Pak. J. Bot., 51(1): 133-142, 2019. DOI: 10.30848/PJB2019-1(35)

PROFILING OF 21 NOVEL MICRORNA CLUSTERS AND THEIR TARGETS IN AN IMPORTANT GRAIN: WHEAT (*TRITICUM AESTIVUM* L.)

HABIBULLAH KHAN ACHAKZAI[1*], MUHAMMAD YOUNAS KHAN BAROZAI[1], ABDUL KABIR KHAN ACHAKZAI[1], MUHAMMAD ASGHAR[2] AND MOHAMMAD DIN[1]

[1]*Department of Botany, University of Balochistan, Quetta, Pakistan*
[2]*Department of Chemistry, University of Balochistan, Quetta, Pakistan*
[*]*Corresponding author's email: habibullahkhan2019@gmail.com*

Abstract

MicroRNAs (miRNAs) are short non-protein coding RNAs made up of 18 to 26 nucleotides and produced in a cell endogenously. Most of them are conserved in nature evolutionally and thus providing a logical basis for the prediction of novel miRNAs and their clusters in many plants. In this research, structural and functional approaches have been combined to make prediction of novel miRNA clusters and their targets in wheat. The total numbers of detected novel miRNA clusters were 21 related to 19 miRNA families in wheat. These families were found as: 160, 396, 399, 414, 530, 2118, 2275, 5049, 5065, 5066, 5067, 5174, 5522, 5568, 6198, 6221, 7742, 7757 and 7778. Various attributes related to these miRNA clusters such as secondary structures, phylogenetic tree and web logo were generated and the minimum free energy (MFE) of the stem-loop structures was also achieved and reported. The mature miRNAs were found in the stem region of the hair-pin structures and, in this regard, 93 targets of the miRNAs were identified as well. The identified targets were various proteins including hypothetical, transporter, metabolism, transcription factor, cell signalling and stress related.

Key words: microRNA clusters, Expressed sequence tags, Wheat.

Introduction

MicroRNAs (miRNAs) are class of RNAs which do not take part in protein coding having short length from 18 to 26 nucleotide (nt) long. Chen *et al*., (2013) revealed that long hairpin precursor is processed to produce mature miRNA sequence. The gene expression is regulated by miRNAs *via* messenger RNA's (mRNA) decomposition or translation (Khvorova *et al*., 2007). RNA polymerase II is responsible for miRNAs gene transcription (Lee *et al*., 2004). The first transcribed products are primary miRNAs (Pri-miRNAs) made up of 1000-nt long chains and one or more than one is present in stem portion as mature miRNAs (Lee & Ambros, 2001). Lee *et al*., (2004) reported that the transcript was covered by specified and adapted 5′ end, large adenosines are attached with 3′ end and merged. Consequently, in the nucleus Pri-miRNAs are processed by microprocessor complex containing Drosha, RNase III, double-stranded RNA (dsRNA) binding protein and Pasha/DGCR8 (Denli *et al*., 2004). In this regard, Vaucheret *et al*., (2012) explained that precursor miRNAs were produced from Pri-miRNAs as a result of complete process. Thus, pre-miRNAs of approximately 90-340-nt long chain are produced in plants. The pre-miRNA hairpin structures transported from nucleus to cytoplasm, are further processed into 18 to 26 nt miRNAs by RNase III protein. Schwarz & Zamore, (2002) reported that RNA induced silencing complex (RISC) modified the less stable miRNA to double stranded pre-miRNAs. Mature miRNA with the help of RISC play important role in order to gene regulation, expression, elongation, inhibition or/and translation by generating messenger RNA (mRNA) complementarity of miRNA within its targeted mRNA (Aukerman & Sakai, 2003; Tang *et al*., 2003). In this regard, miRNAs perform versatile regulatory functions in eukaryotic organisms. Consequently, miRNAs play vital role in different biological processes of all eukaryotic cells, which include development, growth, organogenesis, cell signalling, transgene suppression (Chen, 2003; Allen *et al*., 2005; Yoshikawa *et al*., 2005), defend cells against biotic and abiotic stresses (Sunkar *et al*., 2004) and microbes (Balmer & Mauch-Mani, 2013). Almost all the researchers have concluded that miRNAs are conserved in animals and as well as in plants (Reinhart *et al*., 2002; Barozai *et al*., 2008; Jones-Rhoades *et al*., 2012; Barozai, 2013; Barozai *et al*., 2018).

Clusters of miRNA are uncommon in plants, nevertheless, in animals and human, they have been reported commonly (Yu *et al*., 2006). The mir-3630 family has been reported as cluster in different plant species i.e. *Vitis vinifera* (Pantaleo *et al*., 2010) and *Helianthus annuus* (Barozai *et al*., 2012).

Wheat grain or bread wheat belongs to Poaceae family which is one of the largest monocotyledonous family, containing around 780 genera, and 12,000 species (Christenhusz & Byng, 2016). This family has great ecological, economical and nutritional importance, and wheat together with maize and rice provides 60% of the calories and proteins for human's nutrition (Clayton & Renvoize, 1986; Thomasson, 1988; Watson & Dallwitz, 1992). The plant family improvement depends upon the genetically resistant varieties, seed productivity, modern cultivation and biotic and abiotic stress tolerance. Plant improvement can be assessed by studying its genetic makeup and sowing in different locations (Donmez *et al*., 2001; Khalil *et al*., 2002; Morgounov *et al*., 2010; Sanchez-Garcia *et al*., 2013; Gummadov *et al*., 2015).

This manuscript describes the exploration, profiling and characterization of novel miRNAs clusters in wheat and their targets annotation using different software. Therefore, 21miRNA clusters related to 19 miRNAs families in wheat have been reported. These families were found as: 160, 396, 399, 414, 530, 2118, 2275, 5049,

5065, 5066, 5067, 5174, 5522, 5568, 6198, 6221, 7742, 7757 and 7778. Various attributes related to these miRNAs such as secondary structures, phylogenetic tree and web logo were generated and the minimum free energy (MFE) of the stem-loop structures was also achieved and reported. The mature miRNAs were found in the stem region of the hair-pin structures and, in this regard, 93 targets of the miRNA clusters were identified.

Materials and Methods

Various bioinformatics tools are used as comparative genomics approach to achieve novel and interesting information about miRNAs in plants and animals. For this purpose, prediction of wheat pre-miRNAs and their organ expression were obtained by using Basic Local Alignment Search Tool (BLAST) software, miRNAs secondary structures were predicted by algorithm MFOLD software, stable miRNAs were selected based on their MFE, phylogenetic tree was generated employing clustalw software, conservation analyses were performed by web logo software and targets for the predicted miRNAs were predicted employing psRNAtarget software. In the following paragraphs, basic conduction steps related to this research are going to be described.

Reference miRNAs and prediction of candidate's pre-miRNAs: The reported miRNAs were used as reference miRNAs (miRNAs*) for homology search (Barozai *et al*., 2008). 6,394 plant precursors and matures miRNAs of plant sequences present in the miRBase were subjected to BLAST against wheat Express Sequence Tags (EST) (Kozomara & Griffiths-Jones, 2014). Therefore, 4 plants were used as miRNAs* like; *Oryza sativa* (osa), *Brachypodium distachyon* (bdi), *Sorghum bicolor* (sbi) and *Hordeum vulgare* (hvu). This database is openly available at http://www.mirbase.org/ and the Version (Rfam released 21 June 2014).

To predict conserved miRNAs in wheat, they were subjected to BLAST against the wheat ESTs available in national centre for biotechnology information http://www.ncbi.nlm.nih.gov. Currently, a total number of 1,286,372 ESTs are available for wheat plant in the NCBI dbEST https://www.ncbi.nlm.nih.gov/genbank/dbest/dbest_summary (release 130101, 1 January 2013). The miRNA* sequences both mature and precursor sequences were subjected to BLAST against wheat ESTs using BLASTn program http://blast.ncbi.nlm.nih.gov/Blast.cgi according to a reported procedure (Altschul *et al.,* 1990; Barozai *et al.,* 2008) following 4 mismatches with miRNAs*.

Prediction of wheat secondary structures: The generation of initial potential hairpin structures sequences were predicted by using secondary structure/hairpin sequences generation search tools algorithm MFOLD version 3.6 (Zuker, 2003) available at http://www.bioinfo.rpi.edu/applications/mfold/rna/form1.cgi. Minimum free energy (MFE) of the self-folded stem loop structure ≤ -10Kcal/mol was preferred (Ambros *et al.,* 2003).

Scrutinization and selection of stable miRNAs: Those miRNAs having long length, linear shape, lower MFE were retained because of their high stability and those which do not meet these criteria and having large bulges due to low bond order (false miRNAs) were exclude from the obtained data. The repeated EST was retained once.

Organ expression of miRNA: Classification of a miRNA on the basis of its presence in a tissue or an organ was also attempted. Therefore, organ of expression for the obtained novel miRNAs was obtained from readily available EST at https://www.ncbi.nlm.nih.gov/nucest/.

Phylogenetic and conservation analyses: Phylogenetic analysis of miR-160 was performed by comparing with different monocotyledonous plant precursor related to *Triticum aestivum* (tae), *Brassica napus* (bna), *Carica papaya* (cpa), *Cucumis melo* (cme) and *Brachypodium distachyon* (bdi) by using a software freely available at (http://www.genome.jp/tools-bin/clustalw) and following a procedure reported by Larkin *et al*., (2007). For conservation analysis, sequence logo generator studies were performed by using web logo software (http://weblogo.berkeley.edu/logo.cgi, version 2.8) of different plant species precursors such as: *Oryza sativa* (osa), *Brachypodium distachyon* (bdi) and *Zea mays* (zma) according to a procedure reported previously by Crooks *et al*., (2004).

Targets prediction: Wheat miRNA possible targets were achieved according to a method reported by Dai & Zhao, (2011) employing psRNAtarget software, with a few modification parameters, freely available at (http://plantgrn.noble.org/v1_psRNATarget/) and the targets were further classified in different categories as demonstrated previously by Zhang *et al*., (2006).

Results and Discussions

MicroRNA clusters identification: Wheat miRNA clusters were identified by using the available source namely wheat ESTs 1,286,372 for screening. As a result, 21 novel wheat pre-miRNA clusters were identified after the filtration and compilation process by homology searches. These pre-miRNA clusters were found to belong to 19 miRNA families, which were as: miR-160, 396, 399, 414, 530, 2118, 2275, 5049, 5065, 5066, 5067, 5174, 5522, 5568, 6198, 6221, 7742, 7757 and 7778. These 21 new miRNA clusters are reported for the first time in wheat and are shown in Table 1. Using the same procedure, different miRNA clusters have been identified and reported by Barozai *et al*., (2012) in *Helianthus* (Sun flower) and Din *et al*., (2016) in *Capsicum annuum* (chilli). Ambros *et al*., (2003) have explained the criteria B, C and D for homology searches for profiling of miRNAs in various plants and animals. After applying this formula, it was concluded that all these 21 novel wheat miRNA clusters were fulfilling the imperial formula. Ambros *et al*., (2003) have also supposed that the criteria D is enough for profiling miRNA as a novel candidate.

Table 1. Novel wheat miRNA cluster was regarded that precursor miRNA cluster length (PL), minimum free energy (MFE), mature sequences of the cluster (MS), number of mismatches in the cluster matures (represented in bold and in red colour) NM, mature sequence length (ML), source EST (SE), strand orientation (SO), mature sequence arm (MSA), GC percentage (GC%) and organ of expression (OE).

T. aestivum miRNAs	Source miRNAs	PL	MFE	MS	NM	ML	SE	SO	MSA	GC%	OE
tae-miR160a	osa-miR160b	105	-56.70	UGCCUGGCUCCCUGUAUGCCA GCGUGC**G**AGGAGCCAAGCAUG	0 1	21 21	CJ641547	+	3' 5'	62 67	Seed
tae-miR160b	osa-miR160c	100	-56.50	UGCCUGGCUCCCUGUAUGCCA GCGUGC**GA**GGAGCCAAGCAU**G**	0 3	21 21	CJ641547	+	5' 3'	62 67	Seed
tae-miR160c	osa-miR160d	100	-55.50	UGCCUGGCUCCCUGUAUGCCA GCGUGCGAGGAGCCAAGCAUG	0 0	21 21	CJ641547	-	5' 3'	62 67	Seed
tae-miR396	osa-miR396f	151	-59.63	UCUCCACAGGCUUUCUUGAACU AU**G**GUUCAAGAAAGUCCUUGGAAA	0 1	22 24	CJ776495	+	5' 3'	46 38	Root
tae-miR399	osa-miR399e	372	-93.07	UGCCAAAGGAGAUUUGCCCAG **C**GCCAAAGGAGAUUUGCCCAG UGCCAAAGGAGAUUUGCC**GAA**	0 1 3	21 21 21	HX145551	-	5' 5' 3'	52 57 48	Root
tae-miR414	osa-miR414	147	-37.25	UCAUCCUCAUCAUC**G**UCGUCC UC**G**UCCUCAUCAUC**G**UCGUCC	1 2	21 21	GH721200	-	5' 5'	52 57	Leaf
tae-miR530	osa-miR530	343	-108.43	UGCAUUUGCACCUGCACC**CU** UGCAUUUGCACCUGCACCU**G**	2 1	20 20	CA655063	+	5' 3'	55 55	Leaf
tae-miR2118	osa-miR2118a	356	-112.30	UUC**C**CGAUGCCUCCCAUUCCUA UUC**CU**GAUGCCUC**U**CAU**G**CCUA	1 4	22 22	CJ796550	-	5' 3'	55 50	Root
tae-miR2275	osa-miR2275a	251	-97.30	UUUGGUUUCCUCCAAUAUCU**U**A UUUGGUUUCCUCCAAUAUCUC**G**	1 1	22 22	CA597098	+	5' 3'	32 41	Anthers
tae-miR5049	bdi-miR5049	89	-60.20	**CCC**CUCCG**UU**CCAUAUUACUUGU CAAGUAAUAUGGA**A**CGGAGG**G**AGU	4 2	23 24	HX163922	-	5' 3'	48 46	Root

Table 1 (Cont'd.).

T. aestivum miRNAs	Source miRNAs	PL	MFE	MS	NM	ML	SE	SO	MSA	GC%	OE
tae-miR5065	bdi-miR5065	638	-126.78	UAGGCAAUCCACAUAUACACU UAGGCAAUCCACAUAUACACU	2 2	21 21	CJ700549	+	5' 3'	38 38	Shoot
tae-miR5066	bdi-miR5066	434	-131.08	AAGUGUAUAUGUGGAUUGCCU AAGUGUAUAUGUGGAUUGCCU AAGUGUAUAUGUGGAUUGCCU	2 2 2	21 21 21	CJ700549	-	5' 5' 3'	38 38 38	Shoot
tae-miR5067	bdi-miR5067	166	-64.12	UUAACGACAAUUAAUAUGGAU UCAGCGACAAUUAAUAUGGAU	3 1	21 21	CK210231	-	5' 3'	24 33	Shoot
tae-miR5174	bdi-miR5174	86	-61.30	CCUCCGUUCCAGAAUAGAU-G CAUCCAUUUUGGAACGGAGGG	3 3	20 21	CD927329	-	5' 3'	50 52	Seed
tae-miR5522	osa-miR5522	319	-78.88	AACAAUAGGCAUGAGAGGCAU AACGAUAGGAAUGGGAGGCAU	2 1	21 21	CJ796550	+	5' 3'	43 48	Root
tae-miR5568	sbi-miR5568g	85	-44.00	AAAACGUCUUAUAAUUUGGGG CAAAUUAUAAGAUGUUCUGAU	1 3	21 21	BJ288879	+	5 3	33 24	Seed
tae-miR6198	hvu-miR6198	542	-159.46	GCUCUGUCCUUGGACGGUCAUU GCUCUGCCCUUGGAUGGUCAUU	2 3	22 22	BJ237965	-	5' 3'	55 55	Seed
tae-miR6221	sbi-miR6221c	81	-42.30	UUUUGACCUGUGGCCCCUGCU CUGGGGCCACAUCUCAGAAGC	3 2	21 21	CD491436	-	5 3	57 62	Seedling
tae-miR7742	bdi-miR7742	91	-29.10	GCAUUCGCUCAUUCGCACAGU UGUGUGCUCGAGUGAAUGAGU	2 1	21 21	CA483944	+	5' 3'	52 48	Anther
tae-miR7757	bdi-miR7757	419	-143.94	CACAAAACCUUCAGCUAUCCA GAUAGUUUGAGGUUUUGUUUA	1 4	21 21	CJ725773	+	5' 3'	43 29	Mixed tissue
tae-miR7778	bdi-miR7778	355	-130.94	GAGCAUCAUGUCGGCGUGCGCGAUG CUGCGCCGCGCCUUU-UGACGAG	3 3	25 22	HX115168	-	5' 5'	64 68	Leaf

Wheat miRNA clusters descriptions: The detailed description of wheat miRNA clusters is shown in Table 1. The nt for these pre-miRNA clusters ranged from 81 to 638 with an average of 259. They were divided into seven classes depending on the number of nt with a class interval of 100-nt. 7 miRNA clusters fall into 1-100 nt class which make 33% of the total precursors. Similarly, 4 fall into 101-200 nt, making 19%; 1 falls into 201-300 nt, making 5%; 5 fall into 301-400 nt, making 24%; 2 fall into 401-500 nt, making 10%; 1 falls into 501-600 nt, making 5% and 1 pre-miRNA cluster falls into 601-700 nt class boundary, which is 5% of the total precursors. The number of nts and their class boundaries of pre-miRNA clusters in *Coffea* have been previously explained in the same manner by Farzana *et al*., (2017).

The next parameter shown in the Table 1 is the minimum free energy (MFE), which is an indicator for pre-miRNA clusters hairpin structures stability and their secondary structures constancy (Fig. 1). As there is an indirect correlation between the secondary structure of a miRNA cluster stability and the amount MFE. The MFEs for these pre-miRNA clusters ranged from -159 to -29 with an average of -84 Kcal mol^{-1}. They were divided into four classes depending on the amount of energy they released in Kcal mol^{-1} with a class interval of 50 Kcal mol^{-1}. 1 out of 21 miRNA cluster falls into -200-150 class, which makes 5% of the total precursors. Similarly, 6 out of 21 fall into -149-100, making 29%; 10 out of 21 fall into-99-50, making 47% and 4 out of 21 fall into-49-01 making 19% of the total precursors. The MFEs observed for potato plant have been explained by Barozai & Wahid (2012) using the same procedure.

The characterization of the wheat novel conserved miRNA was achieved for the acceptable range of mismatches between the miRNA* and potential wheat miRNAs. The mature sequences were selected from 0 to 4-nt mismatches with an average of 2-nt mismatch. Consequently, total 44 of matures were obtained in the 21 precursors. Among them, 6 out of 44 miRNAs matched perfectly with the miRNAs*, which was 14% of the total matures and 13 out of 44 miRNAs had one mismatch, 12 out of 44 miRNAs had two and 10 out of 44 miRNAs had three and 3 out of 44 miRNAs had four mismatches with the miRNAs*, which were 29%, 27%, 23% and 7% of the total matures respectively. The number of mismatches in miRNAs of other plants have been previously detected by various researchers such as Baloch *et al*., (2015a; b). Lengths of all the mature miRNA sequences were also counted. As a result, the number of nt for the sequences were obtained over the range of 20-25 with an average of 21 nt. Furthermore, these sequences lengths were divided into six classes based on the number of nt. Therefore, classes of 20, 21, 22, 23, 24 and 25 nt were obtained with the number of mature constituting 8, 66, 19, 2, 3 and 2% of the total matures respectively. Mature miRNA sequences lengths and their number of nt in a number of plants have been measured and counted in the same manner previously (Wang *et al*., 2012; Barozai, 2012; Ghani *et al*., 2013).

The miRNA-414 cluster has been already reported by Han *et al*., (2009), which we have also obtained during the proposed study. But the only discrepancy comes into the source EST number. According to the Han and co-workers (2009), this RNA cluster's source EST number is CJ967189 which is totally different from ours, which is GH721200. Consequently, this cluster has the feature of overlapping in it, whose position has been illustrated by red box in Fig. 1.

Next column of the Table 1 showed another parameter namely strand orientation, which showed that 48% of the ESTs were in plus strand and 52% found in minus strand of the total precursors. The location of wheat mature sequence on a precursor is an important parameter. Out of the total 44 matures, 25 matures, which were 57% of the total matures, were located at 5' arm of the hairpin structure and 19 of them, which were 43% of the total matures, and existed at 3' arm as shown in Fig. 1.

Characterization of the wheat novel miRNA clusters secondary structures is also reflected by their stability, which was performed by engagement of nt in guanine (G)/uracil (U) base pairing or Watson - Crick Model. Most of nt of mature miRNAs showed hydrogen bonding with its counterpart in the opposite arm as illustrated in Fig. 1. Stability of wheat miRNA cluster matures in the hairpin structures could be evaluated by the percentage of the bonded of G/U base pairings with opposite arm of the mature to the un-bond. The study revealed that the range of the % age of the bonding was from 43% to 96% with an average of 49% as shown in Fig. 1. This % bond to free of the mature was divided into four classes with a class interval of ≥10. Thus, 16 out of 44 fall into 90-100% range, which make 36% of the total mature. Similarly, 12 out of 44 fall into 80-89%, making 27%; 7 out of 44 fall into 70-79%, making 16% of the total and 9 out of 44 fall into 43-69% and making 20% of the total mature. Due to high bonding order, small bulges and stability in the clusters, the hairpins exhibited almost linear structures. Stability of a precursor is also dependent on the number of guanine (G)-cytosine (C) content, because GC makes pair with each other by three hydrogen bonding. As the number of GC content increases, the stability of the precursor also increases. Therefore, Wheat miRNA cluster GC content was calculated. The results showed that GC content ranged from 24 to 68% with an average of 49% (Table 1). For convenience, GC content was further divided into three classes with a class interval of 20%. GC content of 12 out of 44 matures fall into 20-40% class, which make 27% of the total matures. Similarly, 23 out of 44 matures fall into 41-60%, making 52% and 9 out of 44 matures fall into 61-80% and making 20% of the total matures. These matures were found to be located on the stem area of the hairpin structures.

Determination of the organ of expression for these precursors is an important parameter. Therefore, the wheat precursor miRNA clusters were found in different organs of wheat plant as: 6 out of 21 in seed (29% of the total), 5 out of 21 in root (23% of the total), 3 out of 21 in leaf (14% of the total), 3 out of 21 in shoot (14% of the total), 2 out of 21 in anther (10% of the total), 1 out of 21 in seedling (5% of the total) and 1 out of 21 in mixed tissue (5% of the total). Using the same procedure, various researchers have determined precursors in different organs of different plants previously (Barozai *et al*., 2014; Baloch *et al*., 2015a; b).

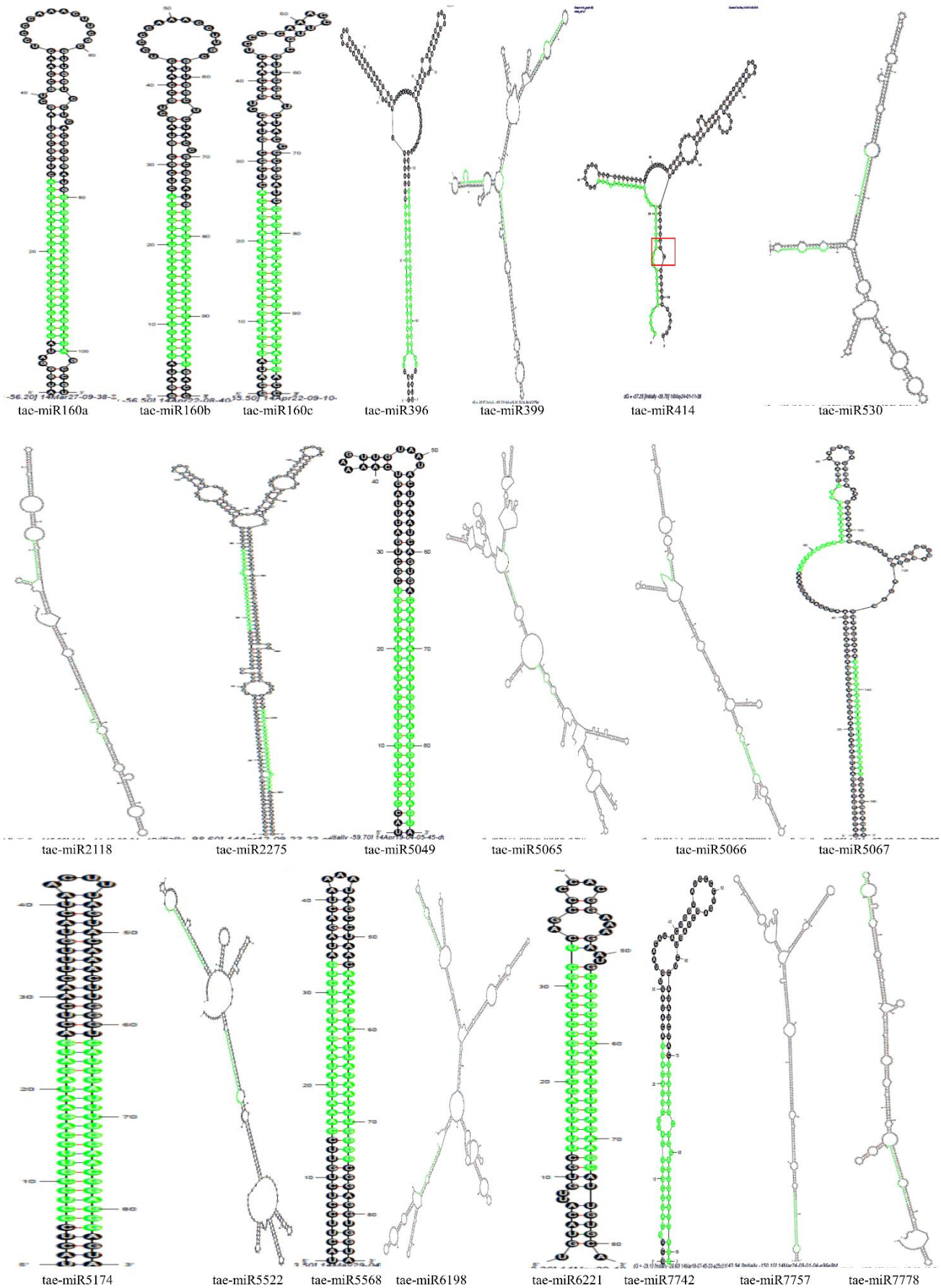

Fig. 1. The novel identified wheat miRNA clusters' secondary structures. These structures were developed through M. fold. These structures clearly were showed the mature miRNAs in the stem region of the stem-loop structures. So, the mature sequence of tae-miRNA414 cluster were highlighted in the red box overlapped in matures nucleotide area.

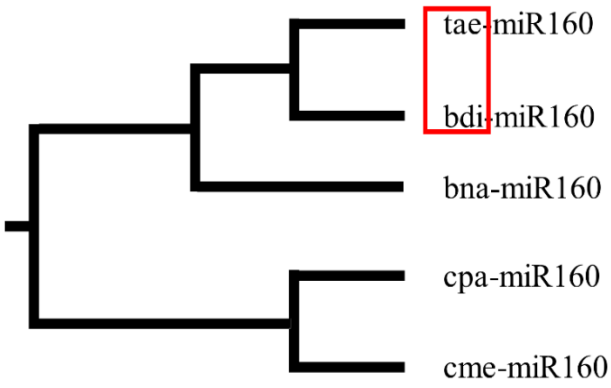

Fig. 2. Wheat miRNA cluster and their phylogenetic analysis. Phylogenetic tree showed that *B. distachyon* (bdi) is closer to *B. napus* (bna) than *C. papaya* (cpa) and *C. melo* (cme). The closer plant species were highlighted in the red box.

Phylogenetic and conservation studies of wheat miRNA clusters: Figs. 2 and 3 show the phylogenetic tree and conservation studies related to wheat miRNA clusters respectively. The red highlighted box in Fig. 2 reveals the closeness of wheat to a grass specie commonly known as purple false brome or stiff brome (*Brachypodium distachyon* (bdi)). The wheat miRNA cluster phylogenetic studies revealed that tae-miRNA160 cluster is close to *Brachypodium distachyon* (bdi) than *B. napus* (bna), *C. papaya* (cpa) and *C. melo* (cme). According to conservation studies of pre-miRNA 160 cluster shown in Fig. 3, the red highlighted box shows the conserved regions of matures related to other plants like; *Oryza sativa, B. distachyon* and *Zea mays*. Similar outcomes have been noted by various specialists previously (Din *et al.,* 2018; Achakzai *et al.,* 2018).

Targets prediction of wheat miRNA clusters: The validation and profiling of wheat miRNA clusters are very important step for targeted genes and their significant functions. For that reason, we implemented strict criteria to explore the targets for them. Bartel (2009) has reported that one miRNA can target a group of proteins. Table 2 shows the targets prediction of miRNA matures. According to the obtained results, single miRNA can target different multiple wheat genes. The number of total targeted proteins were obtained as 93 and divided into different classes based on their functions. These related functions include hypothetical, transportation, metabolism, transcription factor, cell signalling and stress.

The largest number of targeted proteins were obtained as hypothetical proteins shown in first class of Table 2. They were 25 off 93 which make 27% of the total profiled miRNA clusters targeted proteins. A number of researchers have previously reported various targeted hypothetical proteins for different plants miRNAs using the same procedure (Wang *et al*., 2012; Farzana *et al*., 2017; Baloch *et al*., 2018).

In the second class of Table 2, a number of wheat targeted transport proteins by wheat miRNA clusters have been given. These proteins are commonly involved in the transport of many life sustaining materials inside a cell. The examples of transport proteins include ubiquitin carrier protein etc. In this class, 16 off 93 transport proteins were obtained, which make 17% of the total targeted proteins. A number of targeted transport proteins, obtained by the same procedure, in different plants have been previously reported (Din *et al*., 2016; Din *et al*., 2018).

The third class of Table 2 presents the targeted metabolism proteins such as NADH-ubiquinone oxidoreductase, chromosome segregation ATPases, ribosomal RNA large subunit methyltransferase, glutathione synthetase, iron-containing alcohol dehydrogenase etc. Totally, 18 off 93 targeted metabolism proteins were predicted, which make 19% of the total targets. The main function of these targeted proteins is metabolism, which may involve in cell cycle, meristem formation etc (Jin & Martin, 1999; Din *et al*., 2016; Din *et al*., 2018).

The fourth class of the targeted proteins in Table 2 includes all those proteins which are characterized as transcription factors. They assist in the development of plants and are present in almost all plants (Xie *et al*., 2010; Din & Barozai, 2014a; b; Baloch *et al*., 2018). According to the results, 23 off 93 transcription factors were obtained, which make 25% of all the targets. These obtained factors include auxin response factor 22, chlorophyll a-b binding protein, growth-regulating factor 3, DNA-directed and RNA polymerase subunit beta, synthase L of carbamoyl-phosphate ATP-binding chain, synthase L carbamoyl-phosphate etc.

Fifth class of target proteins in Table 2 were predicted as signalling proteins, which are involved in signal transduction in a cell. As a consequence, 6 off 93 signalling proteins were predicted, which make 6% of the total targets. The examples of these obtained proteins include threonine/serine protein phosphatase, alkaline invertase, kinase protein, kinase Wpk4 protein etc. Many researchers have predicted a number of cell signalling targets previously in different plants (Din *et al*., 2016; Farzana *et al*., 2017; Din *et al*., 2018; Ghani *et al*., 2018). According to Curaba *et al*., (2014), miRNAs are involved in signalling pathways of phytohormones through the regulation of their metabolic activities in a cell.

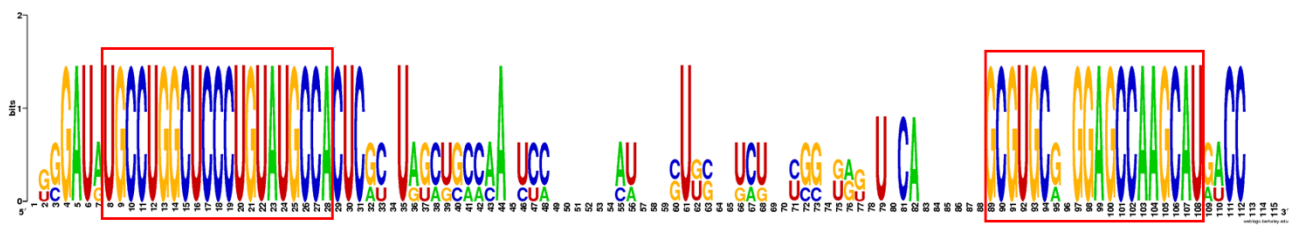

Fig. 3. Wheat miRNA cluster conservation studies. In the red box highlighted portion are showing that miRNA cluster two mature sequences and their conserved nature.

Table 2. Wheat miRNA targets, predicted by ps RNA target and RNA hybrid in term of miRNAs family number, targets accession, description of targets and functions.

Wheat miRNAs	Targets accession	Description of targets	Class	Functions
tae-miR160a, 160b, 160c, 396, 399, 414, 530, 2118, 2275, 5049, 5067, 5522, 5568, 7742, 7757.	TC390640, TC390640, TC390640, CO347612, CK210056, TC425284, TC436890, TC436890, CK201580, CA728335, CA728335, CA627100, CA617047, TC415367, CA627100, TC415367, TC394859, TC452572, CJ658178, TC420551, TC389107, TC413806, CD936138, CK200372, TC439092.	Uncharacterized protein At1g28100.5, Uncharacterized protein At1g28100.5, Uncharacterized protein At1g28100.5, Chromosome chr12 scaffold_36, Uncharacterized protein At1g11850.3, Chromosome undetermined scaffold_114, Os01g0775100 protein, Os01g0775100 protein, Chromosome undetermined SCAF _346, "Predicted protein", "Predicted protein", COP8-like protein, Chromosome undetermined SCAF _31, Chromosome 12 " SCAF _47", COP8-like protein, Chromosome chr12 scaffold_47, "Chromosome9 SCAF 33", Chromosome undetermined SCAF14638, Chromosome chr17 SCAF _16, Chromosome chr1 SCAF_180, Chromosome chr1 SCAF _5, Chromosome chr18 SCAF _96, Os12g0428300 protein, Chromosome undetermined SCAF _43, Chromosome chr3 SCAF _8	First	Hypothetical protein
tae-miR399, 2118, 5522, 5568, 6198, 6221.	TC438970, TC438970, CA597239, CA498385, TC457284, TC391933, CJ639710, BJ237965, TC455801, TC394487, TC457962.	expressed, Ubiquitin carrier protein, Ubiquitin carrier protein, Binding-protein-dependent transport systems inner membrane component precursor, Ribosomal RNA large subunit methyltransferase, Phosphate/phosphoenolpyruvate translocator, TB2/DP1/HVA22 family integral membrane protein, Predicted protein, ABC-type multidrug/protein/lipid transport system, Sec61 alpha subunit, Protein transport protein SEC61 alpha subunit, Ubiquitin carrier protein	Second	Transporter
tae-miR396, 414, 530, 2118, 2275, 5049, 5066, 5568, 6198, 6221, 7742, 7757, 7778.	CK158074, TC387707, TC381895, CV775457, TC381895, TC434827, CA498385, BJ263423, CJ700549, BQ236665, CV522195, TC410387, CO347661, CJ726126, DR732761, TC445096.	NADH-ubiquinone oxidoreductase, ATPase containing von Willebrand factor, Floral homeotic protein, Chromosome segregation ATPases-like protein, Floral homeotic protein, NADH-ubiquinone oxidoreductase chain 4, Ribosomal RNA large subunit methyltransferase, Early light-inducible protein ELIP, Glutathione synthetase, Iron-containing alcohol dehydrogenase, Transketolase, chloroplast, Pyruvate decarboxylase, Xyloglucanendotransglycosylase, Os10g0124300 protein, Os04g0669600 protein, Heavy metal translocating P-type ATPase	Third	Metabolism
tae-miR160a, 160b, 160c, 396, 2118, 2275, 5065, 5066, 5067, 5568, 6198, 7757, 7778.	TC412104, C425315, CK213447, TC412104, TC425315, TC425315, CK213447, CK213447, CJ645826, CJ796550, TC387934, CJ796550, CA692384, CA692384, TC429916, CK166358, CK165162, TC369071, TC427302, TC369071, CK211600, AL821953, TC428203, GH727096, TC456708, CA693035.	factor 8 auxin response, "auxin response factor 22", Chlorophyll a-b binding protein, factor 8 auxin response , factor 22 auxin response, Binding protein chlorophyll a-b, chloroplast precursor, Binding protein chlorophyll a-b, chloroplast pre-cursor, Growth-regulating factor 3, DNA-directed and RNA polymerase subunit beta, Oxysterol-binding protein-like, DNA-directed RNA polymerase subunit beta, DNA binding protein, DNA binding protein, protein-like Zinc finger, Nucleosome assembly protein 1-like protein 2, Nucleosome assembly protein 1-like protein 2, synthase L of carbamoyl-phosphate ATP-binding chain, RNA polymerase sigma factor, synthase L carbamoyl-phosphate, ATP-binding chain, Highlighted protein, CHY zinc finger family protein, expressed, Homeodomain leucine-zipper protein Hox10, Os01g0603100 protein, Homeodomain leucine-zipper protein Hox10, Expressed protein, Auxin response factor 22	Fourth	Transcription factor
tae-miR 2118, 5067, 5568, 7757.	TC371825, TC451647, NP9351614, TC438644, TC438644, TC393805, CD910548.	Threonine/serine protein phosphatase, Alkaline invertase, Kinase protein, kinase Wpk4 protein, Kinase Wpk4 protein, Kinase protein, NB-ARC domain containing protein	Fifth	Cell signaling
tae-miR2275, 5065, 5174, 5522, 6221.	TC401833, CK160979, GH728396, TC411544, TC417457.	NPR1-like 1, OsmC-like protein, Potential autophagy related protein-like, Cold acclimation protein WCOR413, DNAJ heat shock N-terminal domain-containing protein-like	Sixth	Stress related

The last but not least class of targets in Table 2 is stress related. According to Wang *et al*., (2004), the factors such as water deficiency, salinity, high temperatures, oxidative stress and chemical toxicity are stress related. Consequently, 5 off 93 stress related proteins were predicted, which make 5% of the total targets. A few of the found stress related targeted proteins include NPR1-like 1, cold acclimation protein WCOR413, DNAJ heat shock *N*-terminal domain-containing protein-like. Following the same procedure, Din *et al*., (2016), Farzana *et al*., (2017), Din *et al*., (2018) and Barozai *et al*., (2018) have predicted different stress related proteins in different plants.

Conclusions

In this research, structural and functional approaches were employed to predict 21 novel wheat miRNAs clusters, their families and targets. MicroRNA clusters characteristics including secondary structures, phylogenetic tree, web logo and MFE were generated and calculated. Various targets proteins have been predicted for the clusters. The functions and applications of hypothetical and other targeted proteins can be further evaluated. These findings will help to explain the different functions of wheat miRNA clusters. It will also provide bioinformatics approaches for new miRNAs conservation studies.

References

Achakzai, H.K., M.Y.K. Barozai, M. Din, I.A. Baloch and A.K.K. Achakzai. 2018. Identification and annotation of newly conserved microRNAs and their targets in wheat (*Triticum aestivum* L.). *PloS one*, 13(7): e0200033.

Allen, E., Z. Xie, A.M. Gustafson and J.C. Carrington. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 121(2): 207-221.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.,* 215(3): 403-410.

Ambros, V., B. Bartel and D.P. Bartel. 2003. A uniform system for microRNA annotation. *RNA*, 9(3): 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., M.Y.K. Barozai and M. Din. 2015a. Identification and Characterization of 25 and their Targeted Proteins microRNAs in Apricot (*Prunus armeniaca* L.). *The. J. Ani. Plant Sci.,* 25(5): 1466-1476.

Baloch, I.A., M.Y.K. Barozai and M. Din. 2018. Bioinformatics Prediction and Annotation of Cherry (*Prunus avium* L.) microRNAs and their Targeted Proteins. *Turk. J. Bot*., 42.

Baloch, I.A., M.Y.K. Barozai, M. Din and A.K.K. Achakzai. 2015b. Computational Identification of 18 miRNAs and their Targets in three species of Rose. *Pak. J. Bot.,* 47(4): 1281-1285.

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

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

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. Biol.*, 46(8-9): 739-756.

Barozai, M.Y.K., S.Q. Shah, M. Din and R. Muhammad. 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, M.Y.K., Z. Ye, S.R. Sangireddy and S. Zhou. 2018. Bioinformatics profiling and expressional studies of microRNAs in root, stem and leaf of the bioenergy plant switch grass (*Panicum virgatum* L.) under drought stress. *Agr. Gene*, 8: 1-8.

Bartel, D.P. 2009. MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, 136(2): 215-233.

Chen, X. 2003. A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Sci.*, 303: 2022-2025.

Chen, X., W. GAO, J. Zhang, X. Zhang and Z. Lin. 2013. Linkage mapping and expression analysis of miRNAs and their target genes during fibre development in cotton. *B.M.C. Genomics*, 14(1): 706.

Christenhusz, M.J. and J.W. Byng. 2016. The number of known plants species in the world and its annual increase. *Phytotaxa*, 261(3): 201-217.

Clayton, and S.A. Renvoize. 1986. Genera Graminum: Grasses of the World. Kew Bulletin Additional Series 13, Royal Botanical Garden, Kew. Her Majesty's Stationery Office, London. 389 pp.

Crooks, G.E., G. Hon, J.M. Chandonia and S.E. Brenner. 2004. Web-Logo: a sequence logo generator. *Genome Res.*, 14(6): 1188-1190.

Curaba, J., M.B. Singh and P.L. Bhalla. 2014. MiRNAs in the crosstalk between phytohormones signalling pathways. *J. Exp. Bot.*, 65(6): 1425-1438.

Dai, X. and P.X. Zhao. 2011. PsRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.*, 39(suppl-2): W155-W159.

Denli, A.M., B.B. Tops, R.H. Plasterk, R.F. Ketting and G.H. Hannon. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature*, 432(7014): 231-235.

Din, M. and Barozai M.Y.K. 2014a. Profiling and characterization of eggplant (*Solanum melongena* L.) microRNAs and their targets. *Mol. Biol. Rep*., 41(2): 889-894.

Din, M. and M.Y.K. Barozai. 2014b. Profling microRNAs and their targets in an important fleshy fruit: tomato (*Solanum lycopersicum* L.). *Gene*, 535(2): 198-203.

Din, M., M.Y.K. Barozai and A.N. Aziz. 2018. In Silico Profiling and Characterization of Conserved microRNAs in Biofuel Plant Sorghum. *Pak. J. Bot*., 50(6): 2265-2275.

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

Donmez, E., R.G. Sears, J.P. Shroyer and G.M. Paulsen. 2001. Genetic gain in yield attributes of winter wheat in the Great Plains. *Crop Sci*., 41: 1412-1419.

Farzana, B., M.Y.K. Barozai and M. DIN. 2017. Bioinformatics profling and characterization of potential microRNAs and their targets in the genus *Coffea*. *Turk J. Agri.*, 41: 191-200.

Ghani, A., M. Din and M.Y.K. Barozai. 2018. Convergence and Divergence Studies of Plant Precursor MicroRNAs. *Pak. J. Bot*., 50(3): 1085-1091.

Ghani, A., M. Din, I.A. Baloch and M.Y.K. Barozai. 2013. Identifcation of microRNA in 12 plant species of Fabaceae. *Pure Appl. Bio.*, 2(3): 104-115.

Griffiths-Jones, S. 2004. The microRNA Registr8 y. *Nucleic Acids Res.,* 32D: 109-111.

Gummadov, N., M. Keser, B. Akin, M. Cakmak, M. Mert, S. Taner, I. Ozturk, A. Topal, S. Yazar and A.I. Morgounov. 2015. Genetic gains in wheat in Turkey: winter wheat for irrigated conditions. *Crop J.,* 3: 507-516.

Han, Y.S., L.F. Lei, Z.H. Liang, S. Yi, C.A. Jun, L.C. Wen, L.Y. Bo and Z.B. Zhong. 2009. Computational identification of microRNAs and their targets in wheat (*Triticum aestivum* L.). *Sci. China Ser. C-Life Sci.*, 52(11): 1091-1100.

Jin, H. and C. Martin. 1999. Multi functionality and diversity within the plant MYB-gene family. *Plant Mol. Biol.*, 41(5): 577-585.

Jones-Rhoades, M.W. 2012. Conservation and divergence in plant microRNAs. *Plant Mol. Biol*., 80(1): 3-16.

Khalil, I.H., B.F. Carver, E.G. Krenzer, C.T. MacKown and G.W. Horn. 2002. Genetic trends in winter wheat yield and test weight under dual purpose and grain-only management systems. *Crop Sci.,* 42: 710-715.

Khvorova, A., A. Reynolds and S.D. Jayasena. 2007. Functional siRNAs and miRNAs exhibit strand bias. *Cell,* 131(2): 41-49.

Kozomara, A. and S. Griffiths-Jones. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.*, 42(1): 68-73.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson and D.G. Higgins. 2007. ClustalW and Clustal X version 2. *Bioinf.*, 23(21): 2947-2948.

Lee, R.C. and V. Ambros. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Sci*., 294: 862-864.

Lee, Y., M. Kim, J. Han, K.H. Yeom, S. Lee, S.H. Baek and V.M. Kim. 2004. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO J*., 23(20): 4051-4060.

Morgounov, A.I., V. Zykin, I. Belan, L. Roseeva, Y. Zelenskiy, H.F. Gomez-Becerra, H. Budak and F. Bekes. 2010. Genetic gains for grain yield in high latitude spring wheat grown in Western Siberia in 1900-2008. *Field Crops Res.,* 117: 101-112.

Pantaleo, V., G. Szittya, S. Moxon, L. Miozzi, V. Moulton, T. Dalmay and J. Burgyan. 2010. Identification of grapevine microRNAs and their targets using high-throughput sequencing and degraded analysis. *The Plants J*., 62(6): 960-976.

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

Sanchez-Garcia, M., C. Royo, N. Aparicio, J.A. Martin-Sanchez and F. Alvaro. 2013. Genetic improvement of bread wheat yield and associated traits in Spain during the 20th century. *J. Agri. Sci*., 151: 105-118.

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

Sunkar, R. and J.K. Zhu. 2004. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *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.

Thomasson, J.R. 1988. Fossil grasses: 1820-1986 and beyond. In International Symposium on Grass Systematics and Evolution, Washington, DC (USA), 27-31 Jul 1986. Smithsonian Institution Press.

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

Wang, J., X. Yang, H. Xu, X. Chi, M. Zhang and X. Hou. 2012. Identification and characterization of microRNAs and their target genes in *Brassica oleracea*. *Gene*, 505(2): 300-308.

Wang, W., B. Vinocur, S. Oded and A. Arie. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Cell Press*, 9(5): 244-252.

Watson and M.J. Dallwitz. 1992. The Grass Genera of the World. CAB International. Wallingford, UK. 1038 pp.

Xie, T.P., F.A. Freistaedter, C.E. Burklew and B. Zhang. 2010. Identification and characterization of microRNAs and their target genes in tobacco (*Nicotiana tabacum*). *Planta,* 232(6): 1289-1308

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(18): 2164-2175.

Yu, J., F. Wang, G.H. Yang, F.L. Wang, Y.N. Ma, Z.W. Du and J.W. Zhang. 2006. Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. *Bioch. Biop. Res. Comm.*, 349(1): 59-68.

Zhang, B., X. Pan, Q. Wang, G.P. Cobb and T.A. Anderson. 2006. Computational identification of microRNAs and their targets. *C.B.C.*, 30(6): 395-407.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Aci. Res.*, 31(13): 3406-3415.

(Received for publication 20 January 2018)