

# Identification and characterization of new plant microRNAs using EST analysis

Bao Hong ZHANG<sup>1,2,\*</sup>, Xiao Ping PAN<sup>1</sup>, Qing Lian WANG<sup>2</sup>, George P. COBB<sup>1</sup>, Todd A. ANDERSON<sup>1</sup>

<sup>1</sup>The Institute of Environmental and Human Health (TIEHH), and Department of Environmental Toxicology, Texas Tech University, Lubbock, TX 79409-1163, USA

<sup>2</sup>Henan Institute of Science and Technology, Xinxiang 453003, China

## ABSTRACT

Seventy-five previously known plant microRNAs (miRNAs) were classified into 14 families according to their gene sequence identity. A total of 18,694 plant expressed sequence tags (EST) were found in the GenBank EST databases by comparing all previously known *Arabidopsis* miRNAs to GenBank's plant EST databases with BLAST algorithms. After removing the EST sequences with high numbers (more than 2) of mismatched nucleotides, a total of 812 EST contigs were identified. After predicting and scoring the RNA secondary structure of the 812 EST sequences using mFold software, 338 new potential miRNAs were identified in 60 plant species. miRNAs are widespread. Some microRNAs may highly conserve in the plant kingdom, and they may have the same ancestor in very early evolution. There is no nucleotide substitution in most miRNAs among many plant species. Some of the new identified potential miRNAs may be induced and regulated by environmental biotic and abiotic stresses. Some may be preferentially expressed in specific tissues, and are regulated by developmental switching. These findings suggest that EST analysis is a good alternative strategy for identifying new miRNA candidates, their targets, and other genes. A large number of miRNAs exist in different plant species and play important roles in plant developmental switching and plant responses to environmental abiotic and biotic stresses as well as signal transduction. Environmental stresses and developmental switching may be the signals for synthesis and regulation of miRNAs in plants. A model for miRNA induction and expression, and gene regulation by miRNA is hypothesized.

**Keywords:** microRNA, plant, environmental stress, gene regulation, expressed sequence tags.

## INTRODUCTION

MicroRNAs (miRNAs) are a large family of about 21-22 nucleotide non-coding RNAs [1-3]. miRNAs play very important roles in posttranscriptional gene regulation by degradation of target mRNAs or by repression of targeted gene translation in animals, plants, and fungi [1-5]. According to the miRNA Registry Database (Release 3.1, April 2004), 899 miRNA genes have been discovered in various organisms [6]. Out of those, 71 miRNA genes have been predicted and identified in plants, of which 43 are from *Arabidopsis* and 28 from rice (*Oryza sativa*). After searching the DNA databases (<http://www.ncbi.nlm.nih.gov/>), we found 4 more miRNA genes discovered in previous reports [7, 8]; three (microRNA 166a, b and c)

from maize (*Zea mays*) [7], and one (microRNA 166a) from wood tobacco (*Nicotiana glauca*) [8].

Although miRNAs are one of the hottest research topics in biology, the discovery of miRNAs was dominated by chance and serendipity [9]. Currently, most miRNAs were identified by direct cloning of small RNAs [10, 11] or by computational strategies [12-17]. Computational strategies suggest that about 1% of predicted protein-coding genes were miRNAs in metazoan *Drosophila* [10] and in humans [18-20]. Thus, many miRNA genes still await discovery. Computational strategies provide a useful method to predict miRNA genes and their targets, and have been successfully applied in vertebrates [18-21], insects [22], *Arabidopsis* and rice [13, 14, 17, 23, 24]. However, many miRNA genes and their targets, especially in plants, still remain undiscovered. The traditional computational approach used was a bit inefficient and certainly not comprehensive, and it is very difficult to demonstrate the expression of these predicted miRNAs using this approach.

\*Correspondence: Bao Hong ZHANG  
Tel: +1-806-8854567; Fax: +1-806-8854577;  
E-mail: baohong.zhang@ttu.edu

Expressed sequence tags (ESTs) are partial cDNA sequences of expressed genes cloned into a plasmid [25, 26]. EST analysis has proven to be an economically feasible alternative for gene discovery in species lacking a draft genome sequence [26], and many important genes have been found through EST analysis [27-29]. As the number of ESTs dramatically increased and became available in databases, they have become a very powerful tool for discovery of new genes and for solving mysteries in biology [29]. As of June 2004, GenBank (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>) contained 22,165,266 entries in its EST database [30]. The greatest number of plant entries are for wheat (*Triticum aestivum*) (552,245), maize (*Zea mays*) (397,515), barley (*Hordeum vulgare* plus subsp. *Vulgare*) (356,856), soybean (*Glycine max*) (334,668), rice (*Oryza sativa*) (284,006), and *Arabidopsis thaliana* (thale cress) (258,825). This large number of ESTs provides a cost-effective and rapid alternative route toward the discovery and isolation of new genes.

Recently, Jones-Rhoades and Bartel (2004) briefly mentioned that they used *Arabidopsis* microRNA against an EST database to look for potential evidence of miRNAs in other plant species [14]. Bonnet *et al.* (2004) also used an EST database to confirm their 91 new potential microRNAs identified by traditional computational strategies [13]. However, they did not do detailed research in this field; they only focused on identifying new miRNAs using a traditional computational strategy. In this paper, based on the analysis of previously known miRNAs, we used EST analysis and DNA database analysis in detail to identify new potential miRNA genes and their targets. Because the ESTs and expressed genes come from the production of true gene expression, our analysis provides more evidence and confidence in the discovery of new potential miRNAs and their targeted genes. Using this strategy, we identified 338 new miRNA candidates in 60 plant species, and 348 genes associated with 348 ESTs targeted by these miRNA candidates. Based on these results, we formed the hypothesis that stress may play an important role in controlling miRNA expression.

## MATERIALS AND METHODS

### Known plant microRNA analysis

Previously known *Arabidopsis* (*At*) and rice (*Oryza sativa*) (*Os*) miRNA sequences were obtained from the miRNA Registry Database (Release 3.1, April 2004) [6]; maize (*Zm*) and tobacco (*Ns*) miRNA were found in publicly available genome databases by comparing all previously known *Arabidopsis* miRNA to GenBank's DNA databases with BLAST algorithms by Blastn 2.2.9 (May 1, 2004) [31].

Alignments of known plant miRNA were conducted using the DNAMAN software package (Lynnon Corporation, Quebec, Canada

J7V 9M5). Phylogenetic trees were reconstructed using the same DNAMAN software with standard parameters.

### Searching EST databases, carrying out EST analysis, and predicting the secondary structure of EST sequences

The subgroup of Viridiplantae of the publicly available EST databases was searched using Blastn 2.2.9 (May 1, 2004) [31] by comparing all ESTs to all previously known mature *Arabidopsis* miRNAs (20-24 base pairs) listed in the miRNA Registry Database (Release 3.1, April 2004) [6]. Blast parameter settings were as follows: expect 1000; the number of descriptions and alignments were 1000. All Blast results were saved. EST sequences which closely matched (n/n, n-1/n and n-2/n nucleotide matches, where n equals the previously known *Arabidopsis* miRNA length) the previously known *Arabidopsis* miRNAs were chosen, and their secondary structures were predicted and generated using the Zuker folding algorithm with mFold 3.1 [32, 33]. All mFold outputs including information about the number of structures, free energy ( $\Delta G$  kcal/mol), miRNA-like helicity, the number of arms per structure, size of helices within arms, and size and symmetry of internal loops within arms were analyzed. The predicted secondary fold-back structures were manually inspected by comparing to the characteristics of previously known miRNAs. The secondary structure of hairpin stem-loop was assigned a score according to the strategies used by Lai *et al.* (2003) [10]. Briefly, free energy ( $\Delta G$  kcal/mol) and miRNA-like helicity for each individual arm were evaluated and assigned a score. Helicity was calculated +1 for each paired helicity, -1 for each one-nucleotide symmetric loop and -2 for each two-nucleotide symmetric loop, and an overall score was calculated as  $(\text{helical score} + (\text{ABS}(\Delta G)/2)/2)$ . If there was more than one hairpin stem-loop structure for the ESTs containing the miRNAs, each was scored and the hairpin structure with the highest score was considered the miRNA hairpin stem-loop structure.

Predicted miRNAs and their related information were recorded. Closely related EST sequences were blasted against each other and analyzed. If the ESTs had a high similarity (E value less than  $e-100$ ), it indicated that these ESTs were created from the same mRNA, and were then considered as one miRNA.

## RESULTS

### Plant microRNAs are highly conserved

According to the miRNA Registry database (Release 3.1, April 2004) [6], 899 miRNAs have been discovered in various organisms. Out of the 899 genes, 71 miRNA genes have been predicted and identified in plants, of which 43 are from *Arabidopsis* and 28 from rice. After searching the publicly available DNA databases (<http://www.ncbi.nlm.nih.gov/>), we found 4 more miRNA genes discovered by Juarez *et al.* (2004) and McHale and Koning (2004), three from maize [7], and one from wood tobacco (*Nicotiana glauca*) [8].

The four plant species represent both dicots and monocots, two major plant categories. There are different nucleotide identity ratios and related genetic distances among the 75 previously known plant miRNAs (Fig. 1). Some miRNAs are very close to each other, such as miRNA

Identifying new plant microRNAs using EST analysis

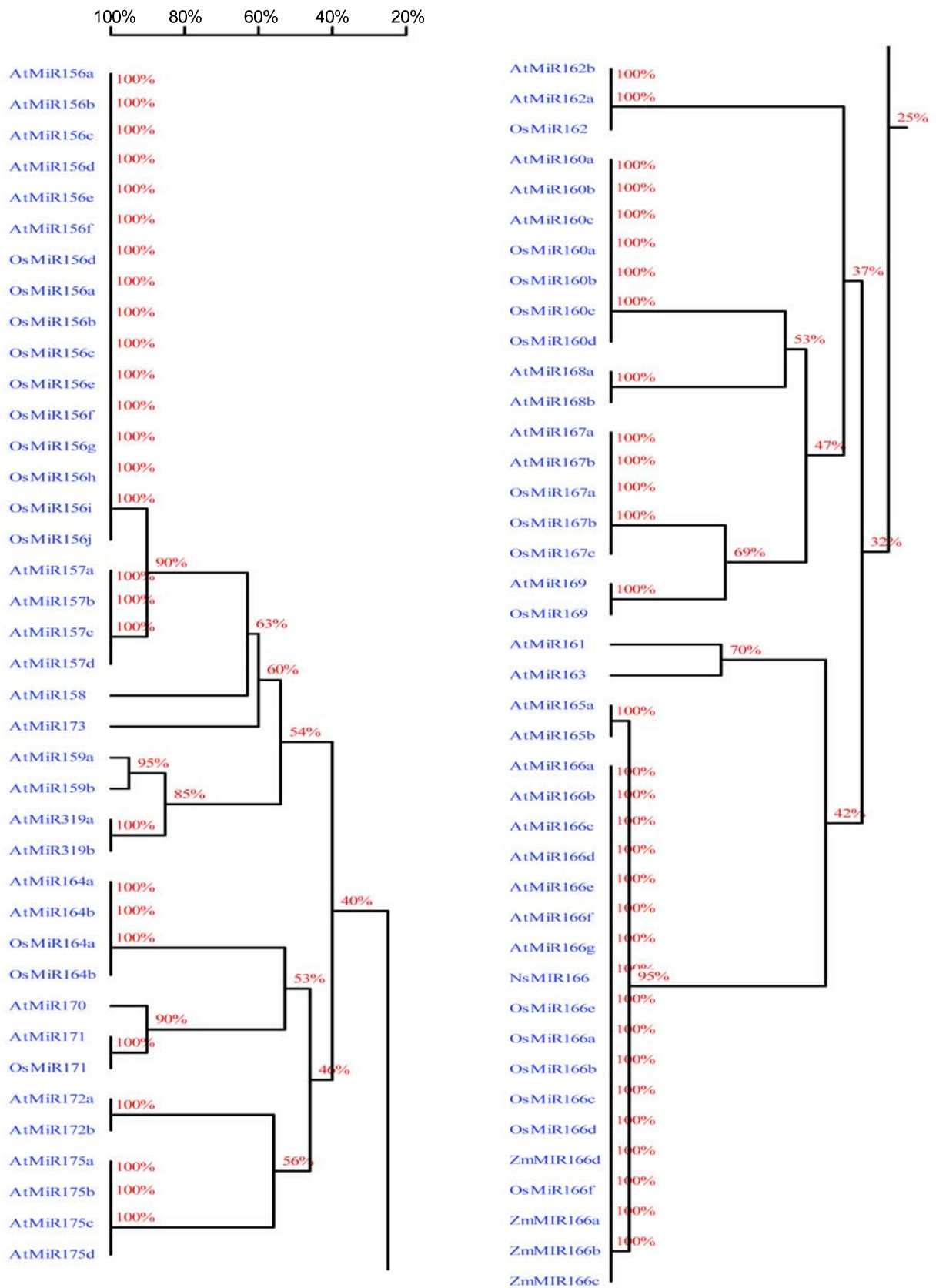
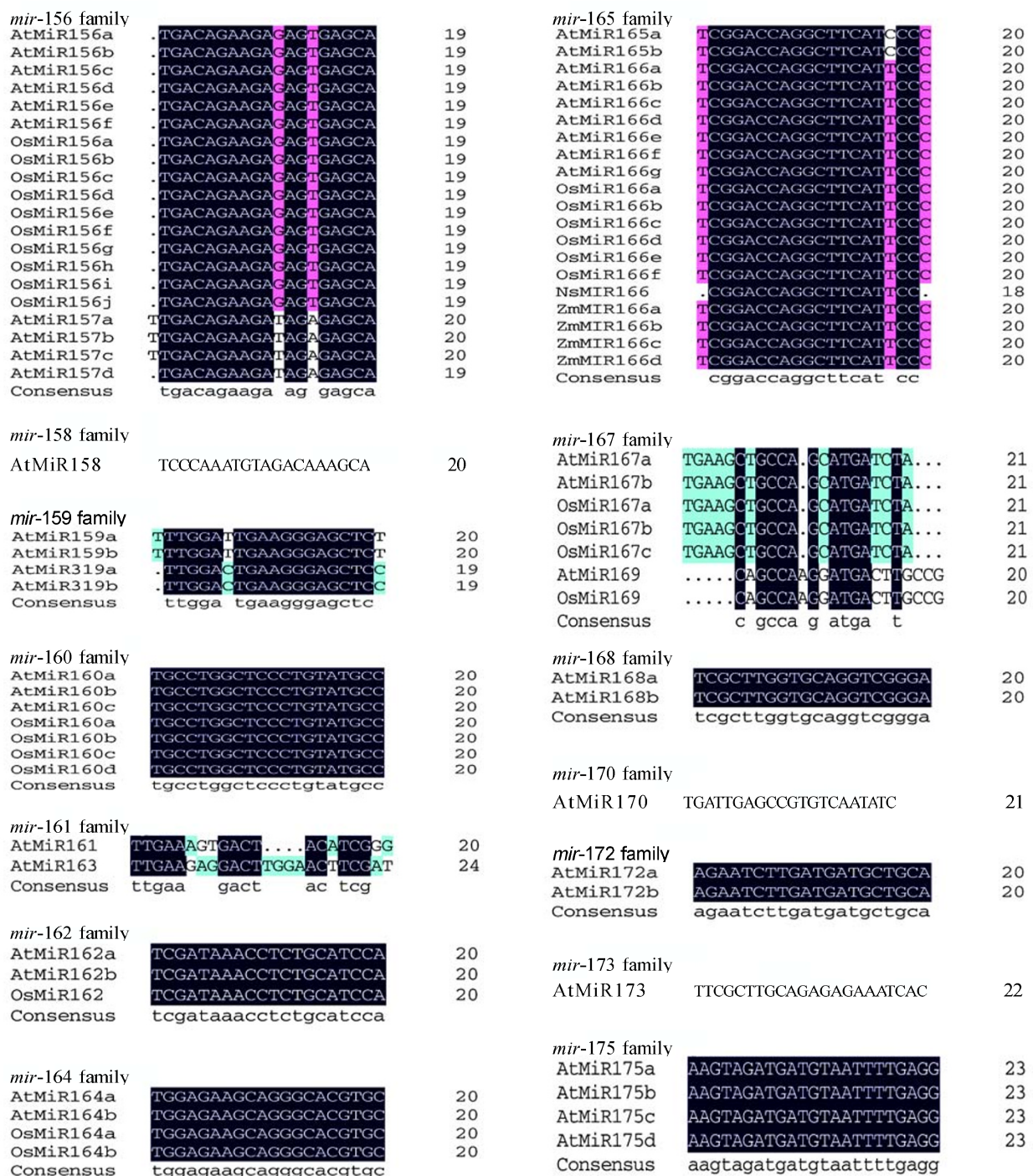


Fig. 1 Nucleotide sequence identity of 75 previously known microRNAs



**Fig. 2** Alignments of 75 previously known plant miRNAs that can be grouped into 14 miRNA families. *Arabidopsis* (*At*) and rice (*Oryza sativa*) (*Os*) miRNA are those identified in the miRNA Registry database (Release 3.1, April 2004); maize (*Zea mays*) (*Zm*) and tobacco (*Nicotiana glauca*) (*Ns*) miRNAs are those found in publicly available genome databases by Blastn against all *Arabidopsis* miRNA genes.

165 and miRNA 166, where there is only 1 nucleotide difference. Others have long genetic distances and low nucleotide identity (Fig. 1), which may be indicative of

their relationship and evolution. The 75 miRNAs were classified into 14 miRNA families based on their nucleotide identity (Fig. 2). In each miRNA family, each member

shares the same or similar gene sequences (Fig. 2). In the miRNAs already reported, only miRNA 166 was found in all four plant species. Fig. 1 and 2 show that these miRNAs from different species are highly conserved, and they share the same DNA sequence. These data suggest that miRNAs may have the same ancestor in very early evolution.

### Identifying new plant microRNAs and their targets

After comparing the gene sequences of 75 previously known miRNAs in different plant species, our results clearly showed that previously known miRNAs are highly conserved (Fig. 1 and 2). In previous reports, several labs have identified some previously known *Arabidopsis* miRNA homologs in rice [1, 2, 11, 34]. Both our results and other reports suggested that different plant species share similar miRNA sequences. Based on these results, in this study, we employed EST analysis to identify new miRNAs in other plant species. Our strategy included the following three steps. First, we searched the EST databases to find ESTs matched with the previously known *Arabidopsis* miRNAs. Then we predicted the secondary structures of the identified ESTs in the first step using RNA mFold software. Finally, we identified new miRNAs.

First, we searched the plant EST databases using Blastn 2.2.9 (May 1, 2004) [31] against all known *Arabidopsis* miRNA genes listed in the miRNA Registry database (Release 3.1, April 2004) [6]. A total of 18,694 Blast hits were found in the databases. To avoid a miscount of the total EST sequences, ESTs with n-2, n-3, n-4, and n-5 mismatched nucleotides with previously known miRNA were chosen and paired with the known miRNA sequences. After removing the EST sequences with high numbers (more than 2) of mismatched hits, we obtained 171, 440, and 201 ESTs with 0, 1, and 2 mismatched nucleotides, respectively (Tab. 1 and 2). This provided a total of 812 ESTs for predicting the RNA secondary structures using RNA mfold software.

EST hits differ in EST databases among previously known miRNAs. miRNAs 156 and 157 received the highest number of EST hits, each more than 200 with 0-2 mismatched nucleotides; the 24 and 18 ESTs containing completely conserved sequences were received for miRNAs 156 and 157, respectively. miRNAs 161 and 163 did not receive any EST hits with 0-2 mismatched nucleotides. Although miRNAs 171 and 166 only received 65 and 32 EST hits, respectively; they received the highest number of

**Tab. 1** Summary of plant EST blast hits to known plant miRNA sequences

miRNA	E value	Blast hits	Number of EST with 0-2 mismatched				ESTs for miRNA with 0-2 mismatched				miRNA
			nuceotides				nuceotides				
			0	1	2	Total	0	1	2	Total	
156	388	1009	24	185		209	22	76		98	82
157	518	1002	18	94	147	259	13	39	1	53	41
158	388	984	0	0	13	13	0	0	4	4	4
159	518	998	11	26	18	55	7	24	4	35	22
160	518	978	11	36		47	8	17		25	14
161	518	984	0	0	0	0	0	0	0	0	0
162	518	996	6	2	11	19	5	2	1	8	8
163	906	994	0	0	0	0	0	0	0	0	7
164	518	998	4	0	6	10	4	0	2	6	6
165	518	976	1	22	1	24	1	17	0	18	2
166	518	987	22	4	6	32	17	4	4	25	23
167	518	984	2	25	1	28	1	18	0	19	15
168	131	998	14	7	24	45	13	7	24	44	17
169	518	995	5	9	3	17	5	9	3	17	11
170	518	822	0	6	43	49	0	6	19	25	21
171	518	993	40	15	10	65	16	10	5	31	24
172	518	1006	6	5	70	81	6	3	39	48	32
173	647	991	0	0	3	3	0	0	1	1	1
319	388	999	7	4	23	34	6	3	7	16	14
Total		18694	171	440	201	812	124	235	105	464	338

**Tab. 2** ESTs with high base pair identity to previously known *Arabidopsis* miRNAs

miRNA	EST without mismatch	EST with 1 nt mismatch	EST with 2 nt mismatch
156	CN132665, CF059894, CF059465, CF059465, CF039276, CF037245, CF036982, CF036852, CF036471, CF036467, CF036182, CF035885, CD835236, CB677777, CB677501, CB674083, CB653304, CB647907, CB643501, CA294779, BI432985, BG300360, BE807821, AU091537	CB675101, D41985, COS29327, COS26369, COS25653, COS23543, COS20756, COS17866, CO414834, AL810223, BP660915, BP657257, CO237265, CO229128, CO169404, CO121346, CO105970, CO097437, CO097436, CO092899, CO081765, CO081202, CO051680, CN948866, CN581492, CN580870, CN578779, CN523533, CN519769, BP056010, BP052650, BP051612, BP051429, BP051130, BP046017, BP045933, BP044777, BP039258, AV777338, CK862588, CR285370, CK762549, CK754211, CK753210, CK751799, CK746364, CK665702, CK441500, CK441500, CK441499, CK441068, CK294165, CK293650, CK290419, CK286971, CK286411, CK286410, CK284079, CK284078, CK275093, CK272946, CK262978, CK254420, CK246692, CK196549, CK119834, CK105751, CK087760, CF836044, CF833064, CF667730, CF664003, CF663934, CF507028, CF444518, CF229561, CF035522, BP132966, CD800420, CD481596, CD476205, CD476172, CD473946, CD472931, CD454302, CD230908, CB970427, CB967554, CB819319, CB818994, CB632048, CB604551, CB093809, CB093704, CB091817, BJ576204, BJ571964, BJ568177, BJ567301, BJ566355, CA799494, CA516324, CA408772, CA408687, CA291432, CA285093, CA263546, CA254724, CA231663, CA228340, CA210204, CA186028, CA182412, CA152290, CA148934, CA091338, CA086406, CA082694, CA078294, CA072223, CA071566, CA032492, CA020690, AU294375, AU293485, BU833512, BU818933, BU667823, BQ968177, BQ874863, BQ506362, BQ696284, BQ586808, BJ478354, BQ111685, BJ322639, BJ317100, AV805456, AV799188, AV788983, AV784285, BM731329, BM731193, BM093898, BM092276, BM080504, BM064926, BF258419, BI470983, BI425943, BG947368, BG947367, BG651519, BG650778, BG650023, BG594131, BG507682, BG447166, BE060620, BG125997, BG040106, BF596468, BF587970, BF421158, BF273577, BF272491, AU162331, BF097000, BE806824, BE595949, BE436147, BE329522, AV558614, AV537797, AV524105, BE057288, AV417559, AW734915, AW679616, AI999403, AI995421, AI738310, AI726550, CK092224, CB674957	
157	CO082782, CO076888, BP039685, BP030242, BP029005, BP129308, BQ916415, BQ119413, BI971210, BI699573, BF187371, BE821022, BE611678, AL368695, AU089181, BE210632, AW756919, AW459710	CN825561, CB330750, CA902638, BJ578566, BJ578544, BJ573618, BJ569965, BJ568024, BJ563000, BJ558059, BJ556392, CA525602, BM536323, BE435668, BE433988, AW933950, COS29327, COS26369, COS20756, BP657257, CO105970, CO097437, CO097436, CO081765, CO081202, CN948866, CN580870, CN523533, BP046017, CK762856, CK751799, CK745155, CK294165, CK293650, CK290419, CK286971, CK286411, CK286410, CK284079, CK284078, CK272946, CK254420, CK246692, CK119834, CK105751, CK087760, BP132966, CD800420, CD472699, CB970427, CB819319, CB818994, CB604551, BJ576204, BJ571964, BJ567301, BJ566355, CA516324, CA408772, CA408687, CA254724, CA228340, CA152290, CA091338, CA082694, CA072223, CA032492, AU294375, AU293485, BU818933, BU667823, BQ968177, BQ874863, BQ111685, AV788983, AV784285, BM731329, BM731193, BF258419, BI425943, BG651519, BG650778, BG650023, BG447166, BF596468, BF587970, BF273577, BF272491, AV558614, AV537797, AV524105, AW696064, AI995421, AI738310, CF059894,	
158	-	-	BI935773, BI934413, AJ319944, BG139397, BG139114, BG136021, BU051914, BQ826645, AI496605, BM309320, BU051914, BQ121631, BQ826645
159	CN007551, CK750935, CF210361, CF208653,	CR289947, CK801265, CK801264, CF433018, CF308674, CF303003, CD725199, CD573488, CD224964, CD224004, CD213304, CD213290, CD204047,	

**Tab. 2** ESTs with high base pair identity to previously known *Arabidopsis* miRNAs (continued-1)

miRNA	EST without mismatch	EST with 1 nt mismatch	EST with 2 nt mismatch
	CB036045, AU251508, BU888582, BM893181, BI265403, AW691937, AJ388855,	CA731881, CA484819, CA261537, CA236922, CA203816, CA193052, CA148454, CA079434, AJ475696, BJ456281, BJ448559, BG050158, BE595318,	
160	CA801322, BU760860, BG882856, CA002787, CF330305, CF327976, CF327535, CF327534, BF293809, CD445121, CO484471,	CD480318, BM887596, AI900826, AI794655, BQ401391, CO106264, CO106199, CO102555, CO095743, BU986263, BJ472971, BJ471957, BJ471544, BJ469735, BJ469445, AV934494, AV934488, AV933114, AV932125, BF622299, AW038480, CB643398, CB643394, CA305796, CNS20271, CF119506, CA266669, CA247079, CA209000, CA067925, CF575341, BE494482, BQ112228, BE921523, CA615738, BQ788860	
161	-	-	-
162	BF003769, BG149136, BI785837, BQ845158, BU081804, CF516290,	CN894119, CA764295,	CF807240
163	-	-	-
164	CK936755, CK113235, CA704421, BU869001	-	CB009586, BF270853, CD929986, CD899685, CB651493, CO076098
165	CB086732	AI737566, AJ502524, AW685461, BI893541, BM092635, BM308126, BQ255147, BQ760548, BQ785760, CA067402, CA067403, CA819984, CD391899, CD405934, CF626917, CK369135, CN126049, CO531986, CO521288, AJ771029, AJ770960, AJ770980	-
166	AI737566, AJ502524, AW685461, BI893541, BM092635, M308126, BQ255147, BQ760548, BQ785760, CA067402, CA067403, CA819984, CD391899, CD405934, CF626917, CK369135, CN126049, O531986, CO521288, AJ771029, AJ770960, AJ770980	BG316028, BJ553847, BI972515, AW694053	
167	CK209908, BI095235	AU239920, AW318298, AW472145, BG509097, BG510970, BJ556759, BJ563257, BM892909, BQ629383, BU548271, BU810017, BU886852, BX928793, CA284319, CA284394, CA287325, CA916400, CD411229, CF309326, CF322848, CF571260, CF630597, CK209889, CK768366, CN472374	CD003107
168	H77158, BE659175, BE661028, AW424354, CB076866, BU871538, BU893331, BU889659, BU888204, BU886509, BU811772, BU809183, CF604588, CN901684	BF097936, BE461111, BE461110, CK254942, CK249947, CK246126, CK243936	CA003609, CA003589, CB667935, CB662543, CB655478, CB629782, CA756101, CA173899, CA156827, CA123434, CF576659, CN145907, CN133168, CD213423, CD210663, CD209437, CD209338, CD207772, CD205823, CD205059, CA826777, BU099153, BG842539, BG842528
169	BF595231, AW201497, CF280503, BQ906497, BU862460,	CA953278, AW596073, BM524615, BM520766, BF598910, BX899561, BX899553, CD960569, BU865420	CN143986, CA148737, CA148650
170	-	CF329292, CD910903, CD441617, CD440809, CA756196, CA009309	AI657239, AI770772, AV419737, AW459328, BE040695, BE365851, BE641457, BG583589, BG598365, BG653310, BG838164, BJ275219, BM892213, BM892515, BQ255243, BQ461013, BX838271, BX839725, BX839739, CA067169, CA069452, CA120921, CA937914, CB087210,

**Tab. 2** ESTs with high base pair identity to previously known *Arabidopsis* miRNAs (continued-2)

miRNA	EST without mismatch	EST with 1 nt mismatch	EST with 2 nt mismatch
			CB307092, CB655995, CB669725, CD439494, CD949754, CD963893, CF069782, CF435529, CF437845, CF437865, CF438830, CF439580, CF440325, CF442989, CF447956, CF603884, CF603900, CK210886, CN524603
171	CN524603, BX839739, BX839725, BX838271, CK210886, CF603900, CF603884, CF447956, CF442989, CF440325, CF439580, CF438830, CF437865, CF437845, CF435529, CF069782, CD439494, CB669725, CB655995, CB307092, CA069452, CA067169, BU635724, BQ461013, BQ255243, BM892515, BG838164, BG653310, BG598365, BG583589, BE641457, BE365851, BE040695, AV419737, AW459328, AI770772, AI657239, CO532905, CO531250, CO129257	AU304757, CA191585, CA139546, CF329292, CD964933, CD910903, CD441617, CD440809, CA756196, CA120921, CA106627, CA104875, CA009309, AI736818, CB087210	BJ275219, CK210439, CA937914, BM892213, CB087210, BP063562, CK297209, CK291560, CK291266, CK289500
172	BU084569, BI320499, CO080218, CO069513, AI484737, BQ114970	CO099417, CO120684, CO090457, BM348229, CF506714	CA922314, CD476336, CR280156, CA755817, CA282304, CA268785, CA218620, CA216531, CA205923, CA139645, BM407544, CK212277, CD895806, CA652598, CA612469, CA609410, CA607806, CF203088, CF511859, CB974375, CB922024, CD527210, CD058684, CB604044, CB603742, BF292190, CK206331, CD920346, CD909382, CD894004, CD885904, CD884790, CA746952, CA639688, BJ281140, BM136378, BJ212004, CN007930, CK205300, CD918238, CD891701, CD890663, CD876238, BE404354, BE423818, BE499159, CD875800, CA687020, CA600558, BJ318547, BJ311202, BJ286231, BJ286200, BJ275253, BJ274634, BJ274033, BJ219453, BJ273777, BJ266612, BJ265295, BJ263258, BJ253503, BJ251456, BJ251357, BJ250738, BJ250684, BJ249723, BJ249680, BJ246862, BJ224436
173	-	-	BE521498, CD534192, CD534180
319	BQ630517, BQ630503, BE475558, CK760683, CK091603, CA483944, CK751513	BJ191979, CF983804, CA184549, CA289676	BM520615, AV527132, BQ609349, AJ613606, BG237979, BQ453148, CA227452, CD925456, CA599744, AJ480356, AJ480355, CA006699, BQ466150, CA149108, CF759104, BQ841783, CB920070, CB919200, BG606141, CN494329, BM338067, AW424565, CA229394





**I. miRNA 166**

*Glycine max*

```

5'-GGGAAUG GGC UGG CGAGGCUUU GGUUCU U
|||||
3'-CCCUUAC UGG ACC UCUUCGGAA UCAAGA G
U- G AG -UA- ACG
    
```

*Glycine max*

```

5'-GGGAAUG GUCUGG CGA GAC CU UCUUGAUC GUGUAG C
|||||
3'-CCCUUAC CGGACC GCU UUG GA GGAACUGG CGUAUC U
U- AG G U UACAUAA UGUU A
    
```

*Hordeum vulgare subsp. vulgare*

```

5'-GGGAAUGA CC GGUC GAGAC GCAU GCG G
|||||
3'-CCCUUACU GG CGAGG CUUUG CGUG UGC C
UC A - G G
    
```

*Ipomoea nil*

```

5'-GGGGAGG GGC UGG CGAUC UUCUUAUCAUC UCUACAUAA AUA ADAUGGAAGAU AAUAU U
|||||
3'-CCCUUAC UGG ACC GCUAC AGUGAG UGUUAG -UA- GUUGAUUUA UAUUAUCUUC UUAUA UAA
U- G AG GU A- ACCC U U
    
```

*Medicago truncatula*

```

5'-GGGAAUG GGC UGG CGAGGCUUU AGUUU ACA AAG C
|||||
3'-CCCUUAC UCG ACC GCUUCGGAA UCAAG UGUU UUC A
U- G AG UA UA AA- UG
    
```

*Saccharum officinarum*

```

5'-CGG CAG UU CAU CCCCA AAGCAAG UCCU UG
|||||
3'-GCC GUC AA GUA GGGU UUCGUUC CCG GU U
G AA UA C CAU CG (18 nt side loop) UUUU UC
    
```

*Sorghum bicolor*

```

5'-G GGAUG GUCUGGUC AGGUCUG GAUUUAAGGA UGA UGC A
|||||
3'-C CCUUAU CGGACCAGG UCUAGAGU CUAAGUUCU AUU AUG G
C U U C AU- U UUA C
    
```

*Zea mays*

```

5'-GGGAAUG GUCUGG CGAGGU GCA AAC AGAUC U
|||||
3'-CCCUUAC CGGACC GUCUA UUG UCGUG A
U- AG GC UC GA- CU
    
```

**J. miRNA 167**

*Arabidopsis thaliana*

```

5'-UGAAGCUC CAG AUGAUCUGUA AUC CA C- U AAACUUCUU A
|||||
3'-ACUUGCAG GUC UACUGGAUCU UGAU AU UUGUGUGUGUGA GUA
C U U GA- AG AU (218 nt side loop) U
    
```

*Glycine max*

```

5'-UGA GCU GC CAGCAUGAUCUAG GGUUAGU GC G
|||||
3'-ACU CGA CG GUCGUACUGGAUC CCAAUUG UG A
C U U ACUC - AU
    
```

*Phaseolus coccineus*

```

5'-UGAAGCUGCCA CAUGAUU CU CCUAU U
|||||
3'-ACUUGACGGU GUACUAGA GA GGAGU G
C UA GA UG
    
```

*Populus tremula x Populus tremuloides*

```

5'-UGAAGCUGC AG AUGAUCUGGU AAUC UGUUA U
|||||
3'-ACUUGCAGC UC UACUGGACUA CUUG ACAAU A
A - C- (122 nt side stem loop) UAUUAUUAU
    
```

*Saccharum sp*

```

5'-UGAAGCUGC AGCAUGAUCUGAUG GU GUUAU U
|||||
3'-ACUUGAUG UCGUACUGGACUAC UA UAUUA G
(70 nt side stem loop) G A
    
```

**K. miRNA 168**

*Glycine max*

```

5'-UGG UUGGUCAGG CGGGA CCGGUUU GCGG GG UG N
|||||
3'-AGU AACUACGUUC GCCCU GGUUAG CCGG CC GC G
C C G GCUG GA
    
```

*Hedyotis terminalis*

```

5'-UCG UUGGUCAGG CG GAUCU UUCAC CACG A
|||||
3'-AGU AACUACGUUC GU CUUGA AAGUG GUGU A
C C A AUUU- (15 nt side stem loop) AUAAA
    
```

**L. miRNA 169**

*Oryza sativa*

```

5'-AGCCAAGGA ACUUGCCGAGU GAU AUUAUG GCU AG U
|||||
3'-UCGGUUCU UGAACGGUUA CUA UAGGUAC CGG UC A
A GU A CCUACC - G
    
```

*Populus tremula*

```

5'-AGCCAAGGA ACUUGCCGAGU UCG CUUUU CUU UGCUU C
|||||
3'-UCGGUUCU UGGACGGUUC AGU GAGAA GAA ACGGA A
A GU CUA A - CUAU UUAU
    
```

*Sorghum bicolor*

```

5'-UAGCCAAGGA GACU GCCUGU CC UUUU CGAAG CA A
|||||
3'-AUCGGUUCU CUGA CGAAUC GG AGAAU GUUUC GU A
- U U GA GAGU A
    
```

**M. miRNA 170**

*Allium cepa*

```

3'-ACUAA UCGAC CAGUCA AAGU CCG CGA UUA AG U
|||||
5'-UGAUU AGCCG GCCAAU CUCA GGC GCC AAU UC G
G AU CG - UAU AUU UA
    
```

*Arabidopsis thaliana*

```

5'-GAUUAUGG UGGUUA UCAGAU UUAU UG C
|||||
3'-CUAUAAACG GCCGAGU AGUCUA GAUG AC A
C U (21st side loop) C AC
    
```

*Glycine max*

```

5'-GAUUAUGG CCGUUAU AGAAAGCA GCUCAAA U
|||||
3'-CUAUAAAC GCCGAGUUA UUUUUUGU UGGGUUUU G
GU C CC U
    
```

*Hedyotis centranthoides*

```

5'-GAU UUGG CCGUUAU GG GGUUG GUUC C
|||||
3'-CUA AAC GCCGAGUUA CU UUAAC CCAG A
U GU G UU- C AAU
    
```

*Hordeum vulgare subsp. vulgare*

```

5'-GAUGUUGGC CG CUA UCAGA CC GCCCG AGG G
|||||
3'-CUAUAAACG GC GAGU AGUCU GG GCGGC UCU C
U C U U CG G AC
    
```

*Oryza sativa*

```

5'-GAUGUUGGC CCGCUCA UCAGA GGCAU GUGA GC AAGCAUG A
|||||
3'-CUAUAAACG GCCGAGU AGUCU UUGUG CACU CG UUCGUC U
U U - UU- AUGC G
    
```

**N. miRNA 171**

*Allium cepa*

```

5'-UGAUU AGCUG GUCAGU UUCA GGC GCU UC A
|||||
3'-ACUAA UGCG CCGUUA GAGU CCG CGG AG C
C G UA GC - AAUUUAUA AU
    
```

*Arabidopsis thaliana*

```

5'-GAUUAUGG UGGUUA UCAGAU UUAU UG C
|||||
3'-CUAUAAACG GCCGAGU AGUCUA GAUG AC A
C U (23rd side loop) C AC
    
```

*Hedyotis centranthoides*

```

5'-GAU UUGG CCGUUAU GG GGUUG GUUC C
|||||
3'-CUA AAC GCCGAGUUA CU UUAAC CCAG A
U GU G UU- C AAU
    
```

*Hordeum vulgare subsp. vulgare*

```

5'-GAUGUUGGC CG CUA UCAGA CC GCCCG AGG G
|||||
3'-CUAUAAACG GC GAGU AGUCU GG GCGGC UCU C
U C U U CG G AC
    
```

*Triticum aestivum*

```

5'-GAUGUUGGC CG CUA UCAGA CCAC GCC GGC G
|||||
3'-CUAUAAACG GC GAGU AGUCU GGUG UGG CGG C
U C U C G (21 nt side loop) C
    
```

**O. miRNA 172**

*Arabidopsis thaliana*

```

5'-GAUUC C AU UG GCAGCA G CAACA UUGGU GU GAAGC GA UCA CCGGUGU U
|||||
3'-UUUGG U UA GC UGUGU C GUUGU AAACA CG UUCU CU AGU GCGGCGU AC
G U GU AAG C- - U- E C- AAG GACAAA U
    
```

*Citrus sinensis*

```

5'-U GCAGGUC UCAAGAUUC CAU CAG AAGGCA AGCAU AAU U
|||||
3'-A CGREGUAG AGUUCUAAAG GUG GUU UUCGU UCGU UUA C
A A ACU UUUUG - U CUU
    
```

*Hordeum vulgare subsp. vulgare*

```

3'-ACGUC AGUA CCUAAG AGGU UGU CCG U
|||||
5'-UACAG UCAU GAAUCC UCCA ACA GGC G
A - AA CU (18th side loop) A
    
```

*Lycopersicon esculentum*

```

5'-AUG GCAU AUCAAGAUUC CGUGA GUU CAAA U GUUAU UAAU U
|||||
3'-UAC CGUA UAGUUCUAAAG GCACU CAA GUUU G CGGUA AUAA U
GU G A AUC GU - AG
    
```

*Solanum tuberosum*

```

3'-ACGUCGUAUAGUUCUAAAG GUAC UUUUA ACCCG AUA U
|||||
5'-UGCAGCAUCAAGAUUC CAUA AAAAA UGGAC UAU A
G A - GAA
    
```

*Vitis vinifera*

```

5'-GAAUC C AU UG GCAGCA G CAACA UUGGU GU GAAGC GA UCA CCGGUGU U
|||||
3'-UUUGG U UA GC UGUGU C GUUGU AAACA CG UUCU CU AGU GCGGCGU AC
G U GU AAG C- - U- E C- AAG GACAAA U
    
```



**Tab. 3** New miRNAs identified from plants

miRNA family	miRNA gene	Species	miRNA sequence*	Length (nt)	Precursor	EST accession	EST length	Strand	Foliar	miRNA		precursor	
										Start	End	Start	End
156	AtMiR 156m	<i>Arabidopsis thaliana</i>	ugacagaagagagagc	20	A1999403	498	+	3'	461	480	213	480	
	AtMiR 156o		245	AV805456	409	+	3'	377	396	152	396		
	AtMiR 156p		47	AV799188	407	+	5'	326	345	326	372		
	AtMiR 156q		59	AV558614	572	+	5'	500	519	500	558		
	AtMiR 156r		326	AV537797	607	+	5'	215	234	215	540		
	AtMiR 156s	218	AI995421	552	+	5'	223	242	223	440			
	BnMiR 156	<i>Brassica napus</i>	ugacagaagagagugagc	20	81	CD835236	694	+	5'	102	121	102	182
	CaMiR 156a		<i>Capsicum annuum</i>	ugacagaagagagagc	20	152	CA516324	508	-	3'	417	436	285
	CaMiR 156b	65		BM064926	474	-	5'	36	55	36	100		
	CsMiR 156a	<i>Citrus sinensis</i>	ugacagaagagagagc	20	56	CK665702	605	-	3'	202	221	166	221
	CsMiR 156b		515	CF836044	785	+	3'	558	577	63	577		
	CrMiR 156	<i>Cycas rumphii</i>	ugacagaagagagagc	20	67	CB091817	671	-	5'	221	240	221	287
	EcMiR 156	<i>Eschscholzia californica</i>	ugacagaagagagagc	20	150	CD481596	691	-	3'	348	367	218	367
	GmMiR 156a	<i>Glycine max</i>	ugacagaagagagugagc	20	83	BE807821	482	+	5'	163	182	163	245
	GmMiR 156b		ugacagaagagagagc	20	147	BM731329	421	-	5'	4	23	4	150
GmMiR 156c	185		BM731193	504	-	5'	67	86	67	251			
GmMiR 156d	76		BI470983	459	-	3'	109	128	53	128			
GmMiR 156e	55		BI425943	559	-	3'	177	196	142	196			
GmMiR 156f	267		BG651519	486	-	3'	373	392	126	392			
GmMiR 156g	209		BG507682	452	-	5'	77	96	77	285			
GmMiR 156h	147		BE806824	416	-	5'	117	136	117	263			
GmMiR 156i	117		AW734915	426	-	3'	174	193	77	193			
HaMiR 156	<i>Helianthus annuus</i>		ugacagaagagagagc	20	362	BQ968177	575	-	5'	354	373	12	373
HsMiR 156	<i>Hordeum vulgare subsp. spontaneum</i>	ugacagaagagagagc	20	178	BJ478354	418	-	3'	325	344	167	344	
HvMiR 156a	<i>Hordeum vulgare subsp. vulgare</i>	ugacagaagagagugagc	20	394	BG300360	624	+	5'	231	250	231	624	
HvMiR 156b		86	CA032492	647	-	5'	523	542	523	608			
HvMiR 156c	178	CA020690	515	-	5'	439	458	281	458				
HvMiR 156d	203	BF258419	902	-	3'	184	203	1	203				
LsMiR 156	<i>Lactuca sativa</i>	ugacagaagagagagc	20	355	BQ874863	654	-	5'	533	552	198	552	
LjMiR 156	<i>Lotus corniculatus var. japonicus</i>	ugacagaagagagagc	20	47	BP051612	510	+	5'	362	381	362	408	
LeMiR 156a	<i>Lycopersicon esculentum</i>	ugacagaagagagagc	20	116	BF097000	482	-	3'	285	304	189	304	
LeMiR 156b		47	BG125997	589	-	5'	25	44	25	71			
LeMiR 156c		263	BE436147	597	-	5'	21	40	21	283			
MdMiR 156a	<i>Malus x domestica</i>	ugacagaagagagagc	20	109	CN580870	505	-	5'	138	157	138	246	
MdMiR 156b	79	CN578779	538	-	3'	166	185	107	185				
NbMiR 156a	<i>Nicotiana benthamiana</i>	ugacagaagagagagc	20	205	CK294165	812	+	5'	516	535	516	720	
NbMiR 156b		169	CK290419	610	+	3'	508	527	359	527			
OsMiR 156m	<i>Oryza sativa</i>	ugacagaagagagugagc	20	86	CB677777	734	+	5'	262	281	262	347	
OsMiR 156n		CB677501	776	+	5'	260	279	260	345				
OsMiR 156o		CB653304	742	+	5'	260	279	260	345				
OsMiR 156p		CB647907	722	+	5'	269	288	269	354				
OsMiR 156q		CB643501	841	+	5'	259	278	259	344				
OsMiR 156r		AU091537	714	+	5'	10	29	10	95				
OsMiR 156s		D41985	412	+	5'	286	305	286	371				
OsMiR 156t		97	CB674083	451	+	5'	322	341	322	418			
OsMiR 156u		ugacagaagagagugagc	20	179	CB675101	270	+	3'	212	231	53	231	
OsMiR 156v		ugacagaagagagugagc	20	133	CB674957	497	+	3'	333	352	220	352	
OsMiR 156w	ugacagaagagagugagc	20	243	CB632048	296	+	3'	237	256	14	256		
OsMiR 156x	ugacagaagagagagc	20	83	AU162331	469	-	3'	398	417	335	417		
PgMiR 156	<i>Picea glauca</i>	ugacagaagagagagc	20	115	CK441500	682	+	3'	627	646	532	646	
PiMiR 156a	<i>Pinus taeda</i>	ugacagaagagagagc	20	160	CF667730	787	-	3'	146	165	6	165	
PiMiR 156b	221	BQ696284	578	-	3'	394	413	193	413				
PpeMiR 156	<i>Populus balsamifera x Populus deltoides</i>	ugacagaagagagagc	20	146	CN519769	797	-	5'	336	355	336	481	
PiMiR 156	<i>Populus tremula</i>	ugacagaagagagagc	20	373	BU818933	577	-	5'	42	61	42	414	
PptMiR 156a	<i>Populus tremula x Populus tremuloides</i>	ugacagaagagagugagc	20	46	CK092224	827	+	5'	414	433	414	459	
PptMiR 156b	ugacagaagagagagc	20	356	CK087760	815	+	3'	455	474	119	474		
PptMiR 156c	85	BU833512	555	-	3'	284	303	219	303				
PaMiR 156	<i>Prunus armeniaca</i>	ugacagaagagagagc	20	114	CB819319	504	+	5'	137	156	137	250	
ScMiR 156a	<i>Saccharum officinarum</i>	ugacagaagagagugagc	20	101	CA294779	738	+	5'	198	217	198	298	
ScMiR 156b		ugacagaagagagagc	20	115	CA086406	691	+	3'	657	676	462	576	
ScMiR 156c		CA210204	735	-	3'	212	231	113	231				
ScMiR 156d		68	CA228340	637	-	3'	391	410	343	410			
ScMiR 156e		327	CA182412	563	-	5'	18	37	18	344			
ScMiR 156f		70	CA148934	689	+	3'	512	531	463	532			
ScMiR 156g		46	CA091338	780	-	3'	163	182	137	182			
ScMiR 156h		403	CA186028	582	-	3'	443	462	60	462			
ScMiR 156i		343	CA078294	1007	-	5'	28	47	28	370			
ScMiR 156j		82	CA072223	704	-	5'	556	575	556	637			

**Tab. 3** New miRNAs identified from plants (continued-1)

miRNA gene family	miRNA Species	miRNA sequence*	Length (nt)	Precurs orhairpi n length (nt)	EST accession number	EST length (nt)	Stra nd	Folb ack arm	miRNA		Precursor	
									Start	End	Start	End
SIMiR 156	<i>Sesamum indicum</i>	ugacagaagagagagagcac	20	81	BU667823	289	+	5'	43	62	43	123
StMiR 156a	<i>Solanum tuberosum</i>	ugacagaagagagugagcac	20	86	BI432985	326	+	5'	53	72	53	138
StMiR 156b	<i>Solanum tuberosum</i>	ugacagaagagagagagcac	20	382	BQ506362	794	+	3'	594	613	232	613
StMiR 156c				85	CK272946	832	+	5'	582	601	582	666
StMiR 156d				171	CK262978	864	+	5'	387	406	387	557
StMiR 156e				72	BQ111685	507	+	5'	423	442	423	494
SbMiR 156a	<i>Sorghum bicolor</i>	ugacagaagagagugagcac	20	84	CN132665	630	+	5'	306	325	306	389
SbMiR 156b		ugacagaagagagagagcac	20	450	BG947368	670	-	5'	78	97	78	527
SbMiR 156b					BG947367	618	-	5'	26	45	26	475
SbMiR 156c				395	CD230908	652	-	5'	130	149	130	524
SbMiR 156d				102	AW679616	690	-	3'	89	108	7	108
SpMiR 156	<i>Sorghum propinquum</i>	ugacagaagagagagagcac	20	125	BF421158	425	-	3'	373	392	268	392
TaMiR 156a	<i>Triticum aestivum</i>	ugacagaagagagagagcac	20	52	BJ322639	640	+	5'	417	436	417	468
TaMiR 156a					CD454302	640	+	5'	474	493	474	525
TaMiR 156b				416	BJ317100	623	-	3'	491	510	95	510
VvMiR 156	<i>Vitis vinifera</i>	ugacagaagagagagagcac	20	47	CD800420	538	+	5'	457	476	457	503
ZmMiR 156m	<i>Zea mays</i>	ugacagaagagagugagcac	20	83	CF039276	479	+	5'	299	318	299	381
ZmMiR 156m					CF037245	511	-	3'	331	350	268	350
ZmMiR 156m					CF036982	592	+	5'	247	266	247	329
ZmMiR 156m					CF036852	598	+	5'	224	243	224	306
ZmMiR 156m					CF036471	579	+	5'	223	242	223	305
ZmMiR 156m					CF036182	372	+	5'	226	245	226	308
ZmMiR 156m					CF035885	553	+	5'	87	106	87	169
ZmMiR 156n				84	CF059465	516	+	5'	221	240	221	304
ZmMiR 156o				99	CF036467	383	+	5'	272	291	272	370
ZmMiR 156p		ugacagaagagagagagcac	20	189	CA408772	619	-	5'	304	323	304	492
ZmMiR 156q					CA408687	636	-	5'	304	323	304	492
ZmMiR 156r		ugacagaagagagugagcac	20	99	CF035522	424	+	5'	274	293	274	372
ZeMiR 156	<i>Zinnia elegans</i>	ugacagaagagagagagcac	20	67	AU293485	732	+	5'	343	362	343	409
AtMiR 157m	<i>Arabidopsis thaliana</i>	uugacagaagagagagagcac	21	50	AV784285	669	+	3'	339	359	310	359
CaMiR 157	<i>Capsicum annuum</i>	uugacagaagagagagagcac	21	155	CA516324	508	-	3'	417	437	283	437
GmMiR 157m	<i>Glycine max</i>	uugacagaagauagagagcac	21	261	BI699573	557	-	5'	223	243	223	483
GmMiR 157n				59	AW459710	417	-	5'	184	204	184	242
GmMiR 157o				57	BE611678	464	-	5'	296	316	296	352
GmMiR 157p				85	BE210632	313	+	5'	207	227	207	291
GmMiR 157q				112	AW756919	420	+	5'	150	170	150	261
GmMiR 157r		uugacagaagagagagagcac	21	147	BM731329	421	-	5'	4	24	4	150
GmMiR 157s				268	BG651519	486	-	3'	373	393	126	393
GrMiR 157a	<i>Gossypium raimondii</i>	uugacagaagauagagagcac	21	83	CO082782	891	+	5'	197	217	197	279
GrMiR 157b				92	CO076888	848	+	5'	169	189	169	260
GrMiR 157c		uugacagaagagagagagcac	21	183	CO105970	783	+	5'	400	420	400	582
GrMiR 157d				89	CO097436	785	+	3'	570	590	502	590
HaMiR 157a	<i>Helianthus annuus</i>	uugacagaagauagagagcac	21	86	BQ916415	492	+	5'	146	166	146	231
HaMiR 157b	<i>Helianthus annuus</i>	uugacagaagagagagagcac	21	364	BQ968177	575	-	3'	354	374	11	374
IbMiR 157	<i>Ipomoea batatas</i>	uugacagaagauagagagcau	21	283	CB330750	758	-	3'	546	566	284	566
InMiR 157a	<i>Ipomoea nil</i>	uugacagaagauagagagcau	21	126	BJ578566	713	+	5'	131	151	131	256
InMiR 157a					BJ578544	727	+	5'	131	151	131	256
InMiR 157b	<i>Ipomoea nil</i>	uugacagaagauagagagcau	21	183	BJ568024	542	+	5'	126	146	126	308
InMiR 157b					BJ573618	495	+	5'	131	151	131	313
InMiR 157b					BJ569965	714	+	5'	116	136	116	298
InMiR 157b					BJ563000	591	-	3'	563	583	401	583
InMiR 157b					BJ556392	592	-	3'	563	583	401	583
LsMiR 157	<i>Lactuca sativa</i>	uugacagaagagagagagcac	21	357	BQ874863	654	-	3'	533	553	197	553
LjMiR 157a	<i>Lotus corniculatus var. japonicus</i>	uugacagaagauagagagcac	21	58	BP039685	546	+	5'	417	737	417	474
LjMiR 157a					BP030242	518	+	5'	396	416	396	453
LjMiR 157b	<i>Lotus corniculatus var. japonicus</i>	uugacagaagagagagagcac	21	86	BP046017	503	-	3'	413	433	348	433
LjMiR 157c		uugacagaagauagagagcau	21	460	CN825561	660	-	3'	556	576	117	576
LeMiR 157a	<i>Lycopersicon esculentum</i>	uugacagaagauagagagcau	21	53	BM536323	331	+	5'	200	220	200	252
LeMiR 157b		uugacagaagauagagagcau	21	430	BE435668	620	-	3'	512	532	103	532
LeMiR 157b					AW933950	633	-	3'	488	508	76	508
LeMiR 157c				123	BE433988	239	+	3'	180	200	78	200
MtMiR 157	<i>Medicago truncatula</i>	uugacagaagauagagagcac	21	111	AW696064	665	+	5'	195	215	195	305
NbMiR 157a	<i>Nicotiana benthamiana</i>	uugacagaagagagagagcac	21	169	CK290419	610	+	3'	507	527	359	527
NbMiR 157b				207	CK294165	812	+	5'	515	535	515	721
NbMiR 157b					CK284079	812	+	5'	521	541	521	727
NtMiR 157	<i>Nicotiana tabacum</i>	uugacagaagauagagagcac	21	131	BP129308	705	+	5'	301	321	301	431
NaMiR 157a	<i>Nuphar advena</i>	uugacagaagagagagagcac	21	140	CK762856	565	-	5'	312	332	312	451

**Tab. 3** New miRNAs identified from plants (continued-2)

miRNA gene family	miRNA	Species	miRNA sequence*	Length (nt)	Precurs	EST	EST	Str	Folb	miRNA		Precursor	
					orhairpi	accession	length	nd	ack	Start	End	Start	End
					n	number		arm					
					(nt)								
	NaMiR 157b			191	CK745155	655	+	5'	441	461	441	631	
	NaMiR 157c			69	CD472699	521	-	3'	258	278	210	278	
	PaMiR 157	<i>Persea americana</i>	uugacagaagagagagac	21	CK751799	634	-	3'	237	257	42	257	
	PeMiR 157	<i>Phaseolus coccineus</i>	uugacagaagauagagacuu	21	CA902638	612	-	3'	480	500	317	500	
	PpcMiR 157	<i>Populus tremula x Populus tremuloides</i>	uugacagaagagagagac	21	CK087760	815	+	3'	454	474	119	474	
	PaMiR 157	<i>Prunus armeniaca</i>	uugacagaagagagagac	21	CB819319	504	+	5'	136	156	136	251	
	SiMiR 157	<i>Sesamum indicum</i>	uugacagaagagagagac	21	BU667823	289	+	5'	42	62	42	124	
	StMiR 157a	<i>Solanum tuberosum</i>	uugacagaagauagagac	21	BQ119413	712	+	5'	137	157	137	227	
	StMiR 157a				BF187371	460	+	5'	137	157	137	227	
	StMiR 157b		uugacagaagagagagac	21	CK272946	832	+	5'	581	601	51	667	
	StMiR 157c			189	CK246692	855	-	3'	806	826	638	826	
	SbMiR 157	<i>Sorghum bicolor</i>	uugacagaagagagagac	21	CN132665	630	+	5'	305	325	305	390	
	ZmMiR 157m	<i>Zea mays</i>	uugacagaagagagagac	21	CA408772	619	-	5'	304	324	304	492	
	ZmMiR 157m				CA408687	636	-	5'	304	324	304	492	
	ZeMiR 157	<i>Zinnia elegans</i>	uugacagaagagagagac	21	AU293485	732	+	5'	342	362	342	410	
158	GaMiR 158	<i>Gossypium arboreum</i>	gcccuaauguagacaagc	20	BE055331	957	+	5'	1	20	1	162	
	LeMiR 158a	<i>Lycopersicon esculentum</i>	auccaauguagacaagca	20	AJ319944	511	-	5'	427	446	426	475	
	LeMiR 158b			74	BI935773	627	+	3'	145	164	92	165	
	PpMiR 158	<i>Physcomitrella patens</i>	ucccaauuauacaagcc	20	BQ826645	604	-	3'	554	573	42	574	
	StMiR 158	<i>Solanum tuberosum</i>	ucccaauuauacaagga	20	BQ121631	663	-	5'	170	189	170	210	
159	GmMiR 159	<i>Glycine max</i>	uuuggauaaggagcucua	21	BM893181	513	+	3'	275	295	121	295	
	HvMiR 159a	<i>Hordeum vulgare</i>	uuuggauaaggagcucug	21	AJ475696	420	+	5'	22	42	22	324	
	HvMiR 159b	<i>Hordeum vulgare subsp. vulgare</i>	uuuggauaaggagcucug	21	BJ448559	622	+	3'	376	396	223	396	
	HvMiR 159b				BJ456281	597	-	5'	202	222	202	375	
	LtMiR 159	<i>Liriodendron tulipifera</i>	uuuggauaaggagcucua	21	CK750935	345	+	5'	1	21	1	345	
	MtMiR 159	<i>Medicago truncatula</i>	uuuggauaaggagcucua	21	BI265403	566	+	3'	481	501	334	501	
	MtMiR 159				AW691937	666	+	3'	476	496	329	496	
	OsMiR 159m	<i>Oryza sativa</i>	uuuggauaaggagcucug	21	CR289947	707	+	3'	402	422	174	422	
	OsMiR 159m				CF303003	461	+	3'	369	389	138	389	
	PgMiR 159	<i>Pennisetum glaucum</i>	uuuggauaaggagcucug	21	CD725199	784	+	3'	431	451	226	451	
	PsMiR 159	<i>Picea sitchensis</i>	uuuggauaaggagcucua	21	CO226626	804	-	5'	560	580	560	752	
	PtMiR 159	<i>Pinus taeda</i>	uuuggauaaggagcucua	21	CO161954	420	+	3'	215	235	38	235	
	PtMiR 159				CO166631	802	-	5'	603	623	603	799	
	ScMiR 159a	<i>Saccharum officinarum</i>	uuuggauaaggagcucug	21	CA261537	540	+	3'	323	343	118	343	
	ScMiR 159a				CA236922	953	+	3'	442	462	236	462	
	ScMiR 159a				CA203816	562	+	3'	443	463	237	463	
	ScMiR 159a				CA148454	685	+	3'	337	357	132	357	
	ScMiR 159a				CA079434	704	+	3'	213	233	8	233	
	ScMiR 159b		uuuggauaaggagcucuu	21	CA238586	636	+	3'	430	450	226	450	
	ScMiR 159c		uuuggauaaggagcucug	21	CA193052	522	+	5'	52	72	52	183	
	SaMiR 159	<i>Schedonorus arundinaceus</i>	uuuggauaaggagcucug	21	CK801265	489	+	3'	367	387	213	387	
	SaMiR 159				CK801264	461	+	3'	363	382	210	382	
	SbMiR 159a	<i>Sorghum bicolor</i>	uuuggauaaggagcucug	21	CF433018	647	+	3'	239	259	37	259	
	SbMiR 159a				CD224964	606	+	3'	453	473	252	473	
	SbMiR 159a				CD224004	694	+	3'	449	469	248	469	
	SbMiR 159a				CD213304	537	+	3'	448	468	247	468	
	SbMiR 159a				CD213290	553	+	3'	448	468	247	468	
	SbMiR 159a				CD204047	669	+	3'	464	484	263	484	
	SbMiR 159b		uuuggauaaggagcucug	21	BE595318	586	+	5'	140	160	140	376	
	SpMiR 159	<i>Sorghum propinquum</i>	uuuggauaaggagcucug	21	BG050158	429	+	5'	7	27	7	398	
	TaMiR 159	<i>Triticum aestivum</i>	uuuggauaaggagcucug	21	CA484819	674	+	3'	460	480	229	480	
	VvMiR 159a	<i>Vitis vinifera</i>	uuuggauaaggagcucua	21	CN007551	661	-	5'	485	505	485	652	
	VvMiR 159a				CF210361	563	+	3'	463	483	312	483	
	VvMiR 159b			557	CF208653	667	-	3'	564	584	28	584	
	ZmMiR 159m	<i>Zea mays</i>	uuuggauaaggagcucug	21	CD573488	604	+	5'	13	33	13	469	
160	GmMiR 160m	<i>Glycine max</i>	ugccugcuccuugauagca	21	CA801322	572	+	5'	23	43	23	102	
	GmMiR 160n		ugccugcuccuugauagca	21	BM887596	554	-	3'	319	339	233	339	
	GmMiR 160o			503	AI794655	529	-	3'	500	520	18	520	
	GrMiR 160a	<i>Gossypium raimondii</i>	ugccugcuccuugauagca	21	CO106264	713	-	5'	157	177	157	262	
	GrMiR 160a				CO106199	715	-	5'	157	177	157	262	
	GrMiR 160a				CO102555	686	-	5'	205	225	205	310	
	GrMiR 160b			293	CO095743	817	-	5'	406	426	406	698	
	HvMiR 160a	<i>Hordeum vulgare subsp. vulgare</i>	ugccugcuccuugauagca	21	CA002787	355	+	3'	324	344	1	344	
	HvMiR 160b		ugccugcuccuugauagca	21	BJ472971	596	-	3'	368	388	342	388	
	HvMiR 160b				BJ471957	479	-	3'	357	377	331	377	
	HvMiR 160b				BJ471544	481	-	3'	357	377	331	377	
	HvMiR 160b				BJ469735	543	-	3'	357	377	331	377	
	HvMiR 160b				BJ469445	560	-	3'	357	377	331	377	

**Tab. 3** New miRNAs identified from plants (continued-3)

miRNA gene family	Species	miRNA sequence*	Length (nt)	Precurs EST orhairpi n length (nt)	EST accession number	EST length	Strand	Fold back arm	miRNA Start	miRNA End	Precursor Start	Precursor End
HvMiR 160b					AV934494	482	-	3'	358	378	331	378
HvMiR 160b					AV934488	412	-	3'	329	349	333	349
HvMiR 160b					AV933114	608	-	3'	321	341	295	341
HvMiR 160b					AV932125	479	-	3'	357	377	331	377
OsMiR 160m	<i>Oryza sativa</i>	ugcuggcucccugauGCCA	21	82	CF330305	650	+	5'	248	268	248	329
OsMiR 160m					CF327976	627	+	5'	254	274	254	335
OsMiR 160n				131	CF327534	574	+	5'	402	422	402	532
OsMiR 160n					CF327535	536	-	3'	221	241	112	241
PgMiR 160	<i>Picea glauca</i>	ugcuggcucccugauGCCA	21	80	CO484471	770	+	5'	269	289	269	348
ScMiR 160	<i>Secale cereale</i>	ugcuggcucccugauGCCU	21	180	BE494482	556	-	3'	357	377	198	377
TtMiR 160a	<i>Triticum turgidum</i>	ugcuggcucccugauGCCA	21	87	BF293809	539	+	5'	259	279	259	345
TtMiR 160b				218	BQ788860	620	-	3'	264	284	67	284
ZmMiR 160m	<i>Zea mays</i>	ugcuggcucccugauGCCA	21	81	CD445121	648	+	5'	285	305	285	365
162 GmMiR 162m	<i>Glycine max</i>	ucgauaaaccucgcauccag	21	367	BU081804	557	+	5'	45	65	45	411
GmMiR 162n				84	CF807240	378	-	5'	300	320	300	383
LsMiR 162	<i>Lactuca sativa</i>	ucgauaaaccucgcauccag	21	89	BQ845158	367	+	5'	240	260	240	328
LjMiR 162	<i>Lupinus luteus</i>	ucgauaaaccucgcauccag	21	80	BG149136	442	+	3'	274	294	215	294
LmMiR 162	<i>Malus x domestica (cultivated apple)</i>	ucgauaaaccuugcauccag	21	88	CN894119	327	+	3'	173	193	106	193
MtMiR 162	<i>Medicago truncatula</i>	ucgauaaaccucgcauccag	21	87	BF003769	478	+	3'	221	241	155	241
OsMiR 162m	<i>Oryza sativa</i>	ucgauaaaccucgcauccag	21	114	CA764295	896	+	3'	425	445	332	445
VvMiR 162	<i>Vitis vinifera</i>	ucgauaaaccucgcauccag	21	85	CF516290	433	+	3'	368	388	304	388
163 OsMiR 163m	<i>Oryza sativa</i>	uugaagaggacuggaacguuc	21	47	CA763268	688	-	3'	621	641	595	641
PpMiR 163	<i>Picea glauca</i>	ugaagaagacuggagcuuca	22	74	CO236262	723	-	3'	654	675	602	675
PpMiR 163				92	CO254401	890	-	3'	691	712	621	712
PsMiR 163	<i>Picea sitchensis</i>	ugaagaagacuggagcuuca	22	93	CO222007	839	-	3'	695	716	624	716
PtMiR 163a	<i>Pinus taeda</i>	ugaagaagacuggagcuuca	22	93	CO165688	908	-	3'	565	586	494	586
PtMiR 163a					CO157878	798	-	3'	573	594	502	594
PtMiR 163a					CO163746	844	-	3'	648	669	578	669
PtMiR 163a					CO163668	831	-	3'	414	435	343	435
PtMiR 163a					CO157806	707	-	3'	293	314	222	314
PtMiR 163a					CF389355	622	-	3'	302	323	231	323
PtMiR 163a					CF669738	784	+	3'	351	372	307	372
PtMiR 163b			22	650	CO160078	735	+	5'	25	46	25	674
PtMiR 163b					CO158895	772	+	5'	101	122	101	750
PpbMiR 163	<i>Populus balsamifera x Populus deltoides</i>	ugaagaagaggauGCCA	21	100	CNS24551	654	+	3'	253	273	174	273
164 CsMiR 164	<i>Citrus sinensis</i>	uggagaagcagggcagugca	21	168	CK936755	674	+	5'	153	173	153	320
GaMiR 164	<i>Gossypium arboreum</i>	uggagaagcagggcaguuua	21	110	BF270853	579	-	3'	94	111	1	111
InMiR 164	<i>Ipomoea nil</i>	uggagaagcagggcagugcg	21	87	BJS77535	704	3'	502	522	436	522	
PtMiR 164	<i>Populus balsamifera subsp. trichocarpa</i>	uggagaagcagggcagugca	21	137	BU869001	621	+	5'	16	36	16	152
PtMiR 164					CK113235	537	+	5'	159	179	159	294
TaMiR 164a	<i>Triticum aestivum</i>	uggagaagcagggcagugca	21	60	CD899685	736	+	5'	615	635	615	674
TaMiR 164b				152	CA704421	476	+	5'	10	30	10	161
TaMiR 164b					CD929986	654	-	3'	171	191	134	191
165 HeMiR 165	<i>Hedyotis centranthoides</i>	ucggaccaggucauucccc	21	165	CB086732	632	+	3'	224	244	80	244
SbMiR 165	<i>Sorghum bicolor</i>	ucggaccaggucauuccca	21	426	CN140010	750	-	3'	726	746	301	726
166 GmMiR 166m	<i>Glycine max</i>	ucggaccaggucauucccc	21	128	CD405934	400	-	5'	112	132	112	239
GmMiR 166m					CD391899	338	-	5'	129	149	129	256
GmMiR 166m					CA819984	385	+	3'	241	261	133	261
GmMiR 166n				409	BQ785760	488	+	5'	47	67	47	455
GmMiR 166o				85	BM308126	468	+	3'	440	460	375	460
GmMiR 166p				106	BI893541	517	+	3'	125	145	39	145
GmMiR 166q		ucggaccaggucauucccg	21	156	BI972515	573	+	3'	306	326	171	326
GmMiR 166r		ucggaccaggucauucccu	21	136	BG316028	415	+	3'	238	258	123	258
HvMiR 166	<i>Hordeum vulgare subsp. vulgare</i>	ucggaccaggucauucccc	21	78	BQ760548	433	+	3'	136	156	79	156
InMiR 166	<i>Ipomoea nil</i>	ucggaccaggucauuccuc	21	169	BJS53847	485	+	3'	310	330	162	330
MtMiR 166a	<i>Medicago truncatula</i>	ucggaccaggucauucccc	21	96	AJ502524	411	+	3'	319	339	244	339
MtMiR 166a				94	BQ255147	607	+	3'	167	187	94	187
MtMiR 166a				97	AW685461	568	+	3'	220	240	144	240
MtMiR 166a				94	AI737566	304	+	3'	133	153	60	153
MtMiR 166b		ucggaccaggucauuccuc	21	171	AW694053	393	+	3'	371	391	221	391
OsMiR 166m	<i>Oryza sativa</i>	ucggaccaggucauucccu	21	97	CA760464	684	-	5'	442	462	442	538
PgMiR 166	<i>Picea glauca</i>	ucggaccaggucauuccuu	21	74	CO476713	810	-	5'	155	175	155	228
SoMiR 166	<i>Saccharum officinarum</i>	ucggaccaggucauucccc	21	127	CA067403	732	+	5'	44	64	44	170
SoMiR 166				131	CA067402	1098	+	5'	44	64	44	172
SbMiR 166	<i>Sorghum bicolor</i>	ucggaccaggucauucccc	21	108	CN126049	707	+	3'	157	177	70	177
TaMiR 166	<i>Triticum aestivum</i>	ucggaccaggucauuccca	21	111	CK204430	801	-	3'	168	188	78	188
TaMiR 166					CK204088	823	-	3'	167	187	77	187
ZmMiR 166m	<i>Zea mays</i>	ucggaccaggucauucccc	21	87	CK369135	655	+	3'	163	183	97	183

**Tab. 3** New miRNAs identified from plants (continued-4)

miRNA gene family	miRNA	Species	miRNA sequence*	Length (nt)	Precursor orhairpin length (nt)	EST accession number	EST length	Strand	Folb ack arm	miRNA		Precursor	
										Start	End	Start	End
	ZmMiR 166n			78		CO531986	680	+	3'	160	180	103	180
	ZmMiR 166o			78		CO521288	766	+	3'	161	181	104	181
	ZmMiR 166p			457		CF626917	661	-	3'	622	642	186	642
167	AtMiR 167m	<i>Arabidopsis thaliana</i>	ugaagcugccagcaugaucug	21		AU239920	619	+	5'	193	213	193	534
	GmMiR 167m	<i>Glycine max</i>	ugaagcugccagcaugaucua	21		BI095235	554	+	5'	405	425	405	486
	GmMiR 167m			64		CD411229	586	-	3'	496	516	453	516
	GmMiR 167n			64		BU548271	462	-	3'	209	229	166	229
	GmMiR 167n					BM892909	489	+	5'	124	144	124	187
	GmMiR 167n					BG509097	498	+	5'	227	247	227	290
	GmMiR 167o		ugaagcugccagcaugaucug	21		BQ629383	549	+	3'	222	242	134	242
	InMiR 167	<i>Ipomoea nil</i>	ugaagcugccagcaugaucug	21		BJ563257	578	+	5'	265	285	265	535
	OsMiR 167m	<i>Oryza sativa</i>	ugaagcugccagcaugaucug	21		CF322848	567	+	5'	129	149	129	228
	OsMiR 167m					CF309326	552	+	5'	102	122	102	201
	OsMiR 167n			90		BX928793	1039	+	5'	277	297	277	366
	PcMiR 167	<i>Phaseolus coccineus</i>	ugaagcugccagcaugaucuu	21		CA916400	283	+	5'	135	155	135	203
	PptMiR 167	<i>Populus tremula x Populus tremuloides</i>	ugaagcugccagcaugaucug	21		BU810017	684	+	5'	256	276	256	664
	SoMiR 167a	<i>Saccharum officinarum</i>	ugaagcugccagcaugaucug	21		CA284394	729	-	3'	629	649	501	649
	SoMiR 167a					CA284319	659	+	5'	123	143	123	271
	SoMiR 167b			139		CA287325	456	+	5'	136	156	136	274
	SMiR 167	<i>Saccharum sp</i>	ugaagcugccagcaugaucug	21		CF571260	403	+	5'	144	164	144	282
	TaMiR 167	<i>Triticum aestivum</i>	ugaagcugccagcaugaucua	21		CK209889	1031	+	3'	397	417	208	417
	ZmMiR 167m	<i>Zea mays</i>	ugaagcugccagcaugaucug	21		CF630597	417	-	3'	365	385	306	385
168	AtMiR 168m	<i>Arabidopsis thaliana</i>	ucgcuuggucagcaggcgaa	21		H77158	464	+	5'	42	62	42	145
	GmMiR 168m	<i>Glycine max</i>	ucgcuuggucagcaggcgaa	21		BE661028	688	+	5'	272	292	272	359
	GmMiR 168m					AW424354	195	+	5'	20	40	20	107
	GmMiR 168m					BE659175	300	-	3'	254	274	187	274
	HtMiR 168	<i>Hedyotis terminalis</i>	ucgcuuggucagcaggcgaa	21		CB076866	639	+	5'	168	188	168	307
	HvMiR 168	<i>Hordeum vulgare subsp. vulgare</i>	ucgcuuggucagcaggcgac	21		CA003609	456	+	5'	127	147	127	192
	HvMiR 168						418	+	5'	128	148	128	193
	LeMiR 168	<i>Lycopersicon esculentum</i>	ucgcuuggucagcaggcgac	21		BF097936	572	+	5'	16	36	16	160
	LeMiR 168					BE461111	551	+	5'	16	36	16	160
	LeMiR 168					BE461110	572	+	5'	16	36	16	160
	OsMiR 168m	<i>Oryza sativa</i>	ucgcuuggucagcaggcgac	21		CB667935	768	+	5'	105	125	105	171
	OsMiR 168m					CB662543	804	+	5'	113	133	113	179
	OsMiR 168m					CB655478	698	+	5'	116	136	116	182
	OsMiR 168m					CB629782	733	+	5'	12	32	12	78
	OsMiR 168m					CA756101	607	+	5'	118	138	118	184
	PbtMiR 168	<i>Populus balsamifera subsp. trichocarpa</i>	ucgcuuggucagcaggcgaa	21		BU871538	508	+	5'	37	57	37	154
	PtMiR 168	<i>Populus tremula</i>	ucgcuuggucagcaggcgaa	21		BU893331	297	+	5'	102	122	102	219
	PtMiR 168					BU889659	436	+	5'	109	129	109	226
	PtMiR 168					BU888204	520	+	5'	103	123	103	220
	PptMiR 168a	<i>Populus tremula x Populus tremuloides</i>	ucgcuuggucagcaggcgaa	21		BU811772	503	+	5'	53	73	53	170
	PptMiR 168a					BU809183	461	+	5'	109	129	109	226
	PptMiR 168b			140		BU886509	514	+	5'	70	90	70	209
	SsMiR 168	<i>Saccharum sp</i>	ucgcuuggucagcaggcgac	21		CF576659	636	+	5'	193	213	193	257
	SoMiR 168	<i>Saccharum officinarum</i>	ucgcuuggucagcaggcgac	21		CA173899	663	+	5'	164	184	164	227
	SoMiR 168					CA156827	708	+	5'	152	172	152	216
	SoMiR 168					CA123434	622	+	5'	184	204	184	248
	StMiR 168	<i>Solanum tuberosum</i>	ucgcuuggucagcaggcgac	21		CK254942	865	+	5'	85	105	85	217
	StMiR 168					CK249947	916	+	5'	12	32	12	144
	StMiR 168					CK246126	980	+	5'	12	32	12	144
	StMiR 168					CK243936	770	+	5'	10	30	10	142
	SbMiR 168	<i>Sorghum bicolor (sorghum)</i>	ucgcuuggucagcaggcgac	21		CN133168	669	+	5'	173	193	173	239
	SbMiR 168					CD213423	572	+	5'	173	193	173	239
	SbMiR 168					CD210663	530	+	5'	183	203	183	249
	SbMiR 168					CD209437	662	+	5'	184	204	184	250
	SbMiR 168					CD209338	549	+	5'	179	199	179	245
	SbMiR 168					CD207772	531	+	5'	193	213	193	259
	SbMiR 168					CD205823	586	+	5'	182	202	182	248
	SbMiR 168					CD205059	581	+	5'	182	202	182	248
	SbMiR 168					CN145907	312	+	5'	95	115	95	161
	VvMiR 168	<i>Vitis vinifera</i>	ucgcuuggucagcaggcgaa	21		CF604588	521	+	5'	52	72	52	161
	ZmMiR 168m	<i>Zea mays</i>	ucgcuuggucagcaggcgac	21		BU099153	532	+	5'	62	82	62	126
	ZmMiR 168m					BG842539	498	+	5'	97	117	97	161
	ZmMiR 168m					CA826777	522	+	5'	62	82	62	126
	ZmMiR 168m					BG842528	494	+	5'	97	117	97	161
169	GmMiR 169m	<i>Glycine max</i>	cagccaaggauagacuugccga	21		BF595231	457	+	3'	386	406	309	406



**Tab. 3** New miRNAs identified from plants (continued-5)

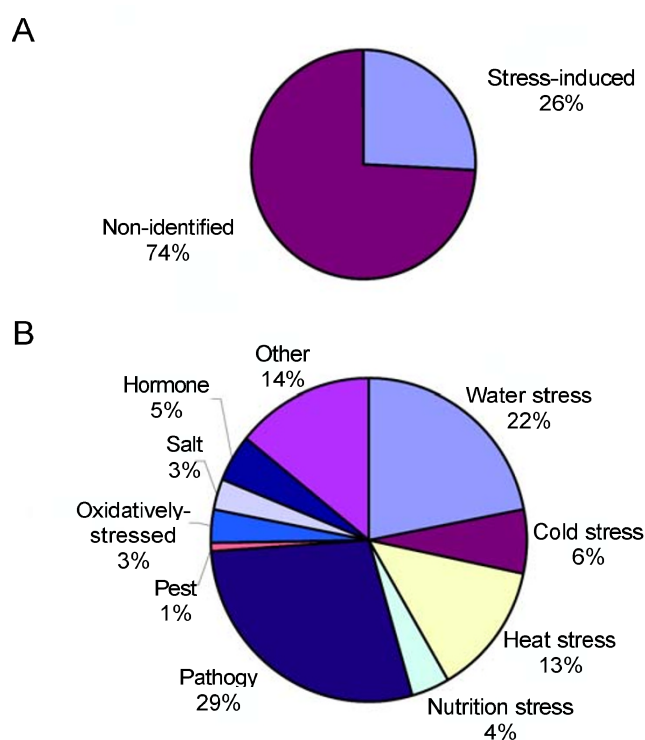
miRNA gene family	Species	miRNA sequence*	Length (nt)	Precurs orhairpi n length (nt)	EST accession number	EST length (nt)	Stra nd	Folb ack arm	miRNA		Precursor		
									Start	End	Start	End	
	GmMiR 169n		60		AW201497	400	+	5'	273	293	273	332	
	GmMiR 169o	cagccaaggagacuugccgg	21	82	CA953278	293	+	5'	90	110	90	171	
	GmMiR 169p		21	101	AW596073	317	+	5'	26	46	26	136	
	GsMiR 169	Glycine soja	cagccaaggagacuugccgg	21	103	BM524615	483	+	5'	63	83	63	165
	GsMiR 169				BM520766	550	+	5'	63	83	63	165	
	GsMiR 169				BF598910	465	+	5'	85	105	85	185	
	OsMiR 169m	Oryza sativa	cagccaaggagacuugccga	21	91	CF280503	479	+	5'	141	191	141	231
	OsMiR 169m				BQ906497	582	+	5'	169	189	169	263	
	OsMiR 169n		cagccaaggagacuugccgg	21	92	BX899561	863	-	3'	410	430	339	430
	OsMiR 169n				BX899553	870	+	5'	198	217	198	289	
	PtMiR 169a	Populus tremula	cagccaaggagacuugccga	21	100	BU862460	379	+	5'	89	109	89	189
	PtMiR 169b		cagccaaggagacuugccga	21	84	BU865420	519	+	5'	164	184	164	247
	SoMiR 169	Saccharum officinarum	uagccaaggagacuugccgg	21	143	CA148737	739	-	3'	643	663	521	663
	SoMiR 169				CA148650	627	+	5'	306	326	306	449	
	SbMiR 169	Sorghum bicolor	agccaaggagacuugccgu	21	92	CN143986	849	+	5'	72	92	72	163
	ZmMiR 169m	Zea mays	cagccaaggagacuugccgg	21	93	CD960569	364	+	5'	221	241	221	313
170	AcMiR 170a	Allium cepa	ugauugagccgggccaauauc	21	89	CF447956	774	-	3'	229	249	161	249
	AcMiR 170b				76	CF442989	808	-	5'	476	496	476	551
	AcMiR 170c				55	CF435529	459	-	3'	215	235	181	235
	AcMiR 170d				215	CF439580	800	-	5'	22	42	22	236
	AcMiR 170d					CF440325	741	-	5'	247	267	247	461
	AcMiR 170d					CF437865	841	-	5'	117	137	117	331
	AcMiR 170d					CF437845	833	-	5'	117	137	117	331
	AtMiR 170m	Arabidopsis thaliana	ugauugagccgggccaauauc	21	92	BX838271	1148	+	3'	125	145	54	145
	AtMiR 170n				172	BX839725	1135	-	5'	944	964	944	1115
	GmMiR 170m	Glycine max	ucauugagccgggccaauauc	21	80	CA937914	421	+	3'	265	285	206	285
	GmMiR 170n		uauugagccgggccaauauc	21	80	BM892213	391	+	3'	195	215	136	215
	HcMiR 170	Hedyotis centranthoides	egaugagccgggccaauauc	21	79	CB087210	578	+	3'	482	502	424	502
	HvMiR 170	Hordeum vulgare subsp. vulgare	ugauugagccgggccaauauc	21	78	CA009309	621	+	3'	394	414	337	414
	OsMiR 170a	Oryza sativa	ugauugagccgggccaauauc	21	99	CF329292	641	+	3'	308	328	230	328
	OsMiR 170b				107	CA756196	623	+	3'	210	230	124	230
	PpbMiR 170	Populus balsamifera x Populus deltoides	ugauugagccgggccaauauc	21	197	CN524603	813	-	5'	121	141	121	417
	StMiR 170	Solanum tuberosum	ugauugagccgggccaauauc	21	124	BG598365	717	-	3'	603	623	500	623
	SbMiR 170	Sorghum bicolor	ugauugagccgggccaauauc	21	54	BE365851	438	-	5'	367	387	367	421
	TaMiR 170a	Triticum aestivum	egaugagccgggccaauauc	21	81	BJ275219	707	-	5'	358	378	358	438
	TaMiR 170b		ugauugagccgggccaauauc	21	94	CD910903	588	+	3'	165	185	92	185
	ZmMiR 170m	Zea mays	ugauugagccgggccaauauc	21	115	CD441617	674	-	5'	530	550	530	644
	ZmMiR 170m					CD440809	620	+	3'	221	241	129	241
	ZmMiR 170n		ugauugagccgggccaauauc	21	79	CD963893	174	-	5'	44	64	44	122
	ZmMiR 170o		ugauugagccgggccaauauc	21	52	CD949754	71	+	5'	8	28	8	59
	ZmMiR 170p		ugauugagccgggccaauauc	21	62	AI657239	555	-	5'	482	502	482	543
171	AcMiR 171a	Allium cepa	ugauugagccgggccaauauc	21	89	CF447956	774	-	3'	229	249	161	249
	AcMiR 171a					CF442989	808	-	3'	476	496	407	496
	AcMiR 171b				115	CF439580	800	-	5'	22	42	22	236
	AcMiR 171b					CF440325	741	-	5'	247	267	247	461
	AcMiR 171b					CF437865	841	-	5'	117	137	117	331
	AcMiR 171b					CF437845	833	-	5'	117	137	117	331
	AcMiR 171c				216	CF435529	459	-	5'	215	235	215	430
	AtMiR 171m	Arabidopsis thaliana	ugauugagccgggccaauauc	21	94	BX838271	1148	+	3'	125	145	54	145
	AtMiR 171n				461	BX839739	1113	-	3'	979	999	539	999
	AtMiR 171n					BX839725	1135	-	3'	944	964	502	964
	AtMiR 171o				339	BU635724	762	-	5'	106	126	106	444
	GcMiR 171	Glycine clandestina	ugauugagccgggccaauauc	21	205	BG838164	855	-	5'	373	393	373	577
	GmMiR 171m	Glycine max	ugauugagccgggccaauauc	21	172	BM892515	567	-	5'	297	317	297	468
	GmMiR 171n		ucauugagccgggccaauauc	21	79	CA937914	421	+	3'	265	285	207	285
	GmMiR 171n					BM892213	391	+	3'	195	215	137	215
	GrMiR 171	Gossypium raimondii	ugauugagccgggccaauauc	21	117	CO129257	489	+	5'	272	292	272	388
	HcMiR 171	Hedyotis centranthoides	gauugagccgggccaauauc	21	79	CB087210	578	+	3'	482	502	424	502
	HvMiR 171	Hordeum vulgare subsp. Vulgare	ugauugagccgggccaauauc	21	78	CA009309	621	+	3'	394	414	337	414
	LeMiR 171	Lotus corniculatus var. japonicus	ugauugagccgggccaauauc	21	83	BP063562	516	-	3'	429	449	367	449
	OsMiR 171m	Oryza sativa	ugauugagccgggccaauauc	21	107	CA756196	623	+	3'	210	230	124	230
	OsMiR 171n				99	CF329292	641	+	3'	308	328	230	328
	ScMiR 171a	Saccharum officinarum	uauugagccgggccaauauc	21	375	CA139546	868	-	5'	420	440	420	794
	ScMiR 171b		ugauugagccgggccaauauc	21	201	CA104875	617	-	3'	545	565	365	565
	StMiR 171	Solanum tuberosum	ugauugagccgggccaauauc	21	124	BG598365	717	-	3'	603	623	500	623
	SbMiR 171	Sorghum bicolor	ugauugagccgggccaauauc	21	55	BE365851	438	-	5'	367	387	367	421
	TaMiR 171a	Triticum aestivum	ugauugagccgggccaauauc	21	94	CD910903	588	+	3'	165	185	92	185
	TaMiR 171b		egaugagccgggccaauauc	21	80	BJ275219	707	-	5'	358	378	358	437



**Tab. 4** ESTs matched by known *Arabidopsis* miRNAs, and identified as new miRNAs in plants

Plant species	EST number in dbEST*	EST contigs identified with known miRNAs	Identified%	miRNA
<i>Aegilops speltoides</i>	4,315	1	0.023	1
<i>Allium cepa</i>	19,582	14	0.071	7
<i>Arabidopsis thaliana</i>	322,641	20	0.006	17
<i>Brassica napus</i>	37,159	1	0.003	1
<i>Capsicum annuum</i>	22,961	3	0.013	3
<i>Citrus sinensis</i>	45,192	5	0.011	5
<i>Cycas rumphii</i>	5,952	1	0.017	1
<i>Eschscholzia californica</i>	8,531	1	0.012	1
<i>Glycine clandestina</i>	933	1	0.107	1
<i>Glycine max</i>	334,668	55	0.016	45
<i>Glycine soja</i>	16,508	3	0.018	1
<i>Gossypium arboreum</i>	39,007	2	0.005	2
<i>Gossypium hirsutum</i>	14,650	1	0.007	1
<i>Gossypium raimondii</i>	63,577	12	0.018	9
<i>Hedyotis centranthoides</i>	5,416	3	0.055	3
<i>Hedyotis terminalis</i>	4,875	1	0.021	1
<i>Helianthus annuus</i>	59,841	3	0.005	3
<i>Hordeum vulgare</i>	356,856	27	0.008	17
<i>Ipomoea batatas</i>	4,168	1	0.024	1
<i>Ipomoea nil</i>	25,899	10	0.037	5
<i>Lactuca sativa</i>	68,188	3	0.004	3
<i>Liriodendron tulipifera</i>	4,282	2	0.047	2
<i>Lotus corniculatus var. japonicus</i>	111,455	6	0.006	5
<i>Lupinus luteus</i>	364	1	0.275	1
<i>Lycopersicon esculentum</i>	150,596	13	0.009	9
<i>Malus x domestica (apple tree)</i>	129,134	3	0.002	3
<i>Medicago truncatula</i>	187,763	9	0.005	8
<i>Nicotiana benthamiana</i>	26,924	5	0.019	4
<i>Nicotiana tabacum</i>	10,945	1	0.009	1
<i>Nuphar advena</i>	6,854	3	0.044	3
<i>Oryza sativa</i>	284,007	41	0.0144	25
<i>Pennisetum glaucum</i>	2,528	1	0.040	1
<i>Persea americana</i>	4,276	1	0.023	1
<i>Phaseolus coccineus</i>	20,120	2	0.010	2
<i>Physcomitrella patens</i>	82,420	2	0.002	2
<i>Picea engelmannii x Picea sitchensis</i>	12,125	1	0.008	1
<i>Picea glauca</i>	55,108	6	0.011	6
<i>Picea sitchensis</i>	12,065	4	0.033	3
<i>Pinus taeda</i>	152,514	13	0.009	5
<i>Populus balsamifera subsp. trichocarpa</i>	26,825	3	0.011	3
<i>Populus balsamifera x Populus deltoides</i>	16,431	3	0.018	3
<i>Populus tremula</i>	31,288	6	0.019	4
<i>Populus tremula x Populus tremuloides</i>	65,981	9	0.014	8
<i>Prunus armeniaca</i>	5,735	2	0.035	2
<i>Saccharum officinarum</i>	246,301	34	0.014	25
<i>Saccharum sp</i>	9,636	2	0.021	2
<i>Schedonorus arundinaceus</i>	2,462	2	0.081	1
<i>Secale cereale</i>	9,194	1	0.011	1
<i>Sesamum indicum</i>	3,328	2	0.060	2
<i>Solanum tuberosum</i>	158,154	17	0.011	13
<i>Sorghum bicolor</i>	190,946	30	0.016	14
<i>Sorghum propinquum</i>	21,387	2	0.009	2
<i>Theobroma cacao</i>	6,557	1	0.015	1
<i>Triticum aestivum</i>	559,149	32	0.006	19
<i>Triticum turgidum</i>	8,714	1	0.011	1
<i>Vitis vinifera</i>	141,616	8	0.006	7
<i>Zea mays</i>	415,211	36	0.009	23
<i>Zinnia elegans</i>	17,529	3	0.017	3
Total	4,650,843	476	0.010	344

\* dbEST release 071604: Summary by Organism - July 16, 2004



**Fig. 4** miRNAs are induced and regulated by environmental biotic and abiotic stresses. **(A)** 26% of ESTs were obtained from stress-induced tissues. **(B)** ESTs containing miRNAs were obtained from tissue by various biotic and abiotic environmental stresses.

Out of the 476 identified EST contigs containing miRNAs, 123, or 25.84 %, were found in stress-induced plant tissues (Fig. 4A). These environmental stresses included abiotic and biotic stresses, such as drought, heat, cold, salinity, pathogen infection, or pests. In the 123 EST contigs, 36 (29%) EST contigs were associated with pathogen infection; 28 (22%) were associated with water stress; and 25 (19%) were associated with temperature stress. Other stresses, such as nutrition deficiency, salinity, and oxidative stress produced 4%, 3%, and 3% EST contigs, respectively (Fig. 4B). Abscisic acid (ABA) is considered a plant stress hormone, and jasmonic acid (JA) and salicylic acid (SA) are believed to be related to plant response to environmental stress. In this study, we found that 6 of the 123 EST contigs were created from the tissue induced by ABA, JA, SA, or other hormones. The other 18 (14%) EST contigs were created from the tissues induced by other abiotic and biotic environmental stresses, such as darkness or oxygen deficiency, wound treatment, acid- or alkaline-treatment.

Among all of the plant ESTs included in the TIGR database, 23% were from different stressed tissues. Here we found that 25.84 % of total ESTs containing potential miRNA were obtained from stress-induced plant tissues. Does this really mean that these microRNAs were induced by stress? More experiments need to be done to confirm this. However, recently three reports have confirmed our

**Tab. 5** Tissue-specific expression of plant miRNAs

miRNAs	Flower	Leaf	Root	Whole seedling	Adult plant	Developing embryos or seeds	Mixed tissue	Stem	Shoot meristem	Fruit, silique and spike	Callus	Somatic embryos	Xylem and phloem	Fiber	Total
156	18	20	8	2	2	11	4	6	13	7	2	1	1		92
157	20	5	3	5	1	3	5	3	1	5	2	1			52
158	2						1							1	4
159	9	6	4	5				2	5		3				34
160	4	10	1	1		3	2	1			4				22
161															
162	2	1	1	2				1		1					8
163	1	1	10						1				1	1	14
164	2		1			3				1					8
165	1	1	1												2
166	2	3	7	5			3	1	1			1	1		24
167	1	3	1	3		5	1	2		1	1				18
168	4	12	6	9		3	1	3			4				41
169		7		3		2		2							14
170	3	4	1			4	8		1	1	1				23
171	3	7	2	1		4	10	3		2	1				31
172	9	12	7	6		5	2	1	2	3					45
173						1									1
319	6	3	1	1		1	1	3		1					16
Total	87	95	54	43	3	45	38	28	24	22	18	3	3	2	449

hypothesis. miRNA 159, 319, 395 and 402 were regulated by GA, cold, water and sulfate starvation stress, individually [11, 14, 37].

Expression of plant miRNAs appears to be developmental or tissue specific. miRNAs are somewhat more strongly expressed in seedlings than in adult plants. Out of the 449 identified EST contigs which contain miRNAs and indicate origin tissues, 43 ESTs were obtained from young seedlings, but only 3 from adult plants (Tab. 5). Early and rapid growth stages may need more miRNA to regulate their gene expression. This is confirmed by recent reports [11, 23, 34]. In previous investigations, most miRNAs have been found to be strongly expressed in floral and leaf tissues [7, 11, 34, 38-41]. This observation was also supported by this study. Out of the 338 newly identified miRNAs in 60 different plant species, the majority were found in flowers and leaves. Fewer were found in roots, developing embryos, seeds or other tissues (Fig. 5 and Tab. 5). This indicates that miRNAs play important roles in plant growth and development other than controlling leaf and floral development [7, 11, 34, 38-41]. However, different miRNAs have different expression patterns. miRNA 157 preferentially expresses in flowers, but miRNAs 160, 163, and 168 preferentially express in leaves rather than in other tissues. miRNAs 156, 159, and 172 preferentially express in both flowers and leaves. miRNA 163 preferentially ex-

presses in roots. Although we obtained EST contigs containing miRNAs from other tissues, it did not appear that these miRNAs preferentially express in those tissues (Tab. 5). However, they may play other roles in those tissues in regulating expression of certain genes. In this study, miRNAs also have been found to be expressed in other tissues, such as roots, vascular tissues, stems, callus, somatic embryos and fibers. This suggests that miRNAs play some role and regulate certain gene expression in these tissues.

In the EST database, ESTs were unequally obtained from different tissues. Some tissues may contribute more ESTs than other tissues. Thus, more experiments need to be conducted to confirm this conclusion. However, this approach gave us more clues to study plant microRNAs, and this strategy alleviates the usually difficult step of predicating the correct tissues and conditions to search for expression evidence in a directed manner.

## DISCUSSION

### miRNAs are widespread and highly conserved in the plant kingdom, and they may have the same ancestor in very early evolution

miRNAs are widespread in plant species and eukaryotes. In previous studies, 71 miRNAs have been reported in *Arabidopsis* and rice. Out of the 71 miRNAs, 43 were

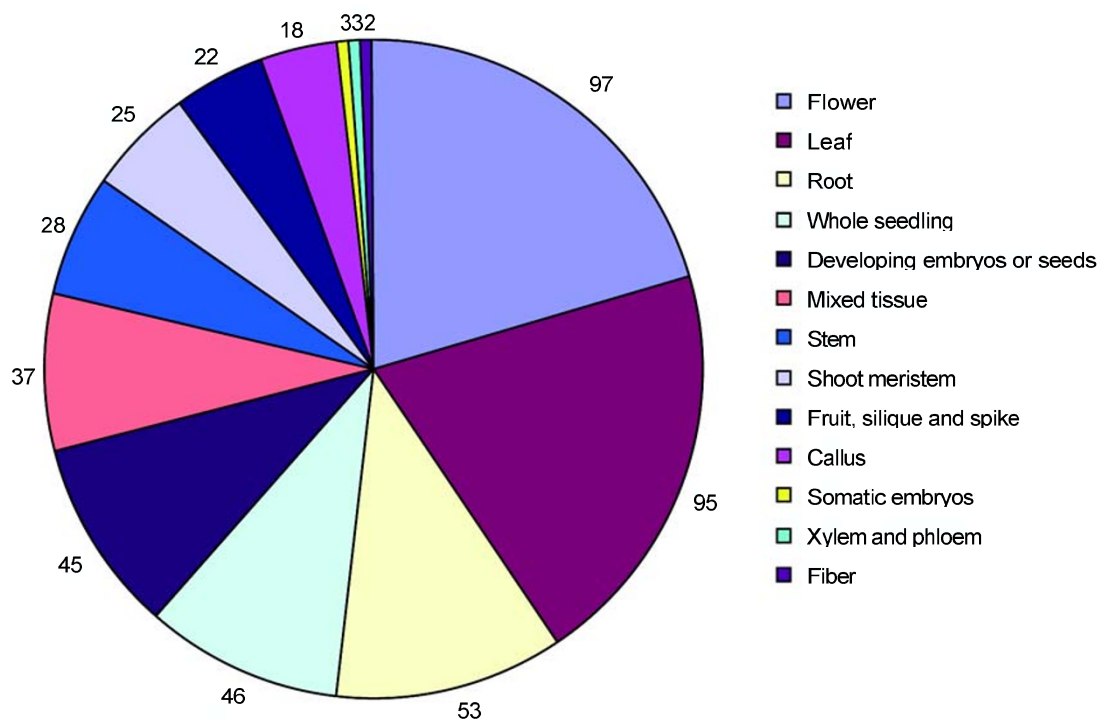


Fig. 5 Preferential expression of miRNAs

predicted and identified in *Arabidopsis*, and 28 in rice (the miRNA Registry database, Release 3.1, April 2004). In this study, we found 338 new miRNAs in 60 plant species using EST analysis and BLAST algorithms. With improvements in the strategy and technology for identifying miRNAs, more miRNAs will be discovered. In metazoans, such as *Drosophila* and humans, computational strategy suggests that miRNAs constitute nearly 1% of predicated protein-coding genes. In comparison with animals, a distinguishing feature of plants is that they are sessile and thus have to cope with, rather than move to avoid, complicatedly adverse environments and invasion of animals [47]. Plants should therefore have more complicated mechanisms to control gene expression and gene regulation. If miRNAs are involved in these mechanisms, there should be more miRNAs in plant genomes than in animal genomes. *Arabidopsis* is a model plant species for genome research; its genome sequence was published in 2000 [42]. The Arabidopsis Information Resource (TAIR) shows that *Arabidopsis* has currently 21,316 proteins (<http://www.arabidopsis.org/tools/bulk/protein/index.jsp>) excluding proteins of unknown subcellular location and chloroplast- and mitochondrion- targeted proteins. If plants have the same 1% miRNAs in their genome, *Arabidopsis* should have 213 miRNA genes. However, only 43 miRNAs have been reported in previous studies, plus the 17 newly discovered miRNAs candidates in this report; 60 miRNAs have been discovered in *Arabidopsis*. This is far from 1.0% of the protein-coding genes, so more miRNAs still await discovery in plants.

miRNAs are highly conserved among different plant species, including dicots and monocots. There is no nucleotide base substitution in the majority of miRNA families from different plant species. Some species only have a few (0-2) nucleotide base substitutions with *Arabidopsis* miRNAs. This suggests miRNAs may have the same ancestor in early evolution. However, some microRNAs may have a very high rate of divergence even if some are highly conserved.

### **Stress may play an important role in miRNAs regulation**

It is a complicated mechanism for both animals and plants to respond to stress. In this specific response, miRNA may be involved. In animals, loss of function of miR 14 increased sensitivity of fly to different stresses [43], and nutrient stress may induce expression of miRNA 234 [19]. In this study, we found 26% of EST contigs containing miRNAs are related to different biotic or abiotic environmental stresses, and out of these ESTs, 22% related to water stress, 19% related to temperature stress, and 29% related to pathogen infection (Fig. 4). These data suggest that environmental stress may play an important

role in miRNA gene expression in plants. This conclusion is also supported by recent reports [11, 14, 37]. Jones-Rhoades and Bartel (2004) showed that sulfate starvation induced overexpression of miRNA 395 [14]. The experiment by Sunkar and Zhu (2004) showed that miRNA 402 is strongly overexpressed under the condition of dehydration, cold, salt stress, or ABA treatment [11]. However, miRNA 319 is induced by cold [11], and miRNA 159 is regulated by GA [37].

Biotic and abiotic stress may function as signals to control and to regulate miRNA genes in plants (Fig. 6). First, these biotic or abiotic stressors form different signaling in plant cells, then induce a Ca<sup>2+</sup> signal or MAPK signal which interacts with a transcriptional factor, and produces down or over expression of the targeted miRNA genes (Fig. 6). The study of specific responses of miRNAs to environmental stress will help us improve plant resistance to environmental stresses, especially disease, drought, and salinity stress.

Another function of miRNAs may be involved in signal transduction. In this study, we found 7 ESTs containing miRNAs are associated with ABA or other hormone signaling. After we searched the NSF 2010 *Arabidopsis* Small RNA Project database (<http://cgrb.orst.edu/smallRNA/>) of small known RNAs and compared all previous known *Arabidopsis* miRNA to ABA signaling related genes with BLAST algorithms, we found miRNAs may be involved in the regulation of the ABA signaling genes.

Although several studies have tried to determine what produces microRNAs, little is known regarding miRNA biogenesis in plants. The data in this study suggest that environmental stresses and developmental switching may function as signaling to induce pri-miRNA transcription by RNA polymerases (Fig. 6). After pri-miRNAs are synthesized, they are processed to miRNA precursors or pre-miRNAs by the double-stranded RNase DICER-LIKE (DCL) [7, 34, 41]. Finally, the cleaved pre-miRNA will be translocated into cytoplasm by a HASTY protein, and cleaved into mature miRNAs (Fig. 6). In the cytosol, miRNA regulate gene expression at posttranscriptional levels by miRNA cleavage or by translational repression [2]. In animals, both of the two regulatory mechanisms are common on repression of target gene expression [2]. In plants, the majority of miRNAs interact with the internal regions of target mRNAs through perfect or near perfect base-pairing to cleave the mRNAs [1, 39, 44]. However, for some miRNAs, such as miRNA 172, although they have perfect or near-perfect complementarity sequence with target mRNAs, they inhibit the expression of target genes by binding to a unique site located within the open reading frame of the target mRNAs rather than miRNA cleavage [38, 45].

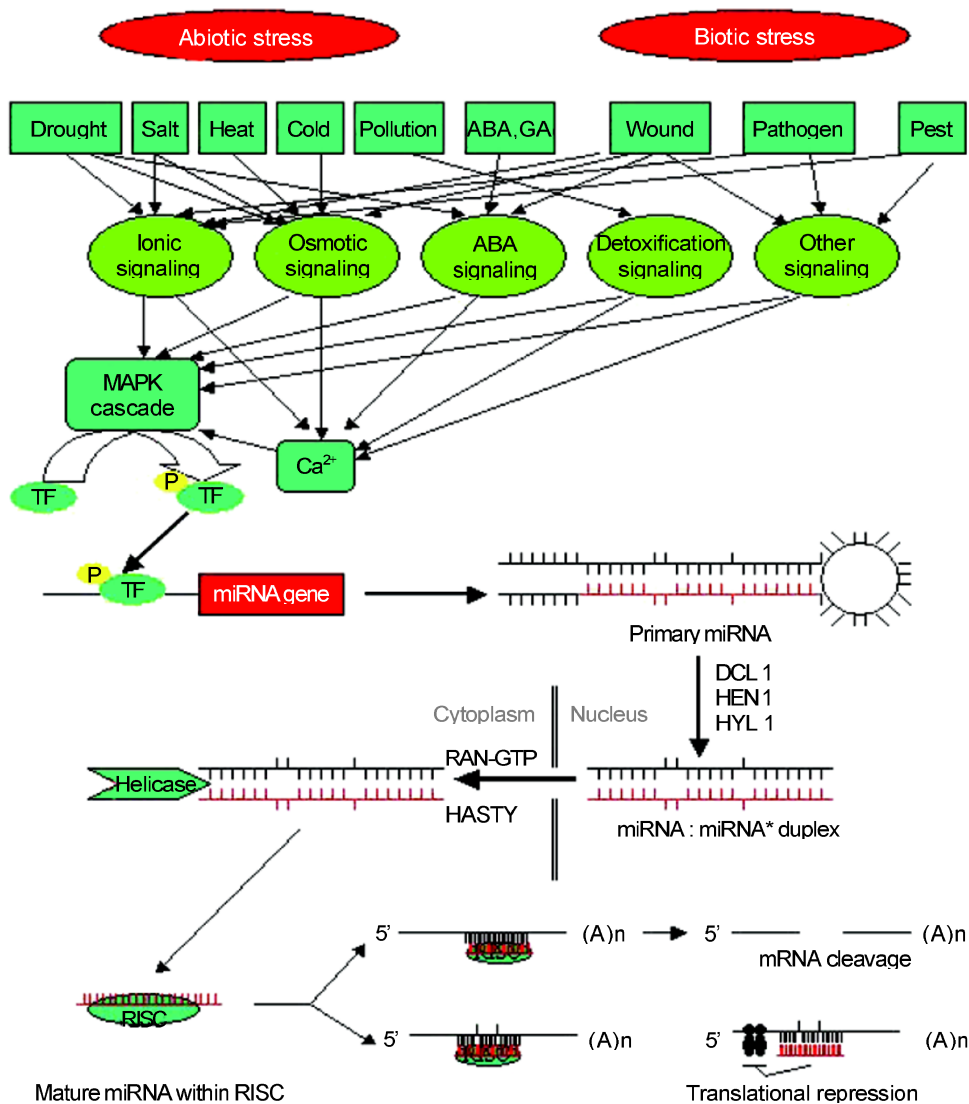


Fig. 6 A model for miRNA biogenesis, regulation of miRNA gene regulation, and the pathway of miRNA function.

**False positives**

Existence of false positives is one problem in many studies. To reduce the number of false positives, we combined multiple methods in this study. This is similar to that described by Bonnet et al. [13]. First, the number of EST hits is dramatically reduced after comparing EST sequences to previously known *Arabidopsis* miRNA sequences. According to the characteristic that plant miRNAs are highly conserved and only a few nucleotides change between plant species, ESTs with only 0-2 mismatched nucleotides with previously known *Arabidopsis* miRNA were considered as potential miRNA candidates. Second, potential ESTs were reduced about 50% by considering the secondary structure parameters based on previous reports [1, 2, 11,

13, 34, 46]. Third, the number of miRNAs was reduced by considering potential miRNA target genes in the DNA database. Fourth, repeated ESTs were removed by comparing similar EST sequences. Fifth, to avoid a miscount of the total EST sequences, ESTs with n-2, n-3, n-4, and n-5 mismatched nucleotides with previously known miRNA were chosen and paired with known miRNA sequences. Sixth, the matched ESTs were blasted against a database of known proteins to kick out the potential ESTs which actually code a protein rather than a structural RNA. After these six steps and other combined strategies, the number of EST hits dramatically reduced from 18,694 to 384. This gave us more confidence in identifying miRNAs.

If these identified miRNAs are confirmed by experi-

mental evidence, such as northern blotting, it will give us more confidence in the newly identified miRNAs and the hypothesis that stress may play an important role in miRNA regulation. However, ESTs are the true products of gene expression, and in some case they may be considered as the result of northern blotting. Actually, Smalheiser [36] successfully predicted the existence of a population of chimeric microRNA precursor-mRNA transcripts expressed in normal human and mouse tissues only by EST analysis. Bonnet *et al.* [13] used the EST database to confirm their 91 potential *Arabidopsis* and rice miRNAs identified by a computational strategy, and several genes have been identified by only EST analysis [27-29]. Thus, we have confidence in these identified miRNAs by EST analysis. EST analysis also gave us more information on miRNAs in other plant species in which it was impossible to identify miRNAs by traditional computational strategy due to lack of a DNA sequence. Finally, EST analysis provides information for designing more new experiments to understand miRNAs [36].

Received, Mar 1, 2005

Revised, Mar 28, 2005

Accepted, Apr 4, 2005

## REFERENCES

- 1 Bartel B, Bartel DP. MicroRNAs: at the root of plant development? *Plant Physiol* 2003; **132**:709-17.
- 2 Bartel DP. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 2004; **116**:281-97.
- 3 Mallory AC, Vaucheret H. MicroRNAs: something important between the genes. *Curr Opin Plant Biol* 2004; **7**:120-5.
- 4 Carrington JC, Ambros V. Role of microRNAs in plant and animal development. *Science* 2003; **301**:336-8.
- 5 Hunter C, Poethig RS. Missing links: miRNAs and plant development. *Curr Opin Genet Dev* 2003; **13**:372-8.
- 6 Griffiths-Jones S. The microRNA Registry. *Nucleic Acids Res* 2004; **32**:D109-11.
- 7 Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP. microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* 2004; **428**:84-8.
- 8 McHale NA, Koning RE. MicroRNA-directed cleavage of *Nicotiana sylvestris* PHAVOLUTA mRNA regulates the vascular cambium and structure of apical meristems. *Plant Cell* 2004; **16**:1730-40.
- 9 Lai EC. microRNAs: Runts of the genome assert themselves. *Curr Biol* 2003; **13**:R925-36.
- 10 Lai EC, Tomancak P, Williams RW, Rubin GM. Computational identification of *Drosophila* microRNA genes. *Genome Biol* 2003; **4**:R42.
- 11 Sunkar R, Zhu JK. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* 2004; **16**:2001-19.
- 12 Adai A, Johnson C, Mlotshwa S, *et al.* Computational prediction of miRNAs in *Arabidopsis thaliana*. *Genome Res* 2005; **15**:78-91.
- 13 Bonnet E, Wuyts J, Rouze P, Van de Peer Y. Detection of 91 potential in plant conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proc Natl Acad Sci U S A* 2004; **101**:11511-6.
- 14 Jones-Rhoades MW, Bartel DP. Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. *Mol Cell* 2004; **14**:787-99.
- 15 Lai EC. Predicting and validating microRNA targets. *Genome Biol* 2004; **5**:A115.
- 16 Li Y, Li W, Jin YX. Computational identification of novel family members of microRNA genes in *Arabidopsis thaliana* and *Oryza sativa*. *Acta Biochim Biophys Sin* 2005; **37**:75-87.
- 17 Rhoades MW, Reinhart BJ, Lim LP, *et al.* Prediction of plant microRNA targets. *Cell* 2002; **110**:513-20.
- 18 Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP. Vertebrate microRNA genes. *Science* 2003; **299**:1540.
- 19 Lim LP, Lau NC, Weinstein EG, *et al.* The microRNAs of *Caenorhabditis elegans*. *Genes Dev* 2003; **17**:991-1008.
- 20 Krol J, Krzyzosiak WJ. Structural aspects of microRNA biogenesis. *IUBMB Life* 2004; **56**:95-100.
- 21 Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell* 2003; **115**:787-98.
- 22 Rajewsky N, Succi ND. Computational identification of microRNA targets. *Dev Biol* 2004; **267**:529-35.
- 23 Wang JF, Zhou H, Chen YQ, Luo QJ, Qu LH. Identification of 20 microRNAs from *Oryza sativa*. *Nucleic Acids Res* 2004; **32**:1688-95.
- 24 Wang XJ, Reyes JL, Chua NH, Gaasterland T. Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biology* 2004; **5**:R65.
- 25 Adams MD, Kelley JM, Gocayne JD, *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 1991; **252**:1651-6.
- 26 Matukumalli LK, Grefenstette JJ, Sonstegard TS, van Tassell CP. EST-PAGE - managing and analyzing EST data. *Bioinformatics* 2004; **20**:286-8.
- 27 Graham MA, Silverstein KAT, Cannon SB, VandenBosch KA. Computational identification and characterization of novel genes from legumes. *Plant Physiol* 2004; **135**:1179-97.
- 28 Jung JD, Park HW, Hahn Y, *et al.* Discovery of genes for ginsenoside biosynthesis by analysis of ginseng expressed sequence tags. *Plant Cell Rep* 2003; **22**:224-30.
- 29 Ohlrogge J, Benning C. Unraveling plant metabolism by EST analysis. *Curr Opin Plant Biol* 2000; **3**:224-8.
- 30 Boguski MS, Lowe TM, Tolstoshev CM. dbEST—database for “expressed sequence tags”. *Nat Genet* 1993; **4**:332-3.
- 31 Altschul SF, Madden TL, Schäffer AA, *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**:3389-402.
- 32 Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of rna secondary structure. *J Mol Biol* 1999; **288**:911-40.
- 33 Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003; **31**:3406-15.
- 34 Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. *Genes Dev* 2002; **16**:1616-26.
- 35 Ambros V. MicroRNA pathways in flies and worms: Growth,



- death, fat, stress, and timing. *Cell* 2003; **113**:673-6.
- 36 Smalheiser NR. EST analyses predict the existence of a population of chimeric microRNA precursor-mRNA transcripts expressed in normal human and mouse tissues. *Genome Biol* 2003; **4**:403.
- 37 Achard P, Herr A, Baulcombe DC, Harberd NP. Modulation of floral development by a gibberellin-regulated microRNA. *Development* 2004; **131**:3357-65.
- 38 Chen XM. A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science* 2004; **303**:2022-5.
- 39 Llave C, Xie ZX, Kasschau KD, Carrington JC. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 2002; **297**:2053-6.
- 40 Palatnik JF, Allen E, Wu XL, et al. Control of leaf morphogenesis by microRNAs. *Nature* 2003; **425**:257-63.
- 41 Park W, Li JJ, Song RT, Messing J, Chen XM. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol* 2002; **12**:1484-95.
- 42 The *Arabidopsis* Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 2000; **408**:796-815.
- 43 Xu PZ, Vernooy SY, Guo M, Hay BA. The *Drosophila* MicroRNA mir-14 suppresses cell death and is required for normal fat metabolism. *Current Biology* 2003; **13**:790-5.
- 44 Llave C. MicroRNAs: more than a role in plant development? *Mol Plant Path* 2004; **5**:361-6.
- 45 Aukerman MJ, Sakai H. Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *Plant Cell* 2003; **15**:2730-41.
- 46 Bartel DP, Chen CZ. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet* 2004; **5**:396-400.
- 47 Zhu JK. Salt and drought stress signal transduction in plants. *Annu Rev Plant Biol* 2002; **53**:247-73.