Cloning and characterization of two cDNAs encoding rice MADS box protein*

YUAN Ziqiang(袁自强), QIAN Xiaoyin(钱晓茵), LIU Jun(刘军), LIU Jiandong(刘建东), QIAN Min(钱 旻) and YANG Jinshui(杨金水)**

(Institute of Genetics, Fudan University, Shanghai 200433, China)

Received July 6, 1999; revised August 9, 1999

Abstract To elucidate the relationship between MADS box gene and rice morphogenesis, RT-PCR with MADS domain specific primers was performed to isolate MADS box gene from young panicle of rice "Zhenshan 97B". Two cDNAs, designated as nmads1 and nmads3, displayed the structure of a typical plant MADS box gene which consists of the MADS domain, I region, K domain, and C-terminal region. Based on sequence homology, nmads1 is classified as a member of GLO subfamily, and nmads3 belongs to AGL2 subfamily. Hybridization analysis revealed that nmads1 and nmads3 were preferentially expressed in rice redifferentiated callus and young panicles but were not in rice seedling. An additional transcript of nmads1 was also found in young panicle of cytoplasmic male-sterile line Zhenshan 97A but was not in its maintenance line Zhanshan 97B.

Keywords: **MADS box gene, rice, cDNA cloning, morphogenesis.**

MADS box genes (the name derived from the initials of the first four identified MADS box genes) exist widely in yeast, human, animals and plants. They encode a family of transcription factors which contain a highly conserved DNA-binding domain of about 60 amino acids in their N terminal, namely MADS box[1]. The MADS box proteins play key roles in several biological processes, all concerning the regulation of cell differentiation. So far, most of the isolated MADS box genes are from plants. They seem to be mainly involved in the genetic control of flower development. Mutation in the gene will result in abnormal flowers[2]. A newly isolated rice MADS box DNA also contributes to flower development[3]. However, MADS box genes have been found in some fern plants which have no flower organ[4]. And when induced by nodule bacteria's infection, the dedifferentiated root cortical cells of some leguminous plants also express MADS box genes[5]. These studies indicate that the function of plant MADS box genes is not restricted to flower morphogenesis. Here we report the characterization of two rice MADS box genes isolated from young panicle by RT-PCR with MADS domain-specific primers. The expression patterns of the two genes in rice at different development stages and in rice callus have also been studied.

* Project supported by the National Natural Science Foundation of China (Grant No. 39770047).

** Corresponding author.

1 Materials and methods

1.1 Rice material

Seeds of the Indica rice (*Oryza sativa L. spp.*) Zhenshan 97B (maintenance line) and Zhenshan 97A (WA-cytoplasmic male-sterile line) were grown in the experimental field of Fudan University.

1.2 Sampling of rice young panicle

When the rice entered reproductive phase, young panicles were collected and examined under a microscope to determine their accurate development stage. The panicles in primordium differentiative phase (panicle size about 2 mm) and in pistil stamen primordium forming phase (panicle size = 0.5—1 cm) were selected.

1.3 Culture of dedifferentiated and redifferentiated callus

The seeds of Zhenshan 97B were shelled, sterilized and inoculated onto callus culture medium (N6 basal medium plus 2 mg/mL 2,4-D and 0.3 mg/mL 6-BA). After about a week, light yellow dedifferentiated calluses were induced on the basis of young shoots and well grown calluses were selected and then inoculated onto the culture medium (N6 basal medium with 0.3 mg/mL 2,4-D and 2 mg/mL 6-BA) for callus redifferentiation. The redifferentiated calluses with green dots and sprouts were obtained after 2—3 weeks of culture.

1.4 Culture of seedlings

The seeds of Zhenshan 97B, sterilized by 10% NaCLO, was germinated and grown in a sand plate. 10-day-old seedlings were collected and frozen in liquid nitrogen and stored in −70℃ until used.

1.5 RT-PCR and cloning of cDNA

Total RNA was prepared according to the procedure described by Sacco et al.[6]. Reverse transcription, PCR amplification, electrophoresis, target cDNA reamplification and purification were carried out based on the procedure described by Cheng et al.[7]. The 5′ primer C used was 5′-GATCAAGAGCATCGAGAA-3′ which encodes the most conserved sequence IKSIEN in MADS domain[8]. Two 3′ oligo(dT) primers were used, Y1, 5′-AAGCTTTTTTTTTTTG-3′, was anchored with a *Hin*dIII restriction site. BT17, 5′-GGATCCTTTTTTTTTTTTTTTT-3′, was not anchored with a *Bam*HI site. The PCR condition was as follows: 1 cycle of 94℃ for 5 min, 54℃ for 2 min, 72℃ for 2 min; 40 cycles of 94℃ for 30 s, 54℃ for 100 s, 72℃ for 120 s; and a final extension at 72℃ for 7 min. The amplified cDNAs were cloned into pGEM-T vector (Promega) and transformed into XL1-blue strains.

1.6 Sequence analysis

The cDNAs were sequenced by an automatic sequencer. Analyses of nucleotide and deduced protein sequence were carried out by GCG Package. Homology comparisons were done with BLAST

mail server of GenBank.

1.7 Southern hybridization

Southern blot transfer, hybridization and washing were performed according to the procedures described by Sambrook et al.[9].

2 Results and analysis

2.1 Isolation of rice cDNAs containing MADS domain from young panicle by RT-PCR

A number of cDNA fragments were amplified from the young panicles of Zhenshan 97B by RT-PCR with two primer pairs: C + Y1 and C + BT17 (fig. 1). A total of 18 cDNA fragments with stronger signals was recovered from sequencing gel and then hybridized with the RNA prepared from Zhenshan 97B young panicles. 2 out of 18 cDNA fragments were cloned and sequenced. The results showed that both of them, designated nmads1 and nmads3, encode a peptide containing MADS box sequence.

2.2 Sequence analysis

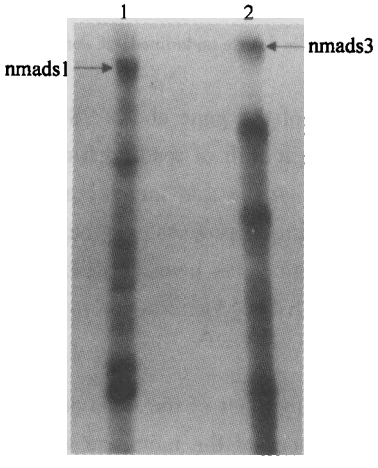

Fig. 1. Amplified cDNAs by RT-PCR. 1, 2: the amplified cDNAs with two primer pairs C + Y1 and C + BT17.

The cDNA nmads1 is 760 bp in size and encodes a peptide of 157 amino acids and cDNA nmads3 is 949 bp in size that encodes a peptide of 236 amino acids (figs. 2 and 3). The 5′ terminus of both nmads1 and nmads3 starts from the sequence of MADS domain-specific primer and 3′ terminus ends at polyA. According to the homology comparison, the initiation codons of the most members of MADS box gene family are located at 7th amino acid upstream of the MADS domain-specific primer[10] (fig. 4), so both nmads1 and nmads3 are nearly full-length cDNAs. Moreover, the gene products of nmads1 and nmads3 are highly homologous to the other members of MADS box proteins and display the typical structure of plant MADS box proteins, which consist of the MADS domain, I region, K domain and c-terminal region[10]. Therefore, nmads1 and nmads3 should be considered as the members of plant MADS box genes (the accession numbers of nmads1 and nmads3 in GenBank are AF095645 and AF095646 respectively).

Based on nucleotide comparison, nmads3 is most highly homologous to wheat TaMADS12 gene. The homology of the nucleotides between them in the three regions, 11—162 bp, 303—506 bp, 608—642 bp are 94%, 80% and 90% respectively. The nmads3 peptide is highly homologous to the four gene products, OsMADS6 (rice), TaMADS12 (wheat), ZAG5 and ZAG3 (corn), and the homology between nmads3 and these produets in the region of 1—219 amino acids are 73%, 70%, 68% and 68% respectively.

The nmads1 is most highly homologous to another rice MADS box gene OsMADS2[11]. The first

```
GATCAAGAGCATCGAGAACTCCACCAACCGCCAGGTGACCTTCTCCAAGCGCAGGAGCGG    60
  I K S I E N S T N R Q V T F S K R R S G
GATCCTCAAGAAGGCCCGCGAGATCAGCGTCCTGTGCGACGCCGAGGTCGGCGTCGTCAT   120
  I L K K A R E I S V L C D A E V G V V I
CTTCTCCAGCGCTGGCAAGCTCTACGACTACTGCTCCCCCAAGACCTCGCTATCAAGAAT   180
  F S S A G K L Y D Y C S P K T S L S R I
CTTGGAGAAGTACCAGACCAATTCCGGAAAGATACTGTGGGATGAGAAGCACAAGAGCCT   240
  L E K Y Q T N S G K I L W D E K H K S L
TAGCGCGGAGATTGATCGAATCAAGAAAGAGAACGATAATATGCAGATTGAGCTCAGGCA   300
  S A E I D R I K K E N D N M Q I E L R H
CTTGAAAGGTGAAGATCTAAACTCTCTGCAGCCCAAAGAGCTCATCATGATTGAGGAGGC   360
  L K G E D L N S L Q P K E L I M I E E A
ACTTGACAATGGGATAGTGAACGTGAATGATAAACTGATGGACCACTGGGAAAGGATAAG   420
  L D N G I V N V N D K L M D H W E R I R
ATGCTGGAAGACGAGAACAAGCTGCTGGCTTTCAAACTGCACCAGCAAGATATAGCGCTG   480
  C W K T R T S C W L S N C T S K I *
AGCGGGAGCATGAGGGATCTTGAGCTTGGGTACCATCCAGACAGGGACTTTGCGGCCCAG   540
ATGCAGATCACCTTCCGCGTGCAGCCCAGCCACCCCAACCTGCAGGAGAACAATTAAGCT   600
GCTAGGTTGCCCCGCCACTTCGATCAGTTATCTCATCCACTGATCCACCACTGGATTGAA   660
TGTCCTAGTGCATTGTCAACTGTATCCCTGTTTTCATGTCTGTTTCGATGAACTATTGAG   720
CATGTCATATGTGAGTTGCTTTGTGTGCCAAAAAAAAAAA                       760
```

Fig. 2. Nucleotide and deduced amino acid sequences of nmads1 cDNA. MADS domain is double underlined, K domain is underlined. The positions of nucleotide is shown on the right.

415 bp of nmads1 shows 99% homology to OsMADS2 and the rest part of the gene shows 98%. At the position of nucleitide 416, nmads1 deleted a 13 bp sequence causing a shift of reading frame and producing a completely different peptide from that of OsMADS2. This might be that nmads1 and OsMADS2 were derived from gene replication and diverged later in the evolution process, or they might be the alternative splicing products of the same gene. The latter explanation has been verified in the study on other plant MADS box genes[12]. The relationship of nmads1 with OsMADS2 needs to be testified by further evidence.

Based on amino acid sequence homology, nmads1 is classified as a member of the GLO subfamily, and nmads3 belongs to the AGL2 subfamily. The amino acid sequences of the members within a subfamily are quite similar(fig.4). The amino acid sequence of nmads1 peptide is 70% identical to the GLO gene product in MADS domain, and nmads3 peptide shows 92% identical to the three gene products, AGL6, ZAG3 and ZAG5 in AGL2 subfamily.

2.3 Expression of nmads1 and nmads3 in different tissues of rice

Figure 5 shows the hybridization result of PCR-amplified product with nmads1 and nmads3. The electrophoresis pattern shows that all lanes containing the products amplified with C + Y1 primers have clear bands, while the amplified products with C + BT17 primers only have high background, suggesting that too many products with different sizes were co-amplified by the unanchored primer BT17. The molecular hybridization reveals that a band about 1.0 kb can be observed in all tissues of rice when hybridized with nmads3, but signal intensity is different from tissue to tissue. As the same RT-PCR condition was used for each tissue, the intensity of hybridization signal could reasonably reveal the genes' expression level. In young panicle, either dedifferentiated or redifferented callus, nmads3 is

```
GATCAAGAGCATCGAGAACAAGATAACCCGGCAGGTCACCTTCTCCAAGCGCCGCAACGG    60
  I  K  S  I  E  N  K  I  T  R  Q  V  T  F  S  K  R  R  N  G
CCTCCTCAAGAAGGCGTACGAGCTCTCCGTGCTCTGCGACGCGGAGGTGGCGCTCATCAT   120
  L  L  K  K  A  Y  E  L  S  V  L  C  D  A  E  V  A  L  I  I
CTTCTCCAGCCGCGGCAAGCTGTACGAATTCGGTAGCGCCGGAATAAACAAGACATTGGA   180
  F  S  S  R  G  K  L  Y  E  F  G  S  A  G  I  N  K  T  L  E
AAAGTACAATAGTTGCTGTTACAACGCTCAAGGTTCAAATAGTGCTCTTGCTGGTGGTGA   140
  K  Y  N  S  C  C  Y  N  A  Q  G  S  N  S  A  L  A  G  G  E
ACATCAGAGCTGGTACCAAGAGATGTCAAGGCTCAAGACTAAGCTTGAATGTCTCCAACG   300
  H  Q  S  W  Y  Q  E  M  S  R  L  K  T  K  L  E  C  L  Q  R
CTCTCAGAGGCACATGCTTGGTGAAGATCTTGGACCATTGAGCATAAAGGAACTGCAGCA   360
  S  Q  R  H  M  L  G  E  D  L  G  P  L  S  I  K  E  L  Q  Q
GCTGGAGAAGCAACTTGAGTACTCACTGTCACAGGCTCGACAACGAAAGACACAAATCAT   420
  L  E  K  Q  L  E  Y  S  L  S  Q  A  R  Q  R  K  T  Q  I  M
GATGGAGCAGGTCGACGATCTTCGCCGGAAGGAACGCCAGCTTGGAGAGCTCAATAAGCA   480
  M  E  Q  V  D  D  L  R  R  K  E  R  Q  L  G  E  L  N  K  Q
ACTGAAAAACAAGCTAGAAGCTGAAGCCGATAGCAGCAACTGCAGATCAGCCATCCAGGA   540
  L  K  N  K  L  E  A  E  A  D  S  S  N  C  R  S  A  I  Q  D
TTCCTGGGTCCATGGCACCGTCGTCAGTGGCGGCAGAGTGCTGAATGCTCAACCACCACC   600
  S  W  V  H  G  T  V  V  S  G  G  R  V  L  N  A  Q  P  P  P
AGATATTGACTGTGAGCCTACTCTGCAAATTGGGTACTATCAATTTGTCCGTCCTGAGCG   660
  D  I  D  C  E  P  T  L  Q  I  G  Y  Y  Q  F  V  R  P  E  R
GCCAATCCAAGAAGCAATGGAGGAGGAGGGGATCAGAACAACAACTTTGTGATGGGATGG   720
  P  I  Q  E  A  M  E  E  E  G  I  R  T  T  T  L  *
CCCCTCTGAAGTCCAAGCTTGCTAATAAAAACGCTGGCGTGCTATAATATATAGTTCGGC   780
AAATGTTGAATCACATGTGTTCTCAGTATGTATATTTCTCCCCTATCCGGTTGTACTCTT   840
AGCACGTACCTATGAGTGTATGTTTGTACGTTATCTATACTCTATGCTAGGCAACCCTAT   900
ACTATTATGGTAAGACGTCTATCTTTTCAATCAAAAAAAAAAAAAAAAA              949
```

Fig. 3. *Nucleotide and deduced amino acid sequences of nmads3 cDNA. MADS domain is double underlined, K domain is underlined. The positions of nucleotide is shown on the right.*

```
                                   MADS domain                                  subfamily
                                   *
MADS-CONSESUS:          MGRGKIEIKR IENKTNRQVT FSKRRNGLLK KAYELSVLCD AEVALIVFSS RGKLYEY
nmads1     Oryza sativa       --S ---S------ -----S-I-- --R-I----- ---GVVI--- A----D-
OSMADS2    Oryza sativa ---------- ---S------ -----S-I-- --R-I----- ---GVVI--- A----D-
OSMADS4    Oryza sativa ---------- ---S------ -----S-I-- --R-IG---- ---GVVI--- A---SD-      GLO
FBP1    Petunia hybrida---------- ---SS----- Y------I-- --K-I----- -R-SV-I-A- S--MH-F
GLO  Antirrhinum majus ---------- ---SS----- Y------IM- --K-I----- -H-SV-I-A- S--MH-F

nmads3     Oryza sativa       --S ----IT---- ---------- ---------- ------I--- ------F
osmads45   Oryza sativa ----RV-L-- ----I----- -A-------- ---------- ------I--N ------F
ZAG3       Zea mays      ----RV-L-- ----I----- ---------- ---------- ------I--- ------F      AGL2
ZAG5       Zea mays      ----RV-L-- ----I----- ---------- ---------- ------I--G ------F
AGL6 Arabidopsis thaliana ----RV-M-- ----I----- ---------- ---------- ------I--- ------F
```

Fig. 4. Comparison of MADS domain conserved sequence of MADS box is on the headline, * indicates the initiation codon of MADS box proteins, amino acids identical to the conserved sequence are indicated by dashes.

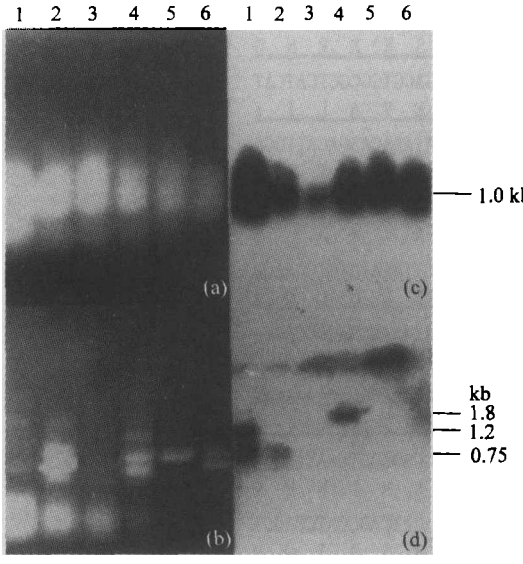

Fig. 5. The electrophoresis of RT-PCR products and the molecular hybridization analysis. (a) and (b) are the electrophoreses of RT-PCR products amplified by C + BT17 and C + Y1 primers respectively, (c) is the hybridization of nmads3 to (a), (d) is the hybridization of nmads1 to (b). 1, Young panicle of Zhenshan 97A (about 2mm in length); 2, young panicle of Zhenshan 97B (about 2mm in length); 3—6 Zhenshan97B (3, ten-day-old seedlings; 4, redifferentiatied callus; 5, dedifferentiated callus; 6, young panicle about 1cm in length).

preferentially expressed; but in seedlings, it is much less expressed. The expression pattern of nmads1 is more complicated. Besides the 1.0 kb band, a fragment about 1.8 kb was detected in redifferentiated callus, and two additional bands with sizes of 1.2 kb and 0.75 kb respectively presented in young panicle, while no band was detected in seedlings.

2.4 Differential expression of nmads1 and nmads3 in rice Zhenshan 97A and Zhenshan 97B

It is interesting to note that the hybridization pattern in young panicle of Zhenshan 97A is remarkably different from that of Zhenshan 97B (fig. 5). Two bands of 0.75 kb and 1.2 kb hybridized with nmads1 can be seen in Zhenshan 97A, while only one band of 0.75 kb appears in Zhenshan 97B, and the signal of the former is much stronger than that of the latter. The expression level of nmads3 in Zhenshan 97A is also higher than that in Zhenshan 97B. According to the fact that the mutated phenotype of flower is the same as that caused by mutation of members in GLO and DEF gene subfamilies detected in wheat cytoplasmic male-sterile line[13], our findings imply that MADS box genes may involve in controlling rice cytoplasmic male-sterile phenotype. It is worth studying further.

The possibility of the hybridization bands involving transcripts of the other MADS box genes can not be ruled out because of the high homology of MADS domain existing in the members of the same subfamily. However, in our study, the hybridization was under the condition of 68℃ stringency, and both the amplified sequences and the probes consisted of large low-homology region except for the MADS domain, so the transcripts of the other MADS box gene , if existed, should be the products of the same gene subfamily. The notable different expressions of nmads1 and nmads3 show that MADS box genes may play important roles in different development processes of rice and in the morphogenesis of young panicle of mail-sterile line and maintenance line.

References

1 Schwarz-Sommer, Z., Huijser, P., Nacken, W. et al., Genetic control of flower development by homeotic genes in *Antirrhinum mojus*, *Science*, 1990, 250: 931.

2 Yanofschy, M. F., Ma, H., Bowman, J. L. et al., The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factor, *Nature*, 1990, 346: 35.

3 Meng, Z., Liu, G. Z., Xu, Z. H. et al., Cloning and characterization of a rice genomic MADS box DNA, *Science in China* (in Chinese), 1998, 28(5): 450.

4 Hasebe, M., Wen, C. K., Kato, M. et al., Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*, *Proc. Natl. Acad. Sci. USA*, 1998, 95: 6222.

5 Heard, J., Dunn, K., Symbiotic induction of a MADS-box gene during development of alfalfa root nodules, *Proc. Natl. Acad. Sci. USA*, 1995, 92: 5273.

6 Sacco, D., Hoge, H., Bisseling, T. et al., Isolation of total and polysome RNA from plant tissue, in *Plant Molecular Biology Manual*, Netherlands: Kluwer Academic Publishers, 1988, B: 1—13.

7 Cheng, N. H., Yang, J. S., Gao, Y. P. et al., Differential display of mRNA between hybrid F1 and its parental inbred lines in maize, *Chinese Science Bulletin*, 1996, 41(5): 451.

8 Fasher, A., Saedler, H., Theissen, G., Restriction fragment length polymorphism-coupled domain-directed differertial display: A highly efficient technique for expression analysis of multigene families, *Proc. Natl. Acad. Sci. USA*, 1995, 92: 5331.

9 Sambrook, J., Fritsch, E. F., Maniatis, T., *Molecular Cloning: A laboratory Manual*. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989, 9.31—9.58.

10 Purugganan, M. D., Rounsley, S. D., Schmidt, R. J. et al., Molecular evolution of flower development: deversification of the plant MADS-box regulatory gene family, *Genetics*, 1995, 140: 345.

11 Chung, Y. Y., Kim, S. R., Kang, H. G. et al., Characterization of two rice MADS box genes homologous to GLOBOSA, *Plant Science*, 1995, 109: 45.

12 Kyozuka, J., Haraurt, R., Peacock, W. J. et al., Eucalyptus has functional epuivalents of the *Arabidopsis* AP1 gene, *Plant Mol. Biol.*, 1997, 35: 573.

13 Ogijara, Y., Futami, K., Tsuji, K. et al., Alloplasmic wheat with *Aegilops crassa* cytoplasm which express photo-period-sensitive homeotic transformation of anthers, show alterations in mitochondrial DNA structure and transcription, *Mol. Gen. Genet.*, 1997, 255: 45.