SCIENCE CHINA
Life Sciences

• RESEARCH PAPER •

September 2012 Vol.55 No.9: 807–817
doi: 10.1007/s11427-012-4367-y

Computational prediction and experimental verification of miRNAs in *Panicum miliaceum* L.

WU YongJun[1], DU JiangFeng[1], WANG XiaoLong[1], FANG XiaoFeng[3], SHAN WeiXing[2] & LIANG ZongSuo[1*]

[1]*College of Life Sciences, Northwest Agriculture & Forestry University, Yangling 712100, China;*
[2]*College of Plant Protection, Northwest Agriculture & Forestry University, Yangling 712100, China;*
[3]*National Institute of Biological Sciences, Beijing 102206, China*

Received January 2, 2012; accepted July 12, 2012

MicroRNAs (miRNAs) are a class of non-coding RNAs that play critical roles in post-transcriptional regulation. Their target genes are involved in a variety of biological processes such as development, metabolism, and stress response. *Panicum miliaceum* L. (*Panicum*) is an important grain crop, but, until now, no miRNAs have been identified in this plant. Using a homology search based on expressed sequence tag (EST) analysis and miRNA precursor secondary structure, a total of 43 new miRNAs were identified. The miRNAs were found to be unevenly distributed among 11 miRNA families. Target analysis using the plant small RNA target analysis server psRNATarget showed that the newly identified miRNAs can potentially regulate 68 target genes. Ten of the 11 miRNA families were annotated as involved in RNA regulation, suggesting they may play an essential role in post-transcriptional regulation in Panicum. Selected miRNAs representing eight of the families were verified by northern blotting, indicating that the prediction method that we used to identify the miRNAs was effective.

microRNA, *Panicum*, prediction, verification, targets

Citation: Wu Y J, Du J F, Wang X L, *et al*. Computational prediction and experimental verification of miRNAs in *Panicum miliaceum* L.. Sci China Life Sci, 2012, 55: 807−817, doi: 10.1007/s11427-012-4367-y

MicroRNAs (miRNAs) are non-coding RNA molecules with important regulatory functions in gene expression. They originate from long self-complementary (stem-loop) precursors (pre-miRNA) consisting of 60–160 nucleotides (nt). The majority of known mature miRNAs are about 21–25 nt long and several steps are involved in their processing and maturation [1,2]. MiRNAs are widely distributed in plant, animal and even virus genomes and most of them have independent genetic loci and show conservation between different species of the same kingdom [3]. MiRNAs can direct the RISC (RNA-induced silencing complex) to post-transcriptionally regulate gene expression through messenger RNA (mRNA) decomposition or translational repression [4]. The target sites on the mRNA that show full complementarity with a miRNA can be specifically degraded; if there are mismatches, the mRNA translation can be inhibited.

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 [5–7]; 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 [8]. Several studies have indicated that miRNAs directly target the transcription factors that regulate plant development as well as specific genes that control various metabolic processes [9–11].

*Corresponding author (email: Liangzs@ms.iswc.ac.cn)

A total of 16772 miRNAs were recorded in the April 2011 release of the miRBase database (http://www.mirbase.org/). The entries included 411 and 491 miRNAs in *Arabidopsis thaliana* and *Oryza sativa*, respectively. Until now, no miRNA has been reported in *Panicum miliaceum* L., a popular grain crop around the world, which has many advantageous attributes such as a short reproductive cycle, drought resistance, and leanness-tolerance [12]. *Panicum* not only serves as a minor cereal in food supply providing humans with a balanced diet, but also is a source of beer, yellow wine and beverage production [13]. The identification of miRNAs in *Panicum* may provide new insights into understanding the mechanisms of the many biological processes that are involved in its cultivation and growth of this important plant.

Three strategies are available currently to identify miRNAs: direct cloning, forward genetics and bioinformatics, and a large number of miRNAs have been identified using these strategies. Direct cloning is the most effective and reliable method, through which the first miRNAs in plants were identified [14]. However, some technical difficulties make direct cloning and sequencing of short RNA molecular hard to implement [15]. For example, miRNAs tend to have highly constrained tissue and time-specific expression patterns; in addition, degradation products of mRNAs and other endogenous non-coding RNAs co-exist with miRNAs and these are sometimes dominant in sample containing small RNAs. These factors make it difficult to validate miRNAs using experiment approaches. Because miRNAs are conserved between different species and pre-miRNAs have a unique secondary structure [3,16–20], these characteristics have been used for miRNA prediction using bioinformatics [21]. Many novel miRNAs have been identified through this kind of strategy [6,9,15,22–26]. Jones-Rhoades and Bartel [9] developed a comparative genomic approach to systematically identify both miRNAs and their targets. Zhang and co-workers [24] identified 338 new candidate miRNAs in 60 plants by scanning a total of 18694 plant expressed sequence tags (ESTs) [27–29]. Later, the same workers [6] reported 30 potential cotton miRNAs based on genomic survey sequence analysis and miRNA secondary structure. Although most recent research has relied on genomic survey sequence analysis to predict miRNAs, predictions based on ESTs can still produce significant results. In this study, we compared *Panicum* EST sequences with all the previously known miRNAs in the miRBase database to identify candidate miRNAs in *Panicum*. To help us understand the biological processes in which they might be involved, we also predicted the potential targets of the candidate miRNAs [30]. To prove their validity, some of the computationally predicted miRNAs were verified by northern blotting hybridization, an effective and widely-used method for detecting miRNAs.

1 Materials and methods

1.1 miRNA and EST sequences datasets

A total of 8273 reported mature miRNAs and their pre-miRNA sequences were obtained from the September 2008 miRBase database (ftp://mirbase.org/pub/mirbase/12.0/). Because some miRNAs show conservation across different kingdoms [31], we included the miRNAs not only from plants such as *Arabidopsis thaliana*, *Zea mays*, *Oryza sativa*, *Glycine max*, and *Sorghum bicolor*, but also from animals like *Homo sapiens*, *Mus musculus*, *Caenorhabditis elegans*, and *Drosophila melanogaster*. To avoid overlap of the miRNAs, any repeated sequences were removed and the remaining sequences were used as the reference miRNAs dataset. A total of 436736 *Panicum* sequences were downloaded from the 2008 release of the NCBI (National Center for Biotechnology Information) GenBank EST databases (http://www.ncbi.nlm.nih.gov/).

1.2 Identification of candidate miRNAs in the *Panicum* EST dataset

BLAST (basic local alignment search tool) [32] was used to search all the reported mature miRNA sequences in the dataset against the downloaded EST sequences. The BLAST parameters were essentially the same as described previously [6,24,26,33,34]; the expected value was set to 1000 and the minimum word size was 7. A stable version of the BLAST tool was downloaded from http://blast.ncbi.nlm.nih.gov/Blast.cgi. Two rules [24,26,33] were adopted during the alignment process: first, less than three mismatches were allowed; and second, the percent identity should be no less than $(L-3+M)/L$, where L is the length of the Hits (see Bio::SearchIO (Dr. Jason Stajich) http://www.bioperl.org/wiki/HOWTO:SearchIO#Abstract), and M is the number of mismatches. A total of 6079 ESTs that potentially contain miRNA sequences were obtained after the BLAST search (Figure 1).

1.3 Identification of pre-miRNA sequences

Plant pre-miRNAs are usually about 70 nt long, while the stem-loop structures of plant pri-miRNAs vary greatly in length (from 100 to 1000 nt) [35]. In this study, we chose 60–250 nt as the length of potential precursors to obtain more precursors and used the RNAShapes software package (http://bibiserv.techfak.uni-bielefeld.de/rnashapes/) [36–38] to compute their structure. We moved a sliding segment of 60–250 nt long along each of the selected EST sequences and extracted any segment that had the same length as the sliding segment. Five base pairs were selected as the scanning window and three base pairs as the stretching window during extraction. The RNAShapes parameters [39] were set as follows: shape probabilities selected, and probability

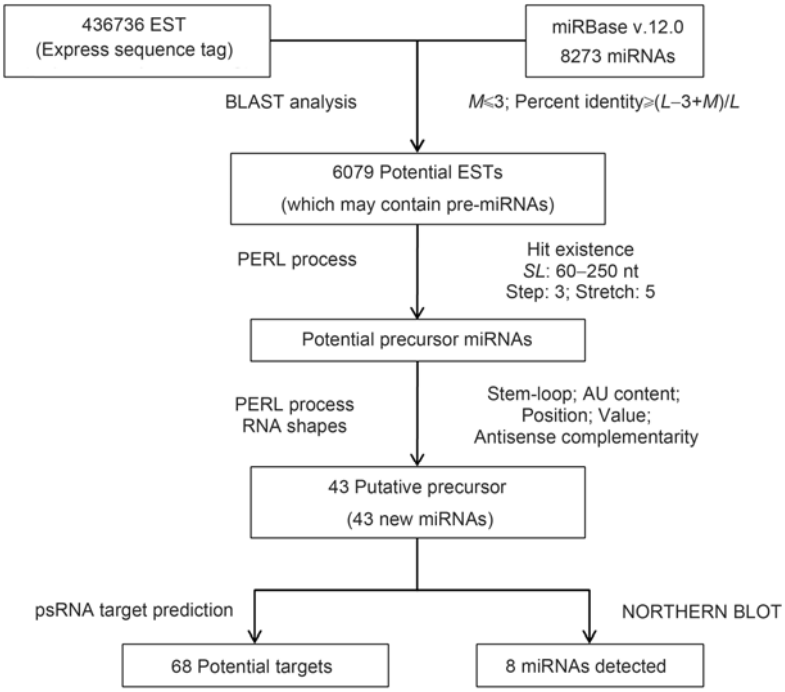

Figure 1 Flowchart for the prediction and verification of the *Panicum* miRNAs. *M*, mismatch number; *L*, length of hits; *SL*, length of the sliding segment; Position, position of the hit in the sliding segment; Value, the absolute value of the difference between the length of mature miRNA and the number of matches.

output filter as 0.5. The following criteria [24,26,40] were adopted to screen potential precursors further: (i) the threshold free energy for the precursor sequences was set to –25 kcal mol^{-1}; (ii) the (A+U) content should be between 52.0% and 68.0%; (iii) the mature miRNA sequence should be on the stem of the precursor sequence; and (iv) the value (the absolute value of the difference between the length of mature miRNA and the number of matches) was less than 5 (Figure 1).

1.4 Prediction of the potential targets

We predicted the potential targets of the newly identified candidate miRNAs using the psRNATarget server at http://www.plantgrn.org/psRNATarget/ [40]. All the parameters [40] except the dataset employed were set to the default. Because genome information of *Panicum* is not available, the dataset was set first as the TIGR Gene Index (http://compbio.dfci.harvard.edu/tgi/tgipage.html) of the species from which the original miRNA sequence came, then as the Arabidopsis Gene Index, and then as the Rice Gene Index (Figure 1).

1.5 Verification of the predicted miRNAs

Two cultivars of *Panicum*, Longmi-5 and Ningmi-9 were planted and harvested after 20 d. Whole-plant total RNA was extracted using Trizol reagent (Tiangen, Beijing, China) following the manufacturer's instruction. The procedure of northern blot was based on a published article [41] with minor modifications. Because of the high sequence similarity of miRNA members within a family, we chose 12 miRNAs that represented 11 families for verification (Table 1).

2 Results and discussion

2.1 Prediction of potential miRNAs

In the plant kingdom, most mature miRNAs are highly conserved across different species [8], and this characteristic makes it possible to predict novel miRNAs based on sequence alignments. The BLAST searches of known miRNAs against the *Panicum* ESTs dataset identified 6079 potential miRNAs-containing ESTs. Further screening of their secondary structures, free energy, and base content, identified a total of 43 potential mature miRNAs in the *P. miliaceum* L. dataset (Table 2). The candidate miRNAs were distributed unevenly among 11 miRNA families. Ten of the miRNAs were members of the miR159 family, and eight were members of the miR169 family, while only one miR171 and one miR1439 were found (Figure 2). This distribution suggested that the *Panicum* miRNAs were as diverse as the miRNAs from *A. thaliana*, *Z. mays* and *O. sativa* (Table 3).

Many miRNA genes are found in multiple locations and the candidate miRNAs that we found were often found in different ESTs with unique gene numbers. Thirteen of the 43 identified miRNAs corresponded to more than one gene number. For example, miR159b-6 matched ESTs with three different gene members, miR444b-1 matched four, and

Table 1 Probes used in the Northern blot assay

Probe	Sequence	Targets
R	TACGATCAGCATTGCTAGCGA	5S rRNA degradation product
EST1	CTGGTAGTAGTGTAGATTGCT	Random EST degradation product
EST2	CATGAAGACCACATAAGGAGC	Random EST degradation product
156	TGTGCTCACTCTCTTCTGTCA	miR156a1/2/3
159	GGGAGCACCCTTCAGTCCAA	miR159a
160	TGGCATACAGGGAGCCAGGCA	miR160-1/2
166	AGGGGAATGAAGCCTGGTCCGA	miR166a/b
167	TAGATCATGCTGGCAGCTTCA	miR167a/b-1/2
169	TAAGCAAGTCATCCTTGGCTA	miR169d
171	TAAGATATTGGCACGGCTCAA	miR171
399	CCGGGCAAATCTCCTTTGGCA	miR399b
444a	AAGCTTGAGGCAACAACTGCA	miR444a-1/2/3
444b	GGCAGCAAGCTTGAGACAACA	miR444b-1/2/3
528	CCCCTCTGCATGCCCCTTCCA	miR528b
1439	TGTACTCACTCCGTTCCAAAT	miR1439

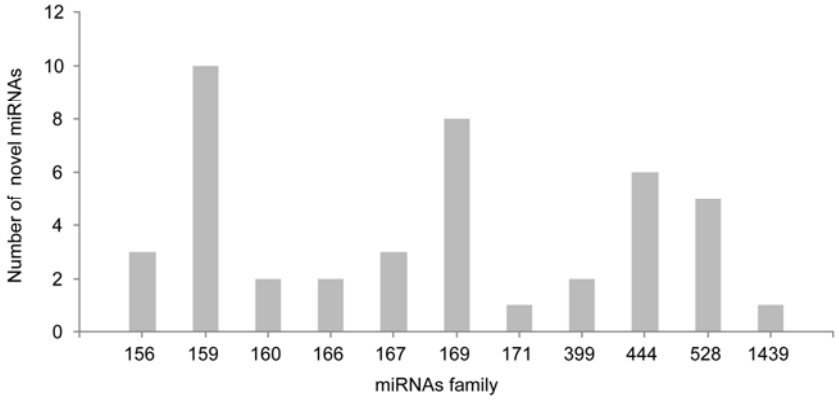

Figure 2 Number of new *Panicum* candidate miRNAs.

miR528a-2 matched six (Table 2, Figure 2). One possible reason for this is that very few studies have been done on the *Panicum* genome [42] and some of the ESTs with distinct gene numbers (Table 2) may turn out to be identical when better assemblies of the genome become available. The apparent differences in the EST sequences may result either from mistakes, repeats in sequencing, or discrepancies between the different groups that generated the sequence [43–47]. Another possibility is that the ESTs correspond to the same miRNA are copies of each other and are expressed under difference circumstances: (i) the copies are functionally complementary to each other, and malfunction of one of the copies is complemented by another; (ii) their expression is tissue- and/or time-specified; and (iii) they are expressed simultaneously.

In contrast to most miRNA precursors in other plants, the precursors in *Panicum* show great diversity. Although the lengths of the miRNA precursors ranged from 84 to 273 nt with an average of 161 nt, approximately 50% of them were 84–126 nt long (Figure 3). The different sizes of the candidate miRNAs within different families suggested that the different precursors might offer unique functions for the regulation of miRNA biogenesis or gene regulation [33]. Their minimal free energy (MFE) ranged from 43.20 to 118.6 kcal and their minimal free energy index (MFEI) ranged from 0.580 to 0.884, which is acceptable. Mature miRNAs are usually located on each arm of the stem-loop hairpin structure; some at the 5′ end of the miRNA precursor sequences, and others at the 3′ end. Among the 43 newly identified *Panicum* miRNAs, 22 were located at the 3′ end and 21 were at the 5′ end (Table 2).

Although all the miRNAs from all the species in the miRBase database were included in our dataset, it turned out that the candidate *Panicum* miRNAs that we identified were only those that were conserved in other plants including *Oryza sativa*, *Triticum aestivum*, *Sorghum bicolor*, *Selaginella moellendorffii*, *Solanum lycopersicum*, *Brassica napus*, *Vitis vinifera*, *Populus trichocarpa* and *Arabidopsis thaliana* (Table 3). Indeed, 36 of 43 candidate Panicum miRNAs showed full complementarity with their counterparts in other plants; the other seven miRNAs had mismatches, especially in the 3′ end. For example, three mismatches were found for miR159a, miR171, and miR1439; and two were found for miR169c and miR169d (Table 2).

Table 2 *Panicum* miRNAs identified by homolog search and secondary structure[a)]

miRNA	Mature sequences	PL	Gene number	Loc	ML	A%	C%	U%	G%	(A+U)%	MFE	MFEI
156a-1	UGACAGAAGAGAGUGAGCACA	98	gi198367980, gi198364716	5′	21	0.235	0.245	0.265	0.255	0.500	55.20	0.767
156a-2	UGACAGAAGAGAGUGAGCACA	84	gi198355601	5′	21	0.262	0.250	0.262	0.226	0.524	44.80	0.723
156a-3	UGACAGAAGAGAGUGAGCACA	84	gi198186083, gi169673860	5′	21	0.214	0.214	0.274	0.298	0.488	53.90	0.884
159a	UUGGACUGAAGGGUGCUCCC	174	gi198319846	3′	20	0.218	0.236	0.276	0.270	0.494	91.30	0.725
159b-1	UUUGGAUUGAAGGGAGCUCUG	257	gi198383357, gi198370413	3′	21	0.175	0.226	0.292	0.307	0.467	105.47	0.580
159b-2	UUUGGAUUGAAGGGAGCUCUG	254	gi198359837, gi198350925	3′	21	0.177	0.232	0.287	0.303	0.465	113.82	0.629
159b-3	UUUGGAUUGAAGGGAGCUCUG	248	gi198347813	3′	21	0.177	0.226	0.286	0.310	0.464	111.17	0.628
159b-4	UUUGGAUUGAAGGGAGCUCUG	259	gi198328585	3′	21	0.178	0.228	0.286	0.309	0.463	116.12	0.628
159b-5	UUUGGAUUGAAGGGAGCUCUG	267	gi198325973	3′	21	0.172	0.236	0.288	0.303	0.461	118.82	0.625
159b-6	UUUGGAUUGAAGGGAGCUCUG	259	gi198068022, gi197980325 gi169675227	3′	21	0.174	0.228	0.290	0.309	0.463	113.22	0.615
159b-7	UUUGGAUUGAAGGGAGCUCUG	273	gi198306925	3′	21	0.168	0.231	0.300	0.300	0.469	118.60	0.621
159b-8	UUUGGAUUGAAGGGAGCUCUG	263	gi198222304	3′	21	0.175	0.236	0.297	0.293	0.471	111.00	0.600
159b-9	UUUGGAUUGAAGGGAGCUCUG	242	gi198383357	3′	21	0.165	0.227	0.293	0.314	0.459	110.02	0.643
160-1	UGCCUGGCUCCCUGUAUGCCA	133	gi198096011	5′	21	0.195	0.278	0.203	0.323	0.398	66.00	0.623
160-2	UGCCUGGCUCCCUGUAUGCCA	128	gi169726518	5′	21	0.156	0.313	0.219	0.313	0.375	71.40	0.714
166a	UCGGACCAGGCUUCAUUCCCCU	117	gi198334060, gi198316752	3′	22	0.171	0.222	0.333	0.274	0.504	66.70	0.855
166b	UCGGACCAGGCUUCAUUCCCCC	189	gi198304610, gi198288637	3′	22	0.148	0.275	0.275	0.302	0.423	101.11	0.738
167a-1	UGAAGCUGCCAGCAUGAUCUA	106	gi198325934	5′	21	0.217	0.255	0.236	0.292	0.453	62.40	0.770
167a-2	UGAAGCUGCCAGCAUGAUCUA	118	gi198143879	5′	21	0.212	0.271	0.229	0.288	0.441	71.10	0.781
167b	UGAAGCUGCCAGCAUGAUCUG	155	gi198318359	5′	21	0.219	0.200	0.290	0.290	0.510	77.34	0.703
169a-1	UAGCCAAGGAUGACUUGCCUG	209	gi198367771	5′	21	0.187	0.234	0.282	0.297	0.469	99.40	0.663
169a-2	UAGCCAAGGAUGACUUGCCUG	208	gi198355088	5′	21	0.183	0.274	0.269	0.274	0.451	107.70	0.709
169a-3	UAGCCAAGGAUGACUUGCCUG	202	gi198347767	5′	21	0.158	0.262	0.277	0.302	0.436	112.50	0.771
169a-4	UAGCCAAGGAUGACUUGCCUG	200	gi198165872	5′	21	0.160	0.270	0.260	0.310	0.420	111.10	0.751
169b-1	UAGCCAAGAAUGACUUGCCUA	175	gi198358934	5′	21	0.229	0.223	0.286	0.263	0.514	91.30	0.730
169b-2	UAGCCAAGAAUGACUUGCCUA	172	gi198341684	5′	21	0.233	0.233	0.285	0.250	0.517	85.60	0.696
169c	UAGUCAAGGAUGACUUGCCU	111	gi198355557	5′	20	0.153	0.243	0.279	0.324	0.432	56.50	0.706
169d	UAGCCAAGGAUGACUUGCUUA	111	gi198351922	5′	21	0.153	0.243	0.279	0.324	0.432	56.40	0.705
171	UUGAGCCGUGCCAAUAUCUUA	117	gi198307426	3′	21	0.205	0.222	0.308	0.265	0.513	51.70	0.638
399a	UGCCAAAGGAGAUUUGCCCGG	104	gi169685067	3′	21	0.183	0.337	0.163	0.317	0.346	64.70	0.744
399b	UGCCAAAGGAGAUUUGCCCAG	166	gi198305723	3′	21	0.235	0.205	0.319	0.241	0.554	65.00	0.575
444a-1	UGCAGUUGUUGCCUCAAGCUU	123	gi197948735	3′	21	0.276	0.228	0.260	0.236	0.537	72.11	0.792
444a-2	UGCAGUUGUUGCCUCAAGCUU	123	gi197948735	3′	21	0.276	0.228	0.260	0.236	0.537	72.11	0.792
444a-3	UGCAGUUGUUGCCUCAAGCUU	116	gi198062967	3′	21	0.267	0.216	0.267	0.250	0.534	70.90	0.834
444b-1	UGUUGUCUCAAGCUUGCUGCC	126	gi198322676, gi198321179 gi198317150, gi198295730	3′	21	0.222	0.222	0.302	0.254	0.524	74.30	0.844
444b-2	UGUUGUCUCAAGCUUGCUGCC	126	gi198289536	3′	21	0.230	0.222	0.302	0.246	0.532	70.40	0.800
444b-3	UGUUGUCUCAAGCUUGCUGCC	126	gi59871869	3′	21	0.230	0.230	0.302	0.238	0.532	72.20	0.820
528a-1	UGGAAGGGGCAUGCAGAGGAG	90	gi198385159, gi198384503	5′	21	0.156	0.256	0.256	0.333	0.411	43.50	0.649
528a-2	UGGAAGGGGCAUGCAGAGGAG	90	gi198369163, gi198356093 gi198349132, gi198336818 gi198307273, gi169725213	5′	21	0.156	0.267	0.244	0.333	0.400	46.00	0.676
528a-3	UGGAAGGGGCAUGCAGAGGAG	88	gi198349905	3′	21	0.159	0.273	0.239	0.330	0.398	44.90	0.670
528a-4	UGGAAGGGGCAUGCAGAGGAG	90	gi198326853, gi198316246 gi198299290, gi198287745	5′	21	0.156	0.267	0.256	0.322	0.411	44.60	0.666
528b	UGGAAGGGGCAUGCAGAGGGG	90	gi198382348	5′	21	0.144	0.256	0.256	0.344	0.400	43.20	0.645
1439	AUUUGGAACGGAGUGAGUACA	263	gi198019034	3′	21	0.350	0.114	0.422	0.114	0.772	105.02	0.691

a) PL, Length of precursor sequence; Loc, location of mature miRNA in the precursor; ML, length of mature sequence; MFE, minimum free energy; MFEI, minimum free energy index. The shaded letters indicate nucleotide mismatches compared with the corresponding sequences in miRBase.

Because miRNAs are believed to occur at a frequency of approximately 0.1% of the total EST sequences of a species [33], the 436736 *Panicum* ESTs in the dataset should contain 43.67 miRNAs. This number is very close to our results, suggesting that the parameters we chose in the prediction process are reasonable and the methods we employed are proper, and, more importantly, the prediction of miRNAs based on ESTs was effective.

2.2 Prediction and analysis of *Panicum* miRNA targets

In this study, we identified a total of 68 potential targets for

Table 3 Total number of miRNAs and their numbers in the corresponding families in *Panicum* and some other plant species

Species	Total number	Family	The 11 miRNAs families identified in *Panicum* and the number of miRNAs in each family										
			156	159	160	166	167	169	171	399	444	528	1439
P. milaceum L.	43	11	3	10	2	2	3	8	1	2	6	5	1
Oryza sativa	491	213	12	6	6	14	10	17	9	11	6	1	1
Triticum aestivum	32	31	0	2	1	0	1	0	1	1	1	0	0
Sorghum bicolor	148	27	9	2	6	11	9	17	11	11	0	1	0
Zea Mays	170	28	12	11	7	14	10	18	14	10	0	2	0
Millet	43	21	0	0	0	0	0	0	0	0	0	0	0
Arabidopsis thaliana	328	251	16	6	6	14	10	29	7	15	0	0	0
Selaginella moellendorffii	58	45	4	1	2	3	0	0	4	0	0	0	0
Solanum lycopersicum	36	23	3	1	1	2	1	4	4	1	0	0	0
Brassica napus	46	17	3	1	0	4	3	13	7	1	0	0	0
Vitis vinifera	163	49	9	3	5	8	5	25	9	9	0	0	0
Populus trichocarpa	233	42	10	6	8	17	8	32	14	12	0	0	0

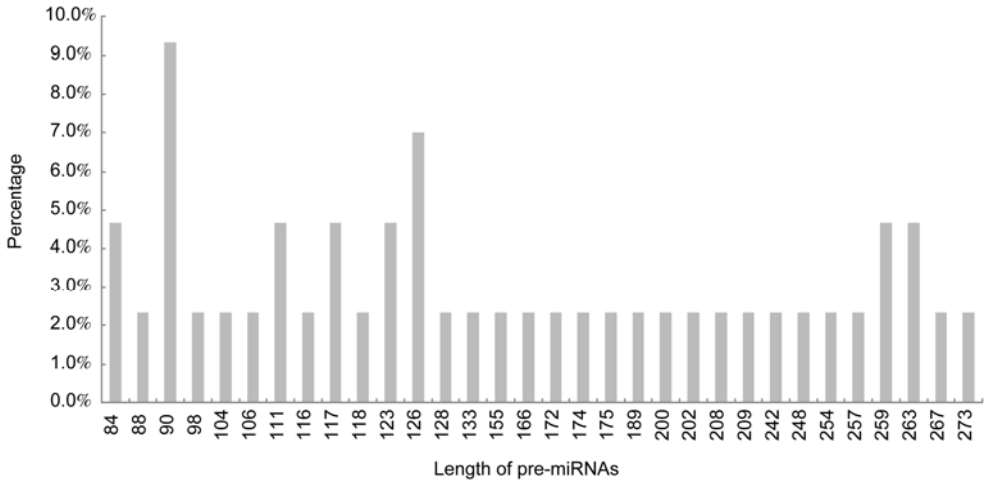

Figure 3 Length distribution of the predicted precursor miRNAs.

the 11 identified miRNA families in *Panicum* based on the fact that miRNAs show perfect or near-perfect complementary to their target mRNA sequences (Table 4). The 68 potential miRNA targets belong to several gene families and have different biological functions that include RNA regulation, metabolism, stress response, RNA processing, signal transduction and transportation.

It is generally accepted that almost all miRNA targets have no more than four mismatches with their corresponding miRNAs in plants [48], and this criterion has been widely adopted to identify miRNA targets in different plant species. In our study, we allowed no more than three mismatches between miRNAs and their corresponding targets and included no more than one mismatch between 2 and 8 nt of the mature miRNAs as a standard.

Among the 11 identified miRNA families, some are known to be involved in multiple functions (for example, miR166, miR399, and miR1439) while others have unique functions (for example, miR171, miR156, and miR160). All the identified families except miR528 are associated with RNA regulation. For example, miR156 targets the Squamosa-promoter binding protein which is a trans-acting factor that binds specifically to the consensus nucleotide sequence 5'-TNCGTACAA-3' [49,50]; and miR159 targets a MYB transcription factor which is a transcription activator [51]. When associated with the BHLH12/MYC1, EGL3, or GL3 transcription factor, MYB protein can promote the synthesis of phenylpropanoid-derived compounds such as anthocyanins and proanthocyanidin [52]. miR160 targets auxin response factors, which are a group of transcriptional factors that bind specifically to the DNA sequence 5'-TGTCTC-3' found in the auxin-responsive promoter elements (AuxREs) [53]. Formation of heterodimers with Aux/IAA proteins may alter their ability to modulate early auxin response genes expression [54]. miR166 targets the homeobox-leucine zipper protein which is a probable transcription factor involved in the regulation of meristem development to promote lateral organ formation and may regulate procambial and vascular tissue formation or maintenance, and vascular development in inflorescence stems [53–55]. Besides these, miR171 targets the F6N15.20 protein, zinc finger (CCCH-type) protein-like, and the Scl1 protein, miR444

Table 4 Predicted targets of the candidate miRNAs identified in *Panicum*

miRNA	Target protein	Target function	Target gene
156a	Squamosa promoter-binding-like protein	Transcription regulation	TC7137 TC28500 TC24952 TC13039 TC36542 TC19607 TC15006 TC42015
	Teosinte glume architecture	Structural protein	TC52769
	SBP transcription factor	Transcription regulation	TC168 TC17992 TC41889
	Diacylglycerol kinase-like	Signal transduction	TC50996
	Methyladenine glycosylase protein-like	DNA processing	TC2901
	Chromosome chr12 scaffold_78, whole genome shotgun sequence		TC34681
159a	Transcription factor GAMYB	Transcription regulation	TC4716
	Inositol 1,3,4-trisphosphate 5/6-kinase	Signal transduction	TC1184
159b	Transcription factor GAMYB	Transcription regulation	TC4716
	Inositol 1,3,4-trisphosphate 5/6-kinase	Signal transduction	TC1184
160	Auxin response factor	Signal transduction	TC52413 TC20541 TC3798
166a	Class III HD-Zip protein	Transcription regulation	TC25111
	Rolled leaf	Unknown	TC40382
	Chromosome chr9 scaffold_7, whole genome shotgun sequence		FL705256
	MJ0042 family finger-like protein	Transcription regulation	FL805767
166b	Class III HD-Zip protein	Transcription regulation	TC25111
	Rolled leaf	Unknown	TC40382
	Chromosome chr9 scaffold_7, whole genome shotgun sequence		FL705256
	MJ0042 family finger-like protein	Transcription regulation	FL805767
167a	Chromosome chr8 scaffold_68, whole genome shotgun sequence		TC49608 TC7980 TC11656 TC38375 TC12105 TC11776 TC32239
167b	Os05g0109600 protein	Transcription regulation	TC23301
	Expressed protein	Unknown	TC39870
	Type I secretion outer membrane protein, TolC family precursor	Signal transduction	FL707871
169a	CCAAT-binding transcription factor subunit B family protein	Transcription regulation	TC32992 TC41994 TC33317 TC50964 TC5643
	Os02g0776400 protein	Transcription regulation	TC17857 TC48466 TC31225
	RAPB protein	Transcription regulation	TC39441
	HAP2 subunit of HAP complex	Structural protein	TC6534 TC13896
	CCAAT-box transcription factor complex WHAP6	Transcription regulation	TC3877
	Os02g0776400 protein	Transcription regulation	TC33524
169b	CCAAT-binding transcription factor subunit B family protein	Transcription regulation	TC32992 TC41994 TC33317 TC50964 TC5643
	Os02g0776400 protein	Transcription regulation	TC17857 TC48466 TC31225
	XAP5 family protein		TC32011
169c	CCAAT-binding transcription factor subunit B family protein	Transcription regulation	TC32992 TC41994 TC33317 TC50964 TC5643
	Os05g0363500 protein	Transcription regulation	FL727055
	Os02g0776400 protein	Transcription regulation	TC17857 TC48466 TC31225
169d	CCAAT-binding transcription factor subunit B family protein	Transcription regulation	TC32992 TC41994 TC33317 TC50964 TC5643
	Predicted protein	Unknown	TC2384
	ABL167Cp	Unknown	TC35322
	Unknow protein	Transcription regulation	TC43271 TC31296
	Os02g0776400 protein	Transcription regulation	TC17857 TC48466 TC31225
	NADH dehydrogenase I chain M	Metabolism	TC9084
	Methionine S-methyltransferase	Transcription regulation	TC15884
171	Scl1 protein	Unknown	FL816450 TC26013
	Os05g0562400 protein	Transcription regulation	FL708626
399a	Expressed protein	Unknown	FE627710
	Tetratricopeptide TPR_2 repeat protein precursor	Transcription regulation	FL689938
	Salt-induced AAA-Type ATPase	Signal transduction	TC26446
399b	Expressed protein	Unknown	FE627710

(*To be continued on the next page*)

(*Continued*)

miRNA	Target protein	Target function	Target gene
	Tetratricopeptide TPR_2 repeat protein precursor	Transcription regulation	FL689938
	Chromosome chr12 scaffold_47		FL802301
444a	MIKC-type MADS-box transcription factor WM30	Transcription regulation	FL808850
	MADS box protein 2	Transcription regulation	TC45658
	BHLH transcription factor PTF1	Transcription factor	TC18118
	Zinc finger (C3HC4-type RING finger) protein-like	Transcription regulation	TC16101
	Non-ribosomal peptide synthetase modules and related proteins		TC4649
	Chromosome chr10 scaffold_179, whole genome shotgun sequence		TC27817
	Chromosome chr7 scaffold_31, whole genome shotgun sequence		TC18635
	Os05g0470700 protein	Transcription regulation	TC41751
444b	MADS box protein 2	Transcription regulation	TC45658
	MIKC-type MADS-box transcription factor WM30	Transcription regulation	FL808850
	Peptidyl-prolyl cis-trans isomerase	Signal transduction	TC1803
	VacJ-like lipoprotein precursor	Transcription regulation	DN148537
528a	Os02g0671800 protein	Transcription regulation	GD041833
	Superoxide dismutase [Cu-Zn]	Transcription regulation	TC8663 TC11635 TC51355
	Chromosome chr5 scaffold_2		TC24045 TC52797 TC22262
	Glyoxalase/Bleomycin resistance protein/dioxygenase domain	Signal transduction	FL746289
	Plasma membrane ATPase 2	Metabolism	FL751986
	N.plumbaginifolia H+-translocating ATPase mRNA	Metabolism	FL815815
528b	Predicted protein	Unknown	GD018502
	Chromosome chr18 scaffold_1, whole genome shotgun sequence		TC49585
1439	DNA (cytosine-5)-methyltransferase 1	DNA processing	TC42883
	50S ribosomal protein L4	Structural protein	TC36575
	Chromosome undetermined scaffold_109, whole genome shotgun sequence		FL703180
	Predicted protein	Unknown	TC24744
	Chromosome undetermined scaffold_77, whole genome shotgun sequence		TC18576
	Chromosome chr6 scaffold_15, whole genome shotgun sequence		TC26456
	Glycine C-acetyltransferase	Signal transduction	FE604006
	FYVE finger-containing phosphoinositide kinase-like	Signal transduction	TC34589
	DNA methylase N-4/N-6 domain protein	DNA processing	FL855194
	Vomeronasal receptor V1RI2	Signal transduction	TC52734
	Os03g0351300 protein	Transcription regulation	TC1806

targets the B1358B12 protein, T6H22.8.1 protein, MADS-box transcription factor, and miR1439 targets the winged helix transcription factor. All these genes are involved in RNA regulation, supporting the idea that miRNAs function in post-transcriptional regulation.

We also found that five miRNA families participate in metabolism. For example, miR166 targets genomic DNA, chromosome 5, TAC clone: K21P3 which has beta-galactosidase activity and participates in the carbohydrate metabolic process [56]; miR169 targets sulfur oxidation protein which is an electron carrier that functions in heme binding and iron ion binding processes; miR528 targets syringolide-induced protein B13-1-1; and miR444 targets the putative P450 hydroxylase which is involved in oxidation reduction and functions as an electron carrier. *P. miliaceum* L. is widely planted in Northern China, where the natural environment is severe. It will be interesting to explore the idea that some of the candidate miRNAs may play roles in the ability of *Panicum* to adapt to environmental stress. One of the candidate miRNAs belonged to the miR1439 family that is involved in signal transduction.

The identified *P. miliaceum* L. miRNAs also target genes that control RNA transport and RNA processing. AT3g25150/MJL12_9 protein, the target of miR399, has nucleic acid binding activity and is thought to be involved in RNA transport. miR1439 targets the tRNA (guanine-

N(1)-)-methyltransferase that is associated with tRNA processing and modification, supporting previous reports that small RNAs are involved in mediating methylation in plants.

2.3 Northern blot analysis

Eight of the 12 selected miRNAs had positive signals with lengths around 21 nt. The miRNAs from three of the families, miR399, miR444, miR1439 failed the verification (Figure 4).

The miR399 family has been reported to target a putative ubiquitin conjugating enzyme (UBC), and miR399 is highly induced in low-phosphate stress conditions [57,58]. This may explain why we did not see a positive signal for miR399 under normal growth conditions. However, miR399 may contribute to the high resistance to stress of *Panicum*.

miR444 targets a MAD box transcription factor that is expressed during inflorescence and fruit development. Therefore, it is possible that miR444 is highly expressed only in the reproductive growth stage [59] and this may be the reason for the high yield of *Panicum* in arid regions.

The function of miR1439 is still not clear; however, high-throughput sequencing carried out by another group indicated that miR1439 might be related to salt-tolerance in rice [60]. Although miR1439 was found at very low frequency (2/80990) in a salt stress library and was not present in the control library, it did have a positive signal in a Northern blot assay in both the salt stress and control libraries [60]. However, perhaps because of the normal growth conditions that we used and/or the species variation, miR1439 was not detected in our Northern blot assay.

The negative controls (probes 5S rRNA, EST1, and EST2) that we used produced no detectable signal. Thus, the Northern blot assay confirmed the effectiveness of our prediction.

3 Conclusion

By searching the publicly available *Panicum* ESTs, 43 potential miRNAs together with their potential targets were identified. 83.7% of the candidate miRNAs are identical to the corresponding sequences in other species (Table 2); however, the secondary structures of their precursors are unique (Figure 5). Many miRNA genes (for example, miR528a-2, miR444b-1, and miR166a) are known to have multiple gene loci and, in our study, they were found in EST sequences with different gene numbers; however, some of these ESTs may turn out to be identical when a better assembly becomes available. A total of 68 potential targets were identified and we found that most of the genes are involved in RNA regulation, suggesting their essential role in the post-transcriptional regulation of biological processes in *P. miliaceum* L. Eight miRNAs out of the 12 that were selected were verified by Northern blot hybridizations, indicating that the method we used to predict miRNA was effective. *Panicum miliaceum* L. is widely grown in northern China, especially in the Loess Plateau. The plant has high nutritional value and strong drought resistance. We expect that an understanding of the role of miRNAs in the metabolic drought resistant mechanisms in the plant will help explain the process from a post-transcriptional level. This study is a start towards this goal. In future studies, we will investigate the function of the candidate miRNAs that were verified by Northern blot more thoroughly through overexpression and RNAi.

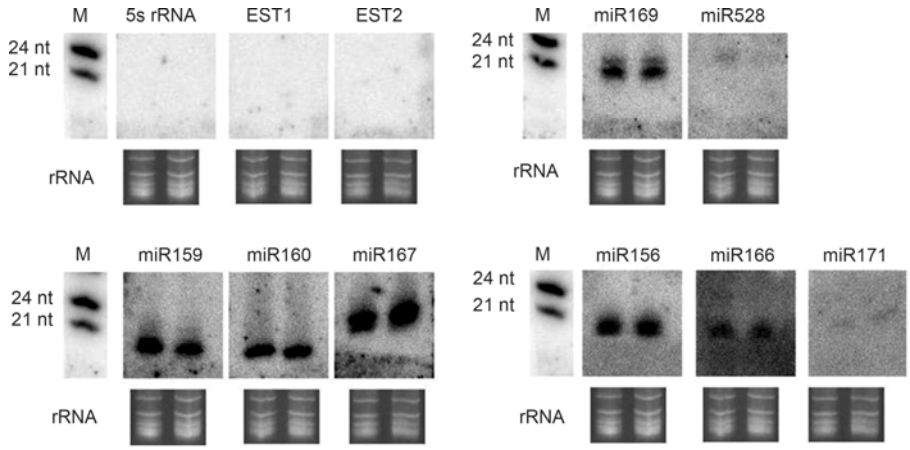

Figure 4 Northern blot hybridization assays. M, RNA marker with 3′-methyl group. rRNA, total RNA sample volume in the different lanes. The name of the miRNA is at the top of each panel. The left lane of each membrane is the Longmi-5 cultivar, the right lane is the Ningmi-9 cultivar. 5s rRNA, EST1, and EST2 are the controls that were used to monitor the effect caused by RNA degradation. No signals were detected from the controls, indicating that degradation had little effect on the results. The miR156, miR159, miR160, miR166, miR167, miR169, miR171, and miR528 candidate miRNAs produced hybrid signals that were located around 21 nt. No detectable signals were seen for miR399, miR444a, miR444b, or miR1439.

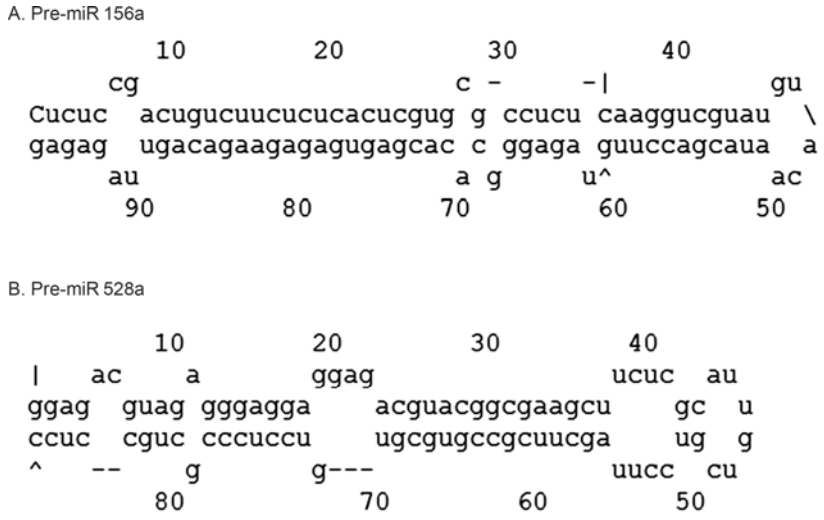

Figure 5 Representative secondary structures of the *Panicum* candidate miRNAs. A, Stem-loop structure of miR156a-1. B, Stem-loop structure of miR528a-3. The actual lengths of the pre-miRNAs may be slightly shorter or longer than shown.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 30600384 and 50809068), and the Young Outstanding Scholar Foundation of Northwest A & F University, China. We thank Wang QinHu for insightful suggestions on some particulars of the study. We are grateful to Professor Hu YinGang for providing the Panicum seeds. We also thank Deng XiLian for access to the high performance computers in the Network Information Center, Northwest A & F University, China.

1 Murchison E P, Hannon G J. MiRNAs on the move: miRNA biogenesis and the RNAi machinery. Curr Opin Cell Biol, 2004, 16: 223–229
2 Lee R C, Feinbaum R L, Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell, 1993, 75: 843–854
3 Lagos-Quintana M, Rauhut R, Lendeckel W, *et al.* Identification of novel genes coding for small expressed RNAs. Science, 2001, 294: 853–858
4 Khvorova A, Reynolds A, Jayasena S D. Functional siRNAs and miRNAs exhibit strand bias. Cell, 2007, 131: 41–49
5 Carrington J C, Ambros V. Role of microRNAs in plant and animal development. Science, 2003, 301: 336–338
6 Zhang B H, Wang Q L, Wang K B, *et al.* Identification of cotton microRNAs and their targets. Gene, 2007, 397: 26–37
7 Zhang B H, Pan X P, Anderson T A. MicroRNA: a new player in stem cells. J Cell Physiol, 2006, 209: 266–269
8 Zhang B H, Pan X P, Cobb G P, *et al.* Plant microRNA: a small regulatory molecule with big impact. Dev Biol, 2006, 289: 3–16
9 Jones-Rhoades M W, Bartel D P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell, 2004, 14: 787–799
10 Saini H K, Griffiths-Jones S, Enright A J. Genomic analysis of human microRNA transcripts. Proc Natl Acad Sci USA, 2007, 104: 17719–17724
11 Fujita S, Iba H. Putative promoter regions of miRNA genes involved in evolutionarily conserved regulatory systems among vertebrates. Bioinformatics, 2008, 24: 303–308
12 Zuloaga F O, Giussani L M, Morrone O. On the taxonomic position of Panicum *aristellum* (Poaceae: Panicoideae: Paniceae). Syst Bot, 2006, 31: 497–505
13 Peiris H, Ibrahim M N M. Nutritive value of guinea grass (*Panicum maximum* Jacq.) and urea supplemented rice straw for cattle. Asian Austral J Anim Sciences, 1995, 8: 83–88
14 Llave C, Kasschau K D, Rector M A, *et al.* Endogenous and silencing-associated small RNAs in plants. Plant Cell, 2002, 14: 1605–1619
15 Grad Y, Aach J, Hayes G D, *et al.* Computational and experimental identification of *C. elegans* microRNAs. Mol Cell, 2003, 11: 1253–1263
16 Lau N C, Lim L P, Weinstein E G, *et al.* An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans.* Science, 2001, 294: 858–862
17 Lee R C, Ambros V. An extensive class of small RNAs in Caenorhabditis *elegans.* Science, 2001, 294: 862–864
18 Mourelatos Z, Dostie J, Paushkin S, *et al.* miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Gene Dev, 2002, 16: 720–728
19 Park W, Li J J, Song R T, *et al.* CARPEL factory, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol, 2002, 12: 1484–1495
20 Reinhart B J, Weinstein E G, Rhoades M W, *et al.* MicroRNAs in plants. Gene Dev, 2002, 16: 1616–1626
21 Yoon S, De Micheli G. Computational identification of microRNAs and their targets. Birth Defects Res, 2006, 78: 118–128
22 Wang X J, Reyes J L, Chua N H, *et al.* Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. Genome Biol, 2004, 5: R65
23 Adai A, Johnson C, Mlotshwa S, *et al.* Computational prediction of miRNAs in *Arabidopsis thaliana*. Genome Res, 2005, 15: 78–91
24 Zhang B H, Pan X P, Wang Q L, *et al.* Identification and characterization of new plant microRNAs using EST analysis. Cell Res, 2005, 15: 336–360
25 Lai E C, Tomancak P, Williams R W, *et al.* Computational identification of *Drosophila* microRNA genes. Genome Biol, 2003, 4: R42
26 Sunkar R, Jagadeeswaran G. *In silico* identification of conserved microRNAs in large number of diverse plant species. BMC Plant Biol, 2008, 8: 37
27 Altinkut-Uncuoglu A, Ercan S, Ertugrul F, *et al.* An EST-SSR marker linked with yellow rust resistance in wheat. Biol Plantarum, 2010, 54: 691–696
28 Cupertino F B, Leal J B, Correa R X, *et al.* Genetic diversity of Eucalyptus hybrids estimated by genomic and EST microsatellite markers. Biol Plantarum, 2011, 55: 379–382
29 Zheng T G, Qiu W M, Fan G E, *et al.* Construction and characterization of a cDNA library from floral organs and fruitlets of Citrus reticulata. Biol Plantarum, 2011, 55: 431–436

30 Rajewsky N, Socci N D. Computational identification of microRNA targets. Dev Biol, 2004, 267: 529–535
31 Arteaga-Vazquez M, Caballero-Perez J, Vielle-Calzada J P. A family of microRNAs present in plants and animals. Plant Cell, 2006, 18: 3355–3369
32 Altschul S F, Gish W, Miller W, *et al.* Basic local alignment search tool. J Mol Biol, 1990, 215: 403–410
33 Zhang B, Pan X, Cannon C H, *et al.* Conservation and divergence of plant microRNA genes. Plant J, 2006, 46: 243–259
34 Dezulian T, Remmert M, Palatnik J F, *et al.* Identification of plant microRNA homologs. Bioinformatics, 2006, 22: 359–360
35 Zhu J K. Reconstituting plant miRNA biogenesis. Proc Natl Acad Sci USA, 2008, 105: 9851–9852
36 Steffen P, Voss B, Rehmsmeier M, *et al.* RNAshapes: an integrated RNA analysis package based on abstract shapes. Bioinformatics, 2006, 22: 500–503
37 Reeder J, Giegerich R. Consensus shapes: an alternative to the Sankoff algorithm for RNA consensus structure prediction. Bioinformatics, 2005, 21: 3516–3523
38 Giegerich R, Voss B,Rehmsmeier M. Abstract shapes of RNA. Nucleic Acids Res, 2004, 32: 4843-4851
39 Zhang B H, Pan X P, Cox S B, *et al.* Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci, 2006, 63: 246–254
40 Dai X B, Zhao P X. psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res, 2011, 39: W155–159
41 Hornstein E, Mansfield J H, Yekta S, *et al.* The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. Nature, 2005, 438: 671–674
42 Yamada-Akiyama H, Akiyama Y, Ebina M, *et al.* Analysis of expressed sequence tags in apomictic guineagrass (*Panicum maximum*). J Plant Physiol, 2009, 166: 750–761
43 Cerda J, Mercade J, Lozano J J, *et al.* Genomic resources for a commercial flatfish, the Senegalese sole (*Solea senegalensis*): EST sequencing, oligo microarray design, and development of the bioinformatic platform Soleamold. BMC Genomics, 2008, 9: 508
44 Kuster H, Becker A, Firnhaber C, *et al.* Development of bioinformatic tools to support EST-sequencing, *in silico*- and microarray-based transcriptome profiling in *mycorrhizal symbioses*. Phytochemistry, 2007, 68: 19–32
45 Li L, Brunk B P, Kissinger J C, *et al.* Gene discovery in the apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. Genome Res, 2003, 13: 443–454
46 Paquola A C, Nishyiama M Y Jr., Reis E M, *et al.* ESTWeb: bioinformatics services for EST sequencing projects. Bioinformatics, 2003, 19: 1587–1588
47 Wang J P, Lindsay B G, Cui L, *et al.* Gene capture prediction and overlap estimation in EST sequencing from one or multiple libraries. BMC Bioinformatics, 2005, 13: 300
48 Rhoades M W, Reinhart B J, Lim L P, *et al.* Prediction of plant microRNA targets. Cell, 2002, 110: 513–520
49 Cardon G H, Hohmann S, Nettesheim K, *et al.* Functional analysis of the *Arabidopsis thaliana* SBP-box gene SPL3: a novel gene involved in the floral transition. Plant J, 1997, 12: 367–737
50 Cardon G, Hohmann S, Klein J, *et al.* Molecular characterisation of the Arabidopsis SBP-box genes. Gene, 1999, 237: 91–104
51 Kranz H D, Denekamp M, Greco R, *et al.* Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis* thaliana. Plant J, 1998, 16: 263–276
52 Stracke R, Werber M,Weisshaar B. The R2R3-MYB gene family in Arabidopsis *thaliana*. Curr Opin Plant Biol, 2001, 4: 447–456
53 Theologis A, Ecker J R, Palm C J, *et al.* Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*. Nature, 2000, 408: 816–820
54 Yamada K, Lim J, Dale J M, *et al.* Empirical analysis of transcriptional activity in the *Arabidopsis* genome. Science, 2003, 302: 842–846
55 Green K A, Prigge M J, Katzman R B, *et al.* CORONA, a member of the class III homeodomain leucine zipper gene family in *Arabidopsis*, regulates stem cell specification and organogenesis. Plant Cell, 2005, 17: 691–704
56 Asamizu E, Sato S, Kaneko T, *et al.* Structural analysis of *Arabidopsis thaliana* chromosome 5. VIII. Sequence features of the regions of 1081958 bp covered by seventeen physically assigned P1 and TAC clones. DNA Res, 1998, 5: 379–391
57 Sunkar R, Zhu J K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell, 2004, 16: 2001–2019
58 Fujii H, Chiou T J, Lin S I, *et al.* A miRNA involved in phosphate-starvation response in *Arabidopsis*. Curr Biol, 2005, 15: 2038–2043
59 Sunkar R, Girke T, Jain P K, *et al.* Cloning and characterization of microRNAs from rice. Plant Cell, 2005, 17: 1397–1411
60 Sunkar R, Zhou X, Zheng Y, *et al.* Identification of novel and candidate miRNAs in rice by high throughput sequencing. BMC Plant Biol, 2008, 8: 25