TÜBİTAK

Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2017) 41: 191-200

doi:10.3906/tar-1612-121

Bioinformatics profiling and characterization of potential microRNAs and their targets in the genus *Coffea*

Farzana BIBI, Muhammad Younas Khan BAROZAI, Muhammad DIN*
Department of Botany, University of Balochistan, Quetta, Pakistan

Received: 30.12.2016 • **Accepted/Published Online:** 22.05.2017 • **Final Version:** 14.06.2017

Abstract: MicroRNAs (miRNAs) are a class of noncoding small endogenous RNAs with lengths of 18 to 26 nucleotides that have been shown to regulate gene expression at posttranscriptional levels by targeting mRNAs for degradation or by inhibiting protein translation. Although thousands of miRNAs have been identified in many species, limited miRNAs have been reported in the genus *Coffea*, a member of the family Rubiaceae that is endemic to tropical Asia and South Africa. The genus *Coffea*, whose seeds are known as coffee beans, is used to make coffee. In this study, we identified 51 potential genus *Coffea* miRNAs, belonging to 51 families, using a well-defined comparative genome-based computational approach in genus *Coffea* expressed sequences tags. These identified miRNAs potentially target 150 protein-coding genes, which can act as transcription factors and take part in multiple biological and metabolic processes, hypothetical proteins, signal transduction, transporters, growth and development, stress-related processes, structural constituents, and disease-related processes. The results of this research may contribute to the understanding of the miRNA-mediated life processes in the genus *Coffea*.

Key words: MicroRNAs, comparative genomics, *Coffea arabica*, *Coffea canephora*, *Coffea racemosa*

1. Introduction

MicroRNAs (miRNAs) are an extensive class of noncoding small endogenous RNAs of 18 to 26 nt in length that are derived from self-complementary fold-back structures of longer precursor sequences (pre-miRNAs) and are generated by Dicer-like 1 (DCL1) in plants (Bartel, 2004). Mature miRNAs inhibit gene expression at posttranscriptional levels by either targeting mRNAs for degradation or inhibiting protein translation. Both processes are accomplished by the complementary base pairing of miRNAs to their target mRNA sequences (Ambros, 2004). In plants, for a majority of cases, miRNAs interact with their targets through perfect or near-perfect base pairing and lead to target mRNA degradation (Jones-Rhoades et al., 2006). Increasing evidence has revealed that miRNAs play an important role in a wide range of development processes in plants, including cell proliferation, stress response, metabolism, inflammation, and signal transduction (Ambros, 2004; Jones-Rhoades et al., 2006; Zhang et al., 2007a; Ali et al., 2016) and cross-kingdom gene regulation (Barozai and Din, 2017). To date, more than 28,645 miRNAs have been identified from 223 species of plants and animals and deposited in the publicly available database miRBase (Release 21) (Griffiths-Jones et al., 2008). The majority of plant miRNAs have been found in species with fully sequenced genomes including 713 from *Oryza sativa*, 401 from *Populus trichocarpa*, 384 from *Arabidopsis thaliana*, 343 from *Solanum tuberosum*, 321 from *Zea mays*, and 241 from *Sorghum bicolor* (Griffiths-Jones et al., 2008). miRNA-related research is continuously growing and miRNAs, along with their functions, are being identified and elucidated using a wide variety of computational tools and experimental methods including direct cloning, deep sequencing, and other approaches. Comparison of miRNAs across multiple plant species has demonstrated that some miRNAs are highly evolutionary conserved from species to species, such as from mosses to higher flowering eudicots in the plant kingdom (Floyd and Bowman, 2004; Zhang et al., 2006a, 2006b). Conservation of miRNA sequences has provided a powerful strategy for identifying miRNAs in other species. Currently, comparative genome-based homolog searches have been used to identify conserved miRNAs in many plant species, including cotton (Barozai et al., 2008), mustard (Xie et al., 2007), soybean (Zhang et al., 2008), wheat (Jin et al., 2008), tomato (Din and Barozai, 2014b), potato (Din et al., 2014), apricot, rose (Baloch et al., 2015a, 2015b), and *Helianthus* (Barozai et al., 2011a). The genus *Coffea* is member of the

* Correspondence: drmdin75@gmail.com

family Rubiaceae, which has 500 genera and over 6000 species (Friedman and Waller, 1983). They have different life forms that are shrubs or trees. The genus *Coffea* is endemic to tropical Asia and South Africa. Its seeds are known as coffee beans and are used to make coffee. Coffee is among the most important and valuable agricultural crops and is a major export product of numerous countries. It also has medicinal properties because of secondary metabolites like polyphenols, caffeine, theobromine, and chlorogenic acids (Ferrazzano, 2009). To our knowledge, although progress has been made on the genus *Coffea*, there is little knowledge about miRNAs in *Coffea* (Rebijith et al., 2013; Akter et al., 2014; Loss-Morais et al., 2014). In this study, we employed a well-defined comparative genome-based homolog search to identify *Coffea* miRNAs. We also investigated the potential functions of predicted *Coffea* miRNAs, particularly in the biological and metabolic processes.

2. Methods and materials

2.1. Fetching of reference miRNA sequences

A total of 2631 known plant miRNAs were downloaded and used as a reference miRNA set for identifying conserved miRNAs in the genus *Coffea*. These miRNAs are all currently available miRNAs deposited in the miRBase database (http://www.mirbase.org/, Release 21: June 2014) (Griffiths-Jones et al., 2008); these miRNAs come from 15 plant species, including *Acacia auriculiformis* (aau), *Arachis hypogeae* (ahy), *Arabidopsis lyrata* (aly), *Arabidopsis thaliana* (ath), *Brassica napus* (bna), *Brassica oleracea* (bol), *Brassica rapa* (bra), *Carica papaya* (cpa), *Cucumes melo* (cme), *Glycine max* (gma), *Hevea brasiliensis* (hbr), *Rehmannia glutinosa* (rgl), *Nicotiana tabacum* (nta), *Solanum lycopersicum* (sly), and *Solanum tuberosum* (stu).

The genus *Coffea* expressed sequences tags (ESTs) and protein databases were obtained from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/, NCBI). Currently, a total of 254,179 ESTs for *Coffea* (*Coffea arabica* 174,275; *Coffea canephora* 69,066; and *Coffea racemosa* 10,838) are available in the NCBI dbEST (release 130101, 1 January 2013). All EST sequences were used for predicting conserved miRNAs as well as for identifying potential miRNA targets.

2.2. Bioinformatics tools used for prediction of miRNAs

The similarity search tool BLASTn and BLASTx programs were employed to identify potential conserved miRNAs and used for removing repeated sequences and protein coding genes, available at https://blast.ncbi.nlm.nih.gov/Blast.cgi (Altschul et al., 1997). Mfold was obtained from http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form (Zuker, 2003) for the prediction of secondary structures. Conservation and phylogenetic analyses were done by the publicly available WebLogo, a sequence logo generator (Larkin et al., 2007), and ClustalW (Crooks et al., 2004) to generate a cladogram tree using neighbor-joining clustering, respectively. For data mining in the identification of miRNAs' targets, psRNATarget, a plant small RNA target analysis server available at http://plantgrn.noble.org/psRNATarget/ (Dai and Zhao, 2011), and RNAhybrid, a miRNA target prediction tool (bibiserv2.cebitec.uni-bielefeld.de/rnahybrid) (Kruger and Rehmsmeier, 2006), were used. The GO database was downloaded from the Gene Ontology website (http://www.geneontology.org/GO.downloads.shtml) for the confirmation of miRNAs' functions.

2.3. Detection of raw sequences of miRNAs from ESTs

Comparative genome-based EST analysis is a well-established approach to find new interesting facts (Barozai and Husnain, 2011; Barozai et al., 2014; Shah et al., 2016) and identify conserved miRNAs in one species using already known miRNAs of another species (Barozai et al., 2011b, 2011c). Since it was developed, EST analysis has been widely used to identify conserved miRNAs in many plant species, including cotton, soybean, radish, tomato, chili, and potato. Thus, BLASTn (Altschul et al., 1990) searches are an ideal tool to identify conserved small RNA sequences, including miRNAs. All of the databases and software were downloaded from the previously mentioned websites. Briefly, we used the BLASTn program to align all known mature miRNA sequences to all genus *Coffea* EST sequences in order to identify potential homologs with a maximum of 4 nt substitutions with the mature reference sequences, including deletion and insertion mutations. After removing the repeated and protein-coding sequences and considering proper secondary structures, only the sequences fitting the following criteria were considered as potential miRNAs in *Coffea*: 1) there were maximum of 4 nt substituted between the EST sequence and the query miRNA sequence; 2) the minimum length of the pre-miRNA was 44 nt; 3) the pre-miRNA could be folded into a perfect stem-loop hairpin secondary structure with the miRNA sitting in one arm of the stem at either the 5' or 3' end; 4) there were no more than 7 nucleotides mismatched between the predicted mature miRNA sequence and its opposite miRNA* sequence (miRNA : miRNA*) in the secondary structure; 5) there were no loops or breaks in the miRNA : miRNA* complex; and 6) the predicted pre-miRNA sequences had a high minimal folding energy (MFE). Using these criteria, we could significantly reduce the total number of sequences for subsequent analyses, ultimately saving time and increasing work efficiency. More importantly, the application of these criteria significantly reduced the total number of false miRNA predictions. The organ of expression for each profiled potential new *Coffea* miRNA was identified from its EST generated organ and/or tissue of coffee plant.

2.4. Prediction of miRNA targets in the genus *Coffea*

Growing evidence has shown that most plant miRNAs function by either perfect or near-perfect binding to complementary sites on their target mRNA sequences (Schwab et al., 2005). This provides a powerful tool for identifying potential miRNA targets simply by aligning and comparing miRNAs with potential target sequences. The criteria for prediction of potential miRNA targets in *Coffea* were similar to those described by Zhang et al. (2008) with some modifications as described by Din et al. (2016). BLASTn was employed as an alignment tool to predict miRNA target sequences. All *Coffea* ESTs were also used to predict potential miRNA targets. In this research, the following criteria were used for identifying potential miRNA targets: 1) no more than four mismatches were allowed between the mature miRNA and its potential target site; 2) no more than two consecutive mismatches were allowed; 3) no mismatches were allowed at positions 10 and 11; and 4) no more than one mismatch was allowed at nucleotide positions 1–9. These criteria significantly reduced the total number of false positive targets (Xie et al., 2010).

3. Results and discussion

3.1. Identification of potential miRNAs in the genus *Coffea*

As reference miRNAs, a total of 2631 miRNAs from 15 plant species deposited in the miRBase database (Version 21, released July 2014) (Griffiths-Jones, 2008) were downloaded and subjected for alignment against 254,179 ESTs of *Coffea* from dbEST (database of ESTs), release 130101, January 2013, at http://blast.ncbi.nlm.nih.gov/Blast.cgi, using the BLASTn program (Altschul et al., 1990). After removing the repeated sequences, and considering proper secondary structures, we were able to identify 51 conserved miRNAs in three species of *Coffea*: 44 in *Coffea arabica*, 6 in *Coffea canephora*, and 1 in *Coffea racemosa*, which belong to 51 miRNA families (Table 1). This indicates that miRNAs widely exist in the genus *Coffea* and further demonstrates that many miRNAs are highly evolutionary conserved among species in the plant kingdom. Similarly, only one miRNA was reported by Akter et al., (2014) in *Coffea arabica*, 58 miRNA sequences belonging to 33 miRNA families from *Coffea canephora* and *Coffea arabica* were reported by Loss-Morais et al. (2014), and 18 miRNAs belonging to 14 miRNA families from *Coffea arabica* were reported by Rebijith et al., (2013). The miRNA families miR-160, -162, -164, -165, -171, -397, -408, -479, and -2111 were reported for only mature miRNA sequences by Loss-Morais et al. (2014) and were also found in this study with pre-miRNA sequences and secondary self-folded stem-loop structures.

3.2. Characterization of microRNAs in the genus *Coffea*

The 51 identified *Coffea* miRNAs belong to 51 families. For all miRNA families, a single member has been identified (Figure 1). Mature miRNA sequences have been shown to be located on either arm of the secondary stem-loop hairpin structure of the potential pre-miRNA. Of the 51 identified *Coffea* miRNAs, 28 (55%) were found to be located on the 3' arm of the stem-loop hairpin structure while 23 (45%) resided on the 5' arm. The length of *Coffea* miRNAs varies from 19 to 24 nt with an average of 21 nt (Figure 1). A majority (59%, or 30 out of 51) of miRNAs are 21 nt in length, followed by 20% of 22 nt, 12% of 24 nt, 4% of 19 and 20 nt respectively, and 2% of 23 nt. The length of *Coffea* pre-miRNA also varies from 44 to 253 nt with an average of 98 nt. However, a majority of the pre-miRNAs are 40 to 82 nt (26 out of 51, or 51%) in length (Table 1), then 83 to 125 = 12 (23%), 126 to 168 = 10 (20%), 169 to 211 = 1 (2%), and 212 to 254 = 2 (4%). Mature miRNA arm locations, length distribution of miRNAs, and their precursor sequences are similar to previous reports in other plant species (Zhang et al., 2006b, 2007b, 2008; Din et al., 2014). MFE is very important for RNAs forming their secondary structures. Generally speaking, the lower the MFE, the more stable the secondary structure of a RNA sequence. The average value of MFEs was –9 to –67 kcal mol^{-1} with an average of –24 kcal mol^{-1} (Table 1). New conserved *Coffea* miRNAs were also characterized for the acceptable range (Zhang et al., 2006a) of mismatches between the reference miRNA and potential *Coffea* miRNAs. The range of mismatches was reported from 0 to 4 nt with an average of 3 nt. A maximum of 39% (20 out of 51) of the miRNAs were observed to have 3 mismatches with their homologs, followed by 4 = 16 (31%), 2 = 11 (22%), 0 = 3 (6%), and 1 = 1 (2%) mismatches. The potential *Coffea* miRNAs were further characterized for Watson–Crick or G/U base pairing. The *Coffea* miRNA secondary structures showed that there were at least 17 nucleotides engaged in Watson–Crick or G/U base pairings between the mature miRNA and the opposite arms (miRNAs*) in the stem region except a few, where the reference miRNAs also have fewer base pairings and secondary structures were not found with large internal loops or bulges. The *Coffea* miRNAs were also characterized in terms of organ of expression, like flower = 27% (14 out of 51), leaves = 25% (13 out of 51), callus = 18% (9 out of 51), suspension cells = 12% (6 out of 51), mixed tissues and seeds = 6% (3 out of 51) each, and fruits and stems = 2% (2 out of 51). Similar ranges of MFEs, numbers of mismatches, Watson–Crick or G/U pairings, and organs of expression have been reported in various plants by different researchers in *Artemisia annua* (Barozai, 2013), chili (Din et al., 2016), tomato (Din and Barozai, 2014b), eggplant (Din and Barozai, 2014a), and potato (Din et al., 2014).

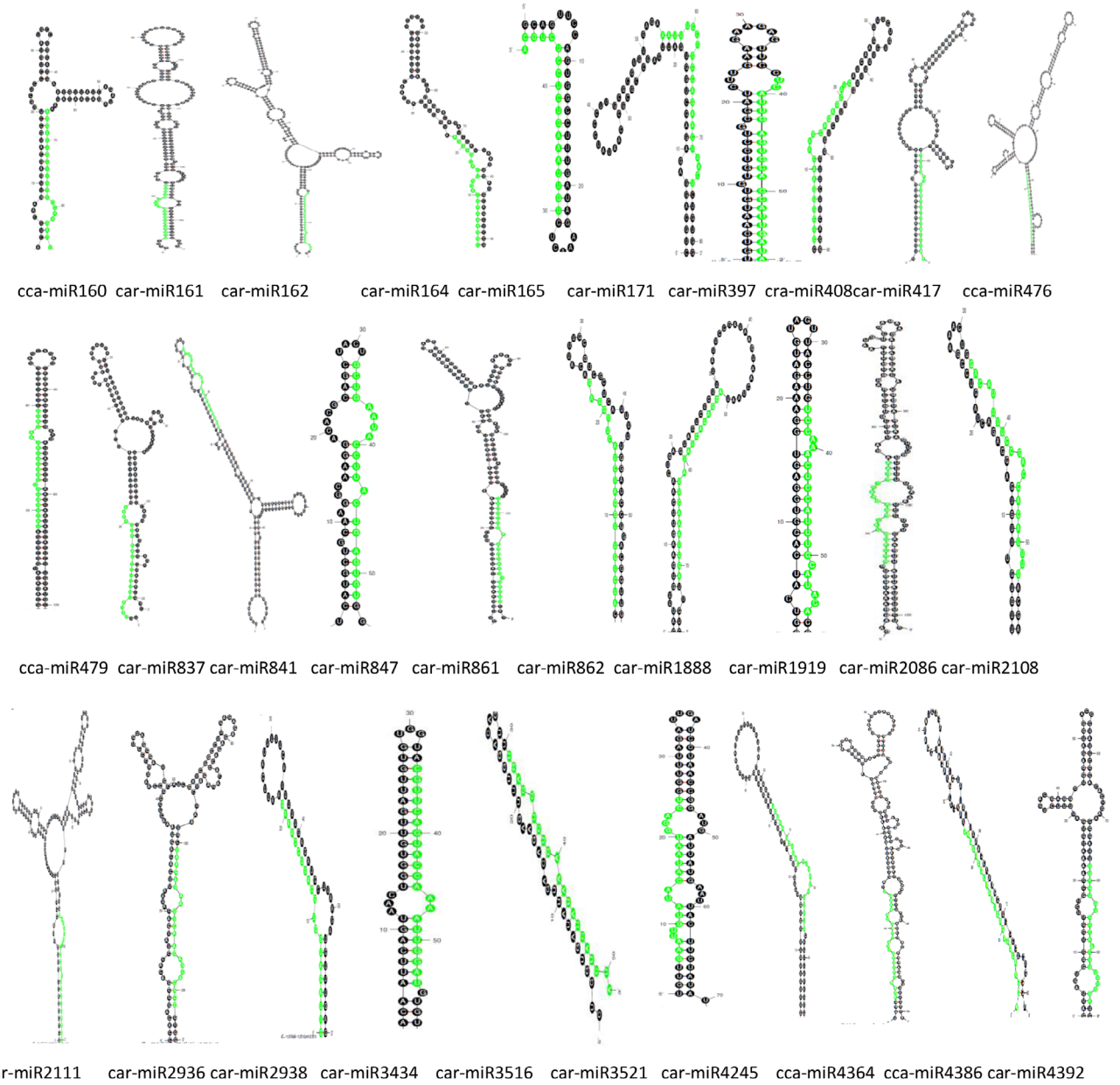

Figure 1. The newly identified *Coffea* miRNAs' secondary structures. The *Coffea* pre-miRNAs' secondary structures were developed through the Mfold algorithm. These structures clearly show the mature miRNAs in the stem portion of the stem-loop structures.

3.3. Validation of genus *Coffea* miRNAs through BLASTx

Checking the validity of *Coffea* miRNAs as noncoding RNAs is an important step. The *Coffea* miRNAs were subjected to BLASTx (Altschul et al., 1997) at NCBI. No homology was found with known proteins, which confirmed the validity of *Coffea* miRNAs as noncoding RNAs.

3.4. Conservation and phylogenetic studies of genus *Coffea* miRNAs

The miRNA of *C. arabica* (mir-171), due to its conserved nature in plants, was selected and subjected to conservation and phylogenetic analyses. *C. arabica* mir-171 showed conservation with *H. centranthoides* (hce), *S. lycopersicum* (sly), and *L. usitatissimum* (lus), as shown in Figure 2.

Table 1. The newly identified conserved genus *Coffea* miRNAs' characterization. *Coffea* miRNAs were characterized in terms of precursor miRNA length (PL), minimum free energy (MFE), mature sequence (MS), number of mismatches (NM), mature sequence length (ML), source EST (SE), mature sequence arm (MSA), GC content percentage (GC%), and organ of expression (OE).

Coffea miRNAs	Ref- miRNA	PL	MFE	MS	NM	ML	SE	MSA	GC%	Organ of expression
cca-miR160	bna-mir160d	86	–24.58	TGCCTGGCTCCCTGTATGCCT	1	21	GT655986	3’	62	Callus
car-miR161	ath-mir161	169	–32.90	TCAATGCATTGTAAGTGATTA	2	21	GT019091	5’	29	Leaves
car-miR162	aly-mir162b	253	–66.72	TCGATAACACCTCTGCATCCAG	0	22	GW470584	3’	50	Leaves
car-miR164	ath-mir164a	92	–30.50	CGGAGAAGCAGGTCACGTGGA	3	21	GT728514	5’	62	Suspension cells
car-miR165	aly-mir165b	48	–15.70	-CGGACCAGGCTTCATCCCA	2	19	GT690348	3’	63	Callus
car-miR171	aly-mir171a	81	–23.80	TGATTGAGCCGCGCCAATATC	0	21	GT708699	3’	52	Flower
car-miR397	ath-mir397a	58	–14.20	TCATTGAGTGCAGCATTGATA	2	21	GW472145	3’	38	Leaves
cra-miR408	ath-mir408a	61	–17.60	CTCGGGAACAAGCAGAGCATC	4	21	GT667061	5’	57	Fruits
car-miR417	ath-mir417	118	–21.29	GAAGCTAGTGAATTTGCTTCCA	2	22	GW439260	3’	41	Callus
cca-miR476	hbr-mir476	230	–61.93	TAATCGTTCTTTGCAAATTC	2	20	DV699198	5’	30	Seeds
cca-miR479	nta-mir479a	101	–42.50	CGTGATACTGGTTGCGGCTCATA	3	23	GT656915	5’	52	Callus
car-miR837	aly-mir837	122	–26.95	TGCAGAACAAGAAACTGATGA	4	21	GW447618	5’	38	Mixed tissues
car-miR841	ath-mir841a	159	–61.20	TTCTACCCACAGGAAACTGAA	4	21	GT713533	3’	43	Flower
car-miR847	aly-mir847	54	–15.30	AGACTCCTCTTCTTCATGCTG	3	21	GW471892	3’	48	Leaves
car-miR861	aly-mir861	140	–37.04	GATAGATATATCTTCAAGAAT	2	21	GW472371	3’	24	Leaves
car-miR862	ath-mir862	60	–20.50	ATCTGTTGGATCTACTTGGAG	3	21	GW464014	5’	43	Flower
car-miR1888	ath-mir1888a	76	–13.90	AAAGTTCAGATTTGTGAAGTT	4	21	GW485195	3’	29	Flower
car-miR1919	sly-mir1919c	58	–16.40	TGTCGCAGATGACTTTCGCAT	2	21	GT705996	3’	48	Flower
car-miR2086	aau-mir2086	122	–28.90	GGCATGCATGGAGAACTGGAA	3	21	GW431296	5’	55	Suspension cells
car-miR2108	gma-mir2108b	58	–18.50	GAAAGGTGTTGTGTTTGTGAG	3	21	GW440369	3’	43	Callus
car-miR2111	cme-mir2111a	153	–28.13	TAATCTGCATCCTGAGGTTCT	2	21	GW447318	3’	43	Mixed tissues
car-miR2936	ath-mir2936	126	–23.34	CTTGAGAGAGAGACCACAGAGA	3	22	GW442445	3’	50	Mixed tissues
car-miR2938	ath-mir2938	60	–11.20	GATCTTTTGAGAGGCTTCCGC	3	21	GW444175	5’	52	Stems
car-miR3434	aly-mir3434	59	–17.20	GCTGATTCTCAGATTTTAAAC	2	21	GW467305	3’	33	Flower
car-miR3516	ahy-mir3516	51	–14.10	GGCTGGTGATATTGACAGAAG	4	21	GW464792	3’	48	Flower
car-miR3521	ahy-mir3521	76	–12.30	TGTTGAATTGTATACATAATG	4	21	GW458324	5’	24	Suspension cells
car-miR4245	aly-mir4245	98	–18.70	ACAAAGTTTTATTATGACAAG	3	21	GW480511	3’	24	Flower
cca-miR4364	gma-mir4364b	161	–37.53	TAACAAGAGTGGGAGAACCTTCTT	3	24	GT653496	5’	42	Callus
cca-miR4386	gma-mir4386	81	–22.80	TTGCAAGTGCTGGAGAGGACTGCA	4	24	EE199732	5’	54	Fruits
car-miR4392	gma-mir4392	100	–19.84	TCTGCGAAAATGTGATTTTCTG	4	22	GT679447	3’	36	Seeds
car-miR5018	ath-mir5018	108	–25.45	TCAAAGATCCACCTTCAGTCCAAT	4	24	GW448486	5’	42	Suspension cells
car-miR5020	ath-mir5020c	44	–15.70	AGGCATGGAAGAAGGTGAGTT	3	21	GT706352	3’	48	Flower buds
car-miR5141	rgl-mir5141	134	–33.70	AGACCCGACGCGACTGACAGATAA	0	24	CF589178	5’	54	Leaves
car-miR5380	gma-mir5380b	164	–33.23	GAAGATGAATGATGATGATGGTGA	3	24	GW463578	5’	37	Flower
car-miR5634	ath-mir5634	74	–21.60	TTTGACTTTCTGAATTTAGGG	4	21	GW470072	3’	33	Leaves
car-miR5640	ath-mir5640	76	–12.90	T-AGGGAAGGAATTAGATTCT	3	19	GT702889	5’	35	Flower
car-miR5653	ath-mir5653	74	–14.60	AGAGTTGAGTTGAGTTGAGTTGGG	4	24	GT688986	3’	46	Callus
car-miR5670	gma-mir5670b	152	–32.34	GGTATCATACCATATTTGCTGC	4	22	GW431927	3’	41	Suspension cells
car-miR5717	bra-mir5717	64	–12.90	GTTTGGAATGTTAGCTGTGGC	4	21	GT712659	5’	48	Flower
car-miR5725	bra-mir5725	70	–17.60	ATTTGCAGCAATCTGATCTTC	4	21	GT690334	5’	38	Callus
car-miR6021	nta-mir6021	151	–39.53	ATGAAAGAGGATGCTATTGGA	3	21	GT692952	3’	38	Leaves
car-miR6151	nta-mir6151f	97	–32.00	CTAGTTTGAGGGATTGGATTGA	4	22	GR981149	3’	41	Flower
car-miR6483	hbr-mir6483	54	–10.80	TCTTGAAGAAATTTTCAGGATC	2	22	GW450554	3’	32	Suspension cells
car-miR8049	stu-mir8049	58	–10.40	GCAGGCTCATGCAGACAGGCA	3	21	GW479302	5’	62	Leaves
car-miR8051	Stu-mir8051	90	–44.60	TCCTATGGTCGAAAGATTCA	2	20	CF588939	3’	40	Leaves
car-miR8140	cpa-mir8140	83	–17.50	TGACTCAAGACTTCAGCTTCA	4	21	GW489082	5’	43	Leaves
car-miR8166	Ath-mir8166	56	–13.50	AGAAAGTGCAGAAAGTTTCTCC	3	22	GW468313	3’	41	Flower
car-miR8170	Ath-mir8170	75	–17.10	TTGCTTAAAGATGTTCTATCC	3	21	GT719527	3’	33	Seeds
cca-miR8172	Ath-mir8172	131	–20.22	ATGGCATCATCTTGTTGGAGGT	3	22	DV685748	5’	45	Leaves
car-miR9408	bol-mir9408	70	–9.25	ATTTCAGCTT-GAGAATGTTGTC	3	22	GW482878	5’	36	Leaves
car-miR9558	bra-mir9558	78	–18.40	AGAGATGTCTGCTTTGGAACA	3	21	GW439671	5’	43	Callus

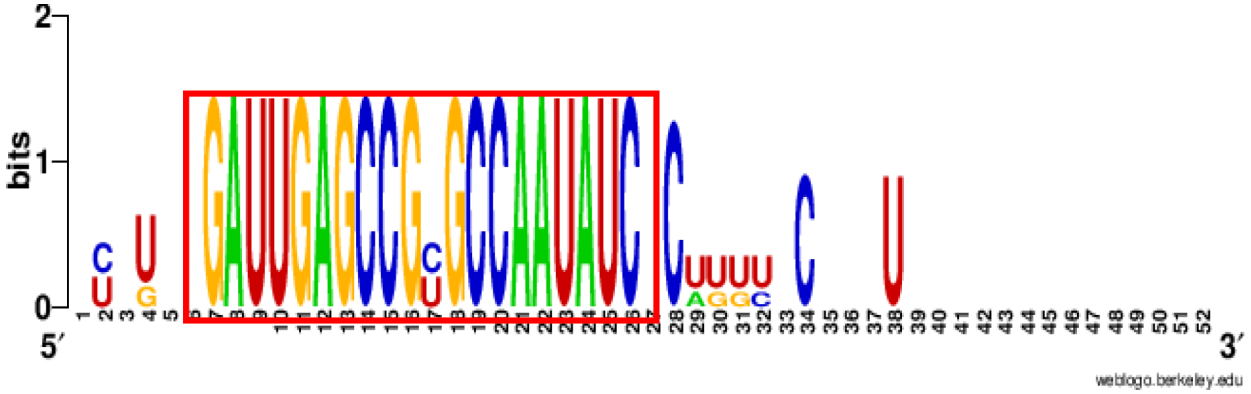

Figure 2. *Coffea* miRNA conservation studies. Alignment of *C. arabica* (car) miRNA (car-mir171) with *H. centranthoides* (hce), *S.lycopersicum* (sly), and *L. usitatissimum* (lus) was generated using WebLogo, a sequence logo generator, showing the conserved nature of mature miRNA sequences. The mature sequences are highlighted in the red box.

Similar findings were reported for *Phaseolus*, carrot, potato, tomato, and chili plants (Barozai et al., 2013a, 2013b; Din et al., 2014, 2016; Din and Barozai, 2014b). The phylogenetic analysis of the same miRNA (mir-171) sequences showed that *C. arabica* is closer to *H. centranthoides* (hce) than *S. lycopersicum* (sly) and *L. usitatissimum* (lus), as illustrated in Figure 3.

3.5. Potential targets of the genus *Coffea*

We adopted a more stringent criterion (Zhang et al., 2006a) with modification by Din et al. (2014) to predict the potential targets for the 51 identified miRNAs in *Coffea*. After carefully considering the aligned results, we identified a total of 150 potential targets (Table 2). In this study, we identified many miRNA targets that are conserved across several plant species, including *Arabidopsis*, rice, poplar, cotton, tomato, potato, chili, eggplant, and corn. Many studies have demonstrated, by experimental and/or computational approaches, that miRNAs target many transcription factors that help control plant development. We also found this class of targets in *Coffea*. MYB transcription factors represent a family of proteins with a conserved MYB DNA-binding domain. This domain was considered to be involved in regulation of secondary metabolism, control of cellular morphogenesis, and regulation of meristem formation and the cell cycle (Jin and Martin, 1999). Our results show that MYB proteins might be the target of miR-3434 in *Coffea*. Another transcription factor, homeobox-leucine zipper protein ATHB-14, is involved in the determination of adaxial-abaxial polarity in the ovule primordium and specifies adaxial leaf fates (Kim et al., 2008). In *Coffea*, we predict that miR-165 directs the regulation of ATHB-14. Aside from MYB and ATHB-14, there are several transcription factor groups that have been detected to be targets of miRNAs, like zinc ion binding (miR-2938, miR-3516), homeobox-leucine zipper protein REVOLUTA (miR-165), and zinc finger protein (miR-5725, miR-8166, miR-9408).

The next profiled *Coffea* miRNAs were observed to regulate metabolic proteins such as serine racemase, caffeic acid 3-O-methyltransferase, vacuolar ATP synthase subunit E, trehalose-6-phosphate synthase, histone-lysine N-methyltransferase, cytochrome C reductase, calmodulin-related protein, alpha-amylase, and glucose-

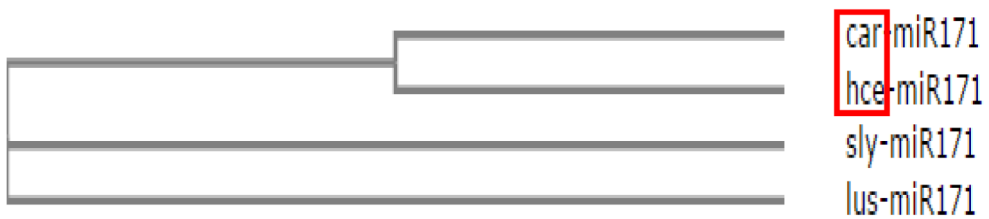

Figure 3. *Coffea* miRNA phylogenetic analysis. Analysis of *C. arabica* (car) miRNA (car-mir171) with *H. centranthoides* (hce), *S. lycopersicum* (sly), and *L. usitatissimum* (lus) was done with the help of ClustalW and a cladogram tree was generated using the neighbor-joining clustering method. The phylogenetic tree showed that *C. arabica* (car) is closer to *H. centranthoides* (hce) than *S. lycopersicum* (sly) and *L. usitatissimum* (lus). The closer plant species are highlighted in the red box.

Table 2. Targets of genus *Coffea* miRNAs as predicted by psRNAtarget and RNAhybrid in terms of miRNA family number, target acc., target description, and function.

Coffea miRNA	Target acc.	Target description	Function
cca-miR160, car-miR161,162, 164, cca-miR165, car-miR417, cca-miR476, cca-miR479, car-miR837, 847, 1919, 2938, 3434, 3516, cca-miR4364, 4386, car-miR4392, 5018, 5141, 5380, 5653, 5717, 5725, 6021, 8049, 8051, 8166, cca-miR8172, car-miR9408, 9558	NP305750, TC362356, TC362295, TC361205, SGN-U619883, TC400535, TC358231, TC401289, NM_001328854, XM_017374849, XM_017238780, SGN-U620162, TC370407, NP1646759, XM_007556415,SGN-U616271, TC374351, BX838928, NP2705764, TC358153, TC386720, NP456897, TC362407, TC371231, NP1659848, SGN-U619489, TC394910, NP232525, TC363392, SGN-U618032, TC374246, SGN-U615369, TC368992, SGN-U625715, SGN-U615687, SGN-U619936, TC373661, SGN-U616929, SGN-U618455, NP13352118, XM_009411299, SGN-U618580, SGN-U615789, XM_017119852, XM_010099976, NM_001036965, XM_014626289, XM_007914167	Putative auxin response factor, auxin response factor 17, auxin response factor 10, auxin response factor 16, pentatricopeptide (PPR) repeat-containing protein, endoribonuclease Dicer homolog, guanine nucleotide exchange factor, NAC5 protein, homeobox-leucine zipper protein REVOLUTA, homeobox-leucine zipper protein ATHB-14, mediator of RNA polymerase II transcription, indoleacetic acid-induced protein 9; transcription factor, similarity to ankyrin-like protein, ATPase/nucleoside-triphosphatase, transcriptional coactivator YAP1, small nuclear ribonucleoprotein D2, ATP binding, vacuolar protein sorting-associated protein 45, domain-containing protein, ubiquitin-protein ligase, zinc ion binding, non-LTR retroelement reverse transcriptase, retrotransposon like protein, MYB transcription factor, oxidoreductase NAD-binding domain-containing protein, zinc ion binding, contains similarity to heat shock transcription factor, similar to RNA helicases, NAC domain-containing protein 78, zinc phosphodiesterase ELAC protein 2, PHD finger protein-related, RNA recognition motif (RRM)-containing protein, ATPase subunit 1, calcium-binding EF hand family protein, homeobox-leucine zipper protein, RNA-binding protein, GDSL-motif lipase/hydrolase-like protein, zinc finger protein, DNA binding, ATP binding, pentatricopeptide repeat-containing protein, ZFN1 (ZINC FINGER PROTEIN 1), F-box family protein, zinc finger protein, isoleucine--tRNA ligase, tRNA synthetase class I (I, L, M, and V) family protein, finger protein 12, transcriptional activator protein acu-15	Transcription factor
car-miR162, 417 , 837, 847, 861, 862, 1919, 2108, 2938, 3434, 4245, cca-miR4364, car-miR4392, 5380, 5634, 5640, 5653, 5670, 5717, 5725, 6021, 6151, 6483, 8051, 8140, 8166, cca-miR8172	EG513934, XM_010326324, XM_015654998, XM_011523105, XM_014029716, SGN-U620027, SGN-U615000, SGN-U615355, SGN-U615009, SGN-U613213, NP237657, SGN-U614620, SGN-U618704, BX837933, TC383691, NP305087, SGN-U615294, TC379708, TC398934, SGN-U615360, SGN-U613503, SGN-U615324, SGN-U616867, SGN-U615295, SGN-U626186, SGN-U620263, NP229721, TC366012, SGN-U615711, SGN-U626543, SGN-U625139, SGN-U620278, XM_016557612, TC397703, SGN-U615916, SGN-U620515, SGN-U618360	Inositol oxygenase 4, serine racemase, putative protein arginine N-methyltransferase, domain containing oxidoreductase, cysteinyl-tRNA synthetase, caffeic acid 3-O-methyltransferase, pectin methylesterase 1, glutamyl-tRNA reductase 1, lipase class 3 family protein, vacuolar ATP synthase subunit E, similarity to mutator-like transposase, trehalose-6-phosphate synthase, nucleotidase family protein, N-acetylglucosamine deacetylase precursor, alpha-glucosidase, putative glucosidase I, ubiquitin-conjugating enzyme variant, glutaredoxin-C5, citrate synthase, ubiquinol-cytochrome C reductase, calmodulin-related protein, aldo/keto reductase AKR, xyloglucan galactosyltransferase, aspartyl protease family protein, alpha-amylase, (+)-delta-cadinene synthase, terpene synthase/cyclase family protein, cysteine proteinase, putative, cysteine synthase, anthocyanidinerhamnosyl-transferase, glucose-6-phosphate dehydrogenase, vitamin K-dependent gamma-carboxylase, monothiol glutaredoxin-S9, stroma ascorbate peroxidase precursor, glucose-6-phosphate 1-dehydrogenase, methionyl aminopeptidase	Metabolism
car-miR161, 171, 397, cca-miR476, car-miR861, 862, 1888, 2936, 2938, 3434, 3521, cca-miR4364, 4386, car-miR5380, 5717, 6021, 6151, 8049, 8140, 8166, 8170, cca-miR8172	SGN-U626881, SGN-U625098, SGN-U626324, SGN-U627786, SGN-U625609, SGN-U614001, SGN-U620177, SGN-U617583, SGN-U619299, BP836641, TC374054, SGN-U618312, TC390997, SGN-U615813, SGN-U615557, NP306739, SGN-U618454, SGN-U614989, TC359052, NP183968, TC375992, SGN-U627832, SGN-U617079, SGN-U618361	Predicted protein, hypothetical protein, predicted protein, hypothetical, hypothetical protein, hypothetical protein isoform 1, hypothetical protein, hypothetical protein, hypothetical protein, uncharacterized protein, uncharacterized protein, hypothetical protein, uncharacterized protein, predicted protein, hypothetical protein, unknown protein, unnamed protein, hypothetical protein, uncharacterized protein, hypothetical protein, protein COBRA precursor, hypothetical protein, unnamed protein, hypothetical protein	Hypothetical
car-miR164, 397, cra-miR408, car-miR2111, 2936, 3434, cca-miR4386, car-miR5640, 8049, 8051, 9558	TC359120, SGN-U613356, SGN-U626870, SGN-U614604, SGN-U613822, TC373116, DR366527, SGN-U619538, SGN-U613972, TC391373, NP224198, XM_012991812, XM_018249558	Phytochrome c, transducin family protein, ubiquitin family protein, protein kinase family protein, 2,3-diketo-5-methylthio-1-phosphopentane phosphatase, MAP kinase kinase, serine/threonine protein phosphatase, microsomal signal peptidase 25 kDa, kinase family protein, calcineurin B-like protein, receptor protein kinase-like, phosphomevalonate kinase, serine/threonine-protein kinase 40	Signal transduction
car-miR837, 3434, cca-miR4386, car-miR5653, 5725, 9558,	XM_014750876, TC362916, SGN-U615371, SGN-U627672, SGN-U614909, XM_010184592, XM_017973497	centromere-associated protein E, kinesin-like protein, cyclic nucleotide-gated ion channel, coatomer protein complex, potassium efflux antiporter, carbonic anhydrase 5B, glycosyltransferase like domain containing 1	Transporter
car-miR162, 164, cca-miR476, car-miR837, 6151, 8051	TC370085, TC379483, TC370018, SGN-U613155, XM_016328266,SGN-U614441, XM_014917224	Expansin-A1 precursor, embryo defective 1381, NAM (no apical meristem)-like protein, Aux/IAA protein, auxin response factor 9, auxin-induced protein, centrosomal protein	Growth and development
car-miR1888, 2086, 2111, 5141, 5380, 5725	SGN-U620022, SGN-U619259, SGN-U628254, TC380075, SGN-U613294, SGN-U614101	Heat shock cognate 70 kDa protein 1, heat shock protein, similar to heat shock protein, dehydration stress-induced protein, chaperone protein, similar to chaperonin-like	Stress-related
car-miR841, 3521, 5020, 5380	SGN-U616689, SGN-U619632, SGN-U619690, SGN-U620019, SGN-U620861	Putative histone H2A, ribosomal protein, 60s acidic ribosomal protein P1, TUA4 (tubulin alpha-4 chain), small nuclear ribonucleoprotein	Structural constituent
car-miR1919, 2111, 2938	NP2698013, SGN-U613205, TC385823	Similar to disease resistance protein, wound-responsive family protein, enhanced disease susceptibility 5	Disease-related

6-phosphate dehydrogenase. Newly predicted *Coffea* miRNA families miR-417, -847, -862, -2108, -4245, -4392, -5634, -5640, -5670, and -6483 were found to target these important proteins (Table 2). miRNA family miR-417 was found to target serine racemase in the N-methyl-D-aspartate receptor, which plays a vital role in metabolism (Talukdar et al., 2017). miR-847 was involved in targeting caffeic acid O-methyltransferase, which is an enzyme known for the biosynthesis process of lignin in plants (Lakshmi and Kalaivani, 2016). During an in vitro study, glucose-6-phosphate dehydrogenase, a cytolasmic enzyme used for maintenance of metabolism (Swartz, 2016), was targeted by miR-6483.

Sixteen percent of the profiled miRNAs of *Coffea* were found to target hypothetical proteins. Such findings were also published earlier (Wang et al., 2012).

Some *Coffea* miRNAs were discovered to target proteins functioning in signal transduction like phytochrome c, transducin family protein, mitogen-activated protein (MAP) kinase, and phosphomevalonate kinase. These proteins were found to be targeted by miRNA families miR-164, -397, -408, -2936, -3434, and -8051, respectively. Such findings were reported by various researchers (Allen et al., 2005; Wang et al., 2012; Ghani et al., 2013). Phytochrome c is an important chromoprotein that is involved in auxin signaling for plant elongation (Favero et al., 2016) targeted by miR-164. Similarly, miR-3434 targeted MAP kinase. Kinase takes part in signaling and immunity of plants (Mithoe et al., 2016), which will help us to regulate cell signaling via miRNAs in plants.

Transporters are proteins involved in the movement of different materials, also targeted by *Coffea* miRNA families. Such proteins were coatomer protein complex, potassium efflux antiporter, and glycosyltransferase-like domain containing 1, which were targeted by miR-5653, -5725, and -9558. Similar results were reported by other researchers (Frazier et al., 2010; Din et al., 2016).

The *Coffea* miRNAs were also found to target the proteins that were involved in plant growth and development. Such miRNA families are miR-162, -476, -837, and -6151, which regulate auxin/IAA protein and auxin-induced protein. In this research, miR-837 targeted auxin-induced protein, which plays a significant role in growth and development (Barozai, 2012a, 2012b, 2012c).

Some *Coffea* miRNA families were also found to target stress-related proteins such as heat shock cognate 70 protein 1, and heat shock protein; structural constituent proteins like putative histone H2A, 60s acidic ribosomal protein P1, and tubulin alpha-4 chain; and disease-related proteins like wound-responsive family protein (Barozai and Wahid, 2012; Wahid et al., 2016). Other researchers also reported such proteins as potential targets of miRNAs (Zhang et al., 2006a; Sunkar and Jagadeeswaran, 2008; Barozai et al., 2013a).

3.6. Conclusions

We have identified 51 potential candidate miRNAs belonging to 51 families from genus *Coffea* EST sequences based on bioinformatics. According to our knowledge based on a literature survey, these miRNA families are reported for the first time. These findings will be helpful to clearly elucidate the functions and processing of miRNAs in *Coffea*. This also proved that the bioinformatics approach for new miRNA identification from plant species whose genomes are not yet sequenced is a powerful technique. The EST-based identification confirmed the miRNA expressions.

References

Akter A, Islam MM, Mondal SI, Mahmud Z, Jewel NA, Ferdous S, Amin MR, Rahman MM (2014). Computational identification of miRNA and targets from expressed sequence tags of coffee (*Coffea arabica*). Saudi J Biol Sci 21: 3-12.

Ali SH, Barozai MYK, Din M, Aziz AN (2016). A review on the sorghum for biofuel and microRNAs. Pure Appl Biol 5: 1064-1100.

Allen E, Xie Z, Gustafson AM, Carrington JC (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121: 207-221.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. J Mol Biol 215: 403-410.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

Ambros V (2004). The functions of animal microRNAs. Nature 431: 350-355.

Baloch IA, Barozai MYK, Din M (2015a). Identification and characterization of 25 microRNAs and their targeted proteins in apricot (*Prunus armeniaca* L.). J Anim Plant Sci 25: 1466-1476.

Baloch IA, Barozai MYK, Din M, Achakzai AKK (2015b). Computational identification of 18 microRNAs and their targets in three species of rose. Pak J Bot 47: 1281-1285.

Barozai MYK (2012a). Identification and characterization of the microRNAs and their targets in *Salmo salar*. Gene 499: 163-168.

Barozai MYK (2012b). In-silico identification of microRNAs and their targets in fibre and oil producing plant Flax (*Linum usitatissimum* L.). Pak J Bot 44: 1357-1362.

Barozai MYK (2012c). The novel 172 sheep (*Ovis aries*) microRNAs and their targets. Mol Biol Rep 39: 6259-6266.

Barozai MYK (2013). Identification of microRNAs and their targets in *Artemisia annua* L. Pak J Bot 45: 461-465.

Barozai MYK, Baloch IA, Din M (2011a). Identification of microRNAs and their targets in *Helianthus*. Mol Biol Rep 39: 2523-2532.

Barozai MYK, Baloch IA, Din M (2011b). Computational identification of microRNAs and their targets in two species of evergreen spruce tree (*Picea*). Proc Wrld Acad Sci E 75: 413-418.

Barozai MYK, Din M (2017). Initial screening of plant most conserved microRNAs targeting infectious viruses: HBV and HCV. In: 2017 14th International Bhurban Conference on Applied Sciences and Technology; 2017. New York, NY, USA: IEEE.

Barozai MYK, Din M, Baloch IA (2011c). Identification of microRNAs in ecological model plant *Mimulus*. J Biophys Chem 2: 322-331.

Barozai MYK, Din M, Baloch IA (2013a). Structural and functional based identification of the bean (Phaseolus) microRNAs and their targets from expressed sequence tags. J Struct Funct Genom 14: 11-18.

Barozai MYK, Husnain T (2011). Identification of biotic and abiotic stress up-regulated ESTs in *Gossypium arboretum*. Mol Biol Rep 39: 1011-1018.

Barozai MYK, Irfan M, Yousaf R, Ali I, Qaisar U, Maqbool A, Zahoor M, Rashid B, Hussnain T, Riazuddin S (2008). Identification of microRNAs in cotton. Plant Physiol Biochem 46: 739-751.

Barozai MYK, Kakar S, Sarangzai AM (2013b). Profiling the carrot (*Daucus carota* L.) microRNAs and their targets. Pak J Bot 45: 353-358.

Barozai MYK, Shah SQ, Din M, Muhammad, R (2014). Codon usage bias and RNA secondary structures analysis for virus resistant genes in *Arabidopsis thaliana* and *Oryza sativa*. Pure Appl Biol 3: 81-91.

Barozai MYK, Wahid AH (2012). In-silico identification and characterization of cumulative abiotic stress responding genes in potato (*Solanum tuberosum* L.). Pak J Bot 44: 57-69.

Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281-297.

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004). WebLogo: a sequence logo generator. Genome Res 14: 1188-1190.

Dai X, Zhao PX (2011). psRNATarget: a plant small RNA target analysis server. Nucl Acids Res 39: 155-159.

Din M, Barozai MYK (2014a). Profiling and characterization of eggplant (*Solanum melongena* L.) microRNAs and their targets. Mol Biol Rep 41: 889-894.

Din M, Barozai MYK (2014b). Profiling microRNAs and their targets in an important fleshy fruit: tomato (*Solanum lycopersicum* L.). Gene 535: 198-203.

Din M, Barozai MYK, Baloch IA (2014). Identification and functional analysis of new conserved microRNAs and their targets in potato (*Solanum tuberosum* L.). Turk J Bot 38: 1199-1213.

Din M, Barozai MYK, Baloch IA (2016). Profiling and annotation of microRNAs and their putative target genes in chilli (*Capsicum annuum* L.) using ESTs. Gene Reps 5: 62-69.

Favero DS, Jacques CN, Iwase A, Le KN, Zhao J, Sugimoto K, Neff MM (2016). Suppressor of phytochrome b represses genes associated with auxin signaling to modulate hypocotyl growth. Plant Physiol 171: 2701-2716.

Ferrazzano GF, Amato I, Ingenito A, De Natale A, Pollio A (2009). Anti-cariogenic effects of polyphenols from plant stimulant beverages (cocoa, coffee, tea). Fitoterapia 80: 255-262.

Floyd SK, Bowman JL (2004). Gene regulation: ancient microRNA target sequences in plants. Nature 428: 485-486.

Frazier TP, Xie F, Freistaedter A, Burklew CE, Zhang B (2010). Identification and characterization of microRNAs and their target genes in tobacco (*Nicotiana tabacum*). Planta 232: 1289-1308.

Friedman J, Waller GR (1983). Caffeine hazards and their prevention in germinating seeds of coffee (*Coffea arabica* L.). J Chem Ecol 9: 1099-1106.

Ghani A, Din M, Baloch IA, Barozai MYK (2013). Identification of microRNA in 12 plant species of Fabaceae. Pure Appl Bio 2: 104-115.

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008). MiRBase: tools for microRNA genomics. Nucl Acids Res 36: 154-158.

Jin H, Martin C (1999). Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol 41: 577-585.

Jin WB, Li NN, Zhang B, Wu FL, Li WJ, Guo AG, Deng ZY (2008). Identification and verification of microRNA in wheat (*Triticum aestivum*). J Plant Res 121: 351-355.

Jones-Rhoades MW, Bartel DP (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14: 787-799.

Jones-Rhoades MW, Bartel DP, Bartel B (2006). MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57: 19-53.

Kim YS, Kim SG, Lee M, Lee I, Park HY, Seo PJ, Jung JH, Kwon EJ, Suh SW, Paek KH et al. (2008). HD-ZIP III activity is modulated by competitive inhibitors via a feedback loop in *Arabidopsis* shoot apical meristem development. Plant Cell 20: 920-933.

Kruger J, Rehmsmeier M (2006). RNAhybrid: microRNA target prediction easy, fast and flexible. Nucl Acids Res 34: 451-454.

Lakshmi K, Kalaivani A (2016). Expression of caffeic acid o-methyltransferase gene involved in lignin biosynthesis of sugarcane. J Sugarcane Res 5: 78-82.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R et al. (2007). ClustalW and ClustalX version 2. Bioinformatics 23: 2947-2948.

Loss-Morais G, Ferreira DC, Margis R, Alves-Ferreira M, Corrêa RL (2014). Identification of novel and conserved microRNAs in *Coffea canephora* and *Coffea arabica*. Gen Mol Biol 37: 671-682.

Mithoe SC, Ludwig C, Pel MJ, Cucinotta M, Casartelli A, Mbengue M, Sklenar J, Derbyshire P, Robatzek S, Pieterse CM et al. (2016). Attenuation of pattern recognition receptor signaling is mediated by a MAP kinase. EMBO Rep 17: 441-454.

Rebijith KB, Asokan R, Ranjitha HH, Krishna V, Nirmalbabu K (2013). In silico mining of novel microRNAs from coffee (*Coffea arabica*) using expressed sequence tags. J Hort Sci Biotech 88: 325-337.

Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005). Specific effects of microRNAs on the plant transcriptome. Dev Cell 8: 517-527.

Shah SQ, Barozai MYK, Din M, Baloch IA, Wahid HA (2016). RNA secondary structure analysis for abiotic stress resistant and housekeeping genes in *Arabidopsis thaliana* and *Oryza sativa*. Pure Appl Biol 5: 476-482.

Sunkar R, Jagadeeswaran G (2008). In silico identification of conserved microRNAs in large number of diverse plant species. BMC Plant Biol 8: 37.

Swartz JR (2016). Control of metabolic flux in cell-free biosynthetic systems. US Patent 20: 115-558.

Talukdar G, Inoue R, Yoshida T, Ishimoto T, Yaku K, Nakagawa T, Mori H (2017). Novel role of serine racemase in anti-apoptosis and metabolism. Biochim Biophys Acta 1861: 3378-3387.

Wahid HA, Barozai MYK, Din M (2016). Functional characterization of fifteen hundred transcripts from Ziarat juniper (*Juniperus excelsa* M.Bieb). Adv Life Sci 4: 20-26.

Wang J, Yang X, Xu H, Chi X, Zhang M, Hou X (2012). Identification and characterization of microRNAs and their target genes in *Brassica oleracea*. Gene 505: 300-308.

Xie F, Frazier TP, Zhang B (2010). Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). Planta 232: 417-434.

Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, Yang ZM (2007). Computational identification of novel microRNAs and targets in *Brassica napus*. FEBS Lett 581: 1464-1474.

Zhang BH, Pan XP, Cannon CH, Cobb GP, Anderson TA (2006a). Conservation and divergence of plant microRNA genes. Plant J 46: 243-259.

Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006b). Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 63: 246-254.

Zhang BH, Pan XP, Stellwag EJ (2008). Identification of soybean microRNAs and their targets. Planta 229: 161-182.

Zhang BH, Wang QL, Pan XP (2007a). MicroRNAs and their regulatory roles in animals and plants. J Cell Physiol 210: 279-289.

Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo TL, Cobb GP, Anderson TA (2007b). Identification of cotton microRNAs and their targets. Gene 397: 26-37.

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