NsMADS1, a Member of the MADS Gene Family from *Nicotiana sylvestris*

Seonghoe Jang and Gynheung An*

Department of Life Science, Pohang University of Science and Technology, Pohang 790-784, Korea

A cDNA clone, *NsMADS1*, was isolated from *Nicotiana sylvestris*. Sequence homology analysis indicated that the gene is a member of the SQUA subfamily. Its transcript was detectable primarily in reproductive organs, especially in sepals and carpels. When the gene was expressed ectopically in tobacco plants, no phenotypic alteration was observed. Thus, *NsMADS1* appears to be a new member of the MADS gene family.

Reproductive success in plants depends on their ability to initiate flower development. Flowering is controlled by various environmental factors, such as photoperiod, temperature, and nutrient availability, and also by programmed genetic factors (Singer and McDaniel, 1986). Recent studies have revealed that the determination of floral meristem and organ identity in angiosperms is controlled by a group of regulatory genes belonging to the MADS family (Yanofsky, 1995; Kang et al., 1997). The MADS box genes also play a major role in controlling a variety of plant development factors.

We have isolated a cDNA clone from *Nicotiana sylvestris*, in which floral primordia formation is induced by long day conditions. The clone, *NsMADS1*, contains an open reading frame of 245 amino acid residues (Fig. 1). The NsMADS1 protein consists of the 56 amino acid (aa) MADS-domain, 35 aa I region, 65 aa K domain, and 88 aa C terminal region, which are found in all plant MADS proteins. The protein is highly homologous to other MADS proteins in the SQUA subfamily, including potato POTM1-1 (82% identity), wild potato SCM1 (78% identity), *Arabidopsis* AP1 (77% identity), and *Antirrhinum* SQUA (65% identity) (Figs. 2 and 3). The NsMADS1 protein contains 51st K and 60th D, which are characteristic of the MADS domain in the SQUA subfamily (Theissen et al., 1996). The SQUA subfamily proteins can be divided into two groups based on conserved sequences in the C terminal region; one contains the GCFA motif and the other contains the RHLN motif (Fig. 3). The GCFA-motif group includes AP1, CAL, and SQUA, which are known to be activated early in flower development to specify the identity of floral meristems (Huijser et al., 1992). NsMADS1 belongs

to the RHLN-motif group, which seems to play diverse roles based on its expression pattern. For examples, the *POTM1-1* transcript was found ubiquitously distributed in plants (Kang and David, 1996) and the *MdMADS2* transcript was preferentially

Figure 1. Nucleotide and deduced amino acid sequences of the *NsMADS1* cDNA. MADS box and K box regions are underlined. A cDNA library was constructed using Lambda ZAPII vector (Stratagene, La Jolla, CA) and mRNA prepared from young floral buds (< 5 mm in length) of *N. sylvestris*. The initial plaque forming unit was 1.6×10^{6}. The probe was *Arabidopsis* APETALA1 (AP1) cDNA which was obtained by PCR using two primers, 5' TCAAAA<u>ATG</u>GGAAGGGGTAG-- GGTTC 3' and 5' CTTCATGCGGCGAAGCAGCCAAGGT 3'. The positions of nucleotides and amino acids are shown on the left and right, respectively. The *Nsi*I site, which was used to generate the gene-specific probe of the 500 bp fragment, is shown in italics. An asterisk (*) indicates a stop codon. The Genbank accession number of *NsMADS1* is AF068725.

*Corresponding author; fax +82-562-279-2199
e-mail genean@postech.ac.kr

A

MADS-box

```
NsMADS1   MGRGRVQLKRIENKINRQVTFSKRASGLLKKAHEISVLCDAEVGLIVFSTKGKLFEY
POTM1-1   *********************************R***********************
SCM1      *********************************R***********************
BpMADS5   *********************************R*********************A**I************
MdMADS2   *********************************R***M*************A**I************
BpMADS3   **************************************************AV****H********
Saap1     ***********************************RA**************A*V***H*******
BoAP1     ***********************************RA**F***********A*V***H*******
AtAP1     ***********************************RA**************A*V***H*******
CAL       *****E*******************RT*****Q*************S****H*******
SQUA      ****K*********************RG*******L*******A****N*******
```

B

K-box

```
NsMADS1   LEHAKLKARLEVLQRNQRHYAGEDLDSLSMKELQNLEHQLDSALKHIRSRKNQLMHESISELQKK
POTM1-1   ****************K**V****E**N*************************************V***Q
SCM1      **N***********EKL*V****E**N*********************A*****************V***Q
BpMADS5   ****************K*FV*******L********Q*********************Y******R*
MdMADS2   ********V***********M***Q**L********Q**************V*Y*********
BpMADS3   M*F*R**GKV*L****H***L*D**E***H*******Q***T****V*T******Y****Q****
Saap1     **YNR***KI*L*E*****L****QAM*S*******Q***T**************D**N***R*
BoAP1     M*YNR***KI*L*E*****L****QAM*P*******C***T**************Y***N***R*
AtAP1     M*YNR***KI*L*E*****L****QAM*P*******Q***T****T******Y***N*****
CAL       M*YSR***KI*L*E*****L**E*EPM*L*D*****Q**ET************N**LNH**R*
SQUA      **YS*****I*L****H***M******M*L**I*S**Q***T**N*T*****LYD******H*
```

C

C-terminal end

```
NsMADS1   YPTAG------DNGEVEG--SSRQQQ--QNTVMPPWMLRHLNG
POTM1-1   *QNTNV----V******GN**Q**GAAN****Q***RHLN*
SCM1      SQNTNV----V******GN**QX*GAAN****Q**VRHLN*
BpMADS5   --------SQQAR*NGR-VDEGTPPHRA*ALL****RHLNQ
MdMADS2   SNYQAIRSSEGIP*DNQQYGDETPTPHRP*MLL*A*IVRHLNE
BpMADS3   L*CLNI-----G*NYQ-EEAPEVRRNELELTLE*IYSC**GCFAT
Saap1     S*FLNM-----G*LYQEEDPMEMRRNDLDLSLE*VYNCN*GCFAA
BoAP1     S*FLNM-----G*LYQEEDQMAMRRNDLDLSLE*VYNCN*GSFAA
AtAP1     S*FLNM-----G*LYQEDDPMAMR-NDLELTLE*VYNCN*GCFAA
CAL       S*FLNM-----G*LYQ*EDQTAMRRNNLDLTLE*IYNY-*GCYAA
SQUA      F*CINV-----GNTY**EGANEDRRNELDLTLDSLYSC**GCFAA
```

Figure 2. Multiple alignment of the amino acid sequences deduced from NsMADS1 and other MADS proteins in the SQUA subfamily. The MADS-box (A), K-box (B), and C-terminal end (C) regions were aligned. Gaps (hyphens) were introduced for the maximum sequence homology. The asterisks (*) indicate identical amino acids to NsMADS1.

present in developing fruit. The results in Figure 4 demonstrate that the *NsMADS1* transcript is present in flowers, especially in sepals and carpels, but not in vegetative organs such as leaves, stems, and roots of mature plants and in seedlings. A similar expression pattern was observed in a rice MADS gene, *OsMADS1*, if palea/lemma were regarded as the first whorl (Chung et al.,1994).

The functional role of the *NsMADS1* gene was investigated by ectopically expressing the gene in transgenic tobacco plants. The cDNA clone was put under control of the CaMV 35S promoter, and the

chimeric molecule (pGA1577) was introduced into tobacco plants by the *Agrobacterium*-mediated transformation technique (An et al., 1988). Among the 10 transgenic plants tested, we did not observe any phenotypic abnormalities (data not shown). RNA blot analysis showed that the *NsMADS1* transcript was detectable in leaves from all of the transgenic plants. Thus, this method was unable to reveal the functional role of the *NsMADS1* gene. It is possible that NsMADS1 may need another organ-specific factor to control the expression of the target gene (Chung et al., 1998). Further studies are required to understand
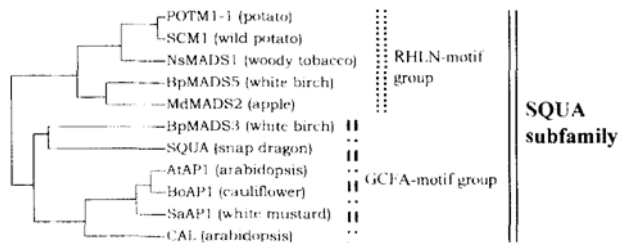
Figure 3. Phylogenic tree for NsMADS1 and other MADS proteins in the SQUA subfamily. Shown is a neighbor-joining method tree based on the amino acid sequences.
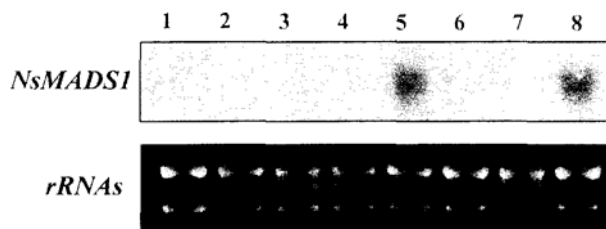

Figure 4. Expression of the *NsMADS1* gene in different organs of tobacco plants. Twenty-five micrograms of total RNA was loaded on each lane. The RNA blot was hybridized with ³²P-labelled *NsMADS1* probe, which was prepared from a cDNA fragment lacking the MADS-box region. 1, 15-day-old seedlings; 2, mature stems; 3, mature leaves; 4, roots from mature plants; 5, sepals; 6, petals; 7, stamens; 8, carpels. Ethidium bromide staining of rRNAs is shown to ensure equal amounts of RNA loading.

the function of the *NsMADS1* gene.

LITERATURE CITED

An G, Ebert PR, Mitra A, Ha SB (1988) Binary vectors. *In* SB Gelvin, RA Schilperoort, eds, Plant Molecular Biology manual, Kluwer Academic Publishers, Dordrecht, Netherlands, pp A3/1-19

Chung YY, Kim SR, Kang HG, Noh YS, Park MC, Finkel D, An G (1994) Early flowering and reduced apical dominance result from ectopic expression of a rice MADS-box gene. Plant Mol Biol **26**: 657-665

Chung YY, Lee KJ, An G (1998) Characterization of a tobacco MADS-box gene homologous to AGL2. Mol Cells **8**: 764-769

Huijser P, Klein J, Lonnig W-E, Meijer H, Saedler H, Sommer H (1992) Bracteromania, an inflororescence anomaly, is caused by the loss of function of the MADS-box gene *squamosa* in *Antirrhinum majus*. EMBO J **11**: 1239-1249

Kang HG, Jang S, Chung JE, Cho YG, An G (1997) Characterization of two rice MADS box genes that control flowering time. Mol Cells **7**: 559-566

Kang SG, Hannapel DJ (1996) A novel MADS-box gene of potato (*Solanum tuberosum* L.) expressed during the early stages of tuberization. Plant Mol Biol **29**: 379-386

Singer SR, McDaniel CN (1986) Floral determination in the terminal and axillary buds of *Nicotiana tabacum* L. Dev Biol **118**: 587-592

Theissen G, Kim JT, Saedler H (1996) Classification and phylogeny of the MADS box multigene family suggest defined roles of MADS box gene subfamilies in the morphological evolution of eukaryotes. J Mol Evol **43**: 484-516

Yanofsky MF (1995) Floral meristems to floral organs: Genes controlling early events in Arabidopsis flower development. Annu Rev Plant Physiol Plant Mol Biol **46**: 167-188