



## Identification of miRNAs and their targets in wheat (*Triticum aestivum* L.) by EST analysis

J. Han, M.L. Kong, H. Xie, Q.P. Sun, Z.J. Nan, Q.Z. Zhang and J.B. Pan

Beijing Key Laboratory for Agricultural Application and New Technique,  
College of Plant Science and Technology, Beijing University of Agriculture,  
Beijing, China

Corresponding authors: J. Han / J.B. Pan  
E-mail: hanjun422@yahoo.com.cn / bacpjb@yahoo.com.cn

Genet. Mol. Res. 12 (3): 3793-3805 (2013)  
Received January 23, 2013  
Accepted August 26, 2013  
Published September 19, 2013  
DOI <http://dx.doi.org/10.4238/2013.September.19.11>

**ABSTRACT.** MicroRNAs (miRNAs) are a newly discovered class of noncoding small RNAs that regulate gene expression by directing target mRNA cleavage or translational inhibition. A large number of miRNAs have been identified in plants. Increasing evidence has shown that miRNAs play multiple roles in plant biological processes. So far, identification of miRNAs has been limited to a few model plant species, whose genomes have been sequenced. Wheat (*Triticum aestivum* L.) is one of the most important cereal crops worldwide. To date, only a few conserved miRNAs have been predicted in wheat. Here, we showed the conserved miRNAs identified in wheat by expressed sequence tag (EST) analysis. All previously known miRNAs from *Arabidopsis*, rice, and other plant species were used in a BLAST search against the wheat EST database to identify novel wheat miRNAs by a series of filtering criteria. By this strategy, we identified 62 conserved miRNAs, belonging to 30 miRNA families, 48 of which were newly discovered in wheat. These newly identified wheat miRNAs may regulate 287 potential targets, which are involved in development, signal transduction, metabolic pathways, disease resistance, ion transportation, and environmental stress response.

**Key words:** MicroRNAs; *Triticum aestivum* L.; Targets; EST analysis

## INTRODUCTION

MicroRNAs (miRNAs) are a class of small, single-stranded, noncoding RNAs that serve as posttranscriptional negative regulators of gene expression in plants and animals (Bartel, 2004). In plants, these 20 to 22 nt small RNAs are endogenously produced from their own genes by a chain of reactions, with the help of enzymes (Bartel, 2004). However, the genes encoding miRNAs are much longer than their mature sequences, and range from several tens to more than 1000 nt. MiRNA genes are first transcribed to the capped and polyadenylated primary miRNAs (pri-miRNAs) by the Pol II or Pol III enzymes (Lee et al., 2004; Borchert et al., 2006). In plants, pri-miRNAs fold into secondary stem-loop structures, which are processed in a two-step manner by a dicer-like enzyme (DCL1) (Bartel, 2004). The pri-miRNA is cleaved by DCL1 to an miRNA precursor (pre-miRNA) with a stem-loop structure, and the mature miRNA sequence is located at one arm of the hairpin structure. Subsequently, the pre-miRNA is further cleaved by DCL1 into an miRNA:miRNA\* duplex, which is then methylated by the RNA methylase HUA ENHANCER 1 (Yu et al., 2005). Finally, the single-stranded mature miRNA is incorporated into a ribonucleoprotein complex known as RNA-induced silencing complex, which identifies target mRNA with perfect (or nearly perfect) complementarity (Bartel, 2004). Studies in plants have revealed the key roles of miRNAs in diverse regulatory pathways, including growth, development, and defense response against various kinds of stress (Ruiz-Ferrer and Voinnet, 2009; Rodriguez et al., 2010; Sunkar et al., 2012).

Currently, 18,227 mature miRNAs have been discovered from various organisms and deposited in a publicly available miRNA database (miRbase, Release 18.0, November 2011). A total of 3580 plant miRNAs were identified and deposited in the current edition of the miRNA registry, most of which were identified from plants with sequenced genomes, including 581 from *Oryza sativa*, 291 from *Arabidopsis thaliana*, 362 from *Glycine max*, 635 from *Medicago truncatula*, 172 from *Zea mays*, 171 from *Sorghum bicolor*, 215 from *Populus trichocarpa*, 142 from *Brachypodium distachyon*, 201 from *Arabidopsis lyrata*, 163 from *Vitis vinifera*, and 37 from *Solanum lycopersicum*. Only a very small amount of miRNAs have been discovered and functionally identified in other plant species with unknown genome sequences, including 44 from *Triticum aestivum*, 21 from *Hordeum vulgare*, 18 from *Vigna unguiculata*, 16 from *Saccharum officinarum*, and 8 from *Phaseolus vulgaris*, among others.

The most challenging problem in understanding plant miRNAs is the identification of more novel miRNAs. Four major approaches have been used for miRNA identification in plants (Zhang et al., 2006a): genetic screening, direct cloning and sequencing of a small RNA library, computational prediction based on whole genome sequences, and expressed sequence tag (EST) analysis. Although the miRNA genes, *lin4* and *let7*, were first found and studied by genetic screening, this method was considered as an inefficient way for miRNA identification owing to its disadvantages of being time-consuming and complicated (Berezikov et al., 2006). Recently, an experimental approach, termed direct cloning, has been described to identify novel miRNAs (Berezikov et al., 2006; Zhang et al., 2006a). Although a large number of miRNAs have been identified by experimental cloning, this method is also expensive and time-consuming, and has difficulty in cloning miRNAs that express at a low level. Additionally, some miRNAs are difficult to find using cloning or deep sequencing, owing to their physical properties, including sequence composition, methylation, and posttranscriptional modifications (Berezikov et al., 2006). Most plant miRNAs are derived from the same gene families, which are strongly conserved and have high sequence identity even between distantly related species.

This conservation makes it possible to identify homologs of known miRNAs in other species (Zhang et al., 2006b). Many programs have been designed to search plant miRNAs using the whole genome sequences. However, the application of this strategy has been greatly limited by the unavailability of whole genome sequences in plants. Compared with the three methods mentioned above, EST analysis is a powerful tool for identifying conserved miRNAs and studying the evolution of conserved miRNAs among various plant species, especially for those whose whole genome sequences are not available (Zhang et al., 2005, 2006b).

Wheat is the most widely grown crop, occupying 17% of all the cultivated land and providing approximately 55% of carbohydrates for human consumption all over the world (Gill et al., 2004). Thus far, only 44 miRNAs have been identified in wheat and deposited in the miRNA database (miRbase), and a large number of wheat miRNAs remain unknown. In this study, all previously known plant miRNAs deposited in the miRNA database were employed to search the wheat EST database in order to predict the potential wheat miRNAs by EST analysis. A total of 62 potential miRNAs were detected, 48 of which were newly discovered in wheat. Based on these newly predicted wheat miRNAs sequences, we did a further BLAST search of the wheat mRNA database and found 278 potential miRNA target genes. Most of these wheat miRNAs were found to be involved in regulating plant development, signal transduction, ion transportation, response to biotic or abiotic stressors, and metabolism.

## MATERIAL AND METHODS

### Mature plant miRNA query sequences and wheat EST sequences

To identify potential wheat miRNAs, a total of 3580 previously known plant mature miRNA sequences from *Oryza sativa*, *Glycine max*, *Zea mays*, *Arabidopsis thaliana*, *Brachypodium distachyon*, *Sorghum bicolor*, *Populus trichocarpa*, *Medicago truncatula*, and other plant species were downloaded from the miRNA registry database (<http://www.mirbase.org/>). The repeated plant miRNA sequences from different species were removed to avoid redundant miRNAs, and the remaining unique miRNA sequences were employed as query sequences for BLAST search against the wheat EST database, which was available online from the National Center for Biotechnology Information (NCBI) GenBank nucleotide databases (<http://www.ncbi.nlm.nih.gov/>).

### Availability of software

BLASTN and BLASTX (<http://www.ncbi.nlm.nih.gov/BLAST/>) were used to search for wheat miRNA homologs and to analyze the protein-encoding sequences, respectively. The software CodonCode Aligner was employed to conduct ClustalW alignment, to discard the repeated ESTs of the same genes. The Zuker RNA folding algorithm software mfold 3.5, available online (<http://mfold.rna.albany.edu/?q=mfold>) (Zuker, 2003), was used to predict the secondary structures of RNAs. The web tool psRNATarget (<http://bioinfo3.noble.org/psRNA-Target/>) was applied to predict the targets of newly identified wheat miRNAs.

### Procedure and screening criteria for wheat miRNA identification

All mature miRNA sequences from other plant species were used to predict poten-

tial wheat miRNAs. We used the same methodology as described by Zhang et al. (2005) to identify new plant miRNAs based on their conservation by EST analysis. The procedure for the prediction of wheat potential miRNAs is illustrated in [Figure S1](#). To avoid repetitious work, the previously known plant miRNAs were screened to remove the repeated miRNA sequences, and the remaining unique miRNA sequences were used as queries for a BLASTN search against the wheat EST database for wheat miRNA homologs. All the parameters were set as follows: maximum target sequences for 1000, expect a threshold for 10; the remaining parameters were default. All wheat EST sequences with no more than four mismatches against the query sequences were saved. These EST sequences were employed to conduct ClustalW analysis to remove the repeated EST sequences of the same gene by the software CodonCode Aligner. The remaining sequences were used for a BLASTX search against the protein database at NCBI to remove the protein-encoding sequences.

The secondary structures of the remaining wheat EST sequences were generated using the software mfold 3.5 (Zuker, 2003), where the default parameters were used to predict secondary structures of the selected sequences. The following criteria (Ambros et al., 2003; Zhang et al., 2006c) were employed to screen the candidates of potential wheat miRNAs or pre-miRNAs: 1) compared with all known plant mature miRNA sequences, no more than four nucleotide mismatches were allowed in potential mature miRNA sequences; 2) the pre-miRNA sequence within an EST sequence could fold into a typical hairpin secondary structure that contained a mature miRNA sequence at one arm of the hairpin structure; 3) no more than six mismatches between the potential mature miRNA sequence and the opposite miRNA\* sequence were allowed; 4) no loops or breaks in miRNA sequences were allowed; 5) the secondary structures of the predicted pre-miRNA should have a higher minimal folding free energy index (MFEI) and negative minimal free energy (MFE) than other small RNAs; and 6) the content of A + U should be within 30 to 70%.

### **Prediction of targets of novel wheat miRNA**

Previous studies have demonstrated that most known plant miRNAs bind to the protein-encoding region of their mRNA targets with perfect or near-perfect sequence complementarity, and miRNAs negatively regulate target gene expression by direct cleavage of target mRNAs or repression of protein translation (Bartel, 2004; Chen, 2004). This characteristic provides an efficient strategy to predict plant miRNA targets by searching for miRNA homologs. Thus far, a large number of plant miRNA targets have been predicted using this approach. In this study, we used the same approach to predict wheat miRNA targets. Targets of newly identified wheat miRNAs were predicted by the web tool psRNATarget, using the *Triticum aestivum* (wheat) DFCI Gene index (TAGI) release 12 as the sequence library for target search, with 3 set as the maximum expectation, no more than four mismatches at complementary sites between miRNA sequences and potential mRNA targets allowed, and no gaps allowed at the complementary sites.

## **RESULTS AND DISCUSSION**

### **Identification of potential wheat miRNAs by EST analysis**

The evolutionary conservation of miRNA families among plant species has provided a

powerful approach to identify conserved miRNAs by EST and/or Genome Survey Sequences analyses in other plant species. Using the methods described above, after searching the NCBI wheat EST database and removing the protein-coding repeated sequences, only the EST sequences in accordance with the screening criteria were selected as miRNA candidates. In this study, 62 potential wheat miRNAs belonging to 30 miRNA families were identified. Given that many wheat miRNAs have been identified by high-throughput sequencing or computational approaches, the mature sequences of the 62 identified miRNAs were compared with published wheat miRNAs (Yao et al., 2007; Dryanova et al., 2008; Jin et al., 2008; Han et al., 2009; Wei et al., 2009; Yin and Shen, 2010). The results indicated that 14 miRNAs had an identical sequence with previously reported wheat miRNAs, including miR159, miR169, miR171, miR319, miR395a-c, miR398, miR408, miR444, miR1120, miR1128, and miR1135. The other 48 miRNAs were newly identified in wheat, including miR393, miR395d, miR397, miR437, miR530, miR774, miR1126, miR1127, miR1435, miR1436a-h, miR1439a-e, miR2593, miR5067a-b, miR5169, miR5175a-c, miR5180a-c, miR5181a-e, miR5183, miR5203a-c, and miR5205a-g (Table 1).

### Characteristics of identified wheat miRNAs

To distinguish miRNAs from other small RNAs, the characteristics of the identified wheat miRNAs were studied, including miRNA mature sequences, length of mature miRNAs, miRNA location, length of precursors, minimal folding free energies, and minimal folding free energy indexes (Table 1). The lengths of identified wheat miRNA precursors varied significantly, from 57 to 211 nt with an average of 108 nucleotides. However, most identified wheat miRNA precursors (82.3%) had 60 to 150 nucleotides, which is similar to the miRNAs in *Arabidopsis thaliana* and *Oryza sativa* (Griffiths-Jones et al., 2006). Despite the great difference in lengths, all the identified wheat miRNA precursors could be folded into the typical miRNA secondary structures (Figure S2). The lengths of these miRNA mature sequences range from 19 to 24 nucleotides, but most (75.8%) were confined to 21 nucleotides, which is quite similar to those mature miRNAs already identified in other model plants. The A and U contents of these predicted wheat miRNA ranges from 33.33 to 76.19%, with an average of 53.93%, which meets the results of previous studies (Zhang et al., 2006c). The negative MFEs of the identified wheat miRNA precursors varied broadly from -27.40 to -98.90 kcal/mol, with an average of -56.76 kcal/mol. Moreover, we also found that the precursors of these miRNAs had high MFEIs, ranging from 0.86 to 1.76 (Table 1), with an average of 1.19, which was significantly higher than those reported for tRNAs (0.64), rRNAs (0.59), and mRNAs (0.62 to 0.66), suggesting that these identified wheat miRNAs are more likely to be true miRNAs (Zhang et al., 2006c). In this study, we also found that 24 mature wheat miRNAs were located on the 3' end of their precursors, whereas the other 38 mature miRNAs were located on the 5' end of their precursors (Table 1 and Figure S2).

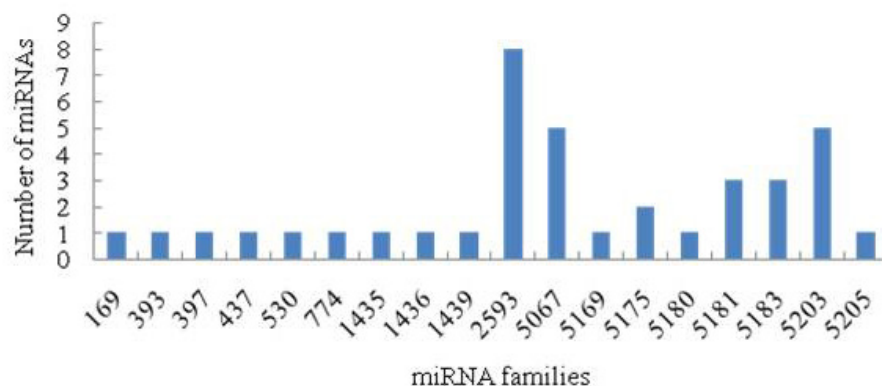
The 48 newly identified miRNAs belonged to 20 miRNA families in wheat. However, the distribution of these miRNAs was not even among the miRNA families (Figure 1), and some miRNA families had more members than the others. MiR393, miR395, miR397, miR437, miR530, miR774, miR1126, miR1127, miR1435, miR2593, miR5169, and miR5183 have only one member, whereas the other eight miRNA families (miR1436, miR1439, miR5067, miR5175, miR5180, miR5181, miR5203, and miR5205) have 2 to 8 members, respectively (Figure 1).



**Table 1.** Identified wheat miRNAs by EST analysis.

miRNAs	EST entries	miRNA mature sequences (5'→3')	L	LM (nt)	LP (nt)	(A+U)%	MFEs (kcal/mol)	MFEIs
miR159	CJ898977	uuuggauugaaggagcucug	3	21	175	52.38	81.00	<b>0.97</b>
miR169	BJ225371	acaggcaaaucuccuuggcua	3	22	112	52.38	60.20	<b>1.13</b>
miR171	BJ275219	cgugauauugcagcgccuaa	5	21	88	47.62	40.60	<b>0.88</b>
miR319	CA483944	aggagcuccuccuucaguccaa	5	21	200	42.86	98.90	<b>0.87</b>
miR393	CA484228	AGAcaaugcgauccuuuggau	5	21	67	54.55	27.40	<b>0.90</b>
miR395a	CK194045	cugaaguuuuugggggaacuc	3	21	66	47.62	32.20	<b>0.93</b>
miR395b	CK194045	augaaguuuuugggggaacuc	3	21	66	52.38	36.00	<b>1.15</b>
miR395c	CK194045	gugaaguuuuugggggaacuc	5	21	211	47.62	98.90	<b>0.86</b>
miR395c	CV763592	gugaaguuuuugggggaacuc	3	21	96	47.62	43.60	<b>0.87</b>
miR395d	CK194045	guuccuGcaagcacuucacga	5	22	211	45.45	98.90	<b>0.86</b>
miR397	GH985139	caucaacgcgucacucaauagg	3	21	104	47.62	77.50	<b>1.42</b>
miR398	CJ711035	uguguucucaggucgccccc	3	21	116	33.33	67.20	<b>0.87</b>
miR408	BE419354	cugcacugccuucuccuggc	3	21	148	33.33	85.60	<b>0.87</b>
miR437	CJ962326	aagucaaaUuuUGcuacuuiu	5	21	180	76.19	59.10	<b>1.38</b>
miR444	DR738640	ugcaguugcugucucaagcuu	3	21	105	52.38	53.20	<b>1.06</b>
miR530	CA655063	ugcauuugcaccugcacCu	5	20	147	45.00	69.20	<b>0.86</b>
miR774	CJ542241	cUuGcaAauuuucaucAcgaa	5	21	160	66.67	58.30	<b>1.09</b>
miR1120	CJ586598	uccgucuccaaauaagaauugu	5	24	90	62.50	46.80	<b>1.39</b>
miR1126	CJ905641	uccaAuauggacuaUGuacggag	5	23	118	56.52	48.80	<b>0.95</b>
miR1127	CK217714	uccCuccguucgaauiac	5	19	105	47.37	72.80	<b>1.32</b>
miR1128	CV775559	uacuacuccuccgucgaaa	5	21	133	47.62	84.80	<b>1.22</b>
miR1135	CJ893172	uccguucgaaauuacugucgag	5	24	107	50.00	65.80	<b>1.23</b>
miR1435	BJ210299	aCaagguuugacuUCaCa	3	20	189	60.00	66.20	<b>0.88</b>
miR1436a	CJ664068	acauuuugggagcggaggagu	3	21	85	47.62	66.50	<b>1.49</b>
miR1436b	CJ550576	acuccuccgucacuuuuAu	5	21	75	52.38	47.50	<b>1.33</b>
miR1436c	CA597465	acuccuccguUccauuuAu	5	21	89	57.14	47.30	<b>1.24</b>
miR1436d	CJ966458	acuccuccgucacuuuuAC	5	21	86	47.62	62.00	<b>1.38</b>
miR1436e	CJ795728	acuccuccguUccUAAaugu	5	21	145	52.38	79.90	<b>1.16</b>
miR1436f	CJ941609	CUauuuugggagcggaggagu	3	21	71	47.62	32.60	<b>0.88</b>
miR1436g	CF133018	aUauuGugggagcggaggagu	3	21	93	47.62	70.60	<b>1.45</b>
miR1436h	CJ590294	acuccuccgucuccaCaauu	5	21	83	42.86	52.50	<b>1.11</b>
miR1439a	CJ829732	UauacucCuccguuccaaaa	5	21	105	57.14	54.90	<b>1.22</b>
miR1439b	CK214157	aGuacucCuccguCccaaaa	5	21	131	47.62	89.00	<b>1.30</b>
miR1439c	DR731536	aaucucCuccguCccaaaa	5	21	130	52.38	74.80	<b>1.21</b>
miR1439d	CJ901160	aaucucCuccguuccaaaa	5	21	92	57.14	58.80	<b>1.49</b>
miR1439e	FR854792	UauacucCucUguuccaaaa	5	21	165	61.90	58.90	<b>0.94</b>
miR2593	CJ550747	GuucuaagguAcauucauuCC	3	21	109	61.90	72.50	<b>1.75</b>
miR5067a	BJ259276	Uuccauuuaguugucgcuca	5	21	120	61.90	55.30	<b>1.21</b>
miR5067b	CJ848204	auccauuuuAaugucgcuca	5	21	68	66.67	34.70	<b>1.53</b>
miR5169	CJ923687	uguucuaUaaacuuugucuaa	3	21	175	71.43	49.80	<b>1.00</b>
miR5175a	CJ829742	aagUauuuagggaacggaggga	3	21	171	57.14	97.20	<b>1.33</b>
miR5175b	CJ558409	agaaauuuUggGacggaggga	3	21	127	52.38	87.80	<b>1.45</b>
miR5175c	BE591362	uccuccgguuccGaaUuAcuu	5	21	107	52.38	69.90	<b>1.37</b>
miR5180a	CJ565048	uaagugucGcagAuuugaacu	5	21	63	61.90	28.80	<b>1.20</b>
miR5180b	DR737458	uaagugucCcGuuuuugaacu	5	21	68	57.14	28.30	<b>0.97</b>
miR5180c	CJ828581	AaagugucGcaguuuuugaacu	5	21	92	61.90	44.70	<b>1.28</b>
miR5181a	CJ624190	uccgauccauuuuagugucg	5	21	86	57.14	51.00	<b>1.38</b>
miR5181b	BQ236191	uccgauccaGaGuaagugucg	5	21	76	47.62	43.30	<b>1.09</b>
miR5181c	CJ517493	uccgauccauuuuagGgucg	5	21	76	52.38	49.00	<b>1.35</b>
miR5181d	CJ665408	cgacacuuuuuuggaucggG	3	21	83	52.38	44.70	<b>1.13</b>
miR5181e	GH725988	cgacacuuuuUgggaucAga	3	21	76	61.90	44.80	<b>1.55</b>
miR5183	GH731872	uauuuuggacaaauCuAagAca	3	21	81	71.40	40.70	<b>1.76</b>
miR5203a	BE470971	acuuuuuuggaUcggaggga	5	21	71	57.14	29.00	<b>0.95</b>
miR5203b	CJ775086	uccuccgUuccaAaauaGgu	5	21	81	52.38	44.30	<b>1.15</b>
miR5203c	CJ835156	uccuccgguccauaUuCaUu	5	21	68	52.38	37.60	<b>1.16</b>
miR5205a	CJ523432	acuaccuccgucuccAaaAuauaag	5	24	83	58.33	53.20	<b>1.54</b>
miR5205b	CV769842	acuccucUgucuccAaaAuauaag	5	24	70	58.33	48.20	<b>1.65</b>
miR5205c	CJ693364	acuccuccguUccuaaAuauaag	5	24	67	58.33	32.20	<b>1.15</b>
miR5205d	CA732203	cuuauaUuAUGgagcggaggagu	3	24	80	54.17	47.60	<b>1.30</b>
miR5205e	CJ571759	cUuacaUuAugggacggaggag	3	23	57	47.83	35.80	<b>1.20</b>
miR5205f	CJ883403	acuUccuccguUccuaaAuauaag	3	24	85	62.50	32.50	<b>1.02</b>
miR5205g	BJ295756	cuuauaUuuCgggacAgagggagu	3	24	83	54.17	48.10	<b>1.26</b>

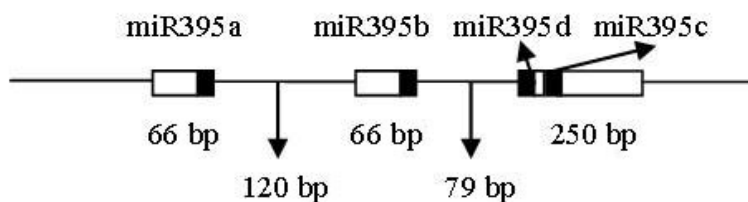
L = location of mature miRNAs; LM = length of mature miRNAs; LP = length of precursors; MFEs = minimal folding free energies; MFEIs = minimal folding free energy indexes. The miRNAs newly identified in wheat are shown in bold, the capital letters in sequences represent the mismatched nucleotides.



**Figure 1.** Family size of newly identified miRNAs in wheat.

### A cluster of four miR395 members in wheat EST sequence CK194045

It has been reported that members of the miR395 family form clusters of various sizes in different plant species, from less than 1 kb in *Oryza sativa* to about 70 kb in *Medicago truncatula* (Li and Mao, 2007). In this study, we found that four miR395 members identified on the same wheat EST sequence (CK194045) formed into a cluster, which spanned less than 1 kb on the wheat genome (Figure 2). Since the cluster is so compact, these miR395 members may be transcribed as a single polycistron.



**Figure 2.** Cluster of four wheat miR395 family members on the EST sequence CK194045. Box = precursor structure. miRNAs within the precursors were shown in black. The size of precursors and intervals were indicated.

### Targets of newly predicted wheat miRNAs

The mature sequences of these newly identified wheat miRNAs were employed to search for their potential regulatory targets, with the searching criteria set as described above in Material and Methods. In this study, we identified a total of 138 potential targets with known functions for the 18 identified miRNA families in wheat (Table 2); 140 additional targets with unknown functions were not listed. The 138 potential targets of wheat miRNAs are involved in different biological functions, including development, signal transduction, metabolic pathways, disease resistance, ion transportation, and responses to various environmental stressors.

**Table 2.** Predicted targets of the newly identified miRNAs in wheat.

miRNAs	Targeted protein	Functions of targets	Targeted genes
miR393	Auxin signaling F-box 1	Transcription factor	CK193017 (1.5)
miR395d	Stress enhanced protein	Stress response	CA622003 (3)
miR397	Phytase	Metabolism	TC443064 (3) TC458643 (3) CA621296 (3)
	Lipid-binding protein		TC422857 (3)
	Armadillo repeat-containing protein	Development	TC403981 (3)
	Lipoxygenase	Metabolism	TC419317 (3)
	Respiratory burst oxidase protein	Metabolism	TC432733 (3)
	Glutathione S-transferase	Metabolism	TC418995 (0)
miR437	Ribulose biphosphate carboxylase	Metabolism	TC437336 (0)
	Vacuolar protein sorting 29		CA718012 (2.5)
	Chitinase-like protein	Metabolism	TC395865 (2.5)
	F-box domain containing protein	Transcription factor	TC440827 (3)
	Subtilisin-chymotrypsin inhibitor		TC379065 (3)
	C2 and GRAM domain protein	Development	CK198414 (3)
miR530	Auxin-induced protein	Signal transduction	DN828913 (2.5) TC399043 (2.5)
	APE/TA/A2-like protein	Transcription factor	TC381895 (3)
miR774	Phosphodiesterase	Metabolism	TC391445 (2.5)
	Zinc finger CCH domain protein	Transcription factor	TC423619 (2.5)
	40S ribosomal protein S3a		CA486365 (3)
	Carbonyl reductase	Metabolism	CA719272 (3)
	Acetyltransferase	Metabolism	TC411773 (3) TC445078 (3) TC384245 (3)
	Serine/threonine protein kinase	Metabolism	TC441963 (3)
	Carboxyl terminal protease	Metabolism	TC412189 (3)
miR1126	NADPH-cytochrome P450 reductase	Metabolism	TC407053 (1.5)
	F-box domain containing protein	Metabolism	CD889246 (3)
miR1435	Copper ion binding protein	Transcription factor	TC425868 (0.5)
	Transmembrane protein 19-like		TC429063 (1)
	5-methylcytosine DNA glycosylase		TC372518 (1.5)
	Acyltransferase	Metabolism	TC381835 (2)
	rRNA methyltransferase	Metabolism	TC402966 (2.5)
	Metal-nicotinamine transporter	Transporter	TC412283 (2.5)
	ABC transporter F family member	Transporter	TC414541 (3)
	Ethylene-induced calmodulin-binding	Stress response	TC380020 (3)
	Terpene synthase	Metabolism	TC379112 (3)
miR1436	Integral membrane protein TerC	Metabolism	CD936383 (0.5)
	Tropinone reductase 2	Metabolism	TC451040 (1)
	Zinc finger CCH domain protein 8	Transcription factor	TC433753 (1)
	Brown planthopper-induced	Stress response	CV775565 (0) CV772007 (0)
	Glutathione S-transferase 1	Metabolism	CA693005 (0)
	Plasma membrane-associated protein		TC430344 (0.5)
	Utp14 protein		TC383983 (0)
	Heavy metal translocating ATPase		CJ961399 (0)

Continued on next page



Table 2. Continued.

miRNAs	Targeted protein	Functions of targets	Targeted genes
miR1439	Serine/threonine -protein kinase	Metabolism	TC433823 (0.5)
	Wall-associated receptor kinase	Disease resistance	CV766965 (0.5)
	C3HC4 zinc finger family protein	Transcription factor	TC412417 (0.5)
	Rust resistance kinase Lr10	Disease resistance	TC370176 (1)
	Glycosyltransferase	Metabolism	CV762246 (0.5)
	Dihydroflavonol-4-reductase	Metabolism	TC381627 (1)
	Ribonuclease H2 subunit C		TC376396 (1)
	bHLH transcription factor	Transcription factor	TC370353 (1)
	Epoxide hydrolase 2-like	Metabolism	CA613387 (0)
	C3HC4 zinc finger family protein	Transcription factor	TC405876 (0)
	Disease resistance protein	Disease resistance	TC453118 (0.5)
	Calmodulin TaCaM2-1	Signal transduction	CK196934 (0.5)
	F-box/FBD/LRR-repeat protein		TC402807 (0)
	Cell wall-associated hydrolase		TC393171 (0)
TetR family	Metabolism	TC412501 (0)	
Receptor protein kinase	Transcription factor	BI323295 (1)	
DNA-methyltransferase	Metabolism	CK154549 (1)	
Transport protein particle component	Transporter	TC411759 (0.5)	
Phosphoglucosyltransferase	Metabolism	BQ579749 (0)	
CHY zinc finger family protein	Transcription factor	AL821953 (1)	
Nitrate transporter	Transporter	TC432197 (2.5)	
Wall-associated receptor kinase	Disease resistance	CJ606356 (2)	
Glucosyltransferase-10	Metabolism	TC449276 (2)	
MAP3K protein kinase	Signal transduction	DR739617 (2)	
Disease resistance protein RPM1	Disease resistance	CJ937950 (2)	
Ribulose-5-phosphate-3-epimerase	Disease resistance	TC421333 (2.5)	
ABC transporter F family member	Metabolism	TC374327 (2.5)	
Transcription factor bHLH35-like	Transporter	TC412942 (0)	
Methylene-tetrahydrofolate reductase	Transcription factor	TC387478 (0)	
Nitrate transporter	Metabolism	CA641079 (0)	
Wd-repeat protein 48	Transporter	CK203241 (0) TC410668 (0)	
Nucleotidyltransferase	Signal transduction	TC459716 (0)	
Lecithine-cholesterol acyltransferase	Metabolism	TC377619 (0)	
RadA-like protein	Metabolism	TC370810 (0.5)	
REF/SRPP-like protein	Metabolism	TC401875 (0.5)	
ALWAYS EARLY 2-like	Metabolism	CK209464 (0.5)	
Phytol kinase 2	Metabolism	TC440105 (0.5)	
Serine/threonine phosphatase type 2c	Metabolism	TC426391 (0.5) TC382066 (0.5)	
F-box domain containing protein	Transcription factor	CJ915608 (0) TC391334 (0.5)	
Protein ROS1-like	Metabolism	TC398706 (0)	
BR11-KD interacting protein 109;	Transcription factor	TC380076 (0.5)	
Abscisic stress ripening-like protein	Signal transduction	TC444618 (1)	

Continued on next page



Plant miRNA targets are more likely to encode transcription factors, and 16 transcription factors for 10 miRNA families were found in our study, which may play important roles in plant growth and development. F-box proteins, acting as auxin receptors mediating Aux/IAA proteins proteasomal degradation and auxin-regulated transcription in the embryogenesis regulation by auxin, were predicted to be targets of miR393 and miR394 in *Arabidopsis thaliana* (Jones-Rhoades and Bartel, 2004). In this study, we found that wheat miR393, miR437, miR1126, miR5175, and miR5205 targeted F-box proteins. The APETALA2 (APE2)-like protein, involved in embryo, endosperm, and seed coat development in *Arabidopsis thaliana* (Ohto et al., 2009), was predicted as the target of wheat miR530, whereas miR172 have been well known to target APE2-like transcription factors in the plant kingdom, such as TARGET OF EAT1 (TOE1), TOE2, and TOE3 in *Arabidopsis thaliana* (Mlotshwa et al., 2006). MiR774, miR1436, and miR1439 targeted zinc finger transcription factors, which are a superfamily of proteins involved in numerous activities of plant growth and development, and are also known to regulate resistance mechanisms for various biotic and abiotic stresses (Feurtado et al., 2011). The basic/helix-loop-helix (bHLH) transcription factors, which control cell proliferation and cell lineage establishment (Toledo-Ortiz et al., 2003), were potential targets for wheat miR1436 and miR5175. MiR1439 also targeted TetR family proteins, and miR5183 targeted the WRKY transcription factor, which plays multiple roles in plant processes such as germination, senescence, and responses to abiotic stresses like drought and cold (Rushton et al., 2010). MiR5205 also targeted the CBS domain containing protein, which plays important roles in response to various stresses such as salinity, drought, cold, high temperature, UV, wounding, and genotoxic stress in *Arabidopsis thaliana* and *Oryza sativa* (Kushwaha et al., 2009).

In addition to targeting transcription factors, some identified wheat miRNAs (Table 2) also preferred to target genes involved in metabolism, such as phytase, lipoxygenase, glutathione *S*-transferase, phosphodiesterase, carbonyl reductase, acetyltransferase, 5-methylcytosine DNA glycosylase, terpene synthase, glycosyltransferase, dihydroflavonol-4-reductase, serine/threonine phosphatase, starch branching enzyme I, serine carboxypeptidase, and so on, which might play vital roles in various metabolic pathways. We also found that several identified wheat miRNAs (Table 2) had complementary sites with auxin-induced protein, calmodulin TaCam2-1, MAP3K protein kinase, Wd-repeat protein 48, abscisic stress ripening-like protein, and Wpk4 protein kinase, all of which may play important roles in signal transduction.

Additionally, there were other predicted targeted genes involved in biotic and abiotic stress resistance. For example, stress-enhanced protein was targeted by miR395d, ethylene-induced calmodulin-binding protein was predicted as the target of miR1435, which might play an important role in plant immunity and ethylene signaling (Nie et al., 2012), and miR1436 and miR5205 targeted brown planthopper-induced resistance protein and jasmonate-induced protein, respectively. Moreover, some genes relating to plant disease resistance were identified as targets of these novel wheat miRNAs, such as miR1436 and miR5067 that targeted wall-associated receptor kinase, which plays a significant role in both the pathogen response and cell expansion during plant development (Kohorn and Kohorn, 2012). Rust resistance kinase Lr10 was also targeted by miR1436, disease resistance protein was predicted as the target of miR1439, disease resistance protein RPM1 was targeted by miR5169, and NB-ARC protein was targeted by miR5205, which is a functional ATPase domain, its nucleotide-binding state of which is proposed to regulate activity of the R protein (van Ooijen et al., 2008).

Another important section of the predicted targets were various kinds of transporters, such as the miR1435-targeted metal-nicotianamine transporter, which is required for the

long-distance transport of iron and manganese in *Oryza sativa* (Ishimaru et al., 2010). ABC transporters were simultaneously predicted as the targets of miR5169 and miR1435, which were reported to transport glutathione conjugates (Klein et al., 2006). MiR5067 and miR5175 targeted the nitrate transporter, and, in addition, miR5175 also targeted the ammonium transporter, whereas the transport protein particle component was identified as the targets of miR1439 and miR5175. Furthermore, several predicted targets were involved in plant development, where armadillo repeat-containing protein, proven to be critical for anther and root development in *Arabidopsis thaliana* (Mei et al., 2012), was predicted to be targeted by miR397. miR437 targeted C2 and GRAM domain proteins, which may play an important role during the late stage of pollen development and germination (Jiang et al., 2005).

Although 3580 plant miRNAs have been identified and deposited in the miRNA database (miRbase), a majority of them were identified from model plant species such as *Arabidopsis thaliana*, *Oryza sativa*, and *Populus trichocarpa*. Owing to its huge and unsequenced genome, only a very small amount of miRNAs have been discovered and functionally identified in wheat. In this study, we identified 48 novel wheat miRNAs, belonging to 18 miRNA families. This work will facilitate research on the functions of miRNAs in wheat, as well as provide help in understanding the regulation of growth and development in wheat. We also believe that the more wheat EST sequences are made publically available, the more wheat miRNAs will be discovered.

## ACKNOWLEDGMENTS

Research supported by the Scientific Research Project of Beijing Educational Committee, China (#PXM2011\_014207\_000037), and National Natural Science Foundation of China (#31301312).

## [Supplementary material](#)

## REFERENCES

- Ambros V, Bartel B, Bartel DP, Burge CB, et al. (2003). A uniform system for microRNA annotation. *RNA* 9: 277-279.
- Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.
- Berezikov E, Cuppen E and Plasterk RH (2006). Approaches to microRNA discovery. *Nat. Genet.* 38 (Suppl): S2-S7.
- Borchert GM, Lanier W and Davidson BL (2006). RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* 13: 1097-1101.
- Chen X (2004). A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science* 303: 2022-2025.
- Dryanova A, Zakharov A and Gulick PJ (2008). Data mining for miRNAs and their targets in the Triticeae. *Genome* 51: 433-443.
- Feurtado JA, Huang D, Wicki-Stordeur L, Hemstock LE, et al. (2011). The *Arabidopsis* C2H2 zinc finger INDETERMINATE DOMAIN1/ENHYDROUS promotes the transition to germination by regulating light and hormonal signaling during seed maturation. *Plant Cell* 23: 1772-1794.
- Gill BS, Appels R, Botha-Oberholster AM, Buell CR, et al. (2004). A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium. *Genetics* 168: 1087-1096.
- Griffiths-Jones S, Grocock RJ, van DS, Bateman A, et al. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34: D140-D144.
- Han Y, Luan F, Zhu H, Shao Y, et al. (2009). Computational identification of microRNAs and their targets in wheat (*Triticum aestivum* L.). *Sci. China C Life Sci.* 52: 1091-1100.
- Ishimaru Y, Masuda H, Bashir K, Inoue H, et al. (2010). Rice metal-nicotianamine transporter, OsYSL2, is required for

- the long-distance transport of iron and manganese. *Plant J.* 62: 379-390.
- Jiang SY, Cai M and Ramachandran S (2005). The *Oryza sativa* no pollen (*Osnop*) gene plays a role in male gametophyte development and most likely encodes a C2-GRAM domain-containing protein. *Plant Mol. Biol.* 57: 835-853.
- Jin W, Li N, Zhang B, Wu F, et al. (2008). Identification and verification of microRNA in wheat (*Triticum aestivum*). *J. Plant Res.* 121: 351-355.
- Jones-Rhoades MW and Bartel DP (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol. Cell* 14: 787-799.
- Klein M, Burla B and Martinoia E (2006). The multidrug resistance-associated protein (MRP/ABCC) subfamily of ATP-binding cassette transporters in plants. *FEBS Lett.* 580: 1112-1122.
- Kohorn BD and Kohorn SL (2012). The cell wall-associated kinases, WAKs, as pectin receptors. *Front Plant Sci.* 3: 88.
- Kushwaha HR, Singh AK, Sopory SK, Singla-Pareek SL, et al. (2009). Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. *BMC Genomics* 10: 200.
- Lee Y, Kim M, Han J, Yeom KH, et al. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23: 4051-4060.
- Li A and Mao L (2007). Evolution of plant microRNA gene families. *Cell Res.* 17: 212-218.
- Mei Y, Gao HB, Yuan M and Xue HW (2012). The *Arabidopsis* ARCP protein, CS11, which is required for microtubule stability, is necessary for root and anther development. *Plant Cell* 24: 1066-1080.
- Mlotshwa S, Yang Z, Kim Y and Chen X (2006). Floral patterning defects induced by *Arabidopsis* APETALA2 and microRNA172 expression in *Nicotiana benthamiana*. *Plant Mol. Biol.* 61: 781-793.
- Nie H, Zhao C, Wu G, Wu Y, et al. (2012). SR1, a calmodulin-binding transcription factor, modulates plant defense and ethylene-induced senescence by directly regulating *NDRI* and *EIN3*. *Plant Physiol.* 158: 1847-1859.
- Ohto MA, Floyd SK, Fischer RL, Goldberg RB, et al. (2009). Effects of APETALA2 on embryo, endosperm, and seed coat development determine seed size in *Arabidopsis*. *Sex Plant Reprod.* 22: 277-289.
- Rodriguez RE, Mecchia MA, Debernardi JM, Schommer C, et al. (2010). Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Development* 137: 103-112.
- Ruiz-Ferrer V and Voinnet O (2009). Roles of plant small RNAs in biotic stress responses. *Annu. Rev. Plant Biol.* 60: 485-510.
- Rushton PJ, Somssich IE, Ringler P and Shen QJ (2010). WRKY transcription factors. *Trends Plant Sci.* 15: 247-258.
- Sunkar R, Li YF and Jagadeeswaran G (2012). Functions of microRNAs in plant stress responses. *Trends Plant Sci.* 17: 196-203.
- Toledo-Ortiz G, Huq E and Quail PH (2003). The *Arabidopsis* basic/helix-loop-helix transcription factor family. *Plant Cell* 15: 1749-1770.
- van Ooijen G, Mayr G, Kasiem MM, Albrecht M, et al. (2008). Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.* 59: 1383-1397.
- Wei B, Cai T, Zhang R, Li A, et al. (2009). Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (*Triticum aestivum* L.) and *Brachypodium distachyon* (L.) Beauv. *Funct. Integr. Genomics* 9: 499-511.
- Yao Y, Guo G, Ni Z, Sunkar R, et al. (2007). Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). *Genome Biol.* 8: R96.
- Yin ZJ and Shen FF (2010). Identification and characterization of conserved microRNAs and their target genes in wheat (*Triticum aestivum*). *Genet. Mol. Res.* 9: 1186-1196.
- Yu B, Yang Z, Li J, Minakhina S, et al. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* 307: 932-935.
- Zhang B, Pan X, Cobb GP and Anderson TA (2006a). Plant microRNA: a small regulatory molecule with big impact. *Dev. Biol.* 289: 3-16.
- Zhang B, Pan X, Cannon CH, Cobb GP, et al. (2006b). Conservation and divergence of plant microRNA genes. *Plant J.* 46: 243-259.
- Zhang BH, Pan XP, Cox SB, Cobb GP, et al. (2006c). Evidence that miRNAs are different from other RNAs. *Cell Mol. Life Sci.* 63: 246-254.
- Zhang BH, Pan XP, Wang QL, Cobb GP, et al. (2005). Identification and characterization of new plant microRNAs using EST analysis. *Cell Res.* 15: 336-360.
- Zuker M (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31: 3406-3415.