Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Gene 397 (2007) 26–37
GENE
www.elsevier.com/locate/gene

Identification of cotton microRNAs and their targets

Baohong Zhang [a,b,c,d,*], Qinglian Wang [a], Kunbo Wang [d], Xiaoping Pan [b], Fang Liu [d], Tenglong Guo [a], George P. Cobb [b], Todd A. Anderson [b]

[a] Henan Institute of Science and Technology, Xinxiang, Henna 453003, PR China
[b] The Institute of Environmental and Human Health, and Department of Environmental Toxicology, Texas Tech University, Lubbock, TX 79409-1163, USA
[c] Department of Biology, East Carolina University, Greenville, NC 27858, USA
[d] Department of Molecular Biology, Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang, Henan 455112, PR China

Received 18 February 2007; received in revised form 21 March 2007; accepted 22 March 2007
Available online 20 April 2007
Received by J.A. Engler

Abstract

No study has been performed on identifying microRNAs (miRNAs) and their targets in cotton although cotton is one of the most important fiber and economic crops around the world. In this study, we found 30 potential cotton miRNAs using a comparative genomic approach based on genomic survey sequence analysis and miRNA secondary structure. These cotton miRNAs belong to 22 miRNA families. Expressed sequence tag (EST) analysis indicated that the predicted miRNAs were expressed in cotton plants. Based on the characteristic that miRNAs exhibit perfect or nearly perfect complementarity with their targeted mRNA sequences, a total of 139 potential miRNA targets were identified in cotton genome. A majority of these targets belong to transcriptional factors which regulate cotton growth and development, including leaf, root, stem, flower, and even fiber development. Those miRNAs may also be involved in other cellular and metabolic processes, such as stress response, signal transduction, and secondary wall synthesis and deposition. Some of the newly identified miRNA targets may be unique to cotton species. In this study, we found that at least 3 miRNA families (miR 396, 414, and 782) target callous synthase, fiber protein Fb23, and fiber quinone-oxidoreductase, suggesting that miRNAs play an important role in cotton fiber differentiation and development.

Keywords: Cotton; microRNA; Homolog; Comparative genomics; Genomic survey sequence; Expressed sequence tag; Posttranscriptional gene regulation

1. Introduction

Cotton is one of the most important economic and fiber crops. There are more than 50 species in the *Gossypium* genus (Fryxell et al., 1992; Wendel and Cronn, 2003). Of them, four species (*Gossypium hirsutum* L., *G. barbadense*, *G. arboreum*, *G. herbaceum*) are widely planted in 70 developed and developing countries, including the U.S., China, and India. It is estimated that more than 180 million people are associated with the cotton fiber industry that annually produces 20 to 30 billion dollars worth of raw cotton (IAC, 1996; Zhang and Feng, 2000). Trade in cotton-related products represents almost 50% of the total $115 billion trade in textiles and the $133 billion trade in clothing (Heijbroek and Husken, 1996). Numerous scientists in many countries have worked on cotton quality and yield (Zhang et al., 2000), and scientists also employ cotton fiber as a unique system to study secondary wall formation and cellulose synthesis (Saxena and Brown, 2005; Jacob-Wilk et al., 2006; Nakashima et al., 2006). Although great progress has been made, several questions remain, including the mechanisms of fiber differentiation and development and cellulous synthesis and deposition. Recently discovered microRNAs (miRNAs)

Abbreviations: 5′RACE: 5′ rapid amplification of cDNA end; Ap2: apetal-like protein; CBF: CCAAT-binding transcription factor; EST: expressed sequence tag; GSS: genomic survey sequence; MFE: minimal folding free energy; MFEI: minimal folding free energy index; miRNA: microRNA; PCR: Polymerase chain reaction; Pol II: RNA polymerase II enzyme; Pol III: RNA polymerase III enzyme; PR: pathogen-related protein; pre-miRNA: microRNA precursor; pri-miRNA: primary microRNA; RISC: RNA-induced silencing complex; SBP: SQUAMOSA-promoter binding protein-like protein; TIGR: Institute for Genome Research.

* Corresponding author. Department of Biology, East Carolina University, Greenville, NC 27858, USA.
E-mail address: zhangb@ecu.edu (B. Zhang).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.03.020

may play important roles in cotton development, especially in cotton fiber differentiation and development.

MicroRNAs are a class of small regulatory RNAs, which negatively regulate gene expression at the posttranscriptional levels by binding target mRNAs for mRNA cleavage or inhibition of mRNA translation (Ambros, 2001; Carrington and Ambros, 2003; Bartel, 2004). Lots of experiments have demonstrated that miRNAs play an important role in multiple biological and metabolic processes in plants and animals (Carrington and Ambros, 2003; Zhang et al., 2007b). In plants, identified miRNA functions include control of tissue (leaf, root, stem, and flower) differentiation and development, phase switch from vegetative growth to reproductive growth, signal transduction, and response to different biotic and abiotic stress (eg. salinity, drought, and pathogens) (Chen, 2005; Zhang et al., 2006d).

In 1993, the first miRNA *lin-4* was accidentally found in *C. elegans* (Lee et al., 1993). However, its functions were not recognized until the early 2000s when an abundant number of additional miRNAs was discovered in worm, fly, and human (Reinhart et al., 2000; Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). As more and more miRNAs were identified in various animals, miRNAs were also discovered in plants (Reinhart et al., 2002). Currently, several hundred miRNAs have been identified in plants by computational and experimental approaches (Griffiths-Jones et al., 2006; Zhang et al., 2006c). Of them, 184 from *Arabidopsis thaliana*, 242 from rice, 215 from *Populus trichocarpa* (Griffiths-Jones et al., 2006), and 118 from maize (Zhang et al., 2006a). However, no report of miRNAs in cotton has been published.

Mature miRNAs are a class of small RNAs with 20–22 nucleotide lengths. However, the genes encoding miRNAs are much longer than their mature sequences, and ranged from several tens to more than 1000 nt. miRNA genes are first transcribed to the capped and polyadenylated primary miRNAs (pri-miRNAs) by Pol II enzyme or by Pol III enzyme (Kurihara and Watanabe, 2004; Lee et al., 2004; Borchert et al., 2006). Then a pri-miRNA is cleaved to a stem loop intermediate called miRNA precursor (pre-miRNA) (Lee et al., 2002; Tang et al., 2003; Kurihara and Watanabe, 2004) with a high minimal folding free energy index (MFEI) (Zhang et al., 2006e). Pre-miRNAs are further cleaved into miRNA:miRNA* (miRNA* is the complementary sequence to miRNA at the opposite arm site) duplex with 2 nt 3′ overhangs (Papp et al., 2003), and miRNA: miRNA* duplex was methylated by HEN1 enzyme (Yu et al., 2005). Finally a mature miRNA enters a ribonucleoprotein complex known as the RNA-induced silencing complex (RISC) and negatively regulates gene expression by targeting mRNA cleavage or inhibiting mRNA translation (Bartel, 2004).

Although the first miRNAs (*lin-4* and *let-7*) were directly cloned by genetic or biochemical approaches (Lee et al., 1993), a majority of identified miRNAs were first predicted by computational approaches and then validated by molecular techniques such as northern blotting (Zhang et al., 2006f). Thus, computational approaches have played important roles in miRNA identification. Several web-based or stand-alone computational programs were developed to predict miRNAs in plants and animals (Zhang et al., 2006f). However, a majority of these programs were based on decoded genome sequences of a few model species. It is difficult to predict miRNAs in species with unsequenced genomes. Evidence indicates that many miRNAs are evolutionarily conserved from species to species within the same kingdom, some from worms to humans in animals (Pasquinelli et al., 2000; Pasquinelli et al., 2003; Altuvia et al., 2005), and from moss to flowering eudicots in plants (Floyd and Bowman, 2004; Zhang et al., 2006c). This suggests a powerful strategy to predict potential miRNAs by using comparative genomics. Weber (2005) found 35 new human and 45 new mouse miRNA candidates by using a homology search (Weber, 2005). Recently, we developed a successful strategy to predict new miRNA candidates by mining out undiscovered miRNAs from the repository of GSSs and ESTs currently available (Zhang et al., 2005). Using this strategy, more than 700 miRNAs were identified in 71 plant species and viruses (Zhang et al., 2006c; Pan et al., 2007), including 118 in maize (Zhang et al., 2006a).

Currently, about 50,000 cotton GSS sequences are deposited in the NCBI GenBank databases. In this study, we compared all of these sequences and the 131 previously known *A. thaliana* miRNAs for identifying potential miRNAs in cotton. Based on these new predicted cotton miRNAs, we also predicted the potential miRNA targets in cotton, especially those related to cotton fiber differentiation, initiation, and development.

2. Methods and materials

2.1. miRNA reference set

To search potential cotton miRNAs, a total of 131 previously known *A. thaliana* miRNAs were defined as a reference set of miRNA sequences. The reason for using *A. thaliana* miRNAs as reference miRNAs is that *A. thaliana* is the most closely related species to cotton in which a large number of miRNAs have been identified and deposited in publicly available databases. The 131 *A. thaliana* mature miRNAs and their precursor sequences were downloaded from the miRNA database (miRBase Sequence Database, http://microrna.sanger.ac.uk; release 9.0, October 2006) (Griffiths-Jones et al., 2006). Although some of these *A. thaliana* miRNAs were initially identified by computational approaches, a majority of them have been validated by experimental approaches including direct cloning, PCR, Northern blotting, and/or 5′ rapid amplification of cDNA end (5′ RACE) (Griffiths-Jones et al., 2006).

2.2. Cotton genomic survey sequences (GSSs), mRNA, cDNA, and ESTs

Cotton genomic survey sequences, mRNA, cDNA, and EST sequences were obtained from the GenBank nucleotide databases from NCBI and cotton nucleotide databases from the Institute for Genome Research (TIGR) at http://www.tigr.org.

2.3. Identifying potential cotton miRNAs using GSS analysis

In our previous work, we developed an efficient strategy for identifying plant and virus miRNAs using GSS analysis as well as EST analysis (Zhang et al., 2005). Using this approach, more

than 700 miRNAs have been identified in plants and viruses (Zhang et al., 2006c; Pan et al., 2007). The same procedure was employed to identify cotton miRNAs. Fig. 1 summarizes the procedure for searching conserved cotton miRNA homologues of previously known *A. thaliana* miRNAs. Briefly, the mature sequences of 131 previously known *A. thaliana* miRNAs were subjected to a BLAST search in the subgroup of Viridiplantae of the publicly available GSS databases using BLASTn 2.2.9 (1 May 2004) (Altschul et al., 1997). Adjusted blast parameter settings were as follows: expect values were set at 1000; the default word-match size between query and database sequences was set at 7; the numbers of descriptions and alignments were raised to 1000. If only a partial previously known *A. thaliana* mature miRNA sequence was aligned to a subjected sequence, the non-aligned parts were manually inspected and compared to determine the number of matching nucleotides. All BLAST results were saved. However, only GSS sequences, which closely matched (no more than 4 mismatches, which also include insertion and/or deletion nucleotides) the previously known *A. thaliana* mature miRNAs, were manually chosen for further consideration. After the conserved sequences were chosen, the whole GSS sequences (containing the conserved miRNA sequences) were selected to predict the secondary structures and identify the miRNA precursor sequences. Although each *A. thaliana* miRNA precursor sequence was also subjected to a BLAST search against the GSS databases of NBCI GenBank, mature miRNA sequences were the focus of the BLAST search due to the fact that only mature miRNAs, rather than miRNA precursors, are conserved in plants (Zhang et al., 2006c).

All selected GSS sequences with no more 4 mismatched and/or deletion/insertion nucleotides with the previously known *A. thaliana* miRNAs were further compared with each other and the mRNA sequences in the NBCI databases to remove protein-coding sequences and repeated sequences. Then, the secondary structures of the remaining sequences were generated using the Zuker folding algorithm with web-based computational software MFOLD 3.1 (Mathews et al., 1999; Zuker, 2003), which was publicly available at http://www.bioinfo.rpi.edu/applications/mfold/rna/form1.cgi. The software default parameters were used in predicting secondary structure of the selected sequences. All MFOLD outputs including free energy (ΔG kcal/mol), the number of nucleotides (A, G, C and U), location of the matched region, and the number of arms per structure were recorded. The minimal folding free energy index (MFEI) for each sequence was calculated as previously reported (Zhang et al., 2006e). In previous studies, we found that miRNA precursor sequences have significantly higher MFEI than other non-coding or coding RNAs, and the RNA sequences are more likely to be miRNAs when the MFEI is greater than 0.85 (Zhang et al., 2006e). To avoid designating other RNAs as miRNA candidates, MFEI was also considered when predicting secondary structures.

In this study, a RNA sequence was considered a miRNA candidate only if it fit all of the following criteria: (1) predicted mature miRNAs had no more than four nucleotide substitutions compared with *A. thaliana* mature miRNAs; (2) a RNA sequence can fold into an appropriate stem-loop hairpin secondary structure; (3) the mature miRNA sat in one arm of the hairpin structure; (4) no more than 6 mismatches between the predicted mature miRNA sequence and its opposite miRNA* sequence in the secondary structure; (5) no loop or break in the miRNA or miRNA* sequences; (6) predicted secondary structure had higher MFEI and negative MFE. After considering these six criteria, the total number of predicted miRNAs was significantly reduced.

2.4. Potential miRNA expression data

Expressed sequence tag (EST) databases have been employed to check the potential expression of miRNA

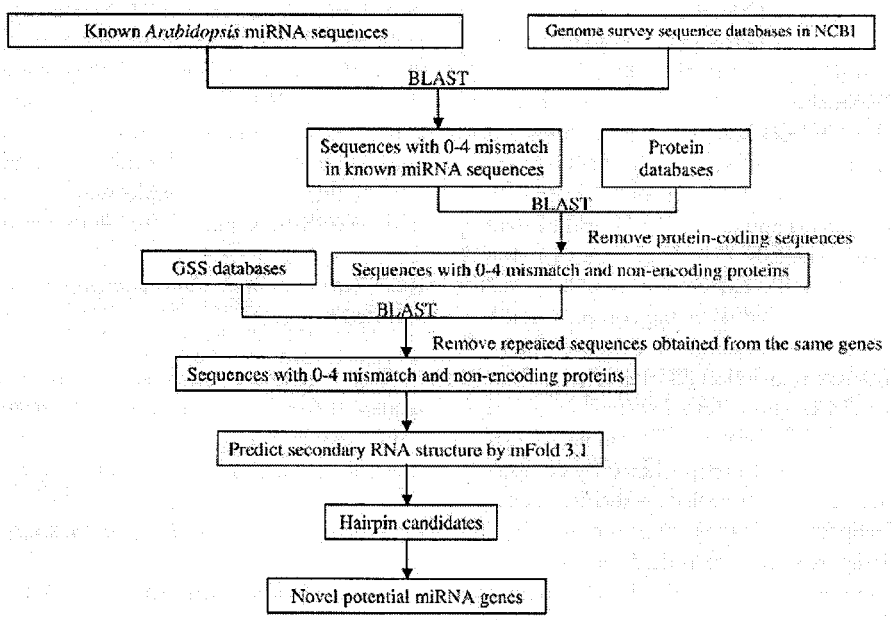

Fig. 1. Schematic representation of the miRNA gene search procedure used to identify cotton homology to known *Arabidopsis thaliana* miRNAs.

Table 1
Cotton miRNAs identified by homolog search and secondary structure

miR	Species	Sequence	ML	A Number		PL		PS	PE	MS	ME	MFE	A	C	G	U	(G + C)%	MFEI
156a	*G. hirsutum*	ugacagaagagagugagcac	20	DX383155, DX383726	–	89	5′	210	298	279	298	45.6	19	23	21	26	49.44	1.04
156b	*G. hirsutum*	guaaugaagagagugagcaa	20	DX524817	–	521	3′	153	673	153	172	131.5	158	90	125	148	41.27	0.61
156c	*G. hirsutum*	uugcagaagaaagugagcuu	21	DX535560	+	53	3′	73	125	105	125	9.6	21	7	8	17	28.30	0.64
162	*G. hirsutum*	gaaguaaaccucugcaucucu	21	DX558318	–	63	5′	23	85	62	82	16.1	22	12	12	17	38.10	0.67
167	*G. hirsutum*	uguaaaugccagcaugaucua	21	DX530709	–	358	3′	83	103	82	439	58.1	113	76	46	123	34.08	0.48
168	*G. hirsutum*	ugcuuggugcaggucgugg	20	DX557005	+	50	3′	1042	1061	1014	1063	20.7	4	11	21	14	64.00	0.65
169a	*G. herbaceum*	uagccaaggaugacuugccug	21	DX401397	+	89	5′	305	325	304	395	41.2	23	24	20	22	49.44	0.94
169b	*G. herbaceum*	uagccaagaaugacuugccug	21	DX401397	+	69	5′	552	572	552	620	37.4	17	17	18	17	50.72	1.07
171	*G. herbaceum*	cugagccgugccaaaaucaau	21	DX398301	+	48	5′	277	297	277	324	17.5	13	11	12	12	47.92	0.76
172a	*G. hirsutum*	ugaaucuugaugaugccaaau	21	DX549579	–	371	5′	307	677	656	676	84.9	156	55	74	86	34.77	0.66
172b	*G. hirsutum*	aucgcaccauuaagauuacu	20	DX544622	–	60	3′	482	541	482	501	10.4	30	8	7	15	25.00	0.69
172c	*G. hirsutum*	ccagcaccauuaagaaucag	20	DX537346	+	573	3′	61	633	614	633	114.4	210	94	68	201	28.27	0.71
390	*G. raimondii*	agucucaggagggauagcuuc	21	DX405484	–	283	3′	372	654	373	393	77.6	73	65	67	78	46.64	0.59
391	*G. hirsutum*	uggcaggagagauagcauag	20	DX538744	+	110	5′	280	389	280	299	23.8	32	26	20	32	41.82	0.52
393	*G. herbaceum*	agaaaagggaucgcauucaau	21	DX400113	–	353	3′	32	384	33	53	63.7	107	70	60	116	36.83	0.49
395	*G. hirsutum*	cugaaguguuuggggaacuc	21	DX543072	–	107	3′	234	340	234	254	51.9	19	25	30	33	51.40	0.94
396a	*G. hirsutum*	ugugauagcuuucuugaacuu	21	DX562190	–	90	3′	580	669	581	601	14.9	23	19	17	31	40.00	0.41
396b	*G. hirsutum*	augcaaagcuuucuugaacuu	21	DX383236	+	99	3′	212	310	290	310	19.9	32	22	13	32	35.35	0.57
400	*G. hirsutum*	uaugagaguauuauauauuu	20	DX549602	+	146	5′	617	762	617	636	16.9	54	0	16	76	10.96	1.06
403	*G. hirsutum*	uuagauucacgcacggauucu	21	DX544824	+	78	5′	1245	1322	1245	1265	15	13	24	13	28	47.44	0.41
407	*G. hirsutum*	cauaaaucauauacuuuucca	21	DX522908	–	133	5′	20	152	132	152	23.5	33	33	19	47	39.39	0.45
413	*G. hirsutum*	cgaaauucucuuguucugca	20	DX547025	+	571	3′	51	621	602	621	96.5	188	120	63	200	32.05	0.53
413	*G. raimondii*	auaguuucucuuguugauuca	21	DX406164	+	108	3′	80	187	167	187	22.9	39	14	23	32	34.26	0.62
413	*G. exiguum*	agaauuucucuuguucugcaa	21	DX394592	+	100	5′	467	566	467	487	18.6	30	14	20	36	34.00	0.55
414a	*G. hirsutum*	ucaucuucaucaucaucgaau	21	DX383852	+	181	3′	353	533	513	533	32.4	46	40	23	72	34.81	0.51
415	*G. hirsutum*	ugaaaagcagaaacagaacaa	21	DX528271	–	89	5′	505	593	573	593	16.2	26	16	14	33	33.71	0.54
417	*G. hirsutum*	auagguagugaauuugauuuu	21	DX547211	+	57	3′	290	347	923	943	5.2	17	0	13	27	22.81	0.40
776	*G. hirsutum*	ucuaagucuucuauuaauguu	21	DX539239	+	266	3′	609	874	854	874	66.1	77	39	61	89	37.59	0.66
779	*G. hirsutum*	uguccucauguugcugcucau	21	DX555223	+	41	3′	118	158	138	158	16.7	8	8	12	13	48.78	0.84
782	*G. hirsutum*	gcauggccuuggauguucuu	20	DX527792	+	141	3′	579	719	700	719	38.7	34	32	26	49	41.13	0.67

ML: mature sequence length; PL: length of pre-miRNAs. Letters in shadow indicate the nucleotide substitution compared with previously known *A. thaliana* miRNAs.

candidates predicted by comparative genomics. In this study, we checked the expression of the identified cotton miRNAs against the NCBI EST database and the TIGR cotton EST database using all identified cotton miRNA homologues.

2.5. Potential miRNA targets

Previous studies demonstrated that miRNA-targeted mR have perfect or near-perfect complementary sites with miR

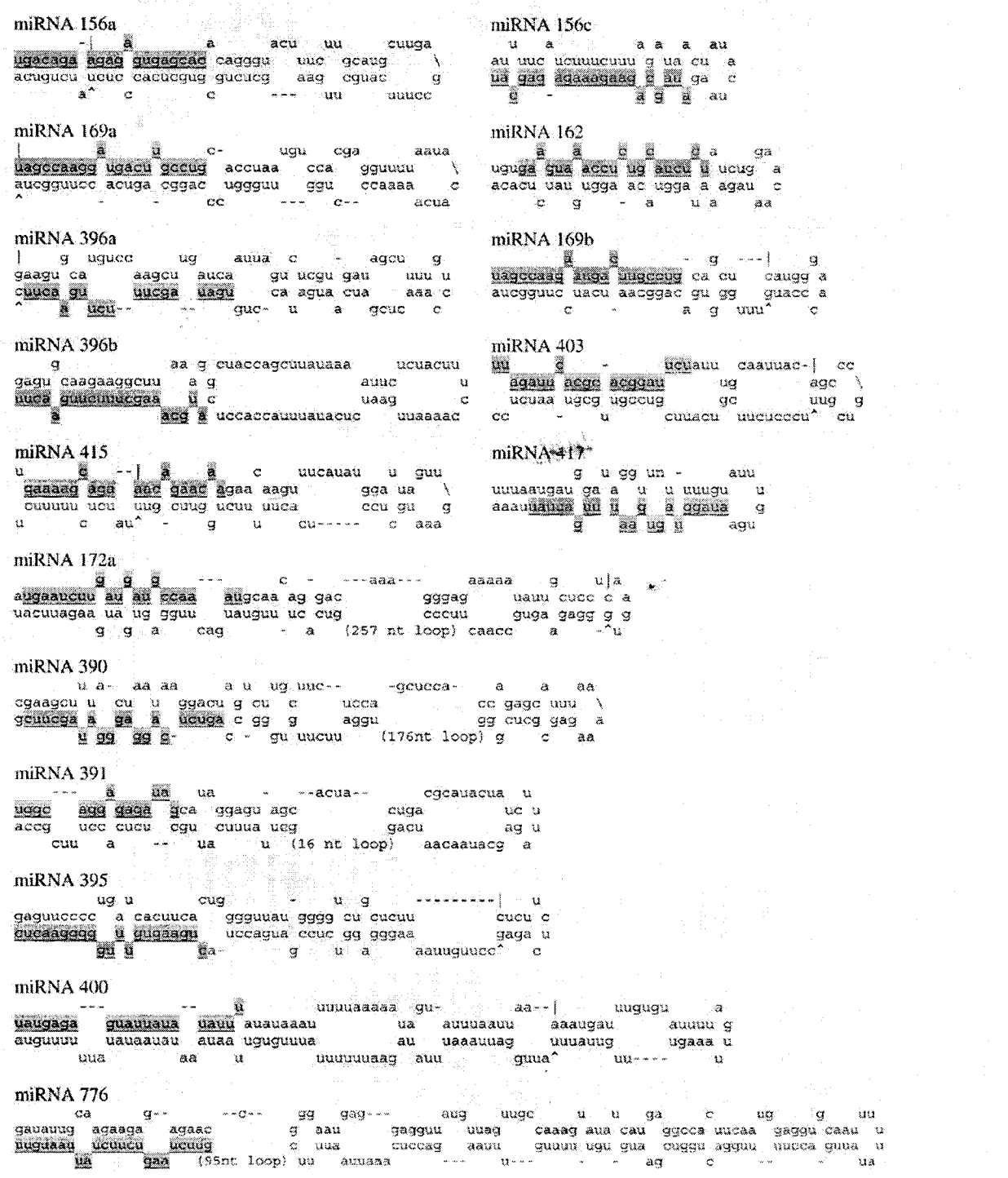

Fig. 2. Predicted hairpin secondary structures of the selected cotton miRNAs identified in this study. Mature miRNA sequences are in shadow. The length of the accurate miRNA precursors may be slightly longer than what is presented here.

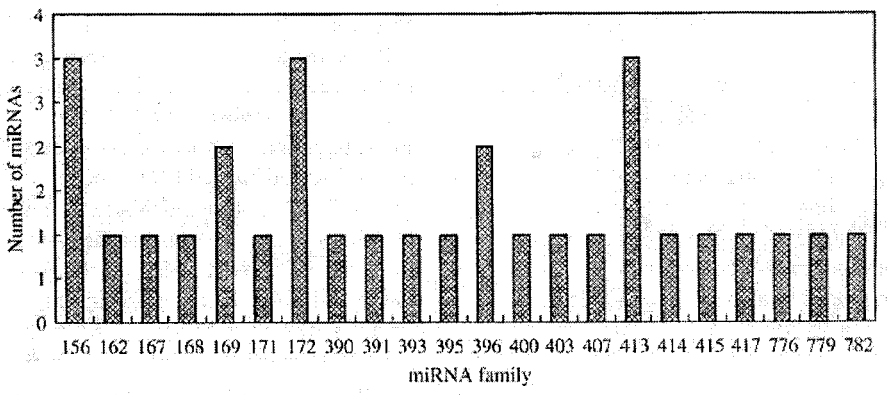

Fig. 3. miRNA family size in cottons.

and miRNAs negatively regulate gene expression by binding to these targeted mRNAs at these complementary sites for direct cleavage of mRNAs or repression of protein translation (Bartel, 2004; Chen, 2004). This suggests a powerful approach to predict miRNA targets in plants by simply using homology search. To date, a majority of miRNA targets in plants have been predicted using this strategy and then confirmed by experimental approaches. In this study, we used the same approach to predict miRNA targets in cotton. The procedure was similar to that described above for predicting cotton miRNA homologues. The only modification was that we used the identified cotton miRNAs to do a BLASTn search in the protein-coding gene databases instead of the GSS database. We tested the identified cotton miRNAs against the GenBank protein-coding nucleotide databases using BLASTn search and the cotton nucleotide databases from TIGR using miRU (Zhang, 2005). The parameters, which used in BLAST search for miRNA targets, include total numbers of mismatched nucleotides between miRNAs and the potential targets and the alignment structures. The conservation of a target site in other plant species was also considered for identifying miRNA targets and removing the false positives. The total number of allowed mismatches at complementary sites between miRNA sequences and potential mRNA targets was no more than four, and no gaps were allowed at the complementary sites.

3. Results

3.1. Potential cotton miRNAs identified

After searching the NCBI genomic survey sequence database, removing the protein-coding, repeated sequences and potential false positives, a total of 30 miRNA were identified

(a)

Species	Common name		miRNA 156
Gossypium hirsutum	Cotton	Eudicots	UGACAGAAGAGAGUGAGCAC
Arabidopsis thaliana		Eudicots	UGACAGAAGAGAGUGAGCAC
Glycine max	Soybean	Eudicots	UGACAGAAGAGAGUGAGCAC
Populus trichocarpa	Black cotton wood	Eudicots	UGACAGAAGAGAGUGAGCAC
Oryza sativa	Rice	Monocots	UGACAGAAGAGAGUGAGCAC
Sorghum bicolor	Sorghum	Monocots	UGACAGAAGAGAGUGAGCAC
Saccharum officinarum	Sugarcane	Monocots	UGACAGAAGAGAGUGAGCAC
Zea mays	Corn	Monocots	UGACAGAAGAGAGUGAGCAC
Physcomitrella patens		Mosses	UGACAGAAGAGAGUGAGCAC

(b)

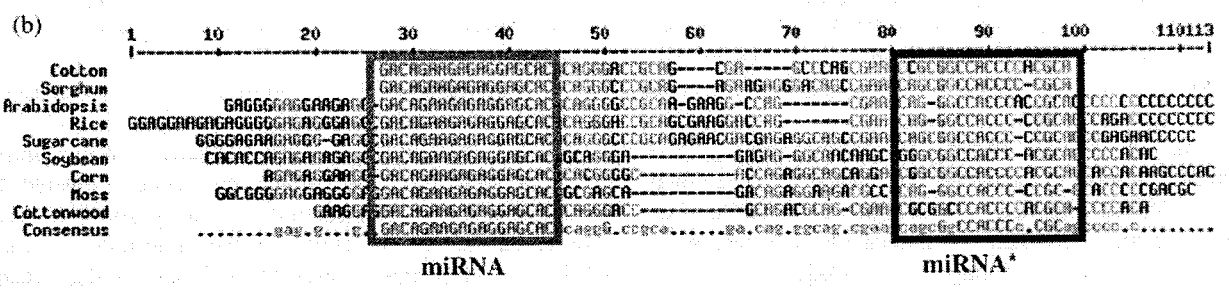

Fig. 4. Comparison of the newly identified cotton miRNA 156 and the miRNA 156s in other plant species deposited in miRBase database. Multiple sequence alignment of (a) mature sequence and (b) precursor sequences of miRNA 156, and (c) secondary structure of each pre-miRNA 156 in different plant species. Mature sequences and secondary structures of miRNA 156 are highly evolutionarily conserved from species to species. However, the other parts of pre-miRNA 156 except the miRNA and miRNA* are less conserved.

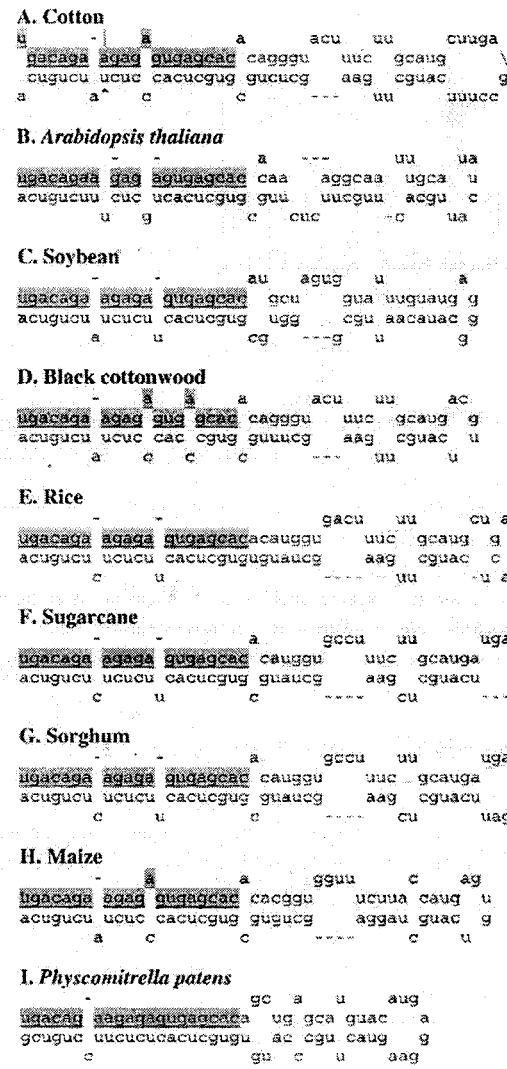

Fig. 4 (*continued*).

in the cotton genomic survey sequence database (Table 1, Fig. 2, and Supplementary material, Fig. S1). These miRNAs belong to 22 miRNA families, suggesting that cotton miRNAs are also as diverse as *A. thaliana* and rice miRNAs. However, the distribution of miRNAs was not even among the miRNA families (Fig. 3). Some miRNA families had more members than others. For example, 3 members were identified in the miRNA 156 and miRNA 172 families while only a single member was found in the majority of miRNA families. One possible reason is that little work has been done on the cotton genome; there are only about 50,000 GSSs in the NCBI database and a majority of the cotton genome sequences have not encoded.

In other plant species, miRNA diversity was not only observed in the numbers within each miRNA family, but also in other aspects, such as the length of miRNA precursors and the location of mature miRNAs. A similar phenomenon was observed in the identified cotton miRNAs. First, the length o miRNA precursors varies in cotton, ranging from 41 to 573 with an average of 177 nucleotides. However, a majority of cotton miRNA precursors have 60–110 nucleotides. Second, the nucleotide content was not even for A, G, C, and U. G and C content ranged from 11.0 to 64.0% with an average of 38.7± 10.4%. Third, mature miRNAs can be located at each arm of the stem-loop hairpin structure; some at the 5′ end of the miRNA precursor sequences, others at the 3′ end. In the 30 identified cotton miRNAs, 17 were located at the 3′ end while 13 were located at the 5′ end.

It is a universal phenomenon for animal miRNAs to be clustered together (Seitz et al., 2004; Tanzer and Stadler, 2004; Altuvia et al., 2005; Tanzer et al., 2005; Yu et al., 2006). For example, conserved miR-17 cluster contains 6 miRNA members in human and other mammal genomes (Lagos-Quintana et al., 2001; Lagos-Quintana et al., 2003; Tanzer and Stadler, 2004). Clustered miRNAs were not frequently observed in plants. To date, only four miRNA clusters have been reported in plants. Jones-Rhoades and Bartel (2004) observed a miRNA 353 cluster in *A. thaliana* which contains three miRNA 395s (miRNA 395a, b, and c) (Jones-Rhoades and Bartel, 2004). Recently, two additional plant miRNA clusters (miRNA 395 and miRNA 393, each contains two miRNAs on the opposite arms of the predicted secondary structure) and the same clusters were identified in *A. thaliana*, rice, and maize (Jones-Rhoades et al., 2004; Zhang et al., 2006c). Another study found that the precursors of miR1219a and miR1219b are located within approximately 200 bp of each other (Talmor-Neiman et al., 2006). For the identified cotton miRNAs, a majority were located in a separate GSS sequence. However, we observed that miRNA 169a and miRNA 169b were clustered together. Unlike the miRNA 395 and miRNA 393 clusters, these two miRNA 169s were located at separate miRNA precursors in the same direction, and they were separated by 155 nt. To our best knowledge, this is the first observation of a cluster for miRNA 169. This miRNA cluster was not observed in model plant species, such as *A. thaliana* and rice, suggesting plant miRNAs evolved quickly. The miRNA 169 cluster may evolve by duplication of miRNAs as observed in animal miRNAs. Accumulating evidence in animals suggests that clustered miRNAs have similar gene expression patterns and are transcribed together as a polycistron. However, the different clustered miRNA pattern between animals and plants may suggest different miRNA evolution mechanisms in these two kingdoms.

In the 30 miRNAs identified in cotton, only three miRNAs (miRNA 156a, 169a, and 395) have the same nucleotide sequences as *A. thaliana* miRNAs; other identified miRNAs have at least a one nucleotide difference compared to *A. thaliana* miRNAs. This suggests that some miRNAs are more conserved than others. However, the same miRNA family is highly evolutionarily conserved and has similar secondary structures in different plant species regardless of eudicots, monocots, or mosses (Fig. 4).

Although 863 plant miRNAs have been identified and deposited in the miRNA database (miRBase, release 9.0,

October 2006), a majority of them were identified from model plant species such as *A. thaliana*, rice, and *P. trichocarpa* (Griffiths-Jones et al., 2006). No single miRNA has been reported in cotton although it is one of the most important fiber and economic crops in the world. Carefully analyzing the species in which a majority of the miRNAs has been identified, one common characteristic is that their genomes have been completely (or nearly) sequenced. In contrast, cultivated cotton is an allotetraploid (AD genome, n=26) species (Brubaker et al., 1999); the components and size of the cotton genome are much bigger than other plant species, especially the model species. Thus, it is more difficult to study cotton miRNAs than other plant species. In this study, we identified 30 miRNAs, belonging to 22 miRNA families, in cotton using comparative genomics which we developed for identifying conserved miRNAs in other plant species. This will facilitate the investigation on the functions of miRNAs in cotton, especially on cotton fiber differentiation and development, the best system for studying cellulose synthesis.

3.2. Characteristics of cotton miRNA expression

ESTs are partially transcribed gene sequences, which have been used to confirm the existence and expression of potential miRNAs predicted by computational approaches in *Arabidopsis*, rice and maize (Bonnet et al., 2004; Zhang et al., 2006a). In this study, we also tested the predicted cotton miRNAs individually against the EST databases of GenBank and the cotton EST databases produced by TIGR. Our BLASTn search results indicated that several predicted cotton miRNAs exist in cotton EST databases, suggesting that these miRNAs were expressed in the cotton genome.

3.3. Cotton miRNA targets

miRNAs regulate gene expression at the posttranscriptional levels by directly cleaving mRNAs (Schwab et al., 2005; Sunkar et al., 2005) or inhibiting protein translation (Aukerman and Sakai, 2003; Chen, 2004). The miRNA-regulated genes control a variety of biological and metabolic processes. For example, in animals, miRNAs regulate developmental timing, stem cell maintenance and differentiation, organ development, signal transduction, disease, and cancer pathogenesis (Carrington and Ambros, 2003; Zhang et al., 2006b; Zhang et al., 2007a); in plants, miRNAs regulate leaf, stem, root, and flower development, phase switch from vegetative growth to reproductive growth, and responses to abiotic and biotic stress (Zhang et al., 2006d). Several studies have indicated that miRNAs directly target transcription factors, which regulate plant development as well as specific genes which control various metabolic processes.

In this study, we identified a total of 139 potential targets for the 22 identified miRNA families in cotton based on the fact that miRNAs show perfect or near-perfect complementarity to their target mRNA sequences (Fig. 5 and Table 2). These 139 potential miRNA targets belong to several gene families and they have different biological functions, including the control of cotton development, signal transduction, phase change, fiber development, and responses to environmental stress. From very beginning, scientists believe that almost all miRNA targets have no more than four mismatches with their corresponded miRNAs in plants (Rhoades et al., 2002), and this criterion has been widely adopted to identify miRNA targets in different plant species by different laboratories. In 2005, Schwab and collaborator suggested an empirical parameter for recognizing

```
SBP (AY779340)(5'-3')              592  acgcucccucucuucuguca  3'
                                        |||| ||||||||||||||
156 (3'-5')                            cacgagugagagaagacagu

Heat-shock protein (AY819767, 5'-3') 659  cuaugcuaucgcgccugcua 678
                                        |||||||||| | |||||°|
391 (3'-5')                            gauacgauagagaggacggu

TC39605 (Katanin)                 1851  guguucaagaaaguuauuaca 1871
                                        ° ||||||||||°|||°|||
396 (3'-5')                            uucaaguucuuucgauagugu

AF085717 (callose synthase)       2835  aguucaggaaugcuauuaug 2854
                                        ||||||°||| |||||°| °°
396 (3'-5')                            ucaaguucuuucgauagugu

TC34607 (fiber protein Fb23)        81  auucggagauguugaggauga 101
                                        |||||° |||| |||°|||||
414 (3'-5')                            uaagcuacuacuacuucuacu

TC37120(Fiber quinone-oxidoreductase)1037 aagagcauugaaggccaugc 1056
                                        ||||°|||° ||||||||||
782 (3'-5')                            uucuuguagguuccgguacg
```

Fig. 5. Predicted miRNA targets and their complementary sites with mRNAs. Each bottom strand depicts the miRNA, and each top strand depicts a miRNA complementary site. Watson–Crick pairing (vertical dashes) and G:U wobble pairing (circles) are indicated.

Table 2
Potential targets of the identified miRNAs in cotton

miRNA	Targeted protein	Target function	Targeted genes or EST homologs of *Arabidopsis* genes[a]
156	Squamosa-promoter binding protein-like protein (SBP)	Transcription factor	TC34708 (1), TC34910 (1), CO092899 (1), AY779340 (3), BG447097(3), TC36356 (1)
	Unknown		TC41412 (3)
162	Serine/threonine protein kinase	Metabolism	TC31417 (3), TC39777 (3)
	Unknown		CO112217 (3), TC40045 (3), CO081163 (3)
167	Unknown		BF274298 (3)
168	Unknown		TC38204 (2.5), TC30572 (3)
169	CCAAT-binding transcription factor	Transcription factor	TC32844 (1.5), TC31014 (2.5), TC40955 (2)
	Nuclear transcription factor Y subunit A-8	Transcription factor	TC29763 (2)
	Unknown protein		CO107474 (2), TC36082 (3), BG442852 (3)
171	mRNA cap methyltransferase	RNA processing	TC40160 (3)
	Unknown		TC36655 (3)
172	AP2 domain containing protein	Transcription factor	TC37805 (3)
	26S proteasome subunit-like protein	Protein degradation	TC29130 (3)
	Small nuclear ribonucleoprotein-like protein	RNA processing	TC40379 (3)
	Unknown protein		CO103466 (2), BG440541 (3), CO089049 (3), CO074155 (3), CO106219 (3), TC39084 (3)
390	Protein phosphatase 2C-like protein	Metabolism	TC38203 (3)
	ATP-dependent peptidase/ATPase/ metallopeptidase/zinc ion binding (VAR2)	Metabolism	TC29295 (3)
	Unknown protein		TC34838 (2.5), CO129000 (3)
391	Aquaporin PIP2.7	Stress response	TC38801 (3), TC35561 (3),
	Calmodulin binding heat shock protein	Stress response	AY819767 (3), TC40030 (2.5)
	Unknown protein		TC35812 (2.5)
393	NADPH oxidase	Metabolism	TC29375 (3)
	26S proteasome regulatory subunit S5A	Protein degradation	TC37709 (3)
	Unknown		CO108800 (2), CO096633 (3), BG444553 (3), CA993875 (3)
395	ATP sulfurylase like protein	Metabolism	CO128820 (2), TC30934 (2)
	Unknown protein		BG443152 (3), TC38540 (2.5), TC38341 (3) ; TC30287 (3), TC34810 (3), TC41629 (3)
396	Callose synthase catalytic subunit	Fiber development	TC37877 (3), AF085717 (3)
	Katanin (microtubule-severing protein)	Fiber development	TC39605 (2), AAP83638 (2), AY324646, AY289200, TC39591 (2)
	NBS resistance protein-like protein	Stress resistance	NP508211 (3), AF469073, AY705376
	Dirigent-like protein	Disease resistance-responsive protein	CD486128 (3), DQ018709, AY560544
	Pathogen-related protein mRNA	Disease resistance response	AY560553, AF305065, AY560552, AF305066, AY588276, AY241395
	Ribonuclease-like protein	Disease resistance response	AY560551, AF416652,
	Unknown proteins		TC41168 (1.5), TC39855 (2.5), AI729912 (2.5), BF274071 (3), BE053871 (2.5), TC38464 (2), TC33397 (2.5)
400	Hypersensitive-induced response protein	Stress response	TC28849 (3),
	UVB-resistance protein-like	Stress response	TC37861 (3)
	Unknown		TC35971 (3)
403	40S ribosomal protein S16		TC28146 (3), X75954
	Unknown protein		CO119475 (3)
407	Unknown		TC28193 (2.5), CO071374 (2.5), TC28818 (3)
413	Phosphoenolpyruvate carboxykinase	Metabolism	TC32496 (3)
	Cyclin D	Cell cycle	TC40773 (3)
	MAP kinase kinase 3	Metabolism	CO129972 (3)
	Methylthioribose kinase	Metabolism	TC33765 (3)
	Unknown		CO124877 (3), TC34191 (3), CD485672 (3)
414	Fiber protein Fb23	Fiber development	TC34607 (3), AY273899
	Splicing factor-like protein		TC39657 (1.5)
	Unknown		CO115005 (1.5), TC33658 (2),
415	Glycosyltransferase	Metabolism	TC32986 (3)
	Translation initiation factor 5A		TC27480 (3)
	Unknown		AW187185 (2.5), BF275332 (2.5)
417	Unknown		TC37761 (3), TC32039 (3), TC31954 (3), TC38461 (3), TC39323 (3), CA993924 (2.5), TC35737 (3)

Table 2 (*continued*)

miRNA	Targeted protein	Target function	Targeted genes or EST homologs of *Arabidopsis* genes[a]
776	Unknown		TC39223 (2.5), CO128782 (2.5), BG445454 (3), TC27932 (3)
779	BEL1-like homeodomain transcription factor	Transcription factor	TC35957 (2.5)
	E6 (Protein kinase)	Fiber development	TC32567 (3), TC32568 (3)
	Unknown		AW728937 (3), TC29921 (3), AW730194 (3), CO109694 (3)
782	Fiber quinone-oxidoreductase	Fiber development	TC37120 (2), AY429443 (2), TC37118 (3), TC37119 (3), TC32628 (3)
	Ferrochelatase	Metabolism	TC29346 (3)
	Unknown		BF271623 (2.5), TC30343 (3), TC34774 (3), BF272717 (3), TC33758 (3), TC35815 (3), CO110625 (3)

[a] The number in brackets stands for the score of miRU.

miRNA targets based on a limited data for 4 miRNAs (miRNA 156, miRNA 159, miRNA 164, and miRNA 319). In their study, they found that no mismatch existed at the positions 10 and 11 of miRNAs between miRNA and their potential targets (Schwab et al., 2005). However, this conclusion has not been widely accepted by the scientific community although it may become a better criterion to identify miRNA targets. Thus, more studies need to be done confirmed this conclusion, particularly for other miRNAs which have not been tested in Schwab's study. In our study, we identify several potential miRNA targets have a mismatch at the position 10 or 11 with their corresponded miRNAs. Some of them are cotton-specific genes. However, more studies (such as detecting mRNA fragments diagnostic of miRNA-directed cleavage) need to be done for testing this conclusion.

In this study, a majority of miRNA targets were found to be transcription factors in *Arabidopsis*, rice, corn and other plant species (Rhoades et al., 2002; Bonnet et al., 2004; Zhang et al., 2006a). In this study, we found that 5 identified cotton miRNA families target a variety of transcription factors, which control cotton development at the tissue level. miRNA 172 targets cotton apetal-like protein (ap2), which controls cotton floral formation and phase change from vegetative growth to reproductive growth. miRNA 156 targets the SQUAMOSA-promoter binding protein-like protein (SBP) transcription factor. miRNA 169 targets CCAAT-binding (CBF) transcription factor and nuclear transcription factor Y subunit A-8. These transcription factors have also been reported as miRNA targets in *Arabidopsis*, rice, and maize except for the nuclear transcription factor Y subunit A-8. In this study, we also found that miRNA 415 and miRNA 779 target transcription factors. Their targets are translation initiation factor 5A and BEL1-like homeodomain transcription factor, respectively.

Three miRNA families (miRNA 391, 396, and 400) may involve cotton response to environmental stress. One of the two miRNA 391 targets is aquaporin PIP2.7, which plays an important role in drought and salt stress. miRNA 391 also targets calmodulin binding heat shock protein. miRNA 396 targets four classes of stress resistance protein: pathogen-related (PR) protein, NBS resistance protein-like protein, dirigent-like protein, and ribonuclease-like protein; all of these four proteins are related to cotton response to pathogen invasion. Two potential targets for miRNA 400 are hypersensitive-induced response protein and UVB-resistance protein. Compared with previous reports in other plant species, miRNAs likely target more abiotic and biotic stress-related proteins in cotton.

Cotton miRNAs also target genes involved in different cellular processes, including metabolic pathways, protein degradation, mature mRNA formation, and signal transduction. miRNA 162 targets serine/threonine protein kinase. miRNA 390 targets protein phosphatase 2C-like protein. miRNA 395 targets ATP sulfurylase like protein. miRNA 413 targets carboxykinase, methythioribose kinase, and MAP kinase. miRNA 415 targets glycosyltransferase. miRNA 779 and 782 target E6 protein kinase and ferrochelatase, respectively. NADPH oxidase may be one target of miRNA 393.

Cotton miRNAs also target mature mRNA formation and protein degradation. mRNA cap methytransferase, small nuclear ribonucleoprotein-like protein, and splicing factor-like protein are potential targets for miRNA 171, 172, and 414, respectively. 26S proteasome regulatory subunit S5A is one potential target for miRNA 393.

Interestingly, at least 3 miRNA families may have cotton-specific unique targets, which are involved in cotton fiber differentiation and development. Callose synthase is considered an important enzyme in cellulous synthesis and fiber development, first cloned by Dr. Brown's lab at the early 2000s (Cui et al., 2001). In this study, we found that the callose synthase catalytic subunit is one target of miRNA 396. Additionally, miRNA 414 and miRNA 782 target fiber protein Fb23 and fiber quinone-oxidoreductase, respectively. Further studying the functions of miRNA in cotton fiber will allow better understanding the molecular mechanisms of cotton fiber differentiation and development.

4. Conclusions

In this study, we identified 30 potential miRNAs and their targets in cotton, one of the most important crops around the world. One miRNA cluster (miRNA 169 cluster) was also observed for the first time in the plant kingdom, suggesting that miRNAs evolved quickly in plants and they may have specific functions in cotton. In addition to the transcription factors which regulate plant development, miRNA targets in cotton also include several genes involved in cotton response to different biotic and abiotic stress, such as pathogen invasion and drought stress. Interestingly, at least three miRNAs (miRNA 396, 414, and 782) target genes involved in cotton fiber formation and

development, suggesting that miRNAs may regulate fiber differentiation and development, one of the most important biological processes in plant development.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2007.03.020.

References

Altschul, S.F., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25, 3389–3402.

Altuvia, Y., et al., 2005. Clustering and conservation patterns of human microRNAs. Nucleic Acids Research 33, 2697–2706.

Ambros, V., 2001. microRNAs: tiny regulators with great potential. Cell 107, 823–826.

Aukerman, M.J., Sakai, H., 2003. Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. Plant Cell 15, 2730–2741.

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297.

Bonnet, E., Wuyts, J., Rouze, P., Van de Peer, Y., 2004. Detection of 91 potential in plant conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. Proceedings of the National Academy of Sciences of the United States of America 101, 11511–11516.

Borchert, G.M., Lanier, W., Davidson, B.L., 2006. RNA polymerase III transcribes human microRNAs. Nature Structural & Molecular Biology 13, 1097–1101.

Brubaker, C.L., Paterson, A.H., Wendel, J.F., 1999. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome 42, 184–203.

Carrington, J.C., Ambros, V., 2003. Role of microRNAs in plant and animal development. Science 301, 336–338.

Chen, X.M., 2004. A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. Science 303, 2022–2025.

Chen, X.M., 2005. microRNA biogenesis and function in plants. FEBS Letters 579, 5923–5931.

Cui, X.J., Shin, H.S., Song, C., Laosinchai, W., Amano, Y., Brown, R.M., 2001. A putative plant homolog of the yeast beta-1,3-glucan synthase subunit FKS1 from cotton (*Gossypium hirsutum* L.) fibers. Planta 213, 223–230.

Floyd, S.K., Bowman, J.L., 2004. Gene regulation: ancient microRNA target sequences in plants. Nature 428, 485–486.

Fryxell, P.A., Craven, L.A., Stewart, J.M., 1992. A revision of *Gossypium* Sect Grandicalyx (Malvaceae), including the description of 6 new species. Systematic Botany 17, 91–114.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., Enright, A.J., 2006. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Research 34, D140–D144.

Heijbroek, A., Husken, H., 1996. The International Cotton Complex. Rabobank, Utrecht, Netherlands.

IAC, 1996. Cotton: Review of World Situation, Monogram by International Advisory Committee, Washington, D.C.

Jacob-Wilk, D., Kurek, I., Hogan, P., Delmer, D.P., 2006. The cotton fiber zinc-binding domain of cellulose synthase A1 from *Gossypium hirsutum* displays rapid turnover *in vitro* and *in vivo*. Proceedings of the National Academy of Sciences of the United States of America 103, 12191–12196.

Jones-Rhoades, M.W., Bartel, D.P., 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Molecular Cell 14, 787–799.

Kurihara, Y., Watanabe, Y., 2004. *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proceedings of the National Academy of Sciences of the United States of America 101, 12753–12758.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T., 2001. Identification of novel genes coding for small expressed RNAs. Science 294, 853–858.

Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., Tuschl, T., 2003. New microRNAs from mouse and human. RNA 9, 175–179.

Lau, N.C., Lim, L.P., Weinstein, E.G., Bartel, D.P., 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294, 858–862.

Lee, R.C., Ambros, V., 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. Science 294, 862–864.

Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. Cell 75, 843–854.

Lee, Y., Jeon, K., Lee, J.T., Kim, S., Kim, V.N., 2002. MicroRNA maturation: stepwise processing and subcellular localization. EMBO Journal 21, 4663–4670.

Lee, Y., et al., 2004. MicroRNA genes are transcribed by RNA polymerase II. EMBO Journal 23, 4051–4060.

Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. Journal of Molecular Biology 288, 911–940.

Nakashima, J., Heathman, A., Brown, R.M., 2006. Antibodies against a *Gossypium hirsutum* recombinant cellulose synthase (Ces A) specifically label cellulose synthase in *Micrasterias denticulata*. Cellulose 13, 181–190.

Pan, X.P., Zhang, B.H., SanFrancisco, M., Cobb, G.P., 2007. Characterizing viral microRNAs and its application on identifying new microRNAs in viruses. Journal of Cellular Physiology 211, 10–18.

Papp, I., et al., 2003. Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant Physiology 132, 1382–1390.

Pasquinelli, A.E., et al., 2003. Expression of the 22 nucleotide *let-7* heterochronic RNA throughout the Metazoa: a role in life history evolution? Evolution & Development 5, 372–378.

Pasquinelli, A.E., et al., 2000. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. Nature 408, 86–89.

Reinhart, B.J., et al., 2000. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403, 901–906.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., Bartel, D.P., 2002. MicroRNAs in plants. Genes & Development 16, 1616–1626.

Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., Bartel, D.P., 2002. Prediction of plant microRNA targets. Cell 110, 513–520.

Saxena, I.M., Brown, R.M., 2005. Cellulose biosynthesis: current views and evolving concepts. Annals of Botany 96, 9–21.

Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., Weigel, D., 2005. Specific effects of microRNAs on the plant transcriptome. Developmental Cell 8, 517–527.

Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., Cavaille, J., 2004. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Research 14, 1741–1748.

Sunkar, R., Girke, T., Jain, P.K., Zhu, J.K., 2005. Cloning and characterization of microRNAs from rice. Plant Cell 17, 1397–1411.

Talmor-Neiman, M., Stav, R., Frank, W., Voss, B., Arazi, T., 2006. Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss. Plant Journal 47, 25–37.

Tang, G.L., Reinhart, B.J., Bartel, D.P., Zamore, P.D., 2003. A biochemical framework for RNA silencing in plants. Genes & Development 17, 49–63.

Tanzer, A., Stadler, P.F., 2004. Molecular evolution of a microRNA cluster. Journal of Molecular Biology 339, 327–335.

Tanzer, A., Amemiya, C.T., Kim, C.B., Stadler, P.F., 2005. Evolution of microRNAs located within *Hox* gene clusters. Journal of Experimental Zoology Part B-Molecular and Developmental Evolution 304B, 75–85.

Weber, M.J., 2005. New human and mouse microRNA genes found by homology search. FEBS Journal 272, 59–73.

Wendel, J.F., Cronn, R.C., 2003. Polyploidy and the evolutionary history of cotton. Advances in Agronomy 78, 139–186.

Yu, B., et al., 2005. Methylation as a crucial step in plant microRNA biogenesis. Science 307, 932–935.

Yu, J., et al., 2006. Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. Biochemical and Biophysical Research Communications 349, 59–68.

Zhang, Y.J., 2005. MiRU: an automated plant miRNA target prediction server. Nucleic Acids Research 33, W701–W704.

Zhang, B.H., Feng, R., 2000. Cotton Resistance to Insect and Pest-Resistant Cotton. Chinese Agricultural Science and Technology Press, Beijing.

Zhang, B.H., Liu, F., Yao, C.B., Wang, K.B., 2000. Recent progress in cotton biotechnology and genetic engineering in China. Current Science 79, 37–44.

Zhang, B.H., Pan, X.P., Wang, Q.L., Cobb, G.P., Anderson, T.A., 2005. Identification and characterization of new plant microRNAs using EST analysis. Cell Research 15, 336–360.

Zhang, B.H., Pan, X.P., Anderson, T.A., 2006a. Identification of 188 conserved maize microRNAs and their targets. FEBS Letters 580, 3753–3762.

Zhang, B.H., Pan, X.P., Anderson, T.A., 2006b. MicroRNA: a new player in stem cells. Journal of Cellular Physiology 209, 266–269.

Zhang, B.H., Pan, X.P., Cannon, C.H., Cobb, G.P., Anderson, T.A., 2006c. Conservation and divergence of plant microRNA genes. Plant Journal 46, 243–259.

Zhang, B.H., Pan, X.P., Cobb, G.P., Anderson, T.A., 2006d. Plant microRNA: a small regulatory molecule with big impact. Developmental Biology 289, 3–16.

Zhang, B.H., Pan, X.P., Cox, S.B., Cobb, G.P., Anderson, T.A., 2006e. Evidence that miRNAs are different from other RNAs. Cellular and Molecular Life Sciences 63, 246–254.

Zhang, B.H., Pan, X.P., Wang, Q.L., Cobb, G.P., Anderson, T.A., 2006f. Computational identification of microRNAs and their targets. Computational Biology and Chemistry 30, 395–407.

Zhang, B.H., Pan, X.P., Cobb, G.P., Anderson, T.A., 2007a. microRNAs as oncogenes and tumor suppressors. Developmental Biology 302, 1–12.

Zhang, B.H., Wang, Q.L., Pan, X.P., 2007b. MicroRNAs and their regulatory roles in animals and plants. Journal of Cellular Physiology 210, 279–289.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31, 3406–3415.