In silico mining of novel microRNAs from coffee (Coffea arabica) using expressed sequence tags

Article *in* Journal of Horticultural Science and Biotechnology · May 2013

DOI: 10.1080/14620316.2013.11512973

5 authors, including:

Rebijith K B
University of Cambridge
33 PUBLICATIONS **113** CITATIONS
SEE PROFILE

Ramasamy Asokan
Indian Institute of Horticultural Research
75 PUBLICATIONS **253** CITATIONS
SEE PROFILE

Ranjitha H Hande
Indian Institute of Horticultural Research
8 PUBLICATIONS **17** CITATIONS
SEE PROFILE

Journal of Horticultural Science & Biotechnology (2013) **88** (3) 325–337

In silico mining of novel microRNAs from coffee (*Coffea arabica*) using expressed sequence tags

By K. B. REBIJITH[1]*, R. ASOKAN[1], H. H. RANJITHA[1], V. KRISHNA[2] and K. NIRMALBABU[3]
[1]Division of Biotechnology, Indian Institute of Horticultural Research' Hessaraghatta Lake (PO), Bangalore - 560 089, Karnataka, India
[2]Department of Biotechnology and Bioinformatics, Kuvempu University, Jnanasahyadri, Shankaraghatta, Shivamogga - 577451, Karnataka, India
[3]Division of Crop Improvement and Biotechnology, Indian Institute of Spices Research (IISR), Marikkunnu, Calicut - 673012, Kerala, India
(e-mail: rebijith@gmail.com)
(Accepted 4 February 2013)

SUMMARY
MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression either by causing mRNA cleavage or by translational repression in plants and animals, respectively. We are beginning to understand the functional roles played by miRNAs in various life-forms to regulate growth and development. In this regard, there have been no reports on miRNAs in coffee, *Coffea arabica*, an economically important cash-crop. The experimental identification and characterisation of spatially and temporally expressed miRNAs have proved to be major challenges. Computational approaches that depend on sequence homology and secondary structure verification are important tools. In the present study, 18 novel miRNAs were identified in *C. arabica* using computational approaches to analyse the 174,275 expressed sequence tag (EST) sequences available in the NCBI-GenBank database. The construction of phylogenetic trees of five large families of miRNAs (those with ≥ 20 miRNAs) clearly indicated that the *C. arabica* miRNAs are evolutionarily conserved regulators of gene expression. Forty-one potential mRNA targets were also identified by homology searches for all 18 miRNAs against the *C. canephora* mRNA dataset using psRNATarget (http://plantgrn.noble.org/psRNATarget/; a plant small RNA target analysis site). Functional characterisation of the 18 newly-identified miRNAs revealed that the majority were involved in transcriptional regulation and signal transduction pathways. Our studies will provide new insights into miRNAs and their functions in *C. arabica,* and have opened-up new avenues for further functional analyses of the control of gene expression.

MicroRNAs (miRNAs) are small, endogenous, non-coding RNAs, approx. 19 – 22 nucleotides (nts) in length, which regulate gene expression in most eukaryotes through mRNA degradation or translational repression (Bartel, 2004). The first miRNA discovered was *lin 4* in *Caenorhabditis elegans* by Lee *et al.* (1993). *Lin 4* regulated the timing of larval development by down-regulating of expression of the lin 14 protein. The first miRNA discovered in plants was in *Arabidopsis thaliana* (Llave *et al.,* 2002; Reinhart *et al.,* 2002; Mette *et al.,* 2002; Park *et al.,* 2002). Almost 4,677 plant miRNAs are now available in miRBase (http://www.mirbase.org), among which 321, 328, 661, and 395 were found in *Zea mays, A. thaliana, Oryza sativa,* and *Glycine max,* respectively (Griffiths-Jones *et al.,* 2006). Plant miRNAs regulate a wide variety of developmental processes such as leaf morphogenesis (Emery *et al.,* 2003; Mallory *et al.,* 2004a), floral differentiation and development (Aukeman and Sakai, 2003; Chen, 2004), and even plant responses to environmental stress (Kasschau *et al.,* 2003; Llave *et al.,* 2004; Sunkar and Zhu, 2004; Fujii *et al.,* 2005; Sunkar *et al.,* 2006). Most plant miRNAs control the expression of translational factors that then regulate important steps during plant development.

*Author for correspondence.

The production of an miRNA involves the transcription of an miRNA gene into a primary RNA by RNA polymerase II. The primary miRNA is then processed in the nucleus by endonucleases such as Drosha and Pasha into a precursor miRNA (pre-miRNA), containing a hairpin structure with 60 – 100 nts. The pre-miRNA is then cleaved into a miRNA: miRNA* duplex by a dicer-like enzyme (DCL-1) in the nucleus and translocated to the cytoplasm by a plant orthologue of Exportin 5. The mature miRNA, with approx. 22 nts, is then released from the pre-miRNA and forms a complex with RISC (the RNA-induced silencing complex) which contains the argonaute protein (Hutvagner and Zamore, 2002). In animals, miRNAs bind to the 3'-untranslated region (UTR), 5'-UTR or even to the coding region of the target mRNA (Ambros, 2004; Bartel, 2004). In plants, miRNAs bind to the protein coding region and cause either mRNA degradation (Llave *et al.,* 2002; Park *et al.,* 2002; Rhoades *et al.,* 2002) or translational repression (Aukeman and Sakai, 2003; Chen, 2004) by perfect, or near-perfect pairing at the "seed region" (i.e., 2 – 8 nts at the 5'-end of the mature miRNA; Lee *et al.,* 2001; Rhoades *et al.,* 2002; Bartel, 2009) that is important for target specificity.

Coffea arabica belongs to the family Rubiaceae and is a globally important cash crop. The identification of

miRNAs in *C. arabica* would increase our understanding of developmental processes, thus providing ways to improve crop and pest management. The three approaches used to identify miRNAs in plants are forward genetics, bioinformatic predictions (Zhang *et al.,* 2005; 2006a,b), or direct cloning and sequencing (Lee and Ambros, 2001; Lagos-Quintana *et al.,* 2001; Chen *et al.,* 2005). We used a bioinformatics approach to identify miRNAs in *C. arabica* using the 174,275 expressed sequence tags (ESTs) available in the NCBI-GenBank database (http://www.ncbi.nlm.nih.gov/).

The objectives of the present study were: (i) to identify and characterise novel miRNAs in *C. arabica;* (ii) to perform a phylogenetic analysis of any newly-identified miRNAs; and (iii) to predict the targets of any newly-identified miRNAs.

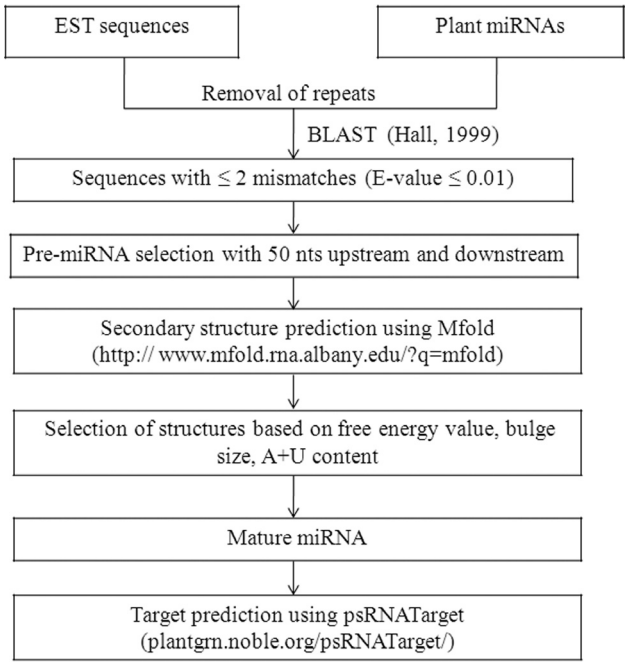

FIG. 1
Steps involved in the prediction of potential miRNAs.

MATERIALS AND METHODS

Retrieval of expressed sequence tags and microRNAs

A total of 174,275 ESTs in *C. arabica,* expressed at various stages of development, were downloaded from the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/). All reported plant miRNAs (4,677), belonging to different plant families were retrieved from miRBase (Release 18.0; http://miRNA.sanger.ac.uk; Griffith-Jones *et al.,* 2006). Redundancy among the miRNAs and ESTs was excluded manually by sequence identity.

Homology searches and potential miRNA identification

EST sequences were aligned with query miRNA sequences using the local BLAST programme in BioEdit (Hall, 1999) for homology searches. The "expectation value" was kept at 0.01, and those sequences having ≤ 2 nt mismatches between the aligned EST and miRNA sequences were selected for subsequent scanning for possible pre-miRNA sequences. All other parameters were kept at their default values.

Secondary structure prediction and naming of miRNAs

Pre-miRNA sequences were extracted from the selected EST sequences with 50 nts upstream and downstream of the first and last residues of the mature miRNA. These pre-miRNA sequences were then submitted to Mfold (Zuker and Stiegler, 1981; Zuker, 2003; http://www.mfold.rna.albany.edu/?q=mfold) to predict their secondary structures. The various criteria considered for secondary structure prediction included: (i) free energy (ΔG) values ≤ –18 kcal mol^{-1}; (ii) a bulge size of not more than 7 nts; (iii) a mature miRNA on one arm of the stem region; and (iv) an A+U content of 30 – 70% (Ambros *et al.,* 2003; Zhang *et al.,* 2006; Xie *et al.,* 2007). Identified miRNAs were named according to miRBase guidelines (Griffiths-Jones *et al.,* 2006).

Phylogenetic analysis

The predicted miRNAs were classified into families according to the Rfam database (Griffiths-Jones *et al.,* 2005). The newly-identified pre-miRNA sequences were then compared with all the plant miRNA sequences

TABLE I
Predicted mature miRNA sequences of C. arabica *with their MFE‡ and MFEI‡ values and A+U content*

miRNA	EST ID†	Start position	End position	Strand	Match extent	MFE‡ (kcal mol^{-1})	A+U content (%)	Predicted miRNA sequence (5'€ 3')	MFEI
car-miR172c	GW446798	79	99	+	20/21	–26.90	60.33	GUAGCAUCAUCAAGAUUCACA	0.560
car-miR5658a	GR996153	492	512	+	20/21	–22.80	55.37	AUGAUGAUGAUGAUGAUGACA	0.422
car-miR5658b	GT020019	103	123	–	20/21	–31.90	57.85	AUGAUGAUGAUGAUGAUGAAU	0.625
car-miR1171a	GW473864	169	191	+	22/23	–33.70	52.03	AGGAGUGGAGUGGAGUGGAGUGG	0.571
car-miR1171b-1	GT680258	116	138	–	23/23	–34.10	60.16	UGGAGUGGAGUGGAGUGGAGUGG	0.695
car-miR1171b-2	GT681553	343	365	+	23/23	–29.10	60.98	UGGAGUGGAGUGGAGUGGAGUGG	0.606
car-miR156f	GW427474	102	123	–	21/22	–22.30	62.30	UUGACAGAAGAGAGAGAGCAUA	0.484
car-miR167g	GT671631	166	187	+	21/22	–39.70	61.48	UGAAGCUGCCAGCAUGAUCUGG	0.844
car-miR4414	GW458140	162	181	+	19/20	–32.30	50.00	UGCUGCUGACUCGUUGGCUC	0.538
car-miR5368	GR988681	248	266	+	19/19	–41.30	47.06	GGACAGUCUCAGGUAGACA	0.655
car-miR2673a	GW478692	238	259	+	20/22	–33.90	55.74	CAUCUUCCUCUUCCUCUUCCUC	0.627
car-miR5532	GT001031	570	591	+	22/22	–26.00	59.02	AUGGAAUAUAUGACAAAGGUGG	0.520
car-miR390b	GW445903	113	133	+	21/21	–21.20	56.20	AAGCUCAGGAGGGAUAGCGCC	0.400
car-miR166c	GT690348	65	85	–	21/21	–24.80	52.07	CCGGACCAGGCUUCAUCCCAG	0.427
car-miR396e	GW491635	611	632	–	20/22	–31.70	47.11	UUCCACAGGCUUUCUUGAACGA	0.495
car-miR398	GW479264	171	191	+	20/21	–33.50	52.07	UGUGUUCUCAGGUCACCCCUU	0.577
car-miR319	GW458822	221	241	+	19/21	–32.50	44.63	AUUGGACUGAAGGGAGCUCCC	0.485
car-miR828a	GW491209	274	295	–	21/22	–20.20	59.02	UCUUGCUCAAAUGAGAAUUCCA	0.404

†EST ID, accession number of the EST in the NCBI database.
‡MFE, minimum free energy; MFEI, minimum free energy index.

deposited in miRBase. Five phylogenetic trees were constructed for large families of miRNAs (i.e., those with ≥ 20 miRNA) using MEGA5.0 (Tamura *et al.,* 2011), to infer evolutionary relationships among members from different species. Neighbor-Joining (NJ) trees were constructed using the Kimura-2-Parameter (K2P) distance model (Kimura, 1980; Saitou and Nei, 1987).

Target prediction

The identification of mRNA targets for miRNAs is important in order to reveal their regulatory functions. The newly-predicted miRNAs were uploaded into psRNATarget, an online web server (http://www.plantgrn.noble.org/psRNATarget/; Grun *et al.,* 2005; Zhang *et al.,* 2005; 2006a) for target prediction. Target predictions were carried out using *C. canephora* unigene sequences with an expectation value of 3.0, with all other parameters set at default values.

Target mRNAs were selected based on the following criteria: (i) mismatches were not allowed in the "seed region"; (ii) ≤ 2 wobble matches were allowed in the "seed region"; (iii) mismatches were not allowed at the cleavage site (positions 10 – 11 of the mature miRNA); and (iv) no gaps were allowed in the "seed region".

RESULTS AND DISCUSSION

In this study, 18 novel miRNAs were identified from *C.*

Table II
Predicted pre-miRNA sequences with their 18 corresponding EST sequences

Pre-miRNA	EST ID†	Start-end positions	Pre-miRNA sequence (5'€ 3')
car-MIR172c	GW446798	29-149	GAUGAUUUAGCUAUGCAUGGAUGGAUCGAUUGACAGUUGUUGUUGCGGGUGUAG CAUCAUCAAGAUUCACAUGCAAAGCCCAGGUGGUAAGCUGUGUUUUUAAUUAAGU GUAUUAGUCUAU
car-MIR5658a	GR996153	442-562	GGGAUUCAAAUCUUACGUAGCAACCAUGAAAUUGUUGCCAGCAAAGGGUGAUGA UGAUGAUGAUGAUGACAAAUUCGGGUGUCAGAUCGAAUGGUCUUUCUUGACUGA UCCUGUGGAGGGG
car-MIR5658b	GT020019	53-173	AAGCCAAUCCCCCAUCGCAAAGAUAAUAUUACUAUUAGUAGAUGAUGAUGAUGAU GAUGAUGAUGAUGAAUAUUGAUGUUGUUGUUACUGCUGCUGCUGCUGCAGAU GCUGGGGGAGG
car-MIR1171b-1	GT680258	66-188	AAUAACAUAUAUCACACACUCUCUCUCUAUAUAUCUAUAUAUCCACCGAGUGGAGU GGAGUGGAGUGGAGUGGAUAAUGCUUCAAAAUGCCCAGCCCUUUUUACAUUGAGG AUUGAGCAUUCA
car-MIR1171a	GW473864	119-241	GGGGUGGGGGCCAUGAAUCUAUUACAACGAGAAGCAUCCUAACUACCACAAGGAG UGGAGUGGAGUGGAGUGGAGGAUUUGGUUUGGAGGGGAAACGCCCUUUAAACAU GGAAUUUGUAAGUA
car-MIR1171b-2	GT681553	293-415	AUGAUACCAUAAAAGCAGUACUCUGUUGUAGUUUGAUGAUACAUAUAUGAUGGA GUGGAGUGGAGUGGAGUGGAGUGGGCUAUAGGUGUGGUGUAUACAUCAAUGAUG GUAUAUAUUGGUAGU
car-MIR156f	GW427474	52-173	UUAUGCACACAAAUCUUAUUGUCAACGCUACUAAUCUAACAGUCCAGGGAUUGAC AGAAGAGAGAGAGCAUACAAAUGAUGAGGGUAAUGAUCUUCAAUUUGUCAAAACG GUUGUAUCGUCU
car-MIR167g	GT671631	116-237	UAUCAUCAAGUUCAUUUAUGUUCUUUCUGAUGUAUUUUGGUCUGAGAGCUUGAA GCUGCCAGCAUGAUCUGGUCAUCUAGCAGCUUCAAUUCGCUCACCAACAAUACAAU CCAAUGAAUUUG
car-MIR4414	GW458140	112-231	AGGAAACUCCAUCUGUCCAACCAAGAUGGUUGGGGUGGGAGAUCGAUCUCUGCU GCUGACUCGUUGGCUCAUGAGCACAUCGAUUGAGUGAUUAACGUUGGAGGCAUUG AAGCAUGCAAG
car-MIR-5368	GR988681	198-316	ACUCUGGAAGAGCUAGAAUUCUAACCUUGUGUCAGGACCUACGGGCCAAGGGAC AGUCUCAGGUAGACAGUUUCUAUGGGGCGUAGGCCUCCCAAAAGGUAACGGAGGC GUGCAAAGGU
car-MIR2673a	GW478692	188-309	UACUAGUGACUCAGAGGGUGAAAGUGACAGUGAAAUUGAAAGUUCAUCUUCAUC UUCCUCUUCCUCUUCCUCAUCAUCCGGGAGUAGUGAUGAUGAUGACGAAGAUGAA GUCGAUGGCAAGG
car-MIR5532	GT001031	520-641	CUACCAUUUCACCACCAAGGCAUGAAUAUGAUAUCUAUCUAGUGUGAUGUAUGGA AUAUAUGACAAAGGUGGGGUAUUUCUAUCGAUCGGUCAUGUCAUAUAGGCCCGAG UUGGACAUCCAA
car-MIR390b	GW445903	63-183	AAGAGAUGAUCUUGAUGAUCAGGUUGAUAAACAGCAUGGGAGGAUCGGUAAAGC UCAGGAGGGAUAGCGCCAUAGAGAGUAAAGUGAUGAUGAUGGCCGUGGUUGUGAU UUGUGAAUAUUG
car-MIR166c	GT690348	15-135	UGAUGCUAUACCAGUGCAGCCAUGAGAAAUAGCAAUGAUUCCAAUGGAAUCCGGA CCAGGCUUCAUCCCAGGCAUUUGGACCCACUCAACAGCAGUUCCAGUGGCCUUUGA UAGAAACUCU
car-MIR396e	GW491635	561-681	AAUUUGGCAGCAUUAUUCGUGGUUCCUGCAACAGAAUGGCCAGGUUGGCCUUCC ACAGGCUUUCUUGAACGAUGGCGGCCCCUGUUCAUGUGCCGCUCACAGUACUUCUG GUCAGCAACUG
car-MIR398	GW479264	121-241	GUGCUGCAGGUUGAUUACAGAUUCCAUGCAAGCAUCAAAUAUGCUGCACUUGUGU UCUCAGGUCACCCCUUUGGGGCAACCCCGUUUUGCUACUGUGAAAUGCAUCUGUAC GUAGCAGUGA
car-MIR319	GW458822	171-291	UAGCCAAUGCAUGGUGUGGGAGCAACUCCUGUCGCUCACUUUGCCCGCCCAUUGG ACUGAAGGGAGCUCCCCGAUGGAUCUCGCCCUUCACUCAUCACCUCAGCGUCCAUC AUAUCUGAUA
car-MIR828a	GW491209	224-345	UGCCGCAGGAGCUCCAGUUGUUCCAUCUGUUAUUUUUGUUGAAUAUAUUUUCUUG CUCAAAUGAGAAUUCCAGUAGUUCUUUAUUUCAUUGUCCGUCCGUCCUGGAAGUC UCCCAGCAAUCA

†EST-ID, accession number of the EST in the NCBI database.

arabica based on sequence homologies, and were characterised using computational approaches. A flow-chart representing the various steps involved is depicted in Figure 1.

Homology searches and secondary structure predictions

An homology search employing all possible miRNAs against all EST sequences yielded 88 potential miRNA-containing sequences. The selected ESTs were then scanned for pre-miRNAs by including 50 nts upstream and downstream from first and last nucleotides of the putitive mature miRNA sequence. Pre-miRNA sequences were then subjected to secondary structure prediction by applying various criteria such as a minimum free energy (MFE), a minimum A+U content, and the presence of the mature miRNA on one arm of the hairpin secondary structure, in order to reduce the number of false positives. Secondary structure predictions yielded 18 mature miRNAs (Table I; Figure 2), along with their pre-miRNA sequences (Table II). The newly-identified miRNAs were named according to miRBase guidelines (Griffiths- Jones *et al.,* 2006).

Characteristics of C. arabica *miRNAs*

The mature miRNAs and pre-miRNAs identified varied in size from 19 – 22 nt, and from 118 – 122 nt, respectively. The minimum free energy values of the identified miRNAs ranged from –41.30 kcal mol^{-1} to –20.00 kcal mol^{-1}.

miRNAs bind more strongly to certain proteins as they have a higher (A+U) content compared to other RNAs (Zhang *et al.,* 2006b; Gupta *et al.,* 2010). Our studies showed that miRNA (A+U) contents ranged from 44 – 62%. According to Qiu *et al.* (2007), every 10^4 ESTs should result in 1.00 – 1.67 miRNAs. In this regard, our study resulted in 18 miRNAs from 174,275, ESTs or an average of 1.09 miRNAs per 10^4 ESTs, similar to earlier studies (Zhang *et al.,* 2006c; Xie *et al.,* 2007). Almost all of our results for *C. arabica* miRNAs agreed with earlier studies in tomato, oil palm, and brinjal (Yin *et al.,* 2008; Nasaruddin *et al.,* 2007; Reddy *et al.,* 2011).

Phylogenetic analysis

Eleven of the 18 newly-predicted miRNAs belonged to 11 Rfam families (Griffiths- Jones *et al.,* 2005), while the other seven (5658a, 5658b, 1171a, 1171b1, 1171b2, 5368, and 5532) did not belong to any of the existing Rfam families. Five of the larger Rfam families (i.e., those with ≥ 20 plant miRNAs) were used in our phylogenetic analysis.

According to Zhang *et al.* (2005), mature and pre-miRNAs are highly-conserved among species in the same kingdom. In this regard, we generated phylogenetic trees for the five Rfam families, MIR-172, MIR-156, MIR-167, MIR-390, and MIR-396 (Figure 3) using the pre-miRNA sequences. The phylogenetic trees showed clustering of *C. arabica* miRNAs with other plant miRNAs and confirmed that both miRNAs and pre-miRNAs can be considered as evolutionarily conserved regulators of gene expression (Zhang *et al.,* 2005).

Target prediction

Compared with their animal counterparts, plant miRNAs recognise a single target site in the coding region to guide mRNA cleavage (Bartel, 2004). The perfect, or near-perfect, complementarity of miRNA and mRNA then leads to RISC-mediated cleavage, which is more common in plants than translational repression, as seen in animals (Rhoades and Bartel, 2004). Predicting the targets for the newly-identified miRNAs is a key step in determining their physiological or biological function. Targets for the newly-identified miRNAs were predicted using psRNATarget (http://www.plantgrn.noble.org/psRNATarget/).

miRNA-regulated genes control a wide range of physiological, biological, and metabolic process in plants (Reddy *et al.,* 2011). In the present study, 41 miRNA targets were identified and are shown in Table III. Our results showed that all predicted miRNA:mRNA duplexes had good "seed region" complementarity (Lewis *et al.,* 2005; Bartel, 2009), with strong alignment matches outside the "seed region" (Vella *et al.,* 2004). G:U wobble creates local distortions in the miRNA:mRNA duplex and thus reduces the effectiveness of the gene silencing mechanism. According to Vella *et al.* (2004), G:U wobble was critical for the down-regulation of *lin-41*. G:U wobble is rare, but tolerable, at various positions in the "seed region" (Didiano and Hobert, 2006; Singh and Nagaraju, 2008). Thus, we allowed up to two G:U wobbles in the "seed region" during mRNA target prediction, as they also provide recognition sites for RNA enzymes and RNA binding proteins (Ghosh *et al.,* 2007). The cleavage site (positions 10-11 in the miRNA) must have a perfect match. Mismatches at the cleavage site reduced both *in vitro* cleavage and *in vivo* phenotypic effects (Mallory *et al.,* 2004b).

Many of the miRNA targets obtained coded for transcription factors involved in plant development and physiological processes (Table IV). Other targets included proteins involved in metabolism, signal transduction, and stress responses. In particular, *car-miR-167g* bound to mRNAs that encode basic Helix-Loop-Helix-family proteins, which bind specifically to the G-box DNA sequence motif CACGTG in *Arabidopsis* (Gabriela *et al.,* 2003).

car-miR-2673a had multiple gene targets involved in nucleic acid binding, DNA repair, helicase activity, proline-rich family proteins (involved in actin polymerisation), and zinc knuckle protein activity. Zinc

TANLE III
Pre-miRNAs and their corresponding Rfam[†] family

SI No.[‡]	Rfam family	Pre-miRNA member
1	MIR-172	car-MIR172c
2	MIR-156	car-MIR156f
3	MIR-167	car-MIR167g
4	MIR-4414	car-MIR4414
5	MIR-2673	car-MIR2673a
6	MIR-390	car-MIR390b
7	MIR-166	car-MIR166c
8	MIR-396	car-MIR396e
9	MIR-398	car-MIR398
10	MIR-159	car-MIR319
11	MIR-828	car-MIR828a

Pre-miRNA-car-MIR5658a, car-MIR5658b, car-MIR171a, car-MIR171b1, car-MIR171b2, car-MIR5368 and car-MIR5532 did not align with any Rfam family.
[†]Rfam, an RNA family database (http://rfam.sanger.ac.uk/; Griffiths-Jones *et al*., 2003).
[‡]SI No., serial number.

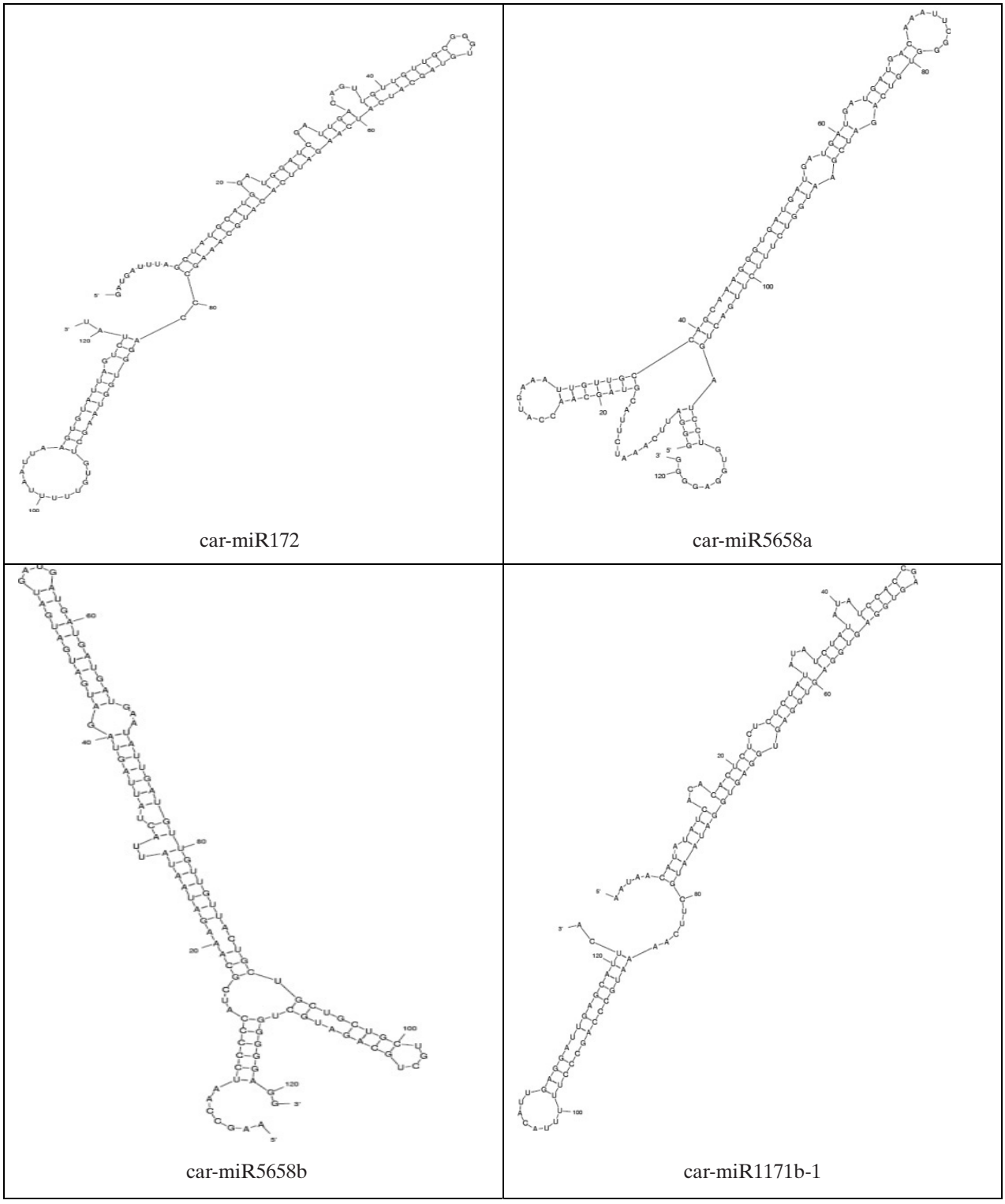

FIG. 2
Hairpin secondary structures in 18 newly-predicted pre-miRNAs from *C. arabica* based on the Mfold programme by Zuker and Stiegler (1981) and Zuker (2003).

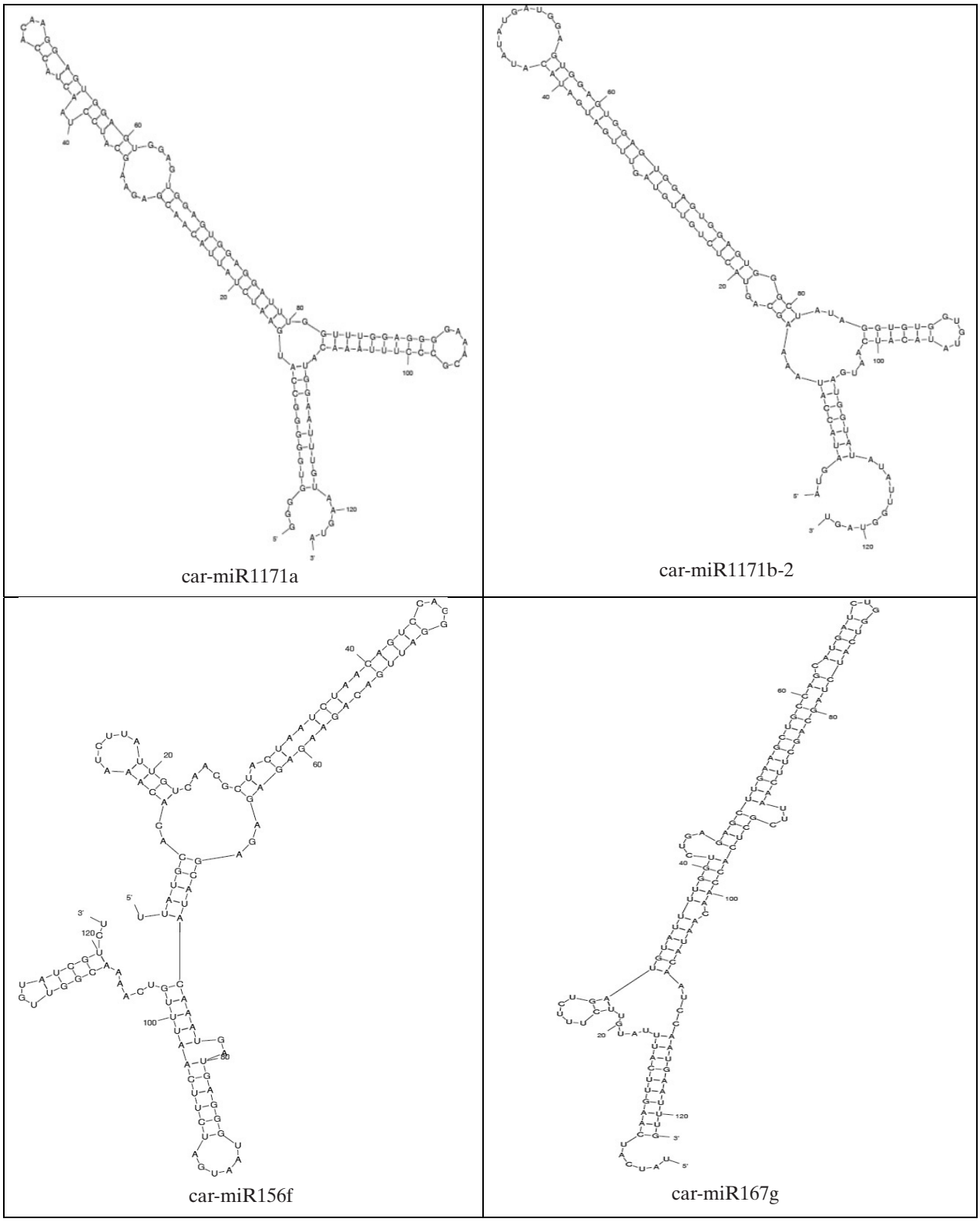

FIG. 2 (continued)

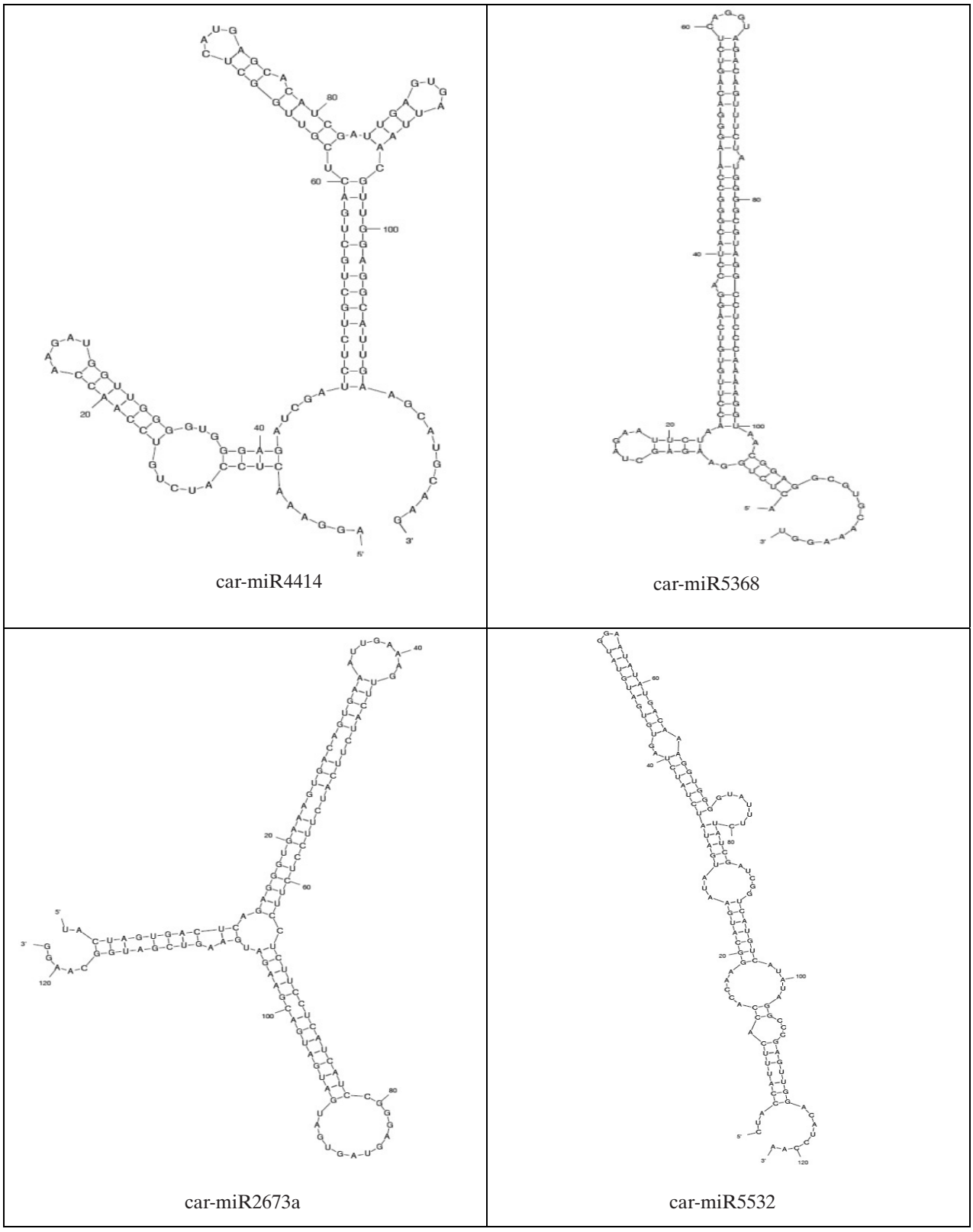

Fig. 2 (continued)

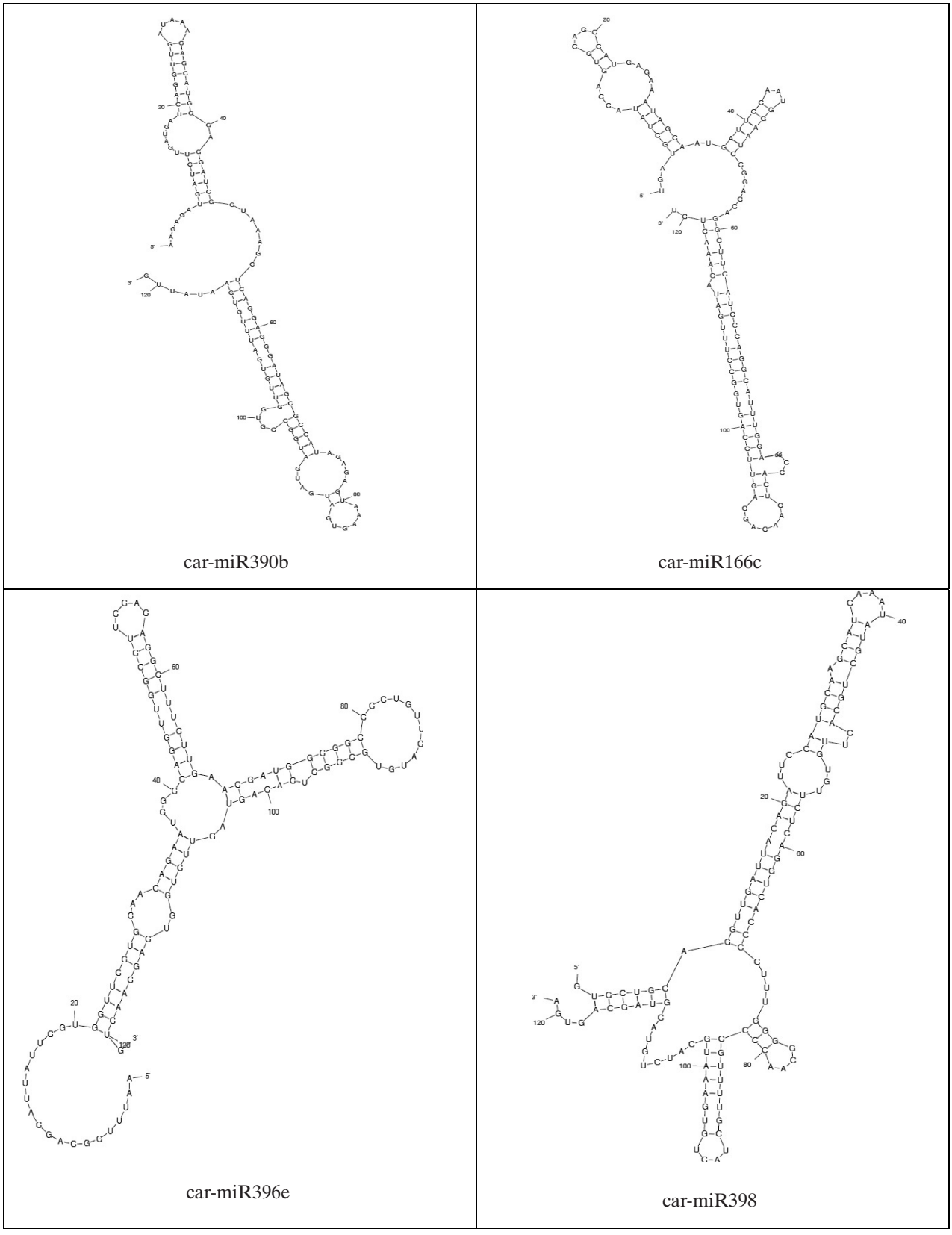

FIG. 2 (continued)

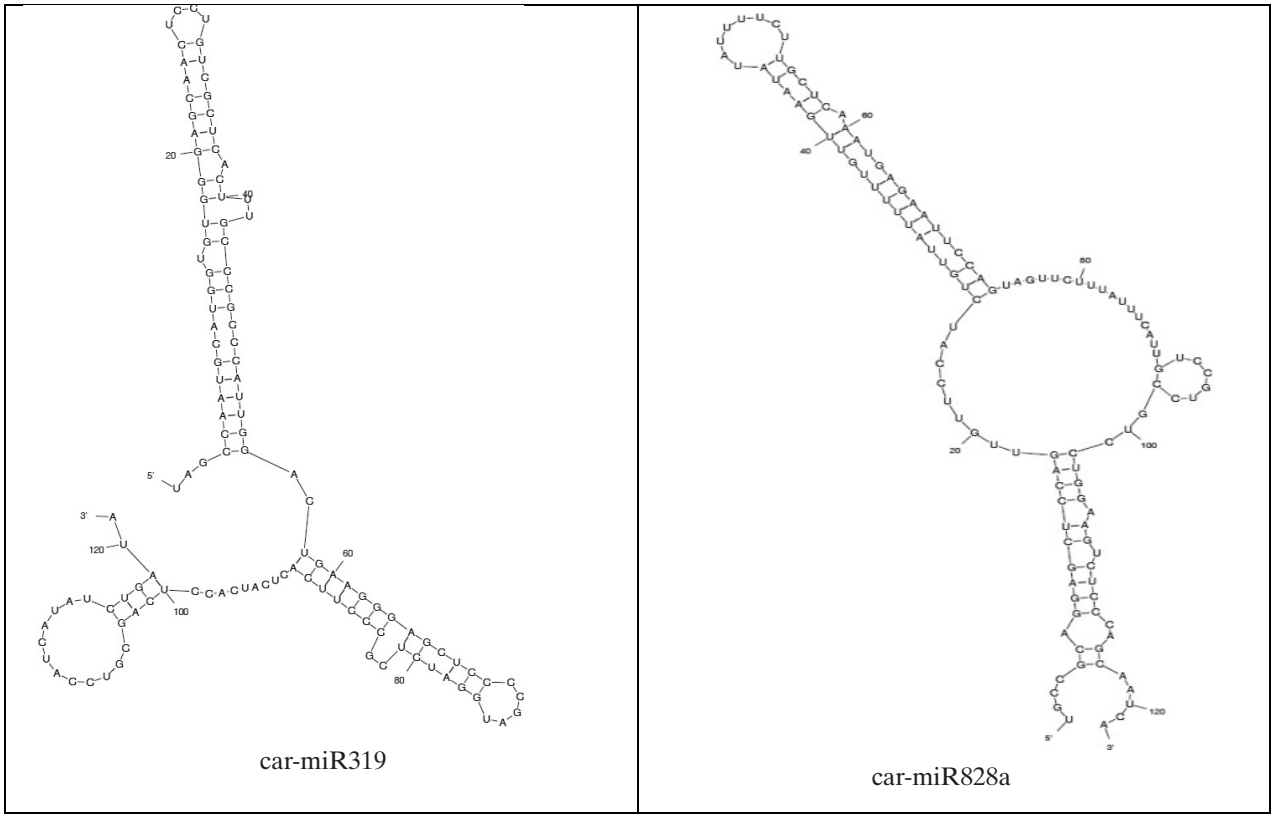

Fig. 2
Hairpin secondary structures in 18 newly-predicted pre-miRNAs from *C. arabica* based on the Mfold programme by Zuker and Stiegler (1981) and Zuker (2003).

knuckle proteins are involved in transcriptional activation, metabolism, cell signalling, and apoptosis.

car-mir1171a also bound to mRNAs encoding ABC transporter-family proteins involved in plant resistance to lead. In contrast, *car-miR-1171b1* and *b2* targeted mRNAs involved in calmodulin binding and carbohydrate metabolism.

In our study, *car-miR-156f* targeted mRNA for the SQUAMOSA promoter binding (SPB) protein, a transcription factor which plays a key role in regulating flowering time and fruit development (Mallory and Vaucheret, 2004; Nasaruddin *et al.,* 2007; Reddy *et al.,* 2011).

car-miR396e bound to the mRNA encoding the transcription activator, GRL 1. *car-miR-5658a* and *b* bound to mRNAs encoding various proteins involved in signal transduction, developmental process, defence responses through protein kinases, myb family proteins, and pathogenesis-related proteins. No target mRNAs were found for four miRNAs (*car-miR-4414*, *car-miR-5354*, *car-miR-828a*, and *car-miR-319a)* probably due to incomplete coverage of all miRNAs in the EST database (Reddy *et al.,* 2011).

Target multiplicity and co-operativity

A single miRNA may have multiple target genes and therefore multiple binding sites in the 3'-UTR (Ghosh *et al.,* 2007). Few of our newly-identified miRNAs showed multiple targets. *car-miR5658a* had a maximum of 18 targets (Table III). Another important characteristic shown by miRNAs is co-operativity, where more than one miRNA can regulate a single target gene. In the present study, no co-operativity was observed among the 18 newly-identified miRNAs from *C. arabica*.

CONCLUSIONS

We have identified 18 novel miRNAs from the 174,275 EST sequences expressed at various developmental stages of *C. arabica*. The potential roles of these miRNAs in *C. arabica* include regulation of transcription and signal transduction pathways. Thus, they may regulate abiotic stress responses, auxin responses, leaf and flower development, and heat shock responses. The full potential of using miRNAs in crop production, protection, and improvement requires further investigation, paying more attention on the whole system than to individual miRNAs and their targets. The results of this study will advance our functional analysis of these miRNA sequences, which probably play important roles in the developmental biology of *C. arabica*.

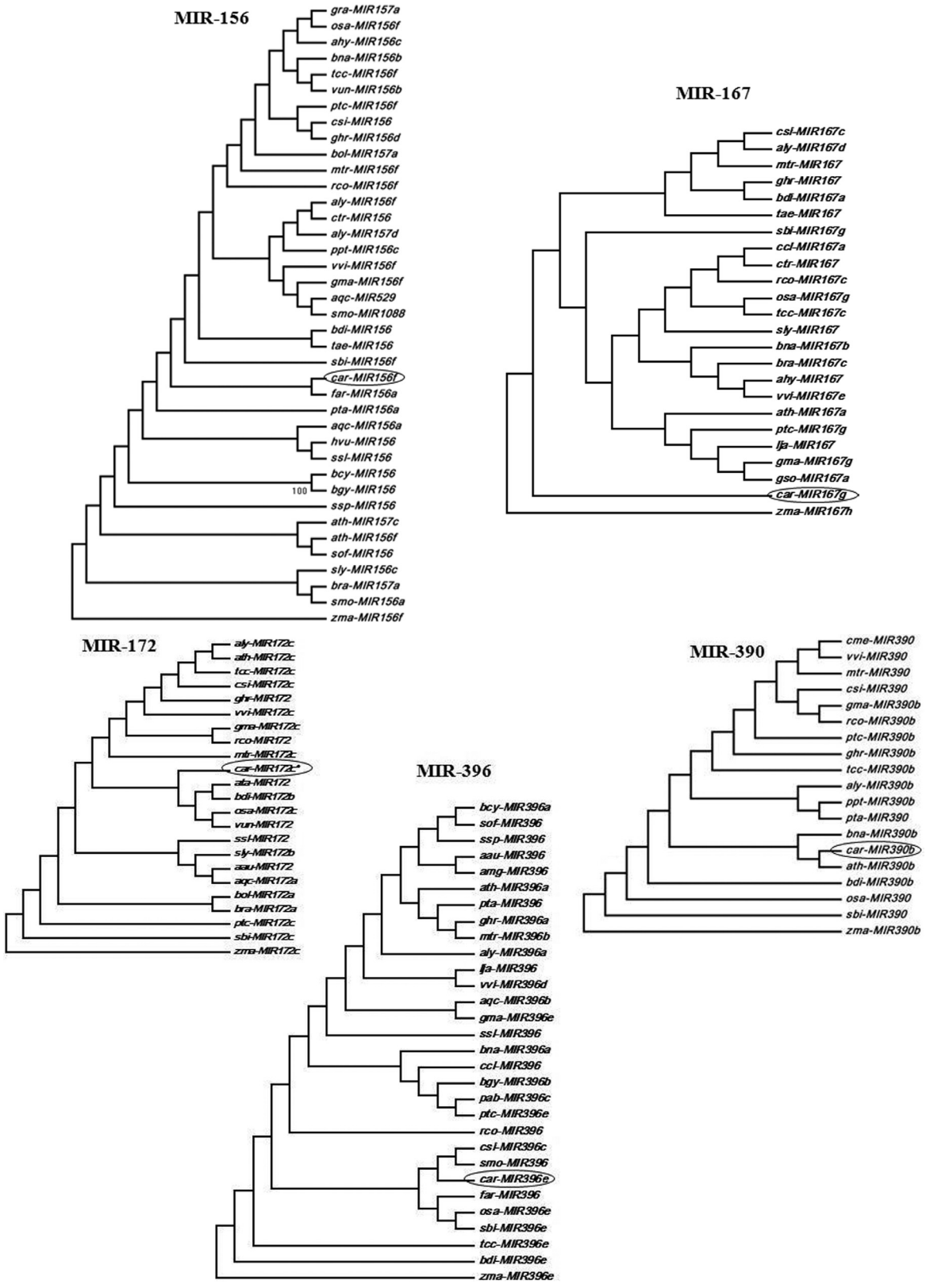

FIG. 3

Five neighbor-joining (NJ) phylogenetic trees (by MEGA. 5. 0) of the five major families (those with ≥ 20 pre-miRNAs) of pre-miRNA sequences (MIR156, MIR167, MIR 172, MIR390, and MIR396) from various members of the plant kingdom. The first three letters indicate the name of the species: car, *Coffea arabica*; gra, *Gossypium raimondii*; osa, *Oryza sativa*; ahy, *Arachis hypogaea*; bna, *Brassica napus*; tcc, *Theobroma cacao*; vun, *Vigna unguiculata*; ptc, *Populus trichocarpa*; csi, *Citrus sinensis*; ghr, *Gossypium hirsutum*; bol, *Brassica oleracea*; mtr, *Medicago truncatula*; rco, *Ricinus communis*; aly, *Arabidopsis lyrata*; ctr, *Citrus trifoliata*; ppt, *Physcomitrella patens*; vvi, *Vitis vinifera*; gma, *Glycine max*; bdi, *Brachypodium distachyon*; tae, *Triticum aestivum*; sbi, *Sorghum bicolor*; far, *Festuca arundinacea*; pta, *Pinus taeda*; aqc, *Aquilegia caerulea*; hvu, *Hordeum vulgare*; ssl, *Salvia sclarea*; bcy, *Bruguiera cylindrica*; bgy, *Bruguiera gymnorhiza*; ssp, *Saccharum* ssp; ath, *Arabidopsis thaliana*; sof, *Saccharum officinarum*; sly, *Solanum lycopersicum*; bra, *Brassica rapa*; smo, *Selaginella moellendorffii*; zma, *Zea mays*; ccl, *Citrus clementina*; lja, *Lotus japonica*; gso, *Glycine soja*; ata, *Aegilops tauschii*; aau, *Acacia auriculiformis*; amg, *Acacia mangium*; pab, *Picea abies*; cme, *Cucumis melo*. Pre-miRNAs from *Coffea arabica* are indicated with black circles.

Table IV

Details of the predicted miRNA targets in Coffea canephora *and annotated protein functions obtained from psRNATarget† (a plant small RNA target analysis server)*

miRNA	Target	E- value			Sequence alignment	Target description
car-miR167g	SGN-U356572	3.0	miRNA	22	GGUCUAGUACGACCGUCGAAGU 1	Basic helix-loop-helix (bHLH) family protein
			Target	338	UCAGUUCAUGAUGGUAGCUUUA 359	
car-miR172c	SGN-U347580	2.5	miRNA	20	CACUUAGAACUACUACGAUG 1	EST with unknown function
			Target	104	UUGAAUCUUGAUGGUGCUAA 123	
car-miR2673a	SGN-U348364	0.5	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	Proline-rich family protein
			Target	345	GAGGAGGAGGAAGAGGAAGAUG 366	
	SGN-U349645	0.5	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	Expressed protein
			Target	390	GAGGAGGAGGAAGAGGAAGAUG 411	
	SGN-U348459	1.5	miRNA	20	CCUUCUCCUUCUCCUUCUAC 1	UBP1 interacting protein 2a (UBA2a), RNA-binding region RNP-1(RRM)
			Target	335	AGAAGAGGAGGAGGAAGAUG 354	
	SGN-U347326	2.0	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	Zinc finger (GATA type) family protein, similar to zinc finger protein
			Target	181	GAGGAGGAGGAAGAGGAAGACG 202	
	SGN-U360137	2.0	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	EST with unknown function
			Target	208	GAGGAGGAGGAAGAGGAAGAGG 229	
	SGN-U354765	2.0	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	Tetratricopeptide repeat (TPR)-containing protein
			Target	473	GAGGAGGAGGAAGAGGAAGAAG 494	
	SGN-U354161	2.0	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	RNA helicase, DNA repair
			Target	147	GACGGAGAGGAAGAGGAAGGUG 168	
	SGN-U348868	2.0	miRNA	22	CUCCUUCUCCUUCUCCUUCUAC 1	Zinc knuckle (CCHC-type) family protein, nucleic acid binding
			Target	754	GAGGAGGAGGAAGAGGAAGAAG 775	
car-miR319	SGN-U354024	2.5	miRNA	20	CCUCGAGGGAAGUCAGGUUA 1	TCP family transcription factor, putative, similar to TFPD
			Target	1013	GGGGGACCCUUCAGUCCAAU 1032	
car-miR398	SGN-U347681	2.5	miRNA	20	UCCCCA-CUGGACUCUUGUGU 1	Plastocyanin-like domain-containing protein
			Target	154	AGGGGUCGACUUGAGAACACA 174	
car-miR5532	SGN-U356488	3.0	miRNA	20	UGGAAACAGUAUAUAAGGUA 1	EST with unknown function
			Target	559	UGCUUUGUCAUAUAUUUUAU 578	
car-miR1171a, b-1, b-2	SGN-U347505	0.0	miRNA	23	GGUGAGGUGAGGUGAGGUGAGGU 1	ABC transporter family involved in resistant to lead
			Target	1832	CCACUCCACUCCACUCCACUCCA 1854	
	SGN-U356219	0.0	miRNA	23	GGUGAGGUGAGGUGAGGUGAGGU 1	EST with unknown function
			Target	543	CCACUCCACUCCACUCCACUCCA 565	
car-miR1171 b-1, b-2	SGN-U355009	2.5	miRNA	20	GAGGUGAGGUGAGGUGAGGU 1	Lactoylglutathione lyase, Calmodulin binding, Involved in CHO metabolism
			Target	1	CUUUACUUCACUUCACUUCA 20	
car-miR156f	SGN-U350056	0.0	miRNA	21	UACGAGAGAGAGAAGACAGUU 1	EST with unknown function
			Target	159	GUGCUCUCUCUCUUCUGUCAA 179	
	SGN-U352279	1.0	miRNA	21	UACGAGAGAGAGAAGACAGUU 1	Squamosa promoter-binding protein-like 9 (SPL9)
			Target	828	GUGCUCUCUCUCUUCUGUCAU 848	
	SGN-U352752	1.0	miRNA	22	AUACGAGAGAGAGAAGACAGUU 1	Squamosa promoter-binding protein-like 4 (SPL4)
			Target	785	UUUGCUCUCUCUCUUCUGUCAC 806	
car-miR396e	SGN-U355115	0.0	miRNA	21	GCAAGUUCUUUCGGACACCUU 1	Expressed protein, identical to transcription activator GRL1
			Target	218	CGUUCAAGAAAGCCUGUGGAA 238	
car-miR5658a	SGN-U359384	1.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA 1	Zinc finger (C3HC4-type RING finger) family protein
			Target	196	GUCAUCAUCAUUAUUAUCAU 215	
	SGN-U348021	1.5	miRNA	20	CAGUAGUAGUAGUAGUAGUA 1	Protein phosphatase 2C family protein
			Target	168	GUCGUCGUCGUCAUCAUCAU 187	
	SGN-U360366	2.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA 1	Expressed protein
			Target	74	AUCAUCAUCAUCAUCAUCAA 93	
	SGN-U359145	2.5	miRNA	21	ACAGUAGUAGUAGUAGUAGUA 1	Encodes a member of the DREB subfamily A-4 of ERF/AP2 transcription factor family
			Target	393	UGUCGUCGUCGUCGUCAUUAU 413	

car-miR5658b	SGN-U359608	2.0	miRNA	21	UAAGUAGUAGUAGUAGUAGUA	1	Expressed protein
			Target	219	AAUCAUCAUCAUCAUCAUCAG	239	
	SGN-U351946	2.0	miRNA	20	AAGUAGUAGUAGUAGUAGUA	1	Membrane HPP family protein
			Target	283	AGCAUCAUCAUCAUCAUCAU	302	
	SGN-U349079	2.0	miRNA	21	UAAGUAGUAGUAGUAGUAGUA	1	1-Aminocyclopropane-1-carboxylate oxidase
			Target	436	AAUCAUCAUCGUCGUCAUCAU	456	
car-miR5658a/b	SGN-U348631	0.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	Expressed protein
			Target	405	GUCAUCAUCAUCAUCAUCAU	424	
	SGN-U355815	0.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	EST with unknown function
			Target	125	GUCAUCAUCAUCAUCAUCAU	144	
	SGN-U347435	0.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	Protein kinase family protein / non phototropic hypocotyl 1-like protein (NPL1)
			Target	601	GUCAUCAUCAUCAUCAUCAU	620	
	SGN-U349342	0.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	Expressed protein, similar to gda-1
			Target	1430	GUCAUCAUCAUCAUCAUCAU	1449	
	SGN-U353789	0.0	miRNA	21	UAAGUAGUAGUAGUAGUAGUA	1	Citrate synthase
			Target	105	AUUCAUCAUCAUCAUCAUCAU	125	
	SGN-U351641	0.0	miRNA	20	AAGUAGUAGUAGUAGUAGUA	1	Major intrinsic family protein
			Target	1062	UUCAUCAUCAUCAUCAUCAU	1081	
	SGN-U350336	0.0	miRNA	21	UAAGUAGUAGUAGUAGUAGUA	1	EST with unknown function
			Target	551	GUUCAUCAUCAUCAUCAUCAU	571	
	SGN-U349144	0.5	miRNA	20	AAGUAGUAGUAGUAGUAGUA	1	Auxin-responsive protein / indoleacetic acid-induced protein 9 (IAA9)
			Target	99	UUCAUCAUCAUCAUCAUUAU	118	
	SGN-U349954	0.5	miRNA	20	AAGUAGUAGUAGUAGUAGUA	1	Auxin-responsive protein / indoleacetic acid-induced protein 9 (IAA9)
			Target	44	UUCAUCAUCAUCAUCAUUAU	63	
	SGN-U357150	1.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	myb family transcription factor (MYB59)
			Target	448	CUCAUCAUCAUCAUCAUCAU	467	
	SGN-U347371	1.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	Pathogenesis-related protein 10
			Target	214	AUCAUCAUCAUCAUCAUCAU	233	
	SGN-U350705	1.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	Expressed protein
			Target	322	AUCAUCAUCAUCAUCAUCAU	341	
	SGN-U359033	1.0	miRNA	20	CAGUAGUAGUAGUAGUAGUA	1	myb family transcription factor (MYB48)
			Target	452	CUCAUCAUCAUCAUCAUCAU	471	
	SGN-U352679	1.0	miRNA	21	ACAGUAGUAGUAGUAGUAGUA	1	Heterotrimeric G protein gamma-subunit (AGG1)
			Target	146	UCUCAUCAUCAUCAUCAUCAU	166	

[†]psRNATarget is a plant small RNA target analysis server (http://www.plantgrn.noble.org/psRNATarget/; Grun *et al*., 2005).

REFERENCES

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

Ambros, V., Bartel, B., Bartel, D. P., Burge, C. B., Carrington, J. C., Chen, X., Dreyfuss, G., Eddy, S. R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G. and Tuschl, T. (2003). A uniform system for microRNA annotation. *RNA*, **9,** 277-279.

Aukeman, M. J. and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. *Plant Cell,* **15,** 2730–2741.

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

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell*, **136,** 215–233.

Bonnet, E., Wuyts, J., Rouze, P. and Van De Peer, Y. (2004). Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proceedings of the National Academy of Sciences of the USA,* **101,** 11511–11516.

Chen, X. (2004). A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. *Science*, **303,** 2022–2025.

Chen, P. Y., Manninga, H., Slanchev, K., Chien, M., Russo, J. J., Ju, J., Sheridan, R., John, B., Marks, D. S., Gaidatzis, D., Sander, C., Zavolan, M. and Tuschl, T. (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes and Development,* **19,** 1288–1293.

Didiano, D. and Hobert, O. (2006). Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nature Structural and Molecular Biology,* **13,** 849–851.

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., Baum, S. F. and Bowman, J. L. (2003). Radial patterning of *Arabidopsis* shoots by class III *HD-ZIP* and *Kanadi* genes. *Current Biology,* **13,** 1768–1774.

Fujii, H., Chiou, T. J., Lin, S. I., Aung, K. and Zhu, J. K. (2005). A miRNA involved in phosphate-starvation response in *Arabidopsis*. *Current Biology,* **15,** 2038–2043.

Gabriela, T. O., Enamul, H. and Peter, H. Q. (2003). The *Arabidopsis* Basic/Helix-Loop-Helix transcription factor family. *The Plant Cell,* **15,** 1749–1770.

Ghosh, Z., Chakrabarti, J. and Mallick, B. (2007). "miRNomics — the bioinformatics of microRNA genes." *Biochemical and Biophysical Research Communications,* **363,** 6–11.

Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. and Eddy S.R. (2003). Rfam: an RNA family database. *Nucleic Acids Research,* **31,** 439–441

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R. and Bateman, A. (2005) Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Research*, **33,** Database Issue, D121–D124.

Griffiths-Jones, S., Grocock, R. J., Van Dongen, S., Bateman, A. and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, **34,** Database Issue, D140–D144.

Grun, D., Wang, Y. L., Langenberger, D., Kristin, C., Gunsalus, K. C. and Rajewsky, N. (2005). MicroRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Computational Biology,* **1(1),** e13.

Gupta, R., Soni, N., Patnaik, P., Sood, I., Singh, R., Rawal, K. and Rani, V. (2010). High AU content: a signature of up-regulated miRNA in cardiac diseases. *Bioinformation,* **5(3),** 132–135.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment 291 editor and analysis 292 program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* **41,** 95–98.

Hutvagner, G. and Zamore, P. D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science*, **297,** 2056–2060.

Jones-Rhoades, M. W. and Bartel, D. P. (2004). Computational identification of plant microRNAs and their targets including a stress-induced miRNA. *Molecular Cell Biology,* **6,** 787–799.

Kasschau, K. D., Xie, Z., Allen, E., Llave, C., Chapman, E. J., Krizan, K. A. and Carrington, J. C. (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function. *Developmental Cell,* **4,** 205– 217

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* **16,** 11–120.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science,* **294,** 853–8.

Lakshmana-Reddy, D. C., Radhika, V., Bharadwaj, A., Khandagale, K. S. and Aswath C. (2012). miRNAs in brinjal (*Solanum melongena*) mined through an *in silico* approach. J*ournal of Horticultural Science & Biotechnology,* **87 (2),** 186–192.

Lee, R. C. and Ambros, V. (2001). An extensive class of RNAs in *Caenorhabditis elegans*. *Science,* **294,** 862–864.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, **75,** 843–854.

Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120,** 15–20.

Llave, C. (2004). MicroRNAs: more than a role in plant development? *Molecular Plant Pathology,* **5,** 361–366.

Llave, C., Kasschau, K., Rector, M. and Carrington, J. (2002). Endogenous and silencing-associated small RNAs in plants. *The Plant Cell,* **14,** 1605–1619.

Mallory, A.C. and Vaucheret, H. (2004). MicroRNAs: Something important between the genes. *Current Opinion in Plant Biology,* **7,** 120–125.

Mallory, A.C., Dugas, D. V., Bartel, D. P. and Bartel, B. (2004a). MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative and floral organs. *Current Biology,* **14,** 1035–1046.

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P.D., Barton, M.K., and Bartel, D. P. (2004b). MicroRNA control of *PHABULOSA* in leaf development: Importance of pairing to the microRNA 59 region. *EMBO Journal*, **23,** 3356–3364.

Mette, M. F., Van Der Winden, J., Matzke, M. and Matzke, A. J. (2002). Short RNAs can identify new candidate transposable element families in *Arabidopsis*. *Plant Physiology,* **130,** 6–9.

Nasaruddin, N. M., Harikrishna, K., Othman, R. Y., Hoon, L. S. and Harikrishna, J. A. (2007). Computational prediction of microRNAs from oil palm (*Elaeis guineensis* Jacq.) expressed sequence tags. *Asia-Pacific Journal of Molecular Biology and Biotechnology,* **15(3),** 107–113.

Park, W., Li, J., Song, R., Messing, J. and Chen, X. (2002). Carpel factory, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana. Current Biology,* **12,** 1484–1495.

Qiu, C. X., Xie, F. L., Zhu, Y. Y., Guo, K., Huang, S.Q., Nie, L. and YANG Z.M. (2007). Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. *Gene,* **395,** 49–61.

Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. and Bartel, D. P. (2002). MicroRNAs in plants. *Genes and Development,* **16,** 1616–1626.

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

Saitou, N. and Nei, N. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* **4,** 406–425.

Singh, J. and Nagaraju, J. (2008). *In silico* prediction and characterization of microRNAs from red flour beetle (*Tribolium castaneum*). *Insect Molecular Biology,* **17,** 427–436.

Sunkar, R. and Zhu, J. K. (2004). Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell,* **16,** 2001–2019.

Sunkar, R., Kapoor, A. and Zhu, J. K. (2006). Post-transcriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by down-regulation of miR398 and important for oxidative stress tolerance. *Plant Cell,* **18,** 2051–2065.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution,* **28,** 2731–2739.

Vella, M.C., Choi, E.Y., Lin, S.Y, Reinert, K. and Slack, F.J. (2004) The *C. elegans* microRNA *let-7* binds to imperfect *let-7* complementary sites from the *lin-41* 3'-UTR. *Genes and Development,* **18,** 132– 137.

Xie, F. L., Huang, S. Q., Guo, K., Xiang, A. L., Zhu, Y. Y., Nie, L. and Yang, Z. M. (2007). Computational identification of novel microRNAs and targets in *Brassica napus. FEBS Letters,* **581,** 1464–1474.

Yin, Z., Li, C., Han, X. and Shen, F. (2008). Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). *Gene,* **414,** 60–66.

Zhang, B. H., Pan, X. P., Wang, Q. L., Cobb, G. P. and Anderson, T. A. (2005). Identification and characterization of new plant microRNAs using EST analysis. *Cell Research,* **15,** 336–360.

Zhang, B. H., Pan, X. and Anderson, T. A. (2006a). Identification of 188 conserved maize microRNAs and their targets. *FEBS Letters,* **580,** 3753–3762.

Zhang, B. H., Pan, X. P., Cannon, C. H., Cobb, G. P. and Anderson, T. A. (2006b). Conservation and divergence of plant microRNA genes. *Plant Journal,* **46,** 243–259.

Zhang, B. H., Pan, X. P., Cannon, C. H., Cobb, G. P. and Anderson, T. A. (2006c). Computational identification of microRNAs and their targets. *Computational Biology and Chemistry,* **30,** 395–407.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research,* **31,** 3406–3415.

Zuker, M. and Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research,* **9,** 133–148.