

RESEARCH ARTICLE

Open Access

Genome-wide characterization of microRNA in foxtail millet (*Setaria italica*)

Fei Yi[†], Shaojun Xie[†], Yuwei Liu, Xin Qi and Jingjuan Yu^{*}

Abstract

Background: MicroRNAs (miRNAs) are a class of short non-coding, endogenous RNAs that play key roles in many biological processes in both animals and plants. Although many miRNAs have been identified in a large number of organisms, the miRNAs in foxtail millet (*Setaria italica*) have, until now, been poorly understood.

Results: In this study, two replicate small RNA libraries from foxtail millet shoots were sequenced, and 40 million reads representing over 10 million unique sequences were generated. We identified 43 known miRNAs, 172 novel miRNAs and 2 mirtron precursor candidates in foxtail millet. Some miRNA*s of the known and novel miRNAs were detected as well. Further, eight novel miRNAs were validated by stem-loop RT-PCR. Potential targets of the foxtail millet miRNAs were predicted based on our strict criteria. Of the predicted target genes, 79% (351) had functional annotations in InterPro and GO analyses, indicating the targets of the miRNAs were involved in a wide range of regulatory functions and some specific biological processes. A total of 69 pairs of syntenic miRNA precursors that were conserved between foxtail millet and sorghum were found. Additionally, stem-loop RT-PCR was conducted to confirm the tissue-specific expression of some miRNAs in the four tissues identified by deep-sequencing.

Conclusions: We predicted, for the first time, 215 miRNAs and 447 miRNA targets in foxtail millet at a genome-wide level. The precursors, expression levels, miRNA* sequences, target functions, conservation, and evolution of miRNAs we identified were investigated. Some of the novel foxtail millet miRNAs and miRNA targets were validated experimentally.

Keywords: miRNA, Foxtail millet, Expression pattern, Mirtron, miRNA*, Targets, Synteny

Background

MicroRNAs (miRNAs) are 20 ~ 22 nucleotide (nt) non-coding small RNAs (smRNAs). Most primary miRNAs (pri-miRNAs) are transcribed from miRNA genes by RNA polymerase II (RNAPol II), while others are transcribed by RNA polymerase III (RNAPol III) [1]. In plants, the nascent pri-miRNA transcripts are first processed into 60 to 500 nt pre-miRNAs by DICER-LIKE1 (DCL1) [2]. Then, the pre-miRNAs undergo a second cleavage in the nucleus, releasing an RNA duplex containing the mature miRNA and the miRNA* sequences. To maintain the correct size and/or to protect them from polyuridylation [2,3], the duplexes are methylated at the 3' ends by HUA ENHANCER [4] and then transported into the cytoplasm by HASTY [5,6]. The mature miRNA is incorporated into an RNA-induced silencing complex (RISC) that contains

the ARGONAUTE (AGO) protein. The complex recognizes partially complementary sequences in the target messenger RNA (mRNA) where it binds, inducing its target mRNA's degradation or a combination of both degradation and translation inhibition [7]. The miRNA* is released and rapidly degraded (had ~9% as many reads as the mature miRNAs) [8]. The presence of the miRNA* has been regarded as a good standard way to reliably annotate a novel miRNA [9].

Mirtrons are a new, recently discovered, type of miRNAs that originate from spliced introns of gene transcripts. Mirtrons reside within the intronic regions of genes and are processed through a Drosha-independent pathway, making them quite distinct from the other miRNAs. The mirtron pathway was first identified in *Drosophila melanogaster*, *Caenorhabditis elegans*, and some mammals [10-12] and has recently been found in *Arabidopsis thaliana* and *Oryza sativa* [13,14]. Recent research shows that the miRNAs from the intronic regions of genes can survive together with their host genes and support their host

* Correspondence: yujj@cau.edu.cn

[†]Equal contributors

State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100193, China

genes by mediating synergistic and antagonistic regulatory effects [15,16].

In plant, miRNAs are known to play crucial roles in many developmental events [17] and regulate target transcripts through two modes of action: degradation and translation inhibition [18,19]. The miRNA degradation occurs through miRNA-guided transcript cleavage in plants [20]. Although the mechanisms involved in the translational inhibition by miRNAs are largely unknown, recent studies from Xuemei Chen's group show that the miRNA translational inhibition occurs at the endoplasmic reticulum (ER) and that ALTERED MERISTEM PROGRAM1 (AMP1) activity allows for the dissection of miRNA-mediated target RNA cleavage from miRNA-mediated target translation inhibition [7].

With the advent of second generation sequencing technology, the rate of miRNA discovery has increased dramatically [21-25]. In dicotyledonous species, 337 mature miRNAs in *A. thaliana*, 401 in *Populus trichocarpa*, and 164 in *Nicotiana tabacum* have been discovered, and, in monocotyledonous species, 713 mature miRNAs in *O. sativa*, 16 in *Saccharum officinarum*, 241 in *Sorghum bicholor*, 42 in *Triticum aestivum* and 321 in *Zea mays* have been reported (miRBase; release 20, June 2013). Although foxtail millet is an important cereal crop, only a few miRNAs in foxtail millet have been reported. Bennetzen *et al.* [26] reported some known miRNAs from the alignment of mature miRNA sequences, but they did not investigate whether or not the corresponding precursors existed in the genome.

Foxtail millet is a diploid C4 panicoid crop species [27]. It is an important grass crop (family Poaceae) that has been planted widely in China. The genome of foxtail millet (Yugu1) has been sequenced [26], making it a possible genetic resource that could be used to investigate plant architecture, genome evolution and physiology in the bioenergy grasses. Because of the important roles that miRNAs play in gene regulatory networks [17], it is important to identify the miRNAs in foxtail millet and to investigate their potential target genes to gain a better understanding of the biological processes in this plant. In the present study, we aimed to characterize the miRNA repertoire of foxtail millet.

Results

Analysis of smRNAs in foxtail millet

To identify miRNAs in foxtail millet, we sequenced smRNAs for shoots (14-day-old) by high throughput Illumina sequencing technology [28]. Nearly 40 million reads of 18 ~ 31 nt, representing over 10 million unique sequences, were generated. In addition, three smRNA datasets (leaf, flower and root) of foxtail millet were downloaded from the Comparative Sequencing of Plant Small RNAs web site [29]. These datasets contained

over 14 million smRNAs of 18 ~ 31 nt, representing over 4 million unique sequences. An analysis of the data from the four datasets (shoot, leaf, flower and root) revealed that a large number of sequences appeared only once. The percentages of these singletons were 84.03% (8,851,929) in shoot, 85.49% (792,474) in leaf, 83.43% (1,972,156) in flower and 82.53% (948,382) in root. In previous reports, only 65% and 82% singletons were found in *Arabidopsis* [8] and rice [22], respectively, suggesting that the smRNAs in foxtail millet are as complex as they are in rice, and more complex than in *Arabidopsis*.

The overall size distribution of the unique reads from four sequencing efforts were very similar, with the 24 nt smRNAs being the most abundant, followed by the 23 nt and 21 nt smRNAs (Figure 1), which differ from the size distribution of total reads (Additional file 1A). A detailed comparison of the smRNAs derived from unique reads (Figure 1) or total reads (Additional file 1A) revealed some features of smRNA species. First, the 24 nt smRNAs dominated the pool of unique species in foxtail millet as observed for many other species such as *A. thaliana* [30] and cucumber [31]. Second, the 21 nt smRNAs replicated in even higher frequencies (ratio of total reads to unique reads) than 24 nt smRNAs in all four tissues. The average times (total reads/unique reads) of 21 nt are nearly 7.8, 9.6, 6.0 and 9.4 and the average times of 24 nt are nearly 2.1, 1.9, 1.9 and 2.0 in shoot, leaf, flower and root, respectively. Third, miRNAs, most of which were 20 to 22 nt in length, were relatively abundant in root and leaf, while siRNAs, most of which were 24 nt long, were relatively more prevalent in flower and shoot. These results were consistent with those reported previously in maize [32,33].

In foxtail millet, smRNA populations in four tissues are extremely complex. Although we have over 54 million reads, sequences of smRNAs have a limited overlap from the four databases (Additional file 1B). Only 79,225 unique sequences appeared in all four tissues and a small portion of the sequences overlapped between two tissues. The smRNAs in each tissue accounted for almost half of all the smRNAs in the combined datasets. This limited overlap indicated that there was a diversity of smRNAs in foxtail millet.

Identification of known miRNAs in foxtail millet

In plants, miRNAs are obtained by the precise excision of ~21 nt smRNAs from the stem of a single-strand, stem-loop precursor. As a result, miRNAs can be identified by looking for a potential fold-back precursor structure that contains the ~21 nt miRNA sequence within one arm of the hairpin. The hairpin must have the lowest free energy of all the alternative folds for that sequence, as predicted by RNA folding programs such as Mfold [34] or RNAfold

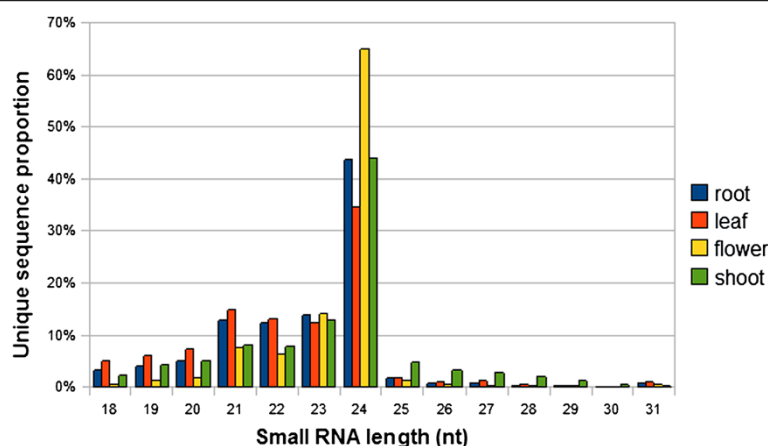


Figure 1 Distribution of the lengths of the unique smRNA sequences. Distribution of the sequence lengths of the smRNAs derived from root, leaf, flower and shoot generated by deep-sequencing. Counts were based on the unique sequences rather than the number of reads per unique sequence.

[21,32]. The minimal folding free energies index (MFEI) can be treated as the main feature to distinguish potential miRNA precursors from other RNAs. It has been reported that more than 90% of miRNA precursors have a MFEI value greater than 0.85, while the MFEI for other RNAs [tRNAs (0.64), rRNAs (0.59) and mRNAs (0.65)] are lower [35]. We used a previously reported workflow [21] to identify miRNAs among the smRNAs from foxtail millet, but with the MFEI values set to 0.85 rather than to 0.15 as was done previously. We identified 215 candidate miRNAs in the foxtail millet smRNA datasets; 137 in shoot, 64 in leaf, 62 in flower and 79 in root. The proportion of miRNA reads (those identified in the study) in each library as compared to the total reads were 3.0% in root, 4.1% in leaf, 1.4% in flower and 1.3% in shoot.

Among the 215 miRNAs, 43 unique mature miRNA sequences belonging to 19 miRNA families were 100% identical to mature miRNAs from 14 well-studied species [26]. We defined the 43 unique mature miRNAs in foxtail millet as known miRNAs (hereafter referred to as sit-miRNAs). There were 24 known miRNAs in root, 23 in leaf, 19 in flower and 35 in shoot tissues (Additional file 2). The 43 sit-miRNAs corresponded to 89 pre-miRNAs, of which 65 had corresponding miRNA* forms. Detailed information about the known sit-miRNAs is shown in Additional file 2.

A summary of the total expression profiles for each of the 19 miRNA families to which the 43 sit-miRNAs belong is shown in Figure 2. The number of miRNA reads were enumerated and normalized against the total count of smRNA reads, reported as reads per million (RPM), for each respective library. The results showed that the total expression of these 19 miRNA families was relatively higher in leaf and root than that in shoot and flower. We noted that the sit-miR156, sit-miR164,

sit-miR166, sit-miR167 and sit-miR172 families showed relatively higher expression (slightly over 1,000 RPM, on average) in one or more of the four tissues. In contrast, the sit-miR319, sit-miR390, sit-miR394, sit-miR399 and sit-miR2118 families showed low expression (less than 100 RPM) levels. In our samples, the following miRNA families showed distinct expression patterns: miR529 is barely detected in roots; miR319 is practically absent in leaves; miR395 and miR397 are practically absent in flowers; miR398 and miR399 are detected only in shoots; and miR2118 is barely detectable in any of the four tissues.

Our data also showed that some members of the same miRNA family had different tissue-specific expression patterns (Figure 3). For example, based on RPM counts, miR156b is highly expressed in flower and shoot but has almost no expression in leaf and root. While the miR156d is highly expressed in leaf and root but has almost no expression in the flower. This should be confirmed by other molecular techniques. Together, these results suggest that members from different families and different members from the same family may have greatly different effects on foxtail millet development.

Novel miRNA identification

Our results showed that in addition to the 43 known mature sit-miRNAs, there were 172 novel miRNAs (hereafter referred to as nov-sit-miRNAs) in the smRNA datasets that have never been reported in other species. Because some of these miRNAs were derived from multiple precursors, these 172 nov-sit-miRNAs corresponded to 212 pre-miRNAs, 68% of which were located in intergenic regions. Only a few of the 212 pre-RNAs were located in introns or in the UTRs of coding genes, which is in

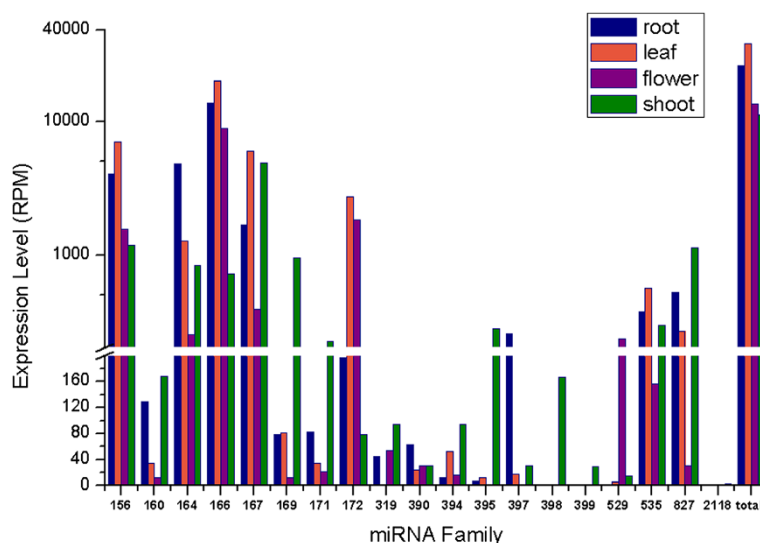


Figure 2 Expression levels of 19 known miRNA families in four tissues from foxtail millet. The expression levels of the miRNA families in each tissue were normalized by the total number of reads in each of the respective libraries. Counts were in reads per million (RPM).

agreement with the reported statistical characteristics of plant pre-miRNAs [3] (Additional file 3).

Among the 172 nov-sit-miRNAs, nine were found in all four libraries, nine were in three libraries, 25 were in two and 129 were in one library. These potentially nov-sit-miRNAs had relatively lower expression levels than the 43 known sit-miRNAs. For example, in leaf, the total expression levels of the 172 nov-sit-miRNAs (2267 RPM) were less than 1% of the total expression levels of the 43 sit-miRNAs (38438 RPM) and, in flower, the total expression levels of all the nov-sit-miRNAs were only 659 RPM. These results were consistent with a study in rice that found that most of the non-conserved miRNAs exhibited tissue-specific expression patterns and had relatively low expression levels compared with known miRNAs [22]. For example, nov-sit-miR43, nov-sit-miR125 and nov-sit-miR159 were expressed only in flowers with expression levels of more than 30 RPM, and nov-sit-miR104 and nov-sit-miR149 were expressed only in roots with expression levels of more than 180 RPM (Additional file 3).

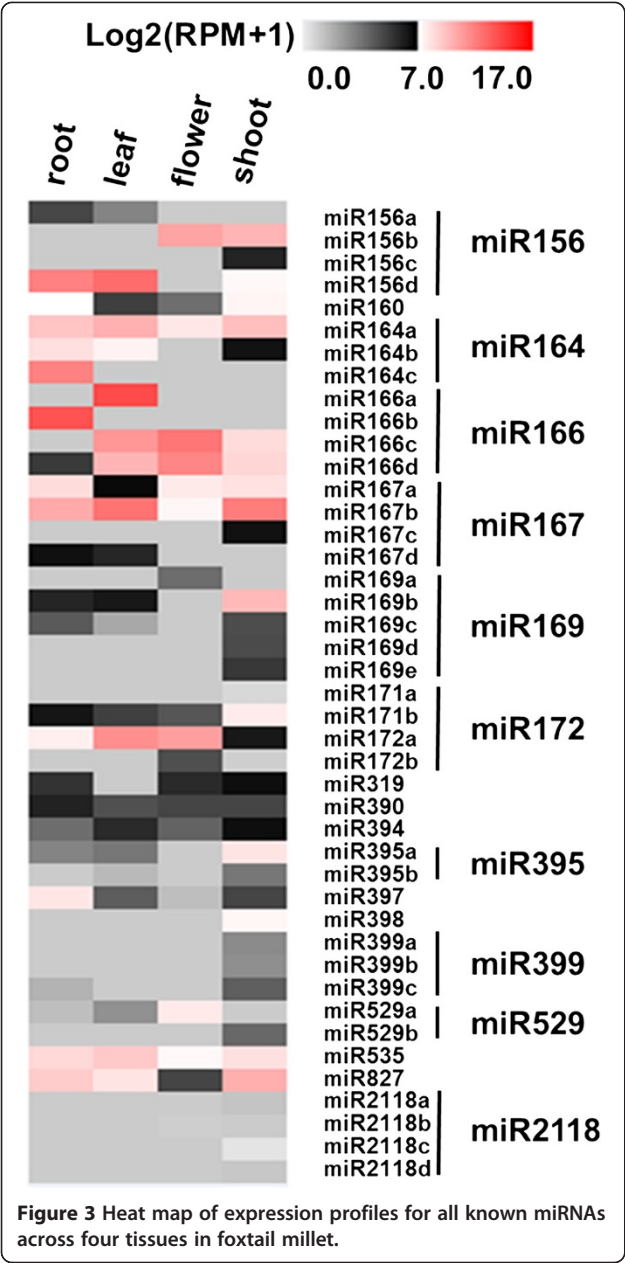
The conservation of miRNA sequences across species could be regarded as powerful evidence to annotate miRNA [9]. We identified conserved miRNA by allowing no more than three mismatches between the mature miRNA sequences in our datasets and the mature miRNA sequences from 14 well-studied species [26] published in miRBase (release 20, June 2013). We identified 33 nov-sit-miRNAs that could be assigned to 14 published miRNA families based on sequence similarity (Additional file 3). These 'conserved' nov-sit-miRNAs generally had higher expression levels than the 'non-conserved' nov-sit-miRNAs. Eight of the nov-sit-miRNAs were validated

experimentally by stem-loop RT-PCR and sequencing (Figure 4A). The presence of a miRNA* could be regarded as additional support for the annotation of the nov-sit-miRNAs [9]. We noted that 87 of the 212 novel pre-miRNAs were found having miRNA*s. A nucleotide composition analysis showed that 50% of the novel pre-miRNAs started with U (Figure 4B and Additional file 3), which is consistent with the statistical characteristic of mature miRNAs [36], and, overall, a higher percentage of A and U compared with C and G was present. Detailed information about novel miRNAs was shown in Additional file 3 and second structure of some novel miRNA precursors were presented in Additional file 4.

To predict mirtrons, we analyzed the secondary structure features of the introns of the foxtail millet genes that were annotated in the Phytozome database (see the Methods section for details). Introns can be folded directly to acquire stem-loop structures similar to the pre-miRNAs. Using RNAfold, we found many introns that had reasonable second structures; that is, a single stem-loop with less than 3 nt overhangs at both ends, and a minimum energy of -25 kcal/mol or less. After mapping the smRNAs in our datasets onto the selected introns using the BLAST algorithm (see the Methods section for detail), we found only a few perfectly matched smRNAs; therefore, only two mirtron precursor candidates could be identified (PAC: 19709968 and PAC: 19675165) (Additional file 5).

Target prediction and expression validated

MiRNAs are known to have diverse expression patterns and play regulatory roles in various developmental and physiological processes [17]. In plants, most miRNAs are known to regulate development by mediating the



destruction of their target mRNAs [37]. The target gene has sites that are almost completely complementary to the miRNA, so, in plants, miRNA target genes can be predicted computationally. To increase the reliability of our target prediction, we used the modified scoring described by Sun *et al.* [38]. We used the sit-miRNA and nov-sit-miRNA sequences in our datasets to search for complementary annotated foxtail millet transcript genes in the Phytozome database [39] and identified 447 candidate target genes for 166 of the foxtail millet miRNAs (Additional file 6).

We randomly picked six target genes to validate using 5'RACE and four were successfully validated (Si016508m, Si005991m, Si016509m and Si034525m) (Table 1). The Si016508m protein, which is encoded by one of the predicted target genes, is a GRAS (GAL, RGA, SCR) transcription factor that has been reported to be involved in development and other processes [40,41]. The Si016508m gene is potentially targeted by five of the sit-miRNAs (sit-miR171a, sit-miR171b, nov-sit-miR15, nov-sit-miR14 and nov-sit-miR49). The Si005991m, Si016509m and Si034525m protein, which are encoded by a predicted target gene (targeted by sit-miR160), is known to have three different domains: a B3 DNA-binding domain, an Aux/IAA-ARF dimerization motif, and a DNA-binding pseudobarrel domain. Previous studies have suggested that for efficient cleavage by miRNAs, base pairing between bases 10 and 11 is essential [42-45]. The 5'RACE results showed that the Si016508m gene had two breakpoints (at positions 1085 bp and 1082 bp). The predicted target sites of these genes showed that sit-miR171a and sit-miR171b guided Si016508m cleavage at 1085 bp, while nov-sit-miR14, nov-sit-miR15 and nov-sit-miR49 guided Si016508m cleavage at 1082 bp (Table 1). To show that these miRNAs, which have validated their targets, are actually expressed, we used the stem-loop RT-PCR method of detection. We detected miRNA160, miRNA171a, nov-sit-miR14,

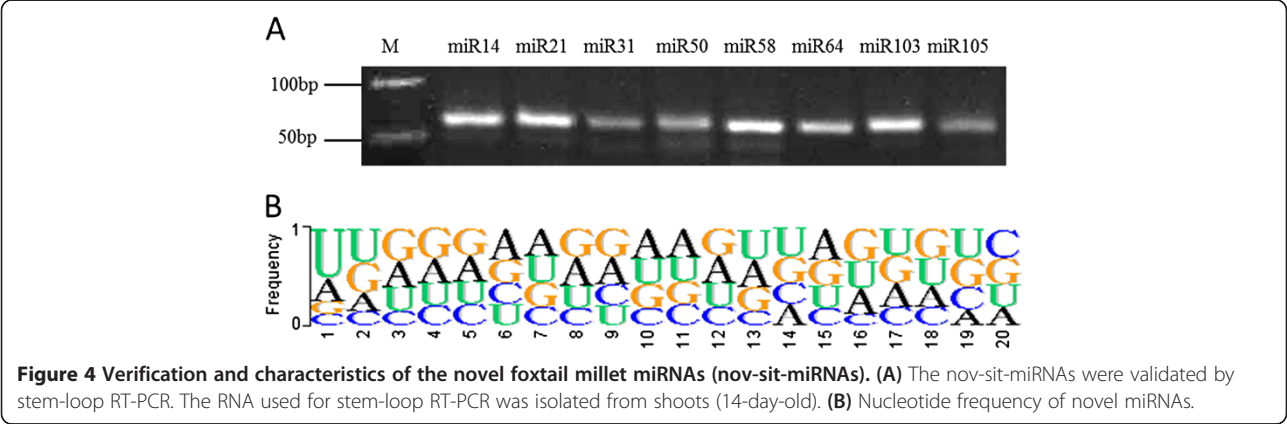


Table 1 Validated miRNA targets

Target ID	miRNA ID	Target start-end	5'-target sequence-3' complementary pattern 3'-miRNA sequence-5'	InterPro description of targets
Si005991m		1756-1776	↓1767(9/10) 5'-AGGCAUACAGGGAGCCAGGCA-3' 3'- ACCGUAUGUCCCU CGGUCCGU-5'	B3 DNA binding domain Aux/IAA-ARF-dimerisation DNA-binding pseudobarrel domain AUX/IAA protein
sit-miR160				
Si016509m		1353-1373	↓1364(9/10) 5'-AGCCAUACAGGGAGCCAGGCA-3' 3'- ACCGUAUGUCCCU CGGUCCGU-5'	AUX/IAA protein B3 DNA binding domain Auxin response factor Aux/IAA-ARF-dimerisation DNA-binding pseudobarrel domain
sit-miR160				
Si034525m		1317-1337	↓1328(6/10) 5'-AGGCAUACAGGGAGCCAGGCA-3' 3'- ACCGUAUGUCCCU CGGUCCGU-5'	AUX/IAA protein B3 DNA binding domain Auxin response factor Aux/IAA-ARF-dimerisation DNA-binding pseudobarrel domain
sit-miR160				
Si016508m		1074-1094	↓1085(7/10) 5'-GAUUAUUGGCGCGCUCAAUCA-3' 3'-CUAUAACCGCGCCGAGUUAGU-5' 3'-CUAUAACCG U GCCGAGUUAGU-5'	Transcription factor GRAS
sit-miR171a				
sit-miR171b				
Si016508m		1071-1091	↓1082(2/10) 5'-AGAGAUUAUUGGCGCGCUCAA-3' 3'-UCUCUAUAACCGCGCCGAGUU-5' 3'- CCUCUAUAACU GCGCCGAGUU-5' 3'-UC ACUAUAACCGAG CCGAGU-5'	Transcription factor GRAS
nov-sit-miR15				
nov-sit-miR14				
nov-sit-miR49				

Mapping of mRNA cleavage sites by RNA Ligase-Mediated 5' RACE. The complementary pattern of miRNA sequences and partial sequences of the target mRNAs are shown in the table. All predicted mismatch base pairings are shown in bold letters. The positions inferred as 5' ends of miRNA-guided cleavage products, as revealed by 5' RACE, and the numbers of sequenced 5' RACE clones corresponding to each site are indicated by vertical arrowheads.

nov-sit-miR15 and nov-sit-miR49 successfully. However, we failed to detect miRNA171b (Additional file 7).

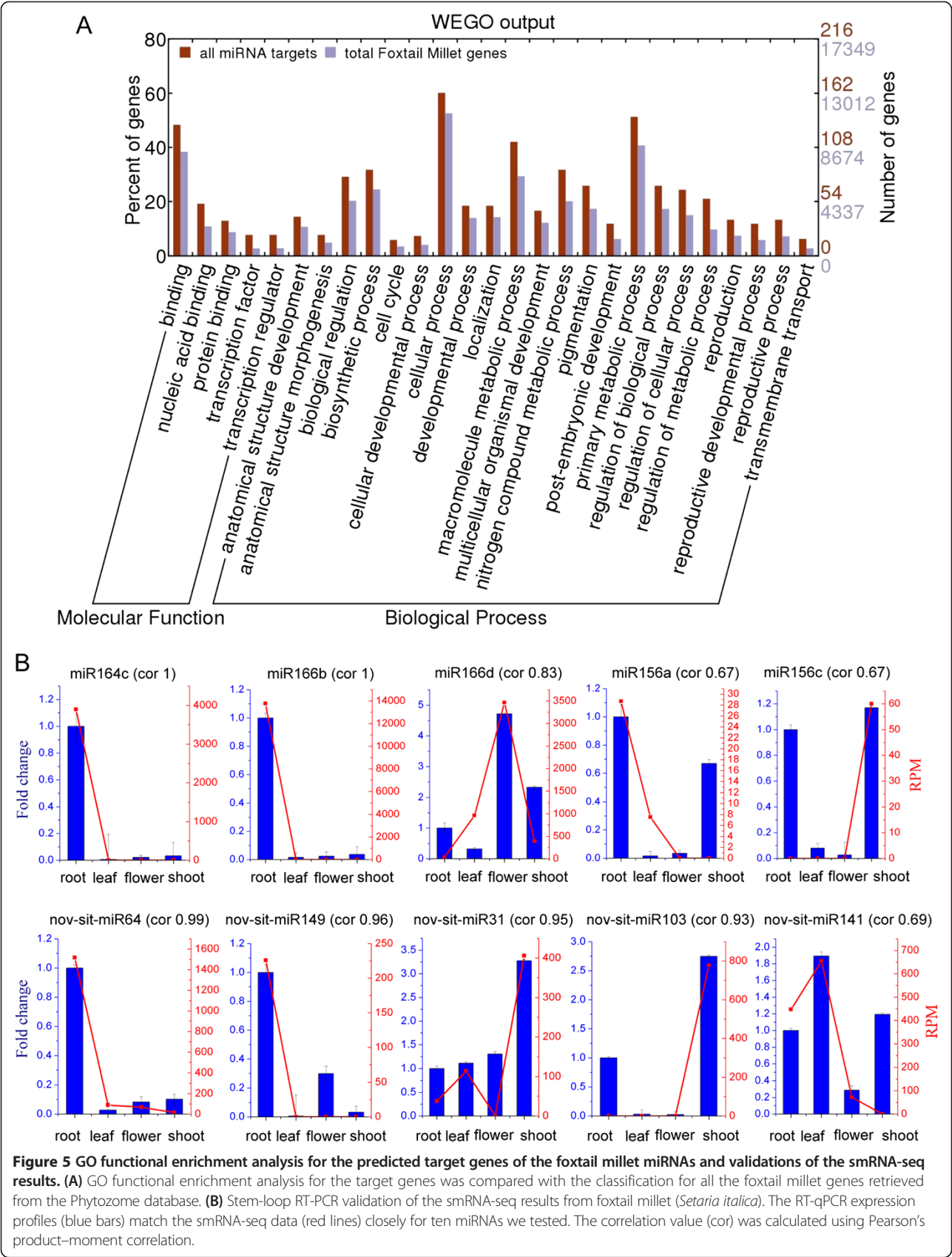
About 79% (351) of the predicted target genes had functional annotations in InterPro, the integrated resource for protein families, domains, and functional sites (Additional file 6). A large proportion of the predicted targets were annotated as transcription factors; for example, the AP2/ERF domain, homeodomain, and CCAAT-binding transcription factors. In addition, some of the predicted targets were annotated as being involved in various classes of molecular functions, including binding proteins (DNA-binding, RNA-binding, ATP-binding and protein-binding), calcium ion transport proteins, and enzymes (protein kinase, synthetase/ligase, acetate-CoA ligase, glycoside hydrolase and oxidoreductase). The functions of predicted miRNA targets indicated that the miRNAs may play multiple roles in gene regulation networks.

To better understand the functional roles of the predicted miRNA target genes in foxtail millet, we analyzed the functional enrichment for all the miRNA targets by gene ontology (GO) [46,47]. The miRNA predicted targets showed enrichment in GO terms from the molecular function and biological process categories, while no enrichment in GO terms was observed in the cellular component category. We found 22 GO terms in the biological process category that showed strong enrichment in biosynthetic process, cellular process,

macromolecule metabolic process, regulation of metabolic process and regulation of cellular process. In the molecular function category, the enriched GO terms included binding, nucleic acid binding, and protein-binding. The GO enrichment analysis showed that the predicted targets of the miRNAs were involved in a wide range of regulatory functions as well as some specific biological processes like metabolism, biosynthesis, and gene expression/transcription (Figure 5A).

To analyze whether or not miRNAs in different tissues have different functions, GO enrichment analyses were conducted for the predicted miRNA targets in each of the four tissues. We found that the functions of the targets were similar in all four tissues; the targets were all enriched in biological regulation, biosynthetic process, macromolecule metabolic process and cellular process. In addition, some of the predicted targets had annotations in the cellular component category, including cell, and cell part in leaf and flower. In the root dataset, the target genes were annotated with significantly fewer GO terms than the target genes in the other three tissues. Details of the GO annotations for the target genes in the four tissues are available in Additional file 8.

To confirm the accuracy and reliability of the smRNA-seq results, we randomly chose ten of the miRNAs (five known miRNAs and five novel miRNAs) for stem-loop RT-PCR validation (Figure 5B). The results showed that



the stem-loop RT-PCR expression profiles of most miRNAs match the smRNA-seq data closely. While there were three miRNAs expression profiles (miR156c, miR156a and nov-sit-miR141) that were not very consistent.

Conservation and synteny with sorghum

Previous reports in *Arabidopsis* [48] and in rice [49] have shown that miRNA gene families evolved from segmental, tandem repeat and whole-genome duplication events. A detailed description of the fate of miRNA genes after whole-genome duplication in maize and sorghum has been published [46].

To date, miRNAs expression and function has been well studied in the major food and feed crops, such as sorghum, maize and rice [46,50-53]. Foxtail millet shares a common ancestor with sorghum and maize (~26 Myr ago). This ancestry is more recent than with rice (~34 Myr ago) [54]. Sorghum has experienced one whole genome duplication and maize was experienced two whole genome duplications since they diverged (~13 Myr ago) [55]. To investigate the evolution of foxtail millet miRNAs, we used sorghum as the reference comparison strain. The syntenic analysis was performed between the foxtail millet and sorghum genes.

Based on the total number of genes in foxtail millet and in sorghum, we found 919 syntenic blocks containing 28,510 pairs of syntenic genes between the two species. In total, 20,090 foxtail millet and 19,877 sorghum genes, and 65 foxtail millet and 57 sorghum pre-miRNAs, were found within the syntenic regions. Nearly all the pre-miRNAs were located in only 47 (5%) of the 919 syntenic blocks, but within these 47 blocks there were 12,757 syntenic gene pairs corresponding to 45% of the total number of syntenic genes in the two species. We identified 69 pairs of syntenic pre-miRNAs, and all the foxtail millet and sorghum chromosomes had blocks that contained at least one of the conserved syntenic pre-miRNAs (Figure 6 and Additional file 9). We noted that 18 of the novel miRNAs that we identified in foxtail millet had syntenic pre-miRNA sequences and conserved mature miRNA sequences (Table 2). Therefore, in some respects, the 18 foxtail millet miRNAs are new members of already known miRNA families. For example, nov-sit-miR58 and nov-sit-miR60 could be considered to be new members of the miR166 family; similarly, nov-sit-miR14, nov-sit-miR49 and nov-sit-miR50 are new members of the miR171 family.

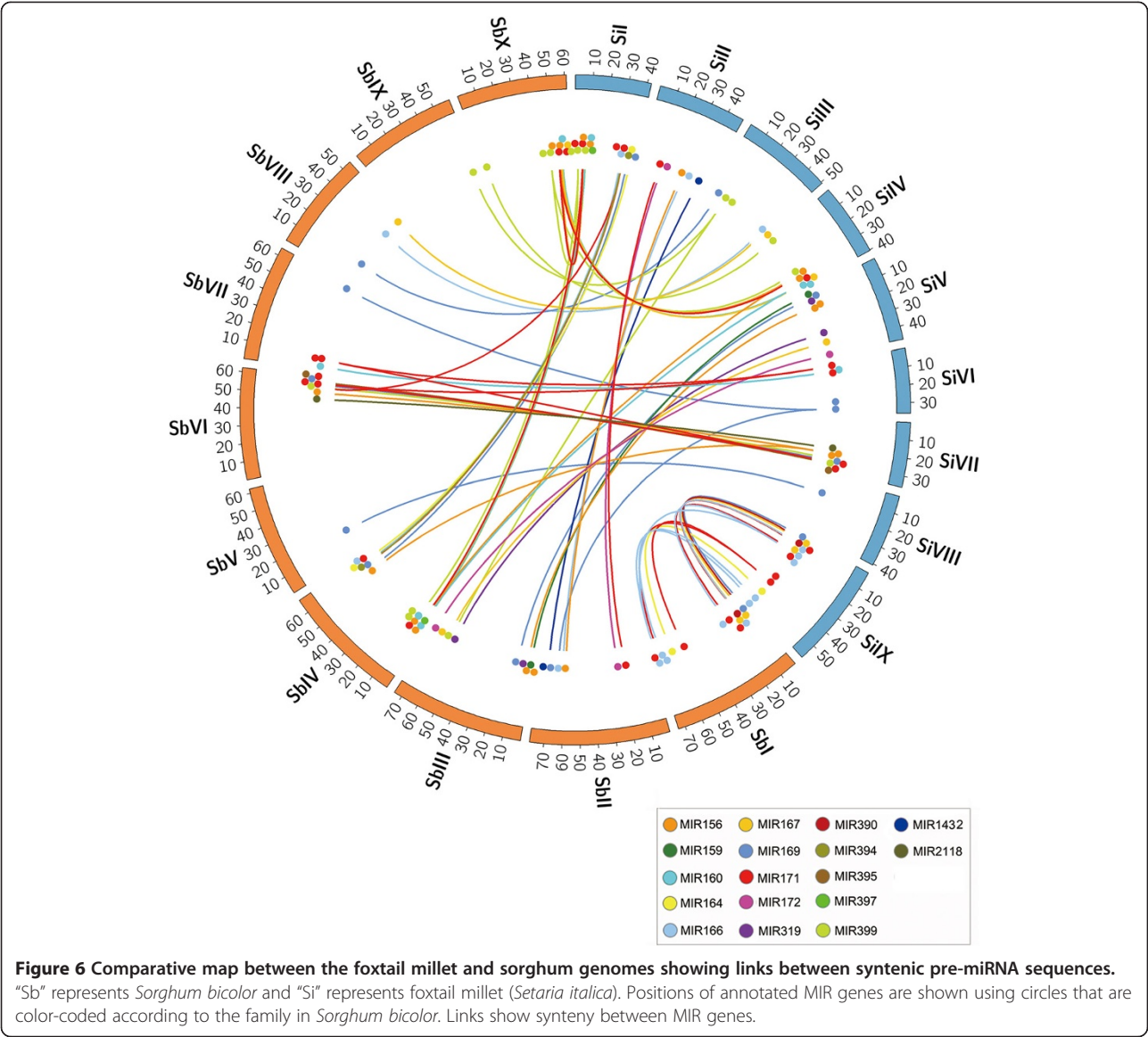
Discussion

Some miRNAs are expressed only in a specific tissue or at a specific site, and an understanding of their unique expression pattern may help in discovering the function of a particular miRNA. To assess the tissue specificity of the miRNAs in our datasets, we investigated the

expression levels of the miRNA in the four foxtail millet tissues (leaf, flower, root and shoot). We found that the expression of the sit-miR172 family was higher in the adult leaf and flower datasets and lower in the young shoot dataset. In maize, the miR172 that is responsible for the transition from juvenile to adult was reported to have the same expression pattern [46]. In addition, compared with the other sit-miRNA families, the sit-miR166 family was highly expressed, especially in leaf (over 20, 285 RPM). In maize, the target gene of miR166 is *rolled leaf1*, which can regulate leaf morphogenesis [17,56]. Similarly, in foxtail millet, the target gene of sit-miR166 was predicted as Si034251m, a homolog of *rolled leaf1*. This finding suggests that there may be a similar mechanism of leaf morphogenesis in foxtail millet as the one reported in maize.

We found that, except for a few of the sit-miRNAs that had relatively high expression levels, most of the sit-miRNAs, especially the nov-sit-miRNAs, had low expression levels in the four studied tissues. It has been suggested that the normal physical functions of an organism require various miRNAs working together for precise regulation and control rather than individual miRNAs working separately [57]. If this suggestion is correct, then the miRNAs with low expression might have an effect on the growth of foxtail millet as great as the miRNAs with high expression. A previous report indicated that most of the miRNA targets in maize are conserved across several plant species [46] and we found that this was also true in foxtail millet. For example, miR156 and miR529 were predicted to target genes that encode SBP-box transcription factors [58-60], and miR164, miR169, miR171, miR172 and miR319 were reported to target No Apical Meristem (NAM) [61,62], CCAAT-binding factor (CBF) [63,64], GRAS transcription factor [65], APETALA2 Ethylene-Responsive Element Binding Proteins (AP2-EREBP) [66,67] and Teosinte branched, Cycloidea, and PCF (TCP) [68,69], respectively. In maize and *Arabidopsis*, miR160 and miR167 were found to target auxin response factor [46,70,71]. In foxtail millet, sit-miR827 had 10 predicted target genes, six were SYG1/Pho81/XPR1 (SPX) proteins, and four were NAD (P)-binding proteins, in agreement with a report that, in maize, miR827 could target NAD (P)-binding and SPX proteins [46]. Most of the predicted target genes of the sit-miRNAs are transcription factors that are known to play important roles in plant growth processes and in the regulation of plant development. This result is also consistent with previous reports in other species [21,72,73].

In *A. thaliana*, the GRF gene family (AtGRF) of proteins that are involved in cell expansion in leaf and cotyledon tissues [74] are regulated by miR396 [75]. We found that nov-sit-miR64 was conserved to miR396 of *Arabidopsis*



and that its predicted target gene contained the same characteristic regions (QLQ (Gln, Leu, Gln) and WRC (Trp, Arg, Cys) domains) as the AtGRF proteins. Compared with other novel miRNAs, the nov-sit-miR64 had a relatively high expression in leaf (88 RPM), suggesting that nov-sit-miR64 may play an important role in leaf; however, nov-sit-miR64 expression was highest in root (1518 RPM) compared with in the other three tissues (less than 90 RPM). These findings suggest that nov-sit-miR64 (or miR396) not only plays an important role in leaf growth, but may also be involved in the development of root. The expression pattern of nov-sit-miR64 was validated by experiment (Figure 5B).

Foxtail millet is considered to be a diploid, tractable model for polyploid biofuel crops like switchgrass (*Panicum virgatum*) and Napier grass (*Pennisetum purpureum*)

[26]. The synthesis of cellulose is a complicated process that depends on carbon fixation, sugar metabolism and transit, and fat metabolism [76-78]. In this study, we also find eight potential miRNA targets that may be involved in the biological synthesis of cellulose and seven miRNAs were predicted to have an effect on the regulation of cellulose biosynthesis (Table 3). For example, sit-miR319 potentially targets four different genes involved in carbohydrate metabolic process, suggesting that sit-miR319 might play a central role in sucrose metabolism and carbon fixation. Cellulose synthase mRNA was predicted to be a target of miR-397 in maize [79], and laccase, which is involved in lignin catabolic process in foxtail millet, was one of the predicted targets of sit-miR397.

The expression pattern of some sit-miRNA families also showed similar trends as the corresponding conserved

Table 2 Mature sequence and precursor sequence conservation of novel miRNA

miRNA precursor	miRNA sequence	Conserved precursor	Conserved miRNA family
nov-sit-MIR105	GCTCACTCCTCTTTCTGTCAGC	sbi-MIR156a/e	miR156
nov-sit-MIR103	GCTCACTTCTCTGTCTGTCAGC	sbi-MIR156f	miR156
nov-sit-MIR16	TTGACAGAAGAGAGCGAGCAC	sbi-MIR156a	miR156
nov-sit-MIR117	CGTGCTCACTACTTTCTGTC	sbi-MIR156g	miR156
nov-sit-MIR60	TCGGACCAGGCTTCATTCCCCA	sbi-MIR166b	miR166
nov-sit-MIR05	TTTCGGACCAGGCTTCATTCC	sbi-MIR166c	miR166
nov-sit-MIR58-2	TCGGACCAGGCTTCATTCCCCCT	sbi-MIR166a/j	miR166
nov-sit-MIR58-3	TCGGACCAGGCTTCATTCCCCCT	sbi-MIR166j	miR166
nov-sit-MIR77-2	TAGCCAAGAATGACTTGCCT	sbi-MIR169i	miR169
nov-sit-MIR32	TTAGCCAAGAATGGCTTGCCCTA	sbi-MIR169q	miR169
nov-sit-MIR14	TTGAGCCGCGTCAATATCTCC	sbi-MIR171h	miR171
nov-sit-MIR49	TGAGCCGAGCCAATATCACT	sbi-MIR171c	miR171
nov-sit-MIR50	TGAGCCGAACCAATATCACTC	sbi-MIR171e/f	miR171
nov-sit-MIR148	AGTGGATGGCGCGGGAGCTAA	sbi-MIR319b	miR319
nov-sit-MIR114	CTGAAGTGTTGGGGAATC	sbi-MIR395a	miR395
nov-sit-MIR123	CGCCAAAGGAGAATTGCCCTG	sbi-MIR399b/h	miR399
nov-sit-MIR101	GGCAGCTCTCCTCTGGCAGG	sbi-MIR399d	miR399
nov-sit-MIR89	GTGCGGTTCTCCTCTGGCATG	sbi-MIR399h	miR399

miRNAs in different plants. For example, the expression levels of sit-miR156, sit-miR164, sit-miR166 and sit-miR167 were comparatively very high, and sit-miR160, sit-miR319, sit-miR390 and sit-miR394 were comparatively very low, in agreement with similar findings in other plants [21,22,80]. Because two of the characteristics (similar expression patterns and conserved targets) of these miRNAs were similar in other plants, we inferred that the regulatory mechanisms and main functions of these miRNA target genes would also be similar, perhaps explaining the similar expression patterns of the conserved miRNAs in different species.

In this work, we examined the expression profiles of ten miRNAs in four tissues, three of which appeared to not

be very consistent with the smRNA-seq data, suggesting that preprocessing and normalizing the data to help in identifying differentially expressed genes is a process that continues to be developed.

Conclusion

By smRNA sequencing, we identified 43 known miRNAs and 172 novel miRNAs in foxtail millet and studied their expression profiles in four different tissues. Potential targets were predicted with strict criteria as described, and four targets were validated by 5'RACE. The functional annotation provided a deeper understanding of the transcription and regulation of the target genes, confirmed many known regulatory mechanisms, and provided a

Table 3 MiRNAs potentially targeting biofuel-related biological processes

miRNA ID	Target ID	InterPro ID	InterPro description	GO annotation
sit-miR169c/e	Si025232m	IPR001220	Legume lectin domain	Carbohydrate binding
sit-miR319	Si022272m	IPR000490	Glycoside hydrolase, family 17	Hydrolase activity, hydrolyzing O-glycosyl compounds, carbohydrate metabolic process
	Si022276m	IPR013781	Glycoside hydrolase, catalytic domain	Catalytic activity, carbohydrate metabolic process, cation binding
sit-miR397	Si001625m	IPR017761	Laccase	Lignin catabolic process, apoplast, hydroquinone: oxygen oxidoreductase activity, oxidation reduction process
nov-sit-miR32	Si025232m	IPR001220	Legume lectin domain	Carbohydrate binding
nov-sit-miR41	Si035473m	IPR001220	Legume lectin domain	Carbohydrate binding
nov-sit-miR77	Si025232m	IPR001220	Legume lectin domain	Carbohydrate binding
nov-sit-miR155	Si020000m	IPR008089	Nucleotide sugar epimerase	Carbohydrate metabolic process, racemase and epimerase activity, acting on carbohydrates and derivatives

window into many more potentially novel pathways. Furthermore, a comparative genomic analysis with sorghum contributed to understanding the evolutionary dynamics of miRNA family expansions and will serve as the basis for future comparative functional genomic studies using syntenic analysis. The identification and characterization of miRNAs from foxtail millet will aid in further research on foxtail millet and other species in the Poaceae family of grasses.

Methods

Plant material and sequencing

Foxtail millet inbred line Yugu1 [26,27] was used in this study. Two smRNA libraries were constructed. The smRNAs extracted from 50 shoots of 14-day-old seedlings were used for each library. Seeds of Yugu1 were germinated on moist paper and incubated at 28°C for 24 h, then transferred to pots filled with 1:1 mix of nutrient soil:vermiculite and grown for 9 days in an illuminating incubator (28°C day/20°C night, 14-h photoperiod, 70% relative humidity). The roots of the seedlings were gently washed, then transferred to 1/4 Hoagland's solution and allowed to grow for an additional 5 days. Briefly, total RNA was isolated using TRIzol reagent (Invitrogen, USA). SmRNA library construction was carried out with an Illumina TruSeq Small RNA sample prep kit (Illumina, USA) according to the manufacturer's instructions. SmRNA was extracted by running total RNA on a 15% PAGE gel (1× TBE, 7 M urea, 15% acrylamide (19:1 acryl:bis-acryl) in 1% TBE at 200 V for 1 hour and excising the bands in the ~18 to 30 nt size range. 5' and 3' adaptors were ligated sequentially to the smRNAs and then amplified by RT-PCR. Samples were prepared for sequencing following the manufacturer's standard protocol (Illumina TruSeq Small RNA sample prep kit) (Illumina, USA) and sequenced on a HighSeq 2000 sequencer (Illumina) to produce 36-bp single reads. The generated raw reads have been deposited in NCBI's SRA database under accession numbers SRA062827.

Identification of foxtail millet miRNA and miRNA*

The foxtail millet genome sequences were downloaded from the Phytozome database (Sitalica_164_hardmasked.fa). To find foxtail millet miRNAs, a previously reported workflow was used [21] with a stricter MFEI [MFE/(precursor length) × 100/(G + C)] cutoff value of greater than 0.85 [35] rather than the previous MFEI threshold of 0.15. First, EMBOSS-einverted [81] was used to identify imperfect inverted repeats from the hard masked genome sequences. Then, the inverted repeats were folded by RNAfold [82] and single loop-stem-loop sequences were selected through an in-house developed Perl script. By checking the secondary structure and copies (less than

10) of these single loop-stem-loop sequences, we identified the miRNA precursor candidates.

Then, for each sequenced RNA sample, short reads were first mapped to all the candidate precursors. The location of the reads was checked to identify and eliminate the reads that mapped on the loop region of the corresponding precursor. Finally, the most abundant reads that were 20 ~ 22 nt in length with no less than 5 reads were regarded as potential miRNAs. A split-screen view of read alignments from a shoot (14-day-old) sample displaying regions of ten miRNA precursors is shown in Additional file 10. These results were displayed using Integrative Genomics Viewer (IGV) [83].

To identify known and conserved miRNAs in foxtail millet, BLASTN was used to match the candidate miRNAs to the mature miRNAs from all 14 of the previously investigated plant species [26] in miRBase Release 20 [84]. The miRNAs that had perfect matches with no mismatches and the same length as the mature plant miRNAs in miRBase were defined as known miRNAs in foxtail millet (sit-miRNAs). The miRNAs that did not have complete matches but had less than three mismatches were defined as conserved miRNA (nov-sit-miRNAs).

To identify miRNA*, the first step of our pipeline is using BLASTN to align smRNAs from our sequencing libraries to miRNA sequences. We identified miRNA* candidates by allowing no more than four mismatches between the smRNA sequences and the miRNA sequences. Then, miRNA precursors were folded by RNAfold and the predicted miRNA* sequences that have 2 nt 3' overhangs to the mature miRNAs were selected by an in-house developed Perl script, which can be obtained by request. Finally, if the sequence of predicted miRNA* can be found among the miRNA* candidates in the BLASTN output, we defined it as miRNA* in our study.

Identification of mirtrons in foxtail millet

To find mirtrons in the foxtail millet, we obtained all the intron sequences from the annotation (Sitalica_164_gene.gff3 and Sitalica_164_gene_exons.gff3) and assembly (Sitalica_164_hardmasked.fa) files (data downloaded from Phytozome) files in the Phytozome database. Because the lengths of plant pre-miRNA are usually between 60 ~ 500 nt, we selected the introns within this length range for secondary structure prediction using RNAfold [82]. Sequences that formed a single stem-loop with 0 to 3 nt overhangs at both ends and with a MFE of less than -25 [13] were retained for further filtering. All the smRNA reads from four foxtail millet tissues (leaf, flower, root, shoot) were mapped onto these introns and if the reads can perfectly mapped to ends of the introns were considered as one of the mature mirtron candidates [14].

Target gene prediction and gene ontology annotation

To predict the target genes of foxtail millet miRNAs more precisely, the foxtail millet gene set (release 164) and annotation data were downloaded from Phytozome. There were nearly 392,553 entries in the transcript sequence file. We used a target prediction method with a modified scoring system as described previously [38]. Basically, targets should fulfill the following criteria: no more than three mismatches between miRNA and target, a position independent score of no more than 3 [85], a position-dependent penalty score of no more than 4 [86], and MFE ratios for miRNA:target duplexes and miRNA:target-binding site duplexes greater than 0.75 [32]. The InterPro resource [87] was used to assign functional annotations to the predicted target genes. We subjected the potential miRNA targets to a functional enrichment analysis using BGI WEGO [46,88]. The Blast2GO software (v2.5.1) [89] with the default parameters was used to obtain the GO terms for each foxtail millet gene. The WEGO online tool (<http://wego.genomics.org.cn/cgi-bin/wego/index.pl>) [90] was used to perform a GO enrichment analysis of the miRNA targets. The Pearson Chi-square test was used for statistical analysis. GO categories that show a significant ($\alpha = 0.05$) enrichment were analyzed and displayed in the output histogram of the WEGO figure.

MiRNA target validation by 5'RACE

cDNA templates were prepared from the total RNA that was extracted from the 14-day-old shoots using a GeneRacer kit (Invitrogen) according to the manufacturer's instructions. For each miRNA target gene, two gene-specific primers (GSP1 and NGSP1) were designed using Primer Premier 5.0 software [91]. These two primers were used for two rounds of PCR and the nested PCR products that were obtained were analyzed on a 1% agarose gel. Positive PCR products were cloned into a pMD19-T vector (TaKaRa, Dalian, China) and transformed to *E. coli* DH10B cells. Sequencing was carried out by Majorbio (Shanghai, China). The primers that were used in this study are listed in Additional file 11.

Validation of novel miRNA by stem-loop RT-PCR

Total RNA from foxtail millet shoot (14-day-old) was extracted as described in miRNA target validation by 5'RACE. PCR was performed as follows: 94°C for 2 min, 40 cycles of 94°C for 15 s and 60°C for 1 min. The PCR products were analyzed by 4% agarose gel and the specificity of the amplification was validated by sequencing the products. The primer design and the stem-loop RT-PCR procedure were performed as described previously [92,93]. All primers used in the study are listed in Additional file 11.

Quantification of microRNAs by stem-loop RT-PCR

Total RNA was isolated from roots, leaves, flowers and shoots (14-day-old) using TRIzol Reagent (Invitrogen, USA) and digested with RNase-free DNase I (Promega, USA) to remove genomic DNA contamination. About 100 ng of DNA-free total RNA was hybridized with a miRNA-specific stem-loop RT primer. The hybridized miRNA molecules were then reverse transcribed into cDNA as described [92,93]. Real-time PCR was performed using SYBR® Premix Ex Taq™ (TaKaRa, Japan) on the 7500 Real Time PCR System (Applied Biosystems, USA) (see Additional file 11 for the primer sequences). All reactions were performed in triplicate. The cycling parameters are as follows: 95°C for 30 sec, 40 cycles of 95°C for 3 sec, 60°C for 30 sec. The endogenous reference, U6 [94], was analyzed on the cDNA template converted from total RNA using U6-specific RT primers (Additional file 11). The $2^{-\Delta\Delta CT}$ method [95] was used to calculate the relative gene expression levels, which were normalized to the expression level of U6. To verify the specificity of the PCR-amplification, some amplicons were cloned and sequenced.

Synten analysis of foxtail millet and sorghum

The protein-coding gene sequences of foxtail millet (release 164) and sorghum (release 79) were downloaded from Phytozome. BLASTN (Evalue = 10^{-3}) was used to identify homologous protein-coding and pre-miRNA genes between foxtail millet and sorghum. (The sorghum pre-miRNA sequences were downloaded from miRBase release 20) DAGchainer [96] was used to identify collinear chains among the homologs. We used Circos plots [97] to show the collinear relationships between the foxtail millet and sorghum pre-miRNA sequences.

Additional files

Additional file 1: Statistics of reads for small RNAs in foxtail millet.

(A) Total reads of 18~31nt small RNAs distribution from four different tissues. (B) Overlap among four sequenced small RNA libraries.

Additional file 2: Detailed information of known miRNAs in foxtail millet.

Additional file 3: Detailed information of novel miRNAs.

Additional file 4: Secondary structure of novel miRNA precursors.

Additional file 5: The detail information about predicted mirtron candidates.

Additional file 6: Potential targets of all miRNAs. Annotations were retrieved from the InterPro database.

Additional file 7: The miRNAs (whose targets have been validated) validated by stem-loop RT-PCR and sequencing.

Additional file 8: GO functional enrichment analysis for predicted target genes of miRNAs in four tissues respectively with comparisons to total foxtail millet genes.

Additional file 9: Syntenic sites harboring MIR genes in foxtail millet and sorghum.

Additional file 10: Split-screen view of read alignments from shoot (14-day-old) sample displaying regions of ten miRNA precursors.

Additional file 11: Primers used in this study.

Abbreviations

AMP1: ALTERED MERISTEM PROGRAM1; AGO: ARGONAUTE; AP2-EREBP: APETALA2 Ethylene-responsive element binding proteins; AtGRF: A. thaliana, the GRF gene family; CBF: CCAAT-binding factor; DCL1: DICER-LIKE1; ER: endoplasmic reticulum; GRAS: GAI, RGA, SCR; IGV: Integrative genomics viewer; MFE: Minimal folding free energies index; miRNAs: microRNAs; mRNA: Messenger RNA; NAM: No apical meristem; nt: nucleotide; pri-miRNAs: primary miRNAs; QLQ: Gln, Leu, Gln; RISC: RNA-induced silencing complex; RNAPol II: RNA polymerase II; RNAPol III: RNA polymerase III; RPM: reads per million; smRNAs: small RNAs; SPX: SYG1/Pho81/XPR1; TCP: Teosinte branched, Cycloidea and PCF; WRC: Trp, Arg, Cys.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JY and SX conceived and designed the research. FY performed the experimental validation and drafted the manuscript. FY, SX and YL analyzed the data. FY and XQ prepared the samples. JY revised and prepared the final manuscript. All authors read and approved the final manuscript.

Acknowledgments

We thank Mingming Kang (China Agricultural University, Prof. Jingjuan Yu's laboratory) and Xiaomei Dong (China Agricultural University, Prof. Jinsheng Lai's laboratory) for useful discussions and their critique of the manuscript. We also would like to thank Jeffrey Bennetzen and Blake's research group for providing the smRNAs of three tissues. This work was supported by the National Basic Research Program of China (Grant No. 2012CB215301) and the National Transgenic Major Program of China (Grant No. 2013ZX08003-002).

Received: 15 April 2013 Accepted: 27 November 2013

Published: 13 December 2013

References

- Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004, **116**(2):281–297.
- Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC: Expression of Arabidopsis MIRNA genes. *Plant Physiol* 2005, **138**(4):2145–2154.
- Bologna NG, Schapire AL, Palatnik JF: Processing of plant microRNA precursors. *Brief Funct Genomics* 2013, **12**(1):37–45.
- Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X: Methylation as a crucial step in plant microRNA biogenesis. *Science* 2005, **307**(5711):932–935.
- Bollman KM, Aukerman MJ, Park M, Hunter C, Berardini TZ, Poethig RS: HASTY, the Arabidopsis ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. *Development* 2003, **130**:1493–1504.
- Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS: Nuclear processing and export of microRNAs in Arabidopsis. *Proc Natl Acad Sci USA* 2005, **102**(10):3691–3696.
- Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, Ji L, Pan Z, Cao X, Mo B, et al: MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* 2013, **153**(3):562–574.
- Rajagopalan R, Vaucheret H, Trejo J, Bartel DP: A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Gene Dev* 2006, **20**:3407–3425.
- Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ, et al: Criteria for annotation of plant MicroRNAs. *Plant Cell* 2008, **20**(12):3186–3190.
- Ruby JG, Jan CH, Bartel DP: Intronic microRNA precursors that bypass Drosha processing. *Nat Lett* 2007, **448**:83–86.
- Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC: Mammalian mirtron genes. *Mol Cell* 2007, **28**(2):328–336.
- Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC: The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. *Cell* 2007, **130**:89–100.
- Joshi PK, Gupta D, Nandal UK, Khan Y, Mukherjee SK, Sanan-Mishra N: Identification of mirtrons in rice using MirtronPred: a tool for predicting plant mirtrons. *Genomics* 2012, **99**(6):370–375.
- Meng Y, Shao C: Large-scale identification of mirtrons in Arabidopsis and rice. *PLoS One* 2012, **7**(2):e31163.
- Tong Y, Peng H, Zhan C, Fan L, Ai T, Wang S: Genome-wide analysis reveals diversity of rice intronic miRNAs in sequence structure, biogenesis and function. *PLoS One* 2013, **8**(5):e63938.
- Lutter D, Marr C, Krumsiek J, Lang E, Theis F: Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. *BMC Genomics* 2010, **11**(1):224.
- He L, Hannon GJ: MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004, **5**(7):522–531.
- Jones-Rhoades MW, Bartel DP, Bartel B: MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 2006, **57**:19–53.
- Llave C, Xie Z, Kasschau KD, Carrington JC: Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science* 2002, **297**(5589):2053–2056.
- Fabian MR, Sonenberg N, Filipowicz W: Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* 2010, **79**:351–379.
- Jiao Y, Song W, Zhang M, Lai J: Identification of novel maize miRNAs by measuring the precision of precursor processing. *BMC Plant Biol* 2011, **11**:141.
- Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C: A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. *Genome Res* 2008, **18**(9):1456–1465.
- Glazov EA, Cottey PA, Barris WC, Moore RJ, Dalrymple BP, Tizard ML: A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* 2008, **18**:957–964.
- Yanik H, Turkas M, Dundar E, Hernandez P, Dorado G, Unver T: Genome-wide identification of alternate bearing-associated microRNAs (miRNAs) in olive (*Olea europaea* L.). *BMC Plant Biol* 2013, **13**:10.
- Kim B, Yu HJ, Park SG, Shin JY, Oh M, Kim N, Mun JH: Identification and profiling of novel microRNAs in the Brassica rapa genome based on small RNA deep sequencing. *BMC Plant Biol* 2012, **12**:218.
- Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, et al: Reference genome sequence of the model plant Setaria. *Nat Biotechnol* 2012, **30**(6):555–561.
- Doust AN, Kellogg EA, Devos KM, Bennetzen JL: Foxtail millet: a sequence-driven grass model system. *Plant Physiol* 2009, **149**:137–141.
- Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ: A large genome center's improvements to the Illumina sequencing system. *Nat Methods* 2008, **5**(12):1005–1010.
- Comparative Sequencing of Plant Small RNAs web site. <http://smallrna.udel.edu/data.php>.
- Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangel JL, et al: High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One* 2007, **2**(2):e219.
- Martínez G, Forment J, Llave C, Pallás V, Gómez G: High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. *PLoS One* 2011, **6**(5):e19523.
- Kang M, Zhao Q, Zhu D, Yu J: Characterization of microRNAs expression during maize seed development. *BMC Genomics* 2012, **13**(1):360.
- Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu BXS, Deng XW: Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. *Plant Cell* 2009, **21**:1053–1069.
- Mathews DH, Sabina J, Zuker M, Turner DH: Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *JMB* 1999, **288**:911–940.
- Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA: Evidence that miRNAs are different from other RNAs. *Cell Mol Life Sci* 2006, **63**(2):246–254.
- Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczynska D, Krzyzosiak WJ: Structural features of MicroRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem* 2004, **279**(42):42230–42239.
- Hake S: MicroRNAs: a role in plant development. *Dispatch* 2003, **13**:R851–R852.

38. Sun YH, Lu S, Shi R, Chiang VL: **Computational prediction of plant miRNA targets.** *Methods Mol Biol* 2011, **744**:175–186.
39. Phytozome v9.1: *Setaria italica*. <http://www.phytozome.net/foxtailmillet.php>.
40. Pysh LD, Joanna Wysocka-Diller W, Camilleri C, Bouchez D, Benfey PN: **The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes.** *Plant J* 1999, **1**(18):111–119.
41. Tian C, Wan P, Sun S, Li J, Chen M: **Genome-wide analysis of the GRAS gene family in rice and Arabidopsis.** *Plant Mol Biol* 2004, **54**(4):519–532.
42. Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T: **Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate.** *EMBO J* 2001, **20**:6877–6888.
43. Haley B, Zamore PD: **Kinetic analysis of the RNAi enzyme complex.** *Nat Struct Mol Biol* 2004, **11**(7):599–606.
44. JM TT: **RISC is a 5' phosphomonoester-producing RNA endonuclease.** *Genes Dev* 2004, **18**(9):975–980.
45. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R: **Control of translation and mRNA degradation by miRNAs and siRNAs.** *Genes Dev* 2006, **20**(5):515–524.
46. Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D: **A genome-wide characterization of microRNA genes in maize.** *PLoS Genet* 2009, **5**(11):e1000716.
47. Goff SA, Ricke D, Lan T, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H: **A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica).** *Science* 2002, **296**(5565):92–100.
48. Maher C, Stein L, Ware D: **Evolution of Arabidopsis microRNA families through duplication events.** *Genome Res* 2006, **16**(4):510–519.
49. Guddeti S, Zhang DC, Li AL, Leseberg CH, Kang H, Li XG, Zhai WX, Johns MA, Mao L: **Molecular evolution of the rice miR395 gene family.** *Cell Res* 2005, **15**(8):631–638.
50. Sunkar R, Girke T, Jain PK, Zhu J: **Cloning and characterization of microRNAs from rice.** *Plant Cell Online* 2005, **17**(5):1397–1411.
51. Lu C, Jeong D, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D: **Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs).** *Proc Natl Acad Sci* 2008, **105**(12):4951–4956.
52. Calviño M, Bruggmann R, Messing J: **Characterization of the small RNA component of the transcriptome from grain and sweet sorghum stems.** *BMC Genomics* 2011, **12**(1):356.
53. Zhang L, Zheng Y, Jagadeeswaran G, Li Y, Gowdu K, Sunkar R: **Identification and temporal expression analysis of conserved and novel microRNAs in Sorghum.** *Genomics* 2011, **98**(6):460–468.
54. Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, Xie M, Zeng P, Yue Z, Wang W: **Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential.** *Nat Biotechnol* 2012, **30**(6):549–554.
55. Schnable JC, Springer NM, Freeling M: **Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss.** *Proc Natl Acad Sci* 2011, **108**(10):4069–4074.
56. Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MC: **microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity.** *Nature* 2004, **428**(6978):84–88.
57. Alvarez-Garcia I, Miska EA: **MicroRNA functions in animal development and human disease.** *Development* 2005, **132**(21):4653–4662.
58. Wang JW, Schwab R, Czech B, Mica E, Weigel D: **Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in Arabidopsis thaliana.** *Plant Cell* 2008, **20**(5):1231–1243.
59. Ling LZ, Zhang SD: **Exploring the evolutionary differences of SBP-box genes targeted by miR156 and miR529 in plants.** *Genetica* 2012, **140**(7–9):317–324.
60. Schwarz S, Grande AV, Bujdosó N, Saedler H, Huijser P: **The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis.** *Plant Mol Biol* 2008, **67**(1–2):183–195.
61. Lu PL, Chen NZ, An R, Su Z, Qi BS, Ren F, Chen J, Wang XC: **A novel drought-inducible gene, ATAF1, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in Arabidopsis.** *Plant Mol Biol* 2007, **63**(2):289–305.
62. Zimmermann R, Werr W: **Pattern formation in the monocot embryo as revealed by NAM and CUC3 orthologues from Zea mays L.** *Plant Mol Biol* 2005, **58**(5):669–685.
63. Ben-Naim O, Eshed R, Parnis A, Teper-Bamnolker P, Shalit A, Coupland G, Samach A, Lifschitz E: **The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA.** *Plant J* 2006, **46**(3):462–476.
64. Combiér JP, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernie T, Ott T, Gamas P, Crespi M, et al: **MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula.** *Genes Dev* 2006, **20**(22):3084–3088.
65. Bolle C: **The role of GRAS proteins in plant signal transduction and development.** *Planta* 2004, **218**(5):683–692.
66. Yin Z, Li C, Han X, Shen F: **Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*).** *Gene* 2008, **414**(1–2):60–66.
67. Kim S, Soltis PS, Wall K, Soltis DE: **Phylogeny and domain evolution in the APETALA2-like gene family.** *Mol Biol Evol* 2006, **23**(1):107–120.
68. Schommer C, Palatnik JF, Aggarwal P, Chetelat A, Cubas P, Farmer EE, Nath U, Weigel D: **Control of jasmonate biosynthesis and senescence by miR319 targets.** *PLoS Biol* 2008, **6**(9):e230.
69. Ori N, Cohen AR, Etzioni A, Brand A, Yanai O, Shleizer S, Menda N, Amsellem Z, Efroni I, Pekker I, et al: **Regulation of LANCEOLATE by miR319 is required for compound-leaf development in tomato.** *Nat Genet* 2007, **39**(6):787–791.
70. Wu M, Tian Q, Reed JW: **Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction.** *Development* 2006, **133**:4211–4218.
71. Mica E, Gianfranceschi L, Pe ME: **Characterization of five microRNA families in maize.** *J Exp Bot* 2006, **57**(11):2601–2612.
72. Cui Q, Yu Z, Pan Y, Purisima EO, Wang E: **MicroRNAs preferentially target the genes with high transcriptional regulation complexity.** *Biochem Biophys Res Commun* 2007, **352**(3):733–738.
73. Dannemann M, Pruber K, Lizano E, Nickel B, Burbano HA, Kelso J: **Transcription factors are targeted by differentially expressed miRNAs in primates.** *Genome Biol Evol* 2012, **4**(4):552–564.
74. Kende H, Kim JH, Choi D: **The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis.** *Plant J* 2003, **36**(1):94–104.
75. Wang L, Gu X, Xu D, Wang W, Wang H, Zeng M, Chang Z, Huang H: **Cui X: miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in Arabidopsis.** *J Exp Bot* 2011, **62**(2):761–773.
76. Peng L, Kawagoe Y, Hogan P, Delmer D: **Sitosterol-beta-glucoside as primer for cellulose synthesis in plants.** *Science* 2002, **295**(5552):147–150.
77. Zhong R, Pena MJ, Zhou GK, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WR, Davill AG, York WS, Ye ZH: **Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis.** *Plant Cell* 2005, **17**(12):3390–3408.
78. Xie F, Frazier TP, Zhang B: **Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*).** *Planta* 2010, **232**(2):417–434.
79. Zhang B, Pan X, Anderson TA: **Identification of 188 conserved maize microRNAs and their targets.** *FEBS Lett* 2006, **580**(15):3753–3762.
80. Zhang R, Marshall D, Bryan GJ, Hornyk C: **Identification and Characterization of miRNA Transcriptome in Potato by High-Throughput Sequencing.** *PLoS One* 2013, **8**(2):e57233.
81. Rice P, Longden I, Bleasby A: **EMBOSS: the European molecular biology open software suite.** *Trends Genet* 2000, **16**(6):276–277.
82. Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL: **The Vienna RNA websuite.** *Nucleic Acids Res* 2008, **36**:W70–W74. Web Server issue.
83. Thorvaldsdóttir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.** *Brief Bioinform* 2013, **14**(2):178–192.
84. Kozomara A, Griffiths-Jones S: **miRBase: integrating microRNA annotation and deep-sequencing data.** *Nucleic Acids Res* 2011, **39**(suppl 1):D152–D157.
85. Jones-Rhoades MW, Bartel DP: **Computational Identification of Plant MicroRNAs and Their Targets, Including a Stress-Induced miRNA.** *Mol Cell* 2004, **14**:787–799.
86. Allen E, Xie Z, Gustafson AM, Carrington JC: **microRNA-directed phasing during trans-acting siRNA biogenesis in plants.** *Cell* 2005, **121**(2):207–221.
87. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Bullard V, Cerutti L, Copley R, et al: **New developments in the InterPro database.** *Nucleic Acids Res* 2007, **35**:D224–D228.
88. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: **Gene ontology: tool for the unification of biology.** *Nat Genet* 2000, **25**(1):25–29.
89. Conesa A, Götz S, García-Gómez JM, Tórtola J, Talón M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21**(18):3674–3676.

90. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L: **WEGO: a web tool for plotting GO annotations.** *Nucleic Acids Res* 2006, **34**(suppl 2):W293–W297.
91. Primer Premier: *Software for PCR Primer Design.* <http://www.premierbiosoft.com/primerdesign/index.html>.
92. Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP: **Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs.** *Plant Methods* 2007, **3**:12.
93. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, et al: **Real-time quantification of microRNAs by stem-loop RT-PCR.** *Nucleic Acids Res* 2005, **33**(20):e179.
94. Turner M, Adhikari S, Subramanian S: **Optimizing stem-loop qPCR assays through multiplexed cDNA synthesis of U6 and miRNAs.** *Plant Signal Behav* 2013, **8**(8):e24918.
95. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method.** *Methods* 2001, **25**(4):402–408.
96. Haas BJ, Delcher AL, Wortman JR, Salzberg SL: **DAGchainer: a tool for mining segmental genome duplications and synteny.** *Bioinformatics* 2004, **18**(20):3643–3646.
97. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA: **Circos: an information aesthetic for comparative genomics.** *Genome Res* 2009, **19**(9):1639–1645.

doi:10.1186/1471-2229-13-212

Cite this article as: Yi et al.: Genome-wide characterization of microRNA in foxtail millet (*Setaria italica*). *BMC Plant Biology* 2013 **13**:212.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

