J Plant Res (2003) 116:71–75
Digital Object Identifier (DOI) 10.1007/s10265-002-0071-5

ORIGINAL ARTICLE

Yoichi Tanabe · Makoto Uchida · Mitsuyasu Hasebe
Motomi Ito

Characterization of the *Selaginella remotifolia* MADS-box gene

Received: June 18, 2002 / Accepted: October 31, 2002 / Published online: December 19, 2002

Abstract Recent progress in plant molecular genetics has revealed that floral organ development is regulated by several homeotic selector genes, most of which belong to the MADS-box gene family. Here we report on *SrMADS1*, a MIKC^c-type MADS-box gene from *Selaginella*, a spikemoss belonging to the lycophytes. *SrMADS1* phylogenetically forms a monophyletic clade with genes of the *LAMB2* group, which are MIKC^c genes of the clubmoss *Lycopodium*, and is expressed in whole sporophytic tissues except roots and rhizophores. Our results and the previous report on *Lycopodium* MIKC^c genes suggest that the ancestral MIKC^c gene of primitive dichotomous plants in the early Devonian was involved in the development of basic sporophytic tissues such as shoot, stem, and sporangium.

Key words Lycophytes · MADS-box gene · MIKC^c-type · *Selaginella remotifolia*

Introduction

The determination and differentiation of floral organs are regulated mainly by members of the MADS-box gene family, which encode transcription factors and are characterized by the well-conserved MADS domain (Shore and Sharrocks 1995). MADS-box genes are phylogenetically divided into two types, I and II, which probably diverged before divergence of plants and metazoans (Alvarez-Buylla et al. 2000). Characteristic of type II MADS-box genes of land plants is the domain structure composed of four regions: MADS (M), the internal region between MADS and K (I), K (K), and the C-terminal region (C) (Ma et al. 1991). The K domain, the next best-conserved domain after MADS, has been reported only from land plant MADS-box genes (Shore and Sharrocks 1995). Comparison of exon-intron structures has revealed two different groups of type II MADS-box genes: MIKC^c and MIKC* (Henschel et al. 2002). All floral homeotic genes are MIKC^c. MIKC^c genes exist in a wide range of land plants including flowering plants and mosses. MIKC* genes are found in clubmosses and mosses (Svensson et al. 2000; Henschel et al. 2002), but the function of MIKC* genes has not been determined fully. MIKC^c genes play important roles not only in floral organ development but also in other developmental processes in sporophytes, such as the switch to flowering (Hartmann et al. 2000; Samach et al. 2000), lateral root differentiation (Zhang and Forde 1998), dehiscence of fruits, siliques (Gu et al. 1998), and others (Theißen et al. 2000).

Hypotheses on the origin and evolution of flowering plant MADS-box genes with such critical functions are based on studies of MADS-box genes in gymnosperms (Shindo et al. 1999; Winter et al. 1999) and ferns (Münster et al. 1997; Hasebe et al. 1998). All previously examined MADS-box genes of the fern *Ceratopteris richardii* are widely expressed in both vegetative and reproductive organs, and do not diverge as in flowering plants. This suggests that recruitment of some MADS-box genes expressed and functioning in specific organs occurred in the flowering plant lineage, which was probably important for the evolution of elaborate plant body plans (Hasebe 1999; Theißen et al. 2000). Characterization of MADS-box genes in other lower land plants is important for testing this hypothesis. Cladistic analyses of morphological data and molecular phylogenetic analyses of vascular plants indicate that lycophytes are the most basal lineage in extant vascular plants (Raubeson and Jansen 1992; Hiesel et al. 1994; Kranz et al. 1995; Kenrick and Crane 1997; Daff and Nickrent 1999;

Y. Tanabe (✉)
Graduate School of Science and Technology, Chiba University, Chiba 263-8522, Japan
Tel. +81-3-54546638
e-mail: ytanabe@moon.sannet.ne.jp

Y. Tanabe · M. Ito
Department of Multi-Disciplinary Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

M. Uchida
Department of Biology, Faculty of Science, Chiba University, Chiba, Japan

M. Hasebe
National Institute for Basic Biology, Okazaki, Japan

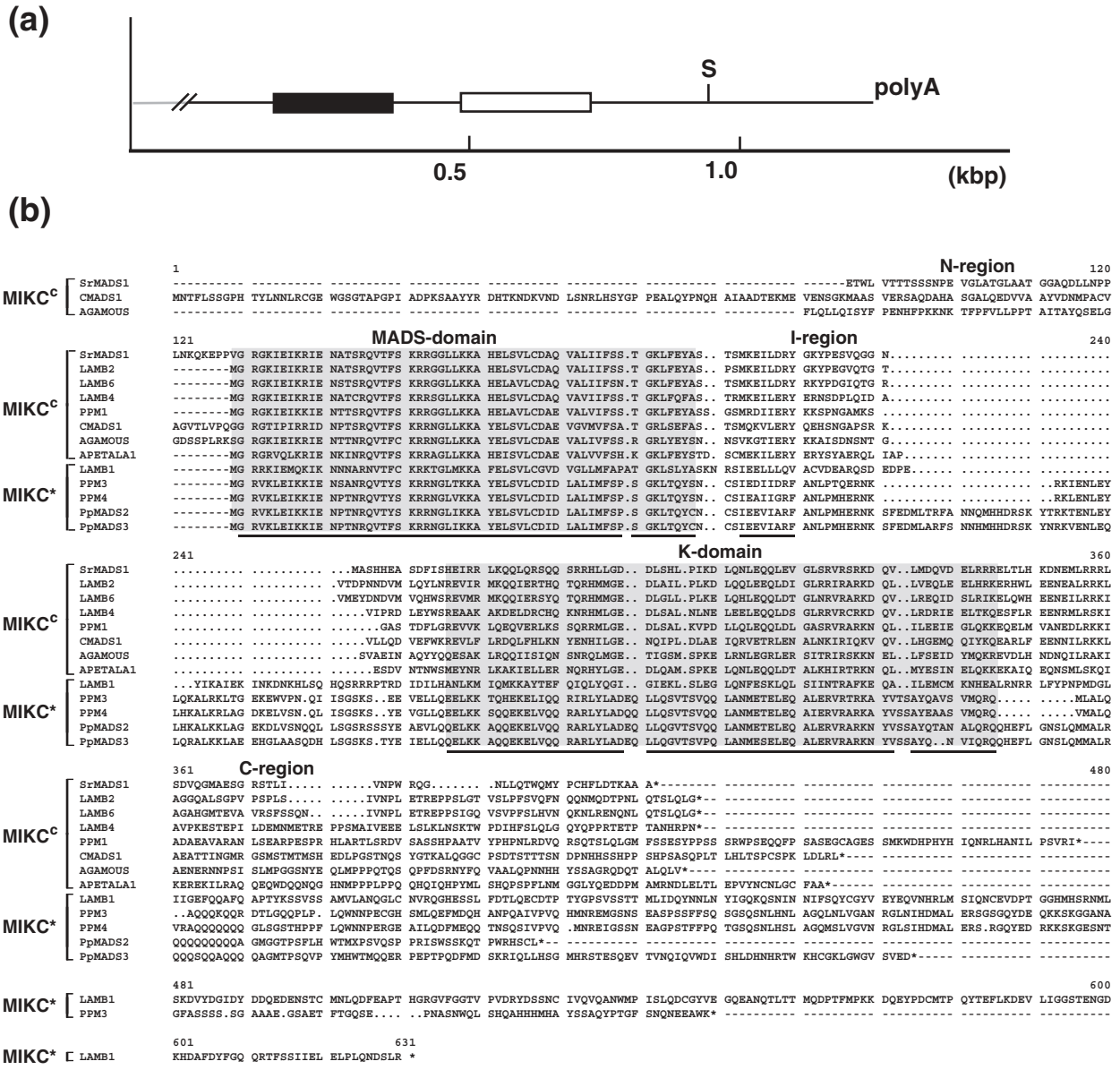


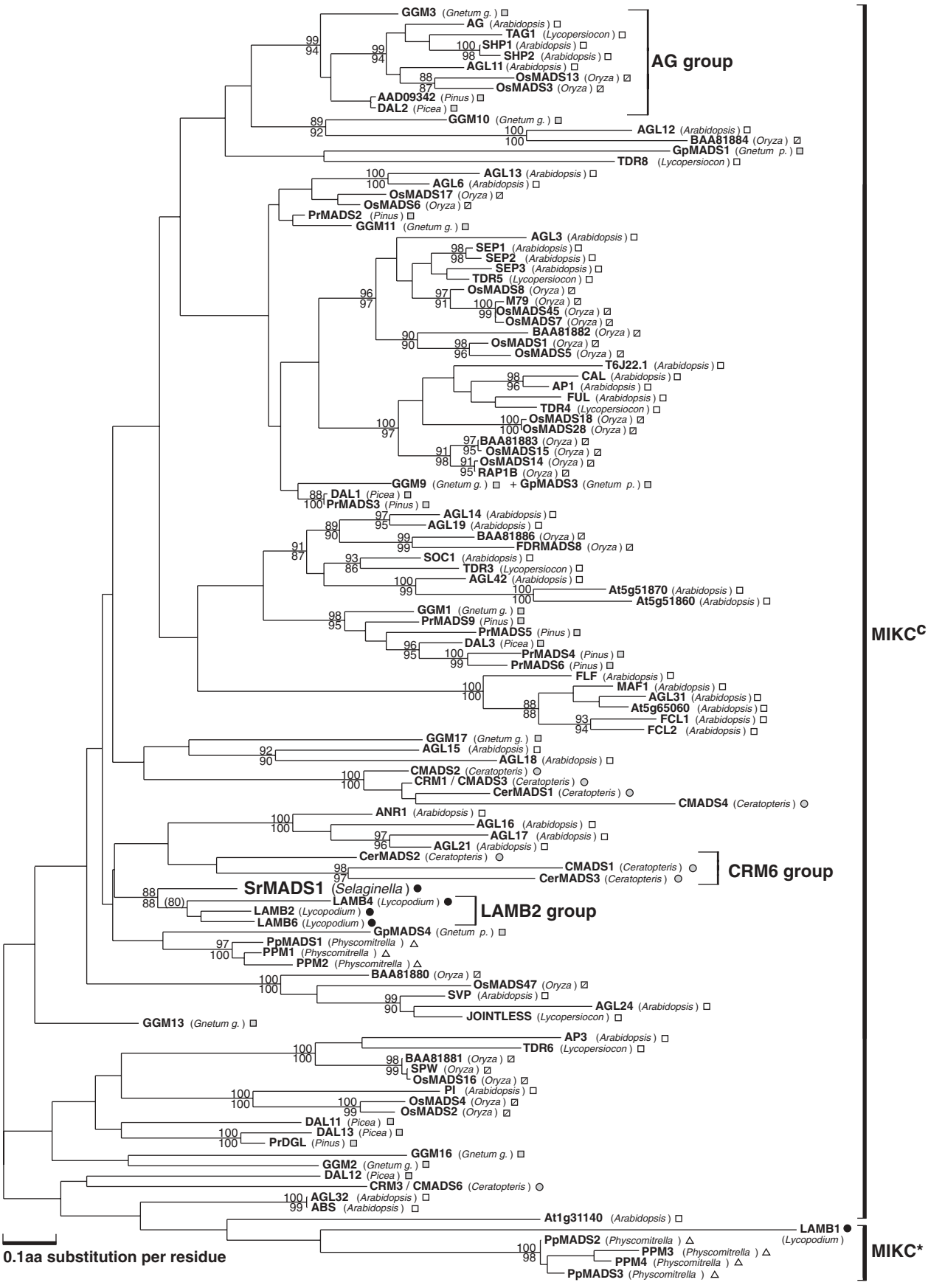
Fig. 1a,b. Cloning of the *SrMADS1* gene. **a** The structure of *SrMADS1* cDNA. The symbol *S* and *polyA* indicate the stop codon and poly(A)+tail, respectively. The *closed rectangle* on the architecture represents MADS-box and the *open rectangle* indicates a K-box. **b** Alignment of deduced amino acid sequences of SrMADS1 and representative MIKC genes of land plants: LAMB1, 2, 4, and 6 (*Lycopodium*), PPM1, 3, and 4, PpMADS2 and 3 (*Physcomitrella*), CMADS1 (*Ceratopteris*), and AGAMOUS and APETALA1 (*Arabidopsis*). Amino acid positions indicated by *bars* were used for the phylogenetic analysis

Pryer et al. 2001). In this present study, a MADS-box gene of the lycophyte spikemoss *Selaginella remotifolia* was cloned and its expression pattern was analyzed in order to determine the evolution of MADS-box genes and plant body plans.

Materials and methods

Selaginella remotifolia Spring was collected at Showanomori Park in Chiba Prefecture, Japan. The samples were

Fig. 2. A gene tree of MIKC-type MADS-box genes based on the maximum likelihood method. The *closed circles* indicate genes from lycophytes (*Selaginella remotifolia* and *Lycopodium annotinum*). *Open square*, *square filled with an angle bar*, *shaded square*, *shaded circle*, and *open triangle* represent dicots (*Arabidopsis thaliana* and *Lycopersicon esculentum*), monocots (*Oryza sativa*), gymnosperms (*Gnetum gnemon*, *G. parvifolium*, *Picea abies*, and *Pinus radiata*), ferns (*Ceratopteris richardii*), and mosses (*Physcomitrella patens*), respectively. The aligned 130 amino acid residues correspond to positions 2–58 (MADS domain), 63–70 (I region), 93–157 (K domain) from the initial methionine of APETALA1. Bootstrap values calculated with RELL method are represented above the nodes and those calculated with the NJ method are shown below. Only values over 85% with both ML and NJ are indicated


GGM3 (Gnetum g.)
AG (Arabidopsis)
TAG1 (Lycopersiocon)
SHP1 (Arabidopsis)
SHP2 (Arabidopsis)
AGL11 (Arabidopsis)
OsMADS13 (Oryza)
OsMADS3 (Oryza)
AAD09342 (Pinus)
DAL2 (Picea)
AG group
GGM10 (Gnetum g.)
AGL12 (Arabidopsis)
BAA81884 (Oryza)
GpMADS1 (Gnetum p.)
TDR8 (Lycopersiocon)
AGL13 (Arabidopsis)
AGL6 (Arabidopsis)
OsMADS17 (Oryza)
OsMADS6 (Oryza)
PrMADS2 (Pinus)
GGM11 (Gnetum g.)
AGL3 (Arabidopsis)
SEP1 (Arabidopsis)
SEP2 (Arabidopsis)
SEP3 (Arabidopsis)
TDR5 (Lycopersiocon)
OsMADS8 (Oryza)
M79 (Oryza)
OsMADS45 (Oryza)
OsMADS7 (Oryza)
BAA81882 (Oryza)
OsMADS1 (Oryza)
OsMADS5 (Oryza)
T6J22.1 (Arabidopsis)
CAL (Arabidopsis)
AP1 (Arabidopsis)
FUL (Arabidopsis)
TDR4 (Lycopersiocon)
OsMADS18 (Oryza)
OsMADS28 (Oryza)
BAA81883 (Oryza)
OsMADS15 (Oryza)
OsMADS14 (Oryza)
RAP1B (Oryza)
GGM9 (Gnetum g.) + GpMADS3 (Gnetum p.)
DAL1 (Picea)
PrMADS3 (Pinus)
AGL14 (Arabidopsis)
AGL19 (Arabidopsis)
BAA81886 (Oryza)
FDRMADS8 (Oryza)
SOC1 (Arabidopsis)
TDR3 (Lycopersiocon)
AGL42 (Arabidopsis)
At5g51870 (Arabidopsis)
At5g51860 (Arabidopsis)
GGM1 (Gnetum g.)
PrMADS9 (Pinus)
PrMADS5 (Pinus)
DAL3 (Picea)
PrMADS4 (Pinus)
PrMADS6 (Pinus)
FLF (Arabidopsis)
MAF1 (Arabidopsis)
AGL31 (Arabidopsis)
At5g65060 (Arabidopsis)
FCL1 (Arabidopsis)
FCL2 (Arabidopsis)
MIKC^C
GGM17 (Gnetum g.)
AGL15 (Arabidopsis)
AGL18 (Arabidopsis)
CMADS2 (Ceratopteris)
CRM1 / CMADS3 (Ceratopteris)
CerMADS1 (Ceratopteris)
CMADS4 (Ceratopteris)
ANR1 (Arabidopsis)
AGL16 (Arabidopsis)
AGL17 (Arabidopsis)
AGL21 (Arabidopsis)
CerMADS2 (Ceratopteris)
CMADS1 (Ceratopteris)
CerMADS3 (Ceratopteris)
CRM6 group
SrMADS1 (Selaginella)
LAMB4 (Lycopodium)
LAMB2 (Lycopodium)
LAMB6 (Lycopodium)
LAMB2 group
GpMADS4 (Gnetum p.)
PpMADS1 (Physcomitrella)
PPM1 (Physcomitrella)
PPM2 (Physcomitrella)
BAA81880 (Oryza)
OsMADS47 (Oryza)
SVP (Arabidopsis)
AGL24 (Arabidopsis)
JOINTLESS (Lycopersiocon)
GGM13 (Gnetum g.)
AP3 (Arabidopsis)
TDR6 (Lycopersiocon)
BAA81881 (Oryza)
SPW (Oryza)
OsMADS16 (Oryza)
PI (Arabidopsis)
OsMADS4 (Oryza)
OsMADS2 (Oryza)
DAL11 (Picea)
DAL13 (Picea)
PrDGL (Pinus)
GGM16 (Gnetum g.)
GGM2 (Gnetum g.)
DAL12 (Picea)
CRM3 / CMADS6 (Ceratopteris)
AGL32 (Arabidopsis)
ABS (Arabidopsis)
At1g31140 (Arabidopsis)
LAMB1 (Lycopodium)
PpMADS2 (Physcomitrella)
PPM3 (Physcomitrella)
PPM4 (Physcomitrella)
PpMADS3 (Physcomitrella)
MIKC*
0.1aa substitution per residue

divided into various tissues and preserved in a deep freezer. Total RNA extraction and 3′ and 5′ RACE were performed as described by Shindo et al. (1999). The nested MADS domain-specific primers, duMADS2–2 (5′-{CAU}$_4$AAR AARGCITAYGARCTIAGYGT-3′) and AllMADS2 (5′-{CAU}$_4$GARYTIWSIGTIYTITGYGAYGC-3′), were used to amplify the MADS-box clone.

To perform RT-PCR expression analysis, complementary DNAs were synthesized from total RNAs extracted from apices in the vegetative stage, strobili, microphylls, stems, and a mixture of rhizophores and roots. The PCR conditions were 1 cycle at 94°C for 1 min, followed by 30 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 1.5 min, and a final step at 72°C for 5 min. A PCR amplification test was performed with the *SrMADS1*-specific internal primers SrMF1 (5′-TCTAAACAAGCAAAAGGAACCGCC-3′) and SrMR1 (5′-CGAAGCATCTCATTGTCCTTGTG-3′). The *Selaginella remotifolia* ortholog of the 6-phosphogluconate dehydrogenase gene *Sr6PGD* (AB086022), which was constitutively expressed in all tissues examined, was used as a positive control. *Sr6PGD* was amplified by PCR with the *Sr6PGD*-specific primers (5′-TTTCTGAGCGG ACTCAAGGAGG-3′) and (5′-TAAGTGTGGGCACCA AAATAGTC-3′).

To construct a phylogenetic tree for MIKC-type MADS-box genes, amino acid sequences were obtained from EMBL/DDBJ/GenBank DNA databases and aligned using Clustal W, version 1.6 (Thompson et al. 1994). The maximum likelihood (ML) distances were calculated with ProtML under the conditions of the JTT model (Jones et al. 1992) and a Neighbor-Joining (NJ) tree was obtained with Njdist (Adachi and Hasegawa 1992–1996). The tree was further analyzed by local rearrangement search using ProtML to obtain the ML tree.

Results and discussion

A cDNA of *SrMADS1* (AB086021), a MADS-box gene, was isolated from *Selaginella remotifolia* by 3′ and 5′ RACE using MADS domain-specific degenerate primers. Other MADS-box genes were not found, despite the use of three other primers and various PCR conditions. Start codons exist in most MIKC genes adjacent to the N-terminal of MADS domains, but putative start codons were not found between 126 bp of the 5′ region to the first nucleotide of the MADS-box of *SrMADS1* (Fig. 1a). Comparison of deduced amino acid sequences of *SrMADS1* with other MIKC^c and MIKC* genes shows *SrMADS1* to be a typical MIKC^c-type MADS-box gene, lacking an extended I region (Fig. 1b).

To examine the phylogenetic relationship of *SrMADS1* to other MADS-box genes, an ML tree was constructed using representative MADS-box sequences from dicots, monocots, gymnosperms, leptosporangiate ferns, lycophytes, and mosses. *LAMB2–6*, which are MIKC^c genes of *Lycopodium* (a lycophyte clubmoss), form a clade as the *LAMB2* group and *LAMB3* and *5* closely related to *LAMB4* and *6*, respectively (Svensson and Engström 2002).

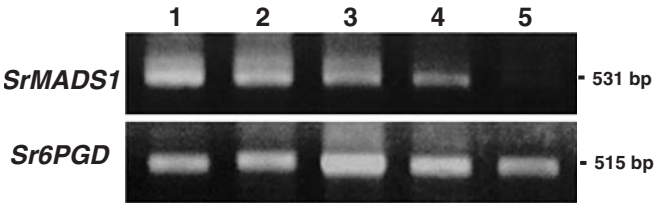


Fig. 3. RT-PCR analysis of *SrMADS1* gene expression in sporophytic tissues. Complementary DNAs were synthesized from RNAs extracted from apices in the vegetative stage (lane 1), strobili (lane 2), microphylls (lane 3), stems (lane 4) and a mixture of roots and rhizophores (lane 5). The *Sr6PGD* gene was used as a positive control

LAMB3 and *5* were excluded from our analysis because they lack the K-box (Svensson and Engström 2002). As shown by the ML tree, *SrMADS1* forms a clade with *LAMB2*, *4*, and *6* with a high bootstrap value, and *SrMADS1* is not closely related to gene(s) of the *LAMB2* group (Fig. 2). Moreover, this tree indicates that *SrMADS1* may be the sister of the *LAMB2* group; the ML tree supports this topology with a bootstrap value of 80%, but the NJ tree does not support this topology (Fig. 2). Unlike *SrMADS1*, genes of the *LAMB2* group do not possess additional amino acids in the N-terminal of the MADS domain. Some genes of the *AG* group and all genes of the *CRM6* group have similar additional amino acids in the N-terminal of the MADS domains, but *SrMADS1* did not cluster with *AG* and *CRM6* clades in the phylogenetic tree (Fig. 2). Thus *SrMADS1* appears to have obtained the N-terminal region originally in the lineage of *SrMADS1*. The expression pattern of *SrMADS1* was assessed by RT-PCR using total RNA extracted from vegetative shoot tips, strobili, microphylls, stems without leaves, and rhizophores with roots. No amplification was detected in the roots or rhizophores, though we were unable to separate these (Fig. 3). Accumulation of *SrMADS1* mRNA in gametophytic tissues is unknown, but *SrMADS1* transcripts preferentially accumulate in whole sporophytic tissues, except in roots and rhizophores. *LAMB2*, *4*, *5*, and *6* are also expressed in a broad range of sporophytic organs of *Lycopodium*, though the expression pattern of *LAMB3* has not been reported (Svensson and Engström 2002). The expression patterns of *SrMADS1* and *LAMB2*, *4*, *5*, and *6* are similar, except that *LAMB2*, *4*, *5*, and *6* are expressed in the roots but *SrMADS1* is not. The difference between these expression patterns and the results of our phylogenetic analysis suggest that the root of *Selaginella* was not derived from the same common ancestral organ as *Lycopodium*, although further analyses of root development processes in lycophytes are necessary to further test this hypothesis.

Previous authors have suggested that the generally expressed fern MIKC^c genes are more primitive than the MIKC^c genes specifically expressed in reproductive organs (Münster et al. 1997; Hasebe et al. 1998). Our results as well as those with the MIKC^c genes of *Lycopodium* (Svensson and Engström 2002) support this hypothesis. Lycophytes appeared during the early Devonian, approximately 400 million years ago, together with the extinct rhyniophytes and zosterophylls, which had dichotomously branch-

ing stems (Kenrick and Crane 1997) and form a sister to the clade including seed plants and ferns. Both *Selaginella* and *Lycopodium* are placed in the lycophytes, but these two genera are phylogenetically distant from each other (Kenrick and Crane 1997). Our results and those of Svensson and Engström (2002) suggest that the ancestral MIKC^c gene of the primitive dichotomous plants in the early Devonian was involved in the development of basic sporophytic tissues such as shoot, stem, and sporangium.

Acknowledgement We thank T. Asakawa, H. Ishikawa, and S. Kurita for critical comments and R. Sano for technical support.

References

Adachi J, Hasegawa M (1992–1996) ProtML 2.3b3 maximum likelihood inference of protein phylogeny. The Institute of Statistical Mathematics, Tokyo

Alvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, Ribas de Pouplana L, Martinez-Castilla L, Yanofsky MF (2000) An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. Proc Natl Acad Sci USA 97:5328–5333

Daff RJ, Nickrent DL (1999) Phylogenetic relationships of land plants using mitochondrial small-subunit rDNA sequences. Am J Bot 86:372–386

Gu Q, Ferrandiz C, Yanofsky MF, Martienssen R (1998) The *FRUITFULL* MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. Development 125:1509–1517

Hartmann U, Hohmann S, Nettesheim K, Wisman E, Saedler H, Huijser P (2000) Molecular cloning of *SVP*: a negative regulator of the floral transition in *Arabidopsis*. Plant J 21:351–360

Hasebe M (1999) Evolution of reproductive organs in land plants. J Plant Res 112:463–474

Hasebe M, Wen CK, Kato M, Banks JA (1998) Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. Proc Natl Acad Sci USA 95:6222–6227

Henschel K, Kofuji R, Hasebe M, Saedler H, Münster T, Theißen G (2002) Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. Mol Biol Evol 19:801–814

Hiesel R, von Haeseler A, Brennicke A (1994) Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. Proc Natl Acad Sci USA 91:634–638

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature 389:33–39

Kranz HD, Miks D, Siegler ML, Capesius I, Sensen CW, Huss VA (1995) The origin of land plants: phylogenetic relationships among charophytes, bryophytes, and vascular plants inferred from complete small-subunit ribosomal RNA gene sequences. J Mol Evol 41:74–84

Ma H, Yanofsky MF, Meyerowitz EM (1991) *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. Genes Dev 5:484–495

Münster T, Pahnke J, Di Rosa A, Kim JT, Martin W, Saedler H, 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

Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409:618–622

Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science 255:1697–1699

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. Science 288:1613–1616

Shindo S, Ito M, Ueda K, Kato M, Hasebe M (1999) Characterization of MADS genes in the gymnosperm *Gnetum parvifolium* and its implication for the evolution of reproductive organs in seed plants. Evol Dev 3:180–190

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

Svensson ME, Engström P (2002) Closely related MADS-box genes in club moss (*Lycopodium*) show broad expression patterns and are structurally similar to, but phylogenetically distinct from, typical seed plant MADS-box genes. New Phytol 154:439–450

Svensson ME, Johannesson H, Engström P (2000) The *LAMB1* gene from the clubmoss, *Lycopodium annotinum*, is a divergent MADS-box gene, expressed specifically in sporogenic structures. Gene 253:31–43

Theißen G, Becker A, Di Rosa A, Kanno A, Kim JT, Münster T, Winter KU, Saedler H (2000) A short history of MADS-box genes in plants. Plant Mol Biol 42:115–149

Thompson J, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Winter KU, Becker A, Münster T, Kim JT, Saedler H, Theißen G (1999) MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. Proc Natl Acad Sci USA 96:7342–7347

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