rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

NRC Research Press

ARTICLE

Characterization of microRNAs and their targets in wild barley (*Hordeum vulgare* subsp. *spontaneum*) using deep sequencing

Fn1

Pingchuan Deng, Jianxin Bian, Hong Yue, Kewei Feng, Mengxing Wang, Xianghong Du, Song Weining, and Xiaojun Nie

Abstract: MicroRNAs (miRNA) are a class of small, endogenous RNAs that play a negative regulatory role in various developmental and metabolic processes of plants. Wild barley (*Hordeum vulgare* subsp. *spontaneum*), as the progenitor of cultivated barley (*Hordeum vulgare* subsp. *vulgare*), has served as a valuable germplasm resource for barley genetic improvement. To survey miRNAs in wild barley, we sequenced the small RNA library prepared from wild barley using the Illumina deep sequencing technology. A total of 70 known miRNAs and 18 putative novel miRNAs were identified. Sequence analysis revealed that all of the miRNAs identified in wild barley contained the highly conserved hairpin sequences found in barley cultivars. MiRNA target predictions showed that 12 out of 52 miRNA families were predicted to target transcription factors, including 8 highly conserved miRNA families in plants and 4 wheat–barley conserved miRNA families. In addition to transcription factors, other predicted target genes were involved in diverse physiological and metabolic processes and stress defense. Our study for the first time reported the large-scale investigation of small RNAs in wild barley, which will provide essential information for understanding the regulatory role of miRNAs in wild barley and also shed light on future practical utilization of miRNAs for barley improvement.

Key words: MiRNA, wild barley, sequencing, RT-PCR.

Résumé : Les microARN (miRNA) sont une classe de petits ARN endogènes qui jouent un rôle de régulateur négatif dans divers processus développementaux et métaboliques chez les plantes. L'orge sauvage (*Hordeum vulgare* ssp. *spontaneum*), en tant qu'ancêtre de l'orge cultivée (*Hordeum vulgare* ssp. *vulgare*), a servi de ressource génétique précieuse pour l'amélioration génétique de l'orge. Afin de dresser l'inventaire des miRNA chez l'orge sauvage, les auteurs ont séquencé une librairie de petits ARN de l'orge sauvage au moyen d'un séquençage Illumina en profondeur. Au total, 70 miARN connus et 18 miRNA potentiellement inédits ont été identifiés. L'analyse des séquences a révélé que tous les miRNA identifiés chez l'orge sauvage contiennent les séquences en épingle qui sont hautement conservées chez les cultivars d'orge. Les prédictions en matière de cibles pour ces miRNA ont montré que les cibles prédites pour 12 des 52 familles de miRNA sont des facteurs de transcription; celles-ci incluent 8 familles de miRNA hautement conservées chez les plantes ainsi que 4 familles de miRNA conservées chez l'orge et le blé. En plus de facteurs de transcription, d'autres cibles prédites incluent des gènes impliqués dans divers processus physiologiques et métaboliques ainsi que dans les réponses aux stress. Cette étude est la première à rapporter une caractérisation à grande échelle des petits ARN chez l'orge sauvage et a aussi jeté un éclairage sur de futures utilisations pratiques des miRNA pour l'amélioration génétique de l'orge. [Traduit par la Rédaction]

Mots-clés : miRNA, orge sauvage, séquençage, RT-PCR.

Introduction

Wild barley (*Hordeum vulgare* subsp. *spontaneum*), as the progenitor of cultivated barley (*Hordeum vulgare* subsp. *vulgare*), originated about 5.5 million years ago in southwest Asia and then spread to the Eastern Mediterranean, Balkans, North Africa, Central Asia, and Tibet (Dai et al. 2012; Hartmann et al. 2006). Wild barley is found to grow over a remarkably wide range of ecological habitats, including deserts and highlands. The long-term coevolution and local adaption to microclimates resulted in wild

Received 21 December 2015. Accepted 1 March 2016.

Corresponding Editor: S. Cloutier.

P. Deng,* J. Bian,* H. Yue, K. Feng, M. Wang, X. Du, S. Weining, and X. Nie. State Key Laboratory of Crop Stress Biology in Arid Areas, College of Agronomy and Yangling Branch of China Wheat Improvement Center, Northwest A&F University, Yangling, Shaanxi 712100, China.

Corresponding authors: Song Weining (email: sweining2002@yahoo.com); Xiaojun Nie (email: small@nwsuaf.edu.cn).
*These authors contributed equally to this work.

 Published at www.nrcresearchpress.com/gen on 18 March 2016.

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

barley having a high level of genetic diversity and a large number of alleles for many economically important traits (Ellis et al. 2000; Fetch et al. 2003; Roy et al. 2010; Shakhatreh et al. 2010; Steffenson et al. 2007). Previous studies demonstrated that wild barley could display unique and multiple adaptive divergence complexes and remarkable incipient sympatric ecological speciation across an average distance of only 200 m (Shen et al. 2013; Yang et al. 2011). Furthermore, wild barley also showed adaptation to extreme climatic conditions, including high temperature, drought, and high soil salinity, thus providing an indispensable genetic resource for the breeding of stress-tolerant varieties of the cultivated crop and held the promise for barley improvement to meet the challenges of increasing food demand and changing global climate in the future (The International Barley Genome Sequencing Consortium 2012).

MicroRNAs (miRNAs) are a class of endogenous, small (19–24 nt), non-coding RNAs and are now well known as one of the most important factors for gene expression regulation in eukaryotic cells (Bartel 2004; Voinnet 2009). The biogenesis of miRNA begins with transcription of the miRNA gene by RNA polymerase II (Pol II), generating a primary transcript (pri-miRNA) that is both capped and polyadenylated. The pri-miRNA molecule is further folded into a stem-loop structure via intramolecular base-pairing (Xie et al. 2005) and cleaved into precursor-miRNA (pre-miRNA) by Dicer-like 1 (DCL1) (Kurihara and Watanabe 2004). Pre-miRNA is cleaved again to release a miRNA/miRNA* duplex that is loaded into cytoplasmic AGO proteins and transported out of the nucleus by HASTY (Park et al. 2005). The post-transcriptional RNA silencing is mediated through complementary binding of miRNA within RNA-induced silencing complex (RISC) to the mRNA, resulting in mRNA cleavage, translational inhibition, or mRNA decay (Axtell et al. 2011). The other strand is known as miRNA star (miRNA*) and in the majority of cases is degraded (Kawamata et al. 2011). Extensive studies have revealed that miRNAs play a vital role in organ development, proliferation, and abiotic and non-abiotic stress response in plants (Datta and Paul 2015; Voinnet 2009; Wang et al. 2007; Zhang 2015).

Given the significance of miRNAs, a large number of studies have been performed to identify miRNAs and their targets in various plant species over the past years (Datta and Paul 2015; Kozomara and Griffiths-Jones 2014; Zhang 2015). Barley (*H. vulgare*) is the world's fourth most important cereal and widely used for animal feed, beer, and human food (Haberer and Mayer 2015). Previous studies attempted to identified miRNAs associated with development and stress response in barley by computational strategies (Dryanova et al. 2008) or by sequencing the small RNA population (Curaba et al. 2012; Deng et al. 2015; Hackenberg et al. 2012, 2013; Kantar et al. 2010; Kruszka et al. 2013; Lv et al. 2012; Ozhuner et al. 2013; Schreiber et al. 2011). However, the miRNAs in wild barley are not well understood at present. There is no wild barley miRNA in miRBase database (version 21.0, June 2014) (Kozomara and Griffiths-Jones 2014) and only one found in plant microRNA database (PMRD: 17 November 2014) (Zhang et al. 2010). Furthermore, recent studies have shown that species-specific miRNAs and the spatio-temporal regulation of conserved miRNAs play important roles in shaping variation of morphological, developmental, and resistance to abiotic or non-abiotic stresses among related species during evolution (Datta and Paul 2015; Sun et al. 2014; Taylor et al. 2014). Therefore, the characterization of miRNAs in wild barley will contribute to future germplasm exploration and utilization of wild barley. In this study, we reported the first global analysis of miRNA expression in the wild barley based on the Illumina sequencing platform. The results from this study might provide insights into composition and abundance of miRNA in wild barley and pave the way to reveal the regulatory role of miRNAs in barley and beyond.

Materials and methods

Plant materials

The wild barley accession 3-25 collected from Mt. Gilboa, Israel, during an expedition trip in 2009 and stored in our laboratory, was grown in a growth chamber at a relative humidity of 75% and 26/20 °C day/night temperature with light intensity of 3000 lx (Lv et al. 2012). Leaves, stems, and roots at four different stages (seedling, jointing, heading, and filling), and spikes at heading and filling stages were collected to construct the small RNA sequencing library.

Small RNA library preparation and sequencing

Total RNA was extracted from the roots, stems, leaves, and spikes using TRIZOL Reagent (Invitrogen, USA) according to the manufacturer's instructions. The RNA quality was examined using gel electrophoresis (28S: 18S > 1.5) and Bioanalyzer (Agilent 2100, RIN ≥ 8.0). One RNA sample was generated by blending equal quantities of samples from different tissues. Then, this RNA sample was sent to Bioteke Corporation (Beijing, China) to construct a small RNA library and sequence on Illumina HiSeqTM 2000 platform, according to the manufacturer's protocol.

Bioinformatic analysis

The raw sequencing reads were pre-processed as described by Schreiber et al. (2011). In brief, after removing the poor-quality reads (reads with N, homopolymer runs longer than 6 bases, no more than 4 bases whose quality score was lower than 10, and no more than 6 bases whose quality score is lower than 13) and adaptor sequences, the sequences with length of 16–30 nt were retained for further analysis. Then, the small RNAs that were not miRNAs were removed as follows: (*i*) wild barley

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

chloroplast genomes (GenBank accession number: gi|521301089 and gi|521301008), allowing no mismatches and separated into chloroplast reads and non-chloroplast reads (nuclear-derived reads); (*ii*) diverse repeat databases TIGR Hordeum Repeats (Ouyang and Buell 2004), TIGR Oryza Repeats (Ouyang and Buell 2004), TIGR Triticum Repeats (Ouyang and Buell 2004), Triticeae repeat sequence database (TREP) (Wicker et al. 2002), and RepBase(Jurka et al. 2005), allowing no mismatches; (*iii*) plant tRNAs, plant snoRNAs, and Rfam database (Version 11.0) (Burge et al. 2013), allowing no mismatches except Rfam with two mismatches; and, (*iv*) wild barley and barley mRNA sequences, allowing two mismatches. After the reads were mapped to a given library, the mapped reads were removed and the remaining reads were used to map to another library.

The remaining clean sRNAs were subjected to miRNA identification using the miRDeep2 software (Friedlander et al. 2012). As the complete genome sequence of wild barley was unavailable, miRNA precursors were determined through mapping to the reference genomes of the barley cultivars Morex, Barke, and Bowman (Mayer et al. 2012) and Tibetan barley (Zeng et al. 2015). All of the miRNAs from barley (miRBase 21.0, June 2014; published studies) and five grass species (*Oryza sativa*, *Triticum aestivum*, *Zea mays*, *Brachypodium distachyon*, and *Aegilops tauschii*) were collected and used as two query data sets for miRDeep2 software, corresponding to 'mature_ref_this_species' and 'mature_ref_other_species', respectively, resulting in three types of results (known miRNA, homologous miRNA, and novel miRNA). The following key criteria were used for miRNA prediction: (*i*) the miRNA and miRNA* should be derived from the opposite stem-arms; (*ii*) there were four or fewer mismatches between the miRNA and other arm of the hairpin including the miRNA*; (*iii*) although some by-products of miRNA processing (such as length heterogeneity, sequential Dicer cleavage, and RNA editing) certainly exist in plants, there should be little heterogeneity in the sequences matching to the non-miRNA or non-miRNA* region of pre-miRNAs; (*iv*) the number of miRNA reads should be greater than 4 in the library; and, (*v*) the free energy value of precursor should be less than –20 Kcal/mol. Homologous miRNAs and novel miRNAs were used as query data sets to search against miRNAs from barley and miRBase for novel variant miRNA identification. Novel variant miRNAs should meet the following criteria as described by Sun et al. (2014): (*i*) miRNA does not perfectly match barley known miRNA (including miRBase 21.0 and published papers); and, (*ii*) miRNA with one or two mismatches to known miRNA from barley or five other grass species.

MiRNA validation by qRT-PCR

To validate the identified miRNAs, the stem-loop real-time RT-PCR was performed to investigate the expression of six randomly selected miRNAs using the previously described method (Varkonyi-Gasic et al. 2007). Total RNAs were extracted from three individual leaves of wild barley at the third leaf stage using TRIZOL Reagent (Invitrogen, USA). Then 1 μL DNase I-treated RNA and 1 μL stem-loop RT primer for each individual miRNA (1 μmol/L) were used in first-strand-strand cDNA synthesis with 0.25 μL SuperScript III Reverse Transcriptase (Invitrogen, 200 U/μL). Real-time RT-PCR was carried out using SYBR Premix Ex *Taq* (Takara) on the ABI PRISM 7300 Real-time PCR System using the following program: 3 min at 95 °C followed by 40 cycles of 5 s at 95 °C and 31 s at 60 °C. At the end of each PCR reaction, a melting curve was added to monitor the amplification specificity and only samples with correct annealing temperature were used for subsequent analysis. The qRT-PCR was performed in four replications. The primers are listed in Table S1[1] and the housekeeping gene 18s RNA was used as reference. Fn2

Results and discussion

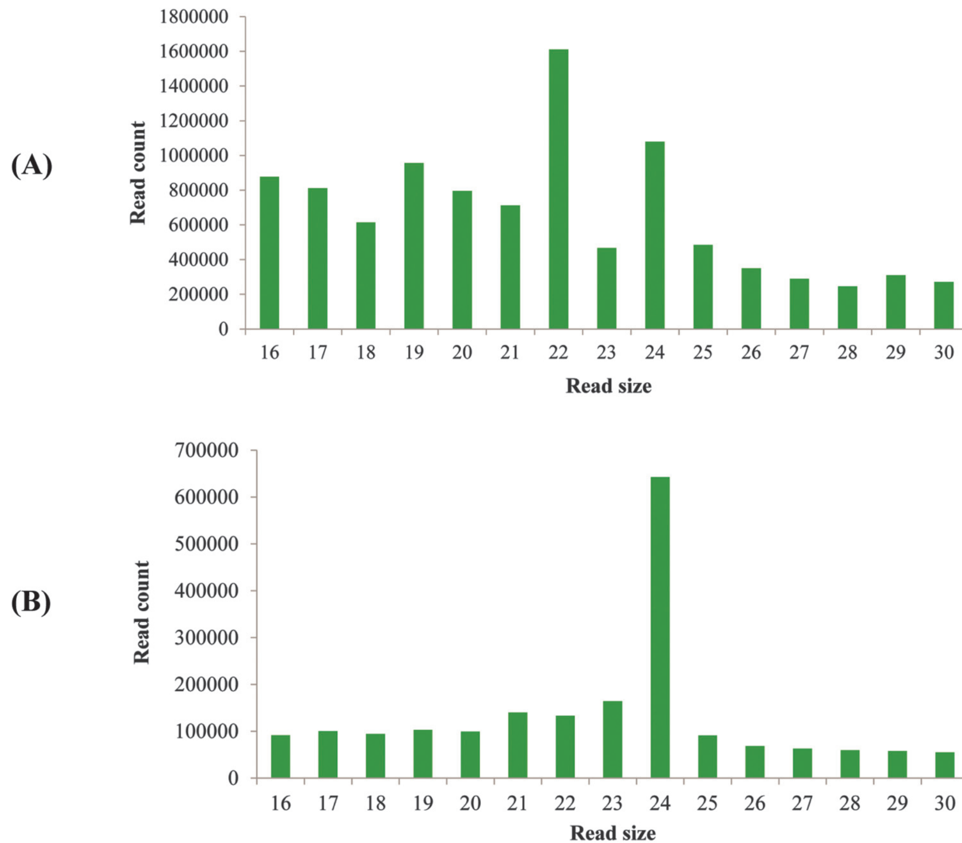
Deep sequencing results of wild barley

To identify miRNAs in wild barley, a small RNA library was constructed by pooling equal quantities of RNAs from different tissues and at various development stages and sequenced on the Illumina HiSeqTM 2000 platform. A total of 19 193 791 raw reads were obtained. After removal of low-quality reads, adaptor sequences, and sequences with lengths outside of 16–30 nt, 9 885 815 clean reads, representing 1 967 628 unique reads, remained (Table S2[1]). Tags removed from raw data mainly focused on low-quality reads, adaptor sequences, sequence lengths less than 16 nt, and sequence lengths greater than 30 nt, accounting for 7.35% (1 411 696), 0.00% (420), 26.62% (5 108 431), and 13.66% (2 622 257), respectively. The most abundant sequence length was 22 nt (Fig. 1A), F1 whereas among total small RNAs 24 nt was the most abundant (Fig. 1B). In general, in plants, size distributions of 24- and 21-nt sRNAs have been prevalent, the majority of which were associated with repeats and miRNAs, respectively (Curaba et al. 2012; Lv et al. 2012; Zabala et al. 2012). A previous study demonstrated that the 22 nt size class had a significant number of repetitions and function differently from the 21- and 24-nt sRNAs (Zabala et al. 2012). Such enrichment of 22-nt sRNAs in our study indicated a possible relevant role of the 22-nt sRNAs in wild barley. Similar results also were observed in wheat (Sun et al. 2014), maize (Zhang et al. 2009), and soybean (Zabala et al. 2012).

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2015-0224.

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

Fig. 1. Length, distribution, and abundance of wild barley small RNAs. (A) Unique reads and (B) total reads. [Colour online.]



A recent study found that many previously annotated miRNAs were the fragments of long noncoding RNAs such as tRNAs, rRNAs, and mRNAs (Taylor et al. 2014). Therefore, the small RNAs that were not miRNAs were removed before miRNA analysis by matching to a number of other sources such as chloroplast genome, repeats database, and Rfam database. First, all the trimmed reads were searched against the wild barley chloroplast genome sequences without mismatch and then separated into chloroplast-derived sRNAs (csRNAs) and nuclear-derived sRNAs. In total, of 3 712 610 reads were identified as csRNAs, accounting for 37.55% of the total reads, of which 84 495 was unique sequence (Fig. 2A). Among them, the csRNAs derived from rRNA (64.64%) were the most abundant, followed by chloroplast tRNA (19.88%), intergenic region (13.99%), and chloroplast gene (1.48%) (Fig. 2B). This result was similar to the result reported by the previous work on the origin of csRNAs in Chinese cabbage (Wang et al. 2011). Furthermore, we also matched all csRNAs with the chloroplast genome of cultivated barley. Our results displayed that more than 99% of the csRNAs from wild barley mapped perfectly to the chloroplast genome.

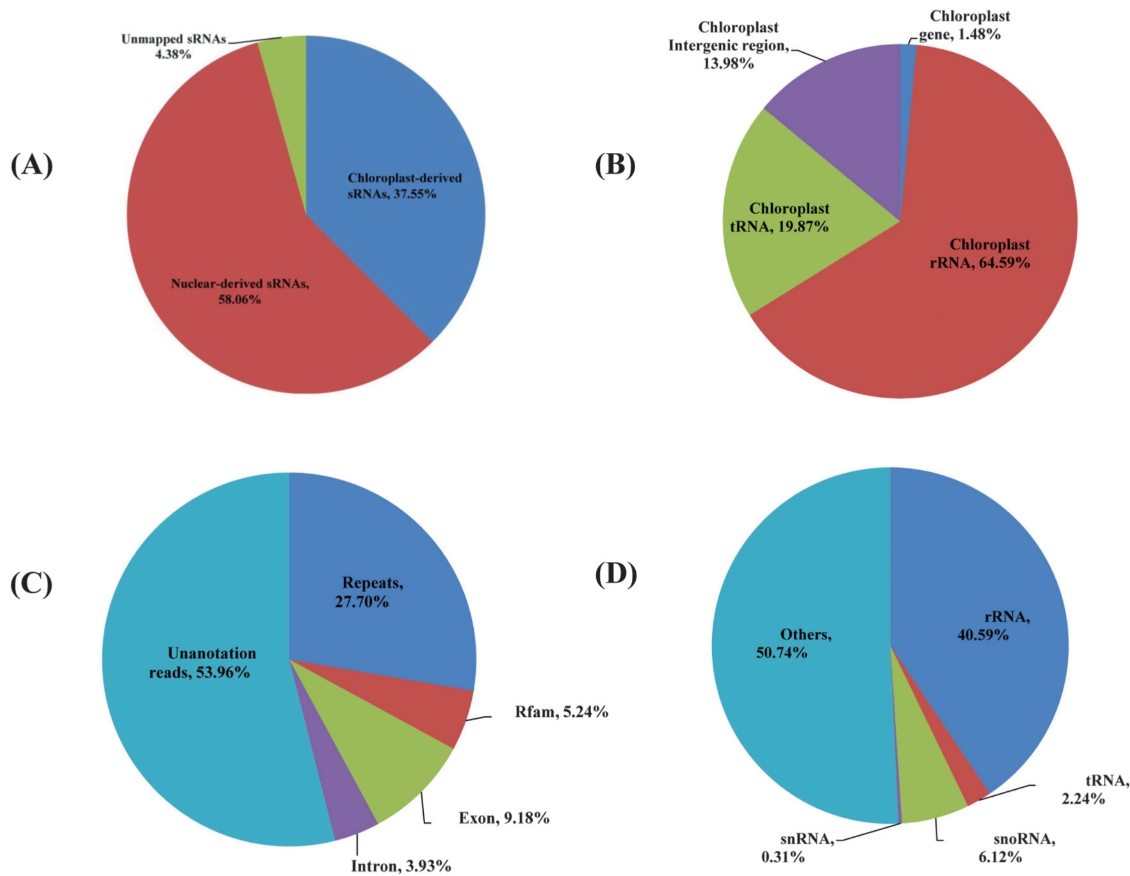
On the other hand, a total of 5 740 120 reads (58.06%), representing 1 615 447 unique reads, matched (no more than two mismatches) to the barley genome sequence (cultivar Morex) and can be considered as nuclear-derived sRNAs (Table S2[1]; Fig. 2A). Since the reference sequence of wild barley was not available, we mapped all nuclear-derived sRNA sequences to the reference genome of barley cultivar Morex for further genomic origin analysis. In this case, we found that more than 60% of the total nuclear-derived sRNAs could perfectly match to the reference genome, accounting for 40% of unique reads. When a mismatch threshold ≤2 was applied, more than 90% of the total nuclear-derived sRNAs could match to the reference genome and represented approximately 80% of unique reads. Non-miRNA reads were further removed by matching the tags against a collection of various plant repeat databases, other small RNA databases, and barley genes as described in Materials and methods. Finally, a total of 2 842 290 reads (521 381 unique reads) were removed, of which 1 709 799 and 323 356 tags were mapped to repeat elements and non-coding RNAs in Rfam, respectively (Fig. 2C). These non-coding RNAs included rRNAs (40.59%), snoRNAs (6.12%), tRNAs (2.24%), snRNAs (0.31%), and unclassified non-conding RNAs (50.74%) (Fig. 2D). After this cleanup, 53.96% of the total nucleotide-derived sRNAs were retained to identify conserved and novel miRNAs.

Identification of known and novel miRNAs in wild barley

Recently, around 271 barley miRNAs, belonging to 155 miRNA families, putatively associated with development and stress response have been identified by sequencing the small RNA population (Curaba et al. 2012; Deng

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

Fig. 2. Distribution of different classes of sRNAs derived from wild barley. (A) Distribution of sRNAs derived from wild barley chloroplast and genome. (B) Distribution of different classes of sRNAs derived from wild barley chloroplast. (C) Distribution of different classes of sRNAs derived from wild barley genome. (D) Distribution of different classes of sRNAs among plant Rfam database.



et al. 2015; Hackenberg et al. 2012, 2013, 2015; Kantar et al. 2010; Kruszka et al. 2013; Lv et al. 2012; Ozhuner et al. 2013; Schreiber et al. 2011; Zhou et al. 2015). We confirmed that 55 miRNAs from 35 families in wild barley shared the same sequences of miRNAs in cultivated barley (Table S3[1]). To find novel variants of known miRNAs, we searched small RNAs with one or two mismatches to known miRNAs in our small RNA library by homolog analysis, resulting in 15 novel variants belonging to 14 known miRNA families (Table S4[1]). In total, 60 known miRNAs, belonging to 38 miRNA families, were identified in wild barley. Among them, miR159 was the most abundantly expressed miRNA, with frequencies of 19 623 and followed by miR166, miR167, and miR156 (Table S4[1]). It has been demonstrated that these four miRNAs targeted MYB transcription factor (Reyes and Chua 2007), homeobox-leucine zipper (Zhu et al. 2011), auxin response factor (Wang et al. 2015), SPL (squamosa promoter-binding-like) transcription factor (Wang et al. 2009), respectively. The high expression of these miRNAs suggested that they play a critical role in wild barley growth and development. A total of seven miRNAs, including two known miRNAs (*hvs*-miR2023a and *hvs*-miR5049a) and five novel miRNAs (*hvs*-miR-n01, *hvs*-miR-n04, *hvs*-miR-n08, *hvs*-miR-n17, and *hvs*-miR-n18) expressed at a very low level (total abundance less than 12), suggesting they might play roles in specific temporal or spatial expression patterns.

Furthermore, among the 20 highly conserved miRNA families in plants (Jones-Rhoades et al. 2006), 16 were present among our reads (miR156, miR159, miR160, miR164, miR166, miR167, miR168, miR171, miR172, miR319, miR390, miR393, miR395, miR396, miR397, and miR399), while the homologues of miR162, miR169, miR394, and miR398 were not found in wild barley. The miR162 family is known to regulate DCL1 expression, which was required for miRNA biogenesis (Xie et al. 2003). However, miR162 has not been found in barley using deep sequencing technologies or precursors searching against barley ESTs that have been deposited in NCBI database. Barley genome analysis also revealed that no homologue could be found in its genome (Mayer et al. 2012), a feature also reported in wheat (Li et al. 2013). These observations indicate that further research is required to study the existence of miR162 in barley. miR169 (targeting NF-YA transcription factor), miR394 (targeting F-box protein), and miR398 (targeting Cu/Zn superoxide dismutase) were recently reported to be

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

Fig. 3. Pre-miRNA stem-loop structures for novel miRNAs identified in this study. Predicted mature miRNA sequences are shown in red.

hvs-n01

hvs-n02

hvs-n03

hvs-n04

hvs-n05

differentially regulated by stress conditions and play important roles in abiotic stress responses, such as high salinity, drought stress, and nitrogen starvation (Liu et al. 2008; Sunkar et al. 2006; Zhao et al. 2011), thus their absence was expected.

EST (expressed sequence tags) and GSS (genome survey sequences) provide a powerful alternative resource to identify miRNA when the complete genome sequence is not available. Therefore, all the wild barley EST and GSS sequences have been screened to identify new miRNAs, which showed low homology to known miRNA as described in Materials and methods. However, there was no miRNA meeting our prediction criteria although there were 18 miRNAs gained by the miRNA prediction pipeline. We further identified new miRNAs by searching the cultivated barley genome and ESTs to extend the information of miRNAs in wild barley. Finally, a total of 18 miRNAs were identified and 55.56% (10/18) of them had the expression of miRNA* (Table S4[1]). The potential mature miRNA sequences showed no similarity to any current entries in miRBase (21.0), indicating that they were novel miRNAs. The length distribution of novel miRNAs ranged from 20 to 24 nt, and the most abundant length was 21 nt (77.78%, 14/18). Details of all predicted miRNAs are given in Table S4[1], and part of their pre-miRNA secondary structures are showed in Fig. 3. To screen the barley specific miRNAs, we explored the presence of 18 novel barley miRNAs across Arabidopsis, soybean, rice, maize, sorghum, barley, and *Brachypodium*. Among them, seven miRNAs (*hvs*-n05, *hvs*-n09, *hvs*-n10, *hvs*-n11, *hvs*-n12, *hvs*-n15, and *hvs*-n18) were found to be barley specific, while the remaining 11 barley miRNAs were shared with wheat but not other species. Furthermore, most of the newly identified miRNAs were infrequently recovered by deep sequencing, with median read frequencies of only 180. In contrast, conserved miRNAs had higher expression counts (slightly over 1265 reads, on average), suggesting that non-conserved miRNAs might express in specific cells or growth conditions with generally low expression level in plants (Rajagopalan et al. 2006). In addition, six miRNAs were selected randomly to validate the results from the bioinformatics-based analysis using stem-loop real-time RT-PCR, with 18S rRNA as reference. Results showed that all the selected miRNAs could be detected by qRT-PCR, supporting their actual presence in wild barley (Fig. 4).

To characterize the nucleotide diversity of miRNA genes across wild and cultivated barley accessions, we examined all the precursors of miRNAs (53 known miRNAs, 14 novel variants, and 18 novel miRNAs) among the current genome assembly of cultivars (Barke, Bowman, Morex, and Tibetan barley). Alignment results suggested that they had highly conserved hairpin sequences, especially within the mature miRNA sequence (Fig. 5). Similar results have also been reported in maize (Zhang et al. 2009). This might be due to the important regulatory function of miRNAs, making them highly conserved across barley accessions. In addition, Lv et al. (2012) conducted a similar analysis of miRNA in cultivated barley (Morex) by pooling equal amounts of RNA from four different tissues. A set of 32 miRNAs (14 novel variants and 18 novel miRNAs) obtained in our study were used as database to search against a small RNA sequencing dataset from cultivated barley (NCBI accession: SRX122678) (Lv et al. 2012). The results indicated that 21 out of those miRNAs (65.63%, including 7 novel variants and 14 novel

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

Fig. 4. Quantitative real-time PCR analysis of miRNA identified in wild barley. The amount of expression was normalized to the level of 16s RNA. [Colour online.]

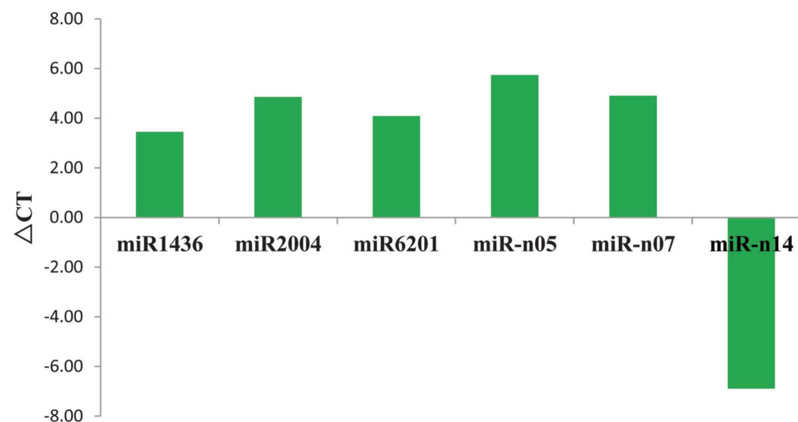
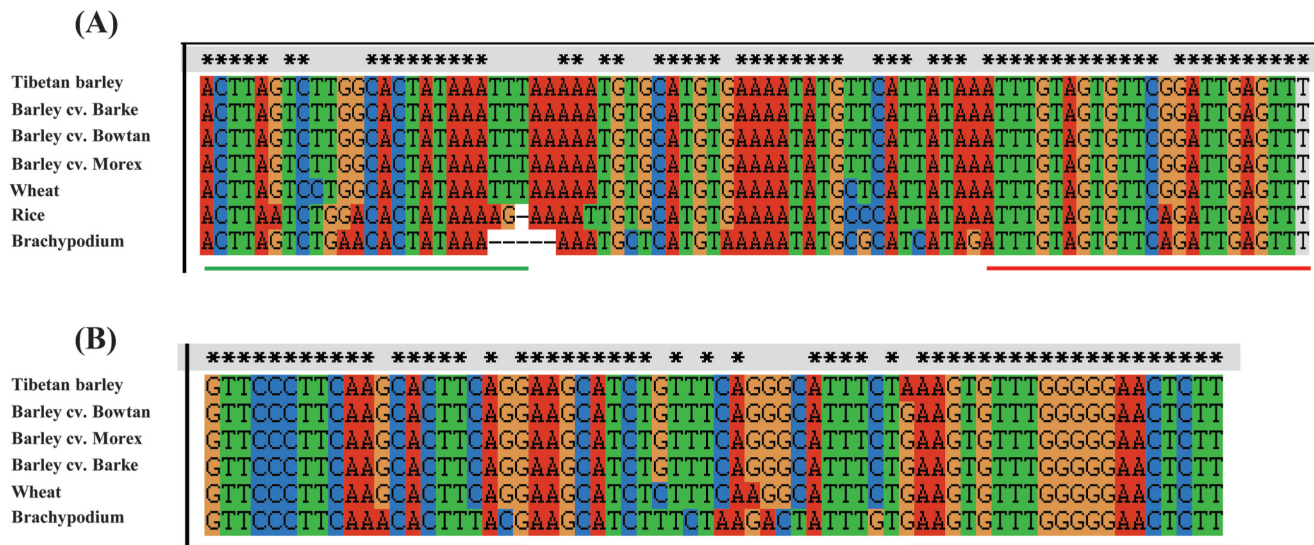


Fig. 5. Partial of syntenic sites harboring miRNA genes in barley and other plants. Shown are the syntenic sites harboring (A) *hvs*-miR1878 and (B) *hvs*-miR395c.

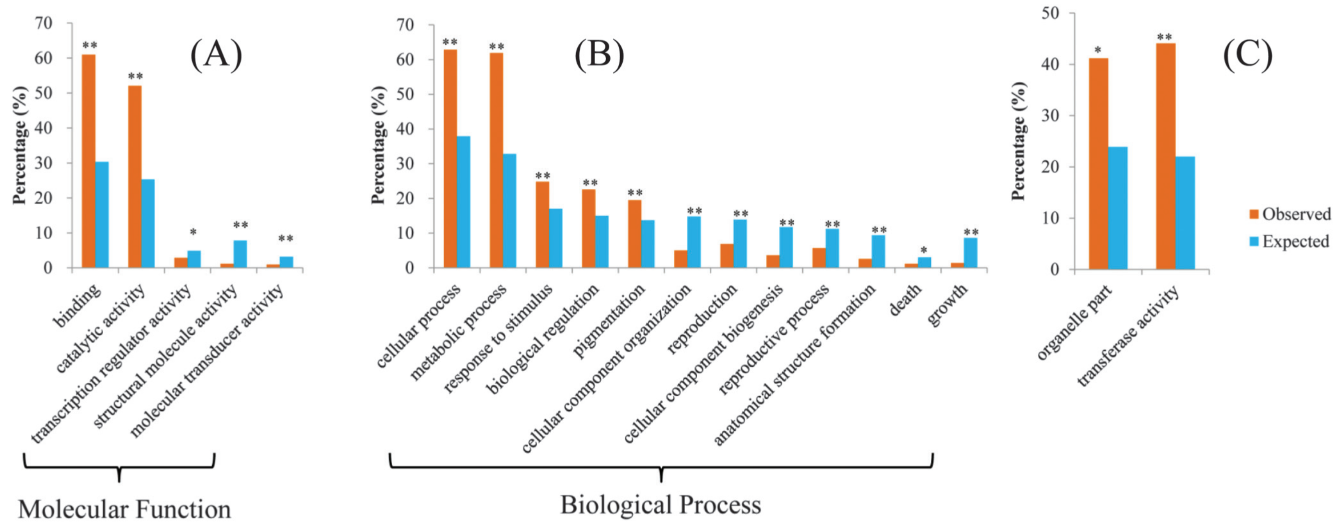


miRNAs) could be detected in cultivated barley (Table S4[1]). Furthermore, the majority of those miRNAs (95.24%, 20/21) displayed significantly different expression levels between those two materials, but it was not possible to associate function with tissue types as both of these two small RNA sequencing datasets were from pooling of RNAs from different tissues. Considering that most of the miRNAs could detect their precursors in different barley genomes, the role of miRNA contributing to phenotype variation between wild barley and cultivar barley might, in most cases, come from their different expression patterns rather than genomic constitution.

Target prediction for wild barley miRNAs

The targets of miRNA were predicted by searching against the HVGI database (Release 12) using psRNA target tool. A total of 708 unique putative genes were targeted by 73 miRNA, belonging to 52 families, with an average of 9.70 (708/73) target genes per miRNA (Table S5[1]). Of those target genes, 61.58% of them had a known function, whereas others were not annotated. We found that 12 out of 52 miRNA families were predicted to target transcription factors, such as SBP, MYB, and HD-ZIP, suggesting that these miRNA families play key roles in post-transcriptional regulation and transcription networks. Among them, eight miRNA-target pairs (*hvs*-miR156 and SBP; *hvs*-miR159 and *GAMyb*; *hvs*-miR160 and Auxin response factor; *hvs*-miR164 and *NAC* transcription factor; *hvs*-miR172 and *APETALA2*-like protein; *hvs*-miR396 and Growth-regulating factor; *hvs*-miR397 and auxin response factor; *hvs*-miR444 and MIKC-type MADS-box transcription factor) were highly conserved among plants. One miRNA-target pair (*hvs*-miR2024 and MIKC-type MADS-box transcription factor) was conserved in wheat (Sun et al. 2014), and the remaining three (*hvs*-miR2004 and WRKY transcription factor; *hvs*-n07 and *ERF* transcriptional factor; *hvs*-n14 and auxin response factor) have only been reported in wild barley. In addition to transcription factors, the predicted target

Fig. 6. Go enrichment analysis of miRNA target genes identified in wild barley. Shown are the significantly enriched GO terms of (A) molecular function and (B) biological process for all miRNAs identified in wild barley; (C) the significantly enriched GO terms for genes targeted by novel miRNAs and the new variants of known miRNAs in wild barley. Observed, numbers of genes observed in this study; Expected, numbers of genes in this same category in the GO enrichment analysis program.



genes were also involved in diverse physiological, metabolic processes, and stress defense. Such targets include pentatricopeptide (PPR) repeat-containing protein (*hvs*-miR1436), heat shock protein 80 (*hvs*-miR396a), superoxide dismutase (*hvs*-miR528), GL2-type homeobox genes (*hvs*-miR6201), cytochrome b (*hvs*-n06), acetolactate synthase (*hvs*-n10), CBF10B (*hvs*-n14), and auxin-responsive protein (*hvs*-n16).

Furthermore, all the putative target genes of wild barley miRNAs were subjected to Gene Ontology (GO) analysis to determine their GO terms, which could provide more information on understanding these miRNAs function (Table S6[1]). About 52.97% (375/708) of the target genes involved in 61 miRNAs (45 conserved miRNA and 16 novel miRNAs) could be annotated and classified into 326 biological processes and 355 molecular functions via Blast2GO tool. Within biological processes, cellular process (GO:0009987) and metabolic process (GO:0008152) were the most common annotations, according for 62.9% and 61.9%, respectively. In addition, there were at least 104 miRNA-target pairs (36 miRNAs and 96 unique genes) associated with stress response (GO:0050896), involving 28 conserved miRNAs (for example, *hvs*-miR159, *hvs*-miR160d, *hvs*-miR164d, and *hvs*-miR172a) and 8 novel miRNAs (for example, *hvs*-n01, *hvs*-n02, *hvs*-n06, and *hvs*-n07). Within biological processes, biological regulation (GO:0065007), pigmentation (GO:0043473), reproductive process (GO:0022414), growth (GO:0040007), and death (GO:0016265) were identified. Within molecular function, binding (GO:0005488) and catalytic activity (GO: 0003824) were the most represented, and at least 27 miRNA-target pairs were associated with regulator activities, such as transcription regulator activity (GO: 0030528), translation regulator activity (GO:0045182), and enzyme regulator activity (GO:0030234). GO function enrichment analysis showed that genes targeted by wild barley miRNAs were significantly enriched for five molecular function (Fig. 6A) and 12 biological process (Fig. 6B). In addition, a total of 11 miRNAs, including 7 novel miRNAs and 4 novel variants of known miRNAs, were not detected in cultivated barley (Table S4[1]). GO function enrichment analysis of these miRNAs displayed that they were significantly enriched for organelle part (GO:0044422) and transferase activity (GO:0016740) (Fig. 6C). Overall, the results of GO analysis revealed that miRNAs and their target genes were closely linked to wild barley development, growth, and stress response.

F6

Conclusion

To our knowledge, this is the first report on the large-scale investigation of small RNAs in wild barley utilizing the next-generation sequencing technique. A total of 90 miRNAs were identified in wild barley, comprising 72 conserved miRNAs and 18 novel miRNAs. To understand the possible function of these miRNAs, 718 unique putative genes were predicted as potential targets for 53 miRNA families. Gene Ontology (GO)-based functional classification showed that these miRNAs and their targets could play important roles in wild barley development, growth, and stress response. This work provided the base for further exploration of miRNAs from wild barley and also shed light on future practical utilization of miRNAs for barley improvement.

Acknowledgements

This research was mainly funded by the National Natural Science Foundation of China (Grant Nos. 31571647 and 31561143005).

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

References

Axtell, M.J., Westholm, J.O., and Lai, E.C. 2011. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. Genome Biol. **12**: 221. doi:10.1186/gb-2011-12-4-221. PMID:21554756.

Bartel, D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell, **116**: 281–297. doi:10.1016/S0092-8674(04)00045-5. PMID:14744438.

Burge, S.W., Daub, J., Eberhardt, R., Tate, J., Barquist, L., Nawrocki, E.P., et al. 2013. Rfam 11.0: 10 years of RNA families. Nucleic Acids Res. **41**: D226–D232. doi:10.1093/nar/gks1005. PMID:23125362.

Curaba, J., Spriggs, A., Taylor, J., Li, Z., and Helliwell, C. 2012. miRNA regulation in the early development of barley seed. BMC Plant Biol. **12**: 120. doi:10.1186/1471-2229-12-120. PMID: 22838835.

Dai, F., Nevo, E., Wu, D., Comadran, J., Zhou, M.X., Qiu, L., et al. 2012. Tibet is one of the centers of domestication of cultivated barley. Proc. Natl. Acad. Sci. U.S.A., **109**: 16969–16973. doi:10.1073/pnas.1215265109. PMID:23033493.

Datta, R., and Paul, S. 2015. Plant microRNAs: master regulator of gene expression mechanism. Cell Biol. Int. **39**: 1182–1190. doi:10.1002/cbin.10502.

Deng, P.C., Wang, L., Cui, L.C., Feng, K.W., Liu, F.Y., Du, X.H., et al. 2015. Global identification of microRNAs and their targets in barley under salinity stress. PloS ONE, **10**(9): e0137990. doi:10.1371/journal.pone.0137990. PMID:26372557.

Dryanova, A., Zakharov, A., and Gulick, P.J. 2008. Data mining for miRNAs and their targets in the Triticeae. Genome, **51**(6): 433–443. doi:10.1139/G08-025. PMID:18521122.

Ellis, R., Forster, B., Robinson, D., Handley, L., Gordon, D., Russell, J., and Powell, W. 2000. Wild barley: a source of genes for crop improvement in the 21st century? J. Exp. Bot. **51**: 9–17. doi:10.1093/jexbot/51.342.9. PMID:10938791.

Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., 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. PMID:17299599.

Fetch, J.T., Steffenson, B., and Nevo, E. 2003. Diversity and sources of multiple disease resistance in *Hordeum spontaneum*. Plant Dis. **87**: 1439–1448. doi:10.1094/PDIS.2003.87.12.1439.

Friedlander, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. 2012. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. **40**: 37–52. doi:10.1093/nar/gkr688. PMID:21911355.

Haberer, G., and Mayer, K.F. 2015. Barley: from brittle to stable harvest. Cell, **162**: 469–471. doi:10.1016/j.cell.2015.07.023. PMID:26232216.

Hackenberg, M., Shi, B.J., Gustafson, P., and Langridge, P. 2012. A transgenic transcription factor (TaDREB3) in barley affects the expression of microRNAs and other small non-coding RNAs PloS ONE, **7**: e42030. doi:10.1371/journal.pone.0042030.

Hackenberg, M., Huang, P.J., Huang, C.Y., Shi, B.J., Gustafson, P., and Langridge, P. 2013. A comprehensive expression profile of microRNAs and other classes of non-coding small RNAs in barley under phosphorous-deficient and -sufficient conditions. DNA Res. **20**(2): 109–125. doi:10.1093/dnares/dss037. PMID:23266877.

Hackenberg, M., Gustafson, P., Langridge, P., and Shi, B.J. 2015. Differential expression of microRNAs and other small RNAs in barley between water and drought conditions. Plant Biotechnol. J. **13**(1): 2–13. doi:10.1111/pbi.12220. PMID:24975557.

Hartmann, A., Kislev, M.E., and Weiss, E. 2006. How and when was wild wheat domesticated? Science, **313**: 296–297. doi:10.1126/science.313.5785.296b. PMID:16869032.

International Barley Genome Sequencing Consortium, Mayer, K.F., Waugh, R., Brown, J.W., Schulman, A., Langridge, P., et al. 2012. A physical, genetic and functional sequence assembly of the barley genome. Nature, **491**: 711–716. doi:10.1038/nature11543. PMID:23075845.

Jones-Rhoades, M.W., Bartel, D.P., and Bartel, B. 2006. MicroRNAs and their regulatory roles in plants. Annu. Rev. Plant Biol. **57**: 19–53. doi:10.1146/annurev.arplant.57.032905.105218. PMID:16669754.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. 2005. Repbase update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. **110**: 462–467. doi:10.1159/000084979. PMID: 16093699.

Kantar, M., Unver, T., and 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.

Kawamata, T., Yoda, M., and Tomari, Y. 2011. Multilayer checkpoints for microRNA authenticity during RISC assembly. EMBO Rep. **12**: 944–949. doi:10.1038/embor.2011.128. PMID: 21738221.

Kozomara, A., and Griffiths-Jones, S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. **42**: D68–D73. doi:10.1093/nar/gkt1181. PMID: 24275495.

Kruszka, K., Pacak, A., Swida-Barteczka, A., Stefaniak, A.K., Kaja, E., Sierocka, I., et al. 2013. Developmentally regulated expression and complex processing of barley pri-microRNAs. BMC Genomics, **14**: 34. doi:10.1186/1471-2164-14-34. PMID: 23324356.

Kurihara, Y., and Watanabe, Y. 2004. Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proc. Natl. Acad. Sci. U.S.A. **101**: 12753–12758. doi:10.1073/pnas.0403115101. PMID:15314213.

Li, Y.F., Zheng, Y., Jagadeeswaran, G., and Sunkar, R. 2013. Characterization of small RNAs and their target genes in wheat seedlings using sequencing-based approaches. Plant Sci. **203–204**: 17–24. doi:10.1016/j.plantsci.2012.12.014. PMID: 23415324.

Liu, H.H., Tian, X., Li, Y.J., Wu, C.A., and Zheng, C.C. 2008. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. RNA, **14**: 836–843. doi:10.1261/rna.895308. PMID:18356539.

Lv, S., Nie, X., Wang, L., Du, X., Biradar, S.S., Jia, X., and Weining, S. 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. PMID:22489137.

Ouyang, S., and Buell, C.R. 2004. The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. Nucleic Acids Res. **32**: D360–D363. doi:10.1093/nar/gkh099. PMID:14681434.

Ozhuner, E., Eldem, V., Ipek, A., Okay, S., Sakcali, S., Zhang, B.H., et al. 2013. Boron stress responsive microRNAs and their targets in barley. PloS ONE, **8**: e59543. doi:10.1371/journal.pone.0059543. PMID:23555702.

Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., and Poethig, R.S. 2005. Nuclear processing and export of microRNAs in Arabidopsis. Proc. Natl. Acad. Sci. U.S.A., **102**: 3691–3696. doi:10.1073/pnas.0405570102. PMID:15738428.

Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. 2006. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Gene. Dev. **20**: 3407–3425. doi:10.1101/gad.1476406. PMID:17182867.

Reyes, J.L., and Chua, N.H. 2007. ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis

rich2/cge-gen/cge-gen/cge99915/cge0295d15z | xppws | S=3 | 4/8/16 | 13:06 | 4/Color Fig: F1-F6 | Art: gen-2015-0224 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

seed germination. Plant J. **49**: 592–606. doi:10.1111/j.1365-313X.2006.02980.x. PMID:17217461.

Roy, J.K., Smith, K.P., Muehlbauer, G.J., Chao, S., Close, T.J., and Steffenson, B.J. 2010. Association mapping of spot blotch resistance in wild barley. Mol. Breeding, **26**: 243–256. doi:10.1007/s11032-010-9402-8.

Schreiber, A.W., Shi, B.J., Huang, C.Y., Langridge, P., and Baumann, U. 2011. Discovery of barley miRNAs through deep sequencing of short reads. BMC Genomics, **12**: 129. doi:10.1186/1471-2164-12-129. PMID:21352554.

Shakhatreh, Y., Haddad, N., Alrababah, M., Grando, S., and Ceccarelli, S. 2010. Phenotypic diversity in wild barley (*Hordeum vulgare* L. ssp. *spontaneum* (C. Koch) Thell.) accessions collected in Jordan. Genet. Res. Crop Evol. **57**: 131–146. doi:10.1007/s10722-009-9457-8.

Shen, Y., Lansky, E., Traber, M., and Nevo, E. 2013. Increases in both acute and chronic temperature potentiate tocotrienol concentrations in wild barley at 'Evolution Canyon'. Chem. Biodivers. **10**: 1696–1705. doi:10.1002/cbdv.201300133. PMID:24078602.

Steffenson, B.J., Olivera, P., Roy, J.K., Jin, Y., Smith, K.P., and Muehlbauer, G.J. 2007. A walk on the wild side: mining wild wheat and barley collections for rust resistance genes. Crop Pasture Sci. **58**: 532–544. doi:10.1071/AR07123.

Sun, F., Guo, G., Du, J., Guo, W., Peng, H., Ni, Z., et al. 2014. Whole-genome discovery of miRNAs and their targets in wheat (*Triticum aestivum* L.). BMC Plant Biol. **14**: 142. doi:10.1186/1471-2229-14-142. PMID:24885911.

Sunkar, R., Kapoor, A., and 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. PMID:16861386.

Taylor, R.S., Tarver, J.E., Hiscock, S.J., and Donoghue, P.C. 2014. Evolutionary history of plant microRNAs. Trends Plant Sci. **19**: 175–182. doi:10.1016/j.tplants.2013.11.008. PMID:24405820.

Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E.F., and Hellens, R.P. 2007. Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Methods, **3**: 12. doi:10.1186/1746-4811-3-12. PMID:17931426.

Voinnet, O. 2009. Origin, biogenesis, and activity of plant microRNAs. Cell, **136**: 669–687. doi:10.1016/j.cell.2009.01.046. PMID:19239888.

Wang, J.W., Czech, B., and Weigel, D. 2009. miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. Cell, **138**: 738–749. doi:10.1016/j.cell.2009.06.014. PMID:19703399.

Wang, L., Yu, X., Wang, H., Lu, Y.Z., de Ruiter, M., Prins, M., and He, Y.K. 2011*a*. A novel class of heat-responsive small RNAs derived from the chloroplast genome of Chinese cabbage (*Brassica rapa*). BMC Genomics, **12**: 289. doi:10.1186/1471-2164-12-289. PMID:21639890.

Wang, L.W., Liu, H.H., Li, D.T., and Chen, H.B. 2011*b*. Identification and characterization of maize microRNAs involved in the very early stage of seed germination. BMC genomics, **12**: 154. doi:10.1186/1471-2164-12-154. PMID:21414237.

Wang, Y., Stricker, H.M., Gou, D., and Liu, L. 2007. MicroRNA: past and present. Front. Biosci. **12**: 2316–2329. doi:10.2741/2234. PMID:17127242.

Wang, Y., Li, K., Chen, L., Zou, Y., Liu, H., Tian, Y., et al. 2015. microRNA167-directed regulation of the auxin response factors, GmARF8a and GmARF8b, is required for soybean (*Glycine max* L.) nodulation and lateral root development. Plant Physiol. **168**(3): 984–999. doi:10.1104/pp.15.00265. PMID:25941314.

Wicker, T., Matthews, D.E., and Keller, B. 2002. TREP: a database for Triticeae repetitive elements. Trends Plant Sci. **7**: 561–562. doi:10.1016/S1360-1385(02)02372-5.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat. Cell Biol. **11**: 228–234. doi:10.1038/ncb0309-228. PMID:19255566.

Xie, Z.X., Kasschau, K.D., and Carrington, J.C. 2003. Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. Curr. Biol. **13**: 784–789. doi:10.1016/S0960-9822(03)00281-1. PMID:12725739.

Xie, Z.X., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., and Carrington, J.C. 2005. Expression of Arabidopsis miRNA genes. Plant Physiol. **138**: 2145–2154. doi:10.1104/pp.105.062943. PMID:16040653.

Yang, Z., Zhang, T., Li, G., and Nevo, E. 2011. Adaptive microclimatic evolution of the dehydrin 6 gene in wild barley at "Evolution Canyon", Israel. Genetica, **139**: 1429–1438. doi:10.1007/s10709-012-9641-1. PMID:22415104.

Zabala, G., Campos, E., Varala, K.K., Bloomfield, S., Jones, S.I., Win, H., et al. 2012. Divergent patterns of endogenous small RNA populations from seed and vegetative tissues of *Glycine max*. BMC Plant Biol. **12**: 177. doi:10.1186/1471-2229-12-177. PMID:23031057.

Zeng, X., Long, H., Wang, Z., Zhao, S., Tang, Y., Huang, Z., et al. 2015. The draft genome of Tibetan hulless barley reveals adaptive patterns to the high stressful Tibetan Plateau. Proc. Natl. Acad. Sci. U.S.A., **112**: 1095–1100. doi:10.1073/pnas.1423628112. PMID:25583503.

Zhang, B. 2015. MicroRNA: a new target for improving plant tolerance to abiotic stress. J. Exp. Bot. **66**: 1749–1761. doi:10.1093/jxb/erv013. PMID:25697792.

Zhang, L., Chia, J.M., Kumari, S., Stein, J.C., Liu, Z., Narechania, A., et al. 2009*a*. A genome-wide characterization of microRNA genes in maize. PLoS Genet. **5**: e1000716. doi:10.1371/journal.pgen.1000716. PMID:19936050.

Zhang, Z., Yu, J., Li, D., Zhang, Z., Liu, F., Zhou, X., et al. 2010. PMRD: plant microRNA database. Nucleic Acids Res. **38**: D806–D813. doi:10.1093/nar/gkp818. PMID:19808935.

Zhao, M., Ding, H., Zhu, J.K., Zhang, F., and Li, W.X. 2011. Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. New Phytol. **190**: 906–915. doi:10.1111/j.1469-8137.2011.03647.x. PMID:21348874.

Zhou, J.P., Cheng, Y., Yin, M.Q., Yang, E.N., Gong, W.P., Liu, C., et al. 2015. Identification of novel miRNAs and miRNA expression profiling in wheat hybrid necrosis. PLoS ONE, **10**(2): e0117507. doi:10.1371/journal.pone.0117507. PMID:25706289.

Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., et al. 2011. Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem. Dev. Cell, **145**: 242–256. doi:10.1016/j.cell.2011.03.024.