

The MADS-domain protein PPM2 preferentially occurs in gametangia and sporophytes of the moss *Physcomitrella patens*

Vanessa Quodt, Wolfram Faigl, Heinz Saedler, Thomas Münster*

Max Planck Institute for Plant Breeding Research, Department of Molecular Plant Genetics, Cologne, Germany

Received 4 April 2007; received in revised form 23 May 2007; accepted 24 May 2007

Available online 5 June 2007

Received by W. Martin

Abstract

To date, the function of MADS-domain transcription factors in non-seed plants remains largely elusive, although a number of genes have been isolated and characterized from a variety of species. In our study we analyzed *PPM2*, a classical MIKC-type MADS-box gene from the moss *Physcomitrella patens*, taking advantage of the unique technical properties *Physcomitrella* offers in terms of efficient homologous recombination. We determined mRNA and protein distribution and performed targeted disruption of the genomic locus for functional analysis of *PPM2*. Despite weak ubiquitous expression, PPM2 protein is mostly found in male and female gametangia and basal parts of developing sporophytes. Therefore, *PPM2* seems to function in both the haploid and the diploid phase of the moss life cycle. This situation reflects an evolutionary transition state of gene recruitment from an ancestral gametophytic generation into a derived sporophytic generation which became dominating in tracheophytes. However, a knock-out of the *PPM2* gene did not cause visible phenotypical changes in the respective structures. The implications of our findings for the understanding of the evolutionary history of MADS-box transcription factors in plants are discussed.

© 2007 Elsevier B.V. All rights reserved.

Keywords: MADS-box gene; Transcription factor; Protein expression; Gene knock-out; Gene recruitment; Land plant evolution

1. Introduction

Members of the MADS-box gene family encode developmentally important transcriptional regulator proteins in animals, plants and fungi. In seed plants MADS-box genes are involved in many different aspects of plant life (for review: Theißen et al., 2000;

Jack, 2001). Well known are MADS-box genes responsible for the definition of organ identity in angiosperm flowers; their functional interplay is summarized in combinatorial models based on the “ABC-model” of floral development (Coen & Meyerowitz, 1991). However, in a typical seed plant like *Arabidopsis thaliana* the so-called MIKC^c MADS-box genes, of which the ABC function genes are typical representatives, constitute less than half of the more than 100 MADS-box genes which have been identified in genome-wide surveys (Kofuji et al., 2003; Pařenicová et al., 2003). To date, the MIKC^c genes are by far the best analyzed MADS-box genes in plants, and a number of mutant phenotypes have been described so far.

MIKC^c MADS-box genes and the transcription factor proteins they encode show a modular organization comprising a MADS-(M-), an intervening (I-), a keratin-like (K-) and a C-terminal (C-) domain (overview in: Kaufmann et al., 2005). The highly conserved MADS-domain is responsible for DNA-binding and dimerization with other MADS-domain proteins. Other determinants of MADS-domain protein-protein interaction capability are the intervening domain, which mainly defines dimerization

Abbreviations: bp, base pair(s); CA_nG, C A/T-rich G motif; cDNA, DNA complementary to RNA; ^c, classic-type; DAPI, 4',6-diamidino-2-phenylindole; DEF, DEFICIENS; DsRED, red fluorescent protein of *Discosoma spec.*; EST, expressed sequence tag(s); GUS, β-Glucuronidase; kb, kilobase(s) or 1000 bp; MADS-box, an acronym designating a conserved DNA sequence encoding a DNA-binding protein domain, with MADS being derived from genes *MCM1*, *AGAMOUS*, *DEFICIENS* and *SRF*; MIKC, MADS-, intervening, keratin-like and C-terminal domains; MYA, million years ago; ORF, open reading frame; PCR, polymerase chain reaction(s); *PpMADS1*, *Physcomitrella patens* MADS-box gene No. 1; *PPMn*, *Physcomitrella patens* MADS-box gene No. *n*; RACE, rapid amplification of cDNA ends; rpm, rounds per minute; RT, reverse transcription; UTR, untranslated region(s); wt, wild type; ::, novel junction (fusion or insertion).

* Corresponding author. Carl-von-Linné-Weg 10, 50829 Cologne, Germany. Tel.: +49 221 5062 122; fax: +49 221 5062 113.

E-mail address: muenster@mpiz-koeln.mpg.de (T. Münster).

specificity, and probably the keratin-like domain. The C-terminal domain is the major structural determinant for the formation of higher order protein complexes and sometimes possesses transcriptional activation domains. To fulfil their regulatory functions, MIKC^c MADS-domain proteins bind as obligate protein-dimers to so-called CARG-box recognition sites (consensus: CC(A/T)₆GG) in the genome (Schwarz-Sommer et al., 1992).

MIKC^c MADS-box gene function and evolution have been intensively studied in a large number of seed plant species. One important finding for the reconstruction of the evolutionary history of the MADS-box gene family was the identification of angiosperm gene orthologs in gymnosperms showing a high degree of functional conservation (Sundström & Engström, 2002). So far MIKC^c MADS-box genes have been characterized from different representatives of lycophytes (Svensson & Engström, 2002; Tanabe et al., 2003), pteridophytes (Münster et al., 1997; Hasebe et al., 1998; Münster et al., 2002), bryophytes (Krogan and Ashton, 2000; Henschel et al., 2002; Hohe et al., 2002) and green algae (Tanabe et al., 2005). However, no evidence for the existence of seed plant gene orthologs could be obtained from non-seed plants. Almost all available data focus on structural and phylogenetic characterization of the genes as well as on their transcript patterns. Information on MADS-domain protein distribution or putative gene functions is completely lacking. An obvious evolutionary trend of MIKC^c MADS-box genes in plants is the increase in gene family size, ranging from just one gene in Charophycean green algae to about 40 in flowering plants (Kofuji et al., 2003; Tanabe et al., 2005). Because no MIKC^c MADS-box genes could be identified in the genomes of the chlorophycean green alga *Chlamydomonas reinhardtii* (Tanabe et al., 2005) and the red alga *Cyanidioschyzon merolae* (Matsuzaki et al., 2004), MIKC^c MADS-box genes very likely originated in a common ancestor of charophytes and land plants at least 500 million years ago (MYA).

Since mutants for MIKC^c MADS-box genes are not available in non-seed plants, a direct functional characterization of these genes in the respective species background has not been possible so far. Moreover, the predominantly broad expression domains of the MIKC^c MADS-box genes in ferns and lycophytes do not provide indirect clues regarding a putative functional context of the genes.

To take advantage of unique technical approaches for detailed protein expression analysis, as well as for direct functional studies, the characterization of MADS-box genes has been initiated in the moss *Physcomitrella patens* (Krogan and Ashton, 2000; Henschel et al., 2002; Hohe et al., 2002). *Physcomitrella* provides the potential to manipulate gene loci directly due to the high frequency of efficient homologous recombination in its genome (overview in: Schaefer, 2002). Based on cDNA cloning and EST sequencing, four MIKC^c MADS-box genes from *Physcomitrella* have been previously characterized (Krogan and Ashton, 2000; Henschel et al., 2002; Hohe et al., 2002). After sequencing of the entire *Physcomitrella* genome, a genome-wide survey of the raw sequence data led to the identification of a total of six different MIKC^c MADS-box gene loci (V.Q. and T.M., data not shown). Analysis of the gene phylogeny

reveals that the *Physcomitrella* MIKC^c MADS-box genes belong to two clearly distinct subclades, both containing three members (data not shown). However, no information is thus far available on MIKC^c MADS-box gene activity and function in *Physcomitrella*.

The following study focusses on the analysis of the *Physcomitrella* MIKC^c MADS-box gene *PPM2*. We report the elucidation of *PPM2* expression both on mRNA and protein level, identifying gametangia and the basal part of the moss sporophyte as major expression domains of *PPM2* protein. We demonstrate intensive splicing in the 5' UTR region of the *PPM2* mRNA and present the analysis of a putative function of the 5' UTR in the control of *PPM2* translation. Additionally, we describe disruption of the genomic locus of *PPM2* and subsequent screening of the resulting transgenic moss lines for phenotypical changes. Finally, we propose a function for *PPM2* and discuss the impact of the presented findings on the understanding of MIKC^c MADS-box gene evolution.

2. Materials and methods

2.1. Transcript expression analysis

Total RNA was isolated from different stages of the moss life cycle according to Logemann et al. (1987). Single stranded cDNA was subsequently produced with Superscript II reverse transcriptase (Invitrogen). RT-PCR was performed for all developmental stages using cDNA derived from about 100 ng of total RNA as template. Fragments of the cDNA of *PPM1*, *PPM2* and *PpMADS1* were amplified using gene specific primers complementary to the C-terminal region of the respective gene. *Actin* cDNAs were amplified accordingly. The bands shown in Fig. 1 represent samples taken after 32 PCR cycles (for *PPM1*, *PPM2* and *PpMADS1*) and 35 PCR cycles (for *actin*). Primer sequences are as follows: *PPM1*-f 5'-

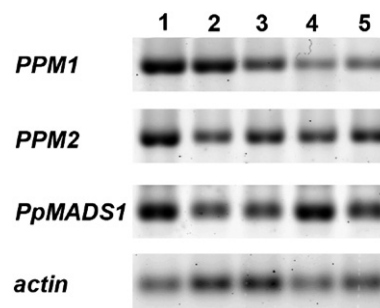


Fig. 1. RT-PCR analysis of transcript patterns of *Physcomitrella patens* MIKC^c MADS-box genes. Fragments of *PPM1*, *PPM2* or *PpMADS1* transcripts were amplified with gene specific primers from cDNAs representing different stages of the moss life cycle (1–5, indicated above the lanes): (1) protonema, 11 days after propagation; (2) protonema with young gametophores, 4 weeks after propagation; (3) mature gametophores, 10 days after induction of the reproductive phase, 8.5 weeks after propagation; (4) gametangia-bearing gametophores 11 weeks after propagation, 4 weeks after induction; (5) gametophores with mature sporophytes, 9 months after propagation. *Actin* was used as control for equal amplification rates of the reactions. Expected sizes of PCR products amplified from *PPM1*, *PPM2* and *PpMADS1* are 516 bp, 559 bp and 489 bp, respectively.

GATTTCCTGGGTCGCGAGGTCGTG-3', PPM1_r 5'-ACCTGCA-CAGCCTTCGGAGGCAGA-3', PPM2_f 5'-GAAGGGCTCGGATGGAATG-CAAAA-3', PPM2_r 5'-ACCTGCACAACCTTCTGGCCCAAC-3', PpMADS1_f 5'-GATTACTTGGTCGGGAGGTGATT-3', PpMADS1_r 5'-ACCTGCACAACCTTCTCCGTGCT-3'.

For the analysis of the *PPM2* 5' UTR, specific primers were designed and used in PCR reactions with 35 cycles (Fig. 4B). The identity of the PCR products was confirmed by sequencing. Primer sequences are as follows: UTR_f 5'-GGCCTCCCC-AGGGTCCTGATC-3', UTR_r 5'-AACCCGACCCTTGGACCGGACTC-3'.

2.2. Moss transformation for gene disruption and gene-reporter fusion

In the DNA construct *PPM2*_{exon7}:*GUS*_{nptII}, the β -Glucuronidase (*GUS*) open reading frame was fused in frame to exon 7 of *PPM2* via *AatII*, followed by a *35S::nptII:ter* selection cassette for resistance against Paromomycin. The core construct is flanked by stretches of DNA homologous to the wild type *PPM2* locus in order to ensure homologous recombination upon transformation. The 5' and 3' flanking regions are 1.7 kb and 0.9 kb, respectively. Fusion of *PPM2* and *DsRED* was achieved by ligating the *DsRED* ORF to exon 7 of *PPM2*. The *aph4* gene, conveying resistance to Hygromycin under control of the *nos* promoter and terminator, was connected downstream. The construct was flanked by 1.5 kb (5') and 0.7 kb (3') of homologous *PPM2* genomic sequence.

For the replacement of the *PPM2* coding region by the *GUS* ORF, a 1.7 kb fragment of the *PPM2* promoter was cloned upstream of the *GUS* ORF, followed by a *35S::nptII:ter* selection cassette and a 1.3 kb 3' flanking region, consisting of the *PPM2* coding region up to intron 4. *PPM2::DEF*_{5'UTR}:*GUS* was synthesized from the resulting construct by replacing the putative 5' UTR region of *PPM2* (-730 to -1) by a PCR product representing the *DEFICIENS* 5' UTR. This led to a shorter 5' flanking homologous region of about 1 kb.

2.3. Transformation procedure and selection

The transformation procedure was performed according to Schaefer (2001). 12 μ g of purified construct DNA was used for the transformation of 400,000 protoplasts via polyethylene-glycol-mediated transfer. After recovery, the protoplasts were cultivated on plant antibiotic-containing medium (Paromomycin 40 μ g ml⁻¹; Hygromycin 15 μ g ml⁻¹) for two rounds of selection.

2.4. Molecular characterization of transgenic moss lines

Putative positive transformants were identified by DNA isolation and PCR as described in the online "PHYSCOmanual 1.2" by Hiwatashi and colleagues (<http://www.nibb.ac.jp/~evodevo/PHYSCOmanual/00Eindex.htm>). For this purpose, young tissue from the edge of each clone was transferred to 30 μ l of PCR buffer, frozen twice in liquid nitrogen, incubated at 68 °C for 10 min and centrifuged 5 min at 5000 rpm. 2 μ l of this DNA preparation was immediately used for PCR analysis.

Correct 5' integration of the constructs at the loci was tested by PCR amplification with one primer complementary to a sequence upstream of the integration site and another primer binding to DNA within the construct. Sequencing of these PCR products confirmed correct integration and, thus, full retention of the respective promoter region (data not shown). Accordingly, 3' integration was additionally demonstrated in some lines (data not shown). The fusion cDNAs of *PPM2*_{exon7}:*GUS* and *PPM2*_{exon7}:*DsRED* were partially amplified by RT-PCR and verified by sequencing. The presence of a *GUS* transcript as well as the absence of a *PPM2* transcript in *PPM2::GUS* and *PPM2::DEF*_{5'UTR}:*GUS* lines was also determined by RT-PCR.

The existence of at least two independent mutant lines in each case, sharing identical reporter gene signals, strongly supports the authenticity of the observed patterns. Additional illegitimate insertions in the genome cannot be completely excluded, but they are known to be rare.

2.5. Semi-thin sections of moss tissue and microscopic documentation

GUS-stained gametophores of the *PPM2*_{exon7}:*GUS* line were embedded in Araldite according to a modified version of the protocol described in Sorensen et al. (2002). After vacuum infiltration and overnight incubation in a 5% glutaraldehyde fixative, the samples were carefully dehydrated through an ethanol series (10%, 30%, 50%, 70%, 90%, 2 \times 100%) on ice. Eosin yellow was added to the last steps (90% and 100%) of the ethanol series at a final concentration of 0.1% to stain all moss tissues lacking a *GUS* signal. This step was crucial to retain visibility of the minuscule gametophores in the synthetic resin. After a 12 h incubation at 4 °C, the samples were transferred to a 1:1 ethanol:propylene oxide mix and infiltrated at 4 °C for 5 h. Subsequently, the mixture was replaced by absolute propylene oxide and kept at 4 °C over night. The gametophores were then exposed to a 1:1 solution of propylene oxide and Araldite mix (prepared from the Epoxy Resin Kit from Agar Scientific) for 2 h, followed by vacuum infiltration and overnight incubation in pure Araldite mix. The next day, the infiltrated gametophores were positioned upside down in conical polyethylene capsules, submerged in freshly prepared Araldite mix and vacuum infiltrated. After polymerization at 60 °C for 48 h, semi-thin sections of typically 800 nm were cut using glass knives. The sections were embedded in fresh Araldite on microscopic slides and examined using dark-field microscopy.

2.6. Histochemical detection of *GUS* activity

The histochemical *GUS* activity was assayed as described by Nishiyama et al. (2000).

2.7. Maintenance of moss cultures

P. patens (Hedw.) B.S.G. was cultivated as described by Schaefer (www2.unil.ch/lpc/docs/pdf/PPprotocols2001.pdf).

3. Results

3.1. *PPM2* mRNA is ubiquitously expressed

In order to analyze the transcript expression pattern of *PPM2*, total RNA was isolated from different stages of the moss life cycle for an RT-PCR approach. For practical reasons, the different moss tissues were harvested together at the different stages of the moss life cycle. As a consequence, the samples represent mixtures of different tissues and organs (with the exception of young protonema) and, therefore, a temporal pattern of MADS-box gene activity was obtained by this method rather than a tissue- or organ-specific pattern. Besides *PPM2*, the expression of two additional *Physcomitrella* MIKC^c genes, *PPM1* and *PpMADS1*, was examined using sets of gene specific primers.

All three MIKC^c genes exhibited a broad expression pattern including protonema, gametophore or sporophyte tissues (Fig. 1).

While *PPM2* and *PpMADS1* transcripts could be detected in similar amounts in all stages, *PPM1* expression was strongest in early protonema and young gametophores and decreased in later stages of the moss life cycle.

3.2. *PPM2* protein is strongly expressed in archegonia, spermatozoids and sporophyte feet

While the temporal expression pattern of *PPM2* transcript was demonstrated by RT-PCR, tissue or organ specific localization of *PPM2* protein was analyzed by *in planta* expression of reporter gene fusions. Since alternative splicing events have been reported in the exons encoding the C-terminal domain of *PPM2* (Krogan and Ashton, 2000; Henschel et al., 2002), we chose to fuse both reporter genes to exon 7 of *PPM2* in order to abolish these events. As a consequence, both fusion proteins should represent all *PPM2* expression domains. Connection of the β -Glucuronidase open reading frame to

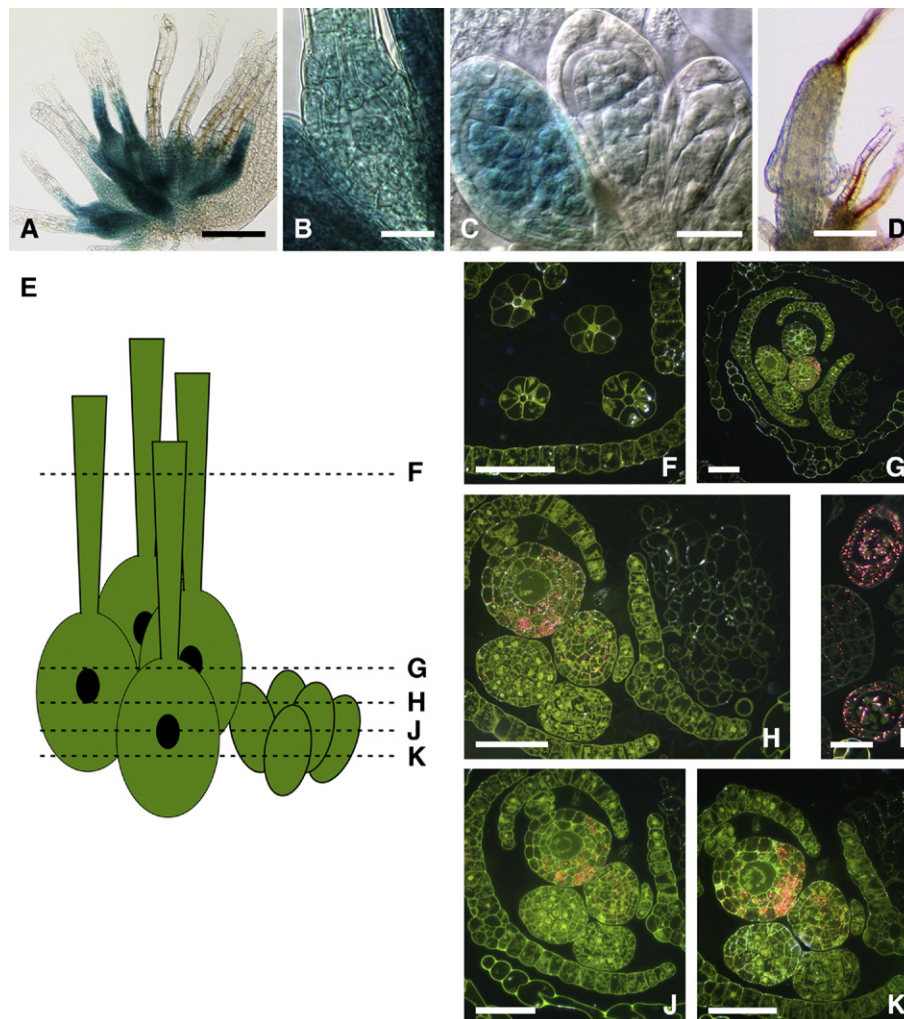


Fig. 2. *In planta* reporter gene studies demonstrating the cellular localization of a *PPM2*_{exon7}:GUS fusion protein in different moss tissues. (A–D) Microscopic images of gametangia and a sporophyte: expression in a bundle of archegonia (A), the ventral part of a single archegonium with egg cell (B), a group of antheridia of different developmental stages (C) and in a young sporophyte (D). (E–K) Semi-thin sections showing the localization of *PPM2*_{exon7}:GUS fusion protein in the apical region of a *Physcomitrella* gametophore bearing both types of gametangia. The images were recorded during dark-field microscopy and display GUS staining in pink. Section through archegonia necks (F), ventral parts of archegonia (G, H, J, K), and empty antheridia (G, H, K) as depicted in the schematic drawing of the organs (E). A section of antheridia containing spermatozoids is shown in (I). The scale bars indicate 200 μ m (A, D), 50 μ m (F–H, J, K) and 20 μ m (B, C, I), respectively.

exon 7 of *PPM2* leads to intense GUS activity in female reproductive organs, spermatozooids and sporophyte feet (Fig. 2A–D) in two independent knock-in lines. Expression in archegonia is mostly restricted to the ventral area, but cell-specific distribution, especially in the egg cell or zygote, cannot be resolved studying entire organs. Spermatozooids show increasing intensity of GUS activity that can be correlated with maturation (Fig. 2C). However, expression of *PPM2*_{exon7}:*GUS* exhibits a high degree of variability according to the level of intensity and distribution among gametophore apices. While in some bundles of gametangia every single organ produces a signal upon staining, in other cases only few or none of the reproductive organs do so (data not shown). This is also true for sporophyte feet where signals vary strikingly in appearance and strength of expression. Semi-thin sections of stained gametophores carrying gametangia reveal that the fusion protein

*PPM2*_{exon7}:*GUS* is found in all of the cell layers composing the ventral part of an archegonium (Fig. 2G, H, J, K). The unequal distribution of GUS product among cells of one section is very likely the result of methodical limitations and obviously depends on the section plane. However, the sections do not contribute to elucidating protein localization in egg cells or zygotes. The delicate cells were likely disrupted during the experimental procedure. GUS activity can also be associated with antheridia carrying mature spermatozooids (Fig. 2I). In contrast, male sexual organs that have already released their sperm cells no longer show any GUS activity (Fig. 2H).

To observe *PPM2* localization *in vivo* with emphasis on the temporal aspects of development on a single gametophore, knock-in moss lines expressing a fusion of *PPM2* with the fluorescent protein DsRED were produced. Analysis of two independent lines did not only confirm the results of the

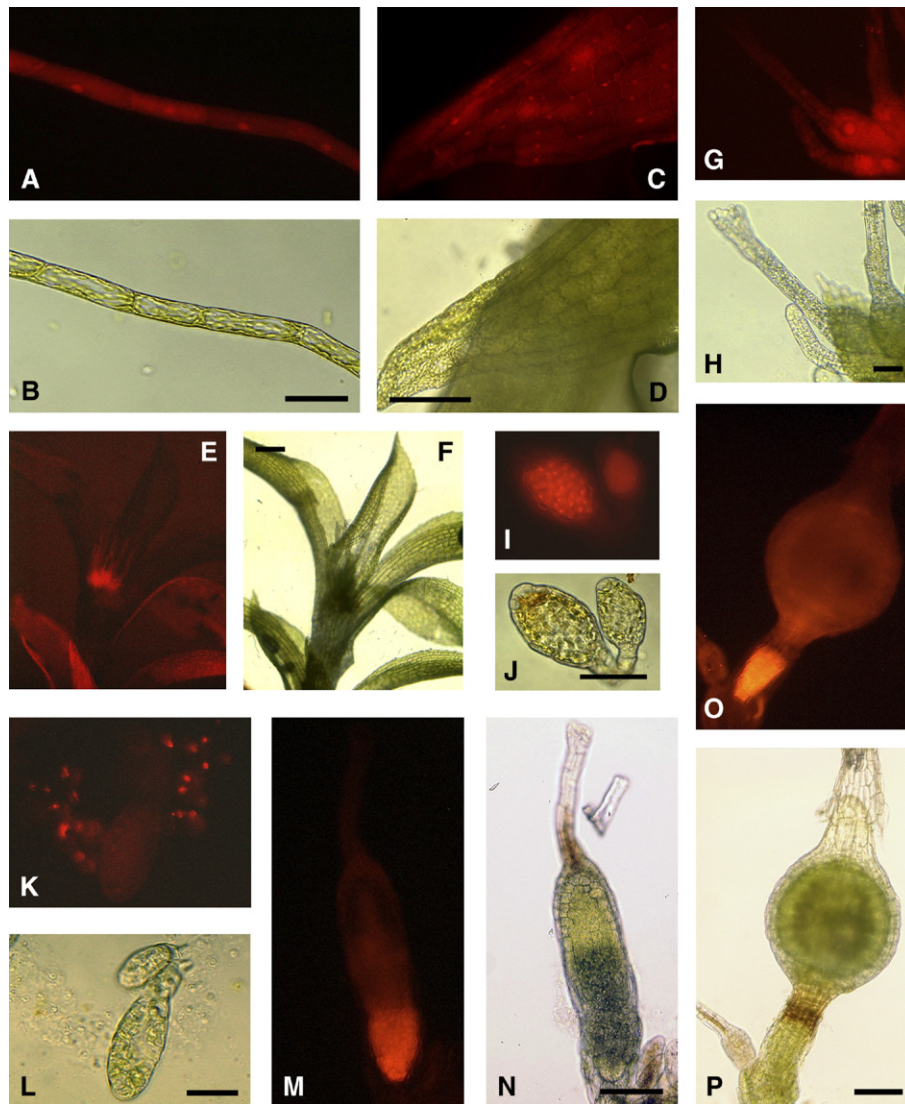


Fig. 3. *In planta* reporter gene studies depicting expression of a *PPM2*_{exon7}:*DsRED* fusion in *Physcomitrella*. (A–P) Fluorescent signal of a *PPM2*_{exon7}:*DsRED* fusion protein in the nucleus and cytoplasm, and the corresponding light microscopic image in protonema (A, B), a leaflet (C, D), the apical region of a gametangia-bearing gametophore (E, F), a bundle of archegonia of different developmental stages (with zygote) (G, H), a young and a fully developed antheridium with spermatozooids (I, J), spermatozooids released from a mature antheridium (K, L), a young sporophyte (M, N) and a fully developed sporophyte (O, P). Scale bars correspond to 200 μ m (F), 100 μ m (D, N, P) and 50 μ m (B, H, J, L), respectively.

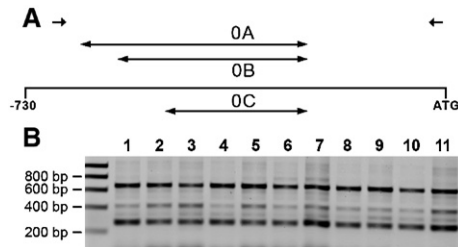


Fig. 4. Structural analysis of the *PPM2* 5' UTR. (A) Schematic representation of the 5' UTR of *PPM2* with the putative transcription start site at -730. Introns 0A, 0B and 0C, which share a 3' splice site at -239 but have different 5' splice sites at -633, -569 and -486, respectively, are shown. The primers used for RT-PCR are indicated as arrows. (B) RT-PCR analysis of the *PPM2* 5' leader from cDNA pools representing different stages of the moss life cycle (1–11). (1) 1 week old protonema, (2) 6 week old protonema, (3) protonema with gametophores, (4) young gametophores, (5) mature gametophores, (6) gametophores 24 h after induction at 17 °C, (7) gametophores 3 weeks after induction at 17 °C, (8) gametophores with gametangia, (9) apex enriched fraction of gametangia-bearing gametophores, (10) young sporophytes and (11) sporophytes. Apparent sizes of the bands are indicated on the left.

PPM2_{exon7}:GUS fusion, but revealed additional aspects of *PPM2* expression (Fig. 3A–P). While strong fluorescent signals in gametangia and sporophyte feet support that *PPM2* is localized in these organs, a weak ubiquitous signal in all the remaining tissues during the complete life cycle was also observed. Apparently, *PPM2* is not only expressed during all stages of the moss life cycle as shown in the RT-PCR (Fig. 1), but also ubiquitously. However, the variability of *PPM2_{exon7}:DsRED* expression in archegonia, spermatozooids and sporophyte feet matches that of *PPM2_{exon7}:GUS* in terms of intensity and distribution. Furthermore, observation of several individual gametophore apices *in vivo* reveals that single reproductive organs or entire bundles regularly lack a fluorescent signal (data not shown). In spite of this fact, the respective apices develop normal sporophytes (without the intense *DsRED* signal in the foot) with viable spores.

While the analysis of lines expressing *PPM2_{exon7}:GUS* does not clarify the localization of *PPM2* protein in egg cells and zygotes, the possibility to observe the *DsRED* reporter gene product *in vivo* allows a thorough examination without the risk of destroying the easily damageable cells. Both the haploid egg cell (not shown) and the diploid zygote (Fig. 3G, H) show a strong fluorescent signal.

Furthermore, the application of protein fusion with *DsRED* allows detection of the subcellular distribution of *PPM2*. The transcription factor clearly allocates within the nucleus (Fig. 3A, C), which was verified by DAPI staining (data not shown), but is also present in the cytosol.

3.3. Alternative splicing of the *PPM2* 5' UTR

The discrepancy between the ubiquitous occurrence of *PPM2* mRNA on the one hand and the strictly tissue-specific protein expression pattern on the other hand suggests that *PPM2* might be under translational control. Since the analyses *via in situ* hybridization and single-cell or tissue-specific RT-

PCR to reveal distinct transcript patterns are not established in *Physcomitrella* yet, it is not possible to clarify directly if *PPM2* regulation is executed on the transcript level. However, it has been reported that many transcription factors are subject to translational control rather than transcriptional control *via* their 5' UTR, where different mechanisms have been described that lead to stalling or dissociation of scanning ribosomes (Hughes, 2006). Therefore, a structural analysis of the *PPM2* 5' UTR was performed to investigate its potential role in translational regulation.

5' RACE experiments revealed a putative transcription initiation site at -730 (in relation to ATG) for *PPM2* mRNA (data not shown). This unusually long 5' UTR exhibits an additional alternatively spliced upstream intron ("intron 0"), a feature that has been associated with translational control of gene expression (Weise et al., 2005). The intron 0 comprises three splice variants A, B and C (Fig. 4A) with a common 3' splice site. Their 5' splice sites differ and thus generate 5' UTR versions of unequal lengths. RT-PCR with cDNAs representing different stages of the moss life cycle demonstrated that all splice variants, including the non-spliced version, exist during all developmental stages in a comparable ratio (Fig. 4B). In order to investigate the composition of 5' UTR splice variants in those tissues with strong *PPM2* expression, the apical parts of gametophores with reproductive organs were harvested to produce an apex-enriched fraction since collecting gametangia alone was technically difficult. Fig. 4B shows that the distribution of 5' UTR versions in the apex-enriched fraction is comparable to all other samples, including gametophores of the same developmental stage that were not enriched. Apparently, there is no major difference between the composition and weighting of splice variants throughout the moss life cycle. Translational control by splicing of the 5' UTR is thus unlikely to be directly responsible for the differential expression pattern of *PPM2*. However, involvement of the *PPM2* 5' UTR in translational regulation by a different mechanism cannot be excluded and was subject to further investigation.

3.4. Replacement of the *PPM2* 5' UTR region has no effect on the *PPM2* protein expression

To elucidate putative regulatory functions in protein synthesis, the 5' UTR region of the *PPM2* genomic locus was replaced by the respective 5' UTR of the homologous *MADS-*

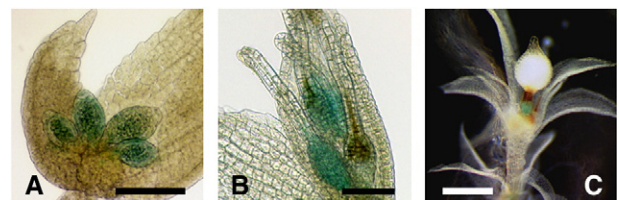


Fig. 5. Functional analysis of the *PPM2* 5' UTR. (A–C) GUS expression pattern resulting from a *DEF_{5'UTR}:GUS* transcript under control of the *PPM2* promoter. The transgenic moss line exhibits GUS signals in antheridia (A), archegonia (B) and sporophytes (C). Scale bars represent 500 μm (C) and 100 μm (A, B), respectively.

box gene *DEFICIENS* of *Antirrhinum majus*. The *DEFICIENS* 5' UTR is known to have no effect on translational control (Zachgo et al., 1995). Therefore, its 5' UTR was selected for a fusion with the coding region of the *GUS* reporter gene. Via homologous recombination, the *DEFICIENS*_{5'UTR}:*GUS* fusion was placed at the *PPM2* locus in the *Physcomitrella* genome directly downstream of the native *PPM2* promoter region (for details see Section 2.2). Six independent moss lines with a correct integration of the *DEFICIENS*_{5'UTR}:*GUS* fusion were identified. As control for the *GUS* expression pattern obtained with the *DEFICIENS* 5' UTR, two independent transgenic moss lines carrying a *PPM2*::*GUS* reporter fusion at the native gene locus were used (see below).

A thorough examination of the transgenic lines revealed that no changes of the *GUS* protein pattern occur in the lines expressing a *DEF*_{5'UTR}:*GUS* fusion mRNA at the *PPM2* locus (Fig. 5). The results clearly demonstrate that *PPM2* is not translationally regulated by its 5' UTR.

3.5. Disruption of the *PPM2* genomic locus caused no obvious phenotypical changes

Targeted integration of a knock-out construct at the genomic locus was used to analyze the function of *PPM2* in moss. Both

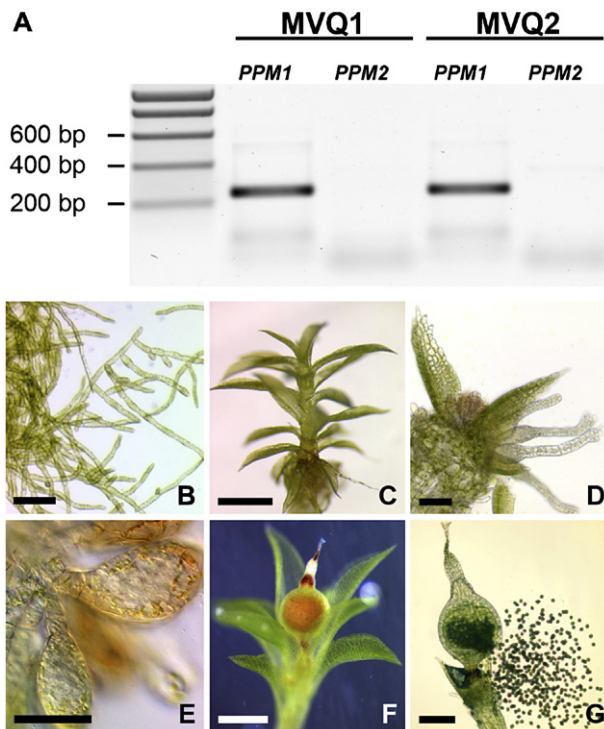


Fig. 6. Analysis of transgenic knock-out moss lines with a disrupted *PPM2* locus. (A) Molecular verification of the targeted *PPM2* gene disruption in the lines MVQ1 and MVQ2. A 33 cycle RT-PCR with cDNA pools representing young protonema and gene specific primers for *PPM1* and *PPM2*. (B–G) Phenotypical characterization of the *PPM2* knock-out lines. The images depict protonema (B), a young gametophore (C), a gametophore apex with male and female gametangia (D), antheridia (E), a sporophyte (F) and a mature sporophyte releasing spores (G). Scale bars correspond to 1000 μ m (C), 500 μ m (F), 200 μ m (B, D, G) and 50 μ m (E), respectively.

genomic PCR and RT-PCR were performed to verify the disruption of the *PPM2* wild type locus in two transgenic lines (Fig. 6A and data not shown). These lines were closely observed for developmental or morphological changes. Based on the *PPM2* expression pattern, the analysis was particularly focused on the (1) development of gametangia and gametes, (2) efficiency of fertilization and (3) early phases of embryo and sporophyte formation.

All moss lines examined did not show any obvious deviations from the wild type morphology (Fig. 6B–G); timing and course of the life cycle did not change. All egg cells and spermatozooids observed looked normal; the mobility of the spermatozooids was not reduced. Also, the number and morphology of sporophytes and developing spores in the sporangia were in the normal range of *Physcomitrella* wild type moss (data not shown).

4. Discussion

4.1. *PPM2* belongs to a small family of MIKC^c MADS-box genes in *Physcomitrella*

Based on the full genome sequence, only six MIKC^c MADS-box genes are present in the moss genome. Phylogeny reconstructions revealed that these genes belong to two subclades, each comprising three members (data not shown). Together with *PPM1* and *PpMADS1*, the gene *PPM2* constitutes the highly supported *PPM2*-like subfamily. The encoded *PPM2*-like proteins are very similar (between 69% and 77% percent sequence identity); moreover, the MADS-domains of *PPM1* and *PPM2* are completely identical.

Compared to MIKC^c gene families in pteridophytes and seed plants, the moss gene family is small. In ferns like *Ceratopteris richardii* more than 30 genes seem to be present in the genome (W. F. and T. M., unpublished data), at least a similar number of genes exist in the gymnosperm *Gnetum gnemon* (whereas this is likely an underestimation due to limited samplings) and about 40 genes have been identified in the eudicotyledonous model plant *A. thaliana* (Kofuji et al., 2003; Pařenicova et al., 2003). On the other hand, only one MIKC^c gene was shown to be present in the genomes of Charophycean green algae (Tanabe et al., 2005). These data suggest a common trend of increasing MIKC^c gene numbers in the evolutionary history of green plants. Furthermore, this implies a correlation between the size of a gene family involved in developmental control (at least in seed plants) and the complexity of the plant organism, which was hypothesized also for other gene families involved in transcriptional regulation, e.g. Class III homeodomain leucine zipper genes (Floyd & Bowman, 2007).

4.2. High levels of *PPM2* protein are restricted to gametangia and basal parts of sporophytes

The virtually ubiquitous *PPM2* transcript pattern obtained by RT-PCR is a typical hallmark of non-seed plant MIKC^c MADS-box genes as shown for pteridophytes and lycophytes (Hasebe et al., 1998; Munster et al., 2002; Svensson and Engstrom, 2002).

The other two *PPM2*-like genes, *PPM1* and *PpMADS1*, also share these broad expression domains. Ubiquitous expression has generally been interpreted as an ancient feature of non-seed plant MIKC^c MADS-box genes, reflecting putative “basal” gene functions (Theißen et al., 2000).

However, in contrast to the transcript pattern, the analysis of *PPM2*-reporter gene fusion lines revealed that the PPM2 protein is indeed present in all cells, but is strongly upregulated in the ventral area of archegonia including egg cell and zygote, in developing and mature spermatozooids and in sporophyte feet (Figs. 2 and 3). The differences between transcript and protein distribution might be caused by restrictions of the methodology. Ubiquitous expression of *PPM2* transcript is mainly temporally resolved, because most of the analyzed plant materials were mixtures of different tissues. Additionally, cells with high PPM2 protein abundance were strongly underrepresented in the harvested samples. Therefore, the RT-PCR might have detected only a bottom-line of minimal, ubiquitous expression, without resolving high transcript abundance in restricted tissues or organs.

All *PPM2*-reporter gene fusion lines show an apparent degree of variability of expression in individual gametophores. Regularly, high expression of the reporter is restricted to only some of the archegonia and antheridia on one apex; in other cases the apical part of a gametophore completely lacks a signal. Observing such gametophores of the *PPM2*_{exon7}:*DsRED* fusion lines, cultivated under the same conditions, revealed that their gametangia generate normal sporophytes producing viable spores. However, these sporophytes did not express the fusion protein in the foot. In many cases the fluorescent signal was completely absent, including the whole gametophore.

Very recently, Singer et al. (in press) published their results on expression studies and gene knock-outs of *PPM1*, *PPM2* and *PpMADS1*. Consistent with our data, their *PPM2* gene disruption lines also lack an obvious phenotype. However, they did not detect any GUS signal in their transgenic lines expressing *GUS* under control of the *PPM2* promoter. We assume that the so far unknown environmental or endogenous parameters determining *PPM2* expression must be different in the respective culture conditions. Therefore, it is plausible that the transgenic moss lines created by Singer and colleagues have been used for GUS staining under conditions that do not require *PPM2* expression.

4.3. *PPM2* protein expression domains are defined on the transcript level

The highly complex structural features of the *PPM2* 5' UTR provide a potential source of differential regulation of expression on the translational level. In this context, a previously unknown intron with different 5' splice sites and a common 3' splice site was identified in the 5' UTR of *PPM2* (Fig. 4).

In spite of this highly complex regulatory potential, replacing the putative 5' UTR completely had no effect on the protein pattern. Thus, the 5' UTR is obviously not responsible for the regulation of PPM2 translation. A potential role of its 3' UTR in

translational regulation can also be excluded: in the *PPM2*-reporter gene fusion moss lines the 3' UTR was derived from the respective fused reporter gene. Taking all experimental data into account, control of PPM2 appears to occur on the transcriptional, not on the translational level.

4.4. *PPM2* represents the transition state between MADS-box gene expression in green algae and tracheophytes

The life cycle of plants features an alternation of a haploid generation, called gametophyte, and a diploid generation, called sporophyte. The fact that sporophyte-dominant plants evolved from gametophyte-dominant ancestors suggests that the precursors of land plant MADS-box genes, like the precursors of many other genes, originated from genes which were active in the haploid generation and were recruited into a functional context in the diploid phase during the evolution of land plants (Nishiyama et al., 2003; Kofuji et al., 2003; Tanabe et al., 2005). To verify this hypothesis, information on a number of informative non-seed plant lineages, which have been analyzed for their MADS-box gene families and the respective gene expression patterns, was collected and scrutinized.

Charophycean green algae display what may be considered as the ancestral state of MADS-box gene expression (Tanabe et al., 2005). In both the unicellular charophycean green alga *Closterium peracerosum-strigosum-littorale* complex and the multicellular representative *Chara globularis* MADS-box gene expression is basically restricted to haploid cells. However, it is upregulated in gametangial cells, which are homologous to the moss gametes (Tanabe et al., 2005). By contrast, transcript abundance in the only diploid cell type, the zygote, seems to be very low. These reported algal MADS-box gene transcript expression patterns partly resemble the *Physcomitrella* PPM2 protein pattern, because both patterns show high expression levels in gametangia.

In lycophytes, a more derived lineage of non-seed plants, MIKC^c MADS-box genes are expressed in both vegetative and reproductive tissues of the sporophytic generation, while data on isolated gametophytes are still lacking (Svensson & Engström, 2002; Tanabe et al., 2003). In pteridophytes, which are the closest sister group of seed plants, MADS-box genes are predominantly active in both vegetative and generative structures of sporophytic tissues; additional gametophytic expression is more exceptional (Münster et al., 1997; Hasebe et al., 1998; Münster et al., 2002). MIKC^c MADS-box genes in seed plants are nearly absent from gametophytic tissues. Out of 38 genes in *A. thaliana* only one, *AGL18*, is detectable in pollen (Kofuji et al., 2003). To date it is unclear whether it has been retained in the haploid generation or was re-recruited from the sporophyte.

Comparing the patterns of MIKC^c MADS-box gene expression in the different non-seed plant lineages and *PPM2* expression in *Physcomitrella*, it seems likely that *PPM2* expression represents the hypothesized transition state of MADS-box gene recruitment from a functional context in the gametophytic phase into a new functional context in the sporophytic phase of the plant life cycle. While *PPM2* expression in the moss gametophores still resembles the

patterns of charophycean MADS-box genes, *PPM2* additionally shows prominent sporophytic expression, the latter being typical of the majority of the homologous genes in lycophytes and pteridophytes. However, the question whether MADS-box gene recruitment had direct impact on the evolution of both the plant life cycle, comprising the two different generations, and the development of complex sporophytes of land plants remains to be answered yet and needs further investigation.

4.5. *PPM2* potentially functions in the definition of metabolic sink tissues in moss gametophyte and sporophyte

Unfortunately, no obvious mutant phenotype has been detected in *PPM2* knock-out lines yet. One possible explanation is a subtle impact of *PPM2* that is not essential for plant development or homeostasis under optimal growth conditions. This scenario is supported by the observed variability or complete lack of protein expression in independent moss samples. On the other hand, the high amino acid sequence similarities among the *PPM2*-like proteins suggest redundant functionality among these proteins, providing a plausible explanation for the lack of a phenotype for the *PPM2* disruption moss lines. All three corresponding genes within the *PPM2*-like subclade are apparently highly comparable in several aspects. As shown for *PPM2*, transcripts of *PPM1* and *PpMADS1* were also found during all developmental stages of the moss life cycle; however, the translational reporter gene fusions revealed distinct protein patterns in both generations. Moreover, the proteins are conspicuously expressed in neighboring tissues, but never co-localize (V. Q., W. F. and T. M., unpublished data). These results suggest that the three proteins do not fulfil redundant functions, even though the ubiquitous weak expression of *PPM2* inevitably produces some overlap. To clarify this matter exhaustively, however, double and triple knock-out mutants are necessary. Furthermore, it cannot be excluded that the three genes of the *PpMADS-S*-like subclade are able to adopt the roles of the genes of the *PPM2*-like subclade.

In which functional context might *PPM2* play a role? Despite the fact that gametangia and the sporophyte foot differ immensely in functionality and structure, they share a high requirement for metabolic energy (“energy sinks”). Sufficient nutrient supply of cells that develop into reproductive cells such as spermatozoids and egg cells, as well as tissues surrounding an egg cell and providing optimal conditions for fertilization, is indispensable. The function of the sporophyte foot in energy supply is even more obvious. In mosses, the diploid generation completely depends on nutrition by the dominant gametophyte (Browning and Gunning, 1979a,b; Renault et al., 1992). A high expression of *PPM2* in the aforementioned tissues of *P. patens* may therefore be involved in the definition of sink tissues to ensure proper development of those organs that are required for the transition into the next generation of the life cycle. The detected variability in *PPM2* protein abundance might be caused by nutrient availability, e.g. in case of starvation *PPM2* protein expression may become stronger to define important metabolic sink organs.

Whether the *PPM2* expression pattern refers to early events in the evolution of land plants cannot be answered yet.

However, due to the fact that nutrient supply and nutrient transport across an apoplastic gap exist in all groups of extant land plants, MADS-box genes might have played an important role in the origin of the underlying molecular mechanisms. Taken together, our findings emphasize that *PPM2* does not provide the typical characteristics of a classical seed plant developmental control gene, e.g. defining the identity of meristems or organs by regulating target genes in those tissues where they are expressed. By contrast, *PPM2* suggests a fundamentally different performance of MADS-box genes in basal non-seed plants. In spite of this revelation, it has to be kept in mind that the expression domains of algal MADS-box genes strongly suggest an ancestral function that is correlated with haploid reproductive organs. Therefore, we conclude that MADS-box genes originated from a common ancestor that fulfilled a basic function that was predominantly, but not exclusively, required in haploid reproductive structures.

Acknowledgments

The authors would like to thank Mitsuyasu Hasebe and Naoki Aono for sharing results prior to publication. We are grateful to Guido Jach for providing the *DsRED* gene and to Peter Huijser and Elmon Schmelzer for help with the microscopic analyses. We thank Daniela Liebsch and Wim Verelst for helpful discussions and comments on the manuscript.

Note: Upon completion of this manuscript, an electronic publication ahead of print by Singer et al. (in press) on the analysis of MIKC^c MADS-box genes in *P. patens* became available. We belatedly included comments on their major conclusions in our discussion.

References

- Browning, A.J., Gunning, B.E.S., 1979a. Structure and function of transfer cells in the sporophyte haustorium of *Funaria hygrometrica* Hedw.: I. The development and ultrastructure of the haustorium. *J. Exp. Bot.* 30, 1233–1246.
- Browning, A.J., Gunning, B.E.S., 1979b. Structure and function of transfer cells in the sporophyte haustorium of *Funaria hygrometrica* Hedw.: II Kinetics of uptake of labelled sugars and localization of absorbed products by freeze-substitution and autoradiography. *J. Exp. Bot.* 30, 1247–1264.
- Coen, E.S., Meyerowitz, E.M., 1991. The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37.
- Floyd, S.K., Bowman, J.L., 2007. The ancestral developmental tool kit of land plants. *Int. J. Plant Sci.* 168, 1–35.
- Hasebe, M., Wen, C.-K., Kato, M., Banks, J.A., 1998. Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. *Proc. Natl. Acad. Sci. U. S. A.* 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.
- Hohe, A., Rensing, S.A., Mildner, M., Lang, D., Reski, R., 2002. Day length and temperature strongly influence sexual reproduction and expression of a novel MADS-box gene in the moss *Physcomitrella patens*. *Plant Biol.* 4, 595–602.
- Hughes, T.A., 2006. Regulation of gene expression by alternative untranslated regions. *Trends Genet.* 22, 119–122.
- Jack, T., 2001. Plant development going MADS. *Plant Mol. Biol.* 46, 515–520.
- Kaufmann, K., Melzer, R., Hei, en, G., 2005. MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347, 183–198.

- Kofuji, R., Sumikawa, N., Yamasaki, M., Kondo, K., Ueda, K., Ito, M., Hasebe, M., 2003. Evolution and divergence of the MADS-Box gene family based on genome-wide expression analyses. *Mol. Biol. Evol.* 20, 1963–1977.
- Krogan, N.T., Ashton, N.W., 2000. Ancestry of plant MADS-box genes revealed by bryophyte (*Physcomitrella patens*) homologues. *New Phytol.* 147, 505–517.
- Logemann, J., Schell, J., Willmitzer, L., 1987. Improved methods for the isolation of RNA from plant tissues. *Anal. Biochem.* 163, 16–20.
- Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., et al., 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature.* 428, 653–657.
- Münster, T., Pahnke, J., Di Rosa, A., Kim, J.T., 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. U. S. A.* 94, 2415–2420.
- Münster, T., Faigl, W., Saedler, H., Theißen, G., 2002. Evolutionary aspects of MADS-box genes in the eusporangiate fern *Ophioglossum*. *Plant Biol.* 4, 474–483.
- Nishiyama, T., Hiwatashi, Y., Sakakibara, K., Kato, M., Hasebe, M., 2000. Tagged mutagenesis and gene-trap in the moss *Physcomitrella patens* by shuttle mutagenesis. *DNA Res* 7, 9–17.
- Nishiyama, T., Fujita, T., Shin, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y., Hasebe, M., 2003. Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8007–8012.
- Pařenicová, L., De Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C., Colombo, L., 2003. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS World. *Plant Cell* 15, 1538–1551.
- Renault, S., Bonnemain, J.L., Faye, L., Gaudillere, J.P., 1992. Physiological aspects of sugar exchange between the gametophyte and the sporophyte of *Polytrichum formosum*. *Plant Phys.* 100, 1815–1822.
- Schaefer, D.G., 2001. Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol.* 4, 143–150.
- Schaefer, D.G., 2002. A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. *Annu. Rev. Plant Biol.* 53, 477–501.
- Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lönnig, W.E., Saedler, H., Sommer, H., 1992. Characterization of the *Antirrhinum* floral homeotic MADS-box gene *deficiens*: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. *EMBO J.* 11, 251–263.
- Singer, S.D., Krogan, N.T., Ashton, N.W., 2007. Clues about the ancestral roles of plant MADS-box genes from functional analysis of moss homologues. *Plant Cell Rep.* doi:10.1007/s00299.007.03312.0.
- Sorensen, A., Guerineau, F., Canales-Holzeis, C., Dickinson, H.G., Scott, R.J., 2002. A novel extinction screen in *Arabidopsis thaliana* identifies mutant plants defective in early microsporangial development. *Plant J.* 29, 581–594.
- Sundström, J., Engström, P., 2002. Conifer reproductive development involves B-type MADS-box genes with distinct and different activities in male organ primordia. *Plant J.* 31, 161–169.
- Svensson, M.E., 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.
- Tanabe, Y., Uchida, M., Hasebe, M., Ito, M., 2003. Characterization of the *Selaginella remotifolia* MADS-box gene. *J. Plant Res.* 116, 71–75.
- Tanabe, Y., Hasebe, M., Sekimoto, H., Nishiyama, T., Kitani, M., Henschel, K., Münster, T., Theißen, G., Nozaki, H., Ito, M., 2005. Characterization of MADS-box genes in charophycean green algae and its implication for the evolution of MADS-box genes. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2436–2441.
- Theißen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Münster, T., Winter, K.-U., Saedler, H., 2000. A short history of MADS-box genes in plants. *Plant Mol. Biol.* 42, 115–149.
- Weise, A., Rodriguez-Franco, M., Timm, B., Hermann, M., Link, S., Jost, W., Gorr, G., 2005. Use of *Physcomitrella patens* actin 5' regions for high transgene expression: importance of 5' introns. *Appl. Microbiol. Biotechnol.* 70, 337–345.
- Zachgo, S., de Anrade Silva, E., Motte, P., Tröbner, W., Saedler, H., Schwarz-Sommer, S., 1995. Functional analysis of the *Antirrhinum* floral homeotic *DEFICIENS* gene *in vivo* and *in vitro* using a temperature-sensitive mutant. *Development* 121, 2861–2875.