Thorben Sprink
Institut für die Sicherheit biotechnologischer Verfahren bei Pflanzen

SPO11 dependent initiation of meiotic double strand breaks in *Arabidopsis thaliana*

Dissertationen aus dem Julius Kühn-Institut

Julius Kühn-Institut
Bundesforschungsinstitut für Kulturpflanzen

Kontakt/Contact:
Thorben Sprink
Julius Kühn-Institut
Federal Research Centre for Cultivated Plants
Institute for Biosafety in Plant Biotechnology
Erwin-Baur-Str. 27
D-06484-Quedlinburg
Germany

Die Schriftenreihe „Dissertationen aus dem Julius Kühn-Institut" veröffentlicht Doktorarbeiten, die in enger Zusammenarbeit mit Universitäten an Instituten des Julius Kühn-Instituts entstanden sind

The publication series „Dissertationen aus dem Julius Kühn-Institut" publishes doctoral dissertations originating from research doctorates completed at the Julius Kühn-Institut (JKI) either in close collaboration with universities or as an outstanding independent work in the JKI research fields.

Der Vertrieb dieser Monographien erfolgt über den Buchhandel (Nachweis im Verzeichnis lieferbarer Bücher - VLB) und OPEN ACCESS im Internetangebot www.jki.bund.de Bereich Veröffentlichungen.

The monographs are distributed through the book trade (listed in German Books in Print - VLB) and OPEN ACCESS through the JKI website www.jki.bund.de (see Publications)

Wir unterstützen den offenen Zugang zu wissenschaftlichem Wissen.
Die Dissertationen aus dem Julius Kühn-Institut erscheinen daher OPEN ACCESS.
Alle Ausgaben stehen kostenfrei im Internet zur Verfügung:
http://www.jki.bund.de Bereich Veröffentlichungen

We advocate open access to scientific knowledge. Dissertations from the Julius Kühn-Institut are therefore published open access. All issues are available free of charge under http://www.jki.bund.de (see Publications).

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation In der Deutschen Nationalbibliografie: detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek (German National Library)
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

ISBN 978-3-95547-023-4
DOI 10.5073/dissjki.2015.006

Herausgeber / Editor
Julius Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, Quedlinburg, Deutschland
Julius Kühn-Institut, Federal Research Centre for Cultivated Plants, Quedlinburg, Germany

Thorben Sprink
Institut für die Sicherheit biotechnologischer Verfahren bei Pflanzen

SPO11 dependent initiation of meiotic double strand breaks in *Arabidopsis thaliana*

Kontakt/Contact:
Thorben Sprink
Julius Kühn-Institut
Federal Research Centre for Cultivated Plants
Institute for Biosafety in Plant Biotechnology
Erwin-Baur-Str. 27
D-06484-Quedlinburg
Germany

Die Schriftenreihe „Dissertationen aus dem Julius Kühn-Institut" veröffentlicht Doktorarbeiten, die in enger Zusammenarbeit mit Universitäten an Instituten des Julius Kühn-Instituts entstanden sind

The publication series „Dissertationen aus dem Julius Kühn-Institut" publishes doctoral dissertations originating from research doctorates completed at the Julius Kühn-Institut (JKI) either in close collaboration with universities or as an outstanding independent work in the JKI research fields.

Der Vertrieb dieser Monographien erfolgt über den Buchhandel (Nachweis im Verzeichnis lieferbarer Bücher - VLB) und OPEN ACCESS im Internetangebot www.jki.bund.de Bereich Veröffentlichungen.

The monographs are distributed through the book trade (listed in German Books in Print - VLB) and OPEN ACCESS through the JKI website www.jki.bund.de (see Publications)

Wir unterstützen den offenen Zugang zu wissenschaftlichem Wissen.
Die Dissertationen aus dem Julius Kühn-Institut erscheinen daher OPEN ACCESS.
Alle Ausgaben stehen kostenfrei im Internet zur Verfügung:
http://www.jki.bund.de Bereich Veröffentlichungen

We advocate open access to scientific knowledge. Dissertations from the Julius Kühn-Institut are therefore published open access. All issues are available free of charge under http://www.jki.bund.de (see Publications).

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation In der Deutschen Nationalbibliografie: detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek (German National Library)
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

ISBN 978-3-95547-023-4
DOI 10.5073/dissjki.2015.006

Herausgeber / Editor
Julius Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, Quedlinburg, Deutschland
Julius Kühn-Institut, Federal Research Centre for Cultivated Plants, Quedlinburg, Germany

SPO11 dependent initiation of meiotic double strand breaks in *Arabidopsis thaliana*

Von der Naturwissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

Thorben Sprink, M.Sc.

geboren am 26.06.1988 in Salzkotten

Referent: Prof. Dr. rer. nat. Thomas Debener

Korreferent: Prof. Dr. rer. nat. Holger Puchta

Tag der Promotion: 29.05.2015

This cumulative thesis comprises the following publications:

Sprink, T., and Hartung, F. (2014). The splicing fate of plant SPO11 genes. Frontiers in Plant Science **5**.

Sprink, T., Metje, J., and Hartung, F. (2015). Plant genome editing by novel tools: TALEN and other sequence specific nucleases. Current Opinion in Biotechnology **32,** 47-53.

Sprink, T., and Hartung, F. (2015). The meiotic functions of SPO11-1 and -2 in *Arabidopsis thaliana* are sequence and to a certain extent species specific.

"Science moves with the spirit of an adventure characterized both by youthful arrogance and by the belief that the truth, once found would be simple as well as pretty."

James D. Watson

Table of contents

Abbreviations ..... - 3 -
Table of Figures ..... - 4 -
Abstract: ..... - 5 -
Zusammenfassung: ..... - 6 -
Schlagwörter: ..... - 7 -
Keywords: ..... - 7 -
Introduction: ..... - 8 -
Double strand break induction and repair during meiosis ..... - 9 -
The synaptonemal complex ..... - 14 -
The evolution and function of SPO11 ..... - 14 -
Marking the break, proteins suitable for marking of double strand breaks ..... - 15 -
Aims of this work: ..... - 18 -
Results and discussion: ..... - 20 -
The function of SPO11-1 and SPO11-2 is sequence and to a certain extent species specific ..... - 20 -
The splicing landscape of *SPO11* is diverse and shows species and sequence specific features ..... - 21 -
SPO11-2 is located on the chromosome during leptotene and early zygotene ..... - 24 -
Co-localization of SPO11-1 and SPO11-2 ..... - 27 -
Western blot analysis of SPO11-2 ..... - 29 -
Production of T-DNA free, SPO11-mutation lines by sequence specific nucleases ..... - 31 -
Introducing mutations with Transcription Activator Like Effector Nucleases (TALENs) ..... - 32 -
Mutation induction using the clustered regularly interspaced short palindromic repeats/Cas9 system ..... - 34 -
Material and Methods: ..... - 40 -
Antibody production ..... - 40 -
Western blot analysis ..... - 40 -
Immunolocalization studies ..... - 41 -
Construction and evaluation of the TALEN constructs ..... - 42 -
Construction and evaluation of the CRISPR/Cas9 constructs ..... - 42 -
References: ..... - 43 -
Acknowledgement: ..... - 54 -
Appendix: ..... - 55 -
Sprink, T., and Hartung, F. (2014). The splicing fate of plant SPO11 genes. .....

Sprink, T., Metje, J., and Hartung, F. (2015). Plant genome editing by novel tools: TALEN and other sequence specific nucleases. .....

Sprink, T., and Hartung, F. (2015). The meiotic functions of SPO11-1 and -2 in *Arabidopsis thaliana* are sequence and to a certain extent species specific. .....

-

Abbreviations

aa:	amino acid
Ath:	*Arabidopsis thaliana*
bp:	base pair
Bra:	*Brassica rapa*
CAP:	cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins
cDNA:	complementary deoxyribonucleic acid
Cpa:	*Carica papaya*
CRISPR:	clustered regularly interspaced short palindromic repeats
D-loop:	displacement loop; hetroduplex joint
DAPI:	4´,6-diamidino-2-phenylindole
DHJ:	double Holliday junction
DNA:	deoxyribonucleic acid
DSB:	double strand break
dsDNA:	double stranded deoxyribonucleic acid
h:	hour
HRM:	high resolution melting
IM:	intramuscular
IV:	intravenous
kDA:	kilodalton
mA:	milliampere
MALDI TOF:	matrix-assisted laser desorption/ionization, time of flight
min:	minute
mya:	million years ago
PAM:	protospacer adjacent motif
PCR:	polymerase chain reaction
RNA:	ribonucleic acid
s:	second
SC:	synaptonemal complex
SDSA:	synthesis dependent strand annealing
sgRNA:	single guide ribonucleic acid
ssDNA:	single stranded deoxyribonucleic acid
SSNs:	sequence specific nucleases
TALEN:	transcription activator-like effector nuclease
PBST:	phosphate buffered saline with Tween 20
TCA:	trichloroacetic acid
T-DNA:	transfer deoxyribonucleic acid
TOPRIM:	topoisomerase-primase
TOPVIA:	topoisomerase VI subunit A
TOPVIB:	topoisomerase VI subunit B
trcRNA:	trans activating clustered regularly interspaced short palindromic repeats ribonucleic acid
µg:	microgram
V:	volt
2D:	two dimensional

Table of Figures

FIGURE 1. MEIOTIC STAGES IN WILD TYPE *A. THALIANA* POLLEN MOTHER CELLS. ..... - 9 -

FIGURE 2. DOUBLE STRAND BREAK INDUCTION AND REPAIR MODEL DURING MEIOSIS. ..... - 13 -

FIGURE 3. CO-IMMUNOLOCALIZATION STUDY OF MEIOSIS SPECIFIC PROTEINS IN MEIOCYTES OF WILD TYPE *A. THALIANA*. ..... - 16 -

FIGURE 4. TIMELY DISTRIBUTION OF DOUBLE STRAND BREAK ASSOCIATED PROTEINS DURING PROPHASE I IN *A. THALIANA*. ..... - 17 -

FIGURE 5. THE SPLICING LANDSCAPE OF CHIMERIC *SPO11* GENES. ..... - 23 -

FIGURE 6. PREDICTED SECONDARY STRUCTURE OF *A. THALIANA* SPO11-1 AND SPO11-2. ..... - 25 -

FIGURE 7. CO-IMMUNOLOCALIZATION STUDY OF ASY1 AND SPO11-2 DURING DIFFERENT MEIOTIC STAGES. ..... - 27 -

FIGURE 8. CO-IMMUNOLOCALIZATION OF SPO11-1 AND SPO11-2. ..... - 28 -

FIGURE 9. WESTERN BLOT ANALYSIS OF PROTEIN SAMPLES FROM *A. THALIANA* FLOWERS USING SPO11-2 ANTIBODY. ..... - 29 -

FIGURE 10. POSITIONAL DISTRIBUTION OF SPO11 SPECIFIC ANTIBODIES. ..... - 31 -

FIGURE 11. MUTATION INDUCTION APPROACH IN *A. THALIANA SPO11-1* AND *-2* USING TRANSCRIPTION ACTIVATOR LIKE EFFECTOR NUCLEASES (TALENS). ..... - 33 -

FIGURE 12. MUTATION INDUCTION APPROACH IN *A. THALIANA* SPO11 USING CLUSTERED REGULARLY INTERSPACED SHORT PALINDROMIC REPEATS/CAS9 (CRISPR/CAS9). ..... - 35 -

FIGURE 13. MUTATION ANALYSES BY HIGH RESOLUTION MELTING AND SEQUENCING IN SPO11-CRISPR APPROACHES. ..... - 37 -

FIGURE 14. A SUBSET OF SANGER SEQUENCING RESULTS FROM THE SPO11-3 CRISPR/CAS9 MUTATION APPROACH. ..... - 38 -

Abstract:

The exchange and reorganization of genetic material between two individuals is an essential function of meiosis. Even though in the last years major findings in the field of meiosis have been achieved, especially in plants some key questions remain concealed. For a better understanding of plant meiosis it is essential to decipher why plants need the activity of two very different meiotic SPO11 proteins. This stands in contrast to metazoa and fungi where a single SPO11 is present and sufficient for a proper meiosis. In *Arabidopsis thaliana* two SPO11, referred to as Ath SPO11-1 and Ath SPO11-2, are necessary in a functional form for the induction of double strand breaks (DSBs) during prophase I of meiosis. In nearly all eukaryotic organisms DSBs ensure on one hand the overall genome stability by enabling correct pairing and distribution of the chromosomes and on the other hand genetic variability by permitting recombination. Without DSBs no physical connection can occur between homologous chromosomes and recombination, pairing, and crossing over are excluded. The absence of these processes leads to a random distribution of the chromosomes during meiosis and to almost complete sterility. Therefore, the evolution, the specific functions, and possible interaction of both different meiotic SPO11 proteins in plants have been analyzed in this thesis. For this purpose database searches were performed and homologs of SPO11 in nearly all kingdoms of life were identified and analyzed. To determine the specific functions of both SPO11 paralogs in *A. thaliana* three non-conserved parts and the well conserved last exon between both meiotic SPO11 were exchanged. Furthermore, the later one was completely deleted, creating chimeric *SPO11* genes. Analyzing these exchanges a sequence specific function for both SPO11 proteins was shown. By exchanging Ath *SPO11-1* and *-2* with related and more ancestral *SPO11*, additionally, species specificity for the respective SPO11 was exposed. Complementation was only possible using SPO11 from the close related (20 mya) species *Brassica rapa.* Partial complementation was achieved when *SPO11-1* cDNA from papaya (*Carica papaya*) was integrated multiple times. By performing these analyses a vast pattern of aberrant spliced forms for *SPO11-1* and *-2* in various species were identified. By examination of the splicing landscape of orthologs from both *SPO11* in its respective source plant and after transformation into *A. thaliana*, furthermore, a species and sequence specificity for the splicing process of *SPO11* have been revealed. Due to analyses performed on the splicing landscape of the chimeric *SPO11* genes a sequence specific distribution of splicing events was identified. These findings indicate that the function and the conserved splicing mechanism of each meiotic active SPO11 paralog is sequence specific and that function of the respective orthologs are species specific. To investigate a possible interaction of both SPO11 during meiosis an antibody against SPO11-2 was produced and a co-immunolocalization study for SPO11-1 and -2 was performed. First findings illustrate that both SPO11 paralogs seem to colocalize during early prophase I. However, SPO11-2 seems to stay longer on the chromatin.

Zusammenfassung:

Der Austausch und die Reorganisation von genetischem Material zwischen zwei Individuen ist eine essentielle Funktion der Meiose. Selbst wenn in den letzten Jahren bedeutsame Erkenntnisse über die Meiose, insbesondere bei Pflanzen, erzielt werden konnten, bleiben einige Schlüsselfragen ungeklärt. Für ein besseres Verständnis der Meiose ist es wichtig, zu entschlüsseln, warum Pflanzen zwei sehr verschiedene meiotisch aktive SPO11 Proteine benötigen. Dies steht im Gegensatz zu vielzelligen Tieren und Pilzen, die nur ein SPO11 besitzen und bei denen dieses für die korrekte Meiose ausreichend ist. In Ackerschmalwand (*Arabidopsis thaliana*) werden zwei SPO11 Proteine benötigt, Ath SPO11-1 und Ath SPO11-2 genannt, um während der Prophase I Doppelstrangbrüche (DSB) zu induzieren. In nahezu allen Eukaryoten sichern DSB auf der einen Seite die Genomstabilität durch eine korrekte Paarung und Verteilung der Chromosomen und auf der anderen Seite die genetische Vielfalt, da sie die Rekombination erst ermöglichen. Ohne DSB gibt es keine Verbindung zwischen homologen Chromosomen und Rekombination, Paarung und chromosomale Überkreuzungen sind daher ausgeschlossen. Dies führt zu einer zufälligen Chromosomenverteilung sowie beinahe vollständiger Sterilität. Deswegen wurden in dieser Arbeit die Evolution, die speziellen Funktionen und ein mögliches Zusammenspiel der beiden verschiedenen SPO11 Proteine in Pflanzen untersucht. Hierfür wurde eine Datenbanksuche durchgeführt und es konnten SPO11 Homologe in nahezu allen biologischen Reichen identifiziert und analysiert werden. Um die speziellen Funktionen der beiden SPO11 Paraloge in *A. thaliana* zu untersuchen, wurden drei nicht konservierte Bereiche sowie das konservierte letzte Exon zwischen beiden SPO11 ausgetauscht. Zusätzlich wurden beide um das letzte Exon verkürzt. Durch diese Änderungen konnte eine Sequenz spezifische Funktion beider SPO11 Proteine aufgedeckt werden. Durch den Austausch von *SPO11-1* und *-2* aus *A. thaliana* mit verwandten und mehr urtümlichen *SPO11* Orthologen konnte zusätzlich eine Speziesspezifität für das jeweilige SPO11 gezeigt werden. Eine Komplementation war nur mit dem nahe verwandten SPO11 aus *Brassica rapa* möglich. Eine Teilkomplementation konnte erreicht werden, indem *SPO11-1* cDNA aus Papaya (*Carica papaya*) mehrfach integriert wurde. Während der Analysen wurde eine erhebliche Anzahl aberranter Formen von gespleißten *SPO11-1* und *-2* in verschiedenen Pflanzen identifiziert. Die Analyse des Spleißens von *SPO11,* sowohl in der Ursprungspflanze als auch nach der Transformation in *A. thaliana* konnte eine Spezies und Sequenzspezifität nachweisen. Zusätzlich zeigte das Spleißmuster der chimären *SPO11* eine Sequenzabhängigkeit der einzelnen Spleißeffekte. Zusammengefasst sind die Funktionen und das konservierte Spleißen von meiotischen SPO11 sequenzspezifisch und die jeweiligen SPO11 Orthologe sind speziesspezifisch. Um ein mögliches Zusammenspiel der beiden SPO11 während der Meiose zu zeigen wurde ein Antikörper gegen SPO11-2 entwickelt. Erste Ergebnisse zeigen eine Kolokalisation beider SPO11 während der frühen Prophase. SPO11-2 scheint jedoch länger an den Chromatiden zu verweilen.

Schlagwörter:

Meiose, SPO11, heterologe Gene

Keywords:

Meosis, SPO11, heterologous expression

Introduction:

The exchange and reorganization of genetic material is an essential player of evolution. In contrast to bacteria most multicellular organisms like plants and metazoa have developed a specialized pathway for the exchange and recombination of genetic material between two individuals called sexual reproduction. This pathway involves the formation of a specialized generative cell type, the so called gametes which are formed by a specific cell division process named meiosis. Meiosis is divided into two division steps. The first division, meiosis I, possesses unique features as the homologous chromosomes pair, recombine, and get separated from each other. The second division, meiosis II is comparable to a mitotic cell division, in which a separation of the sister chromatids takes place (Figure 1). Since there is no DNA replication in between the division cycles, every new forming cell contains only half of the DNA content of its origin (for review see Edlinger and Schlögelhofer, 2011). In a given diploid organism, this process is leading to four genetically unique haploid generative cells. Meiosis is the crucial step that ensures genetic variability in the developing gametes on the one hand and overall genome stability on the other (Roeder, 1997; Zickler and Kleckner, 1998; Paques and Haber, 1999; Knoll and Puchta, 2011, Knoll *et al.,* 2014). A key step during meiosis I is the elongated prophase in which pairing and recombination of the parental chromosomes take place. Prophase I is further subdivided into five stages; leptotene, zygotene, pachytene, diplotene, and diakenisis (Ross *et al.,* 1996; Figure 1). In *A. thaliana* the whole process of meiosis lasts around 33 h (Armstrong *et al.,* 2003). Most of this time the cells are in the interphase (G2) and prophase I (Armstrong *et al.,* 2003; Figure 1).

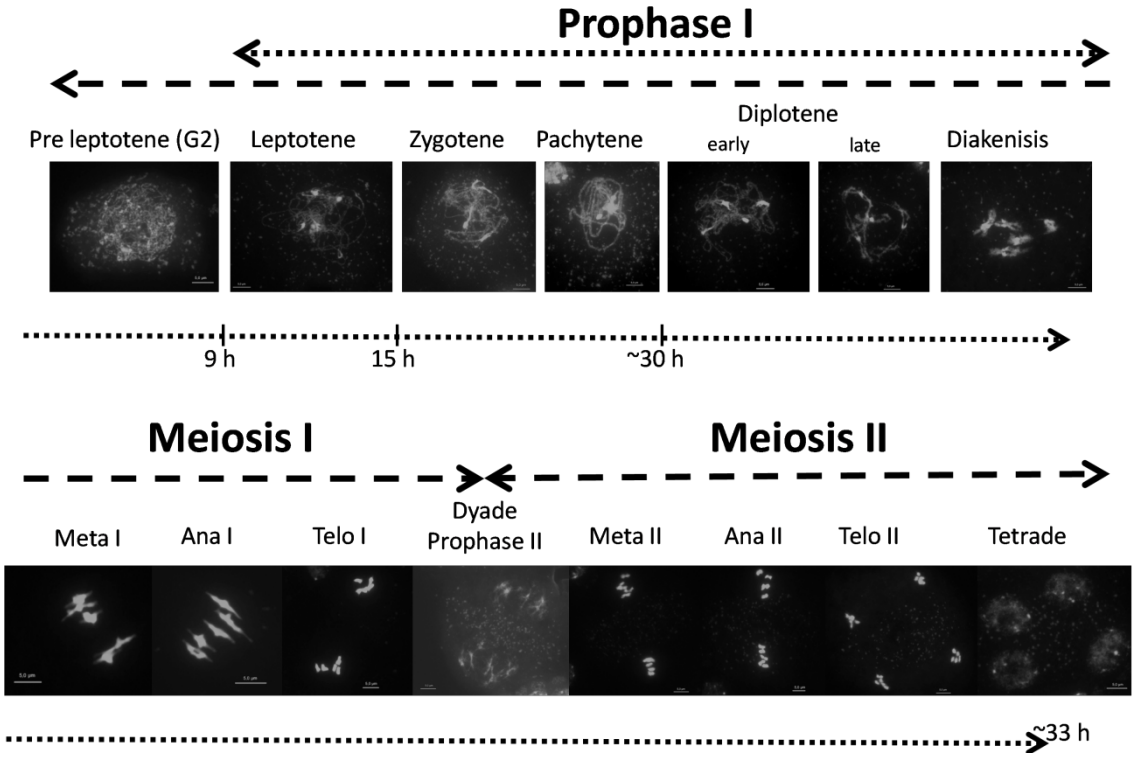


Figure 1. Meiotic stages in wild type *A. thaliana* pollen mother cells.

Fluorescence micrograph of chromosomes from spread wild type *A. thaliana* pollen mother cells in different stages of meiosis I and meiosis II (as indicated) stained with DAPI (4´,6-diamidino-2-phenylindole) diagramed with an estimated timeframe of *A. thaliana* meiosis according to Armstrong *et al.,* 2003. The substages of Prophase I are shown in detail. Meta = Metaphase; Ana = Anaphase, Telo = Telophase. Bar = 5 µm.

Double strand break induction and repair during meiosis

Precise interaction of a meiosis specific protein machinery during prophase I ensures correct pairing, including physical connection via double Holliday Junctions (DHJs), between the homologous chromosomes. Due to this, the stability of the overall chromosome number is guaranteed (reviewed in de Massy, 2013; Jasin and Rothstein, 2013; Keeney *et al.,* 2014). The subsequent resolution of DHJs by the pathway of homologous recombination repair is providing genetic variability in the forming gametes. In most eukaryotic organisms pairing can only be achieved if double strand breaks (DSBs) are initiated by SPO11 during leptotene (Keeney *et al.,* 1997; 1999). In some organisms such as *Drosophila melanogaster* or *Caenorhabditis elegans* the homologous chromosomes are attached by a distinct mechanism that is not involving recombination at all. These organisms also show a lower overall DSB

frequency (Dernburg *et al.,* 1998; McKim *et al.,* 1998; Liu *et al.,* 2002; Colaiacovo *et al.,* 2003; Keeney *et al.,* 2014). However, in most eukaryotes including plants, SPO11 induced breaks and their subsequent repair ensure the physical linkage between the homologous chromosomes. Without these breaks synapsis is inhibited and random distribution of the chromosomes takes place leading to almost complete sterility (Grelon *et al.,* 2001; Stacey *et al.,* 2006; Hartung *et al.,* 2007a).

The eukaryotic SPO11 protein shows homology to the Topoisomerase VI subunit A (TOPVIA) from archaea (Bergerat *et al.,* 1997; Keeney *et al.,* 1997; Grelon *et al.,* 2001; Malik *et al.,* 2007). SPO11 and TOPVIA share seven conserved motifs, a CAP (cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins) domain including a winged helix domain, and a TOPRIM (topoisomerase-primase) domain. Both proteins are able to cleave double stranded DNA constituting a 5´-phosphotyrosyl linkage (Bergerat *et al.,* 1997; Keeney *et al.,* 1997; Malik *et al.,* 2007). This and the following steps of DSB repair are best analyzed in yeast and mouse, where it was shown that SPO11 proteins are released from the break site covalently bound to a short piece of DNA, the so called SPOligo (Neale *et al.,* 2005; Keeney and Neale, 2006; Milman *et al.,* 2009; Garcia *et al.,* 2011; Figure 2). SPOligos are released from the break with a 2 base pair (bp) long 5´-overhang. Each DSB generated by SPO11 is resulting in two SPOligos which are released from the break site and can be used to quantify total DSB levels as well as for identification of genome regions with enhanced DSB formation, so called hotspots (Gerton *et al.,* 2000; Pan *et al.,* 2011; Lange *et al.,* 2011; reviewed in Petes, 2001). The endonucleolytic release of SPOligos leaves free DSB ends which are exonucleolytically resected to gain 3´-single-stranded tails. This process is performed by the MRX/N (MRE11; RAD50 and XRS2/NBS1) complex in combination with SAE2/COM1 (Uanschou *et al.,* 2007; Garcia *et al.,* 2011; for reviews see Czornak *et al.,* 2008; Edlinger and Schlögelhofer, 2011).

It is most likely that the endonuclease activity of MRE11 is mediating the single strand DNA (ssDNA) nick and the exonuclease activity of MRE11 resects the ssDNA towards SPO11. The relatively short piece of ssDNA is further resected by the exonucleases EXO1 and DNA2 in combination with the helicase SGS1 (Mimitou and Symington, 2008; 2009; 2011; Rothenberg *et al.,* 2009; Manfrini *et al.,* 2010; Garcia *et al.,* 2011; reviewed in Edlinger and Schlögelhofer, 2011; Figure 2). DSBs and their processing, especially by the MRN complex, are activating the

kinases ATM or ATR via direct interaction with NBS1 (Williams *et al.,* 2010). ATM activation seems to control a negative feedback circuit that inhibits SPO11 activity as several unrelated studies in mouse, flies, and yeast have recently shown (for review see Keeney *et al.,* 2014). Additionally, ATM is phosphorylating the histone H2AX at the amino acid (aa) serine at position 139, on megabase regions surrounding DSBs (Rogakou *et al.,* 1998; Czornak *et al.,* 2008). Phosphorylated H2AX (γH2AX) forms foci at and near the break sites and accumulates various proteins which are involved in DSBs procession including the MRX/N complex, starting an activation cycle (Czornak *et al.,* 2008).

The ssDNA originating from the exonucleolytic resection by MRE11/EXO1/DNA2 is bound with high affinity by replication protein A (RPA) which is essential for the formation of crossovers in *A. thaliana* (Fanning *et al.,* 2006; Osman *et al.,* 2009; Broderick *et al.,* 2010). RPA loading onto the DNA is a prerequisite for binding of the strand exchange proteins RAD51 and DMC1 which are orthologs of the bacterial RecA protein (for review see Fanning *et al.,* 2006; Broderick *et al.,* 2010). RAD51 and DMC1 in combination with several DNA repair factors and other proteins mediate the formation of heteroduplex joints (D-loops) by strand invasion, DNA elongation, and the capture of a second DNA strand (Petukhova *et al.,* 2000; Shibata *et al.,* 2001; Li *et al.,* 2004; Bleuyard *et al.,* 2005). These D-loops are further dispersed depending on their resolution, either by crossover reactions leading to exchange of genetic material between the parental chromosomes, or by non-crossover reactions. The later ones are resulting in no exchange of genetic material or exchange of only small parts of the chromosomes, *e.g.* gene conversion (Takata *et al.,* 1998; Sonoda *et al.,* 2006; for review see Jasin and Rothstein 2013).

Exchange of genetic material between homologous chromosomes by crossovers can be seen in DAPI (4´,6-diamidino-2-phenylindole) stained meiocytes as chiasmata. The process of crossover/non-crossover decision and the resolution of crossovers are still not fully understood and proteins involved in these processes vary between different phyla. Crossover reactions can only occur if the second end of the DSB is binding to the homologous chromosome and a DHJ is formed. If the extended 3´-end is expelled from the homologous chromosome and annealed with the other end of the DSB, synthesis dependent strand annealing (SDSA) takes place which is leading to a non-crossover reaction (Nassif *et al.,* 1994; Paques and Haber, 1999; Allers and Lichten, 2001; Mimitou and Symington, 2009;

Mannuss *et al.,* 2010). The further processing of DHJs can lead to either crossover or non-crossover reactions depending on their resolution and can also be processed through direct dissolution or transient strand invasion with limited DNA synthesis, followed by dissolution, leading again to a non-crossover reaction (Schwacha and Kleckner, 1995; Allers and Lichten, 2001; Wu and Hickson, 2003; Bishop and Zickler, 2004; Youds and Boulton, 2011). The resolution of a DHJ is performed by a number of protein complexes which are able to reject D-loops or dissolve DHJs. The RTR-complex (RECQ, TOP3; RMI1) is one of these. Other proteins involved in the resolution of DHJs are for example MUS81/GEN1 and YEN1 (Hartung *et al.,* 2007b; Higgins *et al.,* 2008; Svendsen and Harper, 2010; Lorenz *et al.,* 2010; Bauknecht and Kobbe, 2014; for review see Knoll *et al.,* 2014).

It seems that, at least in budding yeast (*Saccharomyces cerevisiae*), the decision between crossover and non-crossover is made quite early at the leptotene/zygotene transition and involves a complex of proteins referred to as ZMM proteins (ZIP1, ZIP2, ZIP3, MER3 and MSH4/5) (Börner *et al.,* 2004; reviewed in Osman *et al.,* 2011). The ZMM proteins appear to stabilize strand invasion during D-loop formation and DHJs, thereby directing them to form crossovers. Mutations in the ZMM proteins, including MSH4/5, which are homologs of the *Escherichia coli* MUTS protein*,* are leading to elimination of most, but not all, crossovers. Two further important proteins, MLH1 and MLH3 which are homologs of the MUTL protein from *E. coli*, also encourage the resolution of DHJs to crossovers rather than non-crossovers. In addition to these proteins another pathways exists which enables the formation of crossovers involving MUS81 and MMS4. A homolog of MUS81 is also involved in crossover formation in *A. thaliana* (de los Santos *et al.,* 2003; Hartung *et al.,* 2007b; Higgins *et al.,* 2008). The final ligation of the DSBs is most probably performed by DNA ligase 4 in combination with XRCC4 (West *et al.,* 2000; Bleuyard *et al.,* 2004; 2006; Bray *et al.,* 2005). A schematic overview of these processes is given in figure 2.

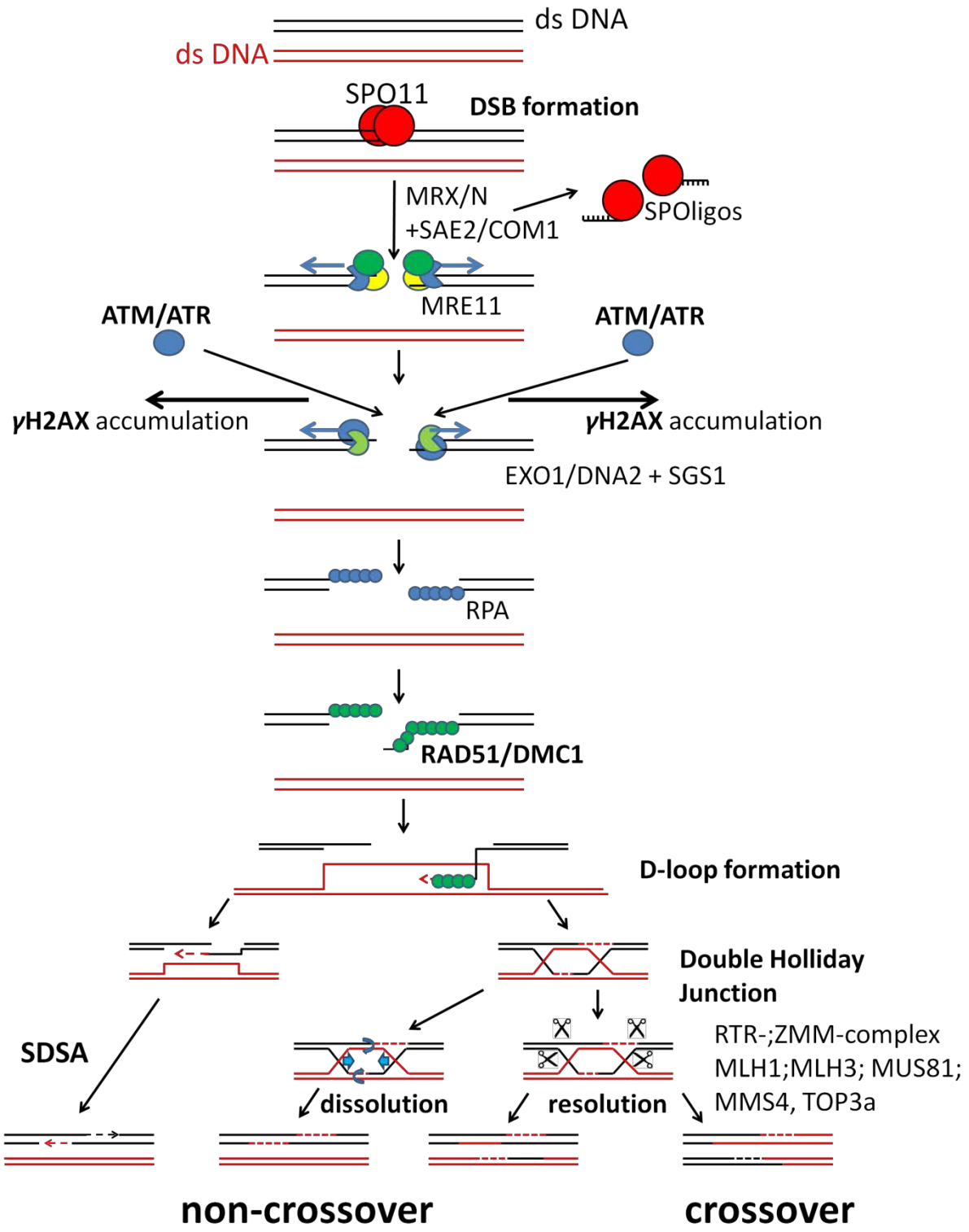


Figure 2. Double strand break induction and repair model during meiosis.

Double strand break (DSB) induction and repair during meiosis. Several key proteins of DSB induction and repair are shown at the stage at which they are active, according to the known processes in model organisms. Due to space limitation not all proteins involved in the corresponding process could be shown and are only referred to in the text. Proteins which are commonly used for marking of DSBs are written in bold. See text for details. ds DNA = double stranded DNA; SDSA = synthesis dependent strand annealing; DSB = double strand break.

The synaptonemal complex

The pairing and synapsis of the homologous chromosomes during prophase I possesses an additional specific feature, the formation of a protein complex between both homologous chromosomes, the so called synaptonemal complex (SC) which includes some of the ZMM proteins (Moses, 1969; reviewed in Zickler and Kleckner, 1999). In *A. thaliana* the two major proteins of the synaptonemal complex are ASY1, a homolog of yeast HOP1 which forms the lateral elements of the SC and ZYP1, a homolog of yeast ZMM protein ZIP1 which forms the central element of the SC (Armstrong *et al.,* 2002; Higgins *et al.,* 2005; Sanchez Moran *et al.,* 2007; 2008). Without ASY1 synapsis is interrupted and the formation of chiasmata is dramatically reduced. In the absence of ZYP1 chiasmata formation is only slightly reduced and recombination occurs not only between homologous chromosomes but also between non-homologous chromosomes (Higgins *et al.,* 2005; Sanchez Moran *et al.,* 2008).

The evolution and function of SPO11

The initiation of the recombination process and synapsis relies in many organisms, like mammals and fungi, on a single SPO11 protein which is sufficient for the induction of meiotic DSBs. Plants stand in contrast to this, since they encode for at least three different SPO11 proteins. In *A. thaliana* two of these have a function in meiosis, referred to as Ath SPO11-1 and Ath SPO11-2 (Keeney *et al.,* 1997; Grelon *et al.,* 2001; Stacey *et al.,* 2006; Hartung *et al.,* 2007a). The third one, Ath SPO11-3, is involved in endoreduplication in somatic cells working in a combination with the second subunit of topoisomerase VI (TOPVIB), but it has no specialized meiotic function (Hartung *et al.,* 2002; Stacey *et al.,* 2006; Simkova *et al.,* 2012). Rice (*Oryza sativa*) possesses furthermore two additional SPO11 proteins, referred to as Osa SPO11-4 and Osa SPO11-5, from which Osa SPO11-4 is believed to play a role in meiosis (An *et al.,* 2011; Shingu *et al.,* 2012).

In yeast and mouse it is proposed that SPO11 forms multimers/homodimers and dimers between two distinct spliced protein variants (Malik *et al.,* 2007; Cole *et al.,* 2010; de Massy, 2013). Differential splicing of *SPO11* was shown in mouse and human and so far in plants only for *SPO11-1* in *A. thaliana* (Romanienko and Camerini-Otero, 1999; 2000; Hartung and

Puchta, 2000). In mouse and human two distinct splice variants for *SPO11* were identified which possess different features concerning the timing of the DSB induction as studies in mouse have shown (Bellani *et al.,* 2010; Kauppi *et al.,* 2011). The process of aberrant splicing seems to be a common feature in meiosis. It was also identified for other meiosis specific proteins, such as DMC1 and MER2, in different species including yeast, mammals and plants pointing towards a conserved mechanism (Habu *et al.,* 1996; Nakagawa and Ogawa, 1997; Terzi and Simpson, 2008). A possible interaction of meiotic SPO11 proteins with other proteins has not been studied in detail. This is one of the reasons why the mechanisms of DSB induction by at least two different SPO11 proteins and their regulation in plants remain cryptic. It is unknown whether and how the different SPO11 proteins act together during meiosis. A theory describes both proteins forming heterodimers which interacts with the DNA and forms DSBs (Hartung *et al.,* 2007a; Malik *et al.,* 2007; de Massy, 2013). Till today no distinct evidence for this hypothesis was found, since no interaction between SPO11-1 and SPO11-2 could be identified. Protein interaction was only shown between SPO11-1 and PRD1 and for SPO11-2 and -3 with TOPVIB (Hartung *et al.,* 2002; De Muyt *et al.,* 2007). But it is known for DSB formation that SPO11-1 as well as SPO11-2 are both needed containing their active tyrosine residue. Mutations of the active tyrosine residue or changes in the conserved motifs of Ath SPO11-1 are leading to a loss of function (Hartung *et al.,* 2007a; Shingu *et al.,* 2010).

Marking the break, proteins suitable for marking of double strand breaks

Information about the localization of a protein and the timeframe in which it is active during meiosis is always resulting in a great gain of knowledge. Co-immunolocalization studies are a suitable tool to achieve such information. Two proteins which are commonly used in these studies for analyzing the different stages of prophase are ASY1 and ZYP1, since their distribution changes during the stages of prophase I. ASY1 appears early in prophase I and ZYP1 appears later but remains longer (Higgins *et al.,* 2005; Sanchez Moran *et al.,* 2008; Kurzbauer *et al.,* 2012; Figure 3). Since DSBs are too tiny to be seen directly during meiosis, the use of DSB associated proteins is a common method to mark the breaks. In principal

most of the proteins mentioned above could be used for this (Armstrong *et al.,* 2002; Higgins *et al.,* 2005; Sanchez Moran *et al.,* 2007).

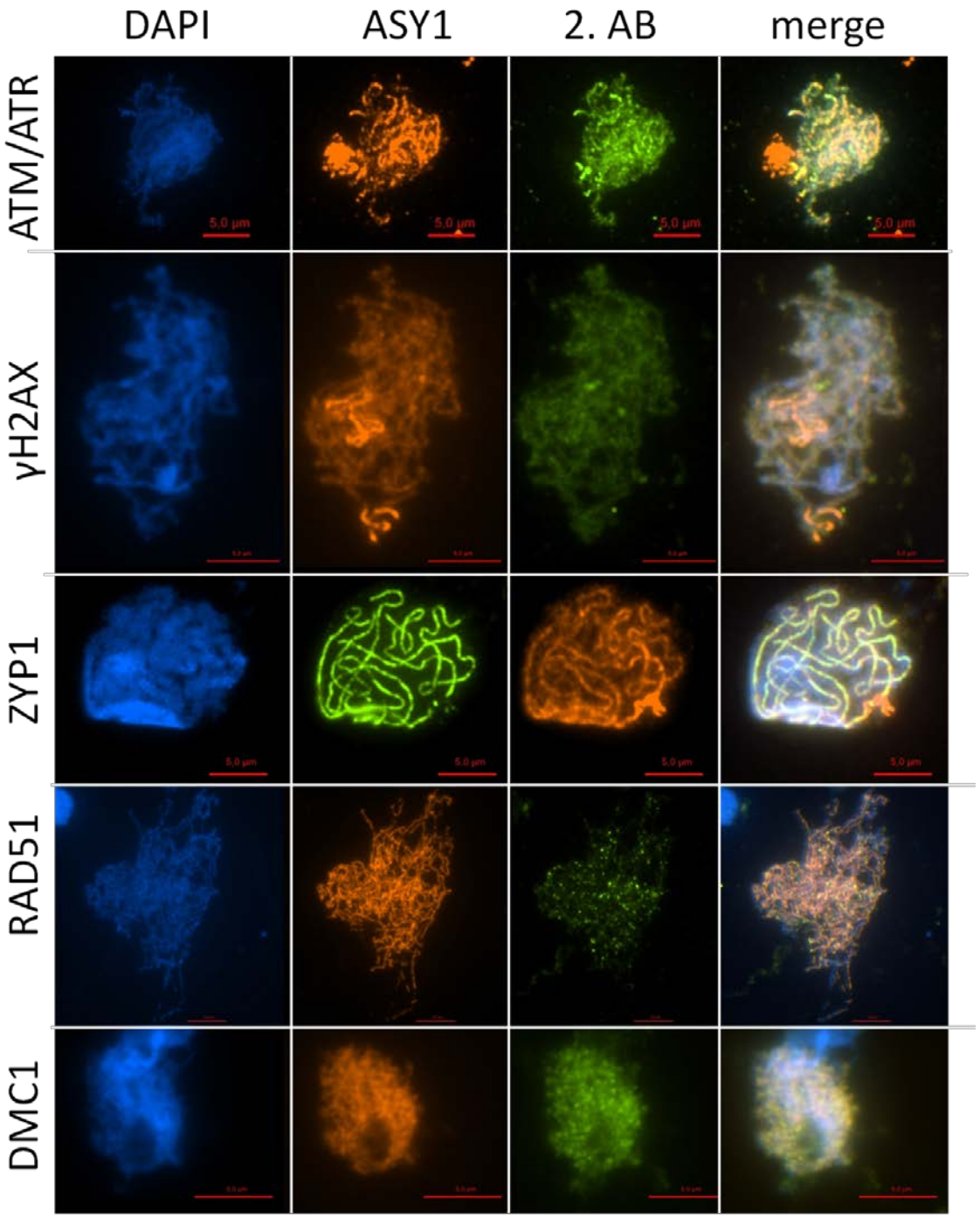


Figure 3. Co-immunolocalization study of meiosis specific proteins in meiocytes of wild type *A. thaliana*.

Meiocytes in different meiotic stages were counterstained with DAPI and immunolocalization of ASY1 as first antibody and various DSBs associated proteins as second antibody (as indicated) was performed. Presented meiotic stages have been chosen according to timeframe in which the corresponding protein can be detected the best. AB = Antibody, red bar = 5 µm

But some of the proteins, like SPO11, are only binding during a short timeframe on the chromosomes, making detection difficult. Commonly used proteins for marking DSBs are those that remain longer at the break site and accumulate there, like γH2AX, ATM and ATR or DMC1/RAD51 (Sanchez Moran *et al.,* 2007; 2008; Kurzbauer *et al.,* 2012; Figure 3; Figure 4). In *A. thaliana,* SPO11-1 can only be localized in a 5 h time frame during preleptotene and leptotene at the chromosomes and with the present antibody SPO11-1 is hard to detect (Figure 4).

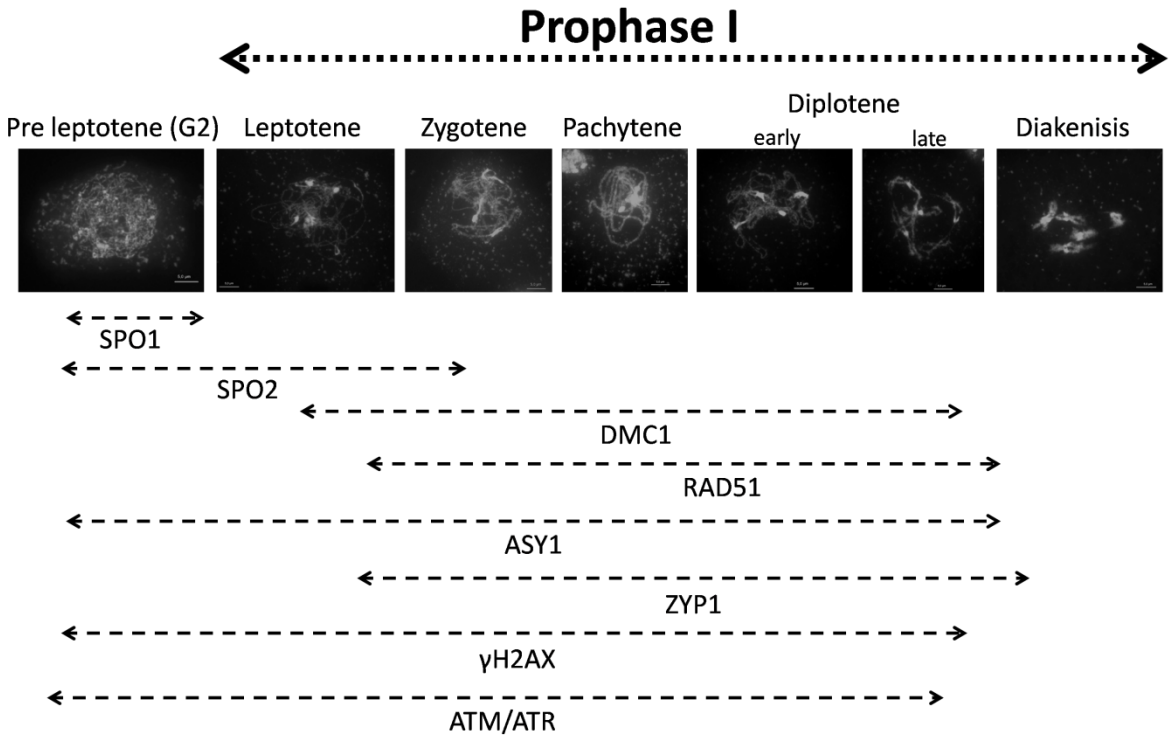


Figure 4. Timely distribution of double strand break associated proteins during prophase I in *A. thaliana*.

Fluorescence micrograph of chromosomes in Prophase I stages, from spread wild type *A. thaliana* pollen mother cells stained with DAPI. Various double strand break associated proteins and the stages in which the corresponding protein is detectable during prophase I are shown, the timely distribution of each protein is indicated as a dashed line. Bar = 5 µm.

So far no one had ever looked at the localization and timing of SPO11-2 during meiosis, since no antibody was available (Sanchez Moran *et al.,* 2008; personal communication).

Aims of this work:

During this work several questions concerning the function, regulation and development of the different SPO11 proteins and their splicing in *A. thaliana* and other organisms were addressed. It is commonly known that the function of meiosis is conserved between eukaryotes, but not for all proteins which are involved in yeast meiosis, orthologs in different clades can be identified. Even if orthologous genes can be found some of these have taken over different functions between miscellaneous species (for reviews see Keeney and Neale, 2006; Cole *et al.,* 2010; de Massy, 2013; Keeney *et al.,* 2014). The growing number of available databases unclosed the opportunity to elucidate the development, distribution, and sequence conservation of different SPO11 proteins throughout the kingdoms of life (Sprink and Hartung, 2014). To evaluate if besides the protein itself also the function of SPO11 is conserved between various plants and still conserved in more diverged *SPO11* genes from green algae and animals, foreign *SPO11* genes were introduced into SPO11-1 and -2 knockout lines. For this purpose genomic DNA as well as complementary DNA (cDNA) was used (Sprink and Hartung, 2015).

Even though derived from a common ancestor and sharing the same conserved motifs, the two meiotic active SPO11 homologs in *A. thaliana* are not redundant and both are needed in a functional form, indicating that the difference between both paralogs is somehow embedded in the less conserved regions of the proteins (Stacey *et al.,* 2006; Hartung *et al.,* 2007a). To gain a deeper understanding, why plants need two very different SPO11 proteins for a functional meiosis and which regions of the proteins are defining the specific functions, three of the non-conserved parts were interchanged between the paralogs, Ath SPO11-1 and -2. Since disruption just before the last exon of Ath SPO11-2 by T-DNA (transfer DNA) insertion is leading to a total loss of function, indicating that this might be an essential part, this last exon was exchanged between both paralogs, too (Stacey *et al.,* 2006; Hartung *et al.,* 2007a). Additionally, both proteins were truncated by the last exon to investigate if the effect of the T-DNA insertion could be mimicked and if the loss of the last exon has such a severe effect on SPO11-1 as it has on SPO11-2 (Sprink and Hartung, 2015). As T-DNA insertions can have multiple effects on the plant and since a mutation line for SPO11-3 is no longer available, new mutation lines for all three SPO11 paralogs in *A. thaliana* were created

and further propagated, to achieve mutation lines without T-DNA insertion and their possible effects. A mutation of the genomic region just around the active tyrosine of the protein was aimed, using the novel sequence specific nucleases (SSNs) TALENs (Transcription activator like effector nucleases) and CRISPR/Cas9 (Clustered regularly interspaced short palyndromic repeats/Cas9) (recently reviewed in Belhaj *et al.,* 2015; Sprink *et al.,* 2015). Furthermore, it has been evaluated if SPO11-2 is binding on the DNA during leptotene as SPO11-1 does, to figure out, if both proteins work together during meiosis. For this purpose SPO11-2 specific antibodies were designed and produced using peptides from the N- and C-terminal part of the protein. Moreover, an additional antibody against the N-terminal region of SPO11-1 was produced, aiming to achieve stronger signals in immunolocalization studies. The SPO11-2 N-terminal antibody was produced in rabbit and mouse, so that the antibody can be used for co-immunolocalization studies together with an antibody against SPO11-1. To gain an idea whether the different splice forms of *SPO11* are also resulting in alternative protein forms a western blot analysis was performed to see if multiple signals in the blot, which might be related to SPO11-2 can be identified.

Results and discussion:

The function of SPO11-1 and SPO11-2 is sequence and to a certain extent species specific

By the construction of four chimeric *SPO11* genes, a sequence specific function of *A. thaliana* SPO11-1 and SPO11-2 was shown. Three non-conserved parts as well as the conserved last exon have been interchanged between *A. thaliana SPO11-1* and *-2*. Additionally, both *SPO11* genes were truncated by the conserved last exon. None of these changes let to a complementation of the sterile phenotype of the SPO11 knockout lines *spo11-1-3* and *spo11-2-3.* A light but significant increase in seed set could only be observed, when the last exon of *SPO11-1* was exchanged with the one from *SPO11-2* (Sprink and Hartung, 2015).

By exchanging *SPO11-1* and *-2* in *A. thaliana* with genes from different species, a species specific function of the respective SPO11 could also be identified. Functional complementation of the sterile phenotype of the *A. thaliana* SPO11 knockout mutants was only achieved, when *SPO11* from a plant belonging to the same family as *A. thaliana,* was used for complementation, such as *SPO11* from *B. rapa.* Both plants diverged around 20 million years ago (mya) (Town *et al.,* 2006). Such a functional complementation could not be observed, when *SPO11* from papaya was integrated which belongs to a different family, but is still part of the order of Brassicales. However, a partial complementation could be accomplished when cDNA of papaya *SPO11-1* was integrated multiple times into the genome of the respective SPO11 knockout. A complementation was neither attained when *SPO11* from a monocot species was used nor when even more diverged *SPO11* genes from mouse and green algae were integrated.

These findings are pointing towards a highly sequence specific function of both SPO11 paralogs in *A. thaliana* as well as a species specific function for the different SPO11 orthologs and are highlighting the essential function of the well conserved last exon (Sprink and Hartung, 2015).

The splicing landscape of *SPO11* is diverse and shows species and sequence specific features

During the identification of SPO11 orthologs in different plants by using commonly available databases retained introns were identified in some of the gathered EST sequences as well as sequences in which exons were skipped. Throughout the determination of the full length cDNA constructs of *SPO11-1* and *-2* from various plants just the same has been observed. It was recognized that the analyzed plants exhibit a pattern of aberrant spliced *SPO11-1* and *SPO11-2* products. Due to these novel findings, splicing of *SPO11* seems to be conserved between the different SPO11 orthologs involved in meiosis throughout evolution (Sprink and Hartung, 2014). In human and mouse two distinct alternative spliced isoforms exist which both possess essential functions during meiosis (Romanienko *et al.,* 1999; Kauppi *et al.,* 2011). Such a conserved alternative spliced form could not be detected for *SPO11-1* or *SPO11-2* in plants. Nevertheless, other putative functional forms were recognized. Unfortunately none of these putative functional forms seem to be conserved between different plants. Additionally, the positional distribution as well as the number of aberrant splicing events is strongly differing between the analyzed plants. An analysis of aberrant splicing of *SPO11-1* in generative and vegetative tissue in several species exposed a difference in the splicing landscape between both tissue types. These findings revealed a tissue specificity of *SPO11* splicing and that splicing of *SPO11* differs between the respective gene sequence and plant species (Sprink and Hartung, 2014).

An evaluation of aberrant splicing of *SPO11* genes from species which were analyzed earlier, introduced into SPO11 knockout lines, sustained the observations made before. The splicing of Ath *SPO11-1* and *-2* showed no remarkable changes when being reintroduced into the respective knockout plants. The splicing of foreign *SPO11* in *A. thaliana* differs from the splicing in its origin plant and, additionally, positional distribution of the splicing events changed depending on the sequence identity between the introduced *SPO11* and Ath *SPO11*. For introduced *SPO11* with higher sequence identity to *SPO11* from *A. thaliana* splicing of prior identified forms, in its initial plant or in *A. thaliana* for the respective homolog, seems more likely (Sprink and Hartung, 2015).

Additionally, the splicing patterns of the chimeric *SPO11* genes (SPO11swap 1 to 4) were evaluated, to gain a deeper insight into the sequence specificity of *SPO11* splicing in *A. thaliana*. Analyzing these patterns revealed that the splicing depends on the nearby sequences, as seen for the N-terminal part of *SPO11-1*, which shows retention of intron number two, even when it is fused to the *SPO11-2* backbone or internal parts of *SPO11-1* are exchanged (Figure 5b), as well as on the overall exon-intron orientation. Especially changes in the overall number of exons and introns as well as changes in the C-terminal part of each gene are leading to novel splice forms (Figure 5). The C-terminal part seems to have an essential function in regulation of splicing especially for *SPO11-2,* since strong changes in the positional distribution of the splicing events have been recognized (Figure 5D-e). Exchange of the conserved last exon is only leading to small changes in the positional distribution of splicing events for *SPO11-1*, but for *SPO11*-2 massive intron retention can be observed (Figure 5e). This finding is endorsing the essential function of the last exon in particular for *SPO11-2* which was shown before by truncation and the exchange of the last exon between both paralogs (Sprink and Hartung, 2015). Since correctly spliced forms could be identified for all interchanged constructs the changes in the exon-intron structure of *SPO11-1* and *-2* as well as the changes of the sequences themselves do not lead to a total disarrangement of the splicing machinery for SPO11 in *A. thaliana*.

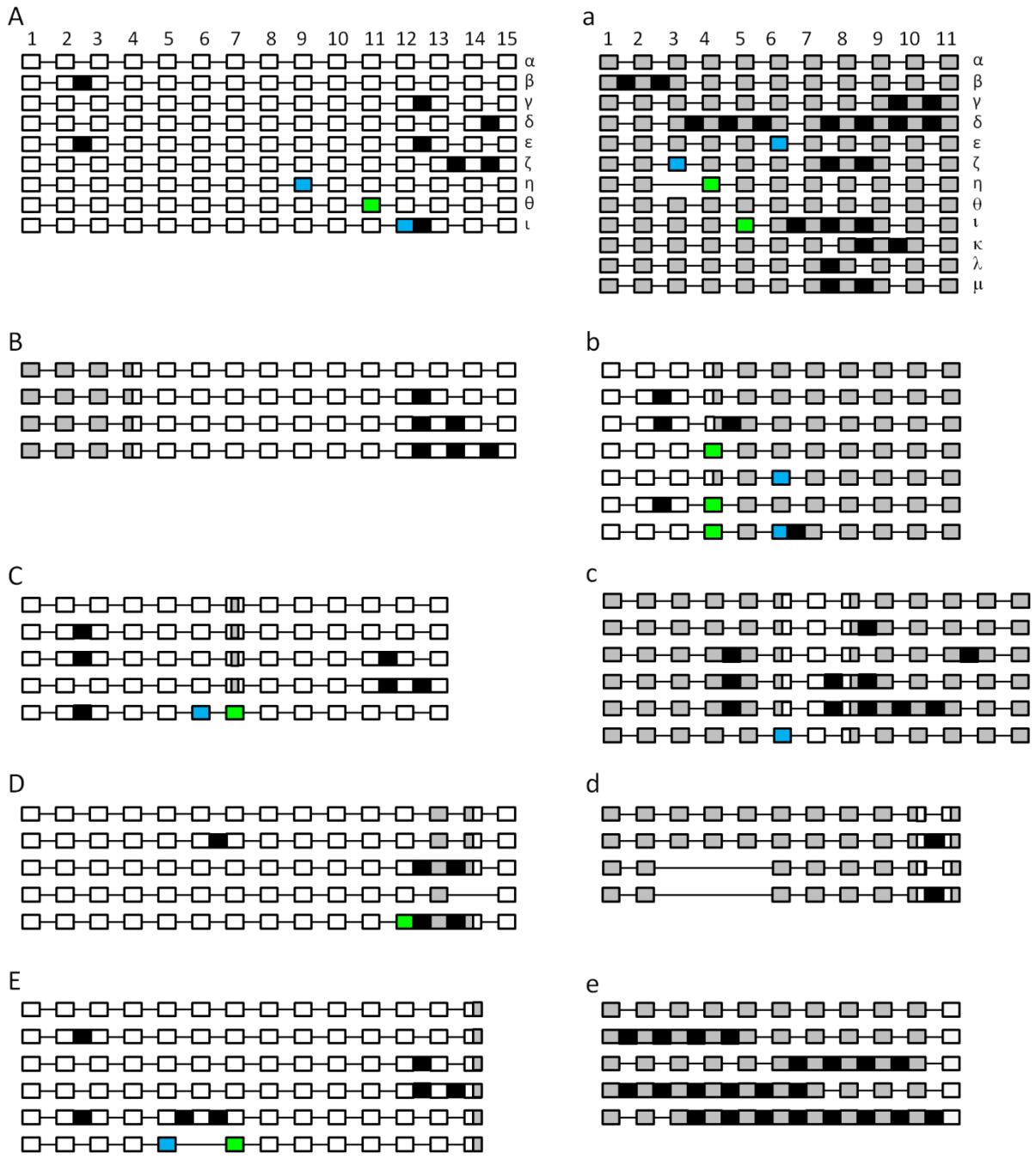


Figure 5. The splicing landscape of chimeric *SPO11* genes.

Schematic non scaled scheme of the different splice forms of *SPO11-1* (**A-E**) and *SPO11-2* (**a-e**). Splice forms of *A. thaliana SPO11-1* (**A**) and *-2* (**a**) as well as SPOswap1 (**B,b**), SPOswap2 (**C,c**), SPOswap3 (**D,d**) and SPOswap4 (**E,e**) are shown. Exons are numbered, *SPO11-1* exons are shown as white, *SPO11-2* exons as grey blocks, exons with parts of *SPO11-1* and *-2* are shown in both colors. Spliced introns are illustrated as black lines, intron retention events are illustrated as black boxes, alternative 5´splice site selection is shown as blue boxes and alternative 3´splice site as light green boxes. In case of exon skipping the corresponding white box is missing. Known splice forms are numbered in greek letters (according to Sprink and Hartung, 2014; 2015).

Taken everything into account, the splicing landscape of *SPO11* shows tissue specificity and depends, like the function of SPO11, also on the respective gene sequences and on the plant species. But since *SPO11* from closely related plants are predominantly spliced in a correct way the splicing of *SPO11* from further diverged plants seem to be less effective. Additionally, previously unknown positional distributions of splice effects in the chimeric *SPO11* genes were identified which had never been observed so far. Summarizing these results let assume that the splicing of *SPO11* is not only embedded in the plain DNA sequence of the respective *SPO11* gene but also in its internal composition, as well as in other factors which remain unknown and are likely to be species specific. Such factors could be splicing factors or other proteins which might have coevolved with SPO11. But the process of alternative splicing seems to be a common mechanism in meiosis since it can be found in various species and for different genes involved in meiosis and it might have regulatory functions. (Engebrecht *et al.,* 1991; Mackey *et al.,* 1997; Romanienko and Camerini-Otero, 1999; Terzi and Simpson, 2008; Sprink and Hartung, 2014). The examination of the splicing patterns of *SPO11* in pollen mother cells in different meiotic stages would be of great interest. Such an analysis could reveal if a correlation between aberrant spliced isoforms and meiotic stages exist in plants comparable to the one found in mouse and human (Kauppi *et al.,* 2011). Furthermore, analyzing the functions and splicing patterns of artificial *SPO11* genes composed of Ath *SPO11* and *SPO11* from other organisms would also be of great interest.

SPO11-2 is located on the chromosome during leptotene and early zygotene

Since no antibody was available against SPO11-2 and it had never been looked at the distribution and timing of SPO11-2 during meiosis a polyclonal antibody was designed and produced to gain such information. For the production of the antibody a 21 aa long peptide localized in the N-terminal part of SPO11-2 has been chosen which is not conserved in SPO11-1. The N-terminal part of the protein was selected since this part seems to be accessible by the antibody in localization studies as a prediction of the secondary structure let assume (Figure 6).

Ath SPO11-1: 41.81 kilodalton

```
MEGKFAISESTNLLQRIKDFTQSVVVDLAEGRSPKISINQFRNYCMNPEADCLCSSDKPKGQEIFTLKKEPQTYR
---------HHHHHHHHHHHHHHHHHHHH-----EEEE-----EEE-----EEEE-----EEEEE------HHHH

IDMLLRVLLIVQQLLQENRHASKRDIYYMHPSAFKAQSIVDRAIGDICILFQCSRYNLNVVSVGNGLVMGWLKFR
HHHHHHHHHHHHHHHH---EEEEEEEEE---------HHHHHHHHHHHHHH--------EEE---EEEE---EEE

EAGRKFDCLNSLNTAYPVPVLVEEVEDIVSLAEYILVVEKETVFQRLANDMFCKTNRCIVITGRGYPDVSTRRFL
----EEEE---------------EE-------EEEEEE-HHHHHHHHH---------EEEE------HHHHHHH

RLLMEKLHLPVHCLVDCDPYGFEILATYRFGSMQMAYDIESLRAPDMKWLGAFPSDSEVYSVPKQCLLPLTEEDK
HHHHH-----EEEEE-----HHHHHHHHHH----H----------EEEEE-------------------HHHH

KRTEAMLLRCYLKREMPQWRLELETMLKRGVKFEIEALSVHSLSFLSEVYIPSKIRREVSSP
HHHHHHHH-------HHHHHHHHHHHH----EEHHHH------HHHHHHHHHHHH----
```

Ath SPO11-2: 43.13 kilodalton

```
MEESSGLSSMKFFSDQHLSYADILLPHEARARIEVSVLNLLRILNSPDPAISDLSLINRKRSNSCINKGILTDVS
------------------------HHHHHHHHHHHHHHHHHHH-------EEEE------------------

YIFLSTSFTKSSLTNAKTAKAFVRVWKVMEICFQILLQEKRVTQRELFYKLLCDSPDYFSSQIEVNRSVQDVVAL
-EEE---EEEE-----HHHHHHHHHHHHHHHHHHHHH---EE-EEEEEEE----------HHHHHHHHHHHHHH

LRCSRYSLGIMASSRGLVAGRLFLQEPGKEAVDCSACGSSGFAITGDLNLLDNTIMRTDARYIIIVEKHAIFHRL
H-------EEEE----EEEEEEEEE----EEEE------------------EE-----EEEEEE—HHHHHH

VEDRVFNHIPCVFITAKGYPDIATRFFLHRMSTTFPDLPILVLVDWNPAGLAILCTFKFGSIGMGLEAYRYACNV
HH---------EEEE------HHHHHHHHHHHHH-----EEEEE-----HHHHHHHHH----------------E

KWIGLRGDDLNLIPEESLVPLKPKDSQIAKSLLSSKILQENYIEELSLMVQTGKRAEIEALYCHGYNYLGKYIAT
EEE-----------------HHHHHHHHHHHHH-----HHHHHHHHHHH----EEHHHH------HHHHHHH

KIVQGKYI
HHHH----
```

Figure 6. Predicted secondary structure of *A. thaliana* SPO11-1 and SPO11-2.

Predicted secondary protein structure and protein mass of the two meiotic SPO11 paralogs in *A. thaliana*. Prediction was made using Jpred V. 3.0 (Cole *et al.,* 2008). Full length protein sequence is shown with subjacent predicted secondary structure. H = alpha helix; E = beta sheet; - = random coil. The 21 amino acids which were used for the production of an N-terminal SPO11-2 antibody are shown in bold red.

The peptide was utilized to immunize two rabbits. The first animal was used for an intravenously (IV) induction of the immunoreactions. In the second animal the immunoreaction was induced by intramuscular (IM) application of the peptide. The blood sera of both animals were utilized for immunolocalization of SPO11-2 in spread preparations of pollen mother cells from wild type *A. thaliana* anthers. With both sera foci could be detected during early prophase I. The serum of the animal which was used for IM induction

of immunoreactions showed a brighter signal, therefore this serum was purified and the cleaned antibodies were used in further studies. Immunolocalization studies using the purified antibody revealed the presence of foci during leptotene and early zygotene on the chromosomes (Figure 7). In *spo11-2-3* no signal on the chromosomes could be detected, indicating that the antibody is specific against SPO11-2 and is not binding SPO11-1 or any other protein in a noticeable amount during meiosis (Figure 7). Around 100 SPO11-2 foci per cell could be identified on the chromosomes during early leptotene which is comparable to the number of SPO11-1 foci found in wild type *A. thaliana* (Sanchez Moran *et al.,* 2007; personal communication). Surprisingly, SPO11-2 can be detected even during zygotene, whereas SPO11-1 cannot be detected during this stage (Sanchez Moran *et al.,* 2007). Nevertheless, in other organisms as mouse and yeast such a behavior of SPO11 staining was observed before (Romanienko and Camerini-Otero, 2000; Prieler *et al.,* 2005). The function of that late detectable SPO11 on the chromosome is still ambiguous since DSB induction is clearly induced earlier by SPO11 during leptotene. So it can only be speculated that the presence of SPO11 in later stages might have other function than cleaving the DNA. Prieler *et al.,* suggested an interaction of these late SPO11 with recombination hotspots. But in plants such an interaction was never shown before and further analyses, like analyses of SPOligo distribution in combination with hotspot identification, have to be made.

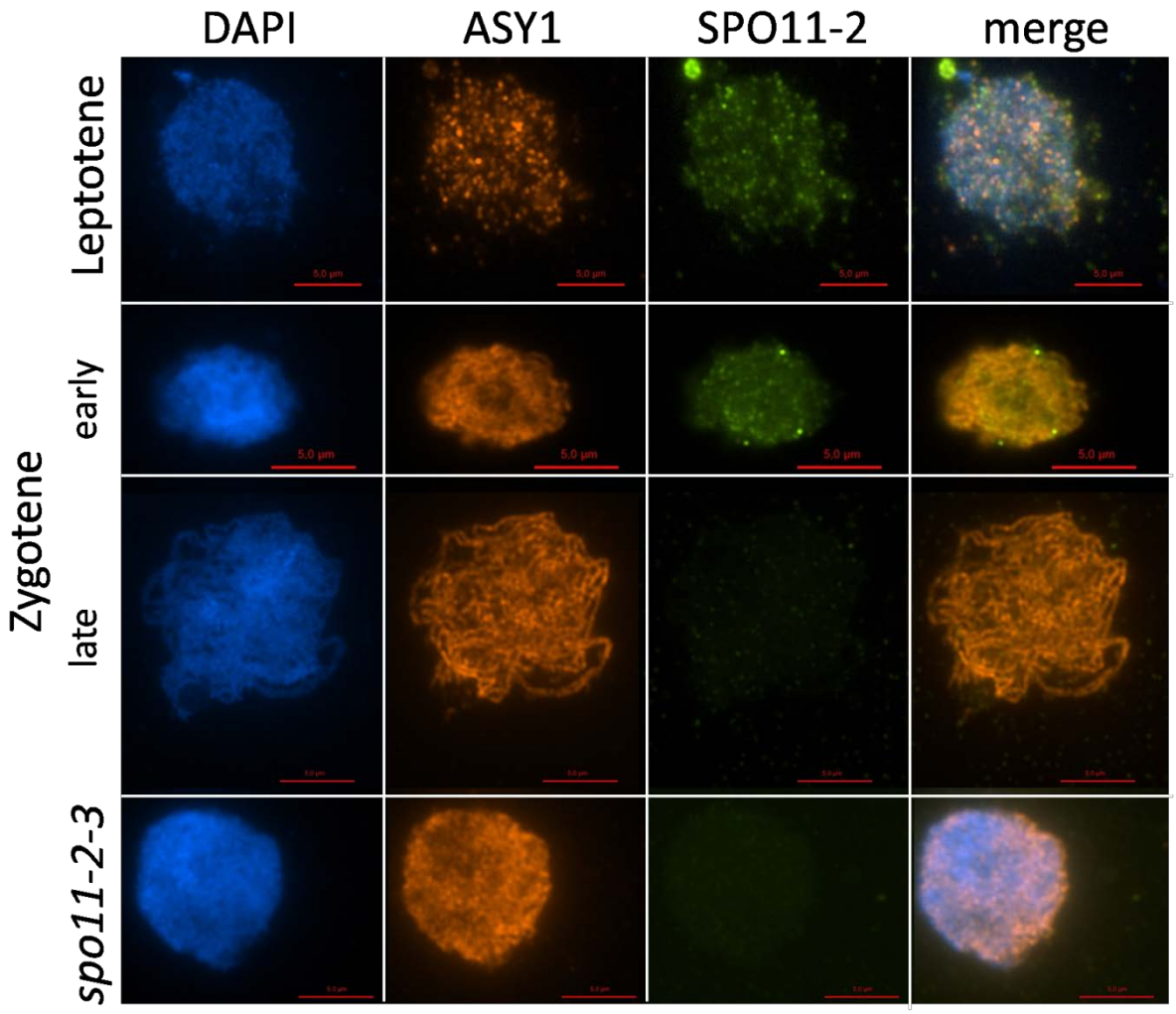


Figure 7. Co-immunolocalization study of ASY1 and SPO11-2 during different meiotic stages.

Meiocytes of wild type *A. thaliana* plants in different meiotic stages (as indicated) and a meiocyte of *spo11-2-3* in leptotene stage were counterstained with DAPI (blue) and immunolocalization of ASY1 (orange) and SPO11-2 (green) was performed. The signal of SPO11-2 is strongest in early leptotene but remains to early zygotene. In late zygotene only very few foci remain, which are hard to distinguish from the overall background. In *spo11-2-3* meiocytes no signal brighter than the overall background can be detected. Red bar = 5 µm.

Co-localization of SPO11-1 and SPO11-2

The fact that both SPO11 proteins in *A. thaliana* can be detected during early leptotene in a comparable number is supporting the theory of both proteins working together. A co-immunolocalization study using antibodies against SPO11-1 and -2 should reveal if both proteins colocalize. For this purpose an antibody against SPO11-2 was produced in mouse,

using the same peptide as used in rabbits. First co-immunolocalization studies have been performed but only a few cells could be analyzed so far. A localization of both proteins proximal to each other was found but the detection is challenging due to differences in signal intensity and the low number of analyzed cells (Figure 8). To propose a meaningful and robust statement further analyses have to be made and much more cells have to be analyzed.

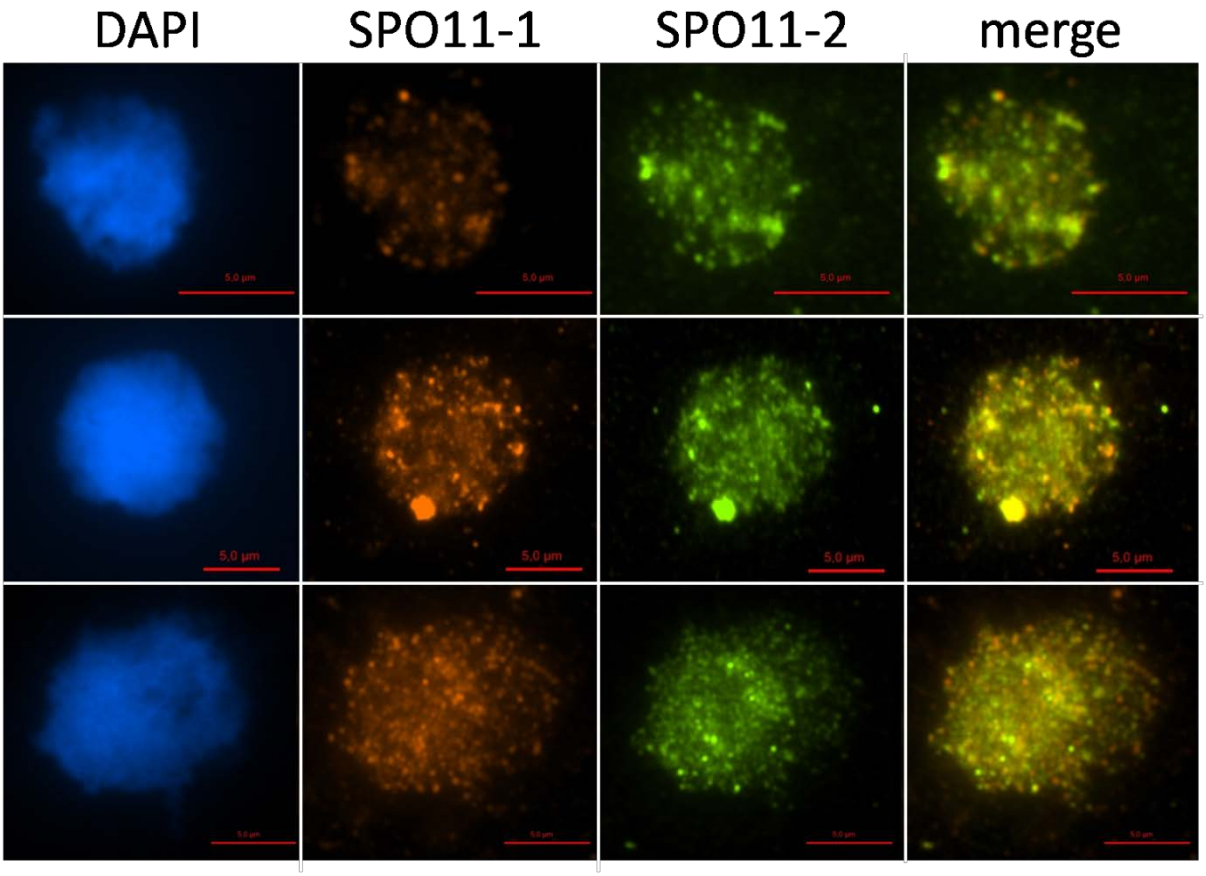


Figure 8. Co-immunolocalization of SPO11-1 and SPO11-2.

Meiocytes of wild type *A. thaliana* plants in preleptotene/leptotene stage were counterstained with DAPI (blue) and co-immunolocalization of SPO11-1 (orange) and SPO11-2 (green) was performed. The signals of both proteins can be detected proximal to each other and in some cases signals overlap, which can be seen in yellow areas in the merged images. Red bar = 5 µm.

Western blot analysis of SPO11-2

To check whether the antibody produced in rabbit is detecting multiple variants of SPO11-2 or other proteins a western blot using two samples of purified total protein from *A. thaliana* wild type flowers was performed (Figure 9). To resolve possible protein complexes half of each sample was boiled at 95°C before separation. The peptide used for immunization of the animals was used as a positive control. The western blot showed multiple signals between 30- and 45 kilodalton (kDa) and two additional signals at around 60 kDa in all samples.

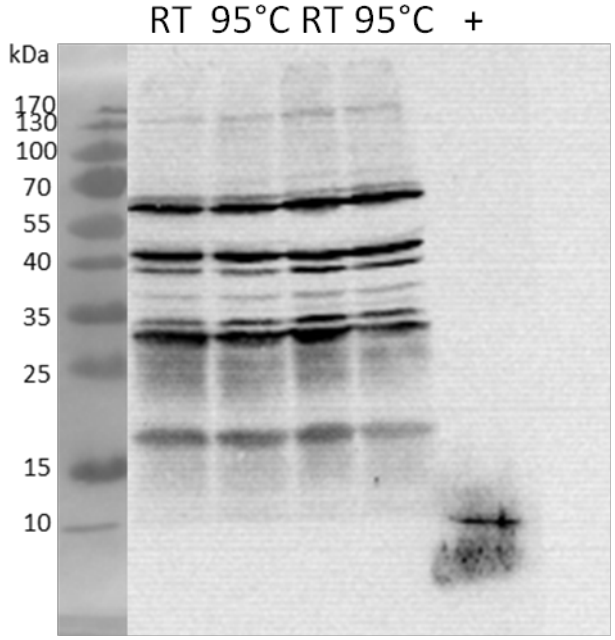


Figure 9. Western blot analysis of protein samples from *A. thaliana* flowers using SPO11-2 antibody.

Western blot of two wild type *A. thaliana* flower protein samples detected with SPO11-2 antibody (1 to 500), before (RT) and after heating at 95°C (95°C). Ten nanogram of the peptide used for the production of the antibody was loaded as positive control (+).

The predicted protein mass of SPO11-2 is 43 kDa indicating that the antibody seems to detect predominantly SPO11-2. The smaller proteins might be alternative spliced forms, as they were spotted before in western blot analyses using a SPO11 specific antibody performed in mouse (Bellani *et al.,* 2010; Lange *et al.,* 2011). The proteins detected at around 60 kDa might be (i) larger isoforms of SPO11-2 (ii) complexes or (iii) unspecific

binding of the antibody. Since heating of the protein samples is not resulting in dissolution of these bands, unspecific binding or the presence of a larger isoform seem to be the more likely hypothesis (Figure 9). Performing two dimensional (2D) electrophoresis followed by protein mass spectroscopy analysis such as MALDI TOF (matrix-assisted laser desorption/ionization, time of flight) could help to identify the proteins which are detected by the antibody and validate the assumption that predominantly SPO11-2 is detected. Such analyses should be done in the future. An additional antibody which is binding at the C-terminal end of the protein, should also detect less protein bands, since many aberrant spliced isoforms are missing the C-terminal end of the gene. An additional western blot analysis using an antibody against SPO11-1 would reveal if such a pattern of protein bands can also be identified for SPO11-1.Since not enough serum containing an antibody against SPO11-1 was available and no antibody against the C-terminal end of SPO11-2 exists, an additional antibody against SPO11-2, using a 21 aa long peptide which is located at the C-terminal end was designed. Furthermore, an additional antibody against SPO11-1, using a 23 aa long peptide which is located in the N-terminal part of SPO11-1, was designed to perform such analyses (Figure 10). The N-terminal part was chosen because of valuable performance of the SPO11-2 antibody and the part seems to be accessible in SPO11-1 like it is in SPO11-2 (Figure 6). Both antibodies are produced in rabbits the same way as the first SPO11-2 antibody. But no analyses could be done using these antibodies so far since they are still in production.

Figure 10. Positional distribution of SPO11 specific antibodies.

Multiple alignment (done with Lasergene V. 12.1.0) of SPO11-1-and-2 from *A. thaliana* (Ath); *B. rapa* (Bra) and *C. papaya* (Cpa). Gaps are represented by dashes; conserved amino acids are shaded in yellow. Conserved motifs are indicated and shown as black boxes. The active tyrosine within motif one is marked with an asterix. Amino acids chosen for the production of antibodies are marked as green (SPO11-1) or blue (SPO11-2) boxes. The already existing antibody against SPO11-1 from the group of Prof. Chris Franklin (Sanchez Moran *et al*., 2007) is marked as an orange box. The amino acid numbering of each protein is shown on the right (according to Sprink and Hartung, 2015)

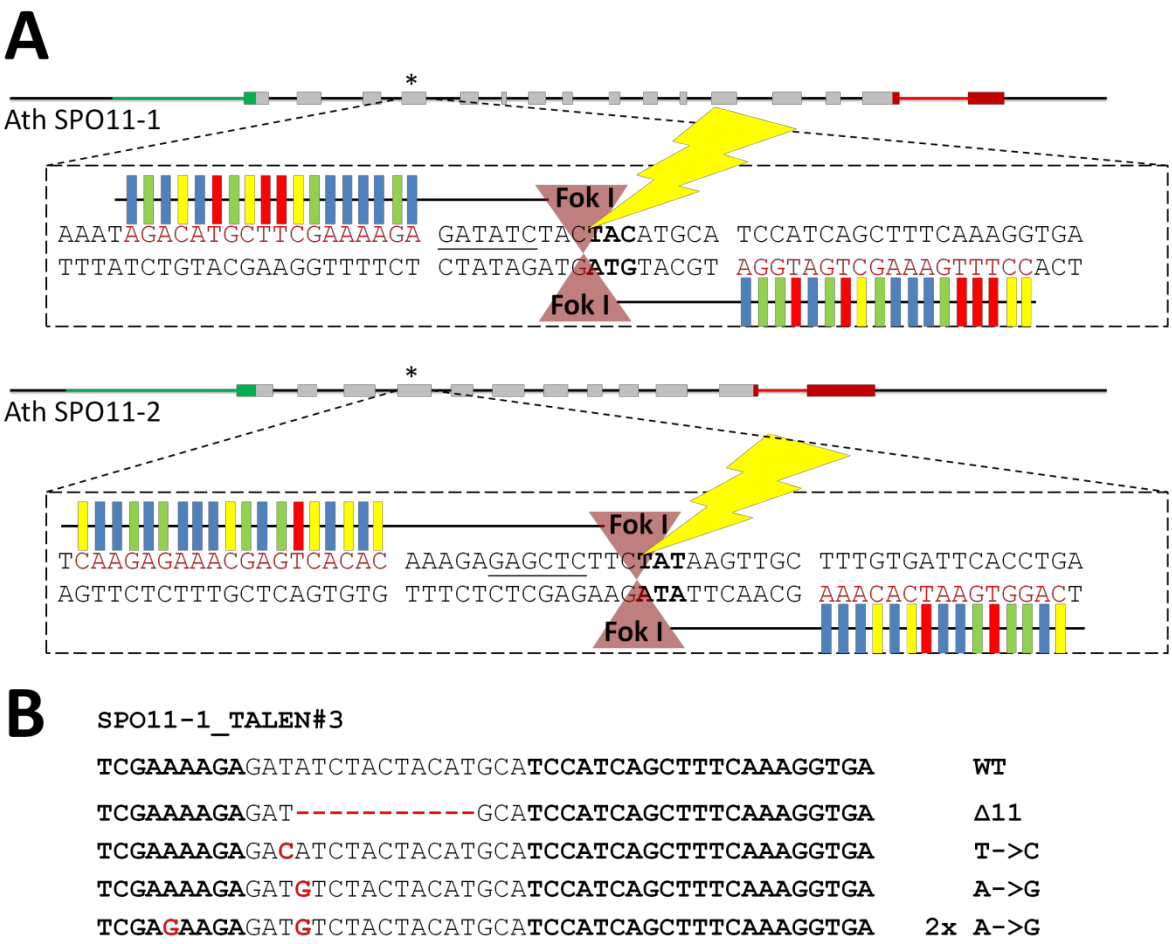
Production of T-DNA free, SPO11-mutation lines by sequence specific nucleases

A phenotypic difference in seed production between the two T-DNA insertion lines *spo11-2-2* and *spo11-2-3* has been observed earlier (Hartung *et al.,* 2007a). For some reason the seed set of *spo11-2-2* is lower compared to *spo11-2-3.* This is hard to explain because the putatively expressed form of SPO11-2 is only seven aa longer in *spo11-2-2* than in *spo11-2-3* and these seven aa are not even part of a conserved domain (Hartung *et al.,* 2007a). To investigate if the difference in seed set is an artifact of the T-DNA insertion or due to remaining truncated protein forms we used the sequence specific nucleases (SSNs)

TALENs and CRISPR/Cas9 in the aim of generating T-DNA free knockout lines, mutated only around the active tyrosine residue. In addition to this a CRISPR/Cas9 construct targeting the genomic region just around the active tyrosine of *SPO11-3* was produced.

Introducing mutations with Transcription Activator Like Effector Nucleases (TALENs)

Two TALEN pairs were designed to target either the genomic region around the active tyrosine of *SPO11-1* or *SPO11-2*. The TALEN pairs were constructed spanning a defined restriction enzyme site (Figure 11A). Screening the offspring revealed a very low overall transformation rate of < 0.1% (SPO11-1 TALEN: 5 out of ~10000; SPO11-2 TALEN: 2 out of ~10000). No heritable mutations could be detected in plants carrying a TALEN pair. Only a very few somatic mutations in plants transformed with the SPO11-1 TALEN pair were identified using restriction enzyme based mutation analysis and sequencing. One deletion of 11 bp was detected at the targeted site all other mutations were single nucleotide substitutions (Figure 11B).



B

SPO11-1_TALEN#3	
TCGAAAAGAGATATCTACTACATGCATCCATCAGCTTTCAAAGGTGA	WT
TCGAAAAGAGAT----------GCATCCATCAGCTTTCAAAGGTGA	Δ11
TCGAAAAGAGACATCTACTACATGCATCCATCAGCTTTCAAAGGTGA	T->C
TCGAAAAGAGATGTCTACTACATGCATCCATCAGCTTTCAAAGGTGA	A->G
TCGAGAAGAGATGTCTACTACATGCATCCATCAGCTTTCAAAGGTGA	2x A->G

Figure 11. Mutation induction approach in *A. thaliana SPO11-1* and *-2* using transcription activator like effector nucleases (TALENs).

(A) Schematic structure of *SPO11-1* and *-2* as well as the TALEN pairs used for double strand break induction (as indicated). The binding site of each TALEN monomer is written in red, the active tyrosine is written in bold and marked with an asterix in the scheme. The spanned restriction enzyme site is underlined. The flash is indicating the predicted cut site of the FokI dimer.

(B) Sanger sequencing results from all analyzed clones obtained from the restriction enzyme based mutation analysis. Sequence of the wild type gene (WT) and all mutated alleles obtained from the mutation identification in the SPO11-1 TALEN approach plant no. 3 are shown. TALEN binding sites are written in bold, mutations are marked in red. The type of the mutation is mentioned at the right site.

It is unclear why the efficiency of the TALEN pairs is quite low in this mutation approaches. But comparable experiments made at the Karlsruher Institute of Technology by the group of Prof. Holger Puchta, using the same vectors, gained in similar results. No mutation could be observed in all approaches (>30 approaches; Alexander Knoll, personal communication). The reason for this remains ambiguous since mutation induction in *A. thaliana* and other plants is possible as several publications have recently shown (Christian *et al.,* 2013; Chen *et al.,* 2014; for recent review see Sprink *et al.,* 2015). One possibility might be that the binary vector pSW5 is not suitable for induction of mutations or that the expression of the TALEN pairs is not leading to a sufficient induction of mutations. Additionally, we might have missed some events by using only restriction enzyme based mutation analysis. But since the group of Prof. Puchta has not found any mutation in their approaches, even in a T7 endonuclease assay, we decided to change the mutation system (Kim *et al.,* 2009).

Mutation induction using the clustered regularly interspaced short palindromic repeats/Cas9 system

One CRISPR/Cas9 construct was designed for each of the three SPO11 paralogs in *A. thaliana* targeting the genomic region around the active tyrosine of each protein (Figure 12).

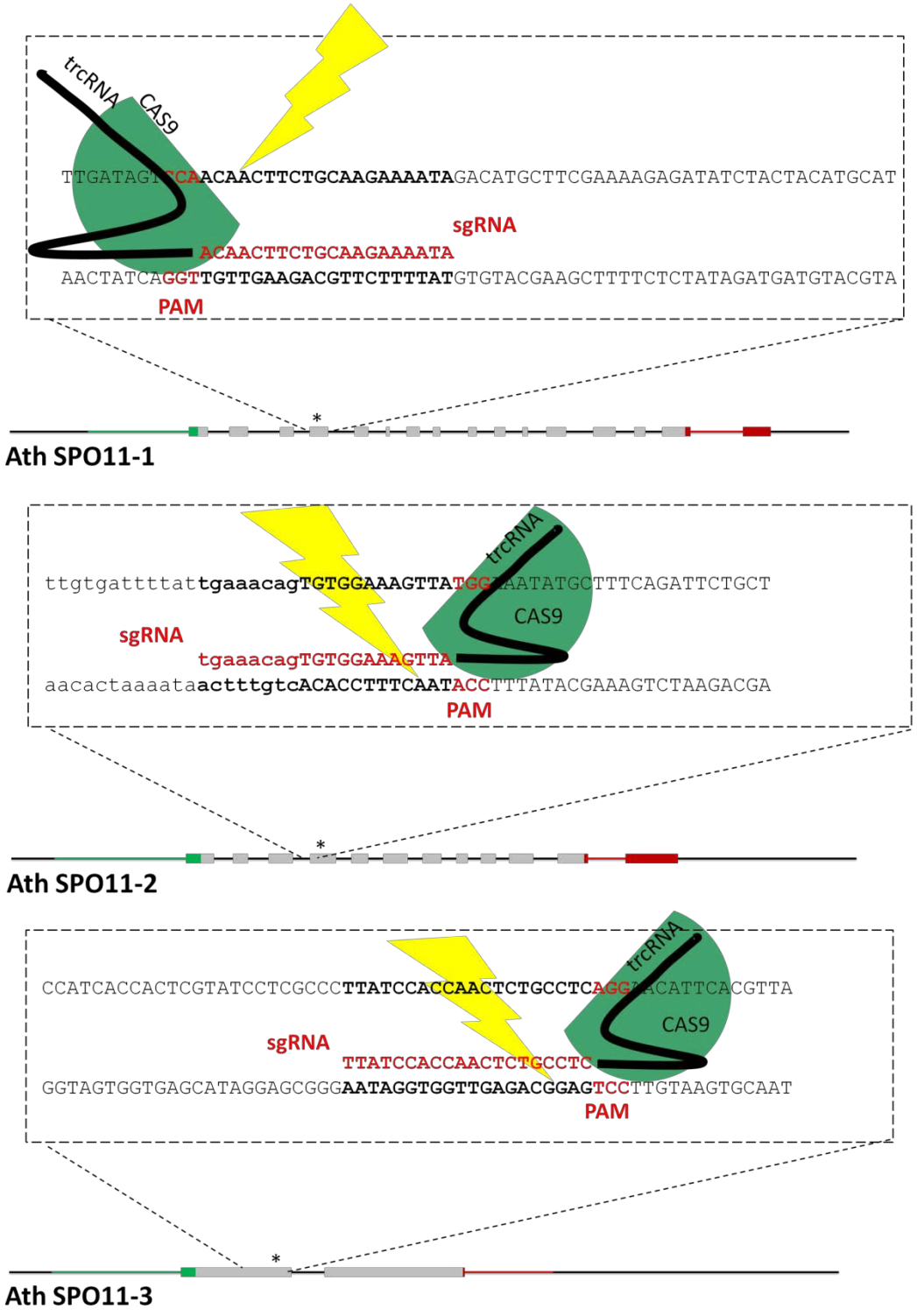


Figure 12. Mutation induction approach in *A. thaliana* SPO11 using clustered regularly interspaced short palindromic repeats/Cas9 (CRISPR/Cas9).

Scheme of the *A. thaliana SPO11* loci showing the binding site of the CRISPR/Cas9 construct with the single guide RNA including the designed gene specific protospacers (bold) and the corresponding protospacer adjacent motifs (red). Nucleotides located in introns are written in lowercase, the ones in exons in capital letters. The active tyrosine residue is indicated with an asterix. The flash is marking the expected site of break induction. sgRNA = single guide RNA; PAM = protospacer adjacent motif, trcRNA = trans activating CRISPR RNA.

To determine the presence of mutation events, induced by the CRISPR/Cas9 system, the progeny of primary transformants were used for high resolution melting (HRM) analysis (Gundry *et al.,* 2003). 20 plants per line and ten lines of each CRISPR approach have been analyzed so far. We could detect mutation in plants of each approach using HRM analysis (Figure 13). Most events could be identified for the SPO11-3 CRISPR approach. The HRM analysis is only providing information if a mutation was induced, but no information about the kind of mutation can be retrieved. For this purpose all lines showing deviations of more than 0.2°C in the HRM curves were further analyzed, by direct sequencing of the PCR reaction. The achieved chromatograms were analyzed and the presence of the second shifted chromatogram differing in intensity, beginning four bp upstream the protospacer adjacent motif (PAM), could be identified in all cases (Figure 13). The sequencing revealed multiple events in which the type of mutation cannot be identified *e.g*. for the plant no. six of SPO11-2 CRISPR line two (Figure 13b). Additionally, single mutations such as an insertion of an adenine in the plant no. 69 of SPO11-3 CRISPR line six were revealed (Figure 13c). The sequenced PCR fragments were further cloned using TA cloning and gained clones were sequenced afterwards to verify the type of mutation. This last step was done so far only for the SPO11-3 CRISPR approach.

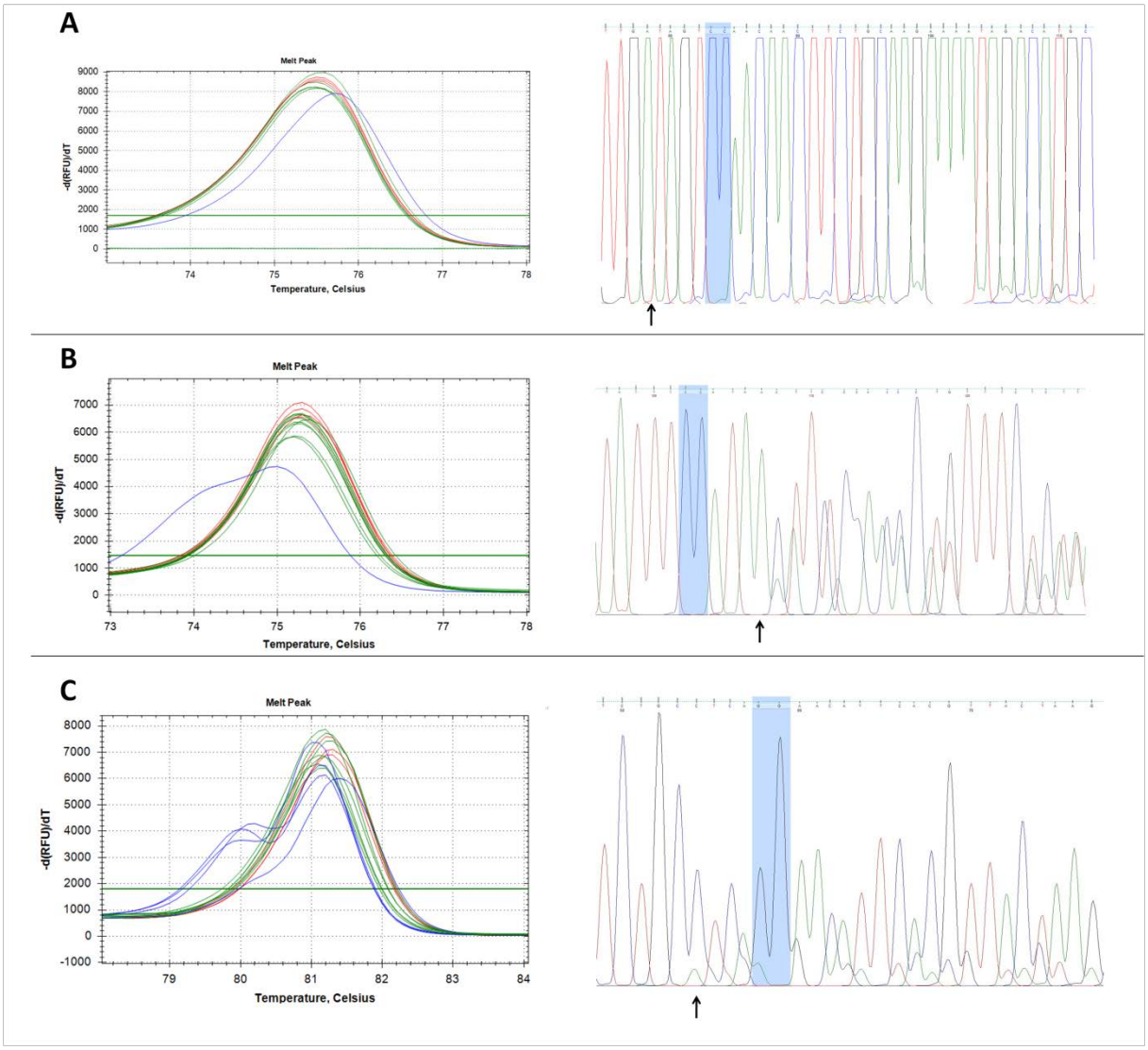


Figure 13. Mutation analyses by high resolution melting and sequencing in SPO11-CRISPR approaches.

High resolution melting analysis and chromatograms of T2 generation of the Cas9 expressing SPO11-1 (**A**), SPO11-2 (**B**), and SPO11-3 (**C**) CRISPR mutation approaches. On the left site melting curves of PCR fragments from wild type controls (red), plants not differing > 0.2°C from wild type control (green) and plants differing > 0.2°C from wild type control (blue) are depicted, showing either a shift in the melting temperature or a second melting peek. On the right site single chromatograms of PCR fragments with shifted meelting curves are shown. The presence of a second chromatogram four base pair upstream of the protospacer adjacent motif can be identified. This region is marked with an arrow.

Multiple mutations in the eight individual SPO11-3 CRISPR plants analyzed so far were detected (Figure 14). The majority of the detected mutations were single insertions of one bp, located four bp upstream of the PAM. In most cases an adenine was integrated, but

also integration of thymine and guanine was observed. Additionally, two deletion events were detected, a one bp and a six bp deletion. Surprisingly a large insertion at the break site, in which parts of the *SPO11-3* gene combined with DNA from unknown origin were integrated, was also identified. Some plants possessed only one type of mutation but in some plants multiple mutation events were identified (Figure 14). These events are most probably the result of independent DSB induction by Cas9 in different cells of a single plant.

Figure 14. A subset of Sanger sequencing results from the SPO11-3 CRISPR/Cas9 mutation approach.

Sequence of the wild type gene (WT) and unique mutated alleles obtained from mutant identification of plants from the SPO11-3 CRISPR approach. Protospacer binding site is written in bold, mutations are marked in red. The line and plant number is written on the left site. The type of the mutation is mentioned at the right site. The mutation was always identified four base pairs upstream of the protospacer adjacent motif. In some individuals multiple mutations were identified, *e.g.* line six plant no. 36.

In the future, analyses of the offspring, the T3 generation, will be done and plants carrying a heterozygous mutation will be identified as it was done for other targets (Fauser *et al.,* 2014). These lines can be further cultivated aiming to identify lines possessing one mutated allele but no T-DNA. These lines have to be further propagated to gain T-DNA free heterozygous mutation lines. The ideal case would be a large deletion or insertion in the gene which can be detected by PCR, such as the one found for plant 36 of line six in the SPO11-3 CRISPR mutation approach (Figure 14). With such a mutation it would be easy to identify plants with a heterozygous altered gene by PCR analysis. If only small mutations can be identified in the progeny it will be hard to distinguish between plants which are heterozygous altered and wild type plants.

Material and Methods:

Antibody production

Secondary structure prediction of both SPO11 paralogs in plants was performed using Jnet Version 3.0 (Cole *et al.,* 2008).

The production of the antibodies was performed by the group of Dr. Frank Rabenstein from the Institute for Epidemiology and Pathogen Diagnostics of the Julius Kuehn Institute, using rabbits from an undefined strain and mice from the BALB/c strain. Immunization of rabbits by IV injection of the corresponding peptide (GenScript USA Inc., NJ; USA) was induced by five injections of the peptide in a two day interval injecting two times 60 µg of peptide in 0,9 % sodium chloride solution, followed by two times 90 µg and one time 120 µg. Three blood samples were taken on a weekly base, starting three weeks after the first injection. IM immunization of rabbits was induced by injection of 400 µg peptide mixed one to one with Freund´s complete adjuvant, followed by two injections of 400 µg peptide with Freund´s incomplete adjuvant after three weeks each. Three blood samples were taken; the first was taken ten days after the last injection, followed by two blood samples taken on a weekly base.

Mice were immunized by subcutaneous injection of 100 µg peptide mixed one to one with Freund´s complete adjuvant, followed by one injection of 100 µg peptide mixed one to one with Freund´s incomplete adjuvant two weeks after the first injection. An additional injection was done one week later, three weeks after the first injection. Only one blood sample was taken one week after the last injection.

Western blot analysis

Proteins were isolated from *A. thaliana* flower tissue using the TCA (trichloroacetic acid) method from Wu and Wang, 1984.

Western blot analysis was performed by Janina Metje at the Max Plank Institute of Biophysical Chemistry in Goettingen.

Two samples of wild type *A. thaliana* flowers were used for western blot analysis. An aliquot of the samples were boiled for 10 min at 95°C to resolve possible protein complexes. Western blotting was performed using the semi-dry method (Towbin *et al.,* 1979). A Schagger gel was used for protein separation yielding in a better resolution of proteins < 80 kDa (Schagger and von Jagow, 1987). The gel and a nitrocellulose membrane were washed in water and afterwards in transfer buffer. Four filter papers were soaked in transfer buffer and a sandwich of two filter papers, gel, membrane, and two filter papers was prepared. Blotting was done at 25 V and 45 mA for 90 min. Subsequently, the membrane was washed with PBST and blocked for 20 min with 5 % (w/v) skim milk in PBST. Incubation with the SPO11-2 antibody diluted 1 to 500 in 5 % (w/v) skim milk in PBST was done over night at 4°C. After incubation the membrane was washed three times with PBST. The membrane was incubated with the second antibody (#170-6515, Bio-Rad Laboratories GmbH, Munich, Germany) diluted 1 : 3000 for 1 h at room temperature, washed three times with PBST and visualized with Western Lightning *Plus* ECL solution (Perkin Elmer; Baesweiler; Germany). Detection and analysis were performed using an Imageready LAS-1000 CCD camera (Fujifilm Europe GmbH; Düsseldorf; Germany) and AIDA software (Fujifilm Europe GmbH; Düsseldorf; Germany).

Immunolocalization studies

Immunostaining was performed as described in Sprink and Hartung 2015, using the following *A. thaliana* specific antibodies, provided by the group of Prof. Chris Franklin in Birmingham. Rat and rabbit anti ASY1 (1:1000, Armstrong *et al.,* 2002); rat and rabbit anti ZYP1 (1:500, Higgins *et al.,* 2005); rat and rabbit anti RAD51 (1:200; Sanchez Moran *et al.,* 2007), rat and rabbit anti DMC1 (1:200; Sanchez Moran *et al.,* 2007) and rabbit anti SPO11-1 (1:100; Sanchez Moran *et al.,* 2007). The Antibodies against ATM/ATR and γH2AX are not *A. thaliana* specific and can be purchased commercially, ATM/ATR (#2851S; Cell Signaling Technology Europe, B.V.;Leiden; The Netherlands) and γH2AX (#07-164; Merck Chemicals GmbH; Schwalbach; Germany). The rabbit and mouse antibody against SPO11-2 produced at the Julius Kuehn Institute was used in a 1 to 100 dilution.

Construction and evaluation of the TALEN constructs

TALENs were assembled followed the Golden Gate cloning protocol by Cermak *et al.,* 2011 and introduced in the binary vector pSW5, based on the binary vector which was used for the SPO11 complementation approaches (Sprink and Hartung, 2015). All components of the TALEN system were received from the group of Prof. Holger Puchta from the Karlsruher Institute of Technology. The target site of the SPO11-1 TALEN pair is harboring an EcoRV restriction enzyme site, the target site of the SPO11-2 TALEN pair a SacI site. After TALEN assembly transformation into wild type *A. thaliana* plants by floral dip method was performed (Clough and Bent, 1998). Screening for positive transformants was done as previously described (Sprink and Hartung, 2015). Mutations were identified by PCR amplification of DNA prior digested for 1 h at 37°C with the restriction enzyme present in the target site of the corresponding TALEN pair to reduce amplification of wild type sequence and enrich the amplification of mutated sequences. PCR fragments were cloned using TA cloning and resulting clones were sequenced afterwards (Sprink and Hartung, 2014).

Construction and evaluation of the CRISPR/Cas9 constructs

CRISPR/Cas9 constructs were assembled and introduced into *A. thaliana* wild type plants followed exactly the protocol by Fauser *et al.,* 2014. Oligonucleotides were ordered from Metabion (Metabion International AG; Planegg/Steinkirchen; Germany).

HRM analyses were performed using kappa HRM fast master mix (VWR International GmbH, Erlangen, Germany), in a final volume of 15 µl according to the instructions of the manufacturer, using the CFX96 cycler from Biorad (Bio-Rad Laboratories GmbH, Munich, Germany). After 50 cycles of PCR amplification a HRM curve was performed in which the temperature was increased in 0.1°C steps, each temperature was kept for 5 s and fluorescence was measured at each step. Every candidate which melting curve was varying more than 0.2°C from the control lines were further analyzed for induced mutations by sequencing, TA cloning and subsequent sequencing of single clones.

References:

Allers, T., and Lichten, M. (2001). Differential timing and control of noncrossover and crossover recombination during meiosis. Cell **106,** 47-57.

An, X.J., Deng, Z.Y., and Wang, T. (2011). OsSpo11-4, a Rice Homologue of the Archaeal TopVIA Protein, Mediates Double-Strand DNA Cleavage and Interacts with OsTopVIB. PLoS One **6,** 14.

Armstrong, S.J., Caryl, A.P., Jones, G.H., and Franklin, F.C. (2002). Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in Arabidopsis and Brassica. J Cell Sci **115,** 3645-3655.

Armstrong, S.J., Franklin, F.C.H., and Jones, G.H. (2003). A meiotic time-course for Arabidopsis thaliana. Sex. Plant Reprod. **16,** 141-149.

Bauknecht, M., and Kobbe, D. (2014). AtGEN1 and AtSEND1, Two Paralogs in Arabidopsis, Possess Holliday Junction Resolvase Activity. Plant Physiol. **166,** 202-216.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N.J., and Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. Current Opinion in Biotechnology **32,** 76-84.

Bellani, M.A., Boateng, K.A., McLeod, D., and Camerini-Otero, R.D. (2010). The Expression Profile of the Major Mouse SPO11 Isoforms Indicates that SPO11β Introduces Double Strand Breaks and Suggests that SPO11α Has an Additional Role in Prophase in both Spermatocytes and Oocytes. Mol. Cell. Biol. **30,** 4391-4403.

Bergerat, A., deMassy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., and Forterre, P. (1997). An atypical topoisomerase II from archaea with implications for meiotic recombination. Nature **386,** 414-417.

Bishop, D.K., and Zickler, D. (2004). Early decision: meiotic crossover interference prior to stable strand exchange and synapsis. Cell **117,** 9-15.

Bleuyard, J.Y., and White, C.I. (2004). The Arabidopsis homologue of Xrcc3 plays an essential role in meiosis. The EMBO journal **23,** 439-449.

Bleuyard, J.-Y., Gallego, M.E., and White, C.I. (2006). Recent advances in understanding of the DNA double-strand break repair machinery of plants. DNA repair **5,** 1-12.

Bleuyard, J.Y., Gallego, M.E., Savigny, F., and White, C.I. (2005). Differing requirements for the Arabidopsis Rad51 paralogs in meiosis and DNA repair. The Plant Journal **41,** 533-545.

Börner, G.V., Kleckner, N., and Hunter, N. (2004). Crossover/Noncrossover Differentiation, Synaptonemal Complex Formation, and Regulatory Surveillance at the Leptotene/Zygotene Transition of Meiosis. Cell **117,** 29-45.

Bray, C.M., and West, C.E. (2005). DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. New Phytol. **168,** 511-528.

Broderick, S., Rehmet, K., Concannon, C., and Nasheuer, H.-P. (2010). Eukaryotic single-stranded DNA binding proteins: central factors in genome stability. In Genome Stability and Human Diseases (Springer), pp. 143-163.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res.**,** gkr218.

Chen, K.L., and Gao, C.X. (2014). Targeted genome modification technologies and their applications in crop improvements. Plant Cell Reports **33,** 575-583.

Christian, M., Qi, Y., Zhang, Y., and Voytas, D.F. (2013). Targeted Mutagenesis of Arabidopsis thaliana Using Engineered TAL Effector Nucleases. G3: Genes|Genomes|Genetics **3,** 1697-1705.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal : for cell and molecular biology **16,** 735-743.

Colaiácovo, M.P., MacQueen, A.J., Martinez-Perez, E., McDonald, K., Adamo, A., La Volpe, A., and Villeneuve, A.M. (2003). Synaptonemal complex assembly in C. elegans is dispensable for loading strand-exchange proteins but critical for proper completion of recombination. Dev. Cell **5,** 463-474.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. Nucleic Acids Res. **36,** W197-W201.

Cole, F., Keeney, S., and Jasin, M. (2010). Evolutionary conservation of meiotic DSB proteins: more than just Spo11. Genes Dev. **24,** 1201-1207.

Czornak, K., Chughtai, S., and Chrzanowska, K.H. (2008). Mystery of DNA repair: the role of the MRN complex and ATM kinase in DNA damage repair. J Appl Genetics **49,** 383-396.

de los Santos, T., Hunter, N., Lee, C., Larkin, B., Loidl, J., and Hollingsworth, N.M. (2003). The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. Genetics **164,** 81-94.

de Massy, B. (2013). Initiation of Meiotic Recombination: How and Where? Conservation and Specificities Among Eukaryotes. Annual Review of Genetics **47,** 563-599.

De Muyt, A., Vezon, D., Gendrot, G., Gallois, J.-L., Stevens, R., and Grelon, M. (2007). AtPRD1 is required for meiotic double strand break formation in Arabidopsis thaliana. The EMBO Journal **26,** 4126-4137.

Dernburg, A.F., McDonald, K., Moulder, G., Barstead, R., Dresser, M., and Villeneuve, A.M. (1998). Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell **94,** 387-398.

Edlinger, B., and Schlogelhofer, P. (2011). Have a break: determinants of meiotic DNA double strand break (DSB) formation and processing in plants. J. Exp. Bot. **62,** 1545-1563.

Engebrecht, J., Voelkelmeiman, K., and Roeder, G.S. (1991). MEIOSIS-SPECIFIC RNA SPLICING IN YEAST. Cell **66,** 1257-1268.

Fanning, E., Klimovich, V., and Nager, A.R. (2006). A dynamic model for replication protein A (RPA) function in DNA processing pathways. Nucleic Acids Res. **34,** 4126-4137.

Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant Journal **79,** 348-359.

Garcia, V., Phelps, S.E., Gray, S., and Neale, M.J. (2011). Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. Nature **479,** 241-244.

Gerton, J.L., DeRisi, J., Shroff, R., Lichten, M., Brown, P.O., and Petes, T.D. (2000). Global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences **97,** 11383-11390.

Grelon, M., Vezon, D., Gendrot, G., and Pelletier, G. (2001). AtSPO11-1 is necessary for efficient meiotic recombination in plants. Embo J. **20,** 589-600.

Gundry, C.N., Vandersteen, J.G., Reed, G.H., Pryor, R.J., Chen, J., and Wittwer, C.T. (2003). Amplicon Melting Analysis with Labeled Primers: A Closed-Tube Method for Differentiating Homozygotes and Heterozygotes. Clin. Chem. **49,** 396-406.

Habu, T., Taki, T., West, A., Nishimune, Y., and Morita, T. (1996). The mouse and human homologs of DMC1, the yeast meiosis-specific homologous recombination gene, have a common unique form of exon-skipped transcript in meiosis. Nucleic Acids Res. **24,** 470-477.

Hartung, F., and Puchta, H. (2000). Molecular characterisation of two paralogous SPO11 homologues in Arabidopsis thaliana. Nucleic Acids Res. **28,** 1548-1554.

Hartung, F., Wurz-Wildersinn, R., Fuchs, J., Schubert, I., Suer, S., and Puchta, H. (2007a). The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in Arabidopsis. Plant Cell **19,** 3090-3099.

Hartung, F., Suer, S., and Puchta, H. (2007b). Two closely related RecQ helicases have antagonistic roles in homologous recombination and DNA repair in Arabidopsis thaliana. Proceedings of the National Academy of Sciences **104,** 18836-18841.

Hartung, F., Angelis, K.J., Meister, A., Schubert, I., Melzer, M., and Puchta, H. (2002). An archaebacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants. Curr. Biol. **12,** 1787-1791.

Higgins, J.D., Buckling, E.F., Franklin, F.C.H., and Jones, G.H. (2008). Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over. The Plant Journal **54,** 152-162.

Higgins, J.D., Sanchez-Moran, E., Armstrong, S.J., Jones, G.H., and Franklin, F.C.H. (2005). The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. Genes Dev. **19,** 2488-2500.

Jasin, M., and Rothstein, R. (2013). Repair of strand breaks by homologous recombination. Cold Spring Harbor perspectives in biology **5,** a012740.

Kauppi, L., Barchi, M., Baudat, F., Romanienko, P.J., Keeney, S., and Jasin, M. (2011). Distinct Properties of the XY Pseudoautosomal Region Crucial for Male Meiosis. Science **331,** 916-920.

Keeney, S., and Neale, M.J. (2006). Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation. Biochem. Soc. Trans. **34,** 523-525.

Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell **88,** 375-384.

Keeney, S., Lange, J., and Mohibullah, N. (2014). Self-Organization of Meiotic Recombination Initiation: General Principles and Molecular Pathways. Annual Review of Genetics **48,** 187-214.

Keeney, S., Baudat, F., Angeles, M., Zhou, Z.-H., Copeland, N.G., Jenkins, N.A., Manova, K., and Jasin, M. (1999). A Mouse Homolog of the< i> Saccharomyces cerevisiae</i> Meiotic Recombination DNA Transesterase Spo11p. Genomics **61,** 170-182.

Kim, H.J., Lee, H.J., Kim, H., Cho, S.W., and Kim, J.-S. (2009). Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. **19,** 1279-1288.

Knoll, A., and Puchta, H. (2011). The role of DNA helicases and their interaction partners in genome stability and meiotic recombination in plants. J. Exp. Bot. **62,** 1565-1579.

Knoll, A., Schropfer, S., and Puchta, H. (2014). The RTR complex as caretaker of genome stability and its unique meiotic function in plants. Front Plant Sci **5,** 33.

Kurzbauer, M.-T., Uanschou, C., Chen, D., and Schlögelhofer, P. (2012). The Recombinases DMC1 and RAD51 Are Functionally and Spatially Separated during Meiosis in Arabidopsis. The Plant Cell Online **24,** 2058-2070.

Lange, J., Pan, J., Cole, F., Thelen, M.P., Jasin, M., and Keeney, S. (2011). ATM controls meiotic double-strand-break formation. Nature **479,** 237-U117.

Li, W., Chen, C., Markmann-Mulisch, U., Timofejeva, L., Schmelzer, E., Ma, H., and Reiss, B. (2004). The Arabidopsis AtRAD51 gene is dispensable for vegetative development but required for meiosis. Proc. Natl. Acad. Sci. U. S. A. **101,** 10596-10601.

Liu, H., Jang, J.K., Kato, N., and McKim, K.S. (2002). mei-P22 encodes a chromosome-associated protein required for the initiation of meiotic recombination in Drosophila melanogaster. Genetics **162,** 245-258.

Lorenz, A., West, S.C., and Whitby, M.C. (2010). The human Holliday junction resolvase GEN1 rescues the meiotic phenotype of a Schizosaccharomyces pombe mus81 mutant. Nucleic Acids Res. **38,** 1866-1873.

Mackey, Z.B., Ramos, W., Levin, D.S., Walter, C.A., McCarrey, J.R., and Tomkinson, A.E. (1997). An alternative splicing event which occurs in mouse pachytene spermatocytes generates a form of DNA ligase III with distinct biochemical properties that may function in meiotic recombination. Mol. Cell. Biol. **17,** 989-998.

Malik, S.B., Ramesh, M.A., Hulstrand, A.M., and Logsdon, J.M. (2007). Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. Mol. Biol. Evol. **24,** 2827-2841.

Manfrini, N., Guerini, I., Citterio, A., Lucchini, G., and Longhese, M.P. (2010). Processing of meiotic DNA double strand breaks requires cyclin-dependent kinase and multiple nucleases. Journal of Biological Chemistry **285,** 11628-11637.

Mannuss, A., Dukowic-Schulze, S., Suer, S., Hartung, F., Pacher, M., and Puchta, H. (2010). RAD5A, RECQ4A, and MUS81 have specific functions in homologous recombination and define different pathways of DNA repair in Arabidopsis thaliana. The Plant Cell Online **22,** 3318-3330.

McKim, K.S., Green-Marroquin, B.L., Sekelsky, J.J., Chin, G., Steinberg, C., Khodosh, R., and Hawley, R.S. (1998). Meiotic synapsis in the absence of recombination. Science **279,** 876-878.

Milman, N., Higuchi, E., and Smith, G.R. (2009). Meiotic DNA Double-Strand Break Repair Requires Two Nucleases, MRN and Ctp1, To Produce a Single Size Class of Rec12 (Spo11)-Oligonucleotide Complexes. Mol. Cell. Biol. **29,** 5998-6005.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature **455,** 770-774.

Mimitou, E.P., and Symington, L.S. (2009). DNA end resection: many nucleases make light work. DNA repair **8,** 983-995.

Mimitou, E.P., and Symington, L.S. (2011). DNA end resection--unraveling the tail. DNA repair **10,** 344-348.

Moses, M.J. (1969). Structure and function of the synaptonemal complex. Genetics **61,** Suppl:41-51.

Nakagawa, T., and Ogawa, H. (1997). Involvement of the MRE2 gene of yeast in formation of meiosis-specific double-strand breaks and crossover recombination through RNA splicing. Genes Cells **2,** 65-79.

Nassif, N., Penney, J., Pal, S., Engels, W., and Gloor, G. (1994). Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol. Cell. Biol. **14,** 1613-1625.

Neale, M.J., Pan, J., and Keeney, S. (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. Nature **436,** 1053-1057.

Osman, K., Sanchez-Moran, E., Mann, S.C., Jones, G.H., and Franklin, F.C.H. (2009). Replication protein A (AtRPA1a) is required for class I crossover formation but is dispensable for meiotic DNA break repair. Embo J. **28,** 394-404.

Osman, K., Higgins, J.D., Sanchez-Moran, E., Armstrong, S.J., and Franklin, F.C.H. (2011). Pathways to meiotic recombination in Arabidopsis thaliana. New Phytol. **190,** 523-544.

Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H.G., Tischfield, S.E., Zhu, X., Neale, M.J., Jasin, M., and Socci, N.D. (2011). A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. Cell **144,** 719-731.

Pâques, F., and Haber, J.E. (1999). Multiple Pathways of Recombination Induced by Double-Strand Breaks in Saccharomyces cerevisiae. Microbiology and Molecular Biology Reviews **63,** 349-404.

Petes, T.D. (2001). Meiotic recombination hot spots and cold spots. Nature Reviews Genetics **2,** 360-369.

Petukhova, G., Sung, P., and Klein, H. (2000). Promotion of Rad51-dependent D-loop formation by yeast recombination factor Rdh54/Tid1. Genes Dev. **14,** 2206-2215.

Prieler, S., Penkner, A., Borde, V., and Klein, F. (2005). The control of Spo11's interaction with meiotic recombination hotspots. Genes Dev. **19,** 255-269.

Roeder, G.S. (1997). Meiotic chromosomes: it takes two to tango. Genes Dev. **11,** 2600-2621.

Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. Journal of biological chemistry **273,** 5858-5868.

Romanienko, P.J., and Camerini-Otero, R.D. (1999). Cloning, Characterization, and Localization of Mouse and Human SPO11. Genomics **61,** 156-169.

Romanienko, P.J., and Camerini-Otero, R.D. (2000). The Mouse Spo11 Gene Is Required for Meiotic Chromosome Synapsis. Mol. Cell **6,** 975-987.

Ross, K.J., Fransz, P., and Jones, G.H. (1996). A light microscopic atlas of meiosis in Arabidopsis thaliana. Chromosome Res. **4,** 507-516.

Rothenberg, M., Kohli, J., and Ludin, K. (2009). Ctp1 and the MRN-complex are required for endonucleolytic Rec12 removal with release of a single class of oligonucleotides in fission yeast. PLoS Genet **5,** e1000722.

Sanchez-Moran, E., Santos, J.L., Jones, G.H., and Franklin, F.C.H. (2007). ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. Genes Dev. **21,** 2220-2233.

Sanchez-Moran, E., Osman, K., Higgins, J.D., Pradillo, M., Cunado, N., Jones, G.H., and Franklin, F.C.H. (2008). ASY1 coordinates early events in the plant meiotic recombination pathway. Cytogenet. Genome Res. **120,** 302-312.

Schagger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Analytical biochemistry **166,** 368-379.

Schwacha, A., and Kleckner, N. (1995). Identification of double Holliday junctions as intermediates in meiotic recombination. Cell **83,** 783-791.

Shibata, T., Nishinaka, T., Mikawa, T., Aihara, H., Kurumizaka, H., Yokoyama, S., and Ito, Y. (2001). Homologous genetic recombination as an intrinsic dynamic property of a DNA structure induced by RecA/Rad51-family proteins: A possible advantage of DNA over RNA as genomic material. Proc. Natl. Acad. Sci. U. S. A. **98,** 8425-8432.

Shingu, Y., Mikawa, T., Onuma, M., Hirayama, T., and Shibata, T. (2010). A DNA-binding surface of SPO11-1, an Arabidopsis SPO11 orthologue required for normal meiosis. Febs J. **277,** 2360-2374.

Shingu, Y., Tokai, T., Agawa, Y., Toyota, K., Ahamed, S., Kawagishi-Kobayashi, M., Komatsu, A., Mikawa, T., Yamamoto, M.T., Wakasa, K., Shibata, T., and Kusano, K. (2012). The double-stranded break-forming activity of plant SPO11s and a novel rice SPO11 revealed by a Drosophila bioassay. BMC Mol. Biol. **13,** 16.

Simkova, K., Moreau, F., Pawlak, P., Vriet, C., Baruah, A., Alexandre, C., Hennig, L., Apel, K., and Laloi, C. (2012). Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. **109,** 16360-16365.

Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y., and Takeda, S. (2006). Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. DNA repair **5,** 1021-1029.

Sprink, T., and Hartung, F. (2014). The splicing fate of plant SPO11 genes. Frontiers in Plant Science **5**.

Sprink, T., Metje, J., and Hartung, F. (2015). Plant genome editing by novel tools: TALEN and other sequence specific nucleases. Current Opinion in Biotechnology **32,** 47-53.

Sprink, T., and Hartung, F. (2015). The meiotic functions of SPO11-1 and -2 in *Arabidopsis thaliana* are sequence and to certain extent species specific. Plant Cell. Under Review

Stacey, N.J., Kuromori, T., Azumi, Y., Roberts, G., Breuer, C., Wada, T., Maxwell, A., Roberts, K., and Sugimoto-Shirasu, K. (2006). Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. Plant J. **48,** 206-216.

Svendsen, J.M., and Harper, J.W. (2010). GEN1/Yen1 and the SLX4 complex: solutions to the problem of Holliday junction resolution. Genes Dev. **24,** 521-536.

Takata, M., Sasaki, M.S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A., and Takeda, S. (1998). Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. The EMBO journal **17,** 5497-5508.

Terzi, L.C., and Simpson, G.G. (2008). Regulation of Flowering Time by RNA Processing. In Nuclear pre-mRNA Processing in Plants, A.N. Reddy and M. Golovkin, eds (Springer Berlin Heidelberg), pp. 201-218.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proceedings of the National Academy of Sciences **76,** 4350-4354.

Town, C.D., Cheung, F., Maiti, R., Crabtree, J., Haas, B.J., Wortman, J.R., Hine, E.E., Althoff, R., Arbogast, T.S., Tallon, L.J., Vigouroux, M., Trick, M., and Bancroft, I. (2006). Comparative genomics of Brassica oleracea and Arabidopsis thaliana reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell **18,** 1348-1359.

Uanschou, C., Siwiec, T., Pedrosa-Harand, A., Kerzendorfer, C., Sanchez-Moran, E., Novatchkova, M., Akimcheva, S., Woglar, A., Klein, F., and Schlögelhofer, P. (2007). A novel plant gene essential for meiosis is related to the human CtIP and the yeast COM1/SAE2 gene. The EMBO journal **26,** 5061-5070.

West, C.E., Waterworth, W.M., Jiang, Q., and Bray, C.M. (2000). Arabidopsis DNA ligase IV is induced by gamma-irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4. The Plant journal : for cell and molecular biology **24,** 67-78.

Williams, G.J., Lees-Miller, S.P., and Tainer, J.A. (2010). Mre11–Rad50–Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. DNA repair **9,** 1299-1306.

Wu, F.-S., and Wang, M.-Y. (1984). Extraction of proteins for sodium dodecyl sulfate-polyacrylamide gel electrophoresis from protease-rich plant tissues. Analytical biochemistry **139,** 100-103.

Wu, L., and Hickson, I.D. (2003). The Bloom's syndrome helicase suppresses crossing over during homologous recombination. Nature **426,** 870-874.

Youds, J.L., and Boulton, S.J. (2011). The choice in meiosis–defining the factors that influence crossover or non-crossover formation. Journal of cell science **124,** 501-513.

Zickler, D., and Kleckner, N. (1998). THE LEPTOTENE-ZYGOTENE TRANSITION OF MEIOSIS. Annual Review of Genetics **32,** 619-697.

Zickler, D., and Kleckner, N. (1999). Meiotic chromosomes: integrating structure and function. Annual review of genetics **33,** 603-754.

Acknowledgement:

I would like to thank everybody who made this work possible:

Especially I would like to thank my supervisor Dr. Frank Hartung for giving me the opportunity to do my PhD in his lab, for the trust in me and my work and all the helpful discussions.

Prof. Holger Puchta and his group at the KIT especially Friederich Fauser, Alexander Knoll and Simon Schiml for providing me with material for genome editing and for the many helpful advices.

Prof. Dr. Thomas Debener for supervising this work.

The group of Prof. Dr. Chris Franklin from the university of Birmingham, especially Dr. Kim Osman and Dr. Eugenio Sanchez Moran, for teaching me microscopic analyses of meiosis and providing me with meiotic specific antibodies.

Dr. Frank Rabenstein and his team for production of SPO11 specific antibodies.

Janina Metje for performing western blot analysis.

All the people in the lab helping me with this work especially Katrin Schulze.

The “deutsche Forschungsgemeinschaft DFG SPP1384“ for funding.

Appendix:

Sprink, T., and Hartung, F. (2014). The splicing fate of plant SPO11 genes. Frontiers in Plant Science **5**.

Sprink, T., Metje, J., and Hartung, F. (2015). Plant genome editing by novel tools: TALEN and other sequence specific nucleases. Current Opinion in Biotechnology **32,** 47-53.

Sprink, T., and Hartung, F. (2015). The meiotic functions of SPO11-1 and -2 in *Arabidopsis thaliana* are sequence and to a certain extent species specific.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 21 May 2014
doi: 10.3389/fpls.2014.00214

The splicing fate of plant SPO11 genes

Thorben Sprink* and Frank Hartung

Biosafety in Plant Biotechnology, Julius Kühn Institute, Quedlinburg, Germany

Edited by:
Changbin Chen, University of Minnesota, USA

Reviewed by:
Paul Fransz, University of Amsterdam, Netherlands
Joann Mudge, National Center for Genome Resources, USA

****Correspondence:***
Thorben Sprink, Biosafety in Plant Biotechnology, Julius Kuehn Institute, Erwin-Baur Str. 27, Quedlinburg 06484, Germany
e-mail: thorben.sprink@jki.bund.de

Toward the global understanding of plant meiosis, it seems to be essential to decipher why all as yet sequenced plants need or at least encode for two different meiotic *SPO11* genes. This is in contrast to mammals and fungi, where only one SPO11 is present. Both SPO11 in *Arabidopsis thaliana* are essential for the initiation of double strand breaks (DSBs) during the meiotic prophase. In nearly all eukaryotic organisms DSB induction during prophase I by SPO11 leads to meiotic DSB repair, thereby ensuring the formation of a necessary number of crossovers (CO) as physical connections between the homologous chromosomes. We aim to investigate the specific functions and evolution of both *SPO11* genes in land plants. Therefore, we identified and cloned the respective orthologous genes from *Brassica rapa*, *Carica papaya*, *Oryza sativa,* and *Physcomitrella patens*. In parallel we determined the full length cDNA sequences of SPO11-1 and -2 from all of these plants by RT-PCR. During these experiments we observed that the analyzed plants exhibit a pattern of alternative splicing products of both SPO11 mRNAs. Such an aberrant splicing has previously been described for Arabidopsis and therefore seems to be conserved throughout evolution. Most of the splicing forms of *SPO11-1* and *-2* seem to be non-functional as they either showed intron retention (IR) or shortened exons. However, the positional distribution and number of alternative splicing events vary strongly between the different plants. The cDNAs showed in most cases premature termination codons (PTCs) due to frameshift. Nevertheless, in some cases we found alternatively spliced but functional cDNAs. These findings let us suggest that alternative splicing of SPO11 depends on the respective gene sequence and on the plant species. Therefore, this conserved mechanism might play a role concerning regulation of SPO11.

Keywords: SPO11, *Arabidopsis thaliana*, alternative splicing, meiosis, double strand breaks

INTRODUCTION

In most eukaryotic organisms the rearrangement of the parental alleles by homologous recombination during meiosis is one essential step leading to genetic diversity. Correct pairing and subsequent homologous recombination in prophase I ensure stability of the chromosome number on the one hand and variability in the developing cells due to crossover resolution resulting in exchange of genetic material between the homologous chromosomes on the other hand. One crucial aspect in the arrangement of the recombination progress is the initial formation of double strand breaks (DSBs) by SPO11. The eukaryotic SPO11, which shows homology to the archaeal Topoisomerase VIA subunit (TOPVIA), is one of the key factors mediating the formation of DSBs in a wide range of organisms (Bergerat et al., 1997; Keeney et al., 1997; Grelon et al., 2001). Without DSBs and their subsequent repair as crossovers there is no physical linkage between the homologous chromosomes and random chromosome distribution would appear (Cole et al., 2010). Like TOPVIA, SPO11 is able to cleave DNA via a 5′ phosphotyrosyl linkage thereby defining the acceptor sites of exchange between the parental genomes (Cole et al., 2010). In contrast to animals and fungi where a single SPO11 is sufficient for meiotic DSB formation, plants encode for at least two SPO11, referred to as SPO11-1 and -2, that are both essential in a functional protein form for DSB formation during meiosis (Keeney et al., 1997; Grelon et al., 2001; Hartung et al., 2007; Shingu et al., 2012). However, the mechanism by which two very different SPO11 proteins in plants induce DSBs specifically during meiosis is still unclear. Our long term aim is to investigate the specific functions, origin and evolution of each SPO11 in the plant kingdom. By analyzing complete genomic sequences of more than 40 plants, we were able to show that all as yet sequenced land plants encode for at least three *SPO11* genes. Two of them, AthSPO11-1 and -2 play a meiotic role. The third one, AthSPO11-3 together with TOPVIB, the second subunit of the topoisomerase, possesses essential functions during somatic development of plant cells but plays no role in meiosis (Hartung et al., 2002a, 2007; Stacey et al., 2006; Simkova et al., 2012).

The phylogenetic analyses of SPO11-1 and -2 in land plants and algae show very clearly that both genes are highly conserved and ancient in the lineage of plants but cannot be found in algae or protists in the same form. An analysis of a high number of available genomic and protein sequences of *SPO11* in virtually all kingdoms of life shows that at least one duplication of the original *SPO11* from archae must have occurred very early preceding the split of animals and plants (Malik et al., 2007; this work). In addition to this, the intron content and localization in the *SPO11* genes from different organisms shows ancestral conservation between animals, fungi, and plants but also dramatic

variations in protists and green algae (Hartung et al., 2002b; this work).

Early investigations of SPO11-1 expression in *Arabidopsis thaliana* exhibited an extensive pattern of alternative splicing, which we were now able to show also for SPO11-2 (Hartung and Puchta, 2000). Analyzing the expression in other plants we could identify various non-functional alternative splicing events for *SPO11-1* and *-2* in *Oryza sativa*, *Brassica rapa*, *Carica papaya,* and *Physcomitrella patens*. Additionally, we found putative functional forms of alternative spliced *SPO11-1* or *-2* for the first time in plants, namely in *B. rapa*, *C. papaya*, *O. sativa,* and *P. patens*. The fact that both SPO11 show such a diversified splicing pattern and that alternative splicing for both SPO11 is conserved between the different species indicates that SPO11 has an ancient complex transcriptional regulation mechanism, most probably involving the non-sense mediated decay pathway as described for other meiotic genes (Chiba and Green, 2009).

MATERIALS AND METHODS

ACCESSION NUMBERS

We sequenced the cDNA of *SPO11-1* and *SPO11-2* from *B. rapa*, *C. papaya*, and *P. patens*. The resulting sequences have been deposited in this order in the NCBI database under accession numbers KF841348, KF841349, KF841350, KF926859, KF926860, and KF926861.

PLANT MATERIAL AND GROWTH CONDITIONS

Arabidopsis (*Arabidopsis thaliana L.*) wild type plants (Col-0) were seeded on a 3:1 mixture of soil and vermiculite spiked with 4 g/l Plantacote (Wilhelm Haug GmbH und Co. KG, Ammerbuch, Germany) as fertilizer and 0, 4 g/l Exemptor (BAYER crop science, Langenfeld, Germany) as an preventive insecticide. Plants were kept under short day conditions (8-h light/16-h dark cycle at 18°C) for 3 weeks and then transferred to a green house under a long day regime (16-h light/8 h- dark at 20°C). Rice (*O. sativa subsp. Japonica*) plants were grown in the greenhouse under a long day regime as well as *B. rapa* var. fastplant. Papaya (*C. papaya L.*) trees were grown in a public tropical greenhouse on loamy soil. *P. patens* gametophores were kindly provided by Gertrud Wiedemann from the group of Ralf Reski (Freiburg, Germany) on solid media.

GENE COMPILATION AND SOURCE OF SEQUENCE DATA

A total of 42 *SPO11-1* and 39 *SPO11-2* sequences from land plants were extracted from different databases using the Arabidopsis and *O. sativa* orthologs as starting point. The databases used were: Phytozome (http://www.phytozome.net), JGI (http://www.jgi.doe.gov), Ensembl plants (http://plants.ensembl.org/index.html), Gramene (http://www.gramene.org/), CoGeBlast (http://genomevolution.org/r/5kv5), and NCBI (http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html). Models predicting not the full length cDNA but only a few assembled ESTs were manually curated by aligning these sequences to annotated *SPO11-1* and *-2* of *A. thaliana* as well as *O. sativa* using MegAlign (DNASTAR Inc. Madison, WI, USA). For some species the ESTs and the cDNA prediction did not cover the whole sequence. In these cases, the corresponding genomic DNA region was screened for possible matches and manually added to the model if possible. To check the accuracy of our prediction, elected coding sequences (CDS) were amplified using Primers covering the whole predicted CDS (Supplemental Table 1). The sequence of each gene was checked by sequencing, using the Sanger method (GATC Biotech AG, Konstanz, Germany). All sequences used for phylogenetic comparisons and their accession codes are listed in Supplemental Tables 2, 3.

RNA ISOLATION AND USED TISSUE

All kits used in this study were used following the instructions of the manufacturer. Total RNA was isolated using the Bio & Sell RNA mini Kit (Bio&Sell e.K., Feucht, Germany). To evaluate the abundance of SPO11 transcripts in generative tissue, fresh young flowers were used for RNA isolation. In the case of *C. papaya*, flowers were stored in RNAshield (Zymo research Europe GmbH, Freiburg, Germany) prior to RNA isolation. To check the abundance in vegetative tissue, leaf material was used. In the case of *C. papaya* no leaf material was available so fruit exocarp tissue was utilized instead. To check expression in *P. patens* 6-week old gametophores were used for RNA Isolation. Isolated RNA was treated with DNase I (Thermo Fisher Scientific, Germany). To check contamination with genomic DNA in the treated RNA, a PCR was performed with RNA as a template. No contamination was found in the RNA samples after DNase treatment (data not shown). cDNA was produced using an anchored oligo dT Primer with the Maxima H Minus Reverse Transcriptase Kit (Thermo Fisher Scientific, Germany) using 2–4 μg of total RNA as a template for the RT-reaction.

MOLECULAR CHARACTERIZATION OF SPO11

Reverse transcribed cDNA was used as a template for a PCR reaction using 50 amplification cycles. The resulting PCR products were purified using the GeneJET PCR purification Kit (Thermo Fischer Scientific, Germany) and cloned into the insTA-cloning vector system (Thermo Fischer Scientific, Germany). Resulting clones were screened in a colony PCR using M13 Primer. Clones differing in the size of their insert were sequenced and analyzed using MegAlign.

RESULTS

IDENTIFICATION OF SPO11 HOMOLOGOUS AMONG THE PLANT KINGDOM

The progress in sequencing and the growing amount of data input into the sequence databases provided us with a powerful tool for the identification of putative homologous proteins in a rapidly growing number of organisms by database searches using common bioinformatics tools such as BLAST-programs (TBLASTN = protein sequence search against the respective genomic sequence). By using known sequences of SPO11 from *A. thaliana* and *O. sativa* we were able to identify orthologs to SPO11-1 and -2 in all publicly available land plant genome assemblies sequenced to date. The identities of the orthologs to SPO11-1 from *A. thaliana* ranges between 95.9% for *Arabidopsis lyrata* to 45.4% for *P. patens*. The identities of the orthologs to SPO11-2 from *A. thaliana* is comparable to the identities found for SPO11-1. For *A. lyrata* the identity is 96.9%

and the least identity is found again for *P. patens* with 47.5% (Supplemental Tables 2, 3). In both cases, the monocotyledonous plants show approximately 10% less identity compared to the dicotyledonous plants representing the earlier split of mono- and dicots (Supplemental Tables 2, 3).

In our database analyses we found orthologs of SPO11-1 and -2 in all land plants with completely sequenced genomes. The conserved gene structure of *SPO11-1* in land plants contains 15 exons and 14 introns in the coding region. This structure has been verified earlier by sequencing of the cDNAs from *A. thaliana* and *O. sativa* (Hartung and Puchta, 2000; Jain et al., 2006). In a large number of cases, the annotation of these orthologs corresponded to the known cDNAs but in several cases the correspondence was incomplete. In virtually all of the latter cases we could perform a manual correction according to the known sequences. In the Asterid *Utricularia gibba* we found that intron number one was missing, clearly indicating an intron loss event in this species. In **Table 1** the predicted position and phase of the introns in relation to their deduced protein sequence is given. All plants with a completely sequenced genome possess *SPO11-2* and show a conserved gene structure concerning the position of the 10 introns in the coding region of *SPO11-2* (**Table 1**). However, we can identify three exceptions. Firstly, *Malus domestica, Prunus persica, Vitis viniferis, Fragaria vesca*, and *Eucalyptus grandis* all miss the first intron so it has most probably been lost in a common ancestor of these species. Secondly, in some rice species a loss event of intron two occurred, as this intron is missing only in *O. sativa* and *O. glaberrima*. This intron loss event must have occurred recently as the close relative *O. brachyantha* contains intron two. Thirdly, the plant *Aquilegia coerulea,* belonging to the *Ranunculaceae*, encodes for a *SPO11-2* gene which does not contain a single intron (Supplemental Figure 1). Most probably this *SPO11-2* gene is a reinserted copy of a fully spliced reverse transcribed mRNA, a mechanism which is also proposed to have resulted in the origin of *SPO11-3* (Hartung et al., 2002b).

Considering all this, it is very clear that SPO11-2 existed before the evolution of land plants that took place approximately 450 mya, exemplarily shown by the *SPO11-2* sequence (genomic and cDNA) of the moss *P. patens*, an extant member of one of the oldest land plant lines (Supplemental Figure 1). However, there is a recognizable gap of conservation considering a second or third *SPO11* gene in green algae and other algae that belong to the heterokontophyta or rhodophyta. All fully sequenced green algae contain a single *SPO11* gene that shows the highest sequence identity to SPO11-3 from land plants. In all of these algae, the second subunit TOPVIB is also present as has been shown earlier by Malik et al. (2007). This indicates that like land plants, algae most probably possess a functional complex of TOPVIA and B. A very interesting feature of the *SPO11-3* gene structure in green and other algae is that this gene possesses a high number of introns (14 in *Chlamydomonas reinhardtii*) that are not correlated to the introns found in plant *SPO11-1* or *-2*, whereas *SPO11-3* in land plants possesses only one intron (whose position is corresponding to intron no. 6 of *CreSPO11-3*) or none at all (Supplemental Figure 2).

Table 1 | Intron localization of *A. thaliana, H. sapiens,* and the SPO11 genes from the two fungi *C. cinerea* (Basidiomycota) and *C. grayi* (Ascomycota) with respect to their corresponding amino acid sequence positions.

Intron no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14			End
Ath pos. (aa)	18.3	51.6	76.3	110.3	135	140.6	164	176.3	192.3	212	222.6	258	298.3	319.6			362
IP SPO11-1	1[a]	2	1	1	0	2	0	1	1	0	2	0	1	2			
Hsa pos. (aa)	43.6	81.6	111.3	133.6	170		199	211.3		248	281.3	294	319.6	357			396
IP SPO11	2	2	1	2	0		0	1		0	1	0	2	0			
Cci pos. (aa)	68.6		99.3	121.6	140	158	187		215	239	261.6		312.6	331	344.6	376.6	401
IP SPO11	2		1	2	0	0*	0		0	0	2		2	0	2	2	
Cgr pos. (aa)			74.3	96.6	133		164		190	215							378
IP SPO11			1	2	0		0		0	0							
Ath pos. (aa)	28	56	99.3		145.3				175	218	249.6	270.6	296.3	339			383
IP-SPO11-2	0[b]	0[c]	1		1				0	0	2	2	1	0			

The numbering of introns was done with respect to the highest number of 14 introns in Arabidopsis SPO11-1. Gaps are included in the other lines to better visualize the conserved intron positions.

[a] *This intron has been lost in Utricularia gibba.*

[b] *This intron has been lost in Fragaria vesca, Malus domestica, Mimulus guttatus, Prunus persica, and Vitis vinifera.*

[c] *This intron has been lost in Oryza brachyantha and Oryza sativa.*

* *This intron number 5 of C. cinerea is in the same conserved position as intron number 5 of Arabidopsis SPO11-1 and H. sapiens but is preceded by a non-conserved intron position (no. 4).*

Color coding: Orange, intron position conserved at least since the split of the plant and animal kingdom, sometimes (8 and 12) lost later on in fungis; Yellow, intron position conserved between H. sapiens (as representative for animals) and two fungal divisions. Abbreviations: IP, Intron position; Ath, Arabidopsis thaliana; Cci, Coprinopsis cinerea; Cgr, Cladonia grayi; Hsa, Homo sapiens.

Malik et al. (2007) performed extensive phylogenetic analyses in which they described a second *SPO11* gene that can be found in chlorophyta (prasinophyceae), rhodophyta, and heterokontophyta and is by its sequence homology most related to *SPO11-2* of plants. However, a meiotic function of the gene has not been demonstrated for any of these organisms so far, and additionally, the gene structure is highly different compared to *SPO11-2* from land plants (Supplemental Figure 2). The *SPO11-2* similar genes of phylogenetically very different algae either possess no intron at all, or a much smaller number of introns in positions that are not correlated to the highly conserved positions found in all land plant *SPO11-2* orthologs (Supplemental Figure 2). Taking all data together, two very early duplications of the original *SPO11-3* (which is orthologous to *TOP6A* from archea) must have occurred, followed by a number of losses in different kingdoms.

This raises the question if SPO11-2 from algae is really orthologous to SPO11-2 from land plants. To address this question, we can use the method of comparison of intron positions which we already developed earlier (Hartung et al., 2002b). In brief, after the alignment of the protein sequences, each intron position is projected onto these sequences which can result in an intron located in between two coding triplets (phase 0) or interrupting a coding triplet after the first or second nucleotide (phase 1 and 2 which results in e.g., amino acid 18.3 or 18.6, respectively). Doing so for all genes, we can clearly see that six intron positions in *SPO11-1* are conserved throughout the animal and plant kingdom, spanning a time frame of almost one billion years (**Table 1**; Hartung et al., 2002b). These introns are number 3, 5, 7, 8, 10, and 12 with respect to the *AthSPO11-1* gene (**Table 1**). The ancient intron positions 8 and 12 most probably have been lost after the divergence of plants and animals/fungis in the fungi kingdom only. Furthermore, even one intron of *SPO11-2* (no. 6) is somehow conserved with respect to fungal *SPO11* which is a single copy *SPO11* (Hartung et al., 2002b). These conserved intron positions cannot be found in the second *SPO11* copy in algae or protists (Supplemental Figure 2). Considering this, we think that the second *SPO11* in protists and algae is an ortholog of plant *SPO11-2* due to its sequence conservation but a lot of changes concerning its gene structure have taken place during evolution (Malik et al., 2007; this work).

ANALYSIS OF SPO11 cDNAs

Based on the obtained database sequences, we designed primer pairs to amplify the whole coding sequence (CDS) of *SPO11-1* and *SPO11-2* from *B. rapa, C. papaya, O. sativa*, and *P. patens*. The predicted models fit the amplified CDSs in all cases. Using preamplified cDNA of the corresponding species, both *SPO11* could be amplified in their full length from *C. papaya*, *B. rapa*, and *A. thaliana*. From *P. patens* and *O. sativa* only *SPO11-1* could be amplified as a full length construct, for *SPO11-2* from *P. patens two* overlapping fragments were amplified, sequenced, and artificially put together afterwards. For *O. sativa* no full length construct of *SPO11-2* could be amplified due to high GC content in the 5′ region (GC > 80%). Every time we tried to evaluate *SPO11-2* all constructs were artificially modified due to a repetitive sequence in the 5′ region. Due to this artificial error *SPO11-2* from *O. sativa* was not further analyzed in detail. In this region the PCR leaped directly from one repetitive sequence to the next, resulting in constructs without a methionine that could not possibly be spliced in a natural way. The structures of the *SPO11-1* and *-2* genes are shown schematically in **Figure 1**. In all cases, *SPO11-1* consists of 15 exons and 14 introns. *SPO11-2* codes for 11 exons interrupted by 10 introns in all cases, except for *O. sativa* and *O. brychyantha* in which intron 2 has been lost. The CDS and protein length of each analyzed SPO11 is shown in **Table 2**.

Full length cDNAs were assembled from the RT-PCR data compared to the genomic sequences in the databases. Astonishingly, in our attempts to amplify the cDNA by RT-PCR for each gene we barely found one clearly distinguishable band. In most cases, more than one band accompanied with a smear was visible in the ethidium bromide stained gel (**Figure 2**). After cloning and sequencing of the PCR-products we were able to identify different alternatively spliced variants for both *SPO11* cDNAs.

PATTERN OF ALTERNATIVE SPLICED SPO11

In the course of analyzing the patterns of alternative splicing events for *SPO11*, different splicing events which lead to putative non-functional proteins could be detected (**Figure 3**). In most cases we found intron retention (IR) mostly leading to a premature termination codon (PTC) and an altered length of the putative proteins. In some cases exon skipping (ES) occurred and we also observed events with altered 5′ or 3′ splice sites (alt 5′ss or alt 3′ss) leading to shorter or longer exons which led to the integration of PTCs in most cases.

When comparing the patterns of alternative splicing events of *SPO11-1* in vegetative and generative tissue we could only detect very few events with a matching pattern in both tissue types (Supplemental Table 4). Furthermore, these patterns are also different between the analyzed plants. We found no conserved alternatively splicing events between two different plants in our analyses, indicating that the events are species and tissue specific.

Analyzing *A. thaliana SPO11-1* (**Figure 3A**), a total of eight alternative splicing events could be found (β-ι). From these, five events were IR (β-ζ), one alt 5′ss (θ), one alt 3′ss (η), and one alt 3′ss combined with IR (ι). All alternative splicing events resulted in altered putative truncated proteins varying from 69 amino acids (aa) to 324 aa in length instead of 362 aa (Supplemental Table 4). For *A. thaliana SPO11-2* (**Figure 3a**), six alternative splicing events could be observed (β-η), three IR events (β-δ), one alt 5′ss (ε), one alt 5′ss combined with IR (ζ), and one alt 3′ss combined with ES (η). Five forms result in PTC and putative truncated proteins ranging from 52 to 305 aa instead of 383 aa. One form missing exon 3 and parts of exon 4 does not contain a PTC and is leading to a putative functional protein of 303 aa (η) (Supplemental Table 4).

The analysis of *SPO11-1* alternative splicing events in *B. rapa* revealed five different forms (β-ζ), which consist of two IR (β,γ), two alt 3′ss (ε,ζ), and one combination of ES with IR (δ) (**Figure 3B**). Leading to one alternative splicing event without PTC where the protein length is shortened by 9 aa (ε). All other events lead to PTC and therefore the putative protein sequences

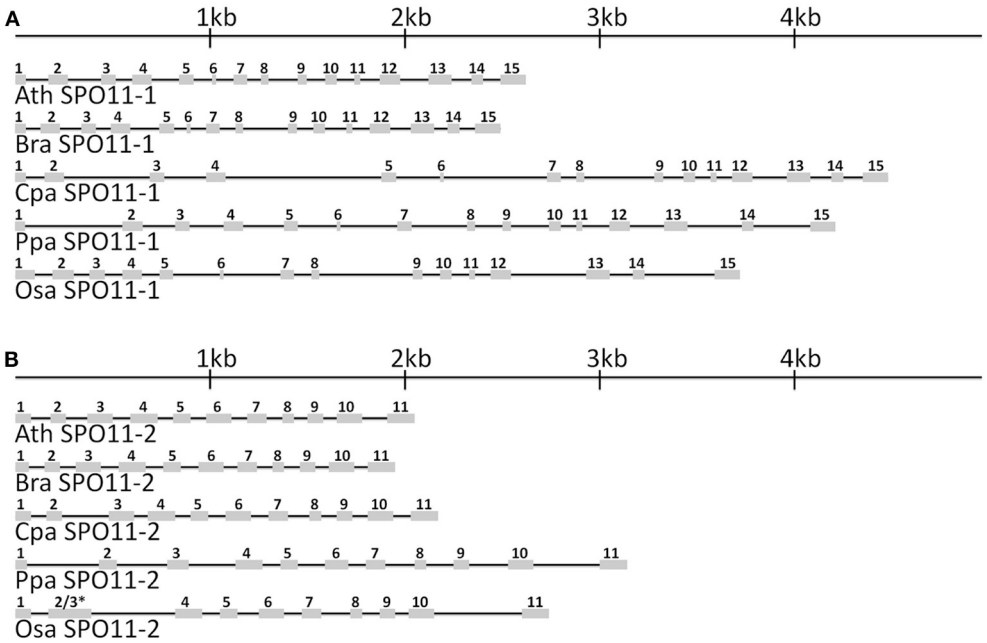


FIGURE 1 | The in-scale exon–intron organization of SPO11-1 (A) and SPO11-2 (B) for five analyzed species. Ath, *Arabidopsis thaliana*; Bra, *Brassica rapa*; Cpa, *Carica papaya*; Ppa, *Physcomitrella patens*; Osa, *Oryza sativa*. Coding regions are represented by gray boxes. The introns are represented by black lines. * Intron 2 has been lost in OsaSPO11-2. For a better comparison exon 2 was marked with 2 and 3 due to their fusion.

Table 2 | Length of the coding sequence and the respective deduced protein length of SPO11-1 and -2 from different species.

Organism	Gene	CDS length (bp)	Protein length (aa)
Arabidopsis thaliana	SPO11-1	1089	362
	SPO11-2	1152	383
Brassica rapa	SPO11-1	1089	362
	SPO11-2	1143	380
Carica papaya	SPO11-1	1086	361
	SPO11-2	1149	382
Oryza sativa	SPO11-1	1146	381
	SPO11-2	1158	385
Physcomitrella patens	SPO11-1	1086	361
	SPO11-2	1113	370

Abbreviations: bp, basepair; aa, amino acid.

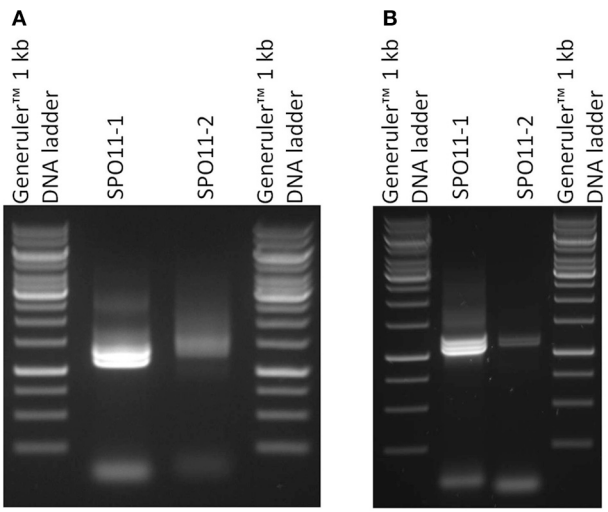


FIGURE 2 | Semiquantitative RT-PCR of SPO11-1 and -2 from *Arabidopsis thaliana* (A) and *Brassica rapa* (B). 1 μl of each cDNA was used for the PCR reaction. In the case of SPO11-1, distinct bands are visible. The lower band represents the α form of SPO11-1, the others are a mixture of other forms. The same holds true for SPO11-2.

were truncated ranging from 82 to 153 aa instead of 362 aa (Supplemental Table 4). In the case of *B. rapa SPO11-2*, *five* alternative splicing events were detected (β-ζ). All of them had one or more IR (**Figure 3b**), four of them with a PTC putatively leading to truncated proteins between 32 and 268 aa length. One IR event, the retention of intron 10 (δ), did not lead to a PTC resulting in an altered putative protein with 404 aa instead of 380 aa (Supplemental Table 4).

The evaluation of the alternative splicing events in *SPO11-1* from *C. papaya* revealed the highest number of 11 alternative splicing events (β-μ), all differing in type (**Figure 3C**). We found IR, ES, alt 5′ and 3′ss as well as all kinds of combinations between those types. All constructs contained a PTC leading to putative truncated proteins ranging from 30 to 210 aa in size, instead of 361 aa (Supplemental Table 4). When looking at *CpaSPO11-2*, five different alternative splicing events were detected (β-ζ). All had IR but also one combination of IR with an alt 3′ss was detected (ζ) (**Figure 3c**). Four events lead to PTC and putative proteins between 97 and 270 aa instead of 382 aa. One event could lead

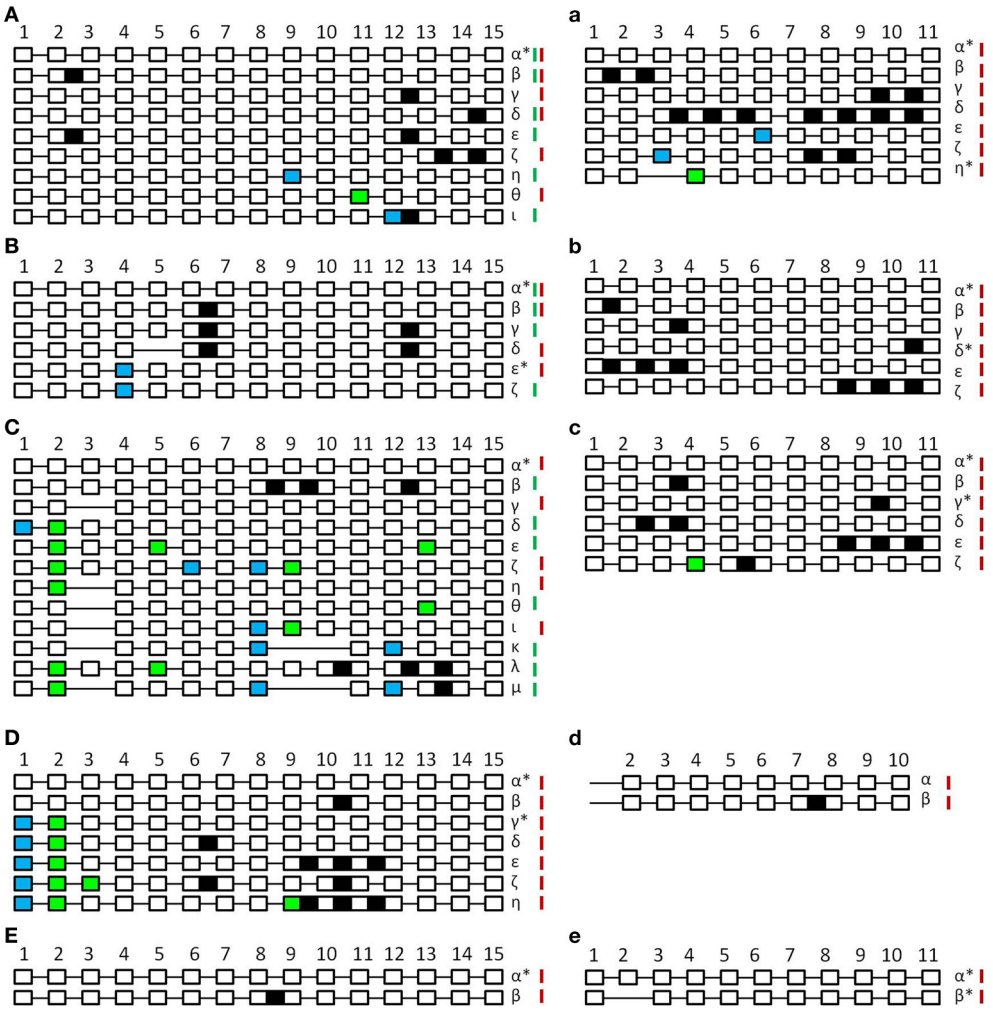


FIGURE 3 | Schematic unscaled schema of the different splice forms of SPO11-1 (A–E) and -2 (a–e) from *Arabidopsis thaliana* (A,a), *Brassica rapa* (B,b), *Carica papaya* (C,c), *Oryza sativa* (D,d), and *Physcomitrella patens* (E,e). Exons are numbered and shown as white blocks, spliced introns as black lines. Intron retention events are illustrated as black boxes, alternative 5′ splice site selection are shown as blue boxes and alterative 3′ splice site selection as light green boxes. In the case of exon skipping the corresponding white box is missing. Splicing forms are named in Greek letters. Splice forms found in generative tissue are marked with a red bar; splice forms found in vegetative tissue are marked with a green bar. Splice forms found in both tissues have both bars. Putative functional forms are marked with an asterisk. Due to high GC content and resulting PCR failure, amplification of OsaSPO11-2 was only possible from exon 2 so exon 1 is not indicated.

to an altered protein with 410 aa in length containing intron 9 (γ) (Supplemental Table 4).

In *O. sativa* we were only able to analyze the alternative splicing events for *SPO11-1*, due to the fact that *SPO11-2* has a very high GC content in the 5′ region of its genomic coding sequence. This high GC content prevented successful amplification of the cDNA up to exon 2. In the case of *SPO11-1* we identified six alternative splicing events (β-η). We found IR as well as a combination of alt 5′ and 3′ss with and without IR (**Figure 3D**). Five of these constructs lead to PTC resulting in altered putative protein lengths between 109 and 237 aa instead of 381 aa. One construct with a shortened exon 1 and 2 did not lead to a PTC (γ) and results in a truncated putative protein with the length of 350 aa (Supplemental Table 4). Despite the problems with PCR amplification, we identified one alternative splicing event (**Figure 3d**), containing intron 7 for *SPO11-2*.

Looking at *P. patens*, we could only find one alternative splicing event for each *SPO11* (**Figures 3E,e**). In *SPO11-1*, intron 8 was retained resulting in a PTC and a putative shortened protein of 181 aa instead of 361 aa (Supplemental Table 4). In *SPO11-2*, exon 2 was skipped without causing a PTC, but generating a putative truncated protein with a length of 342 aa instead of 372 aa (Supplemental Table 4).

The majority of alternative transcripts found in these experiments lead to putative non-functional proteins. Only a small number of alternative transcripts may lead to functional protein forms these transcripts were exclusively found in generative tissue and were outnumbered by the alternative transcripts which contained a PTC.

DISCUSSION

EVOLUTION OF DIFFERENT SPO11 GENES

The time frame of *SPO11* gene evolution remains unclear as a second *SPO11* copy must have arisen very early, most probably by gene duplication and subsequent divergence of the two genes. The most likely scenario is that SPO11-3, which shows by far the best sequence homology to TOPVIA from archaea and additionally is still functional and interacting with TOPVIB in plants, was the ancestor of gene duplications giving rise to other *SPO11* copies (Hartung et al., 2002a; Malik et al., 2007). The phylogenetic sequence homology of SPO11-2 to the second SPO11 found in protists shown by Malik et al. (2007) favors this gene as the first result of duplication and speciation. However, as we could show earlier and sustain here, *SPO11-1* from plants is clearly orthologous to *SPO11* from fungi and animals, indicating a very early appearance of this gene by duplication of *SPO11-3* (Hartung et al., 2002a; Forterre et al., 2007; this work). Therefore, in our opinion a duplication of the ancestral *SPO11-3* must have occurred twice and very early giving birth to *SPO11-1* and *-2* that currently we can find either in animals and fungi (*SPO11-1*) or algae and protists (*SPO11-2*). The organisms that currently only contain *SPO11-1* must have lost the other copies, whereas protists that contain *SPO11-2* and *-3* orthologs have lost only *SPO11-1* (**Figure 4**). Finally, in land plants all known copies of *SPO11* are still encoded and active as we and others have show for all three *SPO11* genes earlier (Grelon et al., 2001; Hartung et al., 2002a,b, 2007; Sugimoto-Shirasu et al., 2002; Stacey et al., 2006). In addition, *SPO11-3* is present together with the second subunit *TOPVIB*, not only in plants but also in all so far investigated green algae and protists, which is not the case in animals and fungi (Malik et al., 2007) (**Figure 4**). This points to a conserved and linked function of both gene products together as we and others have shown for Arabidopsis (Hartung et al., 2002b; Sugimoto-Shirasu et al., 2002).

Nevertheless, the exact evolution and function of two SPO11 in plant meiosis is still enigmatic. We show that both meiotically active *SPO11* genes are undergoing an extremely complicated splicing procedure leading to high numbers of mostly aberrant alternative splice products. Despite the very high conservation of the gene structure for *SPO11-1* and *-2*, whose introns are in virtually 100% identical positions throughout all land plants, the alternative splicing seems to be regulated specifically in each species. It is not clear whether all different splicing forms of *SPO11* found in this study are real alternative spliced transcripts or if some may result from sampling unprocessed pre-mRNAs or genomic DNA contamination. However, there are some clues that the identified alternative splicing patterns are real events. (1) The pattern is found for both *SPO11* in a similar rate and the same as described by Hartung and Puchta (2000), (2) the pattern is conserved between different species, (3) amplification of genomic DNA was not possible (Supplemental Figure 3A) and (4) of the analyzed meiotic genes, only *SPO11-1* and *-2*, *PHS1* and *VIP3* show this pattern (Supplemental Figure 3B). An alternative splicing pattern was described for *VIP3* and *SPO11-1* earlier (Hartung and Puchta, 2000; Terzi and Simpson, 2009). This study is slightly differing in the findings for *SPO11-2* from the study done by Hartung and Puchta (2000), due to the fact, that we now took a closer look especially on *SPO11-2* and used a different protocol for RT-PCR combined with a higher number of PCR cycles. The conservation of alternative splicing between orthologous genes has been described in *A. thaliana* and *O. sativa* (Wang and Brendel, 2006). For this reason, it is not extraordinary that the alternative splicing is conserved not only between *A. thaliana* and *O. sativa* but also between the other analyzed species. Wang and Brendel (2006) also reported that the type of alternative splicing is more conserved than the respective intron which is spliced, also seen for *SPO11-1* and *-2* in this study.

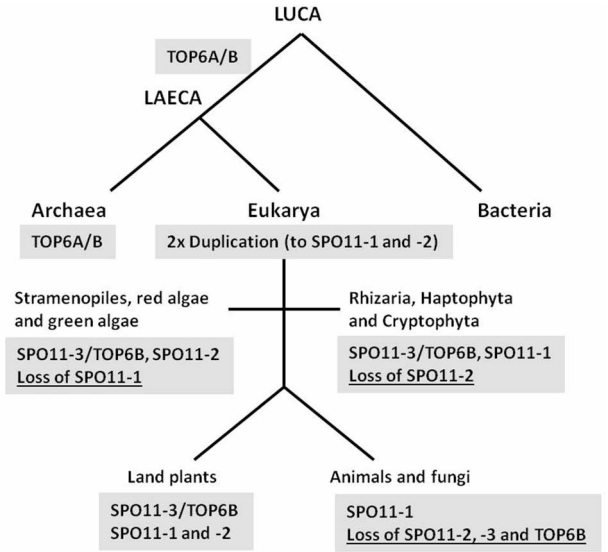


FIGURE 4 | Proposed evolution scheme of SPO11 by two duplications and different loss events. The proposed evolution of the three different SPO11 genes nowadays found in land plants is shown schematically. Whereas bacteria do not possess a topoisomerase 6, LAECA has developed a topoisomerase type 6 from which the subunit TOP6A is orthologous to SPO11-3 in eukaryotes. Two duplication events of SPO11-3 took place after separation of eukarya from archaea resulting in the additional SPO11-1 and SPO11-2 genes. In different phyla loss events of either SPO11-1 or -2 occurred. After separation of the animal and fungal kingdom SPO11-2 and -3 as well as TOP6B must have been lost resulting in the remaining single SPO11 gene present in these two kingdoms. Abbreviations: LUCA, last universal common ancestor; LAECA, last archaeal-eukaryal common ancestor. The term LAECA was taken from Forterre (2013).

Having a look at another kingdom in the eukaryotes previous studies showed also for mouse and human a pattern of alternative spliced transcripts for SPO11 (Shannon et al., 1999). In this previous work various alternative spliced transcripts were identified. Most of them were not further analyzed, but two transcripts variants with the expected size code for functional proteins. These two forms, SPO11-α and SPO11-β differ only in the abundance of exon 2. SPO11-α is missing exon 2 resulting in a shortened protein. The same forms were found in humans (Romanienko and Camerini-Otero, 1999). We were not able to find splicing forms equivalent to SPO11-alpha/beta from mammals due to the fact that the protein sequence in this area has not much homology to SPO11 from plants. But we were able to find other putative functional forms in plants as shown in **Figure 3**. The fact, that

alternative splicing of *SPO11* is also common in other kingdoms, let us suggest that this mechanism is highly conserved and might have a regulating function.

SPO11 AND THE NMD PATHWAY

Many aspects are known to initiate non-sense mediated decay in plants. It was shown that long 3′ untranslated regions (UTRs) as well as an intron in the 3′UTR can trigger the NMD pathway (Kertész et al., 2006). We could previously show that *A. thaliana SPO11-1* and *-2* both harbor an intron in the 3′UTR and show different poly A sites, which sometimes results in long 3′UTRs (Hartung and Puchta, 2000). In this study we determined various poly A sites of *SPO11* in *O. sativa* and *C. papaya* (data not shown) that affect the position of the poly A tail and sometimes lead to long 3′UTRs. Another aspect which may lead to non-sense mediated decay besides a long 3′UTR are upstream open reading frames (uORFs) adjacent to the start codon of the gene (Nyikó et al., 2009). Analyzing the 5′UTR of *A. thaliana SPO11-1* and *-2*, we could identify in both cases long uORFs. For other species such as *C. papaya* and *O. sativa*, such long and adjacent uORF could not be found for both *SPO11*. However, for all analyzed species we were able to identify alternative splicing events that lead to PTCs which are presumed to be targeted by the non-sense mediated decay pathway (for recent review see Reddy, 2007). In plants, many pathways such as the circadian clock and the flowering time are controlled via alternative splicing of core genes (James et al., 2012; Staiger and Brown, 2013). Alternative splicing and various polyadenylation has been reported for *VIP 3* during flower development of Arabidopsis (Terzi and Simpson, 2009). VIP 3 is the Arabidopsis ortholog of SKI 8 from yeast, one of the described direct interaction partners of SPO11 in *Saccharomyces cerevisiae* (Arora et al., 2004). There must be a reason for the conserved alternative splicing of *SPO11-1* and *-2* in plants. One possibility could be that *SPO11* is controlled in a precise way via the pathways of alternative splicing and non-sense mediated decay. The NMD pathway offer a mechanism which is routinely used by mammals and others to regulate gene expression (Lareau et al., 2004; Lejeune and Maquat, 2005). Such effects were observed for mice and men where the splicing of *SPO11* and other meiosis specific genes are regulated during meiosis (Habu et al., 1996; Schmid et al., 2013). It has long been known for yeast that genes which are involved in meiosis show alternative splicing (Engebrecht et al., 1991; Guisbert et al., 2012). Considering that the number of possible NMD candidates in plants are quite similar to the frequency observed for humans, it seems likely that plants may also use non-sense mediated decay and alternative splicing for gene regulation in a comparable way (Lareau et al., 2004; Wang and Brendel, 2006).

While further analyses on the localization of the alternative spliced isoforms need to be done, this study revealed differences in the alternative spliced forms of *SPO11-1* and *-2* between generative and vegetative tissue. Such tissue specific regulation of NMD was shown before. Especially in mammals this has been studied recently (Zetoune et al., 2008; Huang and Wilkinson, 2012) An accurate differentiation between single cell types could give closer insight into the alternative splicing during pre-meiotic and meiotic stages as done for yeast and mammals (Engebrecht et al., 1991; Schmid et al., 2013). The very weak expression especially for *SPO11-2* could make this a challenging task. Up to now little is known about the function of the conserved domains in *SPO11* (Bergerat et al., 1997). A closer look and more information on those domains could contribute to the understanding of the putative function of the alternative spliced isoforms. Investigating $nmd^{-/-}$ mutants could provide us with more information about the potential regulation of *SPO11-1* and *-2* via NMD in Arabidopsis. In previously published studies, *SPO11* mRNA was not captured mostly due to its weak expression and inadequate conditions for the amplification of *SPO11* (Simpson et al., 2008; Kalyna et al., 2012). Taking a closer look at *SPO11* expression in these plants would be of great advantage.

ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) SPP1384. We would like to thank Katrin Schulze for her support during the experiments and Verena Mecklenbrauck and Janina Metje for thorough reading of the manuscript. We also thank the group of Prof. Reski in Freiburg for providing us with *P. patens* gamethophores.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpls.2014.00214/abstract

REFERENCES

Arora, C., Kee, K., Maleki, S., and Keeney, S. (2004). Antiviral protein Ski8 is a direct partner of Spo11 in meiotic DNA break formation, independent of its cytoplasmic role in RNA metabolism. *Mol. Cell* 13, 549–559. doi: 10.1016/S1097-2765(04)00063-2

Bergerat, A., Demassy, B., Gadelle, D., Varoutas, P. C., Nicolas, A., and Forterre, P. (1997). An atypical topoisomerase II from archaea with implications for meiotic recombination. *Nature* 386, 414–417. doi: 10.1038/386414a0

Chiba, Y., and Green, P. (2009). mRNA degradation machinery in plants. *J. Plant Biol.* 52, 114–124. doi: 10.1007/s12374-009-9021-2

Cole, F., Keeney, S., and Jasin, M. (2010). Evolutionary conservation of meiotic DSB proteins: more than just Spo11. *Genes Dev.* 24, 1201–1207. doi: 10.1101/gad.1944710

Engebrecht, J., Voelkelmeiman, K., and Roeder, G. S. (1991). Meiosis-specific rna splicing in yeast. *Cell* 66, 1257–1268. doi: 10.1016/0092-8674(91)90047-3

Forterre, P. (2013). The common ancestor of archaea and eukarya was not an archaeon. *Archaea* 2013:372396. doi: 10.1155/2013/372396

Forterre, P., Gribaldo, S., Gadelle, D., and Serre, M. C. (2007). Origin and evolution of DNA topoisomerases. *Biochimie* 89, 427–446. doi: 10.1016/j.biochi.2006.12.009

Grelon, M., Vezon, D., Gendrot, G., and Pelletier, G. (2001). AtSPO11-1 is necessary for efficient meiotic recombination in plants. *EMBO J.* 20, 589–600. doi: 10.1093/emboj/20.3.589

Guisbert, K. S. K., Zhang, Y., Flatow, J., Hurtado, S., Staley, J. P., Lin, S., et al. (2012). Meiosis-induced alterations in transcript architecture and noncoding RNA expression in S. *cerevisiae*. *RNA* 18, 1142–1153. doi: 10.1261/rna.030510.111

Habu, T., Taki, T., West, A., Nishimune, Y., and Morita, T. (1996). The mouse and human homologs of DMC1, the yeast meiosis-specific homologous recombination gene, have a common unique form of exon-skipped transcript in meiosis. *Nucleic Acids Res.* 24, 470–477. doi: 10.1093/nar/24.3.470

Hartung, F., Angelis, K. J., Meister, A., Schubert, I., Melzer, M., and Puchta, H. (2002a). An archaebacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants. *Curr. Biol.* 12, 1787–1791. doi: 10.1016/S0960-9822(02)01218-6

Hartung, F., Blattner, F. R., and Puchta, H. (2002b). Intron gain and loss in the evolution of the conserved eukaryotic recombination machinery. *Nucleic Acids Res.* 30, 5175–5181. doi: 10.1093/nar/gkf649

Hartung, F., and Puchta, H. (2000). Molecular characterisation of two paralogous SPO11 homologues in *Arabidopsis thaliana*. *Nucleic Acids Res.* 28, 1548–1554. doi: 10.1093/nar/28.7.1548

Hartung, F., Wurz-Wildersinn, R., Fuchs, J., Schubert, I., Suer, S., and Puchta, H. (2007). The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in Arabidopsis. *Plant Cell* 19, 3090–3099. doi: 10.1105/tpc.107.054817

Huang, L. L., and Wilkinson, M. F. (2012). Regulation of nonsense-mediated mRNA decay. *Wiley Interdiscip. Rev. RNA* 3, 807–828. doi: 10.1002/wrna.1137

Jain, M., Tyagi, A. K., and Khurana, J. P. (2006). Overexpression of putative topoisomerase 6 genes from rice confers stress tolerance in transgenic Arabidopsis plants. *FEBS J.* 273, 5245–5260. doi: 10.1111/j.1742-4658.2006.05518.x

James, A. B., Syed, N. H., Bordage, S., Marshall, J., Nimmo, G. A., Jenkins, G. I., et al. (2012). Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. *Plant Cell* 24, 961–981. doi: 10.1105/tpc.111.093948

Kalyna, M., Simpson, C. G., Syed, N. H., Lewandowska, D., Marquez, Y., Kusenda, B., et al. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucleic Acids Res.* 40, 2454–2469. doi: 10.1093/nar/gkr932

Keeney, S., Giroux, C. N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* 88, 375–384. doi: 10.1016/S0092-8674(00)81876-0

Kertész, S., Kerenyi, Z., Merai, Z., Bartos, I., Palfy, T., Barta, E., et al. (2006). Both introns and long 3′-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. *Nucleic Acids Res.* 34, 6147–6157. doi: 10.1093/nar/gkl737

Lareau, L. F., Green, R. E., Bhatnagar, R. S., and Brenner, S. E. (2004). The evolving roles of alternative splicing. *Curr. Opin. Struct. Biol.* 14, 273–282. doi: 10.1016/j.sbi.2004.05.002

Lejeune, F., and Maquat, L. E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr. Opin. Cell Biol.* 17, 309–315. doi: 10.1016/j.ceb.2005.03.002

Malik, S. B., Ramesh, M. A., Hulstrand, A. M., and Logsdon, J. M. (2007). Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. *Mol. Biol. Evol.* 24, 2827–2841. doi: 10.1093/molbev/msm217

Nyikó, T., Sonkoly, B., Mérai, Z., Benkovics, A., and Silhavy, D. (2009). Plant upstream ORFs can trigger nonsense-mediated mRNA decay in a size-dependent manner. *Plant Mol. Biol.* 71, 367–378. doi: 10.1007/s11103-009-9528-4

Reddy, A. S. N. (2007). "Alternative splicing of pre-messenger RNAs in plants in the genomic era," in *Annual Review of Plant Biology* (Palo Alto: Annual Reviews), 267–294.

Romanienko, P. J., and Camerini-Otero, R. D. (1999). Cloning, characterization, and localization of mouse and human SPO11. *Genomics* 61, 156–169. doi: 10.1006/geno.1999.5955

Schmid, R., Grellscheid, S. N., Ehrmann, I., Dalgliesh, C., Danilenko, M., Paronetto, M. P., et al. (2013). The splicing landscape is globally reprogrammed during male meiosis. *Nucleic Acids Res.* 41, 10170–10184. doi: 10.1093/nar/gkt811

Shannon, M., Richardson, L., Christian, A., Handel, M. A., and Thelen, M. P. (1999). Differential gene expression of mammalian SPO11/TOP6A homologs during meiosis. *FEBS Lett.* 462, 329–334. doi: 10.1016/S0014-5793(99)01546-X

Shingu, Y., Tokai, T., Agawa, Y., Toyota, K., Ahamed, S., Kawagishi-Kobayashi, M., et al. (2012). The double-stranded break-forming activity of plant SPO11s and a novel rice SPO11 revealed by a Drosophila bioassay. *BMC Mol. Biol.* 13:1. doi: 10.1186/1471-2199-13-1

Simkova, K., Moreau, F., Pawlak, P., Vriet, C., Baruah, A., Alexandre, C., et al. (2012). Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16360–16365. doi: 10.1073/pnas.1202041109

Simpson, C. G., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., et al. (2008). Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts. *Plant J.* 53, 1035–1048. doi: 10.1111/j.1365-313X.2007.03392.x

Stacey, N. J., Kuromori, T., Azumi, Y., Roberts, G., Breuer, C., Wada, T., et al. (2006). Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J.* 48, 206–216. doi: 10.1111/j.1365-313X.2006.02867.x

Staiger, D., and Brown, J. W. S. (2013). Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell* 25, 3640–3656. doi: 10.1105/tpc.113.113803

Sugimoto-Shirasu, K., Stacey, N. J., Corsar, J., Roberts, K., and McCann, M. C. (2002). DNA topoisomerase VI is essential for endoreduplication in Arabidopsis. *Curr. Biol.* 12, 1782–1786. doi: 10.1016/S0960-9822(02)01198-3

Terzi, L. C., and Simpson, G. G. (2009). "Regulation of flowering time by RNA processing," in *Nuclear pre-mRNA Processing in Plants,* eds A. S. N. Reddy and M. Golovkin (Berlin; Heidelberg: Springer-Verlag), 201–219.

Wang, B.-B., and Brendel, V. (2006). Genomewide comparative analysis of alternative splicing in plants. *Proc. Natl. Acad. Sci.* 103, 7175–7180. doi: 10.1073/pnas.0602039103

Zetoune, A., Fontanière, S., Magnin, D., Anczuków, O., Buisson, M., Zhang, C., et al. (2008). Comparison of nonsense-mediated mRNA decay efficiency in various murine tissues. *BMC Genet.* 9:83. doi: 10.1186/1471-2156-9-83

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 February 2014; accepted: 30 April 2014; published online: 21 May 2014.
*Citation: Sprink T and Hartung F (2014) The splicing fate of plant SPO11 genes. Front. Plant Sci. **5**:214. doi: 10.3389/fpls.2014.00214*
This article was submitted to Plant Genetics and Genomics, a section of the journal Frontiers in Plant Science.

Plant genome editing by novel tools: TALEN and other sequence specific nucleases

- Thorben Sprink[1],
- Janina Metje[2],
- Frank Hartung[1],

- [1] Julius Kühn Institut, Institute for Biosafety in Plant Biotechnology, Erwin Baur-Str. 27, 06484 Quedlinburg, Germany
- [2] Max Plank Institute for Biophysical Chemistry, Research Group Autophagy, Am Fassberg 11, 37077 Göttingen, Germany

http://www.sciencedirect.com/science/article/pii/S0958166914001979

DOI: doi:10.1016/j.copbio.2014.11.010

“Species and sequence specificity of SPO11”

The meiotic functions of SPO11-1 and -2 in *Arabidopsis thaliana* are sequence and to a certain extent species specific.

Thorben Sprink* and Frank Hartung

Julius Kuehn Institute

Institute for Biosafety in Plant Biotechnology

Erwin Baur Strasse 27

06484 Quedlinburg

-Germany-

*Thorben.Sprink@jki.bund.de

+493946 47543

The estimated length of the printed paper will be 13.9 pages

Abstract:

Major findings in the field of plant meiosis have been achieved recently but important questions towards the understanding of meiosis in plants remain concealed. A key question is why plants need two very different meiotic SPO11 proteins and whether these two interact together. In *Arabidopsis thaliana* both meiotic SPO11 are essential in a functional form for double strand break induction. This stands in contrast to metazoa where a single SPO11 is present and sufficient for proper meiosis. We aim to investigate the specific function and evolution of both meiotic SPO11 paralogs in plants. By exchanging three not conserved parts as well as exchange and deletion of the last exon between Ath *SPO11-1* and *-2*, we were able to show a sequence specific function for both SPO11 proteins. By exchanging SPO11-1 and -2 in *A. thaliana* with related and ancestral SPO11 we additionally could show a species specific function of the respective SPO11. Complementation was possible using SPO11 from the close related (20 mya) species *Brassica rapa* and partial complementation with multiple integrated SPO11-1 cDNAs from papaya. These findings indicate that the function of each meiotic active SPO11 paralog is sequence specific and that the respective orthologs are species specific.

Introduction

In most eukaryotic organisms a functional meiosis is the crucial step that ensures overall genome stability on one hand and provides genetic diversity on the other (Roeder 1997; Zickler and Kleckner, 1998; Paques and Haber, 1999; Knoll and Puchta, 2011). Precise interaction of the meiosis specific protein machinery enables correct pairing of the homologous chromosomes including their physical connection via double Holliday Junctions. This is leading to at least one crossover per chromosome. Thereby it guarantees stability of the chromosome number whereas the subsequent resolution of the crossovers by homologous recombination provides genetic variability in the developing gametes. Initial formation of double strand breaks (DSBs) by SPO11 at the leptotene stage of early prophase I is a key aspect that is needed in most organisms for the following meiotic processes. SPO11 induced DSBs facilitate a physical linkage between the homologous chromosomes, without these DSBs and their subsequent repair no pairing occurs and the chromosomes are randomly distributed (Grelon et al., 2001; Stacey et al., 2006; Hartung et al., 2007; Cole et al., 2010; this work).

The eukaryotic SPO11 protein which is a homolog of the topoisomerase VI (TOPVI) subunit A is this key factor introducing DSBs in a wide range of organisms. The *SPO11* gene is quite conserved between various organisms as we and others could demonstrate earlier (Bergerat et al,. 1997; Keeney et al., 1997; Grelon et al.; 2001; Sprink and Hartung, 2014). SPO11 and TOPVIA share seven conserved motifs, including a winged helix and a TOPRIM domain (Malik et al., 2007; de Massy, 2013). Both proteins are able to cleave double stranded DNA constituting a 5´-phosphotyrosyl linkage. The resulting DSBs are defining the acceptor sites of exchange between the parental chromosomes (Keeney et al., 1997; Malik et al., 2007; Cole et al., 2010). Previous work mostly done in yeast and mouse demonstrates that SPO11 proteins are subsequently released bound to a short piece of DNA, the so called SPOligo. These SPOligos are released with a 2 bp long 5´-overhang from the break site. The release is performed by a combined action of several proteins such as the MRX-complex (MRE11, RAD50, XRS2) in *Saccharomyces cerevisiae* or the MRN-complex (MRE11, RAD50, NBS1) in other species including *A. thaliana*, in a combination with SAE2/COM1 (Alani et al., 1990; Cao et al., 1990; Liu et al., 1995; Nairz and Klein, 1997; Prinz et al., 1997; Buhler et al., 2001; Prieler et al., 2004; Neale et al., 2005; Mimitou and Symington, 2008; Rothenberg et al., 2009; Garcia et al., 2011). Following the removal of SPO11 from the break sites different specialized meiotic DNA repair proteins including RPA1, RAD51 and DMC1 in combination with several DNA repair factors mediate strand invasion, DNA elongation and the capture of a second DNA strand followed by subsequent repair and ligation of the break (reviewed in Edlinger and Schlögelhofer, 2011; de Massy, 2013).

In many organisms like mammals and fungi a single SPO11 is present and sufficient for meiotic DSB formation. In contrast to this, plants encode for at least three SPO11 proteins, from which two play a meiotic role in *A. thaliana* referred to as Ath SPO11-1 and Ath SPO11-2. Both are essential in a functional form for DSB formation during meiosis (Keeney et al., 1997; Grelon et al., 2001; Stacey et al., 2006; Hartung et al., 2007). The third one, Ath SPO11-3 possesses pivotal functions during somatic development of plant cells in combination with the second subunit of the topoisomerase VI (TOPVIB), but it has no function in meiosis (Hartung et al., 2002a; 2007; Stacey et al., 2006; Simkova et al., 2012). In rice (*Oryza sativa*) two additional SPO11 proteins have been identified from which one, Osa SPO11-4, is proposed to have a function in meiosis (An et al. 2011; Shingu et al., 2012).

The mechanism of meiotic DSB formation in plants by at least two different SPO11 proteins remains mysterious. To date it is still unclear whether and how they collaborate in meiosis

and which regions of the proteins are defining their specific functions. Therefore, we investigated the specific function, interaction and evolution of SPO11 in the plant kingdom. We and others were able to reveal the evolution of SPO11 in all kingdoms of life and could identify a widely conserved mechanisms of aberrant splicing for both *SPO11-1* and *-2* in numerous plants (Hartung and Puchta, 2000; Hartung et al., 2002b; Malik et al., 2007; Sprink and Hartung, 2014).

Here we addressed several questions concerning the function of the two meiotic SPO11 proteins in plants. First we wanted to evaluate if the function of orthologous *SPO11* genes is conserved between different related plants and if the function is still conserved in more ancestral *SPO11* genes from green algae and animals. By using both genomic DNA and complementary DNA (cDNA) for complementation approaches we were able to check if the aberrant splicing has any effect on the complementation efficiency. In a second approach we investigated which regions of Ath SPO11-1 and -2 are defining the different functions of both proteins by interchanging three less conserved regions between both paralogs. Last but not least we deleted and exchanged over cross the last exon of both genes, since we discovered earlier that disruption by a T-DNA in this part of the respective gene leads to a total loss of function indicating that it might be an essential part (Stacey et al., 2006; Hartung et al., 2007).

Results

The function of AthSPO11 is sequence specific.

As a control for a positive complementation approach we transformed *spo11-1-3* plants with a full genomic construct of *SPO11-1* from *A. thaliana* including 553 bp of the promoter region and 496 bp of the 3´-UTR region (*spo11-1-3*-Ath SPO1g). We did the same for *spo11-2-3*, as we used the full genomic region of Ath *SPO11-2* including 704 bp of the promoter region and 496 bp of the 3´-UTR region (*spo11-2-3*-Ath SPO2g). This architecture of the UTR regions was also used for all other complementation approaches. Like in a previous study done in 2007 most generated lines produced a similar number of seeds as the wild type control (Hartung et al., 2007). We analyzed eight independent *spo11-1-3* Ath SPO1g lines and eight independent *spo11-2-3* Ath SPO2g lines which were homozygous for the respective mutation and carried the wild type genomic complementation construct. From these lines six out of eight (*spo11-1-3*-Ath SPO1) and seven out of eight (*spo11-2-3*-Ath SPO2) were able to fully complement the sterile phenotype of the respective knockout mutant (table 1).

Table 1. Mean seed set of *A. thaliana* plants with an additional full genomic construct of the *A. thaliana* SPO11 (AthSPOg) and without (*SPO11-1 / SPO11-2*) containing none (+/+) one (+/-) or two (-/-) disrupted alleles of endogenous *SPO11*.

	SPO11-1	*SPO11-1*-Ath SPO1g	*SPO11-2*	*SPO11-2*-AthSPO2g
+/+	93.1 ± 9,1	91.2 ± 14.6	93.5 ± 9.2	98.1 ± 22.9
+/-	106.4 ± 9,2	97.4 ± 21.1	105.1 ± 9.8	101.3 ± 22.2
-/-	3.5±0,6	87.8^{***} ± 22.6	4.8 ± 0.8	97.3^{***} ± 10.2

Seed set is shown in % mean ± standard error of the mean (SEM) calculated against combined mean seed set of wild type and heterozygous plants. (Welch´s T-test, P <0.05). *** = highly significant, p-value < 0.001

Analysis of the meiotic stages in the complemented plants showed a distribution and pairing of the chromosomes as in the wild type control. The homologous chromosomes paired in pachytene stage and five bivalents were formed at the diplotene stage of prophase I (Figure 1).

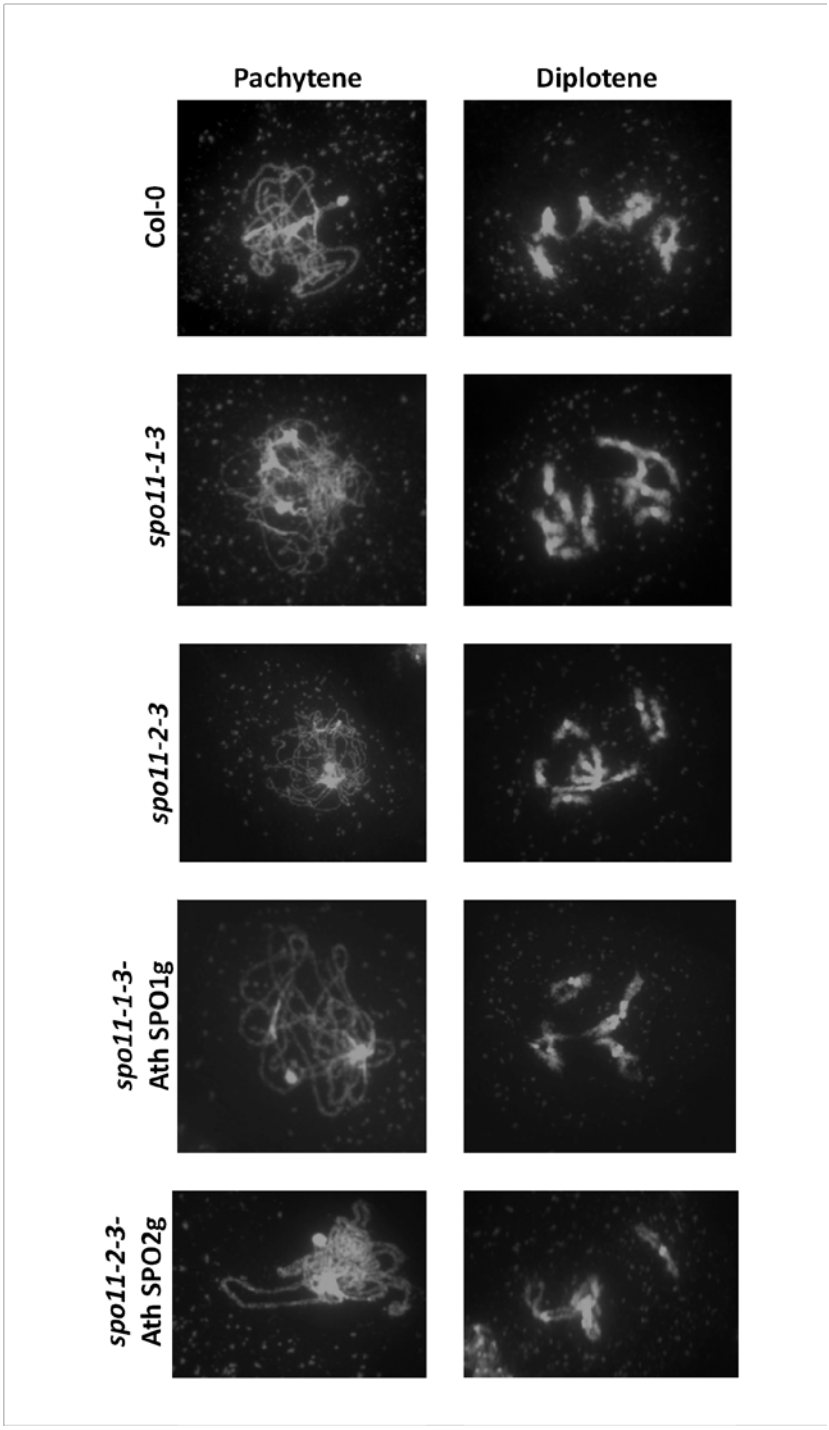


Figure1. Fluorescence micrograph of DAPI-stained nuclei.

Male meiotic chromosomes counterstained with DAPI during pachytene and diplotene stage of prophase I in wild type (Col-0) and *spo11-1-3* and *spo11-2-3* single mutants as well as *SPO11* single mutants transformed with endogeneous *SPO11* genomic DNA (Ath-SPOg) as indicated. Pairing of chromosomes during pachytene stage and formation of five bivalents could be observed in the case of wild type and the successful complemented mutants. The single knockout lines do not show any pairing or formation of bivalents instead ten univalents are formed.

Immunolocalization studies in spread preparations of *spo11-1-3*-Ath SPO11-1 and *spo11-2-3*-Ath SPO11-2 meiocytes revealed a restoration of RAD51 loading onto the DNA just like in wild type control (~150 foci / cell; n=10) (Figure 2).

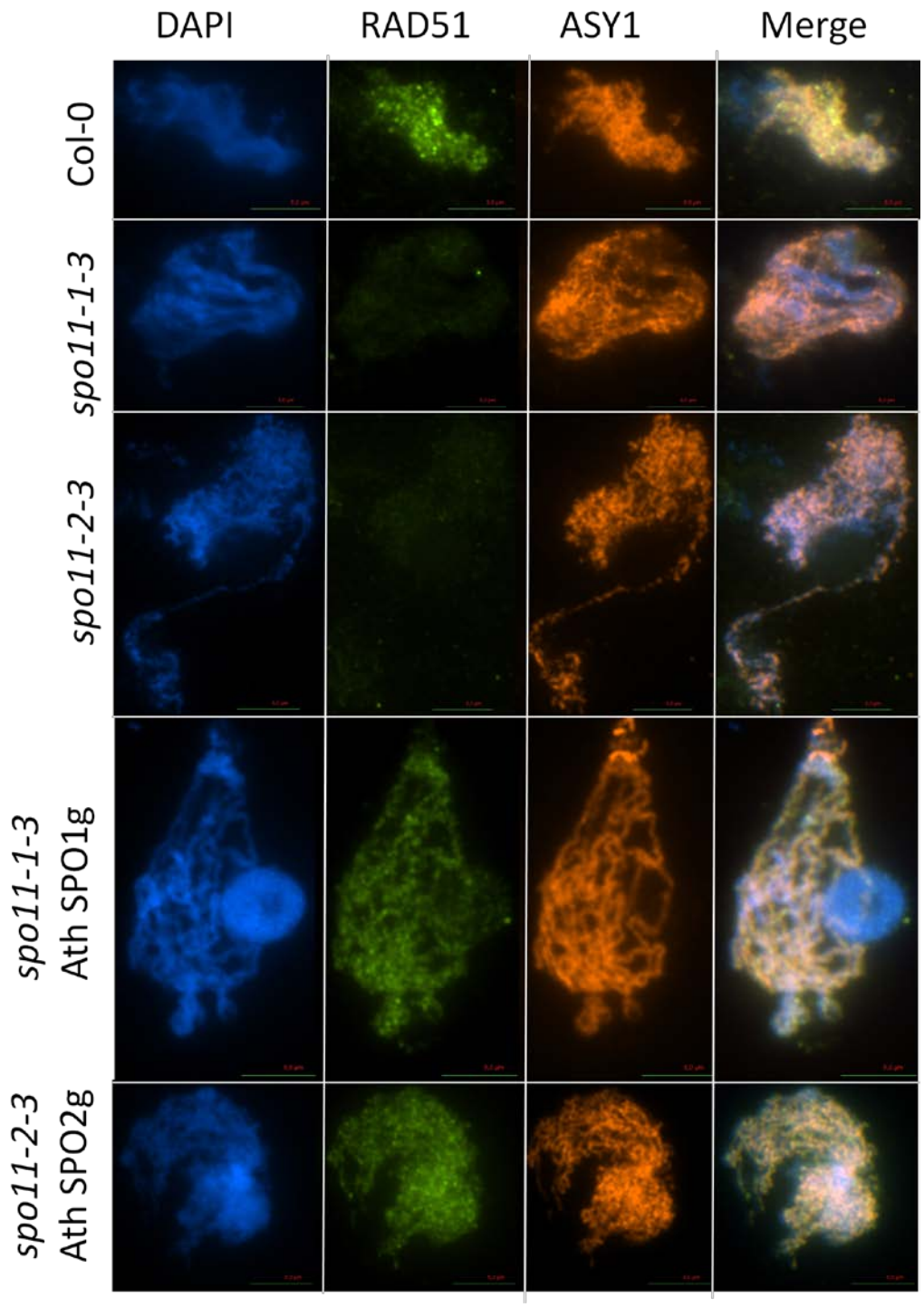


Figure 2. Dual immunolocalization of ASY1 and RAD51 proteins in meiocytes of different SPO11 mutant lines.

Meiocytes were counterstained with DAPI (blue) and an immunolocalization of ASY1 (orange) and RAD51 (green) was performed using rabbit polyclonal antibody against ASY1 and rat polyclonal antibody against RAD51. *A. thaliana* meiocytes in zygotene and pachytene stage of wild type, *SPO11* single knockout lines and complemented single knockouts were used for localization studies (as indicated). Green bars = 5µm.

This stands in contrast to *spo11-1-3* and *spo11-2-3* where RAD51 is present in the cytoplasma of the cells but is not loaded onto the DNA (Figure 2). The additional expression of a respective *SPO11* under its natural promoter in Col-0 wild type plants had no influence on the number of DSBs since the number of RAD51 foci remained the same as in the wild type control ~150 foci/cell (n=10) (Supplementary figure 1).To investigate the sequence specificity of Ath *SPO11* genes we created three constructs with interchanged parts between Ath *SPO11-1* and Ath *SPO11-2* (named SPO1swap1 to 3 and SPO2swap1 to 3). In all cases the endogenous promoter and 3´-UTR of the larger part of the respective chimeric gene was used. In the swapped regions both proteins showed less sequence identity between each other compared to the conserved parts of the proteins (Figure 3). All swaps were performed on the genomic level switching exon as well as intron regions (Supplementary figure 2).For the first construct we swapped the N-terminal part between Ath SPO11-1 and -2 by interchanging the first 81 amino acids (aa) of Ath SPO11-1 with the first 102 aa of Ath SPO11-2 and vice versa (SPO1swap1 and SPO2swap1) (Figure 3 and Supplemental figure 2).

Figure 3. Clustal W alignments of SPO11-1 and -2 from different plants of the order of Brassicales with indicated swaps between Ath SPO11-1 and -2.

Multiple alignment (done with Lasergene V. 12.1.0) of SPO11-1-and-2 from *A. thaliana* (Ath); *B. rapa* (Bra) and *C. papaya* (Cpa). Gaps are represented by dashes; conserved amino acids are shaded in yellow. Conserved motifs are indicated and shown as black boxes. The active Tyrosine within motif 1 is marked with an asterix. Interchanged regions between Ath SPO11-1 and Ath SPO11-2 are indicated and shown as red boxes. Swap 4 represents the last exon. The amino acid numbering of each protein is shown on the right.

We analyzed ten individual *spo11-1-3*-SPO1swap1 lines and four individual *spo11-2-3*-SPO2swap1 lines. None of them showed any evidence for a successful complementation, all analyzed lines showed the same reduced seed set as the non transgenic control (Figure 4).

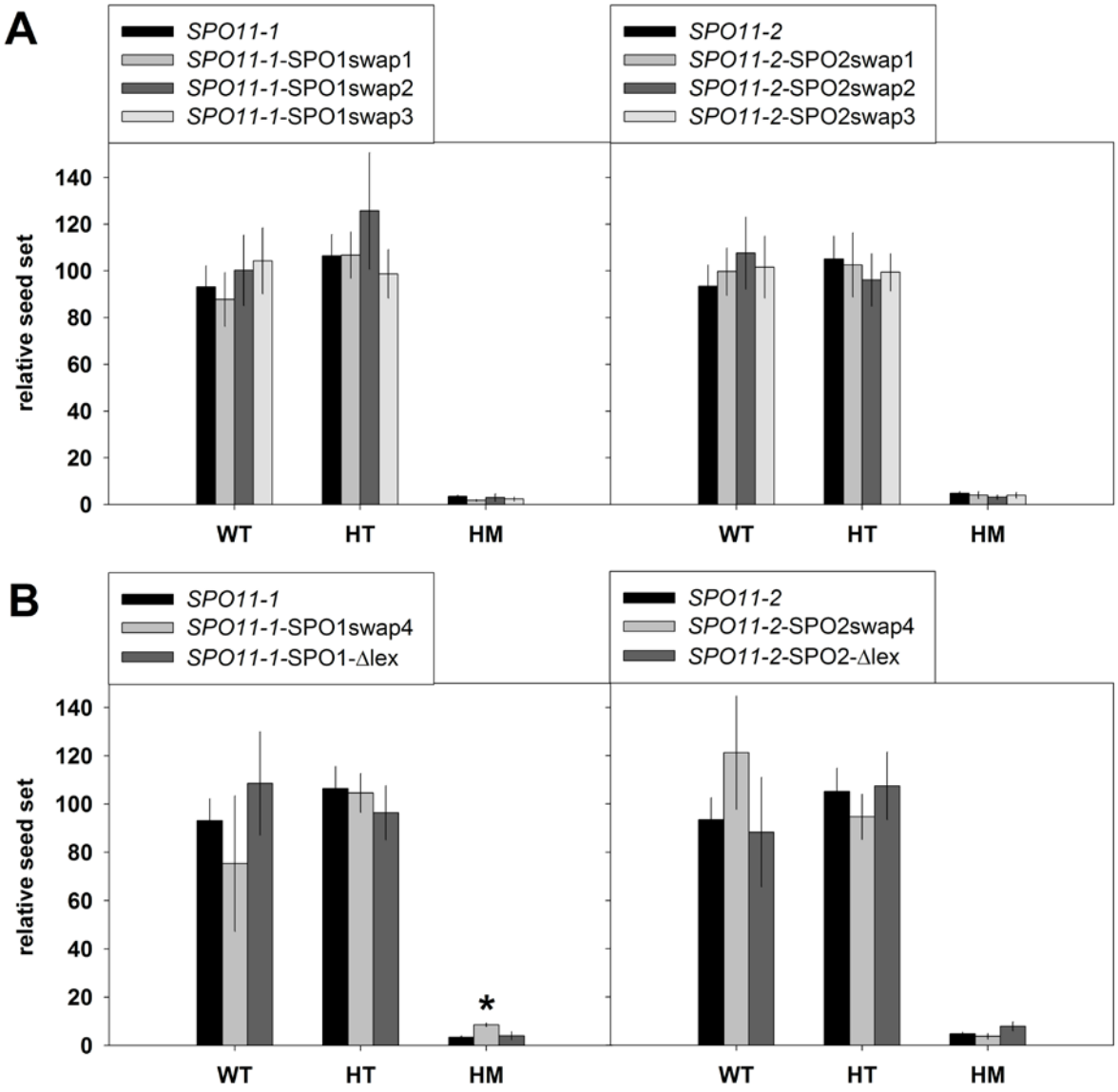


Figure 4. Seed set in complementation experiments using interchanged SPO11 constructs.

Relative seed set of *A. thaliana* plants containing an artificial SPO11 construct and none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11,* compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean (SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; P < 0,05) * = significant. **(A)** *SPO11*-swap1 to 3 plants **(B)** *SPO11*-swap4 and *SPO11* – last exon plants (*SPO11* –Δlex).

Additionally, we also had a look at these constructs when transformed into the opposite genetic background SPO1swap1 in *spo11-2-3* and SPO2swap1 in *spo11-1-3*. Six individual *spo11-1-3*-SPO2swap1 lines and six *spo11-2-3*-SPO1swap1 lines showed all the same result, none of the lines possessed an increased seed set (Supplemental figure 3). The expression of the swapped constructs had no influence on the seed set in the respective heterozygous or wild type plants (Figure 4 and supplementary figure 3).The same holds true for the second swap approach, in which we interchanged 21 aa from Ath SPO11-1 with 23 aa from Ath SPO11-2 and vice versa (SPO1swap2 and SPO2swap2). The interchanged region was located in the middle part of the proteins between motif two and motif three (Figure 4). We tested four individual lines of *spo11-1-3*-SPO1swap2 and six lines of *spo11-2-3*-SPO2swap2 all homozygous for the respective mutation. In none of these lines any induction of seed set could be observed also when transformed in the opposite genetic background (0/6 *spo11-1-3*-SPO2swap2 and 0/3 *spo11-2-3*-SPO1swap2) (Supplementary figure 3). Surprisingly, we could observe sterile siliques in some wild type plants and plants heterozygous for the T-DNA insertion in either *SPO11-1* or *-2* transformed with SPO1swap2. This effect ranged from only a few sterile siliques to nearly complete sterility (Supplemental figure 4). In all of these plants the expression of Ath *SPO11-1* and Ath *SPO11-2* as well as the expression of *SPO1swap2* were analyzed, a silencing of either *SPO11-1 or -2* could not be detected (Supplementary fig 4B). A similar effect of SPO2swap2 on either wild type or heterozygous plants could not be detected.

As third swap approach we interchanged 51 aa of Ath SPO11-1 and 54 aa of Ath SPO11-2 (SPO1swap3 and SPO2swap3) located near the C-terminal end of the proteins including the conserved motif six (Figure 3). We analyzed three individual *spo11-1-3*-SPO1swap3 and four individual *spo11-2-3*-SPO2swap3 lines, as in the other swap construct no change in fertility was observed (Figure 4). Also when transformed in the opposite genetic background both constructs did not lead to any change in fertility, neither in the respective mutant lines (0/3 *spo11-1-3*-SPO2swap3 and 0/2 *spo11-2-3*-SPO1swap3) nor in wild type and heterozygous plants (Supplementary figure 3).

By changing three non conserved parts between both SPO11 paralogs in *A. thaliana* we think that the function of SPO11 is sequences specific, because not any type of positive complementation was observed. If these parts would be functional in the background of the respective other protein, we should have gained DSBs and an increase in seed set, or at least fragmentation of the chromosomes, but none of this could be observed. In contrast to this we

have identified a negative interference in wild type plants for SPO1swap2 leading us to suppose that some swap constructs may bind to the DNA but are not able to cleave it and compete with the endogenous SPO11 proteins.

The last exon codes for a conserved but gene specific function

The C-terminal end seems to have an essential function for Ath SPO11-1 and Ath SPO11-2 since disruption of this part of the protein, namely the last exon, by T-DNA insertion is leading to a complete loss of function for Ath SPO11-2. Additionally, the last exon is quite conserved between both proteins since it is containing the conserved motif seven and ten additional conserved amino acids (Figure 3).

We wanted to elicit if the loss of function due the T-DNA insertion is caused by disrupting the overall structure of the genomic region or if the loss of the last exon alone has the same devastating effect on Ath SPO11-2 as well as on Ath SPO11-1. For this purpose we designed full genomic constructs for Ath *SPO11-1* and Ath *SPO11-2* lacking the last exon but keeping the endogenous promoter. An artificial stop codon (TAG) was introduced just after exon 14 of Ath *SPO11-1* and after exon ten of Ath *SPO11-2.* The gene specific 3´-UTR region was fused to this artificially truncated protein (SPO1-Δlex and SPO2-Δlex) (Supplemental figure 2). We had a look on eight individual *spo11-1-3*-SPO1-Δlex lines and three individual *spo11-2-3*-SPO2-Δlex lines which were homozygous for the respective knockout and were transformed with the truncated protein and none of these lines showed an increased number of seeds (Figure 4).

To address the question whether the function of the last exon is conserved and if it can be exchanged between Ath *SPO11-1* and Ath *SPO11-2* we designed a full genomic construct including the endogenous promoter region and 3´-UTR for both genes in which the last exon was exchanged (SPO1swap4 and SPO2swap4) (Figure 3). We analyzed two individual *spo11-1-3*-SPO1swap4 lines which were homozygous for the respective mutation. Both lines showed a slightly but significant induction of seed production compared to the control knockout line (Figure 4). This effect was not observed for SPO2swap4 where five individual lines were analyzed and none of these lines showed an increase in seed production (Figure 4). After transformation of the constructs in the opposite genetic background also no obvious change in fertility could be observed. When looking at the meiotic figures of *spo11-1-3*-SPO1swap4, we were not able to identify enhanced pairing of the chromosomes in pachytene stage. We neither could find any evidence for the formation of bivalents during diplotene

stage. The same holds true for the meiotic figures of *spo11-2-3*-SPO2swap4 (Supplemental figure 5).

The function of SPO11 is to a certain extend species specific.

To answer the question whether the function of SPO11 is conserved between plants which are related to a different extend we tried to complement the sterile phenotypes of *spo11-1-3* and *spo11-2-3* with genes from species which diverged more recently from *A. thaliana* ~20 million years ago (mya), earlier (~70 mya) and from a monocot species which diverged ~150-200 mya. Additionally, we wanted to elucidate if the function of SPO11 is also conserved between land plants and more ancestral genes from green algae or animals. An overview of the used genes and their sequence identity in respect to *A. thaliana* SPO11-1 and -2 is shown in table 2.

For the first heterologous complementation approach we used the full genomic sequence of *SPO11-1* and *-2* from *Brassica rapa* (*B. rapa)* (*spo11-1-3*-Bra SPO1g and *spo11-2-3*-Bra SPO2g) which is closely related to *A. thaliana* (~ 20 mya;) and shows the highest sequence identity (Town et al., 2006; Table 2).

Table 2. Pairwise comparison of Ath SPO11-1 and Ath SPO11-2 to SPO11 proteins from organism used for the complementation approaches.

Organism	Gene	gDNA (bp)	CDS (bp)	protein (aa)	Ath 1 (%)	Ath 2 (%)
Arabidopsis thaliana	*SPO11-1*	2633	1089	362	100	*30,5*
	SPO11-2	2029	1152	383	*30,5*	100
Brassica rapa	*SPO11-1*	2492	1089	362	**90,3**	30,1
	SPO11-2	1870	1143	380	31,8	**92,3**
Carica papaya	*SPO11-1*	4491	1086	361	72,8	32,2
	SPO11-2	2163	1149	382	30,5	73,8
Oryza sativa	*SPO11-1*	3710	1146	381	58,7	31,4
	SPO11-2	2710	1158	385	29,1	62,8
	SPO11-4	1572	1464	488	**28,9**	**25,1**
Chlamydomonas reinhardtii	*SPO11*	4920	1239	413	44,2	27,8
Mus musculus	*SPO11β*	15877	1188	396	34,1	29,7

One pair alignment of full SPO11 protein sequences from indicated species were performed against *A. thaliana* SPO11-1 and -2 using MegAlign (Lipman Pearson *Ktuple: 2, Gap penalty 4, Gap lenght penalty 12).* Highest and lowest identity of the orthologs is shown in bold and the identity between Ath SPO11-1 and -2 are shown in italics.

Just like in the control experiment we used the endogenous promoter and 3´-UTR region of the corresponding *A. thaliana* gene for all of the complementation approaches. We analyzed three individual *spo11-1-3*-Bra SPO1g lines and six individual *spo11-2-3*-Bra SPO2g lines. From these lines two out of three (*spo11-1-3*-BraSPO1g) and six out of six (*spo11-2-3*-BraSPO2g) were able to fully complement the sterile phenotype of the respective mutant, as seed set was restored just as in the Ath SPOg lines (Figure 5).

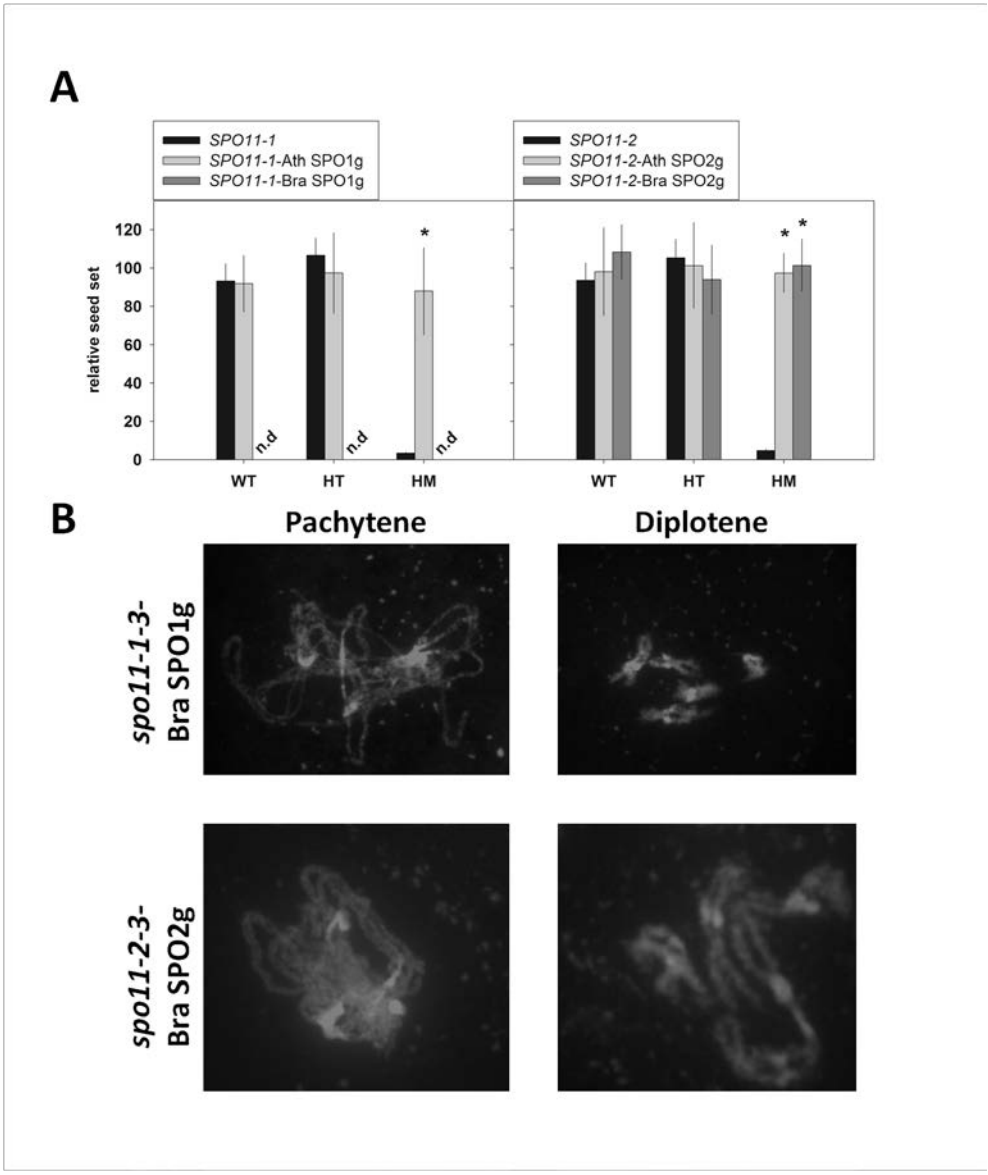


Figure 5. Successful complementation of spo11-1-3 and spo11-2-3 using genomic *SPO11* constructs from *A. thaliana* and *B. rapa.*

(A) Relative seed set of *A. thaliana* plants containing a full genomic construct of *A. thaliana* or *B. rapa* SPO11 and none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11*, compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean (SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; $P < 0,05$) * = significant. Seed set of *spo11-1-3* Bra SPO1g was not determined. **(B)** Fluorescence micrograph of DAPI-stained nuclei from *spo11-1-3* Bra SPO1g and *spo11-2-3* Bra SPO2g anthers, in pachytene and diplotene stage of prophase I. A reconstitution of pairing during pachytene stage as well as the formation of five bivalents could be observed for the successful complemented knockout lines.

The meiotic figures in pollen mother cells from *spo11-1-3*-Bra SPO1g and *spo11-2-3*-Bra SPO2g looked the same as the ones in Col-0 wild type plants. We observed pairing of the homologous chromosomes during pachytene stage as well as the formation of five bivalents during diplotene stage of prophase I (Figure 5). In the later stages of meiosis the distributions of chromosomes between the forming gametes were also just like in wild type control (Supplementary figure 6).

As a second heterologous complementation approach we used the full genomic sequence of *SPO11-1* and *SPO11-2* from papaya (*Carica papaya*) (*spo11-1-3*-Cpa SPO1g and *spo11-2-3*-Cpa SPO2g). *C. papaya* belongs like *A. thaliana* and *B. rapa* to the order of the Brassicales (Supplemental figure 7). Referring to Woodhouse et al. ancestors of *A. thaliana* and *C. papaya* diversified around 72 mya (Woodhouse et al., 2010). We analyzed nine individual *spo11-1-3*-Cpa SPO1g lines and 15 individual *spo11-2-3*-CpaSPO2g lines and none of these lines showed any signs of positive complementation. All of the analyzed lines had the same reduced seed set as their respective knockout control (Figure 6). The meiotic stages of DAPI stained spreads of *spo11-1-3*-Cpa SPO1g and *spo11-2-3*-CpaSPO2g pollen mother cells were also identical to the ones found in the respective knockout plants (Figure 6).

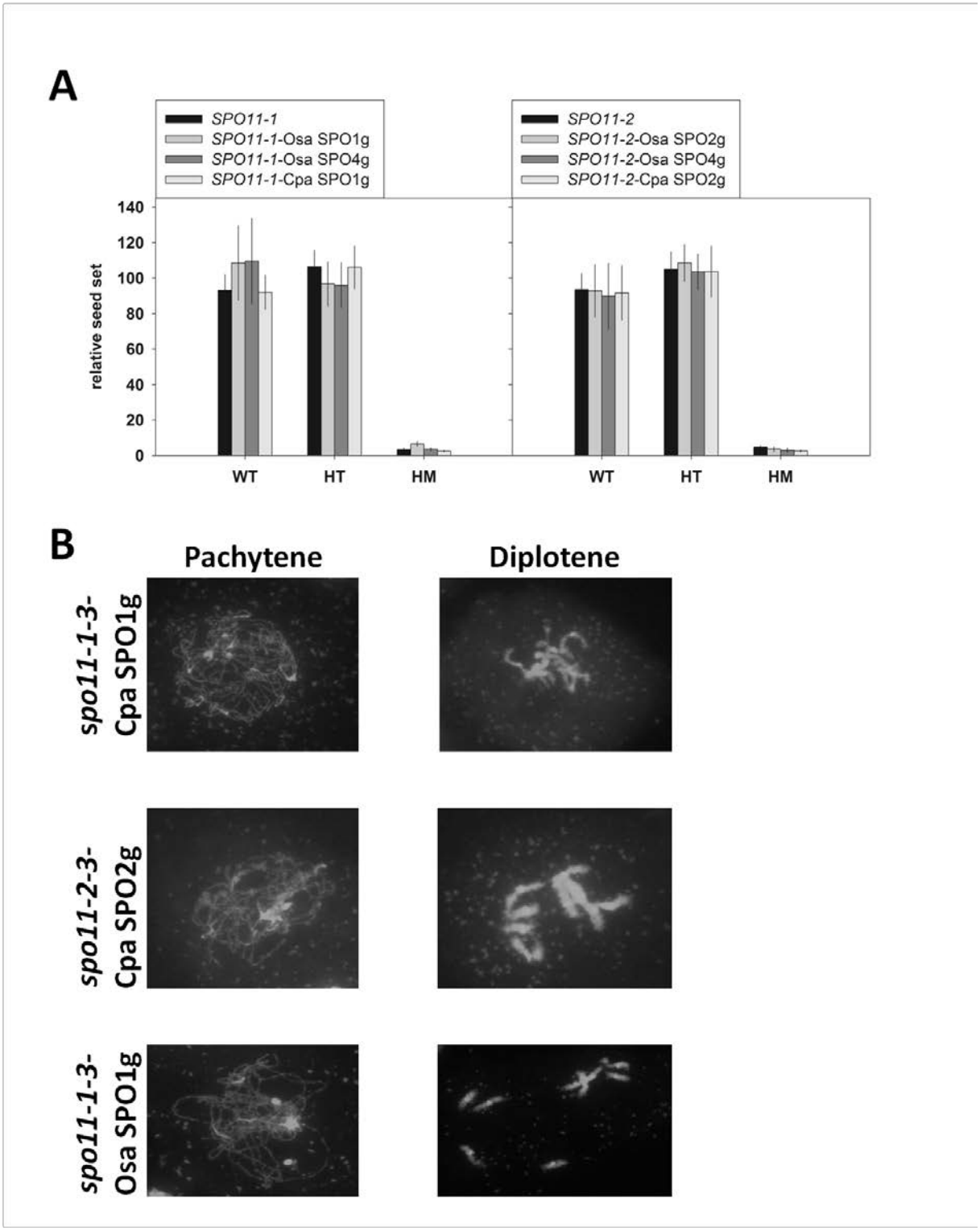


Figure 6. Complementation experiments of *spo11-1-3* and *spo11-2-3* using different genomic *SPO11* constructs from *O. sativa* and *C. papaya*.
(A) Relative seed set of *A. thaliana* plants containing a full genomic construct of *O. sativa* or *C .papaya* SPO11 and none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11*, compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean

(SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; $P < 0,05$).

(B) Fluorescence micrograph of DAPI-stained nuclei from *spo11-1-3*-Cpa SPO1g, *spo11-2-3*-Cpa SPO2g and *spo11-1-3*-Osa SPO1g anthers in pachytene and diplotene stage of prophase I. Neither reconstitution of pairing nor formation of bivalents could be observed for any of the complemented mutants.

In none of the analyzed lines we could identify pairing of the chromosomes during pachytene stage or formation of bivalents during diplotene stage, instead ten univalents were formed and randomly distributed just like it is known from the respective *A. thaliana* mutants.

For the third heterologous complementation approach we used the full genomic region of *SPO11-1* and *-2* from a monocot species, rice (*Oryza sativa*) subspecies japonica for this attempt (*spo11-1-3*-OsaSPO1g and *spo11-2-3*-OsaSPO2g). Monocots and dicots diverged at around 150 to 200 mya (Soltis et al., 2008) and also the sequence of SPO11 changed reasonably between monocots and dicots, 19 additional aa were added to SPO11-1 in monocots. (Table 2, Town et al., 2006, Sprink and Hartung, 2014). We analyzed seven independent *spo11-1-3*-Osa SPO1g lines and five independent *spo11-2-3*-OsaSPO2g lines and none of the lines were able to complement the sterile phenotype of the respective mutant (Figure 6). We also checked the meiotic stages of a *spo11-1-3*-OsaSPO1 line and could not identify any difference to the respective mutant (Figure 6).

In *O. sativa* an additional SPO11 gene has been described which is thought to be involved in meiosis, Osa *SPO11-4* (An et al., 2011). To investigate if Osa SPO11-4, which has all the conserved motifs but shows the lowest sequence identity to Ath SPO11 between all compared SPO11 proteins (Table 2), is able to complement the sterile phenotype of either *spo11-1-3* or *spo11-2-3* we designed two complementation constructs using the respective promoter and 3´-UTR region from either Ath *SPO11-1* or Ath *SPO11-2* (*spo11-1-3*-OsaSPO4g and *spo11-2-3*-Osa SPO4g). We created nine individual lines for *spo11-1-3*-Osa SPO4g and *spo11-2-3*-Osa SPO4g each. None of these showed an increased number of seed set. All lines had a seed set comparable to the respective knockout (Figure 6).

Complementation with ancestral SPO11

In contrasts to land plants and some protists mammals and green algae posses only a single *SPO11* gene therefore loss of the other ones during evolution is very likely (Sprink and Hartung, 2014). The sole SPO11 of green algae shows the highest sequence identity to SPO11-3 from land plants and the SPO11 from animals shows the highest sequence identity to SPO11-1 from land plants (Malik et al., 2007; Sprink and Hartung, 2014).

To investigate if the function of SPO11 is conserved between ancestral genes we tried to complement *spo11-1-3* and *spo11-2-3* plants with SPO11 from the green algae

Chlamydomonas reinhardtii (*spo11-1-3*-Cre SPO and *spo11-2-3*-Cre SPO) as well as SPO11 from mouse (Mus. musculus) (*spo11-1-3*-Mmu SPO and *spo11-2-3*-Mmu SPO). Land plants diverged from algae around 450 mya and the split between plants and animals ranges back to >1000 mya (Wang et al. 1999).

Due to huge differences in gene structure between Cre *SPO11* and Ath *SPO11* as well as many repetitive sequences in the intronic regions of *C. reinhardtii* we used the respective cDNA of Cre *SPO11* for a complementation approach. Just like in the other approaches we used the endogeneous promoter and 3´-UTR region from *A. thaliana.* For the complementation approach with mouse SPO11 we also used the cDNA since the gene structure between Mmu *SPO11* and Ath *SPO11* is quite different as the intron sequences are larger in mouse compared to *A. thaliana.* The β-form of *SPO11* transcript was used for the complementation approach since it is the complete splice form of Mmu *SPO11* including exon 2 (Kauppi et al., 2011). We investigated the seed set of ten individual *spo11-1-3*-Mmu SPO and eight individual *spo11-2-3*-Mmu SPO lines homozygous for the respective knockout. For the complementation approach with Cre SPO11 we analyzed four individual *spo11-1-3*-Cre SPO and 11 individual *spo11-2-3*-Cre SPO11 lines. None of these lines showed an increase in seed production compared to the respective knockout control neither the lines expressing Mmu SPO11 nor the ones transformed with Cre SPO11 (Figure 7).

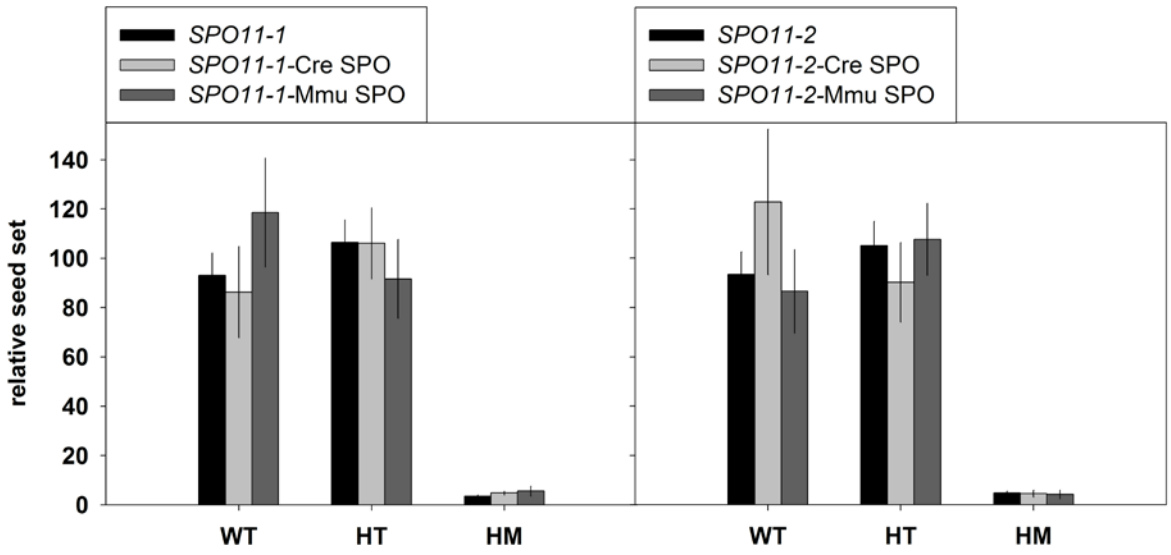


Figure 7. Seed set in complementation experiments of *spo11-1-3* and *spo11-2-3* with *SPO11* cDNA constructs from *C. reinhardtii* (Cre) and *M. musculus* (Mmu).

Relative seed set of *A. thaliana* plants containing a complementary DNA construct of *C. reinhardtii* or *M. musculus* SPO11 and none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11*, compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean (SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; $P < 0,05$).

The splicing landscape of SPO11 homologs changes when transformed in *A. thaliana*

Since we know from previous studies that *SPO11* is differentially spliced in various plants we now had the chance to elicit whether the aberrant splicing is sequence and/or species specific and if the splicing is effected by neighboring sequences (Hartung and Puchta 2000; Sprink and Hartung, 2014). By having a first look on the different splice variants of *spo11-1-3*-AthSPO1g plants we could not detect any unknown splice variants. We identified three differentially spliced transcripts all retaining introns (splice variants β, γ and λ) besides the functionally spliced form (Figure 8).

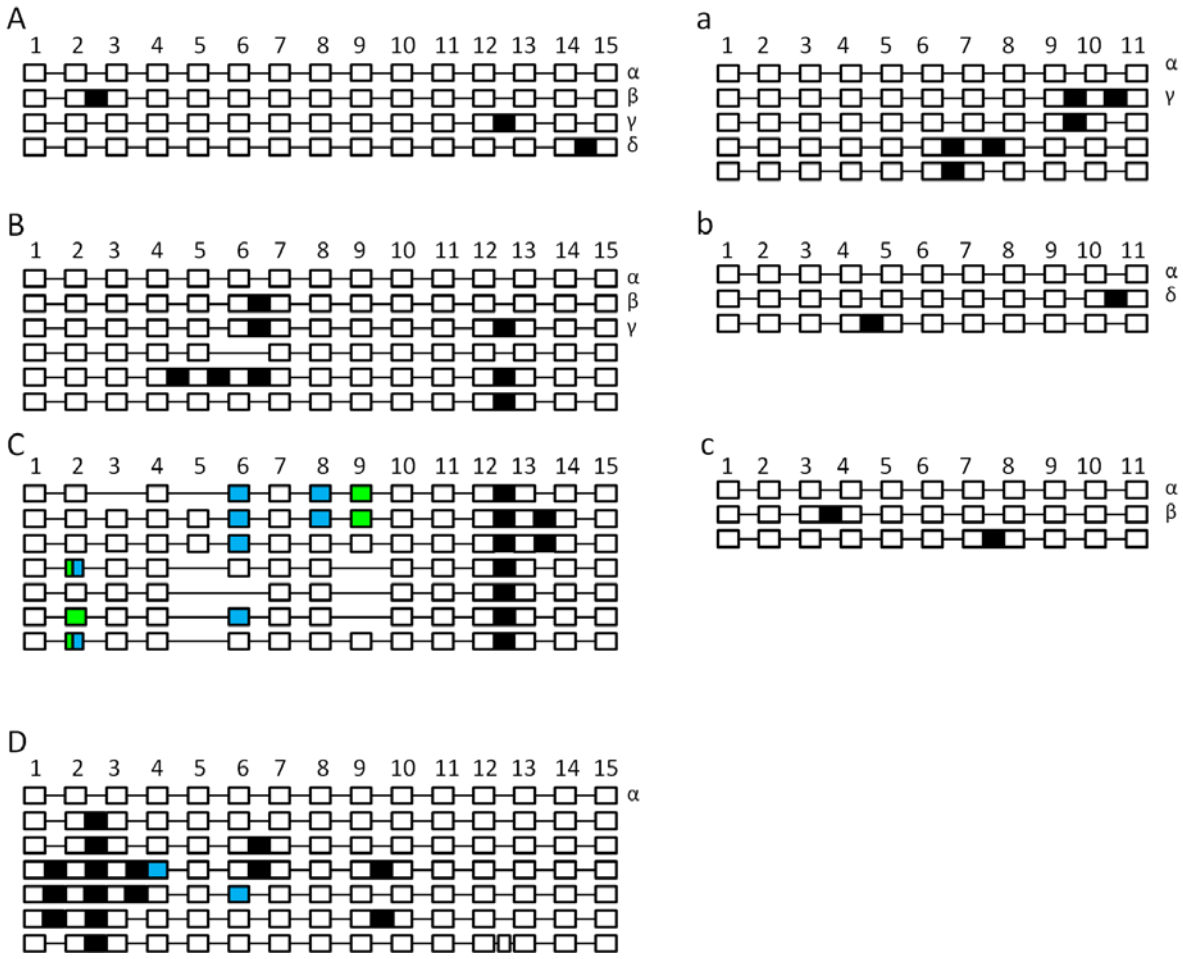


Figure 8. The splicing landscape of SPO11 genes introduced in *A. thaliana.*

Schematic non scaled schema of the different splice forms of various *SPO11-1* (A-D) and *-2* (a-c) genes introduced in *A. thaliana*. Splice forms of *A. thaliana* (A,a), *B. rapa* (B,b), *C. papaya* (C,c) and *O. sativa* (D) are shown. Exons are numbered and shown as white blocks, spliced Introns as black lines. Intron retention events are illustrated as black boxes, alternative 5´splice site selection is shown as blue boxes and alterative 3´splice site selection as light green boxes. In case of exon skipping the corresponding white box is missing, additional exons are not numbered but shown as an additional white box. Known splice forms are named in greek letters new ones are not named (according to Sprink and Hartung, 2014).

When looking at s*po11-2-3*-Ath SPO2g we could identify besides the functional spliced variant four additional transcript variants which all showed intron retention (IR). One of these forms was known (γ) and three were previously unknown (Figure 8) in the wild type control we were also able to identify four previously unknown spliced transcript variants for Ath *SPO11-2* (Hartung and Sprink, 2014) (Supplemental figure 8).

When looking on the aberrant splicing of Bra SPO1g we could detect five aberrant spliced transcript variants besides the correclyt spliced form. We could identify two previously known splice variants from *B. rapa* (β and γ) as well as two previously unknown forms, one IR and one splice form in which exon six was skipped. One form was detected which is identical with the splice variant γ from *A. thaliana* (Hartung and Sprink, 2014). For Bra SPO2g we could detect besides the functional form two additional transcript variants, one previously identified (δ) and one new variant in which intron four was retained (Figure 8). The splicing landscape of Cpa SPO1g is already quite divergent in *C. papaya* and after transformation in *A. thaliana* we never observed any correctly spliced form. We detected seven previously unknown spliced transcript variants all containing intron 12 in a combination with alternative 5´- and/or 3´-splice site selection, exon skipping and retention of additional introns (Figure 8).. The splicing of Cpa SPO2g in *A.thaliana* is resulting in two additional aberrant spliced transcripts beside the functional form, one of the transcript forms was known from Cpa SPO11-2 (β) and one is comparable to one found for Ath SPO11-2 (κ) (Figure 8).

The splicing landscape of Osa *SPO11-1* in *A. thaliana* is also completely disturbed, we were able to identify the correct spliced variant, but besides this we detected at least six aberrant spliced transcript variants all showing IR. Three variants showed additionally alternative 3´splice site selection and one variant created a completely new exon between exon 12 and exon 13 (Figure 8). The splicing of Osa SPO2g could not be analyzed in detail due to an already described problem concerning the genomic sequence (Sprink and Hartung, 2014).

Complementation with cDNA

To investigate if the great amount of aberrant splicing of *SPO11* has an influence on the complementation efficiency and if cDNA might be overcoming this problem, we additionally cloned and transformed the cDNA of different *SPO11* genes in the same way as the genomic constructs. To check this approach, we used cDNA of Ath *SPO11-1* and *SPO11-2* as control.

Like the constructs using the genomic DNA, most generated lines which were homozygous for the respective T-DNA insertion and carried the cDNA complementation construct, produced a similar number of seeds as the wild type control. We analyzed seven independent *spo11-1-3* lines carrying Ath *SPO11-1* cDNA (*spo11-1-3*-Ath SPO1c) and eight *spo11-2-3*-lines carrying Ath *SPO11-2* cDNA (*spo11-2-3*-Ath SPO2c). From these lines four out of seven (*spo11-1-3*-Ath SPO1c) and seven out of eight (*spo11-2-3*-Ath SPO2c) were able to complement the sterile phenotype of their respective homozygous control carrying no complementation construct (Figure 9). The same holds true for *B. rapa* when using cDNA instead of genomic DNA for a complementation approach (*spo11-1-3*-Bra SPO1c and *spo11-2-3*-BraSPO2c). Six out of eight (*spo11-1-3*-BraSPO1c) and six out of seven (*spo11-2-3*-BraSPO2c) lines were able to complement the sterile phenotype of the respective knockout mutant (Figure 9). As we could not detect any functionally spliced transcript variant for Cpa *SPO11-1* in *A-thaliana* we also created complementation lines for Cpa SPO11 using cDNA (*spo11-1-3*-CpaSPO1c and *spo11-2-3*-Cpa SPO2c). We analyzed eight independent *spo11-1-3*-CpaSPO1c lines and to our surprise two of them showed an increased number of seed production (Figure 9).

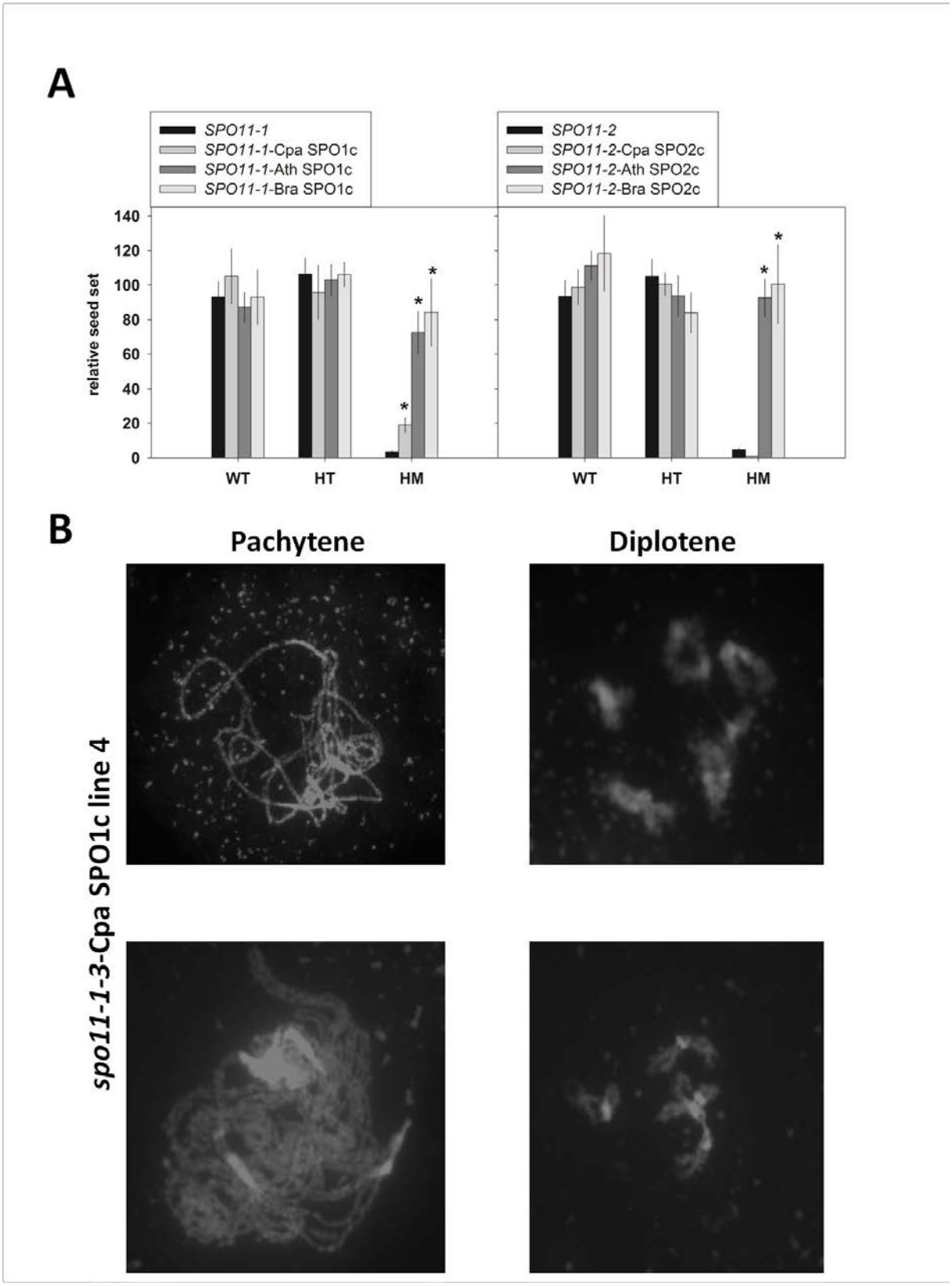

A
SPO11-1
SPO11-1-Cpa SPO1c
SPO11-1-Ath SPO1c
SPO11-1-Bra SPO1c
SPO11-2
SPO11-2-Cpa SPO2c
SPO11-2-Ath SPO2c
SPO11-2-Bra SPO2c
relative seed set
140
120
100
80
60
40
20
0
WT
HT
HM
WT
HT
HM
B
Pachytene
Diplotene
spo11-1-3-Cpa SPO1c line 4

Figure 9. Complementation experiments of *spo11-1-3* and *spo11-2-3* plants with *SPO11* cDNA constructs from *A. thaliana, B. rapa* or *C. papaya*.

(A) Relative seed set of *A. thaliana* plants containing a complementary DNA construct of *A. thaliana, B. rapa* or *C .papaya* SPO11 and none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11*, compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean (SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; $P < 0,05$) * = significant.

(B) Fluorescence micrograph of DAPI-stained nuclei from *spo11-1-3* Cpa SPO1c line 4 anthers in pachytene and diplotene stage of prophase I. A reconstitution of pairing during pachytene stage as well as the formation of five bivalents could be observed in around 20 % of the analyzed nuclei.

Both lines with increased seed set did not show a three to one mendelian segregation for their offspring indicating that more than one copy of Cpa SPO1c was integrated in these lines. We could detect a seed set of around 40% compared to wild type for one of the lines (line four). DAPI stained spreads of pollen mother cells of this line showed pairing of the chromosomes at pachytene stage in ~ 20% of the cells. Formation of bivalents during diplotene stage could also be observed in some cases (Figure 9). Immunolocalization studies in spread meiocytes of *spo11-1-3*-CpaSPO1c line 4 showed that in some cells loading of RAD51 onto the chromosomes was restored (Supplemental figure 9). But the number of foci per cell was highly variable making it impossible to give a meaningful mean. Analyzing the offspring of this line we discovered a ten times higher seed number per silique compared to the respective knockout line (Table 3).

Table.3. The average number of seeds per silique in *A. thaliana* plants with either zero (wild type), one (heterozygous; +/-) or two (homozygous; -/-) disrupted endogenous SPO11 as well as average seed number per silique in *spo11-1-3* Cpa SPO1c line 4 plants and *spo11-2-3* Cpa SPO2c plants.

Strain	Seeds/silique	Siliques analyzed
Col-0 (wild type)	45.5 +/- 3,3	15
spo11-1-3 +/-	44.6 +/- 4.7	15
spo11-1-3 -/-	0.42	150
spo11-1-3 -/- Cpa1c	5.8 +/- 2.4	80
spo11-2-3 +/-	45.9 +/- 3.5	15
spo11-2-3 -/-	0.69	180
spo11-2-3 -/-Cpa2c	0.23	40

Seeds were counted from individual siliques and are given as mean ± standard error of the mean

The siliques were also bigger compared to the respective knockout (Supplemental figure 10). Flowering time of *spo11-1-3*-Cpa1c line 4 was the same as in the respective knockout which is flowering longer compared to wild type and is producing due to the longer flower period more siliques, like it is known from many other sterile mutants. (Garcia et al., 2003; Stacey et al., 2006).

The other six lines of *spo11-1-3*-CpaSPO1c which showed a three to one mendelian segregation showed no increased number of seeds, nor could any pairing of chromosomes be detected in pachytene stage in these lines. The same holds true for *spo11-2-3*-CpaSPO2c. We analyzed eight individual lines, none of these lines showed an increased seed set, nor was the number of seeds per silique enhanced (Figure 9 and Table 3).

Discussion

In most known eukaryotic organisms a proper pairing of homologous chromosomes with subsequent recombination via crossovers is essential for genetic variability as well as correct disjunction of the chromosomes in the first meiotic division. In most so far analyzed eukaryotic organisms SPO11 plays a major role in the induction of DSBs and without those no pairing of chromosomes and subsequent random disjunction occurs. The ancestral "SPO11", TOPVIA from archaea is working in a tetrameric complex, composed of two TOPVIA and two TOPVIB subunits each. The protein complex is able to cut and religate DNA double strands in one process. Fungi and mammals contain only a single homolog of TOPVIA and have lost the second subunit. The SPO11 proteins have kept their ability to cleave double stranded DNA but the resealing of the breaks has been taken over by other proteins such as DNA ligase IV in combination with XRCC4 or XRCC3 (Bleuyard et al., 2004; 2006; Bray et al., 2005; Kozak et al., 2009) . In contrast to mammals and fungi land plants have kept a TOPVIB homolog as well as at least three TOPVIA homologs (in *A. thaliana*, Ath SPO11-1, -2- and -3) (Malik et al., 2007, Sprink and Hartung, 2014). In *A. thaliana* an interaction with TOPVIB has been shown only for SPO11-3 which is not involved in meiosis (Hartung and Puchta, 2002a). Even though a subject of general interest, still no solution could be found for the question why plants need and encode for two meiotic active SPO11. Due to a number of studies performed on Ath SPO11-1 and Ath SPO11-2 we know that the function of both proteins is not redundant (Grelon et al., 2001; Stacey et al., 2007; Hartung et al., 2007). Additionally, we need at least one functional copy of each *SPO11* for proper meiosis in *A. thaliana*. The mutation of the active tyrosine as well as changes in the conserved motifs are leading to a loss of function for Ath SPO11-1 and -2 as well as for Osa SPO11-1, changes in the non conserved part of Ath SPO11-1 do not have such an severe effect (Hartung et al., 2007; Shingu et al., 2010). Studies on SPO11-2 concerning DNA binding efficiency and structural analysis have not been done so far, so we can only assume that the findings achieved for SPO11-1 can be transferred to SPO11-2 since the described

domains seem to be conserved. In this work we addressed the question if the function of each SPO11 is sequence specific and encoded in the respective non conserved parts. By exchanging these parts of the respective SPO11 we hoped to identify regions of SPO11 which are defining the differences between both SPO11 in *A. thaliana*. Since we could not observe any difference in seed set between the complementation approaches using swapped constructs and the respective knockout lines, we have to assume that the species specificity of SPO11 lies in more than one region. Nevertheless, we cannot rule out that the exchange of sequence parts is leading to a disruption of the overall structure of the whole protein or to a disruption of a specific domain. In both cases a functional interaction might be prevented. We haven´t modified the structure of the very conserved TOPRIM domain by exchanging parts of it, which is assumed to span from motif three to motif five within the swapped approaches (An et al., 2011). But the CAP domain including the winged helix doamin is disrupted at least by the first swap and it seems that such an exchange is not functional. Additionally, the winged helix domain which is located ranging from aa 9 to 137 in Ath SPO11-1 and between aa 91 - 163 in Ath SPO11-2 is disturbed, too. Winged helix domains are assumed to mediate protein-protein as well as protein-DNA interaction since they can be found in many proteins that bind DNA such as transcription factors and restriction enzymes (Wah et al., 1997; Cicero et al., 2001). The winged helix domain seems to possess also an essential function for SPO11 since it spans the active tyrosine which is responsible for cutting the DNA. The second swap approach, which is the smallest one, is not harboring any conserved motifs. It showed also no positive complementation but a reasonable negative interference on wild type and heterozygous plants could be observed. Such an effect of partial sterility, but not in such a severe manner, was observed earlier for *SPO11* genes mutated in their active tyrosine residue (Hartung et al., 2007). These findings are indicating that SPO1swap2 seems to bind to the DNA but is not able to cleave it what might have different possible reasons such as i) missfolding of the chimeric SPO11 protein and physical distortion of the DNA/SPO11-1 and -2 cleavage complex ii) the disability of binding interaction partners which are necessary for cleavage iii) tighter binding of interaction partners paired with a missing ability to cleave DNA which results in their sequestration into an inactive complex. In all of these possible cases, the chimeric protein stands in competition to the natural occurring SPO11. Further analyses have to be done to investigate the possible structure of this chimeric protein for example by mutating single amino acids and investigate the binding capacity of DNA like it was done before for SPO11-1 (Shingu et al., 2010). Considering why SPO2swap2 as well as all the other swaps seem not to interfere with the naturally occurring SPO11 we can only

speculate that the overall structure is to strong disrupted which might prevent an effective binding to the DNA and therefore no negative interference occurred.

The last exon seems to have an absolutely essential function for Ath SPO11-2 as we know from the knockout lines *spo11-2-2* and *spo11-2-3* (Stacey et al., 2006; Hartung et al., 2007). We now were able to show that the loss of function observed for Ath SPO11-2 is indead due to the loss of the last exon and not due to a potentially interferring T-DNA insertion. Additionally, we could show that the last exon is also essential for Ath SPO11-1. A functional complementation could not be achieved by exchanging the last exon between both genes but a significant higher number of seeds as in the near sterile mutant was observed when the last exon of Ath SPO11-2 was fused to Ath SPO11-1. Since the conserved domain in the last exon was fully exchanged this is hard to explain, but the best guess that this exchange only worked in one direction is that there are additional functions in the last exon of SPO11-2 compared to SPO11-1. This question can be analyzed by exchanging smaller parts between both exons. Furthermore, we assume that without the last exon cleavage of the DNA is not possible otherwise we would have observed fragmentation of the DNA rather than unbroken univalents in the complemented mutant lines (Supplemental figure5).

Other studies showed before that alteration of single amino acids in the non conserved parts of SPO11, seem to have no negative effect on the DNA binding activity and sometimes even do not alter the cleavage capability of SPO11-1 (Shingu et al., 2010). With our heterologous complementation approach we wanted to evaluate if the function of SPO11 is conserved between orthologous SPO11 genes from organisms which are related to a different extend. Since a fully functional complementation is possible between Ath SPO11 and Bra SPO11 the function seems to be conserved at least in the family of the Brassicaceae. Multiple small changes especially in the N-terminal part of SPO11-1 and -2 seem to have no negative effect on its function, since this is the part where SPO11 from *A. thaliana* and *B. rapa* differ most. The overall structure seems not to be influenced by these small changes as SPO11 from both species must have the conserved domains at the appropriate location for a functional interaction. The subsequent repair of the breaks is conducted as in wild type, indicating that putative interacting factors of Ath SPO11 can also recognize Bra SPO11. If not, fragmentation of the chromosomes should be visible just like it is known from mutants lacking DSBs repair proteins such as MRE11 or RAD51 (Puizina et al., 2004; Li et al,.2005).

A positive complementation approach with Cpa SPO11 is not possible under natural expression conditions using genomic DNA, even though the sequence identity is quite high

(~73%). After evaluating the splicing landscape of Cpa *SPO11-1* in *A. thaliana* we detected a divergent pattern of abbernt spliced isoforms as it is the case in *C. papaya* (Sprink and Hartung, 2014). A functional spliced form of Cpa SPO1g could not be detected in flowers of *A. thaliana*, but the presence of a functional spliced form is not excluded since in *C. papaya* this form is very rare, too. Multiple insertions of Cpa SPO1c were leading to an increased number of seeds. This could have different reasons; the first one is that the binding of Cpa SPO11-1 on the DNA of *A. thaliana* is not effective enough due to changes in the TOPRIM and or winged helix domain. When it is expressed multiple times the loading of Cpa SPO11-1 onto the DNA might be enough to create a sufficient number of breaks ensuring the pairing of DNA in some cells. In other cells there might be an insufficient number of breaks as no pairing is visible. Another possibility is that Cpa SPO11 is binding to the DNA but a break cannot be induced by insufficient binding of partner proteins or improper binding to the respective second SPO11 protein to build up a tetramer. We cannot dismiss the last possibility because a combined expression of Cpa SPO11-1 and -2 was not leading to an enhanced seed set (Supplemental figure 11), but there might be other proteins necessary that coevolved with SPO11 in each plant and therefore cannot recognize the ones from *C. papaya*. One possibility to get a better understanding of what happens is to produce an antibody against Cpa SPO11-1 and have a look on its distribution. If we would see a loading of Cpa SPO11-1 onto the DNA like it is known from Ath SPO11-1 (Sanchez Moran., 2007), the second theory would be the more likely one, if no loading could be observed the first one would be more likely. A positive complementation approach could not be observed for Cpa SPO11-2, the reason is unclear and hard to explain since the sequence identity between SPO11-2 from *C. papaya* and A. *thaliana* is higher than between the orthologous SPO11-1 proteins. One possibility is that we didn´t had enough Cpa SPO11-2 loading onto the DNA to create a break since we haven´t had a line with multiple copies integrated. Creating such a line and having a look on its meiocytes could answer this question. Additionally, creating swapped constructs between Cpa SPO11 and Ath SPO11 would result in a great gain of knowledge enabling us to have a look on which parts of SPO11 need to be conserved and which parts can be edited.

A functional complementation between *O. sativa* and *A. thaliana* was not possible, most probably due to major changes in the protein sequence leading to structural changes. Especially Osa SPO11-1 possesses a major difference to Ath SPO11 since it has 19 additional amino acids located just behind the start codon, a feature that most known monocot SPO11-1 proteins share. Additionally, reasonable changes in the conserved domains occurred as Osa SPO11-1 has a shorter winged helix domain.

A complementation with more ancestral genes failed as well. The previously described *SPO11-4* gene from *O. sativa* which is proposed to have a function in meiosis seems not to be a homolog of either *SPO11-1* or *-2*. The authors described that a fragmentation of the chromosomes was observed in RNAi plants of Osa SPO11-4, this is leading to the assumption that Osa SPO11-4 is more likely involved in DSBs repair rather than in DSBs induction (An et al., 2011)

The analysis of the splicing landscape of the different orthologous plant *SPO11* genes in *A. thaliana* showed that there is a species specific pattern of aberrant splicing for each *SPO11*. Since *SPO11* from closely related plants was spliced predominantly in a correct way, the splicing of *SPO11* from more distant plants seems to be much less effective. New splicing patterns were found especially for *SPO11-1* of *C. papaya* and *O. sativa* which had never been observed so far. This lead us to suggest that the splicing of *SPO11* is not only embedded in the plain sequence of the respective *SPO11* gene but also in other factors which remain unknown and are probably species specific. Taking a closer look on *SPO11* splicing in various plants as well as at different time points during meiosis would be of great interest to gain a better understanding of this putative regulation step of SPO11 by aberrant and maybe also alternative splicing patterns.

Material and Methods

Plant material and growth conditions

For the complementation approaches, the mutant lines *spo11-1-3* (SALK_146172) and *spo11-2-3* (GABI line 749C12) were used. Both mutant lines have been previously described (Alonso et al., 2003; Stacey et al., 2006; Hartung et al., 2007). For propagation and to obtain anthers for evaluation of meiosis in pollen mother cells the plants were grown as previously described (Sprink and Hartung, 2014). For the selection of positive transformation events, seeds from inflorescence transformed with *Agrobacterium tumefaciens* were surface sterilized with 4% sodium hypochlorite, stratificated at 4°C overnight and sown on agar plates containing germination medium (GM = 4.9 g/l Murashige and Skoog including vitamins, 10 g/l sucrose and 0.8 g/l agar (adjusted to pH 5.7 with KOH)). The plantlets were cultivated in a plant culture chamber under controlled conditions of 22°C with 16 h light and 8 h dark.

Molecular characterization of the mutant lines

For genotyping of the mutant plants, DNA was extracted from a small leave of the plantlets. For PCR analysis, first primer pair was used to amplify the sequence that is interrupted by T-DNA in the mutants (SPO11-1: SP1-2 and SP1-R3; SPO11-2: SP2-2 and SP2-RP2) (Supplemental figure 12 and Supplementary data1) The presence of the T-DNA Insertion was checked using a left border specific primers for each line (SALK LBd1 for SPO11-1 and GABI LB1 for SPO11-2;) and a gene specific primer for each SPO11 gene located downstream of the T-DNA insertion (SP1-R3 and SP2-RP2). A full list of the primer sequences and PCR conditions is available as supplementary data (Supplemental data 1). Transformed plants were double checked by growing on media containing 6 mg/l phosphinothricin (PPT) and PCR checked using a primer pair specific for the insertion of the phosphinothricin acetyl transferase (PAT) (Supplementary data 1). To identify the genetic background of the plants transformed with SPO11 from *A. thaliana*, *B. rapa* and *C. papaya* a primer pair located outside the promoter and 3´-UTR region was used (SPO11-1: SP1-3Lr2 and SP15L4; SPO11-2: SP2-(-5) and SP2-R (-4)) (Supplemental figure 12 and Supplementary data 1) to amplify the sequence that is interrupted by the T-DNA, due to high sequence identity between the endogeneous *SPO11* and the introduced paralog *SPO11* gene.

Plasmid construction and plant transformation

Transformation of *A. thaliana* was performed as described (Clough and Bent, 1998). Due to sterility of homozygous SPO11 mutants, plants heterozygous for the T-DNA insertion had been used for transformation. The constructs used for plant transformation are based on the binary plasmid pPZP201 (Hajdukiewicz et al., 1994) with an enhanced multiple cloning site (MCS) and modified as previously described (Bonnet et al., 2013). For the double mutant*spo11-1-3*:SP1$_{Pro}$:Cpa SPO11-1 SP2$_{Pro}$:CpaSPO11-2cDNA the vector was edited, the PPT-resistance cassette under the control of the CaMV 35S gene promoter was exchanged by a gentamycin resistance gene (aaaC1) under the control of the PcUbi4-2 promoter. Plants homozygous for the first event were used for the transformation with the second gene. After selection of transformed plants in the T1 generation by PPT resistance (6 mg/l PPT), the T2 generation was checked for mendelian 3:1 segregation to obtain lines with a single insertion event.

RNA extraction and RT-PCR

All kits used in this study were applied, if not especially mentioned, strictly following the manufacturer's instructions. Total RNA of *A. thaliana*, *B. rapa* and *C. papaya* was isolated from fresh young flowers using the RNA mini Kit from Bio & Sell (Bio&Sell e.K., Feucht, Germany).Total RNA from *Chlamydomonas reinhardtii* was isolated from fresh grown liquid cultures. Isolated RNA was treated with DNase I (Thermo Fisher Scientific, Germany) and afterwards cleaned and concentrated using the GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Fisher Scientific, Germany). To check contamination with genomic DNA in the DNase I treated RNA a PCR was performed with RNA as template using gene specific primers for *SPO11-1*. CDNA was produced using an anchored oligo-dT Primer (VT_{20}) using the Maxima H Minus Reverse Transcriptase Kit (Thermo Fisher Scientific, Germany) using 2-5 µg of total RNA as template for RT-reaction.

Molecular Characterization of SPO11 splice variants

The screening for aberrant spliced SPO11 transcripts were performed as previously described with one exception (Sprink and Hartung, 2014). For analysis of the splice variants of the transgene, plants homozygous for the respective knockout were used for RNA isolation and cDNA production, to ensure that contamination with the endogeneous transcripts is excluded.

Complementation experiments

We generated several constructs to rescue the observed phenotypes of *spo11-1-3 and spo11-2-3* and to check whether a heterologous SPO11 protein is able to complement the sterility. For all complementation approaches the respective promoter and 3´-UTR region of the corresponding *SPO11* was used as described (Hartung et al., 2007). The genomic regions from ATG to Stop of *SPO11-1* and *-2* from *A. thaliana, B. rapa, C. papaya* and *O. sativa* and the genomic region of SPO11-4 from *O. sativa* were amplified using gene specific primers with an 15 bp long attached linker by a high proofreading polymerase (Q5® High-Fidelity DNA Polymerase, New England Biolabs, Ipswich, Massachusetts, USA). Linker sequences were homolog to the corresponding 5´- and 3´-UTR-region from the respective *SPO11* from *A. thaliana* (Supplementary data 1). The corresponding promoter and 3´-UTR regions were added to the heterologous genes and transferred into the binary vector via the homologous

linkers using the In Fusion High fidelity cloning Kit (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France). The cDNA constructs for *SPO11-1* and *SPO11-2* from *A. thaliana*, *B. rapa* and *C. papaya* were amplified using high quality cDNA samples prepared from fresh young flowers also using the same linker primers as for the genomic DNA. The cDNA construct for SPO11 from *C. reinhardtii* was constructed using cDNA from liquid culture kindly provided by Serge Zagermann from the Institute of Horticultural Production Systems Biosystems Engineering (Leibniz University Hannover). The cDNA constructs for mouse (*Mus musculus)* were made from a plasmid containing the mouse *SPO11-β* cDNA which is able to complement *spo11* in mouse kindly provided by Dr. Scott Keeney from the Memorial Sloan Kettering cancer center (NY, USA) (Kauppi et al.,2011). All amplified genes were fully sequenced after construction and before we transformed them into the corresponding heterozygous mutants.

Constructs with interchanged section between Ath SPO11-1 and -2 (Supplemental figure 2) were prepared using segment specific primers with attached linkers. Resulted fragments were fused together by the sites of homology and added to the vector using also the In Fusion high fidelity cloning Kit. Seed set was calculated as mean ± SEM of each genotype by comparing the mean number of seeds from wild type and heterozygous plants of the individual construct with every plant carrying the respective construct.

Preparation of pollen mother cells

The staining of the chromosomes of the pollen mother cells was performed as described (Pawlowski et al., 2013. P 3 ff.). Primary inflorescences were cut just after the first bud had opened and were fixed in ice cold fixative (3:1 ethanol : acetic acid). After 24 h the fixative was exchanged. Flowers were dissected in fixative under a stereo microscope. All buds containing mature pollen were discarded all other buds were washed 3x in 0.01 M citrate buffer (pH 4.5) and digested in a mixture of 0.33% cellulase (C1794, Sigma-Aldrich Chemie GmbH, Taufenkirchen, Germany) and 0.33% pectolyase (P5936, Sigma-Aldrich Chemie GmbH, Taufenkirchen, Germany) in 0.01 M citrate buffer for 90 min at 37°C in a humid chamber. Each flower bud was squashed on a separate slide, mixed with 5µl of 60% acetic acid, briefly stirred and incubated for 45s on a heated plate at 45°C. A ring of fixative was drawn around the droplet and the slide was tilted, afterwards the slide was dried from the back using a hairdryer.

Meiocytes for immunolocalization of meiotic proteins were isolated, with minor changes as described (Pawlowski et al., 2013. P 3 ff.) from immature flower buds, using a stereomicroscope and fine forceps. Inflorescence was cut and immediately put on a damp filter paper. All buds containing mature pollen were discarded from the other buds anthers were dissected transferred onto a clean glass slide (~ 40/slide) and digested in 5 µl of a mixture of 0.4% cytohelicase (C8274, Sigma-Aldrich Chemie GmbH, Taufenkirchen, Germany), 1.5% sucrose and 1% of polyvinylpyrolidone K30 (MW 40,000; Carl Roth GmbH & Co. KG, Karlsruhe, Germany) for 5 min on a heated plate at 37°C in a humid chamber, smashed and digested in additional 5 µl of digestion medium for 5 min. Afterwards mixed with 10 µl of 0.25% of Lipsol (SciLabware Limited, Staffordshire, UK) as spreading medium, incubated for 4 min at 37°C on a heated plate and then fixed for min. 2 h with 4% paraformaldehyde under the fume hood. Immunostaining was done as described using antibodies (ABs) against the meiosis specific protein Ath ASY1 (1:1000) and Ath RAD51 (1:200) (Armstrong et al., 2002; Sanchez Moran et al. 2007; 2008). As secondary ABs, 1:200 goat anti rat conjugated with Alexa488® (112-545-167, Dianova GmbH, Hamburg, Germany) and 1:200 goat anti rabbit conjugated with Cy3 (111-165-144, Dianova GmbH, Hamburg, Germany) ABs were used.

All slides were stained with 7 µl of VECTASHIELD antifading mounting medium (H-1000, Vector Laboratories Inc., Burlingame,California, USA) containing 0.01 mg/ml DAPI (4´6-Diamidin-2-phenylindol). Staining of chromatin and meiotic stages was analyzed using a fluorescence microscope (Nikon ECLIPSE Ni-E microscope, DAPI-5060C; CFI Plan Apochromat 60X/1.4 and 100X/1.45, DS-QiMC camera).

Statistical analyses

Statistical analyses were performed using Welch´s T-Test by comparing the combined relative seed set of wild type and heterozygous plants with the relative seed set of wild type , heterozygous and homozygous plants alone. Before using the combined seed set of wild type and heterozygous plants for calculations, it was analyzed if a difference in seed set could be measured between both. No difference was detected expect for the construct SPO1swap2 which seems to interfere with the endogenous SPO11. Additional tests were performed between plants carrying the respective construct against the respective background line carrying no construct. Relative data were used for the calculation due to seasonal changes of the absolute number of produced seeds and seed number was also varying between plants which were pre cultivated on media compared to plants which were sown directly into soil. A

parallel planting of all lines was not possible due to limitations of the cultivation area. Detailed information on seed set as well as all P-values can be found in supplemental data 2.

Supplemental data:

Supplemental Figure1: Dual immunolocalization of ASY1 and RAD51 proteins in meiocytes of wild type plants expressing an additional wild type SPO11 allele.

Supplemental figure 2. In scale exon-intron organization of swapped (swap1 to 4) and truncated (Δlex) Ath SPO11-1 and Ath SPO11-2 constructs.

Supplemental figure 3. Seed set in complementation experiments of *spo11-1-3* and *spo11-2-3* using interchanged *SPO11* constructs.

Supplemental figure 4. Negative interference of SPO1swap2 on wildtype plants.

Supplemental figure 5. Fluorescence micrograph of DAPI stained nuclei from spo11-swap4 plants.

Supplemental figure 6. Fluorescence micrograph of DAPI stained nuclei from *spo11*-BraSPOg plants.

Supplemental figure 7. Relationship between the used organisms for the complementation approach.

Supplemental figure 8. The Splicing landscape of endogenous SPO11-1 and -2 in *A. thaliana.*

Supplemental figure 9. Dual immunolocalization of ASY1 and RAD51 proteins in meiocytes of *spo11-1-3*-Cpa SPO1c.

Supplemental figure 10. Partial complementation of the sterile phenotype of *spo11-1-3* by the multiple expression of SPO11-1 cDNA from *C. papaya*.

Supplemental figure 11. Seed set in the complementation experiment of *spo11-1-3* Cpa SPO1g with *Spo11-2* cDNA from *C. papaya*.

Supplemental figure 12. In scale exon-intron organization of Ath *SPO11-1* and Ath *SPO11-2.*

Supplemental data 1: Primerlist

Supplemental data 2: Seed set numbers and statistics.

Accession Numbers

Sequence data from this article can be found in the databases under the following accession numbers: At3g13170 (Ath SPO11-1), At1g63990 (Ath SPO11-2); At5g20850 (RAD51); Os03g54091 (Osa SPO11-1); Os08g0156900 (Osa SPO11-2); Os03g0284800 (Osa SPO11-4); Gene ID: 5727367 (Cre SPO11); MGI:1349669 (Mmu SPO); KP895877 (Cpa SPO1); KP895878 (Cpa SPO2)

Acknowledgments:

We thank the group of Chris Franklin, especially Kim Osman and Eugenio Sanchez Moran for providing us with antibodies and giving me the chance to learn the methods in his lab. Additionally, we would like to thank the Institute of Horticultural Production Systems Biosystems Engineering and the Botanical institute of the Leibniz University Hannover for providing us with *Chlamydomonas rheinhardtii* and *O. sativa* and Scott Keeney from the Memorial Sloan Kettering Cancer Center for providing us with mouse cDNA. We also would like to thank Janina Metje for thorough reading of the manuscript and Katrin Schulze for helping with some experiments. Finally we wanted to thank the Deutsche Forschungsgemeinschaft, SPP1384, for funding this project.

Author contributions

F.H conceived the project and designed experiments. T.S designed experiments, performed all the experiments and wrote the article.

References:

Alani, E., Padmore, R., and Kleckner, N. (1990). Analysis of wild-type and rad50 mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. Cell 61, 419-436.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., and Ecker, J.R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301, 653-657.

An, X.J., Deng, Z.Y., and Wang, T. (2011). OsSpo11-4, a Rice Homologue of the Archaeal TopVIA Protein, Mediates Double-Strand DNA Cleavage and Interacts with OsTopVIB. PLoS One 6, 14.

Armstrong, S.J., Caryl, A.P., Jones, G.H., and Franklin, F.C.H. (2002). Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in Arabidopsis and Brassica. Journal of Cell Science 115, 3645-3655.

Bergerat, A., deMassy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., and Forterre, P. (1997). An atypical topoisomerase II from archaea with implications for meiotic recombination. Nature 386, 414-417.

Bleuyard, J.Y., and White, C.I. (2004). The Arabidopsis homologue of Xrcc3 plays an essential role in meiosis. The EMBO journal 23, 439-449.

Bleuyard, J.-Y., Gallego, M.E., and White, C.I. (2006). Recent advances in understanding of the DNA double-strand break repair machinery of plants. DNA repair 5, 1-12.

Bonnet, S., Knoll, A., Hartung, F., and Puchta, H. (2013). Different functions for the domains of the Arabidopsis thaliana RMI1 protein in DNA cross-link repair, somatic and meiotic recombination. Nucleic Acids Res. 41, 9349-9360.

Bray, C.M., and West, C.E. (2005). DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. New Phytol. 168, 511-528.

Buhler, C., Lebbink, J.H., Bocs, C., Ladenstein, R., and Forterre, P. (2001). DNA topoisomerase VI generates ATP-dependent double-strand breaks with two-nucleotide overhangs. The Journal of biological chemistry 276, 37215-37222.

Cao, L., Alani, E., and Kleckner, N. (1990). A pathway for generation and processing of double-strand breaks during meiotic recombination in S. cerevisiae. Cell 61, 1089-1101.

Cicero, M.P., Hubl, S.T., Harrison, C.J., Littlefield, O., Hardy, J.A., and Nelson, H.C. (2001). The wing in yeast heat shock transcription factor (HSF) DNA-binding domain is required for full activity. Nucleic Acids Res. 29, 1715-1723.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal : for cell and molecular biology 16, 735-743.

Cole, F., Keeney, S., and Jasin, M. (2010). Evolutionary conservation of meiotic DSB proteins: more than just Spo11. Genes Dev. 24, 1201-1207.

de Massy, B. (2013). Initiation of Meiotic Recombination: How and Where? Conservation and Specificities Among Eukaryotes. Annual Review of Genetics 47, 563-599.

Edlinger, B., and Schlogelhofer, P. (2011). Have a break: determinants of meiotic DNA double strand break (DSB) formation and processing in plants. J. Exp. Bot. 62, 1545-1563.

Garcia, V., Phelps, S.E., Gray, S., and Neale, M.J. (2011). Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. Nature 479, 241-244.

Garcia, V., Bruchet, H., Camescasse, D., Granier, F., Bouchez, D., and Tissier, A. (2003). AtATM is essential for meiosis and the somatic response to DNA damage in plants. The Plant Cell Online 15, 119-132.

Grelon, M., Vezon, D., Gendrot, G., and Pelletier, G. (2001). AtSPO11-1 is necessary for efficient meiotic recombination in plants. Embo J. 20, 589-600.

Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994). THE SMALL, VERSATILE PPZP FAMILY OF AGROBACTERIUM BINARY VECTORS FOR PLANT TRANSFORMATION. Plant Mol.Biol. 25, 989-994.

Hartung, F., and Puchta, H. (2000). Molecular characterisation of two paralogous SPO11 homologues in Arabidopsis thaliana. Nucleic Acids Res. 28, 1548-1554.

Hartung, F., Angelis, K.J., Meister, A., Schubert, I., Melzer, M., and Puchta, H. (2002a). An archaebacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants. Curr. Biol. 12, 1787-1791.

Hartung, F., Blattner, F.R., and Puchta, H. (2002b). Intron gain and loss in the evolution of the conserved eukaryotic recombination machinery. Nucleic Acids Res. 30, 5175-5181.

Hartung, F., Wurz-Wildersinn, R., Fuchs, J., Schubert, I., Suer, S., and Puchta, H. (2007). The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in Arabidopsis. Plant Cell 19, 3090-3099.

Kauppi, L., Barchi, M., Baudat, F., Romanienko, P.J., Keeney, S., and Jasin, M. (2011). Distinct Properties of the XY Pseudoautosomal Region Crucial for Male Meiosis. Science 331, 916-920.

Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88, 375-384.

Knoll, A., and Puchta, H. (2011). The role of DNA helicases and their interaction partners in genome stability and meiotic recombination in plants. J. Exp. Bot. 62, 1565-1579.

Kozak, J., West, C.E., White, C., da Costa-Nunes, J.A., and Angelis, K.J. (2009). Rapid repair of DNA double strand breaks in Arabidopsis thaliana is dependent on proteins involved in chromosome structure maintenance. DNA repair 8, 413-419.

Li, W., Yang, X., Lin, Z., Timofejeva, L., Xiao, R., Makaroff, C.A., and Ma, H. (2005). The AtRAD51C Gene Is Required for Normal Meiotic Chromosome Synapsis and Double-Stranded Break Repair in Arabidopsis. Plant Physiol. 138, 965-976.

Liu, J., Wu, T.C., and Lichten, M. (1995). The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate. Embo j 14, 4599-4608.

Malik, S.B., Ramesh, M.A., Hulstrand, A.M., and Logsdon, J.M. (2007). Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. Mol. Biol. Evol. 24, 2827-2841.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature 455, 770-774.

Nairz, K., and Klein, F. (1997). mre11S—a yeast mutation that blocks double-strand-break processing and permits nonhomologous synapsis in meiosis. Genes Dev. 11, 2272-2290.

Neale, M.J., Pan, J., and Keeney, S. (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. Nature 436, 1053-1057.

Pâques, F., and Haber, J.E. (1999). Multiple Pathways of Recombination Induced by Double-Strand Breaks in Saccharomyces cerevisiae. Microbiology and Molecular Biology Reviews 63, 349-404.

Pawlowski, Wojciech P., Grelon, Mathilde, Armstrong, Susan (Eds.) (2013) Plant meiosis, Methods in Molecular Biology, Vol. 990; 2013, XV, 238 p. 52 illus., 28 illus. in color.

Prieler, S., Penkner, A., Borde, V., and Klein, F. (2005). The control of Spo11's interaction with meiotic recombination hotspots. Genes Dev. 19, 255-269.

Prinz, S., Amon, A., and Klein, F. (1997). Isolation of COM1, a new gene required to complete meiotic double-strand break-induced recombination in Saccharomyces cerevisiae. Genetics 146, 781.

Puizina, J., Siroky, J., Mokros, P., Schweizer, D., and Riha, K. (2004). Mre11 deficiency in Arabidopsis is associated with chromosomal instability in somatic cells and Spo11-dependent genome fragmentation during meiosis. The Plant Cell Online 16, 1968-1978.

Roeder, G.S. (1997). Meiotic chromosomes: it takes two to tango. Genes Dev. 11, 2600-2621.

Rothenberg, M., Kohli, J., and Ludin, K. (2009). Ctp1 and the MRN-complex are required for endonucleolytic Rec12 removal with release of a single class of oligonucleotides in fission yeast. PLoS Genet 5, e1000722.

Sanchez-Moran, E., Santos, J.L., Jones, G.H., and Franklin, F.C.H. (2007). ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. Genes Dev. 21, 2220-2233.

Sanchez-Moran, E., Osman, K., Higgins, J.D., Pradillo, M., Cunado, N., Jones, G.H., and Franklin, F.C.H. (2008). ASY1 coordinates early events in the plant meiotic recombination pathway. Cytogenet. Genome Res. 120, 302-312.

Shingu, Y., Mikawa, T., Onuma, M., Hirayama, T., and Shibata, T. (2010). A DNA-binding surface of SPO11-1, an Arabidopsis SPO11 orthologue required for normal meiosis. Febs J. 277, 2360-2374.

Shingu, Y., Tokai, T., Agawa, Y., Toyota, K., Ahamed, S., Kawagishi-Kobayashi, M., Komatsu, A., Mikawa, T., Yamamoto, M.T., Wakasa, K., Shibata, T., and Kusano, K. (2012). The double-stranded break-forming activity of plant SPO11s and a novel rice SPO11 revealed by a Drosophila bioassay. BMC Mol. Biol. 13, 16.

Simkova, K., Moreau, F., Pawlak, P., Vriet, C., Baruah, A., Alexandre, C., Hennig, L., Apel, K., and Laloi, C. (2012). Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. 109, 16360-16365.

Soltis, D.E., Bell, C.D., Kim, S., and Soltis, P.S. (2008). Origin and early evolution of angiosperms. Annals of the New York Academy of Sciences 1133, 3-25.

Sprink, T., and Hartung, F. (2014). The splicing fate of plant SPO11 genes. Frontiers in Plant Science 5.

Stacey, N.J., Kuromori, T., Azumi, Y., Roberts, G., Breuer, C., Wada, T., Maxwell, A., Roberts, K., and Sugimoto-Shirasu, K. (2006). Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. Plant J. 48, 206-216.

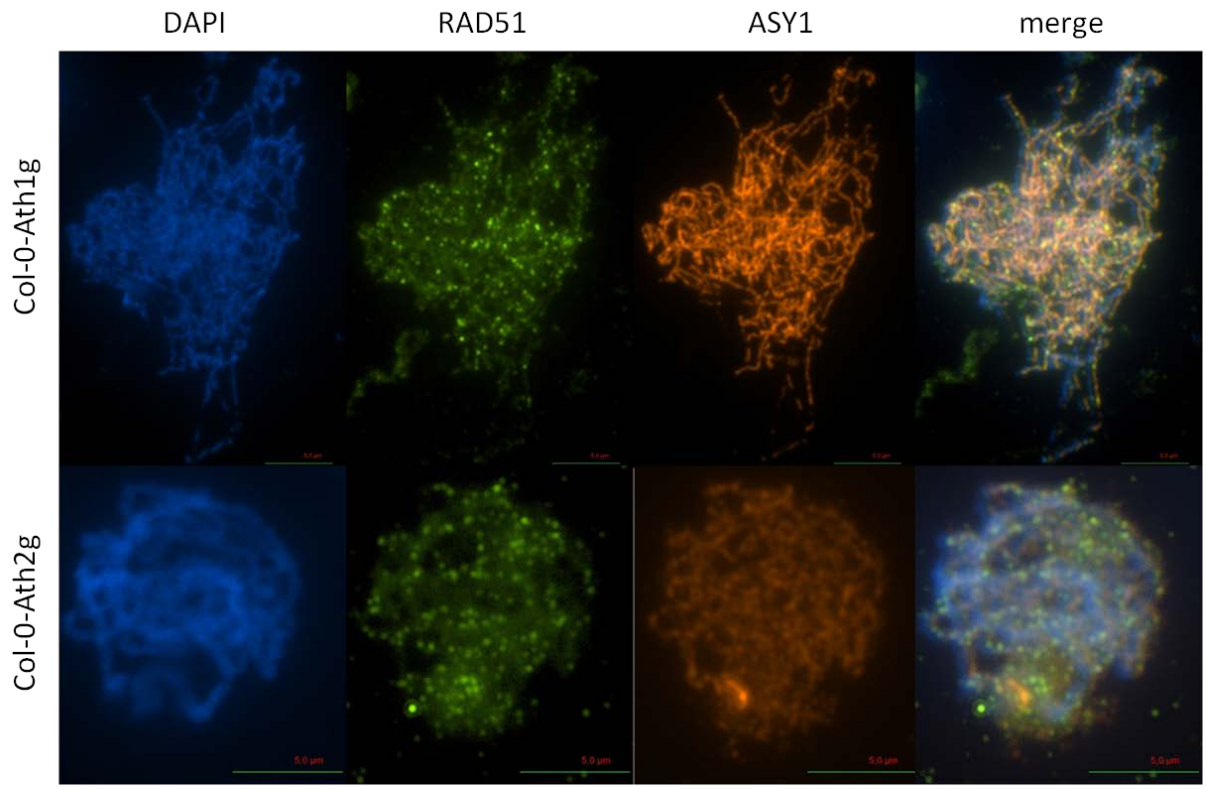
Town, C.D., Cheung, F., Maiti, R., Crabtree, J., Haas, B.J., Wortman, J.R., Hine, E.E., Althoff, R., Arbogast, T.S., Tallon, L.J., Vigouroux, M., Trick, M., and Bancroft, I. (2006). Comparative genomics of Brassica oleracea and Arabidopsis thaliana reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell 18, 1348-1359.

Wah, D.A., Hirsch, J.A., Dorner, L.F., Schildkraut, I., and Aggarwal, A.K. (1997). Structure of the multimodular endonuclease FokI bound to DNA. Nature 388, 97-100.

Wang, D.Y., Kumar, S., and Hedges, S.B. (1999). Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proceedings of the Royal Society B: Biological Sciences 266, 163-171.

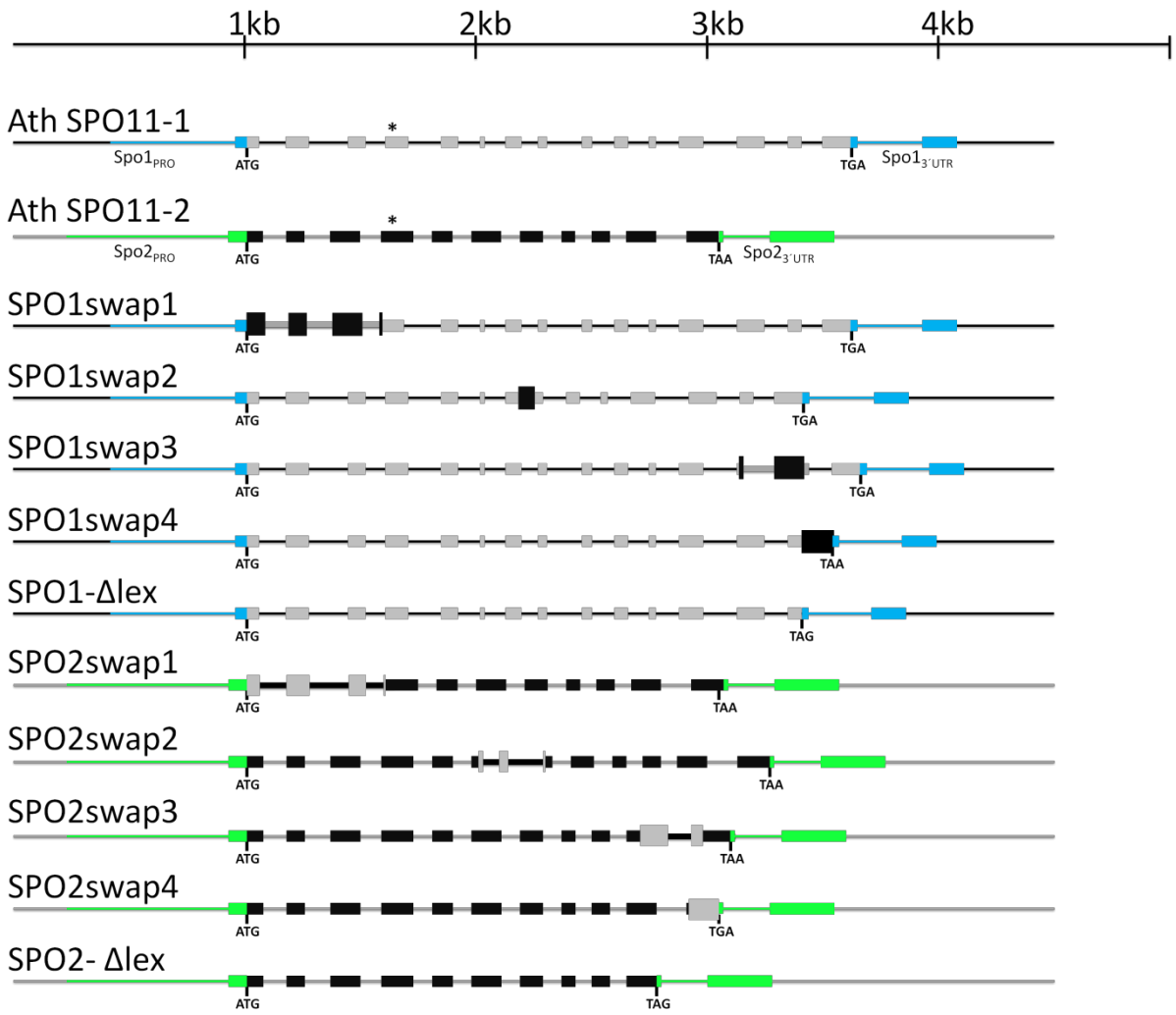
Woodhouse, M.R., Pedersen, B., and Freeling, M. (2010). Transposed Genes in Arabidopsis Are Often Associated with Flanking Repeats. PLoS Genet 6, e1000949.

Zickler, D., and Kleckner, N. (1998). THE LEPTOTENE-ZYGOTENE TRANSITION OF MEIOSIS. Annual Review of Genetics 32, 619-697.



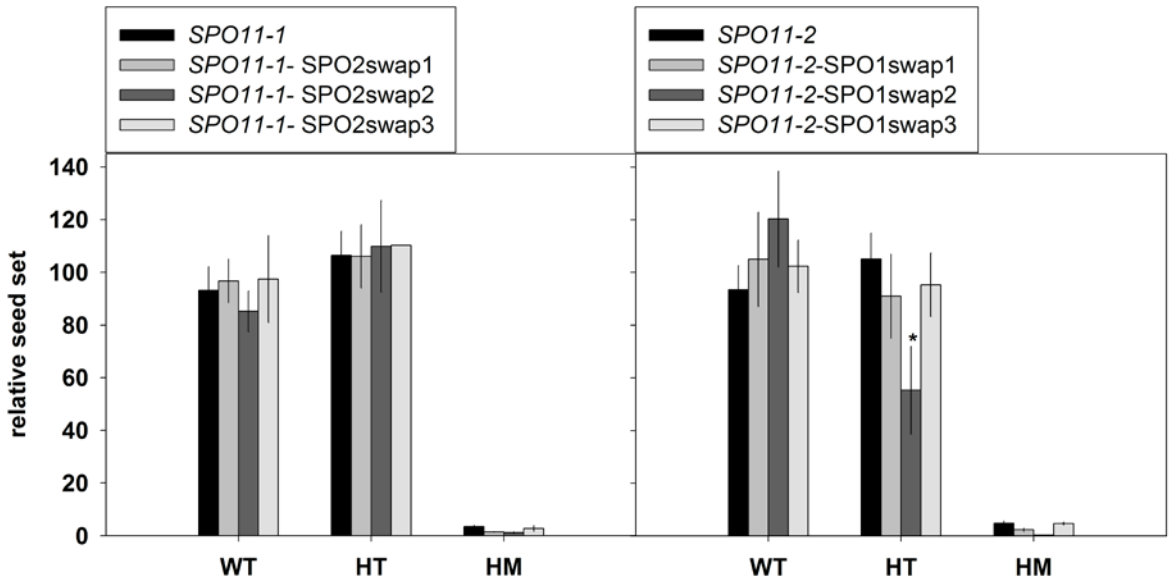
Supplemental figure 1. Dual immunolocalization of ASY1 and RAD51 proteins in meiocytes of wild type plants expressing an additional wild type SPO11 allele.

Meiocytes were counterstained with DAPI (blue) and dual immunolocalization of ASY1 (orange) and RAD51 (green) was performed using rabbit polyclonal antibody against ASY1 and rat polyclonal antibody against RAD51. Meiocytes in zygotene (Col Ath1g) and pachytene (Col Ath2g) stage were used for immunolocalization. Green bar = 5µm.



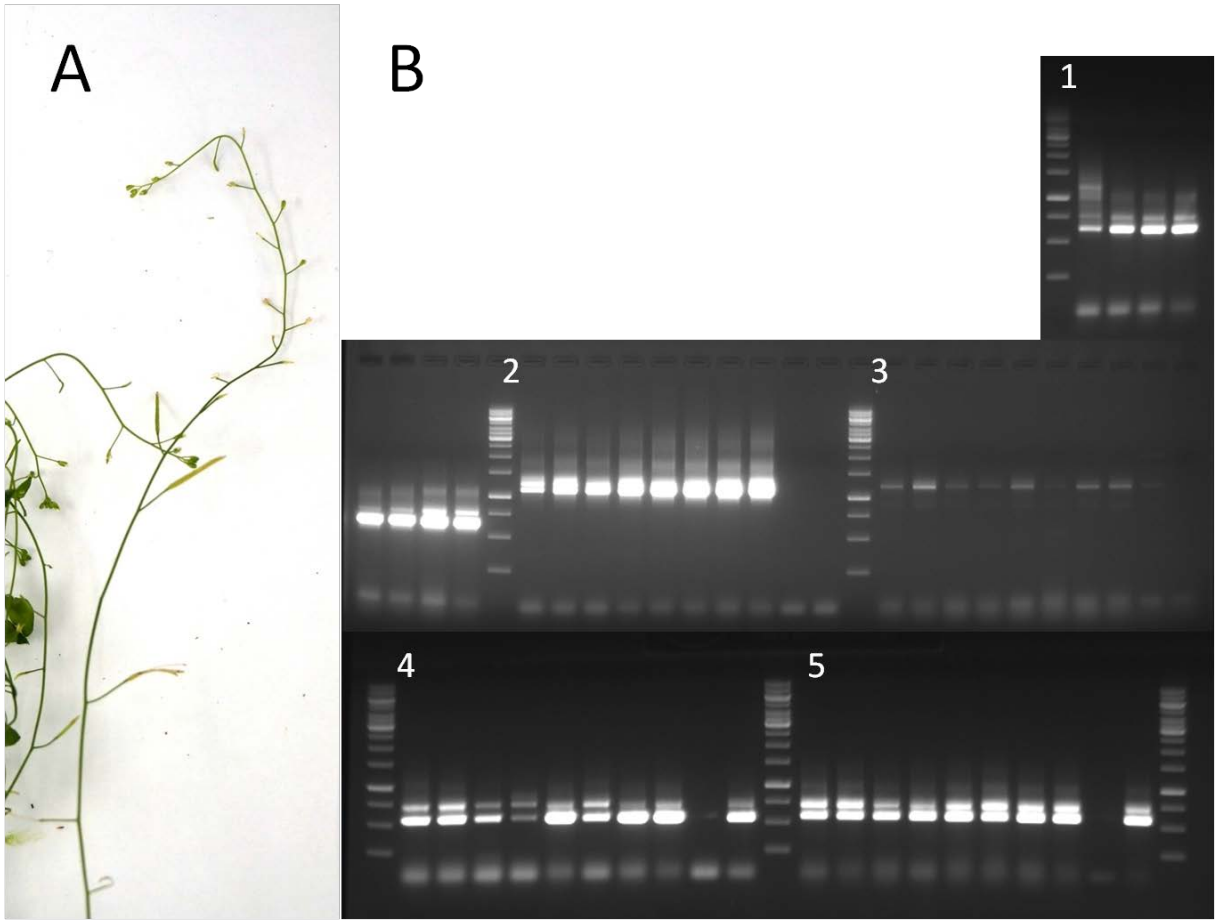
Supplemental figure 2. In scale exon-intron organization of swapped (swap1 to 4) and truncated (Δlex) Ath SPO11-1 and Ath SPO11-2 constructs.

Coding region of Arabidopsis SPO11-1 and -2 are shown as grey and black boxes, introns are represented as black and grey lines. The respective promoter (pro) and UTR regions are shown in blue and green, 3´-UTRs are shown as boxes 3´-introns as lines. The respective start and STOP by its codon (ATG, TAA, TGA ,TAG). The active tyrosine of each SPO11 is indicated by an asterix. Interchanged regions between SPO11-1 and -2 are shown in bold



Supplemental figure 3. Seed set in complementation experiments of *spo11-1-3* and *spo11-2-3* using interchanged *SPO11* constructs.

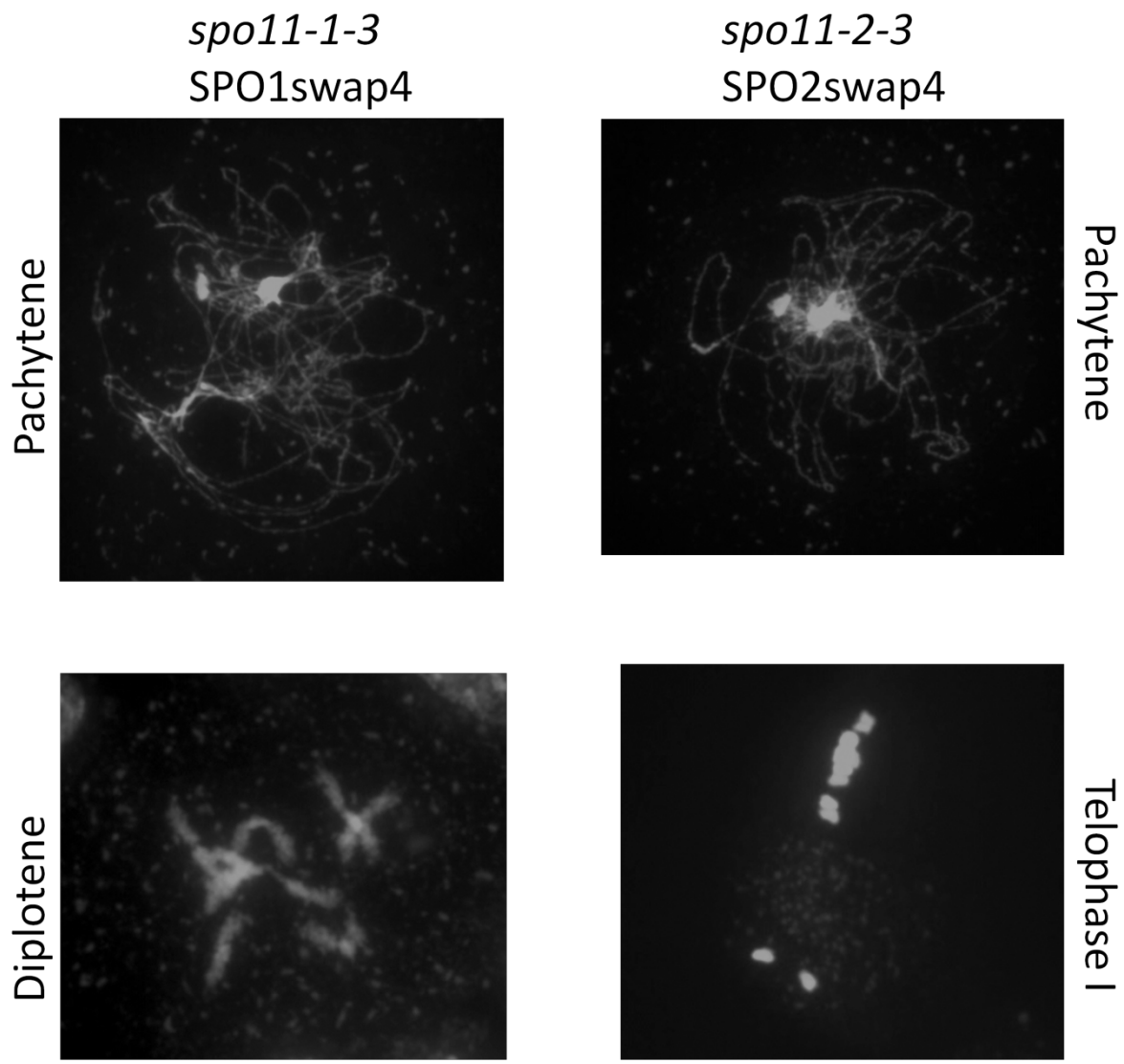
Relative seed set of *A. thaliana* plants containing an artificial SPO11 construct and none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11*, compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean (SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; $P < 0,05$) * = significant



Supplemental figure 4. Negative interference of SPO1swap2 on wildtype plants.

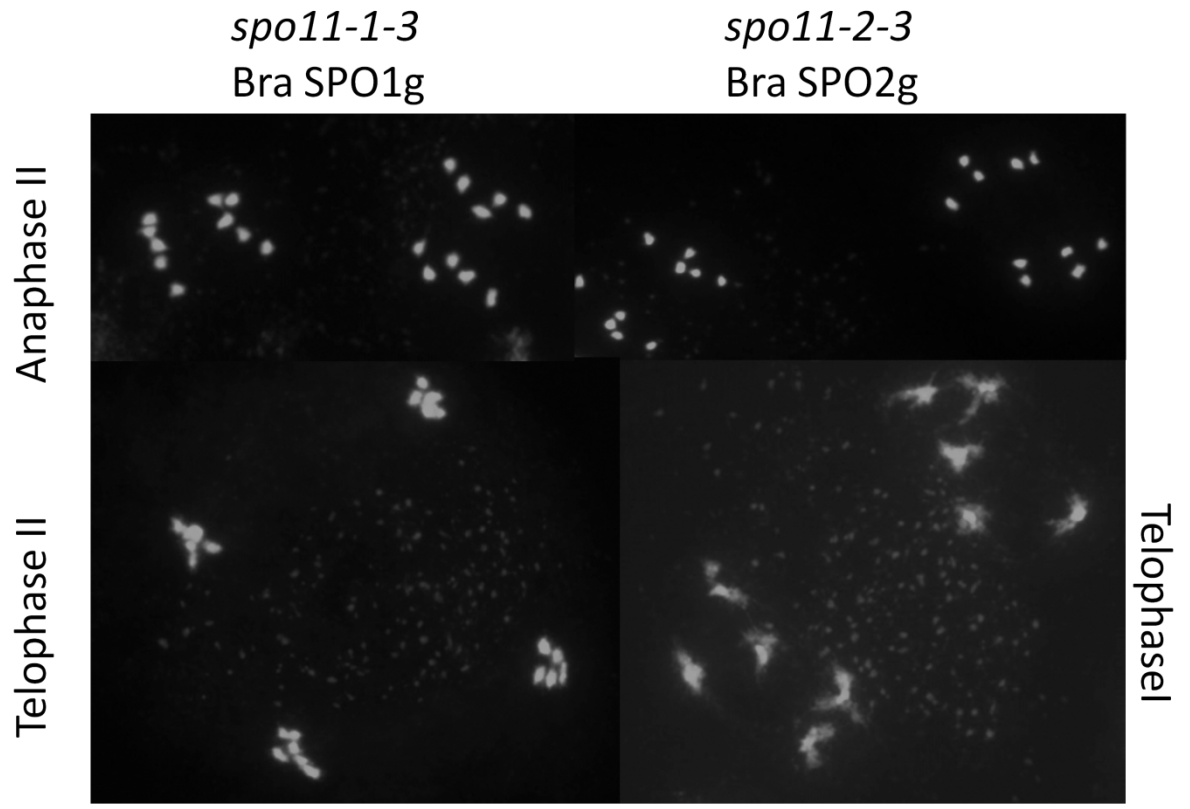
(A) Flowering stem of a wild type plant carrying SPO1swap 2 construct, some siliques are normal developed but many remain small and show abnormal seed set.

(B) Semiquantitative RT-PCR of *SPO11-1* (2, 4, 5), *SPO11-2* (3) and *SPO1swap2* (1, 2) in eight different wildtype and heterozygous plants which showed negative interference. 1µl of cDNA was used for each reaction following primer sets were used: (**1**) SPO1 ATG and SPO2-2 (located in the swap); (**2**) SPO1ATG and SPO1 Stop; (**3**) SPO2ATG and SPO2 Stop; (**4**) SPO1ATG and SPO1Ex8_rev (located in the interchanged region, missing in SPO1swap2); (**5**) SPO1Ex8 and SPO1 Stop.



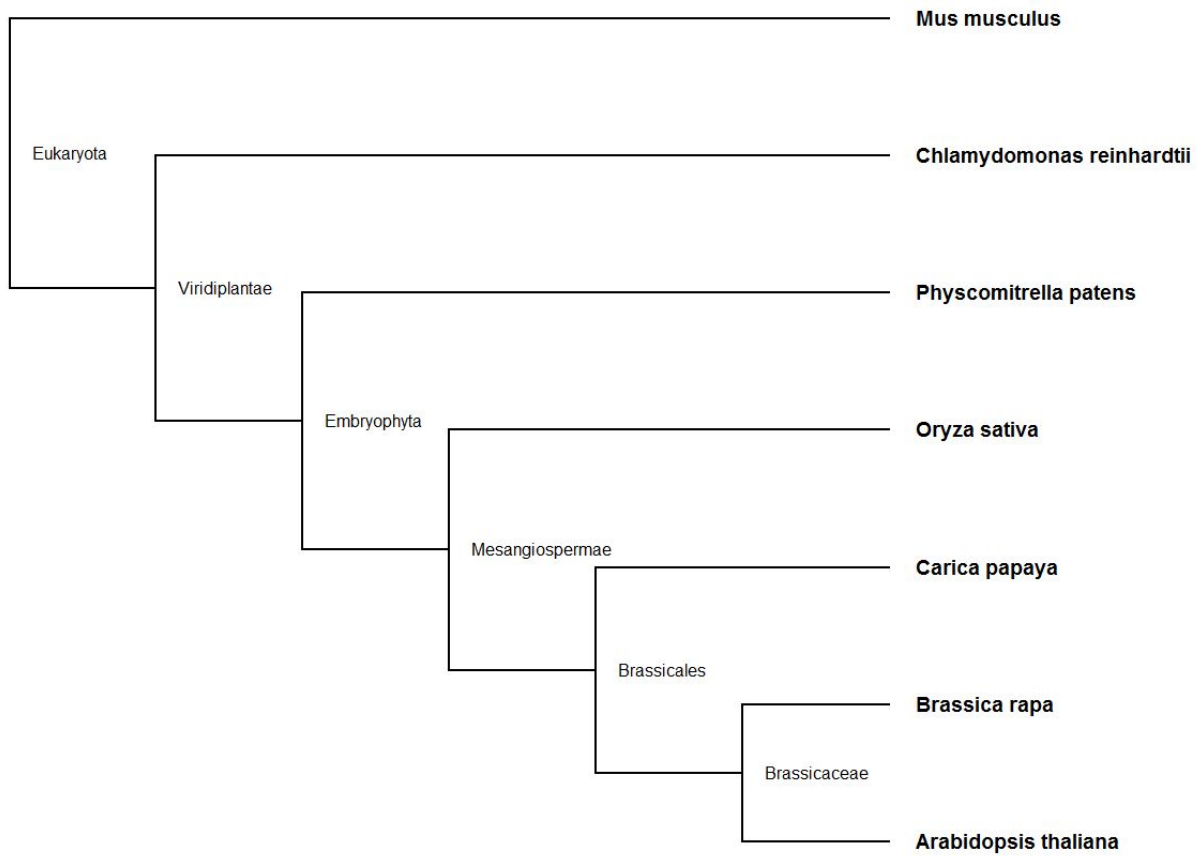
Supplemental figure 5. Fluorescence micrograph of DAPI stained nuclei from spo11-swap4 plants.

Fluorescence micrograph of DAPI-stained nuclei from *spo11-1-3* SPO1swap4 and *spo11-2-3* SPO2swap4 plants in different meiotic stages (as indicated). Neither a pairing of chromosomes and formation of bivalents during diplotene stage nor an ordered distribution of chromosomes during Telophase I could be observed.



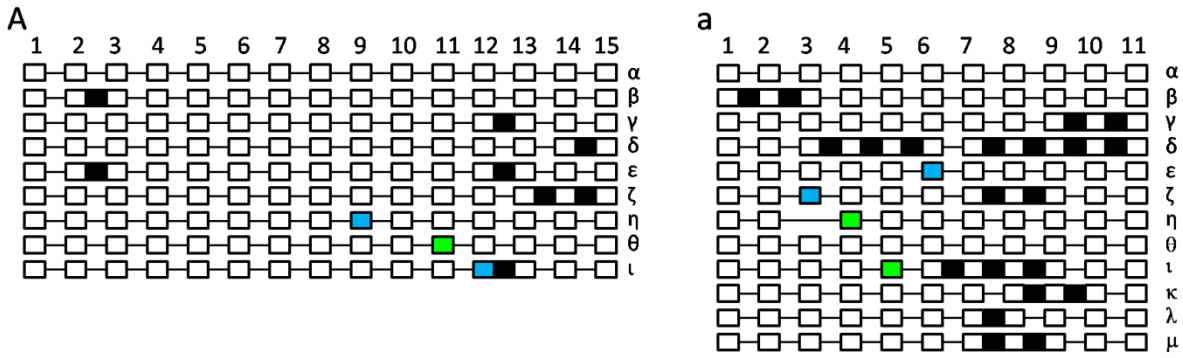
Supplemental figure 6. Fluorescence micrograph of DAPI stained nuclei from *spo11*-BraSPOg plants.

Fluorescence micrograph of DAPI-stained nuclei from *spo11-1-3* Bra SPO1g and *spo11-2-3* Bra SPO2g anthers in different meiotic stages (as indicated). An ordered distribution of the five sister chromatids per developing gamete could be observed in Ana- and Telophase II, The correct distribution of the five homologous chromosomes was observed in Teleophase I.



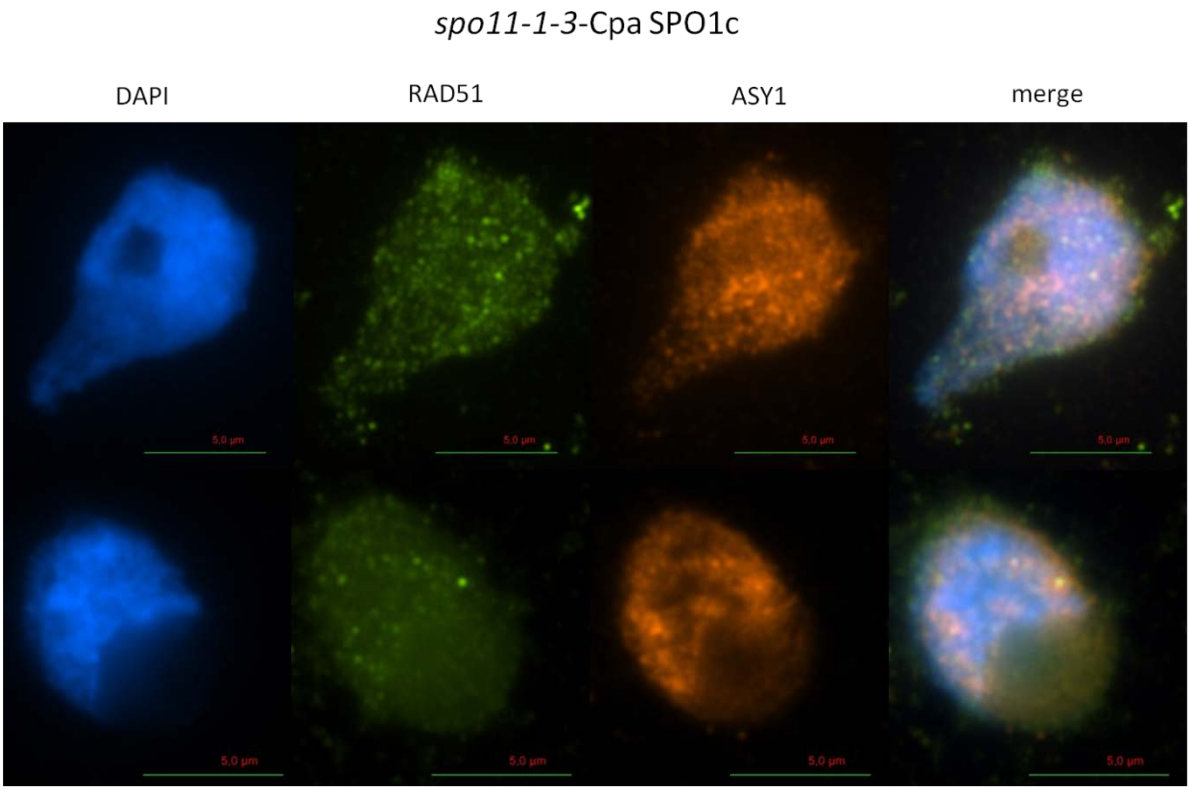
Supplemental figure 7. Relationship between the used organisms for the complementation approach.

Unrooted, unscaled tree, showing the relationship between the different organisms used for the complementation approaches.



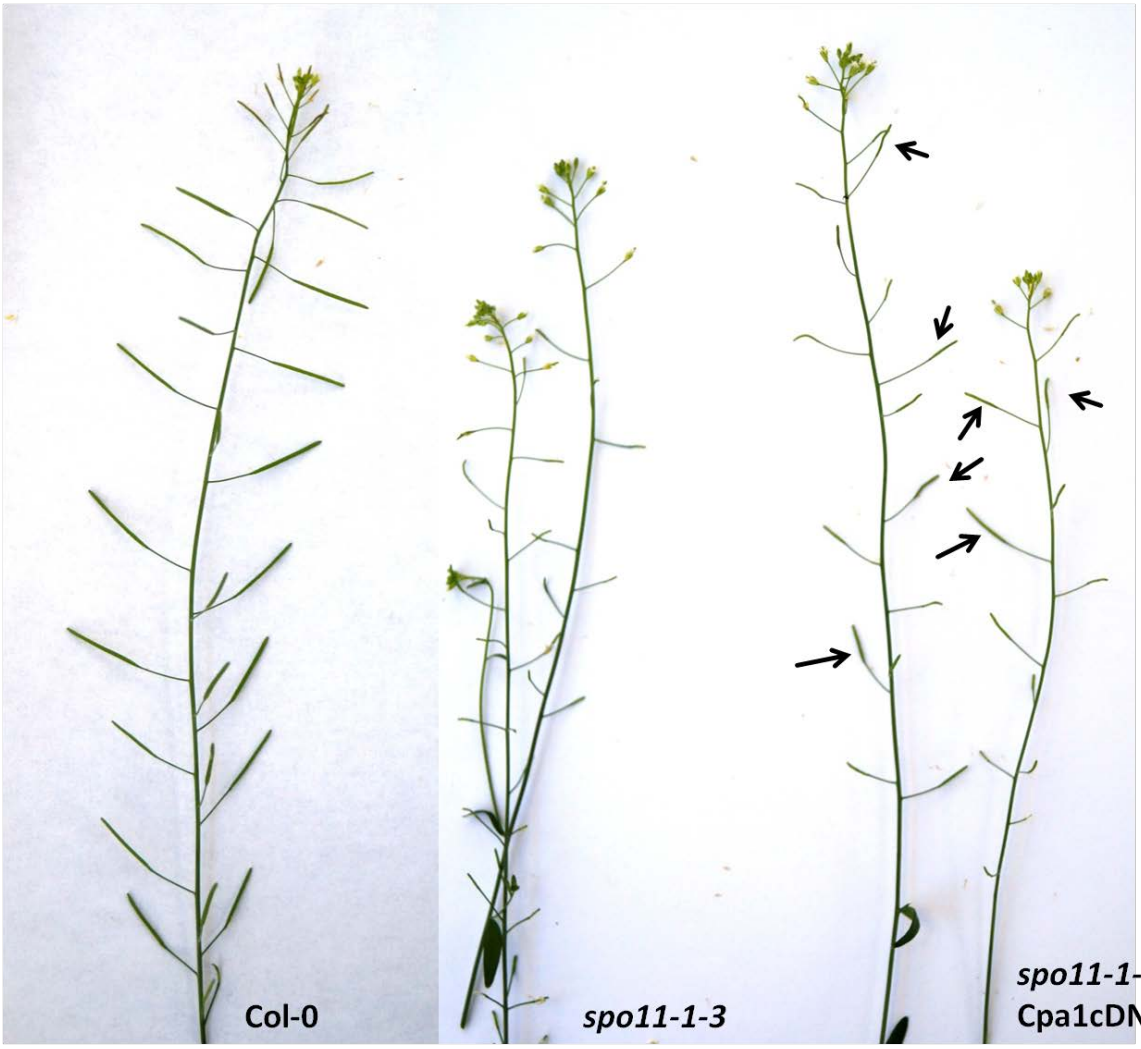
Supplemental figure 8. The Splicing landscape of endogenous SPO11-1 and -2 in *A. thaliana.*

Schematic non scaled shema of the different splice forms of SPO11-1 (A) and -2 (a) from A. *thaliana*. Exons are numbered and shown as white blocks, spliced Introns as black lines. Intron retention events are illustrated as black boxes, alternative 5´splice site selection are shown as blue boxes and alterative 3´splice site selection as light green boxes. In case of exon skipping the corresponding white box is missing. Splice forms are named in greek letters (according to Sprink and Hartung., 2014)



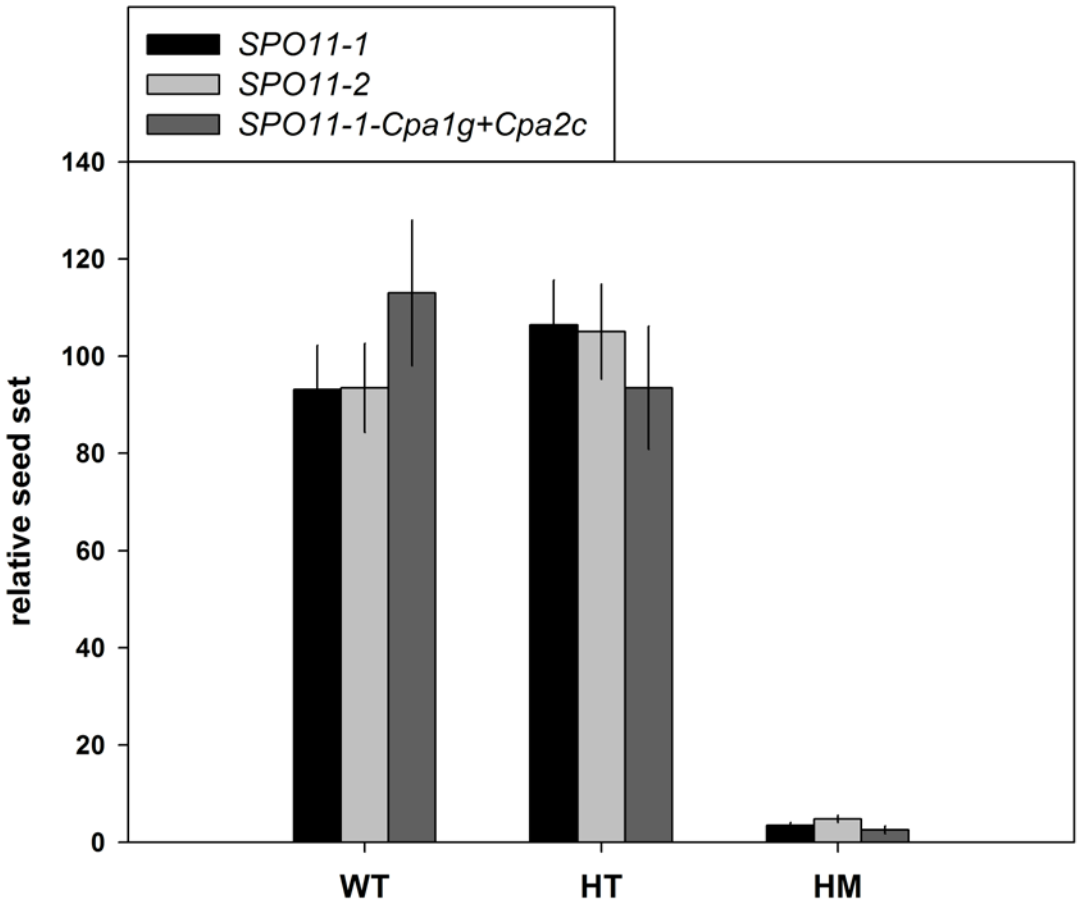
Supplemental figure 9. Dual immunolocalization of ASY1 and RAD51 proteins in meiocytes of *spo11-1-3*-Cpa SPO1c.

Meiocytes were counterstained with DAPI (blue) and dual immunolocalization of ASY1 (orange) and RAD51 (green) was performed using rabbit polyclonal antibody against ASY1 and rat polyclonal antibody against RAD51. Meiocytes in zygotene stage were used for localization. Green bar = 5µm. A restoration of RAD51 loading onto the chromosomes can be observed in



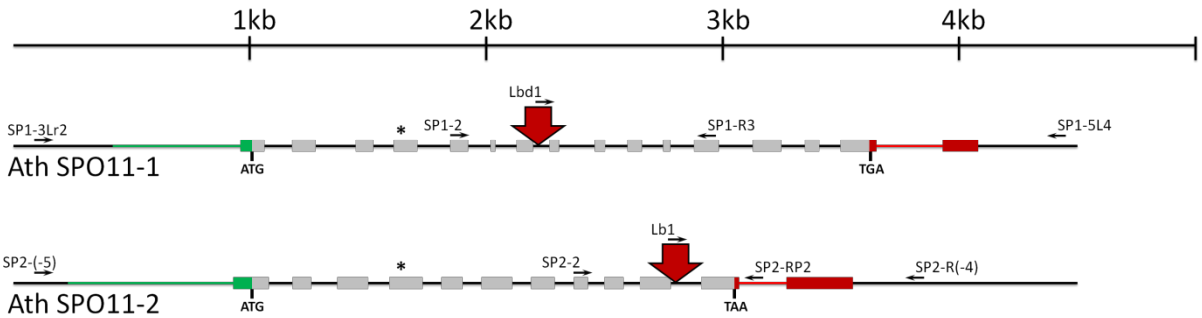
Supplemental figure 10. Partial complementation of the sterile phenotype of *spo11-1-3* by the multiple expression of SPO11-1 cDNA from *C. papaya*.

Flowering stems of wild type (Col-0), *spo11-1-3* and *spo11-1-3* CpaSPO1c are shown. Arrows indicate enlarged siliques in the complementation mutant.



Supplemental figure 11. Seed set in the complementation experiment of *spo11-1-3* Cpa SPO1g with *Spo11-2* cDNA from *C. papaya*.

Relative seed set of *A. thaliana* plants containing an full genomic construct of *C. papaya* SPO11-1 and a complementary DNA construct of *C. papaya* SPO11-2 possesing none (WT) one (HT) or two (HM = homozygous) disrupted alleles of endogenous *SPO11*, compared to the seed set of the respective control carrying no construct (*SPO11-1; SPO11-2*). Seed sets are shown as mean seed set in % ± standard error of the mean (SEM), compared to the respective combined seed set of the corresponding wild type (WT) and heterozygous (HT) plants. (Welch´s T-test; $P < 0,05$)



Supplemental figure 12. In scale exon-intron organization of Ath *SPO11-1* and Ath *SPO11-2.*

Schematic drawing of the Ath *SPO11-1* and *SPO11-2* genes, including the respective promoter region shown as green lines, 5´-UTR regions, represented as green boxes and 3´-UTR regions shown in red boxes including introns (red lines). The T-DNA insertion is depicted as a red arrow. Primers used for genotyping of the respective mutants are shown as black arrows and indicated.


JKI
Julius Kühn-Institut
Bundesforschungsinstitut für Kulturpflanzen