Available online at www.sciencedirect.com
ScienceDirect
Agricultural Sciences in China
2010, 9(11): 1561-1570
November 2010

Genomic Analysis of MicroRNA Promoters and Their Cis-Acting Elements in Soybean

LIU Yong-xin[1], HAN Ying-peng[1], CHANG Wei[1], ZOU Quan[2], GUO Mao-zu[3] and LI Wen-bin[1]

[1] *Key Laboratory of Soybean Biology, Ministry of Education/Soybean Research Institute, Northeast Agricultural University, Harbin 150030, P.R.China*
[2] *School of Information Science and Technology, Xiamen University, Xiamen 361005, P.R.China*
[3] *Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, P.R.China*

Abstract

MicroRNAs (miRNAs) are derived from distinct loci in the genome and play crucial roles in RNA-mediated gene silencing mechanisms that regulate cellular processes during development and stress responses of plants. The miRNAs are approximately 21 nucleotides long and code for the complementary strand to a larger genic mRNA. They are often found within the complementary primary transcript (pri-miRNAs). In the past few years, a growing number of soybean miRNAs have been discovered, however, little is known about the transcriptional regulation of these miRNAs. In this study, promoters and cis-acting elements of soybean miRNAs were analyzed using the genomic data for the first time. A total of 82 miRNAs were located among 122 loci in genome, some were present as double or multiple copies. Five clusters that included ten miRNAs were found in genome, and only one cluster share the same promoter. A total of 191 promoters from 122 loci of the soybean miRNA sequences were found and further analyzed. The results indicated that the conserved soybean miRNA genes had a greater proportion of promoters than that of non-conserved ones, and the distribution of the transcript start sites (TSSs) and TATA-boxes found had different motif styles between conserved and non-conserved miRNA genes. Furthermore, the cis-acting elements 5′ of the TSSs were analyzed to obtain potential function and spatiotemporal expression pattern of miRNAs. The data obtained here may lead to the identification of specific sequences upstream of pre-miRNAs and the functional annotation of miRNAs in soybean.

Key words: miRNAs, soybean, promoter, cis-acting elements, miRNA clusters

INTRODUCTION

MicroRNAs (miRNAs) are endogenous single-stranded non-coding small RNAs with a length of about 21 bp that are encoded within genes either clustered or dispersed in the genome (Bartel 2004; Voinnet 2009). The miRMAs guide the RNA-induced silencing protein complex (RISC) to target sites. RISC negatively regulates post-transcriptionally the expression of genes. Numerous reports have demonstrated the importance of miRNA-mediated regulation in key processes, such as cellular proliferation, apoptosis, differentiation and development and pathogen-host interactions (Chen 2004; Jones-Rhoades and Bartel 2004; Laufs *et al*. 2004; Sunkar and Zhu 2004; He *et al*. 2007; Parker and Sheth 2007; Pillai *et al*. 2007; Liu *et al*. 2008; Carthew and Sontheimer 2009). Over-expressions or knockouts of miRNA genes disturb metabolisms and consequently result in the abnormal phenotypes (Palatnik *et al*. 2003;

Received 4 August, 2010 Accepted 13 October, 2010
LIU Yong-xin, MSc candidate, E-mail: woodcorpse@163.com; Correspondence LI Wen-bin, Professor, Tel: +86-451-55190778, Fax: +86-451-55103336, E-mail: wenbinli@neau.edu.cn

doi:10.1016/S1671-2927(09)60252-2

Guo *et al*. 2005).

The process of miRNA production in plants involves several steps (Bartel 2004; Kurihara and Watanabe 2004). First, the miRNA encoding gene is transcribed in the nucleus as a primary transcript (pri-miRNA), a longer sequence, including more than several hundred nucleotides, produced by PolII. Subsequently, the pri-miRNA is cleaved to an intermediate named miRNA precursor (pre-miRNA) with stem-loop structures. While the pre-miRNAs in animals are transported by exportin 5 from the nucleus into the cytoplasm (Yi *et al*. 2003; Lund *et al*. 2004), the plant miRNA precursors are processed by Dicer-like enzyme 1 (DCL1) in the nucleus (Tang *et al*. 2003; Kurihara and Watanabe 2004; Allen *et al*. 2005) and transported into the cytoplasm by HASTY, a plant ortholog of exportin 5 (Park *et al*. 2005). The single-stranded miRNAs associate with Argonaute (AGO) proteins and complementary mRNAs in a complex termed RNA-induced silencing complex (RISC). The complexes cause cleavage and reduce translation of the complementary mRNAs (Bartel 2004). The majority of plant miRNAs exist as independent transcription units that were encoded in intergenic regions (Lee and Ambros 2001; Jones-Rhoades and Bartel 2004). In contrast, animal miRNAs are predominantly found in intronic regions and may be transcribed as a part of the encompassing gene (Lee *et al*. 2004). Therefore, plant pri-miRNA transcription might differ from that of animals.

Extensive research has shown the existence of 2 043 plant miRNAs by 2010 (Release 14.0, http://www.mirbase.org/, 414 for *Oryza sativa*, 190 for *Arabidopsis thaliana*, 108 for *Medicago truncatula* and 85 for Glycine *max*). The total of 85 soybean miRNAs representing 51 families were reported by recent three papers (Zhang *et al*. 2005; Subramanian *et al*. 2008; Wang *et al*. 2009b). Of them, 34 soybean miRNAs, representing 29 families, had no homologous counterparts in other plant species. Therefore, it was inferred that non-conserved miRNAs in soybean might play species-specific roles in development and stressful responses.

The features of upstream sequences of miRNAs including promoters, transcription start sites (TSS), and diversity of the associated motifs and specific elements, were expected to be critical to regulation of the transcription of miRNA (Xie *et al*. 2005). Without such data understanding the location and range of pri-miRNAs, expression patterns of miRNAs, and miRNA-mediated regulation of pathways or other networks will be impossible. Recently, several studies have focused on the identification of miRNA promoters in plants (Xie *et al*. 2005; Zhou *et al*. 2007; Cui *et al*. 2009). Pri-miRNAs, typically were transcribed by RNA polymerase II. Consequently they have promoter elements, which are similar to those of protein-coding genes (Bartel 2004). It is known that the class II promoters have two parts: the core promoter and upstream element. The core promoter has at least two motifs: a TATA box beginning at approximate position -30 and an initiator centered on the transcription start site (TSS). The results of Allen *et al*. (2005) showed that most of miRNA promoters in *Arabidopsis* contained TATA-boxes in their core promoter regions. However, in animals there were exceptions. For example, promoter of miRNA gene mir-23a-27a-24-2 in human had no common elements like TATA-boxes or initiator elements required for initiating transcription (Lee *et al*. 2004). Such TATA-less promoters were often found within larger constitutively expressed genes (Smale 2001).

Xie *et al*. (2005) identified 63 transcription start site (TSS) in *Arabidopsis* by 5′-RACE amplification. Zhou *et al*. (2007) reported the promoters of intergenic miRNA regions in *Arabidopsis thaliana* and *Oryza sativa*. Both showed that most known miRNA genes had the same types of promoters as protein-coding genes. Cui *et al*. (2009) analyzed the surrounding regions of TSS, the TATA-box of miRNA transcripts and the organization of miRNA clusters in rice. Highly conserved rice miRNAs had a greater proportion of promoters than the non-conserved miRNAs. Presently, miRNA promoter predictions through bioinformatics have proved to efficiently identify regulatory elements underlying miRNA expression (Shahmuradov *et al*. 2005; Wang *et al*. 2006). However, relative to *Arabidopsis* and rice, little is known about promoters associated with miRNAs in soybean by early 2010. Therefore, the aim of this study is to identify: (i) specific sequences or motifs adjacent to pre-miRNAs in soybean by computational methodologies, and to compare highly conserved with non-conserved miRNAs; (ii) independent and co-transcribed miRNAs in order to explore the pattern of miRNA clusters associated with

the upstream specific promoter sequences of the pre-miRNAs; (iii) potential cis-acting element in upstream of the TSS and to infer potential functions from temporal or spatial expressions of the miRNAs.

MATERIALS AND METHODS

Genomic analysis and classification of the predicted soybean miRNA loci

All miRNA sequences of soybean from the miRBase database (ver.14.0, http://www.mirbase.org/, Griffiths-Jones *et al*. 2006), were downloaded. The genome sequence of soybean was downloaded from phytozome (http://www.phytozome.org; Schmutz *et al*. 2010). The predicted miRNA encoding regions in the soybean genome were identified through BLASTN 2.2.19 at identities ⩾97% and mismatch ⩽2 (Altschul *et al*. 1997). Predicted loci were drawn on the chromosomes through Mapchart 2.1 (Voorrips 2002). The soybean miRNAs were sorted into two categories, highly conserved and non-conserved miRNAs, according to the methods described by Cui *et al*. (2009). The miRNAs were also classified into another two categories; intergenic miRNAs (located between two protein-coding genes); and intronic miRNAs (each of them was predicted to overlap a single protein-coding gene) (Griffiths-Jones *et al*. 2006; Saini *et al*. 2007; Cui *et al*. 2009; Schmutz *et al*. 2010).

Potential promoter sequence identification for soybean miRNAs

Sequences in the intergenic regions of 5′ to the pre-miRNA were organized according to the method described previously by Zhou *et al*. (2007). If a pre-miRNA and the gene on the 5′ side were predicted to be transcribed in the same direction and the distance between them was larger than 2 400 bp, the 2 000 bp sequence of 5′ to the pre-miRNA was retrieved. Otherwise, the sequence between 400 bp downstream of the upstream gene and the precursor was used. Similarly, if a pre-miRNA and its upstream gene were predicted to be transcribed in the opposite direction and the distance between them was longer than 4 000 bp, the 2 000 bp sequence upstream of the precursor was obtained. Otherwise, 400 bp of sequence from the precursor and the middle point between the upstream gene and precursor was retrieved.

Prediction of specific sequence motifs 5′ to soybean miRNA genes

TSS and TATA-box like sequences were predicted using TSSP (http://mendel.cs.rhul.ac.uk/mendel.php?topic=gen). The predictions were obtained at the default thresholds of TSSP. The program had been trained and tested on independent sets of well-known promoters (Shahmuradov *et al*. 2005). TSSP of Softberry was adopted for supplemental predictions (http://www.Softberry.com, Shahmuradov *et al*. 2003; Solovyev and Shahmuradov 2003).

Prediction of cis-acting elements in the upstream of the TSS

The potential promoter regions (from TSS to 800 bp of upstream) were analyzed for the potential cis-acting elements and motifs. If an upstream 800 bp of TSS overlapped another TSS, only the sequence between the two TSSs would be used as the potential promoter region, to exclude redundancy. The analysis of cis-acting elements for the putative promoters was performed by PlantCARE database (http://intra.psb.ugent.be:8080/PlantCARE), a database of plant cis-acting regularly elements (Lescot *et al*. 2002).

RESULTS

The localization and clustering of soybean miRNA genes

A total of 82 pre-miRNAs of soybean (that represented 96.5% of the known pre-miRNAs for soybean) were located at 122 loci distributed among all 20 soybean chromosomes. The remaining 3.5% of the predicted pre-miRNAs (gma-MIR2107 copies) were mostly repetitive and/or located in the unanchored sequence scaffolds (Schmutz *et al*. 2010). Consequently, those predicted pre-miRNAs will not be discussed further. The

copy numbers of the predicted pre-miRNAs in genome were shown in Fig. 1. This figure showed that most of them had single copy in genome. The miRNA and the cluster distribution on genome were displayed in Fig. 2. This figure showed that the miRNAs more evenly distributed in genome. The potential clusters of miRNA genes in soybean genome, as well as their genomic organization within 10 kb regions were determined. As a result five gene clusters including ten miRNA genes were identified (Table).

Identification of the promoter regions for the miRNA genes

Upstream 2 000 bp of sequences 5′ to the pre-miRNAs were analyzed. A total of 122 candidate loci from the 82 predicted pre-miRNAs were used to predict promoter regions. There were 191 predicted promoters (Fig. 3). In contrast, seldom promoters were predicted among 122 random sequences with length of 2 000 bp among the regions which were not predicted to be transcribed as control.

The distribution of soybean pri-miRNAs in the genome was examined when the genes were divided into two groups (intergenic versus intronic and conserved versus non-conserved pri-miRNAs; Fig. 4-A, B). When the pri-miRNAs were classified into conserved and non-conserved groups, the numbers of promoters for the pri-miRNAs in each group were considerably different. The promoter numbers for the conserved pri-miRNAs were much higher than for the non-conserved pri-miRNAs. When the pri-miRNAs were classified into intergenic and intronic groups, the numbers of the promoters were relatively consistent in the two groups.

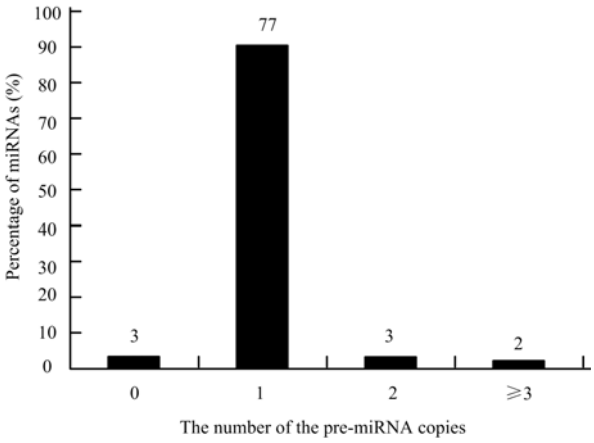

Fig. 1 The copy number of pre-miRNAs distributed in soybean genome.

Table Features of predicted soybean miRNA clusters on the genome

No.	Members in cluster	Chr	LG	Strand	Cluster length (bp)	Distance (bp)	Number of promoters
1	MIR2119, MIR398a	1	D1a	-	387	135	Double
2	MIR2107-9, MIR2107-10	8	A2	+	7 427	7 319	None
3	MIR2107-21, MIR2107-22	15	E	+	2 109	2 001	None
4	MIR2107-28, MIR2107-29	18	G	-	4 856	4 748	None
5	MIR2107-30, MIR2107-31	18	G	-	6 196	6 088	Single

Chr, chromosome number; LG, linkage group designation.

Analysis of the specific sequences in the upstream of pre-miRNA

A total of 191 TSSs were predicted within 2 000 bp upstream of the pre-miRNAs of the 82 miRNA genes. The distribution of the putative TSS positions in the corresponding pri-miRNA sequences was analyzed. The highest peak of predicted TSSs (16.2%) was found within 200 bp upstream of the pre-miRNAs. The weak peaks in the -600 to -800 bp and -1 400 to -1 600 region represented 12.0 and 11.5% of the total TSS predictions, respectively. The TSS numbers slightly tended to decrease from -1 to -2 000 bp (Fig. 5-A).

The TSSs of the conserved miRNA genes had two obvious peaks at the region of the pre-miRNA upstream (-1 to -200 bp and -600 to -800 bp). The TSS numbers tended to decrease gradually from -1 to -2000 bp. However, the TSSs of the non-conserved miRNA genes had the highest frequency peak at the -1 400 to -1600 bp of the pre-miRNA. Furthermore, 35.3% of TSSs were located in the region between -1 200 and -1 600 bp (Fig. 5-A). This phenomenon was not previously reported by 2010.

The distribution of TSSs among the intergenic and intronic miRNA genes were different (Fig. 5-B). The distribution of TSSs in the intronic miRNAs had a peak

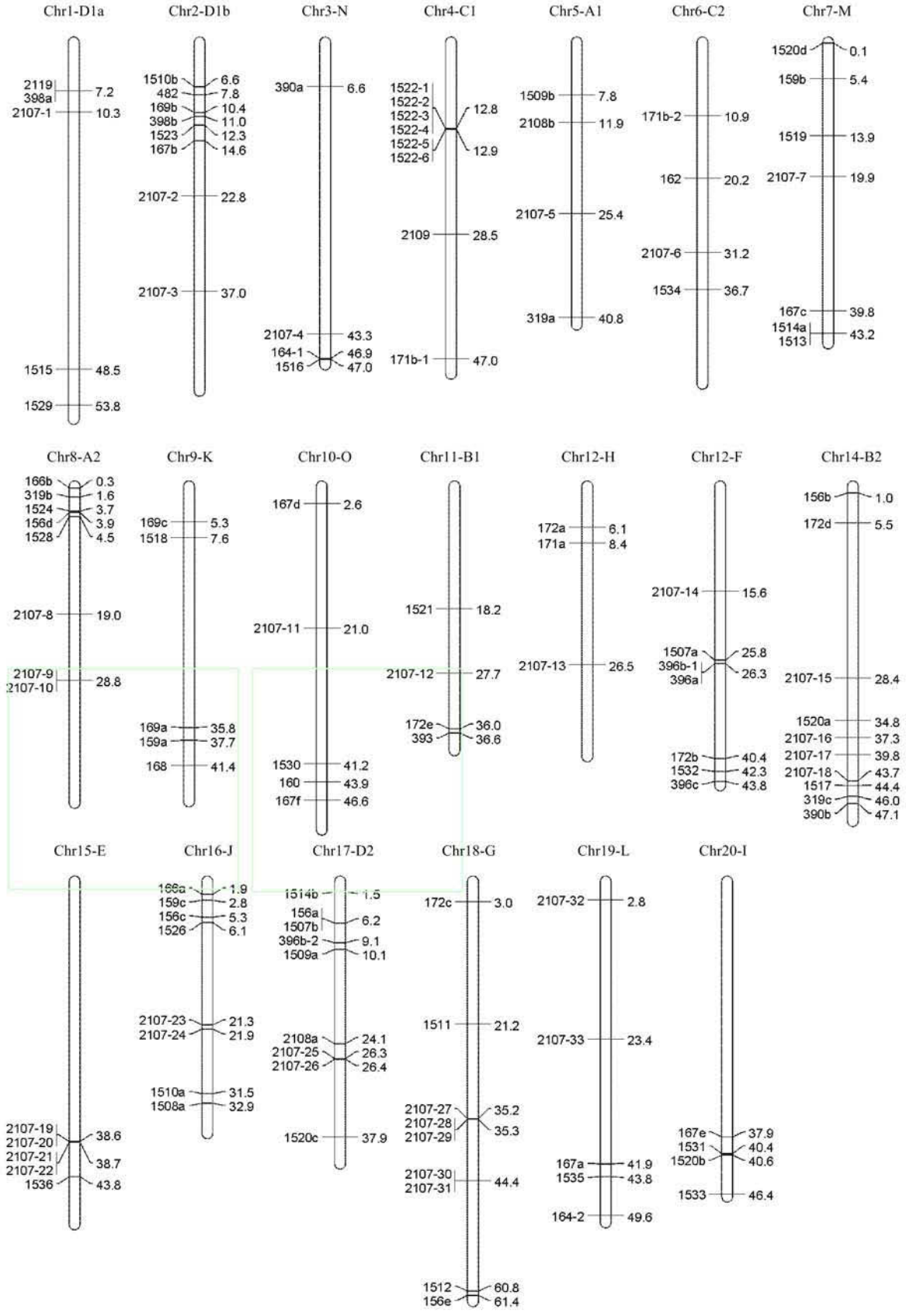

Fig. 2 The miRNA distribution on the soybean chromosomes showing clustering (positions are given in Mb).

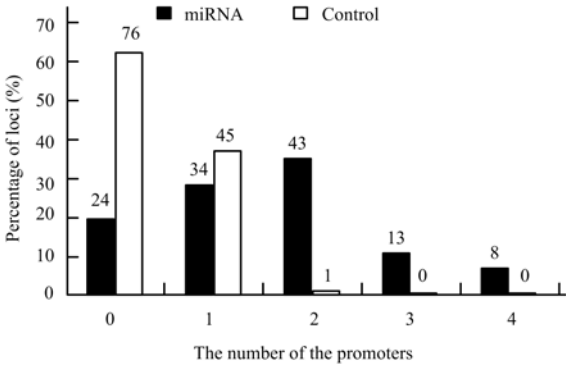

Fig. 3 The predictions of the promoter numbers in the upstream regions of pre-miRNAs compared to their random genome sequences as a control.

in their TSS distribution between the -800 and -1 000 bp region. They represented 15.7% of the total promoters for intronic miRNA genes. In contrast, only 5.7% of the promoters for intergenic miRNA genes were found within the same region.

The TATA-box is a well characterized promoter element. In this study, 182 TATA-boxes were found in 191 potential promoters. The other 9 promoters had no significant TATA-box and fell into the class termed TATA-less promoters, including MIR1509a, MIR1514, MIR1507a, MIR1523, MIR162, MIR393 and MIR2107. The frequency of the TATA-less promoters in non-conserved miRNA genes (7.4%) was two-fold higher than

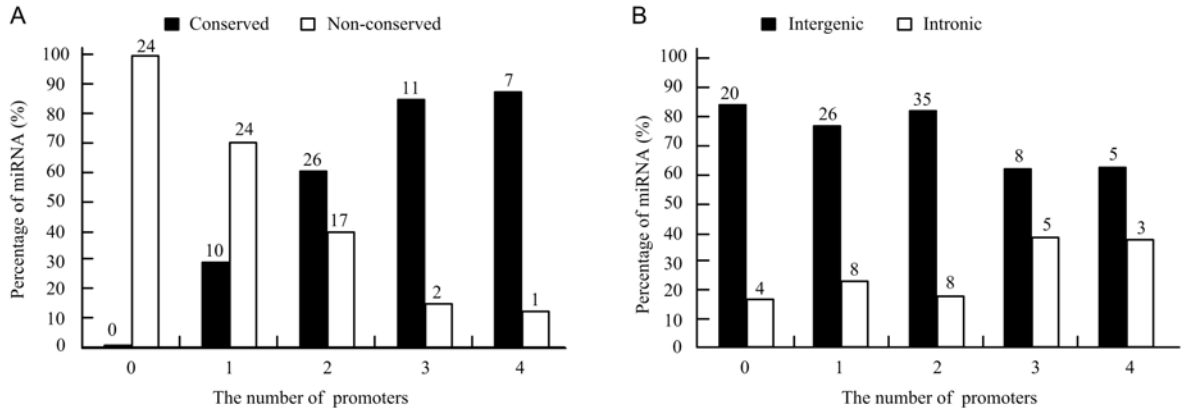

Fig. 4 Distribution of soybean miRNA with the same number of of promoters. A, The percentage of pri-miRNAs in conserved or non-conserved miRNAs. B, The percentage of pri-miRNAs located within the intergenic or intronic regions.

that of the conserved genes (3.3%, Fig. 5-C, D). The distribution of the TATA-boxes in the upstream regions of the pre-miRNA encoding genes showed the similar patterns to those of TSSs. They were mostly located at -35 bp relative to the associated TSS. However, some discrepancies were apparent. The highest peak of TATA-boxes for the total miRNAs occurred in the region from -1 400 to -1 600 bp. Three high density regions of TATA-boxes, included the -1 to -400 bp, -600 to -800 bp, and -1 400 to -1 600 bp regions. Together they contained 49% of the miRNAs gene promoters.

Analysis cis-acting elements in the potential promoters

Cis-acting elements, involved in gene expression, have been analyzed extensively (Zhou *et al*. 2007; Liu *et al*. 2008). To further elucidate the spatial and temporal expression pattern and function of the soybean genes encoding miRNAs, the sequences of 800 bp upstream of the potential TSS of 77 miRNA genes were determined through the PlantCARE analysis (Appendix A). Several known stress-responsive elements were identified including ABA (abscisic acid)-response elements (ABREs); anaerobic induction elements (AREs); MYB binding sites involved in drought inducement (MBSs); heat-stress-responsive elements (HSEs); low-temperature-responsive elements (LTRs); light-responsive elements (LREs); and defense stress-responsive elements (TC-rich repeats). Other regulatory elements, such as those possibly involved in the regulation of response to gibberellic acid (GA), ethylene, salicylic acid (SA), and methyl jasmonate (MeJA), were found.

Most ABA-responsive genes had the conserved

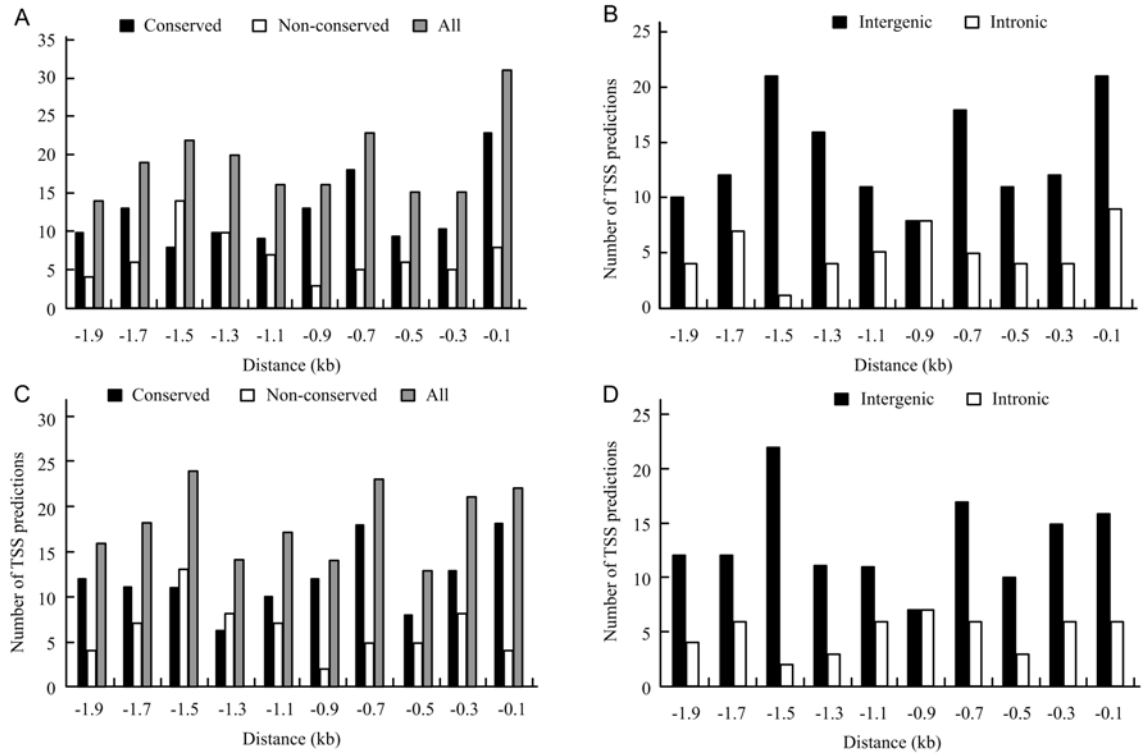

Fig. 5 Genomic distributions of TSSs and TATA-boxes in the upstream regions of the pre-miRNAs sequences. A, TSS predictions in the conserved and non-conserved miRNAs. B, TSS predictions from pri-miRNAs in intronic and intergenic regions. C, TATA-box predictions in the conserved and non-conserved miRNAs. D, TATA-box predictions in intronic and intergenic region.

ABREs in plant promoters, and ABREs were significant cis-elements associated with many genes responsive to abiotic stress in *Arabidopsis* and rice (Mundy *et al*. 1990; Ross and Shen 2006; Liu *et al*. 2008). In this study, ABREs were found in 38 miRNA encoding genes. More than one ABRE was found in 16 of the miRNA encoding genes. This result suggested many of the miRNA encoding genes might be important in ABA-mediated stress-response processes.

Additionally, AREs have been associated with genes responding to hypoxia, low-temperatures and dehydration stresses (Dolferus *et al*. 1994). AREs were found in their predicted promoter regions of 46 putative miRNA encoding genes (Appendix A). This result suggested most of the miRNA encoding genes might be important in hypoxia, low-temperatures and dehydration stress-response processes.

DISCUSSION

Less than two position mismatches on the chromosome were allowed in the analysis of putative miRNA encoding genes to eliminate sequencing errors and sequence differences among cultivars (Schmutz *et al*. 2010). However, three soybean miRNAs (MIR1525, MIR1527, and MIR1508b) could not be aligned with the chromosome. The unsuccessful anchoring of MIR1525 and MIR1527 to any chromosome might be due to the incomplete chromosomes or their gene sequences. The gene MIR1508b had the same genome position as MIR1508a, which might be the same MIR1508 genes from different varieties.

Animal miRNA genes were often clustered and co-transcribed as polycistronic RNAs (Kim 2005), whereas, plant miRNA genes were rarely arranged in tandem (Voinnet 2009). In the present work, five potential clusters of miRNA genes of soybean, that was different from the previously reported five clusters (Zhang *et al*. 2008), were identified (Table). The result of Zhang *et al*. (2008) was based on 69 predicted miRNAs. On the contary, our result was based on 85 verified miRNAs. The different miRNA sources caused the distinct result. Through further analysis of the upstream sequence regions, only one gene cluster identified in this study, MIR2119-398a, shared the same single TSS. The other

four clusters only included one miRNA, MIR2107, and were seldom found to have promoters, suggesting they were pseudogenes (Table).

Three pre-miRNA sequences were unsuitable for the promoter analysis in this study due to the failure to align with the chromosome sequences. Some of the pri-miRNA promoters could not be predicted, probably due to the following reasons: First, the exactly pri-miRNA was unknown, their promoters might be located more than 2 kb upstream of the pre-miRNAs. Second, most of the promoter prediction softwares used homology searches, hence, the pri-miRNA that did not have conserved promoters could not be predicted. Third, some of pre-miRNAs with multiple copies were pseudogenes and so truly did not have promoters.

The miRNA genes in plants and animals were found in diverse genomic locations. In animals, they were either encoded in inter geneic protein-non-coding regions, or within the introns of genes (Altuvia *et al*. 2005). In contrast, the majority of miRNA genes in plants occurred in the intergenic regions. The promoter distributions among intronic miRNA encoding genes was different from that of intergenic miRNAs. Intronic miRNAs were encoded within the transcripts of the host genes, although one third of them might be controlled by their own, rather than a host gene promoter (Martinez *et al*. 2008; Ozsolak *et al*. 2008). The intergenic miRNAs of soybean were located far away from any annotated genes. That, implied independent transcription from their own regulatory elements occurred. The distribution of the promoter motif peaks was similar to the total miRNAs since the intergenic miRNAs were the majority of the total miRNA in soybean. The intronic miRNAs that lie within introns of protein-coding genes and the associated TSSs in the region from -1 400 to -1 600 bp might co-express with the host genes (Lee *et al*. 2004).

TATA-less promoters tended to be found in two classes of genes. The first class was genes that were constitutively active in all cells. Such genes often have no known function or control the common biochemical pathways needed to sustain cellular life. The second class of genes with TATA-less promoters was developmentally regulated genes. Examples included the homeotic genes that control pattern and organ development and the genes that were active during the development of the immune system in mammals (Smale 2001). In this study, most of the TATA-less promoters were derived from non-conserved miRNAs and tended to the second class.

Seven AREs and two ABREs within the predicted promoter regions of three miR393 family members were found in *Arabidopsis* (Liu *et al*. 2008). In the present study, these motifs were not found in soybean miR393 family. However, cis-acting elements of AREs and ABREs were identified in most of the members of miR156 family. MiR156a to miR156e were probably the targets of squamosa-promoter binding protein (SBP) as the promoters contained SBP boxes. SBP was related with membrane protein transport activity and took part in the control of the metabolism of glucose, inorganic salts and ATP production (Rhoades *et al*. 2002; Zhang *et al*. 2008; Wang *et al*. 2009a). In this study, an interesting result is that cis-cating elements of AREs existing more than half of used miRNAs might play important role in anaerobic reaction. Another fact is that the majority of miRNA studied were derived from two papers, in which used soybean tissue was inoculated with rhizobium (Subramanian *et al*. 2008; Wang *et al*. 2009). This result seemed to demonstrate those miRNAs may play important role in nitrogen fixation.

Seven miRNAs of soybean that contained MSA-like cis-acting elements had functions in the cell cycle (Appendix A). Two of them (miR319a, b) were targeted to TCP genes that regulated gene expression in cycling cells (Tremousaygue *et al*. 2003). The miR159 family targeted the mRNAs of MYB proteins, binding to the promoter of the floral meristem identity gene LEAFY (Rhoades *et al*. 2002). The MBS was also found in each promoter of the miR159 family, from which was the inferred the negative feedback loop model. This finding supported the report that miRNAs might play a role in negative feedback loops that controlled their expression (Johnston *et al*. 2005; Megraw *et al*. 2006). The miR160 gene family in *Arabidopsis* was involved in feedback loops (Megraw *et al*. 2006). However, no auxin related regulatory element was found in the miR160 family in this study. Therefore, the apparently orthologous miRNA family judges by sequence similarity may have different regulatory patterns in different species.

In summary, a basic procedure to predict potential

function from the related motifs that were derived from potential promoters region of known soybean miRNAs, was described. A total of 191 promoters from 122 loci of the soybean pri-miRNA sequences were found and further analyzed. Previous reports of miRNAs functions were focused on miRNA targets. In the present study, cis-acting elements within the predicted miRNAs promoter regions were emphasized to provide a method to infer miRNA functions.

Acknowledgements

This study was financially supported by the National High-Tech R&D Program of China (863 Program, 2006AA100104-4), the Project of 948 from Ministry of Agriculture of China (2006-G5), the National Nature Science Foundation of China (30971810, 60932008), the National Basic Research Program of China (973 Program, 2009CB118400), the Postdoctoral Fund in Heilongjiang Province, China (LBH-Z07228), the Foundation Projects of Northeast Agricultural University, China, the Technology Project of Education Ministry of Heilongjiang Province, China (11541025), and the Technology Project of Harbin, China (2009RFQXN085).

Appendix A associated with this paper can be available on hppt://www.ChinaAgriSci.com/V2/En/appendix.htm.

References

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

Altschul S F, Madden T L, Schaffer A A, Zhang J, Zhang Z, Miller W, Lipman D J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389-3402.

Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein M J, Tuschl T, Margalit H. 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research*, **33**, 2697-2706.

Bartel D P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-297.

Carthew R W, Sontheimer E J. 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642-655.

Chen X. 2004. A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science*, **303**, 2022-2025.

Cui X, Xu S M, Mu D S, Yang Z M. 2009. Genomic analysis of rice microRNA promoters and clusters. *Gene*, **431**, 61-66.

Dolferus R, Jacobs M, Peacock W J, Dennis E S. 1994. Differential interactions of promoter elements in stress responses of the *Arabidopsis* Adh gene. *Plant Physiology*, **105**, 1075-1087.

Griffiths-Jones S, Grocock R J, van Dongen S, Bateman A, Enright A J. 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, **34**, 140-144.

Guo H S, Xie Q, Fei J F, Chua N H. 2005. MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *The Plant Cell*, **17**, 1376-1386.

He L, He X, Lowe S W, Hannon G J. 2007. MicroRNAs join the p53 network-another piece in the tumour-suppression puzzle. *Nature Reviews Cancer*, **7**, 819-822.

Johnston Jr R J, Chang S, Etchberger J F, Ortiz C O, Hobert O. 2005. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proceedings of the National Academy of Sciences of the USA*, **102**, 12449-12454.

Jones-Rhoades M W, Bartel D P. 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell*, **14**, 787-799.

Kim V N. 2005. MicroRNA biogenesis: coordinated cropping and dicing. *Nature Reviews Molecular Cell Biology*, **6**, 376-385.

Kurihara Y, Watanabe Y. 2004. *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proceedings of the National Academy of Sciences of the USA*, **101**, 12753-12758.

Laufs P, Peaucelle A, Morin H, Traas J. 2004. MicroRNA regulation of the CUC genes is required for boundary size control in *Arabidopsis* meristems. *Development*, **131**, 4311-4322.

Lee R C, Ambros V. 2001. An extensive class of small RNAs in caenorhabditis elegans. *Science*, **294**, 862-864.

Lee Y, Kim M, Han J, Yeom K H, Lee S, Baek S H, Kim V N. 2004. MicroRNA genes are transcribed by RNA polymerase II. *EMBO Journal*, **23**, 4051-4060.

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, van de Peer Y, Rouze P, Rombauts S. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, **30**, 325-327.

Liu H H, Tian X, Li Y J, Wu C A, Zheng C C. 2008. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA*, **14**, 836-843.

Lund E, Guttinger S, Calado A, Dahlberg J E, Kutay U. 2004. Nuclear export of microRNA precursors. *Science*, **303**, 95-98.

Martinez N J, Ow M C, Reece-Hoyes J S, Barrasa M I, Ambros V R, Walhout A J. 2008. Genome-scale spatiotemporal analysis of caenorhabditis elegans microRNA promoter activity. *Genome Research*, **18**, 2005-2015.

Megraw M, Baev V, Rusinov V, Jensen S T, Kalantidis K, Hatzigeorgiou A G. 2006. MicroRNA promoter element discovery in *Arabidopsis*. *RNA*, **12**, 1612-1619.

Mundy J, Yamaguchi-Shinozaki K, Chua N H. 1990. Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice rab gene. *Proceedings of the National Academy of Sciences of the USA*, **87**, 1406-1410.

Ozsolak F, Poling L L, Wang Z, Liu H, Liu X S, Roeder R G, Zhang X, Song J S, Fisher D E. 2008. Chromatin structure analyses identify miRNA promoters. *Genes & Development*, **22**, 3172-3183.

Palatnik J F, Allen E, Wu X, Schommer C, Schwab R, Carrington J C, Weigel D. 2003. Control of leaf morphogenesis by microRNAs. *Nature*, **425**, 257-263.

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig R S. 2005. Nuclear processing and export of microRNAs in Arabidopsis. *Proceedings of the National Academy of Sciences of the USA*, **102**, 3691-3696.

Parker R, Sheth U. 2007. P bodies and the control of mRNA translation and degradation. *Molecular Cell*, **25**, 635-646.

Pillai R S, Bhattacharyya S N, Filipowicz W. 2007. Repression of protein synthesis by miRNAs: how many mechanisms? *Trends in Cell Biology*, **17**, 118-126.

Rhoades M W, Reinhart B J, Lim L P, Burge CB, Bartel B, Bartel D P. 2002. Prediction of plant microRNA targets. *Cell*, **110**, 513-520.

Ross C, Shen Q J. 2006. Computational prediction and experimental verification of HVA1-like abscisic acid responsive promoters in rice (*Oryza sativa*). *Plant Molecular Biology*, **62**, 233-246.

Saini H K, Griffiths-Jones S, Enright A J. 2007. Genomic analysis of human microRNA transcripts. *Proceedings of the National Academy of Sciences of the USA*, **104**, 17719-17724.

Schmutz J, Cannon S B, Schlueter J, Ma J, Mitros T, Nelson W, Hyten D L, Song Q, Thelen J J, Cheng J, *et al*. 2010. Genome sequence of the palaeopolyploid soybean. *Nature*, **463**, 178-183.

Shahmuradov I A, Gammerman A J, Hancock J M, Bramley P M, Solovyev V V. 2003. PlantProm: a database of plant promoter sequences. *Nucleic Acids Research*, **31**, 114-117.

Shahmuradov I A, Solovyev V V, Gammerman A J. 2005. Plant promoter prediction with confidence estimation. *Nucleic Acids Research*, **33**, 1069-1076.

Smale S T. 2001. Core promoters: active contributors to combinatorial gene regulation. *Genes Development*, **15**, 2503-2508.

Solovyev V V, Shahmuradov I A. 2003. PromH: Promoters identification using orthologous genomic sequences. *Nucleic Acids Research*, **31**, 3540-3545.

Subramanian S, Fu Y, Sunkar R, Barbazuk W B, Zhu J K, Yu O. 2008. Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics*, **9**, 160.

Sunkar R, Zhu J K. 2004. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *The Plant Cell*, **16**, 2001-2019.

Tang G, Reinhart B J, Bartel D P, Zamore P D. 2003. A biochemical framework for RNA silencing in plants. *Genes Development*, **17**, 49-63.

Tremousaygue D, Garnier L, Bardet C, Dabos P, Herve C, Lescure B. 2003. Internal telomeric repeats and 'TCP domain' protein-binding sites co-operate to regulate gene expression in *Arabidopsis thaliana* cycling cells. *Plant Journal*, **33**, 957-966.

Voinnet O. 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell*, **136**, 669-687.

Voorrips R E. 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity,* **93**, 77-78.

Wang Y, Hindemitt T, Mayer K F. 2006. Significant sequence similarities in promoters and precursors of *Arabidopsis thaliana* non-conserved microRNAs. *Bioinformatics*, **22**, 2585-2589.

Wang Y, Hu Z, Yang Y, Chen X, Chen G. 2009a. Function annotation of an SBP-box gene in *Arabidopsis* based on analysis of co-expression networks and promoters. *International Journal of Molecular Sciences*, **10**, 116-132.

Wang Y, Li P, Cao X, Wang X, Zhang A, Li X. 2009b. Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochemical and Biophysical Research Communications*, **378**, 799-803.

Xie Z, Allen E, Fahlgren N, Calamar A, Givan S A, Carrington J C. 2005. Expression of *Arabidopsis* MIRNA genes. *Plant Physiology*, **138**, 2145-2154.

Yi R, Qin Y, Macara I G, Cullen B R. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Development*, **17**, 3011-3016.

Zhang B H, Pan X P, Wang Q L, Cobb G P, Anderson T A. 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Research*, **15**, 336-360.

Zhang B, Pan X, Stellwag E J. 2008. Identification of soybean microRNAs and their targets. *Planta*, **229**, 161-182.

Zhou X, Ruan J, Wang G, Zhang W. 2007. Characterization and identification of microRNA core promoters in four model species. PLoS Computational Biology, **3**, e37.

(Managing editor ZHANG Yi-min)