Plant Signaling & Behavior 7:2, 246–259; February 2012; © 2012 Landes Bioscience

Identification of miRNAs in sorghum by using bioinformatics approach

Amit Katiyar,[1,†] Shuchi Smita,[1,†] Viswanathan Chinnusamy,[2] Dev Mani Pandey[3] and Kailash Chander Bansal[1,†,*]

[1]National Research Centre on Plant Biotechnology; Indian Agricultural Research Institute Campus; New Delhi, India; [2]Division of Plant Physiology; Indian Agricultural Research Institute; New Delhi, India; [3]Department of Biotechnology; Birla Institute of Technology; Mesra; Ranchi; Jharkhand, India

[†]Current affiliation: National Bureau of Plant Genetic Resources; Indian Agricultural Research Institute Campus; New Delhi, India

Keywords: sorghum, miRNA, target genes, miRNA cluster

Abbreviations: GSS, genomic survey sequences; EST, expressed sequenced tag; WGS, whole genome sequence; KEGG, Kyoto encyclopedia of genes and genomes; COG, clusters of orthologous groups of proteins; GO, gene ontology; MFEI, minimal folding free energy index

MicroRNAs (miRNAs) regulate gene expression mainly by post-transcriptional gene silencing (PTGS) and in some cases by transcriptional genes silencing (TGS). miRNAs play critical roles in developmental processes, nutrient homeostasis, abiotic stress and pathogen responses of plants. In contrast to the large number of miRNAs predicted in cereal model plant rice, only 148 miRNAs were predicted in sorghum till date (miRBase release 17). This suggested that miRNAs identified in sorghum is far from saturation. Hence, we developed a bioinformatics pipeline using an in-house PERL script and publicly available structure prediction tools to identify miRNAs and their target genes from publically available Expressed Sequence Tags (EST) and Genomic Survey Sequence (GSS). About 1,379 known and unique plant miRNAs from 33 different crops were used to predict new miRNAs in sorghum. We identified 31 new miRNAs belonging to 10 different miRNA families. We predicted 72 potential target genes for 31 miRNAs, and most of these target genes are predicted to be involved in plant growth and development. These newly identified miRNAs add to the growing database of miRNA and lay the foundation for further understanding of miRNA function in sorghum plant development.

Introduction

Sorghum (*Sorghum bicolor* L.) is an important cereal crop that is highly resistant to drought and heat stress, and is used for food, fodder, and as a raw materials for the production of starch, alcohol and biofuels.[1–3] The extensive agricultural use of sorghum and the emerging demand of sorghum for biofuel production necessitates development of cultivars with higher yields, altered stem reserves, and improved resistance to biotic and abiotic stresses. MicroRNAs (miRNAs) play important roles in development, nutrient acquisition and use, and tolerance to abiotic and biotic stresses.[4] miRNAs are small non-coding RNAs of approximately 21 nucleotides (nt) in length that act mainly in PTGS and in some cases TGS to regulate the expression of their target genes. miRNAsare widespread in all eukaryotes including unicellular green alga.[5]The genes encoding miRNAs, *MIR* genes, are transcribed by RNA polymerase II to produce a primary transcript (pri-miRNA). The stem-loop structure of the pri-miRNA is processed Dicer-Like (DCL) RNase III enzymes (mainly DCL1 in Arabidopsis) to produce ~21 nucleotide long miRNA-miRNA* duplex. The DCL1 catalyzed processing of pri-miRNAs to pre-miRNAs also requires two additional dsRNA-binding proteins namely HYPONASTIC LEAVES1 (HYL1) and SERRATE (SE) in Arabidopsis.[6,7] The HEN1, a methyltransferase, catalyzes 2'-O-methylation of the 3' termini nucleotide in the miRNA-miRNA* duplex. The miRNAs are exported in to cytosol with the help of HASTY (exportin 5). In cytosol, miRNA is loaded into AGO1 containing RISC which catalyzes PTGS.[8,9] Plant miRNAs negatively regulate the transcripts levels of their target genes, and play important roles in plant growth, organ development, cell differentiation and proliferation, cell death, signal transduction, and stress response.[8–11]Identification of miRNAs and their target genes therefore is an important step toward understanding the biological functions of miRNAs. Recently, computational approaches are used wildly as rapid, accurate, and affordable method to identify miRNAs.The computational approaches have been very effective in plants, where miRNA and its target mRNA have often nearly perfectly complementary.[12–14] The earliest miRNAs from plant kingdom were discovered in *Arabidopsis thaliana* in 2002,[15,16] and subsequent miRNAs have been identified in several plants by computational and experimental approaches.[8,17] Conserved nature of mature miRNAs among different species and the unique secondary structure of pri-miRNAs,[15,16,18–21] facilitate miRNA prediction using bioinformatics approaches.[22]

*Corresponding author: Kailash Chander Bansal; Email: kcbansal@nbpgr.ernet.in
Submitted: 08/08/11; Revised: 11/30/11; Accepted: 12/01/11
http://dx.doi.org/10.4161/psb.18914

A comparative genomics approach for the prediction of novel miRNAs and their targets was developed by Jones-Rhoades and Bartel.[13] Identification of several miRNAs from Arabidopsis,[23] rice,[24] corn,[25] cotton,[26] *Medicago truncatula*,[27] soybean,[4] citrus,[28] mustard,[29] wheat,[30] potato,[31] tomato,[32] switchgrass,[33] and sorghum[34] by computational approaches have been reported. The database of Genomic Survey Sequences (GSS) and Expressed Sequenced Tag (EST) are the major resources for identification of miRNAs in most of the plants. Using this approach, more than 700 miRNAs have been identified in plants.[25,35,36] Most research groups prefer to use expressed sequence tags (ESTs) over genome sequence as ESTs provide direct evidence for miRNA expression.[37,38] Several homology based tools (e.g., MIRcheck;[13] miRU[39]) are available for identification of potential miRNA target genes. The microRNAs registry database[40,41] (Release 17) include 232, 491, 234 and 170 miRNAs from *Arabidopsis thaliana*,[12,42–45] *Oryza sativa*,[24,46–49] *Populus trichocarpa*,[50,51] and *Zea mays*,[52] respectively. The number of miRNAs reported in sorghum in miRbase version-17 is only 148. This indicates potential for identification of additional miRNAs in sorghum as currently only less number of miRNAs are reported in sorghum as compared with other plant species. In this study, we used computational pipeline to predict novel miRNAs and their target genes in sorghum. Researchers can further validate these newly predicted miRNAs by using direct sequencing of small RNA libraries or by northern blotting.[53] Validation of functions these miRNAs in sorghum will help understand development and stress responses of sorghum.

Results

Prediction of miRNAs. Plant miRNAs exhibit high degree of conservation within plant kingdom.[54,55] Hence, known miRNAs from one plant species can be used to identify the conserved miRNAs in target species. A total of 2,728 plant miRNAs belong to 33 different plant species were downloaded from microRNA repository miRBase, version 17. From this data set, we omitted previously reported sorghum miRNAs to avoid prediction of previously identified sorghum miRNAs. Multiple sequence alignment was performed to eliminate miRNA with same sequence, and finally we obtained 1,379 non-redundant reference miRNAs belonging to 643 different miRNAs families. These reference miRNAs were used to identify miRNAs from ESTs (240161 sequence) and GSS (799,504 sequences) in sorghum by using an in-house PERL script. Sorghum GSS and ESTs that perfectly matched with reference miRNAs were considered as possible precursors of miRNAs (pri-miRNAs) in this study. Initially, we identified 375 pri-miRNAs (326 from GSS and 51 from ESTs).

Minimization of false positives. Removal of false positive is an essential step in computational prediction of miRNAs. We applied a number of initial filters to new potential pri-miRNAs as suggested by Zhang et al.[56] and Bonnet et al.[57] We aligned initially predicted 375 miRNAs with previously reported 148 miRNAs in sorghum (miRBase, version 17) and eliminated if found. To analyze, whether the predicted putative pri-miRNAs are non protein coding RNAs, BLASTX was performed against NCBI non-redundant protein database and excluded if putative pri-miRNAs with protein coding potential. The candidate miRNA precursors were also aligned with known non-coding RNAs such as tRNA, rRNA, snRNA or snoRNA and discarded if found similar. The candidate miRNA precursors were also aligned with plastid or mitochondrial genomes to eliminate precursors with similarity to these genomes. These primary filtering strategies reduced the number of predicted miRNA precursors in sorghum, and thus we obtained 33 and 53 valid new miRNA precursors from EST and GSS sequences, respectively.

Pri-miRNA structural filter. The 86 putative sorghum miRNA precursors were carefully examined to make sure that they qualify for the updated plant miRNA annotation criteria.[53,58] One important feature that distinguishes miRNAs from other endogenous small RNAs is that pri-miRNA transcript adopts a stem-loop structure and the miRNA is derived from the stem-arm. The miRNA precursors with 250 nt upstream and downstream to the mature miRNA sequence were analyzed for their ability to fold into a stem-loop hairpin structure using the RNAFold program[59,60] and those that fulfilled the hairpin structure criteria described by Jones-Rhoades et al.,[8] were selected as potential candidate precursors for miRNA. Among the 86 putative miRNA precursors screened, only 54 passed initial filters for positional overlaps, secondary structure and orientation of mature miRNA sequence within the respective stem-loop structures. As a result, a total of 54 new miRNAs, nine from EST and 45 from GSS sequence, were predicted in the stem-arm of the stem-loop hairpin structures. All 54 sorghum miRNAs were considered as valid candidates after satisfying the empirical formula for biogenesis and expression of the miRNAs as suggested by Ambroset al.[58] Additionally, we mapped all these 54 miRNA precursors on sorghum genome to eliminate overlapping miRNAs. As a result, we obtained 31 new miRNA precursors mapped to unique genome loci and listed in **Table 1**. The predicted miRNA precursor sequences and hairpin structures are shown in **Figure S1**.

Sorghum miRNAs family. The newly identified miRNAs in this study were assigned to different families of miRNA in sorghum. The family assignment was based on sequence similarity between newly predicted miRNAs and already known miRNAs in other plants including known sorghum miRNAs in the miRBase. To determine the sequence similarity, multiple sequence alignment was performed by using ClustalW[61] and based on cluster analysis (data not shown), predicted 31 miRNAs were assigned to respective MIR family. In this study, 31 newly identified miRNAs were assigned to 10 diverse MIR families namely miR156, miR166, miR167, miR168, miR171, miR390, miR396, miR398, miR399 and miR444 in sorghum (**Fig. 1**). Typically each *MIR* loci produced single precursor but a few *MIR* loci produced two or more precursors, probably due to exon shuffling.

Sequence characteristics of new miRNAs. Among the newly predicted miRNAs, the largest number, i.e., nine miRNAs were assigned to miR166 family followed by six miRNAs assigned to miR396 family. The nucleotide length of these newly identified miRNAs varied from 19 to 24 nt, with an average of 20.91 ± 1.18 nt (**Fig. 2**). The nucleotide length of sorghum pre-miRNA

Table 1. List of 31 sorghum miRNAs identified by comparative genomics and secondary structure analysis

miR Family	miRNA mature Sequence	L*	Chr*	Precursor		PL*	MM*	Arm*	Strand*	MFE	AMFE	MFEI
				Start	End					kcal/ mol		
sbi-MIR156j	UGACAGAAGAGAGUGAGCACA	21	3	3473047	3473132	86	2	5′	-	54.9	63.84	1.25
sbi-MIR156k	UGACAGAAGAGAGUGAGCACA	21	3	3473329	3473501	173	1	5′	-	88	50.87	1.04
sbi-MIR156l	UGACAGAAGAGAGUGAGCACA	21	4	5373507	5373664	158	3	5′	-	79.5	50.32	0.92
sbi-MIR156m	UGCUCUCUGCUCUCACUGUCAUC	23	2	62836711	62836860	150	3	3′	-	81.8	54.53	0.91
sbi-MIR166l	GGAAUGUUGUCUGGUUCAAGG	21	1	17295173	17295276	104	3	5′	-	46.7	44.9	1.09
sbi-MIR166m	UCGGACCAGGCUUCAUUCC	19	1	7426516	7426597	82	2	3′	+	38.6	47.07	0.73
sbi-MIR166n	UCGGACCAGGCUUCAUUCC	19	1	69265255	69265358	104	3	3′	-	59.1	56.83	1.02
sbi-MIR166o	UCGGACCAGGCUUCAUUCC	19	1	17295173	17295276	104	3	3′	-	46.7	44.9	1.09
sbi-MIR166p	UCGGACCAGGCUUCAUUCCCC	21	1	17295156	17295297	142	4	3′	-	66.7	46.97	1.13
sbi-MIR166q	UCGGACCAGGCUUCAUUCCCC	21	1	69265255	69265360	106	3	3′	-	60.7	57.26	1.05
sbi-MIR166r	UCGGACCAGGCUUCAUUCCCC	21	1	7426516	7426597	82	2	3′	+	38.6	47.07	0.73
sbi-MIR166s	UCGGACCAGGCUUCAUUCCCCC	22	1	69265255	69265358	104	4	3′	-	59.1	56.83	1.02
sbi-MIR166t	UCGGACCAGGCUUCAUUCCCCU	22	1	17295156	17295297	142	4	3′	-	66.7	46.97	1.13
sbi-MIR167j	GAUCGUGCUGCGCAGUUUCACC	22	3	64088363	64088485	123	2	3′	-	61.9	50.33	1.11
sbi-MIR167k	UGAAGCUGCCAGCAUGAUCUGA	22	3	64088363	64088485	123	1	5′	-	61.9	50.33	1.11
sbi-MIR168b	CCCGCCUUGCACCAAGUGAA	20	4	2246312	2246408	97	3	3′	-	56.1	57.84	0.84
sbi-MIR168c	GAUCCCGCCUUGCACCAAGUGAAU	24	4	2246328	2246408	81	5	3′	-	52.9	65.31	0.98
sbi-MIR171l	UUGAGCCGUGCCAAUAUCAC	20	7	7609102	7609232	131	1	3′	+	74.1	56.56	0.89
sbi-MIR171m	UUGAGCCGUGCCAAUAUCACG	21	7	7609102	7609232	131	1	3′	+	74.1	56.56	0.89
sbi-MIR390b	CGCUAUCUAUCCUGAGCUCCA	21	1	2870964	2871206	243	2	3′	+	115.5	47.54	1.03
sbi-MIR396f	GUUCAAGAAAGCUGUGGAAGA	21	4	66092395	66092515	121	2	5′	-	55.1	45.55	0.92
sbi-MIR396 g	GUUCAAUAAAGCUGUGGGAAA	21	4	66092521	66092630	110	2	3′	-	42	38.18	0.75
sbi-MIR396h	UCCACAGGCUUUCUUGAACUG	21	4	67655115	67655256	142	2	5′	-	61.8	43.52	0.86
sbi-MIR396i	UCCCACAGCUUUAUUGAACUG	21	4	66092395	66092515	121	2	5′	-	37.5	30.99	0.65
sbi-MIR396j	UCCCACAGCUUUAUUGAACUG	21	4	66092515	66092635	121	2	3′	+	55.1	45.55	0.92
sbi-MIR396k	UCUCCACAGGCUUUCUUGAACU	22	4	67655122	67655251	130	3	5′	-	65.6	50.46	0.99
sbi-MIR398b	GGGGCGGACUGGGAACACAUG	21	2	15190815	15190961	147	2	5′	-	69.6	47.35	0.81
sbi-MIR399l	GGGCAACUUCUCCUUUGGCAGA	22	9	55688233	55688348	116	3	5′	+	49.3	42.5	0.74
sbi-MIR444a	UGCAGUUGUUGUCUCAAGCUU	21	4	53723200	53728533	126	1	3′	-	75.2	59.68	1.27
sbi-MIR444b	UGUUGUCUCAAGCUUGCUGCC	21	4	53723200	53728533	126	3	3′	-	75.2	59.68	1.27
sbi-MIR444c	UUGUGGCUUUCUUGCAAGUUG	21	4	59021719	59021792	74	1	3′	+	22.5	30.4	0.64

*L, length of mature miRNAs; *PL, precursor length; *Arm, location of mature miRNAs on secondary stem-loop structures of pre miRNA sequences; *Strand, miRNAs existence in sense (+) and antisense (-) strand.

(stem-loop) varied from 74 to 243 nt, with an average of 122.58 ± 32.78 nt.The length distribution of miRNAs and their precursor sequences are similar to the previous reports in other plant species.[4,25,26,35] Out of 31 miRNAs, 22 (70.97%) began with a 5' uridine, a characteristic feature of miRNAs. Mature miRNA sequences have been shown to be located on the stem-arm of the secondary stem-loop hairpin structure of the potential pre-miRNA. Out of 31 miRNAs identified, 11 (35.48%) were found to be located on the 5' arm of the stem-loop hairpin structure, while 20 (64.52%) resided on the 3' arm. It is previously reported that microRNA precursors, unlike other non-coding RNAs, have lower folding free energy than random sequence.[14] Minimal folding free energy has been considered as one of important feature in previously described methods for miRNAs identification.[62,63]All newly identified sorghum miRNA precursors have negative minimal folding free energies (MFE), ranging from −22.5 to −115.5 kcal mol^{-1} with an average of −61.05 ± 17.82 kcal mol^{-1} (**Table 1**). MFEs are strongly and positively correlated with their sequence length.[4] To normalize the potential effect of sequence length on MFE and to differentiate miRNAs from other RNAs,[64] we used two energy measurements-namely MFE (AMFE) and minimal folding free energy index (MFEI), and demonstrated that a candidate RNA sequence is more likely to be an miRNA when the MFEI is greater than 0.85. The newly identified sorghum pri-miRNAs had a high MFEI (0.64–1.27), with an average of about 0.96 (**Fig. 3**) which is

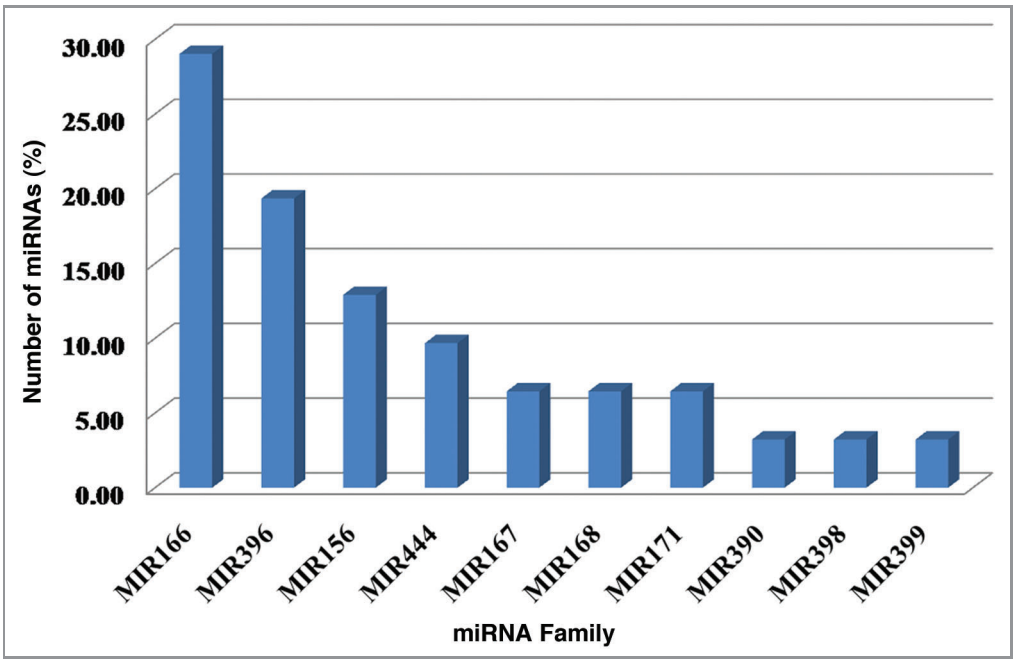

Figure 1. Distribution of 31 newly predicted miRNAs under diverse 10 miRNA families.

significantly higher than that for tRNAs (0.64), rRNAs (0.59), and mRNAs (0.62–0.66).[65]

***MIR* gene clusters in sorghum genome.** In general, miRNA gene clusters are found in both animal and plant genomes. Several studies revealed that some members of *MIR* gene family are physically clustered in plant genomes.[18,66–70] Some clusters are so compact where multiple miRNAs are aligned in the same orientation and transcribed as a polycistronic transcript.[68,71,72] Clusters are conserved across vertebrates: from teleost fish to human.[73] However in plants, only few miRNA cluster have been found.[13,26,56,68,70,74] We mapped all the previously registered 148 sorghum miRNAs and 31 miRNA predictedin this study to examine the potential clusters of *MIR* genes on the sorghum genome. As a result, 24 compact clusters were predicted for 73 sorghum miRNAs, having their genomic organization within 10 kb. The identified miRNA clusters belong to 12 different *MIR* gene families (**Table 2**). Our analysis revealed that chromosome 4 has seven *MIR* gene clusters, while chromosome

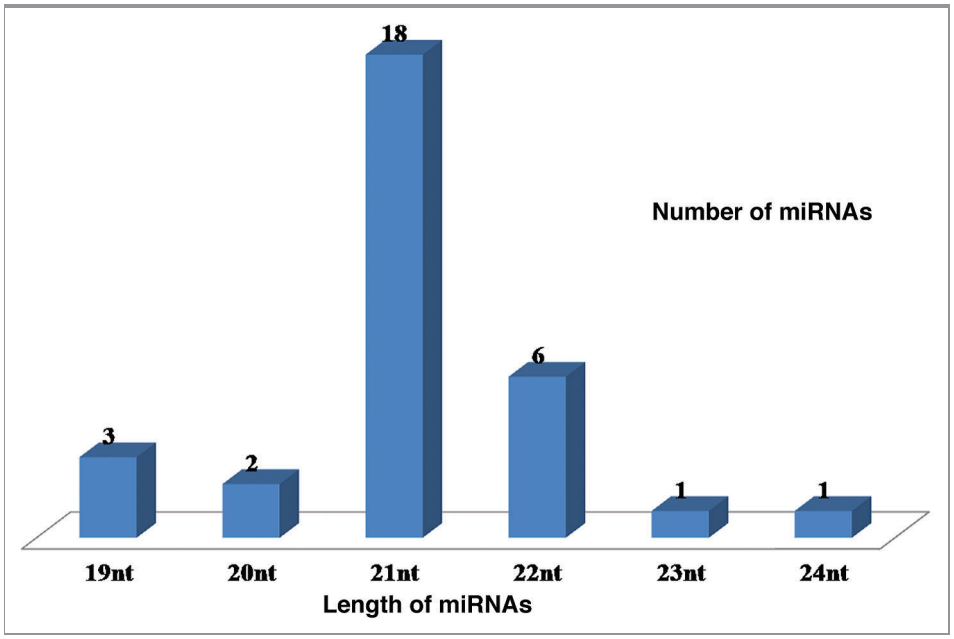

Figure 2. Length distribution of mature miRNAs in sorghum.

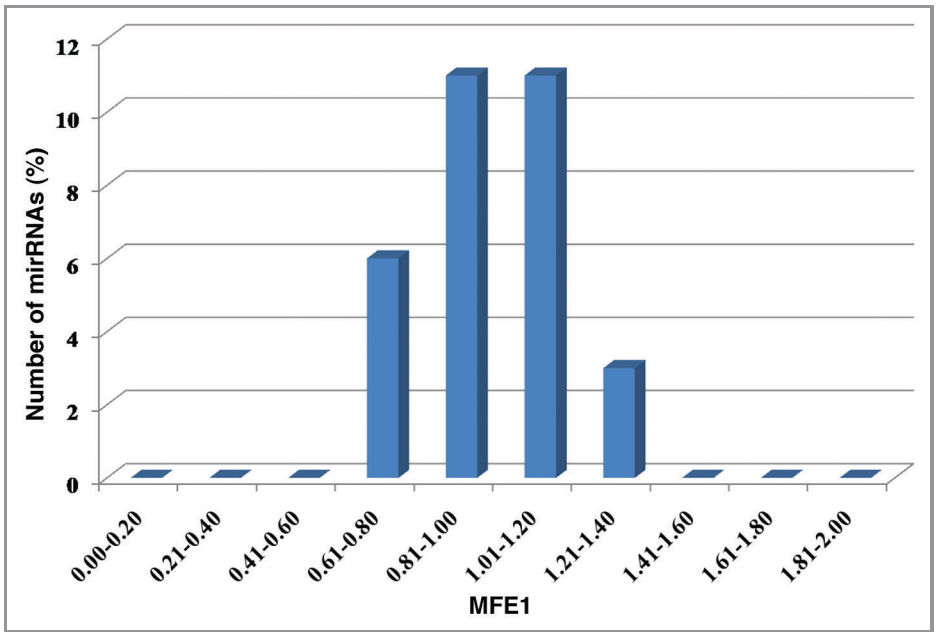

Figure 3. Minimal folding free energy index of pre-miRNAs in sorghum.

Table 2. Sorghum miRNA gene clusters on different chromosomes

Cluster number	Cluster name	miRNAs members	Cluster length$	Distance between miRNAs#	Chromosome
1	MIR390	b*, a	242	b-a: Overlap	Chr.1
2	MIR166	m*, r*, b	82	m-r: Overlap; r-b: Overlap	Chr.1
3	MIR166	p*, t*, a, l*, o*	142	p-t: Overlap; t-a: Overlap; a-l: Overlap; l-o: Overlap	Chr.1
4	MIR166	n*, q*, s*, c	106	n-q: overlap; q-s: Overlap; s-c: Overlap	Chr.1
5	MIR398	b*, a	147	b-a: Overlap	Chr.2
6	MIR156	m*, d	150	m-d: Overlap	Chr.2
7	MIR169	f, g	2984	f-g: 2686	Chr.2
8	MIR156	j*, b, k*, c	454	j-b: Overlap; b-k: 198; k-c: Overlap	Chr.3
9	MIR167	j*, k*, g	124	j-k: Overlap; k-g: Overlap	Chr.3
10	MIR168	b*, a, c*	110	b-a: Overlap; a-c: Overlap	Chr.4
11	MIR156	l*, a	158	l-a: Overlap	Chr.4
12	MIR399	j, k	5412	j-k: 5258	Chr.4
13	MIR444	a*, b*	126	a-b: Overlap	Chr.4
14	MIR166	g, f	407	g-f: 136	Chr.4
15	MIR396	c, f*, i*,a, j*, g*	7351	c-f: 6947; f-i:Overlap; i-a; Overlap; a-j: Overlap; j-g: Overlap	Chr.4
16	MIR396	h*, k*, d	142	h-k: Overlap; k-d: Overlap	Chr.4
17	MIR395	f, c, d, e	805	f-c: 71, c-d: 246, g-h: 89	Chr.6
18	MIR395	a, b, g, h	1014	a-b: 445, b-g: 75, g-h: 94	Chr.6
19	MIR395	i, j, k, l	742	i-j: 87, j-k: 84, k-l: 227	Chr.7
20	MIR171	b, l*, m*	133	b-l: Overlap; l-m: Overlap	Chr.7
21	MIR169	l, m, n	8633	l-m: 5291; m-n: 3063	Chr.7
22	MIR167	i, e	2466	i-e: 2157	Chr.8
23	MIR399	c, e, g, l*	6127	c-e: 1438; e-g: 4311; g-l: Overlap	Chr.9
24	MIR399	f, h	2492	f-h: 2242	Chr.10

#, Distance, Distance (nt) to previous miRNA gene in the cluster; $, Cluster length, Total miRNAs occupied region in one cluster; *, miRNA members, A star mark denote predicted sorghum miRNAs in this study.

1 has four *MIR* geneclusters (**Fig. 4**). We also noticed that no cluster was detected on chromosome 2. This indicates *MIR* gene clusters are common in sorghum. This prediction is similar as previously predicted *MIR* gene clusters in human[75] and soybean.[4] Altuvia et al.[75] confirmed that 42% miRNA genes are placed in clusters in the human genome using a 3 kb threshold or 48% if using 10 kb threshold between two miRNA genes. Zhang and colleges[4] also observed that 16% of the total identified soybean miRNAs genes are arranged in clusters. The sorghum *MIR* gene clusters were found diverse in structure and varies in cluster length from 82 to 8,633 bp, with an average of 1689.54. Here, we observed that each cluster consists of miRNA genes strictly from same gene families. This is in support to the previously given hypothesis that plant *MIR* gene clusters are comprised of homologous members.[13,56,68] Generally, miRNA members in a cluster share high sequence conservation; whereas a regular decrease in sequence similarity suggests that duplication events occurred at various time points. Compact clusters between two miRNAs have been found in various plant species. [25,54,70,76,77] In this article, we identified 31 new miRNAs. Most of which are arranged in compact cluster, and same miRNA family members with high sequence similarity formed a cluster.

Potential target genes for newly predicted miRNAs. The putative target genes of sorghum miRNAs were identified by a perfect or near-perfect sequence complementarity between miRNA and its target transcript. We searched the potential miRNA targets for predicted 31 new miRNAs against mRNA sequence of sorghum by using plant miRNA analysis psRNA Target tool[39] and UEA sRNA plant target prediction tool[78] and as a results, we obtained 72 (of which 49 are unique) potential target genes for 31 newly predicted miRNAs belonging to 10 different miR families (**Table 3**). The sequence alignments of 31 putative miRNAs and their corresponding targets in sorghum are shown in Supplemental Figure 2. We observe that number of targets per miRNA varied and some miRNAs have multiple target genes. For example, miR396 has 13 target genes, whereas miR444 has eight target genes. We noticed that a miRNA family members target the same set of genes, suggesting a functional redundancy amongthe family members. For example, a few members of miR166 family (miR166m-s) target the mRNA of homeobox leucine zipper transcription factor gene (target accession no: CN140010). In contrast, some members of miRNA families (e.g., miR444) have specific target genes. For example, miR444a target to WD-40 repeat family protein (target accession no: CN125113), involved in signal transducer activity. Pathway analysis of predicted target genes revealed that 14 targets are metabolism-associated. Among these, six miRNA (e.g., miR444b, 166 min, 166o, 166p, 166q and 166r) targets genes involved in sulfur metabolism, whereas eight members of miR177 targets genes involved in riboflavin metabolism. Most of the predicted targets of newly identified mRNAs are transcription factors that may have potential role in plant growth and development (**Table 3**). Remaining miRNAs target genes are involved in a broad range of biological functions, such as hydrolase activity (miR156 target EH409419) oligopeptide transporter activity (miR167target CX614408), riboflavin synthase activity (all predicted targets of miR171), kinase activity (miR396), zinc and calcium ion binding (miR444), translation initiation factor activity (miR444) and signal transducer activity (miR444) (**Table 3**). We also observed that when miRNA has more than one target, the potential targets belonged to the same gene family. All predicted miRNA and their targets

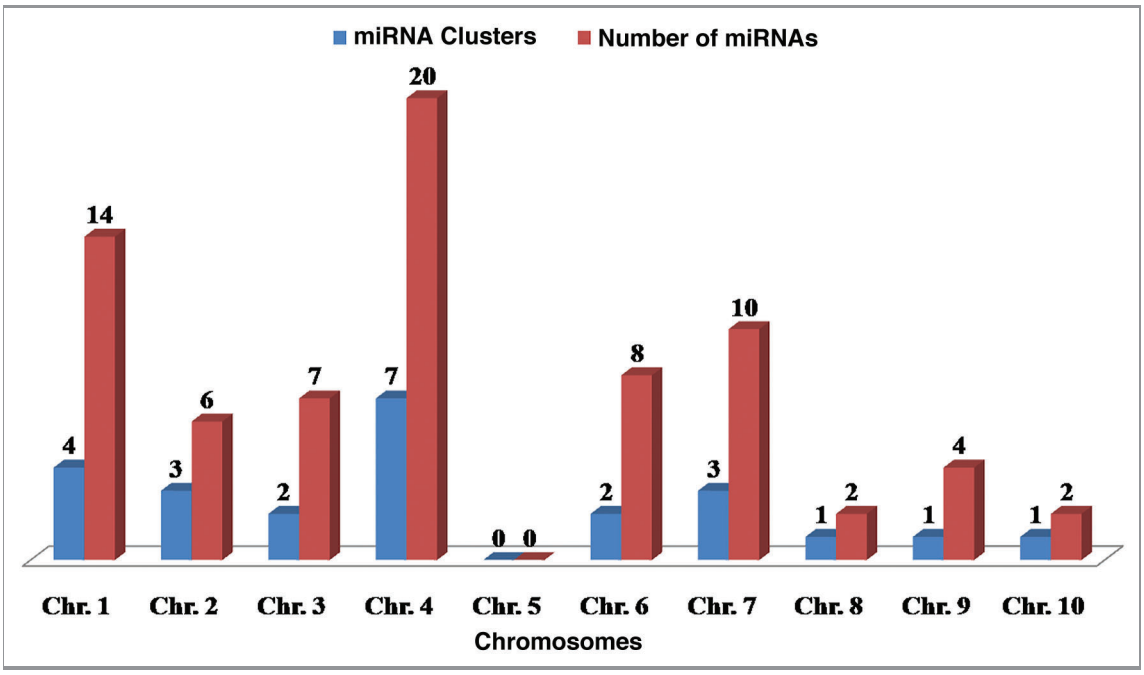

Figure 4. Number of miRNA clusters and occupied miRNAs gene on sorghum chromosomes.

Table 3. Potential target genes and their predicted functions for 31 newly identified miRNAs in sorghum

miRNA Acc.	Target Gene Acc.	Gene Annotation	Target Function	KEGG Pathway	COG Function	EST Expression
sbi-MIR156j	BG947367	Squamosa promoter-binding-like protein 16	DNA Binding Transcription Factor	No Hits Found	No Hits Found	Panicle and Callus
sbi-MIR156j	AW747167	SBP transcription factor	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR156j	EH409419	Hydrolase, α/β fold family protein, expressed	Hydrolase Activity	No Hits Found	Hydrolases or Acyltransferases (α/β hydrolase superfamily)	Ovary and Root
sbi-MIR156j	CF756128	Jumonji Domain Protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	Pollen
sbi-MIR156k	BG947367	Squamosa promoter-binding-like protein 16	DNA Binding Transcription Factor	No Hits Found	No Hits Found	Panicle and Callus
sbi-MIR156k	AW747167	SBP transcription factor	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR156k	EH409419	Hydrolase, α/β fold family protein, expressed	Hydrolase Activity	No Hits Found	Hydrolases or Acyltransferases (α/β hydrolase superfamily)	Ovary and Root
sbi-MIR156k	CF756053	Jumonji Domain Protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	Pollen
sbi-MIR156l	BG947367	Squamosa promoter-binding-like protein 16	DNA Binding Transcription Factor	No Hits Found	No Hits Found	Panicle and Callus
sbi-MIR156l	AW747167	SBP transcription factor	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR156l	EH409419	Hydrolase, α/β fold family protein, expressed	Hydrolase Activity	No Hits Found	Predicted Hydrolases or Acyltransferases (α/β hydrolase superfamily)	Ovary and Root
sbi-MIR156l	CF756128	Jumonji Domain Protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	Pollen
sbi-MIR156m	BG948102	Jumonji Domain Protein	Actin Filament Binding	No Hits Found	No Hits Found	Panicle and Root
sbi-MIR156m	BM325400	Phagocytosis and cell motility protein ELMO1-like	Mysoin II Binding	No Hits Found	No Hits Found	Ovary and Panicle
sbi-MIR166l	CN126049	Calcium channel α-1 subunit	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166l	CD461667	Catalytic Domain of Protein Kinases	Kinase Activity	No Hits Found	Serine/threonine protein kinases	Leaf
sbi-MIR166l	CN135236	CAP22 protein	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166m	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166m	CF074035	Diphosphonucleotide Phosphatase1	3′(2′),5′-Bisphosphate Nucleotidase activity / Inositol or Phosphatidylinositol Phosphataseactivity	Sulfur Metabolism	Inorganic ion Transport and Metabolism	Shoot
sbi-MIR166n	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166n	CD209789	Sodium- and Lithium-Tolerant 1 (SLT1)	Molecular Function Unknown	No Hits Found	No Hits Found	Callus, Leaf, Ovary, Panicle and Root
sbi-MIR166o	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166o	CF074035	Diphosphonucleotide Phosphatase1	3′(2′),5′-Bisphosphate Nucleotidase activity / Inositol or Phosphatidylinositol Phosphataseactivity	Sulfur Metabolism	Inorganic ion Transport and Metabolism	Shoot
sbi-MIR166o	CD209789	Sodium- and Lithium-Tolerant 1 (SLT1)	Molecular Function Unknown	No Hits Found	No Hits Found	Callus, Leaf, Ovary, Panicle and Root

Table 3. Potential target genes and their predicted functions for 31 newly identified miRNAs in sorghum (continued)

miRNA Acc.	Target Gene Acc.	Gene Annotation	Target Function	KEGG Pathway	COG Function	EST Expression
sbi-MIR166p	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166p	CF074035	Diphosphonucleotide Phosphatase1	3′(2′),5′-Bisphosphate Nucleotidase activity / Inositol or Phosphatidylinositol Phosphataseactivity	Sulfur Metabolism	Inorganic ion Transport and Metabolism	Shoot
sbi-MIR166p	CN126049	Calcium channel α-1 subunit	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166q	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166q	CF074035	Diphosphonucleotide Phosphatase1	3′(2′),5′-Bisphosphate Nucleotidase activity / Inositol or Phosphatidylinositol Phosphataseactivity	Sulfur Metabolism	Inorganic ion Transport and Metabolism	Shoot
sbi-MIR166q	CN126049	Calcium channel α-1 subunit	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166r	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166r	CF074035	Diphosphonucleotide Phosphatase1	3′(2′),5′-Bisphosphate Nucleotidase activity / Inositol or Phosphatidylinositol Phosphataseactivity	Sulfur Metabolism	Inorganic ion Transport and Metabolism	Shoot
sbi-MIR166r	CN126049	Calcium channel α-1 subunit	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166r	CN140010	Homeobox-leucine zipper protein	DNA Binding Transcription Factor	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166r	CN126049	Calcium channel α-1 subunit	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR166s	CN126049	Calcium channel α-1 subunit	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR167j	CX614408	Iron-Phytosiderophore Transporter Yellow Stripe	Oligopeptide Transporter Activity	No Hits Found	No Hits Found	Root and Shoot
sbi-MIR167k	CD423596	Auxin Response Factor 9	DNA Binding Transcription Factor / Protein Dimerization Activity	No Hits Found	No Hits Found	Ovary, Panicle and Root
sbi-MIR167k	CN140701	Arv1-like protein	Molecular Function Unknown	No Hits Found	No Hits Found	No Hits Found
sbi-MIR168c	CD209773	PWWP domain-containing protein	Molecular Function Unknown	No Hits Found	No Hits Found	Leaf
sbi-MIR171l	CN129969	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Polen and Root
sbi-MIR171l	CN142205	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Polen and Root
sbi-MIR171l	AW747132	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Polen and Root
sbi-MIR171l	CN131513	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Polen and Root
sbi-MIR171	CN131513	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Polen and Root
sbi-MIR171m	CN129969	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Polen and Root
sbi-MIR171m	BM317608	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Pollen and Root
sbi-MIR171m	BE355399	6,7-dimethyl-8-ribityllumazine synthase	Riboflavin Synthase Activity	Riboflavin Metabolism	Riboflavin Synthase Beta-Chain	Leaf, Ovary, Pollen and Root

Table 3. Potential target genes and their predicted functions for 31 newly identified miRNAs in sorghum (continued)

miRNA Acc.	Target Gene Acc.	Gene Annotation	Target Function	KEGG Pathway	COG Function	EST Expression
sbi-MIR390b	BE363723	ARF-like GTPas	Transcription factor	No Hits Found	No Hits Found	Ovary
sbi-MIR396f	CF480198	Exosome complex exonuclease RRP41	RNA Binding	RNA Degradation	RNase PH	Leaf and pollen
sbi-MIR396 g	CB927788	plant synaptotagmin	Metal Ion Binding	No Hits Found	No Hits Found	Leaf, Panicle and Root
sbi-MIR396 g	CD211715	Ubiquitin carrier protein	Ubiquitin-Protein Ligase Activity	Ubiquitin Mediated Proteolysis	No Hits Found	Callus and Embryo
sbi-MIR396 g	CD222706	Ubiquitin carrier protein	Ubiquitin-Protein Ligase Activity	Ubiquitin Mediated Proteolysis	No Hits Found	Callus and Embryo
sbi-MIR396h	AF466199	Putative Receptor Protein Kinase	Kinase Activity	No Hits Found	Serine/Threonine Protein Kinases	Ovary and Panicle
sbi-MIR396h	CF072738	Growth-regulating factor 1	Kinase Activity	No Hits Found	Serine/Threonine Protein Kinases	No Hits Found
sbi-MIR396h	CD204209	Ankyrin-repeat containing protein	Protein Binding, Protein Kinase Activity, Protein Self-Association and Ubiquitin-Protein Ligase Activity	No Hits Found	Ankyrin repeat proteins	Leaf
sbi-MIR396i	CD234788	Homeodomainleucine zipper protein 16	DNA-binding Transcription Factor Activity	No Hits Found	No Hits Found	Callus, Embryo and Shoot
sbi-MIR396i	CF427893	Homeodomainleucine zipper protein 16	DNA-binding Transcription Factor Activity	No Hits Found	No Hits Found	Callus, Embryo and Shoot
sbi-MIR396j	CD234788	Homeodomainleucine zipper protein 16	DNA-binding Transcription Factor Activity	No Hits Found	No Hits Found	Callus, Embryo and Shoot
sbi-MIR396j	CX612479	Homeodomainleucine zipper protein 16	DNA-binding Transcription Factor Activity	No Hits Found	No Hits Found	Callus, Embryo and Shoot
sbi-MIR396k	AW676947	Growth-regulating factor 1	Protein Binding	No Hits Found	No Hits Found	Root
sbi-MIR396k	AF466199	Putative Receptor Protein Kinase	Kinase Activity	No Hits Found	Serine/Threonine Protein Kinases	Ovary and Panicle
sbi-MIR396k	BM329506	Peptidase family protein	Peptidase Activity	No Hits Found	No Hits Found	Leaf, Pollen and Root
sbi-MIR396k	CD207048	C2 domain-containing protein	Molecular Function Unknown	No Hits Found	No Hits Found	Embryo, Leaf and Shoot
sbi-MIR398b	CF480868	UDP-N-acetylglucosaminetransferase subunit ALG14	Transferase Activity	N-Glycan Biosynthesis	No Hits Found	Pollen and Root
sbi-MIR398b	CF487358	UDP-N-acetylglucosaminetransferase subunit ALG14	Transferase Activity	N-Glycan Biosynthesis	No Hits Found	Pollen and Root
sbi-MIR398b	BG158064	Ent-kaurene oxidase	Oxidoreductase activity,	No Hits Found	Cytochrome P450	No Hits Found
sbi-MIR398b	BM330737	Chloroplast 30S ribosomal protein S3	Structural Constituent of Ribosome	Ribosome	Ribosomal Protein L22	Leaf, Ovary and Panicle
sbi-MIR444a	BM323459	MADS-box transcription factor 57	Transcription Factor Binding	No Hits Found	No Hits Found	No Hits Found

Table 3. Potential target genes and their predicted functions for 31 newly identified miRNAs in sorghum (continued)

miRNA Acc.	Target Gene Acc.	Gene Annotation	Target Function	KEGG Pathway	COG Function	EST Expression
sbi-MIR444a	CD224118	Zinc finger (C3HC4-type RING finger) protein-like	Zinc Ion Binding	No Hits Found	No Hits Found	Callus and Ovary
sbi-MIR444a	CD225619	MADS-box transcription factor 57	Calcium Ion Binding	No Hits Found	No Hits Found	Callus and Embryo
sbi-MIR444a	CN125113	WD-40 repeat family protein	Signal Transducer Activity	No Hits Found	WD40 Repeat Protein	Embryo, Ovary and Root
sbi-MIR444b	BE596704	MADS-box transcription factor 57	Transcription Factor Binding	No Hits Found	No Hits Found	No Hits Found
sbi-MIR444b	BM330337	Putative far-red impaired response protein	Zinc Ion Binding	No Hits Found	No Hits Found	Leaf
sbi-MIR444b	AW564049	Ferredoxin-sulfite reductase precursor	Sulfite Reductase Activity	Sulfur Metabolism	Sulfite ReductaseHemoprotein Beta-Component	Callus, Embryo, Leaf and Pollen
sbi-MIR444c	BM323459	MADS-box transcription factor 57	Transcription Factor Binding	No Hits Found	No Hits Found	No Hits Found
sbi-MIR444c	BM325378	Eukaryotic translation initiation factor 3 subunit A	Translation Initiation Factor Activity	RNA Transport	Chromosome Segregation ATPases	Leaf and Ovary

share high similarity to their orthologs in *Arabidopsis thaliana* and *Zea mays*.

Discussion

Large numbers of miRNAs have been predicted in cereal model plant rice. However, in sorghum, only 148 miRNAs were reported till date (miRBase release 17). This suggested that miRNA identification in sorghum is far from saturation. We used ESTs and GSS to predict miRNAs and predicted 31 new miRNAs, in addition to 148 miRNAs in miRBase release 17. We also predicted 72 potential target genes for the newly identified 31 miRNAs. In consistent with earlier studies, members of a miRNA family target same set of genes, suggesting a functional redundancy of the miRNA family members. We mapped newly identified 31 miRNAs and previously known 148 miRNAs on sorghum genome, and found that several *MIR* genes are arranged in clusters in sorghum genome. Each cluster consists of *MIR* genes belonging to the same family.

Sorghum crop is highly tolerant to drought and heat stress, and the expansion of members of *MIR* gene families, specifically miR169 family, was suggested as one of the probable reasons for adaptation of sorghum to abiotic stresses.[79] Besides miR169 family, expansion of other miRNA families may also contribute to the better adaptation of sorghum. The miR166 gene family and its role in leaf development are evolutionarily conserved in all land plants. The miR166 family regulates the expression of the HD-ZIP III (class-III homeodomain-leucin zipper) gene family that is necessary for proper specification of leaf polarity, in both Arabidopsis and maize.[80] In this study, initially we predicted 26 candidate miRNAs, belong to miR166 family. Later, genome mapping of these miRNAs showed that only nine miRNAs mapped to unique genomic loci. The fact that the copies of Sb-miR166 family member do not vary during the evolution suggests its importance in sorghum. Previous studies also showed that many miRNAs are evolutionarily conserved across animals and plants.[85–87] However, some miRNAs are species-specific.[88] For instance, miR444 family is present in rice but not in Arabidopsis, suggesting that it might be restricted to monocots.[89] Further studies revealed that miR444 family members are conserved only in monocot species (e.g., barley, maize, wheat, sorghum, Brachypodium and sugarcane) but not in dicot species (e.g., Arabidopsis and Populus).[8,89,90] Here, we identified three members of miR444 family in sorghum namely miR444a, b and c which were identical to the previously reported miR444c.1, c.2 and d, respectively in rice.[91] The precursors of these newly identified miR444a, b and c, had a high minimal folding free energy index of 1.27, 1.27 and 0.64, respectively. This is significantly higher than those reported for tRNAs (0.64), rRNAs (0.59), and mRNAs (0.62–0.66), suggesting that pri-miRNA folding of this family is less stable than that of others. It is previously reported that miR444 family members target MADS-box transcription factors that are involved in number of biological functions including developmental processes namely meristem identity, root development, fruit dehiscence, flowering time[94–97] and tolerance to salt and cold stresses.[92,93,98] We also

observed that all three members of miR444 family target to MADS-box transcription factors (target accession no: BM323459 and BE596704) and other target genes with a role in calcium ion binding, translation initiation factor activity, DNA binding transcription factor, sulfite reducates activity and signal transducer activity. The lack of a miR444 homolog and their conserved target gene (MADS box) in dicot families such as Arabidopsis provided strong evidence that miRNA-mediated regulation of MADS box gene is conserved only in monocots and known as 'monocot-specific' family. Further molecular genetic analysis of miR444 family might be helpful to unravel the significance of this monocot-specific family.

EST mining from publically available database could provide evidence for the expression of miRNAs in different tissues. In this study, EST profiles were explored from UniGene database that revealed the expression patterns of miRNAs families in various tissues and at different development stages (**Table S1**). The differential expression pattern of miRNAs suggests their potential role in development of the respective tissues or process in these tissues. Thus, the newly identified miRNAs and their predicted roles form the basis for understanding their role in sorghum plant development and stress adaptation.

Materials and Methods

Reference sequences for miRNA prediction. To identify potential conserved miRNAs, a total 2,728 previously identified plant miRNA from 33 different plant species were obtained from miRBase database (release 17, April 2011) (www.mirbase.org/).[40,41,99] This set include miRNA sequences from *Chlamydomonas reinhardti*[i] (50), *Pinus taeda* (37), *Physcomitrella patens* (229), *Selaginella moellendorffii* (58), *Arabidopsis thaliana* (232), *Brassica napus* (46), *Brassica oleracea* (6), *Brassica rapa* (19), *Carica papaya* (1), *Glycine max* (203), *Lotus japonicus* (3), *Medicago truncatula* (375), *Phaseolus vulgaris* (8), *Vigna unguiculata* (2), *Gossypium arboreum* (1), *Gossypiumherbecium* (1), *Gossypium hirsutum* (34), *Gossypium raimondii* (4), *Aquilegia coerulea* (45), *Malusdo mestica* (1), *Citrus clementine* (5), *Citrus reticulata* (4) *Citrus sinensis* (60), *Citrus trifoliata* (6), *Populus euphratica* (5), *Populus trichocarpa* (234), *Solanum lycopersicum* (36), *Vitis vinifera* (163), *Brachypodium distachyon* (139), *Oryza sativa* (491), *Saccharum officinarum* (16), *Triticum aestivum* (44) and *Zea mays* (170). After removal of the redundant sequences, 1379 miRNAs were used as reference set.

Sorghum EST, GSS and WGS sequence data set. Sorghum expresses sequence tag (EST), genomic survey sequences (GSS) and whole genome sequence (WGS) were obtained from GenBank nucleotide database available at NCBI (www.ncbi.nlm.nih.gov/). This data set contains 240161 nucleotide sequence from EST and 799,504 nucleotide sequences from GSS (Till January 5, 2010).

Non-coding data set. Non coding data set of mRNA were used to discriminate between miRNA and other structural RNAs (e.g., tRNA, rRNA, snRNA and snoRNA). The BLASTN search was performed against pfam (http://rfam.sanger.ac.uk/) database[100] to remove ESTs or GSS having similarity with structural RNAs. The filter for tRNA was also conducted by blast of possible miRNAs precursors against genomic tRNA database (v.2.4.2) (http://gtrnadb.ucsc.edu/blast.html).[101] The parameters for BLAST alignment was fixed as Alignment Program: blastn; Expect: 0.01; Word Size; 11; Database All eukaryotic tRNA.The tRNA genomic data set contain tRNA gene sequences from Arabidopsis, soybean and rice. The snRNA and snoRNA sequences from plant kingdom were also retrieved at random basis from NCBI (www.ncbi.nlm.nih.gov) and mapped with predicted miRNA data set to exclude false positive miRNA precursors in sorghum.

Prediction of secondary structure. To make data non-redundant, including EST, GSS and reference miRNA sequence, multiple sequence alignment was performed by using locally installed ClustalX (version 2.0.12) and web based ClustalW[61] (version 1.83) (www.genome.jp/tools/clustalw) with default parameters. The unique reference miRNA sequences were mapped on EST and GSS sequence by using an in-house PERL script (www.perl.org) and miRNAs with no mismatch were only retained for further analysis. Flanking region of 250 nt base pair upstream and downstream from miRNA sequence from EST and GSS sorghum sequences were extracted and folded using RNAFold version 1.8.4 from the Vienna RNA package[60] (rna.tbi.univie.ac.at/) to find out minimum free energy containing structure. To predict real miRNA precursor triplet-SVM classifier[102] program which is based on support vector machine was used (bioinfo.au.tsinghua.edu.cn/mirnasvm/). This software package needs third-party softwares namelyRNAfold and LibSVM packages. The minimal folding free energy Index (MFEI) was calculated using the following equation: MFEI = [(MFE*/length of the RNA sequence) *100]/(G+C) %. *MFE denotes the negative folding free energies (ΔG).

MicroRNAs target genes. The putative target sites of miRNAs were identified by aligning the miRNA sequences either perfectly or near-perfectly binding to complementary sites on their target mRNA sequences[64] by using Plant Target Prediction Tool available on UEA sRNA ToolKit[78] (srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi) and psRNA Target server[39] (http://plantgrn.noble.org/psRNATarget/) with default parameters; Maximum expectation: 3.0, length for complementarity scoring (hspsize): 20, Target accessibility-allowed maximum energy to unpair the target site (UPE): 25.0, Flanking length around target site for target accessibility analysis: 17 bp in upstream and 13 bp in downstream, Range of central mismatch leading to translation inhibition: 9–11 nt. The rules used in UEA sRNA Tool Kit for target prediction suggested by Allen et al.[103] and Schwab et al.[64] were as follows: (1) No more than four mismatches between the small RNA and the target (G-U bases count as 0.5 mismatches); (2) No more than two adjacent mismatches in the miRNA: target duplex; (3) No adjacent mismatches in positions 2–12 of the miRNA: target duplex (5' of miRNA); (4) No mismatches in positions 10–11 of the miRNA: target duplex; (5) No more than 2.5 mismatches in positions 1–12 of the of the miRNA: target duplex (5' of miRNA); (6) The minimum free energy (MFE) of the miRNA/target duplex should be $\geq$ 74% of the MFE of the miRNA bound to its perfect complement.

Functional analysis of target genes. The functional assignment of predicted target genes were annotated by COGnitor program that compare gene sequence against the Clusters of Orthologous Groups of proteins (COG) database[104] (version 66) (www.ncbi.nih.gov/COG). AmiGO (version 1.8) (amigo.genontology.org) and KEGG (Kyoto Encyclopedia of Genes and Genomes) (www.genome.jp/kegg)pathway analyses were employed to further investigate the biological processes and corresponding metabolic networks regulated by potential miRNAs. All predicted target genes with an e value of 1e−30 were identified by BLASTX searching program[105] against the GO protein and KEGG databases (version 58.0) (Released on April 1, 2011).

Conclusions and Prospective

In this study, we identified 31 new miRNAs in sorghum by analyzing ESTs and GSS. The study revealed that 73 diverse miRNAs (including miRBase, version 17 registered sorghum miRNAs) were arranged into 24 compact clusters on sorghum genome. We also found three members of monocot species-specific MIR444 family, widely involved in regulation of MADS-box transcription factor expression. About 72 potential target genes for 31 individual miRNAs belonging to nine different miRNA families were predicted. We noticed that majority of the predicted target genes were transcription factors, which are involved in the regulation of plant growth and development. The findings from this study will contribute to further understanding the miRNAs function and regulatory mechanisms in sorghum.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Indian Council of Agricultural Research (ICAR) for supporting this work through the ICAR-sponsored Network Project on Transgenics in Crops (NPTC).

Supplementary Material

Supplementary material may be found at:
www.landesbioscience.com/journals/psb/article/18914

References

1. Prasad PVV, Pisipati SR, Mutava RN, Tuinstra MR. Sensitivity of grain sorghum to high temperature stress during reproductive development. Crop Sci 2008; 48: 1911-7; http://dx.doi.org/10.2135/cropsci2008.01.0036
2. Hattori T, Sonobe K, Araki H, Inanaga S, An P, Morita S. Silicon application by sorghum through the alleviation of stress-induced increase in hydraulic resistance. J Plant Nutr 2008; 31:1482-95; http://dx.doi.org/10.1080/01904160802208477
3. Mutegi E, Fabrice S, Moses M, Ben K, Bernard R, Caroline M, et al. Ecogeographical distribution of wild, weedy and cultivated *Sorghum bicolor* (L.) Moench in Kenya: implications for conservation and crop-to-wild gene flow. Genet Resour Crop Evol 2010; 57:243-53; http://dx.doi.org/10.1007/s10722-009-9466-7
4. Zhang B, Pan X, Stellwag EJ. Identification of soybean microRNAs and their targets. Planta 2008; 229:161-82; PMID:18815805; http://dx.doi.org/10.1007/s00425-008-0818-x
5. Zhao T, Li G, Mi S, Li S, Hannon GJ, Wang XJ, et al. A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*. Genes Dev 2007; 21:1190-203; PMID:17470535; http://dx.doi.org/10.1101/gad.1543507
6. Hiraguri A, Itoh R, Kondo N, Nomura Y, Aizawa D, Murai Y, et al. Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in *Arabidopsis thaliana*. Plant Mol Biol 2005; 57:173-88; PMID:15821876; http://dx.doi.org/10.1007/s11103-004-6853-5
7. Lu C, Fedoroff N. A mutation in the *Arabidopsis* HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. Plant Cell 2000; 12:2351-66; PMID:11148283
8. Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol 2006; 57:19-53; PMID:16669754; http://dx.doi.org/10.1146/annurev.arplant.57.032905.105218
9. Chen R, Hu Z, Zhang H. Identification of microRNAs in wild soybean (*Glycine soja*). J Integr Plant Biol 2009; 51:1071-9; PMID:20021554; http://dx.doi.org/10.1111/j.1744-7909.2009.00887.x
10. Sunkar R, Chinnusamy V, Zhu J, Zhu JK. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci 2007; 12:301-9; PMID:17573231; http://dx.doi.org/10.1016/j.tplants.2007.05.001
11. Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell 2009; 136:669-87; PMID:19239888; http://dx.doi.org/10.1016/j.cell.2009.01.046
12. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell 2002; 110:513-20; PMID:12202040; http://dx.doi.org/10.1016/S0092-8674(02)00863-2
13. Jones-Rhoades MW, Bartel DP. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 2004; 14:787-99; PMID:15200956; http://dx.doi.org/10.1016/j.molcel.2004.05.027
14. Bonnet E, Wuyts J, Rouzé P, Van de Peer Y. Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics 2004; 20:2911-7; PMID:15217813; http://dx.doi.org/10.1093/bioinformatics/bth374
15. Park W, Li J, Song R, Messing J, Chen X. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol 2002; 12:1484-95; PMID:12225663; http://dx.doi.org/10.1016/S0960-9822(02)01017-5
16. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. Genes Dev 2002; 16:1616-26; PMID:12101121; http://dx.doi.org/10.1101/gad.1004402
17. Sunkar R, Zhu J-K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 2004; 16:2001-19; PMID:15258262; http://dx.doi.org/10.1105/tpc.104.022830
18. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science 2001; 294:853-8; PMID:11679670; http://dx.doi.org/10.1126/science.1064921
19. Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 2001; 294: 858-62; PMID:11679671; http://dx.doi.org/10.1126/science.1065062
20. Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. Science 2001; 294:862-4; PMID:11679672; http://dx.doi.org/10.1126/science.1065329
21. Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, et al. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev 2002; 16:720-8; PMID:11914277; http://dx.doi.org/10.1101/gad.974702
22. Yoon S, De Micheli G. Computational identification of microRNAs and their targets. Birth Defects Res C Embryo Today 2006; 78:118-28; PMID:16847881; http://dx.doi.org/10.1002/bdrc.20067
23. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, et al. Computational prediction of miRNAs in Arabidopsis thaliana. Genome Res 2005; 15:78-91; PMID:15632092; http://dx.doi.org/10.1101/gr.2908205
24. Li Y, Li W, Jin YX. Computational identification of novel family members of microRNA genes in *Arabidopsis thaliana* and *Oryza sativa*. Acta Biochim Biophys Sin (Shanghai) 2005; 37:75-87; PMID:15685364; http://dx.doi.org/10.1093/abbs/37.2.75
25. Zhang BH, Pan XP, Anderson TA. Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 2006; 580:3753-62; PMID:16780841; http://dx.doi.org/10.1016/j.febslet.2006.05.063
26. Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo TL, et al. Identification of cotton microRNAs and their targets. Gene 2007; 397:26-37; PMID:17574351; http://dx.doi.org/10.1016/j.gene.2007.03.020
27. Zhou ZS, Huang SQ, Yang ZM. Bioinformatic identification and expression analysis of new microRNAs from *Medicago truncatula*. Biochem Biophys Res Commun 2008; 374:538-42; PMID:18662674; http://dx.doi.org/10.1016/j.bbrc.2008.07.083

28. Song C, Jia Q, Fang J, Li F, Wang C, Zhang Z. Computational identification of citrus microRNAs and target analysis in citrus expressed sequence tags. Plant Biol (Stuttg) 2010; 12:927-34; PMID:21040308; http://dx.doi.org/10.1111/j.1438-8677.2009.00300.x
29. Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, et al. Computational identification of novel microRNAs and targets in *Brassica napus*. FEBS Lett 2007; 581:1464-74; PMID:17367786; http://dx.doi.org/10.1016/j.febslet.2007.02.074
30. Jin WB, Li NN, Zhang B, Wu FL, Li WJ, Guo AG, et al. Identification and verification of microRNA in wheat (*Triticum aestivum*). J Plant Res 2008; 121:351-5; PMID:18357413; http://dx.doi.org/10.1007/s10265-007-0139-3
31. Zhang WW, Luo YP, Gong X, Zeng WH, Li SG. Computational identification of 48 potato microRNAs and their targets. Comput Biol Chem 2009; 33: 84-93; PMID:18723398; http://dx.doi.org/10.1016/j.compbiolchem.2008.07.006
32. Yin Z, Li C, Han X, Shen F. Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). Gene 2008; 414:60-6; PMID:18387754; http://dx.doi.org/10.1016/j.gene.2008.02.007
33. Xie F, Frazier TP, Zhang B. Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). Planta 2010; 232:417-34; PMID:20461402; http://dx.doi.org/10.1007/s00425-010-1182-1
34. JiangFeng D, YongJun W, XiaoFeng F, JunXia C, Liang Z, ShiHeng T. Prediction of sorghum miRNAs and their targets with computationalMethods. Chin Sci Bull 2010; 55:1263-70.
35. Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA. Identification and characterization of new plant microRNAs using EST analysis. Cell Res 2005; 15: 336-60; PMID:15916721; http://dx.doi.org/10.1038/sj.cr.7290302
36. Pan XP, Zhang BH, San Francisco M, Cobb GP. Characterizing viral microRNAs and its application on identifying new microRNAs in viruses. J Cell Physiol 2007; 211:10-8; PMID:17167781; http://dx.doi.org/10.1002/jcp.20920
37. Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, et al. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 1991; 252:1651-6; PMID:2047873; http://dx.doi.org/10.1126/science.2047873
38. Matukumalli LK, Grefenstette JJ, Sonstegard TS, Van Tassell CP. EST-PAGE–managing and analyzing EST data. Bioinformatics 2004; 20:286-8; PMID:14734326; http://dx.doi.org/10.1093/bioinformatics/btg411
39. Dai X, Zhao PX. psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res 2011; 39(Web Server issue):W155-9; PMID:21622958; http://dx.doi.org/10.1093/nar/gkr319
40. Griffiths-Jones S. The microRNA Registry. Nucleic Acids Res 2004; 32(Database issue):D109-11; PMID:14681370; http://dx.doi.org/10.1093/nar/gkh023
41. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006; 34(Database issue):D140-4; PMID:16381832; http://dx.doi.org/10.1093/nar/gkj112
42. Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC. Expression of Arabidopsis MIRNA genes. Plant Physiol 2005; 138:2145-54; PMID:16040653; http://dx.doi.org/10.1104/pp.105.062943
43. Lu C, Kulkarni K, Souret FF, MuthuValliappan R, Tej SS, Poethig RS, et al. MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. Genome Res 2006; 16: 1276-88; PMID:16954541; http://dx.doi.org/10.1101/gr.5530106
44. Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Dev 2006; 20:3407-25; PMID:17182867; http://dx.doi.org/10.1101/gad.1476406
45. Moldovan D, Spriggs A, Yang J, Pogson BJ, Dennis ES, Wilson IW. Hypoxia-responsive microRNAs and trans-acting small interfering RNAs in Arabidopsis. J Exp Bot 2010; 61:165-77; PMID:19815687; http://dx.doi.org/10.1093/jxb/erp296
46. Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, et al. A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Res 2008; 18:1456-65; PMID:18687877; http://dx.doi.org/10.1101/gr.075572.107
47. Merchan F, Boualem A, Crespi M, Frugier F. Plant polycistronic precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. Genome Biol 2009; 10:R136; PMID:19951405; http://dx.doi.org/10.1186/gb-2009-10-12-r136
48. Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y. Rice MicroRNA effector complexes and targets. Plant Cell 2009; 21:3421-35; PMID:19903869; http://dx.doi.org/10.1105/tpc.109.070938
49. Xue LJ, Zhang JJ, Xue HW. Characterization and expression profiles of miRNAs in rice seeds. Nucleic Acids Res 2009; 37:916-30; PMID:19103661; http://dx.doi.org/10.1093/nar/gkn998
50. Lu S, Sun YH, Chiang VL. Stress-responsive microRNAs in Populus. Plant J 2008; 55:131-51; PMID:18363789; http://dx.doi.org/10.1111/j.1365-313X.2008.03497.x
51. Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL. Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell 2005; 17:2186-203; PMID:15994906; http://dx.doi.org/10.1105/tpc.105.033456
52. Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, et al. A genome-wide characterization of microRNA genes in maize. PLoS Genet 2009; 5: e1000716; PMID:19936050; http://dx.doi.org/10.1371/journal.pgen.1000716
53. Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, et al. Criteria for annotation of plant MicroRNAs. Plant Cell 2008; 20:3186-90; PMID:19074682; http://dx.doi.org/10.1105/tpc.108.064311
54. Axtell MJ, Bartel DP. Antiquity of microRNAs and their targets in land plants. Plant Cell 2005; 17:1658-73; PMID:15849273; http://dx.doi.org/10.1105/tpc.105.032185
55. Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, et al. High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS One 2007; 2:e219; PMID:17299599; http://dx.doi.org/10.1371/journal.pone.0000219
56. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA. Conservation and divergence of plant microRNA genes. Plant J 2006; 46:243-59; PMID:16623887; http://dx.doi.org/10.1111/j.1365-313X.2006.02697.x
57. Bonnet E, Wuyts J, Rouzé P, Van de Peer Y. Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. Proc Natl Acad Sci U S A 2004; 101:11511-6; PMID:15272084; http://dx.doi.org/10.1073/pnas.0404025101
58. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, et al. A uniform system for microRNA annotation. RNA 2003; 9:277-9; PMID:12592000; http://dx.doi.org/10.1261/rna.2183803
59. Zuker M, Stiegler P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res 1981; 9:133-48; PMID:6163133; http://dx.doi.org/10.1093/nar/9.1.133
60. Hofacker IL. Vienna RNA secondary structure server. Nucleic Acids Res 2003; 31:3429-31; PMID:12824340; http://dx.doi.org/10.1093/nar/gkg599
61. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22:4673-80; PMID:7984417; http://dx.doi.org/10.1093/nar/22.22.4673
62. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, et al. MicroRNA identification based on sequence and structure alignment. Bioinformatics 2005; 21: 3610-4; PMID:15994192; http://dx.doi.org/10.1093/bioinformatics/bti562
63. Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res 2007; 35(Web Server issue):W339-44; PMID:17553836; http://dx.doi.org/10.1093/nar/gkm368
64. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D. Specific effects of microRNAs on the plant transcriptome. Dev Cell 2005; 8:517-27; PMID:15809034; http://dx.doi.org/10.1016/j.devcel.2005.01.018
65. Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA. Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 2006; 63:246-54; PMID:16395542; http://dx.doi.org/10.1007/s00018-005-5467-7
66. Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. EMBO J 2002; 21:4663-70; PMID:12198168; http://dx.doi.org/10.1093/emboj/cdf476
67. Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T. New microRNAs from mouse and human. RNA 2003; 9:175-9; PMID:12554859; http://dx.doi.org/10.1261/rna.2146903
68. Guddeti S, Zhang DC, Li AL, Leseberg CH, Kang H, Li XG, et al. Molecular evolution of the rice miR395 gene family. Cell Res 2005; 15:631-8; PMID:16117853; http://dx.doi.org/10.1038/sj.cr.7290333
69. Tanzer A, Amemiya CT, Kim CB, Stadler PF. Evolution of microRNAs located within Hox gene clusters. J Exp Zool B Mol Dev Evol 2005; 304:75-85; PMID:15643628; http://dx.doi.org/10.1002/jez.b.21021
70. Maher C, Stein L, Ware D. Evolution of Arabidopsis microRNA families through duplication events. Genome Res 2006; 16:510-9; PMID:16520461; http://dx.doi.org/10.1101/gr.4680506
71. Bompfünewerer AF, Flamm C, Fried C, Fritzsch G, Hofacker IL, Lehmann J, et al. Evolutionary patterns of non-coding RNAs. Theory Biosci 2005; 123: 301-69; PMID:18202870; http://dx.doi.org/10.1016/j.thbio.2005.01.002
72. Weber MJ. New human and mouse microRNA genes found by homology search. FEBS J 2005; 272:59-73; PMID:15634332; http://dx.doi.org/10.1111/j.1432-1033.2004.04389.x
73. Tanzer A, Stadler PF. Molecular evolution of a microRNA cluster. J Mol Biol 2004; 339:327-35; PMID:15136036; http://dx.doi.org/10.1016/j.jmb.2004.03.065
74. Talmor-Neiman M, Stav R, Frank W, Voss B, Arazi T. Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss. Plant J 2006; 47:25-37; PMID:16824179; http://dx.doi.org/10.1111/j.1365-313X.2006.02768.x
75. Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, et al. Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 2005; 33:2697-706; PMID:15891114; http://dx.doi.org/10.1093/nar/gki567

76. Maher C, Timmermans MC, Stein L, Ware D. Identifying microRNAs in plant genomes. Proc IEEE CSB 2004; 718–723.
77. Wang S, Zhu QH, Guo X, Gui Y, Bao J, Helliwell C, et al. Molecular evolution and selection of a gene encoding two tandem microRNAs in rice. FEBS Lett 2007; 581:4789-93; PMID:17884044; http://dx.doi.org/10.1016/j.febslet.2007.09.002
78. Moxon S, Schwach F, Dalmay T, Maclean D, Studholme DJ, Moulton V. A toolkit for analysing large-scale plant small RNA datasets. Bioinformatics 2008; 24:2252-3; PMID:18713789; http://dx.doi.org/10.1093/bioinformatics/btn428
79. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, et al. The Sorghum bicolor genome and the diversification of grasses. Nature 2009; 457:551-6; PMID:19189423; http://dx.doi.org/10.1038/nature07723
80. Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MC. microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. Nature 2004; 428:84-8; PMID:14999285; http://dx.doi.org/10.1038/nature02363
81. Floyd SK, Bowman JL. Gene regulation: ancient microRNA target sequences in plants. Nature 2004; 428:485-6; PMID:15057819; http://dx.doi.org/10.1038/428485a
82. Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, et al. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. EMBO J 2004; 23:3356-64; PMID:15282547; http://dx.doi.org/10.1038/sj.emboj.7600340
83. Vaucheret H, Vazquez F, Crété P, Bartel DP. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev 2004; 18:1187-97; PMID:15131082; http://dx.doi.org/10.1101/gad.1201404
84. Guo HS, Xie Q, Fei JF, Chua NH. MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for *arabidopsis* lateral root development. Plant Cell 2005; 17:1376-86; PMID:15829603; http://dx.doi.org/10.1105/tpc.105.030841
85. Ambros V. The functions of animal microRNAs. Nature 2004; 431:350-5; PMID:15372042; http://dx.doi.org/10.1038/nature02871
86. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004; 116:281-97; PMID:14744438; http://dx.doi.org/10.1016/S0092-8674(04)00045-5
87. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, et al. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. Nature 2000; 408:86-9; PMID:11081512; http://dx.doi.org/10.1038/35040556
88. Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D. MicroRNAs and other tiny endogenous RNAs in *C. elegans*. Curr Biol 2003; 13:807-18; PMID:12747828; http://dx.doi.org/10.1016/S0960-9822(03)00287-2
89. Sunkar R, Girke T, Zhu JK. Identification and characterization of endogenous small interfering RNAs from rice. Nucleic Acids Res 2005; 33:4443-54; PMID:16077027; http://dx.doi.org/10.1093/nar/gki758
90. Sunkar R, Girke T, Jain PK, Zhu JK. Cloning and characterization of microRNAs from rice. Plant Cell 2005; 17:1397-411; PMID:15805478; http://dx.doi.org/10.1105/tpc.105.031682
91. Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu J-K. Identification of novel and candidate miRNAs in rice by high throughput sequencing. BMC Plant Biol 2008; 8:25; PMID:18312648; http://dx.doi.org/10.1186/1471-2229-8-25
92. Kuo MH, Nadeau ET, Grayhack EJ. Multiple phosphorylated forms of the *Saccharomyces cerevisiae* Mcm1 protein include an isoform induced in response to high salt concentrations. Mol Cell Biol 1997; 17:819-32; PMID:9001236
93. Lozano R, Angosto T, Gomez P, Payan C, Capel J, Huijser P, et al. Tomato flower abnormalities induced by low temperatures are associated with changes of expression of MADS-Box genes. Plant Physiol 1998; 117:91-100; PMID:9576778; http://dx.doi.org/10.1104/pp.117.1.91
94. Coen ES, Meyerowitz EM. The war of the whorls: genetic interactions controlling flower development. Nature 1991; 353:31-7; PMID:1715520; http://dx.doi.org/10.1038/353031a0
95. Weigel D, Meyerowitz EM. The ABCs of floral homeotic genes. Cell 1994; 78:203-9; PMID:7913881; http://dx.doi.org/10.1016/0092-8674(94)90291-7
96. Riechmann JL, Meyerowitz EM. MADS domain proteins in plant development. Biol Chem 1997; 378:1079-101; PMID:9372178
97. Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Münster T, et al. A short history of MADS-box genes in plants. Plant Mol Biol 2000; 42:115-49; PMID:10688133; http://dx.doi.org/10.1023/A:1006332105728
98. Ding Y, Chen Z, Zhu C. Microarray-based analysis of cadmium-responsive microRNAs in rice (Oryza sativa). J Exp Bot 2011; 62:3563-73; PMID:21362738; http://dx.doi.org/10.1093/jxb/err046
99. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res 2008; 36(Database issue):D154-8; PMID:17991681; http://dx.doi.org/10.1093/nar/gkm952
100. Gardner PP, Daub J, Tate J, Moore BL, Osuch IH, Griffiths-Jones S, et al. Rfam: Wikipedia, clans and the "decimal" release. Nucleic Acids Res 2011; 39(Database issue):D141-5; PMID:21062808; http://dx.doi.org/10.1093/nar/gkq1129
101. Chan PP, Lowe TM. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res 2009; 37(Database issue):D93-7; PMID:18984615; http://dx.doi.org/10.1093/nar/gkn787
102. Xue C, Li F, He T, Liu GP, Li Y, Zhang X. Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinformatics 2005; 6:310; PMID:16381612; http://dx.doi.org/10.1186/1471-2105-6-310
103. Allen E, Xie Z, Gustafson AM, Carrington JC. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 2005; 121:207-21; PMID:15851028; http://dx.doi.org/10.1016/j.cell.2005.04.004
104. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, et al. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res 2001; 29:22-8; PMID:11125040; http://dx.doi.org/10.1093/nar/29.1.22
105. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990; 215:403-10; PMID:2231712