Funct Integr Genomics (2017) 17:171–187
DOI 10.1007/s10142-016-0487-y

ORIGINAL ARTICLE

Wheat miRNA ancestors: evident by transcriptome analysis of A, B, and D genome donors

Burcu Alptekin[1] · **Hikmet Budak**[1,2]

Received: 2 February 2016 / Revised: 6 March 2016 / Accepted: 14 March 2016 / Published online: 31 March 2016

Abstract MicroRNAs are critical players of post-transcriptional gene regulation with profound effects on the fundamental processes of cellular life. Their identification and characterization, together with their targets, hold great significance in exploring and exploiting their roles on a functional context, providing valuable clues into the regulation of important biological processes, such as stress tolerance or environmental adaptation. Wheat is a hardy crop, extensively harvested in temperate regions, and is a major component of the human diet. With the advent of the next generation sequencing technologies considerably decreasing sequencing costs per base-pair, genomic, and transcriptomic data from several wheat species, including the progenitors and wild relatives have become available. In this study, we performed in silico identification and comparative analysis of microRNA repertoires of bread wheat (*Triticum aestivum* L.) and its diploid progenitors and relatives, *Aegilops sharonensis*, *Aegilops speltoides*, *Aegilops tauschii*, *Triticum monococcum*, and *Triticum urartu* through the utilization of publicly available transcriptomic data. Over 200 miRNA families were identified, majority of which have not previously been reported. Ancestral relationships expanded our understanding of wheat miRNA evolution, while *T. monococcum* miRNAs delivered important clues on the effects of domestication on miRNA expression. Comparative analyses on wild *Ae. sharonensis* accessions highlighted candidate miRNAs that can be linked to stress tolerance. The miRNA repertoires of bread wheat and its diploid progenitors and relatives provide important insight into the diversification and distribution of miRNA genes, which should contribute to the elucidation of miRNA evolution of Poaceae family. A thorough understanding of the convergent and divergent expression profiles of miRNAs in different genetic backgrounds can provide unique opportunities to modulation of gene regulation for better crop performance.

Keywords Bread wheat · Poaceae · miRNA · TE-miR · RNA-Seq

This article forms part of a special issue of *Functional & Integrative Genomics* entitled "miRNA in model and complex organisms" (Issue Editors: Hikmet Budak and Baohong Zhang

Electronic supplementary material The online version of this article (doi:10.1007/s10142-016-0487-y) contains supplementary material, which is available to authorized users.

✉ Hikmet Budak
budak@sabanciuniv.edu

Burcu Alptekin
alptekinburcu@sabanciuniv.edu

[1] Molecular Biology, Genetics and Bioengineering Program, Sabanci University, 34956 Istanbul, Turkey

[2] Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717, USA

Introduction

Harvested over 200 million hectares corresponding to 650 million tons of annual production, bread wheat, *Triticum aestivum* L., is the most extensively grown crop following rice and maize (Mochida and Shinozaki 2013; Ani Akpinar et al. 2015). It is an important component of world food consumption by providing around 20 % of human dietary energy (Kurtoglu et al. 2014). Considering its high agronomic value, along with the ever increasing world population, sustainable and steady increases in wheat yields are of vital importance for the food security of next generations (Parry et al. 2011). However, agricultural practices following wheat domestication have largely focused on the selection and breeding of

uniform plants with high yields, considerably narrowing the gene pools of the elite cultivars (Akpinar et al. 2015). Fortunately, extant wild relatives and progenitors of bread wheat still preserve remarkable genetic diversity. These wild gene pools contain diverse alleles that contribute to adaptive processes to different environmental conditions including biotic and abiotic stresses (Akpinar et al. 2012; Budak et al. 2013; Lopes et al. 2015). Exploring these sources of natural diversity as well as effective combination of them with powerful molecular methods (Budak 2010) can provide insights into the cellular mechanisms and associated processes underlying environmental adaptation which may, in turn, be utilized to develop hardy varieties that are high-yielding at the same time.

The large (17 Gb) and complex (80 % repetitive elements in content) nature of the allohexaploid wheat genome (*Triticum aestivum*, $2n = 6\times = 42$, genomic formula AABBDD) is thought to originate from at least two different hybridization events (Chalupska et al. 2008; Marcussen et al. 2014). As the first, the hybridization of two diploid species, *Triticum urartu* ($2n = 2\times = 14$, A^uA^u) and an unknown B genome related organism ($2n = 2\times = 14$, genomic formula SS) occurred about 0.5–0.36 million years ago (MYA) and resulted in the emergence of wild tetraploid emmer wheat *Triticum turgidum ssp. dicocoides* ($2n = 4\times = 28$, AABB) (Dvorák et al. 1993; Akhunova et al. 2010). Around 8000 years ago, another hybridization event between domesticated emmer wheat *Triticum turgidum* ($2n = 4\times = 28$, AABB) and *Aegilops tauschii* ($2n = 2\times = 14$, DD), followed by whole genome duplication gave rise to the hexaploid bread wheat (Jia et al. 2013; Li et al. 2014). Domestication and evolutionary processes, such genome-wide diploidization following polyploidization, reshaped the modern wheat genome; many genes or families are lost, while others are strongly preserved or newly emerged (Murat et al. 2014).

Understanding the structure and organization of the complex wheat genome can be aided by working on its diploid progenitors and relatives. Additionally, diverse gene pools of these diploid germplasms may provide favorable alleles for wheat improvement. With this motivation and the decreasing costs of sequencing technologies, genome and/or transcriptome sequences of several wheat species have become available, delivering important clues into the wheat genome structure and evolution. Sequencing the A genome progenitor of allohexaploid bread wheat, *T. urartu* ($2n = 14$, A^uA^u) revealed that its 4.94 Gb genome consists of 67 % repetitive elements (Ling et al. 2013). Another A genome relative of wheat is *Triticum monococcum* ($2n = 14$, A^mA^m) which has a similar genomic structure, in terms of genome size (5.6 Gb) and content with *T. urartu* (Fricano et al. 2014). While *T. urartu* existed in wild populations until now, *T. monococcum* has undergone domestication and diverged into different subspecies such as *T. monococcum ssp. aegilopoides* (wild winter wheat) and *T. monococcum ssp. monococcum* (domesticated spring wheat) (Fox et al. 2014). Since the genome sizes of the both of these organisms are relatively small with respect to the 17 Gb hexaploid wheat genome, they can be used as simpler surrogates, for which the findings can be translated to the hexaploid wheat through comparative genomics (Brenchley et al. 2012). Furthermore, recent studies revealed that A genome relatives of wheat have a wide repertoire of important alleles for biotic and abiotic stress responses such as disease resistance or salt tolerance (Munns et al. 2012; Saintenac et al. 2013; Zaharieva and Monneveux 2014), which can be utilized in wheat genome improvement programs.

While the A genome progenitor of tetraploid and hexaploid wheat species is fairly known and generally accepted, the identity of the B genome progenitor is much more controversial and remains uncertain. The main challenge is the variation and divergence among the candidate B genome progenitors of wheat (Dvorák and McGuire 1981). In spite of the difficulties, several attempts has been made for the identification of the B genome origin through comparative analyses of nuclear and mitochondrial DNA sequences (Dvořák et al. 1989) and chromosome rearrangements (Jiang and Gill 1994; Devos et al. 1995). Candidate B genome progenitors has been gathered under a section named "Sitopsis" and one of the group members has shown strong evidence to be the progenitor of B genome: *Aegilops speltoides* ($2n = 14$, SS) (Kilian et al. 2007). *Ae. speltoides* is a diploid plant native to Western Asia and Southeastern Europe, presenting a natural source for distinct disease resistance alleles (Sarkar and Stebbins 1956; Haider 2012). Another Sitopsis section member, *Aegilops sharonensis* ($2n = 14$, SS) exhibit remarkable adaptation to various environmental conditions (Bouyioukos et al. 2013). Certain cultivars of *Ae. sharonensis* are resistant to wheat stem rust disease, considered as evidence of wide range genetic variation (Olivera et al. 2007). With a relatively small genome size (7.5 Gb) and high genetic diversity, *Ae. sharonensis* has morphological, phenological, and ecological characteristics that are remarkably different from bread wheat (Eilam et al. 2007).

Despite the controversy on the identity of the B genome progenitor, the D genome donor of bread wheat is clearly defined as *Ae. tauschii* ($2n = 14$, DD). It is an extensively distributed goat grass in Eurasia displaying high range of diversity at phenotypical and molecular levels (Dudnikov and Kawahara 2006). *Ae. tauschii* is a valuable source for development of improved wheat cultivars tolerant to different stress conditions due to its exclusive adaptation ability to many different environmental conditions (Iehisa et al. 2012; Akpinar et al. 2014). The 4.23 Gb sized *Ae. tauschii* genome, containing 65 % repetitive elements, is fairly well characterized (Kumar et al. 2012; Jia et al. 2013; Luo et al. 2013). Several efforts have also focused on the identification of genomic differences, such as Single Nucleotide Polymorphisms (SNPs) between

the D genome of bread wheat and *Ae. tauschii* which may deliver important clues into the evolutionary history of these species, facilitating genetic transfer between two species (Iehisa et al. 2012; Wang et al. 2013).

MicroRNAs (miRNAs) are small (19–24 nt), endogenous, non-coding RNA molecules that are involved in gene expression regulation at the post-transcriptional level (Budak et al. 2015a; Budak and Akpinar 2015). miRNAs have big impact on several processes such as development (Neilson et al. 2007; Curaba et al. 2012), environmental adaptation, stress responses (Trindade et al. 2010; Budak and Akpinar 2011; Budak et al. 2015b) and disease resistance (Navarro et al. 2006; Lucas et al. 2014); therefore, elucidation of the miRNA repertoire as well as an in-depth understanding of their functions and biogenesis hold importance in discovering the complex nature of gene regulation (Sunkar et al. 2008; Budak et al. 2014). In plants, miRNA research has been ongoing for over a decade in several directions. Many studies focused on miRNA-target interactions with the intent of revealing the direct effects on biological processes which may even have immediate impacts on phenotype (Xing et al. 2010). On the other hand, a comprehensive point of view is implemented in miRNA studies by genome/transcriptome-wide miRNA mining approaches (Zhou et al. 2008; Chen and Cao 2015; Liu et al. 2015), unlocking the miRNA contents of various plants (Kurtoglu et al. 2014). Up to now, many different methods have been developed and used for the identification of miRNAs and their specific target genes such as cloning (Wang et al. 2004), genetic screens (Chen 2004), splinted-ligations mediated miRNA detection (Chamnongpol et al. 2010), microarray profiling (Kantar et al. 2011) and computational approaches (Wang et al. 2004; Sunkar et al. 2008; Budak and Akpinar 2011). While most experimental methods are expensive and time consuming, computational miRNA identification approaches from high-throughput genomic or transcriptomic data present as a decent and rapid alternative. Whereas genome-wide in silico miRNA identification can unravel the whole miRNA repertoire of an organism, expression evidence for the putative miRNAs is ultimately required. On the other hand, the utilization of transcriptomic data provides an overview of miRNAs putatively expressed under certain conditions (Akpinar et al. 2015). Even though miRNA prediction tools based on support vector machine algorithms also exist (Kadri et al. 2009; Teune and Steger 2010), homology-based miRNA identification methods, which are based on large scale sequence conservation, are considered as the most powerful and reliable techniques for the mass identification of plant miRNAs (Budak and Akpinar 2015; Budak and Kantar 2015).

In addition to their key roles on gene expression regulation, miRNAs might also have important roles on the stability and organization of polyploid plant genomes. Understanding how the cell regulates multiple copies of the genome may hold importance for interfering genetic regulation (Chen 2007) and many miRNA studies on many economically important polyploid crops such as cotton and sugarcane have focused on this issue (Thiebaut et al. 2012; Xie and Zhang 2015). However, studies which investigate miRNA genes comparatively in diploid and polyploid backgrounds of wild and domesticated varieties are rare to our knowledge (Li et al. 2014). Bread wheat stands as an enticing model for such studies of comparative microRNAome (or miRNAome) analysis, with the presence of multiple genome progenitors. Herein, we computationally investigated the miRNA repertoires of five different diploid wheat progenitors and relatives, *Ae. sharonensis*, *Ae. speltoides*, *Ae. tauschii*, *T. monococcum*, and *T. urartu*, using a homology-based approach on publically available transcriptomic data and compared the related miRNA content with the *T. aestivum* transcriptome. Identification of putatively conserved miRNAs in inter/intra-species level across different members of Poaceae family provides invaluable insight into the evolution of microRNAome with respect to polyploidization and domestication events.

Materials and methods

Construction and pre-processing of transcriptome dataset

In order to identify the miRNA repertoires of wheat diploid relatives, transcriptomic sequences of sixteen *Aegilops sharonensis* accessions (Supplementary File S1), *Aegilops speltoides* accession B-2140016, *Aegilops tauschii*, accession D-2220009, *Triticum monococcum* ssp. *aegilopoides* accession G3116, *Triticum monococcum* ssp. *monococcum* accession DV92 and *Triticum urartu* accession G1812 were obtained from National Center of Biotechnology Information (NCBI) sequence read archive (http://www.ncbi.nlm.nih.gov/sra). Additionally, *Triticum aestivum* cultivar Chinese Spring "leaf" and "root" RNA-sequencing (RNA-Seq) data, obtained from Unité de Recherche Génomique Info (URGI, http://wheat-urgi.versailles.inra.fr/Seq-Repository/RNA-Seq), were included for comparative analysis of miRNAs in diploid and polyploid backgrounds. All transcriptome data is derived from control plants grown under normal conditions, except for *T. monococcum* subspecies which were germinated in the dark and then exposed to 48 h of light (Fox et al. 2014). *T. monococcum* data was concluded to be fairly similar to the remaining datasets and nevertheless included in this study as it enabled us to assess the effects of domestication. In the case of *T. urartu* and two *T. monococcum* subspecies, pre-assembled transcriptomic data is used while for others *de novo* sequence assemblies from raw reads of each library were constructed by Trinity software (Grabherr et al. 2011). Quality trimming and adaptor removal of reads were performed with Trimmomatic (v 0.32) with default parameters "LEADING:5, TRAILING:5,

MINLEN:36" (Bolger et al. 2014). A summary description of the input datasets used in this study is given in Table 1.

All transcriptome assemblies were further evaluated for sequence quality and context. Since the transcriptional machinery of organelles can generate long transcripts which can affect miRNA identification processes, organelle DNA-associated sequences from all transcriptome assemblies were eliminated (Bouyioukos et al. 2013). Chloroplast genome sequences of each organism were obtained from NCBI archive and used as a query for elimination process (Middleton et al. 2014). For the removal of mitochondrial DNA, bread wheat mitochondrial genome sequence (NC_007579.1) was used as the closest data available. Reads were aligned with organellar sequences by BLASTn (version 2. 2. 26; 1E-15, −dust "no") and positive hits (alignment ≥ 95 % of the query and identity ≥ 95 %) were removed. In addition, since transfer ribonucleic acid (tRNA) and ribosomal ribonucleic acid (rRNA) sequences can form hairpin or hairpin-like secondary structures, sequences with positive hits to tRNA or rRNA sequences were also discarded to avoid potential false positives. For each organism, published tRNA and rRNA sequences were obtained from European Nucleotide Archive (http://www.ebi.ac.uk/ena). Assembly sequences aligning with tRNAs/rRNAs were identified (1E-5, −dust "no") and positive hits fulfilling the same criteria defined above were further eliminated.

Reference miRNA set

For the homology-based prediction of putative miRNA repertoires from the transcriptome assemblies of bread wheat and its diploid progenitors and relatives, a specific query which consists of previously identified plant mature miRNA sequences was used. The list of published 4000 unique mature miRNA sequences from 72 different Viridiplantae species were downloaded from the current release of miRBase (v21 June 2014) (Kozomara and Griffiths-Jones 2011) and combined with *T. urartu* mature miRNA sequences identified from (Ling et al. 2013). Unique mature miRNAs identified from *T. urartu* small RNA databases (Ling et al. 2013) were specifically marked as "miR-Sr". All mature miRNA sequences were filtered against redundancy and the remaining 4942 unique mature miRNA sequences were used as query in the homology-based miRNA identification process.

Homology-based in silico miRNA prediction

Homology-based in silico miRNA prediction from all transcriptome assemblies was carried out with a previously described two-step procedure (Lucas and Budak 2012). Briefly, the two-step procedure can be described as follows: (1) preliminary selection of transcriptomic sequences exhibiting sequence homology to a reference miRNA sequence, (2) their subsequent elimination based on the consistency of their stem-loop secondary structures in relation to the general, pre-established preliminary miRNA (pre-miRNA) features (Zhang et al. 2005; Unver and Budak 2009; Kantar et al. 2010). Prediction was employed by the utilization of two in-house Perl scripts "SUmirFind and SUmirFold", previously described in detail (Lucas and Budak 2012; Kurtoglu et al. 2013; Kurtoglu et al. 2014). SUmirFind script utilizes BLASTn alignment algorithm from BLAST+ package and detects candidate miRNA sequences which are aligned to mature miRNAs found in miRNA reference set with less than three mismatches. SUmirFold script further processes these putative sequences in order to generate secondary structures with the utilization of UNAFold version 3.8 (Markham and Zuker 2008) and performs a basic filtering step based on following criteria: (1) In the hairpin structure, mature miRNA and miRNA* sequences can have four and six mismatches, respectively, at most, (2) miRNA* cannot be broken into separate segments or contain a large loop, (3) The GC content of the hairpin should be in the range 24–71 %, (4) The minimum free energy index (MFEI) of the hairpin should be above 0.67. Putative sequences passing the preliminary filters were further evaluated based on pre-miRNA characteristics previously described in (Kurtoglu et al. 2014) with a Java-based semi-automated in-house program. Predicted miRNAs supported by the presence of at least one pre-miRNA sequence possessing the pre-defined stem-loop structure characteristics were

Table 1 Summary list of the transcriptomic datasets used for miRNA identification

Transcriptome assemblies	Tissue	Assembly type	Related project ID
Ae. sharonensis (16 different accessions)	Leaf	Trinity	NCBI archive: PRJEB5340
Ae. speltoides accession B-2140016	Young roots and shoots	Trinity	NCBI archive: PRJNA214743
Ae. tauschii accession D-2220009	Young roots and shoots	Trinity	NCBI archive: PRJNA214744
Triticum aestivum cultivar Chinese Spring	Leaf and root	Trinity	URGI sequence archive: https://urgi.versailles.inra.fr/files/RNASeqWheat/
T. monococcum ssp. *monococcum* accession DV92	Seedling shoots	Pre-assembled	NCBI archive: PRJNA203221
T. monococcum ssp. *aegilopoides* accession G3116	Seedling shoots	Pre-assembled	NCBI archive: PRJNA195398
T. urartu accession G1812	Young roots and shoots	Pre-assembled	NCBI archive: PRJNA191053

recorded and accepted as a proof of putative expression of the related miRNA.

Identification of repetitive elements in putative miRNA-coding sequences

In order to investigate transposable element containing miRNAs, putative pre-miRNA sequences identified from all datasets were separately searched against a publically available repeat library of Poaceae family (MIPS-REdatPoaceae v9.3p, ftp://ftpmips.helmholtz-muenchen.de/plants/REdat/) which contains 34,135 different repeat sequences (Nussbaumer et al. 2013) using RepeatMasker version 4.0.5 (www.repeatmasker.org) at default settings. Sequences covered by repeats more than 50 % of their lengths were recorded as transposable element-related miRNAs, or TE-miRs.

miRNA-target prediction and functional annotation

In silico target prediction of putative miRNA sequences was separately carried out for each miRNA dataset identified from six different organisms using psRNATarget web-tool (http://plantgrn.noble.org/psRNATarget/) at default parameters (Dai and Zhao 2011) against their respective transcriptomes. Sequences marked as putative miRNA targets were retrieved and their functional annotation was performed using Blast2GO (www.blast2go.com) (Conesa and Götz 2008). The initial blast step was performed locally against all non-redundant Viridiplantae (taxid: 33090) proteins (3,485,798) at an *e* value cutoff 10^{-6}, and the subsequent mapping and annotation steps were carried out at default parameters. Gene ontology (GO) terms were further analyzed, recorded, and visualized with multilevel pie graphs.

Results and discussion

miRNA identification from transcriptome assemblies

To explore putatively expressed miRNA repertoires of bread wheat and its diploid progenitors and relatives, transcriptome assemblies were constructed with the Trinity software after quality trimming and adaptor removal of reads by Trimmomatic (v 0.32) at default settings, where pre-assembled sequences were not available. RNA-Seq of 22 different genotypes belonging to six different organisms, *Aegilops sharonensis* (16 different accessions, Supplementary File S1), *Aegilops speltoides* accession B-2140016, *Aegilops tauschii*, accession D-2220009, *Triticum monococcum* ssp. *aegilopoides* accession G3116, *Triticum monococcum* ssp. *monococcum* accession DV9, *Triticum urartu* accession G1812, and *Triticum aestivum* cv. Chinese Spring, yielded assembled sequences ranging from 44 to 223 Mb with an average contig length of 546 to 1847 bases (Table 2). Following further elimination of organellar transcripts and other non-coding RNAs such as tRNAs and rRNAs, homology-based in silico miRNA prediction was performed on each assembly using previously known 4942 unique mature miRNA sequences of the reference miRNA set. Putative miRNAs from each assembly were filtered against redundant miRNAs. Identified miRNAs were named and clustered respect to their homology to known plant miRNAs in mature miRNA level. In case of the equal homology to diverse mature miRNA sequences coming from different families, pre-miRNA homology is considered. Cumulatively, a total of 17,197 unique stem-loops (or pre-miRNAs) corresponding to 259 different miRNA families, of which nine belonging to *T. urartu* had also been previously reported by Ling and his colleagues (Ling et al. 2013), were detected (Supplementary File S2). Since no registered miRNA sequences was detected in miRBase from *Ae. sharonensis*, *Ae. speltoides*, *T. monococcum*, and *T. urartu*, a great majority of identified miRNAs in this study had not been previously reported. It is also important to remark that orthologues of three different *T. urartu* miRNAs (miR-Sr60871, miR-Sr93419, and miR-Sr7354), which are not included in miRBase, were exclusively identified from *T. aestivum* transcriptomes where a similar situation stands for *Ae. tauschii* for four different miRNAs (miR-Sr60871, miR-Sr80912, miR-Sr93419, miR-Sr97354).

In our analysis, the number of identified miRNA families among 22 different transcriptome assemblies of wheat

Table 2 Assembly metrics for the genotypes used in this study. For *Ae. sharonensis*, the average assembly metrics of 16 different accessions is given for simplicity

Transcriptome assemblies	No. of assembled transcripts	Average contig length (b)	Total length (Mb)
Ae. sharonensis (16 different accessions)	177,920	800	143
Ae. speltoides accession B-2140016	80,659	546.38	44
Ae. tauschii accession D-2220009	142,907	804.81	115
Triticum aestivum cv. Chinese Spring	257,230	919.4	168
T. monococcum ssp. *monococcum* accession DV92	120,911	1847	223
T. monococcum ssp. *aegilopoides* accession G3116	117,969	1493	210
T. urartu accession G1812	86,247	1417	122

progenitors and relatives varied in the range of 30 to 89, while this range was 131 to 3046 for the stem-loops, as given in Fig. 1a. Of all datasets, *T. urartu* and *Ae. sharonensis* accession 2172 showed the highest miRNA diversity with the representation of 89 unique miRNA families (1.25 fold of average) (Fig. 1a). Interestingly, *Ae. speltoides* displayed the lowest miRNA diversity and total miRNA stem-loop count which might be related to the relatively small size of transcriptome assembly (Table 2). Despite a slight correlation between the total length of transcriptome assembly and the number of identified miRNA families ($r^2 = 0.64$, type, "Pearson"), associated miRNA stem-loop counts were hardly dependent on assembly size ($r^2 = 0.37$, type, "Pearson"), while it display a positive correlation with miRNA variety of the organism ($r^2 = 0.59$, type, "Pearson"), (Supplementary Figure S1). Even though calculated correlations might be effected from the randomness of dataset, this situation might also arise from differential expression and genomic expansion of some miRNA precursors in different species. For instance; miR1130 family members were represented with 136 putative precursor sequences in *Ae. sharonensis* accession 409 while this number was 57 for *Ae. sharonensis* accession 1192. Since expression of miRNAs in mature and pre-miRNA level highly specific to condition/tissue/time-zone, represented precursor numbers, and related correlations might be altered in different situations.

The miRNA families, miR1130, miR1120, miR1127, miR5049, or miR5181 had the highest numbers of corresponding stem-loops across all organisms. Among these, miR1120 family has already been defined as one of the largest miRNA families in *T. aestivum* (Budak and Akpinar 2015). miR1120, miR1127, and miR1130 have also been reported as important regulators of seed and leaf development in *T. aestivum*, processes crucial to the survival and reproduction of plants (Han et al. 2014). In addition, these miRNA families were detected as conserved among all organisms including different accessions together with 15 other conserved miRNA families (Table 3). In the point of fact; it is a known phenomenon that many miRNA families and their members are evolutionary conserved among species within the same kingdom where they might exist as orthologues or homologues (Llave et al. 2002; Jones-Rhoades and Bartel 2004). The further examination of highly-represented/conserved miRNA families may provide insights about their exclusive functions; hence, it might be suggested that the expression of miRNA families regulating functionally important biological

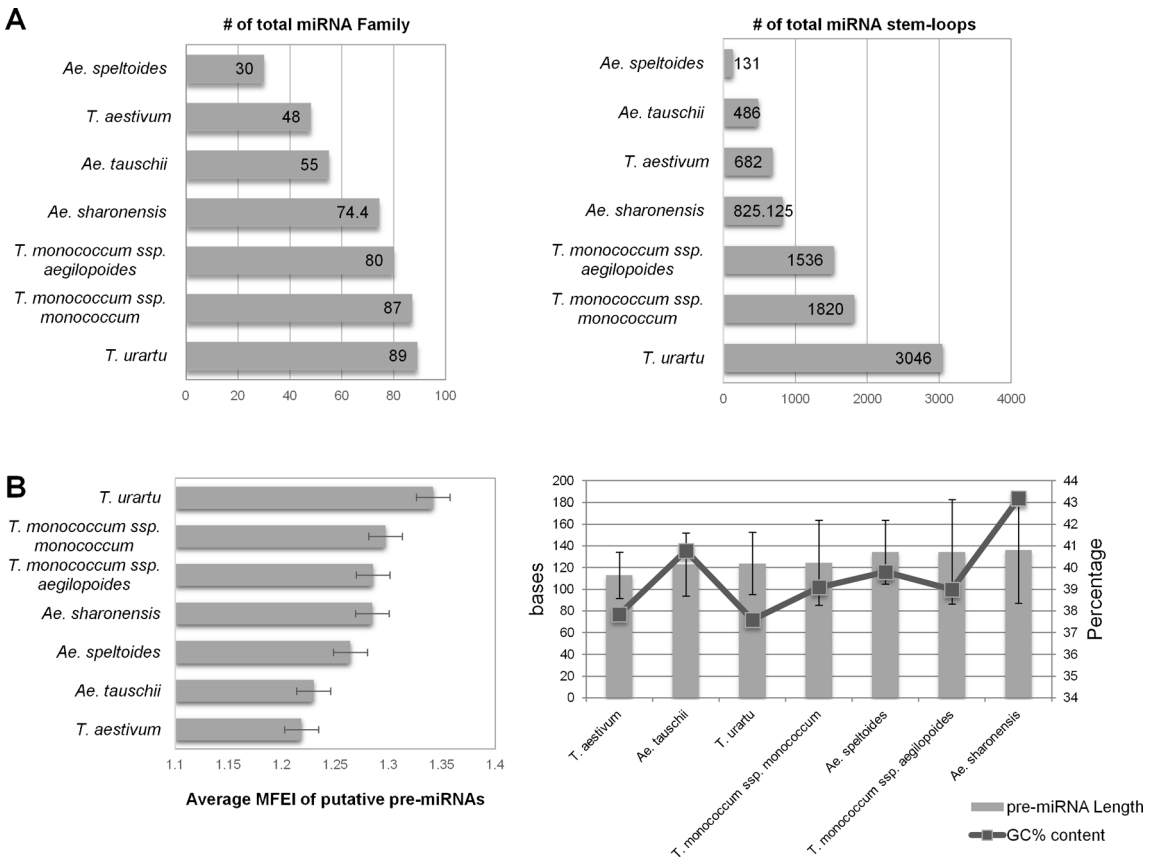

Fig. 1 Summary information regarding the miRNAs identified in this study. **a** The total number of identified miRNA families and corresponding miRNA stem-loops. For the *Ae. sharonensis*, the average of 16 different accession is counted. **b** Main characteristics of putative pre-miRNAs: MFEI, pre-miRNA length, and GC % content. The average values are calculated and represented for each organism

Table 3 List of miRNA families, putatively expressed under normal conditions, in all genotypes (1), in only *Ae. sharonensis* accessions (2), or only in one genotype (3) included in this study

(1) miRNAs families identified from all genotypes	
miR1120, miR1122, miR1127, miR1128, miR1130, miR1135, miR1136, miR1137, miR1436, miR1439, miR5049, miR5164, miR5174, miR5175, miR5180, miR5181, miR5205, miR5568, miR6197, miR819	
(2) miRNA families identified only from all *Ae sharonensis* accessions	
miR1139, miR437, miR5067, miR5070, miR5203, miR6248, miR818, miR-Sr60871	
(3) miRNA families identified exclusively in one genotype	
Ae. sharonensis accession 575	miR2655, miR5168, miR5749, miR6469, miR8040, miR9752
Ae. sharonensis accession 2189	miR1523, miR2592, miR2653, miR5166, miR8029, miR826, miR8661, miR9496
Ae. sharonensis accession 548	miR1510, miR419, miR5079, miR6209, miR7777, miR821, mir825, miR859
Ae. sharonensis accession 1995	miR2102, miR2930, miR6439, miR7693, miR862, miR9737
Ae. sharonensis accession 2205	miR5769, miR6196, miR822
Ae. sharonensis accession 2233	miR2595, miR3433, miR5809, miR779.1, miR858
Ae. sharonensis accession 6793	miR2864.1, miR5337
Ae. sharonensis accession 546	miR5053, miR528, miR535
Ae. sharonensis accession 6856	miR1320, miR7725, miR8141, miR9662
Ae. sharonensis accession 2020	miR472, miR5172, miR6291, miR7505, miR8175, miR904, miR-Sr47011
Ae. sharonensis accession 409	miR1027, miR1448, miR3448, miR3628, miR473, miR5257, miR5304, miR6174, miR7699, miR9664
Ae. sharonensis accession 2172	miR1114, miR417, miR5740, miR6118, miR7495, miR7698, miR829, miR831, miR846, miR855, miR9778
Ae. sharonensis accession 1998	miR1318, miR1432, miR3712, miR391, miR6300, miR774
Ae. sharonensis accession 396	miR1166.1, miR2637, miR2926, miR3440, miR835, miR8588
Ae. sharonensis accession 1192	miR1064, miR2098, miR2913, miR5238, miR5288, miR5380, miR6179, miR6266, miR7538, miR853
Ae. tauschii	miR5527, miR5780, miR5783
T. urartu	miR1437, miR5065, miR5082, miR5160, miR5219, miR5522, miR5832, miR6441, miR7533, miR-Sr60270
T. monococcum ssp. *aegilopoides*	miR5384, miR6193
T. monococcum ssp. *monococcum*	miR9661, miR6233, miR-Sr100772

process for sustainably of plants may tend to be slightly higher and stable under normal conditions respect to other miRNA families taking place in more specific pathways such as stress tolerance.

The characteristics of putative miRNAs

The characteristics of the putative miRNA stem-loops such as the GC % content, pre-miRNA/mature miRNA lengths, and MFEI values were assessed as an indicator of the accuracy of our miRNA identification pipeline. The average mature miRNA length was observed as 21 nucleotides in all of the organisms, as expected, since many of plant-mature miRNAs are ranging from 19–24 nucleotides with a bias towards 21 bases in length (Thakur et al. 2011; Kurtoglu et al. 2014). Animal miRNAs show high conservation of their precursor sequences; however, plant pre-miRNAs are long and diverse, which may be related with the presence of additional regulatory elements in the miRNA genes (Zhang et al. 2006). In our analysis, the average length of putative pre-miRNAs ranged from 110 to 150 in accordance with the previous studies. Interestingly, the average GC % of the identified pre-miRNAs slightly correlated with the average pre-miRNA length across different organisms (Fig. 1b). It is suggested that the pre-miRNA sequences encoded from intronic regions have higher GC contents that may be related to their biogenesis (Zhu et al. 2011b). Hence, high GC contents may indicate that the respective miRNAs originate from intronic sequences. Additionally, intronic miRNA precursor were observed to be lengthy, up to 890 bases (Ramalingam et al. 2014), which may, again, be reminiscent of unique biogenesis routes (Zhang et al. 2006). It is tempting to speculate that higher GC contents may assist in the stability of longer pre-miRNA sequences during miRNA biogenesis, particularly from long intronic miRNA precursors.

Another important criterion for the determination of genuine pre-miRNA secondary structures is the Minimum Folding Energy Index (MFEI). The specific MFEI values are used for

the discrimination of miRNAs from other RNA species, i.e., miRNAs generally have higher MFEIs than other types of RNAs such as tRNAs (0.64), rRNAs (0.59), or mRNAs (0.62–0.66) (Schwab et al. 2005; Kantar et al. 2012). In our analysis, the average of MFEIs of putative pre-miRNAs was remarkably high in all organisms, in agreement with previous studies (Zhang et al. 2007; Jin et al. 2008; Kantar et al. 2012). The highest average MFEI of putative stem-loop structures was highest in *T. urartu* with an average of 1.34 ± 0.23 (Fig. 1b).

miRNAs from A genome-related relatives of bread wheat

The transcriptome sequences of the A genome progenitor of modern wheat, *T. urartu*, suggested the expression of 3046 putative miRNAs belonging to 89 families. Among these, miR1120, miR1127, miR1130, and miR5049 were the most highly expressed families each with more than 200 counts. While none of these families are well-known or well-characterized, the expression of all these four families has previously been shown in wheat (Yao et al. 2007; Han et al. 2014). miR1120 has been linked to developmental processes in both wheat and its close relative barley through expression profiles (Kruszka et al. 2013; Han et al. 2014). However, miR1120 coding region was found to have significant similarities to the TcMar-Stowaway family of DNA transposons in barley, raising the possibility that miR1120 sequence may actually correspond to another class of small non-coding RNAs (Kruszka et al. 2013). Interestingly, miR1127 exhibited seed-specific expression profile in bread wheat, while miR1127* accumulated in the flag leaves, suggesting that the miRNA*, in this case, may have regulatory roles in growth and development (Han et al. 2014). Additionally, miR5049 and its putative target, wpk4 protein kinase, an important signal transducer responsive to various stimuli, have been reported to be drought-inducible (Sano and Youssefian 1994; Ruuska et al. 2008). Given its high putative expression under normal conditions in the A genome progenitor *T. urartu*, further research on this miRNA may unravel intriguing physiological roles. These four miRNA families were also enriched in the transcriptomes of *T. monococcum*, a close relative of the A genome progenitor *T. urartu*.

The two einkorn wheat subspecies, *T. monococcum* ssp. *aegilopoides* and *T. monococcum* ssp. *monococcum*, putatively expressed 80 and 87 miRNA families, respectively, even though the total counts of identified miRNA stem-loops were much fewer than *T. urartu* under similar conditions (1536 and 1820, respectively, vs. 3046). Although 58 miRNA families were commonly identified from all three transcriptomes, putatively expressed miRNA families largely differed from one transcriptome to another. While 11 miRNA families were putatively expressed only in domesticated einkorn wheat, miR6220, miR9772, and miR-Sr80912 families were exclusive to the wild *T. monococcum* ssp. *aegilopoides* and *T. urartu* transcriptomes. miR1030, putatively expressed in domesticated *T. monococcum* ssp. *monococcum* but not in *T. monococcum* ssp. *aegilopoides*, or *T. urartu*, was downregulated in response to drought stress in rice (Zhou et al. 2010). Interestingly, another such miRNA, miR5021, has been indicated in chilling response and dormancy, terpenoid biosynthesis, and lipid metabolism (Barakat et al. 2012; Wang et al. 2012; Fan et al. 2015; Singh et al. 2015). Although these observations are mostly from dicot plants where this miRNA may have specialized roles, terpenoids have recently been indicated in stress tolerance in maize (Vaughan et al. 2015), suggesting that similar networks may also exist across the plant kingdom.

In our analysis, some miRNAs such as miR444, miR5171, miR9653, and miR9654 were detected in two subspecies of *T. monococcum* but not in the A genome progenitor of bread wheat, *T. urartu*. Among these, miR444 was declared as an important regulator of nitrate signaling pathway in rice by experimental characterization under P deprived conditions (Yan et al. 2014). Interestingly, some of these miRNAs were detected as conserved with B and D genome related relatives of bread wheat (certain *Ae. sharonensis* accessions in most cases) as well as itself, while miRNAs such as miR5171 were specific to *T. monococcum* species. Between the two einkorn subspecies, 9 and 18 miRNA families were also identified as unique to the transcriptomes of the wild and domesticated genotypes, respectively. Further characterization of these miRNAs differentially expressed in domesticated and wild A genome lineages will likely expand our knowledge on the molecular mechanisms related with evolution of A genome of bread wheat and consequences of domestication event on miRNA genes.

miRNAs in bread wheat and its progenitors

While the A genome progenitor of bread wheat putatively expressed 89 miRNA families, the D genome progenitor *Ae. tauschii* and the closest identified B genome progenitor, *Ae. speltoides*, revealed the expression of relatively fewer putative miRNA families with 55 and 30 families, respectively, under similar conditions. Through identified 259 different miRNA families, only 18 % corresponding to 48 different miRNA families was detected in *T. aestivum* transcriptomes under normal conditions. Regarding to our results, no *T. aestivum* specific miRNA was detected under normal conditions; however different tissues may exhibit diverse miRNA profiles and detection of new miRNA families might be possible under the same conditions from other tissues such as spike or stem. Since 20 miRNA families were found as conserved across all genotypes used in this study (Table 3), remaining 28 different miRNA families were retained to *T. aestivum* from different progenitors and relatives. Comparative analysis of common

miRNA families across whole genotypes used in this study revealed the unequal contribution of different organism to the miRNA repertoire of bread wheat (Supplementary Figure S2). Many miRNA families from *Ae. speltoides* and *Ae. tauschii* was remained in *T. aestivum* miRNAome while many of the *Ae. sharonensis* miRNAs (80 % of whole identified miRNA families) was identified as "lost". Twenty-five families were detected as 'common' to all three progenitors, of which all but one identified from *T. aestivum*, as well (Fig. 2). Interestingly, miR5050 putatively expressed in *T. urartu*, *Ae. speltoides*, and *Ae. tauschii*, was not detected in *T. aestivum* transcriptome, suggesting that the expression of this miRNA, under normal conditions, might have been lost in bread wheat. Following polyploidization, genome-wide diploizidation acts on polyploid plant genomes to reduce redundancy, which can lead to structural gene loss or functional pseudogenization (Pont et al. 2013; Murat et al. 2014). Diploidization, followed by years of selection for particular traits, might have led to the loss of this miRNA or its expression under normal conditions in the modern bread wheat genome. In a similar study, the expression of this miRNA could not be detected via Northern blotting, while expression of its precursor could be verified in *T. aestivum* Chinese Spring cultivar under normal conditions (Wei et al. 2009). This suggests that the loss of expression in miR5050 in bread wheat is possibly due to a loss at the functional level. Another such miRNA, miR5021, implicated in chilling response, dormancy, terpenoid biosynthesis, and lipid metabolism as mentioned above, is putatively expressed both in *Ae. tauschii* and *Ae. speltoides*, but not in *T. urartu* or *T. aestivum*. Putative expression of this miRNA in the domesticated einkorn *T. monococcum* ssp. *monococcum*, a close relative of *T. urartu*, is curious. These observations indicate that differential expression patterns of miRNAs in related lineages provide valuable insight into the elucidation of miRNA functions, at times in multiple directions, in different genetic backgrounds.

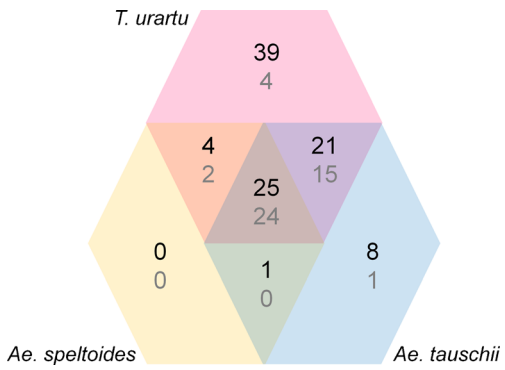

Fig. 2 The number of miRNA families putatively expressed under normal conditions in bread wheat progenitors, *T. urartu*, *Ae. speltoides*, and *Ae. tauschii* (numbers in *black*). *Numbers in gray* indicate the number of families also expressed in *T. aestivum* under similar conditions

In our analyses, miR5062, miR5387, miR6248, and miR9668 family members were identified only from *T. urartu* and *T. aestivum* among the progenitor of bread wheat (*T. urartu*, *Ae. tauschii*, and *Ae. speltoides*). Similarly, miR2120 family was present only in the D genome progenitor *Ae. tauschii* and *T. aestivum* transcriptomes. These miRNAs may have key roles in normal development that they are retained through wheat genome evolution; however, little information is currently available regarding their physiological roles. It should be noted that these miRNAs can still be present in the genomes of the remaining progenitors and expressed under different conditions, preventing their identification from the transcriptome sequences used in this study. In fact, miR5062 has been previously reported from *Ae. tauschii* (Jia et al. 2013). However, the expressional differences in different lineages can point out to potential pathways that can be targeted for crop improvement. For instance, miR2120 was upregulated in rice roots in response to cadmium stress (Huang et al. 2009). The expression of *Ae. tauschii* miR2120 under normal conditions may be linked to environmental adaptation. An in-depth understanding of its role may allow researchers to modulate its expression in *T. aestivum* in order to develop improved varieties. Two miRNAs, miR6180, and miR9653, were putatively expressed only in *T. aestivum* but not in any of the progenitors under normal conditions, suggesting that these miRNA families may be related to important agricultural traits. Aforesaid miRNAs have been predicted in wheat and barley in previous studies; however, their expression has not been experimentally verified (Wei et al. 2009; Lv et al. 2012; Han et al. 2014). Interestingly, small RNA sequencing suggested a seed-specific expression pattern for miR9653, which may indicate a role in seed development (Han et al. 2014). As our knowledge on these miRNAs and the pathways they participate in expands, new candidates may arise to modulate gene expression to achieve better crop performance.

miRNAs putatively expressed in *Ae. sharonensis* genotypes

The miRNA repertoires of 16 different *Ae. sharonensis* accessions collected from different locations in Israel were identified and comparatively assessed. Overall, 9509 unique miRNA stem-loops associated with 237 miRNA families identified from all *Ae. sharonensis* accessions, corresponding to 55.3 % of all identified stem-loops (Supplementary File S2). Accession 396 exhibited the highest number of predicted miRNA stem-loops (973 unique miRNA stem-loops), while accession 2205 had the lowest (602 unique miRNA stem-loops). Additionally, putatively expressed miRNA family variety was highest in accession 2179 with 89 different miRNA

families, corresponding to 847 stem-loops. A total of 28 miRNA families were commonly identified from all *Ae. sharonensis* accessions. Notably, miR1139, miR437, miR5067, miR5070, miR5203, miR6248, miR818, and miR-Sr60871 were putatively expressed in all *Ae. sharonensis* accessions, although they were not necessarily conserved at the interspecies level across all organisms (Table 3). Among these, miR818 targeted serine-threonine kinases in rice and contributed to the regulation of tissue differentiation (Luo et al. 2006). Similarly, the target transcripts of miR6248 were identified as DNA binding proteins and proteins kinases which have widespread key molecular functions within the cell (Liu 2012). Thus, the conservation of these miRNAs might be related to their regulatory roles on important biological processes which are replaced by other essential miRNA pathways in certain genotypes. Interestingly, precursor sequences of miR437 exhibited high sequence similarity to DNA transposons of MITE and TcMar-Stowaway families, which was also observed in maize (Zhang et al. 2009) and sugarcane (Zanca et al. 2010), suggesting that this miRNA may exist as a TE-miR in plants. Curiously, miR437 was not detected in *Ae. speltoides*, *Ae. tauschii*, or *T. aestivum* transcriptomes in this study. The fingerprinting studies on mitochondrial genomes of different Poaceae family members suggested that *Ae. speltoides* and *Ae. sharonensis* were diverged from the same ancestor and further evolved into different branches (El-Shehawi et al. 2012). Hence, the expression of miR437 might have been lost during the evolutionary history in *Ae. speltoides*, as a consequence of TE-related genome reorganization, although we cannot exclude the possibility that this miRNA could also recently evolve in *Ae. sharonensis*. In order to investigate putative miRNA families specifically related to the B genome lineage, we also tried to identify the conserved miRNAs among the both members of Sitopsis section, *Ae. speltoides* and *Ae. sharonensis*; however, no Sitopsis-section related miRNAs could be detected. This situation may arise from the early evolutionary discrimination of the two (El-Shehawi et al. 2012), although we cannot exclude the possibility of miRNAs that remained unidentified in *Ae. speltoides* transcriptome assembly, due to the small size of this dataset. Regarding to Petersen and his colleagues, *Ae. speltoides* corrupt the monophlogenic cluster in the Aegilops genera of Pooideae subfamily suggesting a different clade for the organism (Petersen et al. 2006). Additionally, *Ae. speltoides* showed a closer relationship with genera Triticum which also consists of *T. aestivum*, *T. monococcum*, and *T. urartu* in another study (Escobar et al. 2011); thus it is tempting to speculate that the miRNA analysis of *Ae. speltoides* also provide clues about distinct nature of this organism compared to other Aegilops genera members of Pooideae subfamily.

Of the *Ae. sharonensis* miRNAs, 107 families were exclusive to only one accession (Table 3). Given that all 16 genotypes belong to the same study (PRJEB5340) exposed to the same conditions prior to RNA-sequencing, these miRNAs likely exhibit accession-specific expression profiles that can be related to individual characteristics of different accessions, such as stress tolerance or environmental adaptation as observed across widely distributed *Ae. sharonensis* populations (Bouyioukos et al. 2013). For instance, miR397, putatively expressed exclusively in accession 2233, exhibited differential expression patterns in tolerant and susceptible soybean genotypes in response to rust disease (Kulcheski et al. 2011). Additionally, putative targets of miR7693, another miRNA potentially exclusive to accession 1995, was involved in biological processes, such as "response to pathogen infection", "protection against oxidative stresses", and "disease resistance" in rice (Campo et al. 2013). *Ae. sharonensis* accessions have been observed to display diverse responses against rust infection (Olivera et al. 2007); hence, further investigation of the putatively accession-specific miRNA families and their targets should contribute to the elucidation of molecular pathways underlying stress responses or other beneficial traits observed in wild wheat populations.

Putative wheat miRNA targets

miRNAs regulate gene expression by binding on the complementary sites of target sequences and directing cleavage, decay, or translational inhibition of the target mRNAs (Budak et al. 2015a, b). Identification of the target transcripts hints the physiological pathways the respective miRNAs are involved in. In order to gain insight into the functions of the putative miRNAs identified in this study, target transcripts were predicted using psRNATarget web-tool and subsequently annotated against all Viridiplantae proteins. Gene ontology (GO) annotations were grouped under "biological process (BP)", "molecular function (MF)," and "cellular component (CC)" categories.

Overall, GO annotations of putative targets included "protein modification process" and "cellular component organization" related terms in all organisms with a marked abundance under the BP category, as could be expected by the regulatory roles of miRNAs. For the CC category, "nucleus", "plasma membrane", "cytosol", and "plastid"-related terms were also commonly attributed to putative miRNA targets. Additionally, "nucleic acid binding" were the most dominant MF term among others in all species tested followed by "kinase activity", "transporter activity", "protein and carbohydrate binding", and "enzyme regulator activity" related terms, pointing out to central roles of miRNAs in gene expression.

miRNAs can target transcription factors frequently and contribute the transcriptional regulation in an indirect manner. This property of the miRNAs is recurrently used as biomarker in animal studies, specifically for cancer experiments (Nazarov et al. 2013; Wu et al. 2015). Considering this

important regulatory function of miRNAs, miRNA families targeting transcriptions factor were analyzed in detailed. Respect to our analysis, ratio of miRNA target transcripts which are associated with transcription factors was in the range of 2 to 5 %. Interestingly, *T. aestivum* was displayed the lowest number of targets related with transcription factors (2.45 % of all target transcripts) while wild species such as *T. monococcum* (5.51 % of all target transcripts) represent slightly higher transcription factor range thought the target transcripts. This situation might be associated with the adaptation of wild plant species into the various range of environments which also interface with many stress related conditions, however further investigation of related miRNA families and their target transcripts is necessary. Different important transcription factor family members such as "WRKY", "MADS", or "MYB" were detected as targeted by both conserved and more specific miRNA families such as miR1130 or miR9494. miRNA families targeting transcription factor families might be important in the regulation of many diverse pathways (Qin et al. 2014), their further characterization is necessary in order elucidate our understanding of transcriptional regulation of and miRNA relationship.

Annotations of the putative miRNA targets exhibited marked differences in bread wheat and its progenitors, despite an overall similarity in GO terms that is expected from their close evolutionary relationships. Under the BP category, as shown in Fig. 3, "post-embryonic development"-related terms were evident in bread wheat but did not appear in any of the progenitors, suggesting that this biological process might have gained significance following many years of cultivation. Intriguingly, "photosynthesis" and "biosynthesis" were among the top terms in *Ae. speltoides*, the closest identifiable relative of the B genome progenitor, while these terms were either not as significant or not present at all in A and D genome progenitors, *T. urartu* and *Ae. tauschii*, and also in *T. aestivum*. In contrast, "response to stress"-related terms did not comprise a significant portion of all BP terms in *Ae. speltoides* (Fig. 3). In general, the distributions of BP category annotations were similar in A and D genome progenitors, and *T. aestivum*, while *Ae. speltoides* had fewer terms that also differed in prominence. This may either be due to the relatively smaller set of putative miRNAs identified from *Ae. speltoides* generating an artificial bias in target prediction and annotation, or the actual B genome progenitor of bread wheat may have slightly different miRNA expression patterns than *Ae. speltoides*.

Under the CC category, "plastid"-related terms were the most prominent in all progenitors and *T. aestivum*, making as much as 48 % of all annotations (Supplementary File S3). Plastid related terms were also highly abundant among CC annotations of *Ae. sharonensis* accessions. In *T. urartu*, as well as in two *T. monococcum* subspecies, several targets with "cell wall" and "peroxisome"-related GO terms were observed, in contrast to *Ae. speltoides*, *Ae. tauschii*, and *T. aestivum*, where none of the targets were annotated as cell wall or peroxisome-related. It is tempting to speculate that this observation may point out to ancestral defense mechanisms existing in the A genome lineage that can be exploited in improving wheat varieties against stress factors. Intriguingly, under the MF category, "hydrolase activity" comprised a

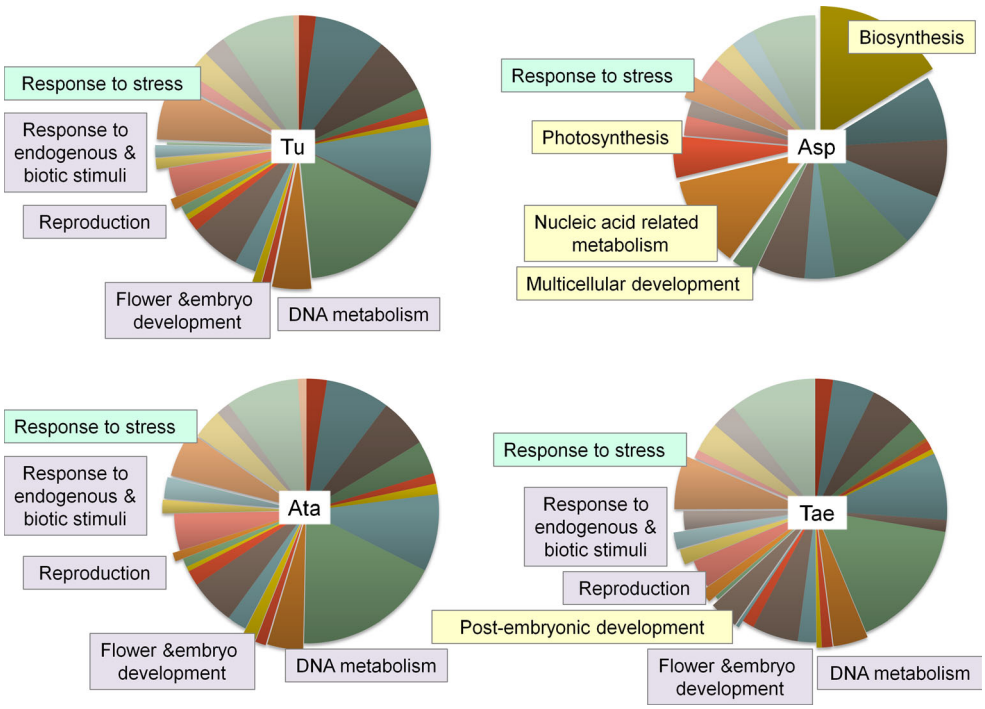

Fig. 3 GO annotations of putative miRNA targets for "biological process" in bread wheat and its progenitors. Annotations unique to one species is highlighted in *pale yellow*, while annotations common to all species or common to Tu, Ata, and Tae are *highlighted in pale green and gray*, respectively, for emphasis. *Other colors* represent diverse "biological process" GO assessments. *Tu*, *T. urartu*, *Asp*, *Ae. speltoides*, *Ata*, *Ae. tauschii*, *Tae*, *T. aestivum*

considerable portion of all target annotations in both *T. aestivum* and *Ae. speltoides*, but not in *T. urartu*, *Ae. tauschii*, or any of the *T. monococcum* subspecies, which had highly similar MF term distributions otherwise. In contrast, "nuclease activity" showed up among only *T. urartu* or *Ae. tauschii* target annotations, and "lipid binding" related terms were exclusive to *T. urartu* (Supplementary File S3).

The Blast2GO annotations of target transcripts of putative *Ae. sharonensis* miRNAs showed that the most of the miRNA targets are located in "plastids", "cytosol", and the "plasma membrane" (Supplementary Figure S1). With respect to the "molecular function" (MF) assessment of GO annotations, most of the miRNA targets were related to "nucleotide binding", "kinase activity", and "protein binding" consistent with previous observation on different plant species (Luo et al. 2006; Liu 2012) (Supplementary Figure S3). In our analysis, *Ae. sharonensis* miRNAs were observed to target an array of biological processes with a marked abundance of "cellular protein modification process", "cellular transport organization", and "transport"-related terms. Additionally, many miRNA target transcripts were involved in stress responses, such as biotic and abiotic stimulus. For instance, Ash_miR1130 and Ash_miR5049 target transcripts were associated with "abscisic acid ripening protein" which is one of the well-known proteins in response to different abiotic (Kalifa et al. 2004; Yang et al. 2005) and biotic stress conditions (Liu et al. 2010). Additionally, Bücker and colleagues detected that this protein acts as a cis-regulating element for miR167 expression in rice or vice versa. *Ae. sharonensis* populations exhibit diverse characteristics against stress conditions (Olivera et al. 2007; Bouyioukos et al. 2013); hence, the elucidation of unique mechanisms underlying stress tolerance of certain subspecies of *Ae. sharonensis* can be aided by further research on such miRNAs. From this point of view, for instance, further analysis and characterization of *Ae. sharonensis* miRNAs targeting the abscisic acid ripening protein may be promising.

Interestingly a few miRNAs, such as miR1130 and miR398, targeted specific proteins associated with cell death such as "programmed cell death protein 4". Among these, miR398 is well characterized in many plants such as *Arabidopsis thaliana* (Sunkar et al. 2006) and *Medicago truncatula* (Trindade et al. 2010) and its differential expression was detected with respect to different stress conditions such as oxidative, salt, or drought stresses (Zhu et al. 2011a). It is a known phenomenon that plants go through apoptosis under different stress conditions, especially under pathogen attack, and regulation of this process hold importance for the survival of cell (Greenberg 1997; Kuzuoglu-Ozturk et al. 2012). Distinct miRNA families might have crucial roles on the regulation of apoptotic/autophagic/necrotic cell death mechanisms under stress conditions and might play significant roles for cell survival.

Repetitive sequences found within putative miRNAs

The contribution of transposable elements (TEs) on the miRNA repertoire of organisms has been fairly well described in many studies (Smalheiser and Torvik 2005; Yao et al. 2007; Piriyapongsa and Jordan 2008; Li et al. 2011). Considering this, we tried to analyze the association of putative miRNAs, identified in this study, with known plant TEs by aligning the pre-miRNA sequences to Poaceae repeat library which contains 34,135 sequences. From 73 to 85 % of the identified stem-loops were mapped against TEs with more than 50 % of their lengths in each plant species; these miRNAs are classified as TE-MIRs (Fig. 4a). The high resemblance of plant miRNAs with repetitive sequences of the genome has been previously shown in several studies; for instance, Piriyapongsa and Jordan revealed that 10 out of 12 *Arabidopsis thaliana* miRNAs were found to be identical to repetitive elements while the same situation was observed for 38 out of 83 miRNAs for *Oryza sativa* (Piriyapongsa and Jordan 2008). Recent studies suggested that miRNA genes

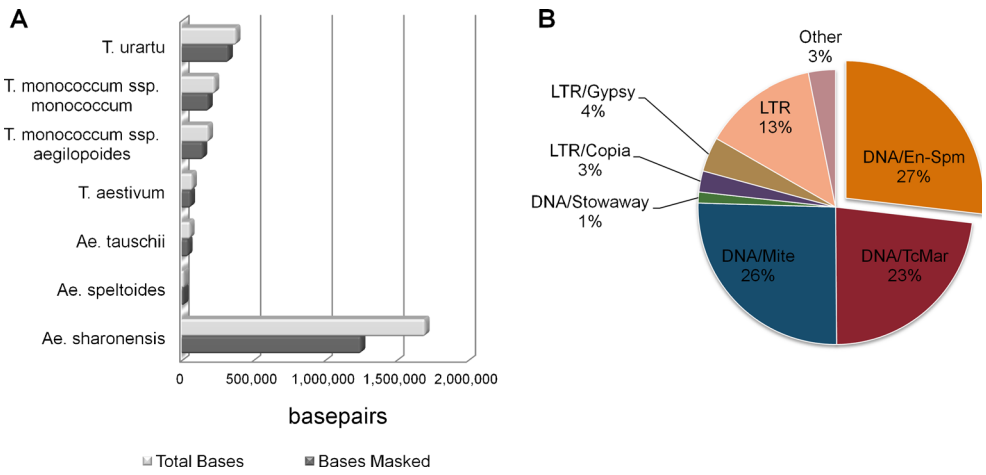

Fig. 4 Repetitive elements found within pre-miRNA sequences. **a** The ratio of repetitive elements to total bases, **b** The distribution of major TE families found within pre-miRNA sequences

might have evolved from imperfect inverted repeats with the help of accumulated mutations (Fahlgren et al. 2007; Feldman and Levy 2012). Additionally, another hypothesis proposed that miRNA genes are directly derived from homologous TEs where they may directly transcribed to miRNAs with the help of related molecular elements (Li et al. 2011). Despite the existing ideas to explain the miRNA-TE relationship, the authentic mechanism(s) of miRNA biogenesis from TEs is yet-to-be discovered. It is also important to note that identification of TE-MIRs is still under controversy, in particular, due to the resemblance of these miRNAs to repeat-related small interfering RNAs (siRNAs). In our pipeline, the required presence of pre-miRNA like stem-loop structures for putative mature miRNAs should prevent misannotation of TE-related siRNAs as TE-MIRs.

In our analysis, putative pre-miRNAs were associated with 667 different repeat sequences including simple repeats, DNA transposons, and retrotransposons. Among 259 different putative miRNA families identified from this study, 111 of them were annotated as TE-MIRs while the most abundant TE-MIRs were from the families, miR5049, miR1130, miR1127, miR1120, and miR5181 families. Additionally, these miRNA families were also detected as the "highly abundant" and constitute approximately 55 % of all identified pre-miRNA sequences across all species; therefore, it is tempting to speculate that TE proliferation may simultaneously increase TE-MIR copies. Despite the presence of both Type I and Type II repeat elements across identified TE-MIRs, the abundances of "Enhancer/Suppressor mutator-like DNA transposons (DNA/En-Spm)", "miniature inverted repeat transposable elements (DNA/MITE)", and "Tcl/Mariner DNA transposon (DNA/TcMar)" families were remarkable (Fig. 4b). Most of the identified TE-MIRs were associated with more than one repeat family, while cases where a TE-MIR uniquely mapped to a single repeat family were also observed. As an example, Ash_miR854 (Ash: *Ae. sharonensis*) was specifically related to $(GGA)_n$ simple repeats, while Ash_miR819 aligned with many DNA transposons from Stowaway, Tcl/Mariner, En-Spm repeat subfamilies. In addition, our observations were in general agreement with previous studies of orthologous miRNAs, such as osa_miR819 which was also associated with DNA/Stowaway family (Piriyapongsa and Jordan 2008). Curiously, contrasting observations on TE-MIRs and the respective repeat families also exist; for instance, miR854 was related to LTR/Gypsy family in rice (Piriyapongsa and Jordan 2008), while this family was associated with $(GGA)_n$ simple repeats in *Ae. sharonensis* in our study.

Several studies have proposed that the miRNA families arise from TE-to-MITE transition during evolutionary processes (Yao et al. 2007; Piriyapongsa and Jordan 2008) where DNA/MITE origins are also linked with the Tcl/Mariner DNA transposon family (Fattash et al. 2013). A recent study from Yu and his colleagues reported that a vernalization-related gene from *T. aestivum*, Vrn-Ala which contains MITE-family repeat elements in its promoter region, is potentially regulated with miR1123 that also carry similar DNA/MITE repeats in its pre-miRNA sequence (Yu et al. 2014). They also successfully validated both the presence of Tae_miR1123 and its unique stem-loop which is transcribed from MITEs with wet-lab tools. These findings, together with our observations, support the hypothesis on TE-related miRNA evolution in plants (Piriyapongsa and Jordan 2008; Li et al. 2011), especially for crop species which have high contents of repetitive elements. Hence, careful handling of repetitive elements in the process of computational miRNA identification holds importance for unraveling of the entire miRNA repertoire of organisms, including TE-MIRs.

Bread wheat is a major constituent of human nutrition. Being a hardy crop, grown extensively in temperate regions, wheat is likely to preserve its agronomic value in the future, as well. However, challenged by global climate changes and the ever increasing world population, stable and sustainable increases in wheat yields are crucial to ensure the food security of upcoming generations. The modulation of miRNA-regulated pathways has become a promising aspect of crop improvement in recent years. Accordingly, several studies have focused on the identification and characterization of these small regulatory molecules in wheat and other crop species. Despite its high agronomic importance, genomics research on wheat has only begun to accelerate due its large, highly repetitive a challenging genome. While genomics-based tools are slowly accumulating to aid breeding programs, wild wheat populations, and progenitors provide a unique resource of favorable alleles and surrogates at the diploid level to study the complex wheat genome. Here, we explored the miRNA repertoires of bread wheat, *Triticum aestivum*, and its diploid progenitors and relatives, *Triticum urartu*, *Aegilops speltoides*, *Aegilops tauschii*, *Triticum monococcum*, and *Aegilops sharonensis* through homology-based in silico miRNA identification from transcriptome sequences. Wheat progenitors provided a unique perspective of miRNA evolution through putatively expressed miRNAs. Two *T. monococcum* subspecies, the wild *T. monococcum ssp. aegilopoides* and the domesticated *T. monococcum ssp. monococcum*, suggested clues into domestication from a miRNA expression point of view. Sixteen *Aegilops sharonensis* genotypes, included in this study, revealed several miRNAs that are putatively expressed in only one genotype, but not the others, which may be related to the high diversity observed in *Ae. sharonensis* populations. In this study, we also observed many miRNAs that were associated with transposable elements. Overall, our observations provide important insights into miRNA evolution in bread wheat and its progenitors and relatives that can be exploited to dissect regulatory pathways of interest or can be targeted for wheat improvement.

Acknowledgments This research was funded by Sabanci University and Montana Plant Science Endowment.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Akhunova AR, Matniyazov RT, Liang H, Akhunov ED (2010) Homoeolog-specific transcriptional bias in allopolyploid wheat. BMC Genomics 11:505. doi:10.1186/1471-2164-11-505

Akpinar BA, Avsar B, Lucas SJ, Budak H (2012) Plant abiotic stress signaling. Plant Signal Behav 7:1450–1455

Akpinar BA, Lucas SJ, Vrána J, et al. (2014) Sequencing chromosome 5D of Aegilops tauschii and comparison with its allopolyploid descendant bread wheat (Triticum aestivum). Plant Biotechnol J 1–13. doi: 10.1111/pbi.12302

Akpinar BA, Kantar M, Budak H (2015) Root precursors of microRNAs in wild emmer and modern wheats show major differences in response to drought stress. Funct Integr Genomics. doi:10.1007/s10142-015-0453-0

Ani Akpinar B, Yuce M, Lucas S et al (2015) Molecular organization and comparative analysis of chromosome 5B of the wild wheat ancestor Triticum dicoccoides. Sci Rep 5:10763. doi:10.1038/srep10763

Barakat A, Sriram A, Park J et al (2012) Genome wide identification of chilling responsive microRNAs in Prunus persica. BMC Genomics 13:481. doi:10.1186/1471-2164-13-481

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. doi:10.1093/bioinformatics/btu170

Bouyioukos C, Moscou MJ, Champouret N et al (2013) Characterisation and analysis of the Aegilops sharonensis transcriptome, a wild relative of wheat in the Sitopsis section. PLoS ONE 8:1–14. doi:10.1371/journal.pone.0072782

Brenchley R, Spannagl M, Pfeifer M et al (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature 491:705–710. doi:10.1038/nature11650

Budak H (2010) Plant genetic resources: effective utilization. Engineering. doi:10.1081/E-EBAF-9780849350504

Budak H, Akpinar A (2011) Dehydration stress-responsive miRNA in Brachypodium distachyon: evident by genome-wide screening of microRNAs expression. Omi A J Integr Biol 15:791–799. doi:10.1089/omi.2011.0073

Budak H, Akpinar BA (2015) Plant miRNAs: biogenesis, organization and origins. Funct Integr Genomics. doi:10.1007/s10142-015-0451-2

Budak H, Kantar M (2015) Harnessing NGS and big data optimally: comparison of miRNA prediction from assembled versus non-assembled sequencing data—the case of the grass *Aegilops tauschii* complex genome. Omi A J Integr Biol 19:407–415. doi:10.1089/omi.2015.0038

Budak H, Kantar M, Yucebilgili Kurtoglu K (2013) Drought tolerance in modern and wild wheat. Sci World J. doi:10.1155/2013/548246

Budak H, Khan Z, Kantar M (2014) History and current status of wheat miRNAs using next-generation sequencing and their roles in development and stress. Brief Funct Genomics elu021–. doi: 10.1093/bfgp/elu021

Budak H, Kantar M, Bulut R, Akpinar BA (2015a) Stress responsive miRNAs and isomiRs in cereals. Plant Sci 235:1–13. doi:10.1016/j.plantsci.2015.02.008

Budak H, Bulut R, Kantar M, Alptekin B (2015a) MicroRNA nomenclature and the need for a revised naming prescription. Brief Funct Genomics elv026. doi: 10.1093/bfgp/elv026

Campo S, Peris-Peris C, Siré C et al (2013) Identification of a novel microRNA (miRNA) from rice that targets an alternatively spliced transcript of the Nramp6 (natural resistance-associated macrophage protein 6) gene involved in pathogen resistance. New Phytol 199:212–227. doi:10.1111/nph.12292

Chalupska D, Lee HY, Faris JD et al (2008) Acc homoeoloci and the evolution of wheat genomes. Proc Natl Acad Sci U S A 105:9691–9696. doi:10.1073/pnas.0803981105

Chamnongpol S, Maroney PA, Nilsen TW (2010) A rapid, quantitative assay for direct detection of microRNAs and other small RNAs using splinted ligation. Methods Mol Biol 667:3–17. doi:10.1261/rna.518107

Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science 303:2022–2025. doi:10.1126/science.1088060

Chen ZJ (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu Rev Plant Biol 58:377–406. doi:10.1146/annurev.arplant.58.032806.103835

Chen M, Cao Z (2015) Genome-wide expression profiling of microRNAs in poplar upon infection with the foliar rust fungus Melampsora larici-populina. BMC Genomics 16:696. doi:10.1186/s12864-015-1891-8

Conesa A, Götz S (2008) Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008:619832. doi:10.1155/2008/619832

Curaba J, Spriggs A, Taylor J et al (2012) miRNA regulation in the early development of barley seed. BMC Plant Biol 12:120. doi:10.1186/1471-2229-12-120

Dai X, Zhao PX (2011) PsRNATarget: a plant small RNA target analysis server. Nucleic Acids Res. doi:10.1093/nar/gkr319

Devos KM, Dubcovsky J, Dvořák J et al (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor Appl Genet 91:282–288. doi:10.1007/BF00220890

Dudnikov AJ, Kawahara T (2006) Aegilops tauschii: genetic variation in Iran. Genet Resour Crop Evol 53:579–586. doi:10.1007/s10722-004-2681-3

Dvorák J, McGuire PE (1981) Nonstructural chromosome differentiation among wheat cultivars, with special reference to differentiation of chromosomes in related species. Genetics 97:391–414

Dvořák J, Zhang H-B, Kota RS, Lassner M (1989) Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. Genome 32:1003–1016. doi:10.1139/g89-545

Dvorák J, Terlizzi P, Zhang HB, Resta P (1993) The evolution of polyploid wheats: identification of the A genome donor species. Genome 36:21–31. doi:10.1139/g93-004

Eilam T, Anikster Y, Millet E et al (2007) Genome size and genome evolution in diploid Triticeae species. Genome 50:1029–1037. doi:10.1139/G07-083

El-Shehawi AM, Fahmi AI, Sayed SM, Elseehy MM (2012) Genetic fingerprinting of wheat and its progenitors by mitochondrial gene orf256. Biomolecules 2:228–239. doi:10.3390/biom2020228

Escobar JS, Scornavacca C, Cenci A et al (2011) Multigenic phylogeny and analysis of tree incongruences in Triticeae (Poaceae). BMC Evol Biol 11:181. doi:10.1186/1471-2148-11-181

Fahlgren N, Howell MD, Kasschau KD et al (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS ONE 2, e219. doi:10.1371/journal.pone.0000219

Fan R, Li Y, Li C, Zhang Y (2015) Differential microRNA analysis of glandular trichomes and young leaves in Xanthium strumarium L. reveals their putative roles in regulating terpenoid biosynthesis. PLoS ONE 10, e0139002. doi:10.1371/journal.pone.0139002

Fattash I, Rooke R, Wong A et al (2013) Miniature inverted-repeat transposable elements: discovery, distribution, and activity 1. Genome 56:475–486. doi:10.1139/gen-2012-0174

Feldman M, Levy AA (2012) Genome evolution due to allopolyploidization in wheat. Genetics 192:763–774. doi:10.1534/genetics.112.146316

Fox SE, Geniza M, Hanumappa M et al (2014) De novo transcriptome assembly and analyses of gene expression during photomorphogenesis in diploid wheat Triticum monococcum. PLoS ONE. doi:10.1371/journal.pone.0096855

Fricano A, Brandolini A, Rossini L et al (2014) Crossability of Triticum urartu and Triticum monococcum wheats, homoeologous recombination, and description of a panel of interspecific introgression lines. G3; Genes|Genomes|Genetics 4:1931–1941. doi:10.1534/g3.114.013623

Grabherr MG, Haas BJ, Yassour M et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652. doi:10.1038/nbt.1883

Greenberg JT (1997) Programmed cell death in plant-pathogen interactions. Annu Rev Plant Physiol Plant Mol Biol 48:525–545. doi:10.1146/annurev.arplant.48.1.525

Haider N (2012) Evidence for the origin of the B genome of bread wheat based on chloroplast DNA. Turkish J Agric For 36:13–25. doi:10.3906/tar-1011-1394

Han R, Jian C, Lv J et al (2014) Identification and characterization of microRNAs in the flag leaf and developing seed of wheat (Triticum aestivum L.). BMC Genomics 15:289. doi:10.1186/1471-2164-15-289

Huang SQ, Peng J, Qiu CX, Yang ZM (2009) Heavy metal-regulated new microRNAs from rice. J Inorg Biochem 103:282–287. doi:10.1016/j.jinorgbio.2008.10.019

Iehisa JCM, Shimizu A, Sato K et al (2012) Discovery of high-confidence single nucleotide polymorphisms from large-scale de novo analysis of leaf transcripts of Aegilops tauschii, a wild wheat progenitor. DNA Res 19:487–497. doi:10.1093/dnares/dss028

Jia J, Zhao S, Kong X et al (2013) Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. Nature 496:91–5. doi:10.1038/nature12028

Jiang J, Gill BS (1994) Different species-specific chromosome translocations in Triticum timopheevii and T. turgidum support the diphyletic origin of polyploid wheats. Chromosom Res 2:59–64. doi:10.1007/BF01539455

Jin W, Li N, Zhang B et al (2008) Identification and verification of microRNA in wheat (Triticum aestivum). J Plant Res 121:351–355. doi:10.1007/s10265-007-0139-3

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14:787–799. doi:10.1016/j.molcel.2004.05.027

Kadri S, Hinman V, Benos PV (2009) HHMMiR: efficient de novo prediction of microRNAs using hierarchical hidden Markov models. BMC Bioinf 10(Suppl 1):S35. doi:10.1186/1471-2105-10-S1-S35

Kalifa Y, Gilad A, Konrad Z et al (2004) The water- and salt-stress-regulated Asr1 (abscisic acid stress ripening) gene encodes a zinc-dependent DNA-binding protein. Biochem J 381:373–378. doi:10.1042/BJ20031800

Kantar M, Unver T, Budak H (2010) Regulation of barley miRNAs upon dehydration stress correlated with target gene expression. Funct Integr Genomics 10:493–507. doi:10.1007/s10142-010-0181-4

Kantar M, Lucas SJ, Budak H (2011) miRNA expression patterns of Triticum dicoccoides in response to shock drought stress. Planta 233:471–484. doi:10.1007/s00425-010-1309-4

Kantar M, Akpinar BA, Valárik M et al (2012) Subgenomic analysis of microRNAs in polyploid wheat. Funct Integr Genomics 12:465–479. doi:10.1007/s10142-012-0285-0

Kilian B, Özkan H, Deusch O et al (2007) Independent wheat B and G genome origins in outcrossing Aegilops progenitor haplotypes. Mol Biol Evol 24:217–227. doi:10.1093/molbev/msl151

Kozomara A, Griffiths-Jones S (2011) MiRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. doi:10.1093/nar/gkq1027

Kruszka K, Pacak A, Swida-Barteczka A et al (2013) Developmentally regulated expression and complex processing of barley pri-microRNAs. BMC Genomics 14:34. doi:10.1186/1471-2164-14-34

Kulcheski FR, de Oliveira LF, Molina LG et al (2011) Identification of novel soybean microRNAs involved in abiotic and biotic stresses. BMC Genomics 12:307. doi:10.1186/1471-2164-12-307

Kumar A, Simons K, Iqbal MJ et al (2012) Physical mapping resources for large plant genomes: radiation hybrids for wheat D-genome progenitor Aegilops tauschii. BMC Genomics 13:597. doi:10.1186/1471-2164-13-597

Kurtoglu KY, Kantar M, Lucas SJ, Budak H (2013) Unique and conserved microRNAs in wheat chromosome 5D revealed by next-generation sequencing. PLoS ONE. doi:10.1371/journal.pone.0069801

Kurtoglu KY, Kantar M, Budak H (2014) New wheat microRNA using whole-genome sequence. Funct Integr Genomics 14:363–379. doi:10.1007/s10142-013-0357-9

Kuzuoglu-Ozturk D, Yalcinkaya OC, Akpinar BA et al (2012) Autophagy-related gene, TdAtg8, in wild emmer wheat plays a role in drought and osmotic stress response. Planta 236:1081–1092. doi:10.1007/s00425-012-1657-3

Li Y, Li C, Xia J, Jin Y (2011) Domestication of transposable elements into microRNA genes in plants. PLoS ONE. doi:10.1371/journal.pone.0019212

Li A, Liu D, Wu J, et al. (2014) mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. Plant Cell 1–24. doi: 10.1105/tpc.114.124388

Ling H-Q, Zhao S, Liu D et al (2013) Draft genome of the wheat A-genome progenitor Triticum urartu. Nature 496:87–90. doi:10.1038/nature11997

Liu Q (2012) Novel miRNAs in the control of arsenite levels in rice. Funct Integr Genomics 12:649–658. doi:10.1007/s10142-012-0282-3

Liu HY, Dai JR, Feng DR et al (2010) Characterization of a novel plantain asr gene, mpasr, that is regulated in response to infection of Fusarium oxysporum f. sp. cubense and abiotic stresses. J Integr Plant Biol 52:315–323. doi:10.1111/j.1744-7909.2010.00912.x

Liu W, Xu L, Wang Y et al (2015) Transcriptome-wide analysis of chromium-stress responsive microRNAs to explore miRNA-mediated regulatory networks in radish (Raphanus sativus L.). Sci Rep 5:14024. doi:10.1038/srep14024

Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science 297:2053–2056. doi:10.1126/science.1076311

Lopes MS, El-Basyoni I, Baenziger PS et al (2015) Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J Exp Bot 66:3477–3486. doi:10.1093/jxb/erv122

Lucas SJ, Budak H (2012) Sorting the wheat from the chaff: identifying miRNAs in genomic survey sequences of Triticum aestivum chromosome 1AL. PLoS ONE. doi:10.1371/journal.pone.0040859

Lucas SJ, Baştaş K, Budak H (2014) Exploring the interaction between small RNAs and R genes during Brachypodium response to Fusarium culmorum infection. Gene 536:254–264. doi:10.1016/j.gene.2013.12.025

Luo YC, Zhou H, Li Y et al (2006) Rice embryogenic calli express a unique set of microRNAs, suggesting regulatory roles of microRNAs in plant post-embryogenic development. FEBS Lett 580:5111–5116. doi:10.1016/j.febslet.2006.08.046

Luo M-C, Gu YQ, You FM et al (2013) A 4-gigabase physical map unlocks the structure and evolution of the complex genome of Aegilops tauschii, the wheat D-genome progenitor. Proc Natl Acad Sci U S A 110:7940–5. doi:10.1073/pnas.1219082110

Lv S, Nie X, Wang L et al (2012) Identification and characterization of microRNAs from barley (hordeum vulgare L.) by high-throughput sequencing. Int J Mol Sci 13:2973–2984. doi:10.3390/ijms13032973

Marcussen T, Sandve SR, Heier L et al (2014) Ancient hybridizations among the ancestral genomes of bread wheat. Science 345:1250092. doi:10.1126/science.1250092

Markham NR, Zuker M (2008) UNAFold: software for nucleic acid folding and hybridization. Methods Mol Biol 453:3–31. doi:10.1007/978-1-60327-429-6-1

Middleton CP, Senerchia N, Stein N et al (2014) Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the triticeae tribe. PLoS ONE. doi:10.1371/journal.pone.0085761

Mochida K, Shinozaki K (2013) Unlocking triticeae genomics to sustainably feed the future. Plant Cell Physiol 54:1931–1950. doi:10.1093/pcp/pct163

Munns R, James RA, Xu B et al (2012) Wheat grain yield on saline soils is improved by an ancestral Na^+ transporter gene. Nat Biotechnol 30:360–364

Murat F, Zhang R, Guizard S et al (2014) Shared subgenome dominance following polyploidization explains grass genome evolutionary plasticity from a seven protochromosome ancestor with 16K protogenes. Genome Biol Evol 6:12–33. doi:10.1093/gbe/evt200

Navarro L, Dunoyer P, Jay F et al (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science 312:436–9. doi:10.1126/science.1126088

Nazarov PV, Reinsbach SE, Muller A et al (2013) Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function. Nucleic Acids Res 41:2817–2831. doi:10.1093/nar/gks1471

Neilson JR, Zheng GXY, Burge CB, Sharp PA (2007) Dynamic regulation of miRNA expression in ordered stages of cellular development. Genes Dev 21:578–589. doi:10.1101/gad.1522907

Nussbaumer T, Martis MM, Roessner SK et al (2013) MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Res. doi:10.1093/nar/gks1153

Olivera PD, Kolmer JA, Anikster Y, Steffenson BJ (2007) Resistance of sharon goatgrass (Aegilops sharonensis) to fungal diseases of wheat. Plant Dis 91:942–950. doi:10.1094/pdis-91-8-0942

Parry MAJ, Reynolds M, Salvucci ME et al (2011) Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. J Exp Bot 62:453–467. doi:10.1093/jxb/erq304

Petersen G, Seberg O, Yde M, Berthelsen K (2006) Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). Mol Phylogenet Evol 39:70–82. doi:10.1016/j.ympev.2006.01.023

Piriyapongsa J, Jordan IK (2008) Dual coding of siRNAs and miRNAs by plant transposable elements. RNA 14:814–821. doi:10.1261/rna.916708

Pont C, Murat F, Guizard S et al (2013) Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. Plant J 76:1030–1044. doi:10.1111/tpj.12366

Qin Z, Li C, Mao L, Wu L (2014) Novel insights from non-conserved microRNAs in plants. Front Plant Sci 5:586. doi:10.3389/fpls.2014.00586

Ramalingam P, Palanichamy JK, Singh A et al (2014) Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. RNA 20:76–87. doi:10.1261/rna.041814.113

Ruuska SA, Lewis DC, Kennedy G et al (2008) Large scale transcriptome analysis of the effects of nitrogen nutrition on accumulation of stem carbohydrate reserves in reproductive stage wheat. Plant Mol Biol 66:15–32. doi:10.1007/s11103-007-9249-5

Saintenac C, Zhang W, Salcedo A (2013) Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. Science 341:783–786. doi:10.1126/science.1239022

Sano H, Youssefian S (1994) Light and nutritional regulation of transcripts encoding a wheat protein kinase homolog is mediated by cytokinins. Proc Natl Acad Sci U S A 91:2582–2586

Sarkar P, Stebbins GL (1956) Morphological evidence concerning the origin of the B genome in wheat. Am J Bot 43:297–304. doi:10.2307/2438947

Schwab R, Palatnik JF, Riester M et al (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell 8:517–527. doi:10.1016/j.devcel.2005.01.018

Singh N, Srivastava S, Sharma A (2015) Identification and analysis of miRNAs and their targets in ginger using bioinformatics approach. Gene. doi:10.1016/j.gene.2015.09.036

Smalheiser NR, Torvik VI (2005) Mammalian microRNAs derived from genomic repeats. Trends Genet 21:322–6. doi:10.1016/j.tig.2005.04.008

Sunkar R, Kapoor A, Zhu J-K (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. Plant Cell 18:2051–2065. doi:10.1105/tpc.106.041673

Sunkar R, Zhou X, Zheng Y et al (2008) Identification of novel and candidate miRNAs in rice by high throughput sequencing. BMC Plant Biol 8:25. doi:10.1186/1471-2229-8-25

Teune J-H, Steger G (2010) NOVOMIR: de novo prediction of microRNA-coding regions in a single plant-genome. J Nucleic Acids. doi:10.4061/2010/495904

Thakur V, Wanchana S, Xu M et al (2011) Characterization of statistical features for plant microRNA prediction. BMC Genomics 12:108. doi:10.1186/1471-2164-12-108

Thiebaut F, Grativol C, Carnavale-Bottino M et al (2012) Computational identification and analysis of novel sugarcane microRNAs. BMC Genomics 13:290. doi:10.1186/1471-2164-13-290

Trindade I, Capitão C, Dalmay T et al (2010) miR398 and miR408 are up-regulated in response to water deficit in Medicago truncatula. Planta 231:705–716. doi:10.1007/s00425-009-1078-0

Unver T, Budak H (2009) Conserved micrornas and their targets in model grass species Brachypodium distachyon. Planta 230:659–669. doi:10.1007/s00425-009-0974-7

Vaughan MM, Christensen S, Schmelz EA et al (2015) Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. Plant Cell Environ. doi:10.1111/pce.12482

Wang X-J, Reyes JL, Chua N-H, Gaasterland T (2004) Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. Genome Biol 5:R65. doi:10.1186/gb-2004-5-9-r65

Wang M, Wang Q, Wang B (2012) Identification and characterization of microRNAs in Asiatic cotton (Gossypium arboreum L.). PLoS ONE. doi:10.1371/journal.pone.0033696

Wang J, Luo MC, Chen Z et al (2013) Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. New Phytol 198:925–937. doi:10.1111/nph.12164

Wei B, Cai T, Zhang R et al (2009) Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (Triticum aestivum L.) and Brachypodium distachyon (L.) Beauv. Funct Integr Genomics 9:499–511. doi:10.1007/s10142-009-0128-9

Wu Q, Qin H, Zhao Q, He X (2015) Emerging role of transcription factor-microRNA-target gene feed-forward loops in cancer (Review). Biomed Reports 611–616. doi: 10.3892/br.2015.477

Xie F, Zhang B (2015) microRNA evolution and expression analysis in polyploidized cotton genome. Plant Biotechnol J. doi:10.1111/pbi.12295

Xing S, Salinas M, Höhmann S et al (2010) miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. Plant Cell 22:3935–3950. doi:10.1105/tpc.110.079343

Yan Y, Wang H, Hamera S et al (2014) MiR444a has multiple functions in the rice nitrate-signaling pathway. Plant J 78:44–55. doi:10.1111/tpj.12446

Yang C-Y, Chen Y-C, Jauh GY, Wang C-S (2005) A Lily ASR protein involves abscisic acid signaling and confers drought and salt resistance in Arabidopsis. Plant Physiol 139:836–846. doi:10.1104/pp.105.065458

Yao C, Zhao B, Li W et al (2007) Cloning of novel repeat-associated small RNAs derived from hairpin precursors in Oryza sativa. Acta Biochim Biophys Sin Shanghai 39:829–34. doi:10.1111/j.1745-7270.2007.00346.x

Yu M, Carver BF, Yan L (2014) TamiR1123 originated from a family of miniature inverted-repeat transposable elements (MITE) including one inserted in the Vrn-A1a promoter in wheat. Plant Sci 215–216:117–123. doi:10.1016/j.plantsci.2013.11.007

Zaharieva M, Monneveux P (2014) Cultivated einkorn wheat (Triticum monococcum L. subsp. monococcum): the long life of a founder crop of agriculture. Genet Resour Crop Evol 61:677–706. doi:10.1007/s10722-014-0084-7

Zanca AS, Vicentini R, Ortiz-Morea FA et al (2010) Identification and expression analysis of microRNAs and targets in the biofuel crop sugarcane. BMC Plant Biol 10:260. doi:10.1186/1471-2229-10-260

Zhang BH, Pan XP, Wang QL et al (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15:336–360. doi:10.1038/sj.cr.7290302

Zhang B, Pan X, Cannon CH et al (2006) Conservation and divergence of plant microRNA genes. Plant J 46:243–259. doi:10.1111/j.1365-313X.2006.02697.x

Zhang B, Wang Q, Pan X (2007) MicroRNAs and their regulatory roles in animals and plants. J Cell Physiol 210:279–289

Zhang L, Chia JM, Kumari S et al (2009) A genome-wide characterization of microRNA genes in maize. PLoS Genet 5, e1000716. doi:10.1371/journal.pgen.1000716

Zhou X, Wang G, Sutoh K et al (2008) Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochim Biophys Acta - Gene Regul Mech 1779:780–788. doi:10.1016/j.bbagrm.2008.04.005

Zhou L, Liu Y, Liu Z et al (2010) Genome-wide identification and analysis of drought-responsive microRNAs in Oryza sativa. J Exp Bot 61:4157–4168. doi:10.1093/jxb/erq237

Zhu C, Ding Y, Liu H (2011a) MiR398 and plant stress responses. Physiol Plant 143:1–9. doi:10.1111/j.1399-3054.2011.01477.x

Zhu S, Jiang Q, Wang G et al (2011b) Chromatin structure characteristics of pre-miRNA genomic sequences. BMC Genomics 12:329. doi:10.1186/1471-2164-12-329