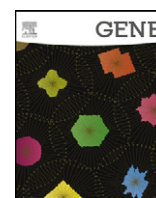


Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Identification of microRNAs and their targets in Finger millet by high throughput sequencing

S. Usha^{a,c}, M.N. Jyothi^{a,c}, N. Sharadamma^b, Rekha Dixit^c, V.R. Devaraj^d, R. Nagesh babu^{a,*}^a Post Graduate Department of Biochemistry, Maharani's Science College for Women, Bangalore 560001, India^b Department of Biochemistry, Indian Institute of Science, Bangalore 560012 India^c Centre for Bioinformatics, Faculty of Biological Engineering, Shobhit University, Meerut, 250110 India^d Department of Biochemistry, Central College Campus, Bangalore University, Bangalore 560001 India

ARTICLE INFO

Article history:

Received 12 May 2015

Received in revised form 3 August 2015

Accepted 5 August 2015

Available online 7 August 2015

Keywords:

cis-regulatory motifs

GO analysis

RT-qPCR

RISC

TFBs

TSS

ABSTRACT

MicroRNAs are short non-coding RNAs which play an important role in regulating gene expression by mRNA cleavage or by translational repression. The majority of identified miRNAs were evolutionarily conserved; however, others expressed in a species-specific manner. Finger millet is an important cereal crop; nonetheless, no practical information is available on microRNAs to date. In this study, we have identified 95 conserved microRNAs belonging to 39 families and 3 novel microRNAs by high throughput sequencing. For the identified conserved and novel miRNAs a total of 507 targets were predicted. 11 miRNAs were validated and tissue specificity was determined by stem loop RT-qPCR, Northern blot. GO analyses revealed targets of miRNA were involved in wide range of regulatory functions. This study implies large number of known and novel miRNAs found in Finger millet which may play important role in growth and development.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Plants have a complex mechanism of gene expression and regulation that influences their development, adaptation and response to biotic and abiotic interactions. Regulation of gene expression is one of the most enigmatic facets of molecular genetics that results in intricate appearance of a biological entity. Scientists have been attempting to elucidate the regulatory mechanisms of gene expression of small non-coding RNAs (Jiang et al., 2013). Several distinct classes of small RNAs have been reported so far including microRNAs (miRNAs) (Khraiwesh et al., 2013). miRNAs are small endogenous 20–24 nt non-coding RNAs derived from single-stranded RNA precursors that can form stem-loop structures and have been proved to play a vital role in wide range of biological process. In plants, after transcription by Pol II or Pol III enzyme into primary miRNA (pri-miRNA), the miRNA gene is processed by Dicer-like (DCL) into a stem-loop that contain miRNA/miRNA* duplex, called an miRNA precursor (pre-miR-NA). Subsequently, the duplex is cleaved from the pre-miRNA and transported from the nucleus into the cytoplasm which combines with Argonaute (AGO)

forming the RNA-induced silencing complex (RISC). Finally, the silencing complex down-regulates targets by either cleaving target mRNAs or repressing the translation process (Voinnet, 2009; Yu and Wang, 2010). Several miRNAs have been identified in plants, and they have been characterized in a wide variety of metabolic and biological processes with important functions (Kulcheski et al., 2011). The first plant miRNAs were identified in *Arabidopsis thaliana* (Park et al., 2002); currently the latest miRBase release (v21, June 2014) contains 46231 mature microRNAs. Earlier microRNAs were identified using bioinformatics analysis or Sanger sequencing; currently high-throughput sequencing platforms are showing significant promise for small RNA discovery and genome-wide transcriptome analysis at single-base pair resolution (Li et al., 2011). High-throughput sequencing was used to identify miRNAs in many plant species, such as rice (Sunkar et al., 2008), alfalfa (Szitty et al., 2008), grape (Pantaleo et al., 2010), tomato (Moxon et al., 2008), orange (Song et al., 2010), soybean (Song et al., 2011), peanut (Zhao et al., 2010), poplar (Klevebring et al., 2009), and black gram (Paul et al., 2014). Moreover, the availability of complete genome sequence for a number of these miRNAs encoding genes from diverse plant species has allowed a better understanding of miRNA evolution in plants (Zhu et al., 2011b). However, an analysis of high throughput sequencing data is a challenge, especially in the absence of native genome sequence (Islam et al., 2015). Finger millet a closely related species to Foxtail millet has its genome sequenced (Bennetzen et al., 2012) and the miRNAs have been identified using high throughput sequencing (Yi et al., 2013). Khan et al. (2014a,b) constructed Foxtail millet miRNA

Abbreviations: GO, gene ontology; miRNA, microRNA; TF, transcription factor; TSS, transcription start site; TFB, transcription factor binding site; NB-ARC, nucleotide binding domain with ARC motif.

* Corresponding author at: Postgraduate Department of Biochemistry, Maharani's Science College for Women, Palace Road, Bangalore 560001, India.

E-mail address: nageshbabur@gmail.com (R. Nagesh babu).

Database (FmMiRNADb: <http://59.163.192.91/FmMiRNADb/index.html>) which provides miRNA marker information and also studied expression profiling under diverse abiotic stress conditions. Novel microRNA-based genetic markers in Foxtail millet has been constructed which promises the genotyping applications at crop improvement (Yadav et al., 2014). Owing to the important roles played by miRNAs in gene regulatory networks, it is important to identify the miRNAs in Finger millet (*Eleusine coracana* L.), a cereal grass grown mostly for its grain and belongs to the family Poaceae. In the present study, we aimed to characterize the miRNA repertoire of Finger millet. First of its study, Illumina sequencing was performed to identify the conserved and novel miRNAs. A total of 98 miRNAs were identified, the expression and tissue specificity were validated by RT-qPCR and Northern blot. The possible targets were predicted, TSS and TFBS were also identified for novel miRNAs.

2. Materials and methods

2.1. Plant material and construction of small RNA library

The Finger millet (*E. coracana* cultivar MR1) was used for the study. Seeds were surface-sterilized with 4% (v/v) sodium hypochlorite and were grown hydroponically on 16-h day/8-h night photoperiod cycle at 25 ± 2 °C for ten days. The samples were frozen in liquid nitrogen and stored at -80 °C for further analysis. Total RNA from the shoot tissue was isolated by RNAiso plus (Takara) according to manufacturers' instructions. Illumina TruSeq Small RNA kit was used to construct small RNA library. Briefly, 1 µg of total RNA was ligated with 3' and 5' adaptors. After reverse transcription and PCR, cDNA construct was sequenced using the Illumina NextSeq 500 (Illumina Inc., USA) (Xcelris, India). The raw data have been submitted to National Center for Biotechnology Information (NCBI-SRA) and can be assessed under SRA accession number (SRR1832812).

2.2. Identification of known and novel miRNAs

After removing undesirable raw reads including low-quality reads, adapter reads and reads either shorter than 15 or longer than 30 nt, remaining sequences were queried against Rfam (12.0) (<http://www.sanger.ac.uk/resources/databases/rfam.html>) database to filter rRNA, tRNA, snRNA, and snoRNA. The unmapped reads from Rfam were BLASTN against miRBase (Release 21) (<http://www.mirbase.org/>) using miRdeep2 (<https://www.mdcberlin.de/rajewsky/miRDeep>) to predict conserved miRNAs. Potential miRNA precursors were identified with Foxtail millet (*Setaria italica*) as reference genome (www.phytozome.net). The secondary structures were predicted using RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) (Gruber et al., 2008). The small RNA sequences which did not share homology with existing miRNA sequences and had characteristic hairpin structures and minimal folding free-energy indices ($MFEI \geq -0.85$) were considered as novel miRNAs.

2.3. Prediction of miRNA targets

The Foxtail millet (*S. italica*) transcripts, downloaded from phytozome v10 (www.phytozome.net) was used to determine the potential target mRNA candidates for miRNAs using psRNATarget with default parameters (<http://plantgrn.noble.org/psRNATarget/>) (Dai and Zhao, 2011). To reduce the false-positive prediction rate, the cut-off threshold was set at 0 to 3.0 points. Thus, sequences with 3.0 points were considered to be miRNA targets. Functional annotations of predicted targets were analyzed using BGI WEGO platform (<http://wego.genomics.org.cn/cgi-bin/wego/index.pl>) (Ye et al., 2006).

2.4. Validation of conserved and novel miRNAs

RT-qPCR was used to validate high throughput sequencing results. 1 µg of total RNA was reverse transcribed using stem-loop primers designed according to Chen et al. (2005) using One Step PrimeScript miRNA cDNA synthesis kit (Takara, Japan). RT-qPCR was performed using SYBR premix ExTaq (Takara, Japan), U6 as internal control and all the primers used are listed (Supplementary material 1). Subsequently, the quantification was carried out using CFX Manager (Biorad). Two biological replicates were used per sample in addition to technical replicates along with a no template control and no RT-enzyme control. The data were analyzed using $2^{-\Delta\Delta CT}$ method and reported as means \pm standard errors (SE) of two biological replicates. The standard deviations of the data were obtained from the three independent experiments with student's t-test ($p \leq 0.05$) using GraphPad Prism v.5. Northern blotting was performed using miRNA northern blot assay kit (Signosis) according to the manufacturer's instructions. Briefly, 5 µg of RNA was fractionated using 15% TBE urea-PAGE, blotted on membranes, and hybridized with biotin labeled antisense RNA (Invitrogen) and U6 probes (Supplementary material 2). The amount of RNA loaded in each well was normalized to the amount of U6 as internal control.

2.5. Functional characterization of novel miRNA genes

The genomic locations and upstream sequences (2 kb) of novel miRNAs were retrieved for TSS prediction. The miRNAs were assumed as independent transcription units to have uniformity. The TSS and TATA-box predictions were made using TSSP web tool (<http://linux1.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter>). Putative promoter sequences from $-1,000$ to -1 of the TSS were retrieved for novel miRNA and used for strong motif identification. The miRNA of Finger millet were scanned for putative transcription factor binding motifs TFBM's using modules (i) PLACE (Plant cis-acting regulatory DNA elements) signal scan search, to identify the known cis-regulatory elements (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>) (Higo et al., 1999) (ii) NSITE (<http://linux1.softberry.com/>) and (iii) MELINA-II a web based tool, to run and compare the motif prediction algorithms at once (<http://melina2.hgc.jp/public/index.html>) (Okumura et al., 2007). The four motif prediction algorithms used in this study were: (a) Consensus, (b) Gibbs Sampler, (c) MD SCAN and (d) MEME (with default parameters). The motifs identified by at least two programs were considered as strong motifs.

3. Results

3.1. Analysis of sRNAs in Finger millet

To identify the known and novel miRNAs in Finger millet small RNA library was prepared and subjected to high throughput sequencing (Quail et al., 2010). The sequencing resulted in 20,512,196 raw reads, after filtering out adapter contaminants, low-quality reads and Rfam contaminants 7,755,160 unique reads were got and the percentage of putative small RNA reads were 73% (Table 1). The most abundant sRNAs range from 21–24 nt (Chi et al., 2011; Xu et al., 2013), and the 21 nt sRNAs represented the most abundant followed by 24 nt which

Table 1
Summary of small RNA reads in Finger millet.

	Read count	Percentage %
Total reads	20,512,196	100
High quality reads	20,143,419	98.2
Adaptor trim sequences	19,001,244	92.6
Mapped to Rfam	4,621,918	22.5
tRNA match	344,933	1.68
rRNA match	574,965	2.80
Unique reads	7,755,160	73

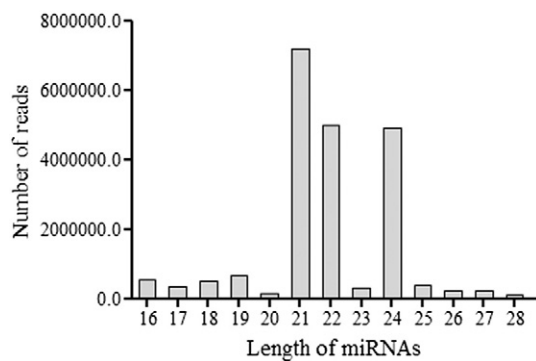


Fig. 1. Length distribution of miRNAs in Finger millet.

was in agreement with previous reports (Morin et al., 2008; Ge et al., 2012) (Fig. 1).

3.2. Identification of conserved and novel miRNAs in Finger millet

Mature miRNAs show more conservancy in plants than their precursor sequence (Bartel, 2004).

Hence, the unique reads were mapped on to the miRBase version 21 using miRDeep2. The miRNAs which shared the perfect homology in the miRBase were considered as conserved miRNAs. The sequence analyses revealed 95 conserved miRNAs belonging to 39 families, many of these identified miRNAs showed conservancy in other plant species. Individual representatives of miRNA families varied with maximum in miR 156 comprising of 11 members and minimum in miR171, 394, 397, 398, 528, 529, 827, 845, 1128, 1511, 1514, 2118, 2199, 2218, 2637, 5054, 5073, 5384, 6300, and 8140 (one each). MEFL was calculated for all the conserved miRNAs, to differentiate from other non-coding RNAs like tRNA, rRNA and mRNA. The average MEFL was found to be -0.809 , the average G + C content for miRNA genes was 48.5% (Supplementary material 3).

We identified 3 novel miRNAs using miRDeep2 by aligning the reads that had no homology with miRBase miRNAs to Foxtail millet genome, and applied prediction criteria for plant miRNAs (Kozomara and Griffiths-Jones, 2014) (Table 2). Eco_N1 was found to be homologous with miR2275 and considered as the new member of the existing miR2275 family identified in Finger millet. The low abundance of novel miRNAs in the data supports the earlier notation of the lower expression levels of novel miRNAs compared with those of conserved miRNAs (Chi et al., 2011). The precursor miRNA candidates were evaluated using RNAfold. The minimum free energy that was required to form the predicted hairpin structure for the precursor was in the range of -31.05 to -42.10 Kcal/mol which is consistent with other reports. The secondary structures were analyzed to validate the novel miRNAs (Fig. 2). Data analysis also revealed the presence of miRNA* sequence for all the three identified miRNAs which further emphasize the presence of these sequences.

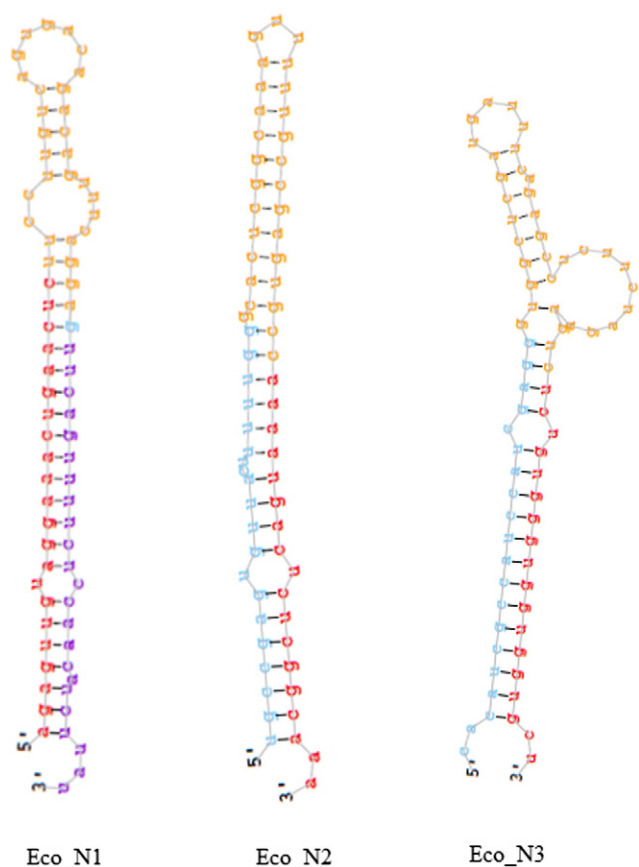


Fig. 2. Secondary structure of novel miRNAs identified in Finger millet.

3.3. Target predictions for miRNAs identified in Finger millet

To understand the importance of the miRNAs putative targets were predicted using psRNA Target with *S. italica* transcripts. Among the targets predicted miR396 and miR 2118 families have the highest number of targets. The predicted target genes (approximately 507 different transcripts) were involved in different biological processes including a large number of gene families. Some of the genes encoded transcription factors, DNA replication proteins, cellular metabolism and other stress response associated proteins. miR156 and miR157 targets SPL protein, miR159 targets MYB and miR166 targets Homeobox-leucine zipper family protein, others include NAC, NB-ARC domain and G2484-1 targeted by miR164 and miR171 respectively. miR396 targets growth regulating factors involved in development and cell proliferation which is shown to interact with KNOTTED1-LIKE HOMEBOX (KNOX) genes, suggests that the interaction might prevent transcription (Kuijt et al., 2014). miR437 targeted DNAJ heat shock protein also known as HSP40s, in soyabean it is shown that DNAJ protein is involved in cell death and disease resistance (Liu and Whitham, 2013). miR2118 targets disease resistance protein rpm1 and LRR and NB-ARC

Table 2
Novel miRNAs in Finger millet.

Novel miRNA	Mature miRNA sequence	miRNA* sequence	Length	Total read count	MFEI	Precursor sequence	Precursor coordinates
Eco_N1	AGAGUUGUAGGAAA CUGAACUC	UUCAGUUUUCUCCA ACAUCUUU	22	2	-1.02	AGAGUUGUAGGAAAACUGAACUCUCCUUGUCAGUCAGACAGUUUCA GGAGUUCAGUUUUCUCCAACUUCUUU	scaffold_6
Eco_N2	AAAAUGACUCUCGG CAAA	UGCCGAGUGUUACU UUUUGG	18	32	-1.25	UGCCGAGUGUUACUUUUUGGGCACUCUGGCAAGUUUUUGCCGAGUGCC AAAAUGACUCUCGGCAAA	scaffold_8
Eco_N3	UCUGUGGGUGGU GGUGCU	CACAUCGCCAUCCA UCGAGG	18	19	-0.82	CACAUCGCCAUCCAUCGAGGGUGGCUCAUGAUUUUCAGAGCCUUCU AGAAGUCUCUGUGGGUGGUGUGCU	scaffold_5

domains-containing disease resistance protein as its targets which emphasize that may be this miRNA is involved in complex regulatory network (Supplementary material 4). Eco_N1 was predicted to target RNA polymerase III, which has DNA-directed RNA polymerase activity and DNA binding as its functions. Studies have shown that RNA Polymerase III detects cytosolic DNA and involved in innate immune responses through RIG-I Pathway (Chiu et al., 2009). However, Eco_N2 and Eco_N3 did not show defined targets (Table 3). The gene enrichment studies were performed by BGI WEGO tool. The GO terms of the targets genes were annotated according to biological processes, molecular functions or involvement as cellular components. Gene annotations showed that majority of the conserved microRNA targeted genes belonged to various biological process including miRNA biogenesis, DNA repair and apoptotic process, novel miRNA targeted gene was involved in molecular function as depicted (Fig. 3). The annotations for all GO terms annotated (Supplementary material 5).

3.4. Expression analysis of conserved and novel miRNAs

We validated the presence of novel miRNAs along with randomly selected known miRNAs by RT-qPCR and Northern blot in shoot and root tissues. miR156, miR157, and miR393 were significantly expressed in shoot; however miR159, miR169, miR396, miR397 and miR398 showed similar expression pattern. Novel miRNAs were found to show same expression patterns in both shoot and root tissues. A significant association between these differentially expressed miRNAs was found by Pearson correlation analysis. A positive co-relation resulted between the expression trends of conserved and novel miRNAs, r and p values were found to be 0.9958 and 0.058368 respectively. Similar expression patterns were also observed in Northern blot as depicted in Fig. 4.

3.5. Characterization of novel miRNA genes

2 kb upstream sequences of novel miRNAs were analyzed for the presence of transcription start site (TSS), TATA box and promoters. Out of three novel miRNA genes TSS and TATA box was identified for two miRNAs (Eco_N1 and Eco_N2). miRNA showed the presence of TSS in the region of –200 to –300 bp. In Eco_N2 two TSS were predicted, one nearer to pre-miRNA start was considered for further analysis. Conserved regulatory motifs were identified using NISTE, 18 regulatory motifs were predicted (Supplementary material 6). Putative promoter motifs were determined using 1 kb sequence upstream of the TSS

using single scan search in PLACE web server (Megraw et al., 2006; Zhang et al., 2006a). Only significant motifs which were present in at least or more than 50% of the analyzed sequences were considered as strong motifs. Eleven conserved motifs were predicted for 2 novel miRNAs, which include WRKY710S, and MYCCONSUSAT (Fig. 5) (Table 4).

4. Discussion

Identification and characterization of miRNA in various plant species is necessary to expedite their regulatory roles. Introduction of high throughput sequencing technology has increased the number of miRNAs identified by allowing sequencing of RNA population on massive scale. Large number of miRNAs have been reported and deposited in miRBase from several model species. To date there is no data on Finger millet whole genome; and it is confront to identify miRNAs. First of its report in this study, an illumina small RNA sequencing was performed to identify the complete set of conserved and novel miRNAs. A large number of unique reads were generated and length distribution showed, the most abundant miRNAs were 21 and 24 nt (Rajagopalan et al., 2006). The data processing, with BlastN search and computational analysis enabled to identify 95 conserved miRNAs from 39 families. A vital part was prediction of novel miRNAs, un annotated reads after conserved miRNA identification were subjected to further bioinformatics analysis and 3 novel miRNAs were predicted using miRDeep2 with Foxtail millet as reference genome. Length of identified novel miRNAs ranged from 18–22 nt and their precursors length ranging from 66–77 nt. Average MEF value was –1.03, much lower than the recommended value of –0.85 (Zhang et al., 2006b). Sequencing frequency in high throughput sequencing can be used as strategy for analyzing relative abundance of miRNAs (Sunkar et al., 2006). miR156 which is found to be largely conserved across plant species was the most abundant family with eleven members and showed high expression levels. A relationship between the degree of evolutionary conservation and the level of expression was observed in conserved miRNAs. Majority of the conserved Finger millet miRNA families identified were conserved evolutionarily across plant species with high levels of expression. The less-conserved miRNA families (miR437, miR528, miR845, miR2199, miR2275, and miR5073) showed lower abundance than the well-conserved miRNA families.

To validate the results, RT-qPCR and Northern blot hybridization were performed, which showed differential expression patterns.

Table 3
Novel miRNA target gene identified using psRNATarget database.

Eco_N1	Target_ID	Target sequence	Expectation (E)	UPE	Target description	GO numbers
AGAGUUGUAGGAAACUGAACUC	Si017911m	ATGGAGAAGGAAGATACAAAGAAGAAAGAGGCTGATGACT CCACTCCTCAGCATCGCCGCAAAGCTGACTTAAATTTGCA CCCCAAGTTCTTCTAAGAAAGACCCAAAGATCATCCTTAA AACGGAACCACAAGAGGAGAACAAGGCCCTGACAATTGAT AAGAAAATGATGCTGACTTGGGAGCTTGCAGAGCAGTTA TGGCCAGGAAGTGGAACTAAGGCTGAGAAACAAGGAACC CTGTCCAGGTTGCATTTGGACGAGCAGACCTTCAATTGC AAGAACTTTTCCCACTCGAAGAAGCTTTTCATCAGACGTAT CTGCTGTAAATACCCAAGGAGCAGCATGAACCATGGGA CTACAACTGCACCGATTATCTGTTACTTACCACTGAGAA GGCCCAACTCTGAAATGCAGAAATCTTGATGAAGAGGA ATTCCGACATTATCCGGCAGAGCCCAAGTTAGTGAAGTGA GTGCAGCTGAAGAACTTGGTCTGATGCAAGGGTAGATACG CCGAGTTGCTCTTCTCCAGTTCTCACTCACTCTTCTTTA CCAAGGCAGGCTGACCCAGATGCTGGCACAATATGAAT ATGAATGCAAAATCCATGGGGACAACAGGAAGAGGCGTC TCGACTCAATCCATGGCTGTGGTCTGAAAGAGCTCCAGATG GCTTCATGGGGAAGATGCTCGTGTACAAGAGTGGTAAAGTG AAGATGACGCTCGGTGATGTGCTTTTGATGTTTACGCTGGCT CGAATTGCATGTTTCTCAGGAAGTTAGCAATTAACACCAG GGAGAAGCACTGCTCGGCATAGGGGAGATCGGAAGCGT GCAATTGTAACCTGATATCAACTCTTGTAGAAATCTATT GAGATTGA	2	20.975	RNA polymerase III RPC4 protein	GO:0003677 GO:0003899 GO:0006383 GO:0005666

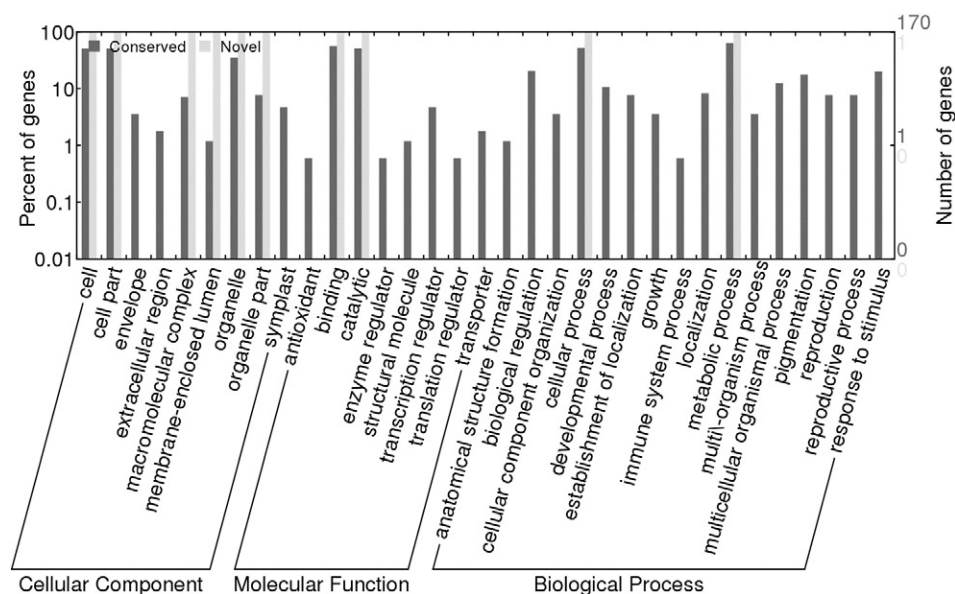


Fig. 3. Gene ontology classification of potential target genes in Finger millet.

miR159 and miR169 showed same expression patterns in root and shoot, studies have revealed that miR169 is responsive to abiotic stress such as drought, cold and salinity in different species (Ding et al., 2009; Li et al., 2013; Tian et al., 2014; Zhang et al., 2012). miR397 and miR398 also showed same expression levels in root and shoot. Ma et al. (2015) showed that in cotton miR397 and miR398 had high expression levels in flower buds, suggesting that these miRNAs may be involved in floral development.

To predict putative miRNA functions, target identification is important (Xie et al., 2006), in this study targets were identified using Fox-tail millet transcripts as reference. miR156/157-Squamosa promoter-binding protein (Reyes and Chua, 2007), miR159-MYB transcription factor (Zhu et al., 2011a), miR160-auxin response factor (ARF), miR164-NAC domain protein, miR166-Homeodomain Leucine Zipper protein III (HDZIP III) (Xie et al., 2003), miR168-AGO protein, miR394-F-box family protein, a well characterized miRNA-target pairs

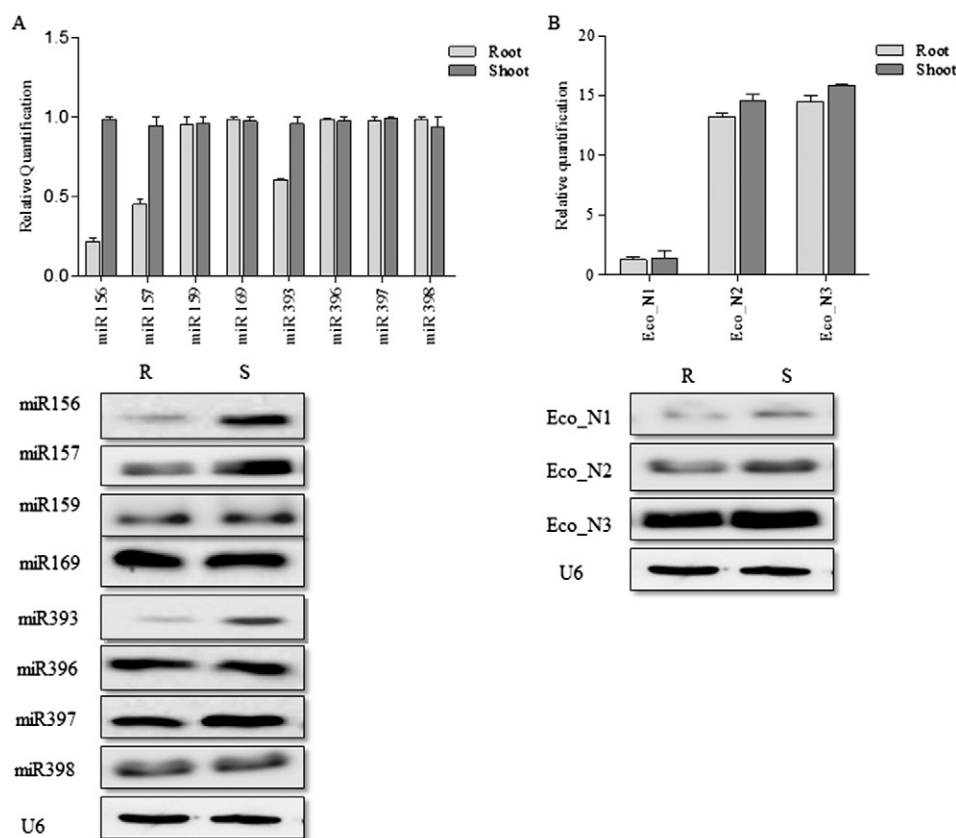


Fig. 4. Tissue specific RT-qPCR and Northern blot analysis for miRNAs in Finger millet A) Conserved miRNAs B) Novel miRNAs.

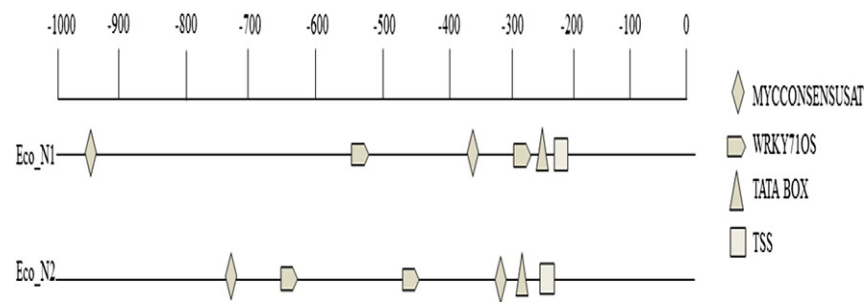


Fig. 5. Positional preferences of TSS, TATA BOX and cis regulatory elements in novel Finger millet miRNAs.

in other plants have also been found in Finger millet. However, a number of widely studied miRNA–target pairs, such as mir162–Dicer-like 1 (DCL1) (Aung et al., 2006), and mir399–E2 ubiquitin conjugating proteins pair were not predicted. miR162 targeted light mediated development protein and miR399 had ATPase as its target. This may be due to non-availability of Finger millet genome. miR444 targeted F-box and HAD proteins, in Rice F-box proteins are involved in temporal and spatial gene expression during panicle and seed development (Chen et al., 2007). miR2118 had NB-ARC domain-containing disease resistance protein as target, NB-ARC domain is a functional ATPase domain, and its nucleotide-binding state is proposed to regulate activity of the R protein and act as molecular switch depending on the nucleotide bound. In rice NB-ARC domain binds to the promoter OsWRKY13 and regulate its expression to achieve disease resistance (Qiu et al., 2007). Eco_N1 was predicted to target RCP4 protein a subunit of DNA-directed RNA polymerase III which is believed to play important role in translation process. With *S. italic* transcripts no hypothetical targets were defined for other two miRNAs. Identification of cis-elements or TFBSs of miRNAs will provide additional knowledge on the regulation of these molecules, which could be predicted through computational tools. Initially 2 kb of novel miRNA upstream sequences were retrieved and scanned for TSS. TSS was predicted using TSSP tool and it was observed that miRNAs had TSS at –200 bp region. The basic transcription elements like TATA box was also predicted and the distance between TSS and TATA box was 33 bp, the results were consistent with the reports of Zhao et al. (2007). 42% of the *Arabidopsis* miRNA genes possessed TATA box which was studied by Xie et al. (2005). All the identified cis-regulatory motifs in the study were conserved, which supports the concept of miRNA conservation among grass species. All the 11 motifs predicted was present in more than 50% of the analyzed sequence, WRKY71OS was one of the motifs identified which is the binding site for WRKY and thought to play role in stress tolerance. Liu et al. (2007) showed that in rice WRKY71OS is a binding

site for OsWRKY71 which was involved in plant defense and up-regulated by several defense signaling molecules, wounding and pathogen infection. MYCCONSUSUSAT was another conserved motif identified which is a binding site for MYC, in *Arabidopsis* MYC recognition site is found in the promoters of the dehydration-responsive gene rd22, which function as cis-acting elements in the drought- and ABA-induced gene expression of rd22 (Abe et al., 2003). cis-regulatory motifs were also identified using MELINA-II and two strong motifs of undefined functions were predicted. Thus the promoters identified in the study may be beneficial for in-depth studies on core promoters in miRNA genes of Finger millet in future.

5. Conclusion



The present study shows, for the first time, a diverse set of miRNAs that are found in Finger millet. The potential target designation suggested important biological function. Validation of high throughput sequencing and functional annotations provided a deeper understanding of the transcription and regulation of the target genes. To gain additional function insights into the miRNA target interaction, the putative targets were subjected to the GO analysis. Gene annotations showed that conserved and novel miRNAs may play an important role in growth and development.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.gene.2015.08.007>.

Authors contributions

Usha conceived and designed experiments. Usha and JMN conducted experiments. NB contributed New Reagents. Usha, RD and VRD analyzed data. SN conducted northern blot studies. NB wrote the manuscript. All authors read and approved the manuscript.

Table 4
Known cis-regulatory elements identified for novel miRNAs in Finger millet.

No	Motif logo/motif	Motif sequence	Algorithm	Function
M1	ACGTATERDI	ACGT	PLACE	Required for etiolation-induced expression of erd1 (early responsive to dehydration) in <i>Arabidopsis</i>
M2	CAATBOX1	CAAT	PLACE	CAAT promoter consensus sequence found in legA gene of pea
M3	CACFTTPPCA1	YACT	PLACE	Component of Mem1 (mesophyll expression module 1)
M4	DOFCOREZM	AAAG	PLACE	Core site required for binding of Dof proteins in maize
M5	EBOXBNNAPA	CANNTG	PLACE	E-box of napA storage-protein gene of Brassica napus
M6	GATABOX	GATA	PLACE	GATA motif in CaMV 35S promoter; Binding with ASF-2
M7	GTGANTG10	GTGA	PLACE	GTGA motif found in the promoter region
M8	MYCCONSUSUSAT	CANNTG	PLACE	MYC recognition site found in the promoters of the dehydration-responsive gene
M9	WRKY71OS	TGAC	PLACE	Binding site of rice WRKY71
M10		AGTAGATG	CONSENSUS	Undefined
M11		ATCTTTTG	CONSENSUS	Undefined

Conflict of interest

The authors declare that they have no conflict of interest. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The research work is supported by the SERB [SR/FT/LS-10/2012] New Delhi, India.

References

- Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2003. *Arabidopsis*AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signalling. *Plant Cell* 15, 63–78.
- Aung, K., Lin, S.I., Wu, C.C., Huang, Y.T., et al., 2006. *pho2*, a phosphate over accumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol.* 141 (3), 1000–1011.
- Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Bennetzen, J.L., Schmutz, J., Wang, H., et al., 2012. Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.* 13 (6), 555–561.
- Chen, C., Ridzon, D.A., Broomer, A.J., et al., 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* 33 (20), 1–9.
- Chen, Yuan, Xu, Yunyuan, Luo, Wei, Li, Wenxuan, et al., 2007. F-Box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* 143 (4), 1467–1483.
- Chi, X., Yang, Q., Chen, X., Wang, J., et al., 2011. Identification and characterization of microRNAs from peanut (*Arachis hypogaea* L.) by high throughput sequencing. *PLoS One* 6, e27530.
- Chiu, Y.H., MacMillan, J.B., Chen, Z.J., 2009. RNA polymerase III detects cytosolic DNA and induces type-I interferons through the RIG-I pathway. *Cell* 138 (3), 576–591.
- Dai, Xinbin, Zhao, Patrick X., 2011. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* 39, W155–W159.
- Ding, D., Zhang, L., Wang, H., Liu, Z., Zhang, Z., Zheng, Y., 2009. Differential expression of miRNAs in response to salt stress in maize roots. *Ann. Bot.* 103 (1), 29–38.
- Ge, A., Shangguan, L., Zhang, X., Dong, Q., Han, J., et al., 2012. Deep sequencing discovery of novel and conserved microRNAs in strawberry (*Fragaria × ananassa*). *Physiol. Plant.* 148, 387–396.
- Gruber, Andreas R., Lorenz, Ronny, et al., 2008. The Vienna RNA Websuite. *Nucleic Acids Res.* 36, W70–W74.
- Higo, K., Ugawa, Y., Iwamoto, M., Korenaga, T., 1999. Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res.* 27 (1), 297–300.
- Islam, Md. Tariqul, Ferdous, Ahlan Sabah, Najnin, Rifat Ara, et al., 2015. High-throughput sequencing reveals diverse sets of conserved, non conserved, and species-specific miRNAs in Jute. *Int. J. Genomics*, 125048 <http://dx.doi.org/10.1155/2015/125048> (14 pp.).
- Jiang, J., Yang, Y., Cao, J., 2013. Identification of microRNAs potentially involved in male sterility of *Brassica campestris* ssp. *chinensis* using microRNA array and quantitative RT-PCR assays. *Cell. Mol. Biol. Lett.* 18 (3), 416–432.
- Khan, Y., Yadav, A., Suresh, B.V., Muthamilarasan, M., Yadav, C.B., et al., 2014a. Comprehensive genome-wide identification and expression profiling of foxtail millet [*Setaria italica* (L.)] miRNAs in response to abiotic stress and development of miRNA database. *Plant Cell Tissue Organ Cult.* 118, 279–292.
- Khan, Yusuf, Yadav, Chandra Bhan, Muthamilarasan, Mehanathan, et al., 2014b. Development of novel microRNA-based genetic markers in foxtail millet for genotyping applications in related grass species. *Mol. Breed.* 34 (4), 2219–2224.
- Khraiwesh, B., Pugalenth, G., Fedoroff, N.V., 2013. Identification and analysis of red sea mangrove (*Avicennia marina*) microRNAs by high-throughput sequencing and their association with stress responses. *PLoS ONE* 8 (4), e60774.
- Klevebrink, D., Street, N.R., Fahlgren, N., et al., 2009. Genome-wide profiling of *Populus* small RNAs. *BMC Genomics* 10, 620.
- Kozomara, A., Griffiths-Jones, S., 2014. MiRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42 (1), D68–D73.
- Kuijt, Suzanne J.H., Greco, Raffaella, Agalou, Adamantia, et al., 2014. Interaction between the GROWTH-REGULATING FACTOR and KNOTTED1-LIKE HOMEODOMAIN families of transcription factors. *Plant Physiol.* 164 (4), 1952–1966.
- Kulcheski, F.R., de Oliveira, L.F., Molina, L.G., Almerao, M.P., Rodrigues, F.A., Marcolino, J., et al., 2011. Identification of novel soybean microRNAs involved in abiotic and biotic stresses. *BMC Genomics* 12, 307.
- Li, H., Dong, Y., Yin, H., Wang, N., Yang, J., Liu, X., et al., 2011. Characterization of the stress associated microRNAs in Glycine max by deep sequencing. *BMC Plant Biol.* 11, 170.
- Li, X., Bian, H., Song, D., Ma, S., Han, N., et al., 2013. Flowering time control in ornamental glloxinia (*Sinningia speciosa*) by manipulation of miR159 expression. *Ann. Bot.* 111 (5), 791–799.
- Liu, J.Z., Whitham, S.A., 2013. Overexpression of a soybean nuclear localized type-III DnaJ domain-containing HSP40 reveals its roles in cell death and disease resistance. *Plant J.* 74, 110–121.
- Liu, X., Bai, X., Wang, X., Chu, C., 2007. OsWRKY71, a rice transcription factor, is involved in rice defense response. *J. Plant Physiol.* 164, 969–979.
- Ma, J., Guo, T., Wang, Q., Wang, K., Sun, R., Zhang, B., 2015. Expression profiles of miRNAs in *Gossypium raimondii*. *J. Zhejiang Univ. Sci. B* 16 (4), 296–303.
- Megraw, M., Baev, V., Rusinov, V., Jensen, S.T., Kalantidis, K., Hatzigeorgiou, A.G., 2006. MicroRNA promoter element discovery in *Arabidopsis*. *RNA* 12, 1612–1619.
- Morin, R.D., Aksay, G., Dolgosheina, E., Ebhardt, H.A., Magrini, V., et al., 2008. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res.* 18, 571–584.
- Moxon, S., Jing, R., Szitty, G., et al., 2008. Deep sequencing of tomatoshort RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res.* 18 (10), 1602–1609.
- Okumura, T., Makiguchi, H., Makita, Y., Yamashita, R., Nakai, K., 2007. Melina II: a web tool for comparisons among several predictive algorithms to find potential motifs from promoter regions. *Nucleic Acids Res.* 35, 227–231.
- Pantaleo, V., Szitty, G., Moxon, S., et al., 2010. Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J.* 62 (6), 960–976.
- Park, W., Li, J., Song, R., Messing, J., Chen, X., 2002. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Current Biology* 12 (17), 1484–1495.
- Paul, S., Kundu, A., Pal, A., 2014. Identification and expression profiling of *Vigna mungo* microRNAs from leaf small RNA transcriptome by deep sequencing. *J. Integr. Plant Biol.* 56 (1), 15–23.
- Qiu, D., Xiao, J., Ding, X., Xiong, M., Cai, M., Cao, Y., Li, X., Xu, C., Wang, S., 2007. OsWRKY13 mediates rice disease resistance by regulating defense-related genes in salicylate- and jasmonate-dependent signalling. *Mol. Plant Microbe Interact.* 20 (5), 492–499.
- Quail, M.A., Kozarewa, I., Smith, F., Scally, A., et al., 2010. A large genome center's improvements to illumine sequencing system. *Nat. Methods* 5 (12), 1005–1010.
- Rajagopalan, R., Vaucheret, H., Trejo, J., Bartel, D.P., 2006. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes* 20 (24), 3407–3425.
- Reyes, J.L., Chua, N.H., 2007. ABA induction of miR159 controlstranscript levels of two MYB factors during *Arabidopsis* seed germination. *Plant J.* 49 (4), 592–606.
- Song, C., Wang, C., Zhang, C., et al., 2010. Deep sequencing discovery of novel and conserved microRNAs in trifoliolate orange (*Citrus trifoliata*). *BMC Genomics* 11 (1), 431.
- Song, Q.X., Liu, Y.F., Hu, X.Y., et al., 2011. Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. *BMC Plant Biol.* 11, 5.
- Sunkar, R., Kapoor, A., Zhu, J.K., 2006. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell* 18 (8), 2051–2065.
- Sunkar, R., Zhou, X., Zheng, Y., Zhang, W., Zhu, J.K., 2008. Identification of novel and candidate miRNAs in rice by high throughput sequencing. *BMC Plant Biol.* 8, 25.
- Szitty, G., Moxon, S., Santos, D.M., et al., 2008. High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. *BMC Genomics* 9, 593.
- Tian, Y., Tian, Y., Luo, X., Zhou, T., Huang, Z., Liu, Y., et al., 2014. Identification and characterization of microRNAs related to salt stress in broccoli, using high-throughput sequencing and bioinformatics analysis. *BMC Plant Biol.* 14 (1), 226.
- Voinnet, O., 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669–687.
- Xie, Z., Kasschau, K.D., Carrington, J.C., 2003. Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. *Curr. Biol.* 13 (9), 784–789.
- Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., Carrington, J.C., 2005. Expression of *Arabidopsis* MIRNA genes. *Plant Physiol.* 138, 2145–2154.
- Xie, K., Wu, C., Xiong, L., 2006. Genomic organization, differential expression, and interaction of SQUAMOSA promoter binding-like transcription factors and microRNA156 in rice. *Plant Physiol.* 142 (1), 280–293.
- Xu, L., Wang, Y., Xu, Y., Wang, L., Zhai, L., et al., 2013. Identification and characterization of novel and conserved microRNAs in radish (*Raphanus sativus* L.) using high-throughput sequencing. *Plant Sci.* 201–202, 108–114.
- Yadav, Chandra Bhan, Muthamilarasan, Mehanathan, Pandey, Garima, Khan, Yusuf, Prasad, Manoj, 2014. Development of novel microRNA-based genetic markers in foxtail millet for genotyping applications in related grass species. *Mol. Breed.* 34 (4), 2219–2224.
- Ye, J., Fang, L., et al., 2006. *Nucleic Acids Res.* 34 (Web service issue), 293–297.
- Yi, F., Xie, S., Liu, Y., Qi, X., Yu, J., 2013. Genome-wide characterization of microRNA in foxtail millet (*Setaria italica*). *BMC Plant Biol.* 13, 212.
- Yu, B., Wang, H., 2010. Translational inhibition by microRNAs in plants. *Prog. Mol. Subcell. Biol.* 50, 41–57.
- Zhang, B.H., Pan, X.P., Cannon, C.H., Cobb, G.P., Anderson, T.A., 2006a. Conservation and divergence of plant microRNA genes. *Plant J.* 46, 243–259.
- Zhang, B.H., Pan, X.P., Cox, S.B., Cobb, G.P., Anderson, T.A., 2006b. Evidence that miRNAs are different from other RNAs. *Cell. Mol. Life Sci.* 63, 246–254.
- Zhang, J.Z., Ai, X.Y., Guo, W.W., Peng, S.A., Deng, X.X., Hu, C.G., 2012. Identification of miRNAs and their target gene using deep sequencing and degradome analysis in trifoliolate orange [*Poncirus trifoliata* (L.) Raf.]. *Mol. Biotechnol.* 51 (1), 44–57.
- Zhao, B., Liang, R., Ge, L., et al., 2007. Identification of drought-induced microRNAs in rice. *Biochem. Biophys. Res. Commun.* 354, 585–590.
- Zhao, C.Z., Xia, H., Frazier, T.P., et al., 2010. Deep sequencing identifies novel and conserved microRNAs in peanuts (*Arachis hypogaea* L.). *BMC Plant Biol.* 10, 3.
- Zhu, H., Hu, F., Wang, R., et al., 2011a. *Arabidopsis* argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* 145 (2), 242–256.
- Zhu, R.S., Li, X., Chen, Q.S., 2011b. Discovering numerical laws of plant microRNA by evolution. *Biochem. Biophys. Res. Commun.* 415, 313–318.