

Characterization of MADS Box Genes from Hot Pepper

Soon-Kee Sung¹, Yong-Hwan Moon, Jae-Eun Chung, Sook-Yi Lee, Hyo Guen Park², and Gynheung An*

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

¹ Biobusiness Research Center, Hansol Institute of Science and Technology, Namyangju 472-850, Korea;

² School of Plant Sciences, Seoul National University, Suwon 441-774, Korea.

(Received January 29, 2001; Accepted March 14, 2001)

The cDNA clone, *CanMADS1*, was isolated from young flower buds of the hot pepper (*Capsicum annuum* L.) by screening a cDNA library using the *OsMADS1* rice MADS-box gene as a probe. We used a yeast two-hybrid screening method to investigate interaction partners of the protein product of *CanMADS1*. A MADS-box gene, *CanMADS6*, was isolated from young flower buds using the region containing the K domain and 15 amino acid residues of the C-terminal region of *CanMADS1* as a bait. *CanMADS1* and *CanMADS6* showed high amino acid sequence similarities to members of the *AGL2* subfamily and the *SQUA* subfamily, respectively. *CanMADS1* transcript was expressed in flower buds and fruits, and the transcription signal was the strongest in the stage of the fruit set (2 d after anthesis). *CanMADS6* showed the same expression pattern as *CanMADS1*. *CanMADS1* and *CanMADS6* were not expressed in leaves. These results suggest that a regulatory role for flower and fruit development of the hot pepper may be accomplished through an interaction of the protein products of the two MADS-box genes, *CanMADS1* and *CanMADS6*.

Keywords: Fruit Development; Hot Pepper; MADS Genes; Yeast Two-Hybrid.

Introduction

MADS (MCM1 AG DEFA SRF) box genes encode a family of highly conserved transcription factors that participate in the signal transduction and developmental control in animals, yeast, and fungi. MADS box proteins have a conserved domain, the MADS domain, which consists of

56 amino acid residues. It is well known that in plants, many MADS-box genes are involved in the control of flower development. The majority of plant MADS-box genes that have been characterized function as floral meristem or organ identity genes (Weigel and Meyerowitz, 1994): *APETALA1* (*API*), *AGAMOUS* (*AG*), *PISTILATA* (*PI*), and *APETALA3* (*AP3*) in *A. thaliana* (Goto and Meyerowitz, 1994; Jack *et al.*, 1992; Mandel *et al.*, 1992; Yanofsky *et al.*, 1990), as well as *SQUAMOSA* (*SQUA*), *PLENA* (*PLE*), *GLOBOSA* (*GLO*), and *DEFICIENS* (*DEF*) in *Antirrhinum majus* (Bradley *et al.*, 1993; Huijser *et al.*, 1992; Sommer *et al.*, 1990; Tröbner *et al.*, 1992).

In addition to floral meristem and organ identity genes, there is a large number of other MADS-box genes in plants whose functions are less well defined. In *Arabidopsis*, at least 17 *AGL* (for *AG-Like*) genes have been isolated (Ma *et al.*, 1991; Mandel and Yanofsky, 1995; Rounsley *et al.*, 1995). It was reported that several MADS-box genes have more subtle functions that are associated with floral meristem and floral organ identity. These genes include: *AGL2* (Flanagan and Ma, 1994), *AGL4* (Savidge *et al.*, 1995), and *AGL9* (Mandel and Yanofsky, 1998) from *Arabidopsis*, *TM5* from tomato (Pnueli *et al.*, 1994), and *FBP2* from petunia (Angenent *et al.*, 1994). In addition to the MADS-box genes that were isolated in annual plants, several MADS-box genes were also isolated from the flowers of fruit trees, including eucalypt and apple (Southerton *et al.*, 1998; Sung *et al.*, 1999; 2000). It was reported that some of the MADS-box genes were expressed in developing receptacles and fruits (Southerton *et al.*, 1998; Sung *et al.*, 2000; Yao *et al.*, 1999).

Plant MADS box proteins consist of a MADS domain, I-region, K domain, and C-terminal region. The conserved MADS domain is required for sequence-specific DNA binding and dimerization (Mizukami *et al.*, 1996; Riechmann *et al.*, 1996a; 1996b; West *et al.*, 1998). The MADS domains bind to the consensus DNA sequences, CARG

* To whom correspondence should be addressed.

Tel: 82-54-279-2176; Fax: 82- 54-279-2199

E-mail: genean@postech.ac.kr

motifs (Huang *et al.*, 1993; 1995; Tilly *et al.*, 1998). Also, in the cases of *AP3* and *PI*, the MADS domain and I-region are needed for nuclear localization of the proteins (McGonigle *et al.*, 1996).

The K domain is the second conserved region carrying 65 to 70 amino acid residues located in the middle of the MADS proteins. It was recently demonstrated that the K domain is required for interactions between MADS proteins (Davies *et al.*, 1996; Fan *et al.*, 1997; Moon *et al.*, 1999a). In experiments using the yeast two-hybrid system, GLO and DEF, the B class proteins of *A. majus* specifically selected each other as a partner for this protein-protein interaction. Furthermore, PLE, the C class protein of *A. majus*, interacted with the products of MADS-box genes expressed at the early (*SQUA*), intermediate (*DEFH72*, *DEFH200*), and late (*DEFH49*) stages of flower development (Davies *et al.*, 1996). In studies using the same system, AG interacted with AGL2, AGL4, AGL6, and AGL9 (Fan *et al.*, 1997). It was also revealed that the B class proteins of rice, OsMADS4 and OsMADS16, interact with each other as a heterodimer (Moon *et al.*, 1999b).

Although various MADS-box genes have been studied in a number of plant species, they have not been reported from hot pepper plants. In this study, we isolated a hot pepper MADS-box gene, *CanMADS1*, from a young hot pepper flower bud cDNA library. We also described the isolation of *CanMADS6*, whose protein product is a protein-protein interaction partner of *CanMADS1*, by the yeast two-hybrid system.

Materials and Methods

Construction of cDNA library and isolation of *CanMADS1*

Total RNA was isolated from young flower buds (0.5 to 2 mm length) of hot pepper (*Capsicum annuum* L.) plants. The mRNA was isolated using the poly(A) Quick mRNA isolation kit (Stratagene, USA). The expression cDNA library was constructed by the standard method according to the manufacturer's protocols. Total cDNAs of a phagemid form of the uni-ZAP cDNA library were obtained by the mass *in vivo* excision method. The initial pfu of the Uni-ZAP cDNA library was 1.24×10^6 .

Hybridization was performed with 2×10^5 plaques using a labeled probe prepared from the MADS domain and I-region of the *OsMADS1* cDNA (Chung *et al.*, 1994). The cDNA clones were rescued by *in vivo* excision using an ExAssist helper phage (Stratagene, USA). Double-stranded DNA was used as a template for the DNA sequence analysis following the manufacturer's instructions (Amersham, UK). An amino acid sequence homology comparison was performed using the BLASTX alignment program (Altschul *et al.*, 1997).

Construction of cDNA library and plasmid for yeast two

hybrid screening The total cDNAs of a phagemid form of the Uni-ZAP cDNA library were digested with *EcoRI* and *XhoI*, and were separated on an agarose gel. DNA fragments of 0.5 to 2.5 kb length were eluted, and the expression cDNA library for yeast two-hybrid screening was made using the HybriZAP vector (Stratagene, USA). The initial pfu of the HybriZAP cDNA library was 4.97×10^6 .

The binding domain vector, pBDGAL4, and the activation domain vector, pADGAL4, were purchased from a commercial supplier (Stratagene, USA). The sequences containing a portion of *CanMADS1* were generated by a polymerase chain reaction (PCR). DNA sequences encoding the following amino acid (aa) residues were amplified: the 90th to 172th aa of *CanMADS1* (*CanMADS1*-KC15). For all constructs, the 5' *EcoRI* site and the 3' *SalI* site were introduced by PCR. The PCR profile used was 30 s at 95°C, 30 s at 57°C, and 30 s at 72°C for a total of 35 cycles. The PCR products were digested with *EcoRI* and *SalI*, ligated to pBDGAL4 or pADGAL4, and transformed into their appropriate hosts.

Yeast two-hybrid screening The yeast strain, YRG-2 (*Mat*, *ura3-52*, *his3-200*, *ade2-101*, *lys2-801*, *trp1-901*, *leu2-3*, *112*, *gal4-542*, *gal80-538*, *LYS::UASGAL1-TATAGAL1-HIS3*, *URA3::UASGAL4* 17mers (x3)-*TATACYC1-lacZ*), was purchased from a commercial supplier (Stratagene, USA). YRG-2 was transformed with pBD/*CanMADS1*-KC15, the binding domain plasmid containing the K domain and 15 amino acid residues of the C-terminal region of *CanMADS1*, according to the modified procedure of the lithium acetate method (Gietz *et al.*, 1992). The transformants were tested for the *HIS3* reporter gene expression, and the absence of growth on a histidine-lacking medium indicated no expression. Using this strain, yeast two-hybrid screening was done as described previously (Moon *et al.*, 1999a). The colonies that turned blue in less than 6 h by filter assay (Breedon and Nasmyth, 1985) were selected for isolation of DNA, and the isolated plasmid DNAs were then retransformed into the YRG-2 yeast strain containing pBDGAL4 or pBD/*CanMADS1*-KC15. Plasmid DNA was recovered from yeast according to the method of Hoffman and Winston (1987) and transformed into an *E. coli* strain, XL1-blue MRF' (Stratagen, USA), by electroporation.

Isolation of the *CanMADS6* cDNA The 5' region of the *CanMADS6* cDNA was isolated by 5'-RACE amplification with the SK primer (5'-CGCTCTAGAACTAGTGGATC-3') and the two *CanMADS6* gene-specific primers (5'-TTGCAAAACCTCAAGCCTGG-3'; nucleotides 353–372 and 5'-TCTCCCTTTCCTTCATCTGC-3'; nucleotides 564–583) that are located within the K domain and the C-terminal region, respectively. The template was floral bud cDNAs of a phagemid form, which were *in vivo* excised from the floral bud cDNA library. The amplified fragment was cloned into pBluescript II SK (–) (Stratagene, USA). Sequences were determined by the ABI automatic sequence system. The *CanMADS6* cDNA was cloned by PCR using *pfu* polymerase (Stratagene, USA). A set of primers was designed to

amplify a near full length clone. The sense primer sequence was 5'-ATTGCCAAACACAGAAAAAAA-3'; nucleotides 1–21, and the antisense primer was 5'-GGTTAACAGCGACACAGCAGC-3'; nucleotide 936–956.

RNA blot analysis Total RNA was isolated by the RNA isolation kit (TRI Reagent, Molecular Research Center) from leaves, young flower buds (0.5 to 2 mm length), flowers (pre-anthesis flower), flowers at fruit set stage (2 d after anthesis), and young fruits (7 to 10 mm length). Total RNA was fractionated on a 1.3% (w/v) agarose gel. After RNA transfer onto a nylon membrane, the blots were prehybridized, hybridized, and washed as previously described (Sung *et al.*, 1999). The 210 bp C-region (nucleotides 590–810) of *CanMADS1* and the 250 bp C-region (nucleotides 620–870) of *CanMADS2* were used as specific probes.

Bacterial strains, plant materials, and plant transformation *Agrobacterium tumefaciens* LBA4404 (Hoekema *et al.*, 1983) was used for transformation of tobacco plants (*Nicotiana tabacum* L. cv. Petit Havana SR1 and cv. Xanthi) by the cocultivation method (An *et al.*, 1988). The cDNA fragment containing the entire coding region of CanMADS1 was inserted between the CaMV35S promoter and the transcript 7 terminator of a binary vector, pGA1530, in the sense orientation. The resulting construct was introduced into the *A. tumefaciens* strain LBA4404. Transgenic tobacco plants were maintained under greenhouse conditions.

Construction of phylogenetic tree Alignment of conceptual amino acid sequences was made using the DNASTAR program CLUSTAL V. Phylogenetic trees were constructed by comparing 170 amino acid sequences comprised of the MADS domain, I-region, and K domain. The analyses employed the neighbor-joining algorithm, as provided by the program MegAlign of the DNASTAR phylogenetic package.

Results

Isolation of CanMADS1 from cDNA library The *CanMADS1* cDNA clone is 1,074 bp long and encodes a putative protein of 245 amino acid residues (GenBank accession No. AF129875). The MADS domain of the cDNA clone is located between the 2nd and 57th amino acids of the protein (Fig. 1A). The second conserved domain, the K box, is located between the residues 92th and 157th. The CanMADS1 protein contains two variable regions: the I-region between the MADS and K domain, and the C-terminal region downstream of the K domain (Purugganan *et al.*, 1995). The C-terminal region of CanMADS1 has one short motif (MNGFIHGWML) that is partially conserved in PrMADS1, EGM3, and AGL2 (Fig. 2A). Based on the amino acid sequence similarity of the entire coding region, *CanMADS1* can be grouped into

A	TA ACC OCT TTT TAT ATA DCC ATC CTC TCT GGA AAA ATT GAA CAA ACA TAG ACA ATA ATG	99
	<u>M</u>	1
	GGA AGA GGA GGA GTT CAG TTG TAG AGA ATA GAA AAG AAA ATA AAC AGA CAA GTC ACT TTT	139
	<u>G R G R Y R L E R I E N K E I N R Q Y T F S</u>	21
	GCT ANG AGA AGA AAT GGA CTT CTT AAG AAA GCT TAT GAA CTT TTT GTC TGT GAT GCT	179
	<u>A K R R R N G L L K R A Y K E L S V L C D A</u>	41
	GAA GTT OCT CTT ATC ATT TTC TCA AAT GGT GGT AAA CTC TAT GAA TTC TGC AGC ACT TCA	239
	<u>R Y A R I I F S N R G R L Y F P C S T S</u>	61
	AGC ATG GGG AAA ACA ATT GAA AAG TAT CAG GCT TGC AGC TAT GCT ACT TGC GAA GCT ANC	299
	<u>S W Y R T I E R Y Q R C S V A T L E A N</u>	81
	CAA TCA OCT ACT GAT ACT CAG AAT AAC TAC CAC GAG TAT CTG AGG CTA AAA GCA AGA GTT	359
	<u>Q S A T D T Q N N Y H E V L R L K A R Y</u>	101
	GAG CTC CTC CAA GGA TCT CAG AGA AAC TTT CTT GGG GAA GAT TTG GGG ACA TTA AGC ACA	419
	<u>R L L Q R H S Q R N P L Q F D I G T L S T</u>	121
	ANG GAC CTG GAG CAG CTT GAG AAT CAA TTA GAG TGC TCC TTA AAG CAA ATC AGG TCA AGG	479
	<u>K D L E Q L E N Q L K S S L K Q I R S R</u>	141
	ANG ACA CAA TTC ATG CTG GAT CAG CTT GAA GAT CTT CAA CAA AAG GAC CAA ATG CTG GCA	539
	<u>N T Q F W L D Q L A D L Q Q R E Q H L A</u>	161
	GAA TCT AAT AGA TTA CTC GGC AGA AAG CTA GAA GAA AGT ACA GCT GGA TTT GCA GTT GGA	599
	<u>E S N R L L R R K L E E S T A G F P Y R</u>	181
	TTG AGT TGG GAA GAT GGA GCT GAT CAA GCT ATG CAT CAA GAT AAT OCT CTC CXC CAC ACT	659
	<u>L S W E D G A D Q A M H G H N H L P R T</u>	201
	GAG GGT TTC TGG CAG OCT CTT GGA TTG CAT TCT TCT GCT GEA CAT TTT GGG TAC AAT OCT	719
	<u>E G F L Q P L G L H S S P P H F G T N P</u>	221
	GTT AAT ACG GAT GAG GTG AAT GCA GCA GCA GCT ACT CAC AAT ATG AAT GGA TTT AAT CTT	779
	<u>V S T D E V N A A A T A A H N M H G P I H</u>	241
	GGA TGG AAG CTT TAA TCT GAT AAT TGG AAG CAT CTA CTC CAT CTT TGG TAC AAC CAA AEA	839
	<u>G W M L *</u>	245
	TAC AGT AAT AAT GTA TTT GCA TGT ATT GTC CTT TTT GTT TTG TTG CAA AET GIA TTT AGA	899
	ACC TTT ATC GAT CAA TGG AGG ACA ATA AAT TTA ATG TAT GEA TAA TGA AGC TGT TGC TAG	959
	<u>CTA CTA GAA CAT TGC AAC CAT CXT GAC TAT GGA TTT CTG CTT GTA TCC CAA TAT GTA ATT</u>	1019
	<u>AAA TGG AAG TGT ATT TTT GTA TCT TTT AAA CTA CTA TTC AGG TTG CAC CTT</u>	1039
B	ATT GGC AAT CAG AGA AAA AAA AAA AAC NAA AAA CAA AEA CTA AGA AAA ATG GGA	69
	<u>M S</u>	2
	AGA GGA GGA GTT CAA TTG AAG AGG AAT GAA AAT AAG ATA AAT AAG CAA GTC ACT TTT TGG	129
	<u>R I R Y Q L H R I R S R I M R Q V T F S</u>	22
	ANG AAG GAA TCT GGT TTG TTS ANG AAA GCT CAT GAG ATC TCT GTC CTT TGT GAT GCT GAA	189
	<u>K H H S Q L L K R A R E R I S Y L I C D A R</u>	42
	GTT GGC TTG ATE GTT TTT TCT TCT AAA GGG AAA CEA TTT GAG TAT TCT ACT GAC TCT TGC	249
	<u>V G L I I Y F S S K G K L F S E I S T D S C</u>	42
	ATG GAA AAG ATT CTT GAG AAG TAT GAA AAG TAC TCA TCT GCT GAG AAG CAG CTT AAT GCA	309
	<u>M R R I L E R Y E R Y S Y A E R Q L N A</u>	82
	ACT GAT GTC GAA AGC GGG AGT TGG ACT TTG GAA CXT GCT AAG CTT AAG GGC AGG CTT	369
	<u>T D V E T P G S N T L E H A K L K A H L</u>	102
	GAG GTT TTG CAA AAG AAC CAA AAG CAT TAT GGG GGA GAA GAC TGG GAC TCA TTG AGT AIG	429
	<u>K V L R R N Q R H Y A G S D I D S L S W</u>	122
	AAA GAG CTT CAG AAT CTG GAG CAG CAA CTC GAT TCT GCT CTT ANG CAC AAT GGA TCA AGA	489
	<u>K F L Q N L E Q Q I D S A L K H I R N R</u>	142
	ANG AAC CAA TTG ATG CAT GAA TCC ATT TCT GAG CTG CAA ANG AAG GAC AAG GGA TTG CAA	549
	<u>K N Q L M H E S I S E L Q K S Q K A L Q</u>	162
	GAA CAA AAC AAC AAT CTT TCA ANG CAG ATT ANG AAG AAG GAG AAA CAA CAG GGC CAG GAG	609
	<u>E Q N N N L S K Q M K E R E K Q L A Q Q</u>	182
	GCT ACT CEE TGG GAG CAA CAG AAC CAT GAC CAT CAC ATC TCA TCT TCA TTT GGT CTG GCA	669
	<u>R T P W K Q Q N H D H L N S S S F G L P</u>	202
	CAT CXC TTT AAC AAC AAT CAC CTA GGG GAA GTA TAT CCA ACT GCA GGA GAG AAT GGA GAA	729
	<u>R P F N N N H L G K V Y P T A G D N G E</u>	222
	GTT GAA GGA TTA TGG GGG CAG CAA CAA CAA AAC GAT GCT GTC ATG CCG CCA TGG ATG CTT	789
	<u>F E C S S R Q Q Q N A A V M P P W M L</u>	242
	GCC GAT CXC AAC GGA TAA AAC TTA TCA AGT ATA AAG AAG GGG ACA ATA CTA ATA TCC TAT	849
	<u>R H L N G *</u>	247
	GTT TTG TTA TCT ACT AAT TGG ATA GAA TAT GTA TAT GTA TGT TAC GTA TTA GAC AAA AEA	909
	<u>CAA CTT TCA TCT CTA ATT AAT TGC GTA TAC ATA TGG CTG CTG TGT GTC TGT TAA CC</u>	926

Fig. 1. The nucleotide sequence and the deduced amino acid sequence of *CanMADS1* (A) and *CanMADS6* (B). The MADS domain is underlined, and the K domain is double underlined. **A.** The C-terminal amino acid sequence conserved in other AGL2 subfamily proteins is indicated in bold and underlined. **B.** Arrowheads and numbers below the amino acid sequences indicate positions of the first amino acid of the fusion proteins selected by the yeast two-hybrid screening method, and the number of selected clones with the same first amino acid, respectively. The C-terminal amino acid sequence conserved in other SQUA subfamily proteins is indicated in bold and underlined. The GenBank accession numbers of *CanMADS1* and *CanMADS6* are AF129875 and AF130118, respectively.

A

```
MGRGRVLEKRIENKINRQVTFKRRNGLLKKAYELSVLCEAEVALIIFSNRGLYEFCS
***K*****
***M*****K**GS**T*-C*INEM**S*QD**K*****V*****PPW**E**P*
***K*****K**GS**T*-C*INEM**S*QD**E*****V*****L**E**P*
***N**L**LDR**K**GSI**V**NKP**KELE**S**R**K**G**Y**N**Q**L**P**
```

CanMADS1
EGM3
PrMADS1
AGL2

```
SSMVKTIKQYRCSYATLEAN-QSATDTQNNYHEVLRKARVELLORSORNFGLGDLGTL
***M*****K**GS**T*-C*INEM**S*QD**K*****V*****PPW**E**P*
***M*****K**GS**T*-C*INEM**S*QD**E*****V*****L**E**P*
***N**L**LDR**K**GSI**V**NKP**KELE**S**R**K**G**Y**N**Q**L**P**
```

CanMADS1
EGM3
PrMADS1
AGL2

```
STKDLQLEQLNQLKQIRSRKTOFMLDQLADLQOREQMLAESNRLLRKLEEST-AGF
NS*E****H***N*****A*****F***XH**HK****V*A**E*WK*****N-TRI
NS*E****H***N*****A*****F***H**HK****V*A**E*WK*****N-TRI
NS*E****R**DG**V**I**Y**S**S**NK****L**T**A**AM**DDMIGVRS
```

CanMADS1
EGM3
PrMADS1
AGL2

B

```
MGRGRVQLRRIENKINRQVTFKRRSGLLKAHEISVLCDAEVGLIVFSSKGLFEYSTD
*****K*****T*****D*****T*****A**
*****K*****T*****D*****T*****A**
*****K*****T*****D*****T*****A**
*****R**K*****G*****L*****A*****N*****
```

CanMADS6
POTM1-1
SLM5
TM4
SQUA

```
SCMERILERYERYSYAERQLNATDVETPGSWLEHAKLKALEVLQORNOHYAGEDLDSI
*****L*****F*****VP**HTS*****K**V*****E**
***K*****T**P**PDSHV*****I**K**H**M**T**
*****F**K**VP**HTS**V*****R*****K**V*****E**
***D*****K*****F*****VSNPEQS**AN**Y**S*****I**L**H**M*****M
```

CanMADS6
POTM1-1
SLM5
TM4
SQUA

```
SMKELONLEOQDLSALKHRSRKNLQMHESIETLQKDKALQEQNNLSKQMKEREKQLA
N*****H*****V**Q**R*****Q**K*****EVA
*L*****F**H**T*****K*****Y**H*****H**T**K**V**K**EKE
*****H*****V*****R*****Q**K*****SAQ
*L**I**S*****T**N**T*****LY**H**I**T**M**A**K**I**K**E**I**
```

CanMADS6
POTM1-1
SLM5
TM4
SQUA

Fig. 2. The deduced amino acid sequence comparisons of CanMADS1 (A) and CanMADS6 (B) with other MADS proteins. The region containing the MADS domain, I-region, and K domain is compared. **A.** Alignment of the amino acid sequences of CanMADS1, EGM3 (*Eucalyptus grandis*; Southerton *et al.*, 1998), PrMADS1 (*Pinus radiata*; Mouradov *et al.*, 1998), and AGL2 (*Arabidopsis thaliana*; Ma *et al.*, 1991). **B.** Alignment of the amino acid sequences of CanMADS6, POTM1-1 (*Solanum tuberosum*; Kang and Hannapel, 1996), SLM5 (*Silene latifolia* syn *Melandrium album*; Hardenack *et al.*, 1994), TM4 (*Lycopersicon esculentum*; Pnueli *et al.*, 1991), and SQUA (*Antirrhinum majus*; Huijser *et al.*, 1992). The MADS box regions are underlined, and the K domains are double underlined. Asterisks indicate identical amino acid residues. Dashes indicate gaps, which were introduced to maximize alignments.

the *AGL2* subfamily (Theißen *et al.*, 1996). The *AGL2* subfamily includes: *AGL2* and *AGL4* of *Arabidopsis* (Ma *et al.*, 1991), *DEFH49* and *DEFH72* of *A. majus* (Davies *et al.*, 1996), *FBP2* of petunia (Angenent *et al.*, 1994), *TM5* of tomato (Pnueli *et al.*, 1994), *EGM1* and *EGM3* of eucalypts (Southerton *et al.*, 1998), *PrMADS1*, *PrMADS2*, and *PrMADS3* of Monterey pine (Mouradov *et al.*, 1998), and *OsMADS1*, *OsMADS5*, *OsMADS7*, and *OsMADS8* of rice (Chung *et al.*, 1994; Kang and An, 1997; Kang *et al.*, 1997). Among these genes, *CanMADS1* was the most homologous to *PrMADS1* (68.0%), *EGM3* (66.9%), and *AGL2* (60.6%) (Fig. 2A).

Yeast two-hybrid screening To identify proteins that interact with CanMADS1, yeast two-hybrid screening was

conducted. The method has proven to be very efficient in identifying interaction partners of MADS proteins. We initially made a fusion between the GAL4 binding domain and the CanMADS1 protein containing the K domain and C-terminal region.

We proceeded to introduce the cDNA expression library, constructed from the mRNA of floral buds, into the YRG-2 yeast strain containing pBD/CanMADS1-KC15. A total of 1.0×10^6 transformants was screened for their ability to grow on a histidine-lacking medium. This initial screening identified 60 colonies, which were subsequently tested for activation of the *LacZ* gene. These experiments resulted in the identification of 23 colonies that activated both *HIS3* and *LacZ*. Plasmid DNAs were prepared from these colonies and retransferred into the YRG-2 strain in order to confirm whether the activation was indeed due to the presence of the fusion protein. We observed that 19 plasmids were able to activate the *LacZ* gene, but only in the presence of pBD/CanMADS1-KC15.

All of the 19 cDNA clones that were selected by the two-hybrid screening method were partial; they lacked the 5' region that encodes for the N-terminal end of the protein. A sequence determination of these clones revealed that all 19 plasmids contained an identical cDNA with different insert sizes (data not shown). This cDNA was designated *CanMADS6*. The 5' region of *CaMADS6* was isolated by PCR using a vector primer and two cDNA specific primers. The *CanMADS6* cDNA clone is 956 bp long and contained a 54 bp 5' untranslated region and an open reading frame of 247 amino acid residues (GenBank accession No. AF130118). The CanMADS6 protein contains a MADS domain, which consists of 56 conserved amino acids (Figs. 1B and 2B). The K domain is present in between the amino acid residues 93th to 158th. The C-terminal region of CanMADS6 has one motif (PPW-MLRHLNG) that is conserved in POTM1-1 (Kang and Hannapel, 1996) and ZAP1 (Mena *et al.*, 1995). As a result of the amino acid sequence comparison, CanMADS6 shows a high sequence similarity to SQUA subfamily genes, such as SLM5 (Hardenack *et al.*, 1994), TM4 (Pnueli *et al.*, 1991), POTM1-1 (Kang and Hannapel, 1996), and SQUA (Huijser *et al.*, 1992).

Expression patterns of *CanMADS1* and *CanMADS6* It has been well established that there is a large number of MADS genes in each plant genome. Therefore, in order to study the expression patterns of *CanMADS1* and *CanMADS6*, the region that does not cross hybridize with other MADS genes was identified by genomic DNA blot analyses. The 220 bp (nucleotides 590–810) and 250 bp (nucleotides 620–870) fragments, which are located at the C-terminal regions of *CanMADS1* and *CanMADS6*, respectively, were used as gene-specific probes. A *CanMADS1* transcript was presented in flower buds and fruits, and the transcript level was the highest in the stage of fruit

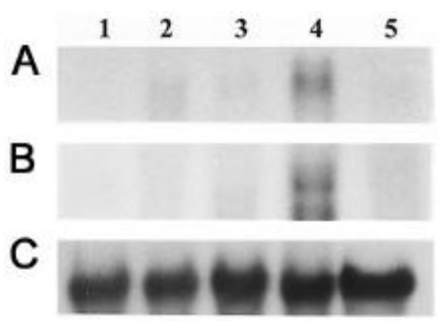


Fig. 3. RNA blot analysis of *CanMADS1* (A) and *CanMADS6* (B) in flower and fruit development. Lane 1, leaves; lane 2, young flower buds (bud length = 0.5–2 mm); lane 3, mature flowers (petals beginning to emerge from sepals); lane 4, flowers at fruit set stage (2 d after pollination); lane 5, young fruit (length = 7–10 mm). Equal loading was confirmed by re-hybridizing with a labeled ribosomal RNA probe (C).

set (2 d after anthesis). *CanMADS6* showed the same expression pattern as *CanMADS1*. *CanMADS1* and *CanMADS6* were not expressed in leaves (Fig. 3).

Ectopic expression of *CanMADS1* We used a transgenic approach to study the influence of the expression of *CanMADS1* on the development of the plants. Because it is not well established on the pepper transformation, we employed a heterologous tobacco system. The *CanMADS1* cDNA was placed under the control of the CaMV35S promoter. Seventeen independent transformants were regenerated from two cultivars, SR1 and Xanthi. Among the seventeen primary transformed lines (T1 generation), we selected one SR1 and three Xanthi plants that expressed high levels of the *CanMADS1* transcript. An ectopical analysis was done in the offspring of self-pollinated primary transgenic lines. The kanamycin resistance of the T2 plant showed a co-segregation with the kanamycin resistance gene; approximately a 75% kanamycin resistance (data not shown). The T2 transgenic plants of the SR1 line, of which all progenies showed kanamycin resistance, flowered 12 to 15 d earlier than the control plants without the transgene (Table 1). The T2 transgenic plants were 15 and 25 cm in height, which is 38 to 55 cm shorter than the control plants (Fig. 4). Similar results were obtained from the three Xanthi lines. The kanamycin-resistant T2 transgenic Xanthi lines flowered 9 to 14 d earlier, and the height was reduced by 14 to 31 cm compared to the control plants.

Discussion

We isolated the *CanMADS1* gene by screening a floral

Table 1. Phenotypes of transgenic tobacco plants with the *CanMADS1*. The progenies from the primary transformed line were analyzed. Progeny carrying the transgenes were identified by visually scoring T₃ seedling for kanamycin resistance.

Cultivar	T1 transgenic line	T2 transgenic line	Days to flowering ^a	Height (cm) ^b
SR1	S1	S1-1	48	15
		S1-2	50	25
		S1-3 ^c	62	63
Xanthi	X1	X1-1	70	58
		X1-2	71	62
		X1-3 ^c	84	79
	X2	X2-1	71	64
		X2-2	69	54
		X2-3 ^c	82	81
	X3	X3-1	69	57
		X3-2	70	62
		X3-3 ^c	80	78

^a Days to flowering include the time from seed germination to the first anthesis.

^b Height was measured when fruits were fully developed.

^c T₂ transgenic line that has not carrying the transgenes was used as control.



Fig. 4. Ectopic phenotypes of transgenic tobacco plant (cv.SR1). Homozygous transgenic line of S1-1 (right) and control plant (left).

bud cDNA library of hot pepper. In addition, we isolated a MADS-box gene, *CanMADS6* which interacts with *CanMADS1* by the yeast two-hybrid screening method.

Using the yeast two-hybrid system, it was revealed that

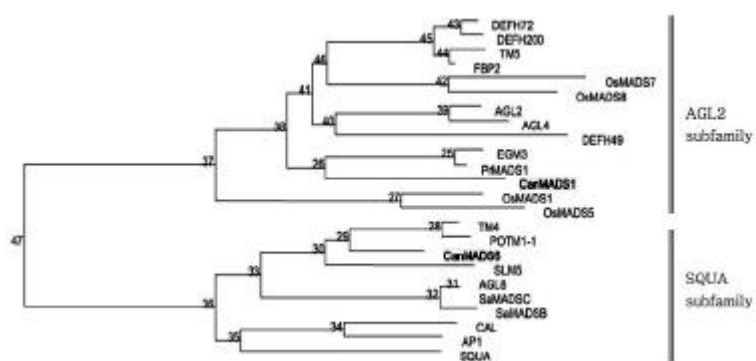


Fig. 5. A phylogenetic tree showing the relationship between the AGL2 subfamily and SQUA subfamily proteins. Hot pepper MADS box proteins are indicated by shaded boxes. The horizontal branches are proportional to the number of base substitutions. The numbers next to the nodes indicate percent divergences, shown for relevant nodes.

C class proteins, and AP1/AGL9 family proteins, interact with each other in *A. thaliana* and *A. majus* (Davies *et al.*, 1996; Fan *et al.*, 1997). In rice, it was demonstrated that the AP1/AGL9 family proteins interact with each other within the subfamily (Moon *et al.*, 1999b). Our results suggest that an AGL2 subfamily protein, CanMADS1, and a SQUA subfamily protein, CanMADS6, can also interact with each other in hot pepper.

The phylogeny of the MADS-box gene family revealed the existence of a distinct gene subfamily, which shares highly related functions and expression patterns (Purgganan *et al.*, 1995; Theißen *et al.*, 1996). Based on the phylogenetic analysis, the *CanMADS1* gene belongs to the AGL2 subfamily. Genes in the AGL2 subfamily show heterogeneous expression patterns, suggesting that they are functionally heterogeneous (Theißen *et al.*, 1996). The expression pattern of *AGL2* in *Arabidopsis* suggests that *AGL2* may play a fundamental role in establishing all floral organs and ovules, seed coats, and developing embryos (Flanagan and Ma, 1994). The protein-protein interaction partner of *CanMADS1*, *CanMADS6*, can be classified into the SQUA subfamily (Fig. 5). *CanMADS6* has high sequence similarities with *AGL8*-homolog genes in the SQUA subfamily, such as *POTM1-1*, *SLM5*, and *TM4* (Theißen *et al.*, 1996). A mutant strain of the *AGL8/FRUITFULL* MADS-box gene displays a lack of coordinated growth of the fruit tissues, suggesting that this gene is required for cellular differentiation during fruit development (Gu *et al.*, 1998). In the present study, both *CanMADS1* and *CanMADS6* were expressed abundantly in the fruit set stage, suggesting their regulatory role in fruit development.

The ectopic expression of *CanMADS1* in heterologous tobacco plants caused early flowering and dwarf phenotypes in transgenic plants (Table 1). These effects of ectopic expression were also reported in other AGL2 subfamily and SQUA subfamily genes. For example, early flowering and dwarfism were observed in transgenic to-

bacco plants ectopically expressing the SQUA subfamily genes of apple (Sung *et al.*, 1999) and AGL2 subfamily of rice, such as *OsMADS1*, *OsMADS5*, *OsMADS7*, and *OsMADS8* (An and An, 2000; Kang and An, 1997; Kang *et al.*, 1997).

Our results, together with the RNA expression patterns and the yeast-two hybrid analysis, suggest that a protein-protein interaction between *CanMADS1* and *CanMADS6* may play a role in the fruit development of the hot pepper.

Acknowledgments This study was supported by a grant (08-03-A-02) from the Ministry of Science and Technology, Korea.

References

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
- An, K. and An, G. (2000) Overriding photoperiod sensitivity of flowering time by constitutive expression of a MADS-box gene. *J. Plant Biol.* **43**, 143–148.
- An, G., Ebert, P. R., Mitra, A., and Ha, S. B. (1988) Binary vectors. Gelvin, S. B. and Schilperoort, R. A. (eds.), pp. A3: 1–19, *Plant Molecular Biology Manual*, Dordrecht, Kluwer Academic Publishers.
- Angenent, G. C., Franken, J., Busscher, M., Weiss, D., and van Tunen, A. J. (1994) Co-suppression of the petunia homeotic gene *fbp2* affects the identity of the generative meristem. *Plant J.* **5**, 33–44.
- Bradley, D., Carpenter, R., Sommer, H., Hartley, N., and Coen, E. (1993) Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the *plena* locus of *Antirrhinum*. *Cell* **72**, 85–95.
- Breedon, L. and Nasmyth, K. (1985) Regulation of the yeast HO gene. *Cold Spring Harbor Symp. Quant. Biol.* **50**, 643–650.
- Chung, Y.-Y., Kim, S.-R., Finkel, D., Yanofsky, M. F., and An, G. (1994) Early flowering and reduced apical dominance re-

- sult from ectopic expression of a rice MADS box gene. *Plant Mol. Biol.* **26**, 657–665.
- Davies, B., Egea-Cortines, M., de Andrade Silva, E., Saedler, H., and Sommer, H. (1996) Multiple interactions amongst floral homeotic MADS box proteins. *EMBO J.* **15**, 4330–4343.
- Fan, H.-Y., Hu, Y., Tudor, M., and Ma, H. (1997) Specific interactions between the K domains of AG and AGLs, members of the MADS domain family of DNA binding proteins. *Plant J.* **12**, 999–1010.
- Flanagan, C. A. and Ma, H. (1994) Spatially and temporally regulated expression of the MADS-box gene *AGL2* in wild-type and mutant *Arabidopsis* flowers. *Plant Mol. Biol.* **26**, 581–595.
- Gietz, D., St Jean, A., Woods, R. A., and Schiestl, R. H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* **20**, 1425.
- Goto, K. and Meyerowitz, E. M. (1994) Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. *Genes Dev.* **8**, 1548–1560.
- Gu, Q., Ferrandiz, C., Yanofsky, M. F., and Martienssen, R. (1998) The *FRUITFUL* MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. *Development* **125**, 1509–1517.
- Hardenack, S., Ye, D., Saedler, H., and Grant, S. (1994) Comparison of MADS box gene expression in developing male and female flowers of the dioecious plant white campion. *Plant Cell* **6**, 1775–1787.
- Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J., and Schilperoort, R. A. (1983). A binary vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature* **303**, 179–181.
- Hoffman, C. S. and Winston, F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* **57**, 267–272.
- Huang, H., Mizukami, Y., Hu, Y., and Ma, H. (1993) Isolation and characterization of the binding sequences for the product of the *Arabidopsis* floral homeotic gene *AGAMOUS*. *Nucleic Acids Res.* **21**, 4769–4776.
- Huang, H., Tudor, M., Weiss, C. A., Hu, Y., and Ma, H. (1995) The *Arabidopsis* MADS-box gene *AGL3* is widely expressed and encodes a sequence-specific DNA-binding protein. *Plant Mol. Biol.* **28**, 549–567.
- Huijser, P. W., Klein, J., Lonig, W.-E., Meijer, H., Saedler, H., and Sommer, H. (1992) Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene *squamosa* in *Antirrhinum majus*. *EMBO J.* **11**, 1239–1249.
- Jack, T., Brochman, L. L., and Meyerowitz, E. M. (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* **68**, 683–697.
- Kang, H.-G. and An, G. (1997) Isolation and characterization of a rice MADS box gene belonging to the *AGL2* gene family. *Mol. Cells* **7**, 45–51.
- Kang, H.-G., Jang, S., Chung, J.-E., Cho, Y.-G., and An, G. (1997) Characterization of two rice MADS box genes that control flowering time. *Mol. Cells* **7**, 559–566.
- Kang, S.-G. and Hannapel, D. J. (1996) A novel MADS-box gene of potato (*Solanum tuberosum* L.) expressed during the early stages of tuberization. *Plant Mol. Biol.* **31**, 379–386.
- Ma, H., Yanofsky, M. F., and Meyerowitz, E. M. (1991) *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev.* **5**, 484–495.
- Mandel, M. A. and Yanofsky, M. F. (1995) The *Arabidopsis AGL8* MADS box gene is expressed in inflorescence meristems and is negatively regulated by *APETALA1*. *Plant Cell* **7**, 763–1771.
- Mandel, M. A. and Yanofsky, M. F. (1998) The *Arabidopsis AGL9* MADS box gene is expressed in young primordia. *Sex. Plant Reprod.* **11**, 22–28.
- Mandel, M. A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1992) Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature* **360**, 273–277.
- McGonigle, B., Bouhidel, K., and Irish, V. F. (1996) Nuclear localization of the *Arabidopsis* *APETALA3* and *PISTILLATA* homeotic gene products depends on their simultaneous expression. *Genes Dev.* **10**, 1812–1821.
- Mena, M., Mandel, M. A., Lerner, D. R., Yanofsky, M. F., and Schmidt, R. J. (1995) A characterization of the MADS-box gene family in maize. *Plant J.* **8**, 845–854.
- Mizukami, Y., Huang, H., Tudor, M., Hu, Y., and Ma, H. (1996) Functional domains of the floral regulator *AGAMOUS*: characterization of the DNA binding domain and analysis of dominant negative mutations. *Plant Cell* **8**, 831–845.
- Moon, Y.-H., Jung, J.-Y., Kang, H.-G., and An, G. (1999a) Identification of a rice *APETALA3* homolog by yeast two-hybrid screening. *Plant Mol. Biol.* **40**, 167–177.
- Moon, Y.-H., Kang, H.-G., Jung, J.-Y., Jeon, J.-S., Sung, S.-K., and An, G. (1999b) Determination of the motif responsible for interaction between the rice *APETALA1/AGAMOUS-LIKE9* family proteins using a yeast two-hybrid system. *Plant Physiol.* **120**, 1193–1203.
- Mouradov, A., Glassick, T. V., Hamdorf, B. A., Murphy, L. C., Marla, S. S., Yang, Y., and Teasdale, R. D. (1998) Family of MADS-box genes expressed early in male and female reproductive structures of Monterey pine. *Plant Physiol.* **117**, 55–61.
- Pnueli, L., Abu-Abeid, M., Zamir, D., Nacken, W., Schwarz-Sommer, Z., and Lifschitz, E. (1991). The MADS box gene family in tomato: temporal expression during floral development, conserved secondary structures and homology with homeotic gene from *Antirrhinum* and *Arabidopsis*. *Plant J.* **1**, 255–266.
- Pnueli, L., Hareven, D., Rounsley, S. D., Yanofsky, M. F., and Lifschitz, E. (1994) Isolation of the tomato *AGAMOUS* gene *TAG1* and analysis of its homeotic role in transgenic plants. *Plant Cell* **6**, 163–173.
- Purugganan, M. D., Rounsley, S. D., Schmidt, R. J., and Yanofsky, M. F. (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. *Genetics* **140**, 345–356.
- Riechmann, J. L., Krizek, B. A., and Meyerowitz, E. M. (1996a) Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins *APETALA1*, *APETALA3*, *PISTILLATA*, and *AGAMOUS*. *Proc. Natl. Acad. Sci. USA* **93**, 4793–4798.
- Riechmann, J. L., Wang, M., and Meyerowitz, E. M. (1996b) DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins *APETALA1*, *APETALA3*, *PISTILLATA*,

- and AGAMOUS. *Nucleic Acids Res.* **24**, 3134–3141.
- Rounsley, S. D., Ditta, G. S., and Yanofsky, M. F. (1995) Diverse roles for MADS box genes in *Arabidopsis* development. *Plant Cell* **7**, 1259–1269.
- Savidge, B., Rounsley, S. D., and Yanofsky, M. F. (1995) Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. *Plant Cell* **7**, 721–733.
- Sommer, H., Beltran, J.-P., Huijser, P. W., Pape, H., Lonig, W.-E., Saedler, H., and Schwarz-Sommer, Z. (1990) *Deficiens*, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*. *EMBO J.* **9**, 605–613.
- Southerton, S. G., Marshall, H., Mouradov, A., and Teasdale, R. D. (1998) Eucalypt MADS-box genes expressed in developing flowers. *Plant Physiol.* **118**, 365–372.
- Sung, S.-K., Yu, G.-H., and An, G. (1999) Characterization of *MdMADS2*, a member of the *SQUAMOSA* subfamily of genes, in apple. *Plant Physiol.* **120**, 969–978.
- Sung, S.-K., Yu, G.-H., Nam, J., Jeong, D.-H., and An, G. (2000) Developmentally regulated expression of two MADS-box genes, *MdMADS3* and *MdMADS4*, in the morphogenesis of flower buds and fruits in apple. *Planta* **210**, 519–528.
- Theißen, G., Kim, J. T., and Saedler, H. (1996) Classification and phylogeny of the MADS-box gene families in the morphological evolution of eukaryotes. *J. Mol. Evol.* **43**, 484–516.
- Tilly, J. J., Allen, D. W., and Jack, T. (1998) The CARG boxes in the promoter of the *Arabidopsis* floral organ identity gene *APETALA3* mediate diverse regulatory effects. *Development* **125**, 1647–1657.
- Tröbner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lonig, W.-E., Saedler, H., Sommer, H., and Schwarz-Sommer, Z. (1992) *GLOBOSA*: a homeotic gene which interacts with *DEFICIENS* in the control of *Antirrhinum* floral organogenesis. *EMBO J.* **11**, 4693–4704.
- Weigel, D. and Meyerowitz, E. M. (1994) The ABCs of floral homeotic genes. *Cell* **78**, 203–209.
- West, A. G., Causier, B. E., Davies, B., and Sharrocks, A. D. (1998) DNA binding and dimerization determinants of *Antirrhinum majus* MADS-box transcription factors. *Nucleic Acids Res.* **26**, 5277–5287.
- Yao, J.-L., Dong, Y.-H., Kvarnheden, A., and Morris, B. (1999) Seven MADS-box genes in apple are expressed in different parts of the fruit. *J. Am. Soc. Hort. Sci.* **124**, 8–13.