シダ類であるリチャードミズワラビ (Ceratopteris richardii) からのMADS遺伝子の単離と系統の解析

J. Phytogeogr. & Taxon.
45 : 83-91, 1997

Rumiko Kofuji*,** and Kazuo Yamaguchi*** : Isolation and Phylogenetic Analysis of MADS Genes from the Fern *Ceratopteris richardii*

小藤累美子*,**・山口和男***：シダ類であるリチャードミズワラビ（*Ceratopteris richardii*）からのMADS遺伝子の単離と系統の解析

Abstract

Plant MADS box genes were first recognized as floral homeotic genes. The MADS box genes are found in many flowering and a non-flowering seed plant and form a large gene family with several subfamilies. Most of these genes are expressed at floral or in similar organs. Considering the existence of the MADS box motif in taxonomically distant organisms, plants, animals and fungi, and its function as a transcriptional factor, it is possible that every plant lineage originally has this motif, participating in functions other than that of the formation of floral organs. In this study, we show the existence of MADS box genes (CerMADS) in a fern, *Ceratopteris richardii*. These genes retain the MADS box, the L region and the K box conserved in the MADS genes from flowering and non-flowering seed plants. They are expressed in sporophytes but not detected in gametophytes. Phylogenetic analyses indicated that these genes do not belong to any subfamilies found in seed plants; they form new subfamilies. The functions of CerMADS genes might be distinct from the known functions in seed plants because of the structural differences between seed plants and ferns.

Key words: cDNA cloning, *Ceratopteris richardii*, evolution, gene phylogeny, MADS gene.

Morphological evolution is one of the interesting subjects in the evolution of land plants. The evolution of morphological characters has been caused by changes of the regulational system. Detailed investigation of the molecular mechanisms are required for us to understand the principle of evolution. MADS genes containing the MADS box encode transcriptional regulators and this gene family is present throughout yeast, animals, and plants. MADS genes, isolated as floral homeotic genes from *Arabidopsis*, AGAMOUS (Yanofsky *et al.* 1990), and *Antirrhinum*, DEFICIENS (Sommer *et al.* 1990), were the first homeotic selector genes in plants. All of previously reported MADS genes of plants have a highly conserved MADS box, a K box downstream of the MADS box and an internal L region (Riechmann *et al.* 1996) between the two boxes. The K box is predicted to form amphipathic helices and may function as the protein dimerization domain (Ma *et al.* 1991). The L region consists of approximately 30 amino acid residues and is considered to determine the DNA binding specificity (Reichmann *et al.* 1996).

Many MADS genes have recently been isolated from various angiosperms including monocotyledons (reviewed in Theißen *et al.* 1996). It has been shown that the MADS gene family is a large group consisting of several subfamilies, including floral homeotic genes (Purugganan *et al.* 1995; Rounsley *et al.* 1995; Hasebe and Banks 1997). The subfamilies related to floral organ

*Division of Life Sciences, Graduate School of Natural Science and Technology, Kanazawa University, Takara-machi 13-1, Kanazawa 920, Japan 〒920 金沢市宝町13-1 金沢大学自然科学研究科生命科学専攻
**Present adress: National Institute for Basic Biology, 38 Nishigounaka, Myo-daiji-cho, Okazaki 444, Japan 〒444 岡崎市明大寺町字西郷中38 基礎生物学研究所
***Institute for Gene Research, Kanazawa University, Takara-machi 13-1, Kanazawa 920, Japan 〒920 金沢市宝町13-1 金沢大学遺伝子実験施設

formation arose before the separation of dicots and monocots (Hasebe and Banks 1997), because these subfamilies exist in the both taxonomic groups. Recently, MADS genes were also discovered which might be functioning in non-floral organs, e.g. root, considering from their expression pattern (Rounsley *et al.* 1995).

Leptosporangiate ferns are a monophyletic group and had diverged from seed plant lineage before the ancestors of gymnosperms and angiosperms divided (Hasebe *et al.* 1993). Ferns do not have floral organs. Most of ferns have diploid sporophytes to wear sporangia in sori on the abaxial surface or margin of the fronds. Sporangia are regarded as being homologous to the nucelli of seed plants. Since MADS genes are present in yeast (Christ and Tye 1991), humans (Norman *et al.* 1988; Yu *et al.* 1992), gymnosperms and angiosperms (Purugganan *et al.* 1995; Tandre *et al.* 1995), it is expected that MADS genes also exist in pteridophytes. If so, it would be interesting to investigate MADS genes in ferns.

One of the homosporous ferns, *Ceratopteris richardii* Brongn. (Lloyd 1974), is a tropical and annual fern. In the sporophytes, the shape of the fronds changes during successive development from vegetative to reproductive sporophytes. Fertile fronds are rolled up to enclose some rows of sporangia on abaxial surface. The gametophytes of *Ceratopteris* have two sexual types, hermaphrodite and male (Hickok *et al.* 1987). *Ceratopteris richardii* was proposed as a model plant, because of its short sexual life cycle (spore to spore), namely less than 120 days under the controlled condition (Hickok *et al.* 1987; Chasan 1992).

In this study, we isolated the cDNA containing MADS box sequence originated from *C. richardii* and showed that several MADS genes are present in *C. richardii*. All of them, named CerMADS, have characteristics in their sequences similar to MADS genes of the seed plants retaining the K box and the L region, but these genes may form new subfamilies.

Materials and Methods

Plant materials

Spores of *C. richardii* Hn strain were germinated on rock wool containing 0.01% of Hyponex 5-10-5 (Murakami Bussan Co.). Two weeks old hermaphroditic gametophytes were transplanted to fresh rock wool containing 0.1% of Hyponex and were self-fertilized following an application of water. The gametophytes used for RNA extraction were a mixture of male and hermaphrodite individuals before fertilization. The sporophytes were cultured under the same conditions over six months.

cDNA cloning and sequencing

Total RNA was extracted using ISOGEN (Nippon Gene). The cDNA synthesis and a rapid amplification of cDNA ends (RACE) method was performed according to the instructions of a 3' Amplifinder RACE Kit (Clontech).

For the selection of the MADS box containing cDNA, we designed two degenerate primers covering the conserved region in the MADS box in-

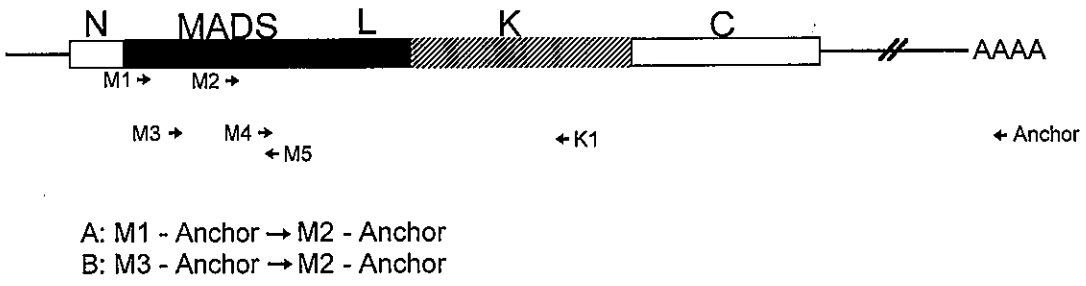

Fig. 1 Schematic structure of the MADS gene and locations of degenerate primers used in the 3' RACE experiments. In this schematic mRNA of the MADS box gene, the boxes indicate an open reading frame. The filled boxes (MADS, L and K) indicate conserved motifs, and the open boxes (N and C) indicate variable regions. The thin line is the untranslated region. Arrows indicate degenerate primers. "Anchor" means the anchor primer of the 3' Amplifinder RACE kit. A and B show successful primer sets.

Table 1. Degenerate primers used for 3' RACE

Name	Sequence (5' to 3')
M1	(G/C)A(A/G)(A/C)TNAA(A/G)(C/A)GGAT(C/A)GAGAAC
M2	AA(A/G)AA(A/G)GCNTA(C/T)GA(G/A)CTNTCNGT
M3	G(C/A)NACNAGCAG(A/G)CA(A/G)GTTACTTT
M4	TCATANAG(C/T)TTNCCTCTTGGA
M5	A(C/T)TN(A/G)TN(A/G)TNTT(T/C)TCTCCAAGAGG
K1	AANA(C/T)TTGN(A/G)(G/T)CTTTCGTGT

cluding SRF and MCM 1. Four degenerate primers were also designed based on the MADS box genes from conifer. The locations and sequences of all degenerate primers are shown in Fig. 1 and Table 1, respectively. The polymerase chain reactions (PCR) were carried out with 30 cycles of 45 sec at 94℃, 45 sec at 55℃ and 2 min. at 72℃. Degenerate primers for nested PCR have an additive *Bam*HI recognized sequence in 5' end. The anchor primer for NN_{-1}-oligo (dT) CDS attached to the 3' Amplifinder RACE Kit has an *Eco*RI recognized sequence at the 5' end. The amplified fragments were purified by agarose gel electrophoresis, digested with *Bam*HI and *Eco*RI, and ligated to pUC 19 digested with the same enzymes.

To amplify the 5' end of MADS cDNA of *C. richardii*, poly $(A)^+$ RNA was directly isolated from fresh leaves using the Dynabeads mRNA DIRECT kit (Dynal) according to the manufacturer's instructions. Reverse transcription (RT) of poly $(A)^+$ RNA followed by PCR (RT-PCR) was performed using the Marathon cDNA amplification kit according to the manufacturer's instructions (Clontech). A double stranded adapter sequence was ligated to the 5' end of the cDNA generated. Nested PCR was subsequently used to amplify specific fragments from the cDNA using two sense adapter primers and two antisense primers specific for each MADS cDNA sequence of *C. richardii*. Specific primers were designed to have at least 46% GC, based on MADS cDNA sequences of *C. richardii* with relatively diverse sequences obtained by the 3' RACE method.

Sequencing analysis was performed with ALF DNA Sequencer (Pharmacia) using Sequencing PRO (Toyobo) and FITC-labeled primers. Nucleotide sequence analyses were done using DNASIS-Mac (Hitachi Software Engineering).

Phylogenetic analyses

One hundred and seventeen deduced amino acid sequences were used for phylogenetic analyses. Five sequences were determined in this study and others were obtained from GENBANK and EMBL. These sequences were aligned using Higgins's multiple alignment program (Higgins *et al.* 1991) of CLUSTAL W (Thompson *et al.* 1994) by the default setting. The phylogenetic tree was constructed from the results of 100 bootstrap replicates using the neighbor-joining method (Saitou and Nei 1987) of CLUSTAL W.

Results

Using total RNA from sporophytes as the template in the 3' RACE experiments, PCR products were detected only with the use of two combinations A and B (Fig. 1) of degenerate primers. No products were obtained in any combination of primers with the use of RNA from gametophytes. Sequencing analyses showed that most cDNA clone obtained in this study had the MADS box. We named the MADS genes from *C. richardii* CerMADS (**Cer**atopteris richardii **MADS**), and numbered them 1 to 6. Since CerMADS 6 was obtained only its MADS box region without the L region and the K box from 3' RACE, no further description of this gene is shown in this study.

Fig. 2 shows the amino acid alignment including CerMADS genes determined in this study. The MADS box and the K box were specified from amino acid sequence similarity. All CerMADS genes have the MADS box, the K box and

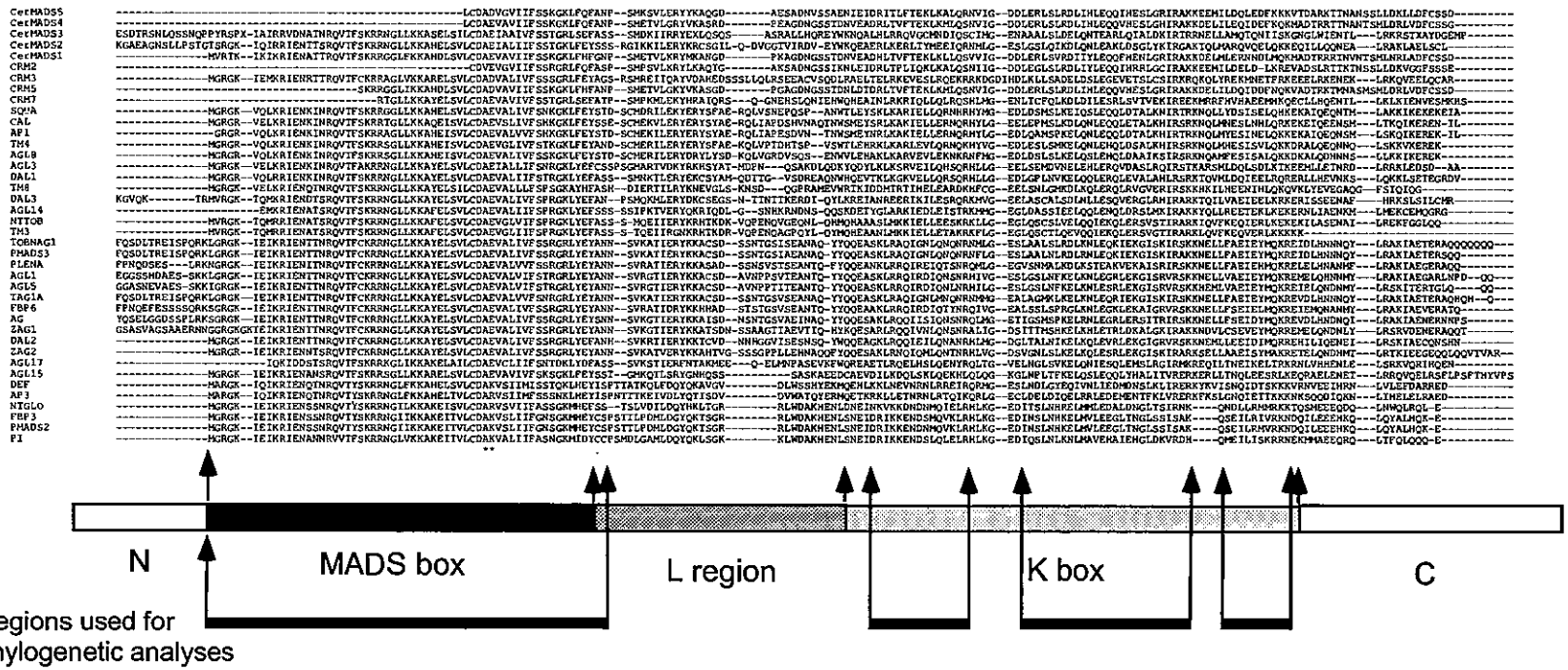

Fig. 2 Alignment of amino acid sequences of MADS genes used in this study. Only 40 sequences are shown. The boxes indicate motifs of MADS proteins corresponding to Fig. 1. The bold line indicates regions used for phylogenetic analyses. Gaps are written by "-". N and C regions are partly omitted from the alignment.

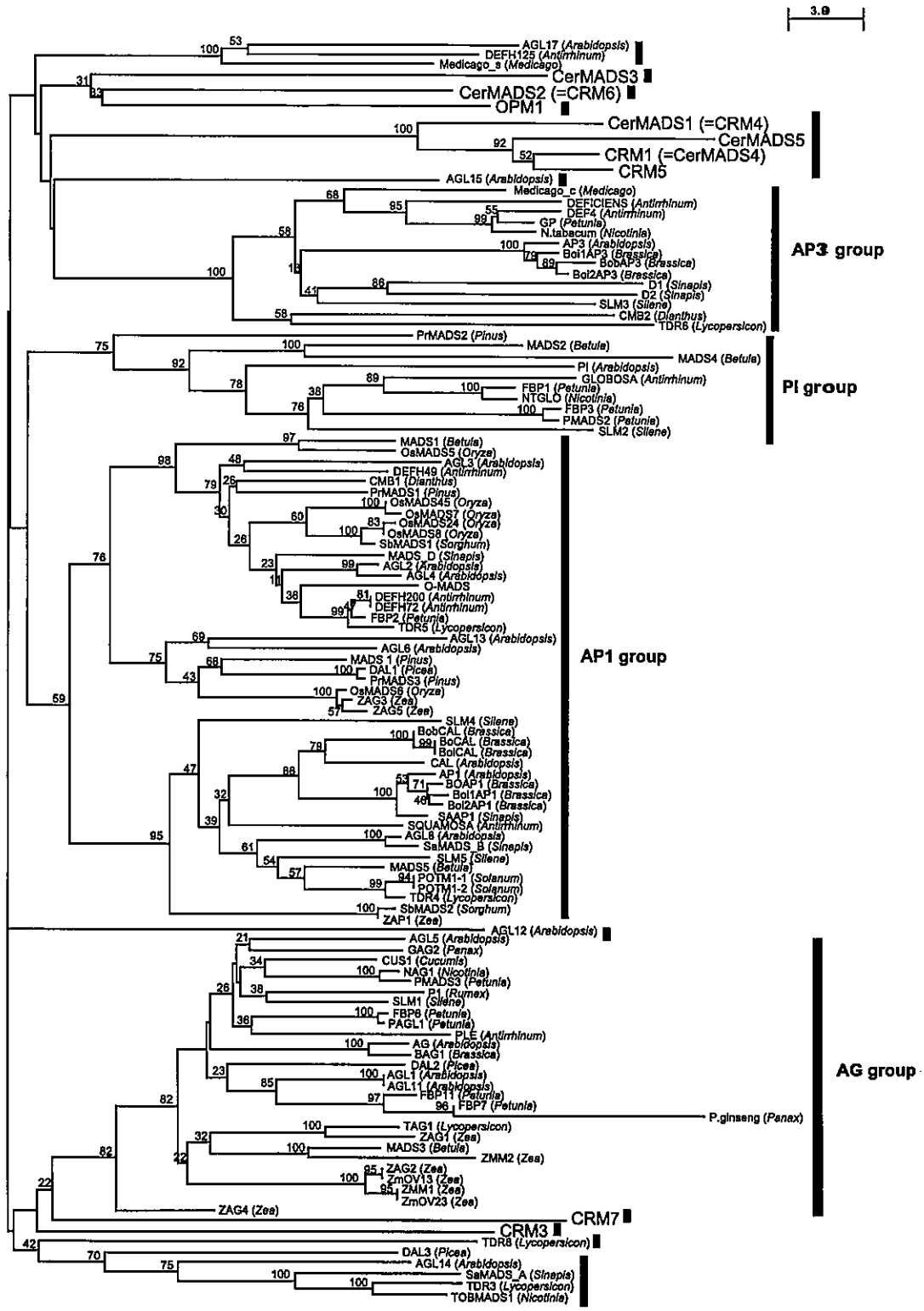

Fig. 3 Phylogenetic tree constructed by 117 MADS genes from plants. The numbers show bootstrap confidential level per 100 times for each node and numbers lower than 10 were omitted. Large letters mean MADS genes isolated from fern. The thick vertical bars indicate recognized subfamilies. The number on a scale bar means percentage divergence per 100 amino acid substitutions.

the internal L region, as do the other plant MADS genes. The amino acid sequence of CerMADS 1 was identical to that of CRM 4 (Münster *et al.* 1997), however, the latter sequence was incomplete in the N terminal portion of the MADS box. Similar case was seen between CerMADS 2 and CRM 6 (Münster *et al.* 1997). In case of CerMADS 4, it was identical to CRM 1, but its N terminal portion of the MADS box was incomplete. The coding sequence of CerMADS 1 presumably starts from the MADS box without any additive N terminal amino acids. Several splicing variations were seen in CerMADS 1 (Kofuji and Yamaguchi, unpublished data). One of them was in the 5' untranslated region like MEF 2 (Yu *et al.* 1992). Similar variations were also observed in other CerMADS genes. In the case of CerMADS 2, two ATG codons are recognized at 225 bp and 231 bp upstream from the MADS box. Such additional N terminal residues are also seen in AGAMOUS and its closely related genes (Bradley *et al.* 1993; Kempin *et al.* 1993; Robert *et al.* 1993; Pnueli *et al.* 1994) but a close sequence similarity between CerMADS 2 and AGAMOUS can not be seen. Concerning CerMADS 3, an ATG codon in frame did not yet appear upstream of the deduced MADS box. This is also similar to AGAMOUS (Yanofsky *et al.* 1990).

Phylogenetic analyses were done using 117 plant MADS gene sequences, including CerMADS genes determined in this study. Nine of them were from the fern, seven were from the gymnosperm and others from the angiosperm. Multiple alignment of deduced amino acid sequences was done using CLUSTAL W by the default option. After that, several parts of alignment (Fig. 2), regions from the MADS box to the K box except both N terminal and C terminal domains (Fig. 1), were used to construct phylogenetic trees. Because variable regions could not align precisely over all genes. A phylogenetic tree was constructed using the neighbor-joining (NJ) method by CLUSTAL W. Bootstrap analysis was done 100 times to examine the consistency of each node. The NJ tree and bootstrap probabilities are shown in Fig. 3. The branch lengths reflect the genetic distance. Fourteen consistent groups were recognized, though any relationships between each group were not clear.

Discussion

The cloning of the CerMADS genes found in this study showed that MADS genes are present in fern in spite of its distant phylogenetic position from seed plants. These genes also retain the K box and the L region in addition to the MADS box, like those of seed plants. Recently, Münster *et al.* (1997) also reported MADS genes (CRM 1 to 7) from *Ceratopteris* and a MADS gene (OPM 1) from *Ophioglossum* in which these regions were conserved. This evidence suggests that seed plant-type MADS genes existed in a common ancestor of ferns and seed plants.

Floral MADS genes of seed plants are clearly classified into three groups by their functions deduced from genetic analyses, class A; APETALA 1 (Mandel *et al.* 1992) and SQUAMOSA (Klein *et al.* 1996), class B; PISTILLATA (Goto and Meyerowitz 1994), APETALA 3 (Jack *et al.* 1992) and DEFICIENS and class C; AGAMOUS and PLENA (Bradley *et al.* 1993). Phylogenetic analyses also show that these are divided to different subgroups. In this study, all MADS genes from fern were constituted other subfamilies rather than any subfamilies of seed plant genes (Fig. 3). Presumable ancestor of these floral homeotic MADS genes could not be determined in genes from fern. For making clear the phylogenetic relationship of each subfamily, more primitive plant, e.g. bryophyte, should be investigated.

Nine MADS genes from fern were divided into six subfamilies. The CerMADS 1 subfamily contains a lot of genes and seems to be stepwise separated into some genes. CerMADS 2 and CerMADS 3 have additional coding regions upstream of the MADS box, and they were located in sister groups, although with low consistency. CRM 3 expressed in both sporophytes and gametophytes (Münster *et al.* 1997) was positioned adjacent to the AP 3 group in a phylogenetic tree constructed by Münster *et al.* (1997) but, in this study, they were detached. CRM 7, CRM 6 (= CerMADS 2) and OPM 1 was in a sister group in Münster et al. (1997) with high consistency (92 %). In this study, CRM 7 was, however, positioned adjacent to the AG (AGAMOUS) group

and OPM 1 belonged to other group.

All CerMADS we isolated were obtained from sporophytes. No RT-PCR products were detected from gametophytes, although equal amounts of RNA were used as the template. This was also the case when specific, not degenerated, primers for CerMADS genes were used (unpublished observation). This conflict might be caused by difference of the growth phase (11 days and 14 days or more) under different cultural conditions.

In this study, the relationship of MADS genes from ferns to those from seed plants, and floral homeotic genes to other subfamilies are still unclear. It is necessary to investigate MADS genes from more distant plants. To explicate the functions of CerMADS genes, further examinations, such as *in situ* hybridization, remain to be done. This study is the first step toward understanding the roles of MADS genes in land plants which have not yet acquired flowers.

We would like to thank Dr. Mitsuyasu Hasebe for valuable discussions and Kunihiko Ueda for reading the manuscript and valuable discussion, especially for the construction of the phylogenetic tree. We are grateful to Dr. Leslie G. Hickok for the gift of the *C. richardii* Hn strains and Dr. Y. Watano for advice. We also thank Dr. S. Sugiura and all members of our laboratory for stimulating discussions and technical assistance.

References

Bradley, D., Carpenter, R., Sommer, H., Hartley, E. 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.

Chasan, R. 1992. Ceratopteris: a model plant for the 90 s. Plant Cell **4**: 113–115.

Christ, C. and Tye, B. -K. 1991. Functional domains of the yeast transcription/replication factor MCM 1. Genes. Devel. **5**: 751–763.

Goto, K. and Meyerowitz, E. M. 1994. Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. Genes. Devel. **8**: 1548–1560.

Hasebe, M. and Banks, J. A. 1997. Evolution of MADS gene family in plants. *In* Iwatsuki, K. and Raven, P. H. (eds.): Evolution and Diversification of Land Plants, pp. 179–197. Springer-Verlag, Tokyo.

Hasebe, M., Ito, M., Kofuji, R., Ueda, K. and Iwatsuki, K. 1993. Phylogeny relationships of ferns deduced from *rbc*L gene sequence. J. Mol. Evol. **37**: 476–482.

Hickok, L. G., Warne, T. R. and Slocum, M. K. 1987. *Ceratopteris richardii*: applications for experimental plant biology. Amer. J. Bot. **74**: 1304–1316.

Higgins, D. G., Bleasby, A. J. and Funchs, R. 1991. Clustal W is a major re-write of the multiple alignment program Clustal V. CABIOS **8**: 189–191.

Jack, T., Brockman, L. L. and Meyerowitz, E. M. 1992. The homeotic gene *APETALA* 3 of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. Cell **68**: 683–697.

Kempin, S. A., Mandel, M. A. and Yanofsky, M. A. 1993. Conversion of perianth into reproductive organs by ectopic expression of the tobacco floral homeotic gene NAG 1. Plant Physiol. **103**: 1041–1046.

Klein, J., Seadler, H. and Hujiser, P. 1996. A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene SQUAMOSA. Mol. Gen. Genet. **260**: 7–16.

Lloyd, R. M. 1974. Systematics of the genus *Ceratopteris* Brongn. (Parkeriaceae) II. Taxonomy. Brittonia **26**: 139–160.

Ma, H., Yanofsky, M. F. and Meyerowitz, E. M. 1991. *AGL 1–AGL 6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. Genes. Devel. **5**: 484–495.

Mandel, M. A., Gustafson-Brown, C., Savidge, B. and Yanofsky, M. F. 1992. Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA 1*. Nature **360**: 273–277.

Münster, T., Pahnke, J., Rosa, A. D., Kim, J. T., Martin, W., Seadler, H. and Theissen, 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 **94**: 2415–2420.

Norman, C., Runswick, M., Pollock, R. and Treisman, R. 1988. Isolation and properties of cDNA clones encoding SRF, a transcription

factor that binds to the c-*fos* serum response element. Cell **55**: 989–1003.

Pnueli, L., Hareven, D., Rounsley, S. D., Yanofsky, M. F. and Lifschitz, E. 1994. Isolation of the tomato *AGAMOUS* gene *TAG 1* 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. 1996. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA 1, APETALA 3, PISTILLATA and AGAMOUS. Proc. Natl. Acad. Sci. USA **94**: 4793–4798.

Robert, J. S., Veit, B., Mandel, M. A., Mena, M. and Hakes, S. 1993. Identification and molecular characterization of ZAG 1, the maize homolog of the Arabidopsis floral homeotic gene AGAMOUS. Plant Cell **8**: 729–737.

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.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**: 406–425.

Sommer, H., Beltrán, J.–P., Hujiser, P., Pape, H., Lönnig, W. –E., Seadler, H. and Schwarzs-Sommer, Z. 1990. Deficiens, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. EMBO J. **9**: 605–613.

Tandre, K., Albert, V. A., Sundas, A. and Engström, P. 1995. Conifer homologues to genes the control floral development in angiosperms. Plant Mol. Biol. **27**: 69–78.

Theiβan,G., Kim., J. T. and Seadler, H. 1996. Classification and phylogeny of the MADS-box multigene family suggest define roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. J. Mol. Evol. **43**: 484–516.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. **22**: 4673–4680.

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A. and Meyerowitz, E. M. 1990. The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. Nature **346**: 35–39.

Yu, Y. –T., Breitbart, R. E., Smoot, L. B., Lee, Y., Mahdavi, V. and Nadal-Ginard, B. 1992. Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. Genes. Devel. **6**: 1783–1798.

摘　要

生物界全体に広く存在するMADS遺伝子は，植物では花器官形態形成遺伝子として知られている。すなわち典型的な双子葉植物が持つ4種類の花器官（がく片，花弁，雄蕊，雌蕊）形成が，MADS遺伝子を中心とする花器官形態形成遺伝子群（ABCの3種類に分けられる）の発現の組合せによって決定されるというものである。MADS遺伝子は，単子葉植物を含めた数多くの被子植物や裸子植物からも単離されており，いくつかのサブファミリーからなる遺伝子ファミリーを形成していることが知られている。それらのほとんどは花器官で発現しているが，最近の研究から，シュート全体や根などの花以外の器官で発現するものも得られてきている。

シダ植物であるリチャードミズワラビ（*Ceratopteris richardii*）を材料としてMADS遺伝子を単離したところ，栄養葉，胞子葉の両方を含む胞子体から6種類のMADS遺伝子（CerMADS）が得られた。これらはいずれもMADS boxの他にL region，K boxという種子植物のMADS遺伝子がもっており，ヒトや酵母のMADS遺伝子にはみられない特徴をそなえていた。このことは，種子植物型のMADS遺伝子が，少なくともシダ類と種子植物の祖先が分かれる前に出現していたことを示唆する。また，これらの遺伝子はいずれも前葉体では発現がみられなかった。

種子植物由来のMADS遺伝子とCerMADSとの関係をあきらかにするため，アミノ酸配列を用いて系統解析をおこなったところ，いずれのCerMADSに関しても，花器官形態形成に関与する遺伝子が属するサブファミリーを含む既存のどの遺伝子と最も近縁であるかはわからかなった。CerMADS 1，4，

5が属するサブファミリーは，リチャードミズワラビ由来の遺伝子のみからなり，系統関係が明確であることがわかった。しかし，他の2種類はいずれの遺伝子とも近縁であるとの結果は得られない。さらに，花器官形態形成に関与するサブファミリーと近縁であることが高い確率で支持されるようなMADS遺伝子は存在していなかった。このことは，これらの遺伝子の系統関係や機能を知るためには，種子植物とシダ類の祖先が分かれるより前にこれらの共通祖先から分かれた植物で，MADS 遺伝子の機能を解析する必要があることを示唆する。

(recieved June 20, 1997; accepted October 29, 1997)