

RESEARCH

Open Access



Integrated microRNA and whole-transcriptome sequencing reveals the involvement of small and long non-coding RNAs in the fiber growth of ramie plant

Yafen Fu¹, Langbo Yi², Fu Li^{1,2}, Jing Rao¹, Xiai Yang¹, Yanzhou Wang¹, Chan Liu¹, Touming Liu^{3*} and Siyuan Zhu^{1*}

Abstract

Background MicroRNAs (miRNAs) and long non-coding RNAs (lncRNAs) are the two main types of non-coding RNAs that play crucial roles in plant growth and development. However, their specific roles in the fiber growth of ramie plant (*Boehmeria nivea* L. Gaud) remain largely unknown.

Methods In this study, we performed miRNA and whole-transcriptome sequencing of two stem bark sections exhibiting different fiber growth stages to determine the expression profiles of miRNAs, lncRNAs, and protein-encoding genes.

Results Among the identified 378 miRNAs and 6,839 lncRNAs, 88 miRNAs and 1,288 lncRNAs exhibited differential expression. Bioinformatics analysis revealed that 29 and 228 differentially expressed protein-encoding genes were targeted by differentially expressed miRNAs and lncRNAs, respectively, constituting eight putative competing endogenous RNA networks. lncR00022274 exhibited downregulated expression in barks with growing fibers. It also had an antisense overlap with the MYB gene, *BntWG10016451*, whose overexpression drastically increased the xylem fiber number and secondary wall thickness of fibers in the stems of transgenic *Arabidopsis*, suggesting the potential association of lncR00022274-*BntWG10016451* expression with fiber growth.

Conclusions These findings provide insights into the roles of ncrRNAs in the regulation of fiber growth in ramie, which can be used for the biotechnological improvement of its fiber yield and quality in the future.

Keywords Ramie, Fiber growth, Small RNA, Long non-coding RNA, MYB transcriptional factor

*Correspondence:

Touming Liu
liutouming@caas.cn
Siyuan Zhu
zhusiyuan@caas.cn

¹Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences, Changsha, China

²College of Biology and Environmental Sciences, Jishou University, Jishou, China

³Yangzhou University, Yangzhou, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Plant fibers are essential for humans as they are a major source of raw materials for the production of paper, textiles, and composites [1]. In addition, fibers are important for plant growth and development because they provide mechanical support to different organs as well as the entire plant body [2]. Fiber formation has gained attention in studies of model species because of the remarkable roles of fibers within plants and their commercial applications. Plant fibers are a type of sclerenchymatous cells with thick secondary cell walls that are mainly composed of cellulose, hemicelluloses (xylan and glucomannan), and lignin [3]. Therefore, fiber growth is mainly involved in the biosynthesis of secondary cell walls [4]. Numerous enzyme-encoding genes are involved in the biosynthesis of secondary cell walls, participating mainly in cellulose biosynthesis and assembly, glucomannan biosynthesis, lignin biosynthesis and polymerization, and the patterned deposition of secondary cell walls [3]. Expression of these enzyme-encoding genes is coordinated at an appropriate time by an NAC-MYB-based transcriptional network [3, 5]. In *Arabidopsis*, at least 10 NAC genes have been identified as the top-level master switches of this regulatory network who regulate the expression levels of the second-level master switches, MYB46 and MYB83; MYB46/MYB83 subsequently modulates the expression levels of downstream transcription factors, precisely controlling the entire secondary wall biosynthesis process [3]. MYB transcription factors play an important role in the regulation of plant cell wall biosynthesis and can be classified into the following four types based on the number of MYB structural domains: 1R-MYB, R2R3-MYB, R1R2R3-MYB, and 4R-MYB [6–10]. R2R3-MYB is the largest subfamily of MYB transcription factors in plants [10, 11]. Hence, genomic identification of R2R3-MYB transcription factors has been performed in several sequenced plants. To date, 126 R2R3-MYB transcription factors have been identified in *Arabidopsis* [12], 108 in grape [13], 100 in sweet orange [14], and 222 in apple [15] plants.

In addition to genes regulating the growth of plant fibers, non-coding RNAs (ncRNAs) also affect the plant fibers. NcRNAs differ from their corresponding mRNAs because their base sequences do not encode proteins [16]. NcRNAs are generally divided into two classes based on the length cutoff: those with less than 200 nucleotides are usually referred to as short or small ncRNAs, including microRNAs (miRNAs), and those with more than 200 bases are known as long ncRNAs (lncRNAs) [17]. Although ncRNAs do not encode proteins, they play crucial roles in plant growth and development, including fiber growth [18]. Genome-wide expression profiling has identified many differentially expressed non-coding miRNAs and lncRNAs that are potentially involved in the regulatory network of secondary wall biosynthesis in various species [19–22]. Functional characterization of *miR397a* revealed that this ncRNA modulates the secondary wall biosynthesis by targeting two laccase genes, *LAC17* and *LAC4*, which are involved in secondary wall biosynthesis in *Citrus* [23]. In *Arabidopsis*, *TCP4* promotes, whereas non-coding *miR319* inhibits the biosynthesis of secondary cell wall by targeting *TCP4* [24].

Ramie (*Boehmeria nivea* L. Gaud) is one of the oldest fiber crops in China, with a cultivation history of 4,700 years in the country [25]. Ramie fibers extracted from the stem bark possess many excellent characteristics [26]. Although significant efforts have been made to understand the mechanisms underlying fiber growth [27–31], the basic mechanisms of miRNAs and lncRNAs involved in secondary cell wall biosynthesis in the fibers of many species, especially those of ramie plant, remain unclear. Therefore, in this study, we characterized and compared the expression profiles of miRNAs and lncRNAs between the bast barks of the top stem (TPS) and middle stem (MPS) sections, where the secondary cell walls of fibers are thick and do not initiate plant growth [21], to identify the ncRNAs associated with fiber growth. Our study provides an important basis for understanding the involvement of ncRNAs in the regulation of ramie fiber growth.

Results

Characterization of miRNA expression profiles in the bast barks

Six bark tissues (three from TPS and three from MPS) were used for small RNA sequencing and yielded a total of 189.0 million clean reads (Table 1). Of these reads, 170.3 million were aligned with known miRNAs by searching a small RNA database, leading to the annotation of 176 ramie miRNAs, and the remaining reads were used for predicting miRNAs, leading to the identification of 202 novel miRNAs (Table S1). Correlation analysis of miRNA expression levels among the six samples revealed a distinct difference between TPS and MPS sample (Fig. S1). Therefore, we compared the expression levels of

Table 1 Statistics of reads from small RNA and mRNA sequencing

Sample	Small RNA sequencing		Whole-transcriptome sequencing	
	Clean reads	mapped	Clean reads	mapped
MPS1	31,414,380	29,391,618	152,009,656	119,859,614
MPS2	31,893,369	28,414,895	152,510,394	115,785,891
MPS3	32,372,805	29,007,238	151,294,888	112,260,807
TPS1	30,979,716	28,579,251	151,382,004	122,664,838
TPS2	32,788,217	29,011,030	151,229,676	112,288,034
TPS3	30,419,934	25,898,874	151,692,564	112,540,713
Total	189,868,421	170,302,906	910,119,182	695,399,897

these 378 miRNAs between the TPS and MPS tissues. The results revealed that 51 and 37 miRNAs exhibited downregulated and upregulated expression, respectively, in fiber-growing MPS tissues (Fig. 1a; Table S2). Five differentially expressed miRNAs were randomly selected for real-time quantitative polymerase chain reaction (qPCR) analysis and their expression differences were further verified (Fig. S2). The results indicated that the differential

expression detected using small RNA sequencing was reliable. These differentially expressed miRNAs may play important roles in controlling ramie fiber growth.

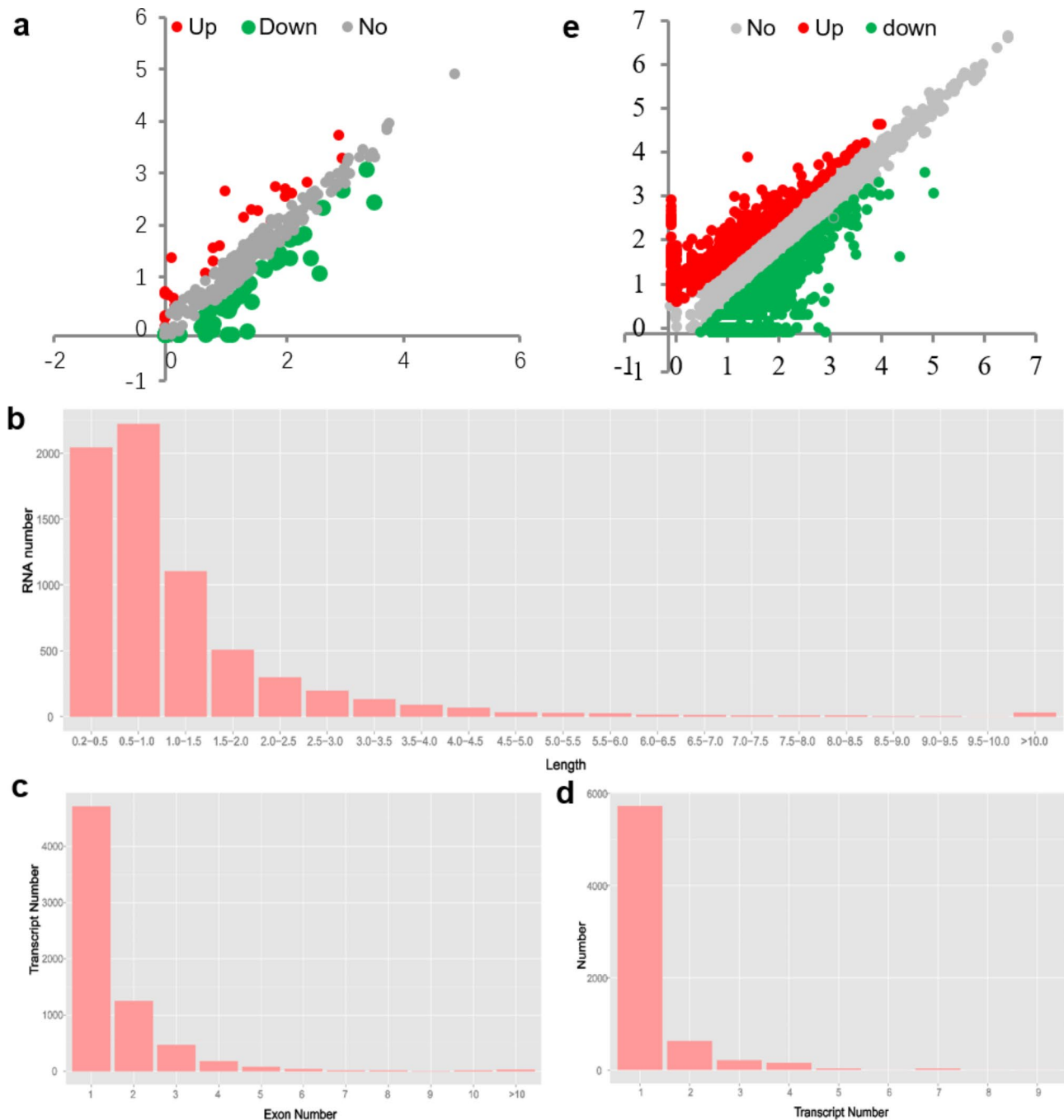


Fig. 1 Characterizing the miRNA and lncRNA in the stem barks of ramie. Comparison of the expression level of miRNAs (**a**) and lncRNAs (**e**) between TPS and MPS libraries. Red dots represent transcripts more prevalent in the MPS library, green dots show those present at a lower frequency in the MPS library, and grey dots indicate transcripts that did not change significantly. **b-d**, Distribution of lncRNAs length (**b**), exon number (**c**), and transcript number (**d**)

Identification of differentially expressed lncRNAs and protein-encoding genes between TPS and MPS via whole-transcriptome sequencing

Whole-transcriptome sequencing is a high-efficient tool for the identification of both protein-encoding genes and ncRNAs [32]. In this study, we performed whole-transcriptome sequencing of barks from TPS and MPS, yielding 910 million sequence reads from libraries of six samples (Table 1). These reads were assembled by aligning them with the ramie genome. After predicting the protein-coding potential of the assembled transcripts, 6,839 lncRNAs were identified in ramie (Fig. S3). The length of the lncRNAs ranged from 200 to 76,889 bp, with an average length of 1217.2 bp, which is far smaller than the length of the protein-encoding transcripts (Fig. 1b) [31]. Only a few exons were observed in the identified lncRNA-transcribing genes (average 1.6 exons), and 5,968 of the 6,839 lncRNA-transcribing genes (~87.2%) had only one or two exons (Fig. 1c), indicating that fewer exons existed in the lncRNA-transcribing genes than in the protein-encoding genes (~6.99 exons) [31]. Of these lncRNA-transcribing genes, 1,107 could yield more than one transcript because of alternative splicing (Fig. 1d). Cluster analysis indicated a distinct difference in the expression levels of these lncRNAs between TPS and MPS samples (Fig. S4). Of these, 1,288 showed differential abundance in the barks of TPS and MPS tissues (Fig. 1e; Table S3). We also randomly selected eight lncRNA-transcribing genes for real-time qPCR analysis and further validated their differences in expression (Fig. S5).

Sequence reads from whole-transcriptome sequencing were used to quantify the expression levels of ramie genes. A distinct difference in gene expression levels between TPS and MPS was observed (Fig. S6). After comparing the transcriptional abundance of ramie genes between the two investigated tissues, 3,308 genes were identified with differential expression (Fig. S7; Table S4). Cellulose is one of the most important components of secondary cell walls, and its biosynthesis is catalyzed by cellulose synthase. There were ten cellulose synthase genes with expression changes in fiber-growing MPS tissues (Table S5), including *BntWG10024039*, an orthologous gene of *Arabidopsis* IRX1 [33]. Notably, these differentially were significantly enriched in Gene Ontology (GO) terms related to secondary cell wall biosynthesis ($P < 0.01$; Fig. S8), including cell wall (GO:0005618), cellulose synthase activity (GO:0016759), and cell wall organization or biogenesis (GO:0071554).

Potential roles of miRNAs in fiber growth

Bioinformatic prediction of the targets of miRNAs identified 683 interacting pairs, consisting of 266 miRNAs and 101 protein-encoding genes (Table S6). Of the

101 protein-encoding targets, 89 were targeted by differentially expressed miRNAs. To explore the role of differentially expressed miRNAs in fiber growth, we investigated the expression profiles of their protein-encoding targets in the stem bark and identified 29 protein-encoding targets with altered expression in the fiber-growing bark (Table S7), including two laccase-encoding genes, *BntWG10014338* and *BntWG10014339*. Laccase is a pivotal enzyme that catalyzes the polymerization of monolignols and thus plays an important role in secondary wall biosynthesis during fiber growth [34, 35]. There was significant upregulation in the expression of *BntWG10014338*, *BntWG10014339* and *BntWG10013420* in the bark of fiber-growing MPS (Fig. 2), indicating their potential role in fiber growth. Notably, both were targeted by miR397a_4, miR397a_5, and miR828a_1, and three of these miRNAs were downregulated in the fiber-growing MPS (Fig. 2). It is likely that the downregulation of miR397a_4 and miR397a_5 caused the accumulation of transcripts of *BntWG10014338* and *BntWG10014339*, thereby promoting fiber growth. In addition, *BntWG10013420* is an ortholog of *AtMYB46* and *AtMYB83*, which are the two master switches for *Arabidopsis* secondary wall biosynthesis (Fig. 3a). We observed that the transcript abundance of *BntWG10013420* was higher in the libraries of MPS than in those of TPS (Fig. 2), which was potentially caused by the downregulated expression of miR828a_1 (Fig. 2), an miRNA with the putative ability to target the transcripts of *BntWG10013420* and to degrade them.

Potential roles of lncRNAs in fiber growth

To understand the function of the identified lncRNAs, we performed bioinformatic prediction of their targets. The results revealed that 1,877 lncRNAs were found to target 2,123 ramie genes, leading in 2,630 interacted pairs (Table S8). Of these interactions, 668, 367, and 1,595 were targeted via the trans-, cis-overlap, and cis-non-overlap. To identify the lncRNAs involved in fiber growth, we investigated the targets of differentially expressed lncRNAs and identified 228 targets with altered expression in MPS (Table S9), including IRX15L-homologous *BntWG10008444*. IRX15L encodes a DUF579-containing protein involved in xylan deposition in the secondary wall [36]. A recent proteomic study revealed that the protein abundance of *BntWG10008444* was significantly downregulated in MPS (Fig. 3b; Li et al., 2021). In this study, we observed a distinct decrease in the transcript abundance of *BntWG10008444* in the MPS libraries (Fig. 3b), further validating the finding of expression changes from the previous proteome analysis. Notably, *BntWG10008444* was a target of differentially expressed lncR00050912 via antisense overlapping.

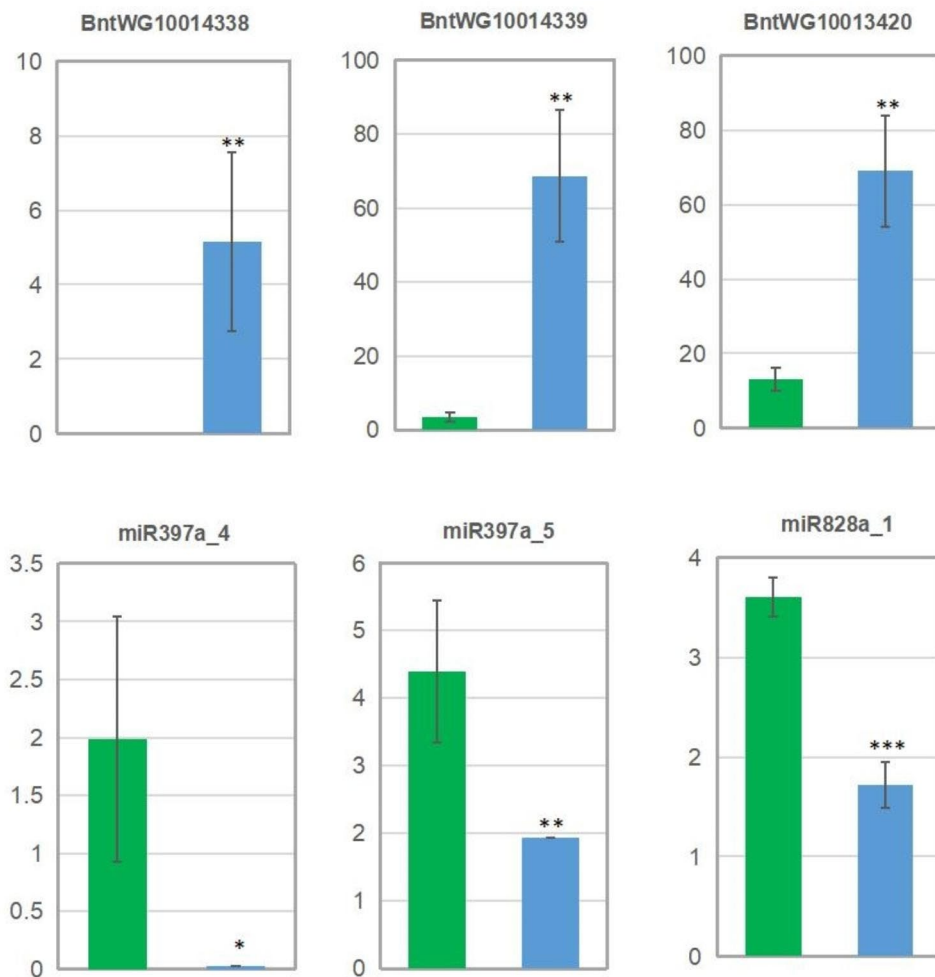


Fig. 2 Comparison of expression level of three miRNAs and their targets in TPS and MPS. The y axis of above three figures indicates the FPKM values, while that of above three figures represents TPM value. The green and blue column represents the value of TPS and MPS, respectively. *, **, and *** indicates that the value of MPS has significant difference comparing with that of TPS at the level of 0.05, 0.01, 0.001, respectively

MYB transcription factors play crucial roles in the regulation of fiber growth in *Arabidopsis* [3]. We identified the MYB gene *BntWG10016451* which was targeted by *lncR00022274*, using an antisense overlapping method. Overexpression of this MYB gene caused a significant increase in fibers and drastically thickened the secondary wall of fibers in transgenic *Arabidopsis* (Fig. 3c and d), indicating its potential role in ramie fiber growth. Notably, significant downregulation in the expression of *lncR00022274* was observed in the fiber-growing MPS (Fig. 3e). *lncR00022274* potentially played a role in fiber growth by interacting with *BntWG10016451* (Fig. 3f); this lncRNA predicted fiber growth by recruiting and silencing the transcript of *BntWG10016451*, whereas in MPS, with growing fibers, the expression of this lncRNA was downregulated, thereby activating *BntWG10016451* to promote fiber growth in ramie.

Identification of the competing endogenous RNA (ceRNA) network associated with fiber growth

Bioinformatics analysis of the interactions between miRNAs and lncRNAs revealed that 131 miRNAs targeted 561 lncRNAs, leading to 866 interacting pairs (Table S10). If miRNAs can target both protein-encoding mRNA and lncRNAs, competition for miRNAs between lncRNAs and mRNA will lead to the formation of a ceRNA network. To identify putative ceRNA networks involved in fiber growth, we identified interacting pairs consisting of differentially expressed miRNAs, lncRNAs, and protein-encoding mRNAs. Finally, we identified eight ceRNAs consisting of 12 differentially expressed miRNAs, 14 differentially expressed lncRNAs, and 14 differentially expressed protein-encoding genes (Fig. 4), indicating that these ceRNAs are associated with fiber growth in ramie. Interestingly, *BntWG10013420*, an orthologous gene of *Arabidopsis* secondary wall biosynthetic *AtMYB46* and *AtMYB83* genes, was associated with one of the ceRNAs.

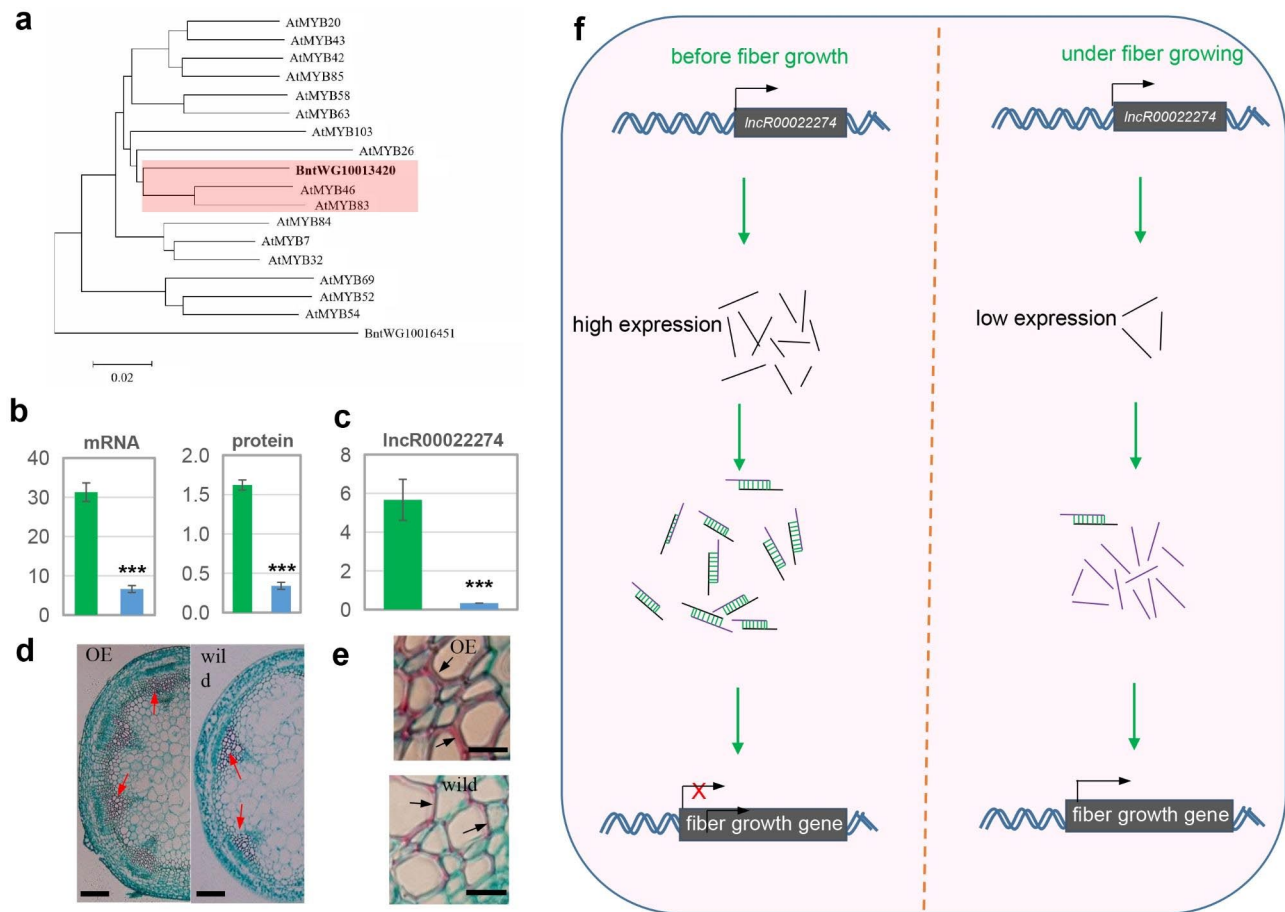


Fig. 3 Evidence for the involvement of lncRNA in fiber growth. **a**, Phylogenetic tree of two ramie MYB proteins and Arabidopsis secondary wall-biosynthetic MYB proteins. **b**, Amount of mRNA and protein abundance of *BntWG10008444* between bark from top (TPS; green) and middle (MPS; blue) portions of stems significantly differed between the two tissues ($P < 0.001$). **c**, Significant difference in the FPKM value of *lncR00022274* was observed between the TPS (green) and MPS (blue) ($P < 0.001$). **d, e**, Light microscopy findings of transacted stems of wild and *BntWG10016451*-overexpressing *Arabidopsis*. Arrows in **d** and **e** indicates the cell wall of fiber cells and xylem regions, respectively. Scale bar = 20 μm (d) and 200 μm (e). **f**, A possible model for explaining the function of *lncR00022274*. The black and purple short lines represent the transcript of *lncR00022274* and its target *BntWG10016451*. Before fiber growth, numerous transcripts of *lncR00022274* anti-overlapped with the MYB gene *BntWG10016451*, thereby to inhibit the translation activity of *BntWG10016451*

Discussion

Numerous efforts have been made to determine the fiber growth mechanism [21, 27, 29, 30], which has greatly improved our understanding of its genetic regulation and lead to the establishment of a putative model for fiber formation [28]. Protein post-translational modifications, such as phosphorylation and ubiquitination, are involved in fiber formation [35, 37], indicating a complex mechanism involved in fiber growth. Although differentially expressed miRNAs and circular RNAs have been identified in the stem bark of growing fibers in ramie [20, 37], the specific roles of these ncRNAs and their targets remain unknown.

In this study, we systemically characterized the expression profiles of miRNAs, lncRNAs, and protein-encoding genes in the fiber-growing bark of stems and identified 29 and 228 protein-encoding genes potentially associated with fiber growth that were targeted by differentially

expressed miRNAs and lncRNAs, respectively. For example, laccase is one of the most important enzymes for fiber growth owing to its involvement in secondary wall biosynthesis [34, 35]. Two *Arabidopsis* laccase genes, *LAC17* and *LAC4*, are targeted by *miR397a* to mediate their functions. Similarly, this study identified two laccase genes with expression changes in the fiber-growing bark, suggesting their involvement in ramie fiber growth; these genes were targeted by two differentially expressed ramie miRNAs, *miR397a_4* and *miR397a_5*. Additionally, we identified a differentially expressed gene that was an ortholog of secondary wall-biosynthetic *IRX15L* [36], which was targeted by differentially expressed *lncR00050912*. Whole-transcriptome sequencing revealed the *BntWG10013420* gene, which has a similar DNA-binding and repression structural domain as the amino acids of the *Arabidopsis AtMYB4* gene. In contrast, the *Arabidopsis AtMYB4* gene has an inhibitory

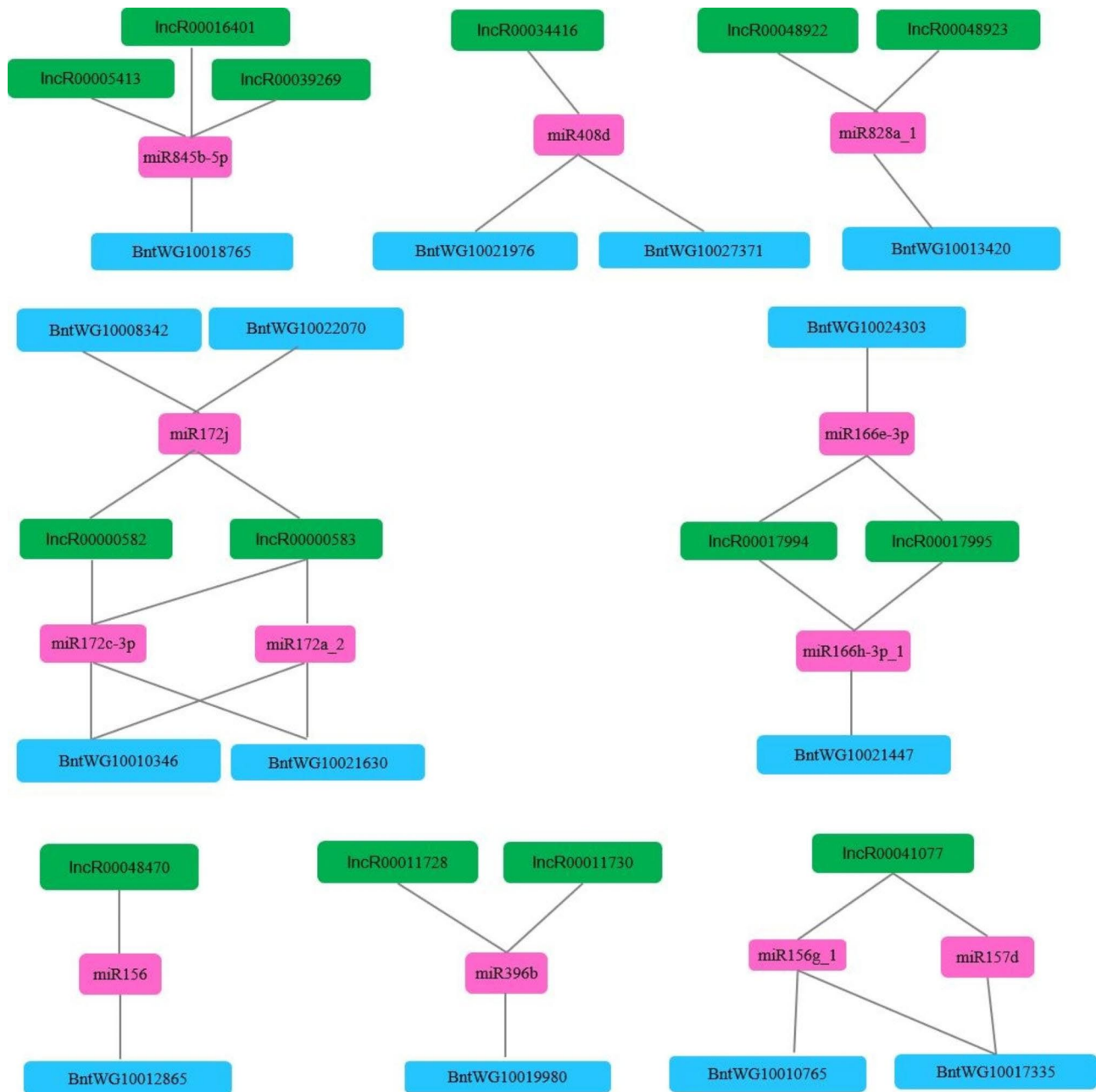


Fig. 4 Putative competing endogenous RNA network consisting of 12 differentially expressed miRNAs (pink), 17 differentially expressed lncRNAs (green), and 16 differentially expressed genes (blue)

effect on secondary wall biosynthesis, particularly in the phenylpropane synthesis pathway of *Arabidopsis thaliana*, blocking lignin biosynthesis [38]. The same phenotype was observed with the overexpression of *TaMYB1D* [39] and *TaMYB4* [40] in wheat, *PvMYB4* [41] in willowherb (*Panicum virgatum*), and *PtMYB156* [42] in poplar, all of which negatively affect secondary cell wall biosynthesis. Loss of function of the rice (*Oryza sativa*) MYB4 transcription factor (Os MYB108) [43] favors lignification. Furthermore, we hypothesized that the *BntWG10013420* gene may be a candidate transcription

factor that inhibits secondary wall biosynthesis in ramie. Therefore, this study provides distinct evidence supporting the involvement of miRNAs and lncRNAs in ramie fiber growth.

MYB transcription factors play pivotal roles in triggering secondary wall biosynthesis, and at least 16 MYB proteins have been reported to be involved in the regulatory network of secondary wall biosynthesis in *Arabidopsis thaliana* [3]. *Whole_GLEAN_10015497* is a ramie MYB gene associated with fiber growth, and its overexpression causes a significant increase in the numbers of

xylem and phloem fibers and thickening of the secondary wall in stem fibers [44]. Interestingly, a previous study revealed that the function of *BntWG10015497* was modulated via ubiquitin modifications; that is, when fibers are growing, a significant reduction in the ubiquitination levels of this MYB protein results in its accumulation [44]. In this study, we identified another MYB gene, *BntWG10016451*, that was associated with the fiber growth of ramie. It was targeted by *lncR00022274* due to an antisense overlap in their sequences, suggesting that the function of *BntWG10016451* is potentially mediated by this lncRNA. These findings indicate that both lncRNA and ubiquitin modifications are involved in the modulation of MYB gene functions, suggesting a complex mechanism for fiber growth.

Conclusions

In this study, we identified 88 miRNAs and 1,288 lncRNAs exhibiting differential expression between the TPS and MPS tissues. These differentially expressed miRNAs and lncRNAs were predicted to target 29 and 228 differentially expressed protein-encoding genes, respectively, resulting in eight putative ceRNA networks. *lncR00022274*, an lncRNA with downregulated expression in MPS, exhibited antisense overlap with the MYB gene, *BntWG10016451*. Notably, overexpression of this MYB gene drastically increased the xylem fiber number and thickened the secondary cell wall of fibers in transgenic *Arabidopsis*, suggesting the potential role of the *lncR00022274*-*BntWG10016451* module in ramie fiber growth.

Methods

Experimental material and RNA isolation

Zhongzhu 1, an elite cultivar that is widely cultivated in China, was selected as the experimental material for this study. Seedlings were planted in the experimental farm of the Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences (Yuanjiang, China) in June, 2016. In 2018, according to the description of Li et al. [37], sections from the top and middle parts of the stems of plant (30-day-old) were separately collected, mixed, and used as TPS and MPS samples, respectively. Three biological replicates of TPS and MPS were collected and immediately frozen in liquid nitrogen. Total RNA was extracted using the E.Z.N.A. Plant RNA kit (OMEGA Bio-Tek, Norcross, GA, USA), according to the manufacturer's protocol.

Library construction and sequencing

Total RNAs from six samples were used to construct the small RNA and cDNA libraries. For whole-transcriptome sequencing, cDNA libraries containing both mRNAs and lncRNAs were constructed using RNAs, except

ribosomal RNAs. Small RNA sequencing was performed using the BGISEQ-500 platform (BGI, Shenzhen, China), and whole-transcriptome sequencing was performed using the Illumina sequencing platform (HiSeq 2500; Illumina, San Diego, CA, USA), according to the manufacturers' instructions. The raw reads for each sample were filtered to generate clean reads for further analysis.

Transcriptome assembly and protein-coding potential analysis

Clean reads sequenced from the cDNA libraries were aligned to the ramie genome (accession ID: PRJNA663427) using the HISAT software [45]. StringTie program [46] was used to assemble the transcripts with default parameters. The protein-coding potential of the transcripts was determined using four approaches: alignment to the PFAM database (default parameters) [47] and prediction using the CPC software [48], txCdsPredict (<http://hgdownload.soe.ucsc.edu/admin/jksrc.zip>), and CNCI [49]. Score thresholds were set to distinguish lncRNAs from mRNAs for the three programs: 0 for CPC and CNCI and 500 for txCdsPredict. Transcripts reported as lncRNAs using at least three of the four prediction methods were confirmed as lncRNAs.

Prediction of miRNAs

To identify miRNAs, clean reads from small RNA sequences were aligned to the miRbase [50] and Rfam [51] databases using the AASRA [52] and CMsearch software [53], respectively. Reads unaligned to these two databases were used to predict novel miRNAs using PIPmiR [54] with default parameters. To determine the homologs of our selected candidates from *Arabidopsis* fiber growth-related MYB/NAC proteins, protein alignment was carried out using the Clustal program [55], and an unrooted phylogenetic tree was constructed using the MEGA 5 software with the neighbor-joining method and a bootstrap test with 1000 replicates [56].

Prediction of miRNA and lncRNA targets

The miRNA targets were predicted using TAPIR [57] and Target Finder [58] with default parameters. Only the targets identified by both tools were considered. lncRNAs execute their functions by targeting mRNAs in cis- or trans-manner [59]. Therefore, to predict the targets of lncRNA, Spearman's and Pearson's correlation coefficients were estimated between protein gene and lncRNA, and only the lncRNA-protein gene pair with $\text{spearman_cor} \geq 0.6$ and $\text{pearson_cor} \geq 0.6$ was considered for subsequent analyses. If the lncRNA and protein gene were co-located in a 10-kb region, the lncRNA was deemed to target the protein gene via cis-regulation. Otherwise, the RNAplex software was used to analyze the trans-acting targets [60], with parameters '-e -30'. Functional

analyses of the lncRNA target genes were conducted using Blast2GO [61]. GO terms with P -value < 0.05 were considered to be enriched [62].

Differential expression analysis

To estimate the expression levels of protein-encoding genes and lncRNAs, clean reads from transcriptome sequencing were aligned to the ramie genome (accession ID: PRJNA663427) using the Bowtie2 software [63], and the read number of each transcript was calculated using the RSEM software [64]. Fragments per kilobase of transcript sequence per million base pairs sequenced (FPKM) values were estimated to measure the expression level of each transcript [65]. For miRNAs, the number of clean tags for each miRNA from small RNA sequencing was calculated and normalized to the number of transcripts per million clean tags (TPM) to measure the expression level of each miRNA [66]. Differentially expressed genes, lncRNAs, and miRNAs were detected using the DEGseq program based on the MA-plot algorithm [67]. To identify the expression differences, P -values adjusted (Q value) for multiple testing using the Benjamini–Hochberg procedure [68] were used to control for false-positives. Sequences were deemed to be significantly differentially expressed if Q value < 0.05 and at least a two-fold change was obtained by dividing the average FPKM/TPM value of the three MPS libraries by that of the three TPS libraries.

Overexpression of *BntWG10016451*

To confirm the functions of *BntWG10016451*, its full-length cDNA was amplified using high-fidelity thermostable DNA polymerase, with the following the primer sequences: 5'-ATGACAAGAGACCCAAAGC-CAA-3' (forward primer) and 5'-TCTTCT-GAAAGAAAGCATATTTGGC-3' (reverse primer). Amplified fragments were ligated to the downstream region of the *CaMV* 35 S promoter in PBI121 vector. After confirming proper ligation via Sanger sequencing, 35 S::*BntWG10016451* was introduced into the *Agrobacterium tumefaciens* strain, GV3101, using heat shock, and the resultant *Agrobacterium* was infiltrated into *Arabidopsis* using the floral dip method [69]. Transgenic *Arabidopsis* plants were grown in a growth chamber under the following conditions: temperature, 22 °C; light, 16-h/8-h light/dark cycle. Finally, stem cell tissue sections prepared from 40-days-old transgenic plants were observed under a transmission electron microscope.

List of abbreviations

ceRNA	competing endogenous RNA
lncRNAs	long non-coding RNAs
miRNAs	microRNAs

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09711-9>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

Not Applicable.

Authors' contributions

Yafen Fu, Langbo Yi, Fu Li and Jing Rao performed the experimental lab works. Xiai Yang supervised the work. Yanzhou Wang performed the field management and collected the sample. Touting Liu and Siyuan Zhu conceived the work and wrote the manuscript. Song Gao revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was financially supported by the National Natural Science Foundation of China (31871678, 32001512), Natural Science Foundation of Hunan Province (2021JJ30769, 2022JJ30649), and the China Agriculture Research System of MOF and MARA (CARS-16).

Data Availability

These sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession number GSE130587. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130587>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Not Applicable.

Consent for publication

Not Applicable.

Received: 7 February 2023 / Accepted: 3 October 2023

Published online: 09 October 2023

References

1. Rai KM, Thu SW, Balasubramanian VK, Cobos CJ, Disasa T, Mendu V. Identification, characterization, and expression analysis of cell wall related genes in sorghum bicolor (L.) Moench, a Food, Fodder, and Biofuel Crop. *Front Plant Sci.* 2016;7:1287.
2. Kang C, Sun F, Yan L, Li R, Bai J, Caetano-Anollés G. Genome-wide identification and characterization of the vacuolar H⁺-ATPase subunit H gene family in crop plants. *Int J Mol Sci.* 2019;20(20):5125.
3. Zhong R, Ye ZH. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol.* 2015;56(2):195–214.
4. Gorshkova T, Brutch N, Chabbert B, Deyholos M, Hayashi T, Lev-Yadun S, et al. Plant fiber formation: state of the art, recent and expected progress, and open questions. *CRC Crit Rev Plant Sci.* 2012;31:201–28.
5. Nakano Y, Yamaguchi M, Endo H, Rejab NA, Ohtani M. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front Plant Sci.* 2015;6:288.
6. Wu TPC, Li B, et al. OsPGIP1-Mediated resistance to bacterial leaf streak in rice is beyond responsive to the polygalacturonase of *Xanthomonas oryzae* pv. *oryzicola*. *Rice (N Y).* 2019;12(1):90.
7. Rosinski JA, Atchley WR. Molecular evolution of the myb family of transcription factors: evidence for polyphyletic origin. *J Mol Evol.* 1998;46:74–83.

8. Jin H, Martin C. Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol Biol*. 1999;41:577–85.
9. Dubos C, Stracke R, Grotewold E, Weissshaar B, Martin C, Lepiniec L. MYB transcription factors in Arabidopsis. *Trends Plant Sci*. 2010;15:573–81.
10. Geethalakshmi S, Barathkumar S, Prabu G. The MYB transcription factor family genes in sugarcane (*Saccharum* sp.). *Plant Mol Biol Reprod*. 2015;33:512–31.
11. Jiang C, Gu X, Peterson T. Identification of conserved gene structures and carboxy-terminal motifs in the myb gene family of Arabidopsis and *Oryza sativa* L. ssp. indica. *Genome Biol*. 2004;5:R46.
12. Stracke R, Werber M, Weissshaar B. The R2R3-MYB gene family in Arabidopsis thaliana. *Curr Opin Plant Biol*. 2001;4:447–56.
13. Matus JT, Aquea F, Arce-Johnson P. Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes. *BMC Plant Biol*. 2008;8:83.
14. Liu C, Wang X, Xu Y, Deng X, Xu Q. Genome-wide analysis of the R2R3-MYB transcription factor gene family in sweet orange (*Citrus sinensis*). *Mol Biol Rep*. 2014;41:6769–85.
15. Cao ZH, Zhang SZ, Wang RK, Zhang RF, Hao YJ. Genome wide analysis of the apple MYB transcription factor family allows the identification of MdoMYB121 gene conferring abiotic stress tolerance in plants. *PLoS ONE*. 2013;8(7):e69955.
16. Yu Y, Zhang Y, Chen X, Chen Y. Plant noncoding RNAs: hidden players in development and stress responses. *Annu Rev Cell Dev Biol*. 2019;35:407–31.
17. Nagano T, Fraser P. No-nonsense functions for long noncoding RNAs. *Cell*. 2011;145(2):178–81.
18. Ha M, Kim VN. (2014). Regulation of microRNA biogenesis. *Nature reviews. Molecular cell biology*, 2014; 15(8), 509–524.
19. Ong SS, Wickneswari R. Characterization of microRNAs expressed during secondary wall biosynthesis in *Acacia mangium*. *PLoS ONE*. 2012;7(11):e49662.
20. Wang J, Huang JS, Hao XY, Feng YP, Cai YJ, Sun LQ. miRNAs expression profile in bast of ramie elongation phase and cell wall thickening and end wall dissolving phase. *Mol Biol Rep*. 2014;41(2):901–7.
21. Chen J, Quan M, Zhang D. Genome-wide identification of novel long non-coding RNAs in *Populus tomentosa* tension wood, opposite wood and normal wood xylem by RNA-seq. *Planta*. 2015;241(1):125–43.
22. Xie F, Jones DC, Wang Q, Sun R, Zhang B. Small RNA sequencing identifies miRNA roles in ovule and fibre development. *Plant Biotechnol J*. 2015;13(3):355–69.
23. Huang JH, Qi YP, Wen SX, Guo P, Chen XM, Chen LS. Illumina microRNA profiles reveal the involvement of miR397a in Citrus adaptation to long-term boron toxicity via modulating secondary cell-wall biosynthesis. *Sci Rep*. 2016;6:22900.
24. Sun X, Wang C, Xiang N, et al. Activation of secondary cell wall biosynthesis by miR319-targeted TCP4 transcription factor. *Plant Biotechnol J*. 2017;15(10):1284–94.
25. Kozłowski R, Batog J, Konczewicz W, et al. Enzymes in bast fibrous plant processing. *Biotechnol Lett*. 2006;28(10):761–5.
26. Jiang W, Song Y, Liu S, Ben H, Zhang Y, Zhou C, Han G, Ragauskas AJ. A green degumming process of ramie. *Ind Crop Prod*. 2018;120:131–4.
27. Liu C, Zeng L, Zhu S, Wu L, Wang Y, Tang S, Wang H, Zheng X, Zhao J, Chen X, Dai Q, Liu T. Draft genome analysis provides insights into the fiber yield, crude protein biosynthesis, and vegetative growth of domesticated ramie (*Boehmeria nivea* L. Gaud.). *DNA Res*. 2018;25:173–81.
28. Li F, Wang Y, Yan L, Zhu S, Liu T. Characterization of the expression profiling of circRNAs in the barks of stems in ramie. *Acta Agron Sinica*. 2021a;47:1020–30.
29. Zeng Z, Li F, Huang R, Wang Y, Liu T. Phosphoproteome analysis reveals an extensive phosphorylation of proteins associated with bast fiber growth in ramie. *BMC Plant Biol*. 2021;21:473.
30. Zeng Z, Zhu S, Wang Y, Bai X, Liu C, Chen J, Zhang T, Wei Y, Li F, Bao Z, Yan L, Wang H, Liu T. Resequencing of 301 ramie accessions identifies genetic loci and breeding selection for fiber yield traits. *Plant Biotechnol J*. 2022;20:323–34.
31. Wang Y, Li F, He Q, Bao Z, Zeng Z, An D, Zhang T, Yan L, Wang H, Zhu S, Liu T. Genomic analyses provide comprehensive insights into the domestication of bast fiber crop ramie (*Boehmeria nivea*). *Plant J*. 2021;107:787–800.
32. Fu X, Zhang X, Qiu J, Zhou X, Yuan M, He Y, et al. 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*. 2019;19:509.
33. Taylor NG, Laurie S, Turner SR. Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis. *Plant Cell*. 2000;12:2529–39.
34. Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Ce'zard L, Le Bris P, et al. Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. *Plant Cell*. 2011;23:1124–37.
35. Zhao Q, Nakashima J, Chen F, Yin Y, Fu C, Yun J, et al. Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. *Plant Cell*. 2013;25:3976–87.
36. Brown DM, Wightman R, Zhang Z, Gomez LD, Atanassov I, Bukowski JP, et al. Arabidopsis genes irregular xylem (IRX15) and IRX15L encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall. *Plant J*. 2011;66:401–13.
37. Li F, Zeng Z, Huang R, Wang Y, Liu T. Identification of proteins associated with bast fiber growth of ramie by differential proteomic analysis. *BMC Genomics*. 2021b;22:865.
38. Ma D, Constabel CP. MYB repressors as regulators of phenylpropanoid metabolism in plants. *Trends Plant Sci*. 2019;24:275–89.
39. Ma QH, Wang C, Zhu HH. TaMYB4 cloned from wheat regulates lignin biosynthesis through negatively controlling the transcripts of both cinnamyl alcohol dehydrogenase and cinnamoyl-CoA reductase genes. *Biochimie*. 2011;93:1179–86.
40. Wei Q, Zhang F, Sun F, Luo Q, Wang R, Hu R, Chen M, Chang J, Yang G, He G. A wheat MYB transcriptional repressor TaMyb1D regulates phenylpropanoid metabolism and enhances tolerance to drought and oxidative stresses in transgenic tobacco plants. *Plant Sci*. 2017;265:112–23.
41. Rao X, Chen X, Shen H, Ma Q, Li G, Tang Y, Pena M, York W, Frazier TP, Lenganhan S, et al. Gene regulatory networks for lignin biosynthesis in switchgrass (*Panicum virgatum*). *Plant Biotechnol J*. 2019;17:580–93.
42. Yang L, Zhao X, Ran L, Li C, Fan D, Luo K. PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Sci Rep*. 2017;7:41209.
43. Miyamoto T, Takada R, Tobimatsu Y, Takeda Y, Suzuki S, Yamamura M, Osakabe K, Osakabe Y, Sakamoto M, Umezawa T. OsMYB108 loss-of-function enriches *p*-coumaroylated and tricin lignin units in rice cell walls. *Plant J*. 2019;98:14290.
44. He Q, Zeng Z, Li F, Huang R, Wang Y, Liu T. Ubiquitylome analysis reveals the involvement of ubiquitination in the bast fiber growth of ramie. *Planta*. 2021;254:1.
45. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
46. Perteau M, Perteau GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015;33:290.
47. Finn R, Bateman A, Clements J, Coggill P, Eberhardt R, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res*. 2014;42:222–30.
48. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007;35:W345–9.
49. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Zhao Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res*. 2013;41:e166–6.
50. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42(Database issue):D68–D73.
51. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate. J. Rfam 12.0: updates to the RNA families database. *Nucleic acids research*. 2014; 43, D130-D137.
52. Tang C, Xie Y, Yan W. AASRA: an anchor alignment-based small RNA annotation pipeline. *BioRxiv*, 2017; 132928.
53. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013;29:2933–5.
54. Breakfield NW, Corcoran DL, Petricka JJ, Shen J, Sae-Seaw J, Rubio-Somoza I, Weigel D, Ohler U, Benfey PN. High-resolution experimental and computational profiling of tissue-specific known and novel miRNAs in Arabidopsis. *Genome Res*. 2012;22:163–76.
55. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947–8.
56. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011;28:2731–9.

57. Bonnet E, He Y, Billiau K, Van de Peer Y. TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics*. 2010;26:1566–8.
58. Fahlgren N, Carrington JC. miRNA target prediction in plants. *Plant MicroRNAs*. 2010; 51–7.
59. Kornienko AE, Guenzl PM, Barlow DP, Pauler FM. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol*. 2013;11:59.
60. Tafer H, Hofacker IL. RNAplex: a fast tool for RNA–RNA interaction search. *Bioinformatics*. 2008;24:2657–63.
61. Conesa A, Göttsch S, Garcíagómez JM, Terol JTalónM, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21:3674–6.
62. Wang T, Zhao M, Zhang X, Liu M, Yang C, Chen Y, Chen R, Wen J, Mysore KS. Novel phosphate deficiency-responsive long non-coding RNAs in the legume model plant *Medicago truncatula* *J Exp Bot*. 2017;68:5937–48.
63. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357.
64. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
65. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28:511.
66. Morrissy AS, Morin RD, Delaney A, Zeng T, McDonald H, Jones S, Zhao Y, Hirst M, Marra MA. Next-generation tag sequencing for cancer gene expression profiling. *Genome Res*. 2009;19:1825–35.
67. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics*. 2010;26:136–8.
68. Ferreira J, Zwiderman A. On the Benjamini–Hochberg method. *The Annals of Statistics*. 2006;34:1827–49.
69. Zhang X, Henriques R, Lin SS, Niu QW, Chua NH. Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat Protoc*. 2006;1:641.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.