Available online at www.sciencedirect.com
ScienceDirect
Gene xx (2007) xxx–xxx
GENE
www.elsevier.com/locate/gene

Comparative analysis of the SBP-box gene families in *P. patens* and seed plants

Maike Riese, Susanne Höhmann, Heinz Saedler, Thomas Münster, Peter Huijser *

Max Planck Institute for Plant Breeding Research, Cologne, Germany

Received 20 March 2007; received in revised form 14 June 2007; accepted 20 June 2007

Received by G. Theissen

Abstract

To come to a better understanding of the evolution and function of the SBP-box transcription factor family in plants, we identified, isolated and characterized 13 of its members from the moss *Physcomitrella patens*. For the majority of the moss SBP-box genes, clear orthologous relationships with family members of flowering plants could be established by phylogenetic analysis based on the conserved DNA-binding SBP-domain, as well as additional synapomorphic molecular characters. The *P. patens* SBP-box genes cluster in four separable groups. One of these consists exclusively of moss genes; the three others are shared with family members of *Arabidopsis* and rice. Besides the family defining DNA-binding SBP-domain, other features can be found conserved between moss and other plant SBP-domain proteins. An AHA-like motif conserved from the unicellular alga *Chlamydomonas reinhardtii* to flowering plants, was found able to promote transcription in a heterologous yeast system. The conservation of a functional microRNA response element in the mRNA of three of the moss SBP-box genes supports the idea of an ancient origin of microRNA dependent regulation of SBP-box gene family members.

As our current knowledge concerning the roles of SBP-box genes in plant development is scarce and the model system *P. patens* allows targeted mutation, the material we isolated and characterized will be helpful to generate the mutant phenotypes necessary to further elucidate these roles.

Keywords: Land plant evolution; Rice; *Arabidopsis*; Transcription factor; AHA-motif; miR156

1. Introduction

In plants, as in all living organisms, transcription factors represent an important level of gene regulation. They enable plants to appropriately regulate growth, differentiation and metabolism, and to respond to endogenous and environmental cues. Particular transcription factors are able to directly interact with DNA in a sequence specific manner. Evolutionary conservation of the respective DNA-binding domains, allow to define over thirty such transcription factor families in the seed plant model *Arabidopsis thaliana*, half of which appear to be unique to plants (Riechmann et al., 2000; Iida et al., 2005). One of these families consists of SBP-domain proteins (Klein et al., 1996) and is in *Arabidopsis* represented by the 17 members of the *SPL* gene family (Cardon et al., 1999).

The SBP-domain encompasses *ca.* 74 amino acid (aa) residues, harbors a nuclear localization signal at its C-terminus and is sufficient to bind DNA involving two zinc-fingers of unusual structure (Klein et al., 1996; Yamasaki et al., 2004, 2006; Birkenbihl et al., 2005).

The corresponding genes, carrying the SBP-domain encoding SBP-box, have found to be highly conserved in green plants, from unicellular algae (Kropat et al., 2005) to mono- and dicotyledonous angiosperms (Cardon et al., 1999; Xie et al., 2006).

Our current knowledge of the regulatory roles SBP-box genes may play in plant development is rather superficial and largely based on a few identified mutant phenotypes. In maize, for instance, mutations of the SBP-box genes *LG1* and *TGA1* uncovered roles in

Abbreviations: aa, amino acid(s); cDNA, DNA complementary to RNA; CFP, cyan fluorescent protein; CRR1, Copper response regulator1; EST, expressed sequence tag; LG1, Liguleless1; miRNA, microRNA; MRE, microRNA responsive element; Myr, million years; RE, responsive element; SBP, SQUAMOSA promoter binding protein; SPL, SQUAMOSA promoter binding protein like; TGA1, teosinte glume architecture1; UTR, untranslated region; YFP, yellow fluorescent protein.

* Corresponding author. Carl-von-Linné-Weg 10, 50829 Cologne, Germany. Tel.: +49 221 5062 170; fax: +49 221 5062 113.
E-mail address: huijser@mpiz-koeln.mpg.de (P. Huijser).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.06.018

leaf and glume development, respectively (Moreno et al., 1997; Wang et al., 2005). In *A. thaliana*, fertility is reduced upon *SPL8* loss-of-function and a role for *SPL8* as a local regulator of GA-mediated signalling has been suggested (Unte et al., 2003; Zhang et al., 2007). Stone et al. (2005) discovered that a reduction of *SPL14* expression increased resistance to the fungal toxin fumonisin B1. In addition, the *spl14* mutant displayed elongated petioles and enhanced leaf serration. Constitutative over-expression of *SPL3* and related SBP-box genes may cause earliness in transgenic *Arabidopsis* lines suggesting a role for these SBP-box genes during the floral transition (Cardon et al., 1997; Wu and Poethig, 2006). Interestingly, these latter genes represent a subfamily of SBP-box genes targeted by the highly similar miRNAs miR156 and miR157 (Rhoades et al., 2002; Schwab et al., 2005; Xie et al., 2006; Wu and Poethig, 2006; Gandikota et al., 2007). This interaction between SBP-box genes and miR156 seems to be of ancient origin as it could also be detected in mosses (Arazi et al., 2005). In addition to the phenotypes obtained from genetic alterations in SBP-box genes, it has recently been shown that an epigenetic mutation of an SBP-box gene causes the Colorless non-ripening phenotype in tomato (Manning et al., 2006). Finally, the only non-seed plant SBP-box gene mutant described to date concerns the *COPPER RESPONSE REGULATOR1* (*CRR1*) in *C. reinhardtii* required for both activating and repressing target genes of a copper- and hypoxia-sensing pathway (Kropat et al., 2005).

In general, the mutant phenotypes obtained so far suggest more physio-developmental roles for SBP-box genes. Furthermore, SBP-box genes in seed plants appear in moderately sized families and with sufficient degrees of similarity between different members, functional redundancy is to be expected. Together, this may lead to less obvious mutant phenotypes, especially under optimal growth conditions.

In order to reduce these restrictions in recognizing and elucidating the molecular genetic mechanisms underlying SBP-box gene actions in plant development, we decided to choose the moss *Physcomitrella patens* as a model system.

In contrast to flowering plants where a reversed situation is encountered, the life cycle of mosses is dominated by a haploid and photoautotrophic gametophyte that supports the diploid sporophytic generation (see Reski, 1998; Cove et al., 2006 for review). As most important difference to flowering plant molecular genetic models, *P. patens* offers the possibility of efficient gene targeting via homologous recombination (reviewed by Schaefer, 2002) as well as the possibility to make double or triple gene knockouts in one step (Hohe and Reski, 2003). Moreover, *P. patens* can fully complete its life cycle when grown *in vitro* and is thus easily accessible for manipulation in different environmental growth conditions.

Mosses and flowering plants are believed to be of monophyletic origin and evolutionary separated for around 500 Myr (Kenrick and Crane, 1997; Nickrent et al., 2000). Although mosses follow relatively simple developmental patterns, they do share many basic morphological features and physiological responses with other land plants, which make them interesting subjects for comparative evolutionary studies.

Here we report the molecular cloning and first characterization of 13 new SBP-box genes from the moss *P. patens*.

Phylogenetic reconstruction based on the conserved DNA-binding domain as well as comparison of additional synapomorphic molecular characters established clear orthologous relationships between the majority of the moss SBP-box genes and those of flowering plants.

2. Materials and methods

2.1. Plant material

P. patens ssp *patens* B.S.G. was grown under standard conditions as described by D.G. Schaefer (http://www.unil.ch/lps/docs/Ppprotocols2001.pdf).

2.2. Construction and screening of cDNA and genomic DNA library

We screened a RAGE (Rapid Analysis of Gene Expression)-pool of digested genomic DNA linked to adapter sequences as described by Henschel (2002) (kindly provided by K. Münster, MPIZ, Cologne, Germany) with the nested RAGE adapter-primers PAP1 (GTAATACGACTCACTATAGGGC) and PAP2 (ACTATAGGGCACGCGTGGT) and SBP-box primers SH65 (GATTACCATCGGCGGCAYAARGTNTG) and SBP1 (CATMGNTTCTGCCAGCAGTG). Based on the isolated sequences, primers were generated and a genomic library of *P. patens* was screened.

For the isolation of cDNAs representing SBP-box genes a library of cDNAs cloned in the phage vector λNM1149 was screened under stringent (2xSSC, 0.1% SDS; hybridization 68 °C, washing 60 °C) and moderate conditions (5xSSC, 0.1% SDS; hybridization 55 °C, washing 58 °C). The library represented all stages of the *P. patens* life cycle including protonema, young and mature gametophores and sporophytes of different stages.

Two SBP-box sequences encompassing only the SBP-box previously identified from a genomic library, and two EST's pp020005015 and pp15003060 from the Freiburg EST collection (Lang et al., 2005), were used for the screen. A segment of all four sequences was amplified by PCR using the following primer pairs: SH191 (TTGGGAAAGAGACATCGGGCAGG) and SH77 (CGGCAGCTTCGTTTGCCCTCGTC) for *PpSBP1*; SH143 (CGGCAGCTTCGTTTGCCCTCGTC) and SH79 (GCTTGATCCTCAACTCGAGGTGTCG) for *PpSBP2*; MR07 (CAAATTGCCGCAGTGAACTTGAGGACG) and MR09 (TGAGGAGCCCGACGAAGATTTG) for *PpSBP3*; SH178 (CATCGTCGACACAAAGTGTGTGAGC) and MR01 (GTCTTAACGCTTCATATCTTGCGAG) for *PpSBP4*. The PCR-fragments were radioactively labelled with α-^{32}P dCTP by a Klenow fill-in reaction and used as probes. A mixture of the labelled PCR-fragments of *PpSBP2* and *PpSBP4* were used in the heterologous screening of the cDNA library.

From the stringent screen we identified the full-length cDNA's of *PpSBP1*, *PpSBP3* and *PpSBP4* and the 5′ part of *PpSBP2*. Screening under moderate conditions identified the 3′ part of *PpSBP7*.

The cDNA was amplified sequenced after sub-cloning into plasmid vector pCR2.1 TOPO using the TOPO TA cloning kit

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

(Invitrogen GmbH, Karlsruhe, Germany) and the phage specific primers MR14 (CCAGTCAACACTTACGCCAAGAG) and MR15 (TCGCCTCCATCAACAAACTTTC).

2.3. Screening of the genomic sequences and isolation of additional SBP-box genes

After release of the genomic sequences of *P. patens*, we screened those with the SBP-domain of AtSPL1 and assembled them with help of an assembling program (available on PHYSCObase; http://moss.nibb.ac.jp/). Based on the genomic sequences we generated primer to amplify the deduced cDNA out of a cDNA pool, containing RNA of all developmental stages. The following primer pairs were used: MR212 (CTGGACATAACCGGCAGTGCTTGCT) and MR197 TCTCTCTGACACCAACAAAGCCCC) for *PpSBP5*; MR233 (TCTCGTCCTTCCCTGTTAC) and MR232 (CTATTTCGAGTTTAGGTAATCGTC) for *PpSBP6*; MR119 (GCAAGTTATTACCAGCAGAGC) and MR335 (ACGTGATTCTAGGCCTGGAC) for *PpSBP7*; MR237 (TTCCCAATACTGATGAGACG) and MR236 (TATTTGGTGCTCCCGTGAAC) for *PpSBP9*; MR207 (CTTGTTGCGGCACTTGCTTG) and MR283 (ACAGAGCCATCCAGTTCACAGTC) for *PpSBP10*; MR291 (CACACACA-CACACACACACG) and MR294 (GGCAGATGGAGGTCATAATAC) for *PpSBP11*; MR242 (AATCTTGTCATCCAGCCG) and MR243 (GCAAGTGTGTTTCCTCCAGTGACA) for *PpSBP12*; MR244 (TTCTAAAAGGAGTCGGCAG) and MR245 (CTTCAGATTACATTCTCGGC) for *PpSBP13*.

2.4. DNA sequencing

The sequencing was done by the MPIZ Automatic DNA Isolation and Sequencing (ADIS) core facility on Applied Biosystems (Weiterstadt, Germany) Abi Prism 377 and 3700 sequencers using BigDye-terminator chemistry.

2.5. Sequence alignments and phylogenetic reconstruction

Multiple alignments of amino acid sequences were generated by the program ClustalW of the MacVector 7.2.2 software package (Accelrys Ltd., Cambridge, UK) using the BLOSUM 30 matrix with an open gap penalty of 10 and an extend gap penalty of 0.05. For the phylogenetic reconstruction two extra residues, one upstream and one downstream, were added to the SBP-domain of 74 aa residues (compare to sequence logo provided by Birkenbihl et al., 2005) resulting in a 76 aa sequence used for alignment. The tree was constructed using the neighbor-joining algorithm of the MacVector 7.2.2 software package.

2.6. Transcriptional activation assay in yeast

The *Saccharomyces cerevisiae* strain AH109 and the GAL4 DNA-binding domain vector pGBKT7 used to test transcriptional activation in yeast were obtained from BD Biosciences as part of the Matchmaker Two-hybrid System 3 (Clontech Inc., Mountain View, USA). The wild-type constructs were generated by PCR-amplification of the cDNA. We used the following primer pairs: MR112 (TATCCATGGGTTGCAGTTCAGATCCATTAG) and MR113 (TATGAATTCAGCCCTTACATCCAACTGTAAG) for *PpSBP2*; MR131 (ATTATGAATTCAGCAGCTAAGCTCACTGCACTG) and MR114 (TATCCATGGCATTGATCCTCACTTGTCATTGCCTTCATC) for *AtSPL1*; SH166 (CGAGAGCCATGGATGAGGTAGGAGC) and o168 (GCGGGATCCTCATCCCAGATTCAAATCAAGTCC) for *AtSPL14* with appropriate cloning sites, *Nco*I/*Eco*RI in the case of *PpSBP1*, *PpSBP2* and *PpSBP3*; *Nco*I/*Bam*HI in the case of *AtSPL1* and *AtSPL14*. For further analysis of the different AHA-motifs, W to R aa residue substitutions were generated using a PCR-based mutagenesis method as described by Lyck et al. (1997). In the following primers used, the codons originally encoding for a W residue have been changed into codons for R (underlined). In addition a translational neutral T to C nucleotid exchange in the primer for *AtSPL1* introduced a diagnostic *Pvu*I restriction site. All mutations are indicated in bold: o175 (GTGGAG**C**GGGATTTGAA**C**GAT**C**GGAAA**C**GGGATG) for *AtSPL1*; o176 (GATGAG**C**GGAATTCGAAGATG**C**G**C**GAT**C**GGGATAG) for *AtSPL14*. Constructs were transformed into yeast strain AH109 and selected on media lacking trypthophane. The ability to grow in the absence of histidine is a test for transcriptional activation of the His reporter gene in yeast. At least five independent transformants were checked in this assay and per construct identical results were obtained.

2.7. Transient assay in moss protoplasts

We transformed moss protoplast according to the protocol described in Schaefer et al. (1991). We used 30 μg per linearized plasmid (see Section 3.5) and cotransfomed them into moss protoplast. After 3 days in darkness, we analyzed the expression of the fluorescent proteins by fluorescent microscopy and measured the intensities from digital images with ImageJ (Rasband, W.S., NIH, USA, http://rsb.info.nih.gov/ij/) as described by Gandikota et al. (2007).

2.8. Accession numbers

The following GenBank accession nos. corresponds to the cloned cDNAs of *PpSBP1* to *-7* and *PpSBP9* to *-13*: AJ968320, AJ968403, AJ968318, AJ968319, EF016491, EF016492, EF016493, EF016494, EF016495, EF647594, EF016496, EF016497. As no cDNA could be isolated for *PpSBP8*, its exon–intron structure and encoded protein were derived from accession fgenesh1_pg.194000034 in the first release of the annotated *P. patens* genome sequence.

3. Results and discussion

3.1. Identification and isolation of SBP-box genes from P. patens

No SBP-box related sequences of *P. patens* were available in the public databases at the time we started our attempt to isolate SBP-box genes from this organism. Initially, therefore, PCR primers were designed based on conserved heterologous SBP-

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

box sequences of the *Arabidopsis SPL* genes (Cardon et al., 1999) and on two *P. patens* SBP-box sequences identified in the Freiburg EST collection (kindly provided by R. Reski, University of Freiburg, Germany; Lang et al., 2005; see Materials and methods). With the help of these primers, fragments of *P. patens* genomic DNA could be amplified and, after sequencing, identified to represent four different SBP-box genes designated as *PpSBP1 to -4* for *Physcomitrella patens SBP-box* gene.

From mid-2005 raw *P. patens* genomic sequences were released in the public databases and, through comparative analysis, additional 9 SBP-box genes (designated *PpSBP5 to -13*) could be identified.

The collected genomic sequence data allowed the isolation of cDNAs covering the complete coding regions for all *PpSBP* genes except *PpSBP8*. The respective cDNAs were obtained either through screening of a phage cDNA library or through direct amplification and cloning from a cDNA pool, both representing mixed poly $(A)^+$ RNA from different developmental stages (see Materials and methods). Their sequences have been deposited in GenBank (http://www.ncbi.nlm.nih.gov/). The failure to isolate *PpSBP8* cDNA is probably due to a low level or absence of expression and also nowadays *PpSBP8* derived sequences could not be identified in the public EST collection (PHYSCObase, http://moss.nibb.ac.jp/).

3.2. Structural organization of SBP-box genes in P. patens

The intron–exon structure of the *P. patens* SBP-box genes, as deduced from the co-linearity of the cDNA and genomic sequences, is depicted in Fig. 1A. As no cDNA sequence could be obtained for *PpSBP8*, its structure is predicted on the basis of sequence similarities to other SBP-box genes. All introns but one have the conserved GT and AG dinucleotides at their termini. In the first intron of *PpSBP3* an AA dinucleotide replaces the common AG at the splice acceptor site. It should be noted that for none of the *Physcomitrella* SBP-box genes the transcriptional start site has been experimentally determined. The true start of the first intron thus remains uncertain. Also, with the exception of *PpSBP1*, *-3*, *-4* and *-7* were cDNAs with poly(A)-tails could be isolated, the end of the last exon remains unknown.

All *Arabidopsis SPL* genes carry an intron at a conserved position within the SBP-box, a feature generally found in SBP-box genes of other flowering plants as well (Xie et al., 2006; own unpublished observations). An intron at the same position is also found in the *P. patens* SBP-box genes with the exception of *PpSBP2* and *PpSBP10*, which lack an intron in the SBP-box. Roy and Penny (2007) obtained that in the evolution of land plants intron-loss is more likely than intron-gain, therefore we conclude that *PpSBP2* and *PpSBP10* lost the intron in the SBP-box. In most known SBP-box genes, for instance in 15 out of the 17 *SPL* genes in *Arabidopsis*, the SBP-box is encoded within the first two exons. However, most of the *P. patens* SBP-box genes do have additional exons upstream of the SBP-box.

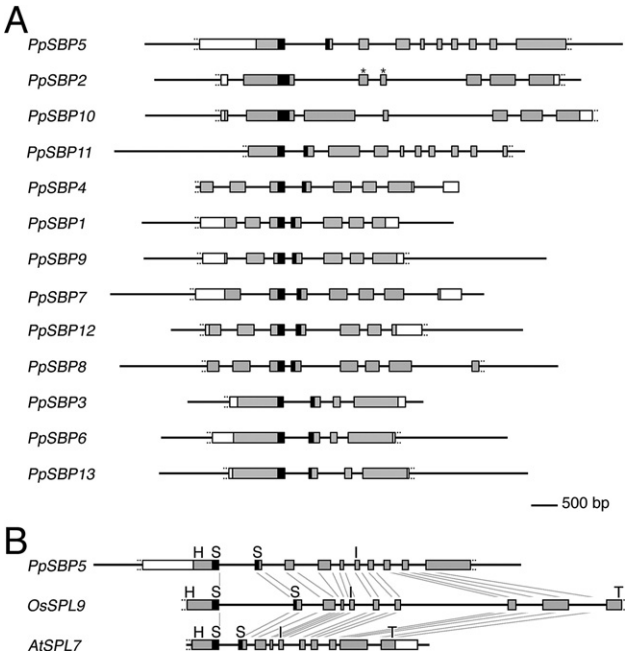

Fig. 1. Genomic organization of *Physcomitrella* SBP-box genes. (A) Schematic representation of the exon–intron structure of the identified *Physcomitrella patens* SBP-box gene loci. Note that the structure of *PpSBP8* is predicted based on its genomic sequence only. (B) Comparison of the genomic organization of *PpSBP5* with other Group I representatives of flowering plants. Exons are represented by boxes. The SBP-box sequences are depicted in black, the remaining coding sequences in gray and the untranslated 5′ and 3′ regions are represented white. The asterisks near two exons of *PpSBP2* indicate uncertainty concerning their prediction due to sequencing problems. General uncertainty with respect to the start of the first exon and the end of the last exon is indicated by short dashed line extensions. bp, base pairs. Coding regions for conserved domains in Group I members are marked by the following letter code in (B): H, AHA-like motif; S, SBP-domain; I, IRPGC-domain; T, transmembrane domain.

3.3. Comparative evolutionary analysis of SBP-box genes between P. patens and seed plants

As representatives of flowering plants for an evolutionary comparison of *P. patens* SBP-box genes, we selected *Arabidopsis* and rice as their fully annotated genomes are available and all of their SBP-box gene family members known (Xie et al., 2006). Except for the conserved DNA-binding domain, the entire set of SBP-domain proteins from *Arabidopsis*, rice and *P. patens* share no further extensive sequences similarities leaving the SBP-domain as the starting point for a phylogenetic reconstruction (Fig. 2). Based on this phylogenetic tree, rooted with *C. reinhardtii* CRR1, the *P. patens* SBP-box genes can be divided over four different groups, in three of which they cluster with genes of *Arabidopsis* and rice. In addition, at least another two groups can be recognized with no *P. patens* representatives.

Not all branches of the phylogenetic tree are supported by high bootstrap values, *e.g.* above 50%. Therefore, the structural properties of the genes and their products within the different groups were analyzed in more detail. Exon–intron structures are to some degree conserved within the different groups, for example within Group I (Fig. 1B). More prominent as molecular synapomorphic characters are particular protein sequence motifs

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

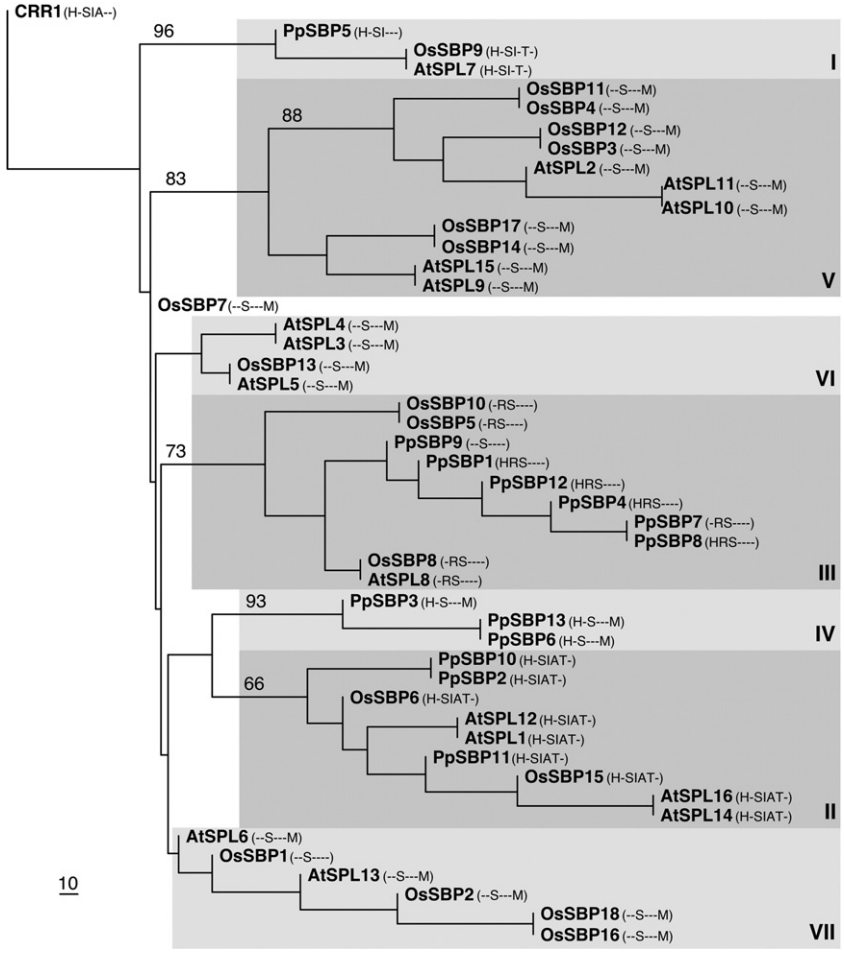

Fig. 2. Reconstruction of the phylogenetic relationships within the SBP-box gene family. The phylogenetic tree is based on the conserved SBP-domain and rooted with CRR1 from *Chlamydomonas reinhardtii* as outgroup. Only bootstrap values over 50% are shown and groups supported by such values and/or additional conserved domains, are boxed in shades of grey and labelled with Roman numbers. Species and gene names are indicated at the end of each branch and followed by a letter code between brackets indicating the presence of conserved sequences characteristic for the different subfamilies. Note that one member of the rice SBP-box gene family, OsSBP19, is excluded from the phylogenetic reconstruction as full-length SBP-domain sequence is not available. Furthermore, AtSPL13 actually represents two fully identical proteins encoded by two perfectly duplicated genes in the genome of *Arabidopsis* that are known as *AtSPL13A* and *-13B*. H, AHA-like motif; R, RTYF-domain; S, SBP-domain; I, IRPGC-domain; A, ankyrin repeat region, T, transmembrane domain; M, MRE.

identified outside of the conserved SBP-domain. That will be described per group in the following sections.

3.3.1. Group I

Group I is somewhat remarkable in that it is represented with single genes in all three species, *i.e. PpSBP5* in *P. patens*, *AtSPL7* in *Arabidopsis* and *OsSBP9* in rice. Within the SBP-domain of the respective proteins, four cysteine residues coordinate the first Zn-ion whereas in all other SBP-domains this is accomplished by three cysteine and one histidine residues (Yamasaki et al., 2004). Furthermore, in these genes the sequences immediate flanking the aforementioned conserved and SBP-box specific intron, encode GKF whereas in all other land plant SBP-box genes this is SRF. Interestingly, in *C. reinhardtii CRR1* this particular SBP-box specific intron is lacking and at the respective position an even different motif, GRF, is encoded. However, analysis of other *C. reinhardtii* SBP-domain proteins (*C. reinhardtii* genome v3.0; http://genome.jgi-psf.org/Chlre3/Chlre3.home.html) showed that also GKF and SRF could already be found like in the land plants. In addition, *C. reinhardtii* SBP-domain proteins show more variation as the residues GRL, SKF, NKF, AKF can also be found. As none of the predicted *C. reinhardtii* SBP-box genes carry an intron at this position, it seems plausible to assume that the presence of an intron has fixed the coding sequence around this respective position in land plants. If in the evolution early land plants gained this intron or *C. reinhardtii* lost it is unknown.

Outside and N-terminal of the SBP-domain, PpSBP5, AtSPL7 and OsSBP9 share some conserved aromatic (W, F, Y), acidic (E, D) and large hydrophobic (L, I, V) aa residues. This amino acid composition is characteristic of many transcriptional activation domains found in mammals, yeast as well as in plants, in which they are referred to as AHA (Aromatic, Hydrophobic, Acidic) motifs (Nover and Scharf, 1997; Döring et al., 2000). Because of this resemblance, the comparable motif conserved in SBP-domain proteins will be referred to as AHA-like. In the middle of the protein a conserved region around the aa residues IRPGC can be found (Fig. 3C). This motif, also found in *C. reinhardtii* CRR1 (Kropat et al., 2005), is only known from SBP-domain proteins and its biological role remains unknown.

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

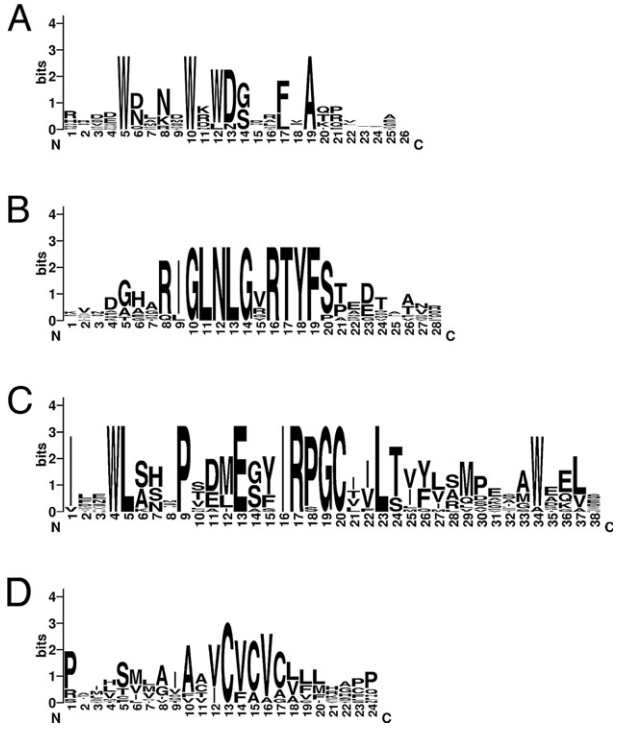

Fig. 3. Protein domains found to be conserved within different groups of SBP-domain proteins. (A) Amino acid sequence logo of the AHA-like1 motif based on the 9 proteins found in Group II. Each logo consists of stacks of symbols, one stack for each position in the sequence. The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each amino acid at that position (Crooks et al., 2004). (B) Amino acid sequence logo of the RTYF-motif based on 9 out of the 10 proteins found in Group III. (C) Amino acid sequence logo of the IRPGC motif based on the 9 proteins found in Group II. (D) Amino acid sequence logo of the transmembrane domain based on all the proteins, except PpSBP5, found in Groups II and I.

3.3.2. Group II

The second phylogenetic discernible group of SBP-box genes consists of the *P. patens* genes *PpSBP2*, *-10* and *-11*, the *Arabidopsis* genes *AtSPL1*, *-12*, *-14* and *-16* and *OsSBP6* and *-15* from rice.

The respective proteins within this group share extensive similarities to those of Group I as recognized in the presence of an AHA-like motif upstream of the SBP-domain (Fig. 3A), an IRPGC motif downstream (Fig. 3C) and a predicted C-terminal transmembrane domain (Fig. 3D). In addition, the members of Group II share an ankyrin repeat region (PROSITE database profile PS50297, http://www.expasy.ch/prosite/). Ankyrin repeat regions are generally known to be involved in protein–protein interactions (Cai and Zhang, 2006) and a similar function in SBP-domain proteins of Group II may be proposed. The interacting protein partners, however, remain unknown. Interestingly, the ankyrin repeat region is not predicted to be present in Group I SBP-domain proteins but is for *C. reinhardtii* CRR1 (Kropat et al., 2005). The C-terminal part of CRR1 is very cysteine-rich and a role in metal binding has been proposed (Kropat et al., 2005). Group II proteins do share some positionally conserved Cysteine residues in their sub-carboxyl terminal region, in particular as part of a sequence in the following referred to as transmembrane-motif (Fig. 3D). If these are involved in metal-ion binding is unknown. Interestingly, the predicted transmembrane alpha-helix (http://www.cbs.dtu.dk/services/TMHMM-2.0/; Krogh et al., 2001) allows the speculation that Group II proteins represent membrane-associated transcription factors (Kim et al., 2006). The transmembrane region can also be recognized in the *Arabidopsis* and rice members of Group I but not in *P. patens* PpSBP5. The single transmembrane domain may have evolved from a region with multiple membrane spanning domains identified in the predicted *C. reinhardtii* SBP-domain protein estExt_fgenesh2_pg.c_50271_Chlr (http://genome.jgi-psf.org/Chlre3/Chlre3.home.html). If the respective SBP-domain proteins indeed represent membrane-tethered transcription factors they probably have to become released by RIP or RUP-like mechanisms (Hoppe et al., 2001) in response to unknown stimuli, in order to enter the nucleus and unfold their transcription regulatory role.

3.3.3. Group III

Whereas *Arabidopsis* is represented with only one, *AtSPL8*, and rice with three SBP-box genes, six *P. patens* genes cluster in Group III. In their N-terminal region the proteins encoded by these genes share a short stretch of conserved aa residues, GLNLGXRTYF (Fig. 3B), previously noticed for AtSPL8 and other flowering plant SBP-domain proteins (Unte et al., 2003). Remarkably, this motif is lacking from the predicted PpSBP9 but can still be found encoded by an intron present at a corresponding position within the gene.

3.3.4. Group IV

In the phylogenetic tree the remaining *P. patens* SBP-box genes *PpSBP3*, *PpSBP6* and *PpSBP13* cluster on a branch without any representatives of *Arabidopsis* or rice (Fig. 2).

It had previously been shown that *PpSBP3* transcripts are targeted by the evolutionary well-conserved miR156 (Arazi et al., 2005) and the corresponding miRNA response element (MRE) can also be found in *PpSBP6 and PpSBP13* (see Axtell and Bartel, 2005, for detection of miR156/157 in other major clades of land plants). Interestingly, the miR156-RE seems to be widely spread among SBP-box genes of seed plants. For instance this MRE is found in 11 out of the 17 members of the *AtSPL* genes and its functionality has been proved for several of them (Chen et al., 2004; Vazquez et al., 2004; Wu and Poethig, 2006; Gandikota et al., 2007). In rice 11 out of the 19 members show also this MRE which functionality was proven by Xie et al. (2006). In fact, all *Arabidopsis* and rice SBP-box genes that are not clustering in any of the previously discussed Groups I, II and III, and marked in Fig. 2 as Group V, VI and VII carry a miR156-RE.

3.3.5. SBP-box gene subfamilies not represented in P. patens

Group I, represented by only one member in each of the three species compared, has not previously been defined by Cardon et al. (1999) but in Section 3.3.2 defined Group II coincides with the SPL1-subfamily and Group III with the LG1-subfamily. Groups V, VI and VII seem not to be represented in *P. patens*

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

but include, respectively, the SPL2-/SPL9- (Group V) and SPL4-subfamilies (Group VI) as previously defined by these authors. A subdivision of Group V in an SPL2-subfamily, redefined through the inclusion of the rice proteins OsSBP3, -4, -11 and -12, and an SPL9-subfamily including OsSBP17 and OsSBP14 seems to find support in some weakly conserved sequences outside the SBP-domain as outlined below.

Close to their N-terminus, the SPL2-subfamily proteins with the exception of AtSPL2, show two conserved tryptophane residues in the sequence WDW, reminiscent of the AHA-like motifs described for Group I and II members. In addition, these proteins, with the exception of OsSBP4 and -11, conserved the sequence LKLGKRTY *ca.* 40 aa residues upstream of their SBP-domain. This latter motif resembles in sequence, as well as in position, the motif depicted in Fig. 3B and found to be characteristic for Group III.

The members of the SPL9-subfamily within Group V lack any positionally conserved tryptophane residues in their N-terminal region. They conserved, however, the motif GLXFGXKIYFE, reminiscent and found at a similar position with respect to the SBP-domain, of the motif depicted in Fig. 3B and characteristic for LG1-subfamily members. Within the phylogenetic tree build on the basis of the SBP-domain, the clades marked as Groups VI and VII (Fig. 2) are not well supported. The genes clustering in Group VI do, however, share some features that allow them to be distinguished from other SBP-box genes. In particular, these genes are all predicted to carry a miR156-RE within the 3′ UTR of their transcripts whereas in other SBP-box genes targeted by this miRNA, the response element is found within the coding region. In addition, the *Arabidopsis* and closely related SBP-domain proteins of other eudicot species, for instance *AmSBP1* and *AmSBP2* from *A. majus* (Klein et al., 1996) and *BpSBP1 B. pendula* (Lännenpää et al., 2004), carry an acidic domain in their N-terminal region. This enrichment of acidic residues however, is not obvious in monocot members.

The remaining SBP-box genes *AtSPL6*, *AtSPL13*, *OsSBP1*, *OsSBP2*, *OsSBP16* and *OsSBP18* brought together in the poorly defined Group VII, seem not to share in sequence any conserved protein domain outside of the SBP-domain. Their transcripts do carry a miR156-RE with the exception of OsSBP1. It should be noted that as the MRE is always found in the same reading frame it results in a small conserved hexad motif, ALSLLS, on the level of the proteins.

3.4. Three variants of an AHA-like motif are present in P. patens SBP-domain proteins

As described in Sections 3.3.1 and 3.3.2, an AHA-like motif has been uncovered upstream of the SBP-domain of all Group I and II proteins. Remarkably, also outside these groups all *P. patens*, with PpSBP9 and PpSBP7 as possible exceptions, but none of the flowering plant SBP-domain proteins carry AHA-like motifs upstream of the DNA-binding domain. In both *PpSBP9* and *PpSBP7* an in-frame AHA-like motif encoding sequence can still be recognized upstream of the putative translational start codons. In PpSBP7 it is however separated from the main ORF through a stop codon. As an AHA-like motif is already present at a comparable position in CRR1, it is well possible that the seed plant proteins lost this motif in evolution.

Upstream of two highly positionally conserved aromatic tryptophane residues present in all AHA-like motifs, often a third aromatic residue can be found with a for the different groups characteristic spacing to the others. For Group I and II, as well as CRR1, the AHA-like motif can be described as W/Y-X_4-WXW, for Group III as Y-X_7-WXW and Group IV as W/Y-X_6-WXW. To determine if these variant AHA-like motifs, hereinafter referred to as respectively AHA-like1 to -3, may indeed act as transcription activation domains, their ability to promote transcription was tested in a heterologous yeast assay.

As Stone et al. (2005) already reported that the N-terminal part of AtSPL14 could activate transcription in yeast, in a first experiment the regions encoding the AHA-like1 motifs of AtSPL1 and AtSPL14, encompassing respectively 69 and 87 aa residues, were subcloned in-frame into a yeast two-hybrid bait vector, carrying the GAL4 DNA-binding domain (GAL4BD; see Materials and methods). In parallel, mutated versions were introduced resulting in the substitution of the three highly conserved tryptophanes into arginine residues. The constructs were transformed into the yeast strain AH109 that in the absence of histidine requires for its growth the activation of a His reporter gene containing a GAL4-responsive upstream activator sequence. As shown in Fig. 4, the GAL4BD fused to the wild-type AtSPL1 and AtSPL14 AHA-like1 motifs allowed the yeast to grow on plates lacking histidine, while the tryptophane mutation constructs like the control (GAL4BD alone) failed to do so (Fig. 4).

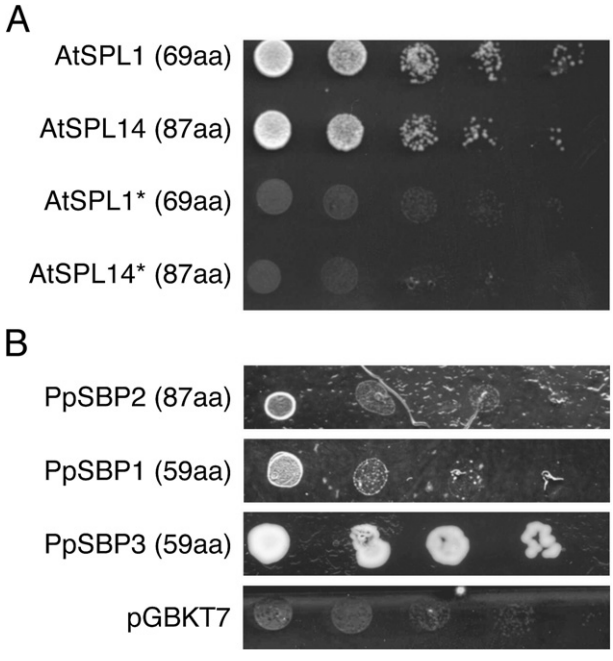

Fig. 4. Transcriptional activation assay in yeast. Dilution series (1:10 per step) of yeast strain AH109 expressing Gal4DB fused to the AHA-like motif of (A): AtSPL1, AtSPL14 and respective mutated versions indicated with an asterisk and (B) PpSBP2, PpSBP1 and PpSBP3. Transcriptional activation, indicated by the growth of the yeast strain on media lacking His, was obtained in the yeast strains expressing the AHA-like motif of AtSPL1, AtSPL14 and PpSBP3.

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

Subsequently, the yeast transactivation assay was repeated with peptides representing the three *P. patens* variants of the AHA-like motifs, respectively 87 aa residues for PpSBP2 AHA-like1, 59 aa residues for PpSBP1 AHA-like2 and 59 aa residues for PpSBP3 AHA-like3. Surprisingly, only the AHA-like3 motif convincingly showed transcriptional activation in yeast (Fig. 4).

These results indicate that at least seed plant AHA-like1 domains are likely transactivation domains and that the highly conserved tryptophane residues are essential for this function. Also the *P. patens* AHA-like3 motif represents a likely activation domain. It cannot, however, be concluded that the *Physcomitrella* AHA-like1 and -2 motifs are not part of transcriptional activation domains as the experiments were conducted in a heterologous system. Besides the requirement for factors possibly absent in yeast, another obvious possibility is that the peptides did not assume the proper structural confirmation. In the case of the *P. patens* PpSBP2 AHA-like1 this seems unlikely as its length is similar to the *Arabidopsis* AtSPL1 and AtSPL14 derived AHA-like1 motifs. In comparison to the *Arabidopsis* AHA-like1 motifs, another reason for the different behaviour of the PpSBP2 AHA-like1 motif in yeast could be the lack of an aromatic residue, *i.e.* phenylalanine, present downstream of the highly conserved tryptophanes. This phenylalanine seems to be conserved in many seed plant AHA-like1 motifs and thus of possible functional relevance.

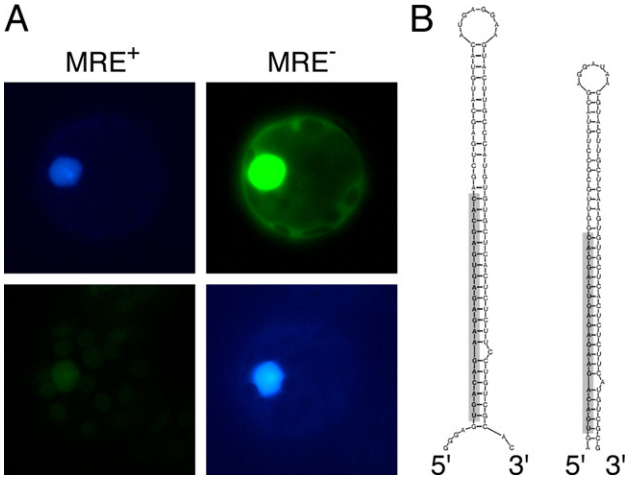

Fig. 5. SBP-box gene miR156-RE functionality assay. (A) Moss protoplasts co-transformed either with *SPL3-UTRwt-CFP* and *SPL3-UTR△4-YFP* (upper row) or with *SPL3-UTRwt-YFP* and *SPL3-UTR△4-CFP* (lower row). Relative fluorescence intensities of the tagged proteins are dependent on the presence (MRE^+) or absence (MRE^-) of a functional miR156-RE, irrespective of the fluorophore. A stronger fluorescence of nuclear localized protein is observed in the case of a non-functional miR156-RE. CFP fluorescence is depicted in blue and YFP in green. (B) Predicted hairpin structures within the putative transcripts of the *PpMIR156A* (left) and *PpMIR156B* (right) loci (predicted by mfold, Zuker et al., 1999). A gray box marks position of the miR156 sequence within the stems of the respective hairpins.

3.5. A seed plant miR156-RE is functional in P. patens

The presence of miR156 as well as the validation of a predicted target, *PpSBP3*, has been demonstrated in *P. patens* (Arazi et al., 2005). To test if a seed plant derived miR156-RE would be functional in *P. patens*, a transient assay was performed in moss protoplasts similar to that conducted by Gandikota et al. (2007). For this purpose expression constructs were used where AtSPL3 was translationally fused to either CFP or YFP as fluorophores and that carried either the wild-type 3′ UTR (*SPL3-UTRwt*) or a 3′ UTR with an MRE altered such that an interaction with miR156 would be abolished (*SPL3-UTR△4*; Gandikota et al., 2007). Subsequently, moss protoplasts were co-transformed with both the constructs *SPL3-UTRwt* and *SPL3-UTR△4* either in combination with CFP and YFP, respectively, or with the fluorophores swapped. After 3 days of transient expression, the constructs carrying the mutated miR156-RE resulted in significant higher fluorescence intensity relative to those with the wild-type miR156-RE (Fig. 5A). We determined the intensity of the fluorophores in 10 protoplasts per transformation (see Materials and methods) and found that the relative intensities of the respective fluorophores increased at least two-fold higher when linked to the mutated miR156-RE (p-values < 0.0001).

Therefore, it is reasonable to assume that a seed plant miR156-RE is able to direct the repression of mRNAs in *P. patens*. Furthermore, this suggests that both seed and non-seed plants use highly similar mechanisms to recognize and repress miRNA targeted mRNA.

Arazi et al. (2005) already proved the presence of miR156 in *P. patens* and the experiment also shows that moss protoplasts most likely express a miR156-encoding locus. Computational analysis of the *P. patens* genome revealed two candidate loci for generating short-hairpin precursors as possible substrates for a DICER-like activity to produce miR156 (Fig. 5B). Transcriptional activity of at least one of these loci could be shown with the help of RT-PCR and primers based on the predicted sequences (Supplementary data).

4. Conclusions

At least 6 subfamilies of SBP-Box genes can be recognized among land plants, three of which are represented by members of both *P. patens* and seed plants. Five of these subfamilies have also been recognized in rice (Xie et al., 2006).

Even after more than 450 Myr of evolutionary separation (Lewis and McCourt, 2004) Group I and II family members still show, in addition to the SBP-domain, extensive and exclusive sequence similarities to *C. reinhardtii* CRR1. As the SBP-domain has not been found outside the plant kingdom, these modern proteins thus probably most closely resemble the earliest form of an SBP-domain protein common to all plants. In addition these proteins conserved an AHA-like transcription activation domain, thereby supporting the notice that original SBP-domain proteins represent genuine transcription factors.

The plant specific SBP-box gene family does not concur with the general observation that conserved gene families become larger from green algae to seed plants (Richardt et al., 2007). We found a comparable number of family members in *C. reinhardtii* (19–22), *P. patens* (13), *A. thaliana* (17) and *O. sativa* (19). However, their distribution over the described subfamilies seems specific for the investigated plants.

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

Comparative analysis of the exon–intron structure of SBP-box genes revealed a high conservation in land plants and a lower degree of conservation between green algae and land plants, as already discussed (Roy and Penny, 2007).

The described analysis marks a starting point to uncover the function of SBP-box genes in *P. patens*. Future analysis will show if the evolutionary conserved molecular features also reflect biological conserved functions in the development of seed and non-seed plants.

Acknowledgements

We are grateful to Dr. Hans Sommer for construction of the *P. patens* cDNA library, Dr. Ulrike Hartmann for providing the AtSPL-yeast-two-hybrid constructs, Dr. Ralf Reski for screening his *P. patens* EST collection for putative SBP-box gene related sequences. In addition we thank the MPIZ Automatic DNA Isolation and Sequencing (ADIS) core facility for DNA sequencing and all the members of our lab for the many discussions and their valuable comments on the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2007.06.018.

References

Arazi, T., Talmor-Neiman, M., Stav, R., Riese, M., Huijser, P., Baulcombe, D.C., 2005. Cloning and charaterization of micro-RNAs from moss. Plant J. 43, 837–848.

Axtell, M.J., Bartel, D.P., 2005. Antiquity of microRNAs and their targets in land plants. Plant Cell 17, 1658–1673.

Birkenbihl, R.P., Jach, G., Saedler, H., Huijser, P., 2005. Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. J. Mol. Biol. 352, 585–596.

Cai, X., Zhang, Y., 2006. Molecular evolution of the *Ankyrin* gene family. Mol. Biol. Evol. 23, 550–558.

Cardon, G.H., Höhmann, S., Nettesheim, K., Saedler, H., Huijser, P., 1997. Functional analysis of the *Arabidopsis thaliana* SBP-box gene *SPL3*: a novel gene involved in floral transition. Plant J. 12, 367–377.

Cardon, G., Höhmann, S., Klein, J., Nettesheim, K., Saedler, H., Huijser, P., 1999. Molecular characterization of the *Arabidopsis* SBP-box genes. Gene 237, 91–140.

Chen, J., Li, W.X., Xie, D., Peng, J.R., Ding, S.W., 2004. Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microRNA in host gene expression. Plant Cell 16, 1302–1313.

Cove, D.J., Bezanilla, M., Harrues, P., Quatrano, R., 2006. Mosses as model system for study of metabolism and development. Annu. Rev. Plant Biol. 57, 497–520.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Döring, P., Treuter, E., Kistner, C., Lyck, R., Chen, A., Nover, L., 2000. The role of AHA motifs in the activator function of tomato heat stress transcription factors HsfA1 and HsfA2. Plant Cell 12, 265–278.

Gandikota, M., Birkenbihl, R.P., Höhmann, S., Cardon, G.H., Saedler, H., Huijser, P., 2007. The miRNA156/157 recognition element in the 3′ UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J. 49 (4), 683–693.

Henschel, K., 2002. Strukturelle und funktionelle Charakterisierung von MADS-Box-Genen aus dem Laubmoos *Physcomitrella patens* (Hedw) B.S.G. Dissertation University of Cologne.

Hohe, A., Reski, R., 2003. A tool for understanding homologous recombination in plants. Plant Cell Rep. 21, 1135–1142.

Hoppe, T., Rape, M., Jentsch, S., 2001. Membrane-bound transcription factors: regulated release by RIP or RUP. Curr. Opin. Plant Biol. 13, 344–348.

Iida, K., et al., 2005. RARTF: database and tools for complete sets of *Arabidopsis* transcription factors. DNA Res. 12, 247–256.

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

Kim, S.-Y., et al., 2006. A membrane-bound NAC transcription factor regulates cell division in *Arabidopsis*. Plant Cell 8, 3132–3144.

Klein, J., Saedler, H., Huijser, 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. 250, 7–16.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

Kropat, J., Tootey, S., Birkenbihl, R.P., Depège, N., Huijser, P., Merchant, S., 2005. A regulator of nutritional copper signaling in *Chlamydomonas* is an SBP domain protein that recognizes the GTAC core of copper response element. Proc. Natl. Acad. Sci. U. S. A. 102, 18730–18735.

Lang, D., Eisinger, J., Reski, R., Rensing, S.A., 2005. Representation and high-quality annotation of the *Physcomitrella patens* transcriptome demonstrates a high proportion of proteins involved in metabolism in mosses. Plant Biol. 7, 238–250.

Lännenpää, M., Janonen, I., Holtta-Vouri, M., Gardemeister, M., Porali, I., Sopanen, T., 2004. A new SBP-box gene *BpSPL1* in silver birch (*Betula pendula*). Physiol. Plant. 120, 491–500.

Lewis, L.A., McCourt, R.M., 2004. Green algae and the origin of land plants. Am. J. Bot. 91, 1535–1556.

Lyck, R., Harmening, U., Höhfeld, I., Treuter, E., Scharf, K.D., Nover, L., 1997. Intracellular distribution and identification of the nuclear localization signals of two plant heat-stress transcription factors. Planta 202, 117–125.

Manning, K., et al., 2006. A naturally occuring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat. Genet. 38, 948–952.

Moreno, A., Harper, L.C., Krueger, R.W., Dellaporta, S.L., Freeling, M., 1997. *Liguleless1* encodes a nuclear-localized protein required for induction of ligules and auricles during maize leaf organogenesis. Genes Dev. 11, 616–628.

Nickrent, D.L., Parkinson, C.L., Palmer, J.D., Duff, R.J., 2000. Multigene phylogeny of land plants with special reference to bryophytes and the earliest land plants. Mol. Biol. Evol. 17, 1885–1895.

Nover, L., Scharf, K.D., 1997. Heat stress proteins and transcription factors. Cell Mol. Life Sci. 53, 80–103.

Reski, R., 1998. Development, genetics and molecular biology of mosses. Bot. Acta 111, 1–15.

Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., Bartel, D.P., 2002. Prediction of plant microRNA targets. Cell 110, 513–520.

Richardt, S., Lang, D., Reski, R., Frank, W., Rensing, S.A., 2007. PlanTAPDB, a phylogeny-based resource of plant transcription-associated proteins. Plant Physiol. 143, 1452–1466.

Riechmann, J.L., et al., 2000. *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Science 290, 2105–2110.

Roy, W.S., Penny, D., 2007. Patterns of intron loss and gain in plants: intron loss-dominated evolution and genome-wide comparison of *O. sativa* and *A. thaliana*. Mol. Biol. Evol. 24, 171–181.

Schaefer, D.G., 2002. A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. Annu. Rev. Plant Biol. 53, 477–501.

Schaefer, D.G., Zryd, J.-P., Knight, C.D., Cove, D.J., 1991. Stable transformation of the moss *Physcomitrella patens*. MGG 226, 418–424.

Schwab, R., Palatnik, J.F., Rieter, M., Schommer, C., Schmid, M., Weigel, D., 2005. Specific effects of microRNAs on the plant transcriptome. Dev. Cell 8, 517–527.

Stone, J.M., Liang, X., Nekl, E., Stiers, J., 2005. *Arabidopsis AtSPL14*, a plant specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. Plant J. 41, 744–754.

Unte, U.S., et al., 2003. *SPL8*, an SBP-box gene that affects pollen sac development in *Arabidopsis*. Plant Cell 15, 1009–1019.

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018

Vazquez, F., Gasciolli, V., Crété, P., Vaucheret, H., 2004. The nuclear dsRNA binding protein *HYL1* is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. Curr. Biol. 14, 346–351.

Wang, H., et al., 2005. The origin of naked grains of maize. Nature 436, 714–719.

Wu, G., Poethig, R.S., 2006. Temporal regulation of shoot development in *Arabidopsis thaliana* by *miR156* and its target *SPL3*. Development 133, 3539–3547.

Xie, K., Wu, C., Xiong, L., 2006. Genomic organization, differential expression and interaction of SPL transcription factors and microRNA156 in rice. Plant Physiol. 142, 280–293.

Yamasaki, K., et al., 2004. A novel zinc-binding motif revealed by solution structures of DNA-binding domains of *Arabidopsis* SBP-family transcription factors. J. Mol. Biol. 337, 49–63.

Yamasaki, K., et al., 2006. An *Arabidopsis* SBP-domain fragment with a disrupted C-terminal zinc-binding site retains its tertiary structure. FEBS Lett. 580, 2109–2116.

Zhang, Y., Schwarz, S., Saedler, H., Huijser, P., 2007. SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in *Arabidopsis*. Plant Mol. Biol. 63, 429–439.

Zuker, M., Mathews, D.H., Turner, D.H., 1999. Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide in RNA biochemistry and biotechnology. In: Barciszewski, J., Clark, B.F.C. (Eds.), NATO ASI Series. Kluwer Academic Publishers.

Please cite this article as: Riese, M. et al. Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene (2007), doi:10.1016/j.gene.2007.06.018