Small RNA and degradome profiling reveals a role for miRNAs and their targets in the developing fibers of *Gossypium barbadense*

Nian Liu[1], Lili Tu[1], Wenxin Tang[1], Wenhui Gao[1], Keith Lindsey[2], Xianlong Zhang[1]*

[1] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei 430070, China

[2] Integrative Cell Biology Laboratory, School of Biological and Biomedical Sciences, University of Durham, South Road, Durham DH1 3LE, United Kingdom

*__Corresponding author:__ Xianlong Zhang

National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei 430070, China

Telephone number: +86-27-8728-0510

Fax number: +86-27-8728-0196

E-mail address: xlzhang@mail.hzau.edu.cn

Running title: miRNAs involved in fiber development

Key words: *Gossypium barbadense*; small RNA sequencing; degradome sequencing; fiber development; miRNA function verification

Word count: Total words: 8978 words including Summary: 175 words; Introduction: 661 words; Results: 2350 words; Discussion: 1838 words; Experimental procedures: 1387 words; Acknowledgments: 52 words; Short legends for supporting information: 139 words; References: 1758 words; Figure legends: 618 words

SUMMARY

microRNAs (miRNAs) are 20-24 nucleotide non-coding small RNAs that play important roles in plant development. The stages of cotton fiber development include initiation, elongation, secondary wall thickening (SWT) and maturation. We constructed seven fiber RNA libraries representing the initiation, elongation and SWT stages. A total of 47 conserved miRNA families and seven candidate miRNAs were profiled using small RNA sequencing. Northern blotting and real-time PCR analyses revealed the dynamic expression of miRNAs during fiber development. In addition, 140 targets of 30 conserved miRNAs and 38 targets of five candidate miRNAs were identified through degradome sequencing. Analysis of correlated expression between miRNAs and their targets demonstrated that specific miRNAs suppressed the expression of transcription factors, SBP and MYB, a leucine-rich receptor-like protein kinase, a pectate lyase, α-tubulin, a UDP-glucuronic acid decarboxylase and cytochrome C oxidase subunit 1 to affect fiber development. Histochemical analyses detected the biological activity of miRNA156/157 in ovule and fiber development. Suppressing miRNA156/157 function resulted in the reduction of mature fiber length, illustrating that miRNA156/157 plays an essential role in fiber elongation.

INTRODUCTION

Cotton is one of the world's most important commercial crops and is a major source of textile fiber. The two most cultivated species are *Gossypium hirsutum* and *Gossypium barbadense*. *G. hirsutum* (Upland cotton) has higher yield potential than *G.*

barbadense (Sea-Island cotton); however, Sea-Island cotton is much better than Upland cotton in fiber quality characteristics, such as length, strength and the micronaire value. Lint fibers differentiate from the epidermal cells of the ovule and can grow to 30 to 60 mm in length in approximately 50 days (Kim and Triplett, 2001). Many studies suggest that plant hormones such as auxin, ethylene and brassinosteroid play significant roles in fiber development at the initiation and elongation stages (Shi *et al.*, 2006; Luo *et al.*, 2007; Zhang *et al.*, 2011). In addition, ROS-mediated Ca^{2+} signaling also effects fiber elongation (Qin and Zhu, 2011). The effects of some transcription factors, such as MYB and HD-ZIP, on fiber initiation have been verified, and the functions of other genes in downstream regulation networks, such as cytoskeleton and cell wall-related genes, directly control fiber elongation and SWT (Li *et al.*, 2007; Pang *et al.*, 2010; Walford *et al.*, 2011; Walford *et al.*, 2012).

miRNAs, as a major type of small RNAs, widely participate in the regulation of plant morphogenesis and development, nutrient uptake and the stress response. For instance, miRNA156 (miR156), with its downstream gene miR172, controls flower timing (Wu *et al.*, 2009). miR172 also regulates floral organ development (Zhao *et al.*, 2007), and miR160 regulates the formation of the root cap (Wang *et al.*, 2005). Moreover, miR395 and miR399 regulate sulfate and phosphate accumulation, respectively, in plants (Chiou *et al.*, 2006; Liang *et al.*, 2010). Under high oxidative conditions, miR398 is suppressed through oxidative signals to enhance stress tolerance in plants (Sunkar *et al.*, 2006). At present, the mechanisms underlying the miRNA-mediated regulation of fiber development remain unclear. One potentially

valuable approach is to explore candidate miRNAs in fiber development through sequencing. In 2008, two miRNAs were cloned from the cotton ovule (0-10 days post-anthesis; DPA) through sequencing (Abdurakhmonov *et al.,* 2008). Using a high-throughput deep sequencing approach, Pang *et al.* (2009) profiled 27 conserved and four candidate miRNAs families in ovules (-3, 0, 3 DPA) and 7 DPA fibers. A total of 18 miRNA targets were also identified through RNA ligase-mediated rapid amplification of cDNA end (RLM-RACE). Studies of *G. hirsutum* cv. Xuzhou142 and the Xuzhou142 *fuzzless-lintless* mutant showed that many miRNAs accumulated differentially between each, suggesting a possible link with fiber development (Kwak *et al.,* 2009). Using these same cultivars, Wang *et al.* (2012) confirmed miRNA expression patterns in ovules (-3, -1, 0, 1, 3 DPA) through Northern blotting analyses, and verified five miRNA targets, which were predicted *in silico*.

Previous studies were focused on miRNA expression during early stage fiber development, and large-scale experimental identification of the miRNA targets has not been carried out. Although predicted miRNA targets could be verified using an RLM-RACE approach, the identification of authentic targets is time-consuming, laborious and difficult to perform in high-throughput. Recently, degradome sequencing has been developed and successfully used in studies with *Arabidopsis*, rice, grapevine and soybean (German *et al.,* 2008; Pantaleo *et al.,* 2010; Zhou *et al.,* 2010; Shamimuzzaman and Vodkin, 2012). Degradome sequencing combines deep sequencing technology and computer analysis to search for miRNA-guided cleaved sites in mRNAs. This allows identification of prospective miRNA targets on a large

scale.

To study the miRNA-mediated regulation of fiber development, we selected samples from ovules (-3, 0, 3 DPA) and fibers (7, 12, 20, 25 DPA) to sequence small RNAs, spanning from initiation stage to SWT stage. We identified 47 conserved miRNAs and seven candidate miRNAs. Several miRNA targets were identified through degradome sequencing, and the expression patterns of miRNAs and targets were used to predict miRNA/target modules that might be involved in fiber development. Transgenic techniques showed that reducing miR156/157 activity reduces mature fiber length, demonstrating function of these miRNAs in cotton fiber development.

RESULTS

Global small RNA profiling during fiber development

To characterize the global expression patterns of miRNAs from the cotton fiber, we constructed seven small RNA libraries using total RNAs extracted from ovules (-3, 0, 3 DPA) and fibers (7, 12, 20, 25 DPA) in *G. barbadense* cv. 3-79. The small RNA libraries were sequenced using high-throughput Illumina Solexa sequencing technology. After removing poor quality reads and adapter sequences, the total redundant reads from the seven libraries ranged from 16.47 to 20.13 million. Less than 4.26 million redundant reads matched the cotton gene index (CGI; reference EST database) in each library. For annotation, the small RNAs were grouped into several classes, including known miRNAs, rRNAs, snRNAs, snoRNAs and tRNAs (Table 1).

All the sequences with lengths between 18 and 26 nt were counted to determine the size distribution. The most abundant size of small RNAs in ovules at -3 to 3 DPA and fibers at 7 to 12 DPA was 24 nt, comprising 48.22% to 58.79% of small RNAs (Figure 1). However, the distribution of small RNAs of 24 nt in size was reduced to 30.67% and 22.15% in fibers at 20 to 25 DPA; clear peaks in abundance of 21 nt miRNAs were observed at late fiber development stages (20 - 25 DPA) (Figure 1).

Identification of conserved and candidate miRNAs in the cotton ovule and fiber

To identify conserved miRNAs in the libraries, small RNAs, which could not be defined as rRNAs, snRNAs, snoRNAs and tRNAs, were aligned against the miRNA sequences deposited in the miRbase 18. The criteria of the blast search required no more than two mismatches with the sequences in miRbase 18. A total of 0.78 million to 2.09 million redundant sequences were identified in the seven libraries as conserved miRNAs, which were clustered into 47 miRNA families (Table 1, S1 and S2). MIREAP software (http://sourceforge.net/projects/mireap/) further predicted 34 conserved miRNA precursors belonging to 18 conserved miRNA families (Figure S1).

Among the 47 conserved miRNA families, the three most abundant miRNA families (Gb-miR156/157, 166 and 167) whose normalized expression levels were greater than 2700 transcripts per million (TPM) clean tags, were highly conserved in mosses, eudicots and monocots (Table S1 and S3). Conversely, 18 relatively young miRNA families, with homologs in only one species, accumulated less than 800 TPM

(Table S1 and S3), consistent with the results of previous studies showing that evolutionarily young miRNAs exhibit lower expression than highly conserved miRNAs (Cuperus *et al.,* 2011).

After removing conserved miRNAs and other cellular small RNAs, we selected unknown small RNAs, which could be mapped to the reference database, to identify previously uncharacterised miRNAs. Based on the structure of the miRNA excised from a stem-loop precursor, seven sequences were annotated as candidate miRNAs through mapping onto stable stem-loop precursors (Table 2). The length of the candidate miRNAs was 20 to 21 nt, matching with the size of Dicer-like protein cleavage products, and the minimal folding free energy index (MFEI) of the seven candidate miRNA precursors ranged from 0.91 to 1.64 (Table 2), which are higher than tRNAs (0.64), rRNAs (0.59) and mRNAs (0.65) (Zhang *et al.,* 2006). MFEI was calculated by the formula (100×minimal folding free energy)/(length×G/C content) to distinguish miRNA from other RNAs. Sequences with MFEI values of more than 0.85 are most likely miRNAs (Zhang *et al.,* 2006). Consistent with these criteria, the seven sequences identified in the present study are authentic candidate miRNAs.

Expression of conserved and candidate miRNAs in fiber development

After identifying 47 conserved miRNA families and seven candidate miRNAs, we further attempted to characterize miRNA expression patterns in ovules (-3, 0 and 3 DPA) and fibers (7, 12, 20 and 25 DPA). The 30 conserved miRNA families and four candidate miRNAs, detected in all libraries with more than 10 TPM at least in one

library, were used to perform cluster analyses.

The seven fiber development libraries were categorized into three groups: ovules (-3, 0 and 3 DPA), early fibers (7, 12 DPA) and late fibers (20, 25 DPA) (Figure 2A). These groups were separately sampled, representing respectively three different stages of fiber development: initiation, elongation and SWT. Based on the hierarchical clustering method, the expression patterns of miRNAs were clustered into four classes (Figure 2A). In class I, the five conserved miRNA families (miR160, 167, 171, 172 and 827) were much more abundant at the initiation stage than at the elongation and SWT stages. The eight conserved miRNA families (miR156/157, 162, 165/166, 169, 390, 2949, 2911 and 3954) and the nmiR3 in class II were reduced from the initiation stage to the elongation stage, followed by a slight increase at the SWT stage. In class III, the abundance of the 13 conserved miRNA families and three candidate miRNAs continuously increased from the initiation stage to the SWT stage. The accumulation of the four conserved miRNA families in class IV dramatically increased with ovule development at the initiation stage and maintained high and stable levels from the elongation to the SWT stage.

Northern blotting analysis was used to further validate the expression of four conserved miRNA families (miR156/157, 165/166, 167 and 172, Figure 2B). The abundances of the four miRNA families were higher at the initiation stage than at the other two fiber development stages. The expression levels of Gb-miR167 and Gb-miR172 gradually reduced from the elongation stage to the SWT stage. However, Gb-miR156/157 accumulation increased from the elongation stage to the SWT stage,

and Gb-miR165/166 was barely detectable at the elongation and SWT stages. The results from the Northern blotting analyses were generally consistent with the cluster analysis data, demonstrating the dynamic expression of different miRNAs during fiber development.

Global identification of conserved and candidate miRNA targets through degradome sequencing

In plants, miRNA predominantly degrades targets by transcript cleavage. Thus, there should be an obvious degradome sequencing signal at the target site compared with other regions of the mRNA. Using a degradome sequencing approach (German *et al.*, 2008), we generated three libraries to identify miRNA targets at the initiation (-3, 0, 3 DPA ovules), elongation (7, 12 DPA fibers) and SWT (20, 25 DPA fibers) stages.

After sequencing three libraries, we obtained 8,089,323 unique reads for the initiation stage, 3,751,257 unique reads for the elongation stage and 2,547,751 unique reads for the SWT stage (Table 3). More than 62% of the unique reads were matched to the cotton EST database (CGI 11), representing 95,021, 87,495 and 81,718 transcripts in initiation, elongation and SWT stage libraries, respectively (Table 3). In total, 140 targets of 30 conserved miRNA families and 38 targets of five candidate miRNAs were identified (Table 4). The detailed annotation of each miRNA target is shown in Table S4. Signature abundance in the position of each target transcript is shown in Figure S2.

Almost one-third of the conserved miRNA targets (48 of 140) in our libraries were

transcription factors, such as SBP, GRAS, AP2, Class III HD-zip, MYB and NAC (Table 4). We also detected other signaling pathway-related gene products. Auxin response factor, auxin-signaling F-box protein and TAS, which are involved in the auxin signaling pathway, were identified as conserved miRNA targets (Table 4). Other conserved miRNA targets, such as calmodulin-binding protein and superoxide dismutase, are involved in the Ca^{2+} signaling pathway triggered through H_2O_2 (Apel and Hirt, 2004; Tang *et al.*, 2014). Moreover, cell wall and cytoskeleton-related genes, such as pectate lyase (PL), UDP-glucuronic acid decarboxylase (UGD) and α-tubulin, were identified as conserved miRNA targets in the libraries. Compared with conserved miRNA targets, many targets of candidate miRNA, such as adenylate translocator (ANT), pentatricopeptide repeat-containing protein and cytochrome C oxidase subunit 1 (CO1), were preferentially located in mitochondria for energy metabolism (Table 4).

Compared with the 21 targets previously identified in cotton fiber through RLM-RACE (Pang *et al.*, 2009; Wang *et al.*, 2012), we identified many more targets not previously found. To assess the reliability of the results, we used a RLM-RACE to verify seven targets identified through degradome sequencing. Equal mixtures of RNA were extracted from the ovules and fibers, independent of the samples of the degradome libraries. The cleaved products from the target sites were detected through degradome sequencing and RLM-RACE (Figure 3). Similar to the RLM-RACE results, the degradome sequencing results showed that the cleavage products of PL, nuclear transcription factor Y, leucine-rich receptor-like protein kinase (RLK) and

transport inhibitor response 1 were precisely mapped from the 9th to 11th position of complementarity from the miRNA 5' end, showing evidence of cleavage through Gb-miR159, 169, 390 and 393, respectively. On the other hand, the RLM-RACE results showed that cleavage products of *MYB*, *ANT* and *CO1* were relatively less abundant and did not map to the 9th to 11th position. The target plot of the data from the degradome analysis also demonstrated that the identified Gb-miR399, nmiR1 and nmiR3 slicing sites in *MYB*, *ANT* and *CO1*, respectively, were not the unique significant cleavage sites observed in the mRNAs.

Expression correlation between miRNAs and their targets

To assess the influence of the miRNAs on their targets, we analyzed the correlation between miRNAs and the identified targets. When a miRNA triggers target mRNA degradation to regulate fiber development, the expression of a target should be negatively correlated with miRNA expression. Quantitative RT-PCR (qRT-PCR) was used to quantify the expression of the target transcripts and their corresponding miRNAs. Seven interesting miRNA/target modules were identified through qRT-PCR analysis (Figure 4).

Previous studies have shown that transcription factors play important roles in fiber development (Walford *et al.*, 2011; Walford *et al.*, 2012). In the present study, the expression of the transcription factor *SBP* (TC253516) gradually decreased from 7 to 25 DPA fibers; conversely, during the same period, the corresponding Gb-miR156/157 abundance increased (Figure 4A). Similarly, Gb-miR399 was also

dramatically accumulated in the fibers (20, 25 DPA). Accordingly, the identified target *MYB* transcription factor was clearly down-regulated (Figure 4F). *RLK* (TC259635) was also identified as the Gb-miR390 target (Figure 3C), which participates in receptor kinase signaling in plants (De Smet *et al.,* 2009). In general, the expression of *RLK* in the ovules (-3, 0, 3 DPA) was lower than in the fibers (7, 12, 20 and 25 DPA). In contrast, the Gb-miR390 accumulation was clearly high in the ovules (-3, 0, 3 DPA) (Figure 4E).

Many cell wall- and cytoskeleton-related genes exhibit fiber preferential expression patterns and affect fiber morphogenesis (Gou *et al.,* 2007; Hovav *et al.,* 2008; Haigler *et al.,* 2012). In the present study, *PL* (DN804324), *UGD* (TC236955) and *α-Tubulin* (TC232723) were identified as the targets of Gb-miR159, Gb-miR164 and Gb-miR167, respectively, in cotton fiber (Figure 3A, Table S4). Gb-miR159 was expressed at higher levels in late fibers (20, 25 DPA) than in early fibers (7, 12 DPA), and the Gb-miR159 target accumulated at lower levels in the late fibers (20, 25 DPA) than in other fiber development periods (Figure 4B). A comparison of the expression levels in the ovules (-3, 0, 3 DPA) with that in the fibers (7, 12, 20, 25 DPA) revealed a negative relationship between Gb-miR164 and its target *UGD* and Gb-miR167 and its target *α-Tubulin* (Figure 4C, D). In addition, when the expression of nmiR3 was dramatically reduced from -3 to 3 DPA in the ovule, the target *CO1* (ES802293) inversely increased during the same period, indicating that nmiR3 may suppress the expression of CO1 at the fiber initiation stage (Figure 3G and 4G).

Suppression of Gb-miR156/157 affects fiber elongation

To verify miRNA activity in cotton ovules and fibers, a sensor containing the miRNA reverse complement sequence in the 3' untranslated region (UTR) of a constitutively expressed GUS gene was constructed (Figure 5A). When an miRNA recognizes the miRNA reverse complement sequence, the GUS signal would be suppressed. A construct without the miRNA reverse complement sequence was constructed as a positive control. Vectors harboring miR156/157, miR165/166, miR172 and positive control sensors were used to transform *G. hirsutum* cv. YZ1 and at least five transformants were obtained. GUS staining showed that the GUS signals of the miR156/157, miR165/166 and miR172 sensors in 0 DPA ovules were barely detectable or dramatically reduced compared with the positive control (Figure 5B). In 25 DPA, the GUS signal of the miR156/157 and miR172 sensors were also much lower than that in the positive control but the miR166 sensor was slightly weaker (Figure 5B). However, in non-fiber tissues, the germinating seeds, there were no obvious differences between the miRNA and the positive control sensors (Figure 5B), suggesting that miR156/157, 166 and 172 were indeed activated in the ovule and fiber.

Interestingly, the fiber length of three miR156/157-sensor transformed plants was lower than that in the positive control plant and wild type plant (Figure 6A, S4 and S5). We speculated that the over expression of miR156/157 sensor might compete with the natural targets that combine with miRNA, leading to attenuated miRNA effect on the natural target. To test our hypothesis, we constructed the miR156/157

target mimicry vector containing a *GbEXPA2* promoter (a fiber preferential promoter, unpublished) to negatively regulate miRNA activity (Figure S3). Because target mimicry was used as a genetic tool to suppress miRNA activity in *Arabidopsis* (Franco-Zorrilla *et al.,* 2007) and the *GbEXPA2* promoter could drive the specific expression of this gene at the fiber elongation stage, the miRNA156/157 activity could be suppressed at the fiber elongation stage. The accumulation of the Gb-miR156/157 target transcript (*SBP*, TC253516) in the transgenic lines was much higher than the one in the wild-type and null segregate control (12 DPA fiber), indicating that miRNA156/157 activity was suppressed (Figure 6B and S4). Moreover, suppressing miR156/157 activity inhibited fiber elongation, eventually leading to reduced mature fiber length compared with the null segregate control (Figure 6C and 6D) and wild-type (Table S5). To further investigate the effect of suppressing miR156/157 on fiber development, the fiber quality of miR156/157 suppressed-lines (MIM156/157), null segregate control and wild type plants were analyzed (Table 5). The upper half mean length of fiber from all the MIM156/157 lines was significantly lower than null segregate control in statistics. And there were little significant difference between null segregate control and wild type in all the tested parameters of fiber quality.

DISCUSSION

Identification and characterization of conserved miRNAs and candidate miRNAs in fiber of *G. barbadense*

We identified 47 conserved miRNA families and seven candidate miRNAs from seven fiber development libraries, covering the fiber development stages (Table 2 and S1). Consistent with the reference criteria (Meyers *et al.,* 2008), 34 conserved miRNA precursors and seven previously uncharacterized miRNA precursors were predicted. Although several miRNAs were identified in *G. hirsutumn* (Abdurakhmonov *et al.,* 2008; Kwak *et al.,* 2009; Pang *et al.,* 2009; Ruan *et al.,* 2009; Li *et al.,* 2012; Romanel *et al.,* 2012; Wang *et al.,* 2012; Yin *et al.,* 2012; Wei *et al.,* 2013; Xue *et al.,* 2013), our study mainly focused on the annotation of miRNAs in *G. barbadense* fibers, with the aim of broadening our understanding of the biological functions of miRNAs in cotton fiber development. Because of the lack of allotetraploid cotton (*G. barbadense*) genomic information, less than 14.47% of the total reads could be matched to the reference EST database (CGI 11.0) (Table 1), so that only 34 conserved miRNAs and seven candidate miRNAs precursors were identified (Figure S1). Considering the 562 candidate miRNA gene loci in the diploid cotton (*G. raimondii*) genome (Li *et al.,* 2012), we propose that many more miRNA genes might be identified after the *G. barbadense* genomic information becomes available.

After identifying these miRNAs, we attempted to further profile their expression in the seven fiber development libraries. Typically, -3 - 3 DPA represented the fiber initiation stage. At 7 - 12 DPA, fibers undergo fast elongation and at 20 - 25 DPA, the secondary wall of fiber cell begins to thicken. Cluster analysis showed differential accumulation of most miRNAs during the three fiber development stages (Figure 2A). Previous microarray analysis demonstrated that most miRNAs were accumulated at

lower levels in 10 DPA fibers than in -3 DPA ovules (Pang *et al.,* 2009). Some of the miRNAs in our results also showed similar expression patterns from the initiation to the elongation stages, while more miRNAs were accumulated at the elongation stage than at the initiation stage. It is possible that the expression pattern of the entire miRNA family is distinct from each individual miRNA. In addition, the sensitivity is different between the results of the microarray and sequencing analyses. The cluster analysis also covered late fiber development. More than a half of miRNAs from the cluster analysis were upregulated from the elongation stage to the SWT stage. Few miRNAs showed stable expression levels in all fiber development stages. The direct relationship between the dynamic expression of miRNAs and the morphogenesis of fibers at different development stages remains unclear and requires further investigation.

Identification of miRNA targets on large scale

Degradome sequencing identifies authentic miRNA targets in a high-throughput manner. This approach has been applied to identify miRNA targets in many plants (German *et al.,* 2008; Pantaleo *et al.,* 2010; Zhou *et al.,* 2010; Shamimuzzaman and Vodkin, 2012). For cotton fibers, we applied degradome sequencing to identify 140 conserved miRNA targets and 38 candidate miRNA targets (Table 4), which is much more than the targets verified through the RLM-RACE approach (Pang *et al.,* 2009; Wang *et al.,* 2012). However, there were still 17 conserved miRNA families and two candidate miRNAs without identified targets in our results. One possible reason is

that the expression of these miRNAs was too low to slice the targets or the abundance of their cleaved targets was too low to detect. Another explanation is that these miRNAs function in their targets primarily at the translational level (Voinnet, 2009). To further evaluate our degradome results, the seven targets identified through degradome sequencing were tested using RLM-RACE (Figure 3). At least four targets were confidently confirmed as authentic miRNA targets, and the cleavage sites of the remaining three targets was also identified through RLM-RACE, but not in the classic miRNA slicing position of the target. Thus, we propose that most of the miRNA targets identified in these results were reliable, and these targets, with a wide variety of functions, will be helpful for the further study of miRNA function in cotton fibers.

As shown in Figures 2 and S2, there are many cleavage sites outside the miRNA target site. Because degradome sequencing was aimed at sequencing all the uncapped transcripts, the cleavage signatures outside miRNA target site may represent other possible type of RNA turnover products (Li *et al.*, 2010). Given that many siRNAs could also guide target cleavage (Carthew and Sontheimer, 2009), it may imply that some unknown siRNAs may also possibly target the transcripts.

Seven miRNA/target comprise a possible regulatory network in fiber development

The analysis of the expression correlation showed that seven miRNAs may suppress their target expression temporally and spatially (Figure 4). The diversified functions of these targets prompted us to investigate the possible regulatory network of miRNA

in fiber development.

Previous studies have revealed that many transcription factors are preferentially expressed in cotton fibers (Samuel Yang *et al.,* 2006). In the present study, the expression of the *SBP* transcript (TC253516) was gradually reduced from the elongation stage to the SWT stage, which negatively correlated with Gb-miR156 expression (Figure 4A). This result indicates that Gb-miR156 may regulate fiber elongation and SWT. However, there was no negative correlation between Gb-miR156 and its target TC253516 at the fiber initiation stage (Figure 4A; Table S4). Thus, it is likely that miRNAs regulate different targets at different fiber development stages. Another interesting transcription factor, MYB, exhibits significant function at the fiber initiation and elongation stages (Pu *et al.,* 2008; Machado *et al.,* 2009; Walford *et al.,* 2011). In the present study, we identified a MYB transcription factor RNA (TC239324) as a Gb-miR399 target (Figure 3E). This transcription factor was expressed lower at the STW stage than at the fiber initiation and elongation stages, but Gb-miR399 abundance dramatically increased at the STW stage, suggesting that suppressing Gb-miR399, at the fiber initiation and elongation stages, might stimulate *MYB* expression to regulate fiber development (Figure 4F).

The expression of Gb-miR390 was much higher at the initiation stage than at the other fiber development stages, which might contribute to suppressing the accumulation of its target *RLK* (TC259635) at the fiber initiation stage (Figure 3C and 4E). In *Arabidopsis*, one RLK family member, CLV1, perceives the CLV3 signal to suppress cell differentiation in the shoot meristem (De Smet *et al.,* 2009). Thus,

Gb-miR390 might specifically suppress RLK expression to promote fiber cell differentiation at the initiation stage.

A miR167 target, *α-tubulin* (TC232723), was identified in the cotton fiber (Table S4). Both Northern blotting and qRT-PCR analyses showed that Gb-miR167 expression was obviously reduced after fiber initiation, thereby increasing the expression of its target (TC232723) at the fiber elongation and SWT stages (Figure 2B and 4D). α-Tubulin is the principal component of microtubules, which participate in plant cell morphology and guide the deposition of cellulose (Mathur and Hulskamp, 2002; Wasteneys, 2004). In cotton, cytoskeleton-related genes (profilins) were more up-regulated in domesticated cotton fibers than in their wild counterparts, reflecting enhanced cotton fiber quality in domesticated cotton (Bao *et al.,* 2011). So, the negative expression correlation between Gb-miR167 and α-tubulin suggests that Gb-miR167 may plays an important role in fiber development.

The cell wall also directly controls fiber morphogenesis and quality (Haigler *et al.,* 2012). Pectate lyase could eliminate de-esterified pectin to loosen the cell wall and knocking down pectate lyase gene expression in cotton suppresses fiber elongation (Wang *et al.,* 2010). Pectate lyase RNA was identified as a target of Gb-miR159 (Figure 3A), and its abundance was higher at the elongation stage than at the SWT stage. In contrast, Gb-miR159 expression was higher at the SWT stage than at the elongation stage (Figure 4B), indicating that Gb-miR159 might modulate the abundance of pectate lyase to promote fiber elongation. Another cell wall-related

gene *UGD* catalyzes UDP-glucuronic acid to UDP-xylose, which is the substrate for producing hemicelluloses. *UGD* was sliced by Gb-miR164, which was overall more abundant at the initiation stage than at the other fiber development stages (Figure 4C; Table S4). Accordingly, the target gene *UGD* (TC236955) was preferentially expressed at the fiber elongation and SWT stages when cell wall-related polysaccharides are largely produced (Figure 4C). Thus, Gb-miR164 might reduce *UGD* gene abundance to modulate cell wall synthesis.

nmiR3 and its target *CO1* (ES802293) are negatively correlated at the fiber initiation stage (Figure 3G and 4G). CO1 is a key oxidase in the respiratory chain, suggesting that nmiR3 might affect fiber cell development through reducing energy production.

We speculate that Gb-miR390/RLK and nmiR3/CO1 modules might be involved in signal transduction and energy supply respectively to affect fiber initiation; Gb-miR156/157/SBP and Gb-miR399/MYB modules may take part in transcription regulation to modulate fiber initiation and elongation; and GbmiR159/PL, 164/UGD and 167/α-tubulin modules might function in fiber elongation and second cell wall thickening by participating in cytoskeleton formation, cell wall synthesis and modification (Figure 7).

Breaking Gb-miR156/157 homeostasis affects fiber elongation

In *Arabidopsis*, miR156 could controls cell number and size through regulating the SBP domain protein (Usami *et al.*, 2009). In addition, miR156, with its target SBP

domain protein, regulates anthocyanin biosynthesis (Gou *et al.,* 2011), and the over accumulation of the anthocyanin/flavonoid metabolic product suppresses fiber elongation (Tan *et al.,* 2013). Previous studies have suggested a function for miR156/157 in fiber development. In the present study, we present evidence to show a role for Gb-miR156/157 in fiber development. First, miR156/157 could suppress the abundance of its target *SBP* in fiber development(TC253516, Figure 4). And the GUS histochemical assay showed that miR156/157 was activated in fiber and ovule but not germinating seed (Figure 5B). Second, suppressing miR156/157 activity at the fiber elongation stage reduced the mature fiber length (Figure 6B - 6D; Table 5 and S5). Although the effect of overexpression of miR156/157 has not been investigated, these results strongly support an important function of miR156/157 in fiber development. Previous studies have demonstrated a lower abundance of miR156/157 in the wild-type ovule than in the *fuzzless-lintless* mutant ovule, suggesting that miR156/157 suppresses fiber development at an early stage (Wang *et al.,* 2012). However, our results show that miR156/157 promotes fiber elongation. Thus, miR156/157 might affect fiber development in a subtle and complicated manner at different fiber development stages. Thus, when miR156/157 homeostasis is disrupted, the normal fiber development process would be disturbed, despite the increasing or decreasing abundance of miR156/157 expression.

In summary, we fully surveyed seven fiber development stages and identified 47 conserved miRNA families and seven candidate miRNAs in *G. barbadense*. The

dynamic expression of a set of characterized miRNAs was negatively correlated with their targets, indicating that the six conserved miRNA families and one candidate miRNA might be involved in the process of fiber development by suppressing the expression of transcription factors, receptor kinases, cytoskeleton proteins, cell wall-related enzymes and oxidases. Using transgenics and histochemical staining assays, we further verified that miR156/157 expression is activated in the ovules and fibers to modulate fiber elongation. Moreover, many targets identified through degradome sequencing could further facilitate functional studies on miRNA-mediated gene regulation in cotton fiber development.

EXPERIMENTAL PROCEDURES

Plant materials and RNA isolation

Gossypium barbadense cv. 3-79 plants were cultivated in an experimental field (Wuhan, Hubei, PR China) using normal farming practices. The bolls were tagged on the day of anthesis, and the stage of pre-anthesis flowers (three days before anthesis, -3 DPA) was estimated based on flower bud size and shape. The bolls and buds were harvested from different developmental stages (-3, 0, 3, 7, 12, 20 and 25 DPA) and stored on ice. The ovules and fibers were carefully excised, immediately immersed in liquid nitrogen and stored at -80℃. Total RNA was extracted from the collected tissues using a modified guideline thiocyanate method (Zhu , 2005).

Small RNA and degradome library construction

Before small RNA and degradome library construction, total RNA extracted from samples was tested through Agilent 2100 bioanalyzer system to guarantee RNA quality.

To construct the seven small RNA libraries, 18 to 30 nt in length of small RNA were isolated on a 15% polyacrylamide gel and ligated to the 5' and 3' RNA adaptors. Purified RNAs were reverse-transcribed to cDNA, followed by PCR amplification to generate the DNA pool. Seven DNA pools from different samples were sequenced on an Illumina Genome Analyzer at the Beijing Genomics Institute (BGI) in Shenzhen.

Three degradome libraries were constructed as previously described (German *et al.*, 2008). Briefly, a 5' RNA adapter was ligated to the cleavage products, which possess a free phosphate at the 5' end. The purified ligated products were reverse-transcribed to cDNA. After amplification, the PCR products were digested using the enzyme *Mme*I and ligated to the 3' adapter. The ligation products were amplified and sequenced on an Illumina Genome Analyzer.

Bioinformatic analysis of sequencing data

The raw reads from the small RNA libraries were first filtered to remove low-quality reads (reads in length < 18 nt, reads with contaminated 5' adaptor, reads with polyA, reads without 3' adaptor) and then trimmed adaptor sequence to get clean reads. The clean sequences were used to search GenBank and the Rfam database to annotate rRNA, tRNA, snRNA and snoRNA. After removing sequences belonging to rRNAs, tRNAs, snRNAs and snoRNAs, the remaining sequences were used to blast against

miRBase 18 (http://www.mirbase.org/) to identify conserved miRNA. Only the sequences that were less than two mismatches with known miRNAs in miRBase were considered as conserved miRNAs. In addition, potential miRNA precursors from ESTs were predicted by mireap (http://sourceforge.net/projects/mireap/) according to default parameters (base pairs of miRNA and miRNA* ≥ 16, number of bulge in miRNA and miRNA* duplex ⩽ 1, size of bulge ⩽ 2，length of precursor ⩽ 300 bp) and hairpin structure was visualized by an RNA hairpin folding and annotation tool (http://srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi). The sequences, which were not annotated to conserved miRNAs or other cellular RNAs, but were mapped to predicted hairpin structural precursors, were identified as candidate miRNAs. The ESTs were from cotton EST CGI 11 (ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/Gossypium).

Perl scripts developed by BGI were used to identify the miRNA targets from three degradome libraries according to German *et al.* (2008). The reference cotton ESTs (CGI 11) and most abundant miRNAs (Table S2) in each miRNA families were inputted to perform alignment. When the alignment score was no more than four, the transcripts were considered as miRNA targets.

Thirty conserved miRNA families and four candidate miRNAs, which were detected in all the libraries and whose normalized expression level were more than 10 TPM at least in one library, were selected to perform cluster analysis. Ovule (-3 DPA) library was considered as control and the miRNA gene expression of the fiber development patterns was calculated using Geneious software based on the

hierarchical clustering method (Sturn *et al.*, 2002).

miRNA Northern blotting analysis

Northern blotting of miRNA was performed as described previously (Pang *et al.*, 2009). Briefly, 20 μg total RNA was separated on a 15% polyacrylamide gel with 8 M urea and transferred on an Immobilon-Ny+ membrane (Merck Millipore, Billerica, MA, USA). The probes were labeled with γ^{32}P-ATP using T4 polynucleotide kinase (New England Biolabs, Ipswich, MA, USA). Hybridization was performed at 37 ℃ overnight in Hybridization Solution (TOYOBO, Osaka, Japan). The membrane was washed at 37 ℃ in a low stringency buffer (1 × SSC, 0.5% SDS) one time, followed in a high stringency buffer (0.2 × SSC, 0.2% SDS) one or more times and exposed using a phosphorimager. The probes are listed in Table S6.

RLM- RACE

The GeneRacer kit (Invitrogen, Carlsbad, CA, USA) was used to perform RLM-RACE. Total RNA (5 μg) from equal mixtures of 0 DPA ovule and fibers (5, 10, 15, and 20 DPA) were ligated to RNA adapter without calf intestinal phosphatase treatment. The cDNAs were transcribed using the GeneRacer Oligo dT primer. The PCR was performed with 5' adaptor primers and 3' gene specific primers according to the manufacturer's instructions.

qRT-RCR analysis

To quantify miRNAs and mRNAs, stem-loop RT-PCR was used and modified from a previous protocol (Varkonyi-Gasic *et al.,* 2007). Briefly, 3 μg total RNA was incubated with 0.4 mM dNTP, 0.05 μM stem-loop primers, 2.5 μM Oligo dT primer, 4 μl of 5×First-strand buffer, 1 μl of DTT (100 mM), 1 μl of RNase inhibitor, l μl of reverse transcriptase (Invitrogen, Carlsbad, CA, USA) in a 20-μl reaction mixture. The reverse-transcription reaction was performed at 16 ℃ for 30 min, followed by 60 cycles at 30 ℃ for 30 s, 42 ℃ for 30 s and 50 ℃ for 1 s. The reaction mixture was incubated at 85 ℃ for 5 min to inactivate the reverse transcriptase.

Real-time PCR was performed using a 7500 real-time system (Applied Biosystems, Foster City, CA, USA) using SsoFast EvaGreen Supermix With Low ROX (Bio-Rad, Hercules, CA, USA). Ubiquitin7 (*GhUBQ7*) was used as endogenous reference gene to reduce the biological and systematic variance. The relative expression levels (R.E.L.) were calculated using the $2^{-\Delta CT}$ method. All primers used are listed in Table S6.

Histochemical assay of GUS activity

Ovules (0 DPA) from transgenic and wild type plants were incubated in the GUS staining solution at 37 ℃ for 4 hours, followed by thorough washing in 75% ethanol according to Deng *et al.* (2012). Stained samples were photographed using a stereomicroscope (Leica Microsystems, Wetzlar, Germany). The staining solution contained 0.9 g/L 5-bromo-4-chloro-3-indolylglucuronide, 50 mM sodium phosphate buffer (pH 7.0), 20% (v/v) methanol and 100 mg/L chloromycetin.

Plasmid construction and genetic transformation

To detect miRNA biological activity, the ORF of the *GUS* gene was cloned into the pGWB402 vector to construct the control sensor plasmid (Nakagawa *et al.*, 2007). miR156/157, 166 and 172 sensor plasmids were constructed using a similar manner; however, the complementary sequences of miR156/157, 166 and 172 were added to the construct through PCR. To suppress miR156/157 function in cotton fibers, we constructed miR156/157 target mimicry vector (MIM156/157). First, the fragment of GbEXPA2 promoter was used to replace CaMV 35S promoter in pGWB402 vector, the vector was named pGWB402-PGbEXPA2 (unpublished). Subsequently, we cloned the 522-bp genomic sequence of *IPS1* in *Arabidopsis*, which includes the ORF and miR399 complementarity motif. The miR399 complementarity motif in *IPS1* was replaced with the miR156/157 complementarity motif through PCR according to the reference (Franco-Zorrilla *et al.*, 2007). The *IPS1* with miR156/157 complementarity motif was cloned downstream of the *GbEXPA2* promoter in the vector (Figure S3). The oligonucleotides for generating the plasmids described above are listed in Table S6.

Agrobacterium tumefaciens (EHA105) with the plasmid was used to transform hypocotyls of *G. hirsutum* cv. YZ1. Then the infected hypocotyls were induced to produce regenerative seedling according to a previously article (Jin *et al.*, 2006).

Fiber quality measurement

Experiment filed and green house were chosen to grow cotton materials in Huazhong Agricultural University, Wuhan, Hubei province. The fiber quality of miR156/157-sensor transformed plants was tested from three plots in experiment field, 2012 and one plot in green house, 2012. The fiber quality of miR156/157 suppressed-lines was tested from three plots in experiment field, 2013; from one plot in experiment field, 2012 and green house, 2013.

The mature fiber length was measured manually with a comb. In addition, mature fiber samples (at least 10 g each sample) were sent to the Center of Cotton Fiber Quality Inspection and Testing, Chinese Ministry of Agriculture (Anyang, Henan province, China) for detailed fiber quality measurements. All the samples for the measurements were collected from the bolls at the same positions on the plant and at the same time.

ACKNOWLEDGMENTS

We are grateful to Nakagawa Tsuyoshi (Shimane University) and Yang Li (Huazhong Agricultural University) for kindly providing plasmids. This work was financially supported by the national Natural Science Foundation of China (NO. 31230056) , National Basic Research Program (NO. 2010CB126001) and the Fundamental Research Funds for the Central Universities (NO. 2013YB06).

SUPPORTING INFORMATION

Figure S1. Hairpin structures of the conserved and candidate miRNA precursors.

Figure S2. Target plots (t-plots) of the validated target mRNAs at the fiber initiation, elongation and SWT stages.

Figure S3. Diagram of miR156/157 target mimicry vector which was used to suppress miRNA activity.

Figure S4. Southern blotting analysis of T1 plants.

Figure S5. Fiber length measurement of Sensor156/157 lines planted in Wuhan, Hubei province (2012).

Table S1. Abundance of conserved Gb-miRNAs and candidate miRNAs in seven fiber development libraries.

Table S2. The most abundant variant of conserved miRNAs from the seven fiber development libraries.

Table S3. Conservation of miRNAs in plants.

Table S4. miRNA targets were identified through degradome sequencing.

Table S5. Fiber quality analysis of miR156/157-suppressed lines (MIM156/157) and wild type planted in experiment field and green house.

Table S6. Primers applied in this study.

REFERENCES

Abdurakhmonov, I.Y., Devor, E.J., Buriev, Z.T., Huang, L., Makamov, A., Shermatov, S.E., Bozorov, T., Kushanov, F.N., Mavlonov, G.T., and Abdukarimov, A. (2008) Small RNA regulation of ovule development in the cotton plant, *G. hirsutum* L. *BMC Plant Biol.* **8**, 93.

Apel, K., Hirt, H. (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.* **55**, 373-399.

Bao, Y., Hu, G., Flagel, L.E., Salmon, A., Bezanilla, M., Paterson, A.H., Wang, Z., and Wendel, J.F. (2011) Parallel up-regulation of the profilin gene family following independent domestication of diploid and allopolyploid cotton (*Gossypium*). *Proc. Natl. Acad. Sci. USA* **108**, 21152-21157.

Carthew, R.W., Sontheimer, E.J. (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642-655.

Chiou, T.J., Aung, K., Lin, S.I., Wu, C.C., Chiang, S.F., and Su, C.L. (2006) Regulation of phosphate homeostasis by MicroRNA in *Arabidopsis*. *Plant Cell*, **18**, 412-421.

Cuperus, J.T., Fahlgren, N., and Carrington, J.C. (2011) Evolution and functional diversification of miRNA genes. *Plant Cell*, **23**, 431-442.

De Smet, I., Voss, U., Jurgens, G., and Beeckman, T. (2009) Receptor-like kinases shape the plant. *Nat. Cell Biol.* **11**, 1166-1173.

Deng, F., Tu, L., Tan, J., Li, Y., Nie, Y., and Zhang, X. (2012) GbPDF1 is involved in cotton fiber initiation via the core cis-element HDZIP2ATATHB2. *Plant Physiol.* **158**, 890-904.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A., and Paz-Ares, J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **39**, 1033-1037.

German, M.A., Pillay, M., Jeong, D.H., *et al.* (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat. Biotechnol.* **26**, 941-946.

Gou, J.Y., Felippes, F.F., Liu, C.J., Weigel, D., and Wang, J.W. (2011) Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell,* **23**, 1512-1522.

Gou, J.Y., Wang, L.J., Chen, S.P., Hu, W.L., and Chen, X.Y. (2007) Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis. *Cell Res.* **17**, 422-434.

Haigler, C.H., Betancur, L., Stiff, M.R., and Tuttle, J.R. (2012) Cotton fiber: a powerful single-cell model for cell wall and cellulose research. *Frontiers in plant science,* **3**, 104.

Hovav, R., Udall, J.A., Chaudhary, B., Rapp, R., Flagel, L., and Wendel, J.F. (2008) Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. *Proc. Natl. Acad. Sci. USA* **105**, 6191-6195.

Jin, S., Zhang, X., Nie, Y., Guo, X., Liang, S., and Zhu, H. (2006) Identification of a novel elite genotype for in vitro culture and genetic transformation of cotton. *Biologia Plantarum,* **50**, 519-524.

Kim, H.J., and Triplett, B.A. (2001) Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. *Plant Physiol.* **127**, 1361-1366.

Kwak, P.B., Wang, Q.Q., Chen, X.S., Qiu, C.X., and Yang, Z.M. (2009) Enrichment of a set of microRNAs during the cotton fiber development. *BMC Genomics,* **10**, 457.

Li, L., Wang, X.L., Huang, G.Q., and Li, X.B. (2007) Molecular characterization of cotton GhTUA9 gene specifically expressed in fibre and involved in cell elongation. *J. Exp. Bot.* **58**, 3227-3238.

Li, Q., Jin, X., and Zhu, Y.X. (2012) Identification and analyses of miRNA genes in allotetraploid *Gossypium hirsutum* fiber cells based on the sequenced diploid *G. raimondii* genome. *J. Genet. Genomics* **39**, 351-360.

Li Y.F., Zheng Y., Addo-Quaye C., Zhang L., Saini A., Jagadeeswaran G., Axtell M.J., Zhang W., Sunkar R. (2010) Transcriptome-wide identification of microRNA targets in rice. *Plant J.* **62,** 742-759.

Liang, G., Yang, F., and Yu, D. (2010) MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* **62**, 1046-1057.

Luo, M., Xiao, Y., Li, X., Lu, X., Deng, W., Li, D., Hou, L., Hu, M., Li, Y., and Pei, Y. (2007) GhDET2, a steroid 5alpha-reductase, plays an important role in cotton fiber cell initiation and elongation. *Plant J.* **51**, 419-430.

Machado, A., Wu, Y., Yang, Y., Llewellyn, D.J., and Dennis, E.S. (2009) The MYB transcription factor GhMYB25 regulates early fibre and trichome development. *Plant J.* **59**, 52-62.

Mathur, J., and Hulskamp, M. (2002) Microtubules and microfilaments in cell

morphogenesis in higher plants. *Curr. Biol.* **12**, R669-R676.

Meyers, B.C., Axtell, M.J., Bartel, B., *et al.* (2008) Criteria for annotation of plant microRNAs. *Plant Cell,* **20**, 3186-3190.

Nakagawa, T., Suzuki, T., Murata, S., *et al.* (2007) Improved Gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Biosci. Biotechnol. Biochem.* **71**, 2095-2100.

Pang, C.Y., Wang, H., Pang, Y., Xu, C., Jiao, Y., Qin, Y.M., Western, T.L., Yu, S.X., and Zhu, Y.X. (2010) Comparative proteomics indicates that biosynthesis of pectic precursors is important for cotton fiber and *Arabidopsis* root hair elongation. *Mol. Cell. Proteomics* **9**, 2019-2033.

Pang, M., Woodward, A.W., Agarwal, V., Guan, X., Ha, M., Ramachandran, V., Chen, X., Triplett, B.A., Stelly, D.M., and Chen, Z.J. (2009) Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber development in allotetraploid cotton (*Gossypium hirsutum* L.). *Genome Biol.* **10**, R122.

Pantaleo, V., Szittya, G., Moxon, S., Miozzi, L., Moulton, V., Dalmay, T., and Burgyan, J. (2010) Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J.* **62**, 960-976.

Pu, L., Li, Q., Fan, X., Yang, W., and Xue, Y. (2008) The R2R3 MYB transcription factor GhMYB109 is required for cotton fiber development. *Genetics*, **180**, 811-820.

Qin, Y.M., and Zhu, Y.X. (2011) How cotton fibers elongate: a tale of linear cell-growth mode. *Curr. Opin. Plant Biol.* **14**, 106-111.

Romanel, E., Silva, T.F., Correa, R.L., Farinelli, L., Hawkins, J.S., Schrago, C.E., and Vaslin, M.F. (2012) Global alteration of microRNAs and transposon-derived small RNAs in cotton (*Gossypium hirsutum*) during cotton leafroll dwarf polerovirus (CLRDV) infection. *Plant Mol. Biol.* **80**, 443-460.

Ruan, M., Zhao, Y., Meng, Z., Wang, X., and Yang, W. (2009) Conserved miRNA analysis in *Gossypium hirsutum* through small RNA sequencing. *Genomics*, **94**, 263-268

Samuel Yang, S., Cheung, F., Lee, J.J., *et al.* (2006) Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. *Plant J.* **47**, 761-775.

Shamimuzzaman, M., and Vodkin, L. (2012) Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. *BMC Genomics*, **13**, 310.

Shi, Y.H., Zhu, S.W., Mao, X.Z., Feng, J.X., Qin, Y.M., Zhang, L., Cheng, J., Wei, L.P., Wang, Z.Y., and Zhu, Y.X. (2006) Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. *Plant Cell*, **18**, 651-664.

Sturn, A., Quackenbush, J., and Trajanoski, Z. (2002) Genesis: cluster analysis of microarray data. *Bioinformatics*, **18**, 207-208.

Sunkar, R., Kapoor, A., and Zhu, J.K. (2006) Posttranscriptional induction of Two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell*, **18**, 2051-2065.

Tan, J., Tu, L., Deng, F., Hu, H., Nie, Y., and Zhang, X. (2013) A genetic and metabolic analysis revealed that cotton fiber cell development was retarded by flavonoid naringenin. *Plant Physiol.* **162**, 86-95.

Tang, W., Tu, L., Yang, X., Tan, J., Deng, F., Hao, J., Guo, K., Lindsey, K., Zhang, X. (2014) The calcium sensor GhCaM7 promotes cotton fiber elongation by modulating reactive oxygen species (ROS) production. *New Phytol.* **202**, 509-520.

Usami, T., Horiguchi, G., Yano, S., and Tsukaya, H. (2009) The more and smaller cells mutants of *Arabidopsis thaliana* identify novel roles for SQUAMOSA PROMOTER BINDING PROTEIN-LIKE genes in the control of heteroblasty. *Development*, **136**, 955-964.

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.

Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, **136**, 669-687.

Walford, S.A., Wu, Y., Llewellyn, D.J., and Dennis, E.S. (2011) GhMYB25-like: a key factor in early cotton fibre development. *Plant J.* **65**, 785-797.

Walford, S.A., Wu, Y., Llewellyn, D.J., and Dennis, E.S. (2012) Epidermal cell differentiation in cotton mediated by the homeodomain leucine zipper gene, GhHD-1. *Plant J.* **71**, 464-478.

Wang, H., Guo, Y., Lv, F., Zhu, H., Wu, S., Jiang, Y., Li, F., Zhou, B., Guo, W., and Zhang, T. (2010) The essential role of GhPEL gene, encoding a pectate lyase, in cell wall loosening by depolymerization of the de-esterified pectin during fiber elongation in cotton. *Plant Mol. Biol.* **72**, 397-406.

Wang, J.W., Wang, L.J., Mao, Y.B., Cai, W.J., Xue, H.W., and Chen, X.Y. (2005) Control of root cap formation by microRNA-targeted auxin response factors in *Arabidopsis*. *Plant Cell*, **17**, 2204-2216.

Wang, Z.M., Xue, W., Dong, C.J., Jin, L.G., Bian, S.M., Wang, C., Wu, X.Y., and Liu, J.Y. (2012) A comparative miRNAome analysis reveals seven fiber initiation-related and 36 novel miRNAs in developing cotton ovules. *Mol. Plant* **5**, 889-900.

Wasteneys, G.O. (2004) Progress in understanding the role of microtubules in plant cells. *Curr. Opin. Plant Biol.* **7**, 651-660.

Wei, M., Wei, H., Wu, M., Song, M., Zhang, J., Yu, J., Fan, S., and Yu, S. (2013) Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton. *BMC Plant Biol.* **13**, 66.

Wu, G., Park, M.Y., Conway, S.R., Wang, J.W., Weigel, D., and Poethig, R.S. (2009) The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell*, **138**, 750-759.

Xue, W., Wang, Z., Du, M., Liu, Y., and Liu, J.Y. (2013) Genome-wide analysis of small RNAs reveals eight fiber elongation-related and 257 novel microRNAs in elongating cotton fiber cells. *BMC Genomics*, **14**, 629.

Yin, Z., Li, Y., Han, X., and Shen, F. (2012) Genome-wide profiling of miRNAs and other small non-coding RNAs in the *Verticillium dahliae*–inoculated cotton roots. *PLoS ONE*, **7**.

Zhang, B.H., Pan, X.P., Cox, S.B., Cobb, G.P., and Anderson, T.A. (2006) Evidence that miRNAs are different from other RNAs. *Cell. Mol. Life Sci.* **63**, 246-254.

Zhang, M., Zheng, X., Song, S., *et al.* (2011) Spatiotemporal manipulation of auxin biosynthesis in cotton ovule epidermal cells enhances fiber yield and quality. *Nat. Biotechnol.* **29**, 453-458.

Zhao, L., Kim, Y., Dinh, T.T., and Chen, X. (2007) miR172 regulates stem cell fate and defines the inner boundary of APETALA3 and PISTILLATA expression domain in *Arabidopsis* floral meristems. *Plant J.* **51**, 840-849.

Zhou, M., Gu, L., Li, P., Song, X., Wei, L., Chen, Z., and Cao, X. (2010) Degradome sequencing reveals endogenous small RNA targets in rice (*Oryza sativa* L. *ssp. indica*). *Front. Biol.* **5**, 67-90.

Zhu L., Tu L., Zeng F., Liu D., Zhang X. (2005) An improved simple protocol for isolation of high quality RNA from *Gossypium spp.* suitable for cDNA library construction. *Acta Agronomica Sinica*, **31**, 1657-1659.

TABLES

Table 1 Distribution of small RNA sequences in the seven fiber development libraries

Type of small RNA	-3 DPA Ovule	0 DPA Ovule	3 DPA Ovule	7 DPA Fiber	12 DPA Fiber	20 DPA Fiber	25 DPA Fiber
Matching CGI11.0	2382069	2576706	3131271	2686708	3362751	3362751	4262742
Known miRNA	2093448	1403450	1736036	776489	1122378	1907607	2090528
rRNA	517336	722919	978571	773837	761398	761398	892907
snRNA	2229	2371	2601	3521	6181	6181	10698
snoRNA	1364	1465	1625	1257	1181	1181	1321
tRNA	213559	225160	220408	783902	1380408	1380408	1755932
Total known small RNAs	2829350	2356119	2940241	2347536	4062026	4062026	4754846
Total reads	16467266	17272495	17131384	20132788	18323896	18323896	18435354

Table 2 Candidate cotton miRNAs identified through small RNA sequencing

Name	miRNA	miRNA*[a]	miRNA loci	Position	Orientation	MFEI
nmiR1	UUCAGAAACCAUCCCUUCCUU	GGAAGGAAUGGUUUCUGAAGC	CO070343	231-358	antisense	1.20
nmiR2	ACAGCUUUAGAAAUCAUCCCU	GGAUGAUUUCUAAAGCUCUAG	CO070343	423-531	antisense	1.64
nmiR3	UCGGACUGGAUUUGUUGACAA	N	CO117073	134-239	sense	1.10
nmiR4	UUACUUUAGAUGUCUCCUUCA	AGGGAAACAUCUAAAGUAAAC	ES816423	113-243	sense	0.94
nmiR5	AAGAGUCAGAUUGCAUUUUG	N	TC264760	431-528	sense	0.91
nmiR6	CAUGACUUUUAGCGGCGUUUG	AGCGUCGCUAAAGGUCAUGAU	TC268435	1143-1209	sense	1.31
nmiR7	UGAAUAUUGUUAAAGUAGAAA	UCUACUUUAACAAUAUUCAUA	BG446822	647-770	sense	1.15

[a]N indicated that sequence was not identified. MFEI is the abbreviation of minimal folding free energy index.

Table 3 Summary data of degradome sequencing from the three libraries

	Initiation stage	Elongation stage	SWT stage
Sequenced (Unique read)	8089323	3751257	2547751
Matched reference database	5078250 (62.78%)	2532917 (67.52%)	1718861 (67.46%)
Represented transcript	95021	87495	81718

Table 4 Overview of miRNA targets in the three degradome libraries

miRNA family	Anotation	Count[a]
Gb-miR156/157	SBP transcription factor	5
	Predicted protein	1
Gb-miR159	Pectate lyase	1
	MYB transcription factor	1
	Predicted protein	1
	No annotation	1
Gb-miR160	Auxin response factor	4
	Ca^{2+}dependent membrane binding protein annexin	1
Gb-miR164	NAC transcription factor	1
	UDP-glucuronic acid decarboxylase	4
Gb-miR165/166	Class III HD-zip transcription factor	6
Gb-miR167	Auxin response factor	11
	α-Tubulin	12
	Proline rich family protein	1
	Predicted protein	2
Gb-miR168	Argonaute protein	1
Gb-miR169	Nuclear transcription factor Y	5
	Predicted protein	1
Gb-miR171	GRAS transcription factor	5
Gb-miR172	AP2 transcription factor	7
	Filament like plant protein	1
	ρ-GTPase activation protein	1
	Transposon protein	1
Gb-miR2911	Cytidine deaminase	4
Gb-miR2950	RCC1 and BTB domain containing protein	1
	No annotation	1
Gb-miR3476	Cytidine deaminase	2
Gb-miR390	TAS3 like	4
	Leucine rich receptor like protein kinase	3
	U5 small nuclear RNA helicase	1
Gb-miR393	Auxin signaling F-box protein	8
Gb-miR394	F-box family protein	2
	Predicted protein	1
Gb-miR395	Reticulon like protein	3
Gb-miR396	26s Proteasome regulatory particle triple-a ATPase	1
Gb-miR397	Laccase	2
Gb-miR398	Superoxide dismutase	2
Gb-miR399	MYB transcription factor	1
Gb-miR403	Calmodulin binding protein	3
	Argonaute protein	1
Gb-miR4414	Membrane bound o-acyltransferase family protein	1
Gb-miR482	Ring/U-box domain containing protein	1

Gb-miR5059	Oligouridylate binding protein	1
	Aluminium induced protein	1
	Pyruvate dehydrogenase β-subunit isoform1	4
Gb-miR5077	Zinc finger family protein	1
	60s Ribosomal protein	2
	Predicted protein	1
Gb-miR530	Zinc finger protein	3
	Translation factor	5
	40s Ribosomal protein	2
Gb-miR535	Transmembrane clptm1 family protein	2
Gb-miR827	60s Ribosomal protein	1
Gb-miR828	MYB like transcription factor 2	2
	Total conserved miRNA targets	140
nmiR1	Adenylate translocator	7
	Pentatricopeptide repeat containing protein	2
	NADH dehydrogenase subunit 9	2
	α-Tubulin	1
	No annotation	1
nmiR2	Pentatricopeptide repeat containing protein	6
nmiR3	Cytochrome c oxidase subunit 1	6
	Ribosomal protein S10	3
nmiR5	Cytoplasmic like protein	2
	Pectin methylesterase inhibitor protein	1
	Stress enhanced protein 1	1
	Temperature induced lipocalin	1
	F-box kelch repeat protein	2
	Cellulose synthase interactive protein 1	1
	No annotation	2
nmiR7	Predicted protein	1
	No annotation	1
	Total candidate miRNA targets	38

[a]Amount of transcripts which were identified as miRNA targets.

Table 5 Comparison of fiber quality parameters between Wild type, Null and miR156/157-suppressed Lines (MIMI156/157) in Hubei province, 2013

Line	Upper half mean length	Uniformity index%	Micronaire	Elongation	Strength (CN/tex)
Wild type	28.11 ± 0.13 a	83.80 ± 0.28 a	4.77 ± 0.64 a	6.63 ± 0.05 a	24.98 ± 0.48 ab
Null	27.58 ± 0.25 a	84.73 ± 0.50 a	4.43 ± 0.60 a	6.70 ± 0.08 a	25.13 ± 1.32 a
MIM156/157-1	25.33 ± 0.38 b	83.31 ± 0.71 a	5.35 ± 0.62 a	6.80 ± 0.19 a	23.99 ± 1.01 ab
MIM156/157-2	25.90 ± 0.68 b	83.30 ± 0.35 a	4.79 ± 0.33 a	6.70 ± 0.12 a	24.10 ± 0.60 ab
MIM156/157-3	25.93 ± 0.79 b	83.08 ± 1.00 a	5.00 ± 0.91 a	6.88 ± 0.16 a	22.28 ± 0.89 b

Null, segregated nontransgenic plants derived from the three transformants MIM156/157-1, 2 and 3. Values are mean ± SD for samples of the Wild type, Null and trangenetic lines in experiment field. In each column, values that are not followed by the same letters are significantly different based on the Tukey's Multiple Comparison Test ($P < 0.05$).

FIGURE LEGENDS

Figure 1. Length distribution of the small RNAs in the seven fiber development libraries.

Figure 2. Expression pattern of conserved miRNA families and candidate miRNAs. A, Hierarchical cluster analysis of differential expression of 30 conserved miRNA families and 4 candidate miRNAs at the seven stages of fiber development. B, Northern blotting analyses of four conserved miRNAs at the seven fiber development stages. 5S RNA was used as loading control. O(-3), O(0) and O(3), ovules harvested at -3, 0, 3 DPA, respectively. F(7), F(12), F(20) and F(25), fibers harvested at 7, 12, 20, 25 DPA, respectively.

Figure 3. The miRNA targets were identified through degradome sequencing and RLM-RACE. Target plots showed signature abundance in the position of target transcripts identified through degradome sequencing. The red dots indicate significant signatures and the red arrows in the target plots show the signatures corresponding to the miRNA cleavage sites. Signature abundance along the mRNA was normalized to the transcripts per 10 million (TP10M) clean tags. The target cleavage sites identified through RLM-RACE, as shown below the target plot. The numbers indicate the cleavage frequency. The black and red arrows in the mRNAs represent the cleavage sites identified through RLM-RACE and degradome sequencing, respectively. Wobble G-U pairs are indicated with circles and no base pairing is indicated with '×'.

Figure 4. Expression correlation between miRNAs and targets. The bars and lines indicate miRNAs and accordingly the target abundance from the qRT-PCR results, respectively, in the seven fiber development libraries. The y-axis on the left and right were used to measure expression level of miRNA and target, respectively. O(-3), O(0) and O(3) indicate ovules harvested at -3, 0, 3 DPA, respectively. F(7), F(12), F(20) and F(25) indicate fibers harvested at 7, 12, 20, 25 DPA, respectively. R.E.L. (relative expression levels) was calculated using *GhUBQ7* as a control. The error bars indicate standard deviation of two biological and three technical replicates.

Figure 5. Detection of the biological activity of the miRNAs in the ovule and fiber. A, The miRNA sensor construct used to detect miRNA biological activity. B, GUS staining images of miRNA sensor construct. Control, the construct without miRNA reverse complement site. Sensor156/157, the construct with Gb-miR156/157 reverse complement site. Sensor166, the construct with Gb-miR166 reverse complement site. Sensor172, the construct with Gb-miR172 reverse complement site. GUS staining image of the non-transforming miRNA sensor construct as a negative control. O (0) and O (25), ovule harvested at 0 and 25 DPA, respectively.

Figure 6. Reduced Gb-miR156/157 expression suppresses fiber elongation. A, The mature fiber length of the samples planted in green house (2012) was measured manually with a comb. Control, transgenic plants containing the construct without the

miRNA reverse complement site. Sensor156/157, transgenic plants of Gb-miR156/157 reverse complement site. B, Relative expression of the Gb-miR156/157 target (TC253516) in fibers (12 DPA). R.E.L., the relative expression levels were calculated using GhUBQ7 as a control. C, The mature fiber length of the samples planted in the experiment field (2013) was measured manually with a comb. Null, segregated nontransgenic plants derived from the three transformants MIM156/157-1, 2 and 3. MIM156/157, plants transformed with the miR156/157 mimicry vector. D, The image of mature fiber. Error bars (A and C) represent standard deviation of samples from at least 20 ovules. Different letters in A and C indicate statistically significant differences at $P < 0.05$ based on ANOVA (Tukey's Multiple Comparison Test). The error bars (B) indicate the standard deviation of three biological replicates.

Figure 7. Summary of the potential miRNA regulatory network in cotton fibers. The round rectangles indicate the three fiber development stages, which are distinct but overlapping to some degree. The T shapes represent negative regulation. The dashed lines indicate hypothetical pathways affecting fiber development.