Identification of conserved miRNA molecules in einkorn wheat (*Triticum monococcum* subsp. *monococcum*) by using small RNA sequencing analysis

Ercan Selçuk ÜNLÜ[1,]* , Sara BATAW[2] , Didem ASLAN ŞEN[2] , Yunus ŞAHİN[2,3] , Nusret ZENCİRCİ[2]

[1]Department of Chemistry, Faculty of Arts and Science, Abant İzzet Baysal University, Bolu, Turkey
[2]Department of Biology, Faculty of Arts and Science, Abant İzzet Baysal University, Bolu, Turkey
[3]Department of Biochemistry, School of Medicine, Altınbaş University, İstanbul, Turkey

Abstract: *Triticum monococcum* subsp. *monococcum* as a first cultivated diploid wheat species possesses desirable agronomic and quality characteristics. Drought and salinity are the most dramatic environmental stress factors that have serious impact on yield and quality of crops; however, plants can use alternative defense mechanisms against these stresses. The posttranscriptional alteration of gene expression by microRNAs (miRNAs) is one of the most conserved mechanisms. In plant species including wheat genomes, miRNAs have been implicated in the management of salt and drought stress; however, studies on einkorn wheat (*Triticum monococcum* subsp. *monococcum*) are not yet available. In this study, we aimed to identify conserved miRNAs in einkorn wheat using next generation sequencing technology and bioinformatics analysis. In order to include a larger set of miRNAs, small RNA molecules from pooled plant samples grown under normal, drought, and salinity conditions were used for the library preparation and sequence analysis. After bioinformatics analysis, we identified 167 putative mature miRNA sequences belonging to 140 distinct miRNA families. We also presented a comparative analysis to propose that miRNAs and their target genes were involved in salt and drought stress control in addition to a comprehensive analysis of the scanned target genes in the *T. aestivum* genome.

Key words: microRNA, wheat, Perl, Mfold, small RNAs

1. Introduction

Wheat is one of the leading global crops, with an annual production of over 615.8 million metric tons. The level of its polyploidy is an important criterion for the classification of wheat species. Even though the time and the location are not clear, wild diploid wheat was spontaneously evolved from its close relative, *Triticum boeticum* Boiss. Wild emmer emerged as a tetraploid wheat form and re-hybridization of this form over time with a diploid close relative resulted in the rise of spelt-like hexaploid wheat. Due to the influence of human practices, wild diploid and tetraploid plants have undergone genetic selection for their useful agronomic traits. This evolution process resulted in the cultivation of diploid (e.g., einkorn) and tetraploid (e.g., emmer) wheat forms. Wheat is classified under the genus *Triticum* of Triticeae (Briggle, 1963), and several species have been characterized with diverse morphological and genetic variations (Curwen-McAdams et al., 2016).

T. monococcum subsp. *monococcum* (einkorn wheat) is a diploid wheat derived from *T. boeoticum* (wild einkorn wheat). It is capable of growing in adverse environmental conditions. It has a high nutritional value and gives acceptable yield on poor soils. Cultivation of einkorn wheat dates back to early times of the first agricultural activities. Since then, it has been cultivated in some provinces of Turkey (Karagöz and Zencirci, 2005), the Balkan countries, and Morocco (Serpen et al., 2008).

Small RNA molecules are noncoding RNA elements with a diverse group of functions. Several classes of small RNAs (e.g., miRNAs, siRNAs, and piRNAs) have been described (Peters and Meister, 2007). MicroRNA (miRNA) molecules are short and single-stranded noncoding RNA molecules acting as posttranscriptional control elements in animals, plants, and fungi (Bartel, 2004; Carthew and Sontheimer, 2009). Biosynthesis of mature miRNA molecules requires a chain of biochemical reactions starting with the transcription process carried out by Pol II or Pol III enzymes, which yields the primary miRNA (Pri-miRNA) molecules. Pri-miRNA is folded into a stem-loop structure that is then systematically digested to produce approximately 21–23-nt-length mature

* Correspondence: esunlu@ibu.edu.tr

527

miRNA molecules (Ritchie et al., 2007). Studies on plants have shown that miRNA molecules have crucial roles in growth, development, and stress resistance processes (Jones-Rhoades et al., 2006; Zhang et al., 2006; Budak and Akpinar, 2015).

miRNAs have been discovered from various organisms, and, to date, a total of 1269 plant miRNAs have been identified and deposited in MirBase (Griffiths-Jones et al., 2008). There are more than a hundred entries available for popular plant species, and for some important plants the unique reads are listed as 600 for *Oryza sativa*, 573 for *Glycine max*, 508 for *Arabidopsis thaliana*, and 158 for *Zea mays* as of 1 March 2018. The numbers drop dramatically for plant species with limited or unknown genome sequences, such as 118 for *Triticum aestivum*, 116 for *Hordeum vulgare*, 102 from *Solanum lycopersicum*, 18 from *Vigna unguiculata*, 16 from *Saccharum officinarum*, and 12 from *Phaseolus vulgaris* (http://www.mirbase.org).

In a recent comprehensive study, 88 miRNA reads for *T. monococcum* subsp. *monococcum* were predicted by the homology-based analysis of putative miRNA sequences from the transcriptome assemblies in the NCBI database (Alptekin and Budak, 2016). The numbers of miRNAs identified from other plant species suggest that more miRNA molecules are yet to be identified from *T. monococcum* subsp. *monococcum*. In particular, we hypothesize that identification of miRNAs involved in stress regulation requires alternative strategies since most of those miRNAs are expressed during stress conditions. In the present study, we pooled the samples of *T. monococcum* subsp. *monococcum* tissues grown under normal conditions and from those subjected to salt and drought stress to increase the number of identified miRNAs, and analyzed the sequences of extracted small RNA molecules to identify the expressed miRNA sequences.

2. Materials and methods

2.1. Sampling of *T. monococcum* subsp. *monococcum* cultures

Einkorn seeds belonging to six different wheat populations were surface-sterilized using 70% ethanol and 30% sodium hypochlorite. Seeds were germinated on half-strength MS solution. The cultures were incubated for 10 days in a growth chamber under controlled conditions at 24 ± 2 °C with a 16-h light and 8-h dark photoperiod before the stress treatment.

After 10 days, the plant samples were grown under control (no treatment), salt stress (100 mM NaCl), and drought stress (0.3 MPa PEG-600) conditions in a growth chamber under the same conditions as defined above (Mahmood et al., 2002). Leaf and root samples from control and treated plants were harvested after 0, 3, 9, 12, and 24 h of the stress application and immediately

frozen in liquid nitrogen. Approximately 4.0-g samples of the pooled wheat tissues from all the treated and control wheat samples were submitted to Source BioScience Plc (Nottingham, UK) for RNA isolation, small RNA library preparation, and sequencing using the Illumina MiSeq next generation sequencing platform.

2.2. Small RNA isolation

Small RNA molecules (<200 nt) were extracted from the pooled samples using a mirVana miRNA Isolation Kit according to the manufacturer's instructions (Life Technologies). The sample was quantified using an Agilent 2100 Bioanalyzer to ensure that the quantity and quality of the submitted material met the specified criteria before progressing through the library preparation.

2.3. Small RNA library construction and sequencing

The library was prepared using a TruSeq Small RNA Sample Preparation Kit. The 3` and 5` adapters were ligated to each end of the RNA molecule, and a reverse transcription reaction was used to create single stranded cDNA. Then cDNA fragments having adapter molecules on both ends underwent 11 cycles of PCR to amplify the amount of prepared material. The resulting library (18.35 nM) was validated with the Agilent 2100 Bioanalyzer. The library was then loaded onto an Illumina MiSeq Flow Cell at a concentration of 8 pM, and the samples were then sequenced using 50-bp paired-end runs.

2.4. Computational sequence analysis

2.4.1. Trimming and collapsing sequences

Before starting blast analysis, the data were cleaned of redundant sequences. First, the sequences were adapted and quality-trimmed using Skewer (version 0.1.12) (Jiang et al., 2014). The trimming parameters were adjusted for the small RNA input. The first processing step of merging identical reads and saving their occurrences (collapsing) was performed in order to provide data in the least redundant way possible and to speed up the classification process.

2.4.2. Profiling of small RNAs

For the general classification of available small RNA molecules, the collapsed sequence data were mapped to the *A. thaliana* genome as a reference from Ensembl (TAIR10) (Kersey et al., 2016) using Bowtie (Langmead et al., 2009) and filtered for known RNA elements. The detailed analysis of the small RNA library sequences received from the Illumina MiSeq platform was analyzed using Perl codes designed by our group as described by Ünlü et al. (2015). Basically, the blast code was generated to analyze small RNA sequences compared with the database generated using formatdb (Altschul et al., 1990) from a total of 30,424 known mature miRNA sequences belonging to 203 different species available at miRBase (Kozomara and Griffiths-Jones, 2011). Given the short

sequences, the code selected sequences having more than a 90% identity to reduce the risk of false positives.

2.4.3. Prediction of secondary structures

For the secondary structure prediction studies, a previously designed Perl code was used to identify precursor miRNA sequences (Ünlü et al., 2015). The code searches for 100% matches for putative mature miRNA sequences in *T. monococcum* subsp. *monococcum* chromosomal scaffold sequences (downloaded from NCBI). When a match is located, the sequence is extracted along with 80 nucleotides upstream and downstream of the located miRNA. Then a prediction of the secondary structures of the extracted miRNA precursor sequences was carried out using the RNA Folding Form application (http://mfold.rna.albany. edu) (Zuker, 2003) with the default software settings. The structural output files in the *ct* file format were uploaded to the Mfold server using the Structure Display and Free Energy Determination application to display the structure (Mathews et al., 1999).

2.4.4. Bioinformatics analysis for the characterization of putative miRNA target genes

To identify the target genes for the predicted miRNAs in this study, the complementary sequence matches were screened in *A. thaliana* and *T. aestivum* genomes using the psRNATarget tool (Dai and Zhao, 2011). We set the parameter to default values except for the maximum expectation being set to 3.0, the length for complementarity scoring (hspsize) being set to 18, and the number of top target genes for each small RNA being set to 50. For the categorization of target genes and the extents of their involvement in stress control, we downloaded the list of *A. thaliana* genes from the TAIR database (http://www.arabidopsis.org) and filtered those annotated as "response to stress" under the GO Slim functional category. Then we compared the list of predicted target genes in terms of whether they belonged to any of the filtered 519 stress-related genes.

For a detailed annotation analysis of miRNA target genes in *T. aestivum*, we downloaded the annotation file (Version 2.2) (Mayer et al., 2014) from the Joint Genome Institute portal that reported the protein-coding gene sequences in *T. aestivum*. We extracted the information for the genes showing the target fingerprints against the identified miRNA sequences. Using simple Perl codes (available from https:// github.com/esunlu/go_cluster_analysis), we clustered the GO annotations for the terms "biological process", "molecular function", and "cellular component" and analyzed the data to obtain a detailed functional categorization.

3. Results

3.1. Prediction of small RNAs

An average of 4.0 g pooled wheat tissues were processed to prepare the small RNA sequencing library. After sequencing, 15,139,448 raw reads were obtained. The reads were processed by trimming adapter sequences, quality filtering, and merging identical reads, yielding 751,647 identical small RNA sequences. The results were further filtered against several databases of known elements in *A. thaliana* in which the largest family of small RNAs was identified including miRNAs, CDS, mRNAs, tRNAs, snoRNAs, ncRNAs, snRNAs, and rRNAs (Figure 1).

From the miRBase database, 30,434 known mature miRNA sequences belonging to 203 different species were obtained and formatted for blast analysis. To reduce the numbers of false positives, we set the parameters to >90% identity and >0.0001 E-value in our blast code. The analysis identified 167 putative mature miRNA sequences belonging to 140 miRNA families (shown in S1 Table). When the sequence lengths were compared, the most abundant read length was 18 nucleotides (24.55%), followed by 21 nucleotides (19.76%) and 19 nucleotides (16.77%) among the total identified miRNAs (Figure 2).

The base distribution analysis at each position of the identified miRNA sequences revealed that uracil and guanine were the most abundant in the first and second positions with 64 and 53 of the sequences, respectively (Figure 3A). In addition, when the base distribution was analyzed against the length of miRNAs, a dominant bias towards uracil (U) at the first nucleotide was found especially for miRNAs with a length of 19–21 nt (Figure 3B).

3.1.1. Validation of *T. monococcum* subsp. *monococcum* miRNAs by secondary structure prediction

Since there are no available sequence data for *T. monococcum* subsp. *monococcum* chromosomes, the *Triticum urartu* chromosomal scaffold sequence data were downloaded from NCBI and used as reference to extract the precursor miRNA sequences. Using the encoded Perl script, 1,455,436 scaffold sequences, with sequence lengths ranging from 50 to 82,078 nucleotides, were searched for 100% positive matches. The extracted sequence frame corresponds to 80 nucleotides upstream of the start of the matching mature miRNA and 80 nucleotides downstream of miRNA.

We were able to extract 111 precursors to be analyzed for characteristic secondary structure folding. Mfold software was used to analyze secondary structures of the extracted pre-miRNA sequences. The default parameters were used to analyze secondary structures of the selected sequences. Seventy-seven of the sequences showed a stem-loop structure that is characterized for pre-miRNA sequences (see S2 Table). It is obvious that completing the assembly of *T. monococcum* subsp. *monococcum* of *T. urartu* chromosomal sequences will enhance the potential of the bioinformatics analysis for the *Triticum* species. The failure to predict of the secondary structure for the remaining 34 *T. monococcum* subsp. *monococcum*
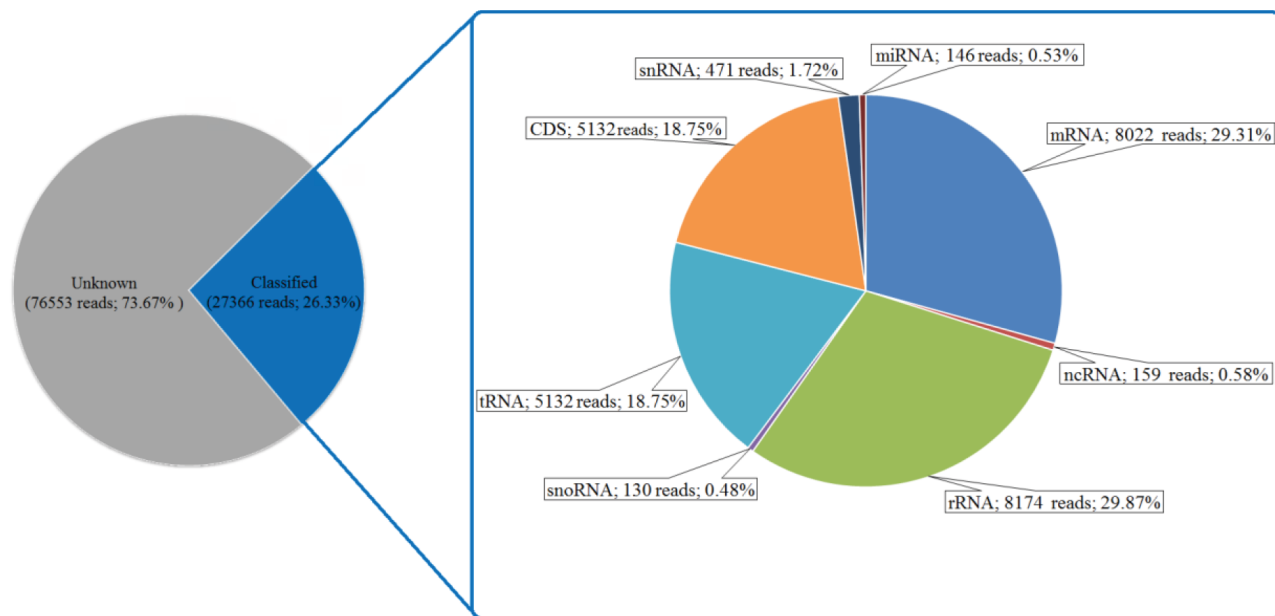
Figure 1. Distribution of small RNA molecules in *T. monococcum* subsp. *monococcum* sequence data.
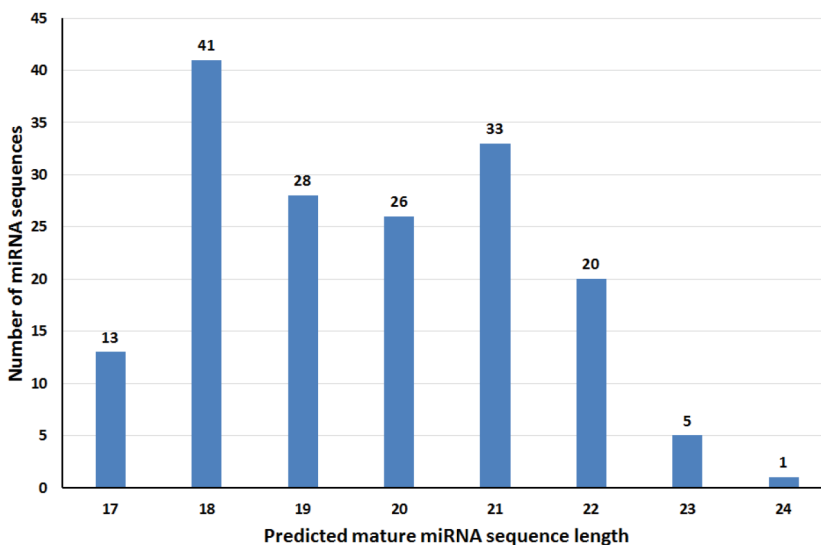
Figure 2. Length distribution of predicted miRNA molecules.

miRNAs was likely to have been due to the incomplete and fragmented nature of the *T. urartu* scaffold sequences used for the analysis.

3.2. Characterization of putative miRNA targets by bioinformatics prediction
Target prediction is an important step to characterize miRNA function. We compared the predicted miRNAs to those verified by experimental analysis in other plants under drought and salt stress conditions (Table 1). In this study, 23 salt stress and 24 drought stress-related miRNAs were identified for *T. monococcum* subsp. *monococcum*.

A comparison of the data among different plants suggests that most of the stress-related miRNAs are common across the compared species. According to the literature data as listed in Table 1, co-expression of 17 identified miRNAs was associated with both salt and drought stress conditions, experimentally (qPCR, northern blot, microarray, etc.). In addition, more than
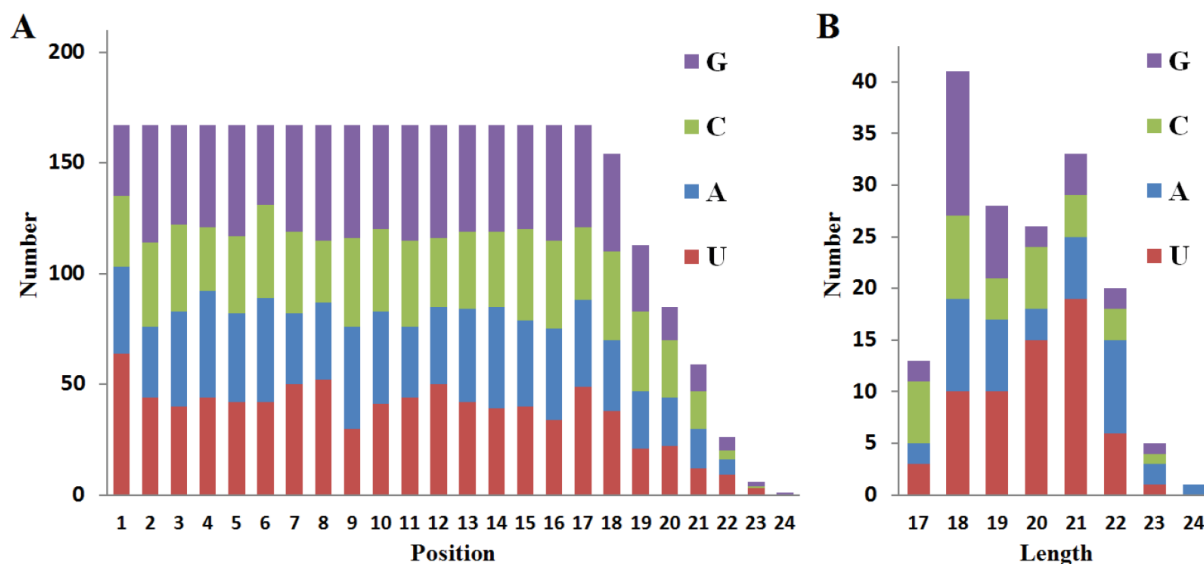
Figure 3. Nucleotide distribution analysis. Analysis for the first nucleotide (A) and positional (B) biases.

half of the salt and drought stress-related miRNAs were conserved in at least two different species (Figure 4). None of the miRNAs controlling either condition was conserved in all five of the analyzed species.

We used the psRNATarget tool to scan possible target genes for sequences of the predicted miRNA families presented in Table 1. For the salt response-related miRNA targets, of the 1435 genes that were identified, 78.33% were proposed to be controlled by cleaving the corresponding transcript. For the drought stress-related miRNA target genes, it is proposed that 79.01% of 1548 genes were controlled by cleaving the corresponding transcript. To enrich our data in terms of the stress responsive gene targets, we compared the identified target genes with *A. thaliana* stress-related genes. In this study, the 22 miRNA families that were identified revealed that 30 target genes were directly related to stress in *A. thaliana*. A list of the names of the miRNAs and the predicted targets is presented in Table 2. To validate whether those *A. thaliana* target genes are conserved in wheat, we carried out a blast analysis for the target gene products containing *T. urartu* proteins. All the proteins are verified in the *T. urartu* genome but five of them are yet to be functionally characterized (Table 2).

We also scanned possible target genes for sequences of the predicted miRNA families in the available *T. aestivum* genome in order to present a more comprehensive putative target list. Screening an EMBL-based reference genome sequence revealed that 113 of the miRNA sequences statistically significantly matched the *T. aestivum* target genes, and 92.90% of 1085 genes were likely to be controlled by cleaving the corresponding transcript. We extracted

the detailed annotation information for the target genes and clustered GO annotations under the terms "biological process", "molecular function", and "cellular component" for 908 putative target genes. We were able to retrieve 14,336 GO term matches, of which 39% were clustered for biological process, followed by 35% for cellular component, and 26% for molecular function clusters (Figure 5). A summary of the functional distribution of the matching GO terms is presented in Figure 6.

4. Discussion

Across the globe, wheat is one of the most demanded crops. In the present study, we aimed to fill an information gap regarding miRNA data for *T. monococcum* subsp. *monococcum*. We carried out a small RNA sequencing analysis to elucidate the miRNA sequences. To increase the number of identified miRNAs, we pooled samples of plants grown under normal, salt, and drought stress conditions. By adopting comparative genomics approaches, we successfully identified 140 distinct miRNA families covering 167 miRNA sequences.

The general sequence profiles of the identified miRNA molecules were similar to those proposed for miRNA characterization studies. Both the first nucleotide bias and the position nucleotide bias observations fit the previously described characteristics of the miRNAs (Lau et al., 2001; Ge et al., 2013). Due to a lack of chromosomal sequence information, we were not able to analyze the secondary structures for all the identified miRNA sequences; however, we did successfully display the structure models for 77 of the 111 analyzed pre-miRNAs extracted from *T. urartu* chromosomal scaffold data.

Table 1. A comparison of the miRNA families identified for *T. monococcum* subsp. *monococcum* in terms of their responsiveness to drought and salt stress in different plant species.

miR	Status of expressional verification									
	T. aestivum		*H. vulgare*		*A. thaliana*		*Z. mays*		*O. sativa*	
	Salt[a]	Drought[f]	Salt[b]	Drought[g]	Salt[c]	Drought[c]	Salt[d]	Drought[i]	Salt[e]	Drought[e]
miR156			√	√	√		√	√	√	√
miR157						√				
miR159		√		√	√		√	√	√	
miR160	√	√					√	√		√
miR164	√		√				√			
miR165					√					
miR166		√	√	√			√	√		√
miR167					√	√	√	√		√
miR168				√	√	√	√	√		√
miR169	√	√	√	√	√			√		
miR171			√		√	√	√			√
miR172	√	√		√						√
miR319					√		√		√	√
miR393				√	√	√				√
miR394					√				√	
miR395	√	√								√
miR396	√				√	√		√		√
miR397										√
miR398								√		
miR408			√			√		√		√
miR444	√			√						
miR528									√	
miR529	√									√
miR530									√	
miR535	√									
miR845										√
miR894								√		
miR1125										√
miR5048				√						
miR5049	√	√		√						

a (Eren et al., 2015), b (Deng et al., 2015), c (Liu et al., 2008), d (Ding et al., 2009), e (Barrera-Figueroa et al., 2012; Zhou et al., 2010), f (Akdogan et al., 2016), g (Hackenberg et al., 2015), i (Wei et al., 2009)
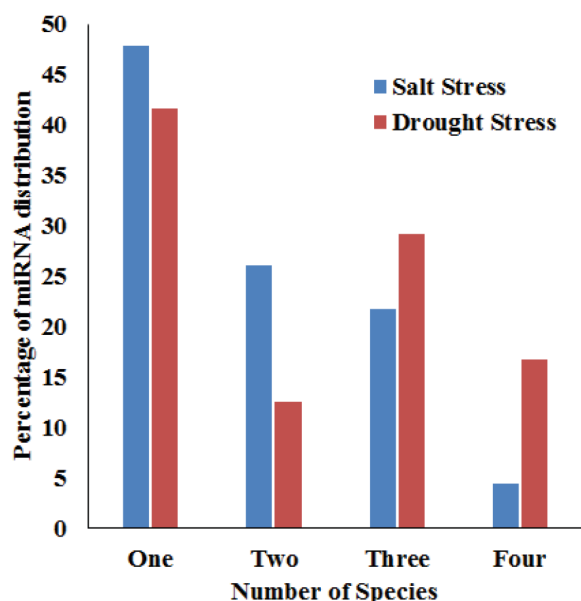
Figure 4. Comparison for number of salt and drought stress related miRNAs among species. Data summarize the conservation level of stress related miRNA among other plant miRNAs that are known to be associated with stress tolerance functions.

We also carried out a bioinformatics analysis for the characterization of the identified miRNA sequences regarding their potential involvement in salt and drought stress regulation. We characterized 23 miRNAs as potential regulators for salt stress and 24 miRNAs as potential regulators for drought stress in *T. monococcum* subsp. *monococcum.* When we analyzed the putative targets for those miRNA sequences, our results showed that 20 target genes and their corresponding miRNAs were identical when compared with the lists for drought and salt stress, except for miR148 and mir845. This suggests that both salt and drought stress are under a common master regulator that controls both conditions, and this fits with a previously proposed model (Deng et al., 2015). It is likely that the *AGO1* gene (predicted as a target for miR168 family) acts as a master regulator for both drought and salt responsive target genes in *T. monococcum* subsp. *monococcum* as previously suggested for *A. thaliana* (Vaucheret et al., 2009). In addition, it was previously shown that this interaction is necessary for a salt stress response in barley, and it is directly related to the miR168 levels under salt stress conditions (Deng et al., 2015). In fact, there is a conserved nature for regulation of miRNA

Table 2. Summary of the target prediction of stress related miRNAs.

| miR family	miR name	Drought	Salt	GenBank (*T. urartu*)	Protein product (*T. urartu*)
miR156	miR156-3p	√	√	EMS63385.1	Argonaute 1B
miR159	miR159-3p	√	√	EMS55264.1	IAA-amino acid hydrolase ILR1-like 5
	miR159-3p	√	√	EMS66412.1	IAA-amino acid hydrolase ILR1-like 3
	miR159-5p	√	√	EMS59656.1	Acetyl-CoA carboxylase
miR165	miR165	x	√	EMS60006.1	3-ketoacyl-CoA synthase 6
miR166	miR166	√	√	EMS47855.1	Putative glutathione S-transferase
	miR166-5p	√	√	EMS67450.1	Zinc finger CCCH domain-containing protein 45
miR168	miR168-5p	√	√	EMS63385.1	Protein argonaute 1B
	miR168	√	√	EMS63385.1	Protein argonaute 1B
	miR168-5p	√	√	EMS48655.1	Proline-rich receptor-like protein kinase PERK13
	miR168-5p	√	√	EMS55864.1	Mitogen-activated protein kinase 17
	miR168	√	√	EMS55864.1	Mitogen-activated protein kinase 17
	miR168-5p	√	√	EMS62275.1	Receptor-like protein kinase
miR169	miR169	√	√	EMS46116.1	D repeat and FYVE domain-containing protein 3
	miR169	√	√	EMS61213.1	Protein TIFY 6B
miR172	miR172	√	√	EMS66886.1	Trihelix transcription factor GT-2
	miR172-5p	√	√	EMS50683.1	Cell division cycle 5-like protein
miR393	miR393	√	√	EMS56796.1	Hypothetical protein
miR395	miR395-5p	√	√	EMS53304.1	Hypothetical protein
	miR395-5p	√	√	EMS68547.1	Alpha-glucan water dikinase, chloroplastic
miR396	miR396-3p	√	√	EMS61897.1	ATP-dependent DNA helicase MPH1
miR398	miR398-3p	√	√	EMS67509.1	Copper chaperone for superoxide dismutase
	miR398-3p	√	√	EMS45437.1	Pectinesterase
miR399	miR399	√	√	EMS47483.1	Disease resistance protein RGA2
	miR399	√	√	EMS48356.1	Hypothetical protein
miR408	miR408	√	√	EMS53029.1	DNA polymerase epsilon catalytic subunit A
miR444	miR444	√	√	EMS60248.1	Hypothetical protein
miR529	miR529	√	√	EMS53316.1	Transcriptional corepressor SEUSS
miR845-5p	miR845-5p	√	x	EMS63076.1	Hypothetical protein
miR5049	miR5049	√	√	EMS46913.1	Pyruvate decarboxylase isozyme 2
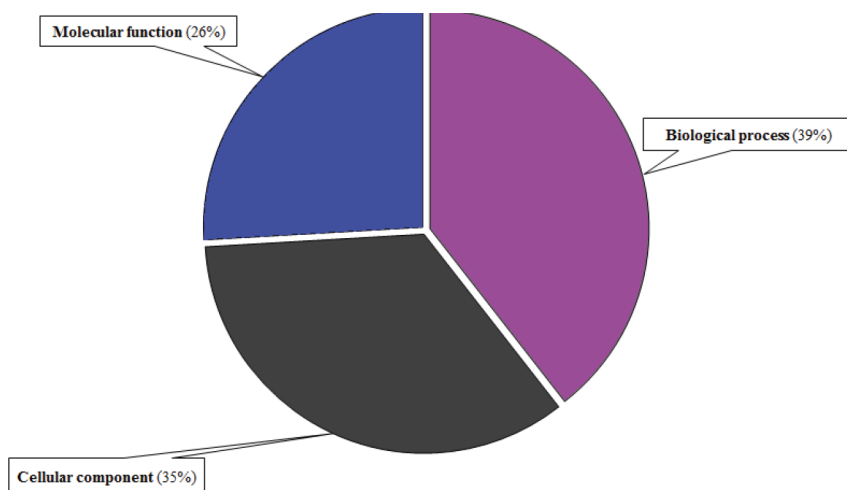
Figure 5. General GO term distributions for miRNAs target genes in the *T. aestivum* genome.
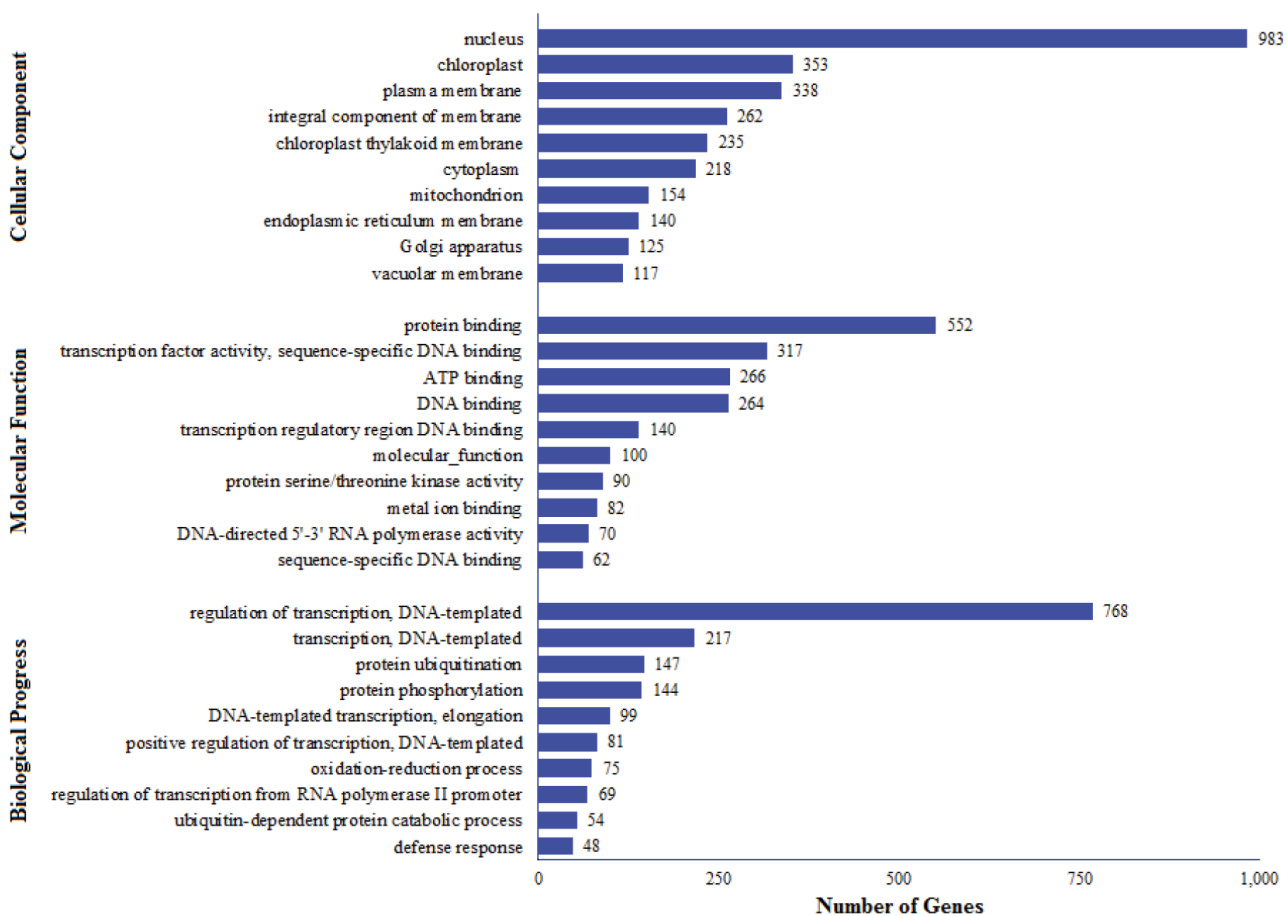

Figure 6. Summarized GO classification of miRNAs target genes in the *T. aestivum* genome. The representation of the number of genes was limited to ten functional classes showing the highest number of genes for each GO term.

machinery especially for stress response (Datta and Paul, 2015). Thus, proposing the involvement of AGO1 in miRNA regulation during stress response would not be misestimating assumption.

In this study, we also carried out a target analysis using the *T. aestivum* genome as a reference. The GO annotation analysis for putative target genes affiliated with the identified miRNAs has a role in protein interactions and the regulation of mRNA levels. Data suggest that the identified miRNAs are involved in transcriptional and posttranscriptional regulatory control. The miRNA/target gene data presented in this study can be used as a reference for comprehensive functional genomics studies.

In conclusion, this study provides a large amount of putative miRNA sequence information for einkorn wheat. As a follow up, the expression profiles of proposed miRNAs and their potential targets under diverse stress conditions should be evaluated in a comprehensive study. In addition, the *T. monococcum* subsp. *monococcum* genome assembly is yet to be completed, and this is the major drawback for detailed molecular and bioinformatics studies on einkorn. Completing the assembly of at least one species among wheat can escalate further molecular studies.

Acknowledgment

References

Akdogan G, Tufekci ED, Uranbey S, Unver T (2016). miRNA-based drought regulation in wheat. Funct Integr Genomics 16: 221-233.

Alptekin B, Budak H (2016). Wheat miRNA ancestors: evident by transcriptome analysis of A, B, and D genome donors. Funct Integr Genomics 17: 171-187.

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

Barrera-Figueroa BE, Gao L, Wu Z, Zhou X, Zhu J, Jin H, Liu R, Zhu JK (2012). High throughput sequencing reveals novel and abiotic stress-regulated microRNAs in the inflorescences of rice. BMC Plant Biol 12: 132.

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

Briggle L (1963). Classification of Triticum species and of wheat varieties grown in the United States. 1st ed. Washington, D.C., USA: United States Department of Agriculture.

Budak H, Akpinar BA (2015). Plant miRNAs: biogenesis, organization and origins. Funct Integr Genomics 15: 523-531.

Carthew RW, Sontheimer EJ (2009). Origins and mechanisms of miRNAs and siRNAs. Cell 136: 642-655.

Curwen-McAdams C, Arterburn M, Murphy K, Cai X, Jones SS (2016). Toward a taxonomic definition of perennial wheat: a new species ×Tritipyrum aaseae described. Genet Resour Crop Evol 64: 1651-1659.

Dai X, Zhao PX (2011). psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res 39: W155-9.

Datta R, Paul S (2015). Plant microRNAs: master regulator of gene expression mechanism. Cell Biol Int 39: 1185-1190.

Deng P, Wang L, Cui L, Feng K, Liu F, Du X, Tong W, Nie X, Ji W, Weining S (2015). Global identification of microRNAs and their targets in barley under salinity stress. PLoS One 10: e0137990.

Ding D, Zhang L, Wang H, Liu Z, Zhang Z, Zheng Y (2009). Differential expression of miRNAs in response to salt stress in maize roots. Ann Bot 103: 29-38.

Eren H, Pekmezci MY, Okay S, Turktas M, Inal B, Ilhan E, Atak M, Erayman M, Unver T (2015). Hexaploid wheat (*Triticum aestivum*) root miRNome analysis in response to salt stress. Ann Appl Biol 167: 208-216.

Ge X, Zhang Y, Jiang J, Zhong Y, Yang X, Li Z, Huang Y, Tan A (2013). Identification of microRNAs in *Helicoverpa armigera* and *Spodoptera litura* based on deep sequencing and homology analysis. Int J Biol Sci 9: 1-15.

Griffiths-Jones S, Saini HK, Dongen S van, Enright AJ (2008). miRBase: tools for microRNA genomics. Nucleic Acids Res 36: D154-8.

Hackenberg M, Gustafson P, Langridge P, Shi BJ (2015). Differential expression of microRNAs and other small RNAs in barley between water and drought conditions. Plant Biotechnol J 13: 2-13.

Jiang H, Lei R, Ding SW, Zhu S (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15: 182.

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

Karagöz A, Zencirci N (2005). Variation in wheat (*Triticum* spp.) landraces from different altitudes of three regions of Turkey. Genet Resour Crop Evol 52: 775-785.

Kersey PJ, Allen JE, Armean I, Boddu S, Bolt BJ, Carvalho-Silva D, Christensen M, Davis P, Falin LJ, Grabmueller C et al (2016). Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res 44: D574-D580.

Kozomara A, Griffiths-Jones S (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39: D152-D157.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.

Lau NC, Lim LP, Weinstein EG, Bartel DP (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294: 858-862.

Liu HH, Tian X, Li YJ, Wu CA, Zheng CC (2008). Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. RNA 14: 836-843.

Mahmood I, Razzaq A, Hafiz I, Kaleem S, Khan A, Qayyum A, Ahmad M (2002). Interaction of callus selection media and stress duration for in vitro selection of drought tolerant callus of wheat. African J Biotechnol 11: 4000-4006.

Mathews DH, Sabina J, Zuker M, Turner DH (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288: 911-940.

Mayer KF, Rogers J, Doležel J, Pozniak C, Eversole K, Feuillet C, Gill B, Friebe B, Lukaszewski AJ, Sourdille P et al (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345: 1251788.

Peters L, Meister G (2007). Argonaute proteins: mediators of RNA silencing. Mol Cell 26: 611-623.

Ritchie W, Legendre M, Gautheret D (2007). RNA stem-loops: to be or not to be cleaved by RNAse III. RNA 13: 457-462.

Serpen A, Gökmen V, Karagöz A, Köksel H (2008). Phytochemical quantification and total antioxidant capacities of emmer (*Triticum dicoccon* Schrank) and einkorn (*Triticum monococcum* L.) wheat landraces. J Agric Food Chem 56: 7285-7292.

Ünlü ES, Gordon DM, Telli M (2015). Small RNA sequencing based identification of miRNAs in *Daphnia magna*. PLoS One 10: e0137617.

Vaucheret H (2009). AGO1 homeostasis involves differential production of 21-nt and 22-nt miR168 species by MIR168a and MIR168b. PLoS One 4: e6442.

Wei L, Zhang D, Xiang F, Zhang Z (2009). Differentially expressed miRNAs potentially involved in the regulation of defense mechanism to drought stress in maize seedlings. Int J Plant Sci 170: 979-989.

Zhang B, Pan X, Cobb GP, Anderson TA (2006). Plant microRNA: a small regulatory molecule with big impact. Dev Biol 289: 3-16.

Zhou L, Liu Y, Liu Z, Kong D, Duan M, Luo L (2010). Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. J Exp Bot 61: 4157-4168.

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

SUPPLEMENTARY DATA
S1 Table. Sequence information for the identified miRNAs.

Predicted miRNA family name	Predicted miRNA name	Sequence length	Copy number	Sequence
miR22	tmo-miR22	22	1	AAGCUGCCAGUUGAAGAACUGU
miR30	tmo-miR30-5p	20	1	UGUAAACAUCCUCGACUGGA
miR99	tmo-miR99	21	1	CACCCGUAGAACCGACCUUGC
miR156	tmo-miR156	20	5	UGACAGAAGAGAGUGAGCAU
	tmo-miR156-3p	21	16	GCUCACCCUCUCUCUGUCAGC
	tmo-miR156-5p	20	1	UGACAGAAGAGAGCGAGCAC
miR157	tmo-miR157	21	1	UUGACAGAAGAUAGAGAGCAC
miR159	tmo-miR159	21	1	UUUGGAUUGAAGGGAGCUCUG
	tmo-miR159-3p	18	1	UUUGCAUGACCGAGGAGC
	tmo-miR159-5p	21	1	AGCUGCUUGUUCAUGGUUCCC
miR160	tmo-miR160	21	1	UGCCUGGCUCCCUGAAUGCCA
miR164	tmo-miR164	21	7	UGGAGAAGCAGGGCACGUGCA
miR165	tmo-miR165	21	1	UCGGACCAGGCUUCAUCCCCC
miR166	tmo-miR166	21	1	UCGGACCAGGCUUCAUUCCCC
	tmo-miR166-3p	17	1	UCGGACCAGGCUUCAAU
	tmo-miR166-5p	18	2	GGUUGUUGUCUGGUUCAA
miR167	tmo-miR167	22	1	UGAAGCUGCCAGAAUGAUCUGA
	tmo-miR167-3p	18	13	UCAUGCUGGAGUUUCAUC
miR168	tmo-miR168	21	2	UCGCUUGGUGCAGGUCGGGAA
	tmo-miR168-3p	18	13	CCCGCCUUGCACCAAGUG
	tmo-miR168-5p	19	1	UUGCUUGGUGCAGAUCGGG
miR169	tmo-miR169	19	1	AGCCAAGGAUGACUUGCCA
	tmo-miR169-3p	19	4	GGCAGUCUCCUUGGCUAGC
miR171	tmo-miR171	21	3	UUGAGCCGUGCCAAUAUCACG
	tmo-miR171-3p	20	2	UUGAGCCGUGCCAAUAUCAC
	tmo-miR171-5p	19	1	UGGUAUUGUUUCGGCUCAU
miR172	tmo-miR172	19	2	GAAUCUUGAUGAUGCUGCA
	tmo-miR172-5p	21	1	GCAGCACCACCAAGAUUCACA
miR181	tmo-miR181-5p	22	1	AACAUUCAACGCUGUCGGUGAG
miR182	tmo-miR182	18	1	GUGGCACUAGUGGAAUUC
miR319	tmo-miR319	19	2	UUGGACUGAAGGGAGCUCC
	tmo-miR319-3p	21	46	CUUGGACUGAAGGGUGCUCCC
	tmo-miR319-5p	20	3	AGAGCGUCCUUCAGUCCACU
miR390	tmo-miR390	21	1	AAGCUCAGGAGGGAUAGCGCC
	tmo-miR390-3p	19	2	GCUAUCUAUCCUGAGCUCC
miR393	tmo-miR393	21	46	UCCAAAGGGAUCGCAUUGAUC
	tmo-miR393-3p	22	1	GAUCAGUGCAAUCCCUCUGGAA
miR394	tmo-miR394	17	5	UUGGCAUUCUGUCCACC
	tmo-miR394-3p	18	1	GUGGGCAUACUGCCAAUG
miR3944	tmo-miR3944-3p	18	1	ACCUUCGGGCUGGCCUGC

S1 Table. (Continued).

miR395	tmo-miR395	21	1	UGAAGUGUUUGGGGGAACUCC
	tmo-miR395-5p	18	1	UGAAGUGUUUGAGGGAAC
miR396	tmo-miR396	20	1	UCCACAGGCUUUCUUGAACU
	tmo-miR396-3p	21	4	GGUCAAGAAAGCUGUGGGAAG
miR397	tmo-miR397	20	12	UUGAGUGCAGCGUUGAUGAA
miR398	tmo-miR398	20	1	UGUGUUCUCAGGUCACCCCU
	tmo-miR398-3p	20	1	GUGUUCUCAGGUCGCCCCUG
miR399	tmo-miR399	21	3	UGCCAAAGGAGAAUUGCCCUG
miR408	tmo-miR408	21	1	UGCACUGCCUCUUCCCUGGCU
miR414	tmo-miR414	17	1	GAGGAUGAUGAGGAUGA
miR444	tmo-miR444	19	1	UGCAGUUGCUGUCUCAAGC
	tmo-miR444-3p	21	8	UGCAGUUGCUGCCUCAAGCUU
miR456	tmo-miR456-5p	18	1	UGCACUGCCUUCAGAGUG
miR466	tmo-miR466-5p	19	1	AACACACACACACACAC
miR479	tmo-miR479	21	1	UGAGCCGAACCAAUAUCACUC
miR529	tmo-miR529	21	1	AGAAGAGAGAGAGUACAGCCC
	tmo-miR529-3p	18	1	GCUGUACCCUCUCUCUUC
miR530	tmo-miR530	20	1	CUGCAUUUGCACCUGCACCU
miR535	tmo-miR535	22	1	UGACAACGAGAGAGGGCACGCG
miR619	tmo-miR619-5p	22	1	GCCUCGGCCUCUCAAAGUGCUG
miR650	tmo-miR650	21	1	CCAUGGUGGAGAUGUCCUGAG
miR706	tmo-miR706	22	1	CCAGGGCUAUACAGAGAAACAC
miR716	tmo-miR716	19	1	CGAGCCCGGGCGGAGCGGC
miR767	tmo-miR767-5p	23	1	UGCACCAUGGUUGUCUGAGCAUG
miR827	tmo-miR827	21	1	UUAGAUGACCAUCAGCAAACA
	tmo-miR827-5p	22	1	UCUGAACUUGUUUUGCUGGUUG
miR845	tmo-miR845-5p	19	1	ACCUUGCUCUGAUACCAAU
miR894	tmo-miR894	20	2	UUCGUUUCACGUCGGGUUCA
miR928	tmo-miR928	17	1	GUGGCUGUGGAAGCUGG
miR1117	tmo-miR1117	19	1	UUAGUACCGGUUCGUGGCA
miR1120	tmo-miR1120	18	1	AUUUUUAUAUUAUGAGAC
miR11214	tmo-miR11214	20	1	UAGUGAUCUAAACGCUCUUA
miR1122	tmo-miR1122	19	1	GUCUAGAUACGGAUGUAUC
miR1125	tmo-miR1125	24	1	AAAUUUAACCAACGAGACCAACUG
miR1131	tmo-miR1131	18	1	CUUUAGUACCGGUUCGUG
miR1133	tmo-miR1133	18	1	AAGUUUUUUCGGACGGAG
miR1135	tmo-miR1135	23	1	CCGUUCGGAAUUACUUGUCGCAG
miR1136	tmo-miR1136	22	2	ACUUGUCGCAGGUAUGGAUAUA
miR1137	tmo-miR1137	18	1	AGUUAGUACAAAGUUGAG
miR1139	tmo-miR1139	22	1	AUGUUACUAGUGUAUGUUACUC
miR1207	tmo-miR1207-5p	18	1	GGGGCAGGGAGGCAGGGA
miR1273	tmo-miR1273	22	1	AAUGAUUCGAUCUCGACUCACU
	tmo-miR1273-3p	18	1	GUCCUGCUCUGUCACCCA

S1 Table. (Continued).

miR1285	tmo-miR1285	20	1	CAGAGGUUGCAGUGAGUGGA
miR1432	tmo-miR1432-5p	20	2	UCAGGAGAGAUGACACCGAC
miR1436	tmo-miR1436	19	1	AUUAUGGGACGGAGGGAGU
miR1520	tmo-miR1520	18	1	CCCAUCACGUGUCAUGUU
miR1584	tmo-miR1584	18	1	AGGAUCAAGGGAAUCGGG
miR1878	tmo-miR1878-3p	23	2	AUUUGUAGUGUUCGGAUUGAGUU
miR2111	tmo-miR2111-5p	21	1	UAAUCUGCAUCCUGAGGUUUA
miR2120	tmo-miR2120	18	1	GAACCGGGACUAAAGAUC
miR2478	tmo-miR2478	18	1	AGAGGGCGUGGGUUCAUA
miR2525	tmo-miR2525	19	1	UUUGAUCCACUUCGCUGUC
miR2538	tmo-miR2538-5p	18	1	AUCCUCUAUUAUUUUAGU
miR2673	tmo-miR2673	18	1	CUUUUCUUCCUCUUCCUC
miR2916	tmo-miR2916	20	1	CAAGAACGAAAGUUGGGGAC
miR2919	tmo-miR2919	19	1	CCUGCCGUCGCUGUGCUUC
miR3348	tmo-miR3348	17	1	CCUCGCCGGGAGGCUCG
miR3630	tmo-miR3630-3p	17	31	AUGGGAAUCUCUCUGAU
miR3682	tmo-miR3682-5p	18	1	AGGAUAACACAGGUAGAA
miR3711	tmo-miR3711	18	3	GCCCUCCUUCUAGCGCCA
miR3885	tmo-miR3885-5p	19	1	UGCUGAGCGGCGGCCGCCG
miR3887	tmo-miR3887-3p	18	1	GGAGAGAUGGCUGUGGAA
miR4922	tmo-miR4922	18	1	UAAAUUGUAUCAUUUUUC
miR4995	tmo-miR4995	21	6	CAUAGGCAGUGGCUUGGUUAA
miR5021	tmo-miR5021	18	1	CUACAAUUUCUUCUUCUU
miR5048	tmo-miR5048	20	1	UAUAUUUGCAGGUUUUAGGU
miR5049	tmo-miR5049	23	1	AGCUGAGACACUUAUUUUGGGAC
	tmo-miR5049-3p	20	1	CAAGUAAUAUGGAUCGGAGG
miR5050	tmo-miR5050	17	1	UUUUGCUGGUUGAACGA
miR5054	tmo-miR5054	18	1	AACCACGUGGCCGUGGGU
miR5056	tmo-miR5056	21	1	UCGGGAGGAAGAACCGGUAAU
miR5059	tmo-miR5059	17	1	CGAGCCUGGGCAGCACC
miR5062	tmo-miR5062	20	2	UGAACCUUGGGGAAAAGCCG
miR5064	tmo-miR5064	20	38	UGAAUUUGUCCAUAGCAUCA
miR5067	tmo-miR5067	18	1	UUCAUAUUAGUUGUCGCU
miR5072	tmo-miR5072	19	4	UUCUGGGUUCGUUCCCCAG
miR5073	tmo-miR5073	23	1	GUUUGGUGAAUCGGAAACAAUUU
miR5076	tmo-miR5076	21	1	UCUUUUUCCUUAAAUGGGAGC
miR5079	tmo-miR5079	22	1	UAUAAUUUGGAUUUGUUAUUUU
miR5082	tmo-miR5082	19	2	GCGAUGAUGGCCGCGCGGG
miR5083	tmo-miR5083	20	1	UAUUUAGUGUUGACCAAAUU
miR5084	tmo-miR5084	20	1	GUGAUCCUCUGCAGUACUGU
miR5096	tmo-miR5096	21	1	AGACAGGGUUUCACCAUGUUG
miR5106	tmo-miR5106	18	1	GGGUCUGUAGCUCAGUUG

S1 Table. (Continued).

miR5141	tmo-miR5141	17	1	CCGUCAGUCGCGUCGGG
miR5169	tmo-miR5169	18	1	UUGACCAAGUUUGUAGAA
miR5174	tmo-miR5174-3p	19	1	UUAUGGAACGGAGGGAGUA
miR5174	tmo-miR5174-5p	19	1	CAAAAACGCUGUUAUAUUA
miR5181	tmo-miR5181	19	1	AACUGCGACACUUAUUAUG
miR528	tmo-miR528-5p	21	2	UGGAAGGGGCAUGCAGAGGAG
miR5368	tmo-miR5368	18	1	GACCCGCGGGCCAAGGGA
miR5387	tmo-miR5387	18	1	CGAACCGGUGCUAAAGGA
miR5503	tmo-miR5503	22	1	AAUGCCUCUAGAAAGAUCCGAA
miR5523	tmo-miR5523	19	1	UAACUAGUAAAUAUGUUCC
miR5532	tmo-miR5532	22	1	UAUGGAAUAUAUGACAAAGGUG
miR5538	tmo-miR5538	22	1	ACUGUUGAGUAACGGCAGCAAG
miR5571	tmo-miR5571-5p	21	1	AUGUGAACCAAGCAAUUCUCA
miR5585	tmo-miR5585-3p	22	1	CCAGGCAAGGUGGCGGGCACCU
miR5658	tmo-miR5658	19	1	GAUGAGAUGAUGAUGAUGA
miR6173	tmo-miR6173	20	1	AUGGGAUUAGAGACCCCAGU
miR6177	tmo-miR6177	20	1	CCAUGGACAGAAGGCACUUA
miR6181	tmo-miR6181	22	1	UGCUCUUCAUGGACUGCGGCGC
miR6182	tmo-miR6182	21	1	GAGUGUGUGAUGGAUGGCUUU
miR6188	tmo-miR6188	19	1	GGAGGAUCGAUGAACCCGG
miR6191	tmo-miR6191	18	1	CUUAGAUUUGUCUAGAUA
miR6198	tmo-miR6198	22	2	CGGCUCUGUCUUGGAUGGUCAU
miR6199	tmo-miR6199	18	1	CCACAGAAUUCUCACAGU
miR6203	tmo-miR6203	21	1	AGGGAUUGCAGGUCUUCUUAA
miR6204	tmo-miR6204	22	1	AGAAAUGGAAAGGAGAAUAAUU
miR6214	tmo-miR6214	20	1	ACGACGACGACGAGCACGAC
miR6219	tmo-miR6219-5p	18	1	UGUAAGAACCGGGACUAA
miR6244	tmo-miR6244	19	1	CCUUGUGGUCGUGGGUUCG
miR6250	tmo-miR6250	20	1	UGCCGCCAAUCUUCUCGGGG
miR6253	tmo-miR6253	19	1	AGGAAAGUGGGCAGUUGGG
miR6300	tmo-miR6300	18	1	GUCGUUGUAGUAUAGUGG
miR6478	tmo-miR6478	20	1	CCGACCUUAGCUCAGUUGGU
miR6621	tmo-miR6621-5p	19	1	AUCUGGUACAACAGCCUGU
miR6874	tmo-miR6874-3p	18	1	UUUACCUAGUUCUGCUGU
miR6981	tmo-miR6981-5p	22	1	AGAGGAGAAGGAAGAAGCUGAA
miR7042	tmo-miR7042-3p	18	1	GUAUCAAGAGAGAAAACA
miR7116	tmo-miR7116-3p	18	1	UCCUUUUUCCUUUGCCUU
miR7398	tmo-miR7398-3p	17	1	CGUAAGAGAAGGGAGAA
miR7757	tmo-miR7757-5p	17	2	CACAAAACCUUCAGCUA
miR8155	tmo-miR8155	17	6	ACCUGGCUCUGAUACCA
miR-B6	tmo-miR-B6-3p	17	1	CGUCUCCGGCGCCGGGU
miR-I5	tmo-miR-I5-3p	18	1	GGAUGAAGAAGACGACGA

S2 Table. Predicted secondary structures for putative miRNA sequences identified in *T. monococcum* subsp. *monococcum*.

Predicted miRNA name	Predicted secondary structure
| tmo-mir-156-3p | ```
   U       -   A       A      GGUU        - G
GG CUGACAGA AGAG GUGAGCAC CGCGGU      UCCUAGCAUG C A
CC GACUGUCU UCUC CACUCGUG GCGUCG      AGGGUCGUAC G G
   -           C   C        C        ----       C C
``` |
| tmo-mir-156-5p | ```
G          A  A     C   CA--    G--   A CCC     GU
 GGAGGCUGAC GA GAGAG GAG    CACG  CGG GG  GCGUC  \
 CCUCCGGCUG CU UUCUC CUC    GUGC  GCC CC  CGCAG  G
G          C  -     -   CCUA    ACA  -  CAC    AG
``` |
| tmo-mir-159 | ```
C     UUU    GAC      CG-        CU--   UG
 UUGUGG  GCAU   CGAG--GAGC   CUUCGAUCC   GGC  \
 GACACC  CGUA   GUUC CUCG    GAAGUUAGG    UCG  A
U     U--    AAA   \      AGG       UUUG    CC
``` |
| tmo-mir-159-3p | ```
   .-GUGCA   G   U     GUUC    UAU  C     A   A
AGGGUUU       GCU CUUG UCAUG    CCAC   CC AUCUCC UUG A
UCUCGAG       CGA GAGC AGUAC    GGUG   GG UAGAGG AGC A
   \ -----    G    C      GUUU   UUC  C     -   A
``` |
| tmo-mir-159-5p | ```
UAUCGA   U  U-  A  G   U     GUUC    UAU  C     A   A
     AGGG UUG  GC GCU CUUG UCAUG    CCAC   CC AUCUCC UUG A
     UCCC AGC  CG CGA GAGC AGUAC    GGUG   GG UAGAGG AGC A
UCGG--   U  UU  C  G   C     GUUU   UUC  C     -   A
``` |
| tmo-mir-160 | ```
GGU   ---      C      G     U  .-GAGA   C GU      C
  GCC   GCUUGC UGGCUCCCU AAUGCCA CC    AGCG \  GGGAGG G
  CGG   CGGACG ACUGAGGGA UUACGGU GG     UCGC G  UCCUCC U
--- CUC      U      G       U \ ----    C --      C
``` |
| tmo-mir-164 | ```
UGGAGA  C  GG      -   U- G  C   -  UCUCC    UC
   AG AG CACGUGCA UGCA GC AGCG GCUC GA   UCCUGCC  C
   UC UC GUGCACGU ACGU CG UCGU CGAG CU   AGGGCGG  G
------  U  UU      C   CC A   -   G -----     CU
``` |
| tmo-mir-166 | ```
.     UU   CGUC    U   GA       C   -  A      .-GAGAG    A
CUUGG  CCG    AUGG UGUC GGGGAAUGA GCC GG UCCGAAA      ACGC \
GGACU  GGU    UACC ACAG CCCCUUACU CGG CC AGGCUUU      UGCG U
U-  AC--    -    UA       U  A -      \ -----    G
``` |
| tmo-mir-166-3p | ```
      UU        A     U   ---      --     A    GA
·GGGGGUUG  GUCUGGUUC AGGUC CCA    CAUACA  UCAUAU CAUG  \
UCCCUAAC   CGGACCAGG UCCAG GGU    GUAUGU  GGUAUA GUAC  G
      UU        C     U   UUA      UA     -    GA
``` |
| tmo-mir-166-5p | ```
UUAA      UU        A     U   ---      --     A    GA
  GGGGGUUG  GUCUGGUUC AGGUC CCA    CAUACA  UCAUAU CAUG  \
  UCCCUAAC   CGGACCAGG UCCAG GGU    GUAUGU  GGUAUA GUAC  G
  ----        UU        C     U   UUA      UA     -    GA
``` |
| tmo-mir-167-3p | ```
   A  -  A  U      G              -  CUAACUC     C
GUGC CC AC AGC GGUGAAGCU CCAGCAUGAUCUGAU GAC       AUGGAU A
CACG GG UG UCG CUACUUUGA GGUCGUACGGACUA CUG       UACCUA G
   A  C  G  U      -              A   -------     A
``` |
| tmo-mir-169 | ```
G       GAA  C       -      A-      ACAG
 GAGAGUG    UG AGCCAAGGAUG ACUUGCC  GCA--GC      A
 CUCUCAU    AC UCGGUUCCUAC UGAACGG  CGU  CG      A
U       UCC  A       C      CC   \   GAAC
``` |
| tmo-mir-169-3p | ```
CCCGA-,      U   U   - -  U-----  AU    --  UU UC
  AGCCAAGGA GACU GCCUG UG GC   GUGA GGGAUC UCU C \
  UCGGUUCCU CUGA CGGAC AC CG   CACU CUCUGG AGG G U
-----       -    -  U A UAUCAU  --     CU UG UC
``` |

S2 Table. (Continued).

| tmo-mir-171 | <pre>   GG   U   --   --   UG   AG A U UG   AGGA
--UGCGAG GAG GAA CGCG GUAUUGG CGGUUCAAUC AG GC GG CCCC   \
  ACGCUC UUC CUU GCGC UAUAACC GCCGAGUUAG UC CG CC GGGG   G
   \   AG   U   UG   AC   GU   CU C - UU   AACA</pre> |
|---|---|
| tmo-mir-172 | <pre>   C C   A   GA-- -   U   U
GGUGCAGCA CA CAAGAUUC CAUCG   UC CGUCG CGUAAAU \
CUACGUCGU GU GUUCUAAG GUAGU   AG GCAGC GUAUUUA A
          A A   A   GGAC C   -   A</pre> |
| tmo-mir-172-5p | <pre>UU A   C C   A   GA-- -   U   U
  GC GGUGCAGCA CA CAAGAUUC CAUCG   UC CGUCG CGUAAAU \
  CG CUACGUCGU GU GUUCUAAG GUAGU   AG GCAGC GUAUUUA A
A- A   A A   A   GGAC C   -   A</pre> |
| tmo-mir-319-3p | <pre>ACCGU---   U   CU-   --   U UUUCU   CC
       CGU GCUUGGA   GA AGGG G   UCGGGCG A
       GUA CGAAUCU   CU UCCC C   AGUCCGC U
GAUGUAUU   C   CUC   CC   U AGCU-   CA</pre> |
| tmo-mir-384-5p | <pre>C   .-A UG AA CCA   ---   UCC- AU   CCGUA
AAGGAG GCA AGC \ GGCAG GUCA UC GUGCUUA   \
UUCCUC CGU UCG G CUGUC CGGU AG CACGAAU   G
-    \ - UG CC CAC   UUA UUAU AC   CUUAA</pre> |
| tmo-mir-393 | <pre>   A   U   U CAUCCA   --   GC
GGGGAAGC UCCAAAGGGAUCGCAU GAUCU UC   UGGU GUUGAUG \
CUCCUUCG AGGUUUUCCUAGCGUA CUAGG AG   ACCA UAACUAC U
        A   -   C CUCG--   AC   UG</pre> |
| tmo-mir-393-3p | <pre>   A U   A   C   -- ------   U
GGA GC AGUGGAGGAUUCCA AGGGAU GCAUUGAUCCAUC UCU   CCG A
CCU CG UCGCCUCUUAAGGU UCCCUA CGUGACUAGGUAG AGA   GGC A
   C U   C   A   CU ACUCGC   C</pre> |
| tmo-mir-397 | <pre>   - C AAAGG C   A   GCGUUGAUG   - UA- C
GGAGGAAG AGA GC   UG CAUUG GUGCA   AACCG UCC   CUC U
CCUCUUUC UCU CG   GC GUAAC CACGU   UUGGC AGG   GAG C
   G   - ----- A   A   ---------   G CCG C</pre> |
| tmo-mir-398 | <pre>   A   CCA   .-A   G-   UC
CGAUCCAGAGG GUG   CUGAGAACAC   AGCGC GGC C
GUUGGGUUUCC CAC   GACUCUUGUG   UCGUG CUG U
          C   UG-   \ -   GA   UU</pre> |
| tmo-mir-399 | <pre>   U- UGUA- ---   AG UG   G AGU
GGCAUGGU--GGCA GCA   CC GGUA UG CGGC UGC   U
UCGUACCA CCGU CGU   GG CCGU GC GCCG ACG   U
    \   CC UAAGA AAA   -- GU   A GUC</pre> |
| tmo-mir-414 | <pre>UG   G   A   ACAAAG   GAA
  GUCGU AUUU UGUCA   CA   G
  CAGUA UAGG GUAGU   GU   A
--   G   A   AGGA-- AUU</pre> |
| tmo-mir-466-5p | <pre>   --   -   A GGA   C
  UGUGUGU UGUGUG UG   UUG A
  ACACACA ACACAC AC   GAC U
AC   C   A AA-   A</pre> |
| tmo-mir-528-5p | <pre>G- A   CG   U G   C -   UGG   UGCUU
  AGC GCAG GUGGAAGGGGCA GCA AGGAG G GCCA   GAGCUU   \
  UCG CGUC UACCUUCUCCGU CGU UCCUC C CGGU   CUCGGA   G
AG G   CU   C G   U U   ---   UCUCC</pre> |
| tmo-mir-530 | <pre>-CAUA U UU U CAG   GGA   AGAGAG   -- AAAG   A
    CU GGU GCA GU GCAAG   AGCU   CCA GC   UGC \
    GA CCA CGU CA CGUUU   UCGA   GGU UG   ACG C
---- U   --   C ---   ACG   GAUAG-   AU GGAA   A</pre> |

S2 Table. (Continued).

tmo-mir-827-5p	<pre>U -- CU AU UAUAUCUCUAGUUCAU CU UUGUUUU C UC
	CGGCA CU CA CG GCAG CUGAAC G UGGUUG \
	GUGGU GA GU GC CGUC GACUUG C ACCAAU A
	- UC AU CU --------------- U- ------- U CU</pre>

tmo-mir-845-5p	<pre>C UAAA .-CCA GG AAAAUCACA--- A UGA
	 AUAAC ACU GGC U CAGA CAAG \
	 UAUUG UGA CCG U GUCU GUUC G
	A UG-- \ --- AA AAGUUAACCAUA C CAU</pre>

tmo-mir-1117	<pre>-- UU AU GGUA
	  UC UAGUACCGGUUCGUGGC GAACC \
	  AG AUCAUGGCCAAGCACCG CUUGG C
	CU UU -- AAAU</pre>

tmo-mir-1120	<pre>·UGAGA C C UC UC UUUUAA
	     GUACUA UC CU GU UAUAAUAUAG G
	     CAUGAU AG GA CA GUAUUAUAUU A
	----- A U GA GA UUUAUA</pre>

tmo-mir-1122	<pre>. U - CA
	AAUUACUUGUCUUGGAUUUGUC AGAUACGGAUGUAUCUAG ACU U
	UUAAUGAACGGAACCUAAACAG UCUAUGCCUACAUAGAUC UGA U
	. U G UU</pre>

tmo-mir-1131	<pre>---------- A CGU C A UCU
	      GGACCUUUAGU CCGGUU GGCACGAAC GGGACUAA GG C
	      UCUGGAAAUCA GGCCAA CCGUGCUUG CCCUGAUU CC A
	UGAUUUCUAA G ACU U A CCA</pre>

tmo-mir-1133	<pre>-GAGAAAA A- C--- C CA CUAA
	    CUU UCC AAGCUUGUCCCU AAA GAUGUAUCUAACA \
	    GAG AGG UUUGAACAGGGA UUU CUACAUAGAUUGU C
	------- GC CUUU A AC AGUU</pre>

tmo-mir-1135	<pre>UUUA- GCA A - GUA
	    UACUCCCUCCGUUCGGAAUUACUUGUC GA AUGGAUGUAUCUAGA C \
	    AUGAGGGAGGCAAGCCUUAAUGAACAG CU UACCUACAUAGAUCU G U
	GUAAG AGC A U AUU</pre>

tmo-mir-1136	<pre> UG GU C UGUAU UACUUCCUUCGUUC AAUUACUUGUCGCAG AUGGAUAUAU UAGA \ AUGAGGGAGGCAAG UUAAUGAGCAGCGUC UACCUAUAUA AUCU U GU UU A UAAUU</pre>
tmo-mir-1137	<pre> A G A - U CCAAA UAAGUGUCUCAA CUU GUAC AACUUUG \ GGUUU AUUCACAGAGUU GAA CAUG UUGAAAU A C - A A C</pre>
tmo-mir-1432-5p	<pre> A A A G AU C--- A GGGUCCUGUG UCAGG GAG UGACACCGAC CCG CGGAUGGGU GGCUU A CUCGGGAUAC AGUCC CUC ACUGUGGUUG GGC GUCUGCCUA CCGGA C A G C A CG CGUA C</pre>

tmo-mir-1436	<pre>CCUCAU G U AAC
	    GUACUCC UCCGUCCCAUAAUAUAAGAGCGUUUUU ACAC \
	    CAUGAGG AGGCAGGGUAUUAUAUUCUUGCAAAAA UGUG A
	------ G C AUC</pre>

tmo-mir-1584	<pre>-------- AGGAGAA AGAGGGAGAG AAGGGAA GGG
	    GCGGCGCU UG GAUC UC U
	    CGUCGUGG GC CUAG AG A
	AGAGGAAG GGUAG-- GAA------- GGGGCAG AGG</pre>

tmo-mir-1878-3p	<pre>UC CA C U GC UG A U G
	 AAUCUUA AAC ACU UU AAACUUAGUCU GCACUAUAAAUUUA AAUG GCAUGU \
	 UUAGAAU UUG UGA AA UUUGAGUUAGG UGUGAUGUUUAAAU UUAC CGUAUA A
	-- C- C U UA CU A U A</pre>

S2 Table. (Continued).

| tmo-mir-2120 | <pre>A- G   AGUAUAA   A- UAGG A
CGU GAGGCCCAUCUGUCCCGGUU   GAACGGGACUAAAG UC   GC U
GUA UUCCGGGUAGACAGGGCCAA   CUUGGCCCUGAUUUC AG   UG U
AA -   CC-----   CC CAA- A</pre> |
|---|---|
| tmo-mir-2538-5p | <pre>A   CAGUGU   UC- UC AUC   AUUUUA UC
UUUGAGAGAG   CCUUCG UUCA U CUCUAU   GU U
AGACUUUUUU   GGAAGC AGGU A GAGAUAA   CA U
-        ACCCU-   UUU UA A--   CGUACC CC</pre> |
| tmo-mir-2673 | <pre>CC U   UUCC------   CAG   GCUGG   U UG
UU UCUUCCUC   UCGGCA   GAC   GUCGCCU AU C
AG GGGAGGAG   AGCCGU   UUG   CGGCGGG UA C
GC C   CACGCCACAU   ---   A----   U UC</pre> |
| tmo-mir-3348 | <pre>CACA   U -   --   C-   ----   .-GC   AUU   AG
   AUAGC CC CCGCG AUCCU GCC   GGGAG   UCGCUGG   CCA G
   UGUCG GG GGCGU UAGGA CGG   CCCUC   AGCGACC   GGU C
C---   U A   AG   AC AACC   \ --   ---   AC</pre> |
| tmo-mir-3630-3p | <pre>UC UG   -- U U A-   ----   AG ACA- G   U
   CA GAAACA AG GA GAU CCA   AGAC GG   UCA AACCC U
   GU UUUUGU UC CU CUA GGU   UCUG UC   AGU UUGGG C
UA --   AG U - AG   AAUA   CA AGAC   A   G</pre> |
| tmo-mir-3682-5p | <pre>-UUU GA A --   UG ---   AA----   UGUGAGAA
   UG GU U   AACC GU UGAGGGACU   UGUGUU   \
   AC CA U   UUGG UA ACUCCCUGA   ACACAA   U
--- AG G UU   CU GUU   AGAUGG   UAGGAUGU</pre> |
| tmo-mir-3711 | <pre>   U   UG U CGAGCCCUCCUUC   C UCUCG-   UCG
CGCGGG GC U G   UAGCG CA   GGU--CAGC C
GUGUCC UG G U   GUCGC GU   CCA GUCG C
    - GU U U-----------   U UAAAAA   \   UCU</pre> |
| tmo-mir-4995 | <pre>-----   GA   AA -U   .-UU   AA GAACG
   UGG AGGGA   G   CAUAGGCAGUGGC   GGUU GG   G
   GCC UCUCU   U   GUAUUCGUCACUG   CCGA CC   A
CGUGU   AG   AG -   \ --   GG ACCCA</pre> |
| tmo-mir-5048 | <pre>CUU   UC UCUUGAC   -   C   UUUUUG- ACC   U
   UGCUUU \   AGACC AAAAUC GUAA   GC   AUG U
   ACGAAA U   UCUGG UUUUGG CGUU   CG   UGC U
---   CU -------   A   A   UAUAUAA AUC   U</pre> |
| tmo-mir-5049 | <pre>G-   AAA A   C   AACU
   CUUUGAU   GU UACUCCCUCC UCCCAAAAUAAGUGUCUCA   \
   GAGAUUA   CA AUGAGGGAGG AGGGUUUUAUUCACAGAGU   A
AG   A-- -   C   CGAA</pre> |
| tmo-mir-5049-3p | <pre>C----------   AGU   -   UC   .-G   CAAUU
        GACA   AAU AUGGA   GGAG   GAGUAU   U
        UUGU   UUA UACUU   CCUU   CUCGUA   U
UGUCCAUGUAU   CUU C   UA   \ -   ACAUA</pre> |
| tmo-mir-5050 | <pre>CGCCGUU   GC- GA AGA   C   -   A
   UUGCUGGUUGAACGACCUCAUCAUG   AC GC   UCU CCU GGC C
   AACGGCCAACUUGCUGGAGUGGUAC   UG CG   AGA GGA CUG U
-------   AGC G- G--   A   G   U</pre> |
| tmo-mir-5054 | <pre>UU-   CU A   AGCCAAAAACA -CUAACC   -   UUC
   UGGA AU UGGA   \   ACG UGGCCGUGGG   \
   ACUU UA ACCU   A   UGC GCUGGCACCC   G
UUC   UU A   AAUUUUACAGU ------   G   CUA</pre> |
| tmo-mir-5064 | <pre>AAAUCC U   UU   U U   AUUCAAU- UG U
      UC GGUUGAA UGUCCAUAGCAUCA CCA CCUACC   GG GC G
      AG CCAACUU AUAGGUAUCGUGGU GGU GGAUGG   CC CG A
GU---- C   CG   C   -   GUGGCAAC GU U</pre> |

S2 Table. (Continued).

| tmo-mir-5067 | ```
          UC    UC             CG   C   C              C    A
GUACUCCUU  GUU  AUAUUAGUUGU  CU AAA GGAUGUAUUUAG ACUU A
UAUGAGGGA  CAA  UAUAAUCAACA  GA UUU CCUACAUAGAUC UGAA A
          GU    GA             AA  C  A            -    U
``` |
| tmo-mir-5073 | ```
UUGGU-    UGAAUCGGAAA        AU
     UUGG              CAAUUUU  \
     AACC              GUUAAAA  A
UUCUUU    UUAAAAAUGGA        AG
``` |
| tmo-mir-5076 | ```
A  .-AAUCAA   AC AUC-    UUCC   C        CGUG
 GC         UAGG  \    CUCCC   UUU AGGAAGAA    A
 CG         AUCC  U    GAGGG   AAA UCCUUUUU    A
- \ ------   UU AGAC      U--- U         CUUA
``` |
| tmo-mir-5079 | ```
UUCG- A    .-U A  UAG           GA   GUAC       CACUAU
    CA UUUAU  UAC UAG  AACAAAUCCAAAUUG AGG    GAAUUCUG      \
    GU AAAUA  AUG GUU  UUGUUUAGGUUUAAU UUC   UUUAGGAU       A
CUAAA  A    \ -  -  UUA            A-  ----        AGCUUCG
``` |
| tmo-mir-5084 | ```
AAA-    AC UCAUA  U  .-AUAGAU    GU        A  CU  AU
   AAAAC \      GA UC        GGUU GAUCCUCUGC GUA  GU  \
   UUUUG U      CU AG        CCAA CUAGGAGAUG CAU  CG  G
CCAC    GU UG--- U  \ ------    UC         A  CG  UA
``` |
| tmo-mir-5141 | ```
AAUUAGA   CUAU   CG   -   UAUCCG     CGCGUCG GU    U     AAAGC
      CCAC    UAA UUG GCU    UCAGU     G CUGCCGC CACUUG     U
      GGUG    GUU AAC CGG    GGUCA     C GAUGGCG GUGAAC     A
-------   AAGC   AU   A   UUAA--     AAA---- UU    U     AAAGU
``` |
| tmo-mir-5174-5p | ```
UUUCAU---- -   CU UA --        C        GAA     A   G
       GC UUAAC AA AC UACUCCUCUGU CCAUAAUAUAA  CGUUUUUG CAUUA \
       CG AGUUG UU UG AUGAGGGAGACA GGUAUUAUAU  GCAAAAAC GUGAU U
UAUUUUAUAC U   U- UC CU        A        GUC     G   G
``` |
| tmo-mir-5181 | ```
·AACCAAUAAAUU      A  C    A      UG    U     AG
       GUACUCC UC GAUCCA AAUAAGUGUCG  GUUUUG ACUAAGGUU   \
       CAUGAGG AG CUAGGU UUAUUCACAGC  CAAAAC UGAUUCUAA   U
-----------      G  A    A      GU    U     CU
``` |
| tmo-mir-5368 | ```
AAGA----- U GA   A--    G UC A CG GG   - GG A   CA
      GC CG UUCU   ACCUU UG AG CC CG CCA AG AC GUCU \
      CG GC GAGG   UGGAA AC UC GG GC GGU UC UG CAGA G
CUUUGGAAA U G-   CAA    A CC C AU GG   A UU A   UG
``` |
| tmo-mir-5387 | ```
A                        U   AG
 GCCCCC--CCUUUAGUACCG UUC   \
 CGGGGG  GGAAAUCGUGGC AAG   C
A       \            C   CA
``` |
| tmo-mir-5523 | ```
AUGCAGA    UAU   U    AUA   UCC      AUGUC     UAU
      UGA  AAC AGUAA   UGU  UCCCC      AACA    U
      ACU  UUG UCAUU   ACA  AGGGG      UUGU    U
GAUAUAA    U--   U    C--   UUA       GAU--    UAU
``` |
| tmo-mir-6182 | ```
CAA    CU ACA-------   UGU---    GG  -      A  U
  UGAC  \          CGAG    GUGAU  A UGGCUUUG GCG \
  AUUG  A          GCUC    CACUA  U ACUGGAAC CGC U
A--    AC AUCAAACAACG   UUCUCC    GU C       C  U
``` |
| tmo-mir-6191 | ```
UCCA  C          G       UGUCUAGAUA            A-    UA
  UC GUCCCAAAUUACUU UCUUAGAUU         UGGAUGUAUCU  ACAC  A
  AG CAGGGUUUAAUGAA AGAAUCUAA         ACCUACAUAGA  UGUG  A
UUAG  A          G       ----------            CC    CA
``` |
| tmo-mir-6198 | ```
GAGA    UUU  - UU      GUAUUGAC     G  U AUG       U
    ACC   UC A CUCACCGA    GGCUCU UCU GG  GUCAUUC G
    UGG   AG U GGGUGGUU    CCGAGG AGA CC  UAGUGAG U
AAAG    UCU  U GG      ACUUCAGU    -  - AA-       U
``` |

S2 Table. (Continued).

| tmo-mir-6203 | ```
AUCC    .-CAA     GAACAAUAAAAUUAA  UU    UU
   AAGG     AGGCGA              AAU  CUUC  U
   UUCC     UCUGCU              UUA  GAGG  U
----    \ ---       AAUUCUUCUGGACG-   GG    CU
``` |
| tmo-mir-6219-5p | ```
U--- G           CGUGUAA            GAGA   U
   GC CCCUUUAGUCCCGGUU   GAACCGGGACUAAAGGG   GGUA U
   CG GGGAAAUCAGGGCCAA   CUUGGCCCUGAUUUCCC   CCAU A
AUUC  -            AAG----             AG--   G
``` |
| tmo-mir-6250 | ```
   CCUG- ACU    U-----     GAUUCU  U  -     UC
GAAAU  CUU  UGCGUU    UCCUCUCUC   UGA AAG UUGGCGGC  C
CUUUG  GAA  GCGCAG    AGGGGAGGG   GCU UUC AACCGCCG  C
   CUAAA  CU-  CUAGAU      ------  C  U    UA
``` |
| tmo-mir-6478 | ```
A  --      CGA    AUCAAAA  G-    GGG
  GC  GCUGGGCUA  GUUGGC      GA  UUCU  \
  UG  UGACUCGAU  CAGCCG      CU  AAGG   A
-  GU        UC-    AACA--- AA    AGG
``` |
| tmo-mir-6874-3p | ```
A----    UU    ACU     CAUUUACCU----    UGC    GAC
   UAUA  UUU   GAGGGAAC            AGUUC   UGUUUU   U
   GUGU  AAA   UUCCCUUG            UUAAG   ACAAAG   U
UUAAC    UU    CUU      UUUAUUUUACUUU    UU-    ACC
``` |
| tmo-mir-8155 | ```
AAAA    --  C  A       .-G GG    GU
   CCUG  GCU UG UACCACU   UU  GGAAU  A
   GGGC  CGA AC GUGGUGG   AA  CUUUA  G
GGGA    AA  U  -      \ -  AA     AU
``` |
| tmo-mir-B6-3p | ```
AC    G  CG        AC  CUCAACAA            --    UUC
  GGC GC  GCU--GGUG  \        CGUCUCC--GGCG  CCGGG   \
  CCG UG  CGA  CCAC  G        GUAGAGG  CCGC  GGCCC   C
--   -  CU   \     UC  A-------      \    AG    UGA
``` |
| tmo-mir-l5-3p | ```
·A   G  CGA GGACG    G  --   - AA-    CCU
  CCAA GU  A    GCGUUGU GU UCUUC UC  GUUG   \
  GGUU CA  G    CGUAGCA CA AGAAG AG  CAGC   G
-   G  AAA AGA--    G  GA    U GUG   UCG
``` |