Journal of Plant Biology, March 1999, 42(1) : 81-84 **PRELIMINARY REPORT**

Identification of MADS Genes from a Brown Alga, *Sargassum fulvellum*

Jongmin Nam[1], Yoo Kyung Lee[2], Jung Hyun Oak[2], Gynheung An[1], and In Kyu Lee[2]*
[1]Department of Life Science, Pohang University of Science and Technology, Pohang 790-784, Korea
[2]Department of Biology, Seoul National University, Seoul 151-742, Korea

The conserved region of numerous MADS genes in gulfweed (*Sargassum fulvellum*) was cloned by PCR with degenerate primers. Analysis of seventy individual clones resulted in the identification of nineteen types of nucleotide sequences. There sequences encode portions of the MADS domain in four distinctive groups. Six clones belong to the AGAMOUS subfamily, ten to AGL2, and two to AGL12. The remaining one clone is distinctive and appears to be diverged from an ancestor of the AGL2 and AP1 groups. There were no A or B class MADS genes. These results suggest that, as found in land plants, MADS genes also play major roles in controlling the development of algae.

Keywords: alga, evolution, MADS gene, transcription factor

In most lower plants, there is a distinct alternation of haploid (n) and diploid (2n) generations. Unlike higher plants, most algae have longer haploid periods and as plants have evolved, the sporophyte generation have remained dominant in the life cycle (Albert, 1999). However, the exact mode and molecular basis of life cycle alternation are not yet known. Studying the regulatory genes that control alternation of the life cycle may help to clarify the issue.

It has been well established that a group of transcription factors that contain a conserved MADS domain plays a major role in controlling the reproductive-organ development of higher plants (Theiben et al., 1996; Tandre et al., 1998). These transcription factors have also been found in nonflowering plants as well as yeast, animals, and fungi (Shore and Sharrocks, 1995; Munster et al., 1997; Hasebe et al., 1998). Recently, it was found that MADS genes also play a role in the vegetative phase, i.e., in root development and repression of flowering (Zhang and Forde, 1998; Sheldon et al., 1999). Although MADS genes are ubiquitous in land plants, their presence in algae has not been exploited yet. Therefore, in this study we investigated MADS genes in a brown alga gulfweed (*Sargassum fulvellum*), which spends its life as a diploid sporophyte, like the higher plants.

Higher plants contain approximately thirty to forty different MADS genes in each plant species. MADS genes can be divided into subgroups based on functional roles and sequence similarity. Each MADS region is highly conserved and contains unique sequences typical to each subfamily.

*Corresponding author; fax +82-2-872-6481
e-mail inkyulee@plaza.snu.ac.kr

Partial clones of the MADS genes from gulfweed were obtained by PCR using gulfweed genomic DNA as a template and a pair of degenerate oligonucleotide primers encoding the peptides EIKRIEN and VLCDAEV, which are the most conserved sequences in the MADS domain. This resulted in amplification of 69 bp sequences (excluding primers) in the MADS region. Sequencing seventy individual clones resulted in the identification of nineteen different types (Fig. 1). As observed in land plants, encode amino acid sequences in the MADS box region were highly conserved. However, eight clones encode MADS domains that have not been reported in other living organisms.

The gulfweed MADS genes can be classified into four groups based on amino acid sequence similarities. The first group consists of six clones that belong to the AG subfamily. In higher plants, the AG subfamily genes determine the identity of gametophyte-producing organs. Since algae also produce gametophytes, it is intriguing to speculate that the alga AG homologs are involved in controlling gametophyte development. Twenty-three amino acid sequences of two gulfweed AG homologs (SfMADS12 and SfMADS13) are identical to those of Arabidopsis AG and rice OsMADS3, and are closely related to several MADS genes in the AG subfamily (Fig. 2). The other four genes (SfMADS14, SfMADS15, SfMADS16, and SfMADS17) showed deviations from other AG genes. In particular, replacement of asparagine with serine in SfMADS14, SfMADS15, and SfMADS16 is unique to algae. The fact that there are at least six AG homologs in gulfweed suggests that algal AG genes may play diverse role in alga development.

The MADS box domain is involved in DNA binding

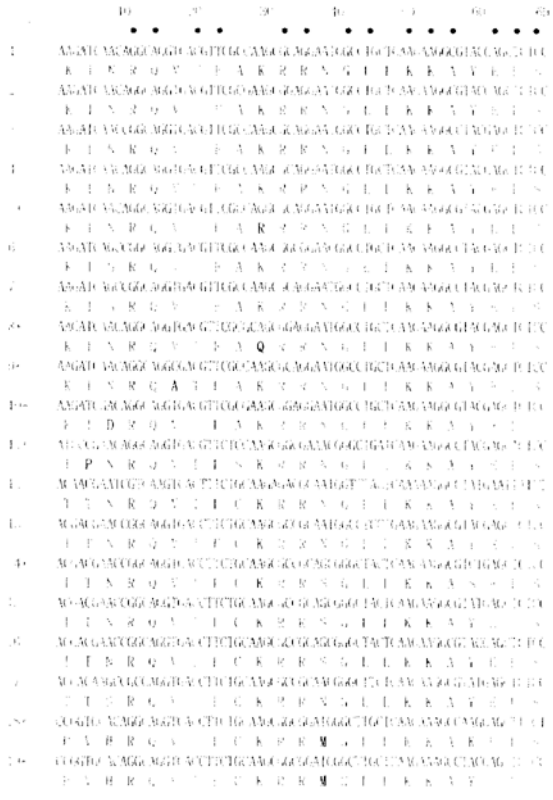

Figure 1. Nucleotide and deduced amino acid sequences of gulfweed MADS domain. The unique residues are marked with shaded box. Asterisks (*) indicate gulfweed MADS genes that encode for the amino acid sequences not found in land plants. Dots (·) indicate conserved amino acid sequences in SfMADS proteins. Forward primer 5'-GARAT(I/G)AARMG(I/G)AT(I/G)GARAAY-3' and reverse primer 5'-ACYTC(I/G)GCRTCRCA(I/G)AG(I/G)AC-3' (where R = A or G; I = inosine; M = A or C; Y = C or T) were utilized for PCR amplification of 69 bp sequences (excluding primers) in the MADS region using gulfweed genomic DNA as a template. PCR conditions were as follows: denaturation at 94°C for 3 min; 6 cycles of touchdown PCR, 94°C for 20 s; annealing (gradually reduced from 56°C to 51°C by 1°C/cycle) for 20 s and 72°C for 30 s; 30 cycles of amplification, 94°C for 20 s, 56°C for 20 s, and 72°C for 30 s; and final extension at 72°C for 5 min. After primary PCR, the corresponding band was reamplified with the following primers containing an enzyme site (*Eco*RI or *Hind*III) at the end; forward primer 5'-CGAGGAATTCGARATIAARMGGATMG-3' and reverse 5'-GGCGAAGCTTACYTCNGCRTCRCANAG-3' (where N = C, A, T, and G). Reamplified products were cut with *Eco*RI or *Hind* III and cloned into pBlueScript SK(-) vector (Stratagene, USA). These clones were sequenced by a BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA).

SRF	KLRRYTTFSKRKTGIMKKAYELS
DNA contacting	* * ** **** * **
Dimerization	* ** ** * * **
AG	TTNRQVTFCKRRNGLLKKAYELS
*Os*MADS3	------------------------
ZMM1	N-S----------------------
*Pr*MADS7	G-S----------------------
PLE	I-----------------------
AGL11	S-----------------------
FBP7	N-----------------------
ZEM2	----------S---------
*Sf*MADS12	------------------------
*Sf*MADS13	------------------------
*Sf*MADS14	----------S------S---
*Sf*MADS15	----------S----------
*Sf*MADS16	----------S----------
*Sf*MADS17	--S---------------------

Figure 2. The comparison of MADS-domains between gulfweed AG homologs and various AG subfamily members. Amino acid residues that are involved in dimerization and DNA contacting are indicated with asterisks (*). Dashes indicate identical residues with Arabidopsis AG.

and protein dimerization. Figure 2 shows that the amino acid residues involved in dimerization are conserved in the AG subfamily (Pellegrini et al., 1995; Theiben et al., 1996). However, some DNA-contacting amino acids are variable. These results imply that selective pressure on protein-protein interaction in the AG subfamily is more stringent than protein-DNA interaction. In the gulfweed AG homologs, substitutions occurred in residues not involved in DNA binding or dimerization.

A majority of the gulfweed MADS genes belongs to the AGL2 subfamily. The twenty-three amino acid sequences encoded by four gulfweed genes (SfMADS1, SfMADS2, SfMADS3, and SfMADS4) were identical to Arabidopsis AGL2 and other AGL2 family genes, including FBP2 and MdMADS1. The other six genes (SfMADS5, SfMADS6, SfMADS7, SfMADS8, SfMADS9, SfMADS10) were split into three distinct branches. Aspartic acid at the third amino acid residue (SfMADS10), alanine at the sixth residue (SfMADS9), and arginine and glutamine at the tenth residue (SfMADS5 and SfMADS8) are quite unusual in land plants. Functional roles of the AGL2 family have not

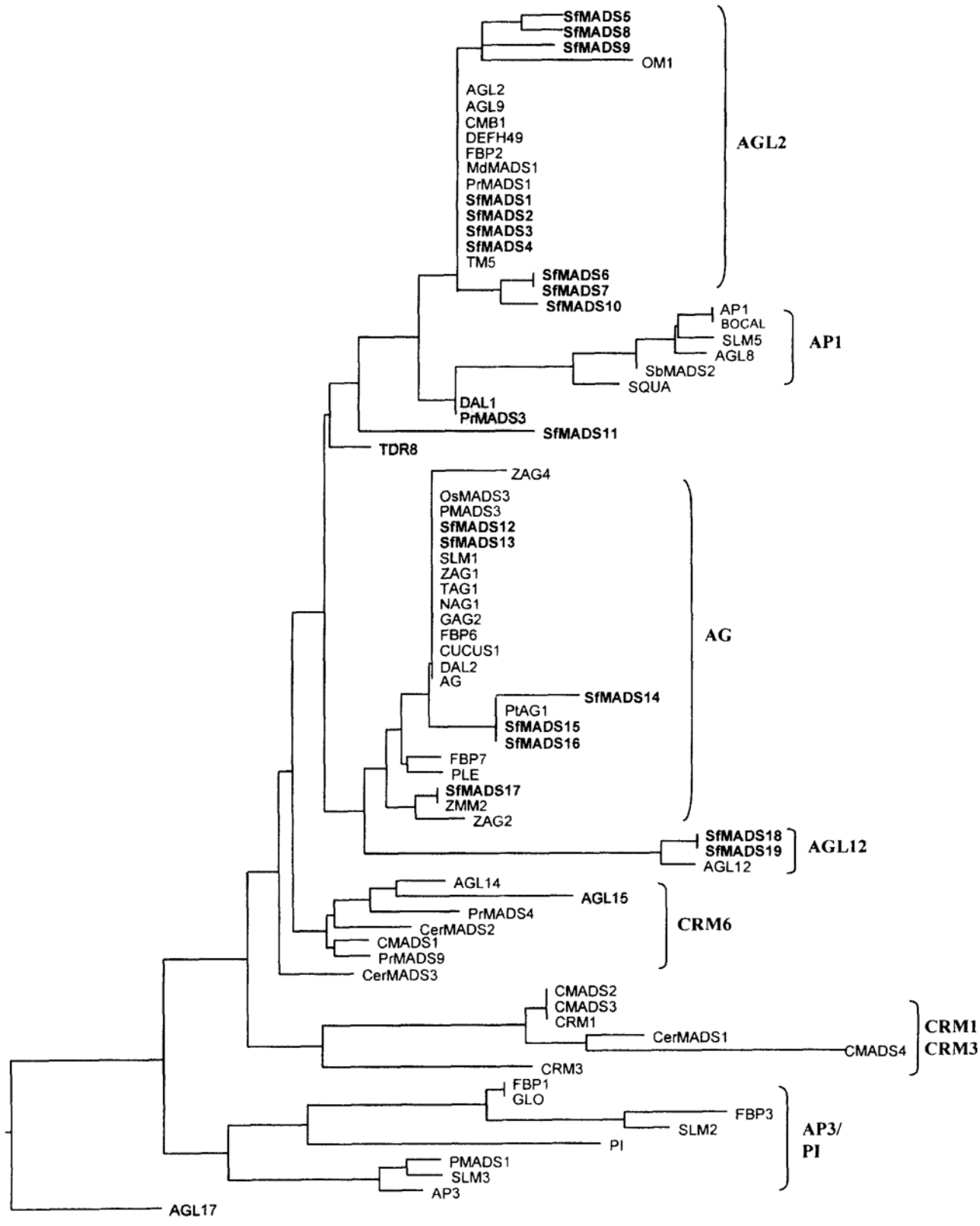

Figure 3. Phylogenetic tree of gulfweed and land plant MADS genes. The tree was established based on the neighbor-joining method using 23 amino acid sequences in the MADS box domain. The horizontal branch length is proportional to the estimated number of amino acids substitutions per residue. Subfamilies are labeled by brackets at the right margin.

been well elucidated, but it is speculated that genes in this family are involved in diverse roles during plant development since AGL2 genes are expressed in a wide variety of organs.

Two types of genes (SfMADS18 and SfMADS19) can be classified into the *AGL12* subfamily. The gulfweed MADS proteins differ from AGL12 by one amino acid at the 13th residue (methionine instead of threonine). It has been reported that *AGL12* was expressed preferentially in nonfloral tissues, such as roots in Arabidopsis (Rounsley et al., 1995). Therefore, gulfweed MADS genes may function in nonreproductive organ development.

The remaining one clone (SfMADS11) is unique and contains four substitutions from AGL2. This gulfweed gene appears to have diverged from an ancestor of AGL2 and AP1 family genes.

To elucidate the evolutionary relationship between the gulfweed MADS genes and other MADS genes from land plants, phylogenetic analysis was carried out using the 23 amino acid sequences within the MADS domain (Fig. 3). Although only 23 amino acid sequences were used for the analysis, the result is overall in agreement with the results of a previous analysis that was carried out with full-length MADS proteins with a few exceptions (Hasebe et al., 1998). In our analysis, we found that DAL1 and PrMADS3 are more closely related to AP1 than AGL2. In addition, TDR8, AGL15, and AGL17 are located differently. Despite their phylogenetic divergence of 380 million years between green plants and brown algae (Pearson, 1995), MADS genes are well conserved. Considering the resemblance of the life cycles of gulfweed and higher plants, this result seems to be relevant. It is proposed that phaeophyta and tracheophyta have common ancestors of AG, AGL2, and AGL12 subfamilies.

To further elucidate structure and function of these regulatory genes, it will be necessary to isolate the full-length cDNAs as well as genomic clones of the gulfweed MADS genes. It will be important to determine whether there are class A and B MADS genes in algae. Such efforts may add valuable information in tracing back plant evolution.

This research was supported in part by grants from the Korea Science and Engineering Foundation (96-0401-06-01-3) and from the Korea Agbiotech Consortium. We are grateful to Yang Suk Kim at the Biological Research Information Center for discussions on phylogenetic analysis.

Received March 15 1999; accepted March 20, 1999.

LITERATURE CITED

Albert VA (1999) Shoot apical meristems and floral patterning: an evolutionary perspective. Trends Plant Sci **39**: 84-86

Hasebe M, Wen CK, Kato M, Banks JA (1998) Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. Proc Natl Acad Sci USA **26**: 6222-6227

Munster T, Pahnke J, Di Rosa A, Kim JT, Martin W, Saedler H, Theiben G (1997) Floral homeotic genes were recruited from homologous MADS- box genes preexisting in the common ancestor of ferns and seed plants. Proc Natl Acad Sci USA **18**: 2415-2420

Pearson LC (1995) The diversity and evolution of plants. CRC Press, Inc, Boca Raton, pp 40-41

Pellegrini L, Tan S, Richmond TJ (1995) Structure of serum response factor core bound to DNA. Nature **376**: 490-498

Rounsley SD, Ditta GS, Yanofsky MF (1995) Diverse roles for MADS box genes in *Arabidopsis* development. Plant Cell **7**: 1259-1269

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES (1999) The *FLF* MADS Box Gene. A repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. Plant Cell **11**: 445-458

Shore P, Sharrocks AD (1995) The MADS-box family of transcription factors. Eur J Biochem **229**: 1-13

Tandre K, Svenson M, Svensson ME, Engstrom P (1998) Conservation of gene structure and activity in the regulation of reproductive organ development of conifers and angiosperms. Plant J **15**: 615-623

Theiben 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

Zhang H, Forde BG (1998) An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. Science **279**: 407-409