Planta (2010) 232:417–434
DOI 10.1007/s00425-010-1182-1

ORIGINAL ARTICLE

Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*)

Fuliang Xie · Taylor P. Frazier · Baohong Zhang

Received: 19 January 2010 / Accepted: 23 April 2010 / Published online: 12 May 2010
© Springer-Verlag 2010

Abstract MicroRNAs (miRNAs) are a class of non-coding small endogenous RNAs with lengths of ~22 nucleotides (nt) that have been shown to regulate gene expression at the post-transcriptional levels by targeting mRNAs for degradation or by inhibiting protein translation. Although thousands of miRNAs have been identified in many species, miRNAs have not yet been identified in switchgrass (*Panicum virgatum*), one of the most important bioenergy crops in the United States and around the world. In this study, we identified 121 potential switchgrass miRNAs, belonging to 44 families, using a well-defined comparative genome-based computational approach. We also identified miRNA clusters and antisense miRNAs in switchgrass expressed sequences tags. These identified miRNAs potentially target 839 protein-coding genes, which can act as transcription factors, and take part in multiple biological and metabolic processes including sucrose and fat metabolism, signal transduction, stress response, and plant development. Gene ontology (GO) analysis, based on these targets, showed that 527 biological processes were involved. Twenty-five of these processes were demonstrated to participate in the metabolism of carbon, glucose, starch, fatty acid, and lignin and in xylem formation. According to pathway enrichment analysis based on Kyoto Encyclopedia of Genes and Genomes (KEGG), 118 metabolism networks were found. These networks are involved in sucrose metabolism, fat metabolism, carbon fixation, hormone regulation, oxidative stress response, and the processing of other secondary metabolites.

F. Xie · T. P. Frazier · B. Zhang (✉)
Department of Biology, East Carolina University,
Greenville, NC 27858, USA
e-mail: zhangb@ecu.edu

Keywords Comparative genome · Gene ontology · MicroRNA · *Panicum* · Pathway · Switchgrass

Abbreviations

DCL1	Dicer-like 1
EST	Expressed sequence tag
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
miRNA	MicroRNA
MFE	Minimal folding free energy
MFEI	Minimal folding free energy index
nt	Nucleotide
pre-miRNAs	miRNA precursor
NCBI	National Center for Biotechnology Information
RISC	RNA-induced silencing complex
SBP	Squamosa-promoter-binding protein-like protein
TAF	TBP-associated factors
TBP	TATA-binding protein

Introduction

MicroRNAs (miRNAs) are an extensive class of non-coding small endogenous RNAs with ~22 nt in length that are derived from self-complementary foldback structures of longer precursor sequences (pre-miRNAs) and are generated by Dicer-like 1 (DCL1) in plants (Bartel 2004). Mature miRNAs inhibit gene expression at the post-transcriptional levels by either targeting mRNAs for degradation or inhibiting protein translation. Both processes are

accomplished by the complementary base pairing of miRNAs to their target mRNA sequences (Ambros 2004). In plants, for a majority of cases, miRNAs interact with their targets through perfect or near-perfect base pairing and lead to target mRNA degradation (Jones-Rhoades et al. 2006). Increasing evidences have revealed that miRNAs play an important role in a wide range of development processes in plants including cell proliferation, stress response, metabolism, inflammation, and signal transduction (Ambros 2004; Jones-Rhoades et al. 2006; Zhang et al. 2007a).

To date, more than 10,000 miRNAs have been identified from 115 species and deposited in the publicly available database miRBase (Release 14) (Griffiths-Jones et al. 2008). The majority of plant miRNAs have been found in species with fully sequenced genomes including 190 from *Arabidopsis thaliana*, 234 from *Populus trichocarpa*, 414 from *Oryza sativa*, 109 from *Zea mays*, 140 from *Sorghum bicolor*, and 108 from *Medicago truncatula* (Griffiths-Jones et al. 2008). miRNA-related research is continuously growing and miRNAs, along with their functions, are being identified and elucidated using a wide variety of computational tools and experimental methods including direct cloning, deep sequencing, and other approaches. Comparison of miRNAs across multiple plant species has demonstrated that some miRNAs are highly evolutionary conserved from species to species, such as from mosses to higher flowering eudicots in the plant kingdom (Floyd and Bowman 2004; Zhang et al. 2006b). Conservation of miRNA sequences has provided a powerful strategy for identifying miRNAs in other species (Pan et al. 2007; Zhang et al. 2005). Currently, comparative genome-based homolog searches have been used to identify conserved miRNAs in many plant species, including cotton (Zhang et al. 2007b), mustard (Xie et al. 2007), soybean (Zhang et al. 2008), wheat (Jin et al. 2008), corn (Zhang et al. 2006a), tomato (Pilcher et al. 2007), potato (Zhang et al. 2009), citrus (Song et al. 2009), and apple (Gleave et al. 2008).

Switchgrass (*Panicum virgatum*) is a warm-season perennial grass commonly grown in the midwest and grassland areas in the United States. Switchgrass has recently received extensive attention due to its huge potential in the development of cellulosic biofuels (Bouton 2007). Actually, switchgrass has become one of the most important dedicated bioenergy crops for North America. Its use in the production of liquid fuel, such as ethanol, is thought of being one of the important potential energy alternatives to replace dependence on fossil fuels (Chen and Schnoor 2009; Uppugundla et al. 2009). Switchgrass has an important economic and environmental value. To our knowledge, although progress has been made on switchgrass, there is little knowledge about miRNAs in switchgrass. In this study, we employed a well-defined comparative genome-based homolog search to identify switchgrass miRNAs. We also investigated the potential functions of predicted switchgrass miRNAs, particularly in the formation of switchgrass biomass and in biofuel-related biological and metabolic processes.

Materials and methods

Sequence databases

A total of 1,699 known plant miRNAs were downloaded and used as a reference miRNA set for identifying conserved miRNAs in switchgrass. These miRNAs are all currently available miRNAs deposited in miRBase database (http://www.mirbase.org/, Release 14: Sept 2009) (Griffiths-Jones et al. 2008); these miRNAs come from 29 plant species, including *A. thaliana*, *O. sativa (rice)*, *P. trichocarpa*, *Brassica napus*, soybean (*Glycine max*), *M. truncatula*, *Physcomitrella patens*, *Saccharum officinarum*, *S. bicolor*, and *Z. mays*.

Switchgrass expressed sequences tags (ESTs) and protein databases were obtained from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/, NCBI). Currently, a total of 436, 535 ESTs and 52 protein sequences are available for switchgrass in the NCBI database. All EST sequences were used for predicting conserved miRNAs as well as for identifying potential miRNA targets.

The GO database was downloaded from the Gene Ontology website (http://www.geneontology.org/GO.downloads.shtml) (Ashburner and Bergman 2005). The KEGG database was obtained from the KEGG website (ftp://ftp.genome.jp/pub/kegg/pathway/) (Kanehisa and Goto 2000).

Software

The alignment tool WATER was employed to identify potential conserved miRNAs and their targets and was downloaded as the EMBOSS package (EMBOSS 6.1.0.1) from the public EMBOSS website (http://emboss.sourceforge.net/) (Smith and Waterman 1981). BLAST 2.2.19 was downloaded from NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/) and used for removing repeated sequences and protein-coding genes. RNAfold was obtained from Vienna RNA Package 1.8.4 (http://www.tbi.univie.ac.at/~ivo/RNA/index.html) (Zuker and Stiegler 1981; Hofacker 2003). MySQL was used for managing all data and was downloaded from the website (http://www.mysql.com/). Perl scripts were developed for data mining in the identification of miRNAs and their targets.

Identification of conserved miRNAs in switchgrass from EST by homologs search

Comparative genome-based EST analysis is a well-established approach to identify conserved miRNAs in one species using already known miRNAs in another species. Since it was developed, EST analysis has been widely used to identify conserved miRNAs in many plant species, including cotton, soybean, oilseed, tomato, and apple. One big issue for current EST analysis is that traditional Blastn searches overlook a lot of potential miRNAs because it is difficult for Blastn searches to align two sequences with deletions, insertions, and gaps within such a small query sequence. Thus, Blastn searches are not an ideal tool to identify conserved small RNA sequences, including miRNAs. To avoid this issue, we adopted WATER to identify potential conserved miRNAs in the bioenergy crop switchgrass.

All of the databases and software were downloaded from the previously mentioned websites. miRNA predication was performed locally using a high performance computer as described in our previous reports. Briefly, we used WATER to align all known mature miRNA sequences to all switchgrass EST sequences in order to identify potential homologs with no >2 nt substitutions, including deletion, and insertion mutations. After removing the repeated and protein-coding sequences and considering proper secondary structure, the sequences only fitting the following criteria were considered potential miRNAs in switchgrass: (1) there were no >2 nt substituted between the EST sequence and the query miRNA sequence; (2) the minimum length of the pre-miRNA was 45 nt; (3) the pre-miRNA could be folded into a perfect stem-loop hairpin secondary structure with the miRNA sitting in one arm of the stem at either the 5′ or 3′ end; (4) there were no more than six nucleotides mismatched between the predicted mature miRNA sequence and its opposite miRNA* sequence in the secondary structure; (5) there were no loops or breaks in the miRNA: miRNA* complex; and (6) the predicted pre-miRNA sequences had a high minimal folding energy (MFE) and MFE index (MFEI). By using these criteria, we could significantly reduce the total number of sequences for subsequent analyses, ultimately saving time and increasing work efficiency. More importantly, the application of these criteria significantly reduced the total number of false miRNA predictions.

Prediction of miRNA targets in switchgrass

Growing evidences have shown that most plant miRNAs function by either perfectly or near-perfectly binding to complementary sites on their target mRNA sequences (Schwab et al. 2005). This provides a powerful way to identify potential miRNA targets simply by aligning and comparing miRNAs with potential target sequences. The criteria for prediction of potential miRNA targets in switchgrass was similar to that described by Schwab and her co-workers but with some modifications (Schwab et al. 2005; Zhang et al. 2008). WATER was employed as an alignment tool to predict miRNA target sequences. Because only a small number of protein-coding genes were reported in switchgrass, all switchgrass ESTs were also used to predict potential miRNA targets. Switchgrass ESTs were Blastx searched against the *Arabidopsis* protein database in order to identify potential protein-coding homolog genes in switchgrass. Repeated protein-coding sequences with an E value of 1e−25 were removed by additional Blastx searches against the switchgrass protein/EST database and the non-redundant database on NCBI. The predicted target genes with unknown function were discarded. In this research, the following criteria were used for identifying potential miRNA targets: (1) no more than four mismatches were allowed between the mature miRNA and its potential target site; (2) no more than one mismatch was allowed at nucleotide positions 1–9; (3) no more than two consecutive mismatches were allowed; and (4) no mismatches were allowed at positions 10 and 11. These criteria significantly reduced the total number of false-positive targets.

Analysis of GO and KEGG pathway

GO and KEGG pathway analyses were employed to further investigate the biological processes and corresponding metabolic networks regulated by potential miRNAs. We constructed database structures on MySQL, which was used for managing and mining KEGG information. All predicted targets with an E value of 1e−30 were identified by Blastx searching against the GO protein database. GO and pathway analyses were performed by using a combined query search against the GO and KEGG databases. In order to investigate the relationship between the miRNAs and the KEGG pathway, we adopted a formula 1 to convert E values into a relation, which is expressed as an integer from 0 to 16. A large relation number indicates a close relationship. The formula is listed as follows:

$$\text{relation} = \text{int}\left(\frac{\text{abs}(\ln(E\ \text{value}))}{\max(\in \text{abs}(\ln(E\ \text{value}))) - \min(\in \text{abs}(\ln(E\ \text{value})))/17}\right) \quad (1)$$

where abs is the absolute value of the function, ln is the natural logarithm function having base e in the set, Max is the function of retrieving maximum value of a set, Min is the function of retrieving minimum value of a set, and Int is the function for returning integer of a parameter.

Results

Identifying potential miRNAs in switchgrass

There, currently, are 1,669 miRNAs from 29 plant species deposited in the miRBase database. After removal of the repeated miRNAs, a total of 755 unique miRNA sequences were collected. These 755 unique miRNAs belong to 321 families. In order to identify as many miRNAs as possible in the bioenergy crop switchgrass, we used all currently available 755 unique plant miRNA sequences as queries. After aligning the 755 plant miRNAs with 436,535 switchgrass ESTs, removing the repeated sequences, and considering proper secondary structure, we were able to identify 121 conserved miRNAs in switchgrass, which belong to 44 miRNA families (Table 1). This indicates that miRNAs widely exist in switchgrass and further demonstrates that many miRNAs are highly evolutionary conserved among species in the plant kingdom.

Characterization of microRNAs in switchgrass

The 121 identified switchgrass miRNAs belong to 44 families with an average of about 3 miRNA members per family. The size of a miRNA family varies from family to family; some miRNA families have a larger number of members, but for a majority of families, only one member has been identified (Fig. 1). The miR-444 family has the largest number (13) of members followed by the miR-414 family with 11 members. Both the miR-169 and the miR-2102 families have seven members. There are three miRNA families (miR-156, miR-167, and miR-531) that each contain six members. In this study, for 19 out of the 44 miRNA families, we only identified one member. The remaining 18 miRNA families have 2–4 members, respectively.

miR-444 was originally identified in rice by a direct cloning approach (Sunkar et al. 2005). Currently, miR-444 has only been identified in two other species, wheat (Yao et al. 2007) and *Brachypodium distachyon* (Unver and Budak 2009), and each species contains only one family member. This suggests that miR-444 only exists in a limited number of plant species and that it is potentially a miRNA specific to monocots. In this study, we identified 13 members of the miR-444 family in switchgrass, which is much larger than the number found in other plant species. Another interesting miRNA is miR-414. Although miR-414 has been identified in *Arabidopsis* (Wang et al. 2004) and rice (Wang et al. 2004), for more than 6 years miR-414 had not been identified in another plant species except moss (*P. patens*) (Fattash et al. 2007). Because *Arabidopsis* and rice belong to two different plant groups (dicots and monocots) and miR-414 was identified in both of these groups as well as in moss, this suggests that miR-414 existed in the plant kingdom several million years ago before the divergence of plants into dicots and monocots. Therefore, miR-414 should be found in almost all plant species. However, it seems that miR-414 is only limited to several specific plant species. In this study, we identified miR-414 in a fourth plant species, switchgrass. Interestingly, miR-414 has more members in switchgrass than in any other species in which miR-414 has been identified. It is unclear what has caused this evolution pattern.

Mature miRNA sequences have been shown to be located on either arm of the secondary stem-loop hairpin structure of the potential pre-miRNA. Of the 121 identified switchgrass miRNAs, 53 (43.8%) were found to be located on the 5′ arm of the stem-loop hairpin structure while 68 (56.2%) resided on the 3′ arm.

The length of switchgrass miRNAs varies from 18 to 24 nt with an average of 20.6 ± 1.0 nt (Fig. 2a). A majority (62 out of 121 or 51.2%) of miRNAs are 21 nucleotides in length. The length of switchgrass pre-miRNA also varies from 46 to 707 nt with an average of 181 ± 135 nt. However, a majority of the pre-miRNAs are 60–139 nt in length (Fig. 2b). The length distribution of miRNAs and their precursor sequences are similar to previous reports in other plant species (Zhang et al. 2006c, 2007b, 2008).

MFE is very important for RNAs forming their secondary structures. Generally speaking, the lower the MFE, the more stable the secondary structure of a RNA sequence. The average value of MFEs was -78.04 ± 60.37 kcal/mol with a range of -10.3 to -395.4 kcal/mol (Fig. 2c). MFEI is a criterion for distinguishing miRNAs from other RNAs. Previous studies have shown that it is more likely to be a potential miRNA if a sequence has a MFEI value >0.85 (Zhang et al. 2006c). For the newly identified 121 switchgrass miRNAs, the average MFEI was 0.86 ± 0.25 with a range of 0.39-1.73 (Fig. 2d).

miRNA clusters in switchgrass

In animals, many miRNAs have been shown to cluster together and have been speculated to have similar expression profiles and functions (Altuvia et al. 2005; Seitz et al. 2004; Tanzer et al. 2005; Tanzer and Stadler 2004; Yu et al. 2006). However, miRNA clusters have rarely been observed in plants. Currently, there are only several clusters have been identified in plants (Jones-Rhoades and Bartel 2004; Talmor-Neiman et al. 2006; Zhang et al. 2006b, 2007b). In this study, due to limited resources, we identified only one miRNA cluster in switchgrass. This cluster includes two miRNAs (miR-2118a and miR-2118b) within the same EST sequence with a distance of 139 nt

Table 1 Switchgrass miRNA identification by homolog search and its secondary structure

miRNA	Query miRNAs	Mature sequences	NM	LM	Location	LP	*C* (%)	MFE	MFEI	EST
pvi-miR156a	ath-miR156d	UGACAGAAGAGAGUGAGCAC	0	20	5′	84	51.19	53.9	1.25	FL903027
pvi-miR156b	ath-miR156d	UGACAGAAGAGAGUGAGCAC	0	20	5′	84	47.62	47.1	1.18	GD028411
pvi-miR156c	ath-miR156d	UGACAGAAGAGAGAGAGCAC	1	20	3′	414	54.59	154	0.68	FL801801
pvi-miR156d	ath-miR156d	UGACAGAAGAGAGAGAGCAC	1	20	5′	424	47.41	156.7	0.78	FL927913
pvi-miR156e	ath-miR156d	UGACAGAAGAGAGUGAGCAU	1	20	5′	118	53.39	62.9	1	GD025154
pvi-miR156f	ath-miR156g	CGACAGAGAGAGAGAGCAC	2	19	5′	355	43.94	100.3	0.64	FL821810
pvi-miR159a	osa-miR159a_1	UUUGGAUUGAAGGGAGCUCUG	0	21	3′	221	53.85	104.6	0.88	FL997282
pvi-miR159b	osa-miR159a_1	UUUGGAUUGAAGGGAGCUCUG	0	21	3′	214	53.74	96.4	0.84	GD036203
pvi-miR160a	bdi-miR160	UGCCUGGCUCCCUGUAUGCC	0	20	5′	81	61.73	46.9	0.94	FL819171
pvi-miR160b	bdi-miR160	UGCCUGGCUCCCUGUAUGCC	0	20	5′	124	63.71	69.4	0.88	FE655198
pvi-miR164a	ath-miR164b	UGGAGAAGCAGGGCACGUGCU	1	21	5′	88	64.77	54.2	0.95	FL699795
pvi-miR164b	ath-miR164c	UGGAGAAGCAGGUCACGUGUG	2	21	3′	198	69.7	95	0.69	FL924649
pvi-miR164c	osa-miR164d	UGGAGAAGGAGCGCACGUGCU	2	21	5′	88	64.77	49.2	0.86	FL699795
pvi-miR166a	sbi-miR166i	UCGGACCAGGCUUCAUUCCC	0	20	3′	104	50	53.4	1.03	GD002178
pvi-miR166b	ath-miR166a	UCGGACCAGGCUUCAUUCCCC	0	21	3′	151	58.94	74.2	0.83	FL954558
pvi-miR167a	ath-miR167b	UGAAGCUGCCAGCAUGAUCUA	0	21	5′	100	55	56.9	1.03	FL992031
pvi-miR167b	gma-miR167c	UGAAGCUGCCAGCAUGAUCUG	0	21	5′	89	44.94	43.9	1.1	GD007307
pvi-miR167c	ath-miR167b	UGAAGCUGCCAGCAUGAUCUG	1	21	5′	122	45.9	58.5	1.05	FL987206
pvi-miR167d	ath-miR167b	UGAAGCUGCCAGCAUGAUCUG	1	21	5′	123	46.34	57.3	1.01	FL981017
pvi-miR167e	ptc-miR167h	UGAAGCUACAACAUGAUCUG	2	20	5′	91	51.65	38.7	0.82	GD032712
pvi-miR167f	ptc-miR167h	UGACGCUGCAAACAUGAUCUG	2	21	3′	145	56.55	58.4	0.71	FL834618
pvi-miR168a	osa-miR168b	GGGCUUGGUGCAGCUCGGGAA	1	21	3′	449	50.78	161.4	0.71	FL904157
pvi-miR168b	ath-miR168a	UCGCUGGUGCAGGUCGGGAA	1	20	5′	292	74.32	145	0.67	FL875971
pvi-miR169a	osa-miR169n	UAGCCAAGAAUGACUUGCCUA	0	21	5′	94	43.62	45.7	1.11	GD008229
pvi-miR169b	ath-miR169m	UAGCCAAGGAUGACUUGCCUG	0	21	5′	94	51.06	45.2	0.94	GD036278
pvi-miR169c	ath-miR169m	UAGCCAAGGAUGACUUGCCUG	0	21	5′	143	55.24	82.3	1.04	GD017574
pvi-miR169d	ath-miR169m	UAGCCAAGGAUGACUUGCCUG	0	21	5′	147	52.38	74.3	0.96	GD018854
pvi-miR169e	bna-miR169e	UAGCCAAGGAUGACUUGCUUA	1	21	5′	86	55.81	48.8	1.02	GD018960
pvi-miR169f	vvi-miR169v	AAGACAAAGGAUGAAUUGCCGG	2	22	5′	433	47.11	131.2	0.64	FE605113
pvi-miR169g	vvi-miR169i	GAGCCAAGGAUGUCUGUCCGU	2	21	3′	428	59.58	182.8	0.72	FL854502
pvi-miR171a	mtr-miR171d	UGAUUGAGCCGUGCCAAUAUC	0	21	3′	86	56.98	41.4	0.84	FL973882
pvi-miR171b	sly-miR171d	UUGAGCCGCGCCAAUAUCUC	1	20	3′	426	69.01	216.2	0.74	FL815972
pvi-miR171c	sly-miR171d	UUGAGCCGCGCCAAUAUCUC	1	20	3′	426	69.25	219.4	0.74	FL869923
pvi-miR171d	sly-miR171c	UAUUGGUGUGGUUCAAUGUGA	2	21	3′	207	38.65	51.2	0.64	DN140671
pvi-miR319	osa-miR319b	UUGGACUGAAGGGUGCUCCC	0	20	3′	170	52.35	84.8	0.95	FL985594
pvi-miR393	bna-miR393	UCCAAAGGGAUCGCAUAUC	2	19	3′	245	47.35	70.1	0.6	FL778463
pvi-miR396	osa-miR396e	UCCACAUGCUUUCUUGAACGG	2	21	3′	264	47.35	88.6	0.71	FE617330
pvi-miR397a	sly-miR397	AUUGAGUGCAGCGUUGAUGA	0	20	5′	71	70.42	57.3	1.15	FE599510
pvi-miR397b	ptc-miR397b	CCAUUGAGUGCUGCUUGAUG	2	20	5′	71	47.89	21.7	0.64	FL787418
pvi-miR398	bol-miR398aA	GGGGUGUCAUAGAACACGGA	2	20	5′	58	70.69	23.9	0.58	FL751004
pvi-miR399a	tae-miR399	UGCCAAAGGAGAAUUGCCC	0	19	3′	120	53.33	65	1.02	GD003499
pvi-miR399b	ath-miR399f	UGCCAAAGGAGAUUUGCCCGG	0	21	3′	76	65.79	45.6	0.91	FE614492
pvi-miR408a	osa-miR408	CUGCACUGCCUCUUCCCUGGC	0	21	3′	172	52.91	80.5	0.88	FL942917
pvi-miR408b	osa-miR408	CUGCACUGCCUCUACCCUGGC	1	21	5′	192	78.13	106	0.71	FL777418
pvi-miR408c	osa-miR408	CUGCCUUCCUCUUCCCUGGC	2	20	5′	166	55.42	46.8	0.51	FL883403
pvi-miR414a	ath-miR414	UCAUCUUCCUCAUCAUCGUCA	1	21	3′	215	47.44	51.5	0.51	FL733313
pvi-miR414b	ath-miR414	UCAUCUUCGUCAUCAUCGUCA	1	21	5′	206	59.71	89.6	0.73	FL721570
pvi-miR414c	ath-miR414	UCAUCAUCAUCAUCAUCAUCA	2	21	3′	396	55.81	135.4	0.61	FL790782

Table 1 continued

miRNA	Query miRNAs	Mature sequences	NM	LM	Location	LP	*C* (%)	MFE	MFEI	EST
pvi-miR414d	ath-miR414	UCAUCAUCAUCAUCAUCAUCA	2	21	3′	75	46.67	13.7	0.39	FL761332
pvi-miR414e	osa-miR414	UCAUCAUCAUCAUCAUCAUCC	2	21	5′	190	39.47	42	0.56	FL939995
pvi-miR414f	ath-miR414	UCAUCAUCUUCAUCAUCGUCA	2	21	3′	318	44.03	81.6	0.58	FL988732
pvi-miR414g	osa-miR414	UCAUCCUCAUCAUCAUCCGUCG	2	22	5′	69	56.52	24.2	0.62	FL947171
pvi-miR414h	ath-miR414	UCAUCCUCAUCAUCAUCGUCU	2	21	3′	288	43.06	64.9	0.52	FL926164
pvi-miR414i	osa-miR414	UCAUCCUCGUCAUCAUCUCC	2	20	3′	109	63.3	47.9	0.69	FL769562
pvi-miR414j	ath-miR414	UCAUCUUCAUCAUCUUCUUCA	2	21	5′	232	42.67	53.4	0.54	FL947785
pvi-miR414k	ath-miR414	UCAUGUUCAGCAUCAUCGUCA	2	21	3′	73	61.64	26.8	0.6	FL798262
pvi-miR437	osa-miR437	AAAGUUAGAGAAGUUUGACUU	0	21	3′	181	27.07	42.7	0.87	FL707816
pvi-miR442	osa-miR442	UGAUUGUAAAUUGCGAGACGAAU	2	23	5′	129	33.33	62.1	1.44	FL805097
pvi-miR444a	osa-miR444f	UGCAGUUGUUGCCUCAAGCUU	0	21	3′	111	42.34	61.8	1.32	FL855387
pvi-miR444b	osa-miR444f	UGCAGUUGUUGCCUCAAGCUU	0	21	3′	58	48.28	20.5	0.73	FL971537
pvi-miR444c	osa-miR444c_1	UGUUGUCUCAAGCUUGCUGCC	0	21	3′	108	45.37	58.9	1.2	FL964016
pvi-miR444d	osa-miR444f	UGCAGUUGUUGUCUCAAGCUU	1	21	3′	96	41.67	50.8	1.27	FL983759
pvi-miR444e	osa-miR444f	UGCAGUUGUUGUCUCAAGCUU	1	21	3′	96	41.67	50.8	1.27	FL964016
pvi-miR444f	osa-miR444f	UGCAGUUGUUGUCUCAAGCUU	1	21	3′	96	41.67	50.8	1.27	FL989890
pvi-miR444g	osa-miR444f	UGCAGUUGUUGUCUCAAGCUU	1	21	3′	96	41.67	50.8	1.27	FL955886
pvi-miR444h	osa-miR444f	UGCAGUUGUUGUCUCAAGCUU	1	21	3′	96	41.67	50.8	1.27	FL985689
pvi-miR444i	osa-miR444c_1	UGUUGCCUCAAGCUUGCUGCC	1	21	3′	123	46.34	72.1	1.27	FL694167
pvi-miR444j	osa-miR444c_1	UGUUGCCUCAAGCUUGCUGCC	1	21	3′	116	46.55	70.9	1.31	FL790300
pvi-miR444k	osa-miR444c_1	UGUUGCCUCAAGCUUGCUGCC	1	21	3′	116	46.55	70.9	1.31	FL790301
pvi-miR444l	bdi-miR444	GUGCUGCCUCAACUUGCUGC	2	20	3′	523	48.57	163.7	0.64	FL803302
pvi-miR444m	osa-miR444c_2	UGCCGUUGAUGUCUCAAGCUU	2	21	3′	630	74.6	306.7	0.65	FL725883
pvi-miR477a	aqc-miR477c	CUCUCCCUACAAGCUCUUCUA	2	21	3′	340	69.12	174.3	0.74	FL952239
pvi-miR477b	aqc-miR477c	CUCUCUCUCAAGUUCUUCUU	2	20	5′	75	41.33	25.7	0.83	FL960944
pvi-miR477c	aqc-miR477d	GUCUUCCUCAAAGGCUUCUA	2	20	5′	46	50	19.6	0.85	FL971489
pvi-miR528	osa-miR528	UGGAAGGGGCAUGCAGAGGAG	0	21	5′	86	60.47	42.7	0.82	FL966875
pvi-miR531a	osa-miR531b	CGCGCCGGGGCUGCAGUGCCG	2	21	5′	136	69.85	57	0.6	FL776075
pvi-miR531b	osa-miR531b	CGCGCCGGGGCUGCAGUGCCG	2	21	3′	121	67.77	50.6	0.62	FL778754
pvi-miR531c	osa-miR531b	CUCACCGGAGCUGCGUGCCG	2	20	5′	47	76.6	35.7	0.99	FL866875
pvi-miR531d	osa-miR531b	CUCGCCGGCGCGCGUGCCG	2	19	3′	103	75.73	61	0.78	FL866963
pvi-miR531e	osa-miR531b	CUCGCCGGGGCGCGUGUCG	2	19	3′	138	76.09	54.9	0.52	FL917916
pvi-miR531f	osa-miR531b	CUCGCCGGGGCUCGUCCG	2	18	3′	216	70.83	95.1	0.62	FL836541
pvi-miR821a	sbi-miR821e	AAAUGAUCAAAAUAAAAGUUG	2	21	5′	77	18.18	11.7	0.84	FL715679
pvi-miR821b	sbi-miR821b	AACUUAUGAACAUAGAAGUUG	2	21	3′	47	40.43	10.3	0.54	FL795568
pvi-miR827	osa-miR827a	UUAGAUGACCAUCAGCAAACA	0	21	3′	127	51.18	63.3	0.97	FL979719
pvi-miR831	ath-miR831	UGAUCUCUUCUUCCUCUUCUUG	2	22	5′	398	69.35	216.4	0.78	DN146798
pvi-miR834	ath-miR834	UGGUAGCAGUAGGGGUUGUAA	2	21	3′	59	59.32	20.5	0.59	FL892777
pvi-miR837	ath-miR837-5p	ACCAUUUCUUGUUCGUUUCA	2	20	5′	180	41.67	52.42	0.7	GD003235
pvi-miR854a	ath-miR854a	GAGGAAGAUAGGGAGGAGGAG	2	21	3′	81	53.09	30	0.7	FL887161
pvi-miR854b	ath-miR854a	GAUGAGGACGGGGAGGAGGAG	2	21	3′	223	71.75	128.7	0.8	FL696643
pvi-miR854c	ath-miR854a	GAUGGGAUGGGGAGGAGGAG	2	20	3′	189	67.2	90.69	0.71	GD015426
pvi-miR864	ath-miR864-5p	UCAGGUAUGAUUGAUGCAAA	2	20	3′	90	41.11	25.9	0.7	GD018310
pvi-miR869	ath-miR869_2	UCUGGUGUUGAGAUUAGUUGAG	2	22	3′	149	43.62	41.5	0.64	FL718154
pvi-miR1128a	tae-miR1128	UACAUACUCCCUCCGUCCCAAA	2	22	5′	340	35.59	101.02	0.83	FL699638
pvi-miR1128b	tae-miR1128	UACUACUCCCUCCGUCUCAAA	2	21	5′	135	32.59	51.9	1.18	FE604006
pvi-miR1132a	tae-miR1132	CAUUUUGGAACGGAAGGAG	1	19	3′	103	37.86	49	1.26	FL864662
pvi-miR1132b	tae-miR1132	AAUUUGGAACGGAAGGAG	2	18	3′	147	36.73	93.3	1.73	FL805977

Table 1 continued

miRNA	Query miRNAs	Mature sequences	NM	LM	Location	LP	*C* (%)	MFE	MFEI	EST
pvi-miR1132c	tae-miR1132	AAUUUGGAACGGAAGGAG	2	18	3′	260	25.38	98.6	1.49	FL855194
pvi-miR1132d	tae-miR1132	CAUUUUGGAACGGAAGUAG	2	19	3′	173	33.53	54.2	0.93	FL963269
pvi-miR1133	tae-miR1133	UAUAUACUCCCUCCGUCCCAAA	2	22	5′	161	34.16	75	1.36	FE612618
pvi-miR1424	osa-miR1424	AUGCACAAAGAUGCUGAUUGU	2	21	5′	48	39.58	16	0.84	FL798174
pvi-miR1436a	osa-miR1436	ACAUUUUAGGACGGAGGGAGU	2	21	3′	162	38.27	79.7	1.29	FL883694
pvi-miR1436b	osa-miR1436	ACAUUUUGGAACGGAGGGAGU	2	21	3′	149	34.23	64.1	1.26	FL863283
pvi-miR1535a	gma-miR1535	CUUAUUUGUGAUGAUGUCU	2	19	5′	178	35.96	45.1	0.7	FL847588
pvi-miR1535b	gma-miR1535	CUUGGUGUGGUGAUGUCU	2	18	5′	271	53.51	99.7	0.69	FL768838
pvi-miR1848a	osa-miR1848	CCUCCCGGCGCGCGUCGUGCA	2	21	5′	106	67.92	45.7	0.63	FL776983
pvi-miR1848b	osa-miR1848	CCUCGCCGGCGCGCCCUGCA	2	20	3′	707	72.56	395.4	0.77	FL802862
pvi-miR1848c	osa-miR1848	CCUCGCCGGCGCUGCGCGUGUA	2	22	5′	99	74.75	46	0.62	DN144104
pvi-miR1852a	osa-miR1852	AUAUGGAUUCAAUGCAGGU	2	19	5′	603	45.27	196.82	0.72	FL998509
pvi-miR1858	osa-miR1858a	GAGCGGAGGCGGAGUGGGGC	2	20	3′	135	71.11	67.7	0.71	FL724130
pvi-miR1875	osa-miR1875	ACAAUGGAGUGAGGUGCAACACAA	2	24	3′	252	46.03	111.6	0.96	FL928066
pvi-miR2102a	osa-miR2102-5p	GGGCAGGCCGCCGCCGCCGC	2	20	3′	204	76.96	105.6	0.67	FL726074
pvi-miR2102b	osa-miR2102-5p	GGGCAGGCCGCCGCCGCCGC	2	20	5′	49	89.8	31.2	0.71	FE617504
pvi-miR2102c	osa-miR2102-5p	GGGCAGGCCGCCGCCUCCAC	2	20	3′	109	81.65	65	0.73	FL812773
pvi-miR2102d	osa-miR2102-5p	GGGCCGCCGCCGCCGCCAC	2	19	5′	55	87.27	38.8	0.81	FL780069
pvi-miR2102e	osa-miR2102-5p	GGGCGCCGCCGCCGCCAC	2	18	3′	286	79.02	149.3	0.66	FL794677
pvi-miR2102f	osa-miR2102-5p	GGGCUGCCGCCGCCGCCAC	2	19	5′	66	80.3	30.3	0.57	FL812607
pvi-miR2102g	osa-miR2102-5p	GUGCAACGCCGCCGCCGCCAC	2	21	5′	269	69.89	140.7	0.75	FL846143
pvi-miR2118a	osa-miR2118g	UUCCCAAUGCCUCCCAUUCCUA	1	22	3′	105	48.57	44.7	0.88	FL967393
pvi-miR2118b	osa-miR2118d	UUCCUGAUGCCUCCCAUGCCUG	1	22	3′	75	52	39.4	1.01	FL967393
pvi-miR2118c	osa-miR2118d	UUCCUGAUGCCUCCUAUGCCUA	1	22	3′	90	51.11	38.3	0.83	FE643345
pvi-miR2911	peu-miR2911	GGCCGGGGGAGGGCGGGA	2	18	5′	549	68.67	278.4	0.74	FL882704

separating the two (Fig. 3). Based on our best knowledge, this is the first time a cluster involving miR-2118ab has been identified in plants.

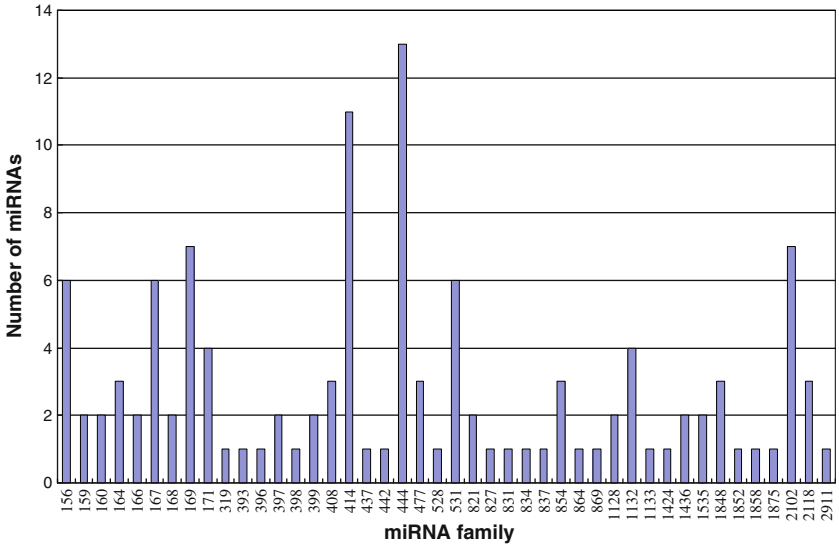
Antisense miRNAs in switchgrass

Antisense miRNAs are another class of miRNAs that were identified for the first time in invertebrates and vertebrates, including fruit flies and humans (Bender 2008; Stark et al. 2008; Tyler et al. 2008). In this case, miRNAs are transcribed and processed from both sense and antisense transcripts derived from the same genomic loci. In our previous study, we identified five pairs of sense and antisense miRNAs in soybeans and these five pairs belonged to three miRNA families (miR-157, miR-162 and miR-396) (Zhang et al. 2008). In this study, we found one more miRNA, miR-164, with an antisense miRNA in switchgrass (Fig. 4). Since the mature sequence and precursor sequences of sense/antisense miRNAs are different, we propose that they might have different targets or implement their functions through different mechanisms in plants (Zhang et al. 2008).

Target prediction

We adopted more stringent criterion (Schwab et al. 2005) to predict the potential targets of the 121 identified miRNAs in switchgrass. After carefully considering the aligned results, we identified at least one target for each miRNA family and a total of 839 potential targets were predicted. In this study, we identified many miRNA targets that are conserved across several plant species, including *Arabidopsis*, rice, poplar, cotton, soybean, and corn.

Many studies have demonstrated, by experimental and/or computational approaches, that miRNAs target many transcription factors that help control plant development. We also found this class of targets in switchgrass. MYB transcription factors represent a family of proteins with a conserved MYB DNA-binding domain. This domain was considered to be involved in regulation of secondary metabolism, control of cellular morphogenesis, and regulation of meristem formation and the cell cycle (Jin and Martin 1999). Our results show that MYB proteins might be the target of miR-156, miR-166, miR-414, and miR-2102 in switchgrass. The general transcription factor



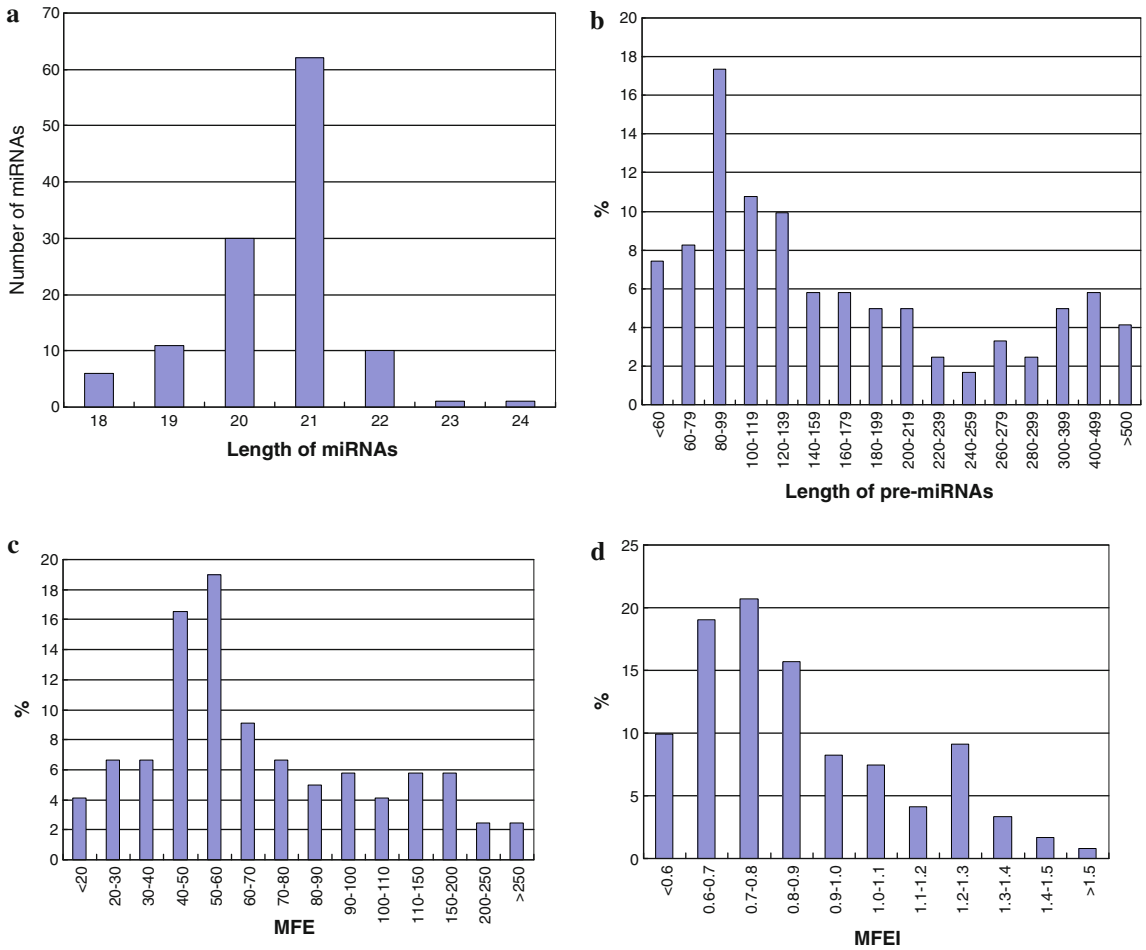


Fig. 1 Size of miRNA families in switchgrass

Fig. 2 Characterization of miRNAs in switchgrass. Distribution of mature miRNA length (**a**), and the length (**b**), MFE (**c**) and MFEI (**d**) of pre-miRNA sequences

TFIID, comprised of the TATA-binding protein (TBP) and a set of 13–14 TBP-associated factors, is employed to the promoters of active, and possibly repressed, genes (Cler et al. 2009). In switchgrass, we predict that miR-2102 directs the regulation of TFIID. Aside from WRKY, MYB, and TFIID, there are several transcription factor groups that

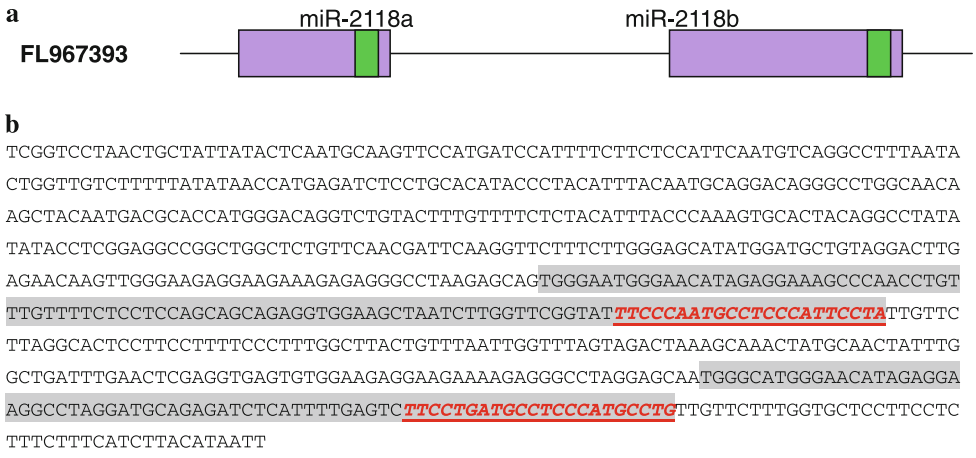


Fig. 3 miRNA-2118a–2118b cluster in switchgrass EST FL967393. **a** A schematic diagram of the organization of the cluster. **b** The EST sequence containing the miRNAs encoded within the cluster. *Shadowed sequences* represent pre-miRNAs; *underlined sequences* represent the mature miRNAs. **c** The predicted secondary structures of miR-2118a and miR-2118b

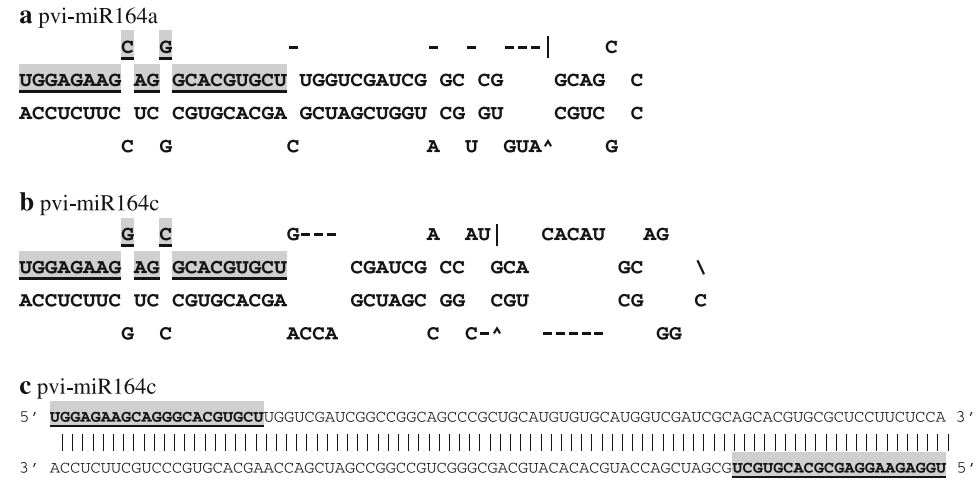


Fig. 4 Sense and antisense miRNAs and their corresponding secondary structures in switchgrass. **a** pvi-miR164a, **b** pvi-miR164c, and **c** alignment of pre-miR164a/164c

have been detected to be targets of miRNAs, like GATA transcription factor (miR-414, miR-2102), BHLH transcription factor (miR-2102), and Zinc finger protein (miR-531, miR-2102, miR-2911).

Current studies suggest that miRNAs are involved in plant response to environmental stresses, such as drought, salinity, low temperature, and nutrient deficiency (Chiou 2007; Jagadeeswaran et al. 2009; Jones-Rhoades and Bartel 2004; Jung and Kang 2007; Lu et al. 2005; Sunkar et al. 2006, 2007; Sunkar and Zhu 2004; Zhang et al. 2005). Several miRNAs also target stress-inducible genes. In this study, we identified that 63 stress-related genes were potential targets of 18 miRNA families, including miR-156, miR-167, miR-169, miR-171, miR-397, miR-408, miR-414, miR-477, miR-531, miR-831, miR-854, miR-1132, miR-1436, miR-1535, miR-1858, miR-2102, miR-2118,

Table 2 Potential targets of the identified miRNAs in switchgrass

miRNA	Target protein	Target function	Accession ID
156	Seed maturation protein	Development	FE658455
156	Cyclin IaZm	Development	FL883896
156	Inositol-3-phosphate synthase	Metabolism	FE628206
156	Phosphonates ABC transport system, substrate-binding periplasmic protein precursor	Metabolism	FE640193
156	Anthocyanidin 5,3-*O*-glucosyltransferase	Metabolism	FL751061
156	Para-nitrobenzyl esterase	Metabolism	FL808072
156	Histidine-containing phosphotransfer protein 5	Metabolism	FL954523
156	CBS domain-containing protein	Metabolism	FL959779
156	Auxin Efflux Carrier family protein	Signal transduction	DN146597
156	Heat shock protein-binding protein	Stress response	FL882624
156	DNA-binding protein	Transcription factor	FE605304
156	MYB-CC type transfactor	Transcription factor	FL914799
164	Sucrose synthase2	Biofuel	FL802595
164	C3 phosphoenolpyruvate carboxylase	Metabolism	FL962781
164	WRKY transcription factor 13	Transcription factor	FL827746
164	ABI3-interacting protein 2	Transcription factor	GD038087
166	Cellulose synthase 11	Biofuel	FL806484
166	G3PX_HORVU RecName	Metabolism	FL789117
166	HOX32_ORYSI RecName	Transcription factor	FL720744
166	Myb-like DNA-binding domain, SHAQKYF class family protein	Transcription factor	FL748909
167	Glycosyl transferase-like protein	Biofuel	FL971577
167	EREB	Metabolism	FL830630
167	Ethylene-responsive protein	Signal transduction	FL738203
167	Calmodulin-binding protein	Signal transduction	FL791460
167	Brassinosteroid insensitive 1-associated receptor kinase 1	Signal transduction	GD013500
167	Cytochrome P450 CYP709H1	Stress response	FL689978
167	Auxin response factor 75	Transcription factor	FL821019
167	DNA-binding protein	Transcription factor	FL861193
167	DHHC zinc finger domain-containing protein	Transcription factor	FL991240
168	Beta-galactosidase 6	Metabolism	FL839194
168	DNA-binding protein	Transcription factor	FL956366
169	Protein pyrophosphatase	Metabolism	FL751239
169	Phospholipid synthesis1	Metabolism	FL967008
169	Aluminum-induced protein-like protein	Stress response	FL785573
171	Dlycerophosphoryl diester phosphodiesterase	Metabolism	DN149012
171	Erwinia-induced protein 2	Stress response	FL889979
171	ABA response element-binding factor	Transcription factor	FL978098
393	Lipase precursor	Metabolism	FL723146
397	D-TDP-glucose dehydratase	Metabolism	FL850129
397	Acetyl-coenzyme A carboxylase	Metabolism	FL957213
397	Peroxisomal membrane protein PMP22	Stress response	FL839966
398	Fiber protein Fb2	Biofuel	FL987047
399	Ragged seedling 2	Development	FE641510
399	Dehydrodolichyl diphosphate synthase 6	Metabolism	GD017081
408	Granule bound starch synthase	Metabolism	FL752682
408	DERL1_ORYSJ RecName	Signal transduction	DN152536
408	Brassinosteroid biosynthesis-like protein	Signal transduction	FE632180
408	P2C03_ORYSJ RecName	Signal transduction	FL714200
408	Negatively light-regulated protein	Stress response	FL963520

Table 2 continued

miRNA	Target protein	Target function	Accession ID
408	Pbs2-type MAP kinase kinase	Stress response	GD029660
414	Triacylglycerol lipase	Biofuel	FE616601
414	Glycosyltransferase	Biofuel	FL767655, FL714018
414	Lipid-binding protein	Biofuel	FL772969
414	Sucrose phosphate synthase1	Biofuel	FL958346
414	Hox3	Development	FE597482
414	Elongation factor 1-delta 1	Development	FE619283
414	Seed maturation protein	Development	FL841365
414	Cell Division Protein AAA ATPase family	Development	GD036313
414	Phosphoenolpyruvate carboxylase kinase 2	Metabolism	DN147290
414	Acyl-coenzyme A dehydrogenase	Metabolism	FE620514
414	Nicotianamine synthase 3	Metabolism	FE625499
414	Cellulose synthase BoCesA5/8	Metabolism	FL696641, FL708975
414	Chalcone synthase C2-Idf-III	Metabolism	FL786796
414	Serine carboxypeptidase 1	Metabolism	FL884382
414	Alcohol dehydrogenase 1	Metabolism	FL894351
414	Plastidic general dicarboxylate transporter	Metabolism	FL911571
414	Catalase isozyme B	Metabolism	FL963301
414	Purple acid phosphatase	Metabolism	FL974192
414	Phosphosulfolactate synthase-related protein	Metabolism	FL974681
414	1,3-beta-glucanosyltransferase gel4 precursor	Metabolism	GD044211
414	HMG1/2-like protein	Signal transduction	FL742196
414	Salicylic acid-induced fragment 1 protein	Signal transduction	FL825470
414	Serine/threonine-protein phosphatase 5	Signal transduction	GD033770
414	Senescence-associated protein DH	Stress response	FE612831
414	Dehydrin COR410	Stress response	FE619312
414	Polyamine oxidase	Stress response	FL728847
414	Na+/H+ antiporter	Stress response	FL747841
414	DnaJ domain containing protein	Stress response	FL752870
414	NADH-ubiquinone oxidoreductase B16.6 subunit	Stress response	FL776761
414	Rho GDP-dissociation inhibitor 1	Stress response	FL787827
414	HSP protein	Stress response	FL805746
414	Cytochrome *c* oxidase	Stress response	FL807065, DN148721
414	SLT1 protein	Stress response	FL904821
414	ALMT1	Stress response	FL941499
414	Phytochrome *C*	Stress response	FL964397
414	G10-like protein	Stress response	FL965333
414	Salt tolerance protein	Stress response	FL986000
414	Protein kinase domain containing protein-2	Stress response	FL996704
414	Two-component response regulator-like PRR73	Transcription factor	FL709463
414	SRF-type transcription factor family protein	Transcription factor	FL792552
414	Zinc ion-binding protein	Transcription factor	FL840723
414	GATA transcription factor 29	Transcription factor	FL993026
414	MYB59	Transcription factor	GD016222, GD043860
414	RNA-binding protein	Transcription factor	GD043601
444	Glucose phosphate isomerase	Biofuel	FL960000
444	Transcription factor AP2D9	Transcription factor	FL892564

Table 2 continued

miRNA	Target protein	Target function	Accession ID
477	Xylanase inhibitor protein 1	Biofuel	FE620894
477	DNA/pantothenate metabolism flavoprotein	Metabolism	FE618817
477	Ferredoxin-dependent glutamate synthase1 precursor	Metabolism	GD022147
477	Chaperone protein dnaJ 16	Stress response	FL991959
477	DNA repair protein RAD23	Stress response	GD002543
477	DNA-binding protein	Transcription factor	FL973241
528	Lipid-binding protein	Biofuel	GD010687
531	Polygalacturonase	Biofuel	FE609184
531	Xylose ABC transport system, permease protein	Biofuel	FE648225
531	Sucrose non-fermenting-related protein kinase 1b	Biofuel	FL697833
531	Glycosyltransferase	Biofuel	FL763392, FL918014
531	Carboxyl-terminal-processing protease	Biofuel	FL804548
531	Serine carboxypeptidase family protein, expressed	Biofuel	FL812381
531	UDPglucose:flavonoid-3-oxy glucosyl transferase	Biofuel	FL855580
531	Carboxylic ester hydrolase	Biofuel	FL858904
531	Cyclin-dependent protein kinase	Cell proliferation	FL850248
531	Seed maturation protein	Development	FL784034
531	Acetyl-coenzyme A synthetase	Metabolism	DN151201
531	Acetyl-CoA acetyltransferase, cytosolic 1	Metabolism	FE632616
531	Catalytic/hydrolase	Metabolism	FL717679, FL811185
531	Guanylate kinase	Metabolism	FL756305
531	Beta-glucanase	Metabolism	FL757700
531	Geranylgeranyl hydrogenase	Metabolism	FL796884
531	Sulfate transporter	Metabolism	FL804701
531	Acyl-coenzyme A oxidase 2	Metabolism	FL851635
531	Alpha-glucosidase-like	Metabolism	FL860818
531	Acetyltransferase	Metabolism	FL948629
531	Fertilization-independent type 1	Metabolism	GD000518
531	Ethylene-responsive factor-like transcription factor ERFL2a	Signal transduction	FE597988
531	Gibberellin receptor GID1L2	Signal transduction	FL709726, FL883283
531	Brassinosteroid LRR receptor kinase	Signal transduction	FL772053
531	GID1-like gibberellin receptor	Signal transduction	FL820906
531	Senescence-associated protein DH	Stress response	DN145067
531	Heat shock protein 17.9	Stress response	FE611944
531	Annexin A4	Stress response	FE652968
531	Peroxidase 52	Stress response	FE658088
531	Proline oxidase	Stress response	FL787729
531	Bifunctional polymyxin resistance arnA protein	Stress response	FL798502
531	Cytochrome P450	Stress response	FL820592, FL818887, FL871843, DN150939
531	Barley mlo defense gene homolog1	Stress response	FL843247
531	Heat shock factor	Stress response	FL859611
531	Drought-induced protein 1	Stress response	FL881778
531	1-aminocyclopropane-1-carboxylate oxidase	Stress response	FL914548
531	Hypersensitive-induced response protein	Stress response	GD026499
531	Zinc finger protein	Transcription factor	FE599186
531	ZF-HD homeobox protein	Transcription factor	FL726694
531	BZIP transcription factor family protein	Transcription factor	FL766685
531	WRKY53-superfamily of TFs having WRKY and zinc finger domains	Transcription factor	FL794693
531	DNA-binding protein	Transcription factor	FL811721

Table 2 continued

miRNA	Target protein	Target function	Accession ID
831	Rp3-like disease resistance protein	Stress response	GD041513
854	Glycoside hydrolases	Biofuel	FL826391
854	Cytosolic aldehyde dehydrogenase RF2D	Metabolism	FL764968
854	Heat shock factor protein HSF30	Stress response	GD001524
864	Beta-D-glucosidase beta subunit precursor	Metabolism	FE631651
1132	Cytochrome P450 CYP87A15	Stress response	FL791995
1132	Translation initiation factor 5A	Transcription factor	FE610694
1436	Peroxidase 2	Stress response	FE618792
1535	Lectin receptor-type protein kinase	Biofuel	GD014688
1535	Cytosolic orthophosphate dikinase	Metabolism	FL812508
1535	Acyl-coenzyme A oxidase 2	Metabolism	GD022207
1535	Cytochrome P450 CYP724B3	Stress response	FL690525
1535	Cytochrome P450 CYP727A4	Stress response	FL978332
1535	Metal tolerance protein A2	Stress response	GD020758
1535	Transcriptional activator FHA1	Transcription factor	FL911436
1848	Glucan endo-1,3-beta-glucosidase GVI precursor	Biofuel	FL760522
1848	AF331854_1 UDP-glucosyltransferase BX8	Biofuel	FL761128
1848	Clathrin assembly protein	Development	FL774398
1848	12-oxo-phytodienoic acid reductase	Metabolism	FL852652
1848	Translation initiation factor eIF-2B gamma subunit	Transcription factor	FL715950
1858	DnaJ-like protein	Stress response	FL767340
2102	UDP-glucose 4-epimerase	Biofuel	FE608058
2102	Pectinesterase inhibitor domain containing protein	Biofuel	FE613265
2102	Glucose-regulated protein 78	Biofuel	FE618495, FL693968, FE636568
2102	Sucrose transporter	Biofuel	FE621301
2102	Polygalacturonase	Biofuel	FE624059
2102	Switch/sucrose non-fermenting 3C	Biofuel	FL720890
2102	Sugar transport protein 14	Biofuel	FL763736
2102	Omega-3 fatty acid desaturase	Biofuel	FL767942
2102	Lipase	Biofuel	FL774689
2102	Lipid-binding protein	Biofuel	FL779887
2102	Glycoside transferase, six-hairpin, subgroup	Biofuel	FL803205
2102	UDP-glucose flavonoid-3-*O*-glucosyltransferase	Biofuel	FL838468
2102	Fatty acid desaturase	Biofuel	FL867087
2102	Fiber protein Fb34	Biofuel	FL886824
2102	Glycosyltransferase	Biofuel	FL912533, FL752079
2102	Fertilization-independent endosperm protein	Development	FL767507
2102	Elongation factor 1-alpha	Development	FL942153
2102	Craniofacial development protein 1	Development	GD038493
2102	Dihydrolipoyllysine-residue succinyltransferase	Metabolism	FE607196
2102	DTDP-4-dehydrorhamnose reductase	Metabolism	FE610159
2102	COLlagen family member (col-3)	Metabolism	FE632928
2102	10-deacetylbaccatin III 10-*O*-acetyltransferase	Metabolism	FE637471
2102	Carbonic anhydrase	Metabolism	FE648577
2102	Betaine-aldehyde dehydrogenase	Metabolism	FE658967
2102	Propionyl-CoA carboxylase beta chain	Metabolism	FL695036
2102	5-enolpyruvylshikimate-3-phosphate synthase	Metabolism	FL698488
2102	BGGP Beta-1-3-galactosyl-*O*-glycosyl-glycoprotein	Metabolism	FL744451
2102	Delta-1-pyrroline-5-carboxylate synthetase 2	Metabolism	FL751667
2102	AF334960_1 ADP-glucose pyrophosphorylase small subunit	Metabolism	FL753821

Table 2 continued

miRNA	Target protein	Target function	Accession ID
2102	4-coumarate coenzyme A ligase 1	Metabolism	FL772150
2102	Cytokinin-*O*-glucosyltransferase 1	Metabolism	FL776852
2102	Beta-fructofuranosidase, insoluble isoenzyme 2	Metabolism	FL809944
2102	Polyphosphoinositide-binding protein Ssh2p	Metabolism	FL816485
2102	Protein disulfide isomerase	Metabolism	FL851305
2102	Esterase	Metabolism	FL860775, FL962600
2102	Cytokinin-*N*-glucosyltransferase 1	Metabolism	FL903962
2102	12-oxo-phytodienoic acid reductase	Metabolism	FL908703
2102	Beta 1,3 glucanase	Metabolism	GD036373
2102	Zeamatin-like protein	Signal transduction	DN151247
2102	Gibberellic acid-insensitive	Signal transduction	FL705645
2102	Gibberellin receptor GID1L2	Signal transduction	FL795017
2102	Cytochrome P450	Stress response	DN150922, FL740970
2102	Erwinia induced protein 1	Stress response	FE599966
2102	Pathogenesis-related protein 5	Stress response	FE628122, FE653825
2102	ZmGR1b	Stress response	FE647928
2102	3-ketoacyl-CoA thiolase 2, peroxisomal precursor	Stress response	FE652020
2102	Super-oxide dismutase	Stress response	FE658372
2102	22.0 kDa class IV heat shock protein	Stress response	FL701547
2102	DnaJ subfamily B member 5	Stress response	FL772394
2102	Cytochrome *c* oxidase chain VI precursor	Stress response	FL828176
2102	Bundle sheath cell-specific protein 1	Stress response	FL849038
2102	Peroxidase	Stress response	FL861710
2102	1-aminocyclopropane-1-carboxylate oxidase	Stress response	FL910091
2102	Disease resistance response protein 206	Stress response	FL931856
2102	TMV response-related protein	Stress response	FL938800
2102	CCCH transcription factor	Transcription factor	FE599812
2102	Myb family transcription factor-related protein	Transcription factor	FE604269, FL796120, FE616043
2102	LysR family transcriptional regulator	Transcription factor	FE605460, FL757656
2102	BHLH transcription factor	Transcription factor	FL704149, FL979384
2102	Auxin-independent growth promoter	Transcription factor	FL757240, FL868592
2102	Zinc finger protein	Transcription factor	FL767042, FL944670
2102	Transcription initiation factor TFIID subunit 10	Transcription factor	FL781881
2102	Transcription factor MYC7E	Transcription factor	FL783822
2102	WRKY53-superfamily of TFs having WRKY and zinc finger domains	Transcription factor	FL794694
2102	Auxin transporter-like protein 1	Transcription factor	FL797094
2102	ABA-responsive protein	Transcription factor	FL797852
2102	DNA-binding protein	Transcription factor	FL811722
2102	GATA transcription factor 25	Transcription factor	FL816529, FL784703
2102	WRKY74-superfamily of TFs having WRKY and zinc finger domains	Transcription factor	FL865645
2102	MYB-like transcription factor DIVARICATA	Transcription factor	FL873993
2102	Dof zinc finger protein 5	Transcription factor	FL885859
2118	Lipid-binding protein	Biofuel	GD013359
2118	Chaperone protein dnaJ 10	Stress response	GD006930
2911	Barley mlo defense gene homolog1/8	Stress response	FL745324, FL705827
2911	Light-induced protein 1-like	Stress response	FL793852

Table 2 continued

miRNA	Target protein	Target function	Accession ID
2911	Salt tolerance-like protein	Stress response	FL892973
2911	Pathogenesis-related protein 4	Stress response	FL949551
2911	DNA-binding protein	Transcription factor	FL805678, FL816153
2911	Zinc finger protein	Transcription factor	FL861610, FL810133

Table 3 GO and KEGG pathway analyses show that miRNAs potentially target biofuel-related biological processes

miRNAs	Biological process	Accession ID for Targets	Total number of target
168, 528, 531, 854, 864, 1,848, 2,102, 2,911	Carbohydrate metabolic process	FL839194, FL744524, FL852328, FL914923, FL697833, FL757700, FL800137, FE609184, FL826391, FE631651, FL760522, FL813014, GD036373, FL809944, FE624059, FL753202	16
2,102	Carbon utilization	FE648577	1
164	Carbon utilization by fixation of carbon dioxide	FL962781	1
531, 1,535, 2,102	Fatty acid beta-oxidation	FL851635, GD022207, FE652020	3
414	Fatty acid beta-oxidation using acyl-CoA dehydrogenase	FE620514	1
531, 834, 1,535, 1,848, 2,102	Fatty acid biosynthetic process	FL697833, FL798882, FL975169, FL961992, DN147396, FL867087, FL767942, FL815939	8
397	Fatty acid elongation	FL957213	1
166, 444, 1,535	Gluconeogenesis	FL789117, FL960000, GD028727	3
2,102	Glucose metabolic process	FL751450	1
531, 2,102	Glucuronoxylan biosynthetic process	FL763392, FL912533	2
166, 444, 854, 1,535, 2,102	Glycolysis	FL789117, FL960000, FL731748, GD028727, FL840658, GD005140, FL700364	7
2,102	Glyoxysome organization	FE652020	1
414, 2,102, 2,911	Lignin biosynthetic process	DN145736, FL760261, FL740970, FL869282, FL763129, FL933730	6
2,102	Lipid biosynthetic process	FL886324	1
414, 531, 531, 1,852, 2,102	Lipid catabolic process	FE616601, FL763088, FL871106, FE618818, FL860775	5
393	Lipid metabolic process	FL723146	1
2,102	Lipid storage	FE597652	1
2,102	Response to sucrose stimulus	FL762749, FE618495	2
408, 2,102, 2,911	Starch biosynthetic process	FL752682, FL753821, DN148651	3
531, 1,535	Starch catabolic process	FL928795, FL822332	2
2,911	Sucrose biosynthetic process	DN148651	1
164, 414	Sucrose metabolic process	FL802595, FL958346	2
2,102	Sucrose transport	FE621301	1
837, 2,102	Xylem and phloem pattern formation	FL972718, FL694096, FL805269	3
1,132, 2,102	Xylem development	FE610694, FL695083	2

and miR-2911. The target genes identified are translated in response to such environmental stresses as heavy metal, high-salinity, pathogen infection, endogenous oxidation stress, and apoptosis. Once a plant suffers stress, several physiological regulation mechanisms are enacted in order to promote homeostasis, like hormone-related signal transduction (Jia et al. 2009). Much evidence has shown that miRNAs play a crucial role in hormone signal

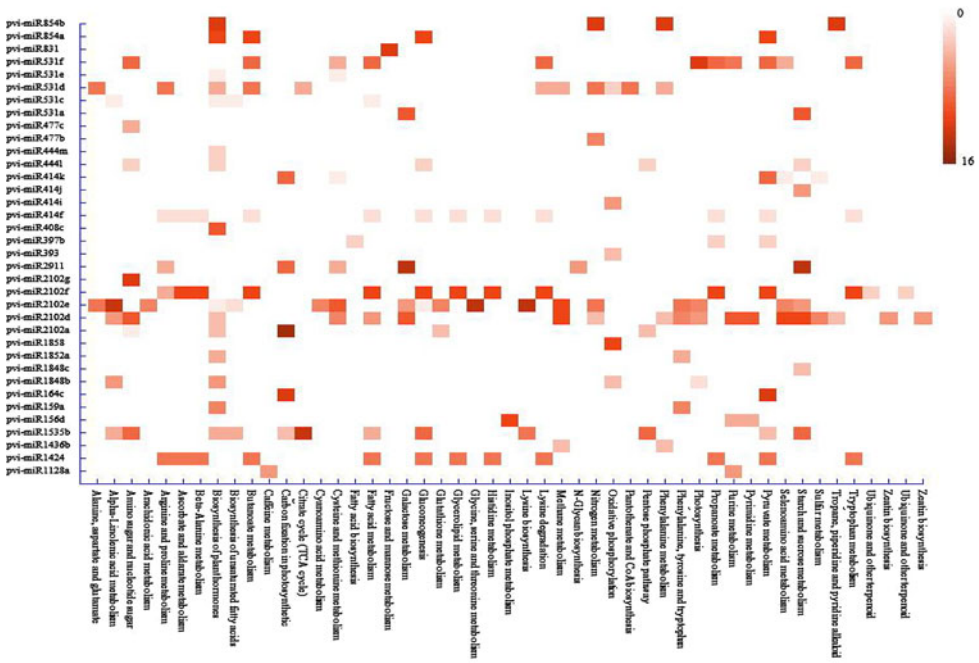


Fig. 5 KEGG pathways among partial predicted targets of the identified miRNAs in switchgrass. Each *column* corresponds to a pathway and each *row* refers to a group of genes targeted by a specific miRNA. *Red* was classified into 17 kinds. The color content class is calculated by a formula with one parameter of *E* value from doing Blastx with protein database ("Methods"). *Dark red* means closer relation between target and pathway as well as a closer relationship with miRNA and pathway

transduction. For instance, miR-398 has been shown to be involved in ABA-meditated salt resistance response in *Arabidopsis* (Jia et al. 2009). Our results show that six miRNAs might target 18 switchgrass genes that are involved in signal transduction of auxin, ethylene, brassinosteroid, salicylic acid, gibberellins, and zeatin.

Because switchgrass is one of the most important bioenergy crops for cellulosic ethanol biofuels, we also asked whether miRNAs target traits relating to biofuel production in addition to those necessary for plant development and the creation of biomass. In this study, we did find that 14 miRNAs potentially target 39 genes that are related to the biofuel regulation network. These miRNA targets genes include those for cellulose and sucrose biosynthesis as well as the genes for glycosyltransferase, xylanase, plygalacturonase, and a sucrose transporter. All of these genes play an important role in cellulose and carbohydrate biosynthesis.

Discussion

Identification of conserved miRNAs in switchgrass

In this study, we identified 121 potential switchgrass miRNAs, belonging to 44 families from a total of 436,535 currently available switchgrass ESTs. This suggests that ~0.0277% of switchgrass ESTs contain potential miRNAs. The ratio is much higher than the previously reported 0.0175% for soybean and 0.010% for other plant species (Zhang et al. 2006b, 2008). Several reasons may contribute to this high ratio for identifying conserved miRNAs from switchgrass EST sequences. One reason is that we employed WATER alignment tool for the first time instead of Blastn alignment tool to identify potential conserved miRNAs. In previous studies, researchers have employed Blastn searches in order to identify potential conserved miRNAs. This is because the Blastn search method is easy to use and people can directly search homologs using the NCBI website. Also, it saved time to use Blastn searches because many problems, such as word-size limitation and point mutants, were avoided when analyzing sequences. Blastn search, however, is more favorable to homolog searches with long sequences as search results with shorter sequences are more time consuming to analyze and the results tend to exclude insertion, deletion, and point mutations. All of the shortcomings of the Blastn search tool can be overcome by the WATER program. WATER uses Smith-Water algorithm tool, which ensures optimal local alignment by exploring all possible alignments and selecting the best (Smith and Waterman 1981). This means that deletion, insertion, and substitution of base alignment between known mature miRNAs and EST sequences can be considered, ultimately enhancing identification efficiency (Zhang et al. 2008). One issue with using the Smith-Water algorithm for alignment is that it is time-consuming with a low program-running speed. The WATER program, however, was able to identify a larger number of potential miRNAs; this could be due to the optimal local alignment feature, which enhances the efficiency of miRNA identification. Another reason for an increase in switchgrass miRNA identification is that we used all known plant miRNAs as a reference set; this would allow for the prediction of more conserved miRNAs using our method. New technologies, such as high through-put sequencing, are

gradually becoming more and more used for identifying novel and conserved miRNAs in plants and over the next few years, the total number of plant miRNAs in the miRBase is expected to rise. The WATER program provides an inexpensive and reliable method that can continue to identify potentially conserved miRNAs in other plant species.

Analysis of potential miRNA targets in switchgrass

Switchgrass is considered to be a main biofuel feedstock and its cellulosic material is normally digested with industrial enzymes in order to facilitate production into ethanol. (Uppugundla et al. 2009). The synthesis of cellulose is a complicated process that depends on carbon fixation, sugar metabolism and transit, and fat metabolism (Peng et al. 2002; Zhong et al. 2005). In this study, we found that 39 potential switchgrass targets were involved in the biological synthesis of cellulose. 14 miRNAs in switchgrass are predicted to have an effect on the regulation of cellulose biosynthesis (Table 2). For example, our results showed that miR-164 in switchgrass potentially targets five different genes including sucrose synthase 2 and C3 phosphoenolpyruvate carboxylase. This suggests that miR-164 might play a central role in sucrose metabolism and carbon fixation. Among the targets predicted to be involved in cellulose biosynthesis, some important proteins such as cellulose synthase, glycosyltransferase, polygalacturonase, sucrose transporter, and fiber protein Fb34 were detected to be targeted by miR-166, miR-414, miR-531, and miR-2102. Cellulose synthase was predicted to be a target of miR-397 in maize (Zhang et al. 2006a). However, we assumed that miR-166 might regulate cellulose synthase in switchgrass.

To better understand miRNA function, the identified target gene set was subjected to analysis by Gene Ontology (GO), which is viewed as a promising method for uncovering the miRNA-gene regulatory network on the basis of biological process and molecular function (Ashburner and Bergman 2005). In this study, miRNAs were detected in 527 biological processes in switchgrass. Our GO biological process enrichment results (Table 3) demonstrate that a total of 19 miRNA families might be involved in 25 biofuel-related metabolic pathways. Ultimately, this data can be used for improving biofuel production from switchgrass. The data also indicates that we can employ miRNA regulation mechanisms to interfere with cellulose biosynthesis genes ranging from carbon fixation to sucrose metabolism. Following GO analysis, we used KEGG to create a pathway enrichment of predicted miRNA target genes. 118 metabolism networks were found to be involved (data not shown). The results reveal that many miRNA family groups of switchgrass take part in the same pathways including amino acid metabolism, carbon metabolism, fatty acid metabolism, sucrose metabolism, sulfur metabolism, and others (Fig. 5). It suggests that several miRNA families co-participate in the same pathways and may play an important role in complicated metabolic pathways by interacting with each other or other components. In this pathway analysis, miR-414, miR-444, miR-531, miR-1535, miR-1848, miR-2102, and miR-2911 might regulate starch and sucrose metabolism in the same pathway network. Carbon fixation in photosynthesis might be collectively regulated by miR-164, miR-414, miR-1535, miR-2102, and miR-2911. These data suggest that biomass yields and photosynthetic efficiency can be improved by regulation of one or all of the miRNAs in a group. Furthermore, there were also several miRNA regulatory groups believed to be involved in fatty acid metabolism, gluconeogenesis, and nitrogen metabolism. Obviously, these results provide a novel idea of utilizing miRNA regulator groups for controlling biomass yields for biofuel production in switchgrass.

Acknowledgments This work was partially supported by a grant from the North Carolina Biotechnology Center (2009-BRG-1202) and a grant from DuPont Science & Engineering.

References

Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H (2005) Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 33:2697–2706

Ambros V (2004) The functions of animal microRNAs. Nature 431:350–355

Ashburner M, Bergman CM (2005) *Drosophila melanogaster*: a case study of a model genomic sequence and its consequences. Genome Res 15:1661–1667

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

Bender W (2008) MicroRNAs in the *Drosophila bithorax* complex. Genes Dev 22:14–19

Bouton JH (2007) Molecular breeding of switchgrass for use as a biofuel crop. Curr Opin Genet Dev 17:553–558

Chen B, Schnoor JL (2009) Role of suberin, suberan, and hemicellulose in phenanthrene sorption by root tissue fractions of switchgrass (*Panicum virgatum*) seedlings. Environ Sci Technol 43:4130–4136

Chiou TJ (2007) The role of microRNAs in sensing nutrient stress. Plant Cell Environ 30:323–332

Cler E, Papai G, Schultz P, Davidson I (2009) Recent advances in understanding the structure and function of general transcription factor TFIID. Cell Mol Life Sci 66:2123–2134

Fattash I, Voss B, Reski R, Hess WR, Frank W (2007) Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution. BMC Plant Biol 7:13

Floyd SK, Bowman JL (2004) Gene regulation: ancient microRNA target sequences in plants. Nature 428:485–486

Gleave AP, Ampomah-Dwamena C, Berthold S, Dejnoprat S, Karunairetnam S, Nain B, Wang Y-Y, Crowhurst RN, MacDiarmid RM (2008) Identification and characterization of primary

microRNAs from apple (*Malus domestica* cv. Royal Gala) expressed sequence tags. Tree Genet Genomes 4:343–358
Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154–D158
Hofacker IL (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31:3429–3431
Jagadeeswaran G, Saini A, Sunkar R (2009) Biotic and abiotic stress down-regulate miR398 expression in *Arabidopsis*. Planta 229:1009–1014
Jia X, Wang WX, Ren L, Chen QJ, Mendu V, Willcut B, Dinkins R, Tang X, Tang G (2009) Differential and dynamic regulation of miR398 in response to ABA and salt stress in *Populus tremula* and *Arabidopsis thaliana*. Plant Mol Biol 71:51–59
Jin H, Martin C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol 41:577–585
Jin WB, Li NN, Zhang B, Wu FL, Li WJ, Guo AG, Deng ZY (2008) Identification and verification of microRNA in wheat (*Triticum aestivum*). J Plant Res 121:351–355
Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14:787–799
Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57:19–53
Jung HJ, Kang H (2007) Expression and functional analyses of microRNA417 in *Arabidopsis thaliana* under stress conditions. Plant Physiol Biochem 45:805–811
Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30
Lu SF, Sun YH, Shi R, Clark C, Li LG, Chiang VL (2005) Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell 17:2186–2203
Pan XP, Zhang BH, SanFrancisco M, Cobb GP (2007) Characterizing viral microRNAs and its application on identifying new microRNAs in viruses. J Cell Physiol 211:10–18
Peng L, Kawagoe Y, Hogan P, Delmer D (2002) Sitosterol-beta-glucoside as primer for cellulose synthesis in plants. Science 295:147–150
Pilcher RLR, Moxon S, Pakseresht N, Moulton V, Manning K, Seymour G, Dalmay T (2007) Identification of novel small RNAs in tomato (*Solanum lycopersicum*). Planta 226:709–717
Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell 8:517–527
Seitz H, Royo H, Bortolin ML, Lin SP, Ferguson-Smith AC, Cavaille J (2004) A large imprinted microRNA gene cluster at the mouse Dlkl-Gtl2 domain. Genome Res 14:1741–1748
Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147:195–197
Song CN, Fang JG, Li XY, Liu H, Chao CT (2009) Identification and characterization of 27 conserved microRNAs in citrus. Planta 230:671–685
Stark A, Bushati N, Jan CH, Kheradpour P, Hodges E, Brennecke J, Bartel DP, Cohen SM, Kellis M (2008) A single Hox locus in *Drosophila* produces functional microRNAs from opposite DNA strands. Genes Develop 22:8–13
Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 16:2001–2019
Sunkar R, Girke T, Jain PK, Zhu JK (2005) Cloning and characterization of microRNAs from rice. Plant Cell 17:1397–1411
Sunkar R, Kapoor A, Zhu JK (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
Sunkar R, Chinnusamy V, Zhu JH, Zhu JK (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci 12:301–309
Talmor-Neiman M, Stav R, Frank W, Voss B, Arazi T (2006) Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss. Plant J 47:25–37
Tanzer A, Stadler PF (2004) Molecular evolution of a microRNA cluster. J Mol Biol 339:327–335
Tanzer A, Amemiya CT, Kim CB, Stadler PF (2005) Evolution of microRNAs located within *Hox* gene clusters. J Exp Zool B Mol Dev Evol 304B:75–85
Tyler DM, Okamura K, Chung WJ, Hagen JW, Berezikov E, Hannon GJ, Lai EC (2008) Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. Genes Develop 22:26–36
Unver T, Budak H (2009) Conserved microRNAs and their targets in model grass species *Brachypodium distachyon*. Planta 230:659–669
Uppugundla N, Engelberth A, Vandhana Ravindranath S, Clausen EC, Lay JO, Gidden J, Carrier DJ (2009) Switchgrass water extracts: extraction, separation and biological activity of rutin and quercitrin. J Agric Food Chem 57:7763–7770
Wang XJ, Reyes JL, Chua NH, Gaasterland T (2004) Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. Genome Biol 5:R65
Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, Yang ZM (2007) Computational identification of novel microRNAs and targets in *Brassica napus*. FEBS Lett 581:1464–1474
Yao YY, Guo GG, Ni ZF, Sunkar R, Du JK, Zhu JK, Sun QX (2007) Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). Genome Biol 8:R96
Yu J, Wang F, Yang GH, Wang FL, Ma YN, Du ZW, Zhang JW (2006) Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. Biochem Biophys Res Commun 349:59–68
Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15:336–360
Zhang BH, Pan XP, Anderson TA (2006a) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 580:3753–3762
Zhang BH, Pan XP, Cannon CH, Cobb GP, Anderson TA (2006b) Conservation and divergence of plant microRNA genes. Plant J 46:243–259
Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006c) Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 63:246–254
Zhang BH, Wang QL, Pan XP (2007a) MicroRNAs and their regulatory roles in animals and plants. J Cell Physiol 210:279–289
Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo TL, Cobb GP, Anderson TA (2007b) Identification of cotton microRNAs and their targets. Gene 397:26–37
Zhang BH, Pan XP, Stellwag EJ (2008) Identification of soybean microRNAs and their targets. Planta 229:161–182
Zhang WW, Luo YP, Gong X, Zeng WH, Li SG (2009) Computational identification of 48 potato microRNAs and their targets. Comput Biol Chem 33:84–93
Zhong R, Pena MJ, Zhou GK, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH 3rd, Darvill AG, York WS, Ye ZH (2005) *Arabidopsis fragile fiber8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell 17:3390–3408
Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res 9:133–148