Identification and characterization of new plant microRNAs using EST analysis

Bao Hong ZHANG^{1, 2, *}, Xiao Ping PAN¹, Qing Lian WANG², George P. COBB¹, Todd A. ANDERSON¹

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

² 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)

*Correspondence: Bao Hong ZHANG

Tel: +1-806-8854567; Fax: +1-806-8854577;

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

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 thorough 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)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 Zucker folding algorithm with mFold 3.1 [32, 33]. All mFold outputs including information about the number of structures, free energy (ΔG 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 (ΔG 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 (helical score + $(ABS(\Delta G)/2)/2$. 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 sylvestris*) [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

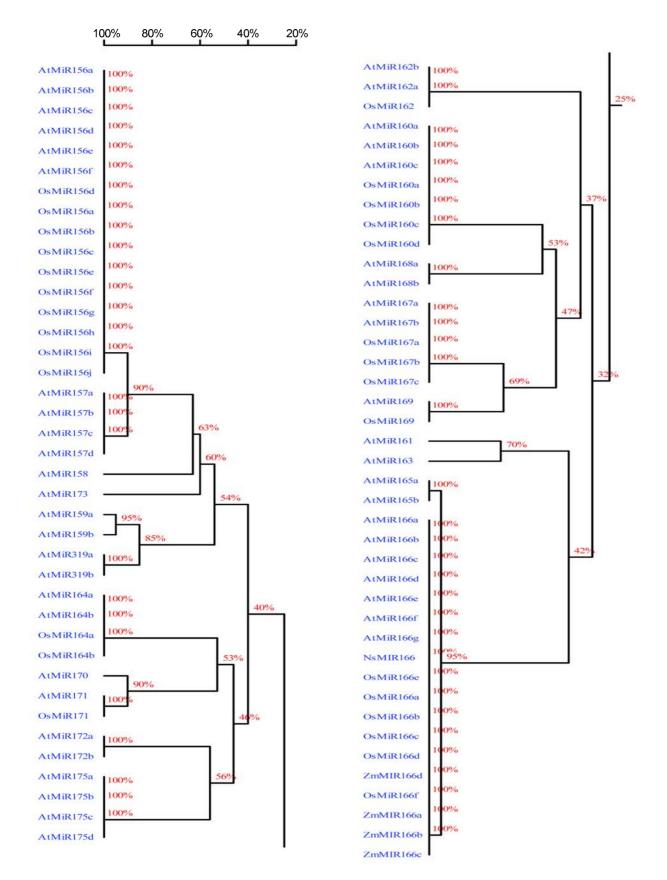


Fig. 1 Nucleotide sequence identity of 75 previously known microRNAs

mir-156 family		<i>mir</i> -165 family	V-2	
AtMiR156a .TGACAGAAGAGA	AG <mark>T</mark> GAGCA 19	AtMiR165a	TCGGACCAGGCTTCATCCCC	20
AtMiR156b .TGACAGAAGAGA	AG <mark>T</mark> GAGCA 19	AtMiR165b	TCGGACCAGGCTTCATCCCC	20
AtMiR156c .TGACAGAAGAGA	AG <mark>T</mark> GAGCA 19	AtMiR166a	TCGGACCAGGCTTCATTCCCC	20
AtMiR156d .TGACAGAAGAGA	AG <mark>T</mark> GAGCA 19	AtMiR166b	TCGGACCAGGCTTCATTCCCC	20
AtMiR156e .TGACAGAAGAGA	AG <mark>T</mark> GAGCA 19	AtMiR166c	TCGGACCAGGCTTCATTCCC	20
AtMiR156f .TGACAGAAGAGA	G <mark>T</mark> GAGCA 19	AtMiR166d	TCGGACCAGGCTTCATTCCCC	20
OsMiR156a .TGACAGAAGAGA	G <mark>T</mark> GAGCA 19	AtMiR166e	TCGGACCAGGCTTCATTCCC	20
	GTGAGCA 19	AtMiR166f	TCGGACCAGGCTTCATTCCC	20
	G <mark>T</mark> GAGCA 19	AtMiR166g	TCGGACCAGGCTTCATTCCC	20
	G <mark>T</mark> GAGCA 19	OsMiR166a	TCGGACCAGGCTTCATTCCCC	20
	GTGAGCA 19	OsMiR166b	TCGGACCAGGCTTCATTCCC	20
	G <mark>T</mark> GAGCA 19	OsMiR166c	TCGGACCAGGCTTCATTCCC	20
	GTGAGCA 19	OsMiR166d	TCGGACCAGGCTTCATTCCC	20
	G <mark>I</mark> GAGCA 19	OsMiR166e	TCGGACCAGGCTTCATTCCC	20
	G <mark>T</mark> GAGCA 19	OsMiR166f	TCGGACCAGGCTTCATTCCC	20
	GIGAGCA 19	NSMIR166	.CGGACCAGGCTTCATTCC.	18
	GAGAGCA 20	ZmMIR166a	TCGGACCAGGCTTCATTCCC	20
AtMiR157b TTGACAGAAGATA		ZmMIR166b	TCGGACCAGGCTTCATTCCC	20
AtMiR157c TTGACAGAAGATA		ZmMIR166c	TCGGACCAGGCTTCATTCCC	20
AtMiR157d .TGACAGAAGATA		ZmMIR166d	TCGGACCAGGCTTCATTCCC	20
Consensus tgacagaaga a		Consensus	cggaccaggcttcat cc	
compensas egacagaaga a	g gugeu			
mir-158 family		mir-167 family		
AtMiR158 TCCCAAATGTAGA	CAAAGCA 20	AtMiR167a	TGAAGCTGCCA. COATGATCTA	21
Atmix158 TCCAAATOIAGA	ICAAAOCA 20		TGAAGCTGCCA.GCATGATCTA	21
<i>mir</i> -159 family			TGAAGCTGCCA.COATGATCTA	21
AtMir159a TTTGGATTGAAG	GGAGCTCT 20	OsMiR167b	TGAAGCTGCCA.CCATGATCTA TGAAGCTGCCA.CCATGATCTA	21
Atmir159b TTTGGATTGAAG		OsMiR167c	TGAAG <mark>CT</mark> GCCA.G <mark>CATGATC</mark> TA	21
AtMiR319a .TTGGACTGAAG			CAGCCAAGGATGACTTGCCG	20
	GGAGCTCC 19	AtMiR169		
Consensus ttgga tgaag	ggagete	OsMiR169	CAGCCAAGCATGACTTGCCG	20
		Consensus	c gcca g atga t	
			5	
mir-160 family		<i>mir</i> -168 family		
AtMiR160a TGCCTGGCTCCC	TGTATGCC 20	AtMiR168a	TCGCTTGGTGCAGGTCGGGA	20
AtMiR160b TGCCTGGCTCCC		AtMiR168b	TCGCTTGGTGCAGGTCGGGA	20
AtMiR160c TGCCTGGCTCCC'		Consensus	tcgcttggtgcaggtcggga	20
OsMiR160a TGCCTGGCTCCC		combenibub	regereggegedggeegggd	
OsMiR160b TGCCTGGCTCCC OsMiR160c TGCCTGGCTCCC				
OsMiR160d TGCCTGGCTCCC'				
Consensus tgcctggctccc		<i>mir</i> -170 family		
		AtMiR170	TGATTGAGCCGTGTCAATATC	21
mir-161 family				
AtMir161 TTGAAAGTGACT	ACATCCCG 20			
		<i>mir</i> -172 family		
AtMiR163 TTGAAGACGACTTGG		AtMiR172a	AGAATCTTGATGATGCTGCA	20
Consensus ttgaa gact	ac tcg	AtMiR172b	AGAATCTTGATGATGCTGCA	20
		Consensus	agaatcttgatgatgctgca	
<i>mir</i> -162 family				
Name of the second s	TOON TOON 20			
AtMiR162a TCGATAAACCTC		mir-173 family		
AtMiR162b TCGATAAACCTC				22
OsMiR162 TCGATAAACCTC	TGCATCCA 20	AtMiR173	TTCGCTTGCAGAGAGAAATCAC	22
Consensus tcgataaacctc	tgcatcca			
54	19212			
		mir-175 family		
mir-164 family		AtMiR175a	AAGTAGATGATGTAATTTTGAGG	23
	GCACGTGC 20	AtMiR175b	AAGTAGATGATGTAATTTTGAGG	23
		AtMiR175c	AAGTAGATGATGTAATTTTGAGG	23
OsMiR164a TGGAGAAGCAGG OsMiR164b TGGAGAAGCAGG		AtMiR175d	AAGTAGATGATGTAATTTTGAGG	23
Consensus tggagaagcagg		Consensus	aagtagatgatgtaattttgagg	
consensus cygagaagCagg	genegege			

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 sylvestris*) (*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

miRNA	E value	Blast hits	Numbe	r of EST	with 0-2 m	ismatched	ESTs for	miRNA w	vith 0-2 mi	smatched	miRNA
				nuc	eotides			nuce	otides		
			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. 1 Summary of plant EST blast hits to known plant miRNA sequences

miRNA	EST without mismatch	EST with 1 nt mismatch	EST with 2 nt mismatch
156	CN132665, CF059894,	CB675101, D41985, CO529327, CO526369, CO525653, CO523543, CO520756,	
	CF059465, CF059465,	CO517866, CO414834, AL810223, BP660915, BP657257, CO237265,	
	CF039276, CF037245,	CO229128, CO169404, CO121346, CO105970, CO097437, CO097436,	
	CF036982, CF036852,	CO092899, CO081765, CO081202, CO051680, CN948866, CN581492,	
	CF036471, CF036467,	CN580870, CN578779, CN523533, CN519769, BP056010, BP052650, BP051612,	
	CF036182, CF035885,	BP051429, BP051130, BP046017, BP045933, BP044777, BP039258, AV777338,	
	CD835236, CB677777,	CK862588, CR285370, CK762549, CK754211, CK753210, CK751799,	
	CB677501, CB674083,	CK746364, CK665702, CK441500, CK441500, CK441499, CK441068,	
	CB653304, CB647907,	CK294165, CK293650, CK290419, CK286971, CK286411, CK286410,	
	CB643501, CA294779,	CK284079, CK284078, CK275093, CK272946, CK262978, CK254420,	
	BI432985, BG300360,	CK246692, CK196549, CK119834, CK105751, CK087760, CF836044,	
	BE807821, AU091537	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, A1999403, A1995421, A1738310,	
		AI726550, CK092224, CB674957	
157	CO082782, CO076888,	CN825561, CB330750, CA902638, BJ578566, BJ578544, BJ573618, BJ569965,	
157		BJ568024, BJ563000, BJ558059, BJ556392, CA525602, BM536323, BE435668,	
	BP039685, BP030242,		
	BP029005, BP129308,	BE433988, AW933950, CO529327, CO526369, CO520756, BP657257,	
	BQ916415, BQ119413,	CO105970, CO097437, CO097436, CO081765, CO081202, CN948866,	
	BI971210, BI699573,	CN580870, CN523533, BP046017, CK762856, CK751799, CK745155,	
	BF187371, BE821022,	CK294165, CK293650, CK290419, CK286971, CK286411, CK286410,	
	BE611678, AL368695,	CK284079, CK284078, CK272946, CK254420, CK246692, CK119834,	
	AU089181, BE210632,	CK105751, CK087760, BP132966, CD800420, CD472699, CB970427,	
	AW756919, AW459710	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,	CR289947, CK801265, CK801264, CF433018, CF308674, CF303003, CD725199,	
	CF210361, CF208653,	CD573488, CD224964, CD224004, CD213304, CD213290, CD204047,	

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

www.cell-research.com | Cell Research, 15(5):336-360, May 2005

miRNA	EST without mismatch	EST with 1 nt mismatch	EST with 2 nt mismatch
	CB036045, AU251508,	CA731881, CA484819, CA261537, CA236922, CA203816, CA193052,	
	BU888582, BM893181,	CA148454, CA079434, AJ475696, BJ456281, BJ448559, BG050158, BE595318,	
	BI265403, AW691937,		
	AJ388855,		
160	CA801322, BU760860,	CD480318, BM887596, AI900826, AI794655, BQ401391, CO106264, CO106199,	
	BG882856, CA002787,	CO102555, CO095743, BU986263, BJ472971, BJ471957, BJ471544, BJ469735,	
	CF330305, CF327976,	BJ469445, AV934494, AV934488, AV933114, AV932125, BF622299,	
	CF327535, CF327534,	AW038480, CB643398, CB643394, CA305796, CN520271, CF119506,	
	BF293809, CD445121,	CA266669, CA247079, CA209000, CA067925, CF575341, BE494482,	
	CO484471,	BQ112228, BE921523, CA615738, BQ788860	
161			
162	BF003769, BG149136,	CN894119, CA764295,	CF807240
	BI785837, BQ845158,		
	BU081804, CF516290,		
163			-
164	CK936755, CK113235,		CB009586, BF270853, CD929986, CD899685,
	CA704421, BU869001		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,	BG316028, BJ553847, BI972515, AW694053	
100	AW685461, BI893541,	5010020, 5000011, 5072210, 111051055	
	BM092635, M308126,		
	BQ255147, BQ760548,		
	BQ785760, CA067402,		
	CA067403, CA819984,		
	CD391899, CD405934,		
	CF626917, CK369135,		
	CN126049, O531986,		
	CO521288, AJ771029,		
	AJ770960, AJ770980		
1/7		AU220020 AW210200 AW472145 DC500007 DC510070 DI55(750	00002107
167	CK209908, BI095235	AU239920, AW318298, AW472145, BG509097, BG510970, BJ556759,	CD003107
		BJ563257, BM892909, BQ629383, BU548271, BU810017, BU886852,	
		BX928793, CA284319, CA284394, CA287325, CA916400, CD411229,	
1/0	1177150 DD/50175	CF309326, CF322848, CF571260, CF630597, CK209889, CK768366, CN472374	C1002(00 C1002500 CD/(7025 CD/(254)
168	H77158, BE659175,	BF097936, BE461111, BE461110, CK254942, CK249947, CK246126, CK243936	CA003609, CA003589, CB667935, CB662543
	BE661028, AW424354,		CB655478, CB629782, CA756101, CA173899
	CB076866, BU871538,		CA156827, CA123434, CF576659, CN145907
	BU893331, BU889659,		CN133168, CD213423, CD210663, CD20943
	BU888204, BU886509,		CD209338, CD207772, CD205823, CD205059
	BU811772, BU809183,		CA826777, BU099153, BG842539, BG842529
	CF604588, CN901684		
169	BF595231, AW201497,	CA953278, AW596073, BM524615, BM520766, BF598910, BX899561,	CN143986, CA148737, CA148650
	CF280503, BQ906497,	BX899553, CD960569, BU865420	
	BU862460,		
170	-	CF329292, CD910903, CD441617, CD440809, CA756196, CA009309	AI657239, AI770772, AV419737, AW45932
			BE040695, BE365851, BE641457, BG583589,
			BG598365, BG653310, BG838164, BJ275219
			BM892213, BM892515, BQ255243, BQ46101
			BX838271, BX839725, BX839739, CA067169
			CA069452, CA120921, CA937914, CB087210

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

niRNA	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,	AU304757, CA191585, CA139546, CF329292, CD964933, CD910903,	BJ275219, CK210439, CA937914, BM892213,
	BX839725, BX838271,	CD441617, CD440809, CA756196, CA120921, CA106627, CA104875,	CB087210, BP063562, CK297209, CK291560,
	CK210886, CF603900,	CA009309, AI736818, CB087210	CK291266, CK289500
	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		
172	BU084569, BI320499,	CO099417, CO120684, CO090457, BM348229, CF506714	CA922314, CD476336, CR280156, CA755817,
	CO080218, CO069513,		CA282304, CA268785, CA218620, CA216531,
	AI484737, BQ114970		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,	BJ191979, CF983804, CA184549, CA289676	BM520615, AV527132, BQ609349, AJ613606,
	BE475558, CK760683,		BG237979, BQ453148, CA227452, CD925456,
	CK091603, CA483944,		CA599744, AJ480356, AJ480355, CA006699,
	CK751513		BQ466150, CA149108, CF759104, BQ841783,
			CB920070, CB919200, BG606141, CN494329,

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

EST hits without nucleotide change. This indicates that the expression of different miRNAs and their number of copies in a plant genome may differ. The higher the number of EST hits, the more copies of the miRNA genes. This also indicates that there is a very high rate of divergence occurring for some miRNAs even if some are highly conserved.

A. miRNA 156 assica napus

- Glycine max S'-UGACAGA AGAGA GUGAGCAC GCU G GU UGUAUG G S'-ACUGUCU UCUCU CACUCGUG UGG C UA ACAUAC G A U CG G GU G Oryza sativa Saccharum officinarum
 Sesamum indicum
 GA
 GA
 A

 U/A
 AG
 CAUUGA
 A

 S⁺
 G CAGAGAG
 GAGAGGACAG
 CCOUGA
 A

 S⁺
 G CAGAGG
 GAGAGGACAG
 CCOUGA
 A

 S⁺
 C GUCUCC
 UCUCUCGUGUU
 UGUUCC
 GUUUUCC
 GUUUCC
 GUUUCC
 GUUUCC
 GUUUCC
 GUUUCC
 GUUUUCC
 GUUUCC
 GUUUCC
 GUUUCC
 GUUUCC
 Solanum tuberosum
 A
 A
 U
 UA
 GU
 U

 Sorghum bicolor
 A
 C
 ACG
 A

 LGACA
 GUGAGGA
 GUGAGGA
 GCAGGA
 GCAU
 V

 '5'
 GAAGAGA
 GUGAGGA
 GGCACGA
 GCAU
 V
 V

 '3'
 UUUCUCU CACUCGUG
 GCCCC
 CUGUA
 V
 U
 U
 U
 U
 U
 V
 U
 V
 U
 U
 U
 U
 V
 V
 U
 V
 U
 U
 U
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 V
 Zea mays A GGU UAC AG 5'- UGACAGA AGAGAGUGAGCAC CACGGU UCCU CAUG 3' - ACUGUCU UCUCUCACUCGUG GUGUCG AGGA GUAC U A C ---- UC- UG B. miRNA 157 mdii Gossypium raim 5. Helianthus annuus
 Heitannus annuus
 AG
 A
 CA

 U
 AGAAGAGACA
 AG
 G
 CAU

 s' ACUGUCUUC UAUCUCUCGUG UC UC ACUUAACGU
 A

 U
 G
 CU
 CUUAACGU
 U GAG AGACLAA COLL ||||||||||||||||||||||||||UUUGGUU GAGGAAUA U ACC $\begin{array}{c} Lycopersion esculentum \\ U C & G & G & C & CUU \\ S' \cdot UGA & AG & AA AUAGA & AG & AUG & A \\ 11111 & 111111 & 11111 & A \\ 3' \cdot ACU & IC & UU & VAUU & UU & VAC & C \\ G & -G & G & G & C & ACC \\ \end{array}$ Nicotiana benhamiana C AG CA CU C -CAAUGAA - UGAGAAG U UAG UG GGG A 5'- UGCU CUU UUUC UUAAA CG GG GCU UGAU GG G CACCAU AAUC G IIII IIII IIII IIII IIII 3'- ACGA GAG GAAG AOUUU GC CC CGA ACUA CC U GUGGUA UUGG C GA A- AC A Gtraibeng U UUGUA- C UA GU AG A Nicotiana tabacum
- Solanum tuberosum
- Solationin funerossium
 U
 U
 AUUUUCA U

 5° UGACAGAAG AUAGAGGACCA AAUGA UGA AUGAU
 UGA AUGAU
 UGA AUGAU

 3° ACUGUCUUC UAUCUCUCGUG UUACU ACU UACCA
 U
 U

 G
 U
 C
 C

C. miRNA 159

F-UGGAGEDECUL AGUCAA GAGG UACU F-ULULULULU LULULULULULU F-ADUUDAAUGA ULUARDUU LUDE AUGA GGG AAU GAGCU CUUAG UAUGGA CCACAG CUACC CA

P. GRACETEU UCA UCCAA AG GOLD CG AGGG UU GC ACU CULG UCAU P. GRACETEU UCA UCCAA AG GOLD CG AGGG UU GC GC CULG UCAU P. LEUCUAUGA AU AGGUU UC UCCAA GC UCCE AG GAGC AGUA P. LEUCUAUGA AU AGGUU UC UCCAA GC UCCE AG GAGC AGUA JU UCA G UCC - U UU G U

Medicana truncatula CUACC CA

COMOCOLA GALANA CONTRACTOR 1111 Pinus taeda

00 Schedonorus arundinaceus

5- GENERGENE AL U LECAL GA GAU CU LIAGA - ACA G U LICE LICE U LIC 1111 GGUG

D. miRNA 160

Glycine max UGGCUCC GUAUGCCAUU UAGAG UCAU GA G 5.00 3' - ACG ACCGAGG UAUGCGGUAG GUUUC AGUA CU C A AG GUUC AGUA CU C Orvra sativa

Picea glauca

 $\begin{array}{cccc} Triticum turgidum \\ C & C & C \\ s^{*} \cdot UCC & UCC & UUAUGCCA & CAUC & CA & CC & U \\ s^{*} \cdot UCC & UCC & UUAUGCCA & CAUC & CAG & CA & ACC & U \\ A & U & C & AG & C & C & U \\ A & U & C & AG & C & C & U \\ \end{array}$

Zea may C 3'- LIGE LIGE LIGUAUGCAC CAU GUA CCAA CG 111 111 111 111 111 111 3'- ACG ACCGAGG ACGUGCCGUG GUG COU GGUU GU A A A G U A A-GC

E. miRNA 162

Glycine max C C UC C AA GU 5'-CUGGAUGCAG GGUU AUCGAUC UUC UG UC 5'-CUGGAUGCAG GGUU AUCGAUC UUC UG UC UG 3'-GACCUACGUC CCAA UAGCUAG AAG AC AG AC U A CU U CA ACU A

Lupinus luteus

 Malux x domestica (cultivated apple)
 AUUU
 GU
 AUUU
 GU
 GU
 AUUU
 GU
 GU
 GU
 AUUU
 GU
 G

F. miRNA 163

F. III.CAR 105 F. Cardana Cu. ACUGU CUG GG C. C. U- UG S. GAGGAA CU. GC CGAC GGAUUG UGGU UGC GA U S. CUUCUG GAC GCG GCGGG CCUGA CUC AA A. CU. ACCU- AA- AA U. UC UC

Pinus taeda

	A GA	C	+UUGGA	UG	GUUA	G	-UCAGCAGA	- A	A
5' -	UGA GAA	UUGGAGC UUCGAC	c	G	A C	A GAAGGAC/	GC	AC	U
	111111	1111111 111111		1	11 1	1.11111111	11	11	C
3' -	ACU CUU	AAUUUUG AAGUUG	G	C	JU G	U UUUCUUG	J C0	UG	C
	A A/	4- U	(Million slide more bound)	GU	AGUA	G	(12 m side men base)	C C	C

G. miRNA 164

G INICIAN LIGHT Triticum aestheum GAAGCA- CA U G GGG S-UGGA GGU CG GCGCC CC U 1111 111 111 111 11 3'-ACCU CCG GC CCCGG GG C GUACAGG AG - A AAG

H. miRNA 165

H. miKNA 105 Hedyotis centramboides A A UU G C C U-S - GOGA UG GUCUGGU CGA GGC AC AC AGAUCC GGAAUCCUUC J - LILL I LILLII III III IIII IIIII S - CCCCU AC CGOACCA CCU CGG UG UG UG UCAGG C - UU G G - A CGU UG UCUUGGAGA A C - UU G G - A CGU UT MANAGE UUUUCCUUUUCUAG

Cell Research, 15(5):336-360, May 2005 | www.cell-research.com

Bao Hong ZHANG et al

I. miRNA 166 Ipomoca nil s-ocoogna cu de c Medicago trancatula UU - CU UC - AGG UU 5'-GGGGAAUG GGE UGG CGAGGCUUU AGUUU ACAA AAG C 3'-CCCCUUAC UGG ACC GCUUCGGAA UCAAG UGUU UUC A U- G AG UA UA UA AA. UG
 Saccharum officinarum
 U
 AC
 GC
 UCC
 UCC
 UGC
 UGC
 UGG
 GG
 CGG
 Sorghum bicolor Sorgium bucou U U CUUGU A CUUGU - UUUG A 5'-G GGAAUG GUCUGGGUUC AGGUCUCG GAUUUAAGGA UGA UGC U 3'-C CUUAC COGACCAGG UCUAGAGU CUUAUU AUG G C UU C AU- U UUA C Zea mays J. miRNA 167 Arabidopsis thaliana Glycine max A - CUUU GGGA GA 5'-UGA GCU GC CAGCAUGAUCUAG GGUUAGU GC G Phaseolus coccineus K. miRNA 168 Glycine max L. miRNA 169

www.cell-research.com | Cell Research, 15(5):336-360, May 2005

M. miRNA 170 Allium cepa Arabidopsis thaliana S' - GAUAUUGGE UGGUUCA UCAGAU
 S' - GAUAUUGGE UGGUUCA UCAGAU
 UUAC UG
 C
 U
 C
 U
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 C
 Glycine max UC A AU 5'-GAUAUUGG CGUUCAAU AGAAAGCA GCUCAAAA \ 1 1 3'-CUAUAACC GCCGAGUUA UUUUUUGU UGGGUUUU G GU C CC U Hedyotis centranthoides C UG UAAC U GUAA 5'- GAU UGG COGUUCAAU GG GGUUG GGUC C 111 1111 111111 11 111 3'- CUA AACC GCGAGUUA CU UUAAC CCGG A U GU G UU C AAUU Hordeum vulgare subsp. vulgare Oryza sativa S*-GAUGUUGGC CGGCUCA UCAGA GGCAU GUGA GC AAGCAUG A 3*-CUAUAACCG GCCGAGU AGUCU UUGUG CACU CG UUCGUGC U U U - AUCG G Orvza sativa

N. miRNA 171 Allium cepa Alliam cepa 5' - UGAUU AGCUG GUCAGU UUCA GGC GCU 11111 1111 1111 1111 111 111 11 3' - ACUAA UCGGC GGGUUA GAGU CGG CGG GG G UA GC AUAUUAUAA AU Hordeum vulgare subsp. Vulgare U A C - A G G A S - GAUGUUGGC CC CUCK UCKAC CC CGCCC A G G HIHIHI I HIHIHI I HIHIHI S - CUAUAACCG G G GAUU AGCU GG G GGGC UCU C U C U U CG G AC

Trancum destrutum
 U A C ----U---- G
 S'-GAUGUUGGC CG CUCA UCAGA CCAC GCC GCC
 GCC G
 ULUUUUUUUUUUUUUUUUUUUU
 S'-CUAUNAACCG GC GAGU AGUCU GGUG UGG
 U C U C G C Intraketogo C

0 miRNA 172

Hordeum vulgare subsp. vulgare G GU GU UU CGGCG U 3°- ACQUC AGUA CGUANG AGGU UGU CCG C 1111 111 111 111 111 3°- UACAG UCAU GAAUCC UCCA ACA GGC G A CU (Itherationer) A Lycopersicon esculentum 5'- JUG GCAU AUCAAGAUUC CGUGA GUUU CAAA U GUUAU UAAUU 3'- JUG CGUU AUCAAGAUUC CGUGA GUUU CAAA U GUUAU UAAUU 3'- UAC CGUA UAGUUCUAAG GCACL CAA GUUU G CGGUA AUUAA U GU G A AUC GU A AUC GU Solanum tuberosum AUAAGU
 AUAAGU
 AUAAGU
 AUAAGU
 AGAGUGUAGUAGUUCUAAG GUAC
 UUUUA ACCCG AUA
 AGAGUGUAGUAGUUCUAAG AUA
 AAAU UGAAC UAU
 AAGU
 AAAU UGAAC UAU
 AAGU
 AAAU GAACAGAUAAAAU AAAU AGAAC AUA
 AA
 GAA

Vitis vinifera

T CANE THE SECOND A AA AA O DO LES COCCOMMA CAMADA DO CAMADA AND C

P. miRNA 319
Glycine max A UC U AG AG G AC UC AA 5°-G GAGCU CUUCAGUCCA AUGGG GAC UAAGAUUCAAUU CU CCG UCAUUA CA UGUUGAGUGUA A 3°-C CUCCA GAAGUCAGGU UGUCC UUG AUUCUAAGUUAA GA GGC AGUAAGU GU ACGACUCAUIAU A C GGU GA AG AAU
Liriodendron tulipifera
A UCCA U UA UCUCC G AC C U U CAAC CUA 5°- GGAGCUCUCUUCAGUCU G AGAUGA GU AGCAAA GCU CUG UCGUUGG UA GAAC CAUC GG G 5°- CCUCGAAGGGAAGUCAGUU UCUGU UCUG UUGGUU CGA GGC AGCAACC AGU UUUG GUAG CCC G C CUCC - UC UGUCC G GU U U U UUA
Triticum aestivum
C ALLOCATION OF A AMA GA AMI C - G ALLOCATION C UNACAU AND C - GO ACCOUNT ALLO AND C - GO ACCOUNT ALLO AND C - CO CO CO CHECKELE CORRECT CORRECT ALLO C - C - C - C - C - C - C - C - C - C
Vitis vinifera
G C CUCCAUC - AAGA
5' - GGGGGUUCCUUU CAG CCAAAACAA GC UGAAG U
3' - UCCUCGAGGGAA GUC GGUUUUGUU CG ACUUC A
- A CCU U AAGU
Physcomitrella patens subsp. patens
U GU U UG CGA CC
5'- GGAGCUCC UUUCGGUCCAA AG GCUG CGGAAGGUGGU C
3'- CCUCGAGG GAAGUCAGGUU UC CGGC GCCUUCCA CCG G
A C UG CC- (76nt side loop) UC

Fig. 3 Examples of the predicated stem-loop structure of newly identified plant miRNA precursors using EST analysis. The miRNA sequences are in grey. The actual size of the precursors may be slightly longer. (A) miRNA 156; (B) miRNA 157; (C) miRNA 159; (D) miRNA 160; (E) miRNA 162; (F) miRNA 163; (G) miRNA 164; (H) miRNA 165; (I) miRNA 166; (J) miRNA 167; (K) miRNA 168; (L) miRNA 169; (M) miRNA 170; (N) miRNA 171; (0) miRNA 172; (P) miRNA 319.

The next step in this study was to predict the secondary structure of all 812 chosen ESTs with 0-2 mismatched nucleotides with previously known Arabidopsis miRNAs, and to identify the new miRNAs and their targeted genes based on the main represented characteristics of miRNAs and their targeted genes. After individually checking and scoring the secondary structure of the 812 chosen ESTs, 464 EST contigs were found to have the patterns of secondary structure conservation resembling those of previously known plant miRNAs (Fig. 3). These 464 ESTs were grouped into 17 families and represented 338 new potential miRNAs in 60 plant species based on their nucleotide sequence similarity as determined by Blastn (Tab. 1 and 3). The other 348 chosen ESTs may be the targets of our newly identified or previously known miRNAs (Tab. 2 and 3).

Out of 171 ESTs without mismatched nucleotides that were compared with previously known *Arabidopsis* miRNAs, 72.5% of ESTs (124) were predicted to have the secondary structure of miRNA and were identified as new potential miRNAs. Only 53.4% and 52.2% of ESTs had the secondary structure of miRNAs and were identified as new potential miRNAs in the ESTs with 1 or 2 nucleotide substitutions relative to their *Arabidopsis* miRNA homologs. This also indicated that plant miRNAs are highly conserved.

To avoid designating small RNAs or fragments of other RNAs as miRNAs, Ambros et al. (2003) developed a combination of criteria to identify and annotate new miRNAs [35]. The combination of criteria includes both expression criteria and biogenesis criteria. The potential miRNAs identified by EST analysis were highly conserved with the known *Arabidopsis* miRNAs, and fit the miRNA criteria of secondary struture. Thus, the newly identified potential miRNAs fit the biogenesis criteria well. ESTs are partial cDNA sequences of expressed genes. If this was counted to expression criteria, all new identified potential miRNAs should be true miRNAs. However, to avoid the possibility of artefactual ESTs [36], a Northern blotting method should be employed to test these new potential miRNAs. However, these newly identified gene sequences are still very good miRNA candidates.

In this study, we identified a total of 464 ESTs containing new potential plant miRNAs. Out of those ESTs containing potential miRNAs, 98, the highest number, belong to the miRNA 156 family; 53 belong to the miRNA 157 family; and 48 belong to the miRNA 172 family. miRNAs 156, 157, 159, and 172 are four well studied plant miRNAs. They have important functions in plant development, especially in leaf and flower development. The expression of miRNAs 161 and 163, and their targeted genes are unclear. We did not find any ESTs with 0-2 base substitutions related to these two Arabidopsis miRNAs. However, we found 31 ESTs with 20 matched nucleotides and 4 nucleotide substitutions related to Arabidopsis miRNA 163, and 12 of the 31 ESTs were predicted to have the secondary structure of miRNA, and were identified as new miRNAs. These 12 ESTs expressed 2 miRNAs and belonged to 3 plant species: Picea glauca, Picea sitchensis, and Pinus taeda.

To avoid ESTs created from the same miRNA precursor, ESTs containing the same length hairpin structure were blasted in dbEST against each other. If the ESTs have high similarity, it indicates that these ESTs were created from the same miRNA sequence and they should be considered as one miRNA. After removing the homezygotic ESTs containing the same miRNA, a total of 338 new potential miRNAs were identified in 60 plant species (Fig. 3). The highest number of new potential miRNAs, 45, was found in soybean (*Glycine max*), followed by rice (*Oryza sativa*) with 25, sugarcane (*Saccharum officinarum*) with 25, maize (*Zea mays*) with 23, and wheat (*Triticum aestivum*) with 19 (Tab. 4). All of these plant species have more than 200,000 EST sequences in the GenBank's EST databases.

Some microRNAs may be induced and regulated by environmental stress; some preferentially express in specific tissues and may be regulated by developmental switching

	miRNA	Species	miRNA sequence*	Length	Precurs	EST	EST	Stra	Folb	miR		precu	
iRNA	gene			(nt)	orhairpi	accession	length	nd	ack	Start	End	Start	E
mily					n length (nt)	number			arm				
56	AtMiR 156m	Arabidopsis thaliana	ugacagaagagagagagcac	20	268	AI999403	498	+	3'	461	480	213	4
	AtMiR 1560	-			245	AV805456	409	+	3'	377	396	152	
	AtMiR 156p				47	AV799188	407	+	5'	326	345	326	
	AtMiR 156q				59	AV558614	572	+	5'	500	519	500	
	AtMiR 156r				326	AV537797	607	+	5'	215	234	215	
	AtMiR 156s				218	AI995421	552	+	5'	223	242	223	
	BnMiR 156	Brassica napus	ugacagaagagagugagcac	20	81	CD835236	694	+	5'	102	121	102	
	CaMiR 156a	Capsicum annuum	ugacagaagagagagagcac	20	152	CA516324	508	-	3'	417	436	285	
	CaMiR 156b				65	BM064926	474	-	5'	36	55	36	
	CsMiR 156a	Citrus sinensis	ugacagaagagagagagcac	20	56	CK665702	605	-	3'	202	221	166	
	CsMiR 156b				515	CF836044	785	+	3'	558	577	63	
	CrMiR 156	Cycas rumphii	ugacagaagagagagagcac	20	67	CB091817	671	-	5'	221	240	221	
	EcMiR 156	Eschscholzia californica	ugacagaagagagagagcac	20	150	CD481596	691	-	3'	348	367	218	
	GmMiR 156a	Glycine max	ugacagaagagagugagcac	20	83	BE807821	482	+	5'	163	182	163	
	GmMiR 156b		ugacagaagagagagagcac	20	147	BM731329	421	-	5'	4	23	4	
	GmMiR 156c				185	BM731193	504	-	5'	67	86	67	
	GmMiR 156d				76	BI470983	459	-	3'	109	128	53	
	GmMiR 156e				55	BI425943	559	-	3'	177	196	142	
	GmMiR 156f				267	BG651519	486	-	3'	373	392	126	
	GmMiR 156g				209	BG507682	452	•	5'	77	96	77	
	GmMiR 156h				147	BE806824	416	-	5'	117	136	117	
	GmMiR 156i				117	AW734915		•	3'	174	193	77	
	HaMiR 156	Helianthus annuus	ugacagaagagagagagcac	20	362	BQ968177	575	-	5'	354	373	12	
	HsMiR 156	Hordeum vulgare subsp. spontaneum	ugacagaagagagagagcac	20	178	BJ478354	418	-	3'	325	344	167	
	HvMiR 156a	Hordeum vulgare subsp. vulgare	ugacagaagagagugagcac	20	394	BG300360	624	+	5'	231	250	231	
	HvMiR 156b		ugacagaagagagagagagcac	20	86	CA032492	647	-	5'	523	542	523	
	HvMiR 156c				178	CA020690	515	-	5'	439	458	281	
	HvMiR 156d				203	BF258419	902	-	3'	184	203	1	
	LsMiR 156	Lactuca sativa	ugacagaagagagagagagcac	20	355	BQ874863	654	-	5'	533	552	198	
	LjMiR 156	Lotus corniculatus var. japonicus	ugacagaagagagagagagcac	20	47	BP051612	510	+	5'	362	381	362	
	LeMiR 156a	Lycopersicon esculentum	ugacagaagagagagagagcac	20	116	BF097000	482	-	3'	285	304	189	
	LeMiR 156b				47	BG125997	589	•	5'	25	44	25	
	LeMiR 156c	M I I I		20	263	BE436147	597	•	5'	21	40	21	
	MdMiR 156a	Malus x domestica	ugacagaagagagagagcac	20	109	CN580870	505	•	5'	138	157	138	
	MdMiR 156b			20	79	CN578779	538	-	3'	166	185	107	
	NbMiR 156a	Nicotiana benthamiana	ugacagaagagagagagcac	20	205	CK294165	812	+	5'	516	535	516	
	NbMiR 156b	0		20	169	CK290419	610	+	3'	508	527	359	
	OsMiR 156m	Oryza sativa	ugacagaagagagugagcac	20	86	CB677777 CB677501	734 776	+	5' 5'	262 260	281 279	262	
	OsMiR 156m						742	+	5'			260	
	OsMiR 156m					CB653304 CB647907		+	5' 5'	260	279	260	
	OsMiR 156m					CB647907 CB643501	722 841	+	5'	269 259	288 278	269 259	
	OsMiR 156m OsMiR 156m					AU091537	841 714	+ +	5'	259	278	239	
	OsMiR 156m OsMiR 156m					D41985	412	+	5'	286	305	286	
	OsMiR 156n OsMiR 156n				97	CB674083	412	+	5'	322	341	322	
	OsMiR 1560 OsMiR 1560		100000000000000000000000000000000000000	20	179	CB674083 CB675101	270	+	3'	212	231	53	
	OsMiR 1560 OsMiR 156p		ugacagaagagagugagcau ugacagaagagagugagcuc	20	133	CB674957	497	+	3'	333	352	220	
	OsMiR 156p OsMiR 156q		ugacagaagagagugagcuc	20	243	CB632048	296	+	3,	237	256	14	
	OsMiR 156q OsMiR 156r		ugacagaacagagugagcac	20	24 <i>3</i> 83	AU162331	296 469		3'	398	417	335	
	PgMiR 156	Picea glauca	ugacagaagagagagagagagcac	20	115	CK441500	682	+	3,	627	646	532	
	PtMiR 156a	Pinus taeda	ugacagaagagagagagagagcac	20	160	CF667730	787		3'	146	165	6	
	PtMiR 156b		"Prendundu Dudindu Derre		221	BQ696284	578		3'	394	413	193	
	PpcMiR 156	Populus balsamifera x Populus deltoids	ugacagaagagagagagcac	20	146	CN519769	797		5'	336	355	336	
	PtMiR 156	Populus tremula	ugacagaagagagagagagcac	20	373	BU818933	577		5'	42	61	42	
	PptMiR 156a	Populus tremula x Populus tremuloides	ugacagaagagagagugagcgc	20	46	CK092224	827	+	5'	414	433	414	
	PptMiR 156b		ugacagaagagagagagagcac	20	356	CK087760	815	+	3'	455	474	119	
	PptMiR 156c				85	BU833512	555		3'	284	303	219	
	PaMiR 156	Prunus armeniaca	ugacagaagagagagagcac	20	114	CB819319	504	+	5'	137	156	137	
	ScMiR 156a	Saccharum officinarum	ugacagaagagagagugagcac	20	101	CA294779	738	+	5'	198	217	198	
	ScMiR 156b		ugacagaagagagagagagagcac	20	115	CA086406	691	+	3'	657	676	462	
	ScMiR 156b		000.0.0.0.0			CA210204	735		3'	212	231	113	
	ScMiR 156c				68	CA228340	637		3,	391	410	343	
	ScMiR 156d				327	CA182412	563	2	5'	18	37	18	
	ScMiR 156e				70	CA148934	689	+	3'	512	531	463	
	ScMiR 156f				46	CA091338	780	2	3'	163	182	137	
	ScMiR 156g				403	CA186028	582	-	3'	443	462	60	
	ScMiR 156g ScMiR 156h				343	CA078294	1007	-	5'	28	402	28	
	Sentity 12011				0.10	211010274	1001	-	*	£0		20	

Tab. 3 New	miRNAs	identified	from	plants

www.cell-research.com | Cell Research, 15(5):336-360, May 2005

	miRNA	Species	miRNA sequence*		n Precurs		EST	Stra		miR		Preca	
	gene			(nt)	-	accession	lengtl	n nd	ack	Start	End	Start	E
amily					n length	number			arm				
	SiMiR 156	Sesamum indicum		20	(nt) 81	BU667823	289	+	5'	43	62	43	1
	StMiR 156a	Solanum tuberosum	ugacagaagagagagagagcac ugacagaagagagugagcac	20	86	BU007823 BI432985	326	+	5, 5,	43 53	72	43 53	1
	StMiR 156b	Solanum tuberosum	ugacagaagagagagagagagcac	20	382	BQ506362	794	+	3,	594	613	232	6
	StMiR 156c		"But "But Bub Bub Bub		85	CK272946	832	+	5'	582	601	582	6
	StMiR 156d				171	CK262978	864	+	5'	387	406	387	5
	StMiR 156e				72	BQ111685	507	+	5'	423	442	423	
	SbMiR 156a	Sorghum bicolor	ugacagaagagagugagcac	20	84	CN132665	630	+	5'	306	325	306	
	SbMiR 156b	0	ugacagaagagagagagagcac	20	450	BG947368	670	-	5'	78	97	78	
	SbMiR 156b					BG947367	618	-	5'	26	45	26	
	SbMiR 156c				395	CD230908	652	-	5'	130	149	130	
	SbMiR 156d				102	AW679616	690	-	3'	89	108	7	
	SpMiR 156	Sorghum propinquum	ugacagaagagagagagcac	20	125	BF421158	425	-	3'	373	392	268	
	TaMiR 156a	Triticum aestivum	ugacagaagagagagagcac	20	52	BJ322639	640	+	5'	417	436	417	
	TaMiR 156a					CD454302	640	+	5'	474	493	474	
	TaMiR 156b				416	BJ317100	623	-	3'	491	510	95	
	VvMiR 156	Vitis vinifera	ugacagaagagagagagagcac	20	47	CD800420	538	+	5'	457	476	457	
	ZmMiR 156m	Zea mays	ugacagaagagagugagcac	20	83	CF039276	479	+	5'	299	318	299	
	ZmMiR 156m					CF037245	511		3'	331	350	268	
	ZmMiR 156m					CF036982	592	+	5'	247	266	247	
	ZmMiR 156m					CF036852	598	+	5'	224	243	224	
	ZmMiR 156m					CF036471	579	+	5'	223	242	223	
	ZmMiR 156m					CF036182	372	+	5'	226	245	226	
	ZmMiR 156m					CF035885	553	+	5'	87	106	87	
	ZmMiR 156n				84	CF059465	516	+	5'	221	240	221	
	ZmMiR 1560				99	CF036467	383	+	5'	272	291	272	
	ZmMiR 156p		ugacagaagagagagagagcac	20	189	CA408772	619	-	5'	304	323	304	
	ZmMiR 156q					CA408687	636	-	5'	304	323	304	
	ZmMiR 156r		ugacagaagagagugggcac	20	99	CF035522	424	+	5'	274	293	274	
	ZeMiR 156	Zinnia elegans	ugacagaagagagagagcac	20	67	AU293485	732	+	5'	343	362	343	
	AtMiR 157m	Arabidopsis thaliana	uugacagaagagagagagcac	21	50	AV784285	669	+	3'	339	359	310	
	CaMiR 157	Capsicum annuum	uugacagaagagagagagcac	21	155	CA516324	508	-	3'	417	437	283	
	GmMiR 157m	Glycine max	uugacagaagauagagagcac	21	261	BI699573	557	-	5'	223	243	223	
	GmMiR 157n				59	AW459710	417	-	5'	184	204	184	
	GmMiR 157o				57	BE611678	464	-	5'	296	316	296	
	GmMiR 157p				85	BE210632	313	+	5'	207	227	207	
	GmMiR 157q				112	AW756919	420	+	5'	150	170	150	
	GmMiR 157r		uugacagaagagagagagcac	21	147	BM731329	421	-	5'	4	24	4	
	GmMiR 157s				268	BG651519	486	-	3'	373	393	126	
	GrMiR 157a	Gossypium raimondii	uugacagaagauagagagcac	21	83	CO082782	891	+	5'	197	217	197	
	GrMiR 157b				92	CO076888	848	+	5'	169	189	169	
	GrMiR 157c		uugacagaagagagagagcac	21	183	CO105970	783	+	5'	400	420	400	
	GrMiR 157d				89	CO097436	785	+	3'	570	590	502	
	HaMiR 157a	Helianthus annuus	uugacagaagauagagagcac	21	86	BQ916415	492	+	5'	146	166	146	
	HaMiR 157b	Helianthus annuus	uugacagaagagagagagcac	21	364	BQ968177	575	-	3'	354	374	11	
	IbMiR 157	Ipomoea batatas	uugacagaagauagagagcau	21	283	CB330750	758	-	3'	546	566	284	
	InMiR 157a	Ipomoea nil	uugacagaagauagagagcau	21	126	BJ578566	713	+	5'	131	151	131	
	InMiR 157a					BJ578544	727	+	5'	131	151	131	
	InMiR 157b	Ipomoea nil	uugacagaagauagagagcau	21	183	BJ568024	542	+	5'	126	146	126	
	InMiR 157b					BJ573618	495	+	5'	131	151	131	
	InMiR 157b					BJ569965	714	+	5'	116	136	116	
	InMiR 157b					BJ563000	591	-	3'	563	583	401	
	InMiR 157b					BJ556392	592	-	3'	563	583	401	
	LsMiR 157	Lactuca sativa	uugacagaagagagagagagcac	21	357	BQ874863	654	-	3'	533	553	197	
	LjMiR 157a	Lotus corniculatus var. japonicus	uugacagaagauagagagcac	21	58	BP039685	546	+	5'	417	737	417	
	LjMiR 157a					BP030242	518	+	5'	396	416	396	
	LjMiR 157b	Lotus corniculatus var. japonicus	uugacagaagagagagagcac	21	86	BP046017	503	-	3'	413	433	348	
	LjMiR 157c		uugacagaagauagagagcau	21	460	CN825561	660	-	3'	556	576	117	
	LeMiR 157a	Lycopersicon esculentum	uugacagaagauagagagcau	21	53	BM536323	331	+	5'	200	220	200	
	LeMiR 157b		uugacagaagauagagagcau	21	430	BE435668	620	-	3'	512	532	103	
	LeMiR 157b					AW933950	633	-	3'	488	508	76	
	LeMiR 157c				123	BE433988	239	+	3'	180	200	78	
	MtMiR 157	Medicago truncatula	uugacagaagauagagggcac	21	111	AW696064	665	+	5'	195	215	195	
	NbMiR 157a	Nicotiana benthamiana	uugacagaagagagagagagcac	21	169	CK290419	610	+	3'	507	527	359	
	NbMiR 157b				207	CK294165	812	+	5'	515	535	515	
	NbMiR 157b					CK284079	812	+	5'	521	541	521	
	NtMiR 157	Nicotiana tabacum	uugacagaagauagagagcac	21	131	BP129308	705	+	5'	301	321	301	
	NaMiR 157a	Nuphar advena	uugacagaagacagagagcac	21	140	CK762856	565	-	5'	312	332	312	

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

martygencmartymartysetsetmartysetmarty<			Tab. 3 New miRN	As identified fro	m pla	ants (continu	ed-2)					
Imageand toNumber10Cristits3031			Species	miRNA sequence*	-									
Nadkii 157 Parear american approximation approxima		gene			(nt)	-		length	nd		Start	End	Start	End
NARR 157b (b) (c) (Tamily					-	number			arm				
PARE 157Personersonmetar Personersonmetar Personerson Persone Personerson Pe		NaMiR 157b					CK745155	655	+	5'	441	461	441	631
PpokePhace non-scrept with environmentageungescapacy inpropersion of a point or main strept with environmentagePhace Name Phate Name		NaMiR 157c				69	CD472699	521	-	3'	258	278	210	278
PANE 157Purpose nonconsiz Popular encodesupper seganging pages2136000000000000000000000000000000000		PaMiR 157	Persea americana	uugacagaagagagagagcac	21	216	CK751799	634	-	3'	237	257	42	257
PartP		PcMiR 157	Phaseolus coccineus	uugacagaagauagagagcau	21	184	CA902638	612	-	3'	480	500	317	500
SARR 157Socumu nalowaunpercapangingapase218810010010 </td <td></td> <td>-</td> <td></td>		-												
Norm Soluti 157														
SARE 157. BUR1371 Gall 400 4 5 1 7 7 5 5 1 7 1														
SARE 17% SARE 17% SARE 15%mgacagaagiiguigagas and 16CACMAGE83848484SARE 157 ZamMR 157m ZamMR 157mZoraganagiiguigagas ZamMR 157m16CACMAGE6101.68.08.08.08.0ZamMR 157m ZamMR 157mZoraganagiiguigagas ZamMR 157m2Modeland61.68.0			Solanum tuberosum	uugacagaagauagagagcac	21	91								
Solke 157: Image: accass acciss glig lig gases: 1 66 1 50 10 50 10 50 10 50 10 50 10				1111032303303030303030232	21	617								
Sordpunctionungacagangingingages2186CN11360609550101010ZaMR 157Zamat Gramungacagangingingages21661055111					21									
ZavRR 157m CAUGAGY 0.000 could of the section of the s			Sorghum bicolor	uugacagaagagagugagcac	21	86		630	+	5'	305	325	305	390
2.80R. 1572.000000000000000000000000000000000000		ZmMiR 157m	Zea mays	uugacagaagagagagagcac	21	189	CA408772	619	-	5'	304	324	304	492
158Gamk 158 a // conjension excluentumaccaanguageacaage inic caanguageacaage20162H00000010000010 <td></td> <td>ZmMiR 157m</td> <td></td> <td></td> <td></td> <td></td> <td>CA408687</td> <td>636</td> <td>-</td> <td>5'</td> <td>304</td> <td>324</td> <td>304</td> <td>492</td>		ZmMiR 157m					CA408687	636	-	5'	304	324	304	492
Lexinx 158. Leconresion exculentum Bioecaangean age 20 50 M19947 51 - 5 4.7 4.8 75. 1099577 627 4.9 75. 75. 1099577 627 4.9 75.		ZeMiR 157	Zinnia elegans	uugacagaagagagagagcac	21	69	AU293485	732	+	5'	342	362	342	410
LAMR 158 Phys.com/rel.p aters: ucceasalguiliacaagă 20 533 B003773 627 + 3" 154 67 SMIR 158 Solumin horosum uucceasalguiliacaată 20 533 B0024645 664 - 5" 57 42 25 159 Gindin 159 Glocien na: uuugganugaaggagecucu 21 175 BM39181 51 4" 5" 12 22 22 22 23 36 H-MIR 159 Hordeam valgare subg. valgare uuugganugaaggagecucu 21 148 8145581 50" 1 2 1 35 LMIR 159 Loticadoron nalipifere uuugganugaaggagecucu 21 148 8102449 70" + 3" 481 50 49 42 14 42 OMIR 159 Dorasitiv uuugganugaaggagecucu 21 249 CR39947 70" + 3" 401 42 42 424 42 PMIR 159 Points indu uuugganugaaggagecucu 21 198 Col161944 40" 5" 50	158								+					
PpMR 158Pycowintrelia paternucccanaguaiiiacangaiCoSSS </td <td></td> <td></td> <td>Lycopersicon esculentum</td> <td>auccaaauguagacaaagca</td> <td>20</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>			Lycopersicon esculentum	auccaaauguagacaaagca	20									
NMRNMRSolution <th< td=""><td></td><td></td><td>Discoursite all a sector</td><td></td><td>20</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></th<>			Discoursite all a sector		20									
19 GmAiR 159 Glycine max uungganungagggage.cui 21 175 MM89318 51 + 3' 2' 2' 2' 3' Alf7566 4'' 3' 2'' 2'' 2'' 3'' Alf7566 4'' 3'' 2'' 1''' </td <td></td> <td>-</td> <td></td>		-												
HvMRiR 159a Hordeum vulgare uuuggauugangagagacucui 21 303 Al47506 420 2 2 2 23 5 LMRIR 159 Hordeum vulgare subsp. vulgare uuuggauugaagagagacucu 21 145 B444850 62 4 5 1 21 145 5 1 21 145 5 4 5 1 21 14 50 50 66 4 3 407 50	159						-							
HvMRI 159 Hordeum vulgare suby, vulgare uuugganugangagagecucu 21 174 B/44559 622 2 376 <td>139</td> <td></td> <td>,</td> <td></td>	139		,											
HvMiR 199 Example and anglamagaaggacucu 21 345 57 202 222 202 375 LMM R 159 Liroidendron hulpifera muaggamagaggacucu 21 345 CK75993 345 4 5 1 21 13 MMMR 159 Medicacy macuality muaggamagagggacucu 21 249 CR39030 461 + 3' 369 399 389 PMM 159 Permisetum glaucum muaggamagagggacucu 21 226 CO75199 784 + 3' 411 420 4			•											
MikiR 159 Medicago runcutula uuugganuganggageucu 21 168 Bl265403 66 + 3' 451 501 344 501 MikiR 159 Oryas saftva uuugganuganggageucu 2 2420 CF80907 707 + 3' 420 80 707 + 3' 420 80 707 + 3' 420 80 707 + 3' 420 420 80 707 + 3' 420 420 80 707 + 3' 420 420 80 707 + 3' 421 60 707 + 3' 421 420 80 707 + 3' 421 420 80 707 + 3' 421 420 80 707 481 501 700 401 40														375
MikiR 159 Overa sative uunggaungaaggageucui 21 24 CRX39947 70 + 3' 476 496 229 496 OsMiR 159m Overa sative uunggaungaaggageucui 21 24 CRX39947 70 + 3' 401 451 35 389 380 381 383 381 383 383 383 381 383 384 383 383 383 384 383 384 383 383 384 383 383 383 384 383		LtMiR 159	Liriodendron tulipifera	uuuggauugaagggagcucua	21	345	CK750935	345	+	5'	1	21	1	345
OxMRR 159m Oryza sativa unuggauugaaggageucu 21 24 24 24 42 12 28 PgMR 159 Peninstem glaucum uuuggauugaaggageucu 21 22 12 26 60 40 4 3<		MtMiR 159	Medicago truncatula	uuuggauugaagggagcucua	21	168	BI265403	566	+	3'	481	501	334	501
OxMiR 159m CR30003 461 + 3 369 389 138 389 PsMR 159 Pennisemm glaucum uuuggauugaaggaggeucu 21 226 CD22466 804 - 5 560 550 70 50 50 40 10		MtMiR 159					AW691937	666	+	3'	476	496	329	496
PgMiR 159 Pennisetum glaucam uuuggauugaaggagucucu 21 226 CD25199 784 + 3' 41 45 5 60 50		OsMiR 159m	Oryza sativa	uuuggauugaagggagcucug	21	249	CR289947	707	+	3'	402	422	174	422
PMIR 159 Picea sitchemsis euugganugaaggageucka 21 193 CO22662 804 - 5 60 580 60 752 PMIR 159 Pinns tacda euugganugaaggageucka 21 193 CO16631 80 - 5 60 53 53 53 PMIR 159 Schift 159a Schift 159a Schift 159a Schift 159a - 53 43 43 33 35 13 337 357 132 35 Schift 159a Schift 159a Schift 159a Schift 159a Schift 159a - 37 43 43 37 <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>+</td> <td></td> <td></td> <td></td> <td></td> <td></td>									+					
PMiR 159 Pinus saeda eungganugaaggageweka 21 198 C01661954 420 + 3' 215 235 38 235 PMIR 159 Sextiri 159a Saccharum officinarum uuugganugaaggageweka 21 226 CA26132 540 + 3' 323 343 118 343 SeMiR 159a Saccharum officinarum uuugganugaaggageweka 21 226 CA26132 540 + 3' 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 188 343 183 343 183 343 183 343 183 343 183 353 35 345 135 353 35 356 35 356 36 36 36 36 36 36 36 36<		-	-											
PMiR 159 Saccharum officinarum uuuggauugaagggagcucui 21 226 CA26637 540 4 3 333 118 343 ScMiR 159a ScMiR 159a ScMiR 159a ScMiR 159a 52 4 3' 337 12 257 ScMiR 159a ScMiR 159a ScMiR 159a ScMiR 159a 337 357 12 257 ScMiR 159a Uuuggauugaagggagcucui 21 225 CA3886 66 4 3' 307 132 357 132 357 132 357 132 357 136 120 225 CA3886 66 4 3' 303 318 143 40 926 460 ScMiR 159a Uuuggauugaagggagcucui 21 132 CA18454 685 47 43' 303 352 13 377 353 363 352 13 377 353 363 362 13 373 352 21 327 355 364 451 453 454 44 454 454 454 454 4														
ScMir 159a Saccharum officinarum uuuggauugaaggagcucu 21 226 CA261537 540 + 3' 32 343 118 343 ScMir 159a ScMir 159b ScMir 159b ScMir 159b ScMir 159b ScMir 159b ScMir 159 ScMir 159b ScMir 159 ScMir 159b			Pinus taeda	cuuggauugaagggagcucca	21	198								
ScMiR 159a CA236922 953 + 3' 442 462 236 462 ScMiR 159a CA203816 5C2 + 3' 443 453 237 463 ScMiR 159a CA03816 5C2 + 3' 433 833 833 ScMiR 159b uuuggauugaaggagcucui 21 225 CA23858 636 + 3' 430 450 226 450 ScMiR 159b uuuggauugaaggagcucui 21 132 CA19092 522 + 5' 52 72 52 183 ScMiR 159 Schedonorus arundinaceus uuuggauugaaggagcucui 21 175 CK801264 461 3' 363 382 213 337 SbMiR 159 Screphum bicolor uuuggauugaaggagcucui 21 222 CF43318 647 4 43' 446 43' 453 448 468 247 468 SbMiR 159a Screphum bicolor uuuggauugaaggagcucui 21 237 BE59518 56 5' 7' 7' 7' 3			Saccharum officinarum	100000000000000000000000000000000000000	21	226								
ScMiR 159a .			Succhar am officinar am	uuuggauugaagggageueug	21	220								
ScMiR 159a														
ScMiR 159b uuuggauugaaggagcucu 21 225 CA23858 636 + 3' 430 <									+				132	357
ScMiR 159c uuuggauugaaggagcucug 21 132 CA193052 522 + 5' 52 72 52 183 SaMiR 159 Schedonorus arundinaceus uuuggauugaaggagacucug 21 175 CK801264 461 + 3' 367 387 213 387 SaMiR 159 Sorghum bicolor uuuggauugaaggagacucug 21 222 CF43301 647 + 3' 363 322 217 35 SbMiR 159a Sorghum bicolor uuuggauugaaggagacucug 21 222 CF43301 647 + 3' 448 468 247 468 SbMiR 159a CD21304 537 + 3' 448 468 247 468 SbMiR 159a uuuggauugaaggagacucug 21 237 BES073 669 + 3' 448 468 247 468 SbMiR 159a uuuggauugaaggagacucug 21 237 BES073 866 + 5' 7' 7' 7' 7'' 7''' 7'''' SpMiR 159a Yriki 159b		ScMiR 159a					CA079434	704	+	3'	213	233	8	233
SaMiR 159 Schedonorus arundinaceus uuuggauugaagggagcucug 21 175 CK 801265 489 + 3' 367 387 213 387 SaMiR 159 Sorghum bicolor uuuggauugaagggagcucug 21 222 CF433018 647 + 3' 363 382 210 382 SbMiR 159a Sorghum bicolor uuuggauugaagggagcucug 21 222 CF433018 647 + 3' 453 473 252 473 SbMiR 159a CD21304 537 + 3' 448 469 248 469 SbMiR 159a CD21304 553 + 3' 448 468 247 468 SbMiR 159a uuuggauugaagggagcucug 21 237 BE595318 586 + 5' 140 160 140 36 362 365 484 484 263 484 SbMiR 159a Uuuggauugaagggagcucug 21 252 CA484819 674 + 3' 460 480 229 480 VvMiR 159 Triticum aestiv		ScMiR 159b		uuuggauugaaaggagcucuu	21	225	CA238586	636	+	3'	430	450	226	450
SaMiR 159 CK 80126 4 61 + 3' 363 382 210 382 SbMiR 159a Sorghum bicolor uuuggauugaagggagcucu 21 222 CF433018 647 + 3' 239 259 37 259 SbMiR 159a Sorghum bicolor uuuggauugaagggagcucu 21 222 CF433018 647 + 3' 433 473 252 273 SbMiR 159a CD22406 606 + 3' 453 474 669 461 - 3' 448 468 247 468 SbMiR 159a CD21300 537 + 3' 448 468 247 468 SbMiR 159a uuuggauugaagggagcucu 21 237 BE595318 586 + 5' 140 160 140 376 SpMiR 159 Sorghum propinquum uuuggauugaagggagcucu 21 252 CA484819 674 + 3' 448 480 229 480 VvMiR 159a Yitis vinifera uuuggauugaagggagcucu 21 252 CA484819 674 + 3' 463 483 312 483 VvMiR 159a Zea mays uuuggauugaagggagcucu		ScMiR 159c		uuuggauugaagggagcucug	21	132	CA193052	522	+	5'	52	72	52	183
SbMiR 159a Sorghum bicolor uuuggauugaagggagcucug 21 22 CF433018 67 + 3' 239 259 37 259 SbMiR 159a CD224964 606 + 3' 453 473 252 473 SbMiR 159a CD224004 606 + 3' 449 469 248 469 SbMiR 159a CD21300 53 + 3' 448 468 247 468 SbMiR 159a CD21300 553 + 3' 448 468 247 468 SbMiR 159a Uuuggaugaaggagcucuu 21 237 BE595318 565 + 3' 468 469 248 469 SbMiR 159 Uuuggaugaaggagcucuu 21 232 BE595318 565 + 3' 460 480 29 480 480 249 480 249 480 249 480 249 480 249 480 249 480 249 480 249 480 249 480 249 450 4			Schedonorus arundinaceus	uuuggauugaagggagcucug	21	175	CK801265	489	+		367	387	213	387
SbMiR 159a CD22496 606 + 3' 453 473 252 473 SbMiR 159a CD224004 694 + 3' 449 469 248 469 SbMiR 159a CD21304 537 + 3' 448 468 247 468 SbMiR 159a CD21304 537 + 3' 448 468 247 468 SbMiR 159a CD21304 537 + 3' 448 468 247 468 SbMiR 159a CD21304 696 + 3' 448 468 247 468 SbMiR 159a uuuggauugaaggaggacucu 21 237 BE595318 566 + 5' 140 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 248 469 24														
SbMiR 159a CD22404 69 + 3' 449 469 248 469 SbMiR 159a CD21304 537 + 3' 448 468 247 468 SbMiR 159a CD21304 637 + 3' 448 468 247 468 SbMiR 159a CD21404 669 + 3' 468 468 247 468 SbMiR 159a CD21404 669 + 3' 468 468 247 468 SbMiR 159b CD21404 669 + 3' 469 484 263 449 468 247 450 SbMiR 159b Unuggaugaggaggeucu 21 322 B605158 265 5' 4''' 4''' 4''' 4'''' 4'''' 4'''' 4''''' 4''''''''''''''''''''''''''''''''''''			Sorghum bicolor	uuuggauugaagggagcucug	21	222								
SbMiR 159a CD21330 S7 + 3' +48 468 247 468 SbMiR 159a CD21320 S5 + 3' +48 468 247 468 SbMiR 159a CD21320 S53 + 3' +48 468 247 468 SbMiR 159a CD21407 660 + 5' 14' 468 247 488 SbMiR 159a Sorbitm propingum uuuggaugaggagcucu 21 237 BE95318 580 + 5' 14' 160 10' 36' SpMiR 159 Sorbitm propingum uuuggaugaggagcucu 21 252 CA4848 674 468 28' 28' 48' 24' 26' 36' 16' 10' 16''' 16'														
SbMiR 159a CD21320 53 + 3' 448 468 247 468 SbMiR 159a CD204047 669 + 3' 464 484 263 484 SbMiR 159b uuuggauugaagggagcucu 21 237 BE59518 586 + 5' 140 160 140 370 SpMiR 159 Sorghum propinguum uuuggauugaagggagcucu 21 322 CA48481 661 + 3' 460 480 237 480 VMiR 159 Triticum aestrium uuuggauugaagggagcucu 21 322 CA48481 661 + 3' 460 480 23 480 VMiR 159 Triticum aestrium uuuggauugaaggagcucu 21 668 CN00751 661 + 3' 463 483 312 483 VMiR 159 Triticum aestrium uuuggauugaaggagcucu 21 668 CN00751 667 5' 65 45 5' 43 31 43 43 43 43 43 32 45 44 44 4														
SbMiR 159a CD20407 669 + 3' 646 484 263 484 SbMiR 159b uuuggaugaaggagacuug 21 237 BE595318 586 + 5' 140 160 140 376 SpMiR 159 Sorghum propinquum uuuggaugaaggagacuug 21 392 BG050158 429 + 5' 7 27 7 398 TaMiR 159 Triticum aestivum uuuggaugaaggagacuug 21 252 CA484819 6'1 + 3' 460 480 252 480 VvMiR 159a Yitis vinifera uuuggaugaaggagacuug 21 168 CN07551 661 - 5' 485 505 485 505 485 505 485 505 485 505 485 505 481 505 481 505 485 505 485 505 481 381 464 480 312 483 312 483 312 483 312 483 312 483 314 461 305 505 55 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>														
SbMiR 159b uuuggauagaggagcucu 21 237 BE595318 586 + 5' 140 160 140 76 SpMiR 159 Sorghum propinquum uuuggauagaggagcucu 21 392 BG051158 249 + 5' 17' 27 7 398 TaMiR 159 Triticum aestivum uuuggauagaggagcucu 21 252 CA484819 6'1 + 3'' 460 480 259 480 VvMiR 159a Vitis vinifera uuuggauagaggagcucu 21 168 CN007551 661 - 5'' 485 505 <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>-</td> <td></td> <td></td> <td></td> <td></td>										-				
SpMiR 159 Sorghum propinquum uuuggauugaaggagacucu 21 392 BG050158 429 + 5' 7 27 7 398 TaMiR 159 Triticum aestivum uuuggauugaaggagacucu 21 252 CA484819 674 + 3' 460 480 229 480 VvMiR 159a Vitis vinifera uuuggauugaaggagacucu 21 168 CN07551 661 - 5' 485 505 485 652 VvMiR 159a Vitis vinifera uuuggauugaaggagacucu 21 168 CN07551 661 - 5' 485 505				uuuggauugaagggagcucug	21	237								
TaMiR 159 Triticum aestivum uuuggauugaaggagacucu 21 252 CA484819 674 + 3' 460 480 229 480 VvMiR 159a Vitis vinifera uuuggauugaaggagacucu 21 168 CN007551 661 - 5' 485 505 485 552 VvMiR 159a Vitis vinifera uuuggauugaaggagacucu 21 168 CN007551 661 - 5' 485 505 485 552 VvMiR 159b Ear mays uuuggauugaaggagacucu 21 1657 CF210361 563 + 5' 483 32 483 ZmMiR 150m Zae mays uuugcaugaaggagacucu 21 457 CD57348 60 + 5' 13 33 13 499 GmMiR 160n GormMiR 160n GormMiR 160n ugccugcucccuguaugceu 21 107 BM87596 554 - 3' 50 138 32 232 399 GmMiR 160a Gossypium raimondii ugccugcucccuguaugceu 21 106 CO106264 713 - 5'			Sorghum propinquum						+	5'	7	27	7	398
VvMiR 159a CF210361 563 + 3' 463 483 312 483 VvMiR 159b 557 CF208653 667 - 3' 564 584 28 584 ZmMiR 159m Zea mays uuuggauugaagggagcucug 21 457 CD573488 604 + 5' 13 33 13 469 160 GmMiR 160m Glycine max ugccuggcucccuguaugcca 21 80 CA801322 572 + 5' 13 33 13 469 160 GmMiR 160n Glycine max ugccuggcucccuguaugcca 21 107 BM87596 529 - 3' 139 23 39 233 39 GmMiR 160a 3' 107 BM87596 529 - 3' 107 157		-	*						+	3'	460	480	229	480
VvMiR 159h VvMiR 159h VvMiR 159h Zea mays uuuggauugaaggagacucu 21 457 CD57348 667 - 3' 54 <td< td=""><td></td><td>VvMiR 159a</td><td>Vitis vinifera</td><td>uuuggauugaagggagcucua</td><td>21</td><td>168</td><td>CN007551</td><td>661</td><td>-</td><td>5'</td><td>485</td><td>505</td><td>485</td><td>652</td></td<>		VvMiR 159a	Vitis vinifera	uuuggauugaagggagcucua	21	168	CN007551	661	-	5'	485	505	485	652
Intersection ZmMiR 159m Zea mays uuuuggauuggaggagcucuu 21 457 CD57348 604 + 5' 13 33 169 160 GmMiR 160m Glycine max ugccuggcucccuguaugce 21 80 CA801322 572 + 5' 13 33 13 160 GmMiR 160m Glycine max ugccuggcucccuguaugce 107 BM887596 554 - 3' 109 339 233 339 GmMiR 160a Gorsypium raimondii ugccuggcucccuguaugceu 107 BM887596 554 - 3' 107 157 262 GrMiR 160a Gorsypium raimondii ugccuggcucccuguaugceu 106 C0106199 715 - 5' 177 177 157 262 GrMiR 160a Gorsypium raimondii ugccuggcucccuguaugceu 21 106 C0106199 715 - 5' 177 177 157 262 GrMiR 160a Grosypium raimondii ugccuggcucccuguaugcea 21 344 CA002787 355 4 5' 106 5' <t< td=""><td></td><td>VvMiR 159a</td><td></td><td></td><td></td><td></td><td>CF210361</td><td>563</td><td>+</td><td>3'</td><td>463</td><td>483</td><td>312</td><td>483</td></t<>		VvMiR 159a					CF210361	563	+	3'	463	483	312	483
160 GmMiR 160m Glycine max ugccuggcucccuguaugcca 21 80 CA801322 572 + 5' 23 43 23 102 GmMiR 160n ugccuggcucccuguaugcca 21 107 BM887596 554 - 3' 319 339 233 339 GmMiR 160a Gosspium raimondii ugccuggcucccuguaugcca 21 107 BM887596 554 - 3' 500 520 18 520 GrMiR 160a Gosspium raimondii ugccuggcucccuguaugcca 21 106 C0106264 713 - 5' 157 177 157 262 GrMiR 160a Gorspium raimondii ugccuggcucccuguaugcca 21 106 C0106264 713 - 5' 157 177 157 262 GrMiR 160a Gorspium raimondii ugccuggcucccuguaugcca 293 C0005743 817 - 5' 157 177 157 262 GrMiR 160b Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcca 21 344 CA002787 355 + 3'									-		564	584	28	
GmMiR 160n ugccuggcucccuguaugccu 21 107 BM88759 554 - 3' 319 339 233 339 GmMiR 160o 503 A1794555 529 - 3' 500 520 18 520 GrMiR 160a Gossypium raimondii ugccuggcucccuguaugccu 21 106 C0106264 713 - 5' 157 177 157 262 GrMiR 160a GrMiR 160a C0106199 715 - 5' 157 177 157 262 GrMiR 160a C0102555 866 - 5' 157 177 157 262 GrMiR 160a C0102555 866 - 5' 157 177 157 262 GrMiR 160b Lgccuggcucccuguaugcca 21 344 605743 817 - 5' 406 698 HvMiR 160b Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcca 21 344 CA002787 355 + 3' 348 342 388 HvMiR 160b Bu471957 479														
GmMiR 1600 503 Al794655 529 - 3' 500 520 18 520 GrMiR 160a Gossypium raimondii ugccuggcucccuguaugceu 21 106 C0106264 713 - 5' 157 177 157 262 GrMiR 160a C010619 715 - 5' 157 177 157 262 GrMiR 160a C0102555 686 - 5' 157 177 157 262 GrMiR 160a C0102555 686 - 5' 157 177 157 262 GrMiR 160a C0102555 686 - 5' 105 177 157 262 GrMiR 160a C0102555 686 - 5' 105 177 157 262 GrMiR 160a Log Mordeum vulgare subsp. vulgare ugccuggcucccuguaugcea 21 344 CA002787 355 + 3' 324 41 344 HvMiR 160b ugccuggcucccuguaugcea 21 47 BJ471957 479 3' 368 <td< td=""><td>160</td><td></td><td>Glycine max</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>	160		Glycine max											
GrMiR 160a Gossypium raimondii ugccuggcucccuguaugccu 21 106 C010626 713 - 5' 157 173 262 GrMiR 160a C010619 715 - 5' 157 177 157 262 GrMiR 160a C010255 686 - 5' 157 173 262 GrMiR 160a C010255 686 - 5' 105 215 210 GrMiR 160b 293 C0095743 817 - 5' 406 426 406 698 HvMiR 160a Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcca 21 344 CA002787 355 + 3' 324 344 1 344 HvMiR 160b ugccuggcucccuguaugcca 21 47 BJ471957 56 - 3' 368 388 342 388 HvMiR 160b BJ471957 479 - 3' 357 377 331 377				ugecuggeucecuguaugecu	21				-					
GrMiR 160a C0106199 715 - 5' 157 177 157 262 GrMiR 160a C0102555 686 - 5' 205 225 205 310 GrMiR 160b 293 C0095743 817 - 5' 406 426 406 698 HvMiR 160a Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcea 21 344 CA002787 355 + 3' 324 344 1 344 HvMiR 160b ugccuggcucccuguaugcea 21 47 BJ472971 596 - 3' 368 388 342 388 HvMiR 160b BJ471957 479 - 3' 357 377 331 377			Gossunium raimondii	110001100011000110110000	21				2					
GrMiR 160a C010255 686 - 5' 205 225 205 310 GrMiR 160b 293 C0095743 817 - 5' 406 426 406 698 HvMiR 160a Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcea 21 344 CA002787 355 + 3' 324 344 1 344 HvMiR 160b ugccuggcucccuguaugcea 21 47 BJ472971 596 - 3' 368 388 342 388 HvMiR 160b BJ471957 479 - 3' 357 377 331 377			oossyptum ratinonati	ugeeuggeueeeuguaugeeu	21	100								
GrMiR 160b 293 C0095743 817 - 5' 406 426 406 698 HvMiR 160a Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcca 21 344 CA022787 355 + 3' 324 344 1 344 HvMiR 160b ugccuggcucccuguaugcca 21 47 BJ472971 596 - 3' 368 388 342 388 HvMiR 160b BJ471957 479 - 3' 357 377 331 377														
HvMiR 160a Hordeum vulgare subsp. vulgare ugccuggcucccuguaugcca 21 344 CA002787 355 + 3' 324 344 1 344 HvMiR 160b ugccuggcucccuguaugcca 21 47 BJ472971 596 - 3' 368 388 342 388 HvMiR 160b BJ471957 479 - 3' 357 377 331 377						293								
HvMiR 160b ugccuggcucccugunugccu 21 47 BJ472971 596 - 3' 368 388 342 388 HvMiR 160b BJ471957 479 - 3' 357 377 331 377			Hordeum vulgare subsp. vulgare	ugecuggcucecuguaugeca	21									
HvMiR 160b BJ471957 479 - 3' 357 377 331 377			÷						-				342	388
HvMiR 160b BJ471544 481 - 3' 357 377 331 377		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					BJ469735	543	-	3'	357	377	331	377
HvMiR 160b BJ469445 560 - 3' 357 377 331 377		HvMiR 160b					BJ469445	560	-	3'	357	377	331	377

Tab. 3 New miRNAs identified from plants (con-	tinued-2)
---	-----------

	miRNA	Species	miRNA sequence*	Length	Precurs	EST	EST	Stra	Folb	miR	NA	Precu	irsor
miRNA	gene			(nt)		accession	length	nd nd	ack	Start	End	Start	End
family					n length	number			arm				
	HvMiR 160b				(nt)	12/02/1/04	492		3'	250	270	221	270
	HVMIR 160b					AV934494 AV934488	482 412	:	3'	358 329	378 349	331 333	378 349
	HvMiR 160b					AV933114	608		3'	321		295	341
	HvMiR 160b					AV932125	479		3'	357	377	331	377
	OsMiR 160m	Oryza sativa	ugccuggcucccuguaugcca	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	Dia d			00	CF327535	536	-	3'	221	241	112	241
	PgMiR 160 ScMiR 160	Picea glauca Secale cereale	ugccuggcucccuguaugcca	21 21	80 180	CO484471 BE494482	770 556	+	5' 3'	269 357	289 377	269 198	348 377
	TtMiR 160a	Triticum turgidum	ugccuggcucccuguaugccu ugccuggcucccuguaugcca	21	87	BF293809	539	+	5'	259	279	259	345
	TtMiR 160b	The second se	ugccuggcucccuguaugccu	21	218	BQ788860	620	-	3'	264	284	67	284
	ZmMiR 160m	Zea mays	ugccuggcucccuguaugcca	21	81	CD445121	648	+	5'	285	305	285	365
162	GmMiR 162m	Glycine max	ucgauaaaccucugcauccag	21	367	BU081804	557	+	5'	45	65	45	411
	GmMiR 162n		ucgaugaaccgcugcauccag	21	84	CF807240	378	-	5'	300	320	300	383
	LsMiR 162	Lactuca sativa	ucgauaaaccucugcauccag	21	89	BQ845158	367	+	5'	240	260	240	328
	LIMiR 162	Lupinus luteus	ucgauaaaccucugcauccag	21	80	BG149136	442	+	3'	274	294	215	294
	LmMiR 162	Malus x domestica (cultivated apple)	ucgauaaaccuuugcauccag	21	88	CN894119	327	+	3'	173	193	106	193
	MtMiR 162 OsMiR 162m	Medicago truncatula Orvza sativa	ucgauaaaccucugcauccag	21 21	87 114	BF003769 CA764295	478 896	+ +	3' 3'	221 425	241 445	155 332	241 445
	OsMiR 162m VvMiR 162	Oryza sativa Vitis vinifera	ucgauaagccucugcauccag ucgauaaaccucugcauccag	21	85	CA764295 CF516290	896 433	+	3'	425 368	445 388	304	388
163	OsMiR 162	Oryza sativa	uugaagaggacuuggaacguuc	21	47	CA763268	433 688	-	3'	621	641	595	641
	PpMiR 163	Picea glauca	ugaagaagacuuggagcuucga	22	74	CO236262	723	-	3'	654	675	602	675
	PpMiR 163	-			92	CO254401	890	-	3'	691	712	621	712
	PsMiR 163	Picea sitchensis	ugaagaagacuuggagcuucga	22	93	CO222007	839	-	3'	695	716	624	716
	PtMiR 163a	Pinus taeda	ugaagaagacuuggagcuucga	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 PtMiR 163a					CO163668 CO157806	831 707	:	3' 3'	414 293	435 314	343 222	435 314
	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	Populus balsamifera x Populus deltoides	augaagaggaguuggaacuua	21	100	CN524551	654	+	3'	253	273	174	273
164	CsMiR 164	Citrus sinensis	uggagaagcagggcacgugca	21	168	CK936755	674	+	5'	153	173	153	320
	GaMiR 164	Gossypium arboreum	uggagaagcagggcacguaaa	21	110	BF270853	579	-	3'	94	111	1	111
	InMiR 164	Ipomoea nil	eggagaagcaggucaegugeg	21	87	BJ577535	704		3'	502	522	436	522
	PtMiR 164 PtMiR 164	Populus balsamifera subsp. trichocarpa	uggagaagcagggcacgugca	21	137	BU869001	621 537	+ +	5' 5'	16 159	36 179	16 159	152 294
	TaMiR 164a	Triticum aestivum	uggagaagcaggucacgugca	21	60	CK113235 CD899685	736	+	5' 5'	615	635	615	674
	TaMiR 164b	In mean acsirian	uggagaagcagggcacgugca	21	152	CA704421	476	+	5'	10	30	10	161
	TaMiR 164b					CD929986	654		3'	171	191	134	191
165	HeMiR 165	Hedyotis centranthoides	ucggaccaggcuucauccccc	21	165	CB086732	632	+	3'	224	244	80	244
	SbMiR 165	Sorghum bicolor	ecggaccaggeuucaucceaa	21	426	CN140010	750	-	3'	726	746	301	726
66	GmMiR 166m	Glycine max	ucggaccaggcuucauucccc	21	128	CD405934	400	-	5'	112	132	112	239
	GmMiR 166m					CD391899	338	-	5'	129	149	129	256
	GmMiR 166m				400	CA819984	385	+	3'	241	261	133	261
	GmMiR 166n				409 85	BQ785760 BM308126	488 468	+	5' 3'	47 440	67 460	47	455
	GmMiR 1660 GmMiR 166p				85 106	BM308126 BI893541	468 517	+ +	3' 3'	440 125	460 145	375 39	460 145
	GmMiR 166p		ucggaccaggcuucauucccg	21	156	BI893541 BI972515	573	+	3,	306	326	171	326
	GmMiR 166r		ucggaccaggeuucauucceg	21	136	BG316028	415	+	3'	238	258	123	258
	HvMiR 166	Hordeum vulgare subsp. vulgare	ucggaccaggcuucauucccc	21	78	BQ760548	433	+	3'		156	79	156
	InMiR 166	Ipomoea nil	ucggaccaggcuucauuccuc	21	169	BJ553847	485	+	3'	310	330	162	330
	MtMiR 166a	Medicago truncatula	ucggaccaggcuucauucccc	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	0	ucggaccaggcuucauuccuc	21	171	AW694053		+	3'	371	391	221	391
	OsMiR 166m	Oryza sativa Bioog alayog	ucggaccaggcuucaaucccu	21	97 74	CA760464	684 810	-	5' 5'	442	462	442	538
	PgMiR 166 SoMiR 166	Picea glauca Saccharum officinarum	ucggaccaggcuucauuccuu ucggaccaggcuucauucccc	21 21	74 127	CO476713 CA067403	810 732	- +	5' 5'	155 44	175 64	155 44	228
	SoMiR 166	sacenar am ogreinar am	asgavaggeuutauuteee	21	127	CA067403	1098	+	5'	44	64	44	172
	SbMiR 166	Sorghum bicolor	ucggaccaggcuucauucccc	21	108	CN126049	707	+	3,	157	177	70	177
	TaMiR 166	Triticum aestivum	ccggaccaggcuucauuccca	21	111	CK204430	801	-	3'	168	188	78	188
	TaMiR 166					CK204088	823	-	3'	167	187	77	187
				21	87	CK369135	655	+	3'	163		97	183

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

	miRNA	Species	miRNA sequence*	-	h Precurs		EST			miR!		Preci	
	gene			(nt)	-	accession	length	nd	ack	Start	End	Start	E
umily					n length	number			arm				
	7				(nt) 78	00621086	(90	+	3'	1/0	180	103	1
	ZmMiR 166n ZmMiR 1660				78 78	CO531986 CO521288	680 766	+	3'	160 161	180	103	1
	ZmMiR 1660 ZmMiR 166p				457	CF626917	661	-	3,		642	186	
67	-	Arabidopsis thaliana	ugaagcugccagcaugaucug	21	342	AU239920	619	+	5'		213	193	
	GmMiR 167m		ugaagcugccagcaugaucua	21	82	BI095235	554	+	5'		425	405	
	GmMiR 167m	olycine max	ugaugeugeugeuugaueuu	21	64	CD411229	586		3,		516	453	
	GmMiR 167n				64	BU548271	462		3'		229	166	
	GmMiR 167n				0.	BM892909	489	+	5'		144		
	GmMiR 167n					BG509097	498	+	5'		247	227	
	GmMiR 1670		ugaagcugccagcaugaucug	21	109	BQ629383	549	+	3'	222	242	134	
	InMiR 167	Ipomoea nil	ugaagcugccagcaugaucug	21	271	BJ563257	578	+	5'	265	285	265	
	OsMiR 167m	Oryza sativa	ugaagcugccagcaugaucug	21	100	CF322848	567	+	5'	129	149	129	
	OsMiR 167m					CF309326	552	+	5'	102	122	102	
	OsMiR 167n				90	BX928793	1039	+	5'	277	297	277	
	PcMiR 167	Phaseolus coccineus	ugaagcugccagcaugaucuu	21	69	CA916400	283	+	5'	135	155	135	
	PptMiR 167	Populus tremula x Populus tremuloides	ugaagcugccagcaugaucug	21	409	BU810017	684	+	5'	256	276	256	
	SoMiR 167a	Saccharum officinarum	ugaagcugccagcaugaucug	21	149	CA284394	729	-	3'	629	649	501	
	SoMiR 167a					CA284319	659	+	5'	123	143	123	
	SoMiR 167b				139	CA287325	456	+	5'	136	156	136	
	SMiR 167	Saccharum sp	ugaagcugccagcaugaucug	21	139	CF571260	403	+	5'	144	164	144	
	TaMiR 167	Triticum aestivum	ugaagcugacagcaugaucua	21	210	CK209889	1031	+	3'	397	417	208	
	ZmMiR 167m	Zea mays	ugaagcugccagcaugaucug	21	80	CF630597	417	-	3'	365	385	306	
68	AtMiR 168m	Arabidopsis thaliana	ucgcuuggugcaggucgggaa	21	104	H77158	464	+	5'	42	62	42	
	GmMiR 168m	Glycine max	ucgcuuggugcaggucgggaa	21	88	BE661028	688	+	5'	272	292	272	
	GmMiR 168m					AW424354	195	+	5'	20	40	20	
	GmMiR 168m					BE659175	300	•	3'	254	274	187	
	HtMiR 168	Hedyotis terminalis	ucgcuuggugcaggucgggaa	21	140	CB076866	639	+	5'	168	188	168	
	HvMiR 168	Hordeum vulgare subsp. vulgare	ucgcuuggugcagaucgggac	21	66	CA003609	456	+	5'	127	147	127	
	HvMiR 168						418	+	5'	128	148	128	
	LeMiR 168	Lycopersicon esculentum	ucgcuuggugcaggucgggac	21	145	BF097936	572	+	5'	16	36	16	
	LeMiR 168					BE461111	551	+	5'	16	36	16	
	LeMiR 168					BE461110	572	+	5'	16	36	16	
		Oryza sativa	ucgcuuggugcagaucgggac	21	67	CB667935	768	+	5'	105	125	105	
	OsMiR 168m					CB662543	804	+	5'	113	133	113	
	OsMiR 168m					CB655478	698 722	+	5' 5'	116	136	116	
	OsMiR 168m					CB629782 CA756101	733 607	+ +	5'	12 118	32 138	12	
	OsMiR 168m PbtMiR 168	Populus balsamifera subsp.	110001100100000000000000000000000000000	21	118	BU871538	508	+	5,	37	57	118 37	
	FULWIR 108	trichocarpa	ucgcuuggugcaggucgggaa	21	110	BU8/1558	508	Ŧ	3	37	57	37	
	PtMiR 168	Populus tremula	ucgcuuggugcaggucgggaa	21	118	BU893331	297	+	5'	102	122	102	
	PtMiR 168	1 opulus iremula	ucgeuuggugeagguegggaa	21	110	BU889659	436	+	5'	102	122	102	
	PtMiR 168					BU888204	520	+	5'		123	103	
	PptMiR 168a	Populus tremula x Populus tremuloides	າເດດຕາມາດຕາເດດລອດກາດອອດລອ	21	118	BU811772	503	+	5'	53	73	53	
	PptMiR 168a	r opinius ir cininte x r opinius ir cinintonices	uegeuuggugeuggueggguu	21	110	BU809183	461	+	5'	109	129	109	
	PptMiR 168b				140	BU886509	514	+	5'	70	90	70	
	SsMiR 168	Saccharum sp	ucgcuuggugcagaucgggac	21	65	CF576659	636	+	5'		213	193	
	SoMiR 168	Saccharum officinarum	ucgcuuggugcagaucgggac	21	65	CA173899	663	+	5'		184	164	
	SoMiR 168	·····				CA156827	708	+	5'		172	152	
	SoMiR 168					CA123434	622	+	5'	184	204	184	
	StMiR 168	Solanum tuberosum	ucgcuuggugcaggucgggac	21	133	CK254942	865	+	5'	85	105	85	
	StMiR 168		0 00 0 00 000 =			CK249947	916	+	5'	12	32	12	
	StMiR 168					CK246126	980	+	5'	12	32	12	
	StMiR 168					CK243936	770	+	5'	10	30	10	
	SbMiR 168	Sorghum bicolor (sorghum)	ucgcuuggugcagaucgggac	21	67	CN133168	669	+	5'	173	193	173	
	SbMiR 168					CD213423	572	+	5'	173	193	173	
	SbMiR 168					CD210663	530	+	5'	183	203	183	
	SbMiR 168					CD209437	662	+	5'	184	204	184	
	SbMiR 168					CD209338	549	+	5'	179	199	179	
	SbMiR 168					CD207772	531	+	5'	193	213	193	
	SbMiR 168					CD205823	586	+	5'	182	202	182	
	SbMiR 168					CD205059	581	+	5'	182	202	182	
	SbMiR 168					CN145907	312	+	5'	95	115	95	
	VvMiR 168	Vitis vinifera	ucgcuuggugcaggucgggaa	21	110	CF604588	521	+	5'	52	72	52	
	ZmMiR 168m	Zea mays	ucgcuuggugcagaucgggac	21	65	BU099153	532	+	5'	62	82	62	
	ZmMiR 168m					BG842539	498	+	5'	97	117	97	
	ZmMiR 168m					CA826777	522	+	5'	62	82	62	
	ZmMiR 168m					BG842528	494	+	5'	97	117	97	
59	GmMiR 169m	Glycine max	cagccaaggaugacuugccga	21	98	BF595231	457	+	3'	386	406	309	

miRMA	miRNA gene	Species	miRNA sequence*	-	Precurs	EST accession	EST length		Folb ack	miR			
family	gene			(nt)	n length (nt)		length	na	arm	Start	End	Start	
	GmMiR 169n				60	AW201497	400	+	5'	273	293	273	
	GmMiR 1690		cagccaaggaugacuugccgg	21	82	CA953278	293	+	5'	90	110		
	GmMiR 169p			21	101	AW596073	317	+	5'	26	46	26	
	GsMiR 169	Glycine soja	cagccaaggaugacuugccgg	21	103	BM524615	483	+	5'	63	83		
	GsMiR 169					BM520766	550	+	5'	63	83	63	
	GsMiR 169					BF598910	465	+	5'	85	105	85	
	OsMiR 169m	Oryza sativa	cagccaaggaugacuugccga	21	91	CF280503	479	+	5'	141	191	141	
	OsMiR 169m					BQ906497	582	+	5'	169	189	169	
	OsMiR 169n		cagecaaggaugaeuugeegg	21	92	BX899561	863	•	3'	410	430	339	
	OsMiR 169n					BX899553	870	+	5'	198	217	198	
	PtMiR 169a	Populus tremula	cagecaaggaugaeuugeega	21	100	BU862460	379	+	5'	89	109	89	
	PtMiR 169b		cagecaaggaugauuugeega	21	84	BU865420	519	+	5'	164	184	164	
	SoMiR 169	Saccharum officinarum	uagccaaggaugacuugccgg	21	143	CA148737	739	-	3'	643	663	521	
	SoMiR 169					CA148650	627	+	5'	306	326	306	
	SbMiR 169	Sorghum bicolor	aagccaaggaugacuugccgu	21	92	CN143986	849	+	5'	72	92	72	
	ZmMiR 169m	Zea mays	cagccaaggaugacuugccgg	21	93	CD960569	364	+	5'	221	241	221	
170	AcMiR 170a	Allium cepa	ugauugagccgcgccaauauc	21	89	CF447956	774	-	3'	229	249	161	
	AcMiR 170b				76	CF442989	808	-	5'	476	496	0 90 0 90 3 63 3 63 3 63 4 141 9 89 9 89 9 80 1 141 9 169 9 80 1 141 1 121 1 211 1 121 1 944 3 900 1 121 1 121 1 121 1 121 1 121 1 121 1 121 1 121 1 121 1 121 1 121 2 244 8 8 9 161 1 129 1 129 1 120	
	AcMiR 170c				55	CF435529	459	-	3'	215	235		
	AcMiR 170d				215	CF439580	800	-	5'	22	42		
	AcMiR 170d					CF440325	741	-	5'	247	267		
	AcMiR 170d					CF437865	841	-	5'	117	137	117	
	AcMiR 170d					CF437845	833	-	5'	117	137	117	
	AtMiR 170m	Arabidopsis thaliana	ugauugagccgcgccaauauc	21	92	BX838271	1148	+	3'	125	145	54	
	AtMiR 170n				172	BX839725	1135	-	5'	944	964	944	
	GmMiR 170m	Glycine max	ucauugageegugeeaauaue	21	80	CA937914	421	+	3'	265	285	206	
	GmMiR 170n		uaauugagccgcgucaauauc	21	80	BM892213	391	+	3'	195	215	136	
	HcMiR 170	Hedyotis centranthoides	cgauugagccgugccaauauc	21	79	CB087210	578	+	3'	482	502	424	
	HvMiR 170	Hordeum vulgare subsp. vulgare	ugauugagccgugccaauauc	21	78	CA009309	621	+	3'	394	414	337	
	OsMiR 170a	Oryza sativa	ugauugagccgugccaauauc	21	99	CF329292	641	+	3'	308	328	230	
	OsMiR 170b				107	CA756196	623	+	3'	210	10 230 1	124	
	PpbMiR 170	Populus balsamifera x Populus deltoides	ugauugagccgcgccaauauc	21	197	CN524603	813		5'	121	141	121	
	StMiR 170	Solanum tuberosum	ugauugagccgcgccaauauc	21	124	BG598365	717		3'	603	623	500	
	SbMiR 170	Sorghum bicolor	ugauugagccgcgccaauauc	21	54	BE365851	438		5'	367	387	367	
	TaMiR 170a	Triticum aestivum	cgauugagccgugccaauauc	21	81	BJ275219	707		5'	358	378	358	
	TaMiR 170b		ugauugagccgugccaauauc	21	94	CD910903	588	+	3'	165	185	92	
	ZmMiR 170m	Zea mays	ugauugageegugeeaauaue	21	115	CD441617	674		5'	530	550	530	
	ZmMiR 170m	-				CD440809	620	+	3'	221	241	129	
	ZmMiR 170n		ugauugagccacgucaauauc	21	79	CD963893	174		5'	44	64	44	
	ZmMiR 170o		ugauugagccacgucaauauc	21	52	CD949754	71	+	5'	8	28	8	
	ZmMiR 170p		ugauugagccgcgccaauauc	21	62	AI657239	555	-	5'	482	502	482	
171	AcMiR 171a	Allium cepa	ugauugagccgcgccaauauc	21	89	CF447956	774		3'	229	249	161	
	AcMiR 171a		0 00 00			CF442989	808		3'	476	496		
	AcMiR 171b				115	CF439580	800		5'	22	42		
	AcMiR 171b					CF440325	741		5'	247	267	521 306 72 221 161 181 22 247 117 117 54 206 136 424 337 230 124 121 500 367 358 92 203 124 121 500 367 358 8 92 204 129 44 8 8 482 161 177 129 14 8 530 129 44 8 8 82 161 177 129 129 129 129 129 129 129 129 129 129	
	AcMiR 171b					CF437865	841		5'	117	137		
	AcMiR 171b					CF437845	833		5'	117	137		
	AcMiR 171c				216	CF435529	459		5'	215	235	1 221 1 221 161 476 5 476 5 181 2 22 2 7 7 117 7 117 5 544 4 944 5 136 2 22 2 117 7 117 5 544 4 337 6 230 0 124 3 500 124 358 5 92 0 530 3 532 2 222 2 207 7 117 7 117 5 539 4 502 5 106 3 373 7 207 5 137 2 <td< td=""><td></td></td<>	
	AtMiR 171m	Arabidopsis thaliana	ugauugagccgcgccaauauc	21	94	BX838271	1148	+	3'	125	145		
	AtMiR 171n		-Gunder of the Sector and the		461	BX839739			3'	979	999		
	AtMiR 171n					BX839725	1135		3'	944	964		
	AtMiR 1710				339	BU635724	762		5'	106	126		
	GeMiR 171	Glycine clandestina	ugauugagccgcgccaauauc	21	205	BG838164	855		5'	373	393		
	GmMiR 171 GmMiR 171m		ugauugagccgcgccaauauc	21	172	BM892515	567	-	5'	297	317		
	GmMiR 171n GmMiR 171n	sy che max	ucauugageegugeeaauaue	21	79	CA937914	421	+	3'	265	285		
	GmMiR 171n GmMiR 171n		- sunnage - gugee aadade	21	.,	BM892213	391	+	3'	195	215		
	GrMiR 1711 GrMiR 171	Gossypium raimondii	ugauugagccgcgccaauauc	21	117	CO129257	489	+	5'	272	215		
	HeMiR 171	Hedyotis centranthoides	gauugageegegeeaauaue	21	79	CB087210	489 578	+	3'	482	502		
	Hemik 171 HvMiR 171	,		21	79 78	CB087210 CA009309	578 621	+	3'	482 394	502 414		
		Hordeum vulgare subsp. Vulgare	ugauugagccgugccaauauc						-		414 449		
	LcMiR 171	Lotus corniculatus var. japonicus	ugauugauccacgccaauauc	21	83	BP063562	516	-	3' 2'	429			
	OsMiR 171m	Oryza sativa	ugauugagccgugccaauauc	21	107	CA756196	623	+	3'	210	230		
	OsMiR 171n	a 1 a .			99	CF329292	641	+	3'	308	328		
	ScMiR 171a	Saccharum officinarum	uaauugageegegecaauaue	21	375	CA139546	868	-	5'	420	440		
	ScMiR 171b		ugauugacccgcgccaauauc	21	201	CA104875	617	-	3'	545	565		
	StMiR 171	Solanum tuberosum	ugauugagccgcgccaauauc	21	124	BG598365	717	-	3'	603	623	500	
	SbMiR 171	Sorghum bicolor	ugauugagccgcgccaauauc	21	55	BE365851	438	-	5'	367	387	367	
	TaMiR 171a	Triticum aestivum	ugauugagccgugccaauauc	21	94	CD910903	588	+	3'	165	185	92	
	TaMiR 171b		cgauugagccgugccaauauc	21	80	BJ275219	707		5'	358	378	358	

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

	miRNA	Species	miRNA sequence*	Length	Precurs	EST	EST	Stra	Folb	miR	NA	Precu	irsor
niRNA	gene			(nt)		accession	length	nd	ack	Start	End	Start	End
amily					-	number			arm				
	T-1 (D 121-				(nt)	CV210/20				007	02/	100	024
	TaMiR 171c ZmMiR 171m	Zea mars	uaauugageegegeeeauaue ugauugageegugeeaauaue	21 21	527 113	CK210439 CD441617		-	3' 5'	906 530	926 550	400 530	926 642
	ZmMiR 171m ZmMiR 171m	Zeu mays	ugauugageegugeeaauaue	21	115	CD441017 CD440809		+	3,	221	241	129	241
	ZeMiR 171	Zinnia elegans	cgauugagccgcgccaauauc	21	65	AU304757			3'	76	96	32	- 96
72	AsMiR 172	Aegilops speltoides	aggaucuugaugaugcugcag	21	54		500	+	5'	141	161	141	194
	AtMiR 172m	Arabidopsis thaliana	agaauccugaugaugcugcag	21	141	BX834115	403	+	5'	181	201	181	321
	AtMiR 172m					BP657662	389	+	5'	188	208	188	328
	AtMiR 172m					BP644108		+	5'	188	208	188	328
	CsMiR 172a	Citrus sinensis	agaaucuugaugaugcugcaa	21	107	CF506714		+	3'	404	424	318	424
	CsMiR 172b	0	agaauccugaugaugcugcac	21	158	CK702444			5'	301	321	301	45
	GmMiR 172m GmMiR 172n	Giycine max	agaaucuugaugaugcugcau	21	120 114	BU084569 BI320499	526 592	+ +	3' 3'	200 384	220 404	101 291	220 404
	GhMiR 172	Gossypium hirsutum	agaauccugaugaugcugcag	21	225	CA997811		+	5'	25	404	251	24
	GrMiR 172a	Gossypium raimondii	agaaucuugaugaugcugcau	21	148	CO080218			3'	600	620	473	62
	GrMiR 172a					CO069513		+	5'	196	216	196	34
	GrMiR 172b		aaaaucuugaugaugcugcau	21	240	CO099417	656	+	5'	351	371	351	59
	HvMiR 172a	Hordeum vulgare subsp. vulgare	agaauccugaugaugcugcag	21	255	BM816977	871	-	5'	437	457	437	69
	HvMiR 172b				90	CB872382		+	3'	298	318	229	31
	HvMiR 172c				137	CD662228		+	5'	148	168	148	28
	HvMiR 172d			21	272	CB878060		-	5'	181	201	181	45
	LeMiR 172 OsMiR 172m	Lycopersicon esculentum Oryza sativa	agaaucuugaugaugcugcau	21 21	106 297	AI484737 CR280455		+	3' 5'	448 159	468 179	363 159	46 45
	OsMiR 172m OsMiR 172n	Oryza sanva	agaauccugaugaugcugcag	21	62		813	+	3,	369	389	328	38
	OsMiR 172n OsMiR 172o				53	CB679168		+	5'	297	317	297	34
	OsMiR 172p				318	CB673867		+	3'	358	378	61	37
	PpeMiR 172	Picea engelmannii x Picea sitchensis	agaauccugaugaugcugcac	21	68	CO212920		+	3'	534	554	487	55
	PgMiR 172	Picea glauca	agaauccugaugaugcugcac	21	248	CK441660		-	5'	64	84	64	31
	PsMiR 172	Picea sitchensis	agaauccugaugaugcugcac	21	68	CO224629	625	+	3'	517	537	470	53
	PsMiR 172					CO217583	628	+	3'	520	540	472	54
	SoMiR 1721	Saccharum officinarum	agaauccugaugaugcugcag	21	300	CA291208	686	•	5'	368	388	368	66
	SoMiR 172b				82	CA300926	571	+	3'	263	283	302	38
	StMiR 172	Solanum tuberosum	agaaucuugaugaugcugcau	21	93	BQ114970		-	5'	492	512	492	58
	SbMiR 172	Sorghum bicolor	agaauccugaugaugcugcag	21	119	CF430424		-	3'	298	318	200	31
	SbMiR 172						661	-	3'	285	305	187	30
	SbMiR 172	71		21	240	CF772781 CF974182		2	5' 3'	139	159	139	38 58
	TcMiR 172 TaMiR 172a	Theobroma cacao Triticum aestivum	agaauccugaugaugcugcag agaauccugaugaugcugcag	21 21	100	CA626451			3,	560 220	580 240	341 141	24
	TaMiR 172a TaMiR 172b	Ir incum destivum	agaauceugaugaugeugeag	21	118	CA486144			3,	354	374	257	37
	TaMiR 172c		aggaucuugaugaugcugcag	21	54	CD920346		+	5'	458	478	458	51
	TaMiR 172c					CK206331		+	5'	245	265	245	29
	TaMiR 172c					CD909382		+	5'	493	513	493	54
	TaMiR 172c					CD894004	614	+	5'	376	396	376	42
	TaMiR 172c					CD885904	549	+	5'	277	297	277	33
	TaMiR 172c					CD884790	620	+	5'	309	329	309	36
	TaMiR 172c					CA746952			5'	233	253	233	28
	TaMiR 172c					CA639688		+	5'	211	231	211	26
	TaMiR 172c					BJ281140		+	5'	351	371	351	40
	TaMiR 172c TaMiR 172c					BM136378		+	5' 5'	107	127 395	107 375	16 42
	VvMiR 172c	Vitis vinifera	agaauccugaugaugcugcag	21	335	BJ212004 CF208353	560 745	+	5' 5'	375 396	416	396	73
	MiZmR 172		agaauccugaugaugcugcag	21	455	CD058685			3'	481	501	47	50
	ZmMiR 172n	2cu mays	agaauccugaugaugcugcau	21	81	BM348229			3'	324	344	264	34
73		Arabidopsis thaliana	uucgcuuggagaganaaaucac	22	85	BE521498		-	3'	286	307	123	30
19	GmMiR 319m	Glycine max	uuggacugaagggagcuccc	20	168	BQ630517	585	+	3'	315	334	167	33
	GmMiR 319m					BQ630503	516	+	3'	315	334	167	33
	GmMiR 319m					BE475558	445	+	3'	364	383	216	38
	GmMiR 319n		uuggagugaagggagcucca	20	166	BQ453148	524	+	3'	191	210	45	21
	GmMiR 319o		uuggacugaaaggagcuccu	20	182	BG237979		+	3'	294	313	132	31
	LtMiR 319	Liriodendron tulipifera	uuggacugaagggagcuccc	20	167	CK760683		+	3'	210	229	63	22
	PpMiR 319	Physcomitrella patens subsp. patens	uuggacugaagggagcucca	20	169	BJ191979		+	3'	443	462	294	46
	PbMiR 319	Populus balsamifera subsp. trichocarpa	uuggacugaagggagcuccc	20	169	CK091603	680	-	5'	96	115	96	26
	SoMiR 319a	Saccharum officinarum	uuggacugaagggugcuccc	20	538	CA184549	607	+	5'	13	32	13	55
	SoMiR 319b		uuggauugaagggagcuccu	20	176	CA229394	554	+	3'	294	313	138	3
	SoMiR 319c		uuggagugaagggagcucca	20	184	CA227452	847	-	5'	238	257	238	42
	TaMiR 319a	Triticum aestivum	uuggacugaagggagcuccc	20	196	CA483944	298	-	5'	96	115	96	29
	TaMiR 319b		uuggagugaagggagcucca	20	184	CD925456	602	-	5'	190	209	190	37
	TaMiR 319c		uuggagugaagggagcucca	20	253	CA599744	644	-	3'	242	261	10	26
	VvMiR 319	Vitis vinifera	uuggacugaagggagcuccu	20	87		220	+	3'	126	145	59	14
		Zea mays	uuggauugaagggagcuccu	20	178	BM338067	520	-	5'	314	333	314	4

* Letters in shadow indicate that the nucleotides are different from the previously known Arabidopsis miRNAs

Tab. 4 ESTs matched by known Arabidopsis	miRNAs, and identified as new miRNAs in plant	s
--	---	---

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

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

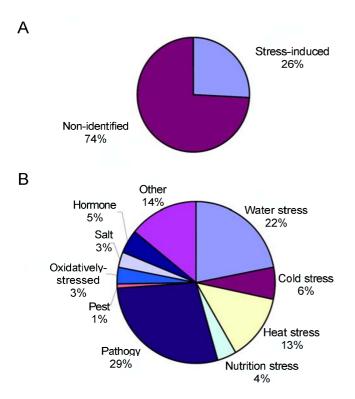


Fig. 4 miRNAs are induced and regulated by environmental biotic and abiotic stresses. **(A)** 26% of ESTs were obtained from stressinduced tisssues. **(B)** ESTs containing miRNAs were obtained from tissue by various biotic and abiotic environmental stresses.

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, acidor 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

miRNAs	Flower	Leaf	Root	Whole		Developing			-	Fruit,	Callus	Somatic	Xylem	Fiber	Total
						embryos or			meristem			embryos			
						seeds				and			phloem		
										spike					
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

Tab. 5 Tissue-specific expression of plant miRNAs

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 expresses 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 embryoes 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

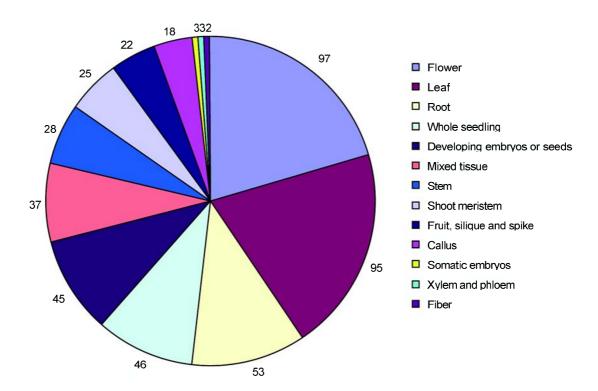


Fig. 5 Preferential expression of miRNAs

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 proteincoding 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 chloroplastand 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^{2+} 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 basepairing to cleave the mRNAs [1, 39, 44]. However, for some miRNAs, such as miRNA 172, although they have perfect or near-perfect complimentarity 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].

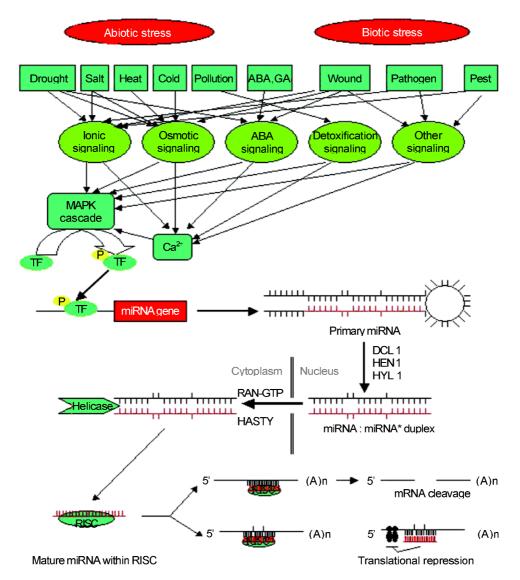


Fig. 6 A model for miRNA biogenesis, regulation of miRNA gene regulation, and the pathway of miRNA function.

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 *Nic-otiana 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.
- Lai EC, Tomancak P, Williams RW, Rubin GM. Computational identification of *Drosophila* microRNA genes. Genome Biol 2003; 4:R42.
- 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, Socci 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.