CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Sydney eScholarship

Cloning and characterization of small RNAs from wheat (*Triticum aestivum* L.)

Sun Qixin, Yao Yingyin, Ni Zhongfu, Guo Ganggang Du Jinkun
Department of Plant Genetics and Breeding, China Agricultural University, Beijing, 100094, China

INTRODUCTION

Small RNAs with the length of 21 to 24 nucleotides function to silence gene expression by multiple mechanisms and are present in diverse eukaryotic organisms [1,2,3]. The two major kinds of small RNA, small interfering RNA (siRNA) and microRNA (miRNA), serve as specificity determinants of transcriptional gene silencing and post-transcriptional gene silencing [3,4]. So far, identification of miRNAs and siRNAs has been limited to a few model plant species such as *Arabidopsis,* rice and *Populus* whose genomes are sequenced. Wheat is one of the most important cereal crops worldwide. To date, only a few conserved miRNAs have been predicted in wheat [5] and no direct cloning and sequencing of small RNAs has been reported in wheat. In this study, by high throughput sequencing of a wheat small RNA library, we have identified 58 miRNAs belonging to 43 miRNA families. We also report the first identification and characterization of a large set of putative endogenous siRNA from wheat.

CLONING AND CHARACTERIZATION OF WHEAT MICRORNAS

In order to identify novel as well as conserved miRNAs in wheat, we generated one small RNA library ranging in size from 18 to 26 nucleotides using pooled RNA isolated from leaves, roots and spikes. Pyrosequencing of the wheat small RNA library was performed at 454 Company, and generated a total of 262,955 sequences. Analysis of these sequences resulted in the identification of 25,453 unique sequences ranging in size from 18 to 26 nucleotides in length. The remaining sequences either had low quality, or contained inserts smaller than 18 nt representing degraded RNA, or were without inserts, and were excluded from further analysis. The majority of the small RNAs were 20 to 24 nt in size, which is the typical size range for Dicer-derived products and the 21 nt size class was predominant.

Identification of conserved miRNAs in wheat: A total of 35 miRNAs belonging to 20 conserved miRNA families were identified [6]. These include miRNA156/157, miR159, miR160, miR164, miR165/166, miR167, miR168, miR169, miR170/171, miR172, miR319, miR390, miR393, miR396, miR397, miR399 and miR408 that are conserved in diverse plant species. We also found miR444 which is a monocot-specific miRNA. The number of times each miRNA is represented in the small RNA library could serve as an index for the estimation of the relative abundance of miRNAs. The large number of miRNA sequences generated in this study could allow us to determine the relative abundance of miRNAs in wheat. The frequency of miRNA family sequences varied from 2 (TamiR390, TamiR396, TamiR397, TamiR399) to 757 times (miR169), indicating that the expression of different miRNA families varied greatly. MiRNAs can be grouped into families based on sequence similarity. Sequence analysis further revealed that 9 conserved miRNA families were represented by more than one member in our library. MiR169 was represented by 5 members, miR156, miR165/166, miR167, miR170/171 and miR172 were represented by 3 members each, and miR159, miR319 and miR168 were represented by 2 members each in the library. High throughput sequencing of the small RNA libraries allowed us to identify the expression levels of each member within a family. Sequence analysis indicated that the relative abundance of certain members within miRNA family varied greatly. For instance, TamiR169b and TamiR169a were represented 365 and 171 times respectively, whereas the other 3 members (TamiR169m, Ta169n and Ta169o) appeared between 38 to 98 times (Fig. 1). These observations indicated that certain members within a miRNA family showed preferential expression which could be attributed to a high level of tissue-specific expression of these members.

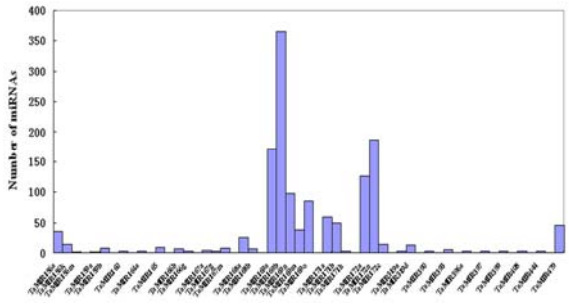

Fig. 1 The frequency of conserved miRNAs present in the sequenced small RNA library

Identification of new monocot-specific and wheat-specific microRNAs: One of the important features that distinguish miRNAs from other endogenous small RNAs is the ability of the miRNA surrounding sequences to adopt a hairpin structure. Since the wheat genome is largely unknown, we have to rely on wheat EST sequences for miRNA surrounding sequences in

predicting the hairpin structures. Our analysis for new miRNAs revealed that 23 sequences that perfectly matched with EST sequences were able to adopt hairpin structures and these formed 23 new miRNA families [6]. The length of these newly identified miRNAs varied from 19 to 24 nt. Ten of the 23 novel miRNAs begin with a 5' uridine, which is a characteristic feature of miRNAs. To determine whether these novel miRNAs are conserved among other plant species, we searched the nucleotide databases for homologs. This analysis indicated that four miRNAs, namely, TamiR506, TamiR510, TamiR514 and TamiR516 are conserved in other monocots such as rice, barley and *Festuca arundanacea*. Hairpin structures can be predicted for these miRNAs from rice, barley and *Festuca arundanacea* using miRNA surrounding sequences obtained from ESTs. These findings suggest that these 4 miRNAs are conserved in monocots and, as they are not present in Arabidopsis or *Populus*, that they are monocot-specific miRNAs.

Expression patterns of conserved and newly identified microRNAs in wheat: To get an insight into the possible stage- or tissue/organ-dependent roles of miRNAs in wheat, Northern blotting and semi-quantitative RT-PCR were performed to examine the expression patterns of 17 TamiRNAs (13 novel and 4 conserved miRNAs) in different tissues, including roots and leaves of seedling, nodal regions, spikes, internodes just below spike, and flag leaf of booting stage. The expression signal of the 4 conserved miRNAs was detected by Northern blotting, and the results indicated ubiquitous expression of TamiR171 in all tissues, whereas TamiR156, TamiR159 and TamiR164 displayed a differential tissue-specific expression pattern. Out of 13 novel wheat miRNAs tested, expression of seven (TamiR502, TamiR507, TamiR509, TamiR512, TamiR513, TamiR514 and TamiR515) could be detected by Northern blotting. Further analysis indicated that TamiR502 seemed to be strongly expressed in internodes, roots and leaves but barely detected in stems and spikes (Fig. 2). TamiR507 and TamiR509 had similar expression patterns;, abundant in roots, moderate in stems and internodes and weak in leaves, spikes and flag leaves. TamiR512 showed tissue-specific expression and was detected only in spikes. TamiR513 and TamiR514 were only expressed in roots. Expression of TamiR515 appeared to be restricted to roots and leaves (Fig.2). The expression of 4 novel wheat miRNAs (TamiR504, TamiR505, TamiR506 and TamiR508) was validated by semi-quantitative RT-PCR (Fig. 3) as these could not be detected using Northern analysis. TamiR505 and TamiR506 had low expression levels in spikes, and TamiR508 was uniformly expressed in stems, internodes and spikes but could not be detected in leaves and roots. TamiR504 showed ubiquitous expression in all of the tissues examined (Fig. 3).

Target predictions for wheat miRNAs: We adopted a set of rules proposed in earlier reports for predicting miRNA targets [7]. To identify potential targets for wheat miRNAs, we searched for the antisense hits in the wheat EST sequences and Unigene sequences, and 30 unigenes were predicted as putative targets for 20 conserved miRNAs [6]. As expected, these target genes were similar or related to the previously validated plant

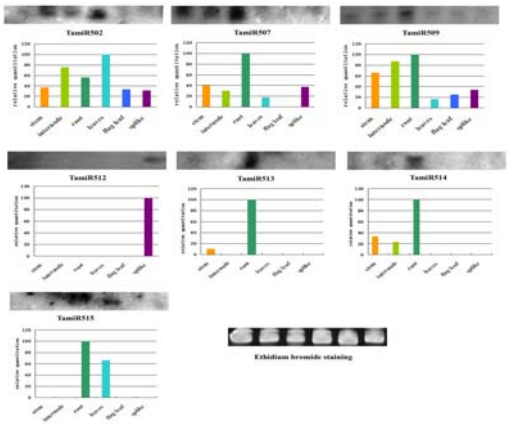

Fig. 2 Expression patterns of novel miRNAs in wheat.

RNA gel blots of low molecular weight RNA from different tissues, including stems, internodes below spikes, leaves, flag leaves, roots and spikes, were probed with labeled oligonucleotides. The tRNA and 5S RNA bands were visualized by ethidium bromide staining of polyacrylamide gels and served as loading controls.

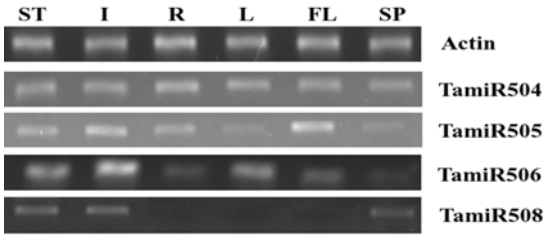

Fig. 3 Semi-quantitative RT-PCR analyses of novel miRNAs in wheat. Relative expression of miRNAs in stems, internodes below spikes, leaves, flag leaves, roots and spikes was analyzed by semi-quantitative RT-PCR. A wheat actin gene was selected to normalize the amount of template added in the PCR reactions. ST, stems; I, internodes below spikes; R, roots; L, leaves; FL, flag leaves; SP, spikes.

miRNA targets in Arabidopsis, rice and *Populus*. Twelve conserved miRNA families have been predicted to target 24 transcription factors. In addition, *TIR1*, *plantacyanin* and *argonaute* were predicated as targets of TamiR393, TamiR408 and TamiR168, respectively, which have been validated in Arabidopsis, rice and *Populus*. Sixteen unigenes were predicted to be putative targets for 12 newly identified miRNAs, among which 10 genes encode proteins with known function. Interestingly, TamiR506 is predicted to target AB182944 encoding *knox1b*, a transcription factor. We also predicted *CRT/DRE binding factor* as a putative target of TamiR507. These two genes have not been previously predicted as putative miRNA targets in any other plant species. In addition, we were unable to predict targets for 11 of the novel wheat miRNAs (miR501, miR503, miR508, miR510, miR511, miR515, miR516, miR517, miR518 miR520 and miR523) by applying the above rules, which could be due to the

limited number of wheat EST sequences available in the databases.

In conclusion, this study led to the discovery of 58 wheat miRNAs comprising 43 miRNA families, of which 20 and 23 belong to conserved and novel wheat miRNA families, respectively. Importantly, we have identified 4 monocot-specific miRNAs. We further show that some of the miRNAs are differentially expressed in tissue or in a developmental stage-dependent manner. In summary, this study provides the first large scale cloning and characterization of wheat miRNAs and their predicted targets, which will serve as a foundation for future functional studies.

CLONING AND CHARACTERIZATION OF WHEAT SIRNAS

In order to probe the diversity of small RNA in wheat, we adapt pyrosequencing performed by 454 company. Briefly, one small RNA library ranging in size from 18 to 26 nt using pooled RNA isolated from leaves, roots and spikes was generated and sequenced, and a total of 25,453 unique sequences ranging in length from 18 to 26 nucleotides was obtained. A total of 2076 putative endogenous siRNA were identified.

Cloning and size distribution of wheat endogenous siRNAs: Endogenous siRNAs in plants can be divided into two classes based on the size and function: the 21nt siRNAs that direct post-transcriptional silencing via mRNA degradation and the 24 nt siRNAs that trigger methylation of homologous DNA leading to TGS [3]. The endogenous siRNAs in wheat we obtained in this study could also be grouped into these size and functional classes. The majority (62%) of the wheat siRNAs were 21 to 24nt in size, which is the typical size range for Dicer-derived products. However, wheat siRNA with the length of 19 to 20nt were also abundant in this study. The mechanism of this pathway needs further investigation.

Location of siRNA in wheat transcripts and their predicted targets: The 2076 siRNAs were mapped to 12949 loci on wheat chromosomes, with an average of 6.23 loci per siRNA. The number of siRNAs that match one and two locations were 311 and 289, respectively, and the remaining (71.09%, 1476/2076) matched wheat ESTs from several to a few hundred times, with the highest being 950 matches. Further investigation revealed that a total of 9887 loci could be mapped to non-coding genes, 2626 were located in the CDS region of protein-coding genes, 360 corresponded to 5′ UTR region and 76 mapped to 3′ UTR region. We were able to predict 4249 trans targets for the 1106 siRNAs. These targets appear to play roles in a broad range of physiological processes. It has been reported that miRNA target sites in plants are usually found in the coding region of mRNA, and that most of the predicted trans targets of siRNA in rice also have a target site in the coding region. Similarly, the number of wheat endogenous siRNA matched to coding region exceeds the numbers matched to 3′ UTR and 5′ UTR regions. Interestingly, 803 wheat siRNAs were found to have more than two complementary sites and 176 had more than 5 complementary sites.

Expression of wheat endogenous siRNA: The number of times each siRNA was represented in the small RNA library could serve as an index for the relative abundance of siRNAs. In this study, we sequenced a total of 25,453 unique sequences ranging in size from 18 to 26 nt in length, and such a large number of siRNA sequences could allow us to determine the relative abundance of siRNAs in wheat. Among the 2076 siRNA sequences, 1571 (75.7%) siRNAs were represented only once, 40 siRNAs were present more than 10 times, and only one siRNA was found with a very high frequency (49 times) in the library, which indicates that most siRNA have a relatively low expression level in wheat as compared to miRNA [6].

In summary, this study provides the first large scale cloning and characterization of wheat siRNAs and their predicted targets, which could serve as a foundation for future functional studies. The identification of wheat siRNAs would also facilitate functional studies of siRNA in wheat, and help to unravel complex gene regulatory networks in eukaryotes.

ACKNOWLEDGEMENTS

This research is supported by National Basic Research Program of China (2007CB109000), 863 Project of China (2007A10Z100)

REFERENCES

1. Jones-Rhoades MW, Bartel DP, Bartel B: 2006, MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57:19-53.
2. Lu, C, Tej, SS, Luo, S, Haudenschild, CD, Meyers, B C, Green, PJ: 2005, Elucidation of the small RNA component of the transcriptome. Science, 309, 1567-1569.
3. Sunkar, R, Girke, T, Zhu, JK: 2005, Identification and characterization of endogenous small interfering RNAs from rice. Nucleic Acids Res., 33, 4443-4454.
4. Hamilton, AJ and Baulcombe, DC: 1999, A species of small antisense RNA in posttranscriptional gene silencing in plants. Science, 286, 950-952.
5. Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA: 2005, Identification and characterization of new plant microRNAs using EST analysis. Cell Res , 15:336-360.
6. Yao Yingyin, Ganggang Guo, Zhongfu Ni, Ramanjulu Sunkar, JinkunDu, Jian-Kang Zhu, Qixin Sun, 2007, Cloning and Characterization of MicroRNAs from wheat (Triticum aestivum L.), Genome Biology 8:R96
7. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D: 2005, Specific effects of microRNAs on the plant transcriptome. Dev Cell , 8:517-527.