The Journal of Animal & Plant Sciences, 25(5): 2015, Page: 1466-1476
ISSN: 1018-7081

IDENTIFICATION AND CHARACTERIZATION OF 25 MICRORNAs AND THEIR TARGETED PROTEINS IN APRICOT (*Prunus armeniaca* L.)

[*1]I. A.Baloch, [1]M. Y. K. Barozai and [1]M. Din

[1]Department of Botany, University of Balochistan, Sariab Road Quetta, Pakistan
Corresponding author's email; iftekharbaloch@gmail.com

ABSTRACT

MicroRNAs (miRNAs) are non-protein coding regulatory RNAs. These are small and endogenous in nature. Their length ranges from 18-26 nucleotides (nt) and they show conserved nature in plants. The conserved nature of miRNAs becomes a logical approach for the hunt of new miRNAs in other organisms by comparative genomics. The miRNAs are reported in large number of plants but very limited number is reported in an important fruit; apricot (*Prunus armeniaca* L). The current study dealt with the identification and characterization of miRNAs and their targeted proteins in apricot. Consequently 25 miRNAs belonging to 24 miRNA families were identified and characterized in *Prunus armeniaca*. All of the identified miRNA families (miR-159, 160, 162,164,166, 167, 172, 390, 393, 394, 395, 396, 408, 477, 482, 535, 3627, 6034, 6266, 6271, 6286, 6482, 6485 and 8123) are being reported for the first time in *Prunus armeniaca*. The precursor miRNAs lengths of newly identified miRNAs range from 51 to 238 nt with an average of 120 nt. The mature miRNA sequences length rang from 20 to 23 nt. A total 33 putative targets for the apricot miRNAs have also been identified. These targets are the significant proteins including transcription factors. Identification of 25 miRNAs and their targeted proteins will be helpful to understand and manage the gene regulation in *Prunus armeniaca*.

Key words: Comparative genomics, Gene regulation, microRNAs, *Prunus armeniaca.*

INTRODUCTION

MicroRNAs (miRNAs), initially named as small temporal RNAs (Hutvagner *et al.,* 2001), are a significant class of non-coding regulatory RNAs. They have endogenous expression in nature and their length ranges from 18-26 nucleotides (nt) (Tang *et al.,* 2003). They play important role in post transcriptional gene regulation (Ambros *et al.,* 2003). The genes of miRNAs are usually transcribed by RNA polymerase II (Lee *et al.,* 2004). The transcription products are named as primary miRNAs (Pri-miRNAs) which may be thousands of nt in length and contain one or more miRNA stem-loops (Lee and Ambros, 2001). This transcript is capped with a specially modified nt at the 5' end, polyadenylated with multiple adenosines at 3' end (Lee *et al.,* 2004) and spliced. The Pri-miRNAs are processed in the nucleus by the microprocessor complex, consisting of the RNase III Drosha, and the double-stranded RNA (dsRNA) binding protein, Pasha/DGCR8 (Denli *et al.,* 2004). After being processed the Pri-miRNA is changed into Precursor-miRNA (Pre-miRNA). The resulting pre-miRNAs are approximately 90-340 nt in length in case of plants. The pre-miRNA hairpins are exported to the cytoplasm where they are further processed into unstable, 18-26 nt miRNA duplex structures by the RNase III protein, Dicer like enzyme (Vaucheret *et al.,* 2012). The less stable of the two strands in the duplex is incorporated into a multiple-protein nuclease complex, the RNA induced silencing complex (RISC), which is also known as microRNA ribonucleoprotein complex: (miRNP) (Schwarz and Zamore, 2002). The RISC with the help of mature miRNA negatively regulates gene expression either by inhibiting translation elongation or by triggering messenger RNA (mRNA) destruction on the basis of the degree of complementarity of miRNA with its target (mRNA) (Aukerman and Sakai. 2003; Tang *et al.*,2003).

In contrast to their name and size, the miRNAs perform mega functions in eukaryotic organisms. They are involved in almost all the cell and biological processes of eukaryotes like; growth and development, organogenesis, (Chen, 2003), transgene suppression (Allen *et al.,* 2005), cell signaling, (Yoshikawa *et al.,* 2005), biotic and abiotic stresses, (Sunkar and Zhu, 2004), and in defense against the attacking microbes (Balmer and Mauch-Mani, 2013). Because of their versatile regulatory functions in eukaryotic organisms, they were named as the mega regulators of eukaryotic genomes (Baloch *et al.,* 2013). Most of the miRNAs are conserved in the animals and plants (Reinhart *et al.,* 2002; Barozai *et al.,* 2008; Rhoades *et al.,* 2011; Barozai, 2012). The conserved nature of the miRNAs becomes a sensible logic for the identification of new orthologues by comparative genomics.

Prunus armeniaca is an important species of family Rosaceae. It is an important fruit and is fragile in nature. As indicated by its botanical name, it is native to Armenia but is cultivated in Pakistan, Turky, Iran, France, Italy and Spain. Almost every part of apricot is utilizable as ingredient in different health and food

products (Kate, 2014). Although the researchers have extensively worked and identified microRNAs in *Prunus persica* (Peach) of the genus "Prunus" e.g. Zhu *et al.,* (2012) identified 47 peach-specific miRNAs, Luo *et al.,* (2013) identified 117 conserved miRNAs and 186 novel miRNA candidates in peach, Zhang *et al.,* (2013) predicted 262 potential microRNAs belonging to 70 miRNA families from the peach but *Prunus armeniaca* (Apricot) remains neglected in this regard. After Surveying the latest releases of microRNA Registry Databases(Version Rfam 20 released June 2013) (Griffiths-Jones, 2004) and Plant MicroRNA database (PMRD**)** (Zhang *et al.,* 2010) it was learned that only 2 miRNAs have been reported in this important species. This created an idea to focus and identify more miRNAs and their targeted proteins in *Prunus armeniaca*. Thus an effort was made utilizing comparative genomics approaches and 25 potential new miRNAs were identified.

MATERIALS AND METHODS

Prediction of candidate pre-miRNA sequences: Use of bioinformatic tools and molecular biology techniques are now a routine and of the most widely used methods for protein interaction studies (Mahmood *et al.,* 2013;Sujay and Angshuman, 2013), molecular diagnosis of pathogens (Razzaq *et al.,* 2013), taxonomic studies (Abou-El-Enain *et al.,* 2014), lineage studies (Zhang and Wang, 2012) prediction of new miRNAs (Baloch *et al.,* 2014) and comparative genomics studies (Barozai *et al.,* 2012; Barozai *et al.,* 2014). This study is also based on comparative genomics approach by applying various bioinformatics tools, same as reported earlier (Barozai, 2012b). The plant pre-miRNAs from the microRNA Registry Database (Version Rfam 20 released June 2013) (Griffiths-Jones, 2004), and PMRD (Zhang *et al.,* 2010) were downloaded and subjected to BLAST against publicly available 15,105 *Prunus armeniaca* ESTs from the database, i.e., dbEST release 130101 at http://blast.ncbi.nlm.nih.gov/Blast.cgi using blastn (Altschul *et al.,* 1990). Adjusted blast parameter settings were as follows: expect values were set at 1000; low complexity was chosen as the sequence filter, database (others: ESTs), organism (*Prunus armeniaca*, taxid*:* 36596), program selection (somewhat similar sequence) and all other parameters were used as default. The FASTA formats of the candidate sequences within the range of 0 - 4 mismatches of mature sequences were saved and single tone EST was created for each miRNA after removing the repeated ESTs of the same gene.

Removal of protein coding sequences: To validate the initial candidate *Prunus armeniaca* miRNAs as non-protein coding, their sequences were subjected for protein homology search against protein database at NCBI using Blastx with default parameter (Stephen *et al.,* 1997). The results were saved and the protein coding sequences were excluded.

Creation of hairpin structures: Zuker folding algorithm, MFOLD (version 3.2) (Zuker, 2003), publicly available at http://www.bioinfo.rpi.edu/applicaions/mfold/rna/form1.cgi, was used to create the hairpin structure of the candidate's sequences. The parameters were adjusted same as reported earlier (Barozai, 2012c). The stem portion of the hairpins were checked for the mature sequences with at least 10 base pairs involved in Watson-Crick or G/U base pairing between the mature miRNA and the opposite strand (miRNA*).

Conservation and Phylogenetic analysis of *Prunus armeniaca* miRNAs: The *Prunus armeniaca* miRNA family MIR-160 conservation and phylogenetic analysis with, *Prunus persica* (ppe), *Oryza sativa* (osa) and *Arabidopsis thaliana* (ath) miRNA orthologues was done with the help of publically available weblogo: a sequence logo generator (Crooks *et al.,* 2004) and ClustalW to generate cladogram tree using neighbor joining clustering method (Larkin *et al.,* 2007) respectively. The results were saved.

Prediction of *Prunus armeniaca* miRNA Targets: We predicted the *Prunus armeniaca* miRNA targets using the same approach as reported earlier (Barozai, 2012d). Briefly, the NCBI Blastn program (Altschul *et al.,* 1990) was applied using newly predicted *Prunus armeniaca* miRNAs as query. The mRNA sequences showing 70% query coverage were selected and subjected to RNA-hybrid, a miRNA target prediction tool (Kruger *et al.,* 2006) for the confirmation of the targets. Only targets having stringent seed site located at either positions 2-7 or/and 8-13 from the 5' end of the miRNA along with the supplementary site and the mfe of the hybridization≤ -15 kcal/mol were selected. The results were saved.

RESULTS AND DISCUSSION

The new conserved *Prunus armeniaca* miRNAs: Total Twenty five new conserved apricots (*Prunus armeniaca*) pre-miRNAs were identified after filtration and completion of the process. The 25 potential Apricot miRNAs belong to 24 families (i-e. miR159, 160, 162,164,166, 167, 172, 390, 393, 394, 395, 396, 408, 477, 482, 535, 3627, 6034, 6266, 6271, 6286, 6482, 6485 and 8123). Most of these microRNAs have also been reported in the *Prunus persica* (peach) with the exception of miR 6034, 6482 and 6485(microRNA Registry Databases, Version Rfam 20 released June 2013 and Plant MicroRNA database) (Griffiths-Jones, 2004; Zhang *et al.,* 2010). Twelve (12) of the newly predicted apricot microRNAs i.e. miR 164a, 167a, 172c, 393b, 396b, 477a, 482c, 535a, 6266, 6271, 6286 and 8123 were predicted by using peach microRNAs as a reference and they showed homology/similarity with that of peach microRNAs suggesting the conserved nature of microRNAs across the species of the same genus.

The empirical formula for biogenesis and expression of the miRNAs, as suggested by Ambros *et al.,* (2003), was used as a criterion to consider the newly predicted 25 Apricot miRNAs as valid candidates. All the identified Apricots pre-miRNAs fulfilled the criteria B, C and D. According to Ambros *et al.,* (2003) only the criterion D is enough for homologous sequences to be validated as new miRNAs in different species. Meyers *et al.,* (2008) further confirmed it in favor of plants miRNA annotation.

***Prunus armeniaca* miRNAs characterization:** As predicted by MFOLD (Zuker, 2003), the minimum folding energy (MFE) of the newly identified *Prunus armeniaca* Pre- miRNAs is one of the key features of miRNAs characterization. The predicted apricot miRNAs in this study have a range from -12 to -72 with an average -32 Kcalmol^{-1}. The pre-miRNAs length ranges from 51 to 238 nt with an average of 120 nt. The mature miRNA sequences lengths range from 20 to 23 nt. Majority (64 % i.e. 16 out of 25) of the *Prunus armeniaca* miRNAs have 21 nt length, followed by 20 nt (24 %), 22 nt (8 %) and 23 nt (4 %). Similar MFE, pre-miRNAs and mature miRNAs lengths ranges were reported in various plants such as carrot (Barozai *et al*., 2013a), tomato (Din and Barozai, 2014a), eggplant (Din and Barozai, 2014b), switchgrass (Xie *et al.,* 2013), flax (Barozai, 2012a) and helianthus (Barozai *et al.,* 2012).The agreements of our results in this study with the previously reported researches strengthens the apricot miRNAs validation. Maximum (48 % i.e 12 out of 25) *Prunus armeniaca* miRNAs are observed to have 4 mismatches with their homologs, followed by 3,2,1 (16% each) and 0 (4%) mismatches. Majority (68% i.e., 17 out of 25) of *Prunus armeniaca* miRNAs are located on the 5' and remaining (32 %) are on the opposite 3' arms of the pre-miRNA secondary structures as illustrated in Figure 1. The predicted miRNA stem-loop structures show that there are at least 9 nucleotides engaged in Watson-crick or G/U base pairings between the mature miRNA and the opposite arms (miRNAs*) in the stem region and the hairpin precursors do not contain large internal loops or bulges. The *Prunus armeniaca* miRNAs characterization such as reference miRNAs (Ref miRNAs), pre-miRNAs length (PL), minimum free folding energies (MFE), maturemiRNA sequences (MS) mature sequence arm (MSA), mature sequence length (ML), number of mismatches (NM), source ESTs (SE), and strand orientation are summarized in Table 1. These findings are similar to the reported works of various groups working on miRNAs (Reinhart *et al.,* 2002; Barozai *et al.,* 2008; Xie *et al.,* 2010;Barozai *et al.,* 2012).To validate these novel miRNAs as strong candidates of miRNAs the relationship between them and known protein is very significant. The *Prunus armeniaca* pre-miRNAs were subjected to Blastx against the protein database at National Center for Biotechnology Information (NCBI) and found no homology with known proteins. This result confirmed our identified pre-miRNAs as strong candidates in *Prunus armeniaca*.

Conservation and Phylogenetic Studies of *Prunus armeniaca* miRNAs: As mir-160 is highly conserved in plants (Jones-Rhoades, 2012), so,the newly identified *Prunus armeniaca* miRNA (mir-160) is subjected for conservation and phylogenic studies. The *Prunus armeniaca* miRNA (mir-160) has shown conservation with *Oryza sativa* (osa), *Prunus persica* (ppe) and *Arabidopsis thaliana* (ath) miRNAs as shown in Figure 2. The Phylogenetic analysis of the same miRNA (mir-160) sequences suggested that the *Prunus armeniaca* is more closed to *Arabidopsis thaliana* as compared to *Oryza sativa* and *Prunus persica* as shown in Figure 3. The results are in agreement with the reported works (Barozai *et al.,* 2008 and Barozai *et al.,* 2012., Ghani*et al.,* 2013).

***Prunus armeniaca* miRNA Targets:** The prediction of novel conserved *Prunus armeniaca* miRNAs targets is a crucial step for validation of miRNAs identified on homology basis. Total 33 targets (Table 2) were annotated for the novel identified *Prunus armeniaca* miRNAs. Almost all of the predicted targets are reported as miRNA targeted proteins in various organisms (Barozai *et al.,* 2008; Xie *et al.,* 2010; Xie *et al.,* 2013).

The transcription factors are the famous and well known class of proteins targeted by miRNAs in almost all plant and animal species (Bonnet *et al.,* 2004; Xie*et al.,* 2010). The novel identified *Prunus armeniaca* miRNAs also target this class of proteins. The predicted *Prunus armeniaca* miRNAs; miR159, 160, 166, 167, 172, 394, 408, 477, 482, 6271, 6286 and 8123putatively target the transcription factors like; S-locus F-box protein 33 (SFB33), S-locus F-box protein 50 (SFB50), R2R3 MYB transcription factor 10, S-locus F-box protein 8, Homeoboxleucine zipper protein (HBLZP), S14 F-box protein, F-box 1, F-box protein 2 and S-locus F-box protein 54 (SFB54). Similar type of transcription factors are reported as miRNAs targets by many research groups in various plants (Xie *et al.,* 2013, Din and Barozai, 2014a,b).

Some important proteins involved in metabolism like; Beta-amylase (AMYB), Polyphenol oxidase pyruvate decarboxylase (PDC), Phytoenedesaturase, Lipoxygenase (LOX), ACC synthase (ACS2), Cinnamate4-hydroxylase (C4H) and Omega-6 fatty acid desaturase (O6FAD) also identified as putative targets of the *Prunus armeniaca* miRNAs (for details Table-2). These findings are in agreements with the targets of plant miRNAs (Xie *et al.,* 2013; Din and Barozai, 2014a, b) previous published work reporting metabolism proteins as putative Other significant proteins playing role in cell signaling such as; Mitogen-activated protein kinase MAPK, cell transport like; Putative sugar transporter and Non-specific lipid transfer protein, biotic and abiotic stress related proteins such as; Powdery mildew

resistance locus O protein, CC-NBS-containing resistance protein and Late embryogenesis-like protein (LEA) are also predicted as putative targets for *Prunus armeniaca* miRNAs (for details Table-2). Almost similar findings were observed by many researchers in various organisms (Chen *et al*., 2012; Ji *et al*., 2012).

CONCLUSION: 25 miRNAs belonging to 24 miRNA families were identified and characterized in *Prunus armeniaca*. The EST based identification is confirmation of their expression. All of the identified miRNA families (miR-159, 160, 162,164,166, 167, 172, 390, 393, 394, 395, 396, 408, 477, 482, 535, 3627, 6034, 6266, 6271, 6286, 6482, 6485 and 8123) are being reported for the first time in *Prunus armeniaca*. This will help us to manage the apricot plant under biotic and abiotic stresses and understand the gene regulation process in apricot.

REFERENCES

Abou-El-Enain, M.M., A.I. El-Nahas, A.Ibrahim, H. Aboel-Atta, H. Latif, S. M. Ahmed ,and M. Atya (2014). Reassessment of the taxonomic relationships between closely related taxa of Papilionoideae. Pure Appl. Biol., 3(1): 32-54.

Allen, E., Z. Xie, J.C. Carrington, and A.M. Gustafson (2005). microRNA directed phasing during transacting siRNA biogenesis in plants. Cell 121: 207–221.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and, D. J. Lipman (1990). Basic local alignment search tool. J. Mol Bio, 215, 403-410.

Ambros, V., B. Bartel, and D.P. Bartel (2003). A uniform system for microRNA annotation. RNA.9: 277-279.

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

Balmer, D, and B.Mauch-Mani (2013). Small yet mighty microRNAs in plant-microbe interactions. MicroRNA, 2(1), 73-80.

Baloch, I. A., and M Din (2014). Bioinformatic hunting of microRNAs. Pure Appl. Biol, 3, 72-80.

Baloch, I. A., M.Y.K. Barozai, and M. Din (2013). MicroRNAs: The mega regulators in eukaryotic genomes. Pure Appl. Biol, 2(3), 83-88.

Barozai, M.Y.K., S.Q.Shah, M.Din, and Raz Muhammad (2014). Codon Usage Bias and RNA Secondary Structures Analysis for Virus Resistant Genes in *Arabidopsis thaliana* and *Oryza sativa*. Pure Appl. Bio., 3(2): 81-91.

Barozai, M. Y. K., Kakar, A. G and Din, M (2012). The relationship between codon usage bias and salt resistant genes in *Arabidopsis thaliana* and Oryza sativa. Pure Appl. Biol, 1(2), 48-51.

Barozai, M.Y.K. (2013). Identification of microRNAs and their targets in *Artemisia annua* L. Pakistan J Bot 45(2):461–465

Barozai, M.Y.K. (2012a). Insilico identification of microRNAs and their targets in fiber and oil producing plant Flax (*Linum usitatissimum* L.) Pakistan J. Bot., 44(4): 1357-1362.

Barozai, M.Y.K. (2012b). Identification and characterization of the microRNAs and their targets in Salmo salar. Gene 499 (1) :163–168

Barozai, M.Y.K. (2012c). The micro RNAs and their targets in the channel catfish (*Ictalurus punctatus*). Mol Biol Rep 39(9):8867–8872

Barozai, M.Y.K. (2012d). The novel 172 sheep (Ovis aries) microRNAs and their targets. Mol Biol Rep 39(5):6259–6266.

Barozai, M.Y.K., I.A. Baloch, and M. Din (2012). Identification of MicroRNAs and their targets in Helianthus. Mol Biol Rep. 39(3):2523–2532

Barozai, M.Y.K., M. Irfan, R. Yousaf, I. Ali, U. Qaisar, A. Maqbool, M. Zahoor, B. Rashid, T. Hussnain, and S. Riazuddin (2008). Identification of micro-RNAs in cotton. Plant Physiol and Biochem., 46(8-9): 739-751.

Bonnet, E., J. Wuyts, P. Rouze, and Y. Van de Peer (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 United States of America, 101(31), 11511-11516.

Chen, L., Y.Ren, Y.Zhang, J.Xu, F.Sun,Z.Zhang, andY. Wang (2012). Genome-wide identification and expression analysis of heat-responsive and novel microRNAs in *Populus tomentosa*. Gene, 504(2), 160-165.

Chen, X.A. (2003). MicroRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Sci 303:2022–2025

Crooks, G.E., G. Hon, J.M. Chandonia, and S.E. Brenner (2004). WebLogo: A sequence logo generator. Genome Research, 14, 1188-1190. doi:10.1101/gr.849004

Denli, A.M., B.B. Tops, H.A. Plasterk, R.F. Ketting, and G.J. Hannon (2004). Processing of primary microRNAs by the Microprocessor complex. Nature. 432(7014):231-5. Epub 2004 Nov 7.

Din, M. and M.Y.K. Barozai (2014a). Profiling microRNAs and their targets in an important fleshy fruit: Tomato *(Solanum lycopersicum).* Gene 535(2):198–203.

Din, M. and M. Y. K. Barozai (2014b). Profiling and characterization of eggplant (*Solanum melongena L.*) micro RNAs and their targets. Molecular Biology Reports 41(2):889–894

Ghani, A., M.Din, I.A. Baloch, and, M.Y.K Barozai (2013). Identification of microRNAs in 12 plant species of Fabaceae. Pure Appl. Biol., 2(3), 104-115.

Table 1. The newly identified *Prunus armeniaca* miRNAs characterization.

Prunus armeniaca miRNAs	Ref miRNAs	PL	MFE kcal/mol	MS with their positions in Precursor	MSA	ML	NM	SE	Strand orientation
par-mir159	mes-mir159b	78	-17.40	47-TTTGGATTGAA**T**GGAGTCTA**T**-67	3’	21	2	CV048007	Minus
par-mir160	mdm-mir160a	128	-29.12	1-TGCCTGGCTCCCTG**C**ATGCCA-21	5’	21	1	CB822973	Plus
par-mir162	ath-mir162b	238	-44.10	201-T**T**GATAAACCTCT**T**CAT**TG**AG-221	3’	21	4	CV047944	Minus
par-mir164	ppe-mir164a	175	-56.90	1-TGGAGAAGCAGGGCACTTGC—21	5’	21	1	CV047356	Minus
par-mir166	mes-mir166i	134	-34.10	5-**A**TGGA**T**CAGGCTTCATT**T**CC**T**-25	5’	21	4	CB821683	Plus
par-mir167	ppe-mir167a	171	-44.90	4-**AC**AAGCTGCCAGC**C**TGATCT**C**-24	5’	21	4	CB819136	Plus
par-mir172	ppe-mir172c	54	- 12.50	6-GGAATC**C**TGATGATGCTGCA**G**-26	5’	21	2	CB823178	Plus
par-mir390	gma-mir390c	130	-27.60	110-CG**G**TC**CC**CA**G**CCTGAGTTTC-129	3’	20	4	CB821515	Plus
par-mir393	ppe-mir393b	91	-20.80	71-TCCAAAGGGATCGCATTG**TTT**-91	3’	21	3	CB823490	Plus
par-mir394	cme-mir394a	65	-15.50	5-TTGGC**G**TTCTGT**G**CACCT**T**C-24	5’	20	3	CV050170	Minus
par-mir395	gma-mir395g	216	-58.60	188-TGAAGTGT**A**T-GGGGAA**T**TT**C**-207	3’	20	4	CV046510	Minus
par-mir396	ppe-mir396b	142	-21.90	3-**A**TC**A**ACA**TT**TTTCTTGAACTT-23	5’	21	4	CB819853	Minus
par-mir408	gma-mir408a	129	-35.20	15-CAGGGGA**CA**A**A**GCAGAG**T**ATG-35	5’	21	4	CV050893	Minus
par-mir477	ppe-mir477a	101	-35.60	2-**G**TTGGGTGCTCTTTTGGG**TGG**-22	5’	21	4	CV049643	Plus
par-mir482	ppe-mir482c	113	-28.60	1-**GTGT**TGGGCTGTTTGGGATG-20	5’	20	4	CB819617	Minus
par-mir482a	mdm-mir482c	141	-72.10	91-TCTTTCCTA**CT**CC**A**CCCATTCC-112	3’	22	3	CV044569	Minus
par-mir535	ppe-mir535a	145	-34.90	5-TGACA**G**A**A**GAGAGAGAGCACGC-35	5’	21	2	CB819319	Plus
par-mir3627	mdm-mir3627b	76	-13.60	1-**GA**GCAGGAGAGAT**T**GCACT**T**-20	5’	20	4	CB819929	Plus
par-mir6034	bna-mir6034	125	-29.10	91-**TG**TGATGT**T**TATAGCTTTGGG-111	3’	21	3	CB820030	Minus
par-mir6266	ppe-mir6266	100	-29.10	10-AAAT**A**CAGGGGCAAAATGATA-30	5’	21	1	CV051795	Plus
par-mir6271	ppe-mir6271	122	-33.80	93-TCAAGATTGAGAGATATAATG-113	3’	21	0	CB820379	Minus
par-mir6286	ppe-mir6286	51	-27.60	1-**A**TTTAACCATTGGATCGAGTTC**A**-23	5’	23	2	CV046689	Plus
par-mir6482	hbr-mir6482	139	-42.30	21-C**T**GGAACTGGTATCAACCCAGC-42	5’	22	1	CV050346	Minus
par-mir6485	hbr-mir6485	71	-13.10	5-T**T**GGATG**C**AG**CC**GAGCATAA-24	5’	20	4	CV051836	Minus
par-mir8123	ppe-mir8123	71	-22.20	11-TGAGCAATGGC**G**CA**G**AG**A**CC**C**-31	5’	21	4	CV049712	Minus

The novel identified *Prunus armeniaca* miRNAs were characterized in terms of PL=Precursor miRNA Length, MFE=Minimum Free Energy, MS=Mature Sequence, MSA=Mature Sequence Arm, ML=Mature sequence Length, NM=Number of Mismatches (shown in bold, blue and enlarged font size), SE=Source EST and strand orientation.

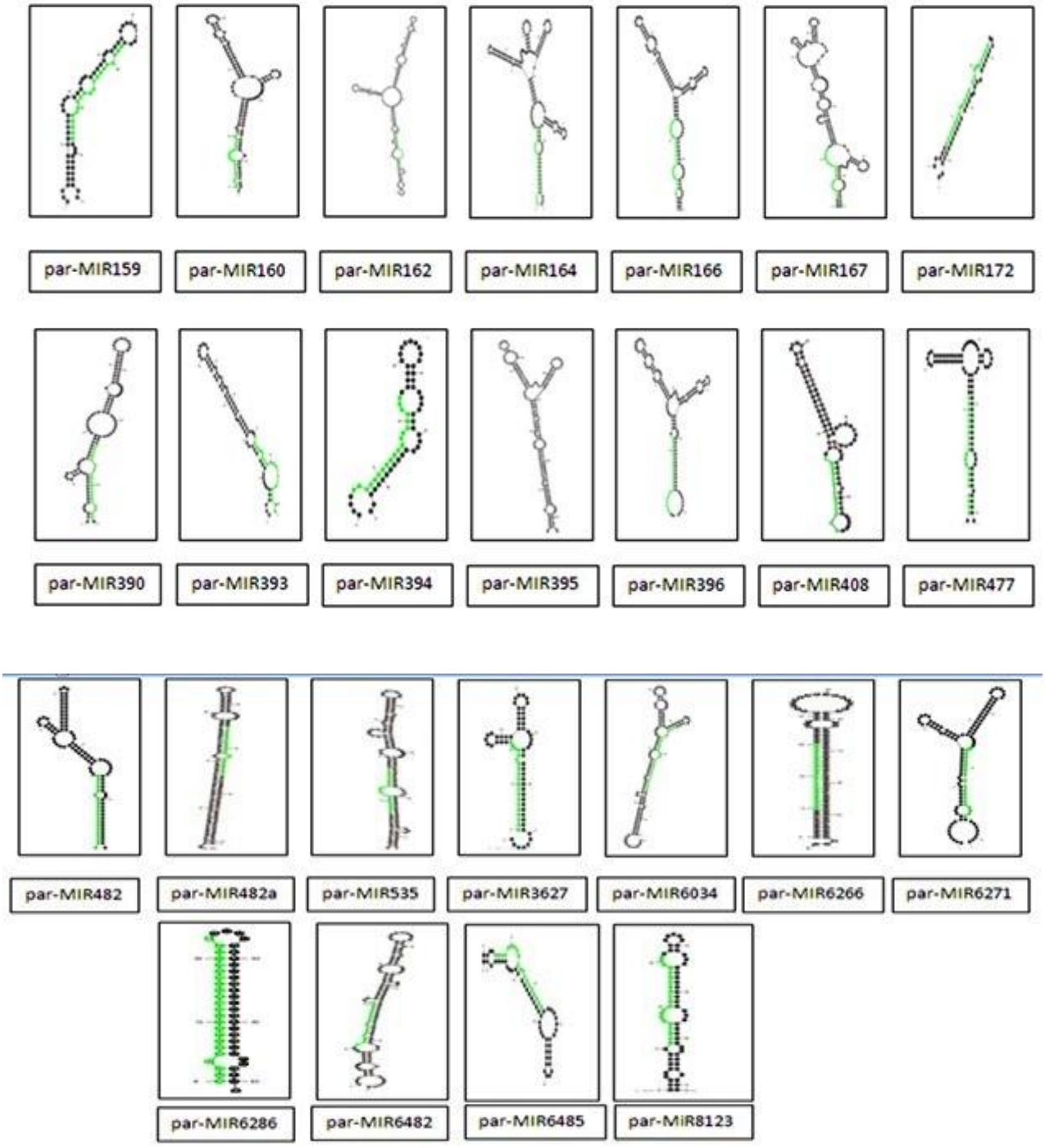

Fig. 1. The novel *Prunus armeniaca* miRNA secondary structures
The *Prunus armeniaca* pre-miRNAs secondary structures are predicted using Mfold algorithm. These structures are clearly showing the mature miRNAs in stem region of the stem-loop structures, highlighted with green.

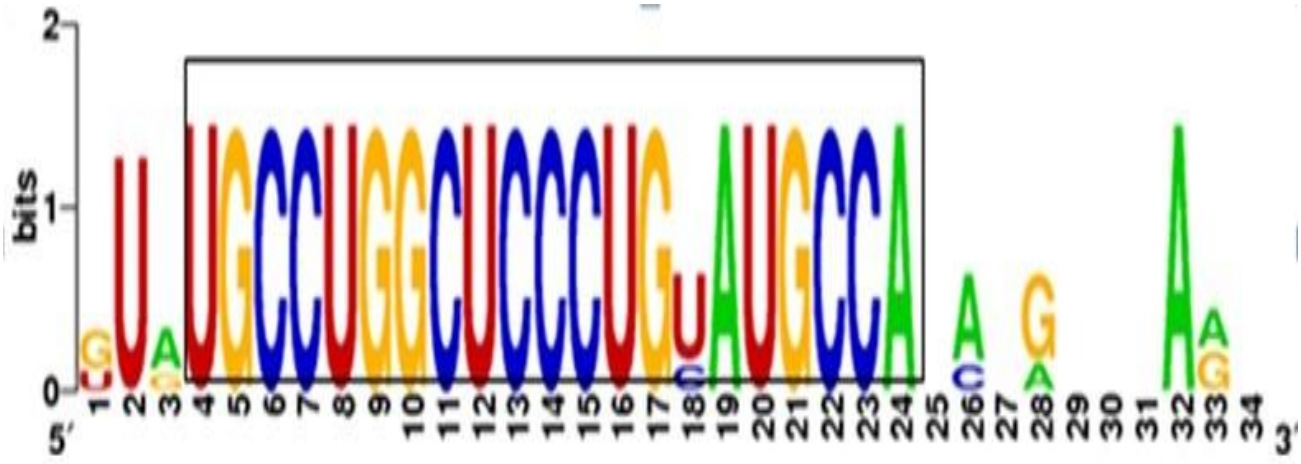

Fig. 2. The *Prunus armeniaca* miRNA conservation studies.

Alignment of the *Prunus armeniaca* pre-miRNA (160) with *Oryza sativa* (osa), *Prunus persica* (ppe) and *Arabidopsis thaliana (ath)*miRNAs using Weblogo: a sequence logo generator, showing miRNA sequences conservation. The conserved mature sequence is highlighted in a box.

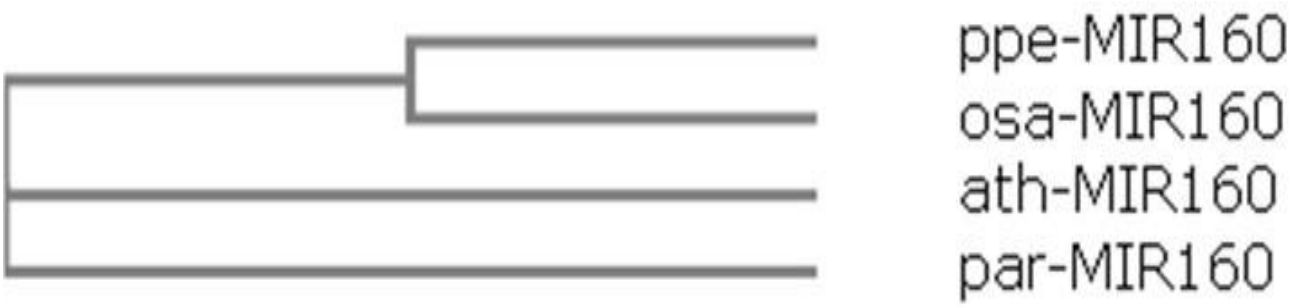

Fig. 3. The *Prunus armeniac a*miRNA phylogenetic analysis.

The Phylogenetic analysis of the *Prunus armeniaca* pre-miRNAs (160) with *Oryza sativa* (osa), *Prunus persica* (ppe) and *Arabidopsis thaliana (ath)* miRNAs, was done with the help of ClustalW and cladogram tree was generated using neighbor joining clustering method. The phylogenetic tree shows that on the basis of pre-miRNA sequences, the *Prunus armeniaca* is more closed to *Arabidopsis thaliana* as compared to *Oryza sativa* and *Prunus persica*

Table 2 .Putative *Prunus armeniaca* miRNA targets. The *Prunus armeniaca* miRNA families and their putative targeted proteins function, Genbank Acc. and RNA-hybrid results are provided. The seed regions 2-7 and 8-13 are shown in red and blue font respectively.

Prunus armeniaca miRNA family	Target Acc	Target Description	RNA-Hybrid Result with MFE of the hybridization
par-miR159	AF139501	Beta-amylase (AMYB)	**mfe: -25.0 kcal/mol** target 5' A U 3' AGACUCCA AUCCAAG UCUGAGGU **UAGGUU**U miRNA 3' UA AAGU 5'
	HQ148064.3	S-locus F-box protein 33 (SFB33)	**mfe: -23.5 kcal/mol** target 5' U UG C 3' G UCC UUCGAUUCAGA C AGG AAGU**UAGGUU**U miRNA 3' UAU UG U 5'
	HQ148074.3	S-locus F-box protein 50 (SFB50)	**mfe: -21.0 kcal/mol** target 5' A UAA CA U 3' GGACUUU UCGA UCCAAG UCUGAGG AGU**U** **AGGUU**U miRNA 3' UA UA 5'
	JX546224.1	SOC1 protein	**mfe: -20.7 kcal/mol** target 5' U UU A 3' AUAG UUCAUUUAGUUUAGG UAUC AG**GUAAGUUAGGUU**U miRNA 3' UG 5'
par-miR160	AF134733	Calcium-binding protein calreticulin	**mfe: -20.7 kcal/mol** target 5' A A GAU UAAAC U 3' GG CG GGGA GUUGGGU CC GU CCCU C**GGUCCG** miRNA 3' A ACGU U 5'
	EU153578.1	R2R3 MYB transcription factor 10	**mfe: -26.6 kcal/mol** target 5' U UUUU 3' UGG UGUGU GGG UCAGGC ACC GUACG CCC **GGUCCG** miRNA 3' U UC U 5'
	AF000952.1	Putative sugar transporter	**mfe: -23.2 kcal/mol** target 5' U UGU U C 3' CAUGC G GGA CUAGGU GUACG C CCU **GGUCCG** miRNA 3' ACC U C U 5'
	GU189299.1	Powdery mildew resistance locus O protein	**mfe: -32.0 kcal/mol** target 5' C UC U 3' GGCAUGC GCCAGGU CCGUACG C**GGUCCG** miRNA 3' A UCCCU U 5'
par-miR162	KF177397	MLO4 (MLO4) mRNA	**mfe: -25.7 kcal/mol** target 5' U GG GGAG C 3' UCAAUGAGGAGGU UA CA AGUUACU**UCUCCA** AU GU miRNA 3' G A A U 5'

par-miR164	AF139496	p78RF mRNA	**mfe: -21.4 kcal/mol** target 5' U CAAG AGGAA C 3' GCAGG CCUGC UUUUCUA CGUUC GGACG **AAGAGG**U miRNA 3' ACG 5′
par-miR166	EU153578	R2R3 MYB transcription factor 10 gene	**mfe: -23.2 kcal/mol** target 5' A CUU AAUUAUUA C 3' AG GUGAAGCCU GAUCUA UC UAC**UUCGGA** **CUAGGU** miRNA 3' CUU A 5'
par-miR167	AF020786.1	Polyphenol oxidase	**mfe: -30.0 kcal/mol** target 5' G AAUAUCC UU CUC A 3' GAGA UCA GGC GGCGGCUUGU CUCU AGU CCG CCG**UCGAAC**A miRNA 3' A 5'
	FJ477418.1	S-locus F-box protein 8	**mfe: -21.1 kcal/mol** target 5' A AA C U A 3' GG AUUA GUUGGC G CUUG UC UAGU **CGACCG** C GAAC miRNA 3' C C U A 5'
par-miR172	HQ148064	S-locus F-box protein 33 (SFB33)	**mfe: -20.7 kcal/mol** target 5' U UCUCCA C A 3' UGC AUU GGGUUCC ACG UAG **CCUAAG**G miRNA 3' G UCGUAG U 5'
par-miR390	EU395434	pyruvate decarboxylase (PDC)	**mfe: -30.5 kcal/mol** target 5' C GU A A 3' ACU GG GGGGACCG UGA CC C**CCCUGGC** miRNA 3' CUU GU GA 5'
par-miR393	AF071894.1	Late embryogenesis-like protein (LEA)	**mfe: -28.6 kcal/mol** target 5' U ACACG G 3' CAAUGUGG CCCUUUGGA GUUACGCU GGGAAACCU miRNA 3' UUU A 5'
	HM234047.1	Non-specific lipid transfer protein	**mfe: -19.0 kcal/mol** target 5' U CUGA U 3' GACAAU UGA CUUUUGGG UUGUUA GCU G**GAAACC**U miRNA 3' U C AG 5'
par-miR394	AY822065.1	Phytoenedesaturase	**mfe: -22.4 kcal/mol** target 5' U GU U U 3' GAAGGUG U GCGUUGA CUUCCAC G **UGCGGU**U miRNA 3' GU UCU 5'
par-miR394	AF139497.1	Homeoboxleucine zipper protein (HBLZP)	**mfe: -21.4 kcal/mol** target 5' U A 3' GGU CACAGAA CCAA CCA **GUGUCU**U GGUU miRNA 3' CUU C GC 5'
par-miR395	EU439430.1	Lipoxygenase (LOX)	**mfe: -25.4 kcal/mol** target 5' A G 3' UCCCAUACACUUU G**GGGUAUGUGAAG** miRNA 3' CUUUAA U 5'
par-miR396	AF184077.1	ACC synthase (ACS2)	**mfe: -19.2 kcal/mol** target 5' C C A 3' UUAAGGAGG UGUUGA AGUUCUUUU **ACAACU** miRNA 3' UUCA U A 5'

par-miR408	EU153578	R2R3 MYB transcription factor 10	**mfe: -24.4 kcal/mol** target 5' G A G GG 3' GUACU CUGCU UUG CC UUG UAUGA GACGA AAC GG GAC miRNA 3' G A G 5'
par-miR477	GQ336821.1	Putative CC-NBS-containing resistance protein	**mfe: -22.7 kcal/mol** target 5' A U 3' CUAAAAG GUACCUAG GGUUUUC CGUGGGUU miRNA 3' GGUG U G 5'
	DQ897929.1	S14 F-box protein	**mfe: -24.1 kcal/mol** target 5' A A CCUAU A 3' UCAUUCAGAA GGU GCUCAGC GGUGGGUUUU UCG UGGGUUG miRNA 3' C 5'
par-miR482	DQ422943.1	F-box 1	**mfe: -24.2 kcal/mol** target 5' C UAA C 3' AUGG UGG GUGGGAAAG UACC ACC CAUCCUUUC miRNA 3' CCU C U U 5'
par-miR535	JX262381.1	NADP-malic protein	**mfe: -28.4 kcal/mol** target 5' A A UUC G A 3' GC UCUCUCUCU UCUGU A CG AGAGAGAGA AGACA U miRNA 3' CGCA G 5'
par-miR3627	HM204477.1	cinnamate 4-hydroxylase (C4H)	**mfe: -27.5 kcal/mol** target 5' U UUG AG A 3' GUGC AA UUUCCUGCUC CACG UU AGAGGACGAG miRNA 3' UU AG 5'
par-miR6034	AF071892.1	Omega-6 fatty acid desaturase (O6FAD)	**mfe: -22.5 kcal/mol** target 5' A UUCCAC G 3' CAAGGUU AACAUCACA GUUUCGA UUGUAGUGU miRNA 3' GG UAU 5'
par-miR6271	DQ422943.1	F-box 1	**mfe: -19.8 kcal/mol** target 5' G A CG C 3' CGUUAUGUC UUCGAUU GG GUAAUAUAG GAGUUAG CU miRNA 3' A AA 5'
par-miR6286	EF053407.1	F-box protein 2	**mfe: -21.8 kcal/mol** target 5' A UCUU A 3' GAAC AUCC UGGUUAAA CUUG UAGG ACCAAUUU miRNA 3' A AGC UU A 5'
par-miR6482	AF134730.1	Mitogen-activated protein kinase MAPK	**mfe: -26.8 kcal/mol** target 5' U A C AGUAC A 3' UUG GGUUGA ACCA GUUCCG GAC CCAACU UGGU CAAGGU miRNA 3' C A C 5'
par-miR6485	AY962818.1	Arginase	**mfe: -20.7 kcal/mol** target 5' U C AUC C 3' UG CUUG CUGCAUUCG AC GAGC GACGUAGGU miRNA 3' AAU C U 5'
par-miR8123	HQ148078.1	S-locus F-box protein 54 (SFB54)	**mfe: -23.9 kcal/mol** target 5' A UA U 3' UUUC GC CAUUGCUC AGAG CG GUAACGAG miRNA 3' CCC A CG U 5'

Griffiths-Jones, S (2004). The microRNA Registry. Nucleic Acids Res., 32D: 109-111.

Hutvagner, G., J. McLachlan, A. E. Pasquinelli, É. Bálint, T. Tuschl and P. D. Zamore (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Sci., 293(5531), 834-838.

Ji, Z., G. Wang, Z. Xie, C. Zhang, and J. Wang (2012). Identification and characterization of microRNA in the dairy goat (Capra hircus) mammary gland by Solexa deep sequencing technology. MolBiol Rep 39(10):9361–9371

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. Plant molecular biology, 80(1), 3-16.

Kate, A. E., U.C. Lohani, N.C. Shahi, J.P. Pandey, A. Sarkar, and K. Chand (2014). Effects of moisture content and feed rate on milling characteristics of wild apricot pits (*Prunus armeniaca* L.). International J. Agricultural and Biological Engineering, 7(1).

Kruger, J. and M. Rehmsmeier (2006). RNAhybrid: MicroRNA target prediction easy, fast and flexible. Nucleic Acids Research, 34, Supplement 2, W451-W454.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, and *et al* (2007). ClustalW and ClustalX version 2.Bioinformatics, 23, 2947-2948.

Lee, R.C. and V. Ambros (2001). An extensive class of small RNAs in *Caenorhabdits elegans*, Sci, 294 (5543), 862–864

Lee, Y., M. Kim, J. Han, K.H. Yeom, S. Lee,S.H.Baek, and V. Kim(2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 23. 4051–4060

Luo, X., Z.Gao, T.Shi,Z. Cheng, Z. Zhang, andZ.Ni (2013). Identification of miRNAs and their target genes in peach (*Prunus persica* L.) using high-throughput sequencing and degradome analysis. PloS one, 8(11), e79090.

Mahmood U.R., W.A. Shamshari, S. Ali, K. Hussain, M. Qasim, andS.A. Bukhari (2013). Interaction studies of heat shockresponsive protein HSA32 and heat shock protein HSP101:an in silico approach. Pure Appl Biol 2: 126–131.

Meyers., B.C., M.J. Axtell, B. Bartel,D.P.Bartel, D.Baulcombe, J.L. Bowman,... and J.K. Zhu (2008). Criteria for annotation of plant MicroRNAs. The Plant Cell, 20, 3186-3190.

Ray, S. and A.Bagchi (2013). Exploring the role of SoxA and SoxX in sulphur oxidation in Allochromatiumvinosum through Protein-protein docking: An in silico approach.Pure Appl. Biol., 2(2): 53-56.

Razzaq, A., M. Irfan, M.Mohsin, and K.A.Malik (2013). Molecular diagnostics of food borne pathogens. Pure Appl. Biol., 2(2): 69-75.

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

Rhoades, M.W (2011). Conservation and divergence in plant microRNAs. Plant Mol Biol., DOI 10.1007/s11103-011-9829-2

Schwarz, D.S and P. D. Zamore (2002). Why do miRNAs live in the miRNP?. Genes &dev, 16(9), 1025-1031.

Stephen, F.A., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman (1997). Gapped BLAST and PSI-BLAST, A new generation of protein database search programs. Nucleic Acids Res., 25: 3389-3402.

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

Tang, G., B. J. Reinhart, D. P. Bartel and P. D. Zamore (2003). A biochemical framework for RNA silencing in plants. Genes dev., 17(1), 49-63.

Vaucheret, H., A. Mallory and G. Lepere (2012). European Patent No. EP 2436768. Munich, Germany: European Patent Office.

Xie, F., C. N. Stewart, F. A. Taki, Q.He, H. Liu and B. Zhang (2013). High-throughput deep sequencing shows that microRNAs play important roles in switchgrass responses to drought and salinity stress. Plant Biotechnology J.

Xie, F., T.P. Frazier and B. Zhang (2010). Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). Planta, 232, 417-434. doi:10.1007/s00425-010-1182-1

Yoshikawa, M., A. Peragine, M.Y. Park and R.S. Poethig (2005). A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. Genes Dev 19:2164–2175.

Zhang, Y., Bai, Y., Han, J., Chen, M., Kayesh, E., Jiang, W., and Fang, J (2013). Bioinformatics prediction of miRNAs in the *Prunus persica* genome with validation of their precise sequences by miR-RACE. J. Plant Physiology, 170(1), 80-92.

Zhang, Z., J. Yu, D. Li, Z. Zhang, F. Liu, X. Zhou and Z. Su (2010). PMRD: plant microRNA database. Nucleic acids research, 38 (suppl 1), D806-D813.

Zhang,T., and Wang, J. (2012). Identification of Putative Hybrids and Natural Lineages in Genus Potamogeton Revealed By Chloroplast and Nuclear DNA Markers. Pure Appl. Bio., 1(1), 1-7.

Zhu H, Xia R, Zhao B, An YQ, Dardick CD, Callahan AM, Liu Z (2012). Unique expression,

processing regulation, and regulatory network of peach (*Prunus persica*) miRNAs" BMC Plant Biol. 12:149.

Zuker, M (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res., 31:3406-3415.