Plant Biotechnology Journal

Plant Biotechnology Journal (2014) **12**, pp. 354–366 doi: 10.1111/pbi.12142

High-throughput deep sequencing shows that microRNAs play important roles in switchgrass responses to drought and salinity stress

Fuliang Xie[1], Charles Neal Stewart Jr[2,3], Faten A. Taki[1], Qiuling He[1], Huawei Liu[1] and Baohong Zhang[1,*]

[1]*Department of Biology, East Carolina University, Greenville, NC, USA*
[2]*Department of Plant Sciences, University of Tennessee, Knoxville, TN, USA*
[3]*BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, USA*

Received 21 June 2013;
revised 22 September 2013;
accepted 3 October 2013.
**Correspondence* (Tel 252-328-2021; fax 252-328-4178; email zhangb@ecu.edu)

Keywords: switchgrass, high-throughput sequencing, microRNA, salt, drought, biofuel.

Summary

MicroRNAs (miRNAs) are an important class of small regulatory RNAs. The goal of this study was to analyse stress-responsive miRNAs in switchgrass (*Panicum virgatum*), the emerging biofuel crop, to facilitate choosing gene targets for improving biomass and biofuel yield. After sequencing three small RNA libraries constructed from control, salt- and drought-treated switchgrass using Illumina sequencing technology, we identified 670 known miRNA families from a total of more than 50 million short reads. A total of 273 miRNAs were identified with precursors: 126 conserved miRNAs and 147 novel miRNAs. Of them, 265 miRNAs were found to have their opposite sequences (miRNA*) with 2-nt overhang on the 3′ end. Of them, 194 were detected in switchgrass transcriptome sequences generated from 31 high-throughput RNA sequencing (RNA-Seq) data sets in NCBI. Many miRNAs were differentially or uniquely expressed during salinity or drought stress treatment. We also discovered 11 miRNA clusters containing 29 miRNAs. These identified miRNAs potentially targeted 28 549 genes with a various function, including transcription factors, stress-response proteins and cellulose biosynthesis-related proteins. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis showed that the identified miRNAs and their targets were classified to 3779 GO terms including 1534 molecular functions, 1851 biological processes and 394 cellular components and were enriched to 147 KEGG pathways. Interestingly, 195 miRNA families and 450 targets were involved in the biosynthesis pathways of carbon, glucose, starch, fatty acid and lignin and in xylem formation, which could aid in designing next-generation switchgrass for biomass and biofuel.

Introduction

Small noncoding RNAs in typical length of 18–40 nucleotides (nt) have been discovered to play crucial roles in a remarkably wide range of biological processes, including cell proliferation, developmental timing and patterning, chromatin modification, genome rearrangement and stress response in plants and animals (Carrington and Ambros, 2003; Finnegan *et al.*, 2003; Xie *et al.*, 2004). To date, these small endogenous RNAs, which function as guide molecules to specify gene repression by base pairing to mRNAs of protein-coding genes (Bartel, 2004), have been categorized to at least six small RNA classes, including microRNAs (miRNAs), natural antisense siRNAs (natsiRNAs) (Kim, 2008), endogenous trans-acting siRNAs (tasiRNAs) (Allen *et al.*, 2005), repeat-associated siRNAs (rasiRNAs) (Djikeng *et al.*, 2001), small scan RNAs (scnRNAs) (Mochizuki and Gorovsky, 2004), heterochromatic siRNAs (hcsiRNAs) (Reinhart and Bartel, 2002) and Piwi-interacting RNAs (piRNAs) (Aravin *et al.*, 2007). Noncoding RNA world is being experienced a revolution that increasing RNA species and their unique regulatory functions are being uncovered, like recent long noncoding RNAs (lncRNAs) (Orom *et al.*, 2010) and competing endogenous RNAs (ceRNAs) (Karreth *et al.*, 2011). To date, according to the newest miRNA database (miRBase, 20 June 2013), a total of 30 424 mature miRNAs have been identified in 206 species, in which there are 7385 mature miRNAs obtained from 72 plant species (Griffiths-Jones, 2004; Griffiths-Jones *et al.*, 2008).

Switchgrass (*Panicum virgatum*), an endogenous prairie grass in eastern North America, is a promising herbaceous lignocellulosic energy crop. Therefore, switchgrass has been the subject of increasing research to increase biomass and decrease recalcitrance of biomass to sugar conversion (Bouton, 2007; Shen *et al.*, 2012). Research includes important studies at the gene regulation level. Recently, overexpression of switchgrass R2R3-MYB transcription factor MYB4 was identified to significantly down-regulate multiple lignin biosynthetic genes, which led to a reduction in lignin and increased fermentable sugar yield and saccharification efficiency (Shen *et al.*, 2012). Interestingly, as a large transcription factor family, MYB transcription factors have been well studied to be targeted by a series of miRNAs in both plants and animals, such as miR159 (Alonso-Peral *et al.*, 2010), miR-150 (Xiao *et al.*, 2007) and miR-29 (Martinez *et al.*, 2011). Furthermore, it has reported that overexpression of miR156 in switchgrass resulted in various morphological alterations and led to an improved biomass production (Fu *et al.*, 2012). Thus, the manipulation of miRNA regulatory function in switchgrass is a promising approach for modifications of lignin and cell wall phenolic ester synthesis to improve bioenergy production efficiency (Shen *et al.*, 2012). miRNAs have also been shown to play key roles in response to various abiotic stresses, including salinity

and drought (Xie *et al.*, 2004); these are all important traits for switchgrass as a biofuel feedstock.

Using bioinformatics approaches, 121 miRNAs and 33 miRNAs have been recently identified in switchgrass (Matts *et al.*, 2010; Xie *et al.*, 2010). The recently draft switchgrass genome (http://www.phytozome.net/panicumvirgatum.php) will provide a useful resources for miRNA discovery. At present, small RNA sequencing (deep sequencing) has emerged as a useful tool for rapid miRNA discovery, in which both conserved and novel miRNAs could be identified, and their abundance profiled simultaneously. In this study, we employed deep sequencing to investigate miRNAs and their expression patterns in switchgrass and to investigate specific stress-specific pathways to enable biofuel feedstock improvements.

Results

High-throughput sequencing of control, salt- and drought-treated small RNA libraries

Three small RNA libraries (control, salinity and drought treatment) yielded a total of 51 953 577 redundant reads and 16 289 218 unique reads (Table 1). There were more than 16–17 million redundant reads and more than 6 million unique reads for each library. To assess overall sequencing similarity between the three libraries, the Jaccard index was computed for the 5000 most abundant small RNAs in each library (Mohorianu *et al.*, 2011). According to pairwise comparisons using the Jaccard index, salinity and drought libraries were more similar (90.8%) when compared with the control library (Table S1). This suggests that many small RNAs would respond similarly to different abiotic stresses based on their overall expression level. The reads of three libraries that mapped back 100% to the switchgrass genome accounted for 72.28–73.75%, representing 57.52% to 58.59% of their unique read counterparts (Table 1). All three libraries displayed similar distributions to other RNA families including rRNA (~2.12% for the unique reads and ~18.0% for the redundant reads), snRNA (~0.04% for the unique reads and ~0.04% for the redundant reads), snoRNA (~0.03% for the unique reads and ~0.03% for the redundant reads) and tRNA (~0.22% for the unique reads and ~3.50% for the redundant reads) (Table 1). For redundant reads, unique reads, matched redundant reads and matched unique reads, all shared similar size distribution between the three libraries and 24 nt being the most frequent length (Figure 1). Overall, the distribution of small RNA abundance and size in switchgrass were consistent with the results reported in *Arabidopsis* (Rajagopalan *et al.*, 2006) and rice (Wei *et al.*, 2011).

Identification of conserved miRNA families in switchgrass

A total of 670 previously known conserved miRNA families were identified in the three small RNA libraries generated from the three treatments. Of these, 478 miRNA families were found in the control, whereas 456 and 482 families were found in the salt-treated and drought-treated samples, respectively. These miRNAs comprised approximately 0.5% of the total unique read sequences with an average sequence rate of 18.48% (Table 1). Most known miRNA reads were 20 nt, while next abundant class was 21 nt long (Figure S1); this is different from the total small RNA population (Figure 1), suggesting that there are many other class of small RNAs except miRNAs in the libraries. There were a total of 37 380, 36 679 and 33 780 miRNA reads in the control, salinity and drought libraries, independently. miR156, miR442, miR1171, miR1423, miR1869, miR2199, miR3436 and miR5140 comprised the top largest miRNA families in the three libraries. Some miRNA families were stably expressed in number in each of the three libraries, for example miR156, miR1171, miR2199 and miR442. However, some miRNA counts diverged sharply. For instance, miR1869 had 1778 reads in the control library and 1399 reads in the salinity library, whereas there was only one read in the drought library.

Known conserved miRNA family expression pattern comparisons

The abundance of 670 conserved miRNA families varied from family to family and from treatment to treatment. The highest abundant miRNA family was miR156 in the three libraries with an average of 118 987 reads per million (RPM) (Table 2 and Data S1). Its expression level was slightly higher in the control compared with those in the salinity and drought treatments. Because both fold change and *P-value* were used as two criteria to determine whether a miRNA was expressed significantly in a sample, miRNA156 was not differentially expressed in the control, salinity and drought treatments. Besides miR156, four additional major miRNA families (miR168, miR167, miR5078 and miR2199) were in the list of top 10 most highly expressed (between 2030 and 28 093 RPM) among the three libraries, but they did not significantly differ. Additionally, miR156 and miR2199 were not only expressed at high levels, but these two families also comprised a large-sized family. Interestingly, the expression of miR3946 and miR1869 was similar in both control and salinity treatments and at a high expression level (between 55 823 and 94 325 RPM); however, both of them had extremely low expression levels in drought treatment (between 2 and 358 RPM).

Table 1 Small RNA categorization in switchgrass*

	Unique (C) (%)	Redundant (C) (%)	Unique (S) (%)	Redundant (S) (%)	Unique (D) (%)	Redundant (D) (%)
Matched	3 632 328 (57.52)	12 908 990 (73.75)	3 843 644 (58.59)	12 665 777 (72.28)	3 784 476 (58.03)	12 278 326 (72.54)
miRNA	32 744 (0.52)	3 793 412 (21.67)	33 447 (0.51)	2 851 615 (16.27)	30 686 (0.47)	2 961 115 (17.49)
rRNA	133 681 (2.12)	3 004 097 (17.16)	141 515 (2.16)	3 282 598 (18.73)	135 410 (2.08)	3 042 605 (17.98)
snRNA	2616 (0.04)	7446 (0.04)	2739 (0.04)	9251 (0.05)	2448 (0.04)	7461 (0.04)
snoRNA	1703 (0.03)	4342 (0.02)	2159 (0.03)	5388 (0.03)	1711 (0.03)	4684 (0.03)
tRNA	14 923 (0.24)	539 293 (3.08)	14 056 (0.21)	765 766 (4.37)	13 657 (0.21)	520 134 (3.07)
unan	6 129 578 (97.06)	10 155 901 (58.02)	6 366 391 (97.04)	10 608 055 (60.54)	6 338 148 (97.18)	10 390 407 (61.39)
Total	6 315 247	17 504 493	6 560 310	17 522 676	6 522 062	16 926 408

*The number represented the raw data generated directly from deep sequencing; C, control; S, salt; D, drought; unan: unannotated; matched, matched to genome.

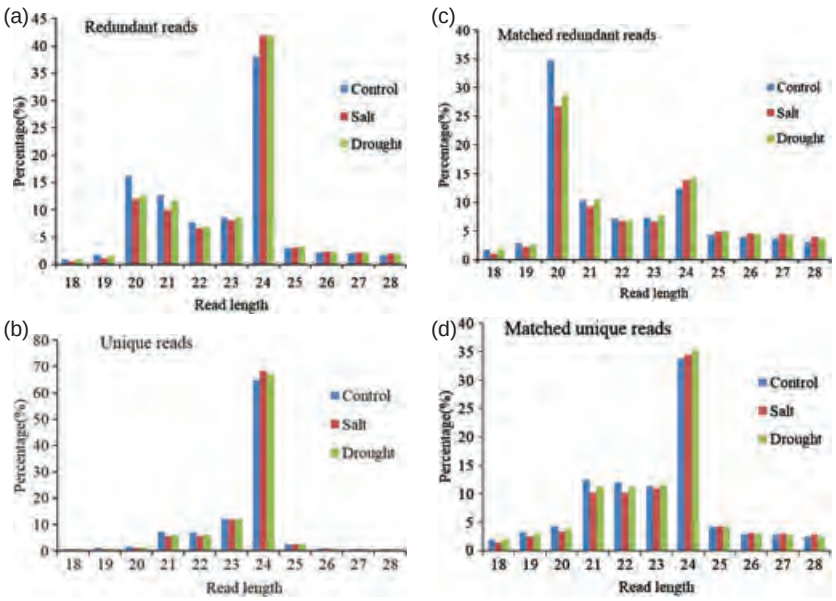

Figure 1 Size distribution of redundant and unique small RNA reads in switchgrass. (a) and (c): Size distribution of redundant sRNA reads from control, salt and drought libraries. (b) and (d): Size distribution of unique sRNA reads from control, salt and drought libraries.

Based on fold change (≥1.0) and *P-value* (≤0.05), 446 of 670 identified conserved miRNA families were significantly differentially expressed among the two treatments and control in switchgrass. On average, 62.8% miRNA families in each library had either high expression or specificity. Compared with salt and drought treatments, the expression of 166 and 158 miRNA families in control library was remarkably higher or specific, respectively (Table S2). Moreover, 54 miRNA families were expressed more abundantly or specifically in control library relative to salt and drought treatments. A total of 145 and 156 miRNA families were expressed specifically in drought treatment when compared with control and salt treatment, respectively. Similarly, a total of 78 and 94 miRNA families were specifically expressed in salinity or drought treatment. For example, miR2671 was only expressed in the salinity treatment and miR838 was detected only in the drought treatment (Table 2 and Data S1). Using heatmap clustering, we classified selected miRNA family groups with similar expression change trends in a single treatment library when compared with other libraries (Figure 2a). For instance, miR164, miR166, miR408 and miR528 were down-regulated by salinity treatment, whereas miR159, miR171, miR157, miR5023, miR5029 and miR1423 were up-regulated by drought treatment (Figure 2a and Data S1).

Identification of miRNA precursors and novel miRNAs

Considering possible sequencing errors, only small RNAs with at least 3 reads (879 568) in one of the three libraries were used to search for miRNA precursors, resulting in 6 643 842 hits on switchgrass genome with 100% nucleotide match. After analysing hairpin structures of miRNA candidates and removing repeated sequences, we were able to identify a total of 273 miRNAs with precursors: 126 conserved miRNAs and 147 switchgrass-specific miRNAs (Table S3 and Data S2). Almost all commonly known plant miRNAs were detected with precursors in the switchgrass genome. This result suggests that these miRNAs are highly conserved and widely exist in the plant kingdom. Based on current known miRNAs from eudicotyledons and monocotyledons deposited in miRBase database, at least 18 potential monocotyledon-specific miRNA families with precursors were found in switchgrass. Certain typical eudicotyledon-specific miRNA families did not exist or were not detected to have precursors in switchgrass. For all newly identified 147 novel miRNAs, we did not find homologs in other plant species, suggesting that these novel miRNA may be switchgrass specific.

In total, 265 of 273 (97.1%) miRNAs were identified to have miRNA* in switchgrass. To further validate the 273 identified pre-miRNAs with transcriptome evidence, switchgrass' EST database and 31 transcriptome sequencing data sets available were combined together and then were assembled. Of 273 miRNAs, 194 (71.1%) were able to be found in the assembled transcriptome sequences: 50 known miRNAs and 144 novel miRNAs (Table S3). The 273 miRNAs belong to 197 miRNA families, in which most of families (171, 62.6%) just have one member. The largest miRNA family is miR395 with 13 members, following by miR156 and miR166 with 9 and 7 members, respectively (Data S2).

miRNAs responsible to salt and drought treatment

Of the 273 identified miRNAs, 109 conserved miRNAs and 81 novel miRNAs were found in all three treatments, accounting for 69.6% of the total (Table S3 and Figure 3). miR1118 and 16 novel miRNAs were specifically expressed in plants subjected to the salinity treatment. There were 17 drought-specific miRNAs, in which 4 miRNAs were conserved and 13 miRNAs are novel. Three conserved miRNAs and 7 novel miRNAs were specific to both salinity and drought treatments. Additionally, three conserved miRNAs and 11 novel miRNAs were found in the control and salinity treatments, but were not expressed in the drought treatment. Similarly, 2 conserved miRNAs (miR1867 and miR479) and 10 novel miRNAs were only detected in control and drought treatments. Thirteen miRNAs (4.8%) were expressed only in control samples and were not detected in either salinity or drought treatment (Table S3 and Figure 3).

Except for the salinity-/drought-specific miRNAs, several well-studied miRNA families were strongly regulated by salinity or drought treatments (Table 2). Of the 208 miRNAs with absolute fold change less than 1 compared with the control, 36 and 44 miRNAs were remarkably down-regulated by salt and drought treatments, respectively. On the contrary, 21 and 29 miRNAs were up-regulated in the salt and drought treatment, respectively

Table 2 The expression of conserved miRNA genes among control (C), salt (S) and drought (D) treatments

	miRNA normalization†			Fold change			Significance		
miRNA	C	S	D	D/C	S/C	S/D	D vs C	S vs C	S vs D
miR156	143047	104822	109091	−0.39	−0.45	−0.06			
miR157	26	19	4618	7.42	−0.49	−7.91	**		**
miR159	63	56	119	0.91	−0.16	−1.06			**
miR160	41	28	61	0.55	−0.55	−1.11			**
miR161	0	3	4	5.3	5	−0.3	**	**	
miR162	11	41	17	0.6	1.84	1.24		**	**
miR164	605	526	552	−0.13	−0.2	−0.07			
miR165	55	10	28	−0.97	−2.33	−1.36		**	**
miR166	7145	6940	7320	0.04	−0.04	−0.08			
miR167	6076	5585	4381	−0.47	−0.12	0.35			
miR168	28093	18076	26926	−0.06	−0.64	−0.57			
miR169	216	203	200	−0.11	−0.09	0.03			
miR170	58	50	49	−0.24	−0.23	0.01			
miR171	68	78	88	0.37	0.21	−0.16			
miR172	177	147	159	−0.15	−0.27	−0.11			
miR390	27	26	24	−0.13	−0.02	0.11			
miR393	66	69	69	0.06	0.06	−0.01			
miR395	27	48	35	0.38	0.83	0.45			
miR396	168	259	169	0.01	0.63	0.62			
miR397	7	7	10	0.59	0.1	−0.49			
miR398	646	678	805	0.32	0.07	−0.25			
miR399	0	0	2	1.44		−1.05	**		*
miR437	2	619	606	7.98	8.01	0.03	**	**	
miR442	719	785	719	0	0.13	0.13			
miR773	0	70	66	12.71	12.79	0.08	**	**	
miR811	37	0	107	1.52	−7.78	−9.3	**	**	**
miR838	0	0	70	6.37		−7.11	**		**
miR893	47	0	0	−12.21	−12.21		**	**	
miR894	1685	925	747	−1.17	−0.86	0.31	**		
miR952	198	2	363	0.87	−6.12	−6.99		**	**
miR1024	16	19	69	2.08	0.23	−1.85	**		**
miR1115	174	0	0	−14.09	−14.09		**	**	
miR1171	503	468	518	0.04	−0.1	−0.15			
miR1423	1485	0	0	−13.62	−12.67		**	**	
miR1441	0	66	0		7.19	8.56		**	**
miR1853	0	60	0		12.57	12.57		**	**
miR1869	3547	3185	0	−14.87	−0.16	14.72	**		**
miR2199	2030	2843	2786	0.46	0.49	0.03			
miR2654	142	1	0	−9.24	−6.29	2.95	**	**	**
miR2671	0	272	0		14.74	12.17		**	**
miR2678	104	0	128	0.3	−9.84	−10.14		**	**
miR2919	1	1	243	7.42	0.4	−7.02	**		**
miR3436	3	183	0	−8.64	5.52	14.17	**	**	**
miR3444	0	1138	17	7.29	13.28	5.99	**	**	**
miR3946	5388	4399	21	−7.99	−0.29	7.7	**		**
miR4387	0	155	0		9.83	9.04		**	**
miR5023	927	0	0	−12.35	−10.82		**	**	
miR5028	69	0	0	−12.75	−12.75		**	**	
miR5029	750	3	2	−8.31	−7.62	0.69	**	**	
miR5077	418	211	187	−1.16	−0.98	0.17	**		
miR5078	5071	4081	4114	−0.3	−0.31	−0.01			
miR5200	0	36	35	11.79	11.83	0.05	**	**	

†All miRNAs with redundant abundance were normalized to transcript expression levels per million reads (RPM). If the original miRNA expression in a library was zero, the normalized expression was adjusted to 0.01 according to a previous report (Murakami *et al.*, 2006). miRNA expression fold change in any two libraries was calculated using the formula, fold change = $\log_2$(treatment 1/treatment 2) (Marsit *et al.*, 2006). A 2 × 2 contingency table was used to perform Pearson's chi-squared test for significance of miRNA expression from two samples. Fold change and *P*-value were combined to determine the final miRNA expression significance. We defined expression difference level as following rules: extremely significant (**) if (fold change ≥ 1 or fold change ≤ −1) and *P-value* ≤0.01; significant (*) if (fold change ≥ 1 or fold change ≤ −1) and 0.05 ≥ *P*-value >0.01; otherwise insignificant.

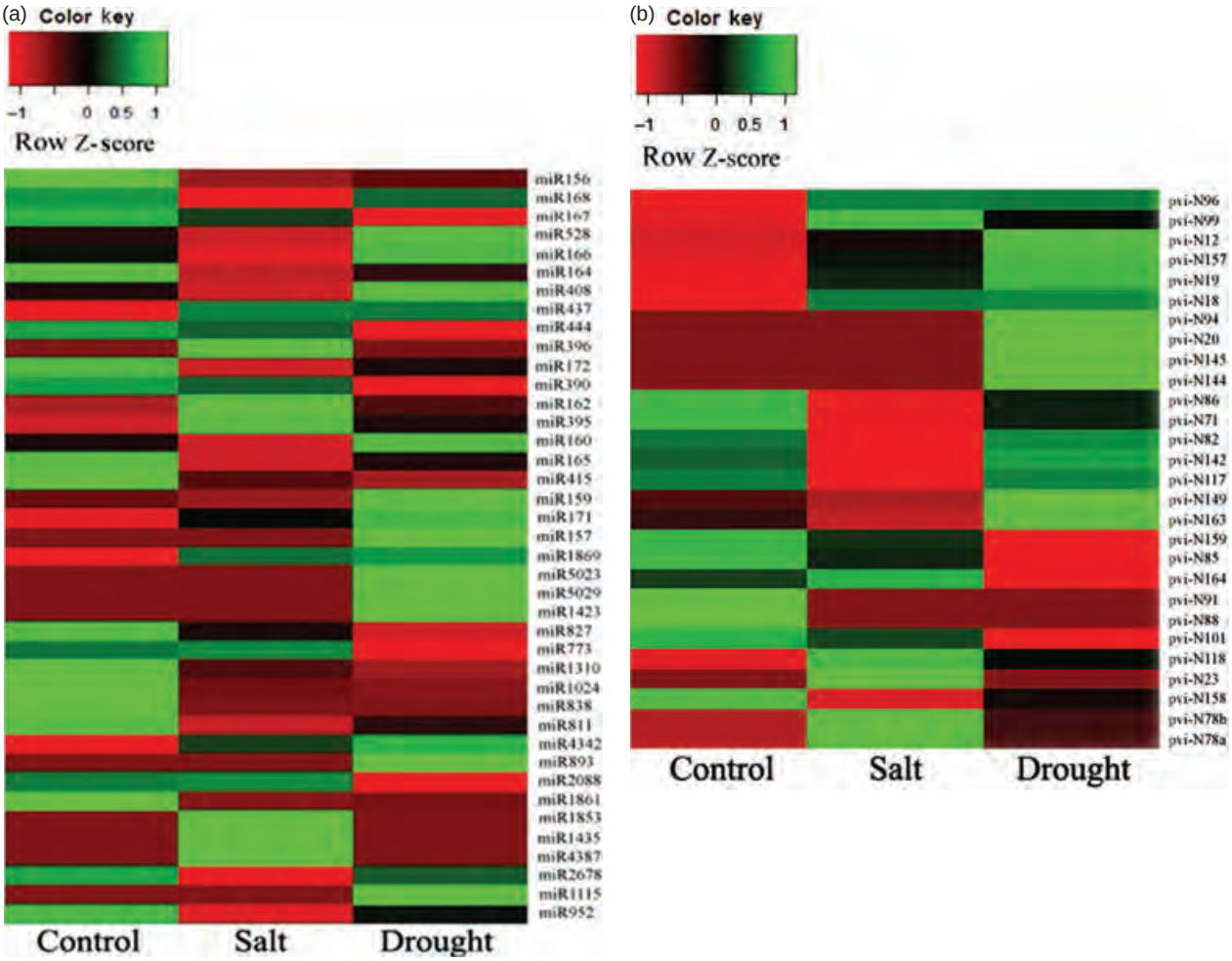

Figure 2 Heatmaps of (a) conserved and (b) novel miRNA read abundance in control, salt and drought libraries in switchgrass. Red: down-regulated; green: up-regulated.

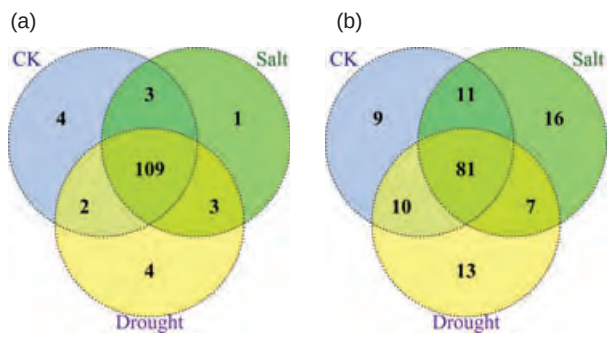

Figure 3 Distribution of miRNAs in control, drought and salinity treatment (a) known miRNAs; (b) novel miRNAs). CK: control.

(Table S4 and Data S2). In addition, at least 108 miRNAs in switchgrass performed considerably different to salt and drought ($P \leq 0.05$). Many novel miRNAs responded to stress of salinity or drought in a similar way (Figure 2b). For instance, newly identified miRNA N12, N157 and N19 were up-regulated in the drought treatment, while 11 novel miRNAs were down-regulated in the salt treatment.

miRNA clusters and its evolution in switchgrass

miRNAs initially derive from a long primary transcript (pri-miRNA), which can be polycistronic and generating one or more hairpin structure sequences (pre-miRNAs). As a result, multiple miRNAs can be located in adjacent loci in the genome to form a miRNA cluster. miRNAs in a cluster might share the same hairpin structure or have independent hairpin structures (Contreras-Cubas *et al.*, 2011). The origin of miRNA clusters might be caused by an initial phase of gene (tandem) duplication, multiple duplications of an entire cluster or random gene insertion by transposable elements (Piriyapongsa *et al.*, 2007; Tanzer and Stadler, 2004).

miR159a was found to be capable of producing a second 21-nt-long miRNA named miR159.2 in *Phaseolus vulgaris* (Contreras-Cubas *et al.*, 2011). Both miR159a and miR159.2 in *Phaseolus vulgaris* come from the 3′ arm of miR159 precursor hairpin structure; however, miR159a and miR159.2 had distinct expression pattern among organs and growth conditions. In this study, 11 miRNA clusters were identified in switchgrass, having 2–4 members located on either plus or minus strand of genomic contigs (Figure 4 and Data S2). Three of them (miR811& N167, miR2923b & N166 and miR159 & N165) might share a common

precursor (Figure 4a–c). However, the three clusters in switchgrass showed that precursor-sharing miRNAs can originate either from same arm or from a different arm. Surprisingly, all precursor-sharing miRNAs could not be co-detected in the drought library, not even their counterpart miRNA*s. Meanwhile, the newly identified miRNA N166 was not detected in the control samples. It indicates that abiotic stress may induce certain miRNA expression in switchgrass. In addition, two miRNA members in the miR159-N165 cluster were co-expressed (Figure 4c). miRNA precursor-sharing mechanism and its expression pattern are so complicated in switchgrass that further investigation is needed.

miR395 can target two families of genes (ATP sulfurylases and sulphate transporter) in *Arabidopsis*, which is up-regulated in sulphate limitation (Liang *et al.*, 2010). Currently, at least 149 miR395s have been found in 20 plant species, including *Arabidopsis*, rice, *Sorghum bicolor* and *Populus trichocarpa*. There are at least two and five miR395 clusters in *Arabidopsis* and rice, respectively. All members in each rice miR395 cluster occur in either the plus or minus strand in the rice genome. Two nearby miRNA members in 4 rice miR395 clusters and 1 rice miR395 cluster span around 144-nt and 301-nt intervals, respectively. All members in each switchgrass miR395 cluster occur on the minus strand and span a 142-nt interval with their adjacent members (Figure 4d–g). However, in *Arabidopsis*, all miRNA members of the miR395 cluster occurred on the same plus or minus strand, and the gap between two adjacent members was about 1328 nt. Compared with the miR395 clusters in *Arabidopsis*, miR395 clusters showed more similarity in chromosome distribution in rice and switchgrass. Because switchgrass is a tetraploid species, there should be between 13 to 48 miR395s in switchgrass if switchgrass had a similar evolution rate as rice. Questions around polyploidization and target gene evolution need to be asked to determine the ancient genomic contribution of miRNAs into extant polyploids; these answers might be parse given the available genome sequences in polyploidy species such as switchgrass and cotton. It would be interesting to delve into questions about intergenomic miRNA-target gene cross-talk among polyploids as well.

Validation of differentially expressed miRNAs in switchgrass by qRT-PCR

Stem-loop qRT-PCR is a reliable method for detecting and measuring miRNA expression levels (Chen *et al.*, 2005). In this study, we employed this technology to validate our deep sequencing data with 3 experimental and biological replicates. According to miRNA profile comparison between deep sequencing and qRT-PCR (Figure 5), almost all tested 12 miRNAs (6 conserved miRNAs and 6 novel miRNAs) treated by salinity and drought share the similar tendency in deep sequencing and qRT-PCR when compared with the control (Figure 5a). Furthermore, the 12 miRNAs expression profile relative to the control exhibited positive correlations between deep sequencing and qRT-PCR (Figure 5b and c).

miRNA target identification

A total of 47 005 conserved miRNA-target pairs and 221 novel miRNA-target pairs were identified from 70 071 switchgrass transcriptome sequences in JGI, representing 6467 (36.6%) unique mature miRNAs and 28 544 (40.7%) nonrepeated genes (Data S3 and S4).

Transcription factors were the major miRNA targets in switchgrass; these transcription factors included squamosa promoter-binding protein-like (SPL), MYB-domain transcription factor (MYB), NAC-domain transcription factor (NAC), APETALA2-like (AP2), homeodomain leucine zipper transcription factor (HD-ZIP), TCP-domain transcription factor (TCP) and GRAS transcription factor (GRAS). Some miRNAs targeted genes associated with cellulose biosynthesis (cellulose synthase- or cellulose synthase-like), sugar metabolism (sucrose synthase and sucrose transporter), starch metabolism (starch synthase), stress response (drought-responsive protein, low-temperature and salt-responsive protein and salt tolerance homolog 2), hormone signal transduction (auxin F-box protein and auxin-responsive GH3 family protein) and development (root hair initiation protein and cell elongation protein) (Data S5). Here, we identified 19 miRNAs (such as miR167 and miR1026) targeting 8 argonaute family proteins; we also identified that miR415 targeted copper/zinc superoxide dismutase 1 and miR156/159 regulated ATPase E1-E2 type family protein. As specific to switchgrass, the 170 identified novel miRNAs in switchgrass also showed cross-functional with some conserved miRNAs for expected targets, such as N117-auxin-responsive factor, N65/N105/pentatricopeptide repeat, N146-NAC, N65-growth-regulating factor and N134-DCL4. This suggests that some switchgrass-specific miRNAs may evolve to co-target certain genes.

At least 194 miRNA families targeted 449 protein-coding genes involved in cellulose and carbohydrate synthesis in switchgrass (Data S5). Thirty-six members of cellulose synthase family, including cellulose synthase 1 and cellulose synthase 6, were targeted by 46 miRNAs. Thirty miRNA families were involved in regulating 23 starch-related enzymes. miR156, miR172, miR1846 and miR1862 might play roles in fatty acid metabolism by targeting fatty acid desaturase 3, long-chain fatty alcohol dehydrogenase family protein, fatty acid desaturase A and fatty acid hydroxylase 1, respectively (Data S5). At least 314 glucose synthesis-related genes, including glucose-6-phosphate dehydrogenase 2, UDP-glucose 6-dehydrogenase family protein, xyloglucan endotransglucosylase/hydrolase family protein, beta glucosidase 31 and ADP-glucose pyrophosphorylase family protein, were predicted to be targeted by 156 miRNA families in switchgrass (Data S5).

A series of drought-related proteins including drought-sensitive 1 and drought-responsive family protein were also identified to be targeted by 23 miRNA families in switchgrass, which included miR156 and miR419. miR156 was inhibited by 0.45- and 0.39-fold by salinity and drought treatment, respectively; however, miR419 was down-regulated and miR4370 was only detected in drought library. It suggests that these three miRNA families might play different roles in switchgrass response to abiotic stress. According to miRNA regulatory network analysis by Cytoscape (http://www.cytoscape.org/), 7 miRNA families and 13 target gene families in switchgrass were identified to be core miRNAs and genes involved in stress response to salt and drought, including miR156, miR2104, miR1858, disease resistance protein (DIRP), drought-responsive family protein (DRRP), early responsive to dehydration stress protein (ERD), stress-responsive alpha–beta barrel domain protein (SRAP) and low-temperature and salt-responsive protein (LTSRP) (Figure 6a and Data S5). Overall, miRNA expression change and miRNA regulatory mechanism are very complicated in switchgrass in response to salt and drought stress. These miRNAs and their targets related to stress response allow us a new insight into understanding stress resistance mechanism in switchgrass.

GO and KEGG pathway analysis

To better understanding the regulatory roles of miRNAs in switchgrass, we performed GO analysis on all identified miRNAs

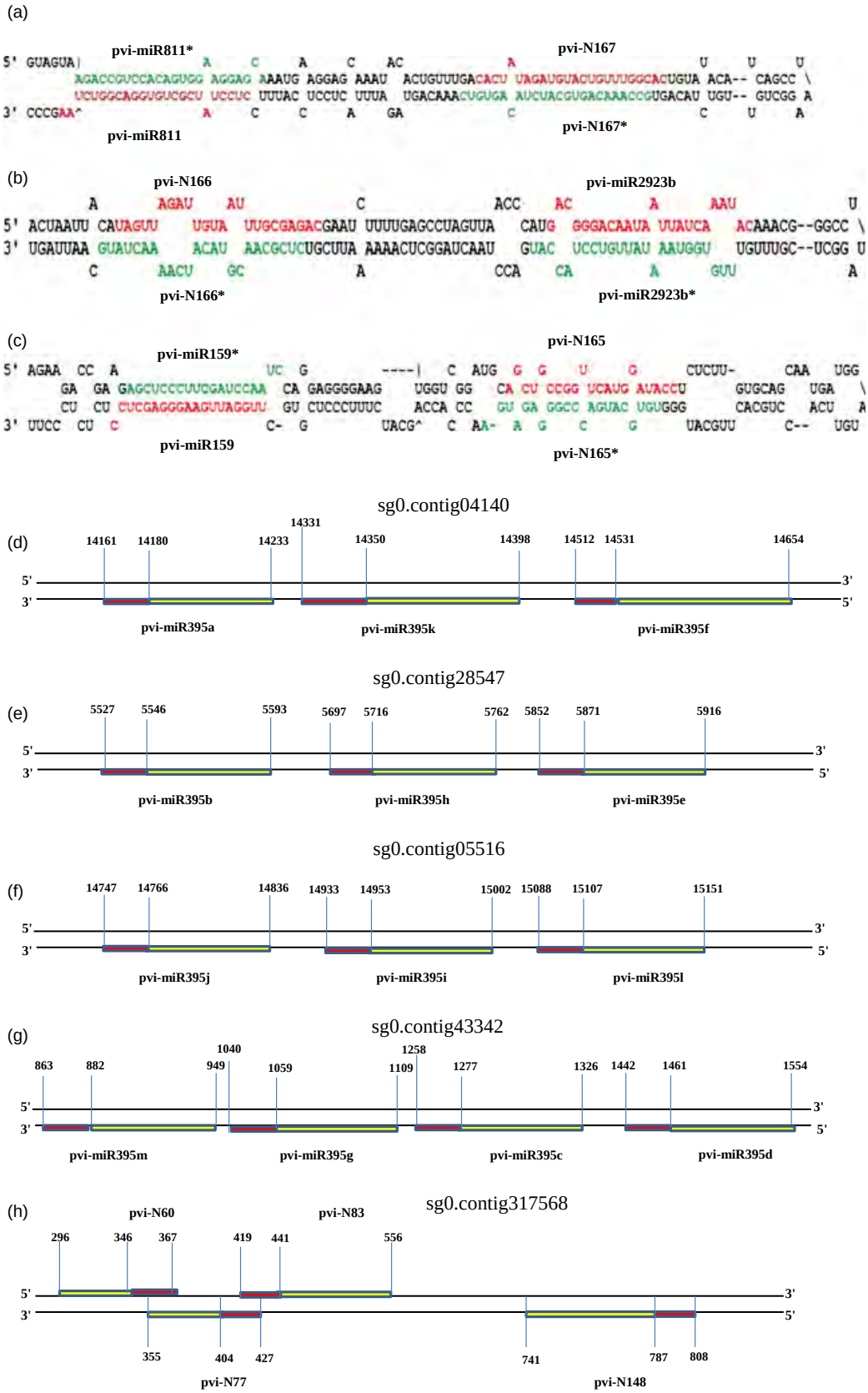

Figure 4 miRNA cluster distribution in switchgrass genome (mature miRNAs are labelled in red colour, and miRNAs were labelled in red and green colour). (a) miRNA N167 and miR811 cluster; (b) miRNA N166 and miR2923b cluster; (c) miRNA N165 and miR159 cluster; (d) miRNA miR395a, miR395k and miR395f cluster; (e) miRNA miR395b, miR395 h and miR395e cluster; (f) miRNA miR395j, miR395i and miR395 l cluster; (g) miRNA miR395 m, miR395 g, miR395c and miR395d cluster; and (h) miRNA N60, N77, N83 and N148 cluster.

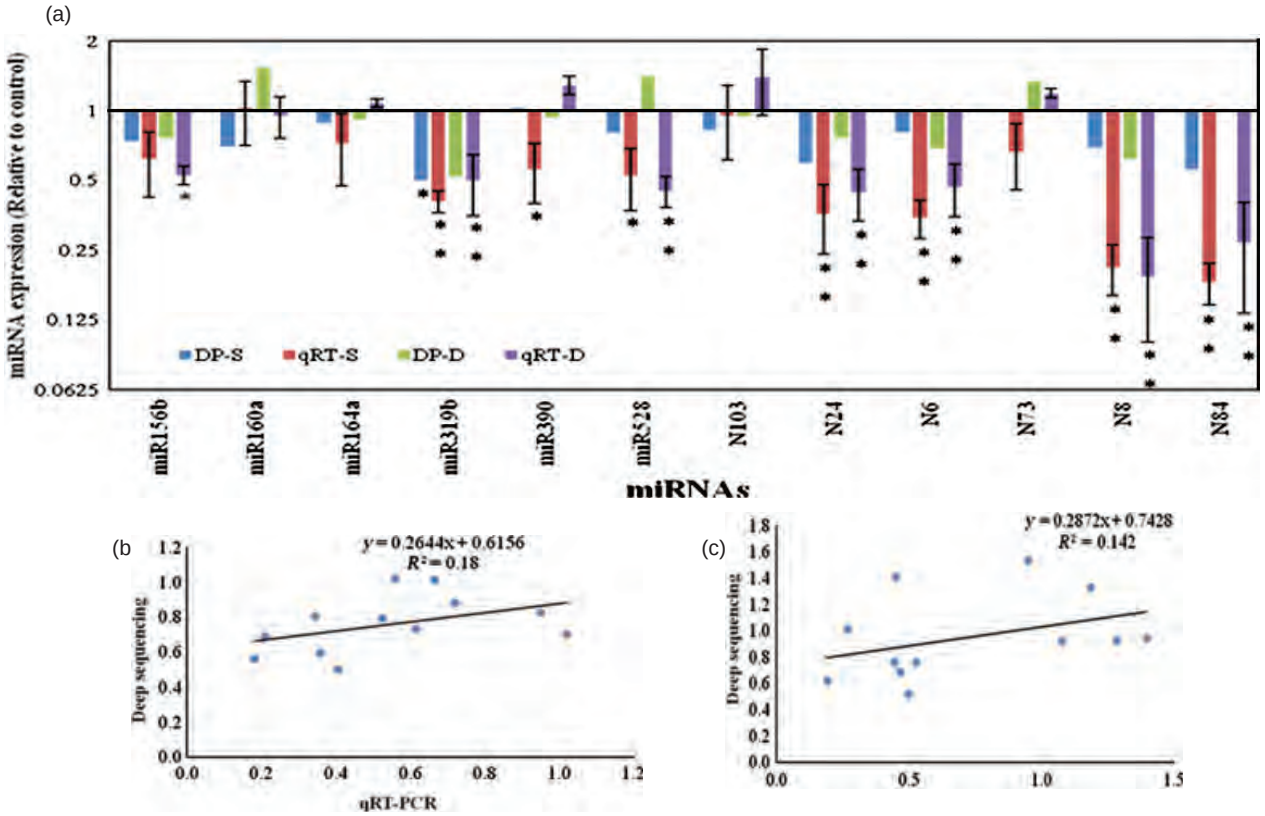

Figure 5 qRT-PCR validation of differentially expressed miRNAs identified by deep sequencing and the comparison of 12 switchgrass miRNAs' expression between deep sequencing and qRT-PCR. (a) miRNA expression in the treatments of salinity and drought stress relative to the control with deep sequencing and qRT-PCR (DP-S: salinity treatment with deep sequencing; qRT-S: salinity treatment with qRT-PCR; DP-D: drought treatment with deep sequencing; and qRT-D: drought treatment with qRT-PCR); (b) scatterplot of miRNA expression of salinity treatment with deep sequencing and qRT-PCR; (c) scatterplot of miRNA expression of drought treatment with deep sequencing and qRT-PCR. *: significant (*P*-value <0.05); **: extremely significant (*P*-value <0.01).

and their targets (Figure S2). A total of 469 miRNA families and their 17 899 targets were able to be categorized to 1851 biological processes, 394 cellular components and 1534 molecular functions. At least 203 miRNA families were shown to be involved in response to salt stress, including miR156, miR164 and miR172, because 507 of their targets were classified into two biological processes of cellular response to salt stress (GO:0071472) and response to salt stress (GO:0009651) (Table 3 and Data S3 and S4). Fifteen miRNA families, such as miR156, miR157 and miR1507, belonged to the GO term, drought recovery (GO:0009819). Additionally, 109 miRNA families were involved in response to oxidative stress based on their biological process categorization (GO:0006979 and GO:0008631). Most importantly, we found that 88 miRNA families were classified to be cellulose synthesis-related process, such as cellulose biosynthetic process (GO:0030244), cellulose catabolic process (GO:0030245), cellulose microfibril organization (GO:0010215) and plant-type cell wall cellulose metabolic process (GO:0052541). The majority of the identified switchgrass miRNAs (92.4%) target the genes with function in the cell part, whereas miRNAs (90.8%) also majorly regulated the genes belonging to the 'bind' in molecular function classification (Figure S2).

We were able to enrich 357 miRNA families and 2837 targets to 147 pathways related to the metabolism of starch, sucrose, glucose, amino acid, nucleotide, protein synthesis and modification, pathogen interaction and hormone signal transduction (Data S3 and S4). According to the annotation of KEGG pathway database, cellulose synthesis exists in a complex pathway, starch and sucrose metabolism, which include at least 50 enzyme members, such as cellulose synthase, UDP-glucose 6-dehydrogenase and glucose-6-phosphatase. miR3468-cellulose synthase-like D5 was involved in starch and sucrose metabolism (ath00500) and amino sugar and nucleotide sugar metabolism (ath00520) (Data S3). Overall, at least 194 miRNA families and their 449 target counterparts in switchgrass that were categorized to 24 pathways showed potential biofuel application because of their participation in the metabolism of cellulose, fatty acid, glucose, starch, sucrose and xylem (Table 3).

Discussion

Differentially expressed miRNAs involved in abiotic stress response and biofuel production

The aberrant expression of many conserved and newly identified miRNAs in switchgrass was induced in drought and salinity treatments, whereby several miRNAs were uniquely expressed in an abiotic stress treatment. miR162, which was induced in both salinity and drought conditions, is predicted to target a series of genes including DCL1, NB-ARC domain-containing disease resistance protein and peroxidase superfamily protein (Data S3). Coincidently, miR162 has been verified to target DCL1 and regulate miRNA biogenesis in *Arabidopsis* (Xie *et al.*, 2003). Thus, overexpressed miR162 may inhibit DCL1 protein synthesis and further affect miRNA biogenesis and lower the total expression of the entire miRNA world given that DCL1 is required during

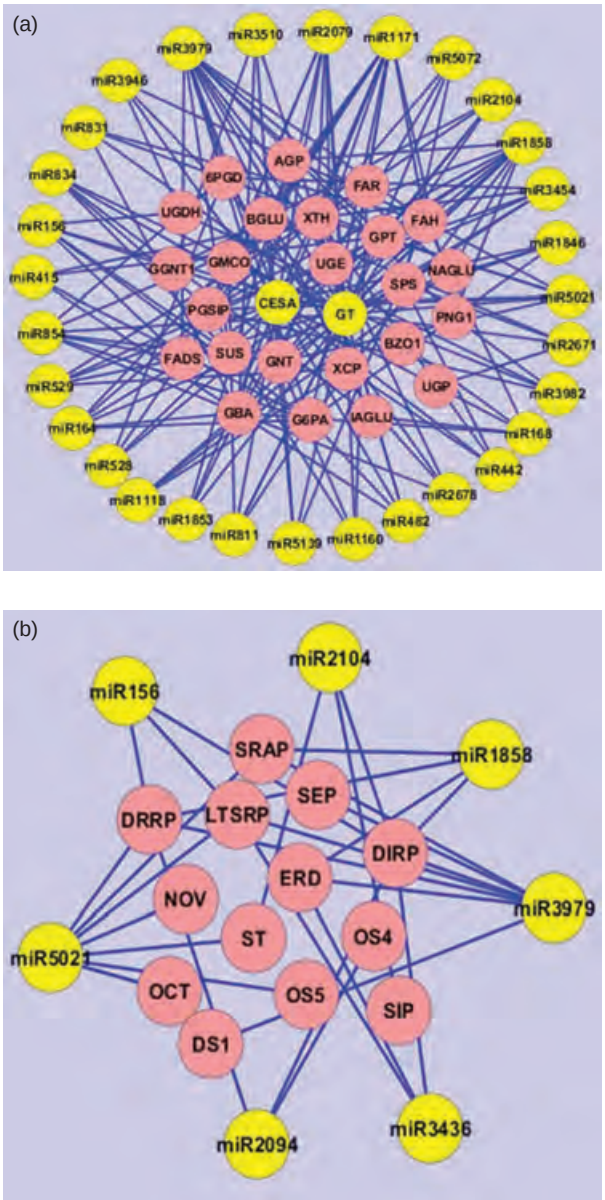

Figure 6 miRNA regulation networks involved in (a) stress response and (b) biofuel in switchgrass. Cellulose synthase (CESA) and UDP-glucosyl transferase (GT) are major two hubs to connect these miRNAs and other targets, such as miR156, miR415, miR3797, miR1858, fatty acid hydroxylase (FAH), beta glucosidase (GBA) and UDP-glucose pyrophosphorylase (UGP). These results can facilitate research to identify gene targets for further biofuel production improvement from switchgrass.

mature miRNA biogenesis. This hypothesis is supported by our sequencing data, in which the total conserved miRNA read was decreased by 5.4% and 4.2% by salinity and drought treatments, respectively. Our previous miRNA expression analysis on switchgrass by qRT-PCR also showed that miR162 was induced under drought treatment (Sun *et al.*, 2012).

Many predicted miRNA targets in switchgrass were already confirmed in model plant species. miR156/157-SPL, miR159/319-MYB, miR160/167-auxin-response factor, miR162-dicer-like 1 (DCL1), miR164-NAC, and miR165/166-HD-ZIP (Rhoades *et al.*, 2002), miR170/171-GRAS, miR171-AP2, miR393-F-box, miR396-growth-regulating factor and miR397/408-laccase (Jones-Rhoades *et al.*, 2006) are well-studied miRNA-target pairs and also identified in switchgrass. However, some well-characterized miRNA-target pairs, such as miR395-ATP sulfurylase, miR168-argonaute protein, miR398-copper superoxide dismutase and miR399-E2 ubiquitin-conjugating protein, were not found in this study (Jones-Rhoades *et al.*, 2006) possibly because of the incomplete switchgrass gene annotation.

At least 14 salt tolerance or salt-responsive proteins were identified to be miRNA targets in switchgrass, including genes for a salt tolerance zinc finger, salt tolerance homolog 2 and low-temperature and salt-responsive protein family (Data S5). Zinc finger proteins are a class of DNA-binding protein transcription factors consisting of zinc finger domain and are involved in many aspects of plant growth and development. Recently, a range of zinc finger proteins were identified to enhance different stress (salt, drought and fungal disease) tolerance in many plant species (Xie *et al.*, 2003). For example, ZFP1, a novel CCCH-type zinc finger protein from cotton, could interact with GZIRD21A and GZIPR5 to improve cotton salt stress tolerance and fungal disease resistance in transgenic tobacco (Xie *et al.*, 2003). The heterologous expression in *Arabidopsis* of a chrysanthemum Cys2/His2 zinc finger protein gene was shown to function as an important regulator involved in the salt and drought stress response in plants (Gao *et al.*, 2011). We predicted that miR2014 potentially targets salt tolerance zinc finger in switchgrass. miR2014 was down-regulated in salt and drought stress with 3.92 and 0.81 fold change (Data S1). The salt-tolerant homolog 2, a B-box protein, was first identified in *Arabidopsis* (Datta *et al.*, 2007). It was reported to activate transcription and positively regulate light-mediated development. In our results, two salt tolerance homolog 2 genes were targeted by miR2079, miR5021 and miR5028. According to the three miRNA families' expression abundance, miR2079 expression was induced in salt and drought treatments by 4.32- and 4.01-fold, respectively. miR5021 was down-regulated in salt and drought exposure; miR5028 was not detected in salt and drought libraries.

Salt and drought induce oxidative stress in plant cells (Zhu, 2002). miR398, a conserved plant miRNA, was shown to target two closely related Cu/Zn superoxide dismutases (cytosolic CSD1 and chloroplastic CSD2) with the purpose of detoxifying superoxide radicals (Sunkar *et al.*, 2006). In a mutant where miR398 was highly expressed, there was enhanced tolerance to high light, heavy metals and other oxidative stresses. In this study, we also observed that miR398 expression was significantly induced in the drought and salinity treatments. In addition, there are a series of other versatile genes that might contribute to stress tolerance, such as early responsive to dehydration stress protein (ERD4), disease resistance protein (TIR-NBS-LRR class and CC-NBS-LRR class) (Yue *et al.*, 2012).

Exploiting miRNAs for increasing biofuel production in switchgrass

Next-generation biofuels such as cellulosic biofuels from switchgrass are sought to decrease the demand of fossil fuels in the transportation sector and help decrease greenhouse gas production (Bouton, 2007). Cellulose is synthesized by cellulose synthase catalysis either from UDP-glucose or from GDP-glucose REFS. Fortunately, a variety of cellulose synthase or its homologs (cellulose synthase 6/A2/A4/A9 and cellulose synthase-like D2/D6/

Table 3 Stress-, resistance- and biofuel-related miRNAs, miRNA targets, GO terms and KEGG pathways in switchgrass

Function type	miRNA families	Targets	Cellular component	Biological process	Molecular function	Pathway
Salt stress	15	14	1	2	2	0
Drought stress	30	20	7	3	1	0
Other stress	19	16	3	1	1	0
Disease resistance	80	86	4	9	4	1
Cellulose	46	36	9	18	5	2
Fatty acid	45	37	10	18	14	2
Glucose	156	314	34	100	89	18
Starch	30	23	5	5	7	1
Sucrose	36	31	11	15	9	1
Transcription factor	182	378	14	129	31	3
Xylem	12	9	6	7	4	0

G3/E1) were found to be targeted by a group of miRNAs. In this study, we also found that at least 194 miRNA families targeted 449 protein-coding genes involved in cellulose and carbohydrate synthesis in switchgrass, which may contribute to biofuel production. Cellulose synthases influence cellulose and cell wall quality and ethanol conversion rate (Chuck *et al.*, 2011; Fu *et al.*, 2012). These results suggest that certain miRNAs could be used to improve cellulose production in switchgrass. Interestingly, two independent research groups recently overexpressed the rice miR156b (OsmiR156) (Fu *et al.*, 2012) and the maize miR156 (*Corngrass 1*) (Chuck *et al.*, 2011) to switchgrass, resulting in increased biomass and starch content and improved switchgrass biofuel conversion rate. miR156s are highly evolutionarily conserved in all plant species with the richest expression abundance (Chen, 2009). Our sequencing results showed that miR156 is the miRNA with the highest expression abundance in switchgrass. It is well known that miR156 regulates leaf development, shoot maturation, phase change and flowering in plants by targeting the squamosa promoter-binding protein-like (SPL) family of transcription factors (Chen, 2009). Chuck and co-worker also reported that overexpression of *Corngrass* 1 in switchgrass could promote juvenile cell wall identities and morphology and prevent flowering. Similarly, miR172 also influences inflorescence and cell fate by targeting APETALA 2 (AP2) and APETALA3 (AP3) (Yue *et al.*, 2012). Therefore, we ask here whether miR172 could be used to promote switchgrass biofuel conversion in a similar way to miR156. In addition, there are other noncellulose cell wall components and growth factors that could be affected by manipulating miRNAs. Here, we identified at least 413 miRNA targets related to the metabolism of glucose, starch, fatty acid, sucrose and xylem (Data S5) and provided more leads for switchgrass bioenergy development. Overall, miRNA regulatory network showed that there are at least 30 miRNA families and 26 target gene families related to biofuel application, which are highly connected to each other (Figure 6b). These results can facilitate research to identify gene targets for further biofuel production improvement from switchgrass.

Experimental procedures

Small RNA libraries preparation and sequencing

Switchgrass cv Alamo seeds were first sterilized with 70% (v/v) ethanol for 60 s, 6% (v/v) bleach for 6–8 min, followed by washing with sterilized water for at least four times. Then, the sterilized seeds were germinated on 1/2 Murashige and Skoog (MS) medium (pH 5.8) containing 0.8% agar under a 16-h light/8-h dark cycle at room temperature for 10 day. The media were supplemented with 0.5% NaCl to simulate salinity stress and with 5% PEG to simulate drought stress. Each treatment was replicated for five times as in five individual plates, and each plate contained 20 seeds. Ten-day-old 'Alamo' seedlings (controls, 0.5% NaCl and 5% PEG treatment) were harvested and immediately frozen in liquid nitrogen according to our previous report (Sun *et al.*, 2012). Total RNA was extracted from each tissue sample using the mirVana miRNA isolation kit (Ambion, Austin, TX) according to the manufacturer's protocol. RNA was quantified and qualified by Nanodrop ND-1000 (Nanodrop technologies, Wilmington, DE). All RNA samples were submitted to BGI (Shenzhen, China) for high-throughput sequencing using Illumina HiSeq high-throughput sequencing platform.

Pipeline of bioinformatics analysis

Three small RNA libraries were generated using deep sequencing for salinity and drought treatments as well no-treated controls. All the raw sequences from the HiSeq were cleaned first, including removing 5′ and 3′ adaptors, and low-quality reads were filtered out. Then, the raw sequences were categorized to unique reads and read counts were also calculated. The top 5000 small RNAs in read abundance were chosen to compute the Jaccard index for assessing the similarity coefficient of the three sequencing libraries (Mohorianu *et al.*, 2011). Clean reads fully matching to other RNAs, including repeat RNA, rRNA, snRNA, snoRNA and tRNA, were excluded by BLASTn-short alignment (blast2.2.26+, ftp://ftp.ncbi.nih.gov/blast/executables/blast+/2.2.26/) against Sanger RNA family database (Rfam 10.1, ftp://ftp.sanger.ac.uk/pub/databases/Rfam) (Gardner *et al.*, 2011). The remaining sequences were further aligned against miRBase (Release 18, http://www.mirbase.org/) (Kozomara and Griffiths-Jones, 2011), allowing up to 3 mismatches. Whether having their opposite sequences (miRNA*) in deep sequencing library with 2-nt overhang is crucial to ensure the validity of a discovered miRNA. However, previous studies have shown that some well-known miRNAs cannot be identified to have miRNA* in a small RNA sequencing data set in some model plant species, such as *Arabidopsis* and rice (Li *et al.*, 2011; Xie *et al.*, 2012). Given this issue, only novel miRNAs and their miRNA*s were required to co-exist in at least one library.

Before miRNA precursor search, sequences falling in the following conditions were discarded: (i) the short reads with greater than 80% A, C, G or T; (ii) reads shorter than 16 nt; (iii) reads containing the minimum homopolymers 7A, 8C, 6G, 7T; (iv)

reads with 10 repeats of a dimer, 6 repeats of a trimer or 5 repeat of a tetramer; (v) sequences with only A + C and no G or T; (vi) sequences with only G and T; (vii) read count less than 3 (Xie *et al.*, 2012). miRDeepFinder (http://www.leonxie.com/deepfinder.php) (Xie *et al.*, 2012) we developed previously was used to analyse miRNA precursors and their targets. MicroRazerS (v1.0, http://www.seqan.de/projects/MicroRazerS/) (Emde *et al.*, 2010) was integrated into miRDeepFinder for rapid alignment of small RNA reads back to switchgrass genome from JGI (http://www.phytozome.net/panicumvirgatum.php). Switchgrass transcriptome genes annotated by JGI were employed to analyse miRNA targets, miRNA function categorization and counterpart pathway enrichment based on databases of GO and KEGG (Xie *et al.*, 2010).

In order to further validate the existence of miRNA precursors identified above, we aligned these miRNAs against switchgrass transcriptome data sets including EST database and transcriptome sequencing data (RNA-seq) in NCBI. We used a total of 31 switchgrass transcriptome data sets assembled by ABYSS (ABySS 1.3.3, http://www.bcgsc.ca/platform/bioinfo/software/abyss) (Simpson *et al.*, 2009) first, and then, the 31 assembled transcriptome sequences were compiled together for further assembly by iAssembler (http://bioinfo.bti.cornell.edu/tool/iAssembler/) (Zheng *et al.*, 2011). In addition, 720 590 switchgrass ESTs available from NCBI were also assembled for detecting miRNA precursors.

Comparison of general miRNA expression profiles in control, salt- and drought-treated switchgrass seedlings

All miRNAs with redundant abundance were normalized to transcript expression levels per million reads (RPM). If the original miRNA expression in a library was zero, the normalized expression was adjusted to 0.01 according to a previous report (Murakami *et al.*, 2006). miRNA expression fold change in any two libraries was calculated using the formula, fold change = $\log_2$(treatment 1/treatment 2) (Marsit *et al.*, 2006). A 2 × 2 contingency table was used to perform Pearson's chi-squared test for significance of miRNA expression from two samples. Fold change and *P*-value were combined to determine the final miRNA expression significance. We defined expression difference level as following rules: extremely significant (**) if (fold change ≥ 1 or fold change ≤ −1) and *P-value* ≤0.01; significant (*) if (fold change ≥ 1 or fold change ≤ −1) and 0.05 ≥ *P*-value >0.01; otherwise insignificant.

miRNA validation and qRT-PCR

Differentially expressed miRNAs identified by deep sequencing were validated using a well-developed stem-looped quantitative real-time PCR (qRT-PCR). Stem-loop qRT-PCR is a reliable method for detecting and measuring the expression levels of miRNAs, which can be used to distinguish a single-nucleotide difference (Chen *et al.*, 2005). The Applied Biosystem TaqMan® microRNA Assays (Foster City, CA) were employed to measure and compare miRNA expression in switchgrass 12-day-old seedlings as our previous reports (Burklew *et al.*, 2012; Frazier *et al.*, 2011; Zhang and Pan, 2009). To validate deep sequencing data and identify miRNAs, we randomly selected 12 switchgrass miRNAs for miRNA expression detection by qRT-PCR: 6 conserved and 6 novel miRNAs. Switchgrass seedlings in three biological replicates for qRT-PCR were treated in the same way with those for deep sequencing. EF-1α and GAPDH were used as reference genes.

Acknowledgements

We greatly appreciate US Department of Energy Joint Genome Institute (JGI, http://www.jgi.doe.gov/) for switchgrass genome and its annotation produced by JGI in collaboration with the switchgrass research community. This project is partially supported by a grant from NIFA. The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

References

Allen, E., Xie, Z., Gustafson, A.M. and Carrington, J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207–221.

Alonso-Peral, M.M., Li, J., Li, Y., Allen, R.S., Schnippenkoetter, W., Ohms, S., White, R.G. and Millar, A.A. (2010) The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in Arabidopsis. *Plant Physiol.* **154**, 757–771.

Aravin, A.A., Hannon, G.J. and Brennecke, J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science*, **318**, 761–764.

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

Bouton, J.H. (2007) Molecular breeding of switchgrass for use as a biofuel crop. *Curr. Opin. Genet. Dev.* **17**, 553–558.

Burklew, C.E., Ashlock, J., Winfrey, W.B. and Zhang, B.H. (2012) Effects of aluminum oxide nanoparticles on the growth, development, and microRNA expression of tobacco (*Nicotiana tabacum*). *PLoS ONE*, **7**, e34783. doi:10.1371/journal.pone.0034783.

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

Chen, X. (2009) Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* **25**, 21–44.

Chen, C.F., Ridzon, D.A., Broomer, A.J., Zhou, Z.H., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R., Lao, K.Q., Livak, K.J. and Guegler, K.J. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **33**, e179. doi:10.1093/nar/gni178.

Chuck, G.S., Tobias, C., Sun, L., Kraemer, F., Li, C., Dibble, D., Arora, R., Bragg, J.N., Vogel, J.P., Singh, S., Simmons, B.A., Pauly, M. and Hake, S. (2011) Overexpression of the maize Corngrass1 microRNA prevents flowering, improves digestibility, and increases starch content of switchgrass. *Proc. Natl Acad. Sci. USA* **108**, 17550–17555.

Contreras-Cubas, C., Rabanal, F.A., Arenas-Huertero, C., Ortiz, M.A., Covarrubias, A.A. and Reyes, J.L. (2011) The Phaseolus vulgaris miR159a precursor encodes a second differentially expressed microRNA. *Plant Mol. Biol.* **80**, 103–115.

Datta, S., Hettiarachchi, C., Johansson, H. and Holm, M. (2007) SALT TOLERANCE HOMOLOG2, a B-box protein in Arabidopsis that activates transcription and positively regulates light-mediated development. *Plant Cell*, **19**, 3242–3255.

Djikeng, A., Shi, H., Tschudi, C. and Ullu, E. (2001) RNA interference in *Trypanosoma brucei*: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. *RNA*, **7**, 1522–1530.

Emde, A.K., Grunert, M., Weese, D., Reinert, K. and Sperling, S.R. (2010) MicroRazerS: rapid alignment of small RNA reads. *Bioinformatics*, **26**, 123–124.

Finnegan, E.J., Margis, R. and Waterhouse, P.M. (2003) Posttranscriptional gene silencing is not compromised in the Arabidopsis CARPEL FACTORY (DICER-LIKE1) mutant, a homolog of Dicer-1 from Drosophila. *Curr. Biol.* **13**, 236–240.

Frazier, T.P., Sun, G.L., Burklew, C.E. and Zhang, B.H. (2011) Salt and drought stresses induce the aberrant expression of microRNA genes in tobacco. *Mol. Biotechnol.* **49**, 159–165.

Fu, C., Sunkar, R., Zhou, C., Shen, H., Zhang, J.Y., Matts, J., Wolf, J., Mann, D.G., Stewart C.N. Jr, Tang, Y. and Wang, Z.Y. (2012) Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol. J.* **10**, 443–452.

Gao, H., Song, A., Zhu, X., Chen, F., Jiang, J., Chen, Y., Sun, Y., Shan, H., Gu, C., Li, P. and Chen, S. (2011) The heterologous expression in Arabidopsis of a chrysanthemum Cys2/His2 zinc finger protein gene confers salinity and drought tolerance. *Planta*, **235**, 979–993.

Gardner, P.P., Daub, J., Tate, J., Moore, B.L., Osuch, I.H., Griffiths-Jones, S., Finn, R.D., Nawrocki, E.P., Kolbe, D.L., Eddy, S.R. and Bateman, A. (2011) Rfam: Wikipedia, clans and the "decimal" release. *Nucleic Acids Res.* **39** (Database issue), D141–D145.

Griffiths-Jones, S. (2004) The microRNA Registry. *Nucleic Acids Res.* **32** (Database issue), 109–111.

Griffiths-Jones, S., Saini, H.K., van Dongen, S. and Enright, A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res.* **36**, D154–D158.

Jones-Rhoades, M.W., Bartel, D.P. and Bartel, B. (2006) MicroRNAS and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **57**, 19–53.

Karreth, F.A., Tay, Y., Perna, D., Ala, U., Tan, S.M., Rust, A.G., DeNicola, G., Webster, K.A., Weiss, D., Perez-Mancera, P.A., Krauthammer, M., Halaban, R., Provero, P., Adams, D.J., Tuveson, D.A. and Pandolfi, P.P. (2011) In vivo identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell*, **147**, 382–395.

Kim, V.N. (2008) Sorting out small RNAs. *Cell*, **133**, 25–26.

Kozomara, A. and Griffiths-Jones, S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39** (Database issue), 152–157.

Li, T., Li, H., Zhang, Y.X. and Liu, J.Y. (2011) Identification and analysis of seven HO-responsive miRNAs and 32 new miRNAs in the seedlings of rice (*Oryza sativa* L. ssp. indica). *Nucleic Acids Res.* **39**, 2821–2833.

Liang, G., Yang, F. and Yu, D. (2010) MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* **62**, 1046–1057.

Marsit, C.J., Eddy, K. and Kelsey, K.T. (2006) MicroRNA responses to cellular stress. *Cancer Res.* **66**, 10843–10848.

Martinez, I., Cazalla, D., Almstead, L.L., Steitz, J.A. and DiMaio, D. (2011) miR-29 and miR-30 regulate B-Myb expression during cellular senescence. *Proc. Natl Acad. Sci. USA* **108**, 522–527.

Matts, J., Jagadeeswaran, G., Roe, B.A. and Sunkar, R. (2010) Identification of microRNAs and their targets in switchgrass, a model biofuel plant species. *J. Plant Physiol.* **167**, 896–904.

Mochizuki, K. and Gorovsky, M.A. (2004) Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. *Genes Dev.* **18**, 2068–2073.

Mohorianu, I., Schwach, F., Jing, R., Lopez-Gomollon, S., Moxon, S., Szittya, G., Sorefan, K., Moulton, V. and Dalmay, T. (2011) Profiling of short RNAs during fleshy fruit development reveals stage-specific sRNAome expression patterns. *Plant J.* **67**, 232–246.

Murakami, Y., Yasuda, T., Saigo, K., Urashima, T., Toyoda, H., Okanoue, T. and Shimotohno, K. (2006) Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene*, **25**, 2537–2545.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., Guigo, R. and Shiekhattar, R. (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell*, **143**, 46–58.

Piriyapongsa, J., Marino-Ramirez, L. and Jordan, I.K. (2007) Origin and evolution of human microRNAs from transposable elements. *Genetics*, **176**, 1323–1337.

Rajagopalan, R., Vaucheret, H., Trejo, J. and Bartel, D.P. (2006) A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* **20**, 3407–3425.

Reinhart, B.J. and Bartel, D.P. (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science*, **297**, 1831.

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

Shen, H., He, X., Poovaiah, C.R., Wuddineh, W.A., Ma, J., Mann, D.G., Wang, H., Jackson, L., Tang, Y., Stewart, C.N. Jr, Chen, F. and Dixon, R.A. (2012) Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol.* **193**, 121–136.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J. and Birol, I. (2009) ABySS: a parallel assembler for short read sequence data. *Genome Res.* **19**, 1117–1123.

Sun, G., Stewart, C.N. Jr, Xiao, P. and Zhang, B. (2012) MicroRNA expression analysis in the cellulosic biofuel crop switchgrass (*Panicum virgatum*) under abiotic stress. *PLoS ONE*, **7**, e32017.

Sunkar, R., Kapoor, A. and 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**, 2051–2065.

Tanzer, A. and Stadler, P.F. (2004) Molecular evolution of a microRNA cluster. *J. Mol. Biol.* **339**, 327–335.

Wei, L.Q., Yan, L.F. and Wang, T. (2011) Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. *Genome Biol.* **12**, R53.

Xiao, C., Calado, D.P., Galler, G., Thai, T.H., Patterson, H.C., Wang, J., Rajewsky, N., Bender, T.P. and Rajewsky, K. (2007) MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb. *Cell*, **131**, 146–159.

Xie, Z., Kasschau, K.D. and Carrington, J.C. (2003) Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Curr. Biol.* **13**, 784–789.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E. and Carrington, J.C. (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2**, E104.

Xie, F., Frazier, T.P. and Zhang, B. (2010) Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). *Planta*, **232**, 417–434.

Xie, F., Xiao, P., Chen, D., Xu, L. and Zhang, B. (2012) miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* **80**, 75–84.

Yue, J.X., Meyers, B.C., Chen, J.Q., Tian, D. and Yang, S. (2012) Tracing the origin and evolutionary history of plant nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes. *New Phytol.* **193**, 1049–1063.

Zhang, B.H. and Pan, X.P. (2009) Expression of MicroRNAs in Cotton. *Mol. Biotechnol.* **42**, 269–274.

Zheng, Y., Zhao, L., Gao, J. and Fei, Z. (2011) iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. *BMC Bioinformatics*, **12**, 453.

Zhu, J.K. (2002) Salt and drought stress signal transduction in plants. *Annu. Rev. Plant Biol.* **53**, 247–273.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Length distribution of redundant (A) and unique (B) conserved known miRNAs in control, salt and drought libraries in switchgrass.

Figure S2 GO ontology classification of identified miRNA families in switchgrass.

Table S1 Jaccard index (%) for similarity of top 5000 most abundant small RNAs among control, salt and drought libraries in switchgrass.

Table S2 Summary of miRNA family comparison among control, salt and drought libraries in switchgrass.

Table S3 Known and novel miRNAs from control, salt and drought libraries in switchgrass.

Table S4 Novel miRNAs expression comparison among control, salt and drought libraries in switchgrass.

Data S1 Conversed miRNAs in switchgrass.
Data S2 Novel miRNAs in switchgrass.
Data S3 miRNA targets for conserved switchgrass miRNAs.
Data S4 miRNA targets for novel switchgrass-specific miRNAs.
Data S5 Gene pathway analysis for switchgrass miRNA targets based on GO and KEGG analysis.