

# Molecular characterisation of two paralogous SPO11 homologues in *Arabidopsis thaliana*

Frank Hartung and Holger Puchta\*

Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstraße 3, D-06466 Gatersleben, Germany

Received December 30, 1999; Revised and Accepted February 15, 2000

DDBJ/EMBL/GenBank accession nos AJ251989, AJ251990

## ABSTRACT

The Spo11 protein of yeast has been found to be covalently bound to double-strand breaks in meiosis, demonstrating a unique role of the protein in the formation of these breaks. Homologues of the *SPO11* gene have been found in various eukaryotes, indicating that the machinery involved in meiotic recombination is conserved in eukaryotes. Here we report on *SPO11* homologues in plants. In contrast to what is known from other eukaryotes, *Arabidopsis thaliana* carries in its genome at least two *SPO11* homologues, *AtSPO11-1* and *AtSPO11-2*. Both genes are not more closely related to each other than to other eukaryotic *SPO11* homologues, indicating that they did not arise via a recent duplication event during higher plant evolution. For both genes three different polyadenylation sites were found. *AtSPO11-1* is expressed not only in generative but also to a lesser extent in somatic tissues. We were able to detect in different organs various *AtSPO11-1* cDNAs in which introns were differently spliced—a surprising phenomenon also reported for *SPO11* homologues in mammals. In the case of *AtSPO11-2* we found that the 3' end of the mRNA is overlapping with a mRNA produced by a gene located in inverse orientation next to it. This points to a possible antisense regulation mechanism. Our findings hint to the intriguing possibility that, at least for plants, Spo11-like proteins might have more and possibly other biological functions than originally anticipated for yeast.

## INTRODUCTION

During meiosis genetic diversity is generated through rearrangement of parental chromosomes via homologous recombination (reviewed in 1–3). A central step in the initiation of the recombination reaction in yeast is the induction of chromosomal double-strand breaks (DSBs) (4,5) defining the acceptor sites of gene conversion. A key actor in the process is the factor that initiates the DSBs. Strong evidence was supplied for yeast that this factor is the Spo11 protein. Spo11p was shown to remain covalently bound at both 5'-termini of DSBs in a

rad50S mutant of *Saccharomyces cerevisiae* (6). Spo11 is coded by a single gene in yeast and is homologous to the subunit A of archaeobacterial topoisomerase VI an atypical type II topoisomerase (7).

A question of major interest is whether the molecular processes during the initiation of meiosis are conserved in all eukaryotes. Besides cytological studies a way to address such a question is to find sequence homologues in the genomes of representative members of other parts of the eukaryotic kingdom. For Spo11 a homologue was first identified in *Schizosaccharomyces pombe* named Rec12 (8). Very recently *SPO11* homologous genes were found in a number of different higher eukaryotes such as *Drosophila melanogaster*, *Caenorhabditis elegans*, mouse and man (9–12). The family of Spo11 proteins is grouped together according to their homology with respect to five conserved protein motifs (motifs I–V) of the archaeobacterial subunit A of Topoisomerase VI (7). Part of these motifs is the proposed Toprim (topoisomerase primase) domain which is conserved in type IA and type II topoisomerases and also other proteins involved in DNA replication and repair (13). From all organisms analysed so far the existence of a single *SPO11* homologue was reported. As expected regarding the function of Spo11 in yeast, it was shown for the majority of cases to be expressed mainly or exclusively during meiosis (10,11). However, in *Drosophila* (9), mouse (11) and man (12) a weak expression in somatic tissue was reported.

Information about proteins involved in the initiation of meiotic recombination in plants is still limited (reviewed in 14–16). A homologue of the meiotic strand exchange protein Dmc1 of yeast was identified in different plant species (17–19). Recently, plant homologues of Mre11 and Rad50, both proteins involved in endonucleolytic degradation of Spo11 containing 5' ends of the meiotic DSBs in yeast (20; Charles White, personal communication), could be isolated from the model plant *Arabidopsis thaliana* (21). In the current study we analysed whether there exists an *Arabidopsis* homologue of *SPO11*. To our surprise we found for the first time in any organism not only one but two genes. AtSpo11-1 and AtSpo11-2 both match perfectly well with the Spo11 family in an alignment of motifs I–V. The deduced proteins are identical between 22 and 35% to their counterparts in other eukaryotes and 28% to each other, arguing strongly against a recent duplication event. In the following we describe the detailed

\*To whom correspondence should be addressed. Tel: +49 39482 5181; Fax: +49 39482 5137; Email: puchta@ipk-gatersleben.de

characterisation of both genes, which led us to further unexpected findings.

## MATERIALS AND METHODS

### Accession numbers

The cDNA sequences for *AtSPO11-1* and *AtSPO11-2* have been deposited in the DDBJ/EMBL/GenBank under the accession numbers AJ251989 and AJ251990.

### Computer analysis

BLAST searches were performed within the Stanford genomic database resources of *A.thaliana* using the Spo11 protein sequence of *C.elegans* and TBLASTN (22). Sequence analysis was performed using DNASTAR (Lasergene) and multiple alignments on the Internet using CLUSTALW (v.1.7). Pairwise comparisons were done with the MEGALIGN program (Lipman–Pearson method) of the DNASTAR package (Lasergene) using a Ktuple of 2 and a gap penalty of 4.

### RNA and DNA isolation

*Arabidopsis thaliana* ecotype columbia RNA was isolated using the Plant RNA Midi Kit (peqlab, Erlangen) following the instructions of the manufacturer. The following tissues were used: 2-week-old liquid root culture, mature rosette leaves, flower shoots, complete flowers or 2–4-week-old sterile grown seedlings. The isolated total RNA was used for enrichment of poly(A)<sup>+</sup>-RNA with the mRNA enrichment Kit (peqlab, Erlangen) following the manufacturer's instructions.

Genomic DNA of *A.thaliana* ecotype columbia was isolated from 2–4-week-old sterile grown seedlings with the Plant DNA Midi Kit (peqlab, Erlangen).

### RT–PCR and RACE

Reverse transcription was performed according to the SMART-protocol from Clontech (Heidelberg) using 5–50 ng of mRNA. One-twentieth of the synthesised SMART-cDNA was preamplified via 18 PCR cycles as described in the SMART-protocol. The non-amplified cDNA was used for different RACE and semi-quantitative RT–PCR experiments. Non-quantitative RT–PCR reactions were performed with two gene-specific primers using one-fiftieth of the preamplified cDNA and the Expand-Long-Template Taq-polymerase (Roche Molecular Biochemicals, Mannheim). After 30 or 35 cycles, respectively, the PCR products were examined by gel electrophoresis and EtBr staining.

RACE reactions were done according to an improved SMART-protocol (23) with nested gene-specific primers using cDNA without preamplification. One nested primer was usually sufficient to obtain clean products for direct sequencing and cloning. Gene specific primers (0.25 μM), 0.25 μM of the heel specific (HS) and 0.05 μM of the heel carrier (HC) primer (for primers see 23) were used in the reactions. Both 3'RACE resulted in PCR products of different sizes which were cloned into the pGEM-T easy vector according to the manufacturer's protocol (Promega, Mannheim).

### Quantitation of mRNA levels

Quantitation of the mRNA amount of *AtSPO11-1*, *AtSPO11-2* and the antisense reading frame was done in series of competitive

RT–PCRs (reviewed in 24) using internal standards. The DNA internal standards, each containing at least one intron to discriminate the amplification product size from the respective cDNA, were preamplified from genomic DNA of *A.thaliana*. The exact amount of the standard DNAs was determined by spectroscopy and diluted to various extents for PCR analysis in series of pilot experiments. Each competitive PCR analysis consisted of four different dilution steps of the internal standard alone and of a combination of appropriate amounts of internal standard and the cDNA population of the respective tissues.

### Direct sequencing and cloning

Sequencing of PCR and RACE products was done with the ABI Prism Big Dye™ Terminator Cycle Sequencing Reaction Kit (PE Applied Biosystems, Weiterstadt) either directly or after cloning (3'RACE) into pGEM T-easy vector (Promega, Mannheim). Before sequencing the templates were purified with the High Pure PCR Product Purification Kit from Roche-Molecular Biochemicals (Mannheim).

### Southern analysis

Genomic DNA (5 μg) of *A.thaliana* ecotype columbia was digested with the respective restriction enzymes overnight. After electrophoresis in 0.9% agarose the gel was denatured, neutralised (25) and transferred overnight onto a Hybond-NX nylon membrane (Amersham, Braunschweig) by capillary transfer using 20× SSC as transfer buffer. After blotting, the membrane was cross-linked with 120 mJ/cm<sup>2</sup> in a Stratalinker 1800 (Stratagene, Amsterdam).

Hybridisation of the Southern blots and signal detection was done with the digoxigenin chemiluminescent labelling and detection system from Roche Molecular Biochemicals (Mannheim). The Dig High Prime labelling Kit was used for labelling the probe, Dig Easy Hyb as hybridisation buffer and CPD Star as chemiluminescent substrate.

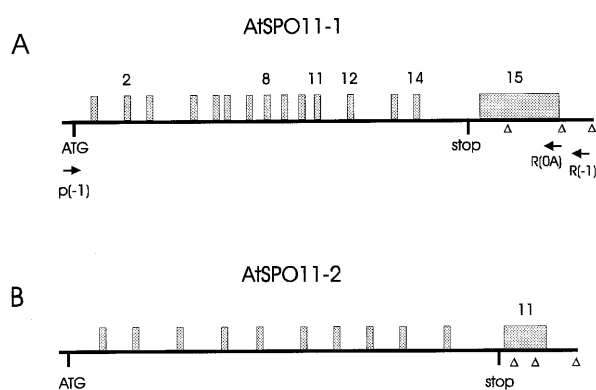
## RESULTS

### Identification of two SPO11 homologues in *A.thaliana*

The rapidly growing sequence database of *A.thaliana* provides a powerful tool to identify putative homologous proteins by database searches with sequence motives of genes of known function from different organisms like yeast, *C.elegans* or mammals. Such database search using TBLASTN with *C.elegans* Spo11 as probe resulted in two significant hits from different chromosomes (accession numbers: AP000375 from chromosome 3 and AC077640 from chromosome 1) showing an identity of >40% in a region of 150 amino acids. The corresponding regions of both BAC sequences were translated and aligned to Spo11 of *C.elegans*, *S.cerevisiae* and *S.pombe*. From this alignment different primers were designed to amplify parts of the genes as cDNA and to perform RACE-experiments.

### cDNA analysis

Using preamplified cDNA from *Arabidopsis* flowers both genes could be amplified via PCR. Overlapping cDNA fragments were sequenced to determine the coding sequence in relation to the genomic BAC clones. To obtain the 5'- and 3'-ends, RACE experiments were performed directly from the RT-reaction as



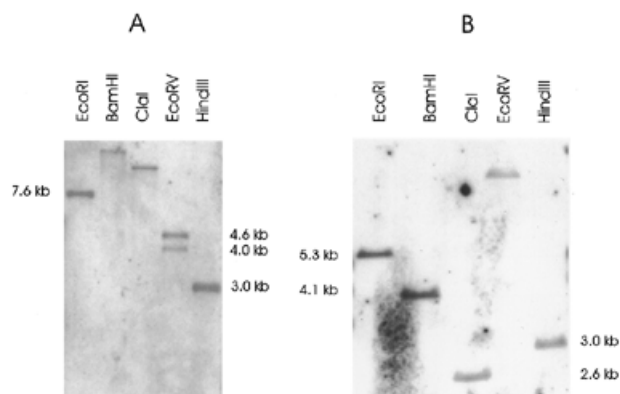
**Figure 1.** Schematic structure of both *A.thaliana* SPO11 homologous genes: Introns are shown as grey boxes above the sequence. Verified polyadenylation sites are represented by open triangles. (A) Primers used for RT-PCR analysis of *AtSPO11-1* are shown as arrows below the schematic sequence.

described in Materials and Methods. The full-length coding sequence of both genes is shown schematically in Figure 1. *AtSPO11-1* consists of 16 exons and 15 introns. The open reading frame (ORF) has a total length of 1089 bp and codes for a putative protein of 362 amino acids (Fig. 1A). The *AtSPO11-2* gene contains 12 exons and 11 introns. The ORF has a total length of 1152 bp and is coding for a deduced protein of 383 amino acids (Fig. 1B). Both full-length cDNA sequences were assembled from the RT-PCR and RACE data in comparison to the genomic DNA. Full-length RT-PCR experiments with preamplified cDNA were also performed, demonstrating the occurrence of the complete mRNAs in the plant (data not shown).

Interestingly, both genes harbour an intron in the 3' untranslated region (UTR) which is not always spliced out. In different RT-PCR and RACE experiments cDNAs containing the intron could be detected. Analysis of the 3'RACE products of *AtSPO11-1* revealed three polyadenylation sites [poly(A) sites]. One is located within the 3'UTR intron and two distal to the intron (Fig. 1A). RT-PCR with primers designed to discriminate between the different poly(A) sites showed a preferential usage of the first poly(A) site within the 3'UTR intron (data not shown). The 3'RACE of *AtSPO11-2* resulted in three distinct bands which after sequencing revealed also three different poly(A) sites (Fig. 1B). Two of them are located in the 3'UTR intron and the third was found to be located 245 nt distal of the 3'UTR intron which was spliced out in the respective cDNA resulting in a deletion of the two other putative poly(A) sites. According to RT-PCR with different specific primer pairs, all three poly(A) sites seemed to be used in approximately equal amounts (data not shown).

#### Both *AtSPO11-1* and *AtSPO11-2* are single copy genes in *Arabidopsis*

Southern analysis of *AtSPO11-1* and *AtSPO11-2* clearly showed that both are represented as single copy genes in *A.thaliana* ecotype columbia (Fig. 2A and B). The banding pattern observed in non-radioactive hybridisation fits perfectly



**Figure 2.** Southern analysis of *A.thaliana* SPO11-1 (A) and SPO11-2 (B). The indicated fragment sizes are calculated from the sequence information of the genomic BAC clones AP000375 for *AtSPO11-1* and AC077640 for *AtSPO11-2*. Both genes show a banding pattern, which is in accordance with a single copy status in the *Arabidopsis* genome.

well with the theoretical calculated one using the sequence information obtained from the respective BAC clones.

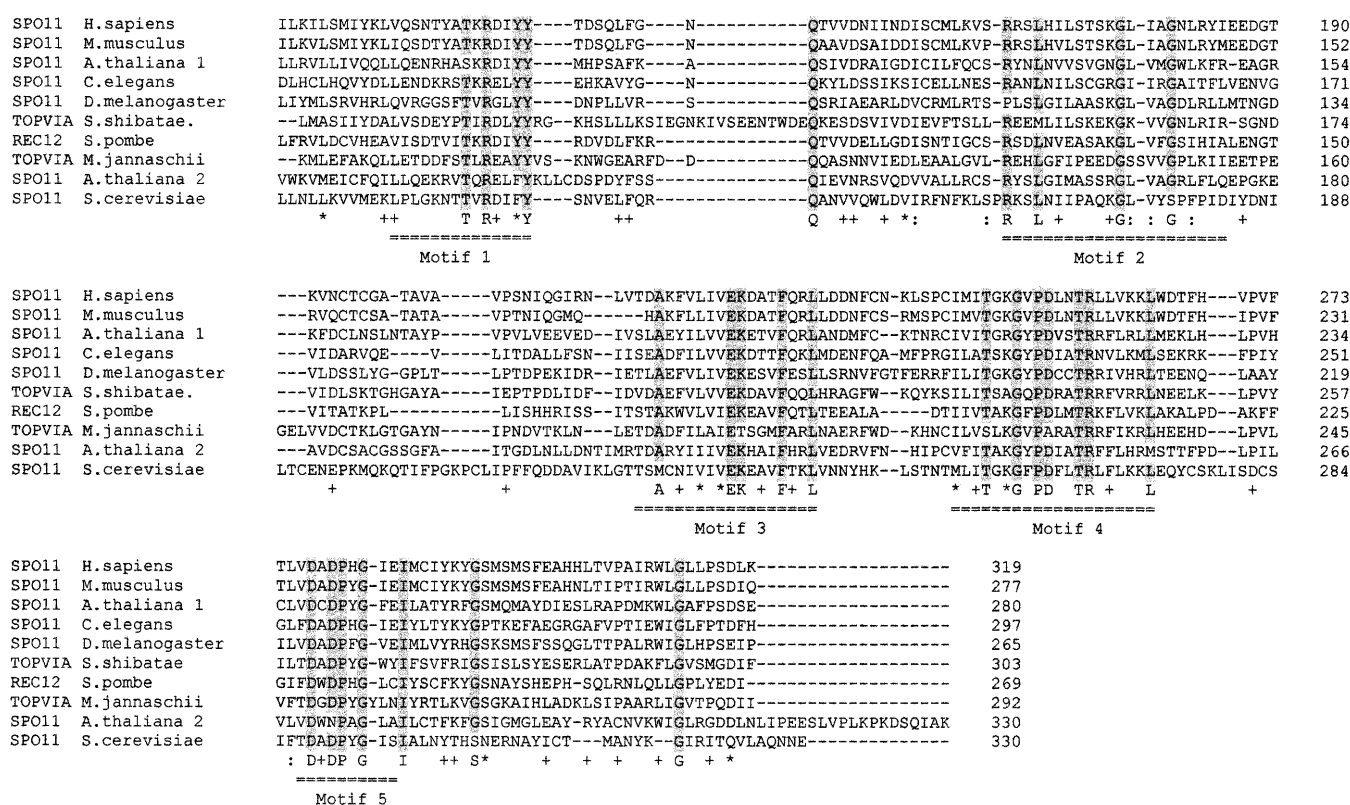
#### Relation of AtSpo11-1 and AtSpo11-2 to other Spo11 homologues

Alignment of the deduced amino acid sequences of *AtSPO11-1* and *AtSPO11-2* with other Spo11 homologues in the region of the most conserved motifs I-V described by Bergerat *et al.* (7) strongly supports our view that both genes of *A.thaliana* are Spo11 homologues (Fig. 3). Both deduced proteins possess all five motifs and showed identities of 26–41% within these motifs. Remarkably, AtSpo11-2 is missing one of the conserved amino acids (Fig. 3, asparagine instead of aspartic acid in position 272) in the so-called Toprim domain (motif V). This domain was found by its conservation in topoisomerases and primases (13) and spans the region of the DNA gyrase motif IV and V. The AtSpo11-2 homologue contains a phenylalanine instead of a second tyrosine within motif I next to the active tyrosine (Fig. 3, position 124 of AtSPO11-2) like the *S.cerevisiae* homologue and in contrast to all other Spo11 homologues (Fig. 3, position 123 of AtSPO11-2).

Pairwise comparison of the complete sequences of AtSpo1-1 and AtSpo11-2 to Spo11 of other organisms resulted in identity values from 21.5 to 34.8% (Table 1). Most remarkably, a direct comparison between AtSpo11-1 and AtSpo11-2 resulted in an identity of 28.3%, indicating that they have not descended from a recent duplication event. These findings are further supported by comparison of the intron positions of both genes with respect to their protein homology alignment. Only four (numbers 6, 7, 10 and 11 of AtSPO11-2) out of 11 possible introns from AtSPO11-2 are at homologous positions.

#### Is *AtSPO11-2* regulated by antisense control?

Database searches with the cDNAs of AtSPO11-1 and AtSPO11-2 using the complete EST database (dbest) and specialised versions for *A.thaliana*, maize and rice gave no significant hits, indicating that the genes are not expressed at high levels in plants. However, surprisingly, three independent ESTs (differing in their length regarding the 5' end) of the



**Figure 3.** Multiple alignment of the DNA gyrase motif I–V from topoisomerase VIA, Spo11 and Rec12. The multiple alignment was done with CLUSTALW (v.1.7; <http://www.ibc.wustl.edu>). Gaps are represented by dashes, amino acids conserved in more than 5 of the 10 sequences are marked by plus sign (6), colon (7) and asterisk (8). Amino acids that are conserved in all of 9 out of 10 positions are shaded in grey and represented by letters. The conserved DNA gyrase motifs I–V are underlined with dashed double lines. The respective amino acid position of each sequence is given on the right.

**Table 1.** Pairwise comparison of both Spo11 proteins of *A.thaliana* to Spo11p from various organisms

	<i>A.thaliana</i> 1	<i>A.thaliana</i> 2
<i>Homo sapiens</i>	34.8	28.2
<i>Mus musculus</i>	33.5	29.1
<i>A.thaliana</i> 1	100	<b>28.3</b>
<i>C.elegans</i>	30.9	26.8
<i>D.melanogaster</i>	32.1	27.1
<i>Sulfolobus shibatae</i>	28.9	23.4
<i>S.pombe</i>	26.3	25.5
<i>Methanococcus jannaschii</i>	27.4	26.6
<i>S.cerevisiae</i>	23.0	21.5

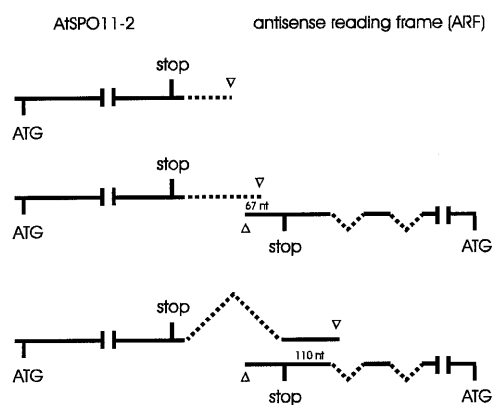
The highest and lowest identities are presented in italics and the identity between AtSpo11-1 and AtSpo11-2 is shown in bold.

*Arabidopsis* database had >98% sequence homology to the 3' end of *AtSPO11-2* in reverse orientation. The reverse sequence had internal gaps in comparison to the *AtSPO11-2* cDNA. An alignment of the EST sequence in comparison to the genomic sequence revealed that the ESTs are from a gene that is expressed in inverse orientation to *AtSPO11-2*. The gene

that we will refer to as antisense reading frame (ARF) has no obvious homology to gene families with known function. There are two introns within the respective range of the ARF resulting in the gaps in the alignment with the 3'UTR of SPO11-2 mRNA. Both genes are overlapping on the mRNA level in different manner, depending on the poly(A) site used during *AtSPO11-2* mRNA processing (Fig. 4). Usage of the first poly(A) site of *AtSPO11-2* is not resulting in an overlap with the ARF mRNA. However, if the second poly(A) site is used, a 67 nt-long overlap occurs and if the third poly(A) site is used the overlapping region is 110 nt long. As our PCR analysis revealed equimolar amounts of mRNAs of Spo11-2 for each of the different poly(A) sites, it seems that the majority of the mRNA produced indeed overlap in antisense with the ARF mRNA.

**Alternatively spliced forms of *AtSPO11-1* are present in different organs**

RT-PCR of *AtSPO11-1* with different primer pairs always resulted in more than one PCR product. We assumed alternative splicing to account for the unexpected gel electrophoresis pattern because *AtSPO11-1* as well as *AtSPO11-2* are single copy genes (see above). To investigate the different splice products we cloned the RT-PCR products from four different tissues using two specific primer pairs [Fig. 1A, primer p(-1) and R(-1) or R(0A)]. The cloning and sequencing of the



**Figure 4.** Possible interactions between *AtSPO11-2* mRNA and the mRNA of an unknown gene in antisense orientation (ARF) next to it. Unspliced introns are shown as dashed lines, spliced introns as curved dashed lines. The polyadenylation sites of both genes are represented by open triangles. Both genes overlap in an antisense manner if polyadenylation sites two and three of *AtSPO11-2* are used.

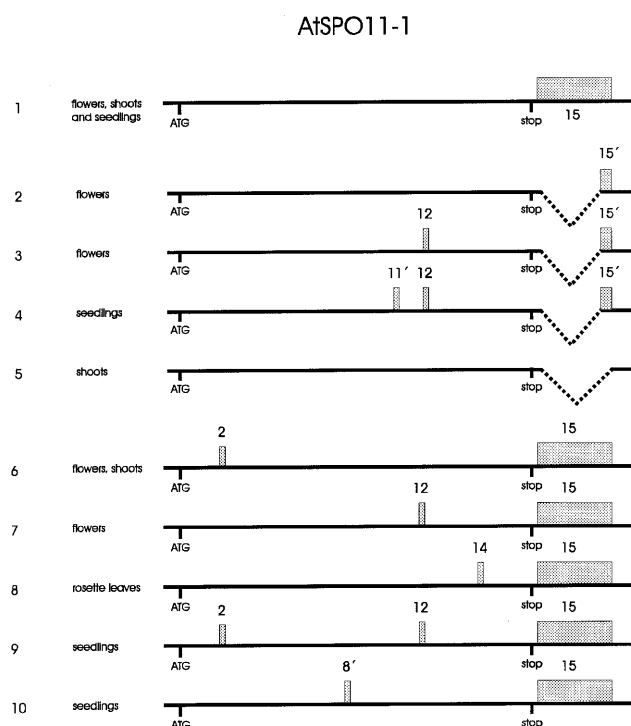
respective cDNAs confirmed their origin from alternative splicing (Fig. 5). Curiously all of the alternatively spliced mRNAs contained premature stop codons.

RT-PCR of the leaves resulted in only one product harbouring intron 14 (Fig. 5, lane 8). The translation of this ORF would result in a truncated protein of 320 instead of 362 amino acids. In shoots three different transcripts were cloned after RT-PCR. Two of them contained the full-length coding sequence but differed with respect to intron 15 which is located behind the stop codon (Fig. 5, lanes 1 and 5). Notably, the splice acceptor site used for processing intron 15 was different to the one we detected in flowers and seedlings (see below). The third product was harbouring intron 2 (Fig. 5, lane 6) which would result only in a 68 amino acid-long protein. The sequencing of the cloned RT-PCR products from flowers showed five different transcripts; two of them identical to the sequences detected in shoots (Fig. 5, lanes 1 and 6). Two additional transcripts were harbouring intron 12 either in the absence or presence of the 3'UTR intron 15 (Fig. 5, lanes 3 and 7). In the third transcript the complete ORF was present without intron 15 (Fig. 5, lane 2). In all cases the splice acceptor site of intron 15 used in flowers differed from the one found in shoots. In seedlings, besides the functional ORF (Fig. 5, lane 1) alternative splicing resulted in three different transcripts harbouring parts of intron 8 (34 nt) or 11 (11 nt) and the full-length introns 12 or 2 and 12 (see Fig. 5, lanes 4, 9 and 10). As the 3'UTR intron 15 seems not to be spliced out in all transcripts, further variations might occur.

In contrast to *AtSPO11-1* we could not detect any splicing variants of *AtSPO11-2* except of intron 11 in the 3'UTR (Fig. 4)

#### Semiquantitative RT-PCR analysis of expression of *AtSPO11-1* and the ARF

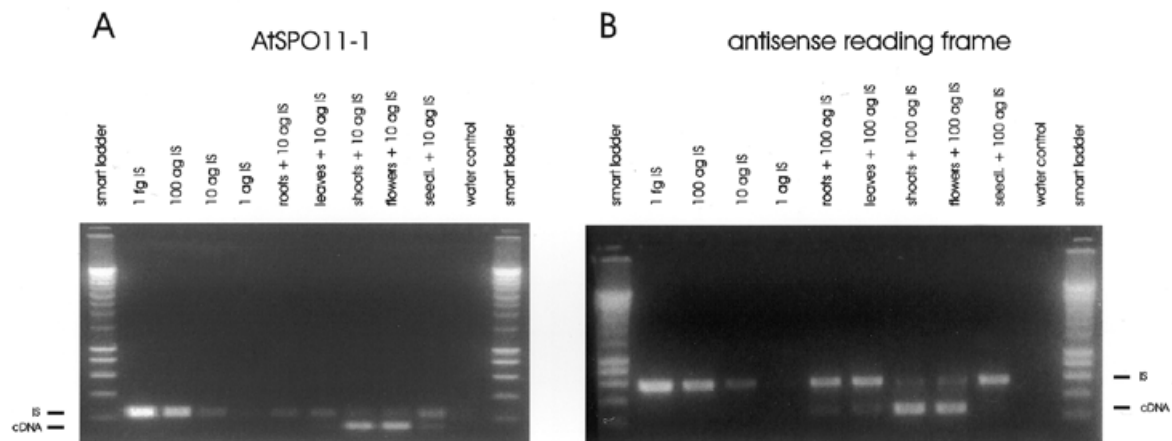
To quantitate the mRNA amount of both *AtSPO11* genes and the ARF gene we performed competitive RT-PCR with internal standards using 5–50 ng of mRNA (for details see



**Figure 5.** Analysis of alternative splicing products of *AtSPO11-1* by RT-PCR and sequencing. All splice products with resting introns would result in truncated proteins due to premature stop codons. Introns are represented by grey boxes and numbered in the same order as in the full gene (Fig. 1). The partially spliced out intron 15 is represented by a curved dashed line. Partially remaining introns are numbered 8', 11' and 15'. The organs from which the cDNAs derived are shown on the left. In number five the intron 15 is slightly longer than in all other lanes because an alternative splice junction was found.

Materials and Methods). In Figure 6, competitive RT-PCRs from *AtSPO11-1* and the antisense EST of *AtSPO11-2* are shown. In both experiments 35 PCR cycles were used. After RT-PCR, 1  $\mu$ l of the undiluted cDNA of *AtSPO11-1* in competition with 10 ag (attogram) internal standard was detectable only in shoots, flowers and as a weak band in young seedlings (Fig. 6A, lanes 7–9). No signal was found in roots and rosette leaves. In comparison to that, the cDNA of ARF was easily detectable in 1:100 dilutions (except of roots where a 1:10 dilution was used) in competition with 100 ag of internal standard. The strongest signals were also found in shoots and flowers (Fig. 6B, lanes 7 and 8). Much weaker signals were found in roots and leaves (lanes 5 and 6) and no signal in seedlings (lane 9). The root cDNA was diluted only 1:10 (or 10  $\mu$ l used in Fig. 6A) because the initial used mRNA amount (5 ng) was 10 times less than in all other cases.

Calculation of the total amount of *AtSPO11* transcripts in 500 ng mRNA of shoots or flowers resulted in a value in the range of 10 fg. The same calculation was done for the ARF resulting in a 1000 times higher value of  $\sim$ 10 pg. For *AtSPO11-2* we found only after 40 PCR cycles in shoots and flowers a signal indicating a mRNA level that was about 10 times lower than *AtSPO11-1* (data not shown).



**Figure 6.** Semiquantitative RT-PCR of *AtSPO11-1* (A) and the ARF gene (B). The origin of the respective samples and the amounts of the internal standard (IS) used are given on top of the lanes. (A) Ten microlitres of root cDNA and 1  $\mu$ l of all other cDNAs were used for RT-PCR together with 10 ag of IS. (B) A 1:10 dilution of root cDNA and 1:100 dilutions of all other cDNAs were used together with 100 ag of IS. (A and B) The lower band in each part represents the cDNA and the upper band the internal standard. The first five bands of the smart ladder are, from bottom to top, 200, 400, 600 and 800 bp and 1 kb.

## DISCUSSION

Spo11 as the key factor to initiate meiotic DSBs is an ancient protein as demonstrated by its homology to the subunit A of archaeobacterial Topoisomerase VI. In eukaryotes the topoisomerase activity seems to have lost its function for sealing DSBs (which resides in subunit B of Topo VI) but is still producing DSBs in a topoisomerase-like manner (6). No homologue of archaeobacterial subunit B was found in eukaryotes so far (7). As the attributed function of Spo11 is to produce meiotic DSBs (6), no sealing step is required before recombination occurs. In yeast Spo11 is covalently bound to the 5' termini of unsealed DSBs (6) and is very probably removed by the Mre11/Rad50/Xrs2 complex (20). Also in yeast Spo11 is essential for DSB initiation and formation of the synaptonemal complex (SC) (6). This seems to be different in *C.elegans* and *D.melanogaster* (9,10) where in SPO11 mutants formation of the SC was unaffected but no meiotic recombination occurred.

In all previously investigated organisms only one SPO11 homologue was detected. Therefore it was surprising to find two homologous genes of SPO11 in *A.thaliana*, which are both expressed on mRNA level. As the *Arabidopsis* genome is not yet completely sequenced, we cannot exclude the possibility that even more homologues occur. The two homologous genes are related to each other in a similar range as the Spo11 proteins of distantly related organisms. There is no clear indication from the sequence comparisons that one of the two genes is equivalent to the Spo11 homologues of other eukaryotes. However, according to the pairwise comparison data AtSpo11-1 seems to be generally closer related to the higher eukaryotic Spo11 homologues than AtSpo11-2. Considering this, and the weak relationship between both genes in intron positions and numbers, they do not seem to have arisen by a recent duplication event during the evolution of higher plants. What is very interesting in this respect is the fact that AtSpo11-2 shows one striking difference to all other Spo11 sequences. The conserved amino acid Asp is exchanged to Asn in position 272 of AtSpo11-2. The Asp 272 together with Asp 270 and Glu 217 is thought to

be responsible for coordination of  $Mg^{2+}$  (26). It is not clear whether the switch to Asn in this  $Mg^{2+}$  coordinating domain will hinder its function.

Not only is the finding of two Spo11 homologues of *Arabidopsis* startling, but the expression of both *AtSPO11-1* and *AtSPO11-2* also has puzzling features. The very complex RT-PCR pattern of *AtSPO11-1* demonstrating at least 10 different splicing products was surprising, as we were not able to detect such patterns with *AtSPO11-2* or any other gene. At the moment it is not clear whether these RNAs are products of an incomplete splicing reaction or real alternative transcripts. However, we used the described experimental setup (SMART and step out PCR according to 23) to isolate cDNAs of a dozen of different genes involved in recombination and repair processes of *A.thaliana* without ever observing such a pattern of RT-PCR (21; F.Hartung and H.Puchta, unpublished observations). We therefore conclude that the occurrence of different splice variants is a unique feature of *AtSPO11-1*. As alternatively spliced cDNAs are also present in somatic and generative tissues, the biological function of this phenomenon is not clear at present. Notably there are indications that a similar phenomenon also occurs in mouse and humans. The authors found cDNAs harbouring introns not only in RACE experiments but also in a mouse testis-specific cDNA-library (6,12; H.Offenberg, personal communication). One intriguing possibility would be that *AtSPO11-1* expression is regulated via a 'nonsense mediated decay' (NMD) pathway (reviewed in 27). The NMD is known to be responsible for a rapid decay of the respective mRNAs due to premature stop codons. Therefore, only if the *AtSPO11-1* protein is required in a specific biological situation would modified processing lead to proper spliced transcripts of *AtSPO11-1*, thereby avoiding the NMD of its mRNA. This could lead to a very fast response resulting in the induction of DSBs. However, from the data presented here we cannot conclude whether different cDNAs of *AtSPO11-1* occur in the same cell or in which parts of different organs the completely processed cDNA of *AtSPO11-1* arise. It is possible that the

complex splicing pattern resembles a mixture of different organs or even different cells within these organs.

The different poly(A) sites of the mRNAs of both SPO11 genes are a further demonstration that in plants polyadenylation is not very strictly connected to a specific poly(A) signal, as seen in animals. In other investigations 14–26 different poly(A) sites were realised within one gene (28; A.Ipsen, personal communication) and great variations in poly(A)-site usage seem to be common for plants (reviewed in 29). The fact that we found by semiquantitative RT-PCR expression of *AtSPO11-1* in different somatic tissues might hint to a putative function in somatic cells. Along with this speculation goes the recent finding that the mutation of the *D.melanogaster* Spo11 homologue, mei-W68, generates a mild somatic hyper-recombination phenotype (9).

To our great surprise we found by EST database search an overlapping antisense transcript to the 3' end of *AtSPO11-2*. As this overlap lies within a region where splicing and/or (differential) processing of *AtSPO11-2* occurs, it is tempting to speculate about a putative regulatory function of this phenomenon. Remarkably, the transcription of ARF is >1000-fold of that of *AtSPO11-2*. The overlapping regions between *AtSPO11-2* and the ARF are long enough to mediate a silencing effect as described for *C.elegans* or *Dictyostelium* (reviewed in 30). Recently the occurrence of antisense RNAs was correlated with post-transcriptional gene silencing in plants (31).

We are aware of the fact that this current report raises more questions than it solves. The conservation of SPO11 homologues between plants and other eukaryotes indicates a conservation of the enzyme machinery for meiotic recombination. However, this might not be the complete story. Why there are two SPO11 homologues in plants? Why there is expression in somatic tissue? Why there are so many different spliced forms of *AtSPO11-1*? Is there a kind of antisense control for *AtSPO11-2* expression? In the end, only genetic approaches in which the functions of the respective gene are eliminated will provide answers to these questions.

## ACKNOWLEDGEMENTS

We like to thank I. Schubert for critical comments on the manuscript and Anja Ipsen, Waltraud Schmidt-Puchta, Andrea Kunze and Hildo Offenbergh for help and discussions.

## REFERENCES

- Zickler,D. and Kleckner,N. (1998) *Annu. Rev. Genet.*, **32**, 619–697.
- Paques,F. and Haber,J.E. (1999) *Microbiol. Mol. Biol. Rev.*, **63**, 349–404.
- Roeder,G.S. (1997) *Genes Dev.*, **11**, 2600–2621.
- Cao,L., Alani,E. and Kleckner,N. (1990) *Cell*, **61**, 1089–1101.
- Sun,H., Treco,D., Schultes,N.P. and Szostak,J.W. (1989) *Nature*, **338**, 87–90.
- Keeney,S., Giroux,C.N. and Kleckner,N. (1997) *Cell*, **88**, 375–384.
- Bergerat,A., de Massy,B., Gabelle,D., Varoutas,P.C., Nicolas,A. and Forterre,P. (1997) *Nature*, **386**, 414–417.
- Lin,Y. and Smith,G.R. (1994) *Genetics*, **136**, 769–779.
- McKim,K.S. and Hayashi-Hagihara,A. (1998) *Genes Dev.*, **12**, 2932–2942.
- Dernburg,A.F., McDonald,K., Moulder,G., Barstead,R., Dresser,M. and Villeneuve,A.M. (1998) *Cell*, **94**, 387–398.
- Keeney,S., Baudat,F., Angeles,M., Zhou,Z.H., Copeland,N.G., Jenkins,N.A., Manova,K. and Jasin,M. (1999) *Genomics*, **61**, 170–182.
- Romanienko,P.J. and Camerini-Otero,R.D. (1999) *Genomics*, **61**, 156–169.
- Aravind,L., Leipe,D.D. and Koonin,E.V. (1998) *Nucleic Acids Res.*, **26**, 4205–4213.
- Puchta,H. and Hohn,B. (1996) *Trends Plant Sci.*, **1**, 340–348.
- Schnable,P.S., Hsia,A.-P. and Nikolau,B.J. (1998) *Curr. Opin. Plant Biol.*, **1**, 123–129.
- Vergunst,A.C. and Hooykaas,P.J.J. (1999) *Crit. Rev. Plant Sci.*, **18**, 1–31.
- Sato,S., Kobayashi,T., Hotta,Y. and Tabata,S. (1995) *DNA Res.*, **2**, 147–150.
- Klimyuk,V.I. and Jones,J.D. (1997) *Plant J.*, **11**, 1–14.
- Doutriaux,M.P., Couteau,F., Bergounioux,C. and White,C. (1998) *Mol. Gen. Genet.*, **257**, 283–291.
- Haber,J.E. (1998) *Cell*, **95**, 583–586.
- Hartung,F. and Puchta,H. (1999) *Plant Physiol.*, **121**, 312.
- Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) *J. Mol. Biol.*, **215**, 403–410.
- Matz,M., Shagin,D., Bogdanova,E., Britanova,O., Lukyanov,S., Diatchenko,L. and Chenchik,A. (1999) *Nucleic Acids Res.*, **27**, 1558–1560.
- Koehler,T., Lassner,D., Rost,A.K., Thamm,B., Pustowoit,B. and Remke,H. (1995) *Quantitation of mRNA by Polymerase Chain Reaction*. Springer Labor Manual, Springer, Berlin.
- Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Nichols,M.D., DeAngelis,K., Keck,J.L. and Berger,J.M. (1999) *EMBO J.*, **18**, 6177–6188.
- Gutierrez,R.A., MacIntosh,G.C. and Green,P.J. (1999) *Trends Plant Sci.*, **4**, 429–438.
- Klahre,U., Hemmings-Mieszczak,M. and Filipowicz,W. (1995) *Plant Mol. Biol.*, **28**, 569–574.
- Rothnie,H.M. (1996) *Plant Mol. Biol.*, **32**, 43–61.
- Oberstrass,J. and Nellen,W. (1997) In Weiss B. (ed.), *Antisense Oligodeoxynucleotides and Antisense RNA*. CRC Press, Boca Raton, FL, pp. 171–195.
- Hamilton,A.J. and Baulcombe,D.C. (1999) *Science*, **286**, 950–952.