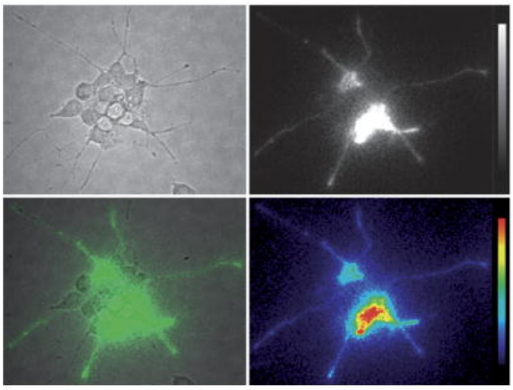
FEBS Letters

The journal for rapid publication
of short reports in molecular biosciences

- Reggie/flotillin membrane microdomains interact with actin
- SAMMA inhibits dendritic cell-mediate HIV transmission
- Real-time bioluminescence imaging in living cells



Published by Elsevier on behalf of the Federation of European Biochemical Societies

Submit to:
http://www.ees.elsevier.com/febsletters

ISSN 0014 5793
Volume 581 Number 24 2 October 2007

FEBS Letters 581 (2007) 4789–4793

Molecular evolution and selection of a gene encoding two tandem microRNAs in rice

Sheng Wang[a], Qian-Hao Zhu[b], Xingyi Guo[a], Yijie Gui[a], Jiandong Bao[a], Chris Helliwell[b], Longjiang Fan[a,*]

[a] *Institute of Crop Science, Zhejiang University, Hangzhou 310029, China*
[b] *CSIRO Plant Industry, Canberra, ACT 2601, Australia*

Received 25 June 2007; revised 31 August 2007; accepted 4 September 2007

Available online 12 Septemper 2007

Edited by Takashi Gojobori

Abstract **It has been shown that overexpression of *MIR156b/c* resulted in a bushy phenotype in maize and rice. Our results indicated that the *MIR156b/c* locus was highly conserved among cereals, but not in dicots and that genome duplication events played an important role in the evolution of the miR156 family. Genetic diversity investigation at the locus indicated that only ~9% of nucleotide diversity observed in wild rice (*O. rufigogon*) was maintained in the cultivated rice and the neutral model was rejected ($P < 0.05$) based on Tajima's *D* and Fu and Li's D^* and F^* tests. To our knowledge, this is the first example of miRNA gene to be targeted by both natural and domestication selection in plants.**

Keywords: MicroRNA; MIR156b/c; Selection; Genome duplication; *Oryza sativa*

1. Introduction

Small RNAs transcriptionally or posttranscriptionally regulate gene expression in eukaryotes. In plants, these small regulatory and noncoding RNAs are classified into microRNAs (miRNAs) and small interfering RNAs (siRNAs). miRNAs are ~21 base-pair (bp) long. They derive from hairpin structured miRNA precursors (pre-miRNAs) that are processed by an RNase III-like enzyme called DCL1 (DICER-LIKE 1) from primary miRNAs (pri-miRNAs) transcribed by RNA polymerase II [1]. miRNAs posttranscriptionally down-regulate gene expression by cleavage or translational repression of target mRNAs. In plants, miRNAs were first found in 2002 [2–5]. To date, 4584 miRNAs have been identified in animals and plants (miRBase, http://microrna.sanger.ac.uk/, Release 9.2). In rice, 242 miRNAs of 62 families have been identified by cloning or computational prediction [6–10].

The miR156 family was one of the first characterized miRNA families in plants [5,11], and has 12 members in rice [5,7]. It is highly conserved in the plant kingdom and has been identified in 45 different plant species [12]. miR156 has been demonstrated to target *SPL* genes, which are plant specific transcription factors containing an SBP box [6,13–15]. Recently, evolution of miRNA families with multiple members (such as miR156, miR169 and miR395) has been investigated in *Arabidopsis*. It was found that duplication events played an important role in diversification and evolution of these miRNA families [16]. Duplication was also one of the main mechanisms involved in the evolution of several rice miRNA families, such as the miR159 and miR395 families [17,18]. Of 12 miR156 family members located on six chromosomes, miR156b and miR156c (miR156b/c hereafter) are tandem miRNAs on chromosome 1. A full-length cDNA (AK110797) encodes both miRNAs. Overexpression of miR156b resulted in reduced plant high and increased number of tillers in rice [13]. In maize the *Corngrass1* (*Cg1*) gene also encodes tandem *MIR156* genes (*MIR156b* and *MIR156c*). The dominant *Cg1* mutant shows dwarfing, multiple-tillers and a bushy phenotype due to overexpression of miR156b/c [19].

Of the six domestication genes identified to date, a notable feature is that five of them encode transcription factors that regulate or target other genes by directly binding to their DNA [20]. Transcription factors, in turn, are major targets of miRNAs. For example, *tga1*, a transcription factor of the *SPL* family known to have had a role in the domestication of maize from teosinte, is one of the targets of *Cg1* [19].

In this study, we show that tandem and whole genome duplication are the driving force for expansion of the miR156 family in rice, and that genomic organization around the *MIR156b/c* locus is highly conserved in cereals. Sequencing the *MIR156b/c* locus in 30 cultivated rice cultivars and 15 wild rice accessions revealed that this locus experienced strong natural and potential domestication selection.

2. Materials and methods

2.1. Plant materials

Forty-five diverse *Oryza* accessions were selected from a wide range of geographical locations to represent a broad range of the genetic diversity within cultivated rice (*Oryza sativa*) and its wild ancestor, *O. rufipogon*. *O. sativa* includes 30 domesticated lines (15 *indica* and 15 *japonica* cultivars). Detailed information of the 45 rice accessions is listed in supplemental Table S1.

2.2. PCR and DNA sequencing

To investigate sequence variation in the *MIR156b/c* locus among the selected rice lines, a ~900bp genomic fragment that covers miR156b and miR156c was amplified from each cultivar/accession using the following pair of primers. Forward primer: 5′ TGGCTAGCTAATCCATGAGA 3′; reverse primer: 5′ TCAGAAATACTTCACAGAGAGAGTACG 3′. Primers were designed based on the genomic sequence

*Corresponding author. Fax: +86 571 86971117.
E-mail address: fanlj@zju.edu.cn (L. Fan).

0014-5793/$32.00
doi:10.1016/j.febslet.2007.09.002

of *japonica* cultivar Nipponbare using Primer3 [21]. The primers were compared to the rice genome sequence (NCBI GenBank) to ensure their specificity.

Genomic DNA was extracted from fresh rice leaves using a cetyltrimethylammonium bromide (CTAB) protocol [22]. PCR reactions were carried out on a thermocycler (Eppendorf) under the following conditions: 95 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 53 °C for 30 s and extension at 72 °C for 90 s, with a final extension at 72 °C for 10 min. PCR products were visualized on 0.8% agarose gel. A product of expected size was amplified from each sample. The amplified products were purified using a glassmilk PCR purification Kit (BioDev-Tech, China). For *O. sativa* cultivars, purified PCR products were directly sequenced on both strands using the forward or reverse primer. For *O. rufipogon*, in which either homozygous or heterozygous individuals might exist, PCR fragments were cloned into pGEM T-easy vectors (PMD19-T, Takara) and sequenced using the forward or reverse primer. At least three independent clones were sequenced. All sequences were deposited into GenBank, with GenBank accession numbers EU004233–EU004275.

2.3. Genome data

The TIGR rice pseudomolecules (release 5) and sorghum (*Sorghum bicolor*) genome segments were downloaded from http://www.tigr.org/tdb/e2k1/osa1/ and http://www.phytozome.net/sorghum, respectively.

2.4. Sequence analysis

MIR156 precursor sequences were aligned using MAFFT (version 5.8) [23] for the construction of phylogenetic trees. Neighbor-joining (NJ) phylogenies based on the Kimura 2-parameter distance matrix were generated by MEGA version 3.1 [24]. Bootstrap confidence values were obtained by 1000 replicates. Over 60% supporting notes were shown. Watterson's estimator of θ [25], the average pairwise nucleotide diversity π [26] and Tajima's D and Fu and Li's D^* and F^* tests of neutrality [27,28] were estimated using DNASP version 4.10.2 [29].

For rice syntenic analysis, the duplicated blocks among different chromosomes available in TIGR's Rice Genome Annotation (http://www.tigr.org/tdb/e2k1/osa1/) were used as guides, and the corresponding genomic sequences containing miR156 family members were downloaded from TIGR. Repeats were filtered using Repeatmasker (tigr_plant_repeat.lib, mips_REdat_4.3_rptmsk.lib, repbase) before the duplicated genomic sequence pairs were aligned using BLASTN. The hits with an E value less than 1e − 20 were used to generate syntenic figures. The same method was used to generate the synthenic relationship between rice and sorghum. The synteny between rice and maize was based on TIGR's syntenic map, in which the genetic map of maize cIBM 2005 was used.

3. Results and discussion

3.1. Evolution of MIR156b/c through duplication events

Twelve *MIR156* genes located on six chromosomes have been identified in rice. *MIR156b* and *MIR156c* are located close to each other on chromosome 1 and are within a single transcript AK110797 (Fig. 1A). *MIR156j* and *MIR156h*, which have the same mature miRNAs but different length precursors, are located at the same genomic region on chromosome 6 (Fig. 1A). Based on analysis of rice genomic sequences, a whole genome duplication event that occurred about 70 million years ago (mya) has been well documented [30–33], which predates the divergence of cereals (~50 mya; [34]) but postdates the divergence of monocots and dicots (~200 mya; [35]). Seventeen large duplicated blocks have been maintained [36]. We found that 11 of 12 *MIR156* loci (except *MIR156a*) are located in the duplicated regions (Fig. 1A). The collinear regions containing *MIR156* genes were further investigated in detail; a perfect syntenic relationship was found for four *MIR156* pairs (*MIR156b/c* and *MIR156l*, *MIR156i* and *MI156e*, *MIR156d* and *MIR156h/j*, and *MIR156f* and *MIR156g*) (Fig. 2A and Figure S1). Only *MIR156k* on chromosome 9 did not have a counterpart at its syntenic region on chromosome 8 (Figure S1). Phylogenetic analysis indicated the *MIRNA*s in each of these four pairs cluster together first (Fig. 1B), similar to the result obtained by Zhang et al. [12] even though not all members of the rice *MIR156* family were included in their analysis. These results suggest that whole genome duplication had a role in evolution of the *MIR156* family.

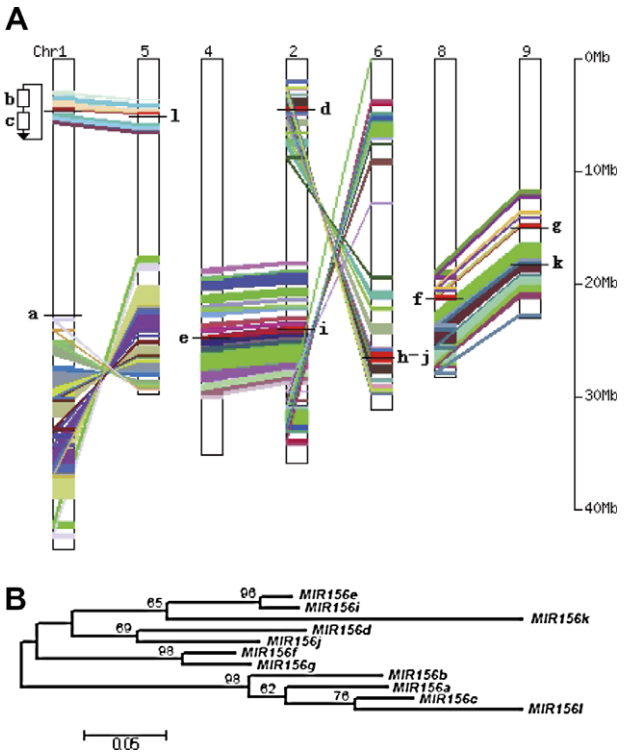


Fig. 1. Genomic localization and phylogenetic tree of the miR156 family in rice. (A) Distribution of miR156 family members and duplication events contributed to evolution of miR156 family in rice genome. a–k refer to miR156a-k (miR156h and miR156j locate at the same genomic position). The transcript of *MIR156b/c* containing miR156b and miR156c is shown, with the arrow indicating their transcription direction. Detailed alignments of duplicated segmental pairs containing miR156 members see Fig. 2 and Fig. S1; (B) Phylogenetic tree of miR156 family. *MIR156* precursor sequences were used to generate the NJ tree.

Phylogenetic analysis indicated that *MIR156b* or *MIR156c* was generated by a tandem duplication event, which occurred before the whole genome duplication event (~70 mya; Fig. 1B). In addition, our findings that the tandem *MIR156b/c* was discovered in sorghum and maize, and the genomic region containing *MIR156b/c* is highly conserved among these cereals (see the following section) also suggest that the tandem duplication event may have occurred before the divergence of cereals (~50 mya). In *Arabidopsis*, both *MIR156b* and *MIR156c* are located on chromosome 4 but are separated by many genes rather than in a tandem configuration [16]. Similarly the tandem *MIR156b/c* was also not found in the dicot plants *Populus trichocarpa* (http://genome. jgi-psf.org),*Medicago truncatula* (http://www.medicago.org) and *Solanum lycopersicum* (http://www.sgn.cornell.edu), for which genomes have been sequenced or are being sequenced. Based on these results, we believe that the tandem duplication of *MIR156b* or *MIR156c* occurred before the divergence of cereals but after the divergence of

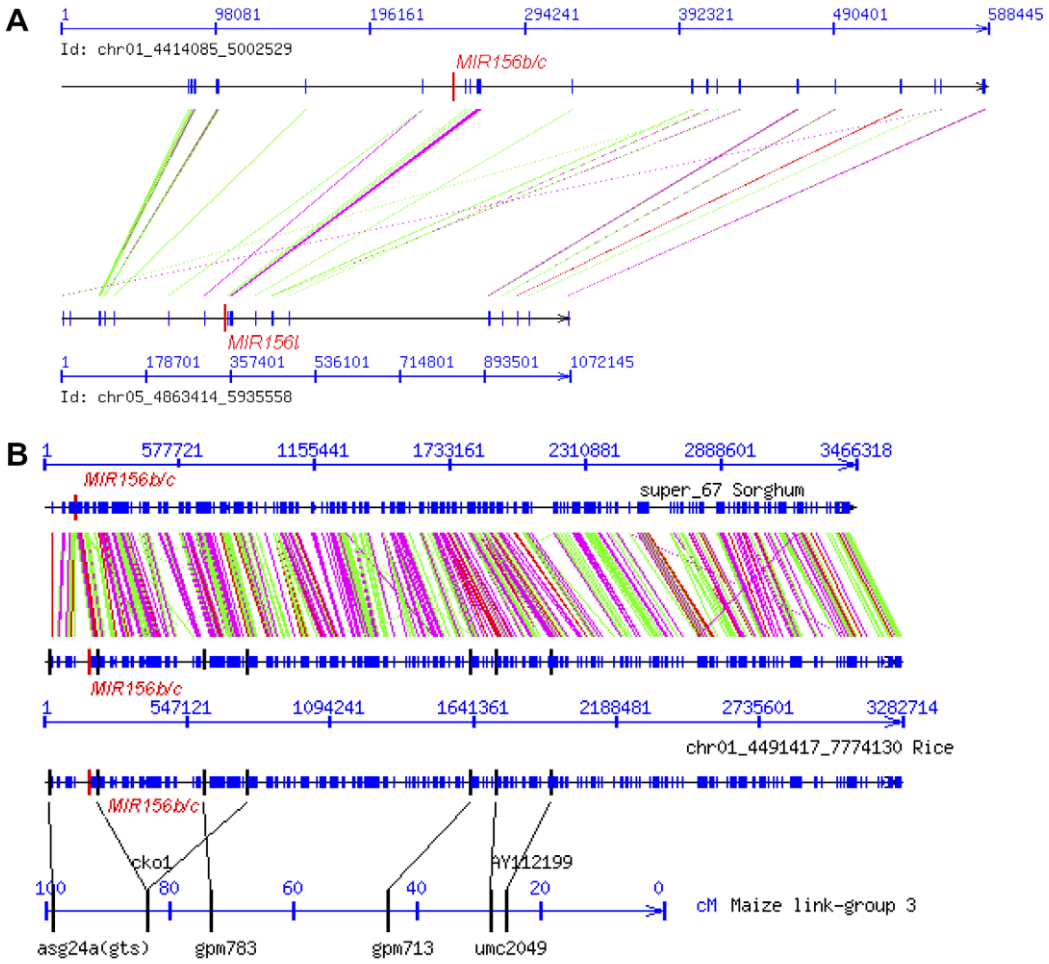


Fig. 2. Genomic synteny of the *MIR156b/c* locus. (A) Alignment between genomic segments containing *MIR156b/c* on chromosome 1 and its duplicated segment on chromosome 5 in rice. Genomic positions of duplicated segmental pairs and miR156 members (vertical red bars) are indicated; (B) Synteny of *MIR156b/c* among rice, maize and sorghum. Synteny between ~3.3 Mb genomic regions containing *MIR156b/c* in rice and sorghum is shown on the top. Maize genetic map (cIBM 2005) with *MIR156b/c* (or *Cg1* gene) mapped to the *cko1* locus is shown at the bottom.

monocots and dicots. Interestingly *MIR156l* is the only *MIR156* member in the syntenic region of *MIR156b/c* on chromosome 5 (Fig. 2A). To see whether there is another *MIR156* member near *MIR156l*, a ~1 Mb genomic syntenic region (TIGR chr5_4863414_5935558, see Fig. 2A) containing *MIR156l* was used for prediction of putative miRNAs using RNA folding software (http://www.tbi.univie.ac.at/~ivo/RNA/). No new putative *MIRNA* gene was found (data not shown), suggesting that the putative *MIRNA* corresponding to *MIR156b* or *MIR156c* on chromosome 5 might have been lost after the whole genome duplication event. Similarly, the paralog of *MIR156k* on chromosome 8 that arose from whole genome duplication was found to be lost (Fig. 1A). The miR156 family of *Arabidopsis* also experienced a large expansion via segmental duplication events and loss of family members [16]. In addition, tandem or segmental duplication events have been shown to have had a role in evolution of other miRNA families in rice, such as miR160, miR162, miR167, miR169, miR171 and miR395 [17,18]. Therefore, as for protein-coding genes, duplication and loss of duplicates may represent one of the main evolutionary routes for birth and death of *MIRNAs* in plants.

Our analysis showed that over half of the duplicated *MIR156* genes have survived after genome or segmental duplication in rice. The retained proportion of duplicated *MIR156* genes is significantly higher than that of protein-coding genes, for which less than 25% of the duplicates are thought to have survived following the genome duplication in rice [30]. This estimation is much lower (3.2–9.5%) if the initial numbers of genes in a duplicated region are considered [37]. The rice *SPL* family, targets of miR156, has also undergone dramatic expansion and has at least 19 members [13], and derived from the genome duplication event (data not shown). Identification and characterization of the functions of the *SPL* genes should shed light on the co-evolution relationship between the miR156 and the *SPL* gene family.

3.2. Highly conserved MIR156b/c among cereals

Large-scale genome sequencing and well documented genetic maps have made the investigation of synteny among cereals possible. A 3.3 Mb genomic alignment containing the *MIR156b/c* locus in rice, maize and sorghum is shown in Fig. 2B. A large sorghum genomic segment (Super_67, ~3.3 Mb) sequenced by the Sorghum Genome Project, DOE-JGI Community Sequencing Program (CSP, http://www.phytozome.net/sorghum), which contains the *MIR156b/c* locus, showed high synteny to the corresponding genomic region of

Table 1
Nucleotide polymorphisms and selection test

Group	Sample size (n)	π (×1000)	θ(×1000)	D	P	D^*	F^*
Cultivated rice	30	0. 43 (0. 16)	1.61 (0. 84)	−2.01*	0.000	−3.24**	−3.34**
Wild species	15	4.81 (0.87)	9.66 (3.92)	−2.09*	0.004	−2.85**	−3.04**

Note: π, the average pairwise nucleotide diversity [20]; θ, Watterson's estimator [19]; D, Tajima's D [21]; D^* and F^*, Fu and Li's D^* and F^* [22]; P, Coalescent simulations. Detailed accessions for two groups see Table S1. $^*P < 0.05$ and $^{**}P < 0.02$.

rice. Maize molecular markers around *Cg1* (or *MIR156b/c*) were also found within the genomic region containing the *MIR156b/c* locus in rice. Genomic and mRNA sequences of *MIR156b/c* have been determined for these three cereals and are available in public databases (GenBank and TIGR). High sequence similarity of the gene among rice, maize and sorghum was observed (Fig. S2). The two highest conserved regions are the stem-loop sequences of *MIR156b* and *MIR156c*. The sequences of the two mature miRNAs are identical in three cereals. However the spacer size between the two miRNA precursors differs, being ~200 bp in maize and ~400 bp in rice. Simple repeats are the major contributor to the difference. These results indicate that the *MIR156b/c* locus is highly conserved at least in rice, maize and sorghum and perhaps in all cereals. The function of *MIR156b/c* also appears to be conserved in rice and maize as the phenotypic changes observed in transgenic plants overexpressing miR156b (rice; [13]) or miR156b/c (maize; [19]) are similar.

MIR156b/c are encoded by a single transcript (full-length cDNA AK110797), suggesting they are co-expressed. However, the mature miR156b and miR156c sequences are identical so it is not possible to determine if both miRNAs are expressed. Co-transcription of *MIRNAs* is not unique to *MIR156b/c*. For example, 24 *MIR395* genes are organized into four compact clusters, each transcribed as a single transcript in rice [17]. Co-transcription of similar or identical *MIRNA* genes might have a dosage effect in plants. In contrast, animal *MIRNA* genes in a cluster usually are not homologs although they are evolutionarily related, such as those in the *miR-17* gene cluster. These non-homologous *MIRNA* genes could regulate multiple functionally related genes simultaneously [38].

3.3. Molecular diversity and selection of MIR156b/c

The conserved function of *MIR156b/c* in maize and rice led us to investigate whether selection forces have acted on it during the origin and domestication of rice [13,19]. To answer this question, genomic sequences of the *MIR156b/c* locus were amplified from 30 rice cultivars and 15 wild rice accessions. Several methods were used to examine whether the nucleotide polymorphism observed in the *MIR156b/c* locus fits the neutral model. For loci selected only by domestication and genetic improvement, a significant selection signal should be detected in cultivated rice but not in wild rice. In contrast, for the loci that are important for both cultivated and wild rice, a selection signal should be detected in both populations [39]. Our results indicated that the *MIR156b/c* locus experienced strong natural selection in *O. rufipogon*, and natural and/or domestication selection in the cultivated rice because the neutral model was rejected ($P < 0.05$) based on significant parameters of Tajima's D test, and Fu and Li's D^* and F^* test (Table 1). Coalescent Simulations of Tajima's D test also supported this result. In view of the importance of the *MIR156b/c* locus in determination of plant appearance in rice and maize and the fact that *O. rufipogon* has a similar appearance to that of cultivated rice [40], it is no surprise to see strong natural selection on the *MIR156b/c* locus in *O. rufipogon*. To know which region of the *MIRNA* gene is the selection target, Tajima's D test was carried out across the *MIR156b/c* locus using sequences amplified from 15 accessions of *O. rufipogon*. The test indicated that a 100–150 bp region (including mature miR156) upstream of the mature miR156b and miR156c in each of their precursors experienced stronger selection. This result implies the importance of mature miRNA sequences and their 5′ flanking regions in miRNA biogenesis and function. Nucleotide diversity of the *MIR156b/c* locus in the cultivated rice was 8.9% of that in *O. rufipogon* according to an estimator of average pairwise nucleotide diversity (π). In other words, *MIR156b/c* lost >90% of its nucleotide diversity during domestication (Table 1). Another estimator (Watterson's θ) also showed the same trend as π. Across the *MIR156b/c* locus, the nucleotide diversity was relatively lower in *MIR156b* and the *MIR156c* precursors. The mature sequences of miR156b and miR156c were extremely conserved among the cultivated and wild rice with no nucleotide mutation or insertions/deletions in the 45 samples from the two groups.

Genetic diversity in neutral (unselected) genes is expected to be reduced only by bottleneck or demographic effects, therefore retaining more diversity than selected genes [39]. A recent investigation on domestication bottlenecks based on 10 neutral loci in rice indicated that domesticated rice maintains about 52.1% (π) of the variability found in its progenitor *O. rufipogon* [41]. In the *MIR156b/c* locus, however, only 8.9% diversity was maintained in the cultivated rice. Therefore the bottleneck effect is not enough to explain this dramatic diversity loss. We believe that the *MIR156b/c* locus was a target of domestication selection, although bottleneck effects could have also narrowed the genetic diversity of this locus in the cultivated rice during domestication process. This result suggests that domestication selection could target not only transcription factors as shown before [20] but their upstream regulators, for example, *MIRNAs*.

Acknowledgements: This work was supported by the National High Technology Research and Development Program of China (2006AA10A102) and 973 Program (2006CB101700). Dr. Fan thanks Pao Zhao-long and Pao Yu-kong Scholarship for Chinese Students Studying Abroad for travel grants.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2007.09.002.

References

[1] 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.

[2] Llave, C., Kasschau, K.D., Rector, M. and Carrington, J.C. (2002) Endogenous and silencing-associated small RNAs in plants. Plant Cell 14, 1605–1619.

[3] Mette, M.F., van der Winden, J., Matzke, M. and Matzke, A.J. (2002) Short RNAs can identify new candidate transposable element families in Arabidopsis. Plant Physiol. 130, 6–9.

[4] Park, W., Li, J., Song, R., Messing, J. and Chen, X. (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr. Biol. 12, 1484–1495.

[5] Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B. and Bartel, D.P. (2002) MicroRNAs in plants. Genes Dev. 16, 1616–1626.

[6] Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B. and Bartel, D.P. (2002) Prediction of plant microRNA targets. Cell 110, 513–520.

[7] Jones-Rhoades, M. and Baretel, D.P. (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol. Cell 14, 787–799.

[8] Wang, J.F., Zhou, H., Chen, Y.Q., Luo, Q.J. and Qu, L.H. (2004) Identification of 20 microRNAs from *Oryza sativa*. Nucleic Acids Res. 32, 1688–1695.

[9] Sunkar, R., Girke, T., Kumar, P. and Zhu, J.-K. (2005) Cloning and characterization of microRNAs from rice. Plant Cell 17, 1397–1411.

[10] Luo, Y.C., Zhou, H., Li, Y., Chen, J.Y., Yang, J.H., Chen, Y.Q. and Qu, L.H. (2006) Rice embryogenic calli express a unique set of microRNAs, suggesting regulatory roles of microRNAs in plant post-embryogenic development. FEBS Lett. 580, 5111–5116.

[11] Xie, Z., 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.

[12] Zhang, B., Pan, X., Cannon, C.H., Cobb, G.P. and Anderson, T.A. (2006) Conservation and divergence of plant microRNA genes. Plant J. 46, 243–259.

[13] Xie, K., Wu, C. and Xiong, L. (2006) Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. Plant Physiol. 142, 280–293.

[14] Kasschau, K., Xie, Z., Allen, E., Llave, C., Chapman, E.J., Krizan, K.A. and Carrington, J.C. (2003) P1/HC-Pro, a viral suppressor of RNA silencing interferes with *Arabidopsis* development and miRNA function. Dev. Cell 4, 205–217.

[15] Gandikota, M., Birkenbihl, R.P., Hohmann, S., Cardon, G.H., Saedler, H. and Huijser, P. (2007) The miRNA156/157 recognition element in the 3′ UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J. 49, 683–693.

[16] Maher, C., Stein, L. and Ware, D. (2006) Evolution of Arabidopsis microRNA families through duplication events. Genome Res. 16, 510–519.

[17] Guddeti, S., Zhang, D.C., Li, A.L., Leseberg, C.H., Kang, H., Li, X.G., Zhai, W.X., Johns, M.A. and Mao, L. (2005) Molecular evolution of the rice miR395 gene family. Cell Res. 15, 631–638.

[18] Jiang, D., Yin, C., Yu, A., Zhou, X., Liang, W., Yuan, Z., Xu, Y., Yu, Q., Wen, T. and Zhang, D. (2006) Duplication and expression analysis of multicopy miRNA gene family members in Arabidopsis and rice. Cell Res. 16, 507–518.

[19] Chuck, G., Cigan, A.M., Saeteurn, K. and Hake, S. (2007) The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. Nat. Genet. 39, 544–549.

[20] Doebley, J. (2006) Unfallen grains: how ancient farmers turned weeds into crops. Science 312, 1318–1319.

[21] Rozen, S. and Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers in: Bioinformatics Methods and Protocols: Methods in Molecular Biology (Krawetz, S. and Misener, S., Eds.), pp. 365–386, Humana Press, Totowa, NJ.

[22] Doyle, J.J. and Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.

[23] Katoh, K., Kuma, K., Toh, H. and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

[24] Kumar, S., Tamura, K. and Nei, M. (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinformatics 5, 150–163.

[25] Watterson, G.A. (1975) On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7, 256–276.

[26] Tajima, F. (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105, 437–460.

[27] Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

[28] Fu, Y.X. and Li, W-H. (1993) Statistical tests of neutrality of mutations. Genetics 133, 693–709.

[29] Rozas, J., Sanchez-De, I., Barrio, J.C., Messeguer, X. and Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

[30] Paterson, A.H., Bowers, J.E. and Chapman, B.A. (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc. Natl. Acad. Sci. USA 101, 9903–9908.

[31] Guyot, R. and Keller, B. (2004) Ancestral genome duplication in rice. Genome 47, 610–614.

[32] Zhang, Y., Xu, G., Guo, X. and Fan, L. (2005) Two ancient rounds of polyploidy in rice genome. J. Zhejiang Univ. Sci. B 6, 87–90.

[33] Wang, X., Shi, X., Hao, B.L., Ge, S. and Luo, J. (2005) Duplication and DNA segmental loss in rice genome and their implications for diploidization. New Phytol. 165, 937–946.

[34] Gaut, B.S. (2002) Evolutionary dynamics of grass genomes. New Phytol. 154, 15–28.

[35] Wolfe, K.H., Gouy, M., Yang, Y.-W., Sharp, P.M. and Li, W.-H. (1989) Date of the monocot–dicot divergence estimated from chloroplast DNA sequence data. Proc. Natl. Acad. Sci. USA 86, 6201–6205.

[36] Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W., Hu, S., Zeng, C., Zhang, J., Zhang, Y., Li, R., Xu, Z., Li, S., Li, X., Zheng, H., Cong, L., Lin, L., Yin, J., Geng, J., Li, G., Shi, J., Liu, J., Lv, H., Li, J., Wang, J., Deng, Y., Ran, L., Shi, X., Wang, X., Wu, Q., Li, C., Ren, X., Wang, J., Wang, X., Li, D., Liu, D., Zhang, X., Ji, Z., Zhao, W., Sun, Y., Zhang, Z., Bao, J., Han, Y., Dong, L., Ji, J., Chen, P., Wu, S., Liu, J., Xiao, Y., Bu, D., Tan, J., Yang, L., Ye, C., Zhang, J., Xu, J., Zhou, Y., Yu, Y., Zhang, B., Zhuang, S., Wei, H., Liu, B., Lei, M., Yu, H., Li, Y., Xu, H., Wei, S., He, X., Fang, L., Zhang, Z., Zhang, Y., Huang, X., Su, Z., Tong, W., Li, J., Tong, Z., Li, S., Ye, J., Wang, L., Fang, L., Lei, T., Chen, C., Chen, H., Xu, Z., Li, H., Huang, H., Zhang, F., Xu, H., Li, N., Zhao, C., Li, S., Dong, L., Huang, Y., Li, L., Xi, Y., Qi, Q., Li, W., Zhang, B., Hu, W., Zhang, Y., Tian, X., Jiao, Y., Liang, X., Jin, J., Gao, L., Zheng, W., Hao, B., Liu, S., Wang, W., Yuan, L., Cao, M., McDermott, J., Samudrala, R., Wang, J., Wong, G.K.-S. and Yang, H. (2005) The genome of *Oryza sativa*: a history of duplication. PloS Biol. 3, e38.

[37] Guo, X., Xu, G., Zhang, Y., Wen, X., Hu, W. and Fan, L. (2006) Incongruent evolution of chromosomal size in rice. Genet. Mol. Res. 30, 373–389.

[38] Li, A. and Mao, L. (2006) Evolution of plant microRNA gene families. Cell Res. 16, 1–7.

[39] Yamasaki, M., Tenaillon, M.I., Bi, I.V., Schroeder, S.G., Sanchez-Villeda, H., Doebley, J.F., Gaut, B.S. and McMullen, M.D. (2005) A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop Improvement. Plant Cell 17, 2859–2872.

[40] Khush, G.S. (1997) Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. 35, 25–34.

[41] Zhu, Q., Zheng, X., Luo, J., Gaut, B.S. and Ge, S. (2007) Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during domestication of rice. Mol. Biol. Evol. 24, 875–888.