-course was used for deep sequencing (0, 1, and 28 days) and named LCK, LT1, and LT2, respectively. Three biological replicates were used for each time point and leaf samples were collected from the nine plants (LCK_a, LCK_b, LCK_c, LT1_a, LT1_b, LT1_c, LT2_a, LT2_b, and LT2_c) for RNA extraction. The collected tissues were immediately frozen in liquid nitrogen and then kept at -80°C for RNA extraction until further processing.

PacBio Iso-Seq Library Preparation and Sequencing

Total RNA was extracted from each sample using an RNeasy Plant mini kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The quality and integrity of RNAs were measured by Nanodrop 2000 (Thermo Scientific) and Agilent 2100 Bioanalyzer (Agilent Technologies, CA, United States). Only RNA samples with 1.8 < OD260/280 < 2.2 and RIN ≥7.0 (RNA Integrity Number) were used for follow-up experiments. The qualified RNA samples were used for further PacBio and BGISEQ library construction, respectively.

For PacBio sequencing library construction, we combined equal amounts of total RNA from the nine replicates, as well as total RNA extracted from root tissues of *S. apetala* after salt treatment for 0, 7 days (200 and 400 mmol/L NaCl), and 14 days (300 mmol/L NaCl). The mixed RNA sample was reverse-transcribed into cDNA using the SMARTer™ PCR cDNA Synthesis Kit. PCR amplification was performed to amplify the cDNAs and the fragments were then subjected to damage repair, end repair, and connect SMRT dumbbell-shaped adapters for a full-length transcriptome library construction. The concentration of SMRTbell library was assessed using a Qubit 3.0 fluorometer with a Qubit™ 1X dsDNA HS Assay kit (Invitrogen, Carlsbad, United States) and the quantified criteria for the library quality was a concentration of >10 ng/μl. Then the qualified full cDNA library was sequenced using the binding kit 2.1 from PacBio Sequel platform at the Beijing Genomics Institute (BGI), China. The raw Iso-Seq data were processed to obtain subread sequences *via* SMRTlink v6.0 software and short-read sequences were finally used for calibration of the consensus sequence to obtain a high-quality sequence.

Transcriptome Sequencing and Differentially Expressed Unigene Identification

After the library was constructed following the methods described by Li et al. (2019), transcriptome sequencing was performed using the BGISEQ-500 sequencing platform with paired-end sequencing (length of 150 bp). Raw reads were filtered by the removal of low-quality sequence fragments, reads containing N blurs, and adapter sequences. The Trinity software (v2.0.6) with default parameters and a minimum contig length of 150 bp was used for assembly generation. Transcript levels were determined from the short-read data through RSEM (Li and Dewey, 2011), with the resulting full-length transcripts used as a reference sequence (ref). The transcript isoform level and gene-level counts were converted into fragments per kilobase of transcript per million mapped reads (FPKM) values. Analysis of differential expression of genes between different treatments was performed using the DESeq R package (1.10.1) with DEGs selected using log2 FC ≥ 1 or ≤ −1, $p < 0.01$ and Q < 0.05.

Small RNA Isolation and BGISEQ Sequencing

Small RNA libraries for the nine individuals of *S. apetala* were constructed according to a previously described procedure (Fehlmann et al., 2016). Briefly, small RNA fragments with a length of 10–30 nt were purified by using a 15% urea-PAGE gel, followed by ligating with adenylated 3′ and 5′ adapters. Then, Adapter-ligated RNAs were reverse-transcribed and the cDNA product was amplified by PCR. The concentration and purity of the PCR yield were quantified by Qubit 2.0 (Invitrogen, Cat No. Q33216). Finally, approximately 20 μg product for each sample was sequenced on a BGISEQ-500 platform, and 50 bp single-end reads were generated.

Identification of Conserved and Novel MicroRNAs

Clean sequencing reads were obtained after discarding the contaminations and low quantity reads. AASRA (Tang et al., 2021) and CMsearch (Nawrocki and Eddy, 2013) were used to map clean reads to the full-length transcript reference sequences and other sRNA databases. To ensure that each unique small RNA was mapped to only one category, we followed a priority rule: miRBase > piRNABank > snoRNA (human/plant) > Rfam > other sRNA. Novel miRNAs were predicted by miRA (Evers et al., 2015) software by evaluating the characteristics of hairpin

structures of miRNA precursors with the assistance of RNAfold web server (http://www.tbi.univie.ac.at/RNA/) (Lorenz et al., 2011).

Abundance Analyses and Target Prediction of MicroRNAs

The abundance of miRNA transcripts was normalized using the transcripts per million (TPM). Differentially accumulating miRNAs for each pair of stress treatments (LT1 vs. LCK and LT2 vs. LCK) were identified *via* DESeq2 (Love et al., 2014) and based on the normalized TPM counts. The settings "|log2Fold change. normalized| ≥ 1" with "$p < 0.01$" and "Q < 0.05" were used as thresholds for determining significant changes in transcript abundances. To increase the accuracy of the results, two types of software, psRobot (Wu et al., 2012) and TargetFinder (Fahlgren and Carrington, 2010), were used for target gene prediction and the intersection of the target gene was chosen as the final results. Functional enrichment analysis for miRNA target genes was performed through screening the GO (www.geneontology.org/) and KEGG Pathway databases (www.genome.jp/kegg/) with "Q < 0.1" as a threshold to determine significant enrichments.

Quantification of MicroRNA Transcript Levels *via* Real-Time Quantitative Polymerase Chain Reaction

cDNAs were synthesized by reverse transcription of total RNA from nine *S. apetala* samples (LCK_a, LCK_b, LCK_c, LT1_a, LT1_b, LT1_c, LT2_a, LT2_b, and LT2_c). RT-qPCRs were carried out on a CFX Connect™ Real-Time PCR Detection System using SYBR green-based real-time PCR with miRNA-specific forward primer and universal reverse primer (**Supplementary Table S1**). The PCR program included an initial denaturation at 94°C for 3 min, and 40 cycles of 20 s at 94°C, and 60°C for 40 s. The specificity of the amplified fragments was checked using the generated melting curve and the $2^{-\triangle\triangle Ct}$ method was used to calculate the abundance of each miRNA against U6 gene (Khraiwesh et al., 2013). All RT-qPCR amplifications were performed in triplicate (technical repetitions) to ensure reproducibility and reliability.

RESULTS

Sequencing and Data Analysis

In this study, the full-length transcriptome of *S. apetala* was sequenced using PacBio isoform sequencing (Iso-Seq) of a pooled RNA sample. A total of 1,702,463 long reads were generated that produced 295,501 unique transcript isoforms with a total length of 418,900,351 base pairs (bp). This formed the *S. apetala* reference sequence database for the identification of known and novel miRNAs in *S. apetala*, as well as for the prediction of corresponding miRNA-target genes.

Nine libraries were constructed for high-throughput small RNA (sRNA) sequencing from NaCl-free (LCK) and NaCl-treated (1 d and 28 days) *S. apetala* leaves. Each library yielded more than 29 million raw sRNA reads (**Table 1**). After trimming adapters contaminants, low-quality tags, RNAs shorter than 18 nt, poly (A) sequences, and other artifacts, 27,207,049 (LCK_a), 27,039,294 (LCK_b), 26,410,458 (LCK_c), 26,298,084 (LT1_a), 26,776,814 (LT1_b), 26,685,055 (LT1_c), 26407348 (LT2_a), 25589700 (LT2_b) and 26356554 (LT2_c) clean reads were acquired for further analysis (**Table 1**). Filtered reads were aligned and mapped to the *S. apetala* full-length transcriptome sequence and the mapping rate for each library varied from 80.83% to 90.95% (**Table 2**). In addition, Pearson's correlation coefficient was calculated based on the abundance of sRNAs, which accounted for a minimum value of 0.889 between samples (**Supplementary Figure S1**). These results indicated the good quality of sequencing data that could be used for the follow-up analyses.

The sRNA size distribution in each library is summarized in **Figure 1**. The most abundant sRNAs ranged from 20 nt to 24 nt in length, which was the typical size range for Dicer-derived products (Wang et al., 2019). For all the nine libraries, 21 nt sRNAs represented the most frequent length, which coincided with the profiles described in other plant species including Populus (Cui et al., 2019), Trifoliate orange (Hu et al., 2022), and grapevine (Chen et al., 2019). The second most abundant class was of 24 nt sRNAs, making up an average of approximately 18.3% of all the sRNAs in the nine libraries (**Figure 1**). The sRNAs were annotated into several different categories (**Table 2**). Of these, an average of 2,653,651 (9.87%), 2,778,522 (10.45%), and 3,176,025 (12.16%) unique sRNAs were annotated as miRNAs in LCK, LT1 and LT2 libraries, respectively. All the libraries showed similar distribution patterns for other RNA families (**Table 2**), including rRNA (~2.78%), snoRNA (~1.77%), snRNA (~1.75%), tRNA (~1.74%) and Rfam other sncRNA (~1.85%) (**Table 2**). Notably, a predominant proportion of sRNA sequences (> 78.54%) were unannotated sRNAs, raising the possibility of the existence of novel miRNAs in *S. apetala* genome.

Identification of Conserved MicroRNAs in *Sonneratia apetala*

To identify the miRNAs conserved in *S. apetala*, we aligned the unique sRNA candidates with all the plant miRNAs in the miRBase 22.1 database. Finally, a total of 114 known mature miRNAs, representing 28 miRNA families were identified in *S. apetala*, which were generated from 97 pre-miRNAs (**Supplementary Data S1**). Of these, 111 conserved miRNAs were found in the control, whereas 106 and 108 miRNAs were found after salt treatment for 1 and 28 days, respectively (**Figure 2A**; **Supplementary Data S1**). The length of known miRNAs varied from 18 nt to 24 nt, with 21 nt miRNAs being the most abundant (**Figure 2**). The next abundant class was 20-nt-long miRNAs, which was different from the total sRNA population (**Figures 1**, **2B**), suggesting the existence of many other types of sRNAs within the libraries. The diversity of conserved *S. apetala* miRNA families could be determined by the number of their members. For example, miR160 and

TABLE 1 | Summary of sequencing data for each sample.

Sample name	Raw tag count	Low-quality tag count	Invalid adapter tag count	PolyA tag count	Short valid length tag	Clean tag count	Q20 of clean tag (%)	Percentage (%)
LCK_a	32000000	253912	438472	193	4100374	27207049	98.9	85.02
LCK_b	31578947	277052	411545	184	3850872	27039294	99	85.62
LCK_c	31168831	258730	436685	374	4062584	26410458	99.1	84.73
LT1_a	29721273	345190	588245	896	2488858	26298084	98.9	88.48
LT1_b	30155684	339962	599595	913	2438400	26776814	98.9	88.8
LT1_c	29906976	336268	574248	944	2310461	26685055	98.8	89.23
LT2_a	31802816	300631	790887	1598	4302352	26407348	98.8	83.03
LT2_b	30868831	320640	621507	844	4336140	25589700	99	82.9
LT2_c	31578947	328989	650973	746	4241685	26356554	99	83.46

miR166 families comprised 6 and 9 members, respectively. Additionally, most of the conserved miRNA families contained more than one member; the two largest families were miR172 and miR396 with 11 miRNA members, followed by miR319 with nine members. However, some conserved miRNA families, such as miR4995, miR530, miR5368, miR5532, miR6300, miR8051, miR827, miR828, and miR8620, had just one member. The number of members for each family is summarized in **Figure 2C**.

Identification of Novel MicroRNAs in *Sonneratia apetala*

Among the remaining unannotated sRNAs, 29 were predicted as putative novel miRNAs in *S. apetala* (**Figure 3A**; **Supplementary Data S2**). The length of the novel mature miRNAs varied from 19 nt to 24 nt, which was consistent with the size of miRNA fragments produced by AGO-guided cleavage (Yang et al., 2020). Of the newly predicted miRNAs, 23 nt and 24 nt accounted for the two major class sizes (**Figure 3B**). Not all of the novel miRNAs were detected in all libraries. Four novel miRNAs (Sap-nmiR4, Sap-nmiR9, Sap-nmiR13, and Sap-nmiR15) were uniquely detected in LCK library, whereas three novel miRNAs, namely, Sap-nmiR11, Sap-nmiR12 and Sap-nmiR14, were found in the LCK and LT1 libraries. Moreover, most novel miRNAs (22) were co-expressed in all three different groups (LCK, LT1, and LT2), accounting for 75.9% of the total. Notably, the abundance of these new miRNAs was relatively low (**Supplementary Data S2**), which was a generic feature of species-specific miRNAs (Feng et al., 2015). The precursor of the 29 potential novel miRNAs ranged in length from 65-nt to 284-nt long, which was concordant with the general length of pre-miRNAs (Lertampaiporn et al., 2013). Furthermore, secondary stem-loop structure of the new miRNAs precursors was also assessed, and their free energies of the thermodynamic ensemble varied from −26.02 kcal/mol to −180.32 kcal/mol with an average value of −80.36 kcal/mol (**Supplementary Figure S2**; **Supplementary File S2**). miRNA star sequences (miRNA*) were detected for each novel miRNA (**Supplementary Data S2**), confirming their authenticity as novel miRNAs.

Accumulation Profiles of Conserved and Novel MicroRNAs

High-throughput sequencing detects the type and abundance of miRNAs (Yang et al., 2020). In this study, miRNA exhibited variable abundances in the nine libraries, with the number of reads ranging from zero to hundreds of thousands, and was exploited as the indicator for assessment of a miRNA's accumulation level. Sap-miR166a-3p represented the highest abundance with an average of 74,614, 169,800 and 271,226 reads in LCK, LT1 and LT2 libraries, independently (**Supplementary Datas S1, S2**). Some miRNAs, such as Sap-miR167d_1 and Sap-miR166d-5p, also exhibited extraordinarily high abundances in the three groups, while other miRNAs, including Sap-miR535a, Sap-nmiR16, and Sap-nmiR28, were moderately accumulated, with total reads varying from 10,080 to 97,452 (**Figure 4A**; **Supplementary Data S1**). Nevertheless, several miRNAs (Sap-miR164b, Sap-miR396b, Sap-nmiR13, Sap-nmiR26, etc) lad extremely low abundances in all the libraries. Additionally, some miRNA counts diverged sharply across control and treatments. For instance, Sap-miR396a-3p_1 had 6,852 reads in the LCK library and 3,520 reads in the LT1 library, while there were only five reads in the LT2 library. Interestingly, the expression of Sap-nmiR20 and Sap-nmiR24 was similar in all the conditions (control and salinity treatments), with a relatively low accumulation level of 455 and 489 reads per million (TPM). Due to the distinction in their abundances among different libraries, we considered that the candidate miRNAs would have diverse functions in response to salt stress in *S. apetala*.

Thereafter, we selected the 10 most abundant miRNAs (5 conserved and 5 novel) in the three groups and analyzed their patterns of accumulation across the control and two salt treatments: these miRNAs clustered into six groups (**Figure 4B**). The first cluster contained two miRNAs, Sap-miR166a-3p and Sap-miR168a-3p, which were quickly induced at 1 day of stress and repressed at 28 days; the second group comprised one miRNA, Sap-nmiR22, which was up-regulated at 1 day and then accumulated at a relatively stable level, whereas Sap-miR396-3p in the third group exhibited an opposite pattern; the fourth class, which included Sap-nmiR16 and Sap-nmiR17, was rapidly repressed at 1 day, then up-regulated at 28 days.

TABLE 2 | Annotations of sRNAs perfectly matching *S. apetala* mRNA transcriptome[a].

Class	0 days			1 days			28 days		
	LCK_a (%)	LCK_b (%)	LCK_c (%)	LT1_a (%)	LT1_b (%)	LT1_c (%)	LT2_a (%)	LT2_b (%)	LT2_c (%)
Clean reads	27207049 (100)	27039294 (100)	26410458 (100)	26298084 (100)	26776814 (100)	26685055 (100)	26407348 (100)	25589700 (100)	26356554 (100)
Total match	23979516 (88.14)	23680675 (87.58)	23064967 (87.33)	21489673 (81.72)	21729758 (81.15)	21570544 (80.83)	23518404 (89.06)	23273943 (90.95)	23753950 (90.13)
Mature (miRNA)	2749347 (10.11)	2602606 (9.63)	2608999 (9.88)	2724558 (10.36)	2811601 (10.50)	2799408 (10.49)	3163093 (11.98)	3125895 (12.22)	3239088 (12.29)
rRNA	834264 (3.07)	833305 (3.08)	857541 (3.25)	672078 (2.56)	691599 (2.58)	682402 (2.56)	669819 (2.54)	670609 (2.62)	727861 (2.76)
snoRNA	536234 (1.97)	533179 (1.97)	522190 (1.98)	470373 (1.79)	476742 (1.78)	469302 (1.76)	408215 (1.55)	398682 (1.56)	418031 (1.59)
snRNA	530600 (1.95)	527374 (1.95)	518043 (1.96)	461389 (1.75)	468199 (1.75)	469981 (1.76)	400266 (1.52)	389494 (1.52)	408576 (1.55)
tRNA	533604 (1.96)	530291 (1.96)	505978 (1.92)	457816 (1.74)	465205 (1.74)	465672 (1.75)	401504 (1.52)	392426 (1.53)	412792 (1.57)
Rfam other sncRNA	559549 (2.06)	557155 (2.06)	552662 (2.09)	480907 (1.83)	490443 (1.83)	482088 (1.81)	432887 (1.64)	432974 (1.69)	446797 (1.70)
Unannotated	21463451 (78.88)	21455384 (79.35)	20845045 (78.92)	21030963 (79.97)	21373025 (79.82)	21316202 (79.87)	20931564 (79.25)	20179620 (78.86)	20703409 (78.54)

[a]*The number represented the abundance of the reads generated directly from deep sequencing; 0 day, 1 day, and 28 days represent leaf treated with 300 mM NaCl for 0, 1, and 28 days, respectively.*

Finally, Sap-miR169d-5p, Sap-nmiR6, and Sap-nmiR28 in the fifth group were repressed at 1 and 28 days, while Sap-miR396b-3p, in the sixth, showed a specific expression profile being up-regulated under the salt treatment. These findings facilitated our comprehension of the modulation of miRNAs in response to salt stress. We examined the transcript levels of these miRNAs with the help of RT-qPCR in leaf tissues of *S. apetala* that were treated with salt for 0, 1, and 28 days. It revealed a concordance with the expression values from high-throughput sequencing ($r = 0.83$, $p =$ 4.3E-04; **Figures 4C,D**).

Differential Expression Analysis and Target Gene Prediction of MicroRNAs Responsible to Salt Stress

To investigate the effect of salinity on *S. apetala* miRNAs, we conducted a differential analysis of accumulations between the libraries treated (LT1 or LT2) and non-treated (LCK) with salt. Expression profiles of miRNAs in two comparisons (LT1 vs. LCK and LT2 vs. LCK) were compared; miRNAs with $\log_2$-fold changes beyond 1.0 or −1.0, and $Q < 0.05$ were considered to be differentially accumulated. In total, 73 miRNAs (56 known and 17 novel) were significantly differentially accumulated in response to salt (**Supplementary Data S3**). Of these, three novel miRNAs, Sap-nmiR9, Sap-nmiR12, and Sap-nmiR14 were markedly differentially accumulated with an absolute value of $\log_2$-ratio (LT/LCK) > 6. Increases in salt treatment time dramatically increased the number of differentially accumulated miRNAs, from 42 (LT1 vs. LCK) to 70 (LT2 vs. LCK) in *S. apetala* leaves. However, only 39 salt stress-regulated miRNAs appeared in both two time-points as compared with the control group. Particularly, 3 and 31 miRNAs were especially significantly altered in samples treated with salt for 1 and 28 days, respectively. Additionally, the number of down-regulated miRNAs varied through the time-course compared to that of up-regulated ones (**Supplementary Data S3**). In LT1 vs. LCK comparison, 18 miRNAs (14 known and 4 novel) were up-regulated, whereas 24 miRNAs (17 known and 7 novel) were downregulated by salt stress. In LT2 vs. LCK comparison, 36 miRNAs (30 known and 6 novel ones) were induced, while 34 miRNAs (23 known and 11 novel ones) were repressed. Among all the differentially accumulated miRNAs, 38 miRNAs (34 known and 4 novel) responded to stress of salinity in a similar way at different time-points during salt treatment.

Furthermore, these salt-regulated miRNAs could be divided into six categories on the basis of their accumulation patterns (**Figures 5A–F**). The miRNAs shown in **Figures 5A,B** did not exhibit significant changes in the accumulations at 1 day, but at 28 days, their modulation varied. Sap-miR169a-5p, Sap-miR169b-5p and Sap-miR160b demonstrated similar patterns of dynamic alterations during salt treatment, as their accumulations changed only after 1 day of salt treatment (**Figure 5C**). These miRNAs were up-regulated in LT1 vs. LCK and LT2 vs. LCK (**Figure 5D**). In contrast, a repression occurred both at 1 and 28 days in the other class (**Figure 5E**). In the last group, Sap-miR172a_2 showed induction at 1 day, but it

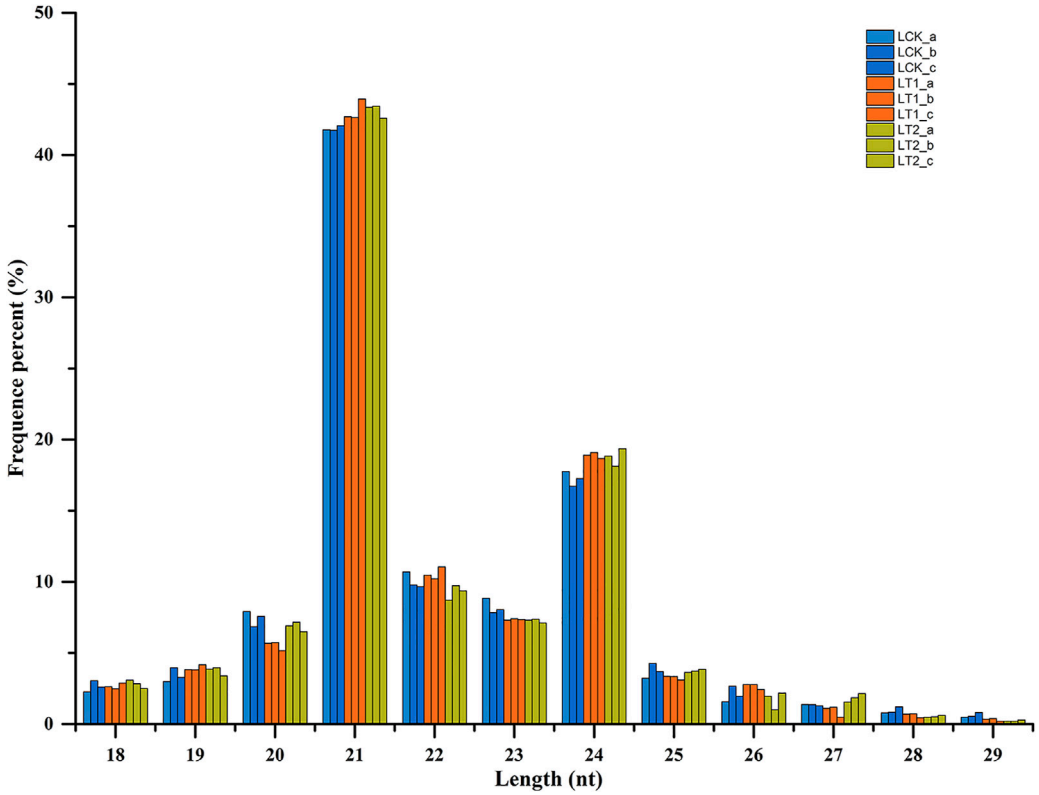


FIGURE 1 | Size distribution of small RNA sequences in different libraries. nt, nucleotides.

was repressed at 28 days, which is unlike the pattern described above (**Figure 5F**). The various alterations of abundances for miRNAs implied that they played various roles under salt stress.

Predicting the targets of miRNAs would be essential for better understanding the functions of these salt-responsive miRNAs. A total of 300 genes were predicted as targets for 67 salt-responsive miRNAs (56 known and 11 novel), corresponding to 409 miRNA-target regulatory interactions for conserved and 62 pairs for novel miRNAs (**Supplementary Datas S4, S5**). We were unable to predict the targets of 6 miRNAs, possibly due to the insufficient *S. apetala* mRNA sequences. Among these expected targets, 25 (8.3%) targets have not been functionally annotated (**Supplementary Data S5**). The remaining annotated target genes participated in a broad spectrum of plant growth and development activities. Many of the predicted target genes were homologous to those encoding essential stress-related transcription factors (TFs), including MYB-domain transcription factor (MYB), APETALA2-like (AP2), homeodomain-leucine zipper transcription factor (HD-ZIP), nuclear transcription factor Y (NFYA), auxin response factor (ARF) family and WRKY transcription factor 22(WRKY22). Moreover, some target genes encoding enzymes or functional proteins that might be involved in plant metabolism, such as photosystem II PsbM protein, UDP-glucuronate decarboxylase (UXS1), protein phosphatase 2C (PP2C), and thioredoxin 1, were also identified. By annotation of targets, most transcripts (171 out of 300) were directly or indirectly implicated in plant salt stress responses, which had the largest proportion of signaling transduction (137, 80.1%), followed by morphological adaption (42, 24.6%), protein turnover (20, 11.7%), basic metabolism adaption (15, 8.8%), ion homeostasis (13, 7.6%), detoxification-related (5, 2.9%), osmotic protection (2, 1.2%) and nutrient modulation (1, 0.6%) (**Supplementary Data S5**). In general, these results implied that miRNAs might be involved in diverse biological processes under salt stress in *S. apetala*.

Co-Expression and Regulation Analysis of *Sonneratia*-Specific MicroRNAs

We mined the alterations in miRNA abundances in other species that were subjected to the same period of salt treatment. A total of 10 miRNAs exhibit a difference when compared to their levels in the leaves of *S. apetala* and *Gossypium hirsutum* Linn. treated with salt for 1 day (Peng et al., 2014) (**Supplementary Table S2**). Out of these, six miRNAs showed inverse trends of accumulation in *S. apetala* and *G. hirsutum* under salt stress conditions. For instance, Sap-miR160a-5p and Sap-miR160b were down-regulated in *S. apetala*, but up-regulated in *G. hirsutum*, whereas Sap-miR166d-5p and Sap-miR396a-3p_2 were up-regulated in *S. apetala* but down-regulated in *G. hirsutum*. Additionally, Sap-miR168a-3p and Sap-miR169v were induced by salt in *S. apetala*, but were not significantly changed in response to salt treatment in *G. hirsutum*. Furthermore, 35 known miRNAs were considered to be *Sonneratia*-specific

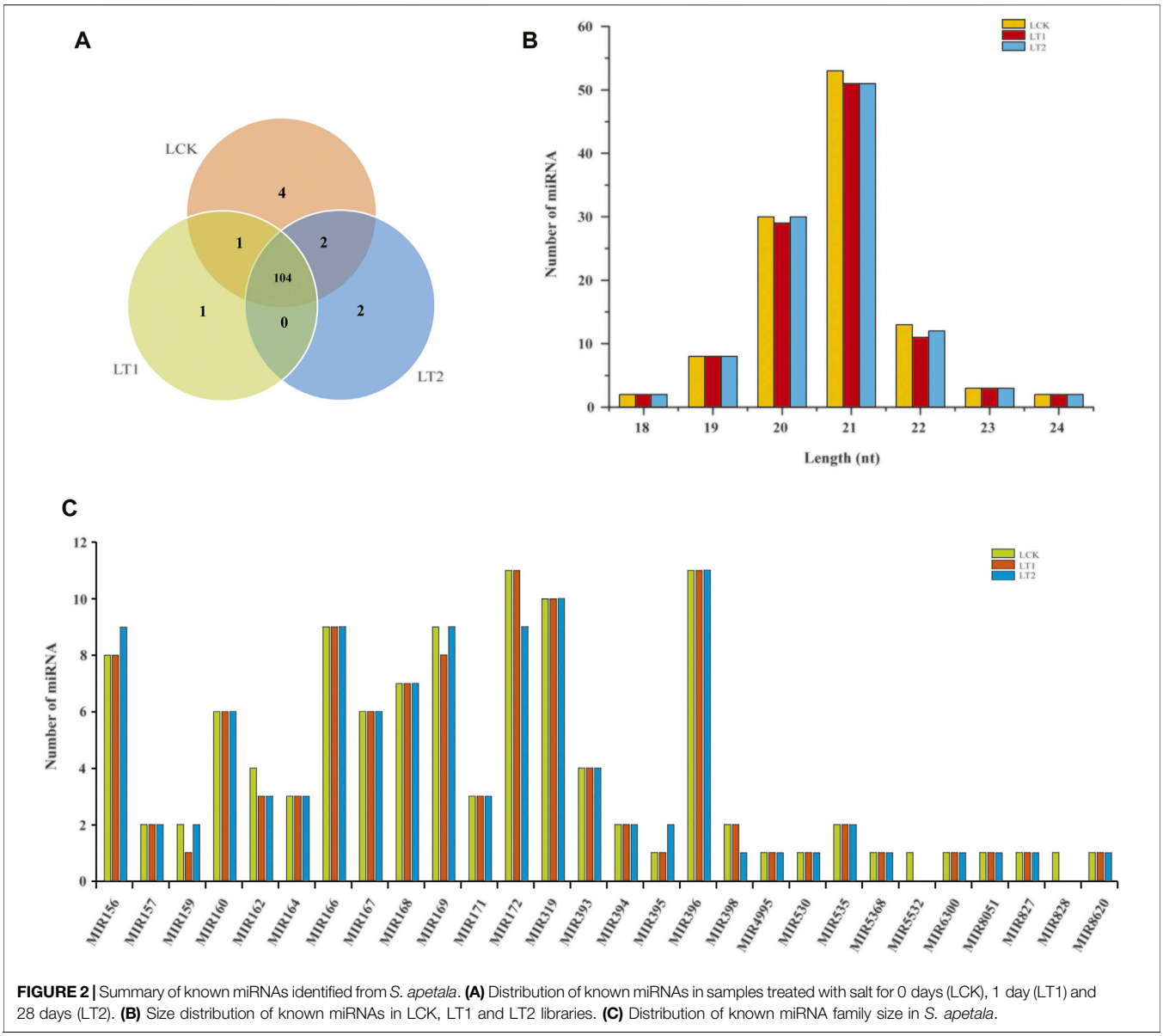


FIGURE 2 | Summary of known miRNAs identified from *S. apetala*. **(A)** Distribution of known miRNAs in samples treated with salt for 0 days (LCK), 1 day (LT1) and 28 days (LT2). **(B)** Size distribution of known miRNAs in LCK, LT1 and LT2 libraries. **(C)** Distribution of known miRNA family size in *S. apetala*.

in genome organization. As showed in **Supplementary Data S6**, Sap-miR166a-3p targeted eight HD-ZIP proteins in *S. apetala* (**Supplementary Data S6**), but Homeobox-leucine zipper family protein (ATHB-15) was identified to be the target in *Populus euphratica* Oliv. (Li et al., 2013). Particularly, these 45 known miRNAs, of which 10 and 35 were specific in expression and genome organization, respectively, as well as the 17 differentially accumulated novel miRNAs were considered to be *Sonneratia*-specific.

Subsequently, RNA-seq comparison of three groups (LCK, LT1, and LT2) for corresponding targets of the 62 *Sonneratia*-specific miRNAs were performed (**Supplementary Data S7–S9**), and differentially expressed target genes that exhibited significant expression correlations with their respective miRNAs ($|r| > 0.75$, $p < 0.01$, $Q < 0.05$) were selected for further exploration. To discover co-accumulated miRNA-target interactions for *Sonneratia*-specific miRNAs, we also focused on their alteration patterns upon salt stress. Particularly, miRNA-target pairs, of which miRNA and targes showed positive correlations and responded to stress in a congruous way were considered to be positively related, whereas targets exhibiting negative correlations and opposite tendency with their referred miRNAs were defined to be negative. Finally, 140 miRNA-target interactions were identified to be co-expressed after salt treatment, representing 34 *Sonneratia*-specific miRNAs and 131 target genes (**Supplementary Datas S8–S10**). To visualize the two-way interaction between miRNA and targets, an internal gene-gene network was constructed according to these miRNA-target pairs (**Figure 6**). Overall, 78 and 91 miRNA-target pairs showed negative or positive interactions independently at 1 and 28 days of salt stress whereas 29 exhibited regulatory relationships at both the time points. It was evident that one miRNA could be co-expressed with

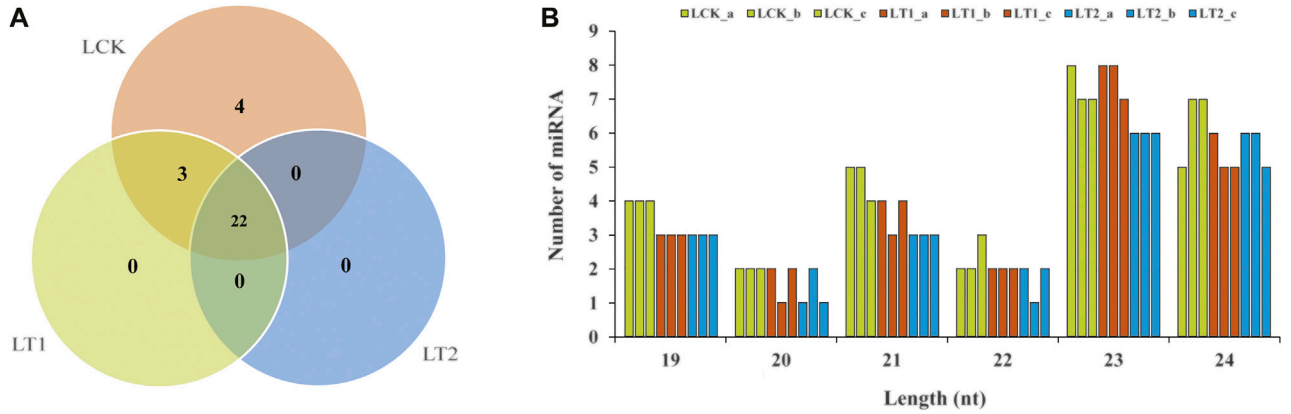


FIGURE 3 | Novel miRNAs identified in *S. apetala* by high-throughput sequencing. **(A)** Distribution of novel miRNAs in control (LCK), 1 day (LT1) and 28 days (LT2) salinity treatment. **(B)** The length distribution of *S. apetala* novel miRNAs. a, b and c represent the three biological repetitions for each treatment.

1-8 target genes, among which the down-regulated Sap-miR396a-5p and Sap-nmiR10 could regulate five and seven mRNAs in LT1 vs. LCK and LT2 vs. LCK comparison, respectively. Simultaneously, Sap-miR169a-5p and Sap-miR169h could positively co-regulate isoform_285675 (NFYA, nuclear transcription factor Y) under early-(1 day) salt stress, while Sap-miR172a_3 and Sap-miR5368 exhibited negative correlations with their common target gene, isoform_105961 (AP2), after a salt-treatment for 28 days. To further probe the possible role of *Sonneratia*-specific miRNAs in response to salt stress, GO-based enrichment analysis was carried out on all the identified target genes in the regulatory interactions. The target genes were significantly mapped to 75 biological processes, 29 cellular components and 16 molecular functions ($p < 0.05$, $Q < 0.1$; **Figure 7A**; **Supplementary Data S11**). With respect to biological processes, the main terms were "biological regulation" (GO: 0065007, 17), "response to stimulus" (GO: 0050896, 16), and "cellular response to stimulus" (GO: 0051716, 15). The predominant terms implicated in cellular components were "cell part" (GO: 0044464, 64), "cell" (GO: 0005623, 64), "intracellular" (GO: 0005622, 61), and "intracellular part" (GO: 0044424, 61). For their molecular functions, the "binding" (GO: 0005488, 67) and "nucleic acid binding" (GO: 0003676, 49) were the most abundant subcategories. These classifications suggested that these miRNA targets, associated with salt tolerance, were primarily related to binding, stimulus, cellular metabolic, and other metabolic processes. When pathway analysis was performed by KEGG annotation, we were able to enrich 63 *Sonneratia*-specific miRNA targets to 34 pathways related to, for instance, metabolism of starch and sucrose (ko00500), spliceosome (ko03040), and fatty acid degradation (ko00071). (**Figure 7B**; **Supplementary Data S12**). The majority of representative pathways for the targets under the "metabolism" category were involved in the phenylpropanoid biosynthesis (ko00940; 6, 17.6%). Two signal transduction and one membrane transport pathways were classified into environmental information processing and thus were related to salt stress resistance. The only pathway categorized under the "organismal systems" (environmental adaptation) category was plant-pathogen interaction (ko04626; 4, 11.8%). The annotations of these target genes may provide new clues into the salt tolerance response in *S. apetala*.

After investigating negatively related miRNA-target pairs, 82 pairwise interactions were selected for further analysis (**Supplementary Data S10**); of these, 17 exhibited negative regulatory relationship in both of the comparisons (LT1 vs. LCK and LT2 vs. LCK; **Supplementary Table S3**). *Basic endochitinase B* (*CHIB*) was inversely regulated (correlation coefficient, r = −0.86) by its respective miRNA. The level of *CHIB* transcripts was up-regulated in leaves treated with salt for 1 and 28 days, while a continuous decrease in Sap-nmiR12 levels was observed after both the salt treatments. Annotation of target genes indicated that a few transcripts were likely to participate in plant salt stress response (**Figure 8**). For instance, a Sap-nmiR6 target was GST, which was involved in "ROS scavenging protection" (Tian et al., 2019); the co-target gene of Sap-miR396a-5p and Sap-nmiR10, *ATP-binding cassette subfamily B member 1* (*ABCB1*), was related to "ion homeostasis" (Tian et al., 2019); Sap-miR160a-5p, Sap-miR160b, Sap-miR164e, Sap-miR393a, and Sap-nmiR16 might play roles in morphological adaption by targeting 2 *ARFs*, UBR4 (Feng et al., 2015), *TIR1* (McDiarmid et al., 2020), and *CESA* (Zaaba et al., 2021), respectively. It was noteworthy that the majority of the negative *Sonneratia*-specific miRNAs (77.4%) regulated the genes with function in stress signal perception and transduction, of which, 15 were predicted to target potential TF genes such as *MYC2*, *HD-ZIP*, and *NFYA*. These *Sonneratia*-specific miRNAs and their targets related to salt stress response allowed us to gain new insights into understanding stress tolerance mechanism in *S. apetala*.

DISCUSSION

Small RNA Sequencing and Identification of MicroRNAs in *Sonneratia apetala*

Salt stress adversely influenced plant growth and development. To cope with salinity, plants adapt themselves through modulating numerous salt-responsive genes at the

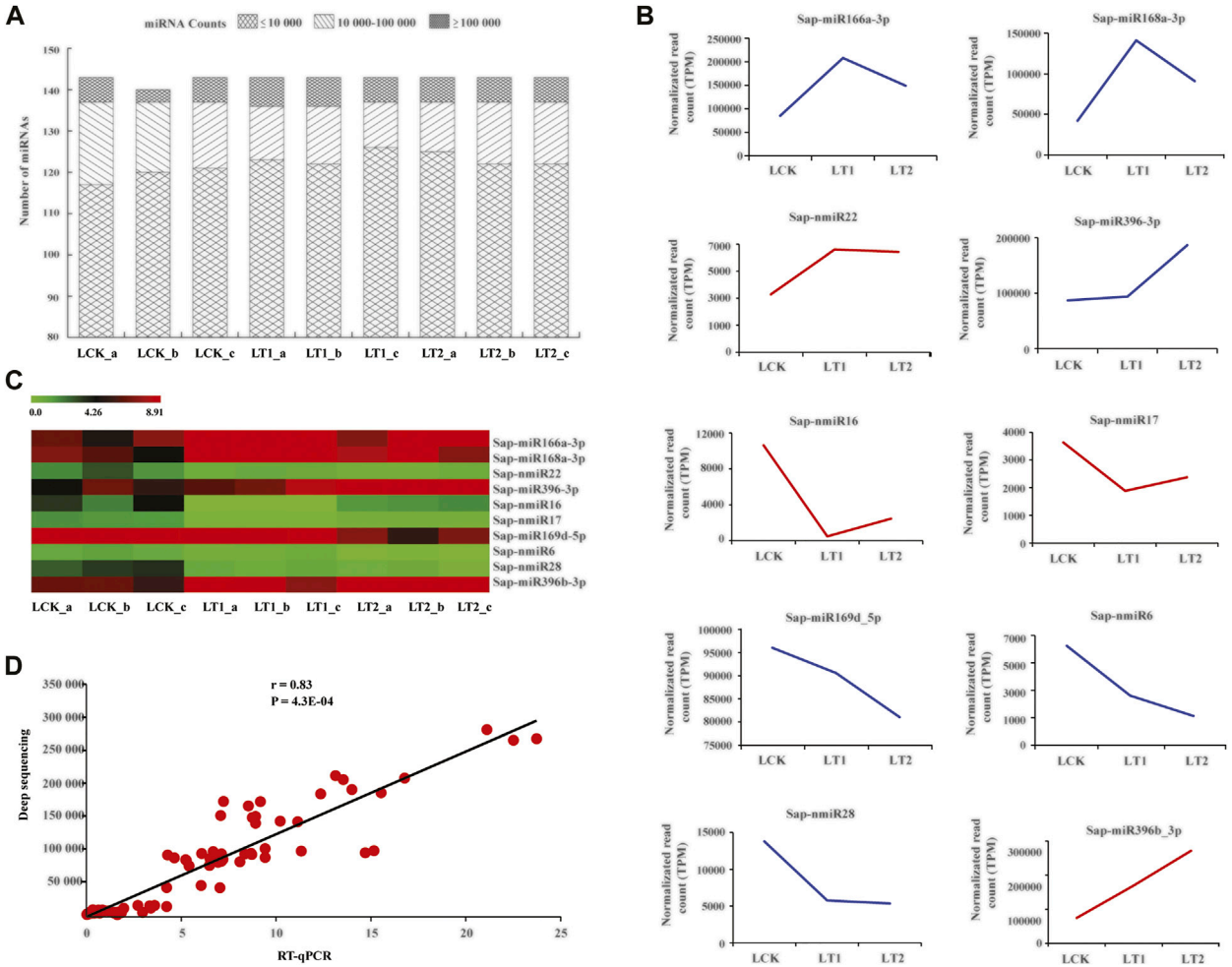


FIGURE 4 | Abundance characteristics of miRNAs obtained from *S. apetala* samples subjected to deep sequencing. **(A)** A global view of miRNA accumulation levels. **(B)** Representative graphs of accumulation profile for 10 most abundant miRNAs. These miRNAs were divided into six categories on the basis of their accumulation patterns, which were distinguished by the colors of lines. **(C)** Relative transcript abundance of the 10 most highly accumulated miRNAs with the help of real-time quantitative polymerase chain reaction (RT-qPCR) in control (LCK), 1 day (LT1) and 28 days (LT2) salinity treatment of *S. apetala*. a, b and c represent the three biological repetitions for each treatment. **(D)** Pearson correlation coefficient (*r*) analysis between accumulation levels by RT-qPCR and deep sequencing for the 10 most abundant miRNAs.

transcriptional and post-transcriptional levels. miRNAs are well known to function as ubiquitous regulators of gene expression and play crucial roles in a plant's response to stress (Wani et al., 2020; Pagano et al., 2021). Increasing number of studies have demonstrated that miRNA-guided gene regulation plays a significant role during the salt stress response of various plant species (Xu et al., 2021). *S. apetala* is a pioneer mangrove species and is widely used for afforestation, which is important for keeping the ecological balance of a coastal zone. Studies of miRNAs related to stress adaptation of mangroves have only been reported in *A. marina* (Khraiwesh et al., 2013), *B. gymnorrhiza* and *K. candel* (Wen et al., 2016) for mangrove plants, and no study has been conducted on *S. apetala* for identification of salt-regulated miRNAs and their target genes.

In this study, more than 29 million sRNA reads were produced by deep sequencing for each library of *S. apetala* in the control and salinity treatments, imparting adequate sequencing depth for further analysis (**Table 1**). The sRNAs of *S. apetala* contained different types of RNAs, including miRNA, rRNA, snoRNA, snRNA, and tRNA (**Table 2**). Majority of the sRNA reads were unannotated, which possibly attributed to the scarcity of genome information on this species, and implied that numerous sRNAs still remain to be identified in *S. apetala*. The annotation performed here was in line with previous reports of other plants such as cotton (Zhan et al., 2021) and *Ipomoea batatas* L. (Yang et al., 2020). Analysis of the length distribution of these sRNA indicated that most were 21-nt and 24-nt (**Figure 1**), which has also been observed in other mangrove species, including *Bruguiera gymnorrhiza* and *Kandelia* (Wen et al., 2016). The maximum abundance of 21 nt sRNAs in all the libraries suggested that sRNAs with this length pattern might play more prevailing roles in the salt stress response of *S. apetala*. The separation in the

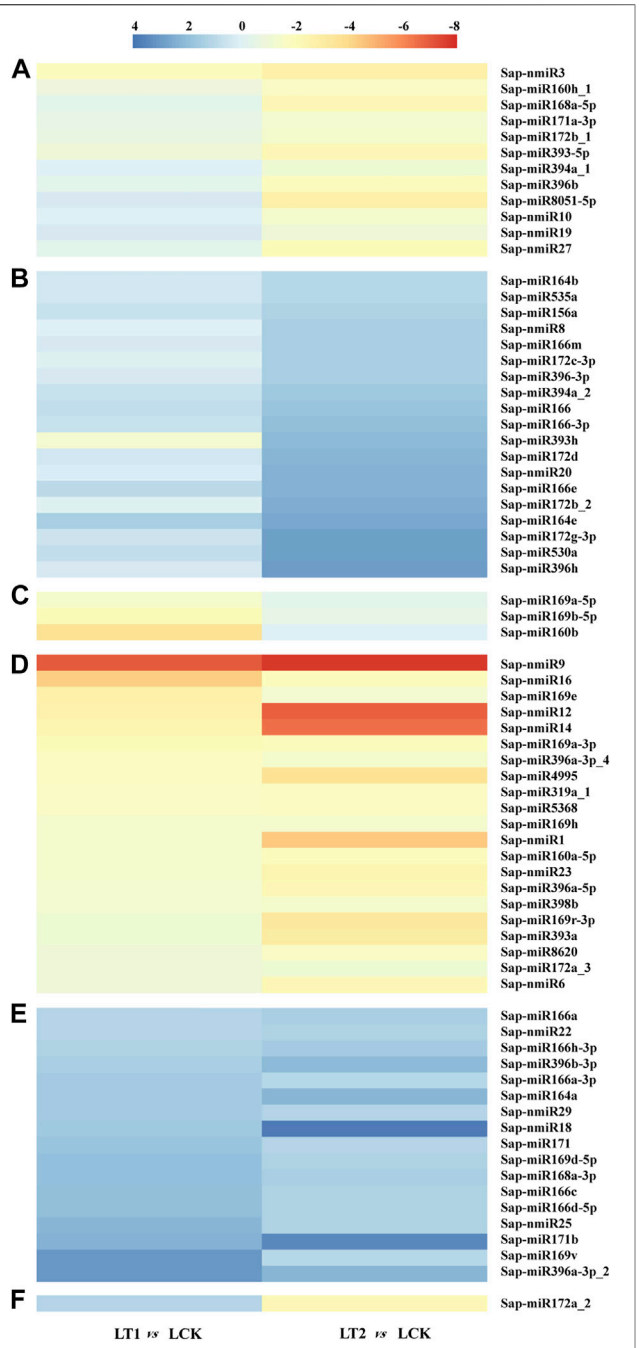


FIGURE 5 | Differentially accumulated *S. apetala* miRNAs after 0 days (LCK), 1 day (LT1) and 28 days (LT2) of salt treatment. Relative abundance level was calculated with $Log_2(TPM_{LT1}/TPM_{LCK})$ and $Log_2(TPM_{LT2}/TPM_{LCK})$. These differentially accumulated miRNAs were divided into six clusters, represented by **(A–F)**.

length of sRNAs may have resulted from the enzymes that process it. For instance, sRNAs processed by DCL1, DCL2, and DCL3 in plants are 21, 22 and 24 nt-long, respectively (Necira et al., 2021).

According to miRNA prediction criteria, a total of 143 miRNAs, including 114 known and 29 novel miRNAs, were detected from the nine small RNA libraries (**Supplementary Datas S1, S2**). Most of the known miRNAs were 21 nt in length, and the novel miRNAs were mainly 23 and 24 nt (**Figures 2B**, **3B**). The importance of length distribution for miRNA lies in the fact that it allows easy interaction with the AGO proteins: In general, the 21 or 22 nt miRNAs tend to function as silencing complexes with AGO1 proteins and cleave the target mRNAs, whereas 24 nt miRNAs guide DNA methylation by binding to AGO4 (Yang et al., 2020). Furthermore, it was inferred that the read copies of miRNAs ranging from zero to hundreds of thousands mirrored their diverse expression levels in *S. apetala*. For example, Sap-miR166a-3p, Sap-miR167d_1 and Sap-miR166d-5p had an extraordinarily high number of reads and averaged to 288,409, indicating that these miRNAs possibly accumulated at a higher level, whereas Sap-nmiR2 and Sap-nmiR19 showed low abundance of less than 100. Overall, the diversity of *S. apetala* miRNAs was also reflected in their highly variable abundances. Notably, known conserved miRNAs usually had relatively higher levels of accumulation when compared with those newly identified (**Supplementary Datas S1, S2**). This was a common characteristic of plant miRNAs and supported by previous studies (Feng et al., 2015). Additionally, a striking divergence also existed in the accumulation patterns of miRNAs at different time points of salt treatment as both, congruously and oppositely regulated miRNAs were detected (**Figure 4**; **Supplementary Datas S1, S2**). Our findings provide a strong basis for further in-depth miRNA studies in *S. apetala*.

Characterization of Salt Stress-Responsive MicroRNAs and Their Targets in *Sonneratia apetala*

Identification of differentially expressed miRNAs and their subsequent functional dissection would aid in the understanding of the response mechanism in *S. apetala* under salt stress. It is well known that salt adaptation is a long-term and dynamic process that involves many morphological, physiological, molecular, and cellular processes (Feng et al., 2020). To probe the dynamic changes of miRNAs during the saline stress, we compared their accumulation from samples treated with salt for 1 and 28 days to that of control libraries (0 days). totally, 73 miRNAs (56 known miRNAs and 17 novel) were found to be regulated by salinity ($p < 0.01$, $Q < 0.05$; **Figure 5**; **Supplementary Data S3**). Earlier studies have demonstrated that miR156, miR159, miR160, miR168, miR169, miR171, miR172, miR393, and miR396 were the major salt stress-regulated miRNAs in plants (Alzahrani et al., 2019). All of these were also identified in our study, indicating the existence of common salt stress-related miRNAs. To the best of our knowledge, some conserved miRNAs were found to be regulated by salt stress for the first time in this study. This includes Sap-miR394a_1 and Sap-miR8051-5p, which maintained a relatively constant level at 1 d of stress but were significantly inhibited at 28 days. Additionally, a few novel salt-responsive miRNAs were identified in *S. apetala* (**Supplementary Data S3**). Several miRNAs, such as Sap-miR395p-3p, Sap-nmiR9

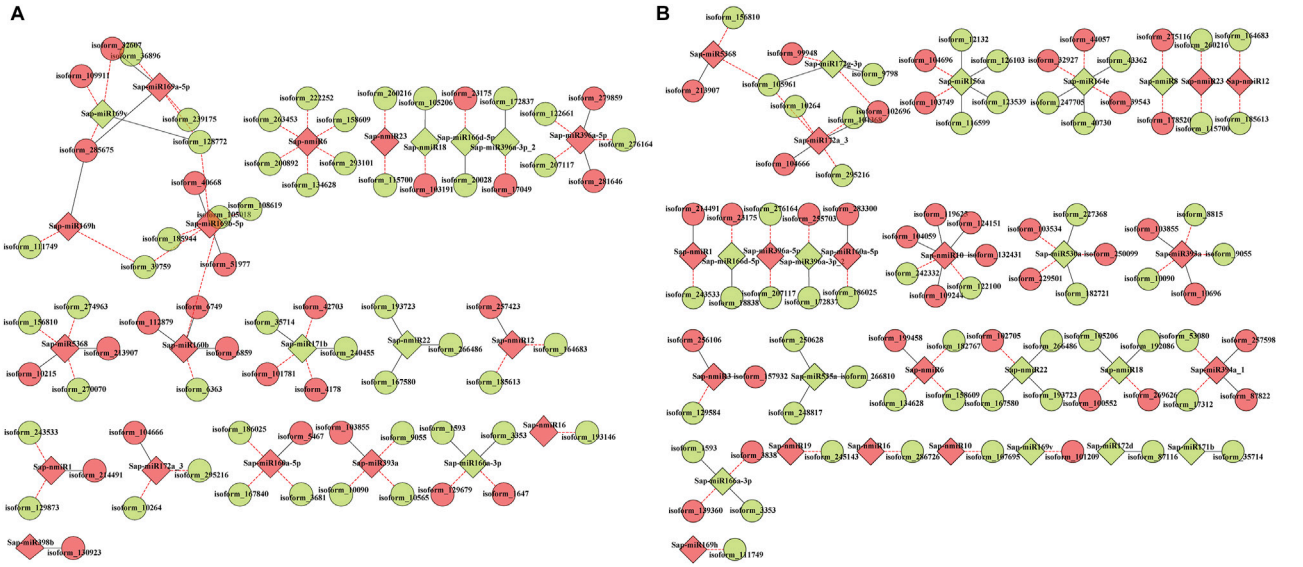


FIGURE 6 | Co-expressed miRNA-target pairs of *Sonneratia*-specific in samples treated with salt for 1 day **(A)** and 28 days **(B)**. Solid lines represent positive correlations and dashed lines represent negative correlations. miRNAs and targets are marked by rhombus and circles, respectively; red/green shading represents down/up regulated miRNAs or target genes.

and Sap-nmiR15, were detected in certain specific libraries, indicating their thorough induction or repression under salt stress. Previous studies have reported that miRNA168 in *Arabidopsis* and maize was coordinately accumulated by salinity, while maize miR167 and *S. linnaeanum*-miR399b were decreased (Yang et al., 2020). Additionally, we noted that some stress-regulated miRNAs, identified in this study, might be fine-tuned across adaptive responses to various stresses. For example, miR169 and miR319 were considered as a bridge, linking plant responses to ABA, drought and salt stress (De la Rosa et al., 2019; Salgado et al., 2021); miR398 was reported to play vital roles in various stresses such as water deficit (Yang et al., 2020), oxidative stress (Li et al., 2020), nutrient deficiency (Islam et al., 2022), salt stress (He et al., 2021), and bacterial infection (Salamon et al., 2021). This was probably due to the shared regulatory genes that are modulated by these miRNAs across distinct stress responses, attempting to reveal complex miRNA-mediated gene regulation was involved in the cross-response to abiotic and biotic stresses. For a further in-depth study of miRNA-mediated stress adaptation in mangroves, we would establish more comparisons to uncover crucial miRNAs that respond to stress. Particularly, miRNA profiles from salinity-exposed libraries would be compared to those of control samples growing for the same period of time.

A crucial step in understanding the potential regulatory roles of these salt-responsive miRNAs is the prediction and analysis of regulation of their targets. A total of 300 target genes for 67 significantly differentially accumulated miRNAs were identified in this study (**Supplementary Datas S4, S5**). The response to salt stress in plants is coupled with a wide range of intracellular processes, including signal sensing and transduction, transcription reprogramming, protein biosynthesis regulation that finally ascertain physiological changes to cope with stress. TFs play prominent roles during acclimation response when plants are subjected to severe environments. Here, 29.3% (88) of the 300 target genes encoded for TF family members, including MYB, AP2, HD-ZIP, NFYA, ARF and WRKY (**Supplementary Datas S4, S5**), which were reported to activate stress-related genes (Baillo et al., 2019; Yoon et al., 2020). Many other target genes encode for enzymes or functional proteins are considered to participate in salt stress response process. In *B. gymnorhiza* and *R. mangle* mangroves, miR172, miR394, and miR396 have been identified to act as important regulators of salt stress adaption via the regulation of stress signaling, oxidative resistance, and defense responses by targeting P-loop containing nucleoside triphosphate hydrolase superfamily protein (NTPase), UDP-glycosyltransferase (UGT) and Rhodanese/Cell cycle control phosphatase superfamily protein (RHOD), respectively (Wen et al., 2016). Herein, *Sonneratia*-specific targets of these miRNAs, involved in stress tolerance, were identified to participate in multiple processes of saline responses (**Supplementary Data S5**). Some targets encoded formetabolic enzymes. Sap-nmiR1 and Sap-nmiR18 could commonly target an *acyl-lipid omega-3 desaturase* (*FDA*). The antisense expression of *Arabidopsis FDA7* gene could reduce tobacco resistance to salinity (Xu et al., 2019). A gene encoding for a ubiquitin-specific protease (UBP) implicated in sugar metabolism (Feng et al., 2015), was the shared target of Sap-miR396a-3p_2 and Sap-nmiR8. Additionally, UBPs could also functions as Na^+/H^+ antiporter regulators, modulating monovalent cations and subsequent pH homeostasis in *Arabidopsis* (Zhou H. et al., 2012). Therefore, Sap-miR396a-3p_2/Sap-nmiR8-targeting *UBP*s may exhibit critical functions in maintenance of Na^+

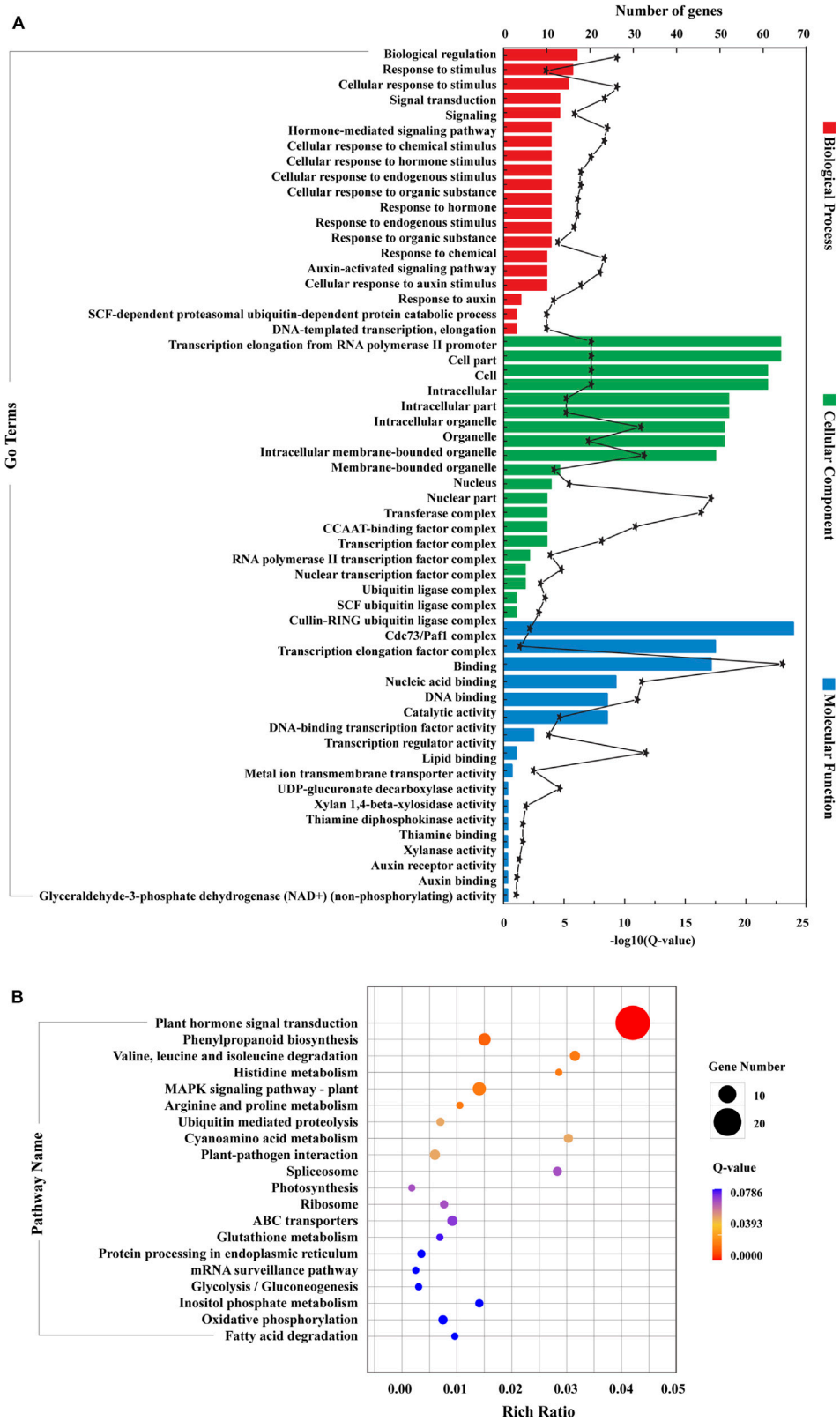


FIGURE 7 | GO **(A)** and KEGG **(B)** classification of target genes that comprised the co-expressed miRNA-target regulatory network. For GO-based enrichment analysis, the 20 most abundant subcategories of each category (biological processes, cellular components and molecular functions) were listed and the first 20 significantly enriched pathways were presented for KEGG analysis. All the enrichment items were listed if the total was less than 20.

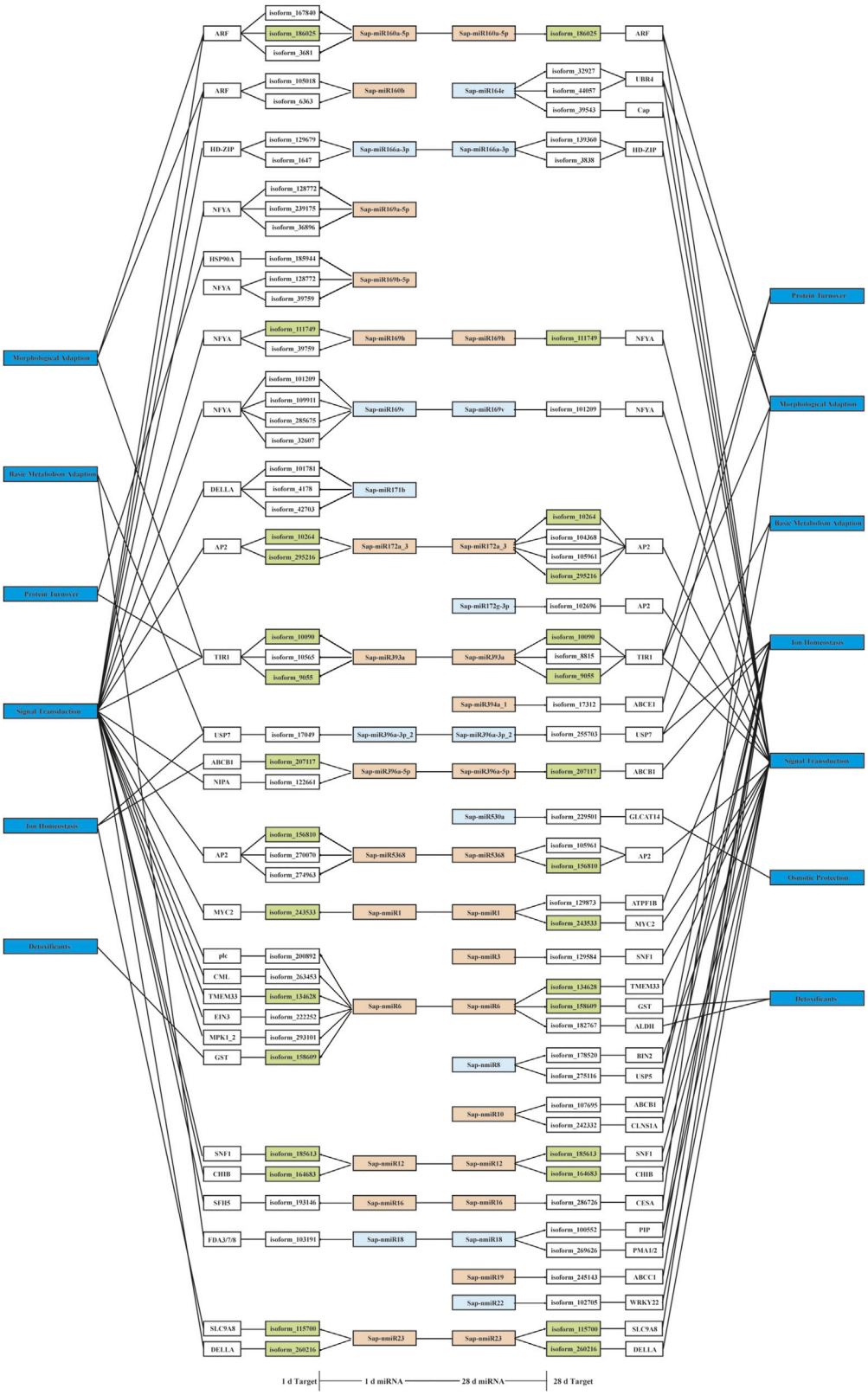


FIGURE 8 | A proposed regulatory network of negatively related *Sonneratia*-specific miRNA-target pairs. The color of the miRNA box indicates the type of expression pattern: orange, down regulated; light blue, up-regulated. Green box indicates that targets were negatively regulated by the respective miRNAs at two time points (1 and 28 days) of salt treatment.

and K^+ homeostasis in *S. apetala*, which is essential for plant survival during exposure to saline stress. Inclusively, Sap-nmiR6 might also play important roles in ion homeostasis by regulating the expression of *nitrate transporter* (*NRT*) under salt conditions. As an essential nutrient required for plant growth, nitrate serve as a signal that modulates plant development, and nitrate uptake is stimulated by salinity (Sa et al., 2019). The uptake and distribution of nitrate at the whole plant level are determined by joint activity of nitrate transporters. Oxidative stress is regarded as the most severe form of salt stress, which results from excessive accumulation of reactive oxygen species (ROS). Herein, several target genes involved in oxidative resistance were found to be cross regulated by miRNAs under saline conditions. Sap-miR4995 is predicted to target *L-ascorbate oxidase homolog* (*AOXH*) and *thioredoxin 1* (*trxA*); these proteins act as significant ROS scavenging enzymes and are involved in modifying intracellular ROS expression levels (Vishwakarma et al., 2015; González et al., 2020). Also, Cu/Zn superoxide dismutase, encoded by *Cu/Zn-SOD*, is widely reported to function in scavenging excess ROS in plants exposed to salinity (Sarkar et al., 2022) and here is predicted to be Sap-miR4995/Sap-nmiR18-regulated. Notably, *Cu/Zn SODs* were identified as targets of miR398 in *B. gymnorhiza* and *A. marina* (Khraiwesh et al., 2013; Wen et al., 2016). The detoxificant enzyme gene, *aldehyde dehydrogenase* (*ALDHs*) was targeted by a novel miRNA (Sap-nmiR6) in *S. apetala*, but by miR399 in *A. marina* (Khraiwesh et al., 2013). These observations indicate a difference in miRNA-mediated salt stress response for different mangrove plants. These miRNA-target pairs may serve as a supplementary resource for translating miRNA-mediated gene regulation for enhancing plant stress tolerance.

Mediation of a Potential Co-Expression Network *via Sonneratia*-Specific MicroRNAs in Response to Salt Stress

Although a large number of stress-related miRNAs were evolutionarily conserved in plants, some miRNAs have variable regulatory patterns across different species. According to the genome and abundance characteristic of salt-responsive miRNAs identified in this study, 62 (45 known and 17 novel) miRNAs potentially modify the acclimation response to salt stress in a species-specific manner (**Supplementary Table S2**; **Supplementary Data S2, S6**). After combining correlation coefficient and accumulation patterns of these *Sonneratia*-specific miRNAs and their target counterparts under salt stress, 34 miRNAs and 131 targets producing 140 miRNA-target interactions were determined to be co-expressed after salt treatment (**Figure 6**; **Supplementary Data S8–S10**). Within this co-expression network, the number of pairwise miRNA-target interactions differed dramatically at two-time points of salt treatment. In LT1 vs. LCK comparison, 78 miRNAs-target pairs (29 positive and 49 negative) were significantly induced under salt stress (**Figure 6A**), whereas 91 miRNAs-target interactions (42 positive and 49 negative) were determined in LT2 vs. LCK comparison (**Figure 6B**). These findings implied that miRNAs, which exhibited concordant positive or negative regulatory relationships with their corresponding targets throughout the time-course of salinity, may have similar response mechanisms in *S. apetala*. Further analysis demonstrated that 111 miRNAs displayed disparate changes in their accumulation in *S. apetala* under early-(1 day) and late-stages (28 days) of salt stress (**Supplementary Data S10**). This differential pattern of miRNA accumulation may be liable to a distinct salt response mechanism at the two-time points. GO enrichment analysis of their putative targets contained many abiotic stress-related categories, such as "response to stimulus", "intracellular" and "nucleic acid binding" (**Figure 7A**). According to KEGG pathways analysis, all the signal transduction pathways were classified into environmental information processing (**Supplementary Data S12**). It is possible that these signal transduction pathways may promote high salt tolerance of *S. apetala*. Further, phenylpropanoid biosynthesis, starch and sucrose metabolism, fatty acid degradation, spliceosome, and other pathways were activated *via* various enzymes under salt stress, which may initiate the production of salt-induced signaling molecules, ion channels, oxidation protectants, ROS, and other stress-related metabolites (Liu et al., 2020).

Generally, miRNAs modulate the specific genes by pairing with mRNAs transcripts, leading to degradation or translation inhibition and therefore their expression profiles should be negatively related (Tyagi et al., 2019). There is no doubt that the negatively correlated miRNA-target pairs should be emphasized. On this basis, we separated out 82 pairwise miRNA-target interactions that exhibited negative correlations for further analysis (**Supplementary Data S10**). Among these, two of the novel miRNAs, Sap-nmiR10 and Sap-nmiR23, exhibited cross-functionality with the known Sap-miR396a-5p and Sap-miR171b for the targets, *ABCB1* and *DELLA*, respectively (**Figure 8**). This indicated that new *Sonneratia*-specific miRNAs could evolve to co-target certain genes with conserved miRNAs and shared common response mechanisms under salt stress. Particularly, 13 miRNAs and their 18 target counterparts (**Supplementary Table S4**) were categorized into the "environmental information processing" category, which were involved in plant hormone signal transduction (ko04075), MAPK signaling pathway–plant (ko04016) and ABC transporters (ko02010). Acclimation response of plants to salt stress requires coordination and integration of multiple phytohormones, including auxin (IAA) (Ribba et al., 2020), ethylene (Wang and Huang, 2019), ABA (Zhu, 2016), and jasmonic acid (JA) (Ali and Baek, 2020). Previous studies have demonstrated that ethylene-induced the detoxification machinery and increased plant salt tolerance (Zhao et al., 2020). As a TF that mediates core ethylene signaling, ethylene-insensitive protein 3 (EIN3) was stabilized by salinity and promoted plant survival to saline stress via the DELLA proteins (Zhao et al., 2020), which was targeted by Sap-miR171b and Sap-nmiR23. Also, auxin participates in multiple processes of salt stress responses, modulating a complex balance of biosynthesis, signaling, ion transport, and finally fine-tuning of physiological changes (Ribba et al., 2020). Notably, auxin

response factors (ARFs) were reported to serve as a point of crosstalk between the two hormones, ethylene and IAA, which not only function in auxin signaling, but also play a vital role in ethylene responses (Ribba et al., 2020). In this study, both Sap-miR160b and Sap-miR160a-5p were down-regulated at the early-stage (1 day) of salt stress, resulting in an increased expression of the target *ARF* (**Supplementary Data S10**). This increase in *ARF* expression may serve crucial functions in salt resistance of *S. apetala*. Sap-miR393a-targeted *TIR1*, encoding an F-box family protein, is a negative regulator in auxin signaling (Fendrych et al., 2018), which was also implicated in ARF-medicated signal transduction. Two out of three target genes, isoform_9055 and isoform_10090, encoding TIR1, exhibited an increased expression until the end of the salt treatment. The accumulation of their miRNA regulator, Sap-miR393a, was down-regulated after 1 and 28 days of salt treatment (**Supplementary Data S10**), indicating a pivotal role of Sap-miR393a-*TIR1* interaction in salt tolerance of *S. apetala*. Further, analysis of transgenic *Arabidopsis* and rice plants had revealed that the overexpression of OsmiR393 led to enhanced sensitivity to salinity and alkaline stresses (Giri et al., 2021). In light of the above observations, we inferred that miR393 was a negative modulator of plant salt tolerance. MAPK cascades are important signaling pathways related to plant responses to salt stress (Kumar et al., 2020). In this study, four target genes, *MYC2*, *WRKY22*, *CHIB* and *MPK1_2*, which were paired by Sap-nmiR1, Sap-nmiR22, Sap-nmiR12 and Sap-nmiR6, respectively, were identified to be associated with MAPK pathway in our KEGG annotation (**Supplementary Table S4**). These results indicated a crucial role for novel miRNAs during the *S. apetala* survival under saline conditions. Further, ABC transporters are known to mediate a H^+-Na^+-Cl^- symport reaction and can improve salt-adapted cells survival during downshifts of extracellular ion concentration (Velamakanni et al., 2009). The co-target of Sap-miR396a-5p and Sap-nmiR10, an *ABCB1,* and Sap-nmiR19-targeted *ABCC1* were induced under salt conditions, which indicated the importance of ABC transporters in ion homeostasis and compatible molecule accumulation during the process of saline tolerance in *S. apetala*. Additionally, target genes involved in basic metabolism of sugar, lipid and amino acid, detoxification, osmotic protection, protein turnover, as well as morphological adaption were also identified in the negatively correlated miRNA-target pairs (**Figure 8**). The discovery of many salt-regulated miRNA indicates that a complicated regulatory network is implicated in response to salinity in *S. apetala*. Further efforts are needed to functionally confirm these putative interactions and probe the regulatory mechanism underlying these interactions in *S. apetala* adaptive response to saline conditions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. These data can be found at: https://ngdc.cncb.ac.cn/gsa/, CRA006866 and CRA006863.

AUTHOR CONTRIBUTIONS

BC designed the experiments; ZD, XZ, YW, FH, and JS performed the experiments; JC and WH collected and analyzed the data; BC wrote the manuscript; and all authors read and approved the manuscript.

FUNDING

This work was supported by the Project of the National Natural Science Foundation of China (31901330), the Competitive Allocation Project of Special Funds for Science and Technology Development in Zhanjiang City (2020A01007), the Scientific Research Fund Project of Hainan University (KYQD(ZR)1830), and the Innovation and Entrepreneurship Training Program for College Students (S202110566024).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.932832/full#supplementary-material

REFERENCES

Ali, M. S., and Baek, K. H. (2020). Jasmonic Acid Signaling Pathway in Response to Abiotic Stresses in Plants. *Int. J. Mol. Sci.* 21 (2), 621. doi:10.3390/ijms21020621

Alzahrani, S. M., Alaraidh, I. A., Khan, M. A., Migdadi, H. M., Alghamdi, S. S., and Alsahli, A. A. (2019). Identification and Characterization of Salt-Responsive microRNAs in *Vicia faba* by High-Throughput Sequencing. *Genes* 10 (4), 303. doi:10.3390/genes10040303

Baillo, E. H., Kimotho, R. N., Zhang, Z. B., and Xu, P. (2019). Transcription Factors Associated with Abiotic and Biotic Stress Tolerance and Their Potential for Crops Improvement. *Genes* 10 (10), 771. doi:10.3390/genes10100771

BIG Data Center Members (2018). Database Resources of the BIG Data Center in 2018. *Nucleic Acids Res.* 46 (D1), D14–D20. doi:10.1093/nar/gkx897

Chen, Q., Deng, B., Gao, J., Zhao, Z., Chen, Z., Song, S., et al. (2019). Comparative Analysis of miRNA Abundance Revealed the Function of Vvi-miR828 in Fruit Coloring in Root Restriction Cultivation Grapevine (*Vitis vinifera* L.). *Int. J. Mol. Sci.* 20 (16), 4058. doi:10.3390/ijms20164058

Cui, J., Lu, W., Lu, Z., Ren, S., Zhao, B., Wang, L., et al. (2019). Identification and Analysis of microRNAs in the SAM and Leaves of *Populus Tomentosa*. *Forests* 10 (2), 130. doi:10.3390/f10020130

Dahibhate, N. L., Roy, U., and Kumar, K. (2020). Phytochemical Screening, Antimicrobial and Antioxidant Activities of Selected Mangrove Species. *Curr. Bioact. Compd.* 16 (2), 152–163. doi:10.2174/1573407214666180808121118

Dasgupta, N., Hazra, A., Bhattacharya, S., Hazra, A., and Das, S. (2017). In Silico screening of Putative miRNAs and Their Targets from a Common Mangrove *Bruguiera Gymnorrhiza*. *Int. J. Mol. Sci. Biol.* 2 (1), 1–18. doi:10.19080/ijcsmb.2017.02.555579

Dassanayake, M., Haas, J. S., Bohnert, H. J., and Cheeseman, J. M. (2009). Shedding Light on an Extremophile Lifestyle through Transcriptomics. *New Phytol.* 183 (3), 764–775. doi:10.1111/j.1469-8137.2009.02913.x

De la Rosa, C., Covarrubias, A. A., and Reyes, J. L. (2019). A Dicistronic Precursor Encoding miR398 and the Legume-specific miR2119 Coregulates *CSD1* and *ADH1* mRNAs in Response to Water Deficit. *Plant, Cell Environ.* 42 (1), 133–144. doi:10.1111/pce.13209

Evers, M., Huttner, M., Dueck, A., Meister, G., and Engelmann, J. C. (2015). miRA: Adaptable Novel miRNA Identification in Plants Using Small RNA Sequencing Data. *BMC Bioinf.* 16 (1), 1–10. doi:10.1186/s12859-015-0798-3

Fahlgren, N., and Carrington, J. C. (2010). "miRNA Target Prediction in Plants," in *Plant MicroRNAs* (Totowa: Humana Press), 51–57. doi:10.1007/978-1-60327-005-2_4

Fehlmann, T., Reinheimer, S., Geng, C., Su, X., Drmanac, S., Alexeev, A., et al. (2016). cPAS-Based Sequencing on the BGISEQ-500 to Explore Small Non-Coding RNAs. *Clin. Epigenet.* 8 (1), 1–11. doi:10.1186/s13148-016-0287-1

Fendrych, M., Akhmanova, M., Merrin, J., Glanc, M., Hagihara, S., Takahashi, K., et al. (2018). Rapid and Reversible Root Growth Inhibition by TIR1 Auxin Signalling. *Nat. Plants* 4 (7), 453–459. doi:10.1038/s41477-018-0190-1

Feng, J., Wang, J., Fan, P., Jia, W., Nie, L., Jiang, P., et al. (2015). High-throughput Deep Sequencing Reveals that microRNAs Play Important Roles in Salt Tolerance of Euhalophyte *Salicornia Europaea*. *BMC Plant Biol.* 15 (1), 1–17. doi:10.1186/s12870-015-0451-3

Feng, X., Xu, S., Li, J., Yang, Y., Chen, Q., Lyu, H., et al. (2020). Molecular Adaptation to Salinity Fluctuation in Tropical Intertidal Environments of a Mangrove Tree *Sonneratia Alba*. *BMC Plant Biol.* 20 (1), 1–14. doi:10.1186/s12870-020-02395-3

Gao, Z., Ma, C., Zheng, C., Yao, Y., and Du, Y. (2022). Advances in the Regulation of Plant Salt-Stress Tolerance by miRNA. *Mol. Biol. Rep.* 1-15. doi:10.1007/s11033-022-07179-6

Giri, J., Parida, S. K., Raghuvanshi, S., and Tyagi, A. K. (2021). Emerging Molecular Strategies for Improving Rice Drought Tolerance. *Curr. Genomics.* 22 (1), 16–25. doi:10.2174/1389202921999201231205024

González, D., Álamos, P., Rivero, M., Orellana, O., Norambuena, J., Chávez, R., et al. (2020). Deciphering the Role of Multiple Thioredoxin Fold Proteins of *Leptospirillum* Sp. In Oxidative Stress Tolerance. *Int. J. Mol. Sci.* 21 (5), 1880. doi:10.3390/ijms21051880

Gupta, K., Mishra, S. K., Gupta, S., Pandey, S., Panigrahi, J., and Wani, S. H. (2021). Functional Role of miRNAs: Key Players in Soybean Improvement. *Phyton* 90 (5), 1339–1362. doi:10.32604/phyton.2021.015239

He, Y., Zhou, J., Hu, Y., Fang, C., Yu, Y., Yang, J., et al. (2021). Overexpression of Sly-miR398b Increased Salt Sensitivity Likely via Regulating Antioxidant System and Photosynthesis in Tomato. *Environ. Exp. Bot.* 181, 104273. doi:10.1016/j.envexpbot.2020.104273

Hu, Z., Liu, A., Gitau, M. M., Huang, X., Chen, L., and Fu, J. (2018). Insights into the microRNA-Regulated Response of Bermudagrass to Cold and Salt Stress. *Environ. Exp. Bot.* 145, 64–74. doi:10.1016/j.envexpbot.2017.10.026

Hu, Z., Wang, F., Yu, H., Zhang, M., Jiang, D., Huang, T., et al. (2022). Effects of Scion-Rootstock Interaction on Citrus Fruit Quality Related to Differentially Expressed Small RNAs. *Sci. Hortic.* 298, 110974. doi:10.1016/j.scienta.2022.110974

Islam, W., Tauqeer, A., Waheed, A., and Zeng, F. (2022). MicroRNA Mediated Plant Responses to Nutrient Stress. *Int. J. Mol. Sci.* 23 (5), 2562. doi:10.3390/ijms23052562

Khraiwesh, B., Pugalenthi, G., and Fedoroff, N. V. (2013). Identification and Analysis of Red Sea Mangrove (*Avicennia Marina*) microRNAs by High-Throughput Sequencing and Their Association with Stress Responses. *PloS one* 8 (4), e60774. doi:10.1371/journal.pone.0060774

Kumar, K., Raina, S. K., and Sultan, S. M. (2020). *Arabidopsis* MAPK Signaling Pathways and Their Cross Talks in Abiotic Stress Response. *J. Plant Biochem. Biotechnol.* 29 (4), 700–714. doi:10.1007/s13562-020-00596-3

Lertampaiporn, S., Thammarongtham, C., Nukoolkit, C., Kaewkamnerdpong, B., and Ruengjitchatchawalya, M. (2013). Heterogeneous Ensemble Approach with Discriminative Features and Modified-SMOTEbagging for Pre-miRNA Classification. *Nucleic Acids Res.* 41 (1), e21. doi:10.1093/nar/gks878

Li, B., and Dewey, C. N. (2011). RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinf.* 12 (1), 1–16. doi:10.1186/1471-2105-12-323

Li, B., Duan, H., Li, J., Deng, X. W., Yin, W., and Xia, X. (2013). Global Identification of miRNAs and Targets in *Populus Euphratica* under Salt Stress. *Plant Mol. Biol.* 81 (6), 525–539. doi:10.1007/s11103-013-0010-y

Li, W., Xu, R., Yan, X., Liang, D., Zhang, L., Qin, X., et al. (2019). De Novo Leaf and Root Transcriptome Analysis to Explore Biosynthetic Pathway of Celangulin V in *Celastrus Angulatus* Maxim. *BMC Genomics* 20 (1), 1–15. doi:10.1186/s12864-018-5397-z

Li, J., Shen, Y., Zhu, J., Liu, S., Zeng, N., and Zhan, X. (2020). miR398 Is Involved in the Relief of Phenanthrene-Induced Oxidative Toxicity in Wheat Roots. *Environ. Pollut.* 258, 113701. doi:10.1016/j.envpol.2019.113701

Li, J., Duan, Y., Sun, N., Wang, L., Feng, S., Fang, Y., et al. (2021). The miR169n-*NF-YA8* Regulation Module Involved in Drought Resistance in *Brassica Napus* L. *Plant Sci.* 313, 111062. doi:10.1016/j.plantsci.2021.111062

Liu, Y., Dai, X. B., Zhao, L. K., Huo, K. S., Jin, P. F., Zhao, D. L., et al. (2020). RNA-Seq Reveals the Salt Tolerance of *Ipomoea Pes-Caprae*, a Wild Relative of Sweet Potato. *J. Plant Physiol.* 255, 153276. doi:10.1016/j.jplph.2020.153276

Lorenz, R., Bernhart, S. H., Siederdissen, C. H., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6 (1), 1–14. doi:10.1186/1748-7188-6-26

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 (12), 550. doi:10.1186/s13059-014-0550-8

McDiarmid, T. A., Kepler, L. D., and Rankin, C. H. (2020). Auxin Does Not Affect a Suite of Morphological or Behavioral Phenotypes in Two Wild-type *C. elegans* Strains. *Micropubl. Biol.*. doi:10.17912/micropub.biology.000307

Mei, M., Wei, J., Ai, W. F., Zhang, L. J., and Lu, X. J. (2021). Integrated RNA and miRNA Sequencing Analysis Reveals a Complex Regulatory Network of *Magnolia Sieboldii* Seed Germination. *Sci. Rep.* 11 (1), 1–12. doi:10.1038/s41598-021-90270-y

Meng, X., Xu, J., Zhang, M., Du, R., Zhao, W., Zeng, Q., et al. (2021). Third-generation Sequencing and Metabolome Analysis Reveal Candidate Genes and Metabolites with Altered Levels in Albino Jackfruit Seedlings. *BMC Genomics* 22 (1), 1–14. doi:10.1186/s12864-021-07873-y

Nawrocki, E. P., and Eddy, S. R. (2013). Infernal 1.1: 100-fold Faster RNA Homology Searches. *Bioinformatics* 29, 2933–2935. doi:10.1093/bioinformatics/btt509

Necira, K., Makki, M., Sanz-García, E., Canto, T., Djilani-Khouadja, F., and Tenllado, F. (2021). Topical Application of *Escherichia* Coli-Encapsulated dsRNA Induces Resistance in *Nicotiana Benthamiana* to Potato Viruses and Involves RDR6 and Combined Activities of DCL2 and DCL4. *Plants* 10 (4), 644. doi:10.3390/plants10040644

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

Pan, J., Li, Z., Wang, Q., Garrell, A. K., Liu, M., Guan, Y., et al. (2018). Comparative Proteomic Investigation of Drought Responses in Foxtail Millet. *BMC Plant Biol.* 18 (1), 1–19. doi:10.1186/s12870-018-1533-9

Pearman, W. S., Freed, N. E., and Silander, O. K. (2020). Testing the Advantages and Disadvantages of Short-And Long-Read Eukaryotic Metagenomics Using Simulated Reads. *BMC Bioinf.* 21 (1), 1–15. doi:10.1186/s12859-020-3528-4

Peng, Z., He, S., Gong, W., Sun, J., Pan, Z., Xu, F., et al. (2014). Comprehensive Analysis of Differentially Expressed Genes and Transcriptional Regulation Induced by Salt Stress in Two Contrasting Cotton Genotypes. *BMC Genomics* 15 (1), 1–28. doi:10.1186/1471-2164-15-760

Ribba, T., Garrido-Vargas, F., and O'Brien, J. A. (2020). Auxin-mediated Responses under Salt Stress: From Developmental Regulation to Biotechnological Applications. *J. Exp. Bot.* 71 (13), 3843–3853. doi:10.1093/jxb/eraa241

Sa, G., Yao, J., Deng, C., Liu, J., Zhang, Y., Zhu, Z., et al. (2019). Amelioration of Nitrate Uptake under Salt Stress by Ectomycorrhiza with and without a Hartig Net. *New Phytol.* 222 (4), 1951–1964. doi:10.1111/nph.15740

Saddhe, A. A., and Kumar, K. (2019). Molecular Cloning, Expression Analysis, and Heterologous Characterization of a Novel Sodium/hydrogen Exchanger from a Mangrove Plant, *Rhizophora Apiculata*. *Plant gene.* 19, 100192. doi:10.1016/j.plgene.2019.100192

Salamon, S., Żok, J., Gromadzka, K., and Błaszczyk, L. (2021). Expression Patterns of miR398, miR167, and miR159 in the Interaction between Bread Wheat (*Triticum aestivum* L.) and *Pathogenic Fusarium Culmorum* and Beneficial *Trichoderma* Fungi. *Pathog* 10 (11), 1461. doi:10.3390/pathogens10111461

Salgado, F. F., Vieira, L. R., Silva, V. N. B., Leão, A. P., Grynberg, P., do Carmo Costa, M. M., et al. (2021). Expression Analysis of miRNAs and Their Putative

Target Genes Confirm a Preponderant Role of Transcription Factors in the Early Response of Oil Palm Plants to Salinity Stress. *BMC Plant Biol.* 21 (1), 1–17. doi:10.1186/s12870-021-03296-9

Sarkar, R. K., Bhowmik, M., Biswas Sarkar, M., Sircar, G., and Bhattacharya, K. (2022). Comprehensive Characterization and Molecular Insights into the Salt Tolerance of a Cu, Zn-Superoxide Dismutase from an Indian Mangrove. *Avicennia Mar. Sci. Rep.* 12 (1), 1–15. doi:10.1038/s41598-022-05726-6

Sharma, P., Al-Dossary, O., Alsubaie, B., Al-Mssallem, I., Nath, O., Mitter, N., et al. (2021). Improvements in the Sequencing and Assembly of Plant Genomes. *Gigabyte*9, 1–11. doi:10.1093/gigascience/giaa14

Si, J., Zhou, T., Bo, W., Xu, F., and Wu, R. (2014). Genome-Wide Analysis of Salt-Responsive and Novel microRNAs in *Populus Euphratica* by Deep Sequencing. *BMC Genet.* 15 (Suppl 1), S6. doi:10.1186/1471-2156-15-S1-S6

Singh, P., Hazra, A., Biswas, S. M., and Chakraborty, S. N. (2020). Identification of Stress-Induced Plant microRNAs and Their Targets from a True Mangrove *Rhizophora Apiculata*-An In Silico Approach. *Int. J. Bioinf. Biol. Sci.* 8 (1), 13–17. doi:10.30954/2319-5169.01.2020.3

Srikanth, S., Lum, S. K. Y., and Chen, Z. (2016). Mangrove Root: Adaptations and Ecological Importance. *Trees* 30 (2), 451–465. doi:10.1007/s00468-015-1233-0

Syed, A., Sarwar, G., Shah, S. H., and Muhammad, S. (2021). Soil Salinity Research in 21st Century in Pakistan: its Impact on Availability of Plant Nutrients, Growth and Yield of Crops. *Commun. Soil Sci. Plant Anal.* 52 (3), 183–200. doi:10.1080/00103624.2020.1854294

Tang, C., Xie, Y. M., Guo, M., and Yan, W. (2021). AASRA: an Anchor Alignment-Based Small RNA Annotation Pipeline. *Biol. Reprod.* 105 (1), 267–277. doi:10.1093/biolre/ioab062

Tian, R., Seim, I., Ren, W., Xu, S., and Yang, G. (2019). Contraction of the ROS Scavenging Enzyme Glutathione S-Transferase Gene Family in Cetaceans. *G3 Genes, Genomes, Genet.* 9 (7), 2303–2315. doi:10.1534/g3.119.400224

Tyagi, S., Sharma, S., Ganie, S. A., Tahir, M., Mir, R. R., and Pandey, R. (2019). Plant microRNAs: Biogenesis, Gene Silencing, Web-Based Analysis Tools and Their Use as Molecular Markers. *3 Biotech.* 9, 413. doi:10.1007/s13205-019-1942-y

Velamakanni, S., Lau, C. H., Gutmann, D. A., Venter, H., Barrera, N. P., Seeger, M. A., et al. (2009). A Multidrug ABC Transporter with a Taste for Salt. *PloS one* 4 (7), e6137. doi:10.1371/journal.pone.0006137

Vishwakarma, A., Tetali, S., Selinski, J., Scheibe, R., and Padmasree, K. (2015). Importance of the Alternative Oxidase (AOX) Pathway in Regulating Cellular Redox and ROS Homeostasis to Optimize Photosynthesis during Restriction of the Cytochrome Oxidase Pathway in *Arabidopsis thaliana*. *Ann. Bot.* 116, 555–569. doi:10.1093/aob/mcv122

Wang, Y. S., and Gu, J. D. (2021). Ecological Responses, Adaptation and Mechanisms of Mangrove Wetland Ecosystem to Global Climate Change and Anthropogenic Activities. *Int. Biodeterior. Biodegr.* 162, 105248. doi:10.1016/j.ibiod.2021.105248

Wang, J., and Huang, R. (2019). Modulation of Ethylene and Ascorbic Acid on Reactive Oxygen Species Scavenging in Plant Salt Response. *Front. Plant Sci.* 10, 319. doi:10.3389/fpls.2019.00319

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: Genome Sequence Archive. *Genomics, Proteomics bioinf.* 15 (1), 14–18. doi:10.1016/j.gpb.2017.01.001

Wang, B., Cheng, D., Chen, Z., Zhang, M., Zhang, G., Jiang, M., et al. (2019). Bioinformatic Exploration of the Targets of Xylem Sap miRNAs in Maize under Cadmium Stress. *Int. J. Mol. Sci.* 20 (6), 1474. doi:10.3390/ijms20061474

Wang, Y., Dai, A., and Tang, T. (2021a). Weak Effect of Gypsy Retrotransposon Bursts on *Sonneratia Alba* Salt Stress Gene Expression. *Front. Plant Sci.* 12, 830079. doi:10.3389/fpls.2021.830079

Wang, Y., Li, X., Wang, C., Gao, L., Wu, Y., Ni, X., et al. (2021b). Unveiling the Transcriptomic Complexity of *Miscanthus Sinensis* Using a Combination of PacBio Long Read-And Illumina Short Read Sequencing Platforms. *BMC Genomics* 22 (1), 1–16. doi:10.1186/s12864-021-07971-x

Wani, S. H., Kumar, V., Khare, T., Tripathi, P., Shah, T., Ramakrishna, C., et al. (2020). miRNA Applications for Engineering Abiotic Stress Tolerance in Plants. *Biologia* 75 (7), 1063–1081. doi:10.2478/s11756-019-00397-7

Wen, M., Lin, X. Q., Xie, M. N., Xie, M. A., Wang, Y. S., Shen, X., et al. (2016). Small RNA Transcriptomes of Mangroves Evolve Adaptively in Extreme Environments. *Sci. Rep.* 6 (1), 27551. doi:10.1038/srep27551

Wu, H. J., Ma, Y. K., Chen, T., Wang, M., and Wang, X. J. (2012). PsRobot: a Web-Based Plant Small RNA Meta-Analysis Toolbox. *Nucleic Acids Res.* 40 (W1), W22–W28. doi:10.1093/nar/gks554

Xu, L., Zeng, W., Li, J., Liu, H., Yan, G., Si, P., et al. (2019). Characteristics of Membrane-Bound *Fatty Acid Desaturase* (*FAD*) Genes in Brassica Napus L. And Their Expressions under Different Cadmium and Salinity Stresses. *Environ. Exp. Bot.* 162, 144–156. doi:10.1016/j.envexpbot.2019.02.016

Xu, T., Zhang, L., Yang, Z., Wei, Y., and Dong, T. (2021). Identification and Functional Characterization of Plant miRNA under Salt Stress Shed Light on Salinity Resistance Improvement through miRNA Manipulation in Crops. *Front. Plant Sci.* 12, 972. doi:10.3389/fpls.2021.665439

Yang, Z., Zhu, P., Kang, H., Liu, L., Cao, Q., Sun, J., et al. (2020). High-throughput Deep Sequencing Reveals the Important Role that microRNAs Play in the Salt Response in Sweet Potato (*Ipomoea Batatas* L.). *BMC Genomics* 21 (1), 1–16. doi:10.1186/s12864-020-6567-3

Yoon, Y., Seo, D. H., Shin, H., Kim, H. J., Kim, C. M., and Jang, G. (2020). The Role of Stress-Responsive Transcription Factors in Modulating Abiotic Stress Tolerance in Plants. *Agron* 10 (6), 788. doi:10.3390/agronomy10060788

Zaaba, N. F., Jaafar, M., and Ismail, H. (2021). Tensile and Morphological Properties of Nanocrystalline Cellulose and Nanofibrillated Cellulose Reinforced PLA Bionanocomposites: A Review. *Polym. Eng. Sci.* 61 (1), 22–38. doi:10.1002/pen.25560

Zhan, J., Diao, Y., Yin, G., Sajjad, M., Wei, X., and Lu, Z. (2021). Integration of mRNA and miRNA Analysis Reveals the Molecular Mechanism of Cotton Response to Salt Stress. *Front. Plant Sci.* 12, 767984. doi:10.3389/fpls.2021.767984

Zhao, Z., Niu, S., Fan, G., Deng, M., and Wang, Y. (2018). Genome-wide Analysis of Gene and microRNA Expression in Diploid and Autotetraploid *Paulownia Fortunei* (Seem) Hemsl. Under Drought Stress by Transcriptome, microRNA, and Degradome Sequencing. *Forests* 9 (2), 88. doi:10.3390/f9020088

Zhao, C., Zhang, H., Song, C., Zhu, J. K., and Shabala, S. (2020). Mechanisms of Plant Responses and Adaptation to Soil Salinity. *Innovation* 1 (1), 100017. doi:10.1016/j.xinn.2020.100017

Zhou, H., Zhao, J., Yang, Y., Chen, C., Liu, Y., Jin, X., et al. (2012). Ubiquitin-specific Protease16 Modulates Salt Tolerance in *Arabidopsis* by Regulating Na^+/H^+ Antiport Activity and Serine Hydroxymethyltransferase Stability. *Plant Cell* 24 (12), 5106–5122. doi:10.1105/tpc.112.106393

Zhou, J., Liu, M. Y., Jiang, J., Qiao, G. R., Lin, S., Li, H. Y., et al. (2012). Expression Profile of miRNAs in *Populus Cathayana* L. And *Salix Matsudana* Koidz under Salt Stress. *Mol. Biol. Rep.* 39 (9), 8645–8654. doi:10.1007/s11033-012-1719-4

Zhu, J. K. (2016). Abiotic Stress Signaling and Responses in Plants. *Cell* 167 (2), 313–324. doi:10.1016/j.cell.2016.08.029

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 09 August 2022
doi: 10.3389/fgene.2022.932207

Identification and Analysis of Long Non-Coding RNAs Related to UV-B-Induced Anthocyanin Biosynthesis During Blood-Fleshed Peach (*Prunus persica*) Ripening

Man Zhang[1,2†], Xiuqi Zhang[3†], Haijing Wang[1,2], Mao Ye[1,2], Yating Liu[1,2], Zhihua Song[3], Tingting Du[3], Hongyan Cao[3], Liqin Song[1,2], Xiao Xiao[1,2], Jianzhen Liu[1,2], Libin Zhang[1,2], Yangbo Song[4], Qing Yang[3], Dong Meng[3] and Junkai Wu[1,2]**

[1]College of Horticulture Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao, China, [2]Hebei Key Laboratory of Horticultural Germplasm Excavation and Innovative Utilization, Qinhuangdao, China, [3]The Key Laboratory for Silviculture and Conservation of Ministry of Education, The College of Forestry, Beijing Forestry University, Beijing, China, [4]College of Agriculture and Animal Husbandry, Qinghai University, Xining, China

OPEN ACCESS

Edited by:
Mahmoud Yaish,
Sultan Qaboos University, Oman

Reviewed by:
Yi Xu,
Nanjing Agricultural University, China
Chunlong Li,
Huazhong Agricultural University, China

***Correspondence:**
Dong Meng
mengdongjlf@163.com
Junkai Wu
mans5@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: 29 April 2022
Accepted: 17 June 2022
Published: 09 August 2022

Citation:
Zhang M, Zhang X, Wang H, Ye M, Liu Y, Song Z, Du T, Cao H, Song L, Xiao X, Liu J, Zhang L, Song Y, Yang Q, Meng D and Wu J (2022) Identification and Analysis of Long Non-Coding RNAs Related to UV-B-Induced Anthocyanin Biosynthesis During Blood-Fleshed Peach (Prunus persica) Ripening. Front. Genet. 13:932207. doi: 10.3389/fgene.2022.932207

Blood flesh is a key fruit trait in peaches (*Prunus persica*) and can be attributed to the accumulation of anthocyanins. The roles of long non-coding RNAs (lncRNAs) have been highlighted by multiple studies in regulating fruit ripening, anthocyanin accumulation, and abiotic stress responses in many flowering plants. Such regulatory functions of lncRNAs in Prunus persica, nonetheless, have not been reported. In this research, we sequenced and analyzed the complete transcriptome of C3-20 (a blood-fleshed peach) fruit at four developmental stages. Analyses of the correlated genes and differentially expressed lncRNA target genes helped to forecast lncRNAs' possible functions. The RNA-seq data were generated using high-throughput sequencing. In total, 17,456 putative lncRNAs, including 4,800 intergenic lncRNAs, 2,199 antisense lncRNAs, and 10,439 intronic lncRNAs were discovered, of which 4,871 differentially expressed lncRNAs (DE-lncRNAs) were annotated in the fruit developmental processes. The target genes of these DE-lncRNAs and their regulatory relationship identifying 21,795 cis-regulated and 18,271 trans-regulated targets of the DE-lncRNAs were in a similar way predicted by us. The enriched GO terms for the target genes included anthocyanin biosynthesis. Flavonoid biosynthesis and plant hormone signal transduction were also included in the enriched KEGG pathways. Co-expression network construction demonstrated that the highly expressed genes might co-regulate multiple other genes associated with auxin signal transduction and take effect in equal pathways. We discovered that lncRNAs, including LNC_000987, LNC_000693, LNC_001323, LNC_003610, LNC_001263, and LNC_003380, correlated with fruit that ripened and could take part in ethylene biosynthesis and metabolism and the ABA signaling pathway. Several essential transcription factors, such as *ERFs, WRKY70, NAC56*, and *NAC72*, may in a similar way regulate fruit ripening. Three DE-lncRNAs, XLOC_011933, XLOC_001865, and XLOC_042291, are involved in UV-B-induced

anthocyanin biosynthesis and positively regulating *UVR8* and *COP10,* were identified and characterized. Our discovery and characterization of XLOC_011933, XLOC_001865, and XLOC_042291 provide a more precise understanding and preliminarily establishes a theoretical framework for UV-B-induced flesh anthocyanin biosynthesis. This phenomenon might encourage more in-depth investigations to study the molecular mechanisms underlying peach flesh coloring.

Keywords: blood-fleshed peach, lncRNAs, UV-B, anthocyanin biosynthesis, ripening

INTRODUCTION

Peach, a small deciduous tree, belongs to the order Rosales, family Rosaceae, subfamily Prunoidae, genus *Prunus*, and *Prunus persica* (L.) Batsch. According to the color of their flesh, there are three main types of peaches in the market, namely, white peach, yellow peach, and red peach. The book *Luoyanghuamuji* contained records of red peaches with blood color. The flesh peach first appeared in Europe in 1,659 (Hedrick, 1917). Accumulation of anthocyanins leads to attractive blood-fleshing in peaches and nectarines (He et al., 2015).

Anthocyanins, which generate characteristic reddish, bluish, and purple hues, are essential determinants of the color of many plant organs (Welch et al., 2008; Escribano-Bailón et al., 2004; Mano et al., 2007; Espley et al., 2007; Deluc et al., 2008). The anthocyanin content is an important indicator of ripening because many fruits and vegetables do not accumulate anthocyanins until they are ripening (Jimenez-Garcia et al., 2013). Anthocyanins also have potential human health benefits and represent a necessary aspect of fruit quality (Martin et al., 2011). The genetics and biochemistry of the anthocyanin accumulation-mediated fruit coloration and its biosynthetic pathway have been well studied. Different anthocyanins include various anthocyanidin aglycones. Among these, six anthocyanidins, Cy, Dp, Pg, Pn, Pt, and Mv, are generally found in most fruits (Macheix et al., 1990). In peach fruits, anthocyanins' predominant component is cyanidin-3-glucoside, with amounts of cyanidin-3-rutinoside. Many botanists have made significant contributions to the study of anthocyanin biosynthesis to develop new varieties with high anthocyanin content (Xie et al., 2011).

Anthocyanin biosynthetic pathways are well known to be conserved in many species (Grotewold, 2006; Lin-Wang et al., 2010; Lim and Ha, 2013). This well-known pathway acts via the phenylpropanoid pathway leading to the formation of anthocyanins, in which a series of enzymes are located on the endoplasmic reticulum's cytoplasmic surface, including *CHS*, *CHI*, *F3H*,*F3′H*,*F3′5′H*, *DFR*, *LDOX,* and *UFGT* (Holton and Cornish, 1995). Several studies have explained the pathway of anthocyanin biosynthesis in detail. However, it is essential to identify other critical players in such complicated regulatory networks.

A variety of environmental factors participate in and influence the biosynthesis process of anthocyanin, resulting in a series of changes in the anthocyanin content (Guo et al., 2008). As we all know, light is a major influencing factor in anthocyanin biosynthesis in the plant kingdom (Albert et al., 2009; Lin-Wang et al., 2011; Azuma et al., 2012; Butelli et al., 2012). Light can positively elevate the fruit anthocyanin content based on its characteristics, including specific light quality and light intensity photoperiod (Ubi et al., 2006; Li et al., 2013). Fruits produce anthocyanins to effectively scavenge the reactive oxygen species (ROS) produced in response to excess UV light. Anthocyanin enhancement in apple skin has been observed after UVB irradiation (Peng et al., 2013; Zoratti et al., 2014). Specific classes of plant photoreceptors, including *PHYs*, *CRYs*, *PHOTs,* and *UVR8*, allow plants to sense the presence of light and thus regulate the biosynthesis of secondary metabolites (Rizzini et al., 2011). Downstream signaling elements of photoreceptors, such as *COP1* and *HY5,* can also be further activated (Stracke et al., 2010; Lau and Deng, 2012). The gene of *COP1* is a key negative regulator of light signal transduction and participates in plant growth under light irradiation. *COP10* is, in a similar way, crucial for photomorphogenesis. *Arabidopsis AtCOP10* was found to be an inhibitor of the transcription factor HY5's COP1-mediated degradation within the nucleus (Osterlund et al., 2000). *HY5* is inconsistent with the two factors mentioned previously and is considered to be a positive regulatory gene with light involvement (Lee et al., 2007). *HY5* is the main gene controlled by *COP1* in dark environments (Osterlund et al., 2000). The expression of *HYH*, which changes the same with the expression of the anthocyanin biosynthesis pathway's regulation of structural genes, is correlated with the anthocyanin content (Zhao et al., 2017). A total of 12 light receptors (*UVR8s*, eight LIGHT-DEPENDENT SHORT HYPOCOTYLS) and four constitutive photomorphogenesis proteins (*COP*) were derived by automated computational analysis in the peach genome.

Long non-coding RNAs (lncRNAs) are characterized by a transcription length of more than 200 nt and not coding proteins (Ma et al., 2013). LncRNAs play essential regulatory roles in various biological processes, such as developmental and environmental regulation (Liu et al., 2012; Wen et al., 2007; Wang M. et al., 2015; Wang T.-Z. et al., 2015; Xin et al., 2011; Li et al., 2007; Boerner and McGinnis, 2012; Tang et al., 2016), *Hippophae rhamnoides* Linn (Zhang et al., 2017), and *Populus* (Liu et al., 2018). LncRNAs were expressed during pollen tube germination and pollen tube growth in plants (Kim and Sung, 2012). Increasing evidence suggests that lncRNAs play essential roles in regulating secondary metabolism (An et al., 2018; Yin et al., 2018). However, the profiles of lncRNAs in fruit trees are not clear. Many research studies have shown that lncRNAs participate in fruit ripening. Another study on tomatoes implicated the silencing of lncRNA1459 and lncRNA 1840,

leading to a conspicuous delay in the wild-type fruit's ripening (Wang et al., 2018). LncRNA1459 functional loss mutant produced by the Cas9 gene-editing technique showed that the tomato ripening process, ethylene production, and lycopene accumulation were primarily repressed, and the expression of the ripening-related genes was significantly altered, providing clues for understanding the function of lncRNA1459 in the ripening of tomato fruit (Li et al., 2014; Zhu et al., 2015). In apples, eight kinds of lncRNAs associated with fruits were tested and found to regulate fruit ripening and glucose metabolism (An N. 2018). RNA sequencing analysis identified DE-lncRNAs in the red and yellow sea buckthorns' fruit, providing a resource for the in-depth study of fruit coloring in ripe fruits (Zhang et al., 2017). In kiwi fruits, lncRNAs associated with fruit development and ripening have been identified (Tang et al., 2016).

Although the role of lncRNAs in various biomolecular processes has been gradually discovered, the regulatory role of lncRNAs in Rosaceae trees is still poorly understood, particularly in peaches. Peaches (*Prunus persica*) are economically important fruit trees with a short maturation phase and a sequenced genome ($2n$ = 16, 225.7 Mb) (Verde et al., 2017). The regulatory mechanism of anthocyanins in peaches with the blood-flesh phenotype has been extensively researched, combining high-quality sequenced genomes. The key genes that regulate the phenotype of the flesh are well known, including *MYB, bHLH, WD40,* and TFs that form the *MYB-bHLH* or *MYB-bHLH-WD40* (*MBW*) complex. However, whether there are other regulatory factors affecting anthocyanin biosynthesis remains to be researched. The present study analyzed blood-flesh transcriptomes of peach fruits at different developmental stages to screen for lncRNAs related to anthocyanin biosynthesis to further understand the regulatory network of the blood-flesh phenotype in peach flesh. In total, 17,456 lncRNAs were identified in the blood-fleshed peach fruit transcriptome dataset. Expression correlation was used between mRNAs and lncRNAs in the peach reference genome. Both positive and negative lncRNA–mRNA pairs were identified. XLOC_011933, XLOC_001865, and XLOC_042291, which are involved in UV-B-induced anthocyanin biosynthesis and positively regulate *UVR8* and *COP10* (constitutive photomorphogenic 10), were identified and characterized. Our investigation served as a reference for later studies exploring the function of lncRNAs in red peaches.

MATERIALS AND METHODS

Sample Collection, RNA Quantification, and Qualification

The blood-fleshed hybrid lines of the *Prunus persica* (L.) Batsch "C3-20" seedling was grown in the experimental station of Hebei Normal University of Science and Technology, Hebei Province, China. Ovary samples were collected 10 days before and 10 days after pollination. The samples were collected from some 5-year-old peaches and designated as follows: 34 DAP, 44 DAP, 54 DAP, 64 DAP, and 74 DAP (DAP refers to days after pollination). The samples we collected, including fruit and ovary, are wrapped in tin foil and stored in a −80°C refrigerator, which is convenient for RNA extraction and physiological indicator measurement.

RNA Isolation, Library Preparation, and Sequencing

Total RNA was isolated by applying the TRIzol reagent (Invitrogen, Carlsbad, CA, United States). The RNA quality was determined by agarose gel electrophoresis. The integrity of the total RNA was also determined. In addition, each sample was constructed with 3 μg RNA. First, ribosomal RNA was removed using a kit (epicenter, United States). Subsequently, the NEBNext®Ultra™ Directed RNA Library preparation kit from Illumina® (NEB, United States) was used.

RNA-Seq Read Processing, Mapping, and Transcriptome Assembly

The authenticity of the quality of sequencing results was checked. We used HISAT2 for the comparative analysis of the reference genome of the filtered reads (Langmead and Salzberg, 2012, 2). StringTie used the results of HISAT2 alignment to splice transcripts, resulting in the smallest possible set of transcripts (Pertea et al., 2016). After screening the lncRNA and TUCP transcripts, the StringTie-EB software was used for the quantitative analysis of the transcripts including mRNA, lncRNA, and TUCP (Trapnell et al., 2010; Pertea et al., 2016).

Identification of lncRNAs

In the initial stage, Cuffmerge software was used to merge the transcripts obtained (Trapnell et al., 2010). Then, the combined transcript set was screened for lncRNA: in the first step, the exon number of transcripts greater than or equal to 2 was selected; in the second step, the length of the transcripts >200 bp was retained; the third step is the screening of known annotations in the transcript; the fourth step is to select transcripts with FPKM greater than or equal to 0.5; and the last step is the screening of coding potential (Kong et al., 2007; Cabili et al., 2011; Lin et al., 2011; Punta, et al., 2012; Sun et al., 2013).

Target Gene Prediction

LncRNAs can interact with target genes in either *cis* or *trans* form, so we adopted two methods to predict the *target* genes of lncRNAs. The *cis* role shows the lncRNA's effect on its neighboring target genes. The coding genes were screened in the 10 kb upstream and downstream regions of lncRNAs. The *trans* role indicates the ability of the lncRNAs to identify each other based on individual expression levels. As the number of samples was less than 25, the Pearson correlation coefficient was used to further analyze the correlation between lncRNAs and the expression level of protein-coding genes. When the absolute value of the correlation coefficient was greater than 0.95, we believed that there was trans-interaction between protein-coding genes and lncRNAs.

GO and KEGG Enrichment Analysis

The GO enrichment analysis method was GOseq and it was performed using the GOseq R package (Young et al., 2010). KEGG is a major public database related to pathways. The pathway significance enrichment analysis was conducted with the KEGG pathway as the unit (Mao et al., 2005). The KEGG (http://

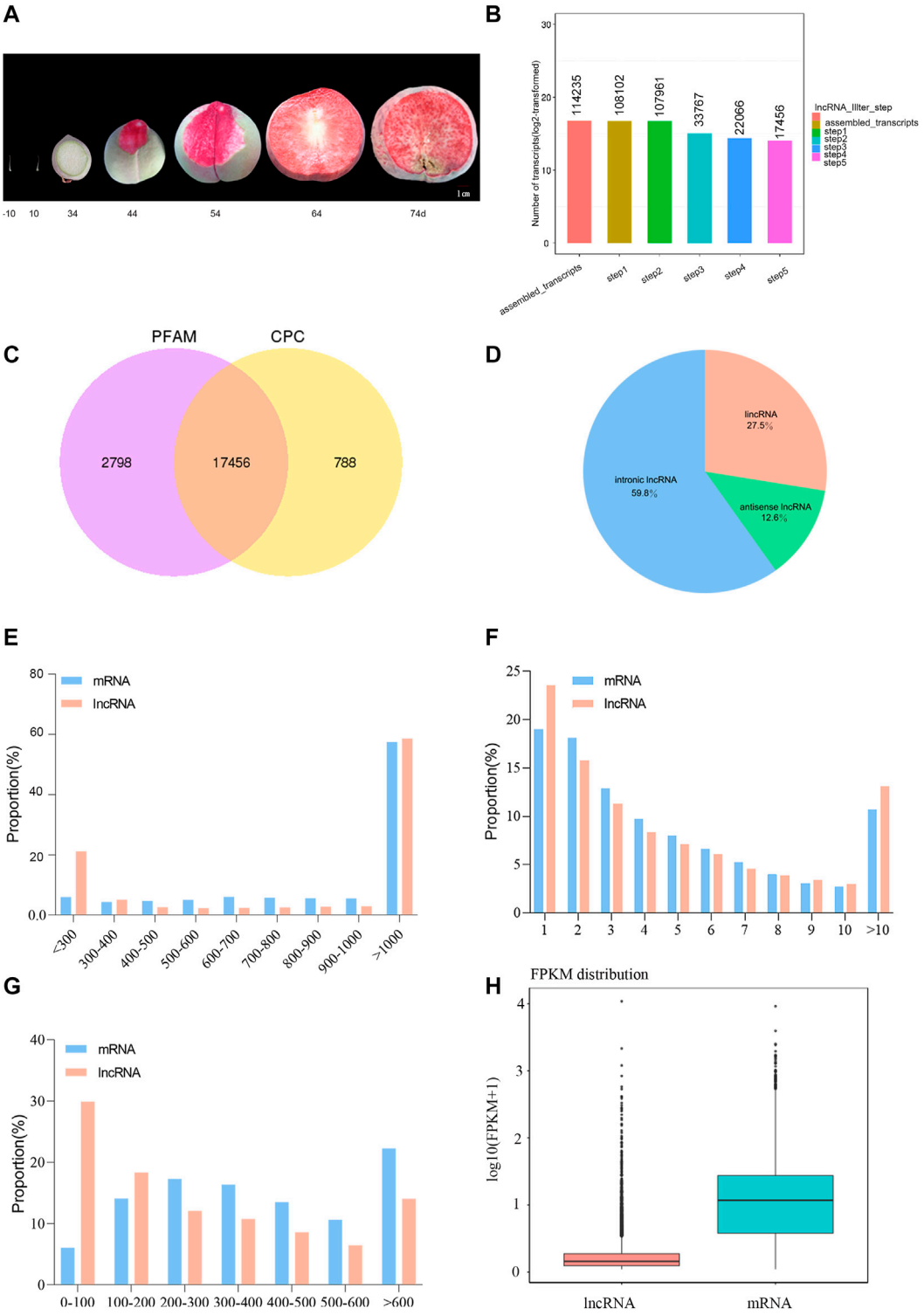


FIGURE 1 | General information on the lncRNAs in peach. **(A)** Phenogram of the peach. **(B)** Basic screening charts for lncRNAs. **(C)** Screening results using CPC and Pfam. **(D)** Composition of different types of lncRNAs. **(E)** Length distribution of peach lncRNAs and mRNAs. **(F)** Distribution of some exons in lncRNAs and mRNAs. **(G)** Distribution of the ORFs in lncRNAs and mRNAs. **(H)** Comparison of expression levels between lncRNAs and mRNAs.

TABLE 1 | Sequencing quality.

Sample name	Raw reads	Clean reads	Percent (%)	Error rate (%)	Clean bases (G)	Q20 (%)	Q30 (%)	GC content (%)
S1_11	103,290,726	100,842,110	97.63	0.02	15.13	97.14	92.75	44.11
S1_12	107,939,668	105,585,708	97.82	0.01	15.84	97.51	93.83	44.45
S1_13	98,090,490	96,262,152	98.14	0.01	14.44	98.03	94.94	44.1
S2_21	105,311,206	103,705,758	98.48	0.01	15.56	97.89	94.63	43.99
S2_22	115,876,152	113,957,100	98.34	0.01	17.09	97.9	94.64	44.16
S2_23	97,715,896	95,926,382	98.17	0.01	14.39	97.99	94.83	44.29
S3_31	104,871,196	102,758,664	97.99	0.01	15.41	97.91	94.66	44.38
S3_32	86,772,210	83,870,212	96.66	0.01	12.58	97.98	94.74	44.36
S3_33	99,187,418	96,244,550	97.03	0.01	14.44	97.96	94.68	44.44
S4_41	110,074,586	106,839,978	97.06	0.01	16.03	98.09	94.97	44.57
S4_42	106,348,746	103,571,376	97.39	0.01	15.54	98	94.77	44.69
S4_43	100,381,514	97,653,486	97.28	0.01	14.65	97.94	94.65	44.67

www.genome.jp/kegg/) enrichment analysis was performed by KOBAS (2.0) software. The interactions between differential lncRNAs and mRNAs were analyzed using Cytoscape software to construct a correlation network diagram (Saito et al., 2012).

Construction of a DE-lncRNA–mRNA Network

The networks of the lncRNA–mRNAs were to elucidate the functions of lncRNAs. Based on the interaction relations in the STRING protein interaction database (http://string-db.org/), the interaction relations of target gene sets (such as the differential gene list) were directly extracted from the database to construct the network for the species contained in the database.

Determination of the Total Anthocyanin Content

We used methanol supplemented with 1% HCl from the plant material. The samples were ground in N2 solution and cultured overnight. Then, after centrifugation for 15 min at 16,000 × *g*, the absorption value of the material at 530 and 657 nm was determined by using a spectrophotometer. Finally, the following formula was used to calculate the anthocyanin content in each period: $Q_{anthocyanins} = (A_{530} - 0.25 \times A_{657}) \times M^{-1}$; $Q_{anthocyanins}$ content of anthocyanins, A_{530} absorption at 530 nm wavelength, A_{657} absorption at 530 nm wavelength, and *M* fresh weight (g) of the tissues (Mehrtens et al., 2005).

Quantitative Real-Time (qRT)-PCR

TransStart Tip Green qPCR SuperMix (Transgen) and CFX Connect were used for qRT-PCR (**Supplementary Table S6**). The reference gene for qRT-PCR was the peach *RPL13* (ribosomal protein L13) gene. The results were unified using the $2^{-\Delta\Delta t}$ calculation method.

Statistical Analysis

All experiments were set up as three replicates, and each data were represented by error lines. All data in this experiment were plotted and analyzed by GraphPad Prism 9, where $p < 0.05$ was * and $p < 0.01$ was ***. The online website (http://www.bioinformatics.com.cn) drawing heatmaps and enrichment analysis diagrams was used.

RESULTS

Library Construction and RNA Sequencing of Different Development Stages of Blood-Fleshed Peach

Ovaries were collected 10 days before pollination and 10 DAP, and peaches were harvested at 34, 44, 54, 64, and 74 DAP (**Figure 1A**). The percentage of clean reads ranged from 96.66–98.48% (**Table 1**). The Q20 and Q30 values were >97 and 92%, respectively, which proved that the quality control data are reliable. The GC content ranged between 43.99 and 44.69%. We could clearly observe that the mapping rate of clean reads was 80.62–90.69% and most of the clean reads (77.7–87.6%) were distributed in the protein-coding region (**Supplementary Tables S1–S2**). A total of 114,235 assembled transcripts were used to screen for lncRNAs. After five basic screening steps (described in **Section 2.4**), 17,456 transcripts were retained, which were used to analyze protein-coding potential using CPC and PFAM (PfamScan) (**Figures 1B,C**). According to their locations in the *P. persica* genome, 4,800 lincRNAs (27.5%), 2,199 antisense lncRNAs (12.6%), and 10,439 intronic lncRNAs (59.8%) were identified (**Figure 1D**). The composition of different lncRNAs is different from that of *Populus* lncRNAs (Liu et al., 2018). The structural analysis of lncRNAs and mRNAs indicated that those in peach fell in the length range of 201–20,369 and 3–15,738 nt, respectively, with corresponding averages of 1,623 and 1,321 nt, respectively. The average transcript length of the lncRNAs was slightly more than that of the mRNAs, although the difference was not as significant as that in poplar (Liu et al., 2018). Most lncRNAs and mRNAs were >1,000 bp long (58.59 and 57.41%, respectively). The distribution of exon numbers showed a similar trend for the lncRNAs and mRNAs (**Figure 1F**). For example, 23.51% of the lncRNAs had one, 15.79% had two, 11.31% had three, 19.02% had

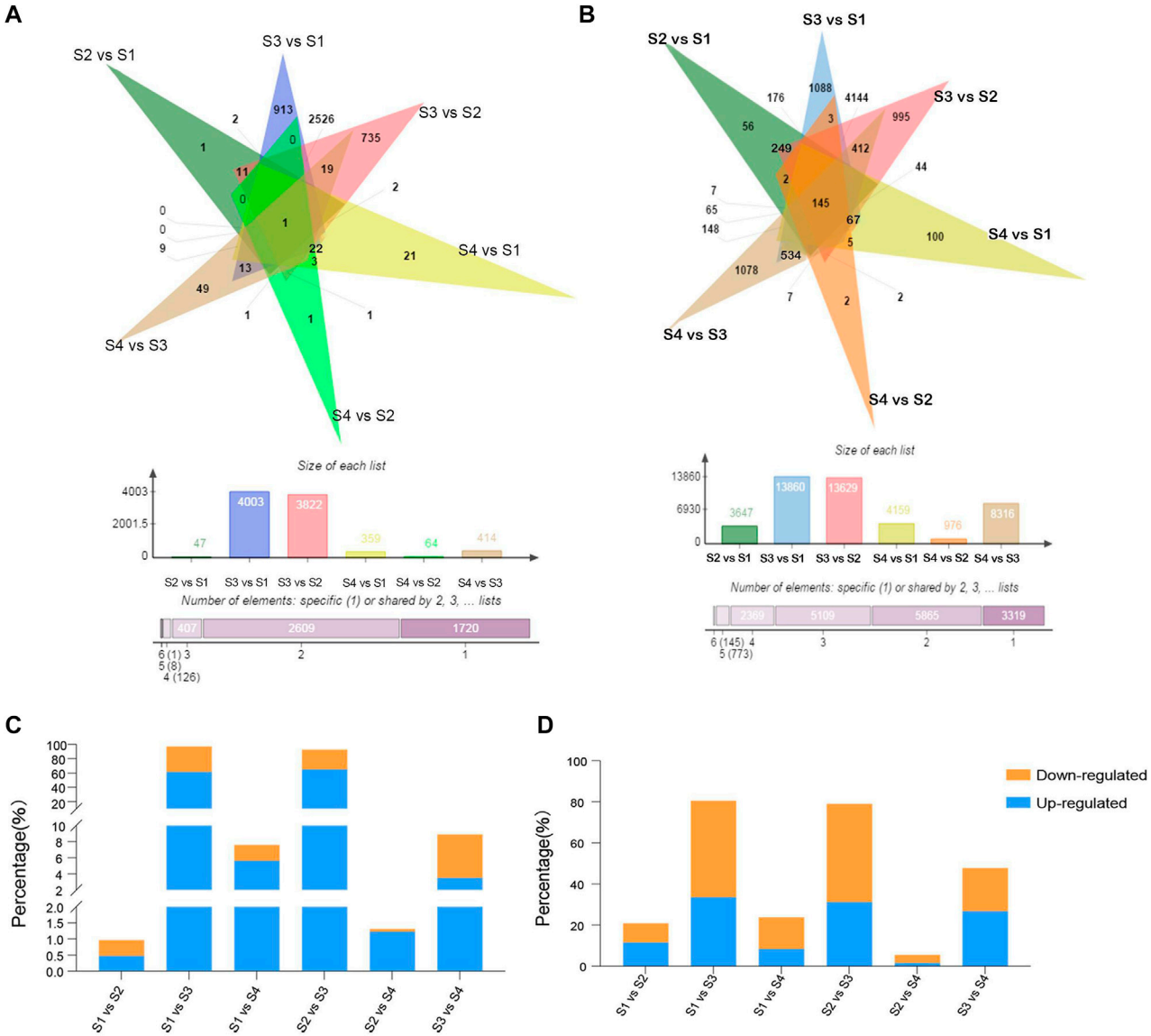


FIGURE 2 | Analysis and identification of differentially expressed lncRNAs (DELs). **(A)** Venn diagram of common and specific DELs. **(B)** Venn diagram of common and specific differentially expressed mRNAs. **(C,D)** Percentages of DEGs and DELs.

one, 18.08% had two, and 12.88% had three exons (**Figure 1F**). The open reading frames (ORFs) of lncRNAs were 22–5,114 nucleotides (nt) in length, of which the majority (29.89%) were ≤100 nt. The ORFs of the mRNAs were 1–5,245 in length, with the majority accounting for 22.64%, being ≥600 nt (**Figure 1G**). The expression level of most lncRNAs was relatively low (**Figure 1H**).

Differentially Expressed lncRNAs

To identify lncRNAs in different developmental stages of blood-fleshed peaches, six comparison groups were analyzed: S1 versus S2, S1 versus S3, S1 versus S4, S2 versus S3, S2 versus S4, and S3 versus S4. Differentially expressed lncRNAs (DELs) were identified for each comparison group and the number of upregulated and downregulated DELs were shown (**Supplementary Figures S1A–F**). In summary, we obtained 4,871 DELs, of which one was common to these groups (**Figure 2A**), and 17,580 DEGs, of which 145 were common to the six comparison groups (**Figure 2B**). Moreover, DEGs were more than DELs in six comparisons, and the percentages of DEGs and DELs in S3 versus S1 and S3 versus S2 were higher (**Figures 2C,D**).

Specifically, 4,515 DELs were differentially expressed in fruits at S3 versus S2, of which 3,168 were upregulated and 1,347 were downregulated. A comparison of the fruits at S4 and S3 revealed significant variations in the expression. A total of 434 DELs,

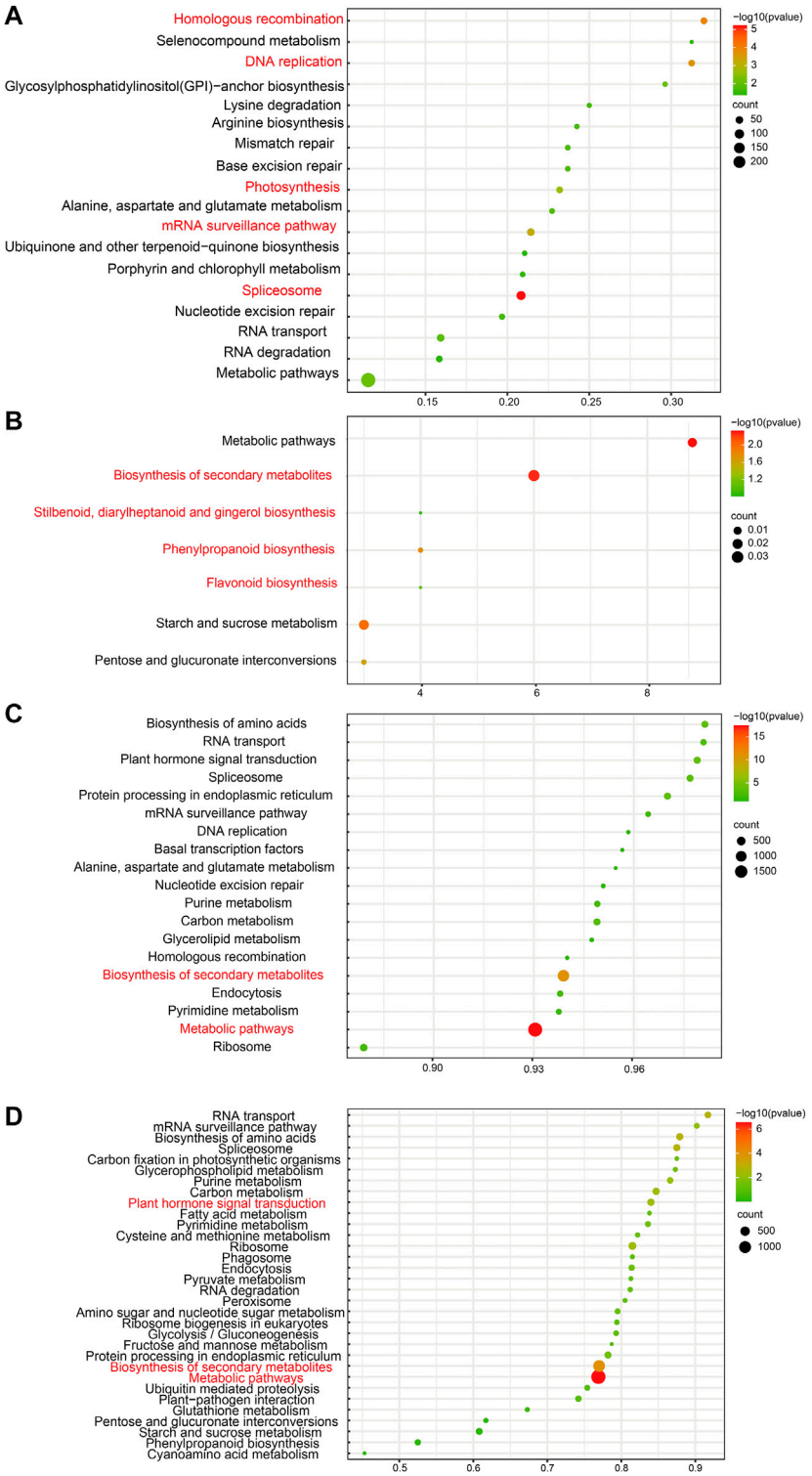


FIGURE 3 | Common and unique enriched KEGG pathways of four developmental stages. **(A)** Common enriched KEGG pathways of four stages with trans-regulation function. **(B)** Common enriched KEGG pathways of four stages with cis-regulation function. **(C)** Unique enriched GO terms of four stages with trans-regulation function. **(D)** Unique enriched GO terms of four stages with cis-regulation function.

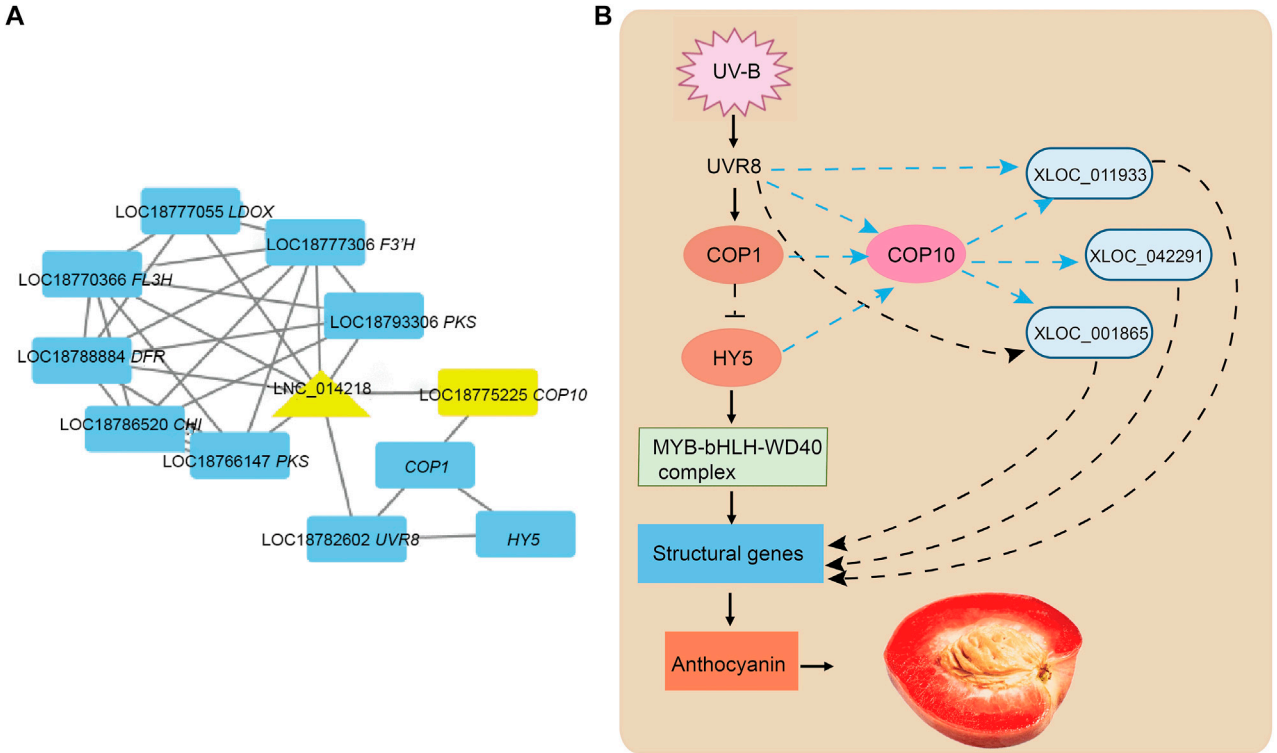


FIGURE 4 | Analysis of XLOC_042291, XLOC_011933, and XLOC_001865 involved lncRNA–mRNA networks associated with UV-B induced in anthocyanin biosynthesis. **(A)** Sub-interaction networks of XLOC_042291, XLOC_011933, and XLOC_001865 from lncRNA–mRNA networks. The triangle represents lncRNAs and the rectangle represents mRNAs. **(B)** Model diagram of UV-light-mediated fruit development of red meat peach. *COP1*, constitutive photomorphogenic 1; *COP10*, constitutive photomorphogenic 10; *HY5*, ELONGATED HYPOCOTYL 5; *UVR8*, UV RESISTANCE LOCUS 8; Ub, ubiquitin.

comprising 170 upregulated and 264 downregulated genes, were identified in the fruits at S3 (**Supplementary Table S3**).

The number of shared and exclusive DELs between the different developmental stages of the blood-fleshed peach is shown by the Venn diagram (**Figure 3A**). Only one of the DELs was expressed in all developmental stages, while one (2.1%) DEL was exclusively expressed between the S2 and S1 stages, 913 (30.5%) between the S3 and S1 stages, 21 (8.7%) between the S4 and S1 stages, 735 (23.2%) between stages S3 and S2, 1 (1.7%) between stages S4 and S2, and 49 (28.9%) between stages S4 and S3. S3 versus S1 stages had the highest number of DELs.

We mapped these lncRNAs and their target transcript mRNAs onto eight chromosomes of the peach genome and found that some lncRNAs were produced at the ends of chromosomes #6 and #3 (**Supplementary Figure S2**).

Expression Correlation of DE-lncRNAs With *Cis*-acting Protein-Coding Genes

We predicted the potential function of lncRNAs in peach fruit to determine whether protein-coding genes could interact with lncRNAs in *cis* or *trans* configurations. Several DE-lncRNAs have been shown to regulate the expression of genes in close proximity (*cis*-acting) or at a distance (*trans*-acting). In total, 67,232 lncRNA–mRNA pairs were co-localized, with 25,846 lncRNAs upstream and 27,008 lncRNAs downstream of mRNAs, 14,375 antisense lncRNAs, and three sense overlapping (10 kb) with *cis* lncRNA targets that were mainly represented in the developmental process. The most relevant GO terms that were associated with unique biological processes and molecular functions contained many key enzymes and important response regulators, consistent with the steps involved in fruit development stages (**Supplementary Table S5**). These analyses identified *cis*-acting lncRNAs with potential regulatory roles.

The main KEGG pathways of *cis*-target genes associated with fruit development include stilbenoid, diarylheptanoid, gingerol, flavonoid, phenylpropanoid, and secondary metabolite biosynthesis, which were enriched with high reliability (**Figure 3B**).

Expression Correlation of DE-lncRNAs With *Trans*-acting Protein-Coding Genes

We also found the lncRNAs and mRNAs with *trans*-regulatory relationships from our analyses. A total of 1,048,575 lncRNA–mRNA were co-expressed and 225,537 and 823,038 lncRNAs were positively and negatively correlated with that of their target mRNAs, respectively.

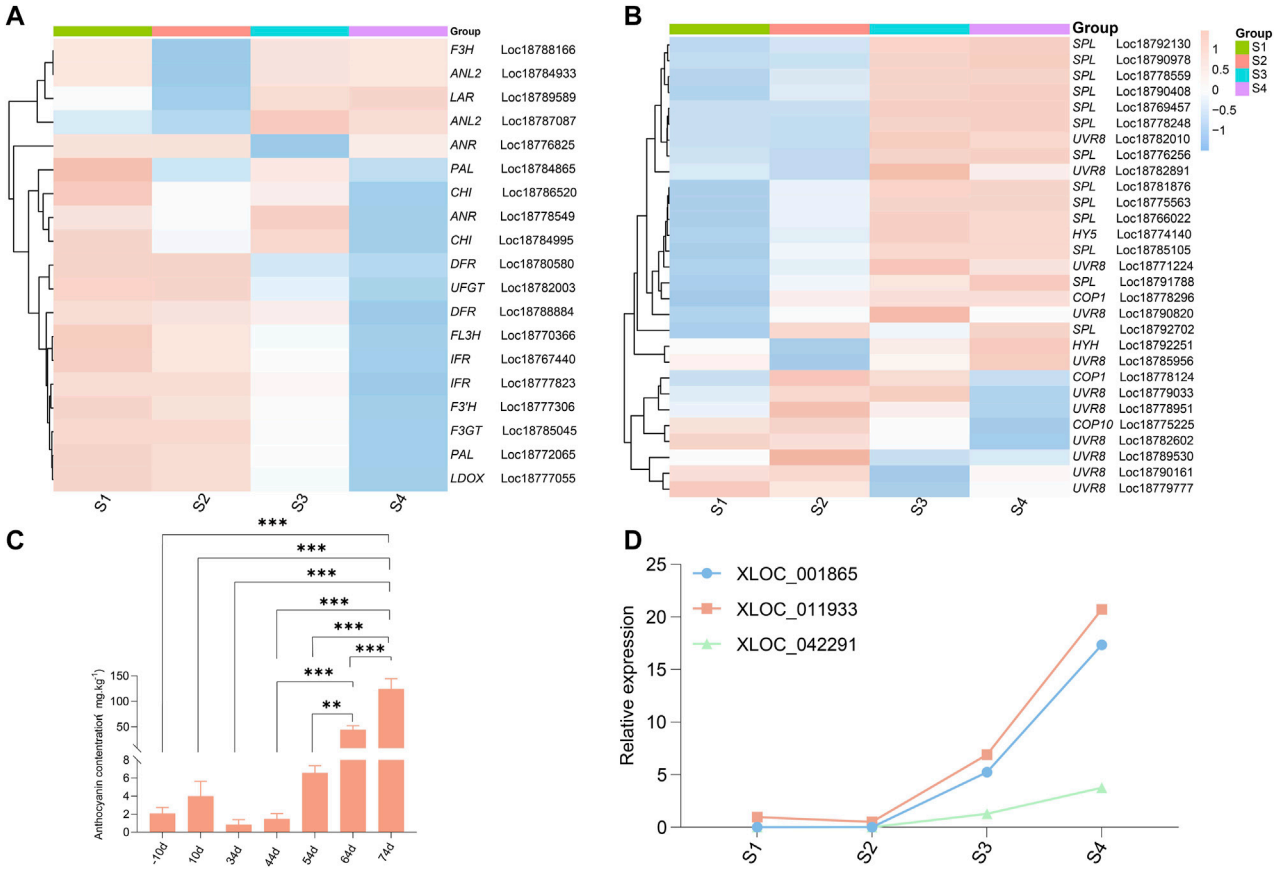


FIGURE 5 | Heatmap and graphical presentations of selected genes during fruit development stages. **(A)** Expression profiles of structural genes involved in the anthocyanin biosynthesis pathway. The color gradient with eight different colors from blue to red corresponds to transcript levels from low to high, with the values representing the log2 (FPKM) values. **(B)** Expression profiles of genes related to UV-light-mediated anthocyanin biosynthesis. The color key at the top of the heatmap represents the relative color scheme of the FPKM values. **(C)** Anthocyanin content in different development stages. **(D)** Graphical representations of XLOC_011933, XLOC_001865, and XLOC_042291 related to anthocyanin biosynthesis based on RNA-seq values in blood-fleshed peach.

Our functional prediction is based on the GO-term enrichment of *trans* lncRNA targets for biological processes, cellular components, and molecular functions. We compared common and uniquely enriched GO terms in the four stages (**Supplementary Table S4**). These results suggested that the DE-lncRNAs' *trans* target genes are involved in a lot of biological processes, such as metabolic processes, cellular processes, and biological regulation, and a variety of molecular functions, such as catalytic activity, binding, transporter activity, and molecular function regulation. These results also suggested that lncRNAs can play important roles in transport and regulation. We identified a series of target genes involved in UV-B-induced anthocyanin biosynthesis in blood-fleshed peaches based on pathway analysis.

Moreover, the KEGG pathway analysis showed that these *trans* target genes of lncRNAs were enriched in spliceosomes, homologous recombination, DNA replication, mRNA surveillance, and photosynthesis pathways (**Figure 3A**). It indicated that lncRNAs take an essential part in RNA splicing and photosynthesis pathways during fruit development and UV-B-induced anthocyanin biosynthesis in the flesh. This phenomenon illustrated the capability of lncRNAs to affect photosynthesis, which is predominantly involved in fruit development and flesh-coloring processes.

Enriched GO terms of the *cis*- and *trans*-target genes of unique DE-lncRNAs in each of the four development stages differed in the number of involved genes. Still, the enrichment mainly showed some biological processes (developmental, cellular, metabolic, and organic substances) and molecular functions, such as catalytic activity and binding. In addition, the number of genes in *cis*-and *trans*-target unique gene GO terms was roughly similar. Unique DE-lncRNAs' trans-target genes enrich KEGG pathways at each stage of development, indicating that there is the highest reliability of metabolic pathways and biosynthesis of secondary metabolites (**Figure 3C**). Unique DE-lncRNAs' cis-target genes are enriched in KEGG pathways at each developmental stage, suggesting that metabolic pathways, secondary metabolite biosynthesis, and plant hormone signal transduction have the highest reliability (**Figure 3D**). These results indicated

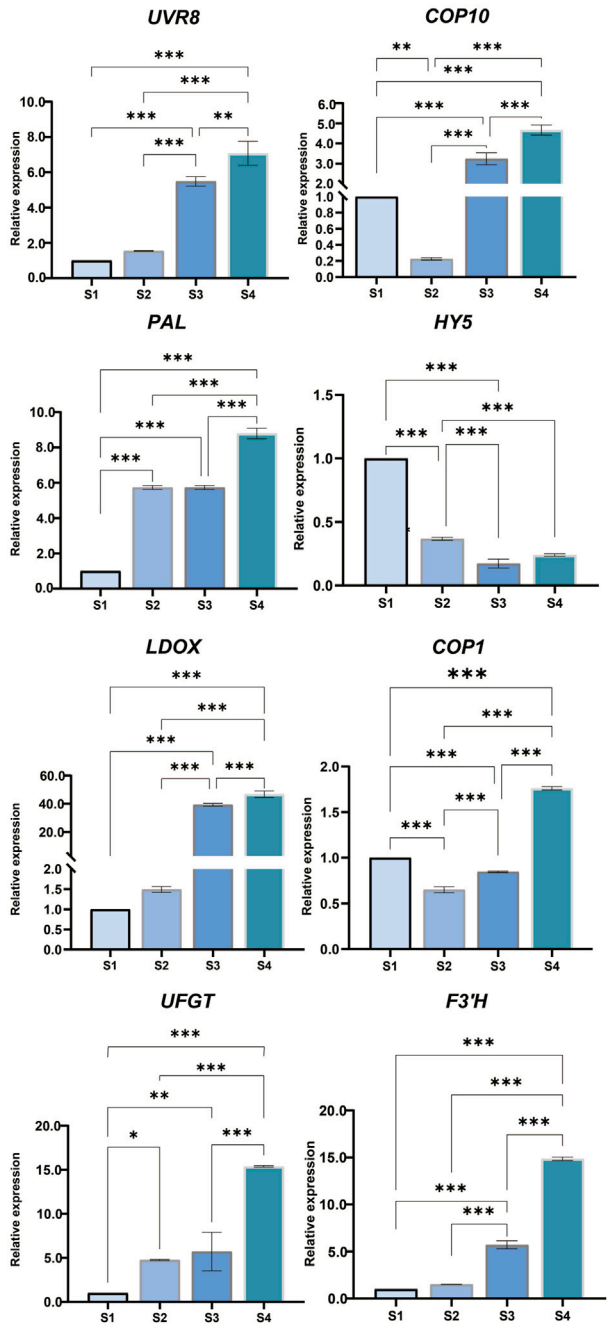


FIGURE 6 | Quantitative real-time (qRT)-PCR was used to determine anthocyanin biosynthesis-related genes *UVR8* (Loc18782602), *UFGT* (Loc18782003), *PAL* (Loc18772065), *LDOX* (Loc18777055), *HY5* (Loc18774140), *F3 'H* (Loc18777306), *COP10* (Loc18775225), and *COP1* (Loc18778124) in peach fruit. The error bar represents the mean + standard error of three repeated measurements (SE) using one-way analysis of variance (ANOVA). $p < 0.05$ is represented by * and $p < 0.001$ is represented by ***.

that unique DE-lncRNAs in each developmental stage may play different roles but are involved in the same fruit development processes.

DE-lncRNAs Participate in lncRNA–mRNA Interactive Networks

Next, we elucidated the function of DE-lncRNAs and the relationship between lncRNAs and mRNAs that were co-expressed and fell <10 kb away from DE-lncRNAs by establishing putative interactive networks using Cytoscape. Three anthocyanin biosynthesis-related lncRNAs (XLOC_011933, XLOC_001865, and XLOC_042291) were identified for further analysis (**Figure 4A**). The transcriptome data of identified lncRNA genes and transcription factors were analyzed. A model diagram on the UV-B regulation of peach anthocyanin synthesis was presented (**Figure 4B**). *COP1* and *COP10* are involved in the degradation of downstream genes, including members of the *HY5*, *HYH,* and *MBW* complex. The identified lncRNAs interacted with some structural genes as critical regulators playing particular roles in UV-B-induced anthocyanin biosynthesis. The results indicated that the three lncRNAs could affect biological processes at different levels.

Analysis of lncRNAs Related to the UV-B-Induced Anthocyanin Biosynthesis of Fruits

We also investigated the cross-talk among lncRNAs, *COP10,* and *UVR8* in UV-B-induced anthocyanin accumulation in the blood-fleshed peach fruit development process. Based on the results of preliminary GO, KEGG, and functional analyses, the expression of the three identified lncRNAs XLOC_011,933, XLOC_001,865, and XLOC_042,291, and some structural genes, such as *PAL, DFR, FL3H, F3GT, LDOX,* and *UFGT,* and plant photoreceptors *UVR8* (UVB photoreceptor), downstream signal elements *COP10, COP1, HY5, HYH* (*HY5* homolog), and *SPL* were analyzed. The results revealed that most structural genes showed similar expression patterns, corroborating the accumulation of anthocyanins, except for *ANL2* (Loc18784933 and Loc18787087), *LAR* (Loc18789589), and *F3H* (Loc18788166) (**Figure 5A**). For the genes that regulate light signaling, only *COP10* (Loc18775225) and *UVR8* (Loc18782602) showed similar expression patterns with the accumulation of anthocyanins, suggesting a relation between their regulation and respective functions (**Figure 5B**).

Our analyses revealed 12 light receptors (*UVR8s*, eight LIGHT-DEPENDENT SHORT HYPOCOTYLS, and four constitutive photomorphogenesis proteins (*COP*)), as derived using automated computational analysis in the peach genome. *UVR8s* and *SPLs* were distributed on different chromosomes and displayed different expression patterns during the four fruit development stages (**Figures 5A,B**). This observation implied that these photoreceptors and signal elements have different functions among molecules, although they belong to the same gene family.

These results suggested lncRNAs' involvement in regulating anthocyanin biosynthesis and fruit flesh pigment in developing blood-fleshed fruits.

TABLE 2 | XLOC_011933, XLOC_001865, and XLOC_042291 regulate anthocyanin biosynthesis in the fruit development of blood-fleshed peach.

lncRNA_ID	mRNA_ID	Predicted product abbreviation	Predicted product	*p*-value
LNC_000563	XM_007202045.2	Naringenin,2-oxoglutarate 3-dioxygenase	*FL3H*	1.21E-07
LNC_000563	XM_007205913.2	Constitutive photomorphogenesis protein 10	*COP10*	2.46E-09
LNC_000563	XM_007211190.2	MYB4	*MYB4*	1.55E-06
LNC_000563	XM_007216787.2	UDP-glucose flavonoid 3-O-glucosyltransferase 3	*UFGT*	5.48E-08
LNC_000563	XM_007217820.2	Ultraviolet-B receptor UVR8	*UVR8*	1.28E-07
LNC_000563	XM_007222255.2	Bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase	*DFR*	3.05E-08
LNC_000563	XM_007208370.2	Phenylalanine ammonia-lyase 1	*PAL*	1.29E-11
LNC_003950	XM_007202045.2	Naringenin,2-oxoglutarate 3-dioxygenase	*FL3H*	1.57E-07
LNC_003950	XM_007205236.2	*MYB5*	*MYB5*	6.47E-07
LNC_003950	XM_007205913.2	Constitutive photomorphogenesis protein 10	*COP10*	1.48E-10
LNC_003950	XM_007208370.2	Phenylalanine ammonia-lyase	*PAL*	5.34E-11
LNC_003950	XM_007210458.2	Leucoanthocyanidin dioxygenase	*LDOX*	8.26E-10
LNC_003950	XM_007215227.2	UDP-glucose flavonoid 3-O-glucosyltransferase 3	*UFGT*	2.21E-06
LNC_003950	XM_007217820.2	Ultraviolet-B receptor UVR8	*UVR8*	5.80E-07
LNC_003950	XM_007217890.2	Anthocyanidin 3-O-glucosyltransferase 2	*F3GT*	3.57E-10
LNC_003950	XM_007222255.2	Bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase	*DFR*	3.84E-08
LNC_014218	XM_007202045.2	Naringenin,2-oxoglutarate 3-dioxygenase	*FL3H*	5.85E-08
LNC_014218	XM_007205913.2	Constitutive photomorphogenesis protein 10	*COP10*	2.32E-09
LNC_014218	XM_007208370.2	Phenylalanine ammonia-lyase	*PAL*	1.53E-12
LNC_014218	XM_007210458.2	Leucoanthocyanidin dioxygenase	*LDOX*	1.48E-10
LNC_014218	XM_007215227.2	UDP-glucose flavonoid 3-O-glucosyltransferase 3	*UFGT*	7.08E-07
LNC_014218	XM_007217820.2	Ultraviolet-B receptor UVR8	*UVR8*	1.12E-07
LNC_014218	XM_007217890.2	Anthocyanidin 3-O-glucosyltransferase 2	*F3GT*	1.95E-12
LNC_014218	XM_007222255.2	Bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase	*DFR*	1.34E-07

The italics refer to the target genes predicted by the three DE-LNcRNAs (XLOC_001865, XLOC_011933, and XLOC_042291).

Examination of the Evolutionary Conservation of lncRNAs

We hypothesized that if lncRNAs are evolutionarily conserved, they could perform similar functions in different species even without coding regions (Kang and Liu, 2015). Naturally, we tested whether the three lncRNAs found in this study are evolutionarily conserved in different plant species. NCBI BLAST analysis of the three lncRNA sequences revealed that XLOC_001865 shares high protein sequence similarities with *Prunus persica* aquaporin *TIP1-2* (LOC18767586), *Prunus mume* aquaporin *TIP1-2-like* (LOC103335008), and *Prunus avium* aquaporin *PIP-type-like* (LOC110757218). Two conserved domains of XLOC_001865 (507 bp), one of which belongs to the major intrinsic protein (*MIP*) superfamily, were conserved with the aforementioned three proteins, implying that this region may have potentially conserved counterparts in the Rosaceae species. Analysis of the gene sequence of XLOC_011933 (4,272 bp) revealed that a 382 bp fragment shared 100% identity with the predicted *Prunus avium* chitinase 2-like (LOC110760813) DNA sequence and revealed an 868 bp repeat sequence at the beginning and interval of the lncRNA sequence. The BLAST analysis revealed that XLOC_042291 (1945 bp) shared no similarity with any known proteins or DNA sequences but with some uncharacterized mRNA or ncRNA in the peach genome. This analysis failed to determine whether the homologous sequences in the other species encoded lncRNAs, suggesting limited evolutionary conservation of the lncRNAs.

Reverse-Transcription Quantitative PCR

The expression trends of some genes, including *COP10* (LOC18775225), *HY5* (LOC18774140), *UVR8* (LOC18782602), *COP1* (LOC18778124), *PAL* (LOC18772065), *F3′H* (LOC18777306), *UFGT* (LOC18782003), and *LDOX* (LOC18777055), were examined. These results were consistent with the trend of FPKM (**Figure 6**).

DISCUSSION

Although extensive studies have described the physiological and molecular aspects of peach development and ripening, few studies have focused on lncRNA-based molecular regulation in controlling anthocyanin biosynthesis and flesh coloration during peach fruit development. In this study, we have attempted to address such missing information by exploring peach development and ripening aspects based on lncRNA-associated mechanisms. No lncRNA has been described as playing a role in peach development; therefore, lncRNAs are associated with flesh color. We performed genome-wide investigations based on sequencing to identify lncRNAs playing a role in *Prunus persica*, thereby providing a new perspective for studying the regulation mechanism of non-coding genes regulatory mechanisms in peach genomes.

Through transcriptome data analysis, stringent filtering criteria mining transcriptome datasets of these lncRNAs were applied. A total of 4,871 differentially expressed lncRNAs and 17,580 DEGs were screened. Also, regulating target genes in two

ways such as *trans* and *cis* could provide valuable methods for identifying the processes involved in lncRNAs and inferring their potential functions. Increasing evidence points out that lncRNAs play indispensable roles in plant responses and the regulation of secondary metabolism (An et al., 2018; Yin et al., 2018). RNA-seq analysis of the blood-fleshed peach revealed numerous lncRNAs involved in various biological processes. In this examination, potential *cis* and *trans* lncRNA target genes and their functions were predicted and analyzed by us according to their relationship with mRNA expression. They were found to be involved in many processes, including flavonoid and anthocyanin biosynthetic pathways.

Studies have shown that the synthesis process of anthocyanin from phenylalanine is complicated, from phenylalanine decomposition to anthocyanin biosynthesis (Jaakola, 2013). It has been reported in many works (Shi and Xie, 2014). We identified three DE-lncRNAs that were significantly involved in anthocyanin biosynthesis pathways. As shown in **Table 2**, XLOC_001,865 was predicted to target *FL3H, COP10, UFGT, UVR8, DFR,* and *PAL*. XLOC_011,933 was predicted to target *PAL, FL3H, LDOX, F3GT, UFGT, MYB5, DFR, UVR8*, and *COP10*. XLOC_042,291 was predicted to target *PAL, FL3H, LDOX, F3GT, DFR, UFGT, COP10,* and *UVR8*. These results indicated a possible regulatory relationship between lncRNAs and anthocyanin biosynthesis structural gene photoreceptors and light signal transduction-related genes. The expression pattern analysis showed a similar pattern for most structural genes, which paralleled the accumulation of anthocyanins (**Figure 5A**). *ANL2* (ANTHOCYANINLESS2), which belongs to the *HD–ZIP* family, has been reported to be possibly involved in the accumulation of anthocyanin in *Arabidopsis* (Kubo et al., 1999), exhibiting an opposite trend with an accumulation of anthocyanins.

In this research, we also identified the conservation of XLOC_001865, XLOC_011933, and XLOC_042291, and found that the evolutionary conservation of lncRNAs could be limited. This result was consistent with the results of lncRNA-related studies (Liu et al., 2012). In addition, we conducted q-PCR verification on the screened structural genes and photoreceptor-related genes and found that their expression trends were consistent with transcriptome data, proving that the transcriptome data were reliable. The results showed that these differentially expressed long non-coding genes played specific roles in the anthocyanin biosynthesis of peach fruits.

Peach is a rich genetic resource, in which the discovery of lncRNAs will bring great convenience to breeding. Rosaceae is a branch of fruit trees, and it is very important to study its fruit. Thus, understanding the lncRNA-mediated network regulation of UV-B-induced anthocyanin biosynthesis would lay the foundation for unraveling the complex molecular genetic mechanisms of positive effects on improving agronomic traits.

CONCLUSION

Color is one of the most essential sensory attributes of fresh fruits, and it influences the choices and preferences of consumers, indicates maturity, and correlates with other quality attributes, such as sugar and acid content and flavor. Water-soluble anthocyanins can produce different colors, such as red, purple, and blue. In the present study, we sought to identify lncRNAs from fruit transcriptomes to identify lncRNAs that function in fruit development. We identified and screened some differentially expressed lncRNAs by the transcriptome analysis. Through fluorescence real-time quantitative PCR (qRT-PCR) experiments, we found that the results were the same as the trend of the transcriptome. This study is the first to analyze and discover the lncRNAs involved in fruit coloration in peaches. The findings from this study may encourage researchers to study peach flesh coloring in detail.

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

JW and DM conceived and designed the project. MY, YL, ZS, TD, and HC participated in the experiment and data analysis. MZ, XZ, and HW drafted the manuscript. LS, XX, JL, LZ, YS, and QY revised the manuscript. Final draft was read and approved by all authors.

FUNDING

This study was supported by the Natural Science Foundation of Hebei Province (C2021407044), the National Natural Science Foundation of China (31922058, 31901281) Outstanding Young Talent Fund in Beijing Forestry University (2019JQ03009).

ACKNOWLEDGMENTS

We are grateful to editage for the embellishment of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.932207/full#supplementary-material

REFERENCES

Albert, N. W., Lewis, D. H., Zhang, H., Irving, L. J., Jameson, P. E., and Davies, K. M. (2009). Light-Induced Vegetative Anthocyanin Pigmentation in Petunia. *J. Exp. Bot.* 60, 2191–2202. doi:10.1093/jxb/erp097

An, N., Fan, S., Wang, Y., Zhang, L., Gao, C., Zhang, D., et al. (2018). Genome-Wide Identification, Characterization and Expression Analysis of Long Non-Coding RNAs in Different Tissues of Apple. *Gene* 666, 44–57. doi:10.1016/j.gene.2018.05.014

Azuma, A., Yakushiji, H., Koshita, Y., and Kobayashi, S. (2012). Flavonoid Biosynthesis-Related Genes in Grape Skin Are Differentially Regulated by Temperature and Light Conditions. *Planta* 236, 1067–1080. doi:10.1007/s00425-012-1650-x

Boerner, S., and McGinnis, K. M. (2012). Computational Identification and Functional Predictions of Long Noncoding RNA in Zea Mays. *PLOS ONE* 7, e43047. doi:10.1371/journal.pone.0043047

Butelli, E., Licciardello, C., Zhang, Y., Liu, J., Mackay, S., Bailey, P., et al. (2012). Retrotransposons Control Fruit-Specific, Cold-Dependent Accumulation of Anthocyanins in Blood Oranges. *Plant Cell* 24, 1242–1255. doi:10.1105/tpc.111.095232

Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative Annotation of Human Large Intergenic Noncoding RNAs Reveals Global Properties and Specific Subclasses. *Genes Dev.* 25, 1915–1927. (TUCP). doi:10.1101/gad.17446611

Deluc, L., Bogs, J., Walker, A. R., Ferrier, T., Decendit, A., Merillon, J.-M., et al. (2008). The Transcription Factor VvMYB5b Contributes to the Regulation of Anthocyanin and Proanthocyanidin Biosynthesis in Developing Grape Berries. *Plant Physiol.* 147, 2041–2053. doi:10.1104/pp.108.118919

Escribano-Bailón, M. T., Santos-Buelga, C., and Rivas-gonzalo, J. C. (2004). Anthocyanins in Cereals. *J. Chromatogr. a* 1054, 129–141.

Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-amma, S., and Allan, A. C. (2007). Red Colouration in Apple Fruit Is Due to the Activity of the MYB Transcription Factor, MdMYB10. *Plant J.* 49, 414–427. doi:10.1111/j.1365-313x.2006.02964.x

Grotewold, E. (2006). The Genetics and Biochemistry of Floral Pigments. *Annu. Rev. Plant Biol.* 57, 761–780. doi:10.1146/annurev.arplant.57.032905.105248

Guo, J., Han, W., and Wang, M. H. (2008). Ultraviolet and Environmental Stresses Involved in the Induction and Regulation of Anthocyanin Biosynthesis: A Review. *Afr. J. Biotechnol.* 7, 4966–4972.

He, H., Allan, A. C., and Han, Y. P. (2015). Molecular Genetics of Blood-Fleshed Peach Reveals Activation of Anthocyanin Biosynthesis by NAC Transcription Factors. *Plant J.* 82, 105–121.

Hedrick, U. P. (1917). *The Peaches of New York. Report of the New York Agricultural Experiment Station for the Year 1916*. Albany, NY: J.B. Lyon Co. doi:10.5962/bhl.title.55218

Holton, T. A., and Cornish, E. C. (1995). Genetics and Biochemistry of Anthocyanin Biosynthesis. *Plant Cell* 7, 1071–1083.

Jaakola, L. (2013). New Insights into the Regulation of Anthocyanin Biosynthesis in Fruits. *Trends Plant Sci.* 18 (9), 477–483. doi:10.1016/j.tplants.2013.06.003

Jimenez-Garcia, S. N., Guevara-Gonzalez, R. G., Miranda-Lopez, R., Feregrino-Perez, A. A., Torres-Pacheco, I., and Vazquez-Cruz, M. A. (2013). Functional Properties and Quality Characteristics of Bioactive Compounds in Berries: Biochemistry, Biotechnology, and Genomics. *Food Res. Int.* 54, 1195–1207. doi:10.1016/j.foodres.2012.11.004

Kang, C., and Liu, Z. (2015). Global Identification and Analysis of Long Non-Coding RNAs in Diploid Strawberry Fragaria Vesca During Flower and Fruit Development. *BMC Genomics* 16 (1), 815–15. doi:10.1186/s12864-015-2014-2

Kim, E.-D., and Sung, S. (2012). Long Noncoding RNA: Unveiling Hidden Layer of Gene Regulatory Networks. *Trends Plant Sci.* 17, 16–21. doi:10.1016/j.tplants.2011.10.008

Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., et al. (2007). CPC: Assess the Protein-Coding Potential of Transcripts Using Sequence Features and Support Vector Machine. *Nucleic Acids Res.* 35, W345–W349. doi:10.1093/nar/gkm391

Kubo, H., Peeters, A. J. M., Aarts, M. G. M., Pereira, A., and Koornneef, M. (1999). ANTHOCYANINLESS2, a Homeobox Gene Affecting Anthocyanin Distribution and Root Development in Arabidopsis. *Plant Cell* 11 (7), 1217–1226. doi:10.1105/tpc.11.7.1217

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

Lau, O. S., and Deng, X. W. (2012). The Photomorphogenic Repressors COP1 and DET1: 20 Years Later. *Trends Plant Sci.* 17, 584–593. doi:10.1016/j.tplants.2012.05.004

Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., et al. (2007). Analysis of Transcription Factor HY5 Genomic Binding Sites Revealed its Hierarchical Role in Light Regulation of Development. *Plant Cell* 19, 731–749. doi:10.1105/tpc.106.047688

Li, L., Eichten, S. R., Shimizu, R., Petsch, K., Yeh, C.-T., Wu, W., et al. (2014). Genome-Wide Discovery and Characterization of Maize Long Non-Coding RNAs. *Genome Biol.* 15, R40. doi:10.1186/gb-2014-15-2-r40

Li, L., Wang, X., Sasidharan, R., Stolc, V., Deng, W., He, H., et al. (2007). Global Identification and Characterization of Transcriptionally Active Regions in the Rice Genome. *PLoS One* 2, e294. doi:10.1371/journal.pone.0000294

Li, Y.-Y., Mao, K., Zhao, C., Zhao, X.-Y., Zhang, R.-F., Zhang, H.-L., et al. (2013). Molecular Cloning and Functional Analysis of a Blue Light Receptor Gene MdCRY2 from Apple (Malus Domestica). *Plant Cell Rep.* 32, 555–566. doi:10.1007/s00299-013-1387-4

Lim, S. H., and Ha, S. H. (2013). Marker Development for Identification of Rice Seed Coat Color. *Plant Biotechnol. Rep.* 7, 391–398.

Lin, M. F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: A Comparative Genomics Method to Distinguish Protein Coding and Non-Coding Regions. *Bioinformatics* 27 (13), i275–i282. (phyloCSF). doi:10.1093/bioinformatics/btr209

Lin-Wang, K., Micheletti, D., Palmer, J., Volz, R., Lozano, L., Espley, R., et al. (2011). High Temperature Reduces Apple Fruit Colour via Modulation of the Anthocyanin Regulatory Complex. *Plant, Cell & Environ.* 34, 1176–1190. doi:10.1111/j.1365-3040.2011.02316.x

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., et al. (2012). Genome-Wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in Arabidopsis. *Plant Cell* 24, 4333–4345. doi:10.1105/tpc.112.102855

Liu, S., Sun, Z., and Xu, M. (2018). Identification and Characterization of Long Non-Coding RNAs Involved in the Formation and Development of Poplar Adventitious Roots. *Industrial Crops Prod.* 118, 334–346. doi:10.1016/j.indcrop.2018.03.071

Ma, L., Bajic, V. B., and Zhang, Z. (2013). On the Classification of Long Non-Coding RNAs. *RNA Biol.* 10, 925–933. doi:10.4161/rna.24604

Macheix, J. J., Fleuriet, A., and Billot, J. (1990). *Fruit Phenolics*. Boca Raton: CRC Press.

Mano, H., Ogasawara, F., Sato, K., Higo, H., and Minobe, Y. (2007). Isolation of a Regulatory Gene of Anthocyanin Biosynthesis in Tuberous Roots of Purple-Fleshed Sweet Potato. *Plant Physiol.* 143, 1252–1268. doi:10.1104/pp.106.094425

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

Martin, C., Butelli, E., Petroni, K., and Tonelli, C. (2011). How Can Research on Plants Contribute to Promoting Human Health? *Plant Cell* 23, 1685–1699. doi:10.1105/tpc.111.083279

Mehrtens, F., Kranz, H., Bednarek, P., and Weisshaar, B. (2005). The Arabidopsis Transcription Factor MYB12 Is a Flavonol-Specific Regulator of Phenylpropanoid Biosynthesis. *Plant Physiol.* 138, 1083–1096. doi:10.1104/pp.104.058032

Osterlund, M. T., Hardtke, C. S., Wei, N., and Deng, X. W. (2000). Targeted Destabilization of HY5 During Light-Regulated Development of Arabidopsis. *Nature* 405, 462–466. doi:10.1038/35013076

Peng, T., Saito, T., Honda, C., Ban, Y., Kondo, S., Liu, J. H., et al. (2013). Screening of UV-B-Induced Genes from Apple Peels by SSH: Possible Involvement of MdCOP1-Mediated Signaling Cascade Genes in Anthocyanin Accumulation. *Physiol. Plant.* 148, 432–444. doi:10.1111/ppl.12002

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-Level Expression Analysis of RNA-Seq Experiments With

HISAT, StringTie and Ballgown. *Nat. Protoc.* 11 (9), 1650–1667. doi:10.1038/nprot.2016.095

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., et al. (2012). The Pfam Protein Families Database. *Nucleic Acids Res.* 40, D290–D301. (Pfam Scan). doi:10.1093/nar/gkr1065

Rizzini, L., Favory, J.-J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E., et al. (2011). Perception of UV-B by the Arabidopsis UVR8 Protein. *Science* 332, 103–106. doi:10.1126/science.1200660

Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., Lotia, S., et al. (2012). A Travel Guide to Cytoscape Plugins. *Nat. Methods* 9, 1069–1076. doi:10.1038/nmeth.2212

Shi, M.-Z., and Xie, D.-Y. (2014). Biosynthesis and Metabolic Engineering of Anthocyanins in *Arabidopsis thaliana*. *Biot.* 8, 47–60. doi:10.2174/1872208307666131218123538

Stracke, R., Favory, J. J., Gruber, H., Bartelniewoehner, L., Bartels, S., Binkert, M., et al. (2010). The Arabidopsis bZIP Transcription Factor HY5 Regulates Expression of the PFG1/MYB12 Gene in Response to Light and Ultraviolet-B Radiation. *Plant Cell Environ.* 33, 88–103. doi:10.1111/j.1365-3040.2009.02061.x

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

Tang, W., Zheng, Y., Dong, J., Yu, J., Yue, J., Liu, F., et al. (2016). Comprehensive Transcriptome Profiling Reveals Long Noncoding RNA Expression and Alternative Splicing Regulation During Fruit Development and Ripening in Kiwifruit (Actinidia Chinensis). *Front. Plant Sci.* 7, 335. doi:10.3389/fpls.2016.00335

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching During Cell Differentiation. *Nat. Biotechnol.* 28, 511–515. doi:10.1038/nbt.1621

Ubi, B. E., Honda, C., Bessho, H., Kondo, S., Wada, M., Kobayashi, S., et al. (2006). Expression Analysis of Anthocyanin Biosynthetic Genes in Apple Skin: Effect of UV-B and Temperature. *Plant Sci.* 170, 571–578. doi:10.1016/j.plantsci.2005.10.009

Verde, I., Jenkins, J., Dondini, L., Micali, S., Pagliarani, G., Vendramin, E., et al. (2017). The Peach v2.0 Release: High-Resolution Linkage Mapping and Deep Resequencing Improve Chromosome-Scale Assembly and Contiguity. *BMC Genomics* 18, 225. doi:10.1186/s12864-017-3606-9

Wang, M., Yuan, D., Tu, L., Gao, W., He, Y., Hu, H., et al. (2015a). Long Noncoding RNA S and Their Proposed Functions in Fibre Development of Cotton (Gossypium spp.). *New Phytol.* 207 (4), 1181–1197. doi:10.1111/nph.13429

Wang, T.-Z., Liu, M., Zhao, M.-G., Chen, R., and Zhang, W.-H. (2015b). Identification and Characterization of Long Non-Coding RNAs Involved in Osmotic and Salt Stress in Medicago Truncatula Using Genome-Wide High-Throughput Sequencing. *BMC Plant Biol.* 15 (1), 1–13. doi:10.1186/s12870-015-0530-5

Wang, Y., Gao, L., Li, J., Zhu, B., Zhu, H., Luo, Y., et al. (2018). Analysis of Long-Non-Coding RNAs Associated with Ethylene in Tomato. *Gene* 674, 151–160. doi:10.1016/j.gene.2018.06.089

Welch, C. R., Wu, Q., and Simon, J. E. (2008). *Recent Advances in Anthocyanin Analysis and Characterization*.

Wen, J., Parker, B. J., and Weiller, G. F. (2007). In Silico Identification and Characterization of mRNA-Like Noncoding Transcripts in Medicago Truncatula. *Silico Biol.* 7, 485–505.

Xie, R., Zheng, L., He, S., Zheng, Y., Yi, S., and Deng, L. (2011). Anthocyanin Biosynthesis in Fruit Tree Crops: Genes and Their Regulation. *Afr. J. Biotech.* 10 (86), 19890–19897. doi:10.5897/ajbx11.028

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:10.1186/1471-2229-11-61

Yin, D.-D., Li, S.-S., Shu, Q.-Y., Gu, Z.-Y., Wu, Q., Feng, C.-Y., et al. (2018). *Identification of microRNAs and Long Non-Coding RNAs Involved in Fatty Acid Biosynthesis in Tree Peony Seeds*.

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.* (GOseq). doi:10.1186/gb-2010-11-2-r14

Zhang, G., Duan, A., Zhang, J., and He, C. (2017). Genome-Wide Analysis of Long Non-Coding RNAs at the Mature Stage of Sea Buckthorn (Hippophae Rhamnoides Linn) Fruit. *Gene* 596, 130–136. doi:10.1016/j.gene.2016.10.017

Zhao, Y., Dong, W., Wang, K., Zhang, B., Allan, A. C., Lin-Wang, K., et al. (2017). Differential Sensitivity of Fruit Pigmentation to Ultraviolet Light Between Two Peach Cultivars. *Front. Plant Sci.* 8, 1552. doi:10.3389/fpls.2017.01552

Zhu, B., Yang, Y., Li, R., Fu, D., Wen, L., Luo, Y., et al. (2015). RNA Sequencing and Functional Analysis Implicate the Regulatory Role of Long Non-Coding RNAs in Tomato Fruit Ripening. *Exbotj* 66, 4483–4495. doi:10.1093/jxb/erv203

Zoratti, L., Karppinen, K., Luengo Escobar, A., Häggman, H., and Jaakola, L. (2014). Light-Controlled Flavonoid Biosynthesis in Fruits. *Front. Plant Sci.* 5, 534. doi:10.3389/fpls.2014.00534

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

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

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 12 August 2022
DOI 10.3389/fgene.2022.893956

OPEN ACCESS

EDITED BY
Kim Yrjälä,
Zhejiang Agriculture and Forestry University, China

REVIEWED BY
Lianfeng Gu,
Fujian Agriculture and Forestry University, China
Zhaoen Yang,
Cotton Research Institute (CAAS), China
Qiang Wei,
Nanjing Forestry University, China

*CORRESPONDENCE
Ying Li,
liying@icbr.ac.cn
Zehui Jiang,
jiangzh@icbr.ac.cn

SPECIALTY SECTION
This article was submitted to RNA, a section of the journal Frontiers in Genetics

RECEIVED 11 March 2022
ACCEPTED 27 June 2022
PUBLISHED 12 August 2022

CITATION
Li Y, Zhang S, Zhang D, Li X, Gao Z and Jiang Z (2022), The miR166–mRNA network regulates vascular tissue differentiation in Moso bamboo. *Front. Genet.* 13:893956. doi: 10.3389/fgene.2022.893956

The miR166–mRNA network regulates vascular tissue differentiation in Moso bamboo

Ying Li[1]*, Shuqin Zhang[1], Deqiang Zhang[2], Xueping Li[1], Zhimin Gao[1] and Zehui Jiang[1]*

[1]National State Forestry and Grassland Administration Key Open Laboratory on the Science and Technology of Bamboo and Rattan, Institute of Gene Science and Industrialization for Bamboo and Rattan Resources, International Centre for Bamboo and Rattan, Beijing, China, [2]National Engineering Laboratory for Tree Breeding, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China

miR166s play an important role in plant tissue differentiation. However, the functions of miR166s in the differentiation of vascular tissue in bamboo have not yet been elucidated. Here, we showed that five miR166s are overexpressed (tags per million reads > 2,000) in underground shoot samples of wild-type (WT) Moso bamboo (*Phyllostachys edulis*) and a thick-walled variant (*P. edulis* "Pachyloen") throughout the developmental process. Potential targets of these miR166s include some genes encoding homeodomain-leucine zipper (HD-ZIP) transcription factors and protein kinases. Cleavage sites for miR166s were identified in seven *PeHD-ZIP* homologs and a protein kinase gene *via* degradome sequencing ($p < 0.05$). Dual-luciferase and transient expression assays confirmed the binding of miR166s to *PeHOXs*. Fluorescence *in situ* hybridization revealed that miR166s were localized to the xylem of the leaf, root, and internode of 2-month-old pot seedlings of WT Moso bamboo. Overall, these findings reveal that miR166s are regulators of vascular tissue differentiation in bamboo. The miR166s identified in our study provide novel targets for bamboo breeding.

KEYWORDS

miR166, vascular tissue, xylem, differentiation, transcription factor, bamboo, seedling

Introduction

Moso bamboo (*Phyllostachys edulis* (Carr.) H. de Lehaie) is a member of the subfamily Bambusoideae in the family Poaceae. Its ability to grow rapidly, coupled with the high quality of the wood of its culms, distinguishes it from rice (*Oryza sativa* L.), other herbaceous plants in the family Poaceae, and dicotyledonous woody and fruit trees. The culm wall of bamboo has a graded hierarchical structure, and vascular bundles (VBs) provide support for this structure (Wei et al., 2019) (Figures 1A–C). The tender shoots of bamboo, which are located at the nodes of well-developed rhizomes (Figures 1D,E), possess a unique taste and nutritional profile due to VBs. VBs provide support to the culm, which is more than 10 m tall; they form a dense network for the large underground rhizomatous system and confer the unique taste and nutrient profile of the shoots.

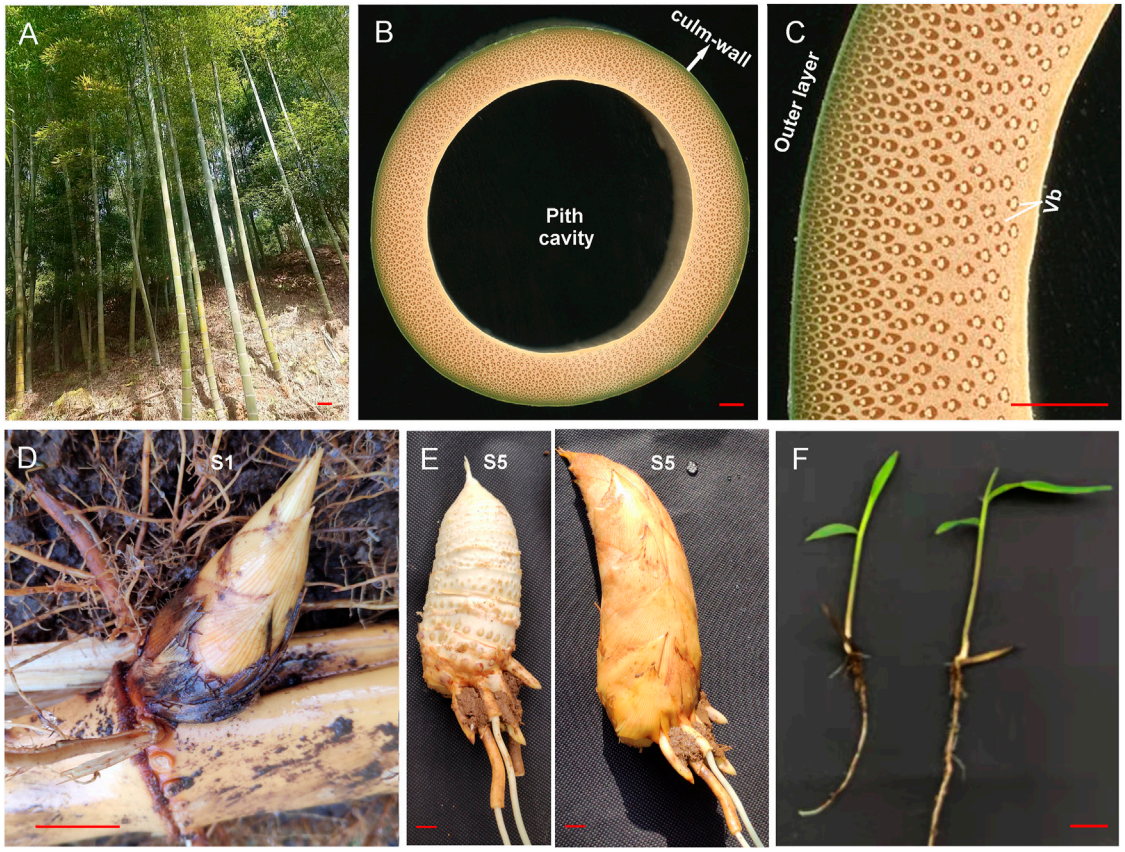


FIGURE 1
Phenotypes of aboveground culms, underground shoots, and seedlings of Moso bamboo. **(A)** Aboveground culms of two-year-old wild-type Moso bamboo plants. **(B)** Cross-sections of the third internode of the mature culm of wild-type Moso bamboo. **(C)** Enlarged image of the local culm wall in **(B)**. **(D)** Underground shoots collected at the S1 stage. **(E)** Underground shoots collected at the S5 stage; the left image shows the edible bamboo shoot with the shell removed. **(F)** Phenotype of two-month-old Moso bamboo seedlings. The scale bar is 10 cm in **(A)** and 1 cm in the other panels. Vb, vascular bundle.

However, the molecular mechanism underlying vascular tissue differentiation remains unclear. The aboveground part of bamboo seedlings is similar to that of rice; however, VBs are densely distributed in parenchyma cells of the internode, and regular changes in the radial direction of the culm wall are often observed, which is in contrast to the morphology of the vascular tissue in rice. Additional studies are needed to elucidate the molecular mechanisms underlying the regulation of the vascular tissues, including the differentiation of vascular tissues in the shoots and seedlings, as such studies could shed light on how variation in regulatory mechanisms drives the formation of specific phenotypes. There is a special need for studies aimed at identifying candidate regulatory elements that affect bamboo biomass and the properties of wood.

In plants, miR166/165 members comprise an important group of microRNAs (miRNAs) that play key roles in meristem development and organ polarity (Byrne, 2006; Zhou et al., 2007) by regulating the expression of various target genes, such as class III homeodomain/leucine zipper (*HD-ZIP III*) transcription factor (TF) genes (Prigge et al., 2005). Argonaute 10 inhibits the activity of miR165/166, which leads to the localized enrichment of *HD-ZIP* transcripts, ensures that the shoot apical meristem is correctly programmed in developing embryos, and supports the formation of the adaxial domain, the vasculature of the cotyledons, and the new leaf primordia (Zhang and Zhang, 2012; Zhou et al., 2015). In rice, miR166 mediates the localization of the RDD1 protein in VBs and sheaths by repressing the expression of the TF gene *rice Dof daily fluctuations 1* (*RDD1*) in the mesophyll (Iwamoto and Tagiri, 2016). During root procambial development in *Arabidopsis*, a network composed of miR166/165, *PHLOEM EARLY DOF 1* and *2* (*PEAR1* and *2*), and the four homologs *DOF6*, *TM O 6*, *OBP2*, and *HCA2*, integrates spatial information of the hormonal domains and miRNA gradients to generate adjacent zones of dividing and quiescent cells, which facilitates radial growth (Miyashima et al., 2019). miRNAs have also been identified in

the aboveground parts of mature bamboo culms (Ge et al., 2017). miRNAs have been shown to play an important role in the tissue differentiation of underground shoots at stages S1–S5 in wild-type (WT) Moso bamboo and its thick-walled (TW) variant (*P. edulis* "Pachyloen"), and miR166s are located in the vascular tissues of Moso bamboo shoots (Li et al., 2022). However, the roles of miR166s in the aforementioned regulatory network remain unclear; furthermore, little is known about their function in vascular tissue differentiation in bamboo seedlings.

Here, we aimed to 1) identify the target genes of miR166s, 2) study the interplay between miR166s and their targets, 3) characterize the localization of miR166s in bamboo shoots and seedlings, and 4) evaluate the implications of localization differences for tissue differentiation. We performed an integrated analysis of miR166s with the transcriptome, miRNAome, and degradome sequencing data using underground shoot samples of the WT Moso bamboo and the TW variant *P. edulis* "Pachyloen" taken throughout five stages of development (WTS1/TWS1–WTS5/TWS5) (Li et al., 2022). We also performed dual-luciferase reporter assays to study the interplay between miR166s and their predicted targets and RNA fluorescence *in situ* hybridization (FISH) to determine the localization of miR166s in leaf, root, and internode samples collected from two-month-old bamboo seedlings (Figure 1F). The results of our study provide new insights into the role of miR166s in vascular tissue differentiation and new candidate regulatory elements that could be used to enhance the biomass and wood properties of Moso bamboo.

Materials and methods

Plant materials

To determine the localization of miR166s, we collected two-month-old seedlings of Moso bamboo that had been maintained in an air-conditioned greenhouse; they were located in the National State Forestry and Grassland Administration Key Open Laboratory of Science and Technology of Bamboo and Rattan, Beijing, China (39°59′ 17.52″ N, 116° 28′46.06″ E). Plants were watered three times a week and maintained at 25 ± 1°C and 55 ± 5% relative humidity under a 16/8 h (light/dark) photoperiod with a light intensity of 1,250 μmol m^{-2} s^{-1}.

Data collection

The transcriptome, small RNA, and degradome sequencing data of underground shoot samples of WT and the TW Moso bamboo at the WTS1/TWS1-WTS5/TWS5 stages were downloaded from the NCBI SRA database using the accession No. PRJNA753616 (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA753616) (Li et al., 2022).

Identification and annotation of miR166s

The downloaded small RNA sequencing reads were screened against various databases, such as Rfam (http://rfam.xfam.org/), Silva (http://www.arb-silva.de/), GtRNAdb (http://lowelab.ucsc.edu/GtRNAdb/), and Repbase (http://www.girinst.org/repbase/), containing data on ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs, small nucleolar RNAs (snoRNAs), and other noncoding RNAs as well as repeat sequences using Bowtie v1.0.0 (Langmead et al., 2009) with the parameters "-v 0". To identify known miRNAs, the unannotated reads were mapped to the Moso bamboo reference genome v2.0 (Zhao et al., 2018), and the successfully mapped reads were compared against miRNAs of close relatives of Moso bamboo (*Brachypodium distachyon* and *Oryza sativa*) in miRBase (Kozomara et al., 2019) (release 22, http://mirbase.org/) using Bowtie v1.0.0., as miRNA annotations to the Moso bamboo genome and Moso bamboo miRNAs in miRbase were lacking (Zhao et al., 2018). Precursor sequences were mapped to the reference Moso bamboo genome using Bowtie v1.0.0.

Unaligned unique reads were used for the prediction of novel miRNAs in miRDeep-P2 (Kuang et al., 2019) following newly updated criteria for plant miRNA annotations (Axtell and Meyers, 2018) as well as the method described by Guo et al. (2020). miRNAs identified with similar sequences were clustered into families.

Expression analysis of miR166s

The abundances of miR166s were normalized to tags per million (TPM) (Li et al., 2010) using the following equation:

$$\mathrm{TPM} = \frac{\mathrm{Readcount} * 1,000,000}{\mathrm{Mapped\ Reads}},$$

where Readcount and Mapped Reads represent the number of reads mapped onto certain miRNAs and those mapped onto all miRNAs in the reference genome v2.0, respectively.

Replicates with high consistency ($r^2 > 0.70$) according to the Pearson's correlation analysis of gene expression (Li et al., 2022) were used in subsequent analyses.

We identified differentially expressed miRNAs (DEmiRs) between two stages for 13 pairwise comparisons of the WT, TW, and WTTW groups using DESeq2 (Love et al., 2014), including TWS2_vs._WTS1/TWS1, WTS3/TWS3_vs._WTS2/TWS2, WTS4/TWS4_vs._WTS3/TWS3, and WTS5/TWS5_vs._WTS4/TWS4 for the WT/TW group and TWS1–5_vs._WTS1–5 for the WTTW group. miRNAs were considered differentially expressed according to the following criteria: TPM > 1, $|\log_2(\mathrm{FC})| \geq 1.00$, and false discovery rate (FDR)-corrected p-value < 0.05 in one of two pairwise

comparisons. Comparisons of derived DEmiRs were referred to as "A_vs._B", for example, S01_vs._S02 corresponds to the comparison of DEmiRs between S01 and S02.

Target prediction and annotation

Potential miR166 targets were identified using psRNATarget (Dai et al., 2018) and the default parameters of Schema V2 (2017 release) (Dai et al., 2018), with the only difference being that a threshold value of 3 was used.

Annotations of these sequences were performed by conducting queries against several databases, including the Pfam (http://pfam.xfam.org/), Swiss-Prot (http://www.uniprot.org/), Clusters of Orthologous Genes (http://www.ncbi.nlm.nih.gov/COG/), Eukaryotic Orthologous Groups (http://www.ncbi.nlm.nih.gov/KOG/), Gene Ontology (http://www.geneontology.org/), Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/), and NR (ftp://ftp.ncbi.nih.gov/blast/db/) databases, using BLAST v2.2.26 (Altschul et al., 1997).

Expression analysis of predicted target genes

Downloaded transcriptome sequencing reads were mapped to the Moso bamboo reference genome v2.0 using HISAT2 v2.0.4 (Kim et al., 2015) with the parameters "--dta -p 6 --max-intronlen 5000000" to identify known genes. The reads mapped to the unannotated regions were then assembled using StringTie v1.3.4d (Pertea et al., 2015) with the parameters "--merge -F 0.1 -T 0.1" to identify novel genes.

Gene expression levels were estimated using fragments per kilobase of transcript per million fragments mapped (FPKM) (Florea et al., 2013) using the following formula:

$$\mathrm{FPKM} = \frac{\text{cDNA Fragments}}{\text{Mapped Fragments (Millions)} \times \text{Transcript Length (kb)}}.$$

Replicates with high consistency ($r^2 > 0.70$) according to the `Pearson's correlation analysis (Li et al., 2022) were used in subsequent analyses.

DESeq2 was used to identify DEGs (Love et al., 2014). DEGs were identified using the following criteria: FPKM > 1, $|\log_2(\mathrm{FC})| \geq 1.00$, and FDR-corrected p-value < 0.01 in one of two pairwise comparisons. Comparisons of derived DEGs were referred to as "A_vs._B", for example, WT1_vs._WT2 corresponds to DEGs between WTS1 and WTS2.

Degradome analysis of miR166–mRNA pairs

Downloaded degradome sequencing reads were searched against all other noncoding RNA sequences from Rfam, with the exception of miRNAs, using Bowtie v1.0.0; reads aligned to rRNAs, tRNAs, snoRNAs, and repeats were removed. The clean reads obtained were mapped to the Moso bamboo reference genome (v2.0), and a maximum of one mismatch was allowed. The miRNA cleavage sites were predicted using the CleaveLand (v4.4) pipeline (Addo-Quaye et al., 2009); only category 0, 1, and 2 results were retained to reduce the incidence of false positives.

Sequence analysis and phylogenetic tree construction

Gene structures of the predicted targets of miR166s were analyzed using bambooGDB (http://www.bamboogdb.org). Domains in the target protein sequences were detected using SMART (http://smart.embl-heidelberg.de). Phylogenetic trees were constructed using MEGA v5 (Tamura et al., 2011) with the neighbor-joining method and under the pairwise deletion and Poisson correction models. Bootstrapping was conducted using 1,000 replicates.

Luciferase reporter assay

Mimics of miRNAs, mRNAs 200 bp in length, and mutant-type (MUT) mRNAs with site-directed mutations in the predicted splicing sites were artificially synthesized (Sangon Biotech, Shanghai, China). They were then cloned into the pmirGLO vector (GeneCreate, Wuhan, China) between *Nhe*I and *Xho*I (TaKaRa, Nojihigashi, Japan); the control group was transfected with empty plasmid.

After confirmation by sequencing, 293T cells were co-transfected with MUT and WT vectors containing negative control (NC) mimics and miRNA mimics (GeneCreate), respectively. Cells were collected and processed 48 h after transfection using the Luciferase Reporter Assay Kit (RG005, Beyotime, Shanghai, China) as per the manufacturer's instructions. The outcomes were quantified in each well as the proportion of the firefly luciferase/*Renilla* luciferase activity. Three independent experiments were performed. The sequences of NC and miRNA mimics as well as those of WT and MUT plasmids, are listed in Supplementary Table S1-1.

Transient expression assays in *Nicotiana benthamiana* leaves

DNA and RNA extraction

Total DNA was isolated from the leaves of two-month-old Moso bamboo seedlings using the DNeasy Plant Mini Kit (Tiangen, Beijing, China) per the manufacturer's instructions. Total RNA was isolated from the roots, stems, and leaves of two-month-old Moso bamboo seedlings using the Tiangen RNAprep Pure Plant Kit (Tiangen) per the manufacturer's instructions. After treatment with RNase-free DNase I (Tiangen), 1 μg of RNA was reverse-transcribed using TransScript® II One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech).

Transient expression assay

Transient expression assays were performed following a previously described method (Wang et al., 2021) with slight modifications. Fragments of *PeHOX10* and *PeHOX32* with the predicted target sites were fused in the pCAMBIA1302 vector with the GFPuv sequence (Yuan et al., 2021) driven by the 35S promoter. The stem-loop sequence of ped-miR166a-3p was cloned and inserted into the pCAMBIA1300 vector after the 35S promoter. 35S::PeHOX10-uvGFP, 35S::PeHOX32-uvGFP, and 35S:ped-miR166a-3p constructs were transformed into *Agrobacterium*. These strains were co-infiltrated into the leaves of *N. benthamiana*. GFP signals were detected 48 h later under the UV light. Primers are shown in Supplementary Table S1-2.

RNA fish

FISH assays were carried out to track the localization of miR166s in the cells and tissues of Moso bamboo seedlings. Paraffin-embedded tissues and sections were prepared following the method in Wei et al. (2017). The sections were first dewaxed in water, incubated with a mixed solution of 30% (w/v) H_2O_2 and methanol (1:10, v/v) at room temperature for 10 min, and then covered by proteinase K at 37°C for 20 min.

Hybridization was conducted using the FISH kit (C007, Shanghai Gefan Biotechnology Co., Ltd., Shanghai, China) as per the manufacturer's protocol. Briefly, the tissues were washed with phosphate-buffered saline (PBS), fixed in 4% (w/v) paraformaldehyde, and acetylated in acetic anhydride, followed by incubation with fluorescein isothiocyanate (FITC)-labeled probes (Abiocenter, Beijing, China) overnight at 65°C. They were then washed with 2 × SSC buffer, formamide/4 × SSC (1:1, v/v), and PBS buffer. Cell nuclei were counterstained with 4,6-diamidino-2-phenylindole (Invitrogen), followed by washing with PBS and antifade buffer (Beyotime, Shanghai, China). Images were taken using a Nikon Eclipse Ci-L microscope with a FITC fluorescence filter. The probe sequence was as follows: 5′-TCGGACCAGGCTTCA TTCCCCC-3′.

Statistical analyses

Statistical analyses were performed using SPSS 19.0 software (IBM, Armonk, NY, United States). All data in this study were presented as mean ± standard deviation, and all statistical tests were two-sided. Student's *t*-tests were conducted to evaluate the significance of differences between groups, and the threshold for statistical significance was $p < 0.05$.

Results

A total of 487 predicted target genes of seven miR166s were detected

We detected seven miR166s (Supplementary Table S2) according to the small RNA sequencing data from the underground shoots of WT and TW plants (Li et al., 2022). The length of all seven miR166s was 21 bp, and four (57.14%) had a 5′ uridine residue as the first base. The average GC content of miR166s was 57.14%. Their secondary structures varied (Figure 2), and the minimum free energy ranged from −68.30 kcal/mol (ped-miR166f) to −49.60 kcal/mol (ped-miR166h-3p) (Supplementary Table S2).

A total of 487 target genes were predicted for all seven miRNAs (Supplementary Table S2). Of these, 25 were TFs from 11 families. The most common TF families were HB-HD-ZIP (homeobox-homeodomain-leucine zipper; 10, 2.05%), NAC (NAM [no apical meristem], ATAF1/2 [*Arabidopsis* transcription activation factor1/2], and CUC2 [cup-shaped cotyledon2]; 5, 1.03%), and MYB-related (v-myb avian myeloblastosis viral oncogene homolog-related, 2, 0.41%) families. In addition, 100 and 16 genes encoding protein kinases (PKs) and transcriptional regulators (TRs) were identified, respectively.

Five miR166s were overexpressed in underground bamboo shoots

All seven miR166s were expressed (TPM > 1) in at least one of the samples in WT and TW groups, and eight (88.89%) miRNAs were expressed in all five stages. Five and three miRNAs were overexpressed (TPM > 2,000) and highly overexpressed (TPM > 10,000) throughout all five stages in WT and TW plants, respectively (Supplementary Table S2).

A total of 423 miRNA-TF pairs comprised five overexpressed miR166s and 234 predicted targets (Supplementary Table S3), among which 25, 16, and 9 encoded TFs, PKs, and TRs,

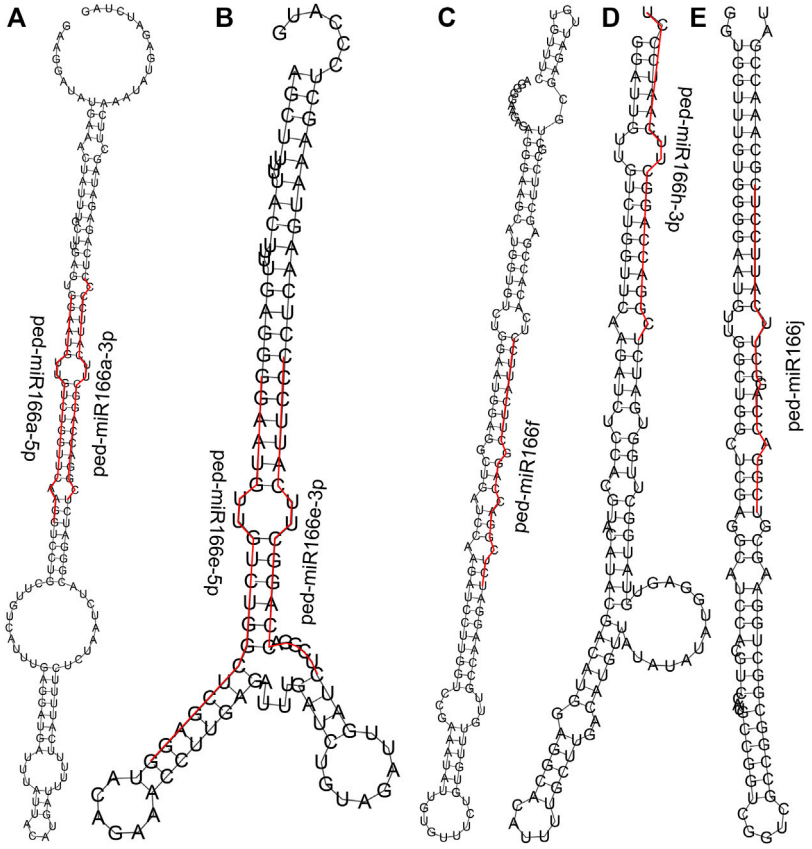


FIGURE 2
Predicted secondary structures of seven members of the miR166 family. Sequences of mature miR166s are denoted by the red line. Secondary structures predicted for ped-miR166a-3p and ped-miR166a-5p **(A)**, ped-miR166e-3p and ped-miR166e-5p **(B)**, ped-miR166f **(C)**, ped-miR166h-3p **(D)**, and ped-miR166j **(E)**.

respectively. ped-miR166j was detected in most pairs (105), followed by ped-miR166f (98), ped-miR166e-3p (83), and ped-miR166h-3 (70).

Nearly all five overexpressed miR166s shared targets from the HB-HD-ZIP TF family (Figure 3). The targets also included genes from other TF families, such as NAC, MYB, and AP2/ERF TFs. The most common PKs detected in miRNA–mRNA pairs were from the receptor-like kinase (RLK)/Pelle class (13).

Two miR166s were differentially expressed in 13 pairwise comparisons

Only two (22.22%) miR166s, ped-miR166c-5p and ped-miR166h-3p, were differentially expressed (TPM > 1, FDR < 0.05, and $|\log_2(FC)| \geq 1$ in one of two pairwise comparisons) among 13 pairwise comparisons (Supplementary Table S2). ped-miR166c-5p showed differential expression in the WTS5_vs._WTS4 comparison in the WT group and at S5 in the WTTW group. ped-miR166h-3p was differentially expressed in the WTS3_vs._WTS2 comparison in the WT group.

Target genes were identified for all seven members that were differentially expressed in at least one of the 13 comparison groups (Figure 4). Most of these genes were upregulated and/or downregulated in the TWS2_vs._TWS1 comparison. Among 235 targets of five overexpressed miR166s, 71 (30.21%) were differentially expressed (TPM > 1, FDR < 0.05, and $|\log_2(FC)| \geq 1$ in one of the two pairwise comparisons) among 13 pairwise comparisons (Supplementary Table S3).

A total of 23 miRNA–mRNA pairs were verified by degradome sequencing

A total of seven cleavage sites were identified for 23 miRNA-mRNA pairs, which included three miR166s and eight targets, and eight cleavage sites were identified ($p < 0.05$) in WT and TW groups (Figure 5A, Supplementary Table S4). Seven targets had HOX, START, and MEKHLA domains, and they were clustered into a

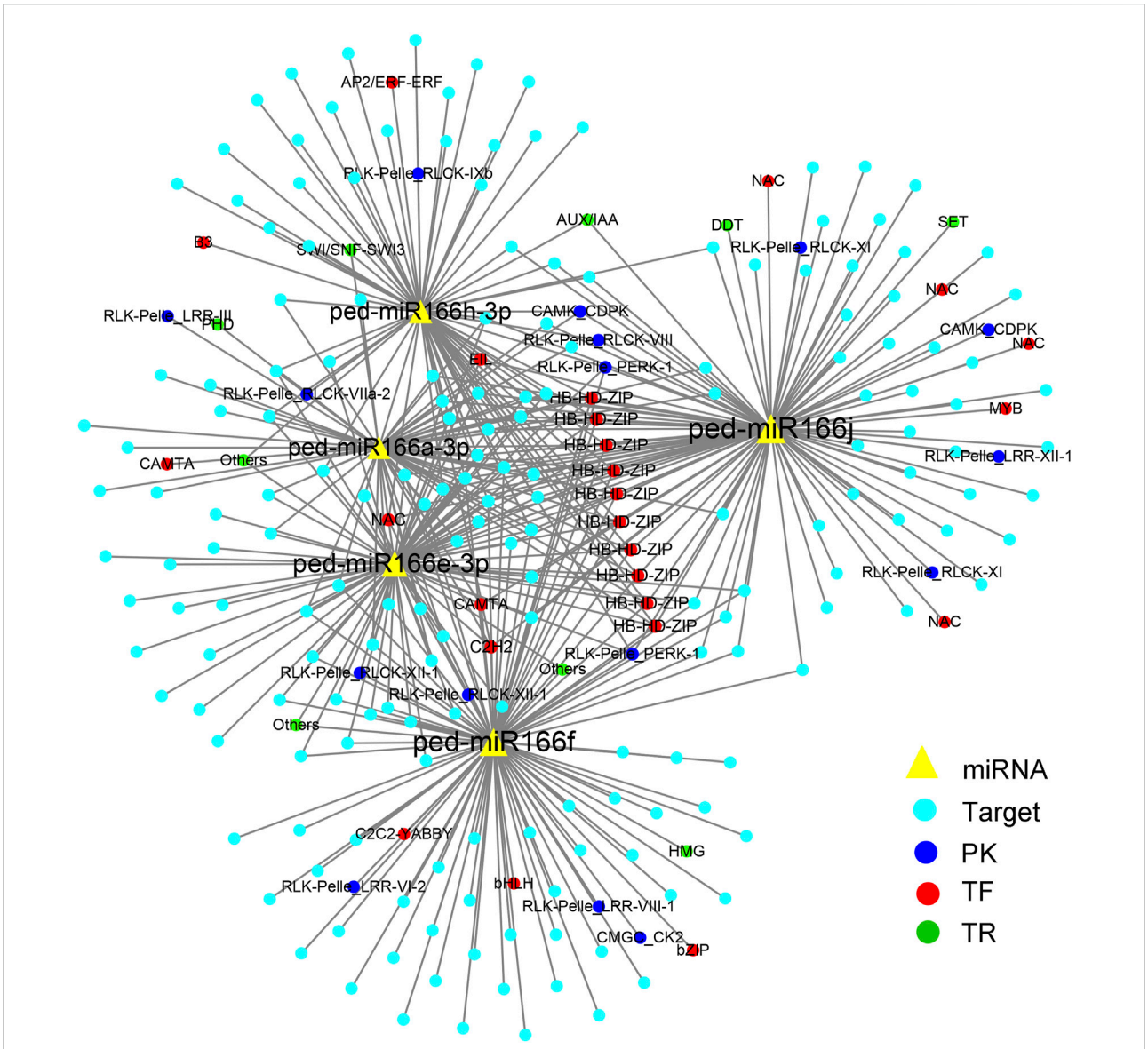


FIGURE 3
Cytoscape regulatory networks of five overexpressed miR166s and their predicted targets. Yellow triangles indicate miRNAs; blue solid circles indicate target genes. Targets that encode protein kinases are indicated by deep blue circles, transcription factors are indicated by red circles, and transcriptional regulators are indicated by green circles.

single group (Figures 5B,C). HOX is a DNA-binding factor involved in the transcriptional regulation of key developmental processes, and START (StAR-related lipid-transfer) is a lipid-binding domain in the StAR and HD-ZIP proteins as well as other signaling proteins. The MEKHLA domain is present in the 3′ end of plant HD-ZIP III homeobox genes, which encode DNA-binding factors involved in the transcriptional regulation of key developmental processes. The S_TKc domain is present in the serine or threonine-specific kinase subfamily of phosphotransferases.

Several miR166s within the same family cleave the same target. For example, *PH02Gene00975*, which encodes a putative *PeHOX10* homolog of rice, had the same cleavage site for ped-miR166a-3p, ped-miR166h-3p, and ped-miR166j in both WT and TW groups (Supplementary Table S4-1). The number of miR166–mRNA pairs differed among groups (Supplementary Table S4-2,3). A total of 21 and 17 cleavage sites were identified in the WT and TW groups, respectively. *PH02Gene20408* and *PH02Gene34818* had the same cleavage site for ped-miR166h-3p and ped-miR166j in the WT group, whereas no cleavage site was detected for these miRNAs in the TW group. *PH02Gene32772* had the same cleavage site for ped-miR166h-3p and ped-miR166j only in the TW group.

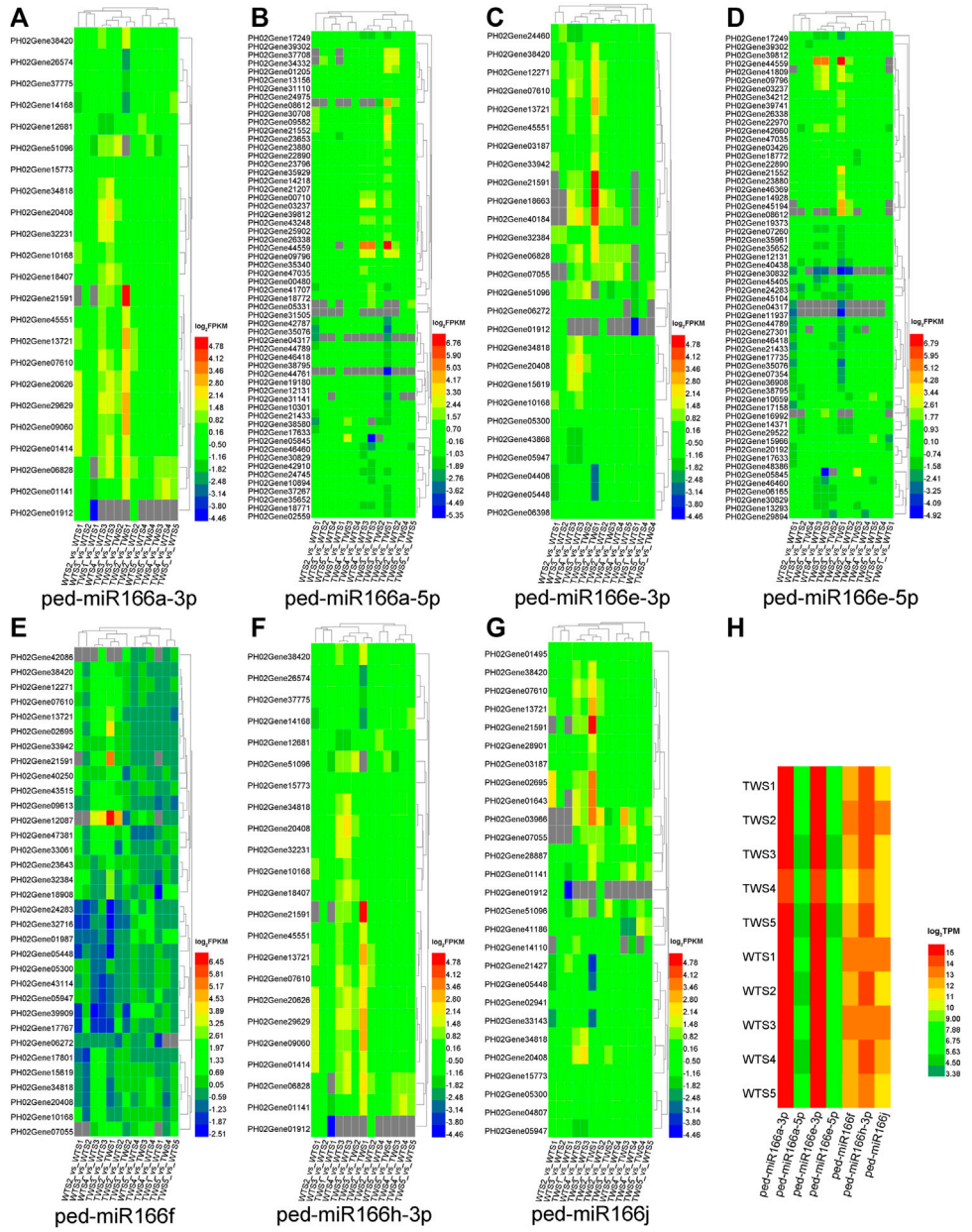


FIGURE 4
Expression patterns of seven miR166s and their target genes differentially expressed in 13 pairwise comparisons of the WT, TW, and WTTW groups. Both the wild-type (WT) and the thick-walled (TW) groups comprise four pairwise comparisons between two successive stages (WTS2/TWS2_vs_WTS1/TWS1, WTS3/TWS3_vs_WTS2/TWS2, WTS4/TWS4_vs_WTS3/TWS3, and WTS5/TWS5_vs_WTS4/TWS4), and the WTTW group comprises five pairwise comparisons between two stages (TWS1–5_vs_WTS1–5). **(A–G)** Expression patterns of the target genes of seven miR166s. **(H)** Expression patterns of seven miR166s. Gray squares indicate miRNAs and genes with low expression (TPM ≤ 1 or FPKM ≤ 1).

miR166s specifically bound to the *PeHOX* transcriptional factor genes

Dual-luciferase reporter assays revealed that three miR166s were specifically bound to *PeHOX10* and *PeHOX32*. The luciferase activity of the miRNA mimics + mRNA-WT group was lower than that of the NC mimics + mRNA-WT group ($p < 0.001$) in the transfected cells (Figures 6A,B). However, no significant differences were observed between the two mutant groups.

Transient expression assays revealed that ped-miR166a-3p bound to *PeHOX10* and *PeHOX32*. Differential fluorescence

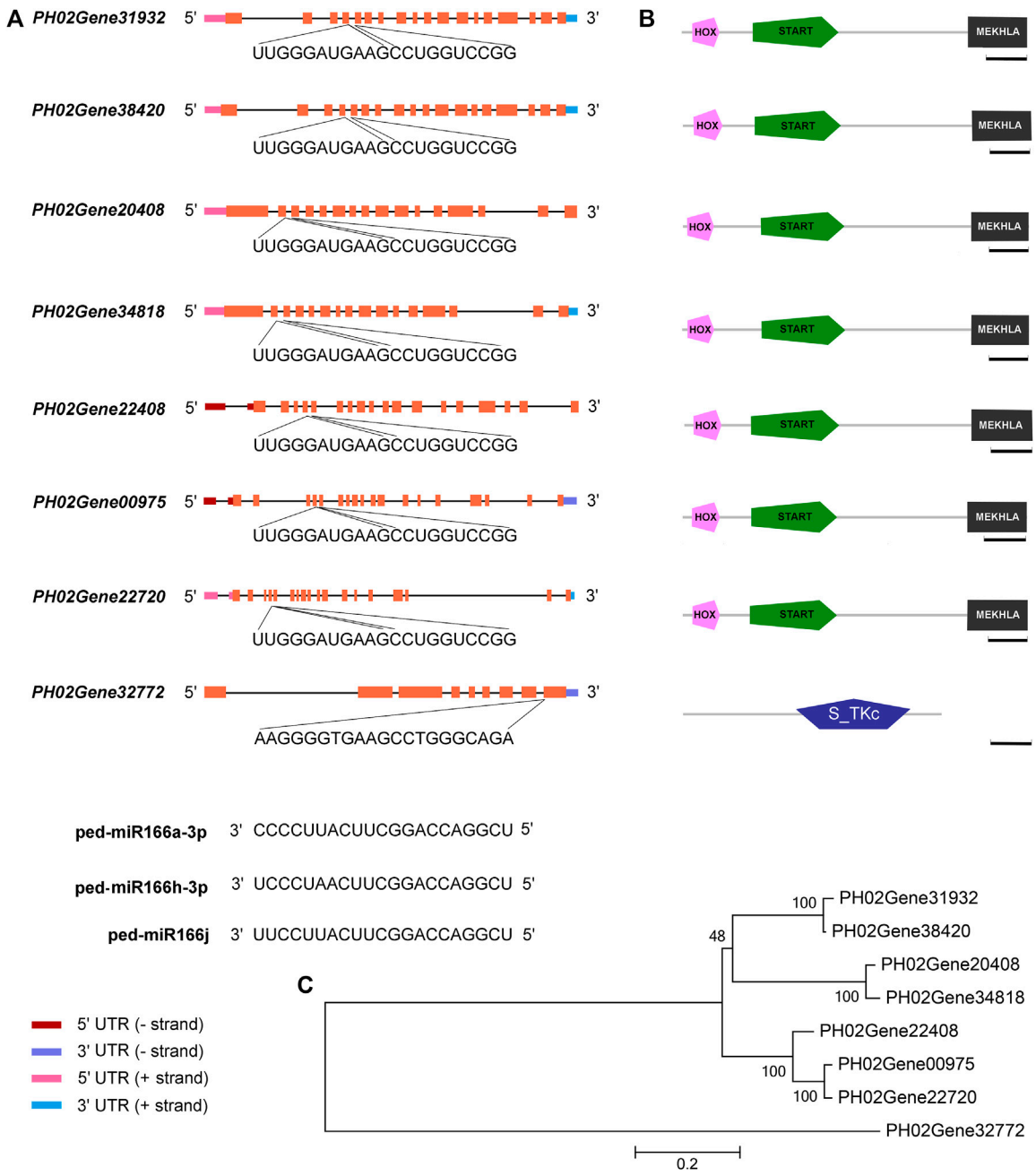


FIGURE 5
miR166s and their predicted targets in Moso bamboo. **(A)** Sequences of three miR166s, seven *HD-ZIP III* genes, and a protein kinase gene. **(B)** *HD-ZIP III* genes encode proteins with the following domains: an N-terminal MEKHLA domain, a START domain, and a C-terminal HOX domain. **(C)** Relationships among eight target genes. Diagrammatic representation of the target genes. Exons correspond to boxes, and introns correspond to lines. The HOX, START, and MEKHLA domains are the three structural domains in the coding sequences. Following them are the sequences of the binding sites in the eight targets and miR166s. Bar = 100 amino acids.

intensity was observed on two sides of the same leaf of *N. benthamiana* under the UV light (Figures 6C–F). The fluorescence intensity on the left side (Empty vector + PeHOXs-eYGFPuv) of treated leaves was higher than that on the right side (miR166 + PeHOXs-eYGFPuv).

miR166s were localized to the xylem

The VBs were distributed in various tissues of bamboo seedlings, and the morphological characteristics varied among organs (Figure 7). In the roots, VBs consisted of approximately a

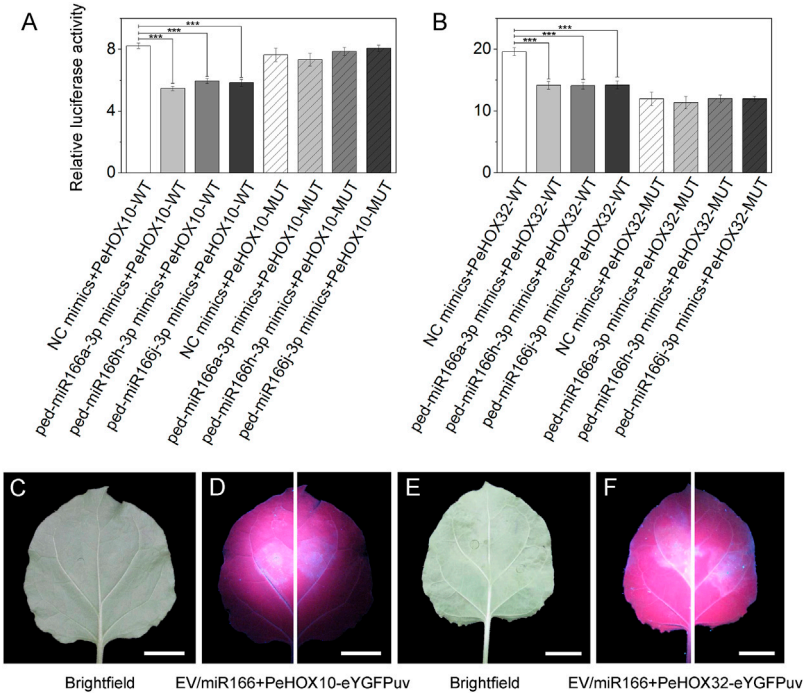


FIGURE 6
Verification of the correlation between three miR166s and two target genes using dual-luciferase reporter and transient expression assays. **(A–B)** Dual-luciferase reporter assays confirmed that miR166s are bound to *PeHOXs*. The three miRNAs were ped-miR166a-3p, ped-miR166h-3p, and ped-miR166j-3p. The two target genes were *PeHOX10* (*PH02Gene00975*) and *PeHOX32* (*PH02Gene31932*). Error bars indicate the standard deviations of three technical repeats. The significance of differences among groups was determined using *t*-tests. $*p < 0.05$; $**p < 0.01$; and $***p < 0.001$. **(C–F)** Transient expression assays confirmed that miR166s are bound to *PeHOXs*. Differential fluorescence intensity was observed between two sides of the same leaf under the UV light, where an empty vector (EV)/miR166 + PeHOXs-eYGFPuv was injected. Bar = 1 cm.

dozen small sieve tubes surrounding seven vessels. In the stems, large VBs were located around strips and had well-developed xylem; more than ten VBs were divided into two rings. VBs in the inner ring were more developed than those in the outer ring, which were mainly distributed in the mechanical tissues. VBs were also located in the leaves and leaf sheaths, and the VB sheaths included thick-walled cells.

Differential fluorescence intensity was observed in various tissues of the internode, leaf, and root of bamboo seedlings in merged images after false-positive signals from the black controls were excluded (Supplementary Figure S1). The fluorescence intensity of miR166 was higher than that of the 4,6-diamidino-2-phenylindole 2 HCl (DAPI) stain in the epidermis, xylem, and VB sheath of the internode (Figures 7A–D). ped-miR166a-3p was localized to the adaxial epidermis, abaxial epidermis, sclerenchyma, VB, and the VB sheath of the leaf (Figures 7E–H). ped-miR166a-3p was localized to the lateral root, epidermis, cortex, xylem, and pericycle of the root (Figure 7I–L). The fluorescence intensity of miR166 was lower than that of the DAPI stain of the phloem of the internode and root. These results confirmed that ped-miR166a-3p was specifically expressed in the xylem of the internode, leaf, and root of two-month-old Moso bamboo seedlings.

Discussion

We assessed the expression patterns of miR166s and their predicted targets and studied their interplay during tissue differentiation in Moso bamboo seedlings.

miR166s interact with *PeHOXs* and *RLKs* during tissue differentiation

Posttranslational modifications are a major mechanism by which miRNAs repress the expression of TFs, and TFs work synergistically or antagonistically to regulate the expression of various downstream target genes. miRNAs are important parts of the miRNA–TR–plant hormone regulatory network, which regulates complex processes with distinct developmental outcomes. In Arabidopsis, HD-ZIP TFs, which are regulated by miR165/166, play multiple, possibly interdependent, roles in

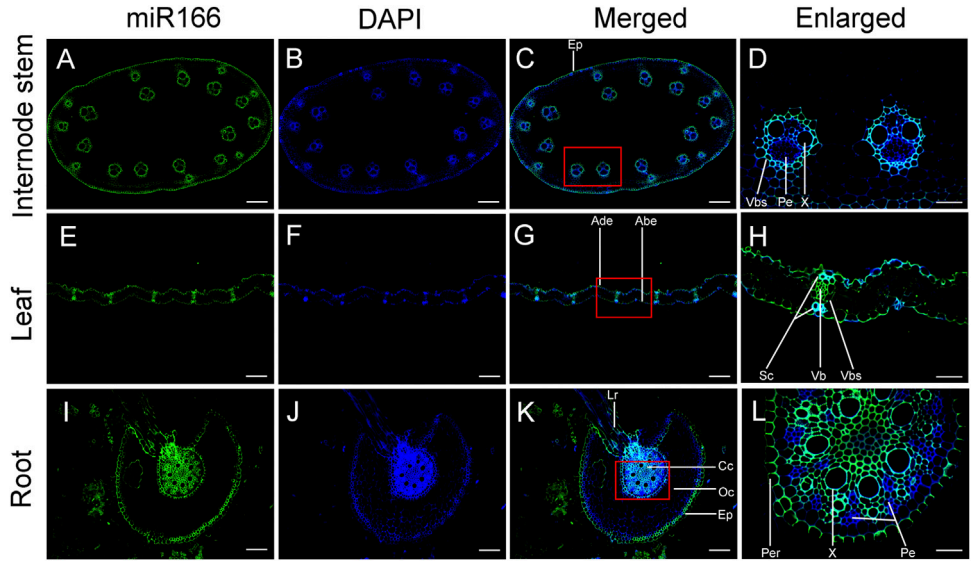


FIGURE 7
Verification of the localization of miR166s to vascular tissues of the internode, leaf, and root of two-month-old Moso bamboo seedlings by fluorescence *in situ* hybridization (FISH). ped-miR166a-3p was localized to the vascular tissues of the internode **(A–D)**, leaf **(E–H)**, and root **(I–L)**. Ep, epidermis; Vb, vascular bundle; Vbs, vascular bundle sheath; Pe, phloem; X, xylem; Ade, adaxial epidermis; Abe, abaxial epidermis; Sc, sclerenchyma; Lr, lateral root; Cc, central cylinder; Co, cortex; Per, pericycle. Bar = 100 μm.

plant development (Byrne, 2006). In our study, HD-ZIP TFs were detected in most miR166–mRNA pairs in the regulatory network. Dual-luciferase reporter assays confirmed that miR166a/h/j bound to *PeHOX10* and *PeHOX32*, both of which are bamboo *HD-ZIP III* homologs of rice. This finding indicates that the same target could be degraded by more than one miR166. miR166s likely affect the function of the meristem by regulating the expression of *PeHOXs* and contributing to distinct developmental outcomes in the underground shoots. This also indicates that miR166s play a key role in tissue differentiation and possibly underlie the morphological innovation of the developing tissues in Moso bamboo shoots.

miR166s affected the expression of other TF genes that participate in diverse processes, including immune and stress responses, light and hormonal signaling, and development, by specifically binding to *cis*-acting elements in the promoters of their target genes. This finding indicates that the miRNA–TF–mRNA regulatory network and miR166s play key roles in tissue differentiation in Moso bamboo. PKs in the RLK/Pelle class were detected in most miRNA–mRNA pairs (13). The RLK/Pelle PKs function in various processes, such as cell wall interactions, disease resistance, and developmental control (Gish and Clark, 2011), indicating that miR166s might affect the organization of the cell wall and regulate developmental processes. miR166s also target genes encoding TRs involved in plant hormone signal transduction. For example, both ped-miR166h-3p and ped-miR166j target *PH02Gene01141*, which encodes an AUX/IAA protein. AUX/IAA family proteins mediate auxin signaling and play important roles in plant growth and development. This finding indicates that genes related to signal transduction pathways and kinase activity during tissue differentiation are particularly important in developing bamboo shoots.

miR166s regulate the development of vascular tissues

Moso bamboo is distinguished from other monocot herbs by its tall culm, luxuriant leaves, and its complex and vast system of underground rhizomes. Strong and unimpeded vascular flow is needed to ensure that a source–sink balance is achieved. Unlike dicots, bamboo plants do not possess secondary cambia, and the multiple layers of vascular tissue in the culms provide resistance against external mechanical stress. Thus, VBs in Moso bamboo is different from and more complex than those in other dicotyledons. Li et al. (2022) found that both miR166h-3p and miR166a-3p are localized to the cambium of underground shoots at the S1 stage and in the vascular tissue of shoots at the S3 stage. This finding indicates that miR166s play an important role in vascular tissue differentiation in bamboo shoots. ped-miR166a-3p was localized to the xylems in the vascular tissue of the leaf, stem internode, and central cylinder of the root elongation

zone. Therefore, ped-miR166a-3p promotes the differentiation of the vascular tissue in various organs of bamboo seedlings.

The TW Moso bamboo has a thick culm wall, with an increased number of VBs and a much narrower pith cavity compared with the WT Moso bamboo. The fact that the apical meristem of the shoot at the S1 stage in the TW Moso bamboo produces more vascular tissue as opposed to pith cells compared with the WT Moso bamboo (Wei et al., 2017) might be explained by developmental differences between WT and TW Moso bamboos. In our study, the greatest number of target DEGs regulated by miR166s was observed in the TWS2_vs._TWS1 comparison. Additionally, the expression profiles of these target DEGs varied in the WTS2_vs._WTS1 and TWS2_vs._TWS1 comparison, and this might have contributed to variation in the phenotypes between WT and TW Moso bamboos. Although only two miR166s were differentially expressed, target DEGs were detected for all seven miR166s. Five miR166s were overexpressed, and three wmiR166s were highly overexpressed. Small changes in the expression levels of these genes thus might be sufficient to upregulate and downregulate the expression of target genes.

The target DEGs in the TWS2_vs._TWS1 comparison varied among the seven miR166s. Expression patterns of DEGs and miR166s in the TWS2_vs._TWS1 comparison were inconsistent. This might be explained by the fact that miR166s and other miRNAs affected the expression of these DEGs; that is, the miR166s might specifically bind to them to control their expression at certain stages, and they might participate in different biological pathways. These findings indicate that the miRNA–mRNA regulatory network is complex. Additional analyses and experiments are needed to verify the roles of miR166s in this regulatory network. Bamboo shoots can reach 18.3 m in only 3 months (Khalil et al., 2012); the large amount of water needed for cell elongation during this process is supplied by an unobstructed transport system. The molecular regulatory network and regulatory mechanisms might become more sophisticated during the developmental process in Moso bamboo, as the rhizomes of vegetative organs mainly reproduce asexually. This also might increase the complexity of the regulatory network, given that the xylem and phloem differ in both structure and function. miR166s might play specific regulatory roles in the vascular tissues during seedling growth, especially in the xylem.

Potential roles of miR166s in xylem patterning

miR165/6 functions as a mobile and morphogenic signal that determines the fates of cells during development. For example, miR165/6 is specifically expressed in the endodermis of the root, but a gradient of miR165/6 towards the center of the vasculature leads to an opposing gradient in the expression of its target genes (Miyashima et al., 2011; Miyashima et al., 2019). Furthermore, AGO1 preferentially recruits small RNAs with 5′-terminal uridine to construct RNA-induced silencing complexes (RISCs) to regulate gene expression at the transcriptional and posttranscriptional levels (Mi et al., 2008). Fan et al. (2022) showed that miR165/6, which is produced in the endodermis, moves to the vasculature to determine the fates of xylem cells in *Arabidopsis* roots. They also suggested that the cytoplasmic AGO1 loading of mobile miRNAs is a key step for promoting miRNA cell-to-cell movement. In this study, three overexpressed miR166s had a 5′ uridine residue as the first base, and ped-miR166a-3p was localized to the endodermis and the xylem in the roots of bamboo seedlings according to RNA FISH. These miR166s might be recruited by bamboo AGO1 to construct miRISCs that regulate the expression of *HD-ZIP* genes; they might also serve as local and long-distance signals in xylem patterning in multiple organs in Moso bamboo.

Moso bamboo shoots differ morphologically from the internodes and roots; they develop from underground vegetative buds located at the nodes of the rhizomes, which are morphologically similar to the roots but have nodes similar to the aboveground culm. In our study, ped-miR166a-3p, which has a 5′ uridine residue as the first base, was localized to the xylem rather than the phloem of the internode and the root. Li et al. (2022) detected ped-miR166a-3p in the cambium and xylem of Moso bamboo shoots at S1 and S3 stages, respectively. This indicates that miR166 is associated with xylem patterning and that there might be a relationship between the development of the shoot, root, and internode. miR166s in Moso bamboo might move between the xylem of the vasculature and function as mobile morphogenic signals that determine the fates of various types of cells during the development process. Additional experiments are needed to clarify the roles of miR166s in xylem patterning.

Conclusion

We characterized changes in the profiles of miR166s and their predicted targets. Our findings indicate that miR166s play key roles in the differentiation of vascular tissue in Moso bamboo. The localization of ped-miR166a-3p to the xylem in the stems and roots of two-month-old seedlings indicates that it plays a key role in the development of the xylem in Moso bamboo. Only ped-miR166a-3p could be localized using RNA FISH. Consequently, its interaction with other miR166s and target genes could not be determined. Additional approaches are needed to confirm the interaction between miR166s and their predicted targets. The miR166s identified in our study could be used to improve the properties of bamboo wood and the edible quality of bamboo shoots. The findings of our study reveal a complex regulatory network, in which miR166s target *PeHOXs*, *PeNACs*, and other genes to control vascular tissue differentiation in bamboo.

Data availability statement

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

Author contributions

YL designed the experiments, conceived the study, designed and performed the experiments, analyzed the data, and wrote the first draft of the manuscript; DZ, XL, and ZG provided technical assistance to YL, and SZ performed the experiments using the paraffin-embedding method.

Funding

This work was supported by the Special Funds for Fundamental Research Funds of the International Center for Bamboo and Rattan (No. 1632019025), the National Natural Science Foundation of China (NSFC) (No. 31901371), and the National Key Research & Development Program of China (No. 2021YFD2200503).

Acknowledgments

We thank Chenglei Zhu (Interactional Center for Bamboo and Rattan) for help with collecting plant materials.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.893956/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Signals of the blank control in Moso bamboo seedlings by fluorescence in situ hybridization. Bar = 20 μm expression.

SUPPLEMENTARY TABLE S1
Sequences of miR166 mimics, target gene fragments, and primers used for dual-luciferase reporter and transient expression assays.

SUPPLEMENTARY TABLE S2
Seven miR166s were identified from the small RNA sequencing data using the underground shoot samples of wild-type (WT) Moso bamboo (Phyllostachys edulis) and a thick-walled (TW) variant (P. edulis 'Pachyloen') downloaded from NCBI. Five and three miRNAs were overexpressed (TPM > 2,000) and highly overexpressed (TPM > 10,000) throughout all developmental stages in WT and TW groups, which are indicated in bold and italics, respectively. NA: not available.

SUPPLEMENTARY TABLE S3
A total of 487 predicted targets of seven miR166s. The transcription factor, protein kinase, and transcriptional regulator targets are in bold, red, and italics, respectively.

SUPPLEMENTARY TABLE S4
A total of 423 miRNA–mRNA pairs comprising five overexpressed (TPM > 2,000) miRNAs and their predicted targets. The transcription factor, protein kinase, and transcriptional regulator targets are in bold, red, and italics, respectively.

SUPPLEMENTARY TABLE S5
A total of 23 miR166–mRNA pairs were confirmed by the degradome sequencing dataset.

References

Addo-Quaye, C., Miller, W., and Axtell, M. J. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics* 25, 130–131. doi:10.1093/bioinformatics/btn604

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

Axtell, M. J., and Meyers, B. C. (2018). Revisiting criteria for plant microRNA annotation in the era of big data. *Plant Cell* 30, 272–284. doi:10.1105/tpc.17.00851

Byrne, M. E. (2006). Shoot meristem function and leaf polarity: The role of class III *HD-ZIP* genes. *Plos Genet.* 2, e89. doi:10.1371/journal.pgen.0020089

Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46, W49–W54. doi:10.1093/nar/gky316

Fan, L., Zhang, C., Gao, B., Zhang, Y., Stewart, E., Jez, J., et al. (2022) Microtubules promote the non-cell autonomous action of microRNAs by inhibiting their cytoplasmic loading onto ARGONAUTE1 in *Arabidopsis*. *Dev. Cell* 57, 995–1008. doi:10.1016/j.devcel.2022.03.015

Florea, L., Song, L., and Salzberg, S. L. (2013). Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000Res* 2, 188. doi:10.12688/f1000research.2-188.v2

Ge, W., Zhang, Y., Cheng, Z., Hou, D., Li, X., and Gao, J. (2017). Main regulatory pathways, key genes and microRNAs involved in flower formation and development of moso bamboo (*Phyllostachys edulis*). *Plant Biotechnol. J.* 15, 82–96. doi:10.1111/pbi.12593

Gish, L. A., and Clark, S. E. (2011). The RLK/Pelle family of kinases. *Plant J.* 66, 117–127. doi:10.1111/j.1365-313X.2011.04518.x

Guo, Z., Kuang, Z., Wang, Y., Zhao, Y., Tao, Y., Cheng, C., et al. (2020). PmiREN: a comprehensive encyclopedia of plant miRNAs. *Nucleic Acids Res.* 48, D1114–D1121. doi:10.1093/nar/gkz894

Iwamoto, M., and Tagiri, A. (2016). MicroRNA-targeted transcription factor gene *RDD1* promotes nutrient ion uptake and accumulation in rice. *Plant J.* 85, 466–477. doi:10.1111/tpj.13117

Khalil, H., Bhat, I., Jawaid, A., Zaidon, D., and Hadi, H. (2012). Bamboo fiber reinforced bio composite: a review. *Mat. Des.* 42, 353–368. doi:10.1007/978-981-33-4795-3_8110.1016/j.matdes.2012.06.015

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

Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 47, D155–D162. doi:10.1093/nar/gky1141

Kuang, Z., Wang, Y., Li, L., and Yang, X. (2019). miRDeep-P2: accurate and fast analysis of the microRNA transcriptome in plants. *Bioinformatics* 35, 2521–2522. doi:10.1093/bioinformatics/bty972

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25. doi:10.1186/gb-2009-10-3-r25

Li, B., Ruotti, V., Stewart, R. M., Thomson, J. A., and Dewey, C. N. (2010). RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* 26, 493–500. doi:10.1093/bioinformatics/btp692

Li, Y., Zhang, D., Zhang, S., Lou, Y., An, X., Jiang, Z., et al. (2022) Transcriptome and miRNAome analysis reveals components regulating tissue differentiation of bamboo shoots. *Plant Physiol.* 188, 2182–2198. doi:10.1093/plphys/kiac018

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

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., et al. (2008). Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. *Cell* 133, 116–127. doi:10.1016/j.cell.2008.02.034

Miyashima, S., Koi, S., Hashimoto, T., and Nakajima, K. (2011). Non-cell-autonomous microRNA165 acts in a dose-dependent manner to regulate multiple differentiation status in the *Arabidopsis* root. *Dev* 138, 2303–2313. doi:10.1242/dev.060491

Miyashima, S., Roszak, P., Sevilem, I., Toyokura, K., Blob, B., Heo, J. O., et al. (2019). Mobile PEAR transcription factors integrate positional cues to prime cambial growth. *Nature* 565, 490–494. doi:10.1038/s41586-018-0839-y

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

Prigge, M. J., Otsuga, D., Alonso, J. M., Ecker, J. R., Drews, G. N., and Clark, S. E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell* 17, 61–76. doi:10.1105/tpc.104.026161

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

Wang, K., Zhang, Y., Zhang, H., Lin, X., Xia, R., Song, L., et al. (2021). MicroRNAs play important roles in regulating the rapid growth of the Phyllostachys edulis culm internode. *New Phytol.* 231. doi:10.1111/nph.17542

Wei, Q., Jiao, C., Guo, L., Ding, Y., Cao, J., Feng, J., et al. (2017). Exploring key cellular processes and candidate genes regulating the primary thickening growth of Moso underground shoots. *New Phytol.* 214, 81–96. doi:10.1111/nph.14284

Wei, X., Zhou, H., Chen, F., and Wang, G. (2019). Bending flexibility of moso bamboo (*Phyllostachys edulis*) with functionally graded structure. *Materials* 12. doi:10.3390/ma12122007

Yuan, G., Lu, H., Tang, D., Hassan, M., Li, Y., Chen, J., et al. (2021). Expanding the application of a UV-visible reporter for transient gene expression and stable transformation in plants. *Hortic. Res.* 8, 234. doi:10.1038/s41438-021-00663-3

Zhang, Z., and Zhang, X. (2012). Argonautes compete for miR165/166 to regulate shoot apical meristem development. *Curr. Opin. Plant Biol.* 15, 652–658. doi:10.1016/j.pbi.2012.05.007

Zhao, H., Gao, Z., Wang, L., Wang, J., Wang, S., Fei, B., et al. (2018). Chromosome-level reference genome and alternative splicing atlas of moso bamboo (*Phyllostachys edulis*). *Gigascience* 7. doi:10.1093/gigascience/giy115

Zhou, G., Kubo, M., Zhong, R., Demura, T., and Ye, Z. (2007). Overexpression of miR165 affects apical meristem formation, organ polarity establishment and vascular development in *Arabidopsis*. *Plant Cell Physiol.* 48, 391–404. doi:10.1093/pcp/pcm008

Zhou, Y., Honda, M., Zhu, H., Zhang, Z., Guo, X., Li, T., et al. (2015). Spatiotemporal sequestration of miR165/166 by *Arabidopsis* argonaute10 promotes shoot apical meristem maintenance. *Cell Rep.* 10, 1819–1827. doi:10.1016/j.celrep.2015.02.047

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 17 August 2022
doi: 10.3389/fgene.2022.957505

Identification of Nitrogen Starvation-Responsive miRNAs to Reveal the miRNA-Mediated Regulatory Network in *Betula luminifera*

Yan Lin†, *Sasa Chu*†, *Xiaoshan Xu*, *Xiao Han*, *Huahong Huang*, *Zaikang Tong** and *Junhong Zhang**

State Key Laboratory of Subtropical Silviculture, School of Forestry & Bio-technology, Zhejiang A&F University, Hangzhou, China

OPEN ACCESS

Edited by:
Yuepeng Song,
Beijing Forestry University, China

Reviewed by:
Chunlong Li,
Huazhong Agricultural University, China
Jinhui Chen,
Hainan University, China

***Correspondence:**
Zaikang Tong
zktong@zafu.edu.cn
Junhong Zhang
zhangjunhong@zafu.edu.cn

†*These authors have contributed equally to this work*

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *31 May 2022*
Accepted: *24 June 2022*
Published: *17 August 2022*

Citation:
Lin Y, Chu S, Xu X, Han X, Huang H, Tong Z and Zhang J (2022) Identification of Nitrogen Starvation-Responsive miRNAs to Reveal the miRNA-Mediated Regulatory Network in Betula luminifera. Front. Genet. 13:957505. doi: 10.3389/fgene.2022.957505

Because of the immobility, plants encounter a series of stresses, such as varied nutrient concentrations in soil, which regulate plant growth, development, and phase transitions. Nitrogen (N) is one of the most limiting factors for plants, which was exemplified by the fact that low nitrogen (LN) has a great adverse effect on plant growth and development. In the present study, we explored the potential role of microRNAs (miRNAs) in response to LN stress in *Betula luminifera*. We identified 198 miRNAs using sRNA sequencing, including 155 known and 43 novel miRNAs. Among them, 98 known miRNAs and 31 novel miRNAs were differentially expressed after 0.5 h or 24 h of LN stress. Based on degradome data, 122 differential expressed miRNAs (DEmiRNAs) including 102 known miRNAs and 20 novel miRNAs targeted 203 genes, comprising 321 miRNA–target pairs. A big proportion of target genes were transcription factors and functional proteins, and most of the Gene Ontology terms were enriched in biological processes; moreover, one Kyoto Encyclopedia of Genes and Genomes term "ascorbate and aldarate metabolism" was significantly enriched. The expression patterns of six miRNAs and their corresponding target genes under LN stress were monitored. According to the potential function for targets of DEmiRNAs, a proposed regulatory network mediated by miRNA–target pairs under LN stress in *B. luminifera* was constructed. Taken together, these findings provide useful information to elucidate miRNA functions and establish a framework for exploring N signaling networks mediated by miRNAs in *B. luminifera*. It may provide new insights into the genetic engineering of the high use efficiency of N in forestry trees.

Keywords: miRNA, nitrogen starvation, target genes, Betula luminifera, sRNA sequencing

INTRODUCTION

Nitrogen (N) is one of the most indispensable elements for plant growth and development, which is considered the main limiting macronutrient (Perez-Tienda et al., 2014). N participates in a variety of biological processes, such as protein synthesis, nucleic acid synthesis, chlorophyll synthesis, ATP synthesis, and phytohormone synthesis, and acts as a signal (Zhao et al., 2015). Nevertheless, the majority of soil does not satisfy the plant's needs (Perez-Tienda et al., 2014). Although a large amount of fertilizers are applied, relatively low efficiency for fertilizer is commonly observed, leading to

environmental pollution and even harm to human health (Ahmed et al., 2017). Therefore, when a plant encounters N starvation, it is urgent to understand the molecular regulatory mechanism of adaption. Plants often grow in barren soil with limited N, especially forestry trees that are usually planted in infertile soil. *Betula luminifera* with rapid growth and excellent wood properties, as a pioneer tree, often grows in barren soil with limited N source, which restricts significantly the growth and productivity. It is urgent to elucidate the adaption mechanism of N starvation in this forestry tree species, which might help to make breeding strategies improve tolerance and adaptation to N starvation in *B. luminifera* and other forestry trees.

Amounting evidence showed that plants have evolved a complex but systematic adaptive mechanism to adapt the N starvation, such as gene transcription and regulation (Liu et al., 2017; Ye et al., 2019), morphological changes (Wang et al., 2019), and photosynthetic characteristics and metabolic pathways (Liu et al., 2019; Zhao et al., 2020). Among them, increasing studies suggest that miRNAs play an important role in response to N starvation, by regulating the expression of target genes, triggering the change of physiological and biochemistry characteristics in many species (Zuluaga and Sonnante, 2019; Li L. J. et al., 2021).

miRNA is a class of 18–24-nt endogenous non-coding RNA, which is processed from the *MIRNA* gene with stem-loop structure by Dicer-like (DCL) (Jones-Rhoades et al., 2006). Plant miRNAs are usually highly complementary with their target genes, triggering the target mRNA cleavage by RNA-induced silencing complex. However, the translational inhibition is occurred in the target genes, when the seed region of miRNAs harbors relatively low complementary with target mRNA (Brodersen et al., 2008). In 2002, miRNA was first found in plants, which was proved to exert various effects on plants (Llave et al., 2002). Previous studies showed that miRNAs are shown to take part in the regulation of growth and development, as well as acclimatization to adversity stress. For instance, miR396e and miR396f are two important regulators of grain size and plant architecture in rice (Miao et al., 2020). The upregulated miR160, miR167, and miR393 might be participated in the regulation of primary root length and lateral root number under N deficiency (Li L. J. et al., 2021). Cui et al. (2020) showed that miR1885 dynamically regulates both innate immunity and plant growth and responds to viral infection, by targeting both TIR-NBS-LRR and a photosynthesis-related gene for negative regulation through distinct modes of action (Cui et al., 2020). Amounting studies have demonstrated that miRNAs help plants to adapt to drought stress (Bhat et al., 2020), cold stress (Yang et al., 2017), heat stress (Pan et al., 2017), and heavy metal toxicity (Celik and Akdas, 2019). In recent years, several studies showed that miRNA–target modules played important roles in response to nutritive stress. For instance, the miR396e/f-GRF4/6/8 module takes part in the regulation of yield, N assimilation, and utilization in rice, and mir396e/f mutants have higher N content and yield in N-deficient conditions (Zhang et al., 2020). TamiR2275 was induced by N starvation stress, which regulated N acquisition and photosynthetic function when plants are challenged by N deprivation (Qiao et al., 2018). Our previous study demonstrated that a potential regulatory mechanism of the miR169c-*NFYA10* module is involved in the low-nitrogen (LN) stress response of *B. luminifera* (Cheng et al., 2021). However, the expression profiles of miRNAs in response to LN stress are waiting to be elucidated in *B. luminifera*.

In recent years, the widespread high-throughput sequencing strategy was used for sRNAome (Wen et al., 2022). The ability to detect low-expressed miRNAs contributes to the tremendous progress of miRNAs, especially in non-model plant species (Dos Santos et al., 2019; Li H. et al., 2021). Furthermore, the combination of sRNA sequencing and degradome sequencing promotes the identification of miRNA target genes (Charles Addo-Quaye, 2008). In our previous study, we identified 114 known miRNAs and defined 49 targets for 26 miRNA families in *B. luminifera* (Zhang et al., 2016), and 44 miRNAs and 71 corresponding target genes were identified in response to heat stress (Pan et al., 2017). Therefore, in the present study, we sequenced and analyzed sRNA libraries and roots and shoots under control, 0.5 h N-starved, and 24 h N-starved treatments, respectively. We tested two hypotheses: 1) The abundance of miRNAs is affected by N starvation, and these miRNAs might have regulatory roles in N starvation adaptation in *B. luminifera* by regulating the abundance of their corresponding target genes. 2) The expression patterns of N starvation-responsive miRNAs change with stress duration.

MATERIALS AND METHODS

Plant Cultivation and Low-Nitrogen Treatments

The *B. luminifera* G49[#] seedlings were propagated *in vitro* in 1/2 MS at the temperature of 25 ± 2°C with a 16-h photoperiod. After 2-month growth, seedlings were transferred to a mixed substance consisting of perlite, peat, and clay loam in proportion to 1:1:2, under 120 μmol m^{-2}•s^{-1} light for 2 weeks. Then, 27 uniformly growing individuals were selected and randomly divided into two groups, namely, CK and LN. Seedlings in the CK group were cultivated using half-strength liquid MS media including 15 mm KNO_3, whereas the seedlings in the LN group were cultivated using half-strength MS media containing 0.03 mm KNO_3 (the only source of N), as described previously (Cheng et al., 2021). Then, the leaves and roots were sampled individually after 0.5 and 24 h after the CK and LN treatments, at 11:00 a.m., respectively. Three randomly selected seedlings were pooled together as one replicate, with three replicates per group, and samples were flash frozen in liquid N and stored at −80°C until RNA extraction.

The Expression Patterns of Low-Nitrogen-Responsive Marker Genes

Total RNA was extracted using TRIZOL (Invitrogen, Carlsbad, CA, United States) and was then reverse transcribed to cDNA using PrimeScript RT Reagent Kit (TaKaRa, Dalian, China). To verify the effectiveness of LN stress treatment, two marker genes *NRT2.5* and *NRT2.7* were quantified via qRT-PCR

(**Supplementary Table S1**). The large ribosomal subunit 39 (*RPL39*) was used as a reference gene to normalize the expression levels of marker genes (Cheng et al., 2021).

sRNA Library Construction, Sequencing, and Bioinformatics

The construction and sequencing of the sRNA library were accomplished by LC Bio (Hangzhou, China) with Illumina Hiseq 2500. After the quality assessment of sequencing data, the retained data were processed using ACGT101-miR (LC Sciences, Houston, Texas, United States), including removing adaptor dimers, junk sequences, and low complexity sequences and filtering out mRNA, Rfam, and Repbase. Then, the remained sequences were blasted against miRBase21.0 and genome, and the hairpin structure of predicted miRNA precursor sequences was validated using RNAfold software (http://rna.tbi.univie.ac. at/cgi-bin/RNAfold. Cgi).

Analysis of Differentially Expressed miRNAs

The reads of miRNAs from the control, 0.5 h, and 24 h of LN treatment were normalized to tags per million. The combination of log2 ratio and Fisher's exact test was used to identify the differential expressed miRNAs (DEmiRNAs). Only when the absolute value of log2 ratio ≥ 1 and $p \leq 0.05$, the miRNA was significantly differentially expressed; however, in other cases, there was no significant difference in expression (Xie et al., 2015).

Target Prediction of DEmiRNAs and Functional Classification

Target prediction of DEmiRNAs both known and novel miRNAs was performed using Targetfinder (http://targetfinder.org/), and the mismatch score was less than four. Based on our published degradome sequencing data of *B. luminifera* (GEO: GSE80074), we used CleaveLand 3.0.1 pipeline to identify and classify the target genes of DEmiRNAs. To further elucidates the potential regulation of target genes, GO (Gene Ontology) analysis was conducted using agriGO2.0 (http://bioinfo.cau.edu.cn/agriGO/). In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation was conducted using a database (http://www.genome.jp/kegg/kegg1.html).

The Expression Patterns of miRNAs and Their Target Genes

Total RNA was extracted as described above and was reversely transcribed with Mir-X miRNA First-Strand Synthesis Kit (TaKaRa, Dalian, China). qRT-PCR was conducted using SYBR Premix EX Taq Kit (TaKaRa). Gene *U6* was used as a reference gene to normalize the expression levels of miRNAs. The RNA was reversely transcribed to cDNA using PrimeScript RT Reagent Kit (TaKaRa), which was used to quantify the target genes. The *RPL39* was used as a reference gene, and the relative expression levels of genes were calculated *via* the $2^{-\Delta\Delta Ct}$ method. The qRT-PCR primers are shown in **Supplementary Table S1**.

RESULTS

The Expression Patterns of Low-Nitrogen-Responsive Marker Genes in *B. luminifera*

To confirm the physiological status of the *B. luminifera* seedlings (i.e., to confirm whether the plants were subjected to LN stress), the expression levels of the marker genes *BlNRT2.5* and *BlNRT2.7* were analyzed. *BlNRT2.5* and *BlNRT2.7* as HATS, responsible for nitrate transport and absorption under LN, are indicators of whether the plants significantly underwent LN stress, as reported previously (Wang et al., 2012). In both root and shoot, the expression level of *BlNRT2.5* was increased significantly after 0.5 h of LN treatment and reached the peak value after 24 h of LN treatment (**Figure 1A**). For the *BlNRT2.7*, the expression level was induced 10-fold after 0.5 h of LN treatment and remained high level after 24 h of N starvation in the root, whereas the expression level was unchanged until 24 h of LN treatment in the shoot (**Figure 1B**).

sRNA Profiles in Response to Low-Nitrogen Stress

To investigate the response of miRNAs to N starvation, six mixed sRNA libraries (each library was a pool of RNAs from three biological samples) were generated from roots and shoots under LN stress for 0, 0.5, and 24 h, respectively, and then were sequenced via Solexa high-throughput sequencing technology. After the remove of adaptor and junk sequences, a total of 66,031,208 high-quality reads were obtained and then mapped to known RNAs, resulting in 2,305,703 (461,778), 5,012,446 (1,028,698), 2,663,870 (514,935), 3,475,390 (785,402), 4,953,286 (1,468,323), and 1,322,911 (319.385) redundant (unique) valid reads, which were used to identify miRNAs (**Supplementary Table S2**).

Length Distribution of sRNAs Under CK and Low-Nitrogen Conditions

Most of the sRNA sequences ranged from 18 to 24 nt, representing 97.87% (96.58%) of redundant (unique) valid reads, which was consistent with the characteristic of the DCL1 cleavage product. In the root, 24-nt sRNAs had the highest proportion of sRNAs in the unique valid reads (**Figure 2A**). Under LN stress, the length distribution patterns of sRNAs in the root were altered, which was exemplified by the fact that the ratio of 24-nt sRNAs increased from 26.31% (0 h) to 40.42% (0.5 h) and then recovered to 23.24% (24 h). A similar phenomenon was observed in the shoot, which was exemplified by the fact that 24-nt sRNAs accounted for 32.26% in 0 h and increased to 45.49% after 0.5 h of LN stress and then reduced to 12.20% (**Figure 2B**). However, the 18-nt sRNAs had the highest proportion, accounting for 21.66% in the shoot after 24-nt LN stress.

For redundant valid reads, 21-nt sRNAs were the most abundant class of sRNAs in the root, and the ratio of 24-nt

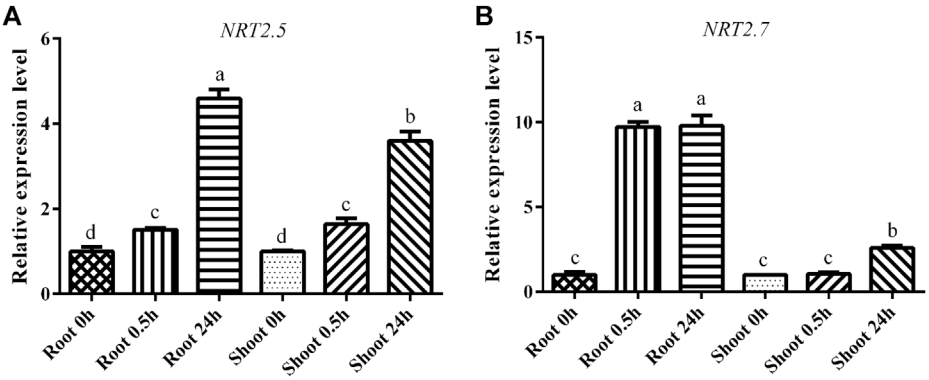


FIGURE 1 | The expression levels of marker genes under low nitrogen (LN) treatment. **(A)** The expression level of *BlNRT2.5*; **(B)** The expression level of *BlNRT2.7* under LN treatment. The reference gene was *BlRPL39*, each sample had three replicates, and error bars were standard deviation.

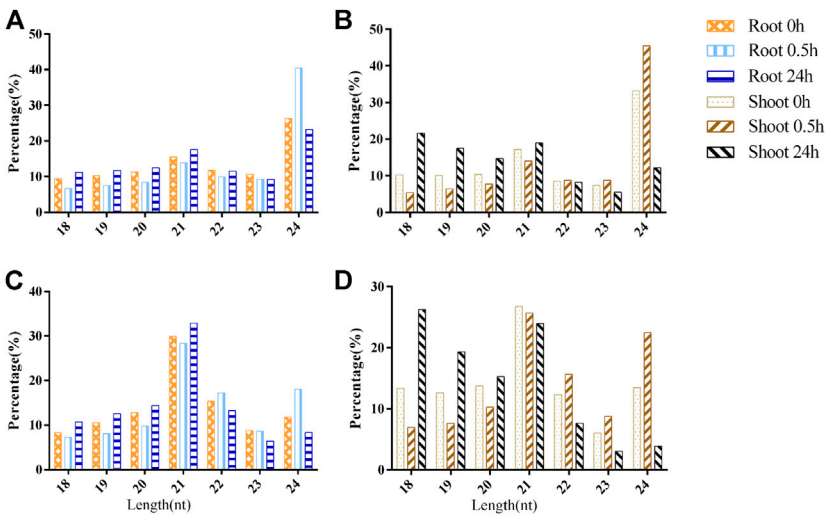


FIGURE 2 | The length distribution of sRNAs under low-nitrogen stress. **(A)**. Unique valid reads in root; **(B)**. Unique valid reads in shoot; **(C)**. Redundant valid reads in root; and **(D)**. Redundant valid reads in shoot.

sRNAs increased after 0.5 h of LN treatment (**Figure 2C**). In the shoot, 21-nt sRNAs were also the most abundant class, accounting for 26.76 and 25.69% in the 0 and 0.5 h of LN treatment, respectively. However, the ratio of 18-nt sRNAs reached the peak level after 24 h of LN stress (**Figure 2D**).

Identification of Known and Novel miRNAs in *B. luminifera*

A total of 155 known miRNAs from 55 families were identified, and 43 novel miRNAs from 32 families were discovered, resulting in 198 miRNAs from 87 families. Among them, 113 known miRNAs were identified with according precursors, whereas other 42 known miRNAs lacked the corresponding precursors, and miR166 was the most abundant miRNA family (**Supplementary Table S3**).

miRNA Expression Response to Low-Nitrogen Stress

In the shoot, 76 known miRNAs from 42 families and 22 novel miRNAs from 20 families were recognized as DEmiRNAs after 0.5 h and/or 24 h N-starved treatment, whereas 98 known miRNAs from 46 families and 31 novel miRNAs from 24 families were differentially expressed in the root (**Supplementary Table S4**). To be precise, 54 and 69 DEmiRNAs were identified in the shoot, whereas 105 and 72 DEmiRNAs were detected in the root after 0.5 and 24 h N-starved treatment, respectively (**Figure 3A**). Among them, 22 and 44 DEmiRNAs were specifically expressed in the shoot after 0.5 and 24 h treatment, respectively, whereas 48 and 38 DEmiRNAs were specifically identified in the root (**Figure 3B**, **Supplementary Table S4**).

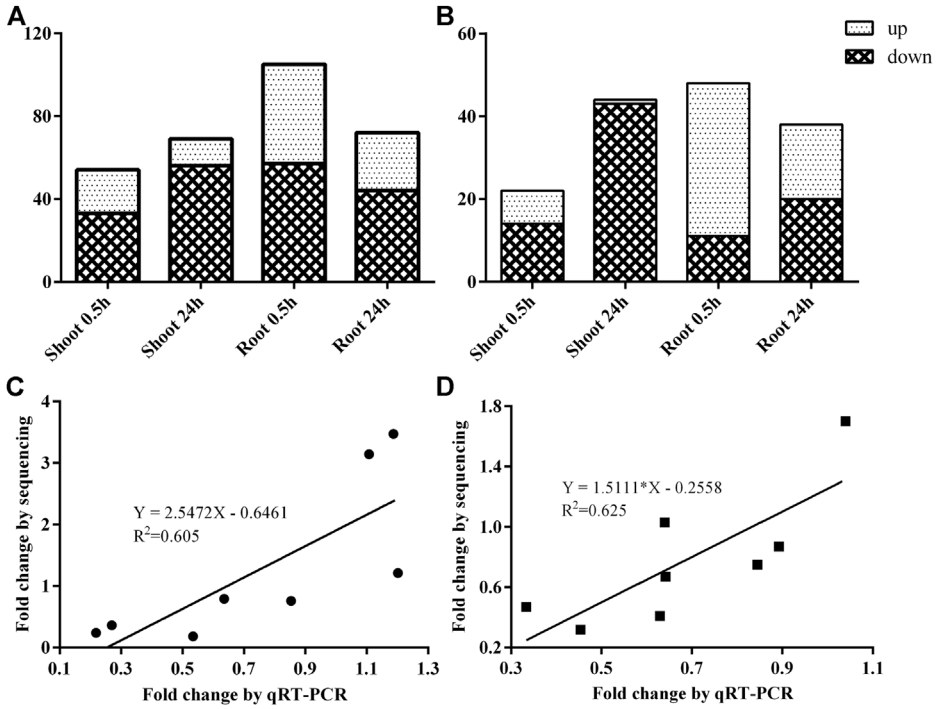


FIGURE 3 | The identification and validation of DEmiRNAs after 0.5 and 24 h N-starved treatment. **(A)** Total number of DEmiRNA in shoot and root; **(B)** The specific DEmiRNA in shoot and root; **(C)** The validation of DEmiRNAs after 0.5 h LN treatment; and **(D)** The validation of DEmiRNAs after 24 h of low nitrogen treatment.

To validate the results of the sRNA sequencing, eight DEmiRNAs were selected randomly for qRT-PCR analysis. Although the fold-changes of DEmiRNAs between qPCR and sRNA sequencing were different, similar expression tendencies were observed (**Figures 3C,D**).

Target Genes Regulated by DEmiRNA

A total of 167 miRNAs were predicted to target 836 genes, involving 1,094 miRNA–target pairs. Based on degradome data, 122 miRNAs including 102 known miRNAs and 20 novel miRNAs targeted 203 genes, comprising 321 miRNA–target pairs (**Supplementary Table S5**). These targets were divided into five categories: categories 0–4 (**Figure 4**). For known miRNAs, most of the target genes were annotated as transcription factors. For instance, *SPL13A* (Squamosa promoter-binding protein-like, *SPL*) was targeted by three miR156, *NAC021* (NAM, ATAF, CUC, *NAC*) by miR164a, *NFYA2/7* (*Nuclear* Transcription Factor Y Subunit A, *NFYA*) by five miR169, *RAP2–7* (Ethylene-responsive transcription factor, *RAP*) by miR172a/b, and *GRF4/6* (Growth-Regulating Factor, *GRF*) by miR396a/b. A big proportion of target genes for novel miRNAs were proteins or enzymes. For example, *NSP2* (Nodulation-Signaling Pathway 2 protein) was targeted by blu-miR1-3p, *CMTA2* (CalModulin-binding Transcription Activator 2-like) by blu-miR5-3p, *OPT3* (OligoPeptide Transporter 3) by blu-miR7-3p, and FBX (F-box/kelch-repeat protein) by blu-miR8-3p.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Pathway Analyses of Target Genes

GO was used to predict the function of target genes. A total of 665 target genes were annotated with 1,115 GO terms (**Supplementary Table S6**). Among them, a large proportion (54.08%) of target genes were classified into biological processes, one-third of genes for molecular function (327 terms), and 169 items belonging to cellular components. Of these, 13 GO terms were extremely significantly enriched ($p < 0.01$), including leaf development, auxin-activated signaling pathway, cell differentiation, transcription factor activity, DNA binding, and nucleus (**Figure 5A**). One KEGG term "ascorbate and aldarate metabolism" (ko00053) was significantly enriched ($p < 0.05$) (**Figure 5B**).

Expression Patterns of miRNAs and Their Target Genes Under Low-Nitrogen Stress

The expression patterns of six miRNAs and their target genes were monitored after 0.5 and 24 h of LN treatment (**Figure 6**). The expression levels of miR156a and miR166b-5p were initially decreased after 0.5 h of N starvation but increased after 24 h of LN treatment, whereas the expression patterns of blu-miR1-3p and blu-miR8-3p showed an opposite trend. However, miR164a and miR167a exhibited continuous decreasing expression patterns under LN treatment. To determine the relationships between

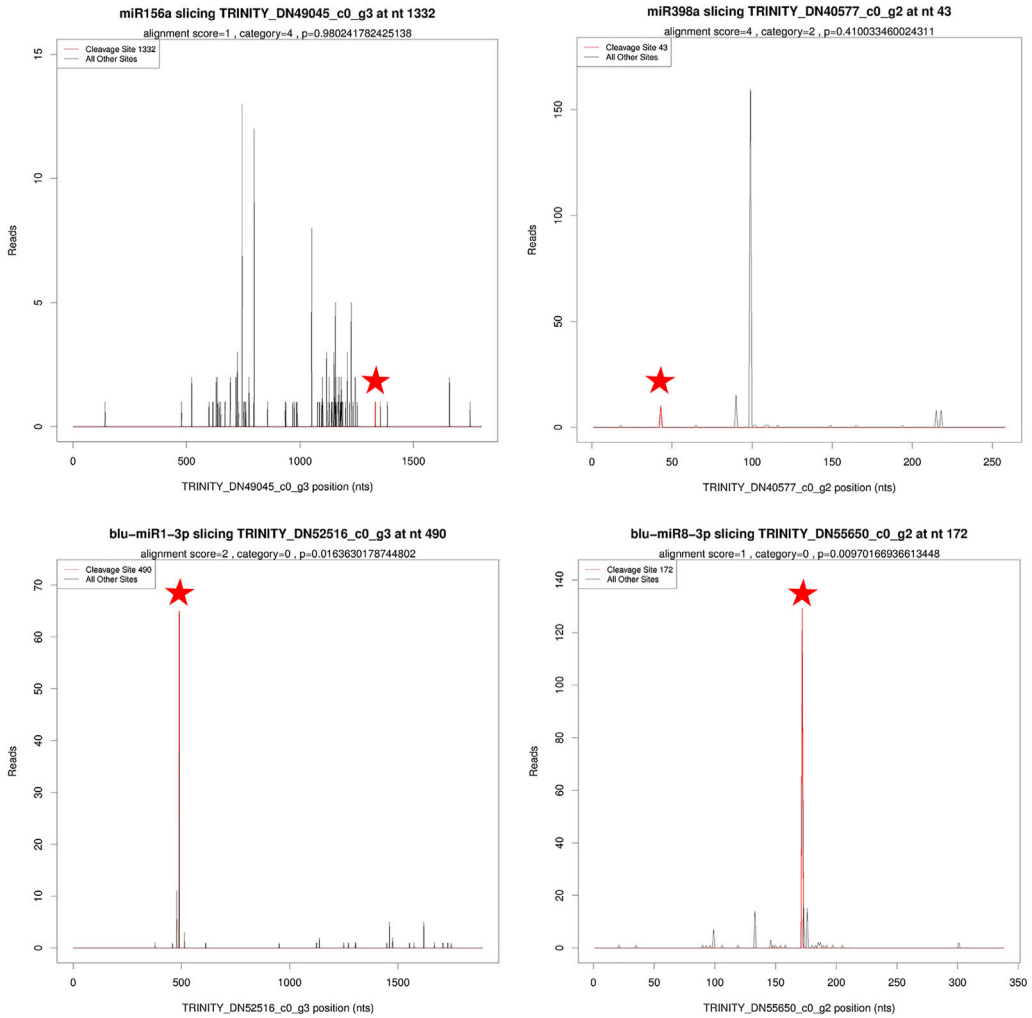


FIGURE 4 | T-plots of miRNA targets based on degradome sequencing. Asterisks were used to mark cleavage sites of target genes.

miRNAs and target genes, we examined the expression levels of target genes under LN treatment using qRT-PCR (**Figure 6**). The overall expression levels of target genes were unchanged or decreased slightly after 0.5 h of LN stress but were induced significantly by 24 h of LN stress.

DISCUSSION

N is one of the most important restrictive elements in plant growth and development, and nitrate is the major form for most land plants (Glass et al., 2002). The limited nitrate in soil restricts plant growth; thus, it is of great importance to understand how plants respond to nitrate deficiency. Amounting studies showed that miRNAs played important roles in response to N starvation, such as *Arabidopsis* (Zhao et al., 2011), wheat (Liu et al., 2021), potato (Tiwari et al., 2020), and rice (Yu et al., 2018).

Length Distribution of sRNAs Varied Under Low-Nitrogen Stress

Several plant species, such as *Arabidopsis*, cotton, and Chinese white poplar, the most redundant sRNAs, were 21-nt sized. In the present study, 21-nt sRNAs were the most abundant class of sRNAs in both root and shoot of *B. luminifera* (**Figures 2C,D**). A similar phenomenon was observed in our previous study (Zhang et al., 2016), whereas the 24-nt class of sRNAs was the most dominant in heat-stressed *B. luminifera* (Pan et al., 2017). It might be attributed to developmental- and/or stress-specific differences. In the *Chrysanthemum nankingense*, 21- and 24-nt sRNAs were the most abundant classes, and the proportion of 24-nt sRNA in the root was higher than that of 21-nt sRNA, but it was the opposite in leaves (Song et al., 2015). In potatoes, 24-nt sRNAs accounted for the largest proportion in all classes, followed by 21-nt sRNAs (Tiwari et al., 2020). In maize, the proportion of 21-nt sRNA was the highest under all treatments of shoots, which was the same in the root under control treatment,

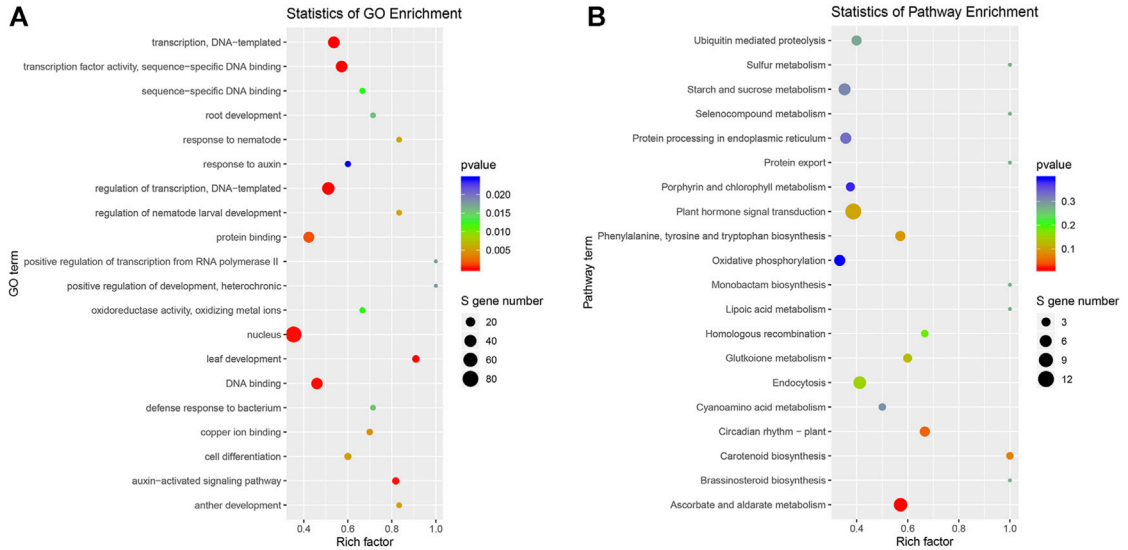


FIGURE 5 | Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of genes targeted by DEmiRNAs. **(A)** GO enrichment analysis; **(B)** KEGG pathway enrichment analysis of genes targeted by DEmiRNAs.

but the sRNA of the root was mainly distributed in 24 and 22 nt under LN treatment (Zhao et al., 2012). Moreover, the ratio of 24-nt sRNAs increased in the root after 0.5 h of LN treatment (**Figure 2C**), and the ratio of 18-nt sRNAs reached the peak level after 24 h of LN stress in the shoot (**Figure 2D**). It was consistent with the major peak at 18 nt in *B. luminifera* roots under LP stress (Zhang et al., 2021). A previous study showed that the components of sRNAs with the size of 18–19 nt originated from tRNAs, and the accumulation of tRNA fragments in Pi-starved (Hsieh et al., 2009) and nitrate-starved *Arabidopsis* (Liang et al., 2012). Moreover, tRNA halves were detected in the phloem sap of *Cucurbita maxima*, which inhibited translational activity *in vitro*, indicating that these tRNA halves coordinate the metabolic status between source and sink tissues, acting as long-distance signals (Zhang et al., 2009). Loss-Morais et al. (2013) showed that tRNA fragments can be incorporated into argonaute (AGO) complexes, thus regulating gene expression post-transcriptionally, acting as miRNAs (Loss-Morais et al., 2013). Therefore, these 18-nt sRNAs induced by LN stress might play important regulatory roles in response to N starvation in *B. luminifera*, which need to be further studied.

miRNA Expression Response to Nitrogen Starvation

Previous studies have identified a series of LN-responsive miRNAs. For instance, 102 miRNAs from 42 families in maize (Zhao et al., 2012) and 404 and 628 putative miRNAs were identified in potato roots and shoots, respectively (Tiwari et al., 2020). Although three studies on miRNA of *B. luminifera* have been conducted (Zhang et al., 2016; Pan et al., 2017; Zhang et al., 2021), the LN-responsive miRNAs were waiting to be identified. In the present study, a total of 198 miRNAs from 87 families were identified, of which 113 known miRNAs were identified with precursors. Thus, it enriches the miRNA information of *B. luminifera*.

miR166 was the most abundant miRNA family, and miR166a accounted for the largest proportion, which was consistent with the result of Chinese cabbage (Ahmed et al., 2019), whereas miR166a-3p was the most abundant miRNA in our previous study (Pan et al., 2017). The inconsistency might relate to the alternative cleavage mechanism depending on the plant species, growth stage, and/or stress. For example, the miR535 family had the highest abundance, and miR156, miR166, and miR168 families also had high expression levels in bananas (Zhu et al., 2019). Several miRNAs were specifically expressed in some samples (**Supplementary Table S4**), such as 156 d-3p in a 0.5 h LN-treated shoot. A similar phenomenon was observed in potatoes, where five conserved miRNAs were specifically expressed in roots and nine in shoots. It indicated that the accumulation of some miRNAs is dependent on plant species, tissues, and/or stresses. In the present study, more downregulated DEmiRNAs were observed under LN stress (**Figure 3**). A similar phenomenon was observed in sorghum and durum wheat, which was exemplified by the fact that the number of downregulated DEmiRNAs was more than that of upregulated DEmiRNAs under LN stress (Liu et al., 2021; Zhu et al., 2021). The GO analysis of DEmiRNA targets showed that leaf development, auxin-activated signaling pathway, cell differentiation, and transcription factor activity were enriched (**Figure 5A**), and the KEGG term "ascorbate and aldarate metabolism" (ko00053) was significantly enriched (**Figure 5B**). It suggested that miRNAs play important roles in leaf development, hormone signaling pathway, cell differentiation, ascorbate and aldarate metabolism, and transcriptional regulations in response to LN stress in *B. luminifera*.

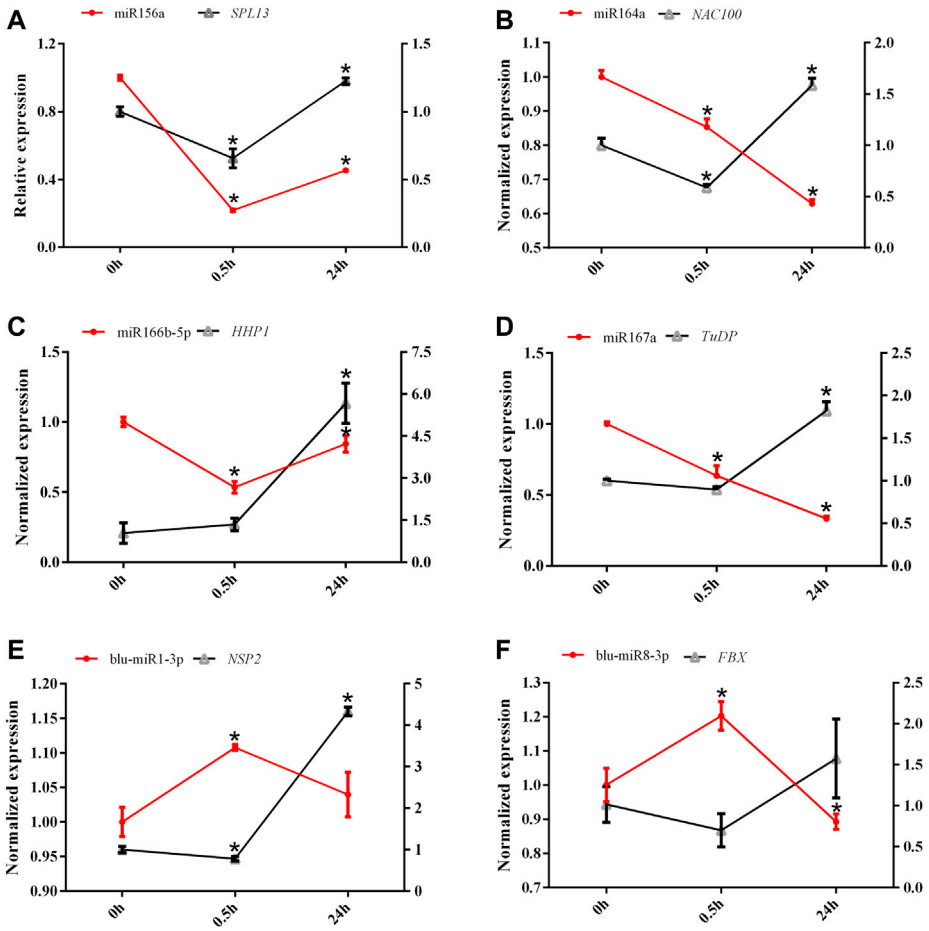


FIGURE 6 | The expression levels of six miRNAs and their predicted target genes under low-nitrogen stress. The left *y* axis represents the expression level of miRNA, and the right *Y* axis represents the expression level of the predicted target gene. All expression was normalized by *BlRPL39*, and three replicates were conducted. **(A)** miR156a-*SPL13*; **(B)** miR164a-*NAC100*; **(C)** miR166b-5p-*HHP1*; **(D)** miR167a-*TuDP*; **(E)** miR1-3p-*NSP2*; **(F)** miR8-3p-*FBX*.

Expression Patterns of miRNAs and Targets in *B. luminifera* Under Low-Nitrogen Stress

Under LN stress, plants change their phenotypes and transcriptional activity (Curci et al., 2017). Amounting studies showed that miRNAs play essential roles in response to LN stress by cleaving target mRNAs or inhibiting translation of mRNAs. In our study, a series of miRNAs and validated target genes based on degradome sequencing might take part in response to environment starvation; thus, a putative model on miRNA–target interactions under LN stress was proposed (**Figure 7**). The biogenesis of miRNAs was processed by DCL1, which was regulated by miR162 (Xie et al., 2003). Then, the mature miRNA entered the AGO to regulate target genes. AGO1 was proved to be targeted by miR168, which was verified in this study. In addition, AGO2 was targeted by miR403, as described in our previous study (Pan et al., 2017). Therefore, the miRNA biogenesis and function were feedback regulated by miR162, miR168, and miR403, when the *B. luminifera* encounter the LN stress, thus helping plants to adapt to N starvation condition. In the present study, most of the target genes for known miRNAs were TFs, such as *ARFs*, *HD-Zips*, *NFYAs*, and *SPLs* (**Supplementary Table S5**), whereas some targets were protein-coding genes, such as Disease resistance protein (RGA), Threonine Synthase (*TS1*) and Laccase (*LACs*). These miRNA–target modules played important roles in transcriptional regulation, auxin response, energy metabolism, biosynthesis, and signal transduction.

The aging pathway in flowering regulation is controlled mainly by miR156. *SPLs* were targeted by miR156, and overexpression of *SPL7* and *SPL8* promotes flowering, and vice versa (Gou et al., 2019). miR156 was induced by LN stress, and the miR156-SPL3/SPL12 module directly activates *MADS50* in the node to regulate crown root development in rice (Shao et al., 2019). In general, miR156 is induced by LN in *Chrysanthemum nankingense* (Song et al., 2015). In our study, miR156s could be divided into two groups based on expression patterns, which was exemplified by the fact that one group was downregulated

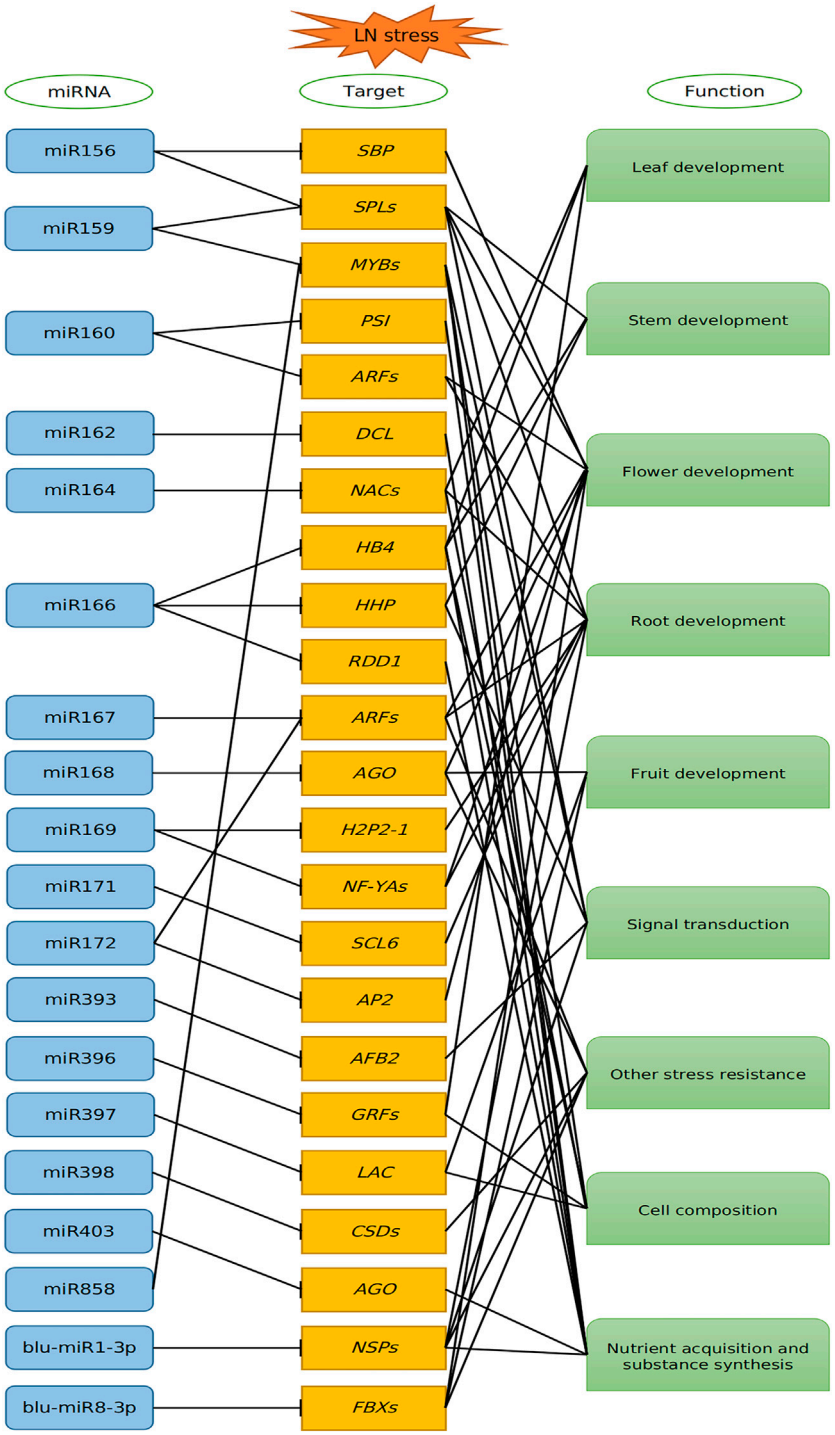


FIGURE 7 | Proposed model of the miRNA-mediated regulatory network in response to low nitrogen in *B luminifera*. The miRNAs negatively regulated their target genes, and these targets are involved in leaf, stem, root development, signal transduction, stress resistance, cell composition and nutrient acquisition, and substance synthesis.

continuously during the treatment period, whereas the other group was upregulated after 0.5 h treatment and then downregulated after 24 h, which was consistent with the miR156 expression in poplar (Bao et al., 2019). miR156a was repressed under LN stress, and *SPL13* was downregulated after 0.5 h but was upregulated after 24 h of LN stress, which might promote the transition from the vegetative to the reproductive stage under LN stress in *B. luminifera*.

A previous study showed that miR164 played a negative role in adventitious root development by downregulating the expression of *NAC1* (Li J. et al., 2021). A NAC gene *EjNACLl47* had a positive effect on cell expansion and organ enlargement in *Eriobotrya japonica* (Chen Q. et al., 2021). Besides, *NAC* is related to leaf senescence; for instance, *GmNAC81* was upregulated during the leaf senescence, and the overexpression of *GmNAC81* accelerated leaf senescence by accumulating H_2O_2 (Ferreira et al., 2020). Leaf senescence is a process of recycling nutrients for other organs (Lim et al., 2007). In the present study, the expression level of miR164 was decreased under LN stress, whereas the target *NAC100* was inhibited after 0.5 h and then was induced significantly after 24 h of LN stress (**Figure 6**). It suggested that miR164 might play an important role in leaf senescence mainly by targeting *NAC100*, to accelerate the N recycling under LN conditions in *B. luminifera*.

The Class III homeodomain Leu zipper (HD-Zip III) gene family plays essential roles in plant growth and development. For example, overexpression of *HOX32* (an HD-Zip III family member) resulted in curled leaf, reduced leaf angle, semi-dwarf phenotype, and reduced photosynthetic capacity. OsHox32 may regulate plant architecture and leaf development by regulating *YABBY* genes in rice (Li et al., 2016). HD-Zip III including *Hox32* targeted by miR166 and the OsmiR166b-OsHox32 module plays important roles in plant growth and development and plant architecture by regulating the cell wall-related gene expression (Chen H. et al., 2021). However, miR166b-5p was validated to target *HHP1* (Heptahelical protein) in *B. luminifera*. Previous studies showed that *HHP* participated in the signal transduction pathway by cooperating with *MYB* to transduce hormone and environmental signals to cope with cold stress (Lee and Seo, 2015). Therefore, we speculated that the inhibited miR166b-5p to ensure the accumulated target *HHP1* thus took part in the signal transduction pathway to cope with LN stress in *B. luminifera*.

blu-miR1-3p was validated to target *NSP* (nodulation-signaling pathway), an GRAS (GAI, RGA, SCR)-type transcription factor, which was a constituent of the Myc signaling pathway (Delaux et al., 2013). NSP is generally related to nodule symbiosis (Singh and Verma, 2021) and also participated in other stress responses, such as drought stress (Yao et al., 2019). Liu et al. (2011) showed that *NSP1* and *NSP2* are indispensable for strigolactone (SL) biosynthesis both in the legume and rice, which was exemplified by the fact that NSP2 is essential for the conversion of orobanchol into didehydro-orobanchol, dependent on DWARF27, a gene essential for SL biosynthesis (Liu et al., 2011). Therefore, we speculated that novel miRNA blu-miR1-3p might participate in SL biosynthesis by targeting the NSP gene, when *B. luminifera* encounters N starvation. However, the response mechanism of blu-miR1-3p under LN stress still needs to be further studied.

CONCLUSION

Integrated analysis of sRNA-Seq and degradome-Seq was conducted to elucidate the molecular mechanisms in response to LN stress in *B. luminifera*. sRNA sequencing identified 98 known miRNAs, and 31 novel miRNAs were differentially expressed after 0.5 or 24 h of LN stress. A total of 122 DEmiRNAs targeted 203 genes, comprising 321 miRNA–target pairs. Thus, a miRNA-mediated network was proposed. Taken together, these findings provide useful information to elucidate miRNA functions and establish a framework for exploring N signaling networks mediated by miRNAs in *B. luminifera* and other woody plants.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the CNGB Sequence Archive (CNSA) of China National GeneBank DataBase (CNGBdb) repository, accession number CNP0003144.

AUTHOR CONTRIBUTIONS

JZ and ZT conceived and designed the experiments. YL, JZ, and SC analyzed the sRNA sequencing data. HH, ZT, and JZ contributed reagents and materials. YL, XX, and XH performed the qRT-PCR experiment. YL, JZ, and SC wrote and revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (31971677), Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding (2021C02070-10), and National Key Research and Development Program of China (2021YFD2200304-2).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.957505/full#supplementary-material

REFERENCES

Addo-Quaye, C., Eshoo, T. W., and Axtell, D. P. M. J. (2008). Endogenous siRNA and miRNA Targets Identified by Sequencing of the *Arabidopsis* Degradome. *Curr. Biol.* 18, 758–762. doi:10.1016/j.cub.2008.04.042

Ahmed, M., Rauf, M., Mukhtar, Z., and Saeed, N. A. (2017). Excessive Use of Nitrogenous Fertilizers: an Unawareness Causing Serious Threats to Environment and Human Health. *Environ. Sci. Pollut. Res.* 24, 26983–26987. doi:10.1007/s11356-017-0589-7

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

Bao, H., Sun, F. S., Xu, Q. H., and Wang, Y. W. (2019). Differential Expression of 10 miRNAs in Poplar under Low Nitrogen Stress and the Identification and Analysis of Target Genes. *Mol. Plant Breed.* 17, 771. doi:10.13271/j.mpb.017.000771

Bhat, K. V., Mondal, T. K., Gaikwad, A. B., Kole, P. R., Chandel, G., and Mohapatra, T. (2020). Genome-wide Identification of Drought-Responsive miRNAs in Grass Pea (*Lathyrus Sativus* L.). *Plant gene.* 21, 100210. doi:10.1016/j.plgene.2019.100210

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

Çelik, Ö., and Akdaş, E. Y. (2019). Tissue-specific Transcriptional Regulation of Seven Heavy Metal Stress-Responsive miRNAs and Their Putative Targets in Nickel Indicator castor Bean (*R. Communis* L.) Plants. *Ecotoxicol. Environ. Saf.* 170, 682–690. doi:10.1016/j.ecoenv.2018.12.006

Chen, H., Fang, R., Deng, R., and Li, J. (2021a). The OsmiRNA166b-OsHox32 Pair Regulates Mechanical Strength of Rice Plants by Modulating Cell Wall Biosynthesis. *Plant Biotechnol. J.* 19, 1468–1480. doi:10.1111/pbi.13565

Chen, Q., Jing, D. L., Wang, S. M., Xu, F., Bao, C. Y., Luo, M., et al. (2021b). The Putative Role of the NAC Transcription Factor EjNACL47 in Cell Enlargement of Loquat (*Eriobotrya Japonica* Lindl.). *Horticulturae* 7, 90323. doi:10.3390/horticulturae7090323

Cheng, L. L., Wu, F. M., Lin, Y., Han, X., Xu, X. S., Zhang, Y. T., et al. (2021). A miR169c-NFYA10 Module Confers Tolerance to Low-Nitrogen Stress to *Betula Luminifera*. *Industrial Crops Prod.* 172, 113988. doi:10.1016/j.indcrop.2021.113988

Cui, C., Wang, J.-J., Zhao, J.-H., Fang, Y.-Y., He, X.-F., Guo, H.-S., et al. (2020). A Brassica miRNA Regulates Plant Growth and Immunity through Distinct Modes of Action. *Mol. Plant* 13, 231–245. doi:10.1016/j.molp.2019.11.010

Curci, P. L., Aiese Cigliano, R., Zuluaga, D. L., Janni, M., Sanseverino, W., and Sonnante, G. (2017). Transcriptomic Response of Durum Wheat to Nitrogen Starvation. *Sci. Rep.* 7, 1176. doi:10.1038/s41598-017-01377-0

Delaux, P. M., Bécard, G., and Combier, J. P. (2013). NSP1 Is a Component of the Myc Signaling Pathway. *New Phytol.* 199, 59–65. doi:10.1111/nph.12340

Dos Santos, T. B., Soares, J. D. M., Lima, J. E., Silva, J. C., Ivamoto, S. T., Baba, V. Y., et al. (2019). An Integrated Analysis of mRNA and sRNA Transcriptional Profiles in Coffea Arabica L. Roots: Insights on Nitrogen Starvation Responses. *Funct. Integr. Genomics* 19, 151–169. doi:10.1007/s10142-018-0634-8

Ferreira, D. O., Fraga, O. T., Pimenta, M. R., Caetano, H. D. N., Machado, J. P. B., Carpinetti, P. A., et al. (2020). GmNAC81 Inversely Modulates Leaf Senescence and Drought Tolerance. *Front. Genet.* 11, 601876. doi:10.3389/fgene.2020.601876

Glass, A. D. M., Britto, D. T., Kaiser, B. N., Kinghorn, J. R., Kronzucker, H. J., Kumar, A., et al. (2002). The Regulation of Nitrate and Ammonium Transport Systems in Plants. *J. Exp. Bot.* 53, 855–864. doi:10.1093/jexbot/53.370.855

Gou, J., Tang, C., Chen, N., Wang, H., Debnath, S., Sun, L., et al. (2019). SPL 7 and SPL 8 Represent a Novel Flowering Regulation Mechanism in Switchgrass. *New Phytol.* 222, 1610–1623. doi:10.1111/nph.15712

Hsieh, L.-C., Lin, S.-I., Shih, A. C.-C., Chen, J.-W., Lin, W.-Y., Tseng, C.-Y., et al. (2009). Uncovering Small RNA-Mediated Responses to Phosphate Deficiency in *Arabidopsis* by Deep Sequencing. *Plant Physiol.* 151, 2120–2132. doi:10.1104/pp.109.147280

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

Lee, H. G., and Seo, P. J. (2015). The MYB 96- HHP Module Integrates Cold and Abscisic Acid Signaling to Activate the CBF - COR Pathway in Arabidopsis. *Plant J.* 82, 962–977. doi:10.1111/tpj.12866

Li, H., Meng, H., Sun, X., Deng, J., Shi, T., Zhu, L., et al. (2021a). Integrated microRNA and Transcriptome Profiling Reveal Key miRNA-mRNA Interaction Pairs Associated with Seed Development in Tartary Buckwheat (*Fagopyrum Tataricum*). *BMC Plant Biol.* 21, 132. doi:10.1186/s12870-021-02914-w

Li, J., Zhang, H., Zhu, J., Shen, Y., Zeng, N., Liu, S., et al. (2021b). Role of miR164 in the Growth of Wheat New Adventitious Roots Exposed to Phenanthrene. *Environ. Pollut.* 284, 117204. doi:10.1016/j.envpol.2021.117204

Li, L. J., Li, Q., Davis, K. E., Patterson, C., Oo, S., Liu, W. Y., et al. (2021c). Response of Root Growth and Development to Nitrogen and Potassium Deficiency as Well as microRNA-Mediated Mechanism in Peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* 12, 695234. doi:10.3389/fpls.2021.695234

Li, Y.-y., Shen, A., Xiong, W., Sun, Q.-l., Luo, Q., Song, T., et al. (2016). Overexpression of OsHox32 Results in Pleiotropic Effects on Plant Type Architecture and Leaf Development in Rice. *Rice* 9, 46. doi:10.1186/s12284-016-0118-1

Liang, G., He, H., and Yu, D. (2012). Identification of Nitrogen Starvation-Responsive microRNAs in *Arabidopsis thaliana*. *PLoS One* 7, e48951. doi:10.1371/journal.pone.0048951

Lim, P. O., Kim, H. J., and Gil Nam, H. (2007). Leaf Senescence. *Annu. Rev. Plant Biol.* 58, 115–136. doi:10.1146/annurev.arplant.57.032905.105316

Liu, H., Able, A. J., and Able, J. A. (2021). Nitrogen Starvation-Responsive MicroRNAs Are Affected by Transgenerational Stress in Durum Wheat Seedlings. *Plants (Basel)* 10, 50826. doi:10.3390/plants10050826

Liu, W., Kohlen, W., Lillo, A., Op Den Camp, R., Ivanov, S., Hartog, M., et al. (2011). Strigolactone Biosynthesis in Medicago Truncatula and Rice Requires the Symbiotic GRAS-type Transcription Factors NSP1 and NSP2. *Plant Cell.* 23, 3853–3865. doi:10.1105/tpc.111.089771

Liu, W., Sun, Q., Wang, K., Du, Q., and Li, W. X. (2017). Nitrogen Limitation Adaptation (NLA) Is Involved in Source-to-sink Remobilization of Nitrate by Mediating the Degradation of NRT 1.7 in Arabidopsis. *New Phytol.* 214, 734–744. doi:10.1111/nph.14396

Liu, Y., Li, M., Xu, J., Liu, X., Wang, S., and Shi, L. (2019). Physiological and Metabolomics Analyses of Young and Old Leaves from Wild and Cultivated Soybean Seedlings under Low-Nitrogen Conditions. *BMC Plant Biol.* 19, 389. doi:10.1186/s12870-019-2005-6

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

Loss-Morais, G., Waterhouse, P. M., and Margis, R. (2013). Description of Plant tRNA-Derived RNA Fragments (tRFs) Associated with Argonaute and Identification of Their Putative Targets. *Biol. Direct* 8, 6. doi:10.1186/1745-6150-8-6

Miao, C., Wang, D., He, R., Liu, S., and Zhu, J. K. (2020). Mutations in MIR 396e and MIR 396f Increase Grain Size and Modulate Shoot Architecture in Rice. *Plant Biotechnol. J.* 18, 491–501. doi:10.1111/pbi.13214

Pan, Y., Niu, M., Liang, J., Lin, E., Tong, Z., and Zhang, J. (2017). Identification of Heat-Responsive miRNAs to Reveal the miRNA-Mediated Regulatory Network of Heat Stress Response in *Betula Luminifera*. *Trees* 31, 1635–1652. doi:10.1007/s00468-017-1575-x

Pérez-Tienda, J., Corrêa, A., Azcón-Aguilar, C., and Ferrol, N. (2014). Transcriptional Regulation of Host NH^{4+} Transporters and GS/GOGAT Pathway in Arbuscular Mycorrhizal Rice Roots. *Plant Physiology Biochem.* 75, 1–8. doi:10.1016/j.plaphy.2013.11.029

Qiao, Q., Wang, X., Yang, M., Zhao, Y., Gu, J., and Xiao, K. (2018). Wheat miRNA Member TaMIR2275 Involves Plant Nitrogen Starvation Adaptation via Enhancement of the N Acquisition-Associated Process. *Acta Physiol. Plant* 40, 183. doi:10.1007/s11738-018-2758-9

Shao, Y., Zhou, H.-Z., Wu, Y., Zhang, H., Lin, J., Jiang, X., et al. (2019). OsSPL3, an SBP-Domain Protein, Regulates Crown Root Development in Rice. *Plant Cell.* 31, 1257–1275. doi:10.1105/tpc.19.00038

Singh, J., and Verma, P. K. (2021). NSP1 Allies with GSK3 to Inhibit Nodule Symbiosis. *Trends Plant Sci.* 26, 999–1001. doi:10.1016/j.tplants.2021.07.001

Song, A., Wang, L., Chen, S., Jiang, J., Guan, Z., Li, P., et al. (2015). Identification of Nitrogen Starvation-Responsive microRNAs in *Chrysanthemum Nankingense*. *Plant Physiology Biochem.* 91, 41–48. doi:10.1016/j.plaphy.2015.04.003

Tiwari, J. K., Buckseth, T., Zinta, R., Saraswati, A., Singh, R. K., Rawat, S., et al. (2020). Genome-wide Identification and Characterization of microRNAs by Small RNA Sequencing for Low Nitrogen Stress in Potato. *PLoS One* 15, e0233076. doi:10.1371/journal.pone.0233076

Wang, J., Song, K., Sun, L., Qin, Q., Sun, Y., Pan, J., et al. (2019). Morphological and Transcriptome Analysis of Wheat Seedlings Response to Low Nitrogen Stress. *Plants (Basel)* 8, 40098. doi:10.3390/plants8040098

Wang, Y.-Y., Hsu, P.-K., and Tsay, Y.-F. (2012). Uptake, Allocation and Signaling of Nitrate. *Trends Plant Sci.* 17, 458–467. doi:10.1016/j.tplants.2012.04.006

Wen, Z., Hong, Y., Qiu, Z. L., Yang, K., Hou, Q. D., Qiao, G., et al. (2022). Identification of miRNAs Mediating Shoot Growth of Grafted Sweet Cherry through Small RNA and Degradome Sequencing. *Sci. Hortic.* 291, 110557. doi:10.1016/j.scienta.2021.110557

Xie, F., Wang, Q., Sun, R., and Zhang, B. (2015). Deep Sequencing Reveals Important Roles of microRNAs in Response to Drought and Salinity Stress in Cotton. *J. Exp. Bot.* 66 (3), 789–804. doi:10.1093/jxb/eru437

Xie, Z., Kasschau, K. D., and Carrington, J. C. (2003). Negative Feedback Regulation of Dicer-Like1 in *Arabidopsis* by microRNA-Guided mRNA Degradation. *Curr. Biol.* 13, 784–789. doi:10.1016/s0960-9822(03)00281-1

Yang, Y., Zhang, X., Su, Y., Zou, J., Wang, Z., Xu, L., et al. (2017). miRNA Alteration Is an Important Mechanism in Sugarcane Response to Low-Temperature Environment. *Bmc Genomics* 18, 833. doi:10.1186/s12864-017-4231-3

Yao, Z., Li, X., and Asimuguli, A. (2019). Analysis of NSP Family Gene Sequence and Expression in Response to Drought Stress in *Arabidopsis thaliana*. *Mol. Plant Breed.* 17, 1169. doi:10.13271/j.mpb.017.001169

Ye, J. Y., Tian, W. H., and Jin, C. W. (2019). A Reevaluation of the Contribution of NRT 1.1 to Nitrate Uptake in Arabidopsis under Low-Nitrate Supply. *FEBS Lett.* 593, 2051–2059. doi:10.1002/1873-3468.13473

Yu, C., Chen, Y., Cao, Y., Chen, H., Wang, J., Tian, F., et al. (2018). Overexpression of miR169o, an Overlapping MicroRNA in Response to Both Nitrogen Limitation and Bacterial Infection, Promotes Nitrogen Use Efficiency and Susceptibility to Bacterial Blight in Rice. *Plant & Cell. Physiology* 59, 1234–1247. doi:10.1093/pcp/pcy060

Zhang, S., Sun, L., and Kragler, F. (2009). The Phloem-Delivered RNA Pool Contains Small Noncoding RNAs and Interferes with Translation. *Plant Physiol.* 150, 378–387. doi:10.1104/pp.108.134767

Zhang, J. H., Lin, Y., Wu, F. M., Zhang, Y. T., Cheng, L. J., Huang, M. H., et al. (2021). Profiling of MicroRNAs and Their Targets in Roots and Shoots Reveals a Potential MiRNA-Mediated Interaction Network in Response to Phosphate Deficiency in the Forestry Tree *Betula Luminifera*. *Front. Genet.* 12. doi:10.3389/fgene.2021.552454

Zhang, J., Huang, M., Liang, J., Pan, Y., Cheng, L., Wu, J., et al. (2016). Genome-wide Mining for microRNAs and Their Targets in *Betula Luminifera* Using High-Throughput Sequencing and Degradome Analyses. *Tree Genet. Genomes* 12, 99. doi:10.1007/s11295-016-1047-2

Zhang, J., Zhou, Z., Bai, J., Tao, X., Wang, L., Zhang, H., et al. (2020). Disruption of MIR396e and MIR396f Improves Rice Yield Under Nitrogen-Deficient Conditions. *Natl. Sci. Rev.* 7, 102–112. doi:10.1093/nsr/nwz142

Zhao, M., Ding, H., Zhu, J. K., Zhang, F., and Li, W. X. (2011). Involvement of miR169 in the Nitrogen-Starvation Responses in Arabidopsis. *New Phytol.* 190, 906–915. doi:10.1111/j.1469-8137.2011.03647.x

Zhao, M., Guo, R., Li, M., Liu, Y., Wang, X., Fu, H., et al. (2020). Physiological Characteristics and Metabolomics Reveal the Tolerance Mechanism to Low Nitrogen in *Glycine Soja* Leaves. *Physiol. Plant.* 168, 819–834. doi:10.1111/ppl.13022

Zhao, M., Tai, H., Sun, S., Zhang, F., Xu, Y., and Li, W.-X. (2012). Cloning and Characterization of Maize miRNAs Involved in Responses to Nitrogen Deficiency. *PLoS One* 7, e29669. doi:10.1371/journal.pone.0029669

Zhao, W., Yang, X., Yu, H., Jiang, W., Sun, N., Liu, X., et al. (2015). RNA-Seq-Based Transcriptome Profiling of Early Nitrogen Deficiency Response in Cucumber Seedlings Provides New Insight into the Putative Nitrogen Regulatory Network. *Plant Cell. Physiology* 56, 455–467. doi:10.1093/pcp/pcu172

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

Zhu, Z., Li, D., Cong, L., and Lu, X. (2021). Identification of microRNAs Involved in Crosstalk between Nitrogen, Phosphorus and Potassium under Multiple Nutrient Deficiency in Sorghum. *Crop J.* 9, 465–475. doi:10.1016/j.cj.2020.07.005

Zuluaga, D. L., and Sonnante, G. (2019). The Use of Nitrogen and its Regulation in Cereals: Structural Genes, Transcription Factors, and the Role of miRNAs. *Plants (Basel)* 8, 294. doi:10.3390/plants8080294

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

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

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fgene.2022.958641

OPEN ACCESS

EDITED BY
Yun Zheng,
Kunming University of Science and Technology, China

REVIEWED BY
Yong-Fang Li,
Henan Normal University, China
Pingchuan Deng,
Northwest A&F University, China

*CORRESPONDENCE
Camilla Alves Santos,
csantos.alvess@gmail.com
Marília Gaspar,
gasparmarilia@sp.gov.br

SPECIALTY SECTION
This article was submitted to RNA, a section of the journal Frontiers in Genetics

RECEIVED 31 May 2022
ACCEPTED 22 August 2022
PUBLISHED 27 September 2022

CITATION
Santos CA, Moro CF, Salgado I, Braga MR and Gaspar M (2022), Noncoding RNAs responsive to nitric oxide and their protein-coding gene targets shed light on root hair formation in Arabidopsis thaliana.
Front. Genet. 13:958641.
doi: 10.3389/fgene.2022.958641

Noncoding RNAs responsive to nitric oxide and their protein-coding gene targets shed light on root hair formation in *Arabidopsis thaliana*

Camilla Alves Santos[1]*, Camila Fernandes Moro[2], Ione Salgado[1], Márcia Regina Braga[1] and Marília Gaspar[1]*

[1]Laboratório de Ecofisiologia e Bioquímica de Plantas, Núcleo de Conservação da Biodiversidade, Instituto de Pesquisas Ambientais, São Paulo, SP, Brasil, [2]Programa de Pós-Graduação em Biologia Celular e Estrutural, Universidade Estadual de Campinas, Campinas, SP, Brasil

An overview of the total *Arabidopsis thaliana* transcriptome, described previously by our research group, pointed some noncoding RNA (ncRNA) as participants in the restoration of hair-root phenotype in *A. thaliana rhd6* mutants, leading us to a deeper investigation. A transcriptional gene expression profiling of seedling roots was performed aiming to identify ncRNA responsive to nitric oxide (GSNO) and auxin (IAA), and their involvement in root hair formation in the *rhd6* null mutant. We identified 3,631 ncRNAs, including new ones, in *A. thaliana* and differential expression (DE) analysis between the following: 1) GSNO-treated *rhd6* vs. untreated *rhd6*, 2) IAA-treated *rhd6* vs. untreated *rhd6*, 3) GSNO-treated *rhd6* vs. IAA-treated *rhd6*, and 4) WS-2 vs. untreated *rhd6* detected the greatest number of DE genes in GSNO-treated *rhd6*. We detected hundreds of *in silico* interactions among ncRNA and protein-coding genes (PCGs), highlighting MIR5658 and MIR171 precursors highly upregulated in GSNO-treated *rhd6* and wild type, respectively. Those ncRNA interact with many DE PCGs involved in hormone signaling, cell wall development, transcription factors, and root hair formation, becoming candidate genes in cell wall modulation and restoration of root hair phenotype by GSNO treatment. Our data shed light on how GSNO modulates ncRNA and their PCG targets in *A. thaliana* root hair formation.

KEYWORDS

cell wall modulation, miRNA, lncRNA, ncRNA–PCG interaction, auxin, *A. thaliana*

Introduction

Transcriptome studies in eukaryotes revealed that more than 90% of the genome is transcribed, with a diverse set of transcripts corresponding to noncoding RNAs (ncRNAs) (Chekanova et al., 2007; Kapranov et al., 2007). Noncoding RNAs are functional and low protein-coding potential RNA molecules, which can be classified according to their

function, location, and length. Depending on their length, ncRNAs can be divided into small ncRNAs (sRNAs) (20–30 nt), which are commonly found as transcriptional and translational regulators, medium-sized ncRNAs (50–200 nt), and also long ncRNAs (lncRNAs) (>200 nt). The last one is usually involved in other processes, such as splicing, gene inactivation, and translational regulation (Wang et al., 2017; Yu et al., 2019).

The advent of high-throughput sequencing technologies has facilitated the identification and characterization of many lncRNAs in plants (Fukuda et al., 2020), such as *Oryza sativa* (Li et al., 2020), *Solanum lycopersicum* (Zhu et al., 2015), and *Arabidopsis thaliana* (Lasky et al., 2014). Recent studies indicated that thousands of lncRNAs are extensively distributed in different regions of the plant genomes, including introns, intergenic regions, natural antisense transcripts (NAT), pseudogenes, and retrotransposons in protein-coding genes (PCGs) (Lee 2012; Wang et al., 2017; Kopp and Mendell 2018). NATs are a specific group of lncRNA which are complementary to PCGs, showing tissue-specific responses to biotic or abiotic stresses (Jannesar et al., 2020; Zhao et al., 2020). Long noncoding RNAs are also involved in translational regulation and post-translational modification, thereby regulating protein phosphorylation, ubiquitination, and acetylation, and modulating tissue gene expression during developmental stages and in response to external stimuli (Kim and Sung 2012; Zhang et al., 2019). There is an increasing number of evidence showing that plant lncRNAs have key roles in genomic imprinting, cell differentiation, epigenetic regulation, and stress responses (Zhang et al., 2011; Di et al., 2014; Liu et al., 2015).

To date, the best-studied ncRNAs in plants are sRNAs, including microRNAs (miRNAs) and nucleolar RNA (snoRNA), which play important roles in transcriptional and post-transcriptional regulation of gene expression (Bardou et al., 2014). Some miRNA precursors are well studied in *A. thaliana*, and their relationship with the nitrogen (N) metabolism is known, such as miR160, miR167, miR171, miR393, miR169, miR826, and miR5090 (Gifford et al., 2008; Wang et al., 2010; Liang et al., 2012). Despite the knowledge achieved regarding N-responsive miRNAs already identified, many still remain uncharacterized.

Nitric oxide (NO) is a gaseous signaling molecule originated mainly from the nitrate metabolism in plants through nitrate/nitrite reductase activities (Salgado et al., 2009; Salgado et al., 2017). NO plays a broad role in the regulation of developmental processes in plants (Mur et al., 2013). We reported previously that NO-donor S-nitrosoglutathione (GSNO), but not auxin (IAA), restored the wild-type root transcriptome profile in *rhd6* (*root hair defective 6*) mutants. NO modulates the expression of a large number of genes related to cell wall composition and metabolism, as well as those encoding ribosomal proteins, DNA and histone-modifying enzymes, and proteins involved in post-translational modification (Moro et al., 2017). A glimpse of the total *A. thaliana* transcriptome described by Moro et al. (2017) suggested that some ncRNA could also be involved in the restoration of wild root hair phenotype, leading us to investigate this hypothesis deeper. Considering that studies reporting NO-responsive ncRNAs in *A. thaliana* are still incipient, here, we performed the mapping of *A. thaliana* RNA-seq reads previously generated by Moro et al. (2017) against the reference genome followed by a *de novo* assembly. After that, we carried out a transcriptional gene expression profiling of *A. thaliana* seedling roots with the purpose of identifying ncRNA genes responsive to GSNO and IAA treatments in the *rhd6* null mutant when compared to the wild type (WS-2). We also investigated the putative interactions between ncRNA–ncRNA and ncRNA–PCG pairs, using the root NO-responsive PCGs for *A. thaliana* described by Moro et al. (2017), aiming to detect PCGs targeted and modulated by those ncRNA.

Materials and methods

Sampling

Plant cultivation and treatment were performed by Moro et al. (2017). In brief, seeds of *A. thaliana rhd6* mutant and its respective Wassilewskija (WS-2) wild ecotype, obtained from the Arabidopsis Biological Resource Center (ABRC), were germinated in petri dishes containing nutritive medium with the addition of 1 mM *S*-nitrosoglutathione (GSNO; Enzo Life Sciences), 50 nM indole-3-acetic acid (IAA; Sigma-Aldrich), or deionized water. The culture plates were kept in growth chambers at 22°C under a 12-h photoperiod at a light intensity of 85 μmol m^2 s^{-1} (μƐ). Roots from 5-day-old WS-2, *rhd6*, and IAA- and GSNO-treated *rhd6* seedlings were collected and immediately frozen in liquid nitrogen for RNA stabilization. For RNA isolation, four biological replicates of each condition/treatment were used, totalizing 16 samples sequenced as described in the study by Moro et al. (2017).

Sequence data analysis

Paired-end libraries were prepared as described in the TruSeq RNA Sample Prep Protocol (Illumina, San Diego, CA, United States). Indexed DNA libraries were normalized, pooled, and sequenced in the paired-end mode in two lanes using an Illumina HiSeq SQ sequencer. All raw reads are available at the Sequence Read Archive (SRA-NCBI) under the accession number SRP285694 (BioProject PRJNA666227) (Moro et al., 2017). The quality of the raw data generated after sequencing was checked in the FastQC software (version 0.10.1) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The *reads* were filtered for Phred quality (QS) 26 (sequence average) and 30 (sequence

edges), and a minimum length of 65 bp, using the SeqyClean package (v.1.9.9) (https://github.com/ibest/seqyclean). This program was also used to remove contaminant sequences from primers, adaptors, and vectors using the Univec database (https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/).

Clean reads of each sample were mapped to the *A. thaliana* genome (TAIR 10.1—GCA_000001735.4) using STAR (Dobin et al., 2013), and the mapped reads were *de novo* assembled and quantified using StringTie (Kovaka et al., 2019). Novel transcripts were named as "MSTRG" by StringTie. TransDecoder package (http://transdecoder.sourceforge.net/) was used in the evaluation of transcript coding potential. We aligned all transcripts against NCBI-nr and Uniprot/UniRef90 (The UniProt Consortium, 2017) using BLASTx with an e-value cutoff of 1^{e-3}. Once a sequence was identified as a potential coding one, it was excluded from the subsequent ncRNA analysis. We downloaded miRNA precursors from Rfam (Kalvari et al., 2018) in order to identify lncRNAs acting as miRNA precursors. Putative lncRNAs were also aligned to miRbase (Kozomara and Griffiths-Jones, 2010) sequences using BLASTn, and those showing cutoff >90% identification accuracy and e-value < 1^{e-1} were identified as probable miRNA precursors. The putative miRNA precursors and conserved lncRNAs were separated and excluded from the remaining dataset to prevent elimination in the next steps. The mature form of miRNAs was identified by a manual curation in all the miRNA fasta sequences, and those containing 20–22 nucleotides in length were classified as mature (miR). To discriminate between lncRNAs and sRNA transcripts, we aligned all remaining transcripts against the RNAcentral database (https://rnacentral.org/). The remaining transcripts were classified as candidate *A. thaliana* lncRNAs. To identify putative transposon sequences in lncRNAs, known transposon sequences of *A. thaliana* from the TAIR10 database (https://www.arabidopsis.org/) were downloaded. Long noncoding RNA sequences were aligned to the *A. thaliana* known transposons, and only those matching with at least 90% identity and e-value < 1^{e-1} were selected and classified as probable transposon sequences.

Next, we sought for natural antisense transcripts (NAT). After the alignment of lncRNA and PCGs using BLASTn, those NAT sequences showing ≥90% identity and no gap region longer than 150 bp were classified as putative lncRNA–NAT pairs. Finally, LncTar software (Li et al., 2015) was employed to confirm the annealing potential of the BLAST-predicted pairs.

Differential expression analysis

To identify differentially expressed (DE) ncRNAs among the different treatments of *A. thaliana* seedlings, we used the prepDE.py3 script present in StringTie to extract gene count information from the program output. The ncRNA gene counts were used as input in the DESeq2 package (Love et al., 2014) (Bioconductor/R). Independent comparisons between samples from 1) GSNO-treated *rhd6* vs. untreated *rhd6*, 2) IAA-treated *rhd6* vs. untreated *rhd6*, 3) GSNO-treated *rhd6* vs. IAA-treated *rhd6*, and 4) WS-2 vs. untreated *rhd6* were performed. Normalization was carried out by adjusting the data distribution according to a negative binomial distribution, followed by removing the contigs with a base mean < 5. The adjusted *p*-value for each gene was calculated using the BH method (Benjamini and Hochberg, 1995), and only those with FDR < 0.05 were considered significant differentially expressed genes (DEGs).

RNA pair interaction evaluation and network analysis

With the purpose of predicting lncRNA and PCG interaction, we used LncTar software. A cutoff value of −0.13 was used for the normalized free energy (ndG), which reflects the relative stability of internal base pairs in the paired RNAs. For testing the potential interaction between sRNA–lncRNA and sRNA–PCG, we used the heuristic mode in IntaRNA for searching (Mann et al., 2017). The lncRNA–PCG interaction types were identified by FEELnc software (Wucher et al., 2017) using the FEELnc_classifier.pl module. The output.gft file from StringTie *de novo* assembly and *A. thaliana* genome reference.gtf were used in the analysis. The interaction types could be classified as genic (lncRNA overlaps a PCG from the reference genome) and intergenic (lncRNA does not overlap a coding region). In addition, the genic lncRNA interactions could be classified as the following subtypes: 1) containing (lncRNA contains the mRNA partner), 2) nested (lncRNA is contained in the mRNA partner), and 3) overlapping (lncRNA partially overlaps PCG partner).

A protein–protein interaction (PPI) network with the PCG targeted by DElncRNA was constructed in Cytoscape v. 3.8.2 (Shannon et al., 2003). Proteins were queried against the STRING (https://string-db.org) database for the identification of interactions among them. The proteins were represented as circles (nodes) and the interactions among them were represented as lines (edges). The constructed hubs were submitted to enrichment analysis in Cytoscape, focusing on cell wall and nitrogen metabolism-related proteins. In addition, co-expression networks were analyzed and generated in CoExpNetViz (Tzfadia et al., 2015), using DElncRNA and DE CPG read counts as input, and selecting DElncRNA as bait nodes and default parameters. The co-expression networks produced were visualized in CoExpNetViz plugin in Cytoscape. We also performed a gene set enrichment analysis (GSEA) aiming to identify ncRNA gene classes with common expression patterns. The analysis was carried out in GSEA software v.4.2.3 (Subramanian et al., 2007) using 1,000 permutations, "phenotype" as permutation type, and FDR < 0.25, as recommended in the manual.

FIGURE 1
Differentially expressed gene (DEG) distribution of the four independent and paired comparisons performed. **(A)** Number of upregulated DEGs identified in the comparisons: GSNO-treated *rhd6* vs. untreated *rhd6*, IAA-treated *rhd6* vs. untreated *rhd6*, GSNO-treated *rhd6* vs. IAA-treated *rhd6*, and WS-2 vs. untreated *rhd6* (FDR < 0.05). **(B)** Noncoding RNA type distribution in each comparison. GSNO-treated *rhd6* vs. IAA-treated *rhd6* comparison is not shown due to the few DEGs found. **(C1)** Expression values detected for DEGs in IAA-treated *rhd6* vs. untreated *rhd6* and **(C2)** GSNO-treated *rhd6* vs. untreated *rhd6* comparisons. The *x*-axis shows the log2FC values, and positive values for upregulated and negative values for downregulated for IAA and GSNO treatments, respectively. The *y*-axis shows the DEGs names. **(D)** Left: union of DEGs identified downregulated in untreated *rhd6* mutants, and upregulated in IAA and GSNO treatments, revealing genes activated by each treatment. Right: union of DEGs identified upregulated in untreated *rhd6* mutants, and downregulated in IAA and GSNO treatments, revealing genes deactivated by each treatment.

Results

Transcriptome overview

In our previous study with root hair *rhd6* mutants of *A. thaliana*, the RNA-seq approach identified 32,841 protein-coding genes, of which 6,670 were differentially expressed in response to GSNO and IAA (Moro et al., 2017). The ncRNA identification and evaluation of their expression profiles in the *A. thaliana* null *rdh6* mutant were performed using the same 16 libraries constructed by Moro et al. (2017), with four replicates of each: IAA-treated *rhd6*, GSNO-treated *rhd6*, untreated *rhd6*, and untreated WS-2. Here, we generated a total of 685,859,082 processed and cleaned reads, of which 658,226,876 (96%) were mapped against the *A. thaliana* reference genome (Supplementary Table S1). Transcripts with the evidence of protein-coding potential were separated and classified as the coding section of transcriptome, previously analyzed by Moro et al. (2017). After the removal of putative protein-coding transcripts, we identified 4,323 ncRNAs in the transcriptome of *A. thaliana*. Therefore, rRNA and tRNA were filtered out from the analysis, yielding a total of 3,631 ncRNA transcripts. The ncRNA genes were distributed in lncRNAs (>200 bp), with 2,768 being lncRNAs (125 novel) and 64 natural antisense transcripts (NAT). Among ncRNA <200 bp were

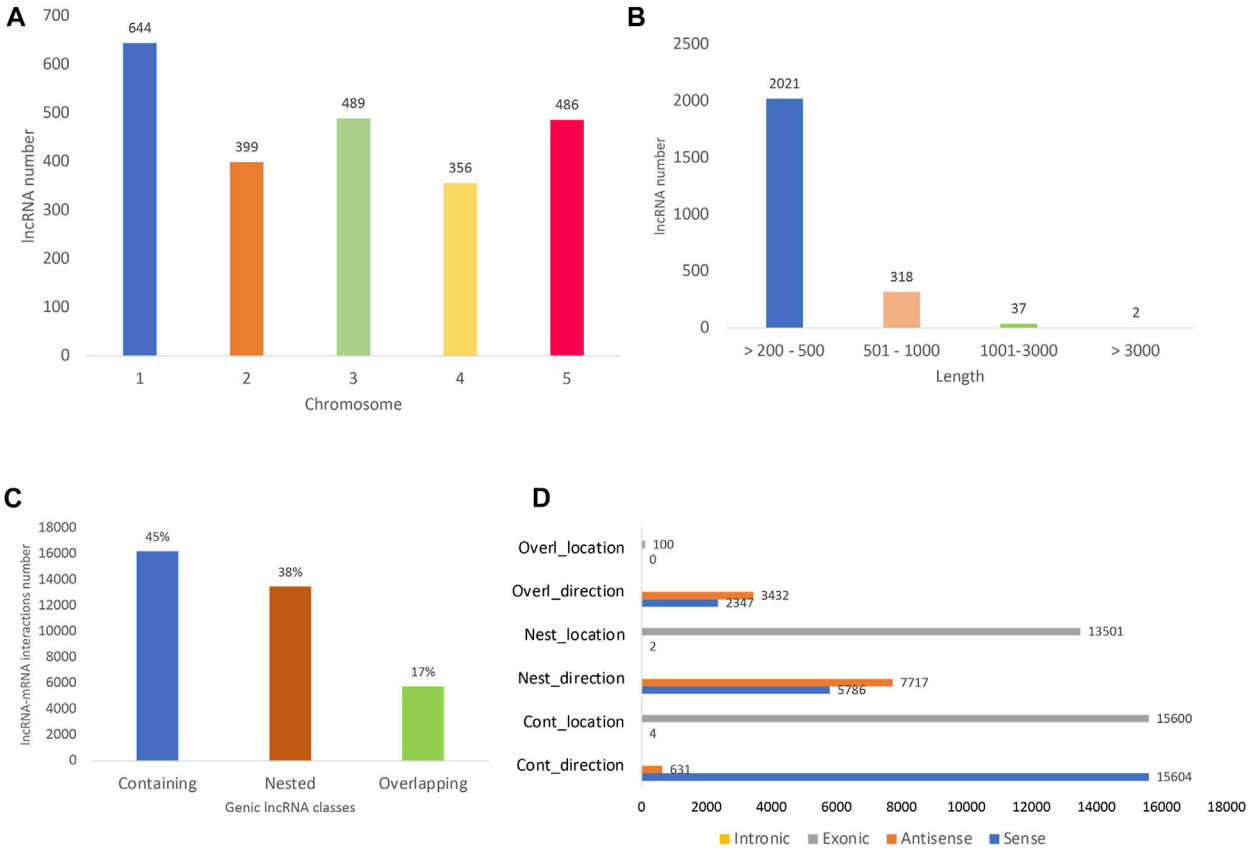


FIGURE 2
Long noncoding RNA (lncRNA) distribution. **(A)** Number of long noncoding RNAs already known (with AT identifiers) along the five *A. thaliana* chromosomes. **(B)** Length of all lncRNA identified, including those with AT identifiers and novel ones (the isoforms were not considered in the count). **(C)** Distribution of lncRNA and their respective differentially expressed protein-coding gene (PCG) targets. The interacting lncRNA–DEmRNA pairs were located mostly in genic regions (99.8%) and distributed in the following subtypes: containing (lncRNA contains the mRNA partner), nested (lncRNA is contained in the mRNA partner), and overlapping (lncRNA partially overlaps PCG partner). **(D)** lncRNA–PCG interaction direction and location, being mostly in sense direction and located in exonic regions. The *x*-axis shows the number of interactions detected and the *y*-axis shows the interaction direction and location. The number of interactions detected as intronic are very low compared to exonic ones, reason why no yellow bars could be observed for intronic interactions in the figure.

170 microRNAs (miRNAs) (77 presenting the mature form), 255 small nucleolar RNAs (snoRNAs), and 100 small RNA (sRNAs). The 100 sRNAs could not be classified neither in miRNA nor in snoRNA after annotation, remaining as sRNAs. We also identified 244 ncRNAs already known in *A. thaliana* genome, which could not be classified, being named as "other RNA" and 14 pseudogenes after RNA central and Rfam queries (Supplementary Table S2). Among the lncRNAs and pseudogenes, 85 and seven sequences are transposons, respectively.

We used all the ncRNA genes as a background reference for the differential expression analysis. Four independent and paired comparisons were performed between the following groups of *A. thaliana* seedlings: 1) GSNO-treated *rhd6* vs. untreated *rhd6*, 2) IAA-treated *rhd6* vs. untreated *rhd6*, 3) GSNO-treated *rhd6* vs. IAA-treated *rhd6*, and 4) WS-2 vs. untreated *rhd6* (FDR < 0.05) (Supplementary Table S3). The GSNO-treated *rhd6* showed the greatest number of differentially expressed genes (DEGs) (Figure 1A), with 45 upregulated ncRNA genes. From these, 39 had their ncRNA type identified after annotation, mostly being lncRNAs (Figure 1B). The expression values (log2FC) varied widely in IAA and GSNO-treated *rhd6* (Figures 1C1,C2). Among the 30 ncRNA genes downregulated between WS-2 and untreated *rhd6* seedlings, 13 are upregulated by GSNO, one by IAA, and three by both treatments (Figure 1D—left). On the other hand, considering the 17 DEGs upregulated between WS-2 and untreated *rhd6*, eight are downregulated by GSNO and one by both compounds, while no DEGs were detected to be downregulated only by IAA (Figure 1D—right).

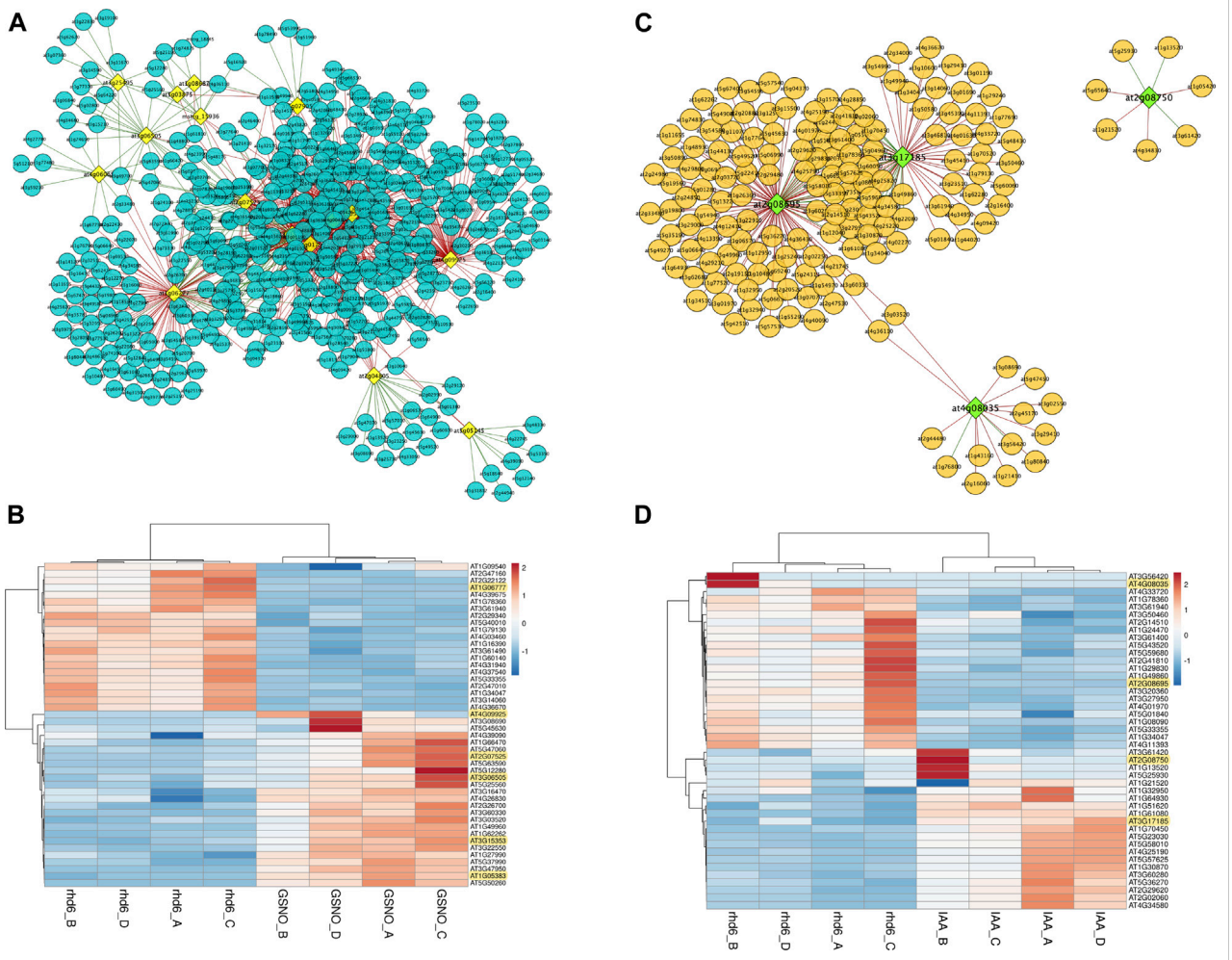


FIGURE 3
Results from co-expression analysis between differentially expressed (DE) long noncoding RNAs (DElncRNA) and differentially expressed protein-coding genes (PCGs). **(A)** Co-expression network generated for DElncRNA and DE PCGs which present similar expression patterns for GSNO treatment. The DElncRNA are represented in yellow and the DE PCG in blue. The lines represent the connections (edges) among DElncRNA and DE PCGs pairs, which may be correlated (green) or anticorrelated (red). **(B)** Heatmap showing the top 10 most correlated and anticorrelated PCGs and how they vary their expression patterns in the same or opposite direction to DElncRNAs, which are highlighted in yellow. All DElncRNA in the figure are correlated with DE PCG, except for AT1G06777. **(C)** IAA treatment co-expression network. The DElncRNA are represented in green and the DE PCG are represented in orange. Green lines represent correlated gene pairs, while red lines are anticorrelated gene pairs. **(D)** Top 10 heatmap with the most correlated and anticorrelated PCGs, and DElncRNAs are highlighted in yellow. AT2G08695 and AT4G08035 DElncRNAs are anticorrelated, while DEmRNA AT3G17185 and AT2G08750 are correlated. In Figures **(B,D)**, the respective color keys in the right side represent the z-scores.

Long noncoding RNA interaction with protein-coding genes upon GSNO and IAA treatments

All the differentially expressed protein-coding genes (DE PCG) mentioned in this study were obtained from the same transcriptome analyzed here and previously reported by Moro et al. (2017). The lncRNA identified here are widely distributed in *A. thaliana* 1–5 chromosomes, with the greatest number found in chromosome 1 (27%) (Figure 2A). Long noncoding RNAs showed an average length of 274 bp, mostly ranging from 200 to 500 bp (85%) (Figure 2B). In terms of lncRNA and PCG interaction, we detected almost only genic interactions (99.8%). The genic lncRNA interactions identified were classified as the following subtypes: containing (45%), nested (38%), and overlapping (17%) (Figure 2C). Here, we identified the lncRNA–mRNA interactions mostly in sense direction (Figure 2D) and located in exonic regions.

We sought for interaction between the DElncRNA (lncRNA differentially expressed) identified in this work and the DE PCG described by Moro et al. (2017). Considering all the DElncRNA genes identified here and

their respective PCG targets, 34 DE PCGs participate in biological processes related to the regulation of the nitrogen compound metabolic process. *WRKY40*, *SZF1*, *MYC2*, *WRKY48*, *AIB*, *MYB48*, and *NAC102* are potential candidate key genes in GSNO signaling pathways, under the modulation of DElncRNA identified here and usually interacting among each other (Supplementary Table S4). In addition, *PRP3*, *EXT13*, *RHS19*, and *MOP10* are included in the other set of 34 PCGs mostly involved in cell wall organization, biogenesis, and degradation (Supplementary Figure S1). We also analyzed how the expression pattern of DElncRNA identified here and DE PCGs varied under GSNO and IAA treatments, using the co-expression network analysis. Genes in a co-expression network may be positively (expression profile among DElncRNAs and DE PCGs rise or fall together among samples) or negatively (expression profiles vary in opposite directions among samples) correlated. Considering GSNO treatment, we detected 459 DE PCGs interacting with 16 key DElncRNA (Figure 3A; Supplementary Table S5). Although the connections among DElncRNA and DE PCGs are mostly negatively correlated, we observed some DElncRNA positively interacting with DE PCGs, such as AT3G25495, AT4G06065, AT2G04805, and AT5G05145 (Figure 3A). Differentially expressed lncRNA genes, AT1G06777, AT4G09925 (MIR5658 precursor), AT3G07525, AT3G06505, AT3G15353, and AT1G05383, are those with the strongest correlation values, with only AT1G06777 being negatively correlated (Figure 3B).

On the other hand, regarding IAA treatment, our analysis showed a total of 164 DE PCGs following the expression profiles of four main DElncRNA genes, highlighting AT2G08695 and AT3G17185 as those with the greatest number of connecting genes in the network (Figure 3C; Supplementary Table S5). AT2G08695 and AT4G08035 are the main DElncRNAs negatively correlated with DE PCGs, while AT3G17185 and AT2G08750 are positively correlated (Figure 3D). We observed 117 and 165 exclusive DE PCGs fluctuating their expression profiles according to DElncRNAs in GSNO and IAA treatments, respectively. Many biological roles such as development, hormone signaling, and protein modification were among those identified in DE PCGs for both treatments, highlighting the transcription factor group as the most represented one. A total of 11 and 29 cell wall-related genes were also detected as modulated by DElncRNA for GSNO and IAA treatments, respectively (Supplementary Table S5).

We also identified some ncRNA >200 bp that could not be annotated as lncRNA, although with important roles in the NO metabolism. After LncTar analysis, which calculates the minimum free energy joint structure of two RNA molecules based on base pairing, we identified the ncRNA AT3G25795 and the novel MSTRG.17591 as key DEncRNAs in GSNO treatment. AT3G25795 appears upregulated in the GSNO-treated and WS-2 seedlings (average log2FC = 3.62), and interacting with DE PCGs related to root hair formation (*EXT13*) and transmembrane protein (*MUL8*), both upregulated in GSNO-treated seedlings. Novel MSTRG.17591 also is upregulated in GSNO-treated and WS-2 individuals (average log2FC = 2.69), targeting the PCG involved in protein degradation (*ATG8E*), which is induced in GSNO-treated and inhibited in IAA-treated seedlings (Supplementary Table S6). In addition, we also detected the long noncoding RNA AT4G09925 (MIR5658 precursor), highly upregulated in GSNO-treated and WS-2 seedlings (log2FC = 4.88 and 5.02, respectively) and downregulated in IAA-treated seedlings (log2FC = −5.23). Furthermore, we identified a total of 64 NATs in *A. thaliana* transcriptome expressed under GSNO and IAA treatments, along with 13 lncRNA–NAT, not differentially expressed, overlapping with DE PCG with biological roles related to development, transcription factors, calcium regulation, protein degradation, heat shock protein, and root hair formation (Supplementary Table S2).

miRNA interaction with lncRNA and protein-coding genes on GSNO- and IAA-treated seedlings

We also sought for ncRNA <200 bp or sRNA, such as miRNA and snoRNA. These RNAs are known to target other ncRNAs, such as lncRNAs and also PCGs. Here, we identified 15 differentially expressed sRNA, mostly under GSNO treatment in the *rhd6* mutant, such as the microRNAs MIR171 and miR398 (Supplementary Table S3). We investigated the interaction between DEsRNA and DElncRNA, in which six DEsRNA with the highest interaction energies were all differentially expressed under GSNO treatment and also targeting DElncRNA responsive to GSNO treatment (Figure 4A; Supplementary Table S6). Three DEsRNA, AT2G14878 (sRNA), AT5G48412 (sRNA), and AT1G62035 (MIR171), are upregulated in GSNO-treated *rhd6* and WS-2, and are potentially involved in the root hair phenotype restoration in null *rhd6* mutants.

In addition, we evaluated the interaction between DEsRNA and DE PCG, selecting those loci with the highest interaction energies (<−20 kcal/mol) and a greater number of PCG targets. Here, we highlight one more time AT2G14878, AT5G48412 and AT1G62035 (MIR171), all responsive in GSNO-treated *rhd6* and WS-2 seedlings. Their respective targeted PCGs show, among the top GO terms identified for biological processes (BP), cell wall organization, root hair development/differentiation, and regulation of nitrogen compound metabolic process (Figure 4B). On the other hand, we detected TAS3 (AT3G17185), a small interference RNA (siRNA), upregulated in IAA-treated *rhd6* and interacting with small auxin upregulated RNA 6 (*SAUR6*) (Supplementary Tables S3, S6).

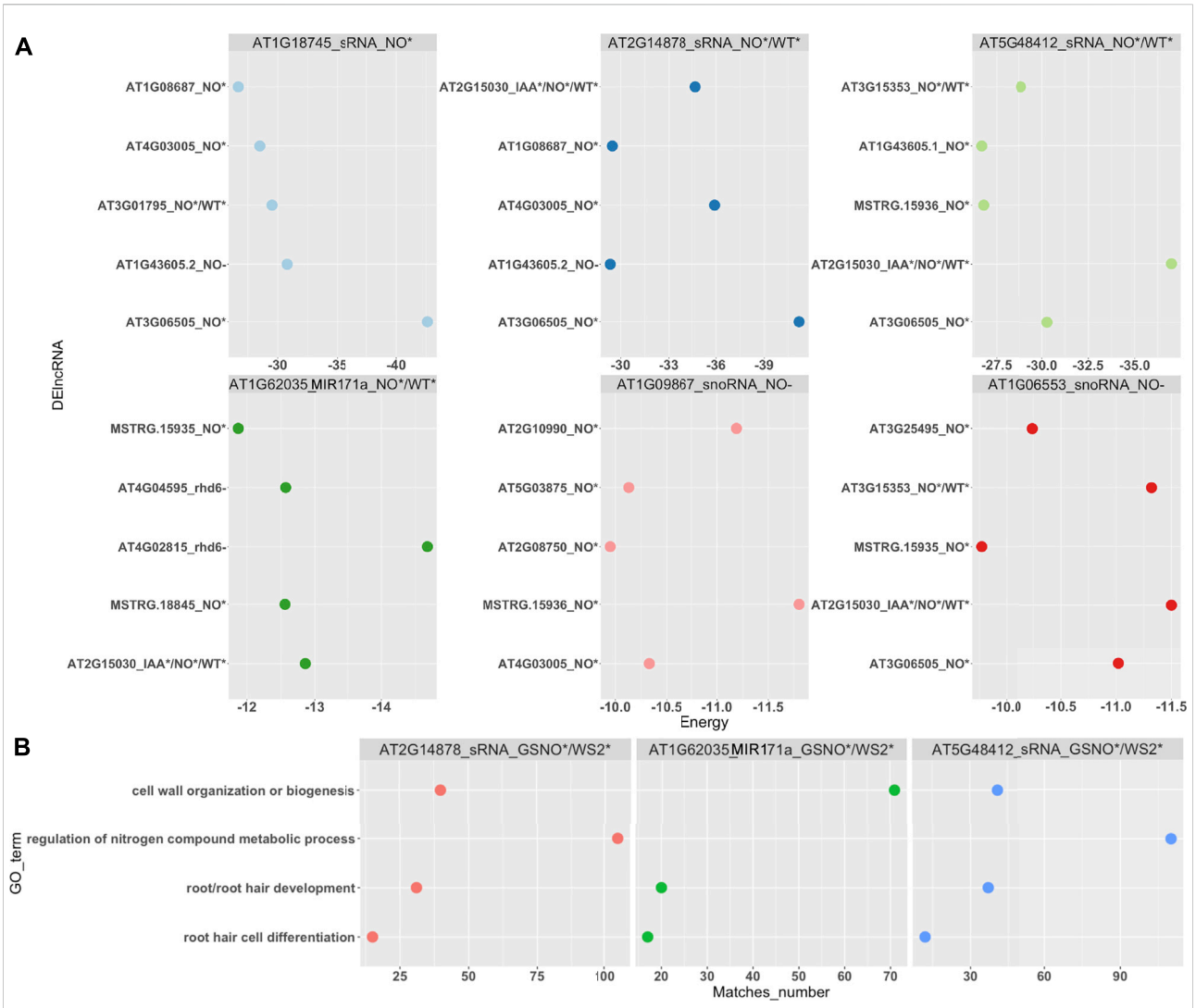


FIGURE 4
Small RNA (sRNA) interaction with DElncRNAs and DE PCG. As sRNA, we consider sRNA (could not have their types identified by annotation), nucleolar RNA (snoRNA), and microRNA (miRNA). **(A)** Top six DElncRNA–sRNA pairs with the strongest interaction energy (negative values). Three of them are differentially expressed in GSNO-treated and WS-2 individuals, potentially acting in the WS-2 root hair formation WS-2 phenotype. The *x*-axis shows the interaction energy values, whereas the *y*-axis represents the DElncRNA genes and their differential expression condition. **(B)** sRNA and DE PCG interaction. Top three DE PCG with the greatest number of Gene Ontology (GO) biological process (BP) terms matches. The BPs are mostly related to nitrogen metabolism and cell wall organization and root development. The *x*-axis represents the BP matches number and the *y*-axis represents the most common BP terms identified.

Nitric oxide central importance in the restoration of the wild root hair phenotype in *A. thaliana*

Based on the determinant role of NO in *A. thaliana* root hair formation and in the restoration of wild root hair phenotype (Moro et al., 2017), we investigated which types of DEncRNA, in general, may potentially be involved in this process. For this purpose, we overlapped the DEGs identified in GSNO-treated *rhd6* vs. untreated *rhd6* and WS-2 vs. untreated *rhd6*, resulting in 25 common DEGs for the two comparisons, nine being downregulated and 16 upregulated for GSNO-treated and WS-2 seedlings (Figure 5A). Aiming to identify genes with similar expression patterns in recovering the wild-type root hair phenotype in the *rhd6* mutant, we used GSEA (FDR < 0.25). First, we analyzed the similar expression pattern between GSNO-treated *rhd6*/WS-2 vs. untreated *rhd6* groups DEGs, identifying 36 out of 45 GSNO-treated upregulated genes among the enriched ones (Figure 5B). Our results suggest some DEncRNAs that contribute to restore the root hair phenotype of the hairless *rhd6* mutant. In addition, we swept the whole ncRNA transcriptome, considering differentially

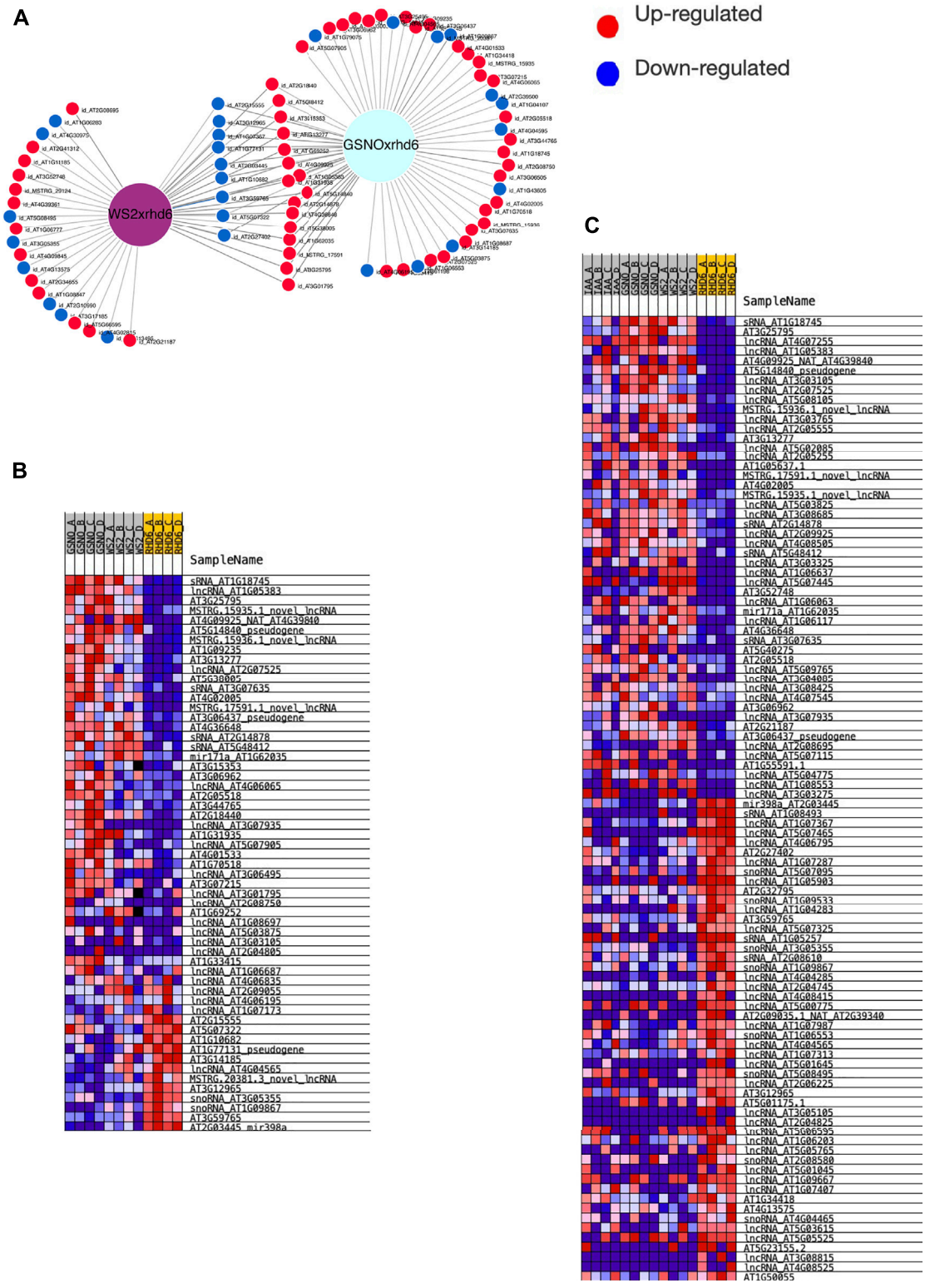


FIGURE 5
Differentially expressed genes in GSNO-treated and WS-2 seedlings and their potential participation in the restoration of root hair phenotype in *rhd6* mutants. **(A)** Intersection of differentially expressed genes (DEGs) detected in the comparison between GSNO-treated *rhd6* vs. untreated *rhd6* and WS-2 vs. *rhd6* reveals 25 genes in common for GSNO-treated *rhd6* and WS-2 seedlings, of which nine are downregulated and 16 are upregulated. **(B)** Gene set enrichment analysis (GSEA) performed for GSNO-treated *rhd6*/WS-2 vs. untreated *rhd6* groups. Figure shows DE genes are enriched and contribute with similar expression patterns for the WS-2 root hair formation (RHF) phenotype. **(C)** GSEA performed for GSNO-treated/IAA-treated/WS-2 vs. *rhd6* groups. Figure shows the top 50 genes contributing the most in GSNO- and IAA-treated for the recovery of RHF phenotype observed in WS-2.

expressed genes and those in which differential expression was not detected. The GSEA was employed to compare the groups of GSNO-treated *rhd6*/IAA-treated *rhd6*/WS-2 vs. untreated *rhd6* (Figure 5C). Therefore, a new set of genes potentially contributing to root hair phenotype restoration in GSNO-treated and IAA-treated *rhd6* seedlings were revealed, bringing 10 upregulated and two downregulated genes detected in GSNO-treated seedlings (Supplementary Table S7). As shown in Figures 5B,C, MIR171 and MIR5658 precursors, and the novel lncRNAs MSTRG 15935, 15936, and 17591 are listed as key DE ncRNAs involved in the root hair phenotype restoration in *A. thaliana rhd6* mutants.

Discussion

With the purpose of identifying ncRNAs and assessing their expression profiles in the *A. thaliana* null *rdh6* mutant, we used the same 16 libraries constructed by Moro et al. (2017), four of each abovementioned condition. We addressed the lncRNA genes expressed in *rhd6 A. thaliana* root seedlings, aiming to verify which lncRNAs were activated (upregulated) or deactivated (downregulated), and how GSNO and IAA exposure modulate the lncRNA and their respective mRNA target gene expression. In addition, we evaluated the putative interactions between ncRNA–ncRNA and ncRNA–PCG pairs upon GSNO and IAA treatments, in three major groups: 1) lncRNA and PCGs, 2) miRNA and lncRNA, and 3) miRNA and PCGs.

Among the 3,631 noncoding RNAs analyzed in the present study, we identified many ncRNA potentially involved in the restoration of the root hair phenotype in the *rhd6* mutant by GSNO. The co-expression network analysis between DElncRNA and DE PCG targets revealed close to 460 DE PCG interacting positively or negatively with 16 central DElncRNA upon GSNO treatment, being one of them not yet identified for *A. thaliana* (MSTRG 15936). Among the lncRNA interactions with positive values is MIR5658, a long noncoding NAT detected here upregulated in GSNO-treated and WS-2 seedlings, and downregulated in IAA-treated seedlings (Supplementary Table S3). MIR5658 has already been identified in different plant species as involved in plant development, hormone signaling, and tolerance to abiotic stress (Biniaz et al., 2022). This miRNA precursor directly upregulates the expression of AT3G25290, a member of auxin-responsive gene family, by targeting its promoter. This activation may be involved in the development and growth of *A. thaliana* (Yang et al., 2019). In our analysis, the MIR5658 precursor appears among the most differentially expressed lncRNAs, upregulated in GSNO-treated *rhd6* (Figure 3B), and seems to be an important regulator that represses the expression of numerous PCGs, but not AT3G25290 (Supplementary Table S5). This precursor is also known for controlling the expression of transcription factors, as those related to growth and development in *A. thaliana*, such as *GRAS* (gibberellic-acid insensitive) and *ERF* (ethylene responsive factor) (Rakhmetullina et al., 2021), mostly upregulated in the study by Moro et al. (2017). In addition, we identified some cell wall-related genes (xyloglucan endotransglucosylases, expansins, and arabinogalactans), whose expression levels varied along with DElncRNAs and were regulated by GSNO treatment (Supplementary Table S5). Xyloglucan endotransglucosylases can act in the degradation and loosening of cell wall, resulting in abnormal root hair formation and growth (Cavalier et al., 2008; Hayashi and Kaida, 2011). Moreover, arabinogalactans and expansins are closely involved in cell wall morphogenesis processes, as cell differentiation and cell wall expansion (Ellis et al., 2010; Lin et al., 2011). Furthermore, other 13 lncRNA–NAT not differentially expressed were detected, overlapping with DE PCG and acting in many biological roles, such as root hair formation (Supplementary Table S2). All of those PCGs are responsive to GSNO treatment and are mostly downregulated, such as WRKY transcription factor 61 (*WRKY61*), belonging to a protein family required for a myriad of biological events related to plant defense, stress, and development (Jiang et al., 2017; Singh et al., 2019).

In addition, we detected microRNAs MIR171 and miR398 differentially expressed in our transcriptome. According to Yan et al. (2022), the signaling pathway of MIR171 in root development is still unknown. Since this microRNA was upregulated in GSNO-treated roots with a similar WS-2 expression pattern, and considering its relevant role in root development (Yan et al., 2022), our data suggest that NO could be one of the signaling molecules implicated in restoring the root hair phenotype in *rhd6* mutant through MIR171 regulation. Among other biological processes, MIR171 regulates root hair differentiation, by targeting some protein-coding genes as those from the scarecrow-like family (SCL) and scarecrow (SCR) GRAS domain transcription factors (Singh et al., 2021). According to our present and previous data (Moro et al., 2017), MIR171 along with *SCL8* and *SCRL7* is upregulated in GSNO-treated *rhd6* seedlings. The interaction between DEsRNA and DElncRNA pairs were also tested under both treatments. Three DEsRNA, AT2G14878 (sRNA), AT5G48412 (sRNA), and AT1G62035 (MIR171) are upregulated in GSNO-treated *rhd6* and WS-2, and are potentially involved in the root hair phenotype restoration in null *rhd6* mutants. As observed in Figure 4, those DEsRNA target some novel DElncRNA for *A. thaliana*, upregulated in both GSNO-treated and WS-2 seedlings, indicating those pairs as candidates in the restoration of normal root hair formation in *rhd6* seedlings. When considering the interactions between DEsRNA and DE PCG pairs (energy < −20 kcal/mol), we again highlight the

three DEsRNA abovementioned, all responsive in GSNO-treated *rhd6* and WS-2 seedlings, and involved in cell wall organization, root hair development, and regulation of the nitrogen metabolism. Interestingly, AT2G14878 was identified as one of the 2,006 genes producing mobile RNAs in *A. thaliana*, which are systemically delivered to distant tissues, being transported in both directions, from root to shoot and from shoot to root (Thieme et al., 2015).

Small interference RNA (siRNA) TAS3 (AT3G17185) was upregulated in IAA-treated *rhd6* and interacting with small auxin upregulated RNA 6 (*SAUR6*) (Supplementary Tables S3, S6). TAS3 is known to suppress gene expression by post-transcriptional gene silencing in plants, orchestrating lateral root (LR) formation in *A. thaliana* by the modulation of miR390, and auxin response factors (*ARF*), as part of the auxin-mediated molecular network (Marin et al., 2010; Meng et al., 2010; Yoon et al., 2010). Yoon et al. (2010) shed light on the role of the miR390/TAS3/ARF pathway in the detection of auxin concentration and LR development. In our previous study, we identified *ARF9*, *ARF16*, and *ARF17* upregulated in IAA- and GSNO-treated *rhd6* seedlings (Moro et al., 2017), agreeing with our recent detection of TAS3 induced upon IAA treatment. Another important miRNA acting in root development is miR160 (Liang et al., 2012), also modulating the expression of *ARF16* and *ARF17*, which were upregulated in GSNO-treated *rhd6* seedlings in the study by Moro et al. (2017).

According to our findings (Figures 5B,C), MIR5658 and MIR171 precursors were upregulated in GSNO-treated *rhd6* and WS-2 seedlings. Along with the novel lncRNAs MSTRG 15935, 15936, and 17591, they are the key ncRNAs interacting with DE PCGs to restore the wild-type root hair phenotype. A much clearer and more determinant influence of GSNO was observed in the *A. thaliana* root hair noncoding transcriptome when compared to IAA, which is in line with previous PCG data shown in the study by Moro et al. (2017).

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/, SRP285694.

Author contributions

CM performed the experiments. CS performed ncRNA transcriptome analyses and wrote the article. MB, IS, and MG designed and supervised the research. CS, MB, and MG contributed to data interpretation, discussion of the results, and article preparation.

Funding

This study was supported by: FAPESP fellowships 2020/11908-7 and 2011/13220-3 and grants 2017/50341-0 and 2019/15095-3. CNPq grants 311860/2013-3 and 479930/2013-9.

Acknowledgments

Authors would like to thank Universidade de São Paulo, Instituto de Biociências (IB-USP) for providing high-performance computational infrastructure and Felipe Rodrigues da Silva for the RNA-seq protein coding genes data analysis by Moro et al. (2017).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.958641/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Long noncoding PCG target genes and how their gene products interact among each other in a protein–protein interaction (PPI) network. The DE PCGs with red and green circles are those GO enriched and with gene products related to nitrogen metabolism and cell wall organization, respectively.

SUPPLEMENTARY TABLE S1
A. thaliana transcriptome overview. Number of raw reads generated, reads processed, and cleaned and reads remaining after mapping against thale cress reference transcriptome.

SUPPLEMENTARY TABLE S2
Noncoding RNA (ncRNA) types identified in *A. thaliana* transcriptome. The ncRNA types are distributed in long noncoding RNA (lncRNA), longer than > 200 bp; natural antisense transcripts (NATs); small RNA (sRNA), including sRNA (no type could be identified by annotation), micro-RNA (miRNA), and nucleolar RNA (snoRNA); and other RNA which noncoding type could not be identified by annotation. The loci containing transposon sequences are also shown.

SUPPLEMENTARY TABLE S3
Differentially expressed genes (DEGs) identified in the four independent and paired comparisons performed: GSNO-treated *rhd6* vs. untreated *rhd6*, IAA-treated *rhd6* vs. untreated *rhd6*, GSNO-treated *rhd6* vs. IAA-treated *rhd6*, and WS-2 vs. untreated *rhd6* (FDR < 0.05). Upregulated and downregulated genes for each condition are shown, along with log2FC and false discovery rate (FDR) values.

SUPPLEMENTARY TABLE S4
Results from long noncoding PCG target genes GO enrichment performed in STRING database. Some of the nitrogen metabolism related genes (red circles from Supplementary Figure S1) shown in protein–protein interaction (PPI) network are highlighted in yellow.

SUPPLEMENTARY TABLE S5
Results from the co-expression analysis performed among differentially expressed genes long noncoding RNA (DElncRNAs) and differentially expressed protein-coding genes (DE PCGs) for GSNO- and IAA-treated seedlings. The edge sheets show gene pairs correlation (positive values) or anticorrelation (negative values). The node sheets show DElncRNA (bait node) and DE PCG (family node) identification. Additionally, the list of exclusive DE PCG identified for GSNO and IAA treatments are provided along with their biological role.

SUPPLEMENTARY TABLE S6
Details from 1) DElncRNA and DEmRNA, 2) DEsRNA and DElncRNA, and 3) DEsRNA and DEmRNA interactions analysis performed. Stronger interactions are represented by smaller negative values (kcal/mol).

SUPPLEMENTARY TABLE S7
Gene set enrichment analysis (GSEA) results for 1) GSNO-treated/WS-2 vs. IAA-treated/untreated *rhd6* and 2) GSNO/IAA-treated/WS-2 vs. untreated *rhd6* groups. The enriched genes are shown and those upregulated for GSNO are highlighted in yellow, whereas those downregulated are highlighted in green.

References

Bardou, F., Ariel, F., Simpson, C. G., Romero-Barrios, N., Laporte, P., Balzergue, S., et al. (2014). Long noncoding RNA modulates alternative splicing regulators in arabidopsis. *Dev. Cell* 30, 166–176. doi:10.1016/j.devcel.2014.06.017

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* 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Biniaz, Y., Tahmasebi, A., Afsharifar, A., Tahmasebi, A., and Poczai, P. (2022). Meta-analysis of common and differential transcriptomic responses to biotic and abiotic stresses in *Arabidopsis thaliana*. *Plants* 11, 502. doi:10.3390/plants11040502

Cavalier, D. M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., et al. (2008). Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* 20, 1519–1537. doi:10.1105/tpc.108.059873

Chekanova, J. A., Gregory, B. D., Reverdatto, S. V., Chen, H., Kumar, R., Hooker, T., et al. (2007). Genome-Wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell* 131, 1340–1353. doi:10.1016/j.cell.2007.10.056

Di, C., Yuan, J., Wu, Y., Li, J., Lin, H., Hu, L., et al. (2014). Characterization of stress-responsive lncRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. *Plant J.* 80, 848–861. doi:10.1111/tpj.12679

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

Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-Proteins: Key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi:10.1104/pp.110.156000

Fukuda, M., Fujiwara, T., and Nishida, S. (2020). Roles of non-coding RNAs in response to nitrogen availability in plants. *Int. J. Mol. Sci.* 21, 8508. doi:10.3390/ijms21228508

Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., and Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl. Acad. Sci. U. S. A.* 105, 803–808. doi:10.1073/pnas.0709559105

Hayashi, T., and Kaida, R. (2011). Functions of xyloglucan in plant cells. *Mol. Plant* 4, 17–24. doi:10.1093/mp/ssq063

Jannesar, M., Seyedi, S. M., Moazzam Jazi, M., Niknam, V., Ebrahimzadeh, H., and Botanga, C. (2020). A genome-wide identification, characterization and functional analysis of salt-related long non-coding RNAs in non-model plant Pistacia vera L. using transcriptome high throughput sequencing. *Sci. Rep.* 10, 5585. doi:10.1038/s41598-020-62108-6

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.* 59, 86–101. doi:10.1111/jipb.12513

Kalvari, I., Nawrocki, E. P., Argasinska, J., Quinones-Olvera, N., Finn, R. D., Bateman, A., et al. (2018). Non-coding RNA analysis using the Rfam database. *Curr. Protoc. Bioinforma.* 62, e51. doi:10.1002/cpbi.51

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316, 1484–1488. doi:10.1126/science.1138341

Kim, E.-D., and Sung, S. (2012). Long noncoding RNA: Unveiling hidden layer of gene regulatory networks. *Trends Plant Sci.* 17, 16–21. doi:10.1016/j.tplants.2011.10.008

Kopp, F., and Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172, 393–407. doi:10.1016/j.cell.2018.01.011

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

Kozomara, A., and Griffiths-Jones, S. (2010). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39, D152–D157. doi:10.1093/nar/gkq1027

Lasky, J. R., Des Marais, D. L., Lowry, D. B., Povolotskaya, I., McKay, J. K., Richards, J. H., et al. (2014). Natural variation in abiotic stress responsive gene expression and local adaptation to climate in *Arabidopsis thaliana*. *Mol. Biol. Evol.* 31, 2283–2296. doi:10.1093/molbev/msu170

Lee, J. T. (2012). Epigenetic regulation by long noncoding RNAs. *Science* 338, 1435–1439. doi:10.1126/science.1231776

Li, J., Ma, W., Zeng, P., Wang, J., Geng, B., Yang, J., et al. (2015). LncTar: A tool for predicting the RNA targets of long noncoding RNAs. *Brief. Bioinform.* 16, 806–812. doi:10.1093/bib/bbu048

Li, W., Li, K., Zhang, Q., Zhu, T., Zhang, Y., Shi, C., et al. (2020). Improved hybrid de novo genome assembly and annotation of African wild rice, Oryza longistaminata, from Illumina and PacBio sequencing reads. *Plant Genome* 13, e20001. doi:10.1002/tpg2.20001

Liang, G., He, H., and Yu, D. (2012). Identification of nitrogen starvation-responsive MicroRNAs in *Arabidopsis thaliana*. *PLoS ONE* 7, e48951. doi:10.1371/journal.pone.0048951

Lin, W.-D., Liao, Y.-Y., Yang, T. J., Pan, C.-Y., Buckhout, T. J., and Schmidt, W. (2011). Coexpression-based clustering of Arabidopsis root genes predicts functional modules in early phosphate deficiency signaling. *Plant Physiol.* 155, 1383–1402. doi:10.1104/pp.110.166520

Liu, J., Wang, H., and Chua, N.-H. (2015). Long noncoding RNA transcriptome of plants. *Plant Biotechnol. J.* 13, 319–328. doi:10.1111/pbi.12336

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

Mann, M., Wright, P. R., and Backofen, R. (2017). IntaRNA 2.0: Enhanced and customizable prediction of RNA–RNA interactions. *Nucleic Acids Res.* 45, W435–W439. doi:10.1093/nar/gkx279

Marin, E., Jouannet, V., Herz, A., Lokerse, A. S., Weijers, D., Vaucheret, H., et al. (2010). miR390, Arabidopsis TAS3 tasiRNAs, and their Auxin Response Factor targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell* 22, 1104–1117. doi:10.1105/tpc.109.072553

Meng, Y., Ma, X., Chen, D., Wu, P., and Chen, M. (2010). MicroRNA-mediated signaling involved in plant root development. *Biochem. Biophys. Res. Commun.* 393, 345–349. doi:10.1016/j.bbrc.2010.01.129

Moro, C. F., Gaspar, M., Silva, F. R. da, Pattathil, S., Hahn, M. G., Salgado, I., et al. (2017). S-nitrosoglutathione promotes cell wall remodelling, alters the transcriptional profile and induces root hair formation in the hairless root hair defective 6 (*rhd6*) mutant of *Arabidopsis thaliana*. *New Phytol.* 213, 1771–1786. doi:10.1111/nph.14309

Mur, L. A. J., Mandon, J., Persijn, S., Cristescu, S. M., Moshkov, I. E., Novikova, G. V., et al. (2013). Nitric oxide in plants: An assessment of the current state of knowledge. *AoB PLANTS* 5, pls052. doi:10.1093/aobpla/pls052

Rakhmetullina, A., Zielenkiewicz, P., Pyrkova, A., Uteulin, K., and Ivashchenko, A. (2021). Prediction of characteristics of interactions of miRNA with mRNA of GRAS, ERF, C2H2 genes of *A. thaliana*, *O. sativa* and *Z. mays*. *Current Plant Biol.* 28.

Salgado, I., de Oliveira, H. C., and Braga, M. R. (2009). *Nitrate Reductase-Deficient Plants: A Model to Study Nitric Oxide Production and Signaling in Plant Defense Response to Pathogen Attack*. Nitric Oxide Plant Physiol, 89

Salgado, I., Oliveira, H. C., and Gaspar, M. (2017). Plant nitric oxide signaling under environmental stresses. *Mech. Plant Horm. Signal. Stress* 1, 345

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

Singh, A., Singh, P. K., Sharma, A. K., Singh, N. K., Sonah, H., Deshmukh, R., et al. (2019). Understanding the role of the WRKY gene family under stress conditions in pigeonpea (Cajanus cajan L.). *Plants* 8, 214. doi:10.3390/plants8070214

Singh, P., Dutta, P., and Chakrabarty, D. (2021). miRNAs play critical roles in response to abiotic stress by modulating cross-talk of phytohormone signaling. *Plant Cell Rep.* 40, 1617–1630. doi:10.1007/s00299-021-02736-y

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., and Mesirov, J. P. (2007). GSEA-P: A desktop application for gene set enrichment analysis. *Bioinformatics* 23, 3251–3253. doi:10.1093/bioinformatics/btm369

The UniProt Consortium (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* 45, D158–D169. doi:10.1093/nar/gkw1099

Thieme, C. J., Rojas-Triana, M., Stecyk, E., Schudoma, C., Zhang, W., Yang, L., et al. (2015). endogenous arabidopsis messenger RNAs transported to distant tissues. *Nat. Plants* 1, 15025–15029. doi:10.1038/nplants.2015.25

Tzfadia, O., Diels, T., De Meyer, S., Vandepoele, K., Aharoni, A., and Van de Peer, Y. (2015). CoExpNetViz: Comparative Co-expression networks construction and visualization tool. *Front. Plant Sci.* 6, 1194. doi:10.3389/fpls.2015.01194

Wang, J., Meng, X., Dobrovolskaya, O. B., Orlov, Y. L., and Chen, M. (2017). Non-coding RNAs and their roles in stress response in plants. *Genomics Proteomics Bioinforma.* 15, 301–312. doi:10.1016/j.gpb.2017.01.007

Wang, L., Mai, Y.-X., Zhang, Y.-C., Luo, Q., and Yang, H.-Q. (2010). MicroRNA171c-Targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in Arabidopsis. *Mol. Plant* 3, 794–806. doi:10.1093/mp/ssq042

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

Yan, X., Liu, X., Hong, C. U. I., and Zhao, M. (2022). The roles of microRNAs in regulating root formation and growth in plants. *J. Integr. Agric.* 21, 901–916. doi:10.1016/s2095-3119(21)63818-2

Yang, G., Li, Y., Wu, B., Zhang, K., Gao, L., and Zheng, C. (2019). MicroRNAs transcriptionally regulate promoter activity in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* 61, 1128–1133. doi:10.1111/jipb.12775

Yoon, E. K., Yang, J. H., Lim, J., Kim, S. H., Kim, S.-K., and Lee, W. S. (2010). Auxin regulation of the microRNA390-dependent transacting small interfering RNA pathway in Arabidopsis lateral root development. *Nucleic Acids Res.* 38, 1382–1391. doi:10.1093/nar/gkp1128

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, M., Zhao, H., Xie, S., Chen, J., Xu, Y., Wang, K., et al. (2011). Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20042–20047. doi:10.1073/pnas.1112186108

Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., et al. (2019). Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* 20, 5573. doi:10.3390/ijms20225573

Zhao, S., Zhang, X., Chen, S., and Zhang, S. (2020). Natural antisense transcripts in the biological hallmarks of cancer: Powerful regulators hidden in the dark. *J. Exp. Clin. Cancer Res.* 39, 187. doi:10.1186/s13046-020-01700-0

Zhu, B., Yang, Y., Li, R., Fu, D., Wen, L., Luo, Y., et al. (2015). RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. *J. Exp. Bot.* 66, 4483–4495. doi:10.1093/jxb/erv203

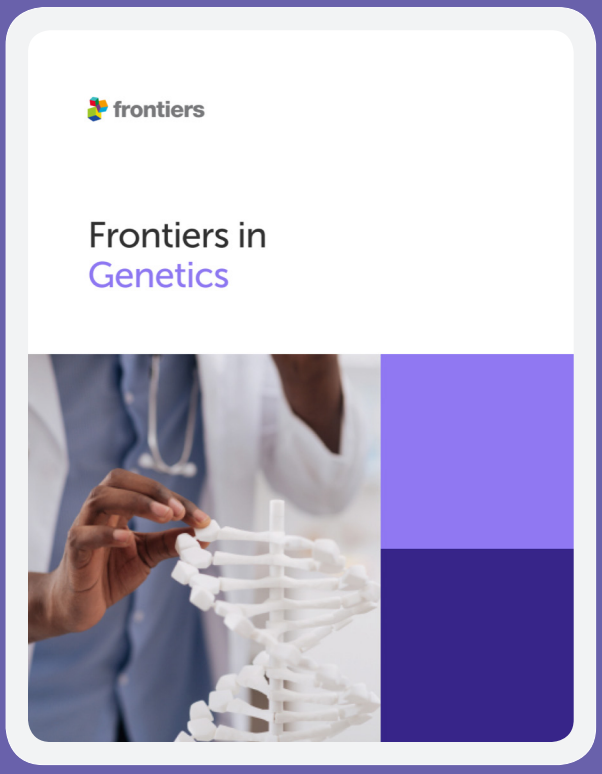
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact