universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Comparative study of plant SR genes regulation and function in *Arabidopsis thaliana* and *Physcomitrella patens*“

verfasst von / submitted by

Monika Maronova

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktorin der Naturwissenschaften (Dr.rer.nat.)

Wien, 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 091 490
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Molekulare Biologie
Betreut von / Supervisor:	ao. Univ.-Prof. Dr. Andrea Barta

Acknowledgements

I would like to thank Andrea for giving me the oportunity to get to know scientific work in her lab, for her advices and encouragement on my journey. I am immensly grateful to Maria, for her kind guidance, patience, inspiration and constant support. I am thankful to all members of the lab for creating a very friendly athmosphere and for help. I wish to thank to our Dundee collaborators Prof John Brown, Dr Craig Simpson and other lab mambers for very fruitful and exciting cooperation. I would like to thank to my revieweres for taking time and effort to read and evaluate my thesis.

Many thanks to my friends for listening, understanding and support. My gratitude goes also to my family for their love, patience and cooperation, especially to my husband Juraj, who did maximum to grant me plenty of time for writing, helped me and encouradged me steadily and constatly.

Abstrakt

Alternatives Spleißen (AS) ist ein hochregulierter Prozess, der mehrere mRNA-Varianten aus einem einzigen Gen erzeugt und zur Transkriptom- und Proteom-Diversität und Feinabstimmung der Genexpression in höheren Eukaryoten wesentlich beiträgt. Einer der Hauptakteure in AS sind SR-Proteine, eine Familie von evolutionär konservierten Spleißfaktoren. Arabidopsis-Genom kodiert achtzehn SR-Proteine und ihre Regulierung und exakte Funktionen sind unklar.

Um eine komplexe Regulierung von AS zu *sezieren*, ist es von großer Wichtigkeit, die Veränderungen in AS unter bestimmten Bedingungen präzise zu überwachen, aber solche Methoden sind in Pflanzen nicht gut etabliert. Die vorliegende Arbeit beschreibt die Entwicklung und Anwendung eines neuartigen hochauflösenden (HR) RT-PCR-Screens, um Veränderungen in mehreren bekannten AS-Ereignissen unter den gewünschten Bedingungen und genetischen Hintergründen zu messen. Das System erwies sich als reproduzierbar, präzise und auch in der Lage, neuartige AS-Transkripte zu erkennen.

Während meiner Dissertation konzentrierte ich mich auf ein pflanzenspezifisches SR-Protein At-RS31. HR-RT-PCR-Screen-Analyse der At-RS31-Überexpression (ÜE) Linie zeigte an, dass At-RS31 AS von Genen, die beteiligen sich am RNA-Metabolismus und -Spleißen, Transkription und DNA-Reparatur, moduliert. Interessanterweise beeinflusste ÜE von At-RS31 das Spleißmuster des DNA-Reparaturgens *UVH1/ATRAD1* und ähnlich wie bei *uvh1* Mutanten waren At-RS31- ÜE -Pflanzen gegenüber der UV-C-Exposition überempfindlich. Erhöhte Levels von At-RS31 verschärfte Effekte von mehreren anderen abiotischen Belastungen und förderten eine frühe Seneszenz, was auf eine wichtige Rolle und eine enge Regulierung der At-RS31-Expression für das Überleben der Pflanzen in wechselnder Umgebung hindeutet.

Die Kenntnis der in Arabidopsis gewonnenen Pflanzen-SR-Gene wurde durch das Studium des Moos Physcomitrella patens ergänzt. Obwohl evolutionärer entfernter Organismus, besitzt Physcomitrella SR-Gene mit hoch konservierter Struktur und AS zu Arabidopsis. Die Charakterisierung von AS in drei orthologen SR-Genen enthüllte ein extensives Spleißen im langen Intron von Pp-RS27 und mehrfache Spleißvarianten von zwei Genen der Pp-RS2Z-Subfamilie. Bemerkenswerterweise zeigte die Analyse von AS-Varianten unter verschiedenen Wachstumsbedingungen, dass AS von Pp-RS27

Veränderungen in Reaktion auf UV-C und Licht gegenüber Dunkel zeigt, ähnlich die Beobachtungen in seinem Orthologe At-RS31, was auf einen hoch konservierten Mechanismus zur Regulierung von AS in Antwort auf externe Anregungen.

Abstract

Alternative splicing (AS) is a highly regulated process that generates multiple mRNA variants from a single gene and contributes greatly to the transcriptome and proteome diversity and fine tuning of gene expression in higher eukaryotes. One of the key players in AS are SR proteins, a family of evolutionary conserved splicing factors. *Arabidopsis* genome encodes eighteen SR proteins and their regulation and exact functions are unclear.

To dissect complex regulation of AS it is of great importance to be able to precisely monitor changes in AS under specific conditions, however such methods are not well established in plants. Present thesis describes the development and application of a novel high resolution (HR) RT-PCR screen to measure changes in multiple known AS events at desired conditions and genetic backgrounds. The system was proven to be reproducible, accurate and also able to detect novel AS transcripts.

During my dissertation I concentrated on a plant-specific SR protein At-RS31. HR RT-PCR screen analysis of At-RS31 overexpression (OE) line indicated that At-RS31 modulates AS of genes involved in RNA metabolism and splicing, transcription and DNA repair. Interestingly, OE of At-RS31 affected splicing pattern of the DNA repair gene *UVH1/ATRAD1* and, similarly to *uvh1-1* mutant, At-RS31 OE plants were hypersensitive to UV-C exposure. Increased levels of At-RS31 aggravated effects of several other abiotic stresses and promoted an early senescence, suggesting important role and tight regulation of *At-RS31* expression for plant survival in changing environment.

Knowledge of plant SR genes gained in Arabidopsis was complemented with the study of the moss *Physcomitrella patens.* Although evolutionary distant organism, it possesses SR genes with highly conserved structure and AS to Arabidopsis. Characterisation of AS in three orthologous SR genes uncovered extensive splicing in the long intron of *Pp-RS27* and multiple splice variants of two genes of Physcomitrella RS2Z subfamily. Remarkably, analysis of AS patterns under different growth conditions revealed that AS of *Pp-RS27* displays changes in response to UV-C and light vs dark similar to those observed in its orthologue *At-RS31*, indicating a highly conserved mechanism to regulate AS in response to external cues.

Contents

Abbreviations

Abbreviation	Definition
ABA	Abscisic acid
AS	Alternative splicing
bp	Base pairs
CaMV	Cauliflower mosaic virus –
cDNA	Complementary deoxyribonucleic acid
CHX	Cycloheximide
CK	Cytokinin
DAG	Days after germination
DNA	Deoxyribonucleic acid
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
EST	Expressed sequence tags
FS	Fully-spliced
hnRNPs	Heterogeneous nuclear ribonucleoproteins
HR RT-PCR	High-resolution reverse transcription polymerase chain reaction
IR	intron retention
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
kb	kilobase
Mb	megabase
mRNA	Messenger ribonucleic acid
NMD	Nonsense-mediated decay
NPTII	Neomycin phosphotransferase II
nt	Nucleotides
PEG	Polyethylene glycol
pre-mRNA	Precursor messenger ribonucleic acid
PTC	Premature termination codon
RNA	Ribonucleic acid
RNA pol II	Ribonucleic acid polymerase II
RNA-Seq	Ribonucleic acid high-throughput sequencing
RNP	Ribonucleic acid recognition motif
RRM	Ribonucleic acid recognition motif domain
RS	Rregion rich in arginines and serines
RT-PCR	Reverse transcription polymerase chain reaction
snRNA	Small nuclear ribonucleic acid
snRNP	Small nuclear ribonucleic particle
SR protein	Serine-Arginin rich protein
UTR	Untranslated region
WT	Wild type

Chapter 1

Introduction

1.1. Pre-mRNA splicing

Eukaryotic nascent pre-mRNA has to go through several processing steps to become a mature messenger RNA (mRNA). One of the pre-mRNA processing steps is splicing.

Splicing is a molecular process of precise removal of intervening segments (introns) and joining of the expressed segments (exons) of the primary transcript. It was discovered by in 1977 in adenoviral mRNAs (Berget et al., 1977; Chow et al., 1977). Splicing was soon after the discovery observed in other organisms, predominantly in eukaryotes, where it occurs in majority of the genes.

Intron removal is carried out by a large complex of proteins and RNAs, called the spliceosome that is considered as one of the most intricate molecular machines in the cell (Nilsen 2003). Systematic studies to dissect of the composition and action of the spliceosome have been made in humans and yeast (Wahl et al., 2009; Will and Lührmann, 2011). Numerous orthologues of human spliceosomal components have been identified in the Arabidopsis genome (Barta et al., 2012), suggesting that the basic mechanism of intron removal is highly conserved among animals and plants.

Two types of spliceosomes have been found, differing in their composition and the splicing signals by which they recognize the introns. The major spliceosome (the U2-type) is in charge of splicing of about 98% of the splicing events, and the minor spliceosome (U12-type) recognizes and processes 1-2% of the introns Minor spliceosome recognizes U12 introns with splice sites that are different from the standard consensus, but highly conserved at the 5' splice site and at the branch point. The composition of minor spliceosome differs from the major spliceosome to the large extent (Sharp and Burge, 1997; Tarn and Steitz, 1996).

1.2. Splicing by the major spliceosome

Introns are defined by 5’ and 3’ splice sites (Figure 1.1A), generally represented as a pair of short degenerate consensus sequences where only the terminal nucleotides are almost invariable, i.e. GU dinucleotide at the 5’ end and a AG dinucleotide at the 3’ end of the intron. Additional *cis*-elements, required for the intron recognition and spliceosome assembly, are the branch-point sequence, which includes an adenine nucleotide important for lariat formation, and the polypyrimidine tract, rich in cytosine and uracile, situated between the branch-point and the 3’ splice site (Figure 1.1A).

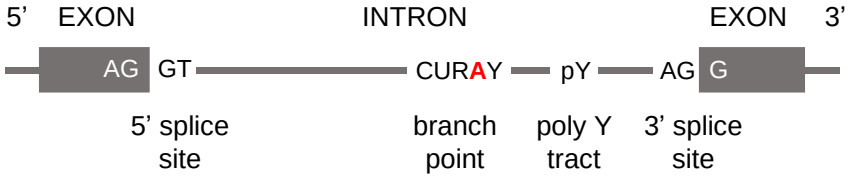
Intron removal by the spliceosome is carried out in two successive nucleophilic transesterification reactions (Figure 1.1B). In the first reaction, the phosphate group of the GU dinucleotide at the 5' splice site is attacked by the 2' hydroxyl group of the adenosine branch-point nucleotide, present in the intron. Exon is cleaved from the intron, and the 5′ end of the intron is ligated to the branch-point adenosine. Next, the 3' hydroxyl group of the detached upstream exon attacks the phosphate group of the 3' splice site, resulting in joining of both exons and release of the intron lariat, which is afterwards degraded (Black, 2003).

The major spliceosome is a highly dynamic ribonucleoprotein machinery, which recognizes the splice sites, assembles and repeatedly rearranges its composition and conformation in a stepwise manner during the catalysis of the splicing (Wahl et al., 2009; Will and Lührmann, 2011).

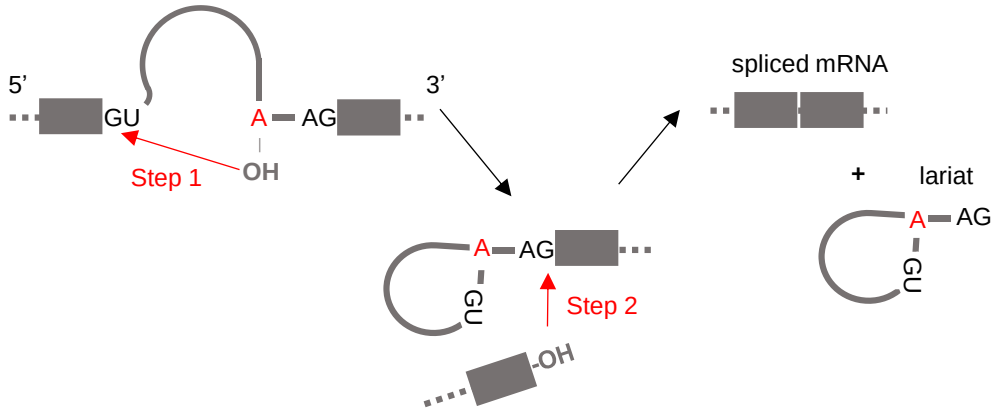
The major spliceosome is composed of five different uridine-rich small nuclear RNA (snRNA) molecules (U1, U2, U4, U5 and U6) and several hundreds of proteins. Each snRNA associates with distinct proteins, forming small nuclear ribonucleoprotein (snRNP) complexes. Five snRNPs form the core of the spliceosome and are responsible for the

Figure 1.1. Splicing by the major spliceosome.
A. Splicing signals of the U2-type intron, recognised by the major spliceosome. Exons are depicted as black boxes, introns as lines. Red A – branch-point adenosine, R – purine, Y - pyrimidine. B. Splicing progresses in two transesterification reactions. Exons are shown as boxes, introns as lines, dashed lines depict the intron currently not involved in the splicing. C. Illustration of the spliceosomal rearrangements during splicing, depicting the core snRNP complexes. Non-snRNPs complexes and auxiliary proteins are not shown. Exons are represented as boxes, introns as lines. Dashed lines depict the nascent transcript at the 3’ end, pol II is RNA polymerase II. Adapted from (Will and Lührmann, 2011).

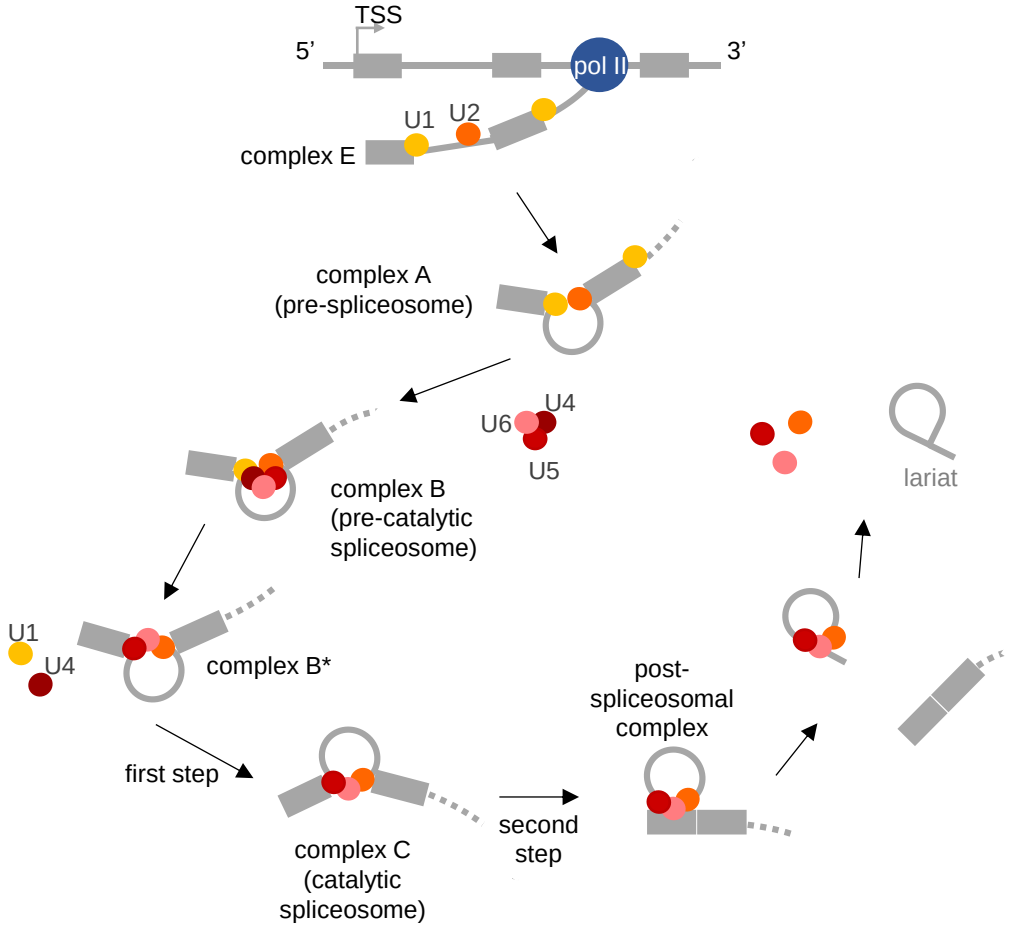
A.


5’
EXON
INTRON
EXON
3’
AG
GT
CURAY
pY
AG
G
5’ splice site
branch point
poly Y tract
3’ splice site

B.


5’
GU
A
AG
3’
Step 1
OH
spliced mRNA
+
lariat
A
AG
GU
A
AG
GU
Step 2
OH

C.


TSS
5’
pol II
3’
U1
U2
complex E
complex A
(pre-spliceosome)
U4
U6
U5
complex B
(pre-catalytic
spliceosome)
lariat
U1
U4
complex B*
post-
spliceosomal
complex
first step
complex C
(catalytic
spliceosome)
second
step

recognition of splice sites and branch-point sequences and for the catalysis of the splicing reaction. Around 170 non-snRNP proteins and spliceosome-associated factors are also dynamically incorporated within the spliceosome. Most of the RNA-RNA interactions between the snRNP and the transcripts are weak, because the splicing sites and the branch point are degenerate, and the splicing factors provide the stabilization of the spliceosome core components on the mRNA.

The spliceosomal cycle (Figure 1.1C) is initiated by the splice site recognition through base pairing of the U1snRNP with the 5' splice site and U2 auxiliary factor (U2AF) with the 3' splice site, resulting in the formation of the complex E. The branch-point sequence is recognised by the U2 snRNP, which eventually interacts with the U1 snRNP and forms the pre-spliceosome (complex A), where the both splice sites get to the proximity. Next, the pre-assembled tri-snRNP (U4-U6·U5 tri-snRNP), consisting of U4, U5 and U6 snRNPs, is attached onto the U2 defining the pre-catalytic complex B. Complex B alters its conformation and composition by releasing the U1 and U4, thereby becoming an activated complex B *, in shape to mediate the first splicing step of the splicing reaction, the cleavage of the exon/intron boundary and forming the intron lariat. Now, rearrangement to the complex C takes place that comprises the downstream exon and the lariat intermediate and facilitates the second catalytic step. After the cleavage of 3' splice site and the splicing of the exons, the post-spliceosomal complex is formed that eventually dissociates and releases the joined exons and the lariat. Intron lariat is degraded and the snRNPs are recycled.

Spliceosome is composed of snRNP complexes and additional about 170 non-snRNP proteins and it is particularly protein-rich molecular machinery. Purification of the spliceosome complexes at different stages revealed dozens of spliceosome-associated factors (Behzadnia et al., 2007; Deckert et al., 2006; Jurica and Moore, 2003). Majority of them are identified in animals and their role in splicing is known. For instance, proteins with helicase (e.g. DEAD box helicases) and isomerase activity provide RNA-RNA interactions within the spliceosome. Large proportion of auxiliary factors is in charge for coupling of the splicing with the transcription. Cross-talk between the spliceosome and pre-mRNA processing enables their coordination and mutual regulation (Kornblihtt et al., 2013).

1.3. Alternative splicing

The usage of particular splice sites in all the gene transcripts was termed constitutive splicing. Different splice site selection during the splicing process leads to the variability in inclusion or exclusion of exons and introns, or their parts, into the mature mRNA in a process called alternative splicing (AS) (Figure 1.2). As a consequence, one gene can produce alternative mRNA products, thereby contributing to the transcriptome and proteome diversity of the cell. Though AS was discovered virtually together with splicing itself, it was considered as an interesting but rare mechanism of expression regulation (Berget et al., 1977; Chow et al., 1977).

Through development of RNA-seq approach, which allows to sequence the whole transcriptome, the estimate of AS frequency has increased to about 95% of multi-exon genes in humans (Barash et al., 2010; Pan et al., 2008) and 61% in Arabidopsis (Marquez et al., 2012), suggesting alternative splicing as a significant contributor to the diversification of the message encoded by a single gene (Kornblihtt et al., 2013).

Transcript isoforms can be generated by different types of AS events (Figure 1.2). Alternative usage of 5’ (donor) or 3’ (acceptor) splice sites leads to shortening or extension of the adjacent exon. Exon skipping or alternative (cassette) exon refer to events where the entire exon is either included or skipped in different transcripts. Inclusion of the entire intron into the mRNA is termed intron retention (IR). Recently identified exitron splicing occurs inside protein-coding exons and removes part of the exonic sequence (Marquez et al., 2012, 2015).

Alternative splicing changes the sequence of the mRNA, thus influences the gene expression. Altered transcripts can code for different protein isoforms that might differ in structure, function, localization and interaction abilities. Proteins with modular structure, such as many transcription and splicing factors, can potentially lack one or more domains, responsible for particular activity localization or interaction (Reddy et al., 2013). Prior to the translation, alternative splicing can control the mRNA localization, transport and translational efficiency.

The stability and turnover of alternative mRNAs can be controlled by alternative splicing, if the transcripts with incorporated premature termination codons (PTCs) are generated. Transcript with PTC can be directed for degradation by nonsense-mediated mRNA decay (NMD) pathway. About 13–18% of intron-containing Arabidopsis genes are

regulated by alternative splicing coupled to NMD (Drechsel et al., 2013; Kalyna et al., 2012).

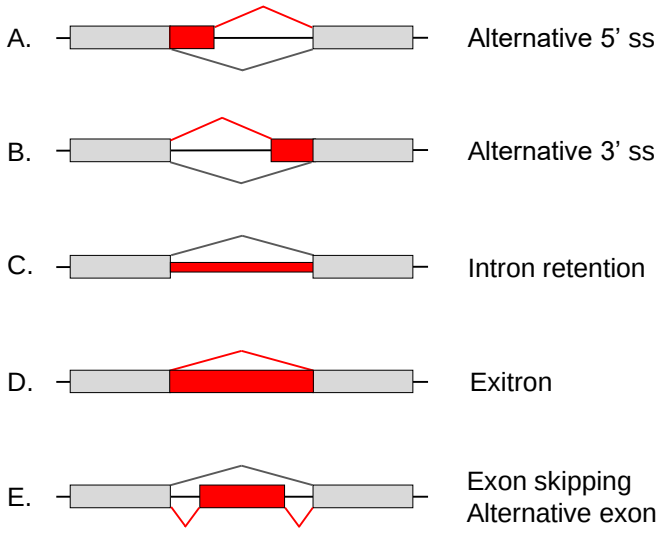


Figure 1.2. Classification of alternative splicing events.
A. and B. Alternative 5' and 3' splice site, respectively C. Intron retention; D. Exitron; E. Exon skipping and alternative inclusion of an exon. Exons are grey boxes and introns are black lines. RNA segments preserved or removed due to AS are shown as red boxes. Carets represent splicing events. Adapted from (Marquez et al., 2015; Syed et al., 2012).

1.4. Alternative splicing regulators: *cis*-acting elements and *trans*-acting factors

Splice sites can be strong or weak, according to the degree of similarity to consensus motifs. The extent to which a splice site is used is assumed to increase as its strength increases (Lim and Burge, 2001). Constitutively spliced exons have generally stronger splice sites than alternatively spliced introns, which more likely have at least one weak splice site that is less efficiently recognised (Kornblihtt et al., 2013; Lim and Burge, 2001).

Splice sites, branch-point sequence and polypyrimidine tract carry insufficient information for initiation of the spliceosome assembly, therefore additional *cis*- and *trans*-acting elements participate to enhance or repress selection of a particular splice site (Black, 2003).

Cis-regulatory elements are short (5-10 nt) sequences that contribute to the recognition of exon-intron boundaries and the regulation of splicing. They can be situated both in exons and introns and provide enhancement or silencing of the splice site usage,

hence their terminology: exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs) (Cartegni et al., 2002; Singh and Valcárcel, 2005). These elements serve as binding sites for *trans*-acting factors, a group of proteins that function as activators or repressors of splicing by controlling the spliceosome assembly. They can recruit or stabilize the spliceosomal components or inhibit other splicing regulators by protein-protein interactions (Figure 1.3). *Cis*-acting elements often occur in clusters, and both enhancers and silencers can be present within the same mRNA molecule, creating a complex composition of regulatory sequences (Black, 2003).

The *trans*-acting factors are mainly represented by two protein families, serine/arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNP). SR proteins are involved in the constitutive pre-mRNAs splicing as well as in regulation of alternative splicing. They were shown to be important also in post-splicing activities, for instance in mRNA nuclear export, subcellular localization and translation (Long and Caceres, 2009). SR proteins possess modular structure that is composed of one or two conserved RNA recognition motifs (RRMs) at the N-terminus, responsible for interacting with the pre-mRNA targets. C-terminal RS region is rich in serine/arginine dipeptides and promotes protein-protein interactions that mediates the spliceosome assembly (Black, 2003). The RS region can be extensively phosphorylated on serine residues and phosphorylation status controls sub-cellular localisation and activity of the SR proteins (Zhou and Fu, 2013).

SR proteins most commonly bind to enhancer sequences and promote the splicing. A small number of SR proteins act as negative splicing regulators by binding to splicing silencer sequences (Black, 2003; Long and Caceres, 2009).

The hnRNP family is a structurally diverse group of RNA-binding proteins, composed of various types of RNA-binding domains, such as RRM, K Homology (KH) domain and glycin-rich regions (RGG), and protein interaction domains (Martinez-Contreras et al., 2007). They act in constitutive and alternative splicing, as well as in other processes of RNA biogenesis, e.g. the transcription, telomere maintenance, DNA binding and repair DNA replication, chromatin remodelling (Han et al., 2010). In regulation of splicing, hnRNPs predominantly bind to splicing silencers and inhibit splicing thereby exert mostly antagonistic effects to SR proteins (Figure 1.3).

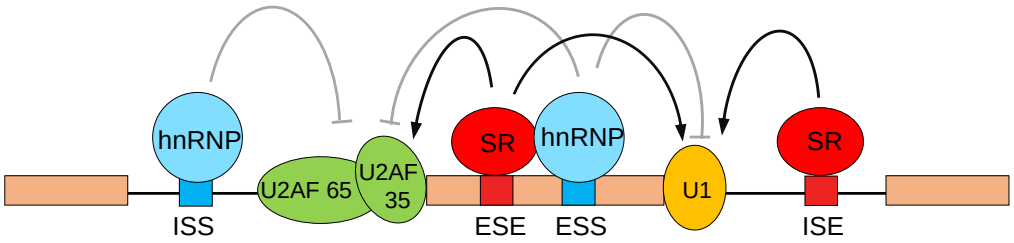


Figure 1.3. Alternative splicing regulation by *cis*-acting elements and *trans*-acting factors. Pre-mRNA splicing is governed by *cis*-acting elements (exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs) and two main families of *trans*-regulatory splicing regulators, serine/arginine-rich proteins (SR) and heterogeneous nuclear ribonucleoproteins (hnRNP). Splicing regulators influence U1 snRNP and U2AF (yellow and green, respectively) that bind to the splice sites adjacent the alternative exon by stimulating (black arrows) or inhibitory (gray lines) effects on the recognition and use of the particular splice site. Adapted from Kornblihtt et al. (2013).

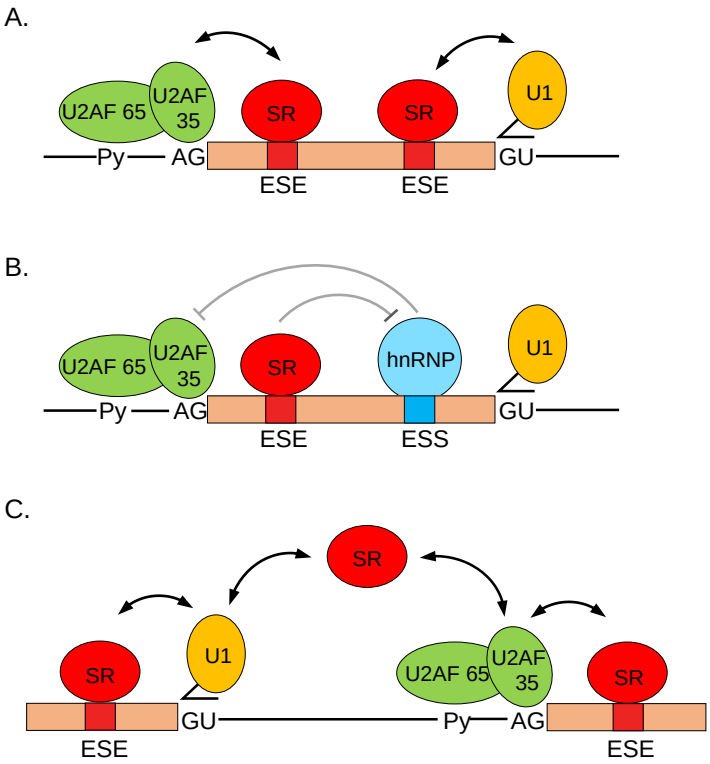


Figure 1.4. Mechanisms of splice site selection facilitated by SR proteins.
A. Recruitment model and exon definition. SR proteins bind to ESE elements and recruit snRNPs to the splice sites. B. Competition model. ESSs are recognized by hnRNPs that inhibit 3' splice site selection by U2AF. SR proteins bound to ESEs can compete and block the action of the splicing repressors and promote splice site selection. C. Intron definition by intron bridging interactions. SR proteins bind factors recognizing the splice sites and juxtapose them. Adapted from Long and Caceres (2009).

The majority of alternative splicing events are governed by the combination of *cis*-elements and the relative abundance and activity of SR factors and hnRNP proteins. The result of this competition influences the ultimate splicing decision (Figure 1.3).

Three mechanisms how SR proteins mediate recognition of exon-intron boundaries are described in Figure 1.4. In the recruitment model, ESE-bound SR proteins recruit and stabilize interactions between the spliceosomal components and the core splice sites, specifically U1 snRNP at the 5' slice site and 35 kDa subunit of the U2AF (U2AF35) at the 3' splice site, in a process known as exon definition (Figure 1.4A). In the inhibitor model, a member of the hnRNP family bound to ESS, inhibits the recognition of the splice site by the spliceosome. SR protein, recognizing the ESE, antagonizes the action of the hnRNP (Figure 1.4B). In the third model, SR proteins do not interact directly with the RNA, instead of that they form a network of protein interactions across intron to bring together the 5' and 3' splice site in the process called intron bridging, or intron definition (Figure 1.4C) (Graveley et al., 2001; Ibrahim et al., 2005; Long and Caceres, 2009).

Accessibility of *cis*-acting elements to splicing factors also plays a significant role in AS regulation. Pre-mRNA secondary structures affect the splicing decision by preventing access of RNA-binding proteins to specific sequence or by changing the relative distances among *cis*-elements (Warf and Berglund, 2010).

1.5. Alternative splicing in plants

Elucidation of specific aspects of pre-mRNA splicing in plants was largely hindered by the lack of *in vitro* splicing system, which was available for yeast and mammals and used to study splicing and its regulation. Nevertheless, plant introns were successfully spliced in HeLa nuclear cell extracts (Brown et al., 1986; Hartmuth and Barta, 1986), but human introns were not processed in plant cells (Barta et al., 1986). This suggested that the core components of the splicing machinery might be shared among plants and animals, but certain specificity in the plant introns architecture and splicing, compared to animals, should exist. The consensus sequences for the 5′ and 3' splice sites and the branch point in plants introns are comparable to core splicing signals in animals (Reddy, 2007; Simpson et al., 2002).

Similarly to humans, majority (about 80%) of plant genes possess at least one intron (Reddy, 2007). Interestingly, plant introns are in average much shorter, around 160 nt

(Iwata and Gotoh, 2011; Marquez et al., 2012) compared to animal introns, which are in average 5 kb long (Sakharkar et al., 2004).

The release of the whole genome sequence of Arabidopsis allowed to identify numerous RNA-binding proteins and orthologues of human splicing factors in the Arabidopsis genome (Lorković and Barta, 2002) which suggests that the basic mechanism of intron removal is highly conserved among metazoans and plants.

The estimates of alternative spicing frequency in Arabidopsis have increased substantially from 1.2% (Zhu et al., 2003) to 61% detected by RNA-seq (Marquez et al., 2012; R. Zhang et al., 2017). Further RNA-seq analyses showed that up to 70% of multi-exonic genes in *Amborella trichopoda* exhibit alternative splicing (Chamala et al., 2015). This suggests that, similarly to animals, alternative splicing in plants represents an important means of gene expression regulation.

Various developmental and environmental signals induce changes to the expression and alternative splicing of splicing factors, inducing adjustments to the transcriptome. These alterations further influence the expression of target genes that carry out changes in the plant growth and the modifications to cope with the environmental conditions (Figure 1.5) (Filichkin et al., 2015; Simpson et al., 2008; Staiger and Brown, 2013; Syed et al., 2012).

Alternative splicing of many genes changes during the development and AS transcripts are expressed in specific tissues and at particular times. In animals, tissue-specific alternative splicing is essential, for instance, for the genes involved in neuronal differentiation, such as neuron-specific splicing factors NOVA in humans (Ule et al., 2005), or in a larvae transitions in worm development (Rukov et al., 2007).

In plant growth and development, alternative splicing is crucial, for example, for the control of the flowering time, and for the regulation of the circadian clock components (Staiger and Brown, 2013; Streitner et al., 2012). Arabidopsis lines with altered levels of splicing factors showed various range of morpholohical phenotypes, indication the importance of alternative splicing for the correct development (Kalyna et al., 2003; Lopato et al., 1999; Reddy and Shad Ali, 2011).

Plants, as sessile organisms, are constantly subjected to various environmental changes. Plants have evolved physiological and morphological adaptations as well as stress-tolerance strategies to cope with the advert environmental cues such as fluctuation in light intensity and quality, diverse plant pathogens, drought, flooding, heat, cold, salinity, and DNA-damaging UV irradiation.

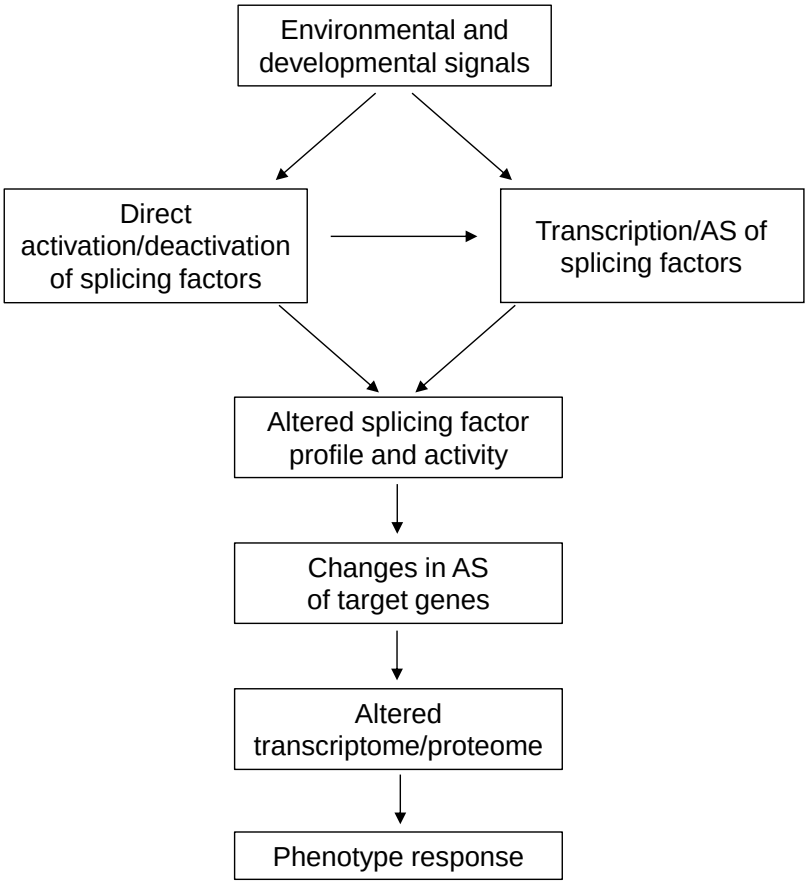


Figure 1.5. Regulation of the gene expression in response to environmental and developmental cues by alternative splicing. (Syed et al., 2012)

Studies of transcriptome changes in plants upon stress-treatment have been focused mainly on the role of transcription factors and less attention was given to the contribution of alternative splicing and splicing factors in stress adaptation. Accumulating genome-wide studies and reports on individual genes indicate that alternative splicing plays crucial role in regulation of this process.

Plant regulatory genes and genes related to stress were shown to be overrepresented in plant AS databases (Duque, 2011; Kazan, 2003; Wang and Brendel, 2006). Studies of ESTs, RT-PCR analyses and deep sequencing studies revealed that stress conditions promote global change in AS profiles indicating participation of AS in stress response processes (Chang et al., 2014; Filichkin et al., 2010; Iida et al., 2004; Palusa et al., 2007; Wu et al., 2014).

Overall increase of IR in plants induced by stress treatment was detected in several genome-wide studies (Chang et al., 2014; Filichkin et al., 2015; Wu et al., 2014). For example, when subjecting *P. patens* protonemata to the heat, numerous proteins

responsible for pre-mRNA splicing responded by increased IR events (Chang et al., 2014). Increased intron retention upon heat stress was demonstrated also in mammals (Shalgi et al., 2014), suggesting that splicing repression in response to heat seems is a general strategy for different eukaryotes.

SR proteins play pivotal role in regulation of alternative splicing. High-light, methyl viologen, salt and heat were shown to have specific effect on relative expression of three AS isoforms of the SR gene *At-SR30* (Filichkin et al., 2010; Tanabe et al., 2007).Wide range of abiotic stresses (NaCl, heat, cold, glucose, hydrogen peroxide) and hormones (ABA, auxin, cytokinin, methyl jasmonate) were shown to impact on expression and alternative splicing of Arabidopsis SR proteins. Salinity, heat and cold strongly affected AS in plenty of SR genes (Palusa et al., 2007). Sequence analysis of AS variants revealed incorporation of PTC into the majority of alternative transcripts and extensive coupling to NMD (Kalyna et al., 2012; Palusa and Reddy, 2010), suggesting that AS and NMD control SR protein abundance in response to various environmental cues.

Light represents crucial energy source for plants, and moreover it is one of the most important environmental signals, triggering many biological processes (H. Zhang et al., 2017).

Genome-wide studies using RNA-seq revealed that light modulates thousands of genes in plants (Shikata et al., 2014; Wu et al., 2014). In Physcomitrella, about 9% of genes responded to the light by alteration of their AS, among them multiple splicing related proteins, including SR proteins and hnRNPs (Wu et al., 2014). Interestingly, light was able to modulate AS of SR protein At-RS31 and increase abundance of protein-coding mRNA1, while in the dark AS transcripts were more pronounced (Petrillo et al., 2014a, 2014b).

Splicing factors and SR proteins in particular are very important mediators of response to environmental factors; however their precise functions and dynamic regulation are still barely understood.

1.6. Plant SR protein family

As indicated in the previous section, plant SR genes are extensively alternatively spliced and their AS splicing patterns are modulated by environmental and developmental cues, indicating their important role in proper plant growth and stress-response (Duque, 2011; Kalyna et al., 2006; Palusa et al., 2007; Reddy and Shad Ali, 2011). Plant SR

proteins have similar domain structure as found in animals suggesting that their roles in pre-mRNA splicing are conserved. However, plant SR protein family comprises almost twice as many SR proteins as identified in humans (12) (Manley and Krainer, 2010), e.g. 18 in Arabidopsis and 22 in rice (Barta et al., 2010). It is likely that the higher complexity originated in several rounds of genome duplications (M. Kalyna and Barta, 2004; Reddy and Shad Ali, 2011; Richardson et al., 2011).

Plant SR proteins can be divided into six subfamilies according to their protein structure (Figure 1.6) (Barta et al., 2012; Kalyna and Barta, 2004). Three subfamilies have their orthologues in humans, whereas other three are specific to plants (RS, RS2Z and SCL) and are considered to fulfil plant-specific functions. Subfamilies are characteristic by high number of paralogous genes, up to four in one subfamily. However, it is not clear

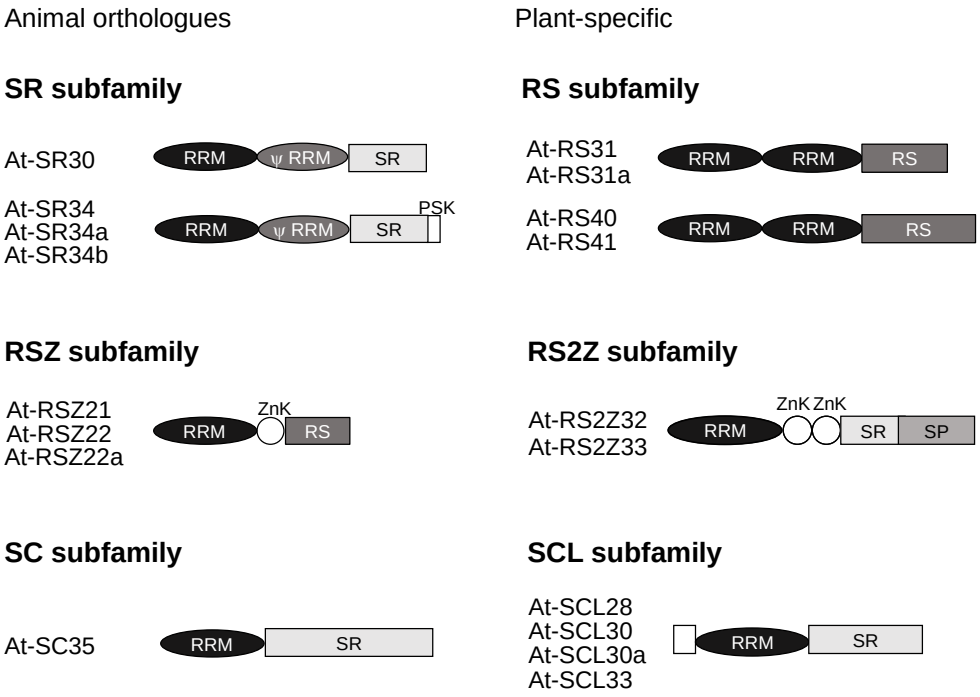


Figure 1.6. SR proteins of *A. thaliana*. Eighteen SR proteins are distributed in to six subfamilies according to their homology. Left column represents SR proteins with orthologues in metazoan, right column illustrates SR proteins specific for the plant kingdom.
Structural domains are termed according to the amino acid specificities: RRM comprise RNA recognition motif, ψRRM is RRM that contains a distinct motif characteristic for SR homologs; SR is rich in serine/arginine dipeptides; RS is rich in arginines and serines; ZnK represents zinc knuckle domain; SP is rich in serines and prolines; PSK is region rich in proline, serine and lysine. SCL proteins possess a plant-specific N-terminal extension. Adapted from (Barta et al., 2010).

whether such closely related proteins differ in their activities and regulation or if they have redundant function (Barta et al., 2012; Kalyna and Barta, 2004).

Some of the AS events in SR genes are highly conserved and can be found also in evolutionary distant land plants, such as mosses (Iida and Go, 2006; Kalyna et al., 2006). Strong preservation together with specific regulation suggests very important function of different AS isoforms of SR proteins, but their precise biological significance remains to be elucidated.

1.7. Moss *Physcomitrella patens* as a model organism for comparative studies of gene functions

Moss *Physcomitrella patens* is non-vascular plant with an interesting evolutionary position, between photosynthetic algae and vascular land plants. *P. patens* is an ancient lineage that diverged from the tracheophytes about 450 million years ago, thus representing one of the earliest plants that moved from aquatic environment to the land (Rensing et al., 2008).

P. patens exhibits alternation of generations, but unlike higher plants, the haploid gametophytic generation is dominant to diploid sporophytic stage. Whereas gametophyte is represented by the majority of the plant tissue, sporophyte is reduced to small spore-producing capsule (Cove, 2005). A haploid spore germinates into filamentous juvenile gametophyte called protonema that consists of two cell types. Chloronema cells are the ones emerging after germination or regeneration from protoplasts or other tissues. Chloronema has assimilation function due to cells with many large chloroplasts and differentiates into fast-growing filaments called caulonemata. Caulonemata can develop buds, which subsequently produce more complex structures, gametophores. Gametophores (leafy shoots) consist of small leaf- and root-like structures (rhizoids). At the apex of gametophores form female and male generative organs that produce gametes. Fertilization generates a diploid zygote that divides and creates multicellular meiotically active sporophyte, producing haploid spores, thus completing the life cycle of *P. patens* (Cove, 2005; Prigge and Bezanilla, 2010; Reddy and Shad Ali, 2011).

Moss is easy to grow and maintain (Cove, 2005; Cove et al., 2006), it enables efficient gene targeting by protoplast transformation (Knight et al., 2002; Schaefer et al., 1991) and recessive phenotypes can be directly recognised due to mostly haploid life cycle.

Despite the evolutionary distance and differences in the physiology and morphology, fundamental genetic, physiological and developmental traits are similar in non-vascular Physcomitrella and in vascular plants. These include elaborated signalling pathways of phytohormones, tolerance to abiotic stresses and DNA-repair mechanisms (Rensing et al., 2008).

The haploid genome of *P. patens* was estimated to 480 Mbp with nearly 36,000 predicted gene models. Average gene structure and exon/intron size is comparable to Arabidopsis (Marquez et al., 2012; Rensing et al., 2008). RNA-seq genome-wide studies estimate the frequency of alternative splicing to 50% of the genes expressed in protonema, with intron retention as a prevalent AS event, likewise in Arabidopsis (Chang et al., 2014; Wu et al., 2014).

P. patens genome encodes at least ten SR proteins, according to the reports of Richardson (2011) and Rauch (2014). They can be distributed into six subfamilies, and many of them are alternatively spliced (Iida and Go, 2006; Kalyna et al., 2006; Rauch et al., 2014; Richardson et al., 2011). Evolutionary analysis of plant-specific RS, RS2Z and SCL subfamilies showed strong conservation of exon-intron positions relatively to the protein domains (Iida and Go, 2006; Kalyna et al., 2006). Moreover, highly conserved alternative splice sites and their surrounding sequences are present in the characteristic long intron, situated between two RNP motifs of the RRM domain in the genes from RS subfamily of dicot and monocot orthologues and in RS2Z orthologues from Arabidopsis to *P. patens* (Kalyna et al., 2006), suggesting existence of conserved mechanisms regulating AS in these genes.

P. patens as a member of an ancestral group of land plants is an invaluable model organism for comparative studies of SR gene functions and evolution of alternative splicing in plants.

1.8. Overall aims of the thesis

The main focus of this thesis was to uncover the roles of plant SR proteins and alternative splicing in regulation of plant development and stress response.

Studies of alternative splicing require tools to accurately measure the ratio of alternatively spliced products. Due to the lack of such tools in plants, we aimed to develop

a platform to monitor the AS changes in multiple genes under various growth conditions and in order to identify alternatively spliced targets of the SR proteins.

We have been particularly interested to elucidate functions of plant-specific SR proteins as they might have evolved to cope with environmental changes and to control plant-specific biological pathways. Specifically, we focused on the members of RS and RS2Z plant-specific families in Arabidopsis and Physcomitrella. In a comparative study of these two species, we intended to get insights on the evolution of SR protein functions in response to external cues as well as on the conservation of the AS regulation mechanisms.

Chapter 2

Development and implementation of the method for monitoring changes in alternative splicing in multiple genes

With raising understanding of AS importance in many plant functions and cell processes and its complex spatio-temporal regulation, a great demand on methods for detection and quantification of AS variants under specific conditions performed in a large scale emerged.

High-throughput technologies provide expression microarrays, designed to measure the total level of gene expression, what makes them incapable to discriminate different splice isoforms. Tiling arrays probe for both coding and non-coding regions, providing a global view on transcriptome, including AS isoforms. For example, Arabidopsis tiling arrays were able to discover new cold-responsive genes that changed in AS. Though, precise quantification of particular transcripts was not performed and relatively high gene expression was required to identify low abundant isoforms (Leviatan et al., 2013). A step forward in the respect of AS quantification was made with the splicing sensitive microarrays that have been successfully used in mammals (Johnson et al., 2003; Pan et al., 2004, 2006; Sugnet et al., 2006; Xing and Lee, 2005), however they are not available for plants.

Whereas previous knowledge of AS events is a prerequisite for designing of splicing sensitive microarrays, high-throughput RNA sequencing (RNA-seq) is able to identify new mRNA isoforms and also it is very sensitive to detect low abundance transcripts. Accurate quantification of multiple similar transcript isoforms is problematic, because of potential mis-assembly of short reads, which might lead to missing of the authentic transcripts or to generation of incorrect mRNA models. Good quality transcriptome databases and subsequent validation and quantification of assemblies by RT-PCR-based medium-throughput methods is critical for performance capacity of RNA-seq analysis (Marquez et al., 2012; Zhang et al., 2015, 2017).

In terms of high quantification accuracy of individual splice variants, the real-time RT-PCR (RT-qPCR) is highly precise method (Vandenbroucke et al., 2001). Its application for the AS change measurement requires specific primers or probes for each of splice variants,

which may not always be possible to design to reliably distinguish and amplify particular AS isoforms. Although PCR products can be precisely measured, for the comparison of multiple variants within one sample and additionally, under various conditions, the normalisation and evaluation becomes intricate, cost- and time-consuming. RT-qPCR may serve preferably as a validation method than a large-scale quantification platform.

Traditional RT-PCR, despite its qualitative rather than quantitative features, has been used in combination with ethidium bromide agarose gel electrophoresis widely as a versatile tool for analysis of alternative splicing (Llorian and Smith, 2012). Modifications, like radiolabelling of the primers might add quantification capacity to the method.

In collaboration with the lab of J.W. Brown, Dundee, Scotland we have developed and applied a novel high resolution (HR) RT-PCR screen designed to monitor 90 splicing events in 89 genes in parallel. Alternative splicing events were selected from published reports, plant alternative splicing databases and EST libraries. Alternatively spliced products were amplified with fluorescently labelled primers and separated of a capillary DNA-analyser that accurately measures abundance of each PCR product in the reaction. Afterwards, we calculated the relative abundance (expressed ratio) of two expected splice isoforms within each reaction and monitored how the ratio changes in dependence of various factors applied.

The aims of the published works were to develop and optimize the method and to test the accuracy and reproducibility of the system. Moreover, we expected to detect statistically significant changes in AS of genes included in the screen as a reaction on growth of wild type plants at light vs. dark conditions as well as in different organs (leaf, flower, root) of wild type plants grown under short- or long-day conditions. By including two transgenic lines stably overexpressing SR proteins At-SR30 (a homolog of human SF2/ASF) (Lopato et al., 1999) and At-RS2Z33 (a plant-specific SR protein from RS2Z subfamily) (Kalyna et al., 2003; Lopato et al., 2002) into the analysis, we intended to detect alternative splicing events modulated by these splicing factors.

The system has proven to be reproducible and accurate to measure even small changes in AS ratio and also proved high resolution to identify small differences (1 nt) in the size of the RT-PCR products, representing splice variants. Numerous significant changes ($P \leq 0.1$) emerged under the applied conditions and genetic backgrounds. Moreover, due to the appearance of new bands of unexpected sizes, the system showed ability to detect putative new splice variants, a large portion of which was confirmed by conventional sequencing as genuine novel isoforms (not shown).

As the study of SR proteins is the subject of my doctorate project, it was interesting to uncover several AS events modulated by overexpression of At-SR30 and At-RS2Z33. The most prominent changes in AS were detected in DNA- and RNA-interacting genes, transcription factors, DNA repair gene *UVH1/ATRAD1*, and genes involved in regulation of flowering time.

Due to successful utilization of the system, the HR RT-PCR screen has been used by us and other groups to study AS in various conditions and plant lines with changed expression of splicing factors (Simpson et al., 2008a, 2008b, 2010), cap-binding complex (CBC) (Raczynska et al., 2010), hnRNP-like RNA binding proteins (Streitner et al., 2012), circadian clock genes (James et al., 2012; Sanchez et al., 2010). Moreover, it has been extended to analyse AS in about 300 different Arabidopsis genes (Simpson et al., 2012) and used to monitor the fate of AS transcripts and their sensitivity to NMD (Kalyna et al., 2012), to identify AS events regulated by nuclear speckle RNA-binding protein upon auxin treatment (Bardou et al., 2014) and to detect AS events modulated by *At-RS31* splicing factor under light/dark conditions (Petrillo et al., 2014a, 2014b).

HR RT-PCR has been successfully used to validate data obtained by RNA-seq analysis whereby assemblies of more than 90% of 586 AS transcripts for 256 genes were confirmed by HR RT-PCR (Marquez et al., 2012). Later it was utilized to validate RNA-seq transcript data quantitatively and contributed substantially to the construction of a comprehensive Arabidopsis reference transcript dataset 1 (AtRTD1) (Zhang et al., 2015). HR RT-PCR was also critical for testing, experimental validation and development of the AtRTD2, a high quality transcriptome for expression and AS analyses (Zhang et al., 2017).

The method was further applied to analyse AS in other species, such as potato, barley and soybean (Calixto et al., 2016; Hancock et al., 2014; Syed et al., 2015).

2.1. Publication

Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts.

Craig G. Simpson, John Fuller, Monika Maronova, Maria Kalyna, Diane Davidson, Jim McNicol, Andrea Barta and John W. S. Brown

The Plant Journal (2008) 53, 1035–1048.

Author's contributions: Preparation of the plant material, preparation of the RNA, optimization of the method and conducting the high resolution RT-PCR screen.

The Plant Journal (2008) **53**, 1035–1048 doi: 10.1111/j.1365-313X.2007.03392.x

TECHNICAL ADVANCE

Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts

Craig G. Simpson[1,*], John Fuller[1], Monika Maronova[2], Maria Kalyna[2], Diane Davidson[1], Jim McNicol[3], Andrea Barta[2] and John W. S. Brown[1]

[1]*Genetics Programme, SCRI, Invergowrie, Dundee DD2 5DA, UK,*

[2]*Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria, and*

[3]*Biomathematics and Statistics Scotland, SCRI, Invergowrie, Dundee DD2 5DA, UK*

Received 15 October 2007; accepted 28 November 2007.

[*]For correspondence (fax +44 1382 562426; e-mail craig.simpson@scri.ac.uk).

Summary

Alternative splicing (AS) increases the proteomic and functional capacity of genomes through the generation of alternative mRNA transcripts from the same gene. AS is now estimated to occur in a third of Arabidopsis and rice genes, and includes genes involved in the control of growth and development, responses to stress and signalling. Regulation of AS reflects the interactions between positive and negative *cis* sequences in the precursor messenger RNA and a range of *trans*-acting factors. The levels and activities of these factors differ in different cells and growth conditions. To identify changes in AS in multiple genes simultaneously, we have established a reproducible RT-PCR panel that can analyse 96 alternative splicing events and accurately measure the ratio of alternatively spliced products. This procedure detected statistically significant changes in AS in different plant organs, in plants grown under different light and day-length conditions, and in plants overexpressing splicing factors. The system provides a convenient, medium-throughput means of monitoring changes in AS in multiple genes. It can readily be applied to much larger or targeted sets of gene transcripts to generate information on the significance and regulation of AS in plant growth and development, specific processes and responses to external stimuli.

Keywords: RT-PCR, splice-site selection, arginine/serine-rich (SR) proteins, non-sense-mediated decay, gene expression.

Introduction

Precursor messenger RNA (pre-mRNA) splicing is the excision of intron sequences from mRNAs mediated by the spliceosome. The accuracy of splicing depends on the recognition of intron and exon signals by protein factors that determine spliceosome assembly. Alternative splicing (AS) produces more than one mRNA from a single gene through the selection and utilization of alternative splice sites in the pre-mRNA (Black, 2003; Blencowe, 2006; Graveley, 2001; Matlin *et al.*, 2005). AS is widespread in higher eukaryotes, with up to 74% of human multi-exon genes containing one or more AS events (Brett *et al.*, 2002; Johnson *et al.*, 2003; Modrek and Lee, 2002). The main consequences of AS are changes in protein structure/function and regulation of gene expression. The inclusion or exclusion of whole or parts of exons or introns can alter protein domains, activity, localization, interactions with other proteins or substrates, and post-translational modification (Black, 2003; Graveley, 2001; Kriventseva *et al.*, 2003; Lareau *et al.*, 2004; Maniatis and Tasic, 2002; Stamm *et al.*, 2005). As such, AS increases the protein complexity of an organism, and gene expression can be affected directly by AS of transcription and splicing factors that modulate DNA recognition and binding, transcriptional complex assembly, and pre-mRNA splicing. In addition, AS can affect mRNA stability and turnover, because many alternatively spliced transcripts contain premature termination codons (PTCs) that are potential

substrates for non-sense-mediated decay (NMD), which recognizes PTCs and targets the mRNA for degradation (Maquat, 2004). Although around 35% of human alternatively spliced transcripts contain PTCs (Lewis *et al.*, 2003), NMD may only be involved in regulating transcript levels of a small number of these transcripts (Pan *et al.*, 2006).

The mechanisms of selection of alternative splice sites involve the recognition of *cis*-acting splicing signals by RNA binding factors to enhance or suppress the use of a particular splice site (Black, 2003; Maniatis and Tasic, 2002; Matlin *et al.*, 2005; Smith and Valcárcel, 2000). Signal sequences are either classical intron splice site or polypyrimidine tract sequences, or sequences found in introns or exons called exonic/intronic enhancers or suppressors (ESE/ISE and ESS/ISS, respectively). Splice-site choice is determined by virtue of their position relative to competing splice sites, and/or through interactions or interference with other proteins or complexes. Factors involved in alternative splice site regulation are arginine/serine-rich (SR) proteins and heterogenous nuclear ribonucleoproteins (hnRNPs) or specific regulatory proteins, which affect the AS of particular gene transcripts or sets of transcripts. SR and hnRNP proteins are families of RNA-binding proteins that have multiple roles associated with splicing, mRNA transport and translation. SR proteins usually contain one or two RNA binding domains of the RNA recognition motif (RRM) type and a reversibly phosphorylated arginine- and serine-rich domain (Fu, 1995; Graveley, 2000). They generally recognize ESE/ISE sequences to promote selection of associated splice sites. On the other hand, hnRNP proteins also contain RNA-binding motifs (RRM and KH domains), but generally interact with splicing sequences to interfere and inhibit selection of associated splice sites, such that SR and hnRNP proteins often act antagonistically (Black, 2003; Maniatis and Tasic, 2002; Smith and Valcárcel, 2000). Therefore, the relative levels of alternatively spliced transcripts from a wide range of genes reflect the different levels of the interacting factors in different cells, tissues or under different conditions, and is referred to as combinatorial control (Smith and Valcárcel, 2000). A second level of regulation is exemplified by the control of AS by specific factors, the expression of which is often restricted to particular cells, tissues or developmental stages. Recent evidence suggests that such factors regulate the AS of genes with related functions, thereby coordinately controlling AS in multiple genes involved in the same biochemical or developmental pathway (Blencowe, 2006; Matlin *et al.*, 2005; Ule *et al.*, 2005).

Bioinformatic estimates of the number of plant genes that undergo AS have increased from 7% to 36.9% (Brett *et al.*, 2002; Campbell *et al.*, 2006; Chen *et al.*, 2007; Iida *et al.*, 2004; Wang and Brendel, 2006). An experimental approach that examined sequences encoding predicted genes on chromosome 2 of Arabidopsis by generating cDNA sequences showed a similar level, with at least 29% of genes showing AS (Xiao *et al.*, 2005). Therefore, although AS appears to occur less frequently in plants than in animal systems, it is clearly significant and affects up to a third of plant genes.

Alternatively spliced genes are involved in a range of plant functions, such as growth and development, signal transduction, responses to biotic and abiotic stress, disease resistance, flowering time, the circadian clock, metabolism and physiology, with regulatory and stress-response genes being particularly well represented (Balasubramanian *et al.*, 2006; Dinesh-Kumar and Baker, 2000; Egawa *et al.*, 2006; Iida *et al.*, 2004; Jia *et al.*, 2004; Jordan *et al.*, 2002; Kazan, 2003; Larkin and Park, 1999; Ner-Gaon *et al.*, 2004; Zhou *et al.*, 2003). Patterns of AS across the genome also change in response to environmental factors and stresses, and in particular temperature (Balasubramanian *et al.*, 2006; Iida *et al.*, 2004). The link with stress and development involves transcriptional, post-transcriptional (including AS) and post-translational effects on various splicing factors. For example, plant SR protein genes have differential transcription patterns, undergo AS and the proteins are phosphorylated (Gao *et al.*, 2004; Gupta *et al.*, 2005; Iida and Go, 2006; Isshiki *et al.*, 2006; Kalyna and Barta, 2004; Kalyna *et al.*, 2003, 2006; Lazar and Goodman, 2000; Lopato *et al.*, 1999; Tanabe *et al.*, 2007). The phosphorylation state of SR proteins in animals regulates their activity in modulating splicing patterns by affecting their ability to bind RNA or interact with other factors in response to external signals (Stamm, 2002). In plants, different stresses can affect the levels of SR protein transcripts and proteins, which in turn affect the AS of downstream genes (Fung *et al.*, 2006; Lazar and Goodman, 2000; Tanabe *et al.*, 2007). In addition, consistent with their influence over the AS of a wide range of genes, overexpression of SR proteins and lammer kinase (a kinase involved in phosphorylating SR proteins; Savaldi-Goldstein *et al.*, 2003) induces many developmental and growth defects, and directly effects the altered AS of pre-mRNAs of particular genes (Kalyna *et al.*, 2003; Lopato *et al.*, 1999). Thus, environmental and developmental cues affect the transcriptional and AS regulation of these factors, which in turn modulate AS to control metabolic and developmental processes (Balasubramanian *et al.*, 2006; Fang *et al.*, 2004; Golovkin and Reddy, 1998; Iida *et al.*, 2004; Kalyna *et al.*, 2003; Lopato *et al.*, 1999, 2002).

Our understanding of AS in plants is not as well developed as in animal systems, because there are far fewer publicly available expressed sequence tags (ESTs), compared with human and mouse, and a lower depth of full-length sequencing of cDNAs, such that many events have not yet been detected and have incomplete annotation (Xiao *et al.*, 2005). For the vast majority of the reported alternatively spliced plant genes there is little or no information on functional differences among the proteins produced, or any mechanistic understanding of how alternative splice sites

are selected. In addition, nothing is known about higher order combinatorial control or coordination of the AS of sets of genes. The advances made in understanding AS in human and mouse are based not only on genomic and bioinformatic analyses of the large numbers of EST and cDNA data, but also on extensive RT-PCR and high-throughput exon microarrays. To begin to address some of the principles in plant AS, we have established a new AS-RT-PCR-based panel that directly measures changes in 90 AS events simultaneously. The utility of this approach was demonstrated by detecting significant changes in splicing in seedlings grown under different conditions and in transgenic lines overexpressing plant SR splicing factors.

Results

Alternative splicing events represented on an RT-PCR panel

For the 96-well format, 90 AS events from 89 genes were selected, along with six control genes that either do not undergo AS or undergo well-documented AS. Events were selected from published reports and by screening databases at the Riken Genomic Science Centre (http://rarge.gsc.riken.jp/a_splicing/index.pl), the Institute for Genomic Research (TIGR; http://www.tigr.org/tdb/e2k1/ath1/splicing_anomalies.html), the Hangzhou Genomics Institute (Zhou *et al.*, 2003) and, latterly, the Alternative Splicing in Plants (ASIP) database (http://www.plantgdb.org/ASIP). To demonstrate the utility of the system, a range of 5′ and 3′ AS events were selected that had a minimum of 11 nt between the alternative splice sites, to increase the probability that alternative splice site usage was regulated. Of the 90 events, 31 used alternative 5′ splice sites and 59 used alternative 3′ splice sites (Table S1). The majority of events involved AS in the coding region, but events involving 5′ and 3′ untranslated region (UTR) sequences were also included (Table S1). The events were not selected on the basis of the biological function of the genes themselves. As controls, RT-PCR was carried out over intronless regions of transcripts of ubiquitin 3, ribosomal protein L12C and RSp31, constitutively spliced introns of actin 11 and *FY*, and the U12-dependent intron of *LUMINDEPENDENS* (Lewandowska *et al.*, 2004). AS controls were rubisco activase, which is expressed constitutively and alternatively spliced with a ratio of approximately 1:1 in all tissues studied so far, and SR protein genes, the AS of which is well documented (Kalyna *et al.*, 2006). These were included both as controls and to monitor changes in their AS in response to growth conditions.

Optimization of AS RT-PCR panel

The procedure for measuring AS changes was optimized by assessing different RNA extraction and clean-up methods, starting RNA concentrations, PCR clean-up methods and developing efficient data extraction for statistical analyses. Briefly, total RNA was isolated from 50–100 mg of ground Arabidopsis seedlings and plant organs. Following DNase treatment, RNA was purified on Qiagen columns. An RT reaction was performed using oligo dT primers, and then aliquoted into a 96-well plate. Each well contained gene-specific primers, one of which was fluorescently labelled. Primers were designed to generate PCR products of between 60 and 700 nt, and to amplify across a region undergoing AS to directly detect different AS products in the same reaction. PCR was carried out for 24 cycles (within the linear amplification range), giving an estimate of transcript levels (see below). One hundredth of the reaction mix was separated on an automated capillary DNA sequencer (ABI3730; resolution $\pm$ 0.5 nt), and the detected bands were analysed using GeneMapper software (designed for fragment/genotype analysis). RT-PCR product size and peak area were assembled in a database, and the expected alternatively spliced products were extracted. Quantification of each alternatively spliced PCR product allowed the determination of the ratio of alternatively spliced products in each reaction, and this value was used to compare treatments to identify changes in AS. At least three biological repetitions were carried out, which allowed statistical analyses to be performed.

Optimal PCR cycle numbers were determined by analysing the ratios of AS products for 24 different gene transcripts using 20–26 PCR cycles. With different cycle numbers, the ratio of alternatively spliced products was constant for highly expressed genes from 20–24 cycles (Figure 1a,b), whereas for genes with low levels of expression, products were difficult to detect with 20 cycles, but amplification ratios were constant in the 22–26-cycle range (Figure 1c), such that, overall, 22–24 cycles was optimal. The maximum standard error (SE) of the mean in these experiments was $< \pm 5\%$, with the majority (22 out of 24 samples) having an SE of $< \pm 3\%$. Experimental variation was determined by RT-PCR at 24 cycles of the 24 AS events, and comparing the ratios of products obtained with the same RNA in three separate experiments. The maximum standard error was $\pm$ 1.5%, with half of the AS events having a standard error of $< \pm 0.5\%$ (data not shown). Thus, the system is quantifiable and highly reproducible. Experiments with the full RT-PCR panel were therefore carried out at 24 cycles, and used three or four biological repetitions. The significance of detected changes in AS was determined statistically for each series of experiments.

Initial experiments analysed RNA from seedlings grown under light and dark conditions, and demonstrated the ability of the RT-PCR system to detect a number of different splicing phenotypes (Figure 2). Firstly, alternatively spliced transcripts were detected for three-quarters of the AS events in the conditions used (70/89 events). Secondly, significant changes in AS were detected between light- and dark-grown

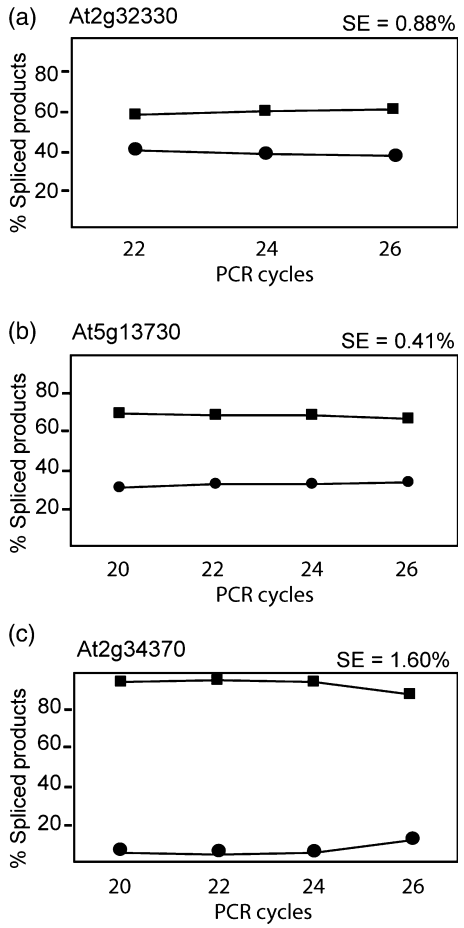


Figure 1. Quantification of the relative levels of alternatively spliced products with different PCR cycle regimes.
Representative graphs of PCR products for different primer pairs are shown for genes with low (a) and higher (b and c) expression levels over 22–26 cycles (a) or 20–26 cycles (b and c). Circles and squares represent the percentage of the total spliced products of the two different alternatively spliced forms. The standard error (SE) between the different cycles for three technical samples is shown.

seedlings for 10 of the 89 events (see below). Thirdly, even within the relatively limited number of genes in the panel and the size of regions amplified, some primer pairs showed major, unexpected RT-PCR products. Twenty of the primer pairs detected products in only one treatment, which may reflect differential transcription in light and dark. For a small number of genes either no transcripts (8/89) or only a single transcript (11/89) was detected (Figure 2). Lack of detection was not because of inefficient priming, because the primers amplified products from DNA. Lack of detection therefore reflects low overall expression levels, or expression in particular cells or tissues relative to the plant material used to extract RNA (all of the genes where transcripts were not detected have low expression, as judged by MPSS; http://mpss.udel.edu/at). Similarly, the lack of detection of an alternatively spliced product may reflect the low abundance of the alternatively spliced transcript in the particular conditions used, its occurence in a small number of cells relative to the plant material used for RNA extraction, or an AS event has been annotared by virtue of a rare EST. Finally, in two cases, the amplified products were not of the expected sizes. Manual inspection of the gene sequences identified 5′ and 3′ splice sites, which would generate the new products and would probably reflect errors in annotation.

Alternative splicing in light- and dark-grown seedlings

The AS RT-PCR panel was used to monitor AS in RNAs extracted from four biological repetitions of 10-day-old seedlings grown in 16-h light/8-h dark conditions, and etiolated seedlings grown in the dark. Although the AS events on the RT-PCR panel were not selected on the basis of biological function, nine genes showed statistically significant changes ($p < 0.01$) in the ratios of AS transcripts between light and dark treatments (Table 1; Figure 3). Eight of the nine events showed an increased use of the distal site in the light-grown material. The consequences of the AS events showing significant changes were altered N-terminal or C-terminal sequences, the addition/loss of amino acids or the generation of premature termination codons, which potentially lead to either production of truncated proteins or NMD (Table 1; Figure 3). In some cases, AS occurred in the 5′ or 3′ UTR, such that the coding sequence of the mRNA appeared to be unaffected. These changes could potentially affect the stability or translatability of the mRNAs.

Two of the AS events led to differential inclusion of putative signal peptides. In carboxypeptidase (At1g71696) transcripts, the alternative intron in exon 3 (Figure 3) removed the authentic translation start site, and altered the N-terminal sequence of the protein, thereby including a putative signal peptide. This gene is alternatively spliced in two different regions (Figure 3), but only the AS of exon 3 responded to the dark/light conditions. Similarly, AS of the tyrosyl tRNA synthetase (At1g10590) transcripts generates a new translation start codon, producing a putative signal peptide. Interestingly, two of the genes that showed significant changes were SR proteins, which are modulators of AS (Figure 3).

Alternative splicing in different organs of plants grown under long- and short-day conditions

Significant changes in AS were observed among different organs (root, leaf and whole flowers) in plants grown under different day length regimes using three biological repetitions. The results are presented in terms of significant changes in AS between organs (Table 2) and between long and short days (Table 3). Seventeen AS events showed significant changes between different organs, of which three also showed changes in both long- and short-day

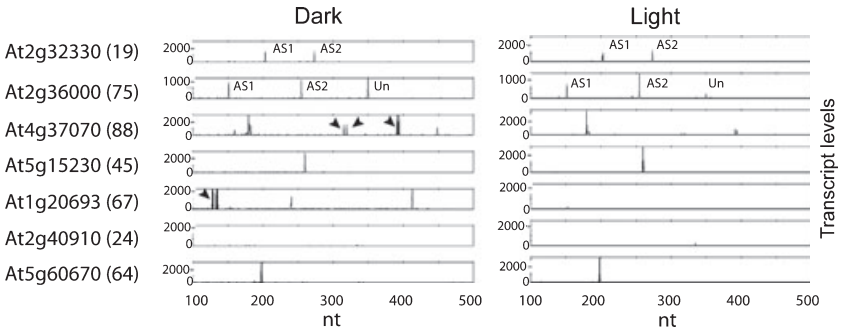


Figure 2. Electropherograms of RT-PCR products of various genes illustrating the range of splicing phenotypes.
Transcript analysis of seven different genes in dark- and light-grown plants showing different splicing phenotypes: no change in levels of alternative splicing (AS) transcripts (At2g32330, primer pair 19); change in ratios of AS transcripts (At2g36000, 75); change in ratios of AS transcripts plus unexpected products (arrowheads) (At4g37070, 88); no AS detected (At5g15230, 45); detection of transcripts in one treatment only (At1g20693, 67); no transcripts detectable (At2g40910, 24); and actin control (At5g60670, 64). Key: AS1 and AS2, RT-PCR product representing alternatively spliced mRNAs; Un, unspliced product; arrowheads, novel products. Numbers on the *x*-axis represent the size markers in nt; numbers on the *y*-axis represent relative fluorescence, reflecting transcript abundance.

Table 1 Genes showing significant ($P \leq 0.1$) alternative splicing changes between light- and dark-grown conditions

	Alternative splicing events				Dark (%)[a]		Light (%)[a]		
Gene identifier	Event	Position	Dist. (nt)	Consequence	Distal	Proximal	Distal	Proximal	*P*-value
At5g13730	Alt 3′ss	IVS2	14	Truncated protein – PTC	**58**	42	**69**	31	0.001
At1g71696	Alt 3′ss[b]	IVS3	23	Change in N-terminal	49	51	70	30	0.026
	Alt3′ss	IVS11/Ex12	31	sequence – signal peptide Truncated protein – PTC	0	**100**	0	**100**	
At2g37340	Alt 3′ss	IVS2	218	Truncated protein – PTC	**83**	17	**94**	6	0.033
At2g43160	Alt 3′ss	Ex2	18	Alternative 5′ UTR sequences	6	94	9	91	0.038
At1g10590	Alt 3′ss	IVS1	45	Alternative ATG adds 14 aa to N-terminal – putative signal peptide	91	9	97	3	0.049
At3g57520	Alt 3′ss	Ex14	34	Alternative stop codon – changes and shortens C-terminal end	0.5	99.5	2	98	0.053
At2g36000	Alt 5′ss	Ex1/3′ UTR	104	Change in C-terminal end	44	56	38	62	0.058
At3g03890	Alt 5′ss	IVS11	20	Change in C-terminal end	98	2	99	1	0.061
At2g21620	Alt 5′ss	IVS 1	18	In frame ± 6aa	90	10	97	3	0.100

Genes: At5g13730 – plastid-encoded RNA polymerase sigma subunit D; At1g71696 – carboxypeptidase; At2g37340 – RSZ33; At2g43160 – clathrin-binding protein (Epsin); At1g10590 - tyrosyl-tRNA synthetase isolog; At3g57520 – imbibition protein homolog; At2g36000 – mitochondrial transcription termination factor-related; At3g03890 – pyridoxamine 5′ phosphate oxidase-related, FMN-binding; At2g21620 – universal stress protein RD2 (auxin regulated). Abbreviations: aa, amino acids; Ex, exon; IVS, intervening sequence or intron; PTC, alternative splicing event generates a premature termination codon truncating the protein.
[a]The results are the means from four different experiments using biological repetitions. For AS events that generate truncated proteins, the figures in bold represent the percentage splicing value for the splicing event that generates the full-length protein as annotated (TAIR).
[b]The alternative 3′ splice site occurs in an alternative intron, which is spliced or unspliced in different transcripts to include or exclude a signal peptide sequence.

treatments. Increased use of both distal and proximal 5′ and 3′ splice sites was observed among the AS events. The genes showing significant changes in AS included AS factors (SR proteins), transcription factors, and a putative nucleic acid binding protein. Some of the largest changes were increased use of the distal 3′ splice site of At5g13730 in leaf, and of At1g09140 in leaf and flower; whereas in At4g35450 there was an increased use of the distal 5′ splice site in root (Table 2). The carboxypeptidase gene (At1g71696) showed significant changes in splice site usage among all three organs tested under long-day conditions. In short days, the transcriptional regulator, Sir2 (At5g09230) and At4g12790 had greatly increased usage of the distal 3′ splice site in leaf and flower/root respectively (Table 2). Use of the alternative splice site in exon 2 of At4g12790 either introduces premature stop codons into the authentic open reading frame (ORF) or, by selection of a second reading frame, generates part of the authentic ORF with a shorter, different N-terminal sequence that encodes a putative signal peptide. Six genes showed significant differences in AS between short- and long-day grown plants in particular organs (Table 3). Of these, three (At4g16845, At2g36000 and At2g37340) also showed significant differences among different organs (Table 2).

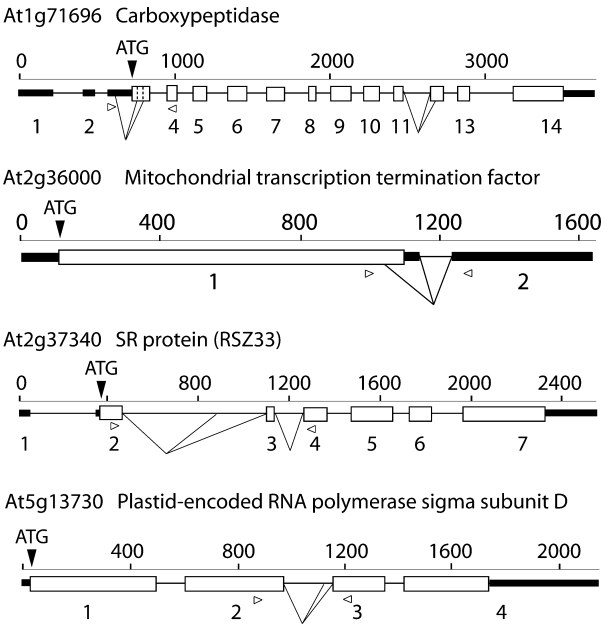


Figure 3. Gene structures and alternative splicing events of selected genes, with significant changes in alternative splicing (AS) between light- and dark-grown seedlings.
The size of the gene, number of exons and translation start site (ATG) are indicated. The analysed AS events are indicated by lines below the gene, and the positions of the primers used are shown by open arrowheads. Key: white boxes, coding exons; black boxes, non-coding, untranslated region (UTR) exons; lines, introns.

Characterization of novel alternative splicing events

Twenty-three primer pairs amplified major RT-PCR products in addition to the expected products. To show that these products arose from the gene in question and represented novel AS events, we cloned products from 20 primer pairs and sequenced around 300 clones. Although not all RT-PCR products were recovered, we identified many of the new products as bona fide AS events. For example, analysis of a MADS box protein AGL31 FLM (At5g65050) identified a novel, major 328-bp product (AS3, Figure 4a). This new alternatively spliced product retained intron 3, whereas introns 2 and 4 were removed. Sequencing also identified minor products, showing utilization of a novel 5′ splice site in the upstream intron (†, Figure 4a) and a novel 3′ splice site 4 nt downstream of the authentic 3′ splice site in intron 4 (*, Figure 4a). The sizes of some of the novel RT-PCR products identified AS transcripts, which are now annotated in TAIR. For example, AS3 of the transcriptional regulator Sir2 family protein (At5g09230) represents a novel exon skipping event, which leads to the production of differently sized proteins through the presence and absence of different translational start codons (Figure 4b). Finally, in a few cases, AS was highly complex with many of the RT-PCR products being identified by either sequencing or being represented in TAIR. For example, the gene encoding an unknown protein (At3g53270) generates eight different 5′ UTR sequences, five of which were annotated in TAIR, and three of which were characterized here by cloning and sequencing (Figure 4c). These AS events included the use of alternative 5′ and 3′ splice sites, intron retention and exon skipping (Figure 4c). For the purpose of demonstrating the utility of the AS RT-PCR panel, only the results for the original selected AS event are presented.

Alternative splicing in transgenic plants overexpressing SR proteins

SR proteins are important splicing factors in both constitutive splicing and AS. We have shown previously that in lines overexpressing SR proteins, the AS of specific transcripts is affected (Kalyna *et al.*, 2003; Lopato *et al.*, 1999). We examined changes in AS in transgenic lines overexpressing the plant SR proteins RSZ33 and SRp30 using three biological repetitions in each case. RSZ33 is a plant-specific SR protein, whereas SRp30 is a plant orthologue of animal SF2/ASF. The SR protein genes are themselves alternatively spliced, with the AS mainly confined to the long intron within each of the genes, and the major AS isoforms contain premature stop codons (Figure 5). Overexpression of the SR proteins caused significant changes in the AS patterns of some of the AS events on the panel, consistent with the role of SR proteins in regulating AS (Table 4; Figures 5 and 6). Overexpression of SRp30 affected the splicing pattern of its own pre-mRNAs (Figure 5), consistent with previously reported changes (Lopato *et al.*, 1999) demonstrating the accuracy and reliability of the system. Besides the autoregulation of SRp30, overexpression of SRp30 and RSZ33 induced significant changes in the transcripts of sixteen genes on the RT-PCR panel (Table 4; Figure 6). Firstly, overexpression of RSZ33 significantly affected the AS of another SR protein, RSp31 (as previously detected; Kalyna *et al.*, 2006), and RNA-interacting proteins, including an hnRNP-like protein (At5g66010). The reduction in use of the distal 3′ splice site in RSp31 (At3g61860) could potentially reduce the levels of this protein substantially, as this AS product encodes the protein (Table 4; Figure 6). As with the growth-condition treatments and different organs, overexpression of SR proteins caused use of either distal or proximal splice sites, leading to changes in amino acid sequence or the inclusion of premature termination codons. For example, overexpression of SRp30 led to increased the use of the distal 3′ splice site in At5g41150 (RAD1), but decreased the use of the distal 3′ splice site in At4g12790.

Discussion

Recent analyses have demonstrated the widespread occurrence of AS in plants, and the range of essential functions in which it is involved (Campbell *et al.*, 2006; Chen *et al.*, 2007;

Table 2 Genes showing significant ($P \leq 0.1$) alternative splicing (AS) changes in different organs

					Leaf (%)[a]		Flower (%)[a]		Root (%)[a]		
Gene identifier	Event	Position	Dist. (nt)	Consequence	Distal	Prox.	Distal	Prox.	Distal	Prox.	*P*-value
Long days											
At1g71696	Alt 3′ss[b]	IVS3	23	Change in N-terminal sequence – signal peptide	59	41	70	30	44	56	0.000
At5g13730	Alt 3′ss	IVS2	14	Truncated protein – PTC	**77**	23	**60**	40	nd	nd	0.001
At4g16845	Alt 5′ss	IVS5	22	Truncated protein – PTC	96	**4**	97	**3**	91	**9**	0.002
At2g37340	Alt 3′ss	IVS2	218	Truncated protein – PTC	**92**	8	**94**	6	**97**	3	0.006
At5g65050	Alt 3′ss	Ex 4	33	In frame ± 11 aas	9	91	10	90	17	83	0.008
At1g09140	Alt 3′ss	IVS10	338	Truncated protein – PTC towards C-terminal	**71**	29	**66**	34	**48**	52	0.039
At4g35450	Alt 5′ss	IVS1	44	Alternative 5′ UTR sequences	69	31	68	32	80	20	0.047
At3g10770	Alt 5′ss	Ex2	24	In frame ± 8 aas	6	94	8	92	5	95	0.055
At3g14230	Alt 5′ss	Ex1	12	In frame ± 4 aas	98	2	96	4	98	2	0.060
At1g23970	Alt 5′ss	IVS3	15	In frame ± 5 aas	15	85	20	80	21	79	0.076
At2g36000	Alt 5′ss	Ex1/3′ UTR	104	Change in C-terminal end	36	64	38	62	49	51	0.082
Short days											
At1g78810	Alt 5′ss	IVS3	56	Small change in C-terminal	74	26	80	20	72	28	0.001
At2g37340	Alt 3′ss	IVS2	218	Truncated protein – PTC	**96**	4	**94**	6	**99**	1	0.004
At4g12790	Alt 3′ss	5′ UTR	126	uORF	**33**	67	**47**	53	**47**	53	0.013
At2g39730	Alt 5′ss	IVS6	11	Alternative stop codon – changes and shortens in C-terminal end	50	50	54	46	52	48	0.049
At1g52500	Alt 3′ss	Ex8	158	Alternative stop codon – changes and shortens in C-terminal end	86	14	77	23	92	8	0.052
At5g09230	Alt 3′ss	Ex2	39	Alternative 5′ UTR sequences	41	59	29	71	28	72	0.054
At3g53270	Alt 3′ss	IVS1	96	Alternative 5′ UTR sequences	63	37	73	27	58	42	0.072
At3g10770	Alt 5′ss	Ex2	24	In frame ± 8 aas	8	92	11	89	5	95	0.076
At3g57520	Alt 3′ss	Ex14	34	Alternative stop codon – changes and shortens C-terminal end	1.5	98.5	0.7	99.3	0.5	99.5	0.081
At5g65050	Alt 3′ss	Ex4	33	In frame ± 11 aas	13	83	11	89	16	84	0.084

Genes: At1g71696, carboxypeptidase; At5g13730, RNA polymerase sigma subunit D (Sig D); At4g16845, vernalization 2 protein (VRN2); At2g37340, RSZ33; At5g65050, MADS-box protein AGL31 FLM; At1g09140, SF2/ASF-like (SRp30); At4g35450, ankyrin repeat-containing protein; At3g10770, unknown function, single-stranded nucleic acid binding R3H domain containing; At3g14230, transcription factor EREBP like; At1g23970, unknown protein (Auxin regulated); At2g36000, mitochondrial transcription termination factor-related; At1g78810, expressed protein; At4g12790, hypothetical ATP binding protein; At2g39730, rubisco activase; At1g52500, formamidopyrimidine–DNA glycosylase; At5g09230, transcriptional regulator (Sir2); At3g53270, unknown; At3g57520, imbibition-protein homolog. Abbreviations: aa, amino acids; Ex, exon; IVS, intervening sequence or intron; PTC, alternative splicing event generates a premature termination codon truncating the protein; IVS, intervening sequence or intron.

[a]The results are the means from three different experiments using biological repetitions. For AS events that generate truncated proteins, the figures in bold represent the percentage splicing value for the splicing event that generates the full-length protein, as annotated (TAIR).

[b]The alternative 3′ splice site occurs in an alternative intron that is spliced or unspliced in different transcripts to include or exclude a signal peptide sequence.

Kazan, 2003; Wang and Brendel, 2006; Xiao *et al.*, 2005;). The function of AS in the regulation of expression is well established in animal systems, and the growing number of characterized examples in plants demonstrates its importance. Major issues concerning AS in plants are the mechanisms of alternative splice site selection, the functions of different proteins produced from the same gene, other consequences of AS on expression, such as NMD or translational blocks, and combinatorial and coordinated control of AS. A prerequisite for many of these areas is the ability to study multiple AS events at the same time, to identify changes in patterns of AS in different cells, tissues and organs, and in plants grown under different conditions. Here, we demonstrate that the AS RT-PCR system allows the analysis of multiple AS events simultaneously, and detects significant changes in AS in different organs, in seedlings grown under light/dark and long-/short-day conditions, and following overexpression of SR proteins. The advantages of this system are that it allows direct visualization of RT-PCR products, clearly identifies product sizes, distinguishes different AS products, quantifies products, thereby allowing the determination of the ratio of products (differential splice site use) within the same reaction, and detects novel products within the amplification range. In addition, although we have concentrated on alternative 5′ and 3′ splice site usage, and have only reported on the original defined events, the

Table 3 Genes showing significant ($P \leq 0.1$) alternative splicing (AS) changes between long- and short-day-grown plants

					Long (%)[a]		Short (%)[a]			
Gene identifier	Event	Position	Dist. (nt)	Consequence	Distal	Prox.	Distal	Prox.	Organ	*P*-value
At4g16845	Alt 5′ss	IVS5	22	Truncated protein – PTC	91	**9**	95	**5**	Root	0.007
At1g20693	Alt 5′ss	IVS3	104	Truncated protein – PTC	**98**	2	**95**	5	Root	0.007
At5g53300	Alt 3′ss	IVS1	27	Alternative 5′ UTR sequences	99	1	92	8	Leaf	0.016
At2g36000	Alt 5′ss	Ex1/3′ UTR	104	Change in C-terminal end	38	62	44	56	Flower	0.017
At2g37340	Alt 3′ss	IVS2	218	Truncated protein – PTC	**92**	8	**96**	4	Leaf	0.030
At2g37340	Alt 3′ss	IVS2	218	Truncated protein –-PTC	**97**	3	**99**	1	Root	0.038
At5g43910	Alt 3′ss	IVS9/Ex10	31	Truncated protein – PTC	39	**61**	42	**38**	Flower	0.089

Genes: At4g16845, vernalization 2 protein (VRN2); At1g20693, HMG box domain containing; At5g53300, ubiquitin-conjugating enzyme; At2g36000, mitochondrial transcription termination factor-related; At2g37340, RSZ33; At5g43910, Ribokinase. Abbreviations: aa, amino acids; Ex, exon; IVS, intervening sequence or intron; PTC, alternative splicing event generates a premature termination codon truncating the protein.
[a]The results are the means from three different experiments using biological repetitions. For AS events that generate truncated proteins, the figures in bold represent the percentage splicing value for the splicing event that generates the full-length protein, as annotated (TAIR).

system is clearly capable of detecting exon skipping and intron-retention events within a detectable size range, and quantifying multiple RT-PCR products.

Of the 89 events included on the AS RT-PCR panel, 36 showed a significant change in AS in at least one of the different treatments, organs or overexpressing lines with some transcripts, such as the SR protein, RSZ33 and the mitochondrial transcription termination factor (At 2g36000), showing significant changes in all four experiments. In a number of cases, the degree of change in AS, although significant, appeared relatively small. However, the experiments were conducted on whole organs, seedlings or plants, and thereby reflect the average level of alternatively spliced transcripts across many different cell types, such that small changes in AS may have significant biological effects in particular cells. In animal systems, tissue- or cell type-specific AS is an important mechanism for modulating function, in a highly regulated manner, in different cells or at specific developmental changes. Similarly, some genes have different levels of overall transcription under different conditions, such that the absolute levels of AS transcripts can vary greatly. The biological relevance of the AS events in this situation will require the integration of transcriptional data from microarrays with changes in AS ratios, and knowledge of the functions of the resultant proteins.

The changes in protein sequence resulting from AS ranged from the addition or loss of a small number of amino acids to major changes in N- and C-terminal sequences of proteins. Three examples involved the inclusion of putative signal peptides, which will clearly affect the localization of the protein. For well-characterized genes, such as those encoding SR proteins, the effects of AS on the domain structure of resultant proteins can be determined. However, in the majority of cases no protein domain information was available to predict changes in protein function. One of the more common outcomes was to alter the frame of the coding region generating premature termination codons in alternatively spliced isoforms. If actively turned over by NMD, the observed level of the PTC-containing transcript would be reduced. More than half of the genes where AS generated a PTC had levels of the PTC-containing transcripts lower than 20% of the total. This may suggest degradation by NMD but remains to be demonstrated, particularly as some plant PTC-containing mRNAs appear to be long-lived, including those of SR proteins (Kalyna and Barta, 2006). In addition, recent work in animal systems suggests that the number of genes regulated by NMD may be relatively small (Pan *et al.*, 2006). Finally, some of the significant changes involved AS in 5′ and 3′ UTRs. Around 20% of plant AS occurs in UTRs (Wang and Brendel, 2006). In animal systems, AS in 5′ UTRs is often coupled to the use of differentially regulated promoters and transcription start sites affecting the N-terminal exon sequence (Mironov *et al.*, 1999). Here, AS in 5′ UTRs either altered the N-terminal sequence of proteins or altered upstream open reading frames (uORFs), which can potentially interfere with or stimulate the translatability of the mRNA (Meijer and Thomas, 2002). Likewise, AS in the 3′ UTR can generate different 3′ exons, and can be associated with the use of differential polyadenylation sites.

Despite the increasing number of ESTs, many AS events in plants remain to be discovered (Xiao *et al.*, 2005). The overall number of AS events makes it highly unlikely that each event is regulated by specific factors, but is instead regulated by the relative levels of general splicing and RNA binding factors in cells. Orthologues of SR proteins, splicing factors and hnRNP proteins have been identified in plants (Lorkovic and Barta, 2002; Wang and Brendel, 2004), but a direct effect on AS of pre-mRNAs of other genes has only been demonstrated for a

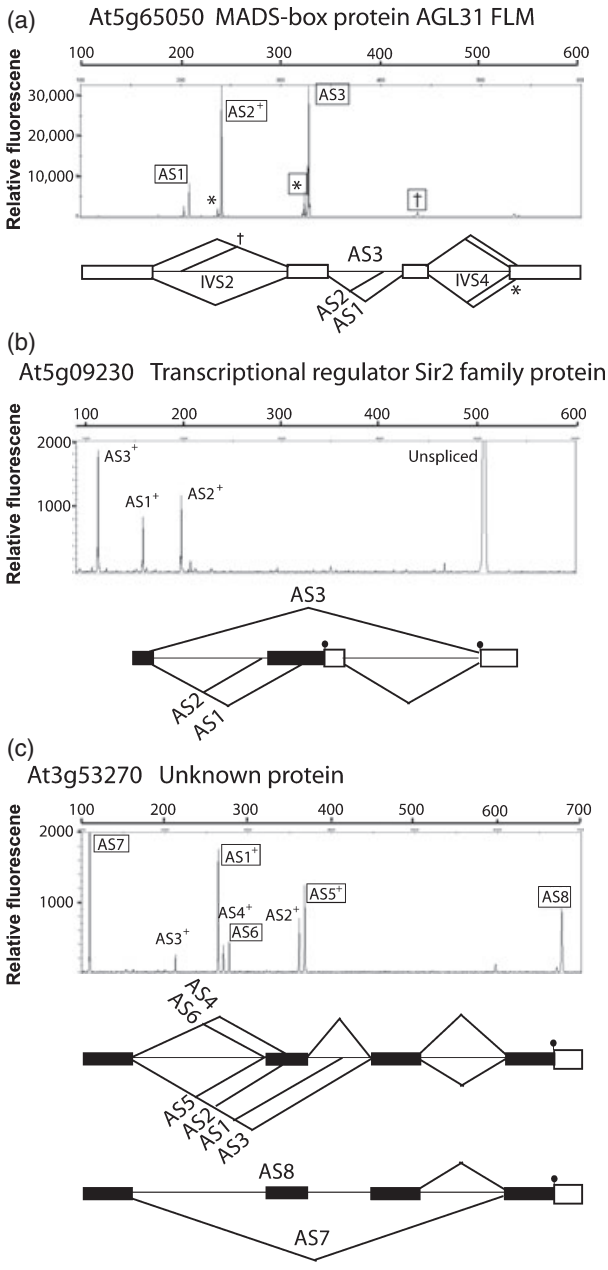


Figure 4. Characterization of novel alternative splicing events.
Unexpected RT-PCR products were characterized by cloning and sequencing and/or comparison with TAIR (version 7). Transcript analysis of (a) At5g65050, (b) At5g090230 and (c) At3g53270. Electropherograms and schematic transcript structure of major RT-PCR products are labelled from AS1 to AS8. †A minor product using a novel alternative 5′ splice site is labelled in (a); *two products using a novel alternative 3′ splice site are labelled in (a), and were identified by sequencing. +AS events that have a TAIR model. Boxed AS labels, * and † are AS events established by cloning and sequencing, and include novel splicing events. Transcript abundance is measured in relative fluorescence units. Splicing events are shown by lines above and below the gene structures. Open boxes represent protein coding exons, black boxes represent untranslated region (UTR) exons and lines are introns. Small black discs indicate translation starts.

small number of SR proteins (Isshiki *et al.*, 2006; Kalyna *et al.*, 2003; Lopato *et al.*, 1999). Transgenic plants overexpressing SR proteins show major pleiotropic effects on growth and development (e.g. embryo development, bilateral symmetry, retarded growth of meristems and seedlings, and stomatal development) (Kalyna *et al.*, 2003; Lopato *et al.*, 1999). It is likely that such effects are the result of the inappropriate AS of many genes. We have shown here, even within the limited number of AS events analysed, that plants overexpressing SR proteins, and light/dark and short-/long-day conditions alter the splicing of transcripts of other SR proteins, RNA-interacting proteins and transcription factors, all of which can affect gene expression patterns. Thus, developmental stage, cell type, or external stimuli or stress can alter the relative levels of RNA-interacting proteins (including SR and hnRNP proteins), both transcriptionally and post-transcriptionally to modulate the splicing patterns of downstream, target genes. One of the challenges is to understand how the relative levels of such regulatory proteins bring about the combinatorial control to determine the splicing of subsets of pre-mRNAs.

Exon arrays have been used in animal systems to quantify the levels of thousands of exons and alternative exons in different cells, tissues and growth conditions (Pan *et al.*, 2004, 2006). Exon arrays provide an important means of analysing AS globally, and are very appropriate to animals where the most common type of AS is the inclusion or skipping of whole exons. In plants, the prevalence of exon skipping is much lower, with intron retention being the most common form of AS, followed by selection of alternative 5′ and 3′ splice sites (Ner-Gaon *et al.*, 2004; Wang and Brendel, 2006). This most likely reflects different gene structures, particularly with respect to the smaller intron size of plants compared with animals. In plants, whole genome tiling arrays offer another possibility to analyse aspects of AS, but will require further optimization combined with validation (Mockler and Ecker, 2005). Such arrays have been used to analyse intron retention in Arabidopsis, and although able to readily detect retained introns, the detection of exon skipping was unreliable, and many 5′ and 3′ AS events, spanning relatively short distances, would not be detected (Ner-Gaon and Fluhr, 2006).

Given our current level of knowledge of AS in plants and the number of potentially undescribed AS events, the RT-PCR system presented here provides a robust means to monitor significant changes in AS. However, it is important to note that discovery of AS events in plants is essential to improve our understanding of AS. We have shown here that many new AS events exist, even in the limited regions amplified in these experiments. In addition, by cloning and sequencing RT-PCR products across the whole-gene regions of a number of Arabidopsis genes, we have

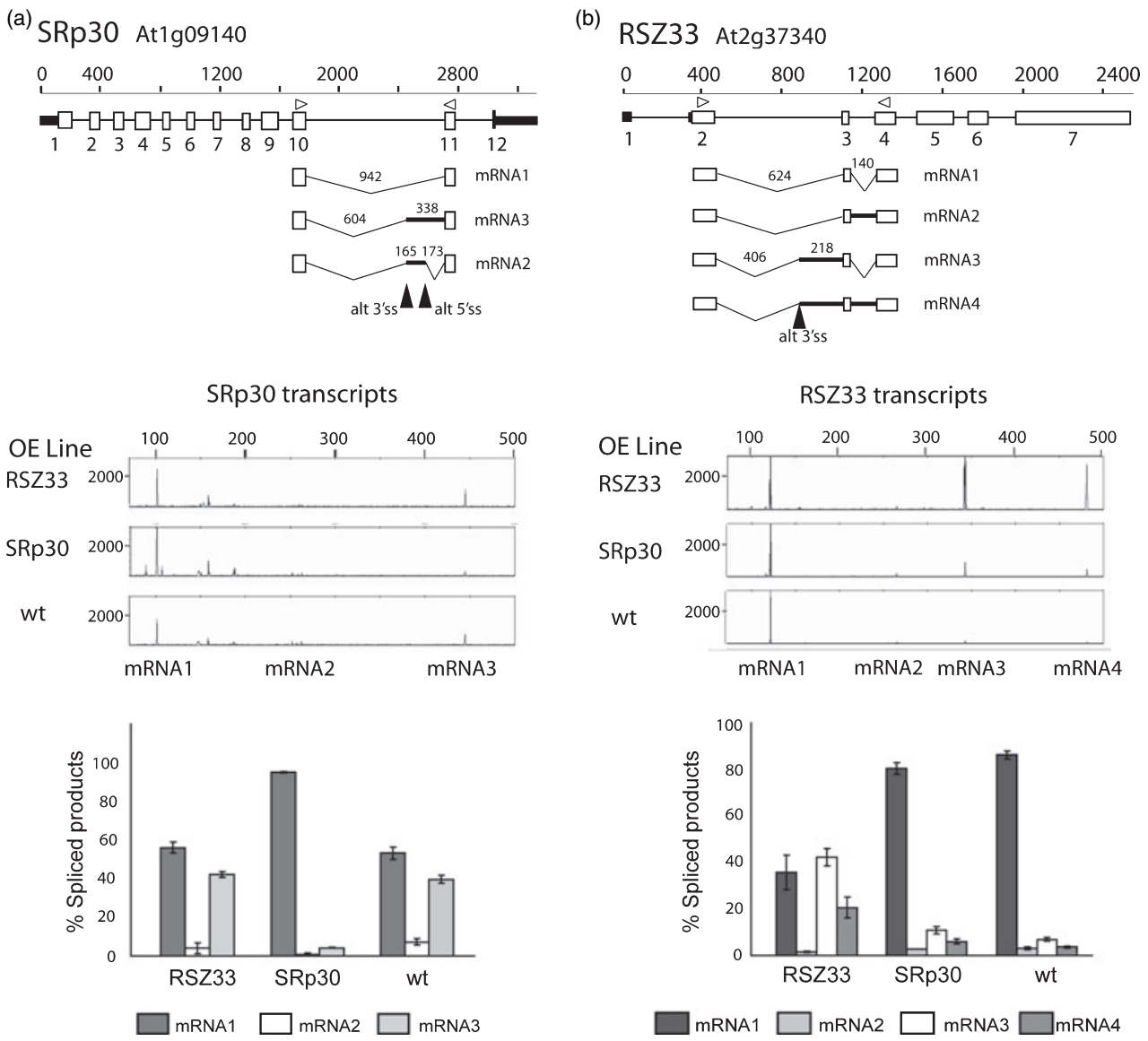


Figure 5. Gene structure and alternative splicing (AS) of SRp30 and RSZ33 in plants overexpressing SRp30 and RSZ33.
AS events detected by RT-PCR are shown below the gene structures for SRp30 (a) and RSZ33 (b). Electropherograms and histograms demonstrate significant changes in the ratios of alternatively spliced transcripts in transgenic lines overexpressing RSZ33 and SRp30, compared with wild-type (wt) seedlings.

also detected novel, unannotated events (CS and JB; unpublished data). This approach or high-throughput sequencing approaches will be required to systematically identify AS events in genes and pathways. To fully understand the principles of AS in plants and, in particular, to address combinatorial control and co-ordinated regulation of AS in plants, the analysis of large numbers of AS events simultaneously is required. The system here lends itself to expansion to analyse several hundred events in a wide range of conditions, to create specific panels for genes expressed at low levels (allowing higher PCR cycles), and to create pathway- or process-specific AS RT-PCR panels.

Experimental procedures

Plant material and growth conditions

Arabidopsis thaliana (wild-type Columbia and overexpressing SR lines) seeds were surface sterilized by incubating in 30% v/v sodium hypochlorite, 0.02% v/v Triton X-100 for 5 min, followed by three washes in sterile 0.02% v/v Triton X-100. Sterilized seeds were resuspended in sterile 0.1% agar solution and plated out on MS10 media in Petri dishes. Plates were sealed with micropore tape, wrapped in foil and placed at 4°C for 72 h. After stratification, plates were placed in the growth room under light conditions (16-h day, 8-h night) at 23°C. For dark-grown seedlings, plates were exposed to light for 6 h before wrapping in foil, and were left in the growth room as for the light-treatment plates. Seedlings were harvested

Table 4 Genes showing significant ($P < 0.1$) alternative splicing (AS) changes induced by overexpression of the SR proteins RSZ33 and SRp30

	Alternative splicing events				Wild type (%)[a]		OE RSZ33 (%)[a]		OE SRp30 (%)[a]		
Gene identifier	Event	Position	Distance	Consequence	Distal	Prox.	Distal	Prox.	Distal	Prox.	*P*-value
At1g09140	Alt 3′ss	IVS10	338	Truncated protein – PTC towards C-terminal	**57**	43	**57**	43	**96**	4	0.000
At3g61860	Alt 3′ss	IVS2	398	Truncated protein – PTC	**97**	3	**42**	58	**97**	3	0.000
At5g41150	Alt 3′ss	IVS5	53	Truncated protein – PTC	**82**	18	**80**	20	**90**	10	0.000
At2g37340	Alt 3′ss	IVS2	218	Truncated protein – PTC	**93**	7	**45**	55	**88**	12	0.000
At4g12790	Alt 3′ss	5′ UTR	126	uORF	51	49	50	50	29	61	0.000
At1g04400	Alt 3′ss	IVS 1	67	Alternative 5′ UTR	98	2	98	2	93	7	0.000
At2g32330	Alt 3′ss	Ex2	71	Truncated protein – PTC	37	**63**	33	**67**	43	**57**	0.001
At5g04430	Alt 3′ss	IVS5	63	In frame ± 21 aas	67	33	66	34	62	38	0.002
At5g65050	Alt 3′ss	IVS3	33	In frame ± 11 aas	16	84	18	82	15	85	0.003
At5g66010	Alt 3′ss	Ex3	77	Truncated protein – PTC	36	**64**	34	**66**	29	**61**	0.011
At3g54790	Alt 3′ss	Ex2	237	Removes annotated ATG – alternative ATG but also uORF	8	**92**	11	**89**	13	**87**	0.016
At3g53270	Alt 3′ss	IVS1	96	Alternative 5′ UTR sequences	78	22	75	25	79	21	0.034
At3g55630	Alt 5′ss	IVS5	66	In frame ± 22 aas	7	93	8	92	6	94	0.044
At3g57520	Alt 3′ss	Ex14	34	Alternative stop codon – changes and shortens C-terminal end	0.7	99.3	0.5	99.5	0.3	99.7	0.069
At4g30480	Alt 3′ss	Ex4	45	Alternative stop codon – changes and shortens C-terminal end	99.3	0.7	98.8	1.2	99.1	0.9	0.078
At2g36000	Alt 5′ss	Ex1/3′ UTR	104	Change in C-terminal end	37	63	36	64	41	59	0.089

Genes: At1g09140, SF2/ASF-like splicing modulator (SRp30); At3g61860, RSp31 RS-rich splicing factor; At5g41150, repair endonuclease (RAD1); At2g37340, RSZ33; At4g12790, hypothetical ATP binding protein; At1g04400, crytochrome 2 apoprotein (CRY2)(PHH1); At2g32330, tRNA *His* guanylyltransferase; At5g04430, putative RNA binding protein; At5g65050, MADS-box protein AGL31 FLM; At5g66010, putative RNA binding protein, hnRNP-like; At3g54790, putative armadillo repeat, U-box containing; At3g53270, unknown protein; At3g55630, tetrahydrofolylpolyglutamate synthase; At3g57520, imbibition protein homolog; At4g30480, tetratricopeptide repeat (TPR); At2g36000, mitochondrial transcription termination factor-related. Abbreviations: aa, amino acids; Ex, exon; IVS, intervening sequence or intron; OE, overexpression line; PTC, alternative splicing event generates a premature termination codon truncating the protein.

[a]The results are the means from three different experiments using biological repetitions. For AS events that generate truncated proteins, the figures in bold represent the percentage splicing value for the splicing event that generates the full-length protein, as annotated (TAIR).

after 10 days, just prior to the development of true leaves, and were then flash frozen in liquid nitrogen and stored at −80°C. For long- and short-day-grown plantlets, seed was sown directly onto compost and incubated at 4°C for 72 h before transfer to a Snijders Microclima 1000 growth cabinet, maintained at 20°C and 70% humidity. Long-day plants were grown under 16 h of light (37 500 lx) and 8 h of dark, whereas short-day plants were grown under 10 h of light and 14 h of dark. Root, leaf and whole flower tissue was isolated from flowering plants. Transgenic lines overexpressing RSZ33 (Kalyna *et al.*, 2003) and SRp30 (Lopato *et al.*, 1999) (SRp30, cDNA; RSZ33, genomic) were grown as above.

RNA extraction and RT-PCR

RNA was extracted from 50–100 mg of ground tissue, as described for the RNeasy Plant Mini RNA isolation kit (Qiagen, http://www.qiagen.com). First-strand cDNA synthesis was performed on 5 μg of extracted RNA using 'Ready-to-go you-prime' first-strand beads (Amersham, http://www.amersham.com). The RNA was incubated for 10 min at 65°C, followed by chilling on ice for 2 min, and was then mixed with 2 μM oligo $d(T)^{18}$ and beads containing all the components for reverse transcription. The reaction mix was incubated for 1 min at room temperature (21°C), mixed gently and further incubated at 37°C for 1 h. The reverse transcription reaction was diluted to a final volume of 100 μl, and 1 μl was aliquoted into a 96-well reaction plate along with 1 × PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl and 3 mM $MgCl_2$), 0.2 mM of dATP, dCTP, dGTP and dTTP (Promega, http://www.promega.com), 1.5 μM of each of the AS gene-specific primers and Taq DNA Polymerase (Roche, http://www.roche.com). The PCR reaction was incubated for 30 min at 48°C prior to performing the standard PCR reaction: 94°C for 2 min, followed by 24 cycles of 94°C for 15 sec, 50°C for 30 sec, 70°C for 1 min and completed with 10 min at 72°C (Perkin Elmer 9700, http://www.perkinelmer.com). All the reaction components apart from the primers and the cDNA were mixed prior to adding to each of the 96 PCR reactions. Specific alternative splice oligonucleotides (MWG) were selected to identify the expected alternatively spliced products, and gave products that were measurable between 60 and 700 bp. The forward primer was labelled with 6-carboxyfluorescein (6-FAM) to visualize the RT-PCR products. A list of the primer sequences is shown in Table S1.

Splicing analysis

A 25-μl volume of the labelled RT-PCR products from the 96 simultaneous RT-PCR reactions was purified using the minElute 96-well UF PCR purification process (Qiagen). Purified RT-PCR product (1 μl) from each of the 96 reactions was suspended in 10 μl Hi Di Formamide (Applied Biosystems, http://www.appliedbiosystems.com) with 0.05 μl of GeneScan −500 LIZ internal size standard. RT-PCR fragments were separated on a 3730 DNA Analyzer (Applied Biosystems), and were then collected and analysed using GENEM-

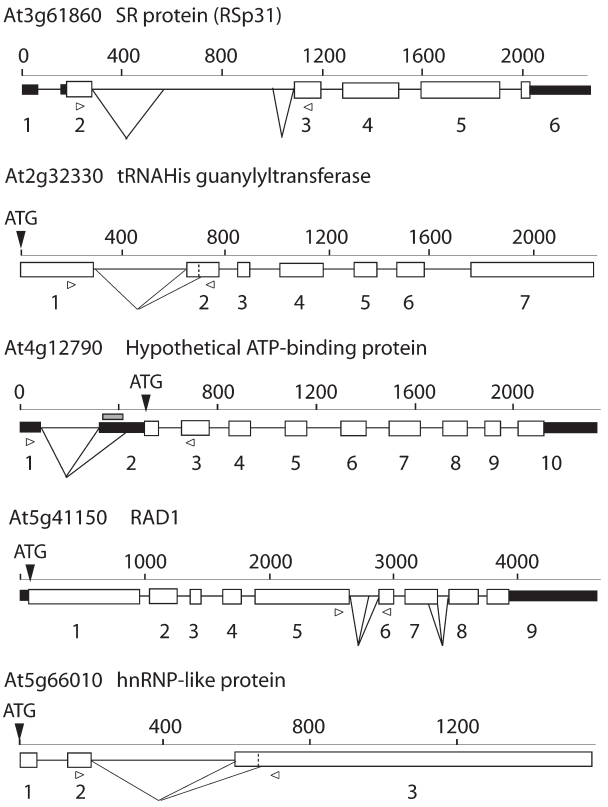


Figure 6. Gene structures and alternative splicing (AS) events of selected genes with significant changes in AS in plants overexpressing RSZ33 or SRp30.
The size of the gene, number of exons and translation start site (ATG) are indicated. The analysed AS events are indicated by lines below the gene structures, and the positions of the primers used are shown by open arrowheads. Key: white boxes, coding exons; black boxes, non-coding, untranslated region (UTR) exons; grey box, upstream open reading frame; lines, introns.

APPER software (Applied Biosystems). RT-PCR products were accurately identified with the ± 0.5 nt resolution of the ABI 3730. The relative fluorescent peak areas for RT-PCR products with expected sizes for the alternatively spliced products were extracted, and a ratio for the AS events was calculated by dividing the value for the spliced product by the sum of the values for the alternatively spliced products. For an accurate measure of AS, the above procedure was replicated with RNA from three technical samples, for the optimization experiments described in Figure 1, and three or four biological samples for all other experiments. Mean AS efficiencies with standard errors were calculated. Means were compared by analysis of variance for the overexpression lines, light/dark and long-/short-day treatments; results were compared by analysis of variance and *P* values were generated. AS events with significant variation ($P < 0.10$) were selected.

Cloning and sequencing of RT-PCR products

RT-PCR products were cloned either from individual bands separated by polyacrylamide gel electrophoresis or by cloning of the RT-PCR reaction into pGEM-T Easy (Promega). Clones were grown in 96-well format, and plasmid DNA was extracted by Millipore's Multi-screen Plasmid Mini-preparation protocol (Millipore, http://www.millipore.com). Sequencing was performed on plasmid templates using 1/16th reactions of ABI Big Dye v3.1, and using M13 forward and reverse primers. Samples were run through a capillary ABI3730 DNA sequencer, and the resulting sequence was analysed on SEQUENCHER v4.5 software (GeneCodes, http://www.genecodes.com).

Acknowledgements

The work was supported by a grant-in-aid from the Scottish Government, and by the European Alternative Splicing Network of Excellence (EURASNET).

Supplementary Material

The following supplementary material is available for this article online:
Table S1. Panel of monitored alternative splicing events.
This material is available as part of the online article from http://www.blackwell-synergy.com
Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Balasubramanian, S., Sureshkumar, S., Lempe, J. and Weigel, D. (2006) Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. *PLoS Genet.* **2**, 980–989.

Black, D.L. (2003) mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* **72**, 291–336.

Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. *Cell*, **126**, 37–47.

Brett, D., Pospisil, H., Valcárcel, J., Reich, J. and Bork, P. (2002) Alternative splicing and genome complexity. *Nat. Genet.* **30**, 29–32.

Campbell, M.A., Haas, B.J., Hamilton, J.H., Mount, S.M. and Buell, C.R. (2006) Comprehensive analysis of alternative splicing in rice and comparative analysis with Arabidopsis. *BMC Genomics*, **7**, 327–343.

Chen, F.-C., Wang, S.-S., Chaw, S.-M., Huang, Y.-T. and Chuang, T.-J. (2007) Plant gene and alternatively spliced variant annotator. A plant genome annotation pipeline fro rice gene and alternatively spliced variant identification with cross-species expressed sequence tag conservation from seven plant species. *Plant Physiol.* **143**, 1086–1095.

Dinesh-Kumar, S.P. and Baker, B.J. (2000) Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. *Proc. Natl Acad. Sci. USA*, **97**, 1908–1913.

Egawa, C., Kobayashi, F., Ishibashi, M., Nakamura, T., Nakamura, C. and Takumi, S. (2006) Differential regulation of transcript accumulation and alternative splicing of a *DREB2* homolog under abiotic stress conditions in common wheat. *Genes Genet. Syst.* **81**, 77–91.

Fang, Y., Hearn, S. and Spector, D.L. (2004) Tissue-specific expression and dynamic organisation of SR splicing factors in Arabidopsis. *Mol. Biol. Cell*, **15**, 2664–2673.

Fu, X.D. (1995) The superfamily of arginine/serine-rich splicing factors. *RNA*, **1**, 663–680.

Fung, R.W., Wang, C.Y., Smith, D.L., Gross, K.C., Tao, Y. and Tian, M. (2006) Characterization of alternative oxidase (AOX) gene

expression in response to methyl salicylate and methyl jasmonate pre-treatment and low temperature in tomatoes. *J. Plant Physiol.* **163**, 1049–1060.

Gao, H., Gordon-Kamm, W.J. and Lyznik, L.A. (2004) ASF/SF2-like maize pre-mRNA splicing factors affect splice site utilization and their transcripts are alternatively spliced. *Gene*, **339**, 25–37.

Golovkin, M. and Reddy, A.S. (1998) The plant U1 small nuclear ribonucleoprotein particle 70K protein interacts with two novel serine/arginine-rich proteins. *Plant Cell*, **10**, 1637–1648.

Graveley, B.R. (2000) Sorting out the complexity of SR protein functions. *RNA*, **6**, 1197–1211.

Graveley, B.R. (2001) Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* **17**, 100–107.

Gupta, S., Wang, B.-B., Stryker, G.A., Zanetti, M.E. and Lal, S.K. (2005) Two novel arginine/serine (SR) proteins in maize are differentially spliced and utilize non-canonical splice sites. *Biochim. Biophys. Acta*, **1728**, 105–114.

Iida, K. and Go, M. (2006) Survey of conserved alternative splicing events of mRNAs encoding SR proteins in land plants. *Mol. Biol. Evol.* **23**, 1085–1094.

Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, K., Toyoda, T., Konagaya, A. and Shinozaki, K. (2004) Genome-wide analysis of alternative splicing in Arabidopsis thaliana based on full-length cDNA sequences. *Nucleic Acids Res.* **32**, 5096–5103.

Isshiki, M., Tsumoto, A. and Shimamoto, K. (2006) The serine-arginine-rich protein family in rice plays important roles in constitutive and alternative splicing of pre-mRNA. *Plant Cell*, **18**, 146–158.

Jia, Y., del Rio, H.S., Robbins, A.L. and Louzada, E.S. (2004) Cloning and sequence analysis of a low temperature-induced gene from trifoliate orange with unusual pre-mRNA processing. *Plant Cell Rep.* **23**, 159–166.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141–2144.

Jordan, T., Schornack, S. and Lahaye, T. (2002) Alternative splicing of transcripts encoding Toll-like plant resistance proteins – what's the functional relevance to innate immunity? *Trends Plant Sci.* **7**, 392–398.

Kalyna, M. and Barta, A. (2004) A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel functions? *Biochem. Soc. Trans.* **32**, 561–564.

Kalyna, M., Lopato, S. and Barta, A. (2003) Ectopic expression of atRSZ33 reveals its function in splicing and causes pleiotropic changes in development. *Mol. Biol. Cell*, **14**, 3565–3577.

Kalyna, M., Lopato, S., Voronin, V. and Barta, A. (2006) Evolutionary conservation and regulation of particular alternative splicing events in plant SR proteins. *Nucleic Acids Res.* **34**, 4395–4405.

Kazan, K. (2003) Alternative splicing and proteome diversity in plants: the tip of the iceberg has just emerged. *Trends Plant Sci.* **8**, 468–471.

Kriventseva, E.V., Koch, I., Apweiler, R., Vingron, M., Bork, P., Gelfland, M.S. and Sunyaev, S. (2003) Increase in functional diversity by alternative splicing. *Trends Genet.* **19**, 124–128.

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.

Larkin, P.D. and Park, W.D. (1999) Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. *Plant Mol. Biol.* **40**, 719–727.

Lazar, G. and Goodman, H.M. (2000) The Arabidopsis splicing factor SR1 is regulated by alternative splicing. *Plant Mol. Biol.* **42**, 571–581.

Lewandowska, D., Simpson, C.G., Clark, G.P., Jennings, N.S., Barciszewska-Pacak, M., Lin, C.-F., Makalowski, W., Brown, J.W.S. and Jarmolowski, A. (2004) Determinants of plant U12-dependent intron splicing efficiency. *Plant Cell*, **16**, 1340–1352.

Lewis, B.P., Green, R.E. and Brenner, S.E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl Acad. Sci. USA*, **7**, 189–192.

Lopato, S., Kalyna, M., Dorner, S., Kobayashi, R., Krainer, A.R. and Barta, A. (1999) atSRp30, one of two SF2/ASF-like proteins from *Arabidopsis thaliana*, regulates splicing of specific plant genes. *Genes Dev.* **13**, 987–1001.

Lopato, S., Forstner, C., Kalyna, M., Hilscher, J., Langhammer, U., Indrapichate, K., Lorkovic, Z.J. and Barta, A. (2002) Network of interactions of a novel plant-specific Arg/Ser-rich protein, atRSZ33, with atSC35-like splicing factors. *J. Biol. Chem.* **277**, 39989–39998.

Lorkovic, Z.J. and Barta, A. (2002) Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Res.* **30**, 623–635.

Maniatis, T. and Tasic, B. (2002) Alternative pre-mRNA splicing and proteome expression in metazoans. *Nature*, **418**, 236–243.

Maquat, L.E. (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* **5**, 89–99.

Matlin, A.J., Clark, F. and Smith, C.W.J. (2005) Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol.* **6**, 386–398.

Meijer, H.A. and Thomas, A.M. (2002) Control of eukaryotic protein synthesis by upstream open reading frames in the 5′-untranslated region of an mRNA. *Biochem. J.* **367**, 1–11.

Mironov, A.A., Fickett, J.W. and Gelfand, M.S. (1999) Frequent alternative splicing of human genes. *Genome Res.* **9**, 1288–1293.

Mockler, T.C. and Ecker, J.R. (2005) Applications of DNA tiling arrays for whole-genome analysis. *Genomics*, **85**, 1–15.

Modrek, B. and Lee, C. (2002) A genomic view of alternative splicing. *Nat. Genet.* **30**, 13–19.

Ner-Gaon, H. and Fluhr, R. (2006) Whole-genome microarray in *Arabidopsis* facilitates global analysis of retained introns. *DNA Res.* **13**, 111–121.

Ner-Gaon, H., Halachmi, R., Savaldi-Goldstein, S., Rubin, E., Ophir, R. and Fluhr, R. (2004) Intron retention is a major phenomenon in alternative splicing in Arabidopsis. *Plant J.* **39**, 877–885.

Pan, Q., Shai, O., Misquitta, C. *et al.* (2004) Revealing global regulatory features of mammalian alternative splicing using a quantitative microarray platform. *Mol. Cell*, **16**, 929–941.

Pan, Q., Saltzman, A.L., Kim, Y.K., Misquitta, C., Shai, O., Maquat, L.E., Frey, B.J. and Blencowe, B.J. (2006) Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated decay to control gene expression. *Genes Dev.* **20**, 153–158.

Savaldi-Goldstein, S., Aviv, D., Davydov, O. and Fluhr, R. (2003) Alternative splicing modulation by a LAMMER kinase impinges on developmental and transcriptome expression. *Plant Cell*, **15**, 926–938.

Smith, C.W. and Valcárcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.* **25**, 381–388.

Stamm, S. (2002) Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Hum. Mol. Genet.* **11**, 2409–2416.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. and Soreq, H. (2005) Function of alternative splicing. *Gene*, **344**, 1–20.

Tanabe, N., Yoshimura, K., Kimura, A., Yabuta, Y. and Shigeoka, S. (2007) Differential expression of alternatively spliced mRNAs of Arabidopsis SR protein homologs, atSR30 amd atSR45a, in response to environmental stress. *Plant Cell Physiol.* **48**, 1036–1049.

Ule, J., Ule, A., Spencer, J. *et al.* (2005) Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* **37**, 844–852.

Wang, B.-B. and Brendel, V. (2004) The ASRG database: identification and survey of Arabidopsis thaliana genes involved in pre-mRNA splicing. *Genome Biol.* **5**, R102.

Wang, B.B. and Brendel, V. (2006) Genome-wide comparative analysis of alternative splicing in plants. *Proc. Natl Acad. Sci. USA*, **103**, 7175–7180.

Xiao, Y.-L., Smith, S.R., Ishmael, N., Redman, J.C., Kumar, N., Monaghan, E.L., Ayele, M., Haas, B.J., Wu, H.C. and Town, C.D. (2005) Analysis of cDNAs of hypothetical genes on Arabidopsis chromosome 2 reveals numerous transcript variants. *Plant Physiol.* **139**, 1323–1337.

Zhou, Y., Zhou, C., Ye, L., Dong, J., Xu, H., Cai, L., Zhang, L. and Wei, L. (2003) Database and analyses of known alternatively spliced genes in plants. *Genomics*, **82**, 584–595.

2.2. Publication

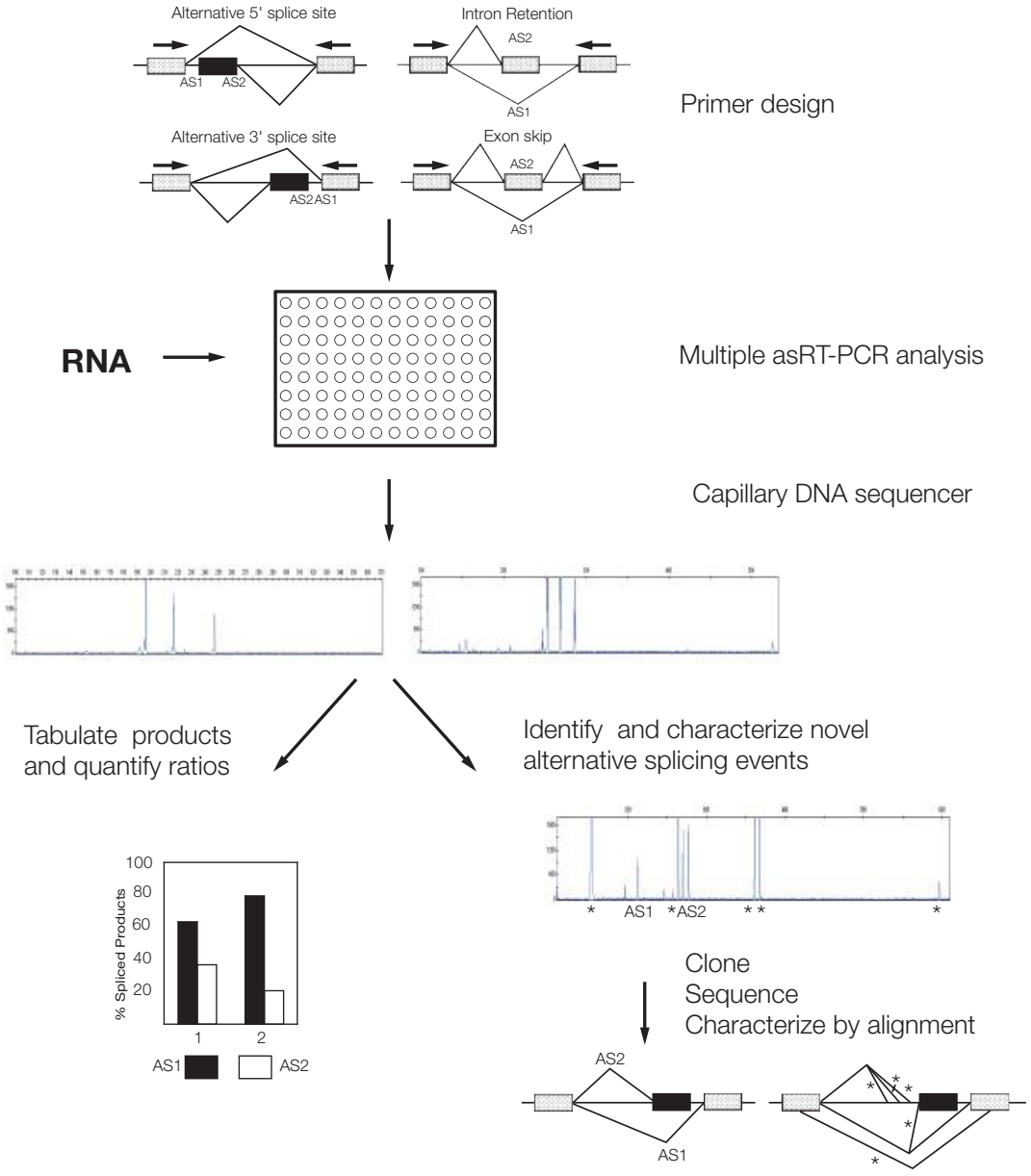
Monitoring changes in plant alternative splicing events

Craig G. Simpson, Naeem Hasan Syed, Sujatha Manthri, John D. Fuller, Monika Maronova, Branislav Kusenda, Maria Kalyna, Andrea Barta, and John W.S. Brown

in: Stamm, S., Smith, C.W.J., Lührmann, R. (Eds.), In "Alternative pre-mRNA Splicing: Theory and Protocols" (2012). Wiley-VCH Verlag GmbH & Co. KGaA, pp. 248–259.

Author's contributions: Preparation of the plant material, preparation of the RNA, optimization of the method and conducting the high resolution RT-PCR screen.

Monitoring Changes in Plant Alternative Splicing Events



Outcome:
System to measure changes in alternative splicing and identify and characterize new alternative splicing events.

Question answered:
How does alternative splicing change between treatments?
What are the most significant splicing events?
What novel alternative splicing events exist?
What is the influence of different factors on splicing?

23
Monitoring Changes in Plant Alternative Splicing Events

Craig G. Simpson, Naeem Hasan Syed, Sujatha Manthri, John D. Fuller, Monika Maronova, Branislav Kusenda, Maria Kalyna, Andrea Barta, and John W.S. Brown

Abstract

In order to address the key aspects of alternative splicing (AS) in plants, a high-resolution RT-PCR system has been established to monitor changes in multiple AS events. The system currently detects AS events from a range of about 300 different *Arabidopsis* genes, and represents different types of AS event: alternative 5′ and 3′ splice sites, exon skipping, and intron retention. The AS RT-PCR (AS-RT-PCR) protocol allows the accurate determination of splicing isoforms and the quantification of changes in AS of all the events, simultaneously. The protocol is currently being used to address the influence on AS different growth conditions, such as abiotic stress and of mutants of splicing and other mRNA biogenesis factors. The procedure has also been used to identify novel AS events that are being characterized.

23.1
Theoretical Background

23.1.1
Alternative Splicing in Plants

Alternative splicing (AS) is less well-characterized in plants than in animal systems. This is due partially to the relatively lower number of expressed sequence tags (ESTs) available, the poor depth of sequencing for transcripts of some plant genes, genomic contamination, and the poor detection of rare and aberrant events [1]. Plants demonstrate many examples of complex AS that have key roles in the development of plants and their response to cellular, biotic, and abiotic factors [2]. Yet, recent reports have suggested that the degree of AS in plants is underestimated [3,4]. As ever-more examples of important and functional AS are reported, interest in AS as a regulator of plant gene expression continues to grow. The key challenges in plant AS are the discovery of AS events, the measurement of changes in AS, determination of the function of different alternatively spliced isoforms and, finally, how AS is regulated. To date, the majority of progress has been made on the characterization of AS factors and, in particular, of the family of SR proteins [5].

The current level of knowledge of AS events in, for example, *Arabidopsis thaliana*, precludes the development of genomic-scale exon or splice junction arrays. Whole-genome tiling arrays are capable of detecting some AS events in *Arabidopsis* and, in particular, intron-retention events [6]. However, whilst the next-generation sequencing (454; Solexa, SOLiD) has the power to identify novel AS events and different AS isoforms, at present the protocols and software to interpret AS on a quantitative basis are still undergoing development (see Chapters 50 and 51, Guigó, Zhang) [7]. Therefore, in order to start addressing the fundamental question of the regulation of AS in plants, and to acquire an estimate of the extent of AS under different conditions, a

Alternative pre-mRNA Splicing: Theory and Protocols, First Edition. Edited by Stefan Stamm, Chris Smith, and Reinhard Lührmann.

system was developed that allowed the simultaneous monitoring of multiple AS events from *Arabidopsis*. This system, which is based on RT-PCR, separates the AS isoforms on a capillary DNA-sequencing machine, thereby allowing the high-resolution identification of isoforms and quantification of their relative abundance [8]. This system is, in principle, similar to the high-throughput system described for human AS in terms of its detection and quantification (see Chapter 22, Klinck). Apart from certain technical differences (e.g., RNA quality estimation, cycle number, automation), the downstream analysis for plant AS events is currently complicated by the number of novel alternatively spliced products that are detected and which require characterization, but this reflects the lower level of knowledge regarding AS events in plants. By exploiting the availability of transgenic lines overexpressing splicing factors, and viable knock-out mutants of splicing factors and other mRNA biogenesis factors, the AS RT-PCR system can be used to address the regulation of plant AS. In addition, by quantifying the changes in AS, it should become possible to resolve how AS can influence the expression of specific genes under different conditions and stresses, of different organs/cells and developmental stages, and between different ecotypes.

23.1.2 Splicing Analysis

Currently, RT-PCR is a widely used protocol to amplify large amounts of DNA product derived from target RNAs. The protocol uses an RNA-dependent DNA polymerase (reverse transcriptase; RT) to form a cDNA copy, and this is followed by PCR amplification using gene-specific DNA oligonucleotides and a thermostable Taq DNA polymerase to produce multiple copies of the target sequence. For genes which produce alternatively spliced RNAs, primers can be designed to amplify different alternatively spliced sequences at the same time. The system developed here used oligonucleotide primer pairs where the upstream primer is end-labeled with fluorescent tags (e.g., carboxyfluorescein; 6-FAM); this allows the RT-PCR products to be separated by using capillary DNA-sequencing machines, such as the ABI 3730 (Applied Biosystems). The fluorescent peak areas for the alternatively spliced isoforms were extracted and analyzed using GeneMapper software (Applied Biosystems) that has been developed for genotyping and fragment analysis. This provides exquisite sensitivity and product size accuracy that allow, first, the identification of AS isoforms, and second, quantification of the ratio of the spliced products observed.

23.1.3 Establishing the RT-PCR Conditions

To enable quantification, it is important not to use too many PCR cycles when amplifying the products. The RT-PCR conditions were determined in control experiments using gene-specific primers for genes known to produce alternatively spliced isoforms, and different cycle numbers ranging from 20 to 28 cycles. In this way, it was established that 24 PCR cycles was in the exponential amplification range for mRNA transcripts of numerous genes with a range of expression levels [8,9]. For optimum resolution, the primers were designed to generate RT-PCR products of between 100 and 600 bp. Separation on the ABI sequencer limits accurate size measurement to products of about 600 bp.

23.1.4 Characterizing Novel Alternatively Spliced Products

The AS RT-PCR procedure was established to detect and measure AS isoforms of genes that were known to be alternatively spliced. However, it became apparent at a very early stage that the sensitivity of the system detected many unexpected RT-PCR products. With a relatively limited knowledge of AS events, these novel products potentially represented unidentified AS events within the region of the gene studied.

It was, therefore, necessary to characterize as many novel AS events as possible. Many such events result in a mRNA with a premature termination codon (PTC), which usually leads to destruction of the mRNA by a mechanism termed nonsense-meditated decay (NMD). As translation is a precondition for NMD, the use of cycloheximide (an inhibitor of translation) or mutants of the NMD pathway [e.g., in up-frameshift (UPF) proteins), leads to an upregulation of PTC-containing mRNAs. The results of these experiments indicated that some novel RT-PCR products were potentially turned over by NMD. To increase the abundance and the probability of successful cloning of these products, RT-PCR reactions from *upf* mutants or cycloheximide-treated plants were used in cloning experiments.

23.1.5
Identifying AS Events in Genes of Interest

The number of novel unannotated AS events detected in the AS RT-PCR system clearly demonstrated that far more AS was occurring than would be expected from current annotations. This situation is comparable with human systems, where advances in detection by array analysis or high-throughput sequencing increased the number of known AS events [10]. Therefore, when studying the expression of specific genes, it is important to consider the likelihood of AS events which are not yet annotated. Therefore, the protocol was modified for RT-PCR and cloning from the AS RT-PCR system to assess routinely those genes of interest for the degree of AS, and to identify such events by cloning. This protocol has been utilized to assess AS in genes belonging to the same developmental pathway, in order ultimately to develop RT-PCR panels specific for individual pathways and processes. For example, AS is known to have a significant role in the flowering time of plants. To characterize in detail AS in flowering-time genes, primers were designed to amplify along the length of various genes, and the resulting RT-PCR products cloned and sequenced.

23.2
Protocols

Protocol 1: Multiple RT-PCR AS Panel Using 96-Well Plates

RNA Extraction

- Plant material: Total RNA is extracted from up to 100 mg of any selected tissue, but for mutant or growth condition analysis, usually 5-week-old leaves are used.
- The RNA is extracted using the RNeasy Plant Mini Kit (Qiagen Cat. No. 74904), following manufacturer's instructions (see Qiagen RNeasy manual).
- The RNA concentrations are determined using Nanodrop (LabTech). The RNA preparations are DNase-treated to remove DNA using DNase RQ1 treatment, according to the manufacturer's instructions (Promega).

First Strand cDNA Synthesis

- 5μg of total RNA (sufficient for 100 PCR reactions; i.e., one 96 well-plate) are added to sterile distilled water to a volume of 31 μl, incubated for 10 min at 65 °C, and chilled on ice for 2 min.
- The sample is transferred to a microfuge tube with "Ready-to-go You-prime" first strand beads (Amersham; Cat No. 27-9264-01) (DO NOT MIX!!) and 2 μl of oligo $d(T)_{18}$ (1 μg μl^{-1}) are added. The mixture is left at room temperature for 1 min.
- The sample is gently mixed with the end of a pipette, centrifuged briefly to collect the sample, which is incubated at 37 °C for 1 h. Sterile distilled water is added to provide a final volume of 100 μl.

PCR

Reagents:

- Roche Taq and Buffer. (Cat. No. 11 146 173)
- Promega dNTPs: made to 1.25 mM (Cat. No. U1240)
- Primers (MWG - Eurofins): 100 µM stock
- PCR plates (Thermo-Fast 96, Semi skirted; Millipore Cat. No. AB-0900).

1) For each PCR reaction the following are required:

	×1	×100
10 × buffer	10 µl	1000 µl
dNTPs	16 µl	1600 µl
Primer F (100 µM)	1.5 µl added separately. (Stocks held in 96-well plate)	
Primer R (100 µM)	1.5 µl added separately. (Stocks held in 96-well plate)	
Taq	0.5 µl	50 µl
SDW	69.5 µl	6950 µl

2) For a 96-well plate reaction, the complete first strand reaction mix is added to the 100× PCR reaction mix. Of this mixture, 97 µl is then added to each well of a 96-well plate containing gene-specific primers (1.5 µl forward primer labeled with 6-FAM, and 1.5 µl of reverse primer) to give a total PCR reaction volume of 100 µl.
3) The samples are mixed by vortexing, centrifuged briefly to collect the samples on the bottom of the well, and then placed on a PCR machine (Perkin Elmer 9700), using the following program:

1 cycle	94 °C for 2 min
24 cycles	94 °C for 15 s
	50 °C for 30 s
	70 °C for 1 min
1 cycle	70 °C for 10 min
Store at 4 °C	

Separation and Analysis of the Spliced Products

The materials required at this stage are LIZ Size standard (Cat. No. 4322682) and Hi Di Formamide (Applied Biosystems Cat. No. 4311320).

1) The labeled RT-PCR product from the RT-PCR reactions is mixed with Hi Di Formamide and a LIZ dye-labeled size marker. For the 96 reactions in the 96-well plate, the following mix should be prepared:

	×1	×100
500 LIZ Size standard	0.05 µl	5 µl
Hi Di Formamide	8.95 µl	895 µl

2) The mix is aliquoted into a 96-well plate, and 1 µl of each purified RT-PCR reaction (see Section 23.2.2.1) is added.
3) The remaining sample is stored at −20 °C for downstream cloning and sequencing if required (see below).
4) Samples are separated on an ABI3730 DNA Analyzer set up for fragment analysis.
5) Peak size and area data are analyzed with GeneMapper v3.5, or with freely available PeakScanner software (Applied Biosystems).
6) RT-PCR products are accurately identified with ±1 bp resolution. The relative fluorescent peak areas for RT-PCR products with expected sizes for the alternatively spliced products are extracted, tabulated, and a ratio for the AS events is calculated by dividing the value for each alternatively spliced product by the sum of the values for the alternatively spliced products.
7) For an accurate statistical measurement of AS ratios, three biological repeats are performed for all experiments. The mean AS ratios (with standard errors) are calculated for three separate biological repetitions.
8) The mean ratios are compared by an analysis of variance between wild-type plants and the different treatments, stresses, or mutant plants. The significance of the differences between treatments is determined for two-way comparisons by using Student's *t*-test, and for multiple treatments by an analysis of variance (ANOVA). AS events with significant variation ($p < 0.1$) are routinely selected.

Protocol 2: Cloning and Characterization of Novel Alternatively Spliced Transcripts

RT-PCR Purification and Ligation into pGEM-T Vector

RT-PCR reactions from the AS RT-PCR protocol were purified to remove all products below 150 bp, using Agencourt AMPure beads (Beckman Coulter Genomics). A number of DNA-purification kits may be used, the AMPure beads were found to be optimal for the removal of low-molecular-weight DNA products:

1) 9 µl of AMPure beads are mixed with 5 µl of RT-PCR reaction, mixed gently by pipetting, and left at room temperature for 3–5 min to facilitate binding.
2) The RT-PCR reaction/bead mixes are placed on a magnetic plate and the cleared supernatant is removed and discarded. The beads with bound DNA are washed twice with 70% ethanol and left on bench to dry completely.
3) 20–40 µl of TE buffer (10 mM Tris–HCl, pH 8, 1 mM EDTA) is added to each tube and mixed to release the purified DNA from the beads. The tubes are replaced on the magnetic plate to separate the DNA solution from the magnetic beads, and the DNA solution is then removed and used for downstream ligations.
4) Depending on the size of the RT fragment, 50–100 ng of DNA is ligated into the pGEM-T vector in a final volume of 10 µl containing 5 µl of 2× ligation buffer, 1 µl (50 ng) pGEM-T vector, and 1 µl of T4 DNA ligase lacking exonuclease activity (pGEM-T kit; Promega Cat. No. A3600).

PCR Amplification of Ligation Reactions of RT-PCR Products in pGEM-T

Each ligation reaction was tested using T7 and SP6 promoter primers flanking the cloning site in the pGEM-T vector:

1) PCRs are performed in a 20 µl reaction containing 2 µl from the ligation reaction, 1× Qiagen PCR buffer containing 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 10 pmol of each primer (T7 and SP6), and 1 unit of Hotstar Taq polymerase (Qiagen Cat. No. 203203).

2) The PCR is performed as follows:

1 cycle	15 min at 95 °C
25 cycles	95 °C for 30 s,
	55 °C for 30 s,
	72 °C for 60 s
1 cycle	72 °C for 7 min
Store at 4 °C	

3) The PCR products are resolved on a 2% agarose gel stained with ethidium bromide.

Transformation and Colony PCR

1) Each transformation is performed using 50 µl of JM109 $>10^8$ cfu μg^{-1} competent cells (Promega Cat. No. L2001) and 4 µl of the ligation reaction in a 50 ml sterile Falcon tube (this improves transformation efficiency) and heat-shocked at exactly 42 °C for 45 s.
2) Transformed cells are revived by adding 950 µl SOC medium to each tube and shaking at 37 °C for 1.5 h at 150 rpm.
3) 100 µl of culture are plated onto two or three LB plates containing ampicillin (100 µg ml^{-1}) and supplemented with 0.5 mM IPTG and 80 µg ml^{-1} X-Gal. The plates are then placed in a 37 °C incubator overnight for at least 16 h.
4) Colony PCRs are performed on selected white colonies with the same conditions used for PCR amplification of the ligation reactions above.
5) PCR products are resolved on 2% agarose gels stained with ethidium bromide and selected colonies are sequenced.

Sequencing

PCR products from selected colonies were purified using Qiagen PCR purification plates (Cat. No. 28181), and used as a template for dideoxy sequencing using the T7 or SP6 primers.

1) A 1/16th sequencing reaction is performed in 96-well plates with: 0.5 or 0.25 µl of Big Dye Terminator reagent, 2 µl 5× Big Dye reaction buffer, 1 µl M13F (10 µM), 4 µl template, and sterile distilled water added to a final volume of 10 µl. The PCR is performed as follows:
   - One cycle for 1 min at 96 °C
   - 25 cycles each at 95 °C for 10 s, 50 °C for 5 s, and 60 °C for 4 min
   - Store at 4 °C.
2) The sequencing reactions are cleaned according to the manufacturer's guidelines. 2.5 µl of 125 mM EDTA (pH 8.0) was added to the 10 µl reaction, mixed by vortexing, followed by the addition of 30 µl of 95% ethanol and incubated for 15 min at room temperature.
3) The plate is centrifuged at 3000 × *g* for 30 min at 4 °C, and then inverted onto a tissue and re-centrifuged for a further 10 s at 100 × *g*.
4) The samples are washed with 150 µl of 70% ethanol, mixed, and centrifuged at 3000 × *g* for 10 min at 4 °C. Removal of the final liquid is achieved by inverting on a tissue, and then re-centrifuging for 10 s at 100 × *g*. The washing step is repeated before leaving the samples to air-dry at room temperature.
5) Samples are suspended in 10 µl Hi Di formamide (Applied Biosciences), and sequencing then performed on an ABI 3730.

Sequence Alignment and AS Discovery

All sequences derived from ABI 3730 runs were trimmed to remove vector sequence, and the gene-specific primer(s) used for RT-PCR were identified. Sequences were aligned either using Geneseqer [11] to align the sequence against reference genomic sequence, or by ClustalW [12].

Protocol 3: Assessing AS in Genes of Interest by Cloning Full-Length cDNA Transcripts

RNA was extracted from *upf3-1* mutant seedlings and reverse transcription carried out as described in Protocol 1 (Sections 23.2.1.1 and 23.2.1.2). Gene-specific primers were designed to amplify full-length gene sequences. For genes larger than 2 kb, overlapping primer pairs were designed. To enable directional cloning, the forward PCR primer was designed to contain CACC additional bases at the 5′ end of the primer for cloning into pENTR/D-TOPO vector (Invitrogen). The reverse primer was designed without including the native stop codon of the gene of interest for future transfer into destination vectors.

Cloning Full-Length cDNAs

1) PCR is performed using Platinum *Pfx* DNA polymerase (Invitrogen) with 200 ng of DNA, 1 µl of 5 µM forward and 1 µl of 5 µM reverse primer, 1 µl of 10 mM dNTP mixture, 1 µl of 50 mM $MgSO_4$, 5 µl of 10× amplification buffer, 0.4 µl Platinum *Pfx* DNA polymerase, and distilled water to a final volume of 50 µl.
2) The PCR cycling is performed as follows:
   - 30 cycles each at 94 °C for 15 s, 55 °C for 30 s, and 68 °C for 1 min kb^{-1}
   - One cycle at 72 °C for 7 min
   - Store at 4 °C

Transformation and Multiple Plasmid Preparations

Materials required:

Solution 1	50 mM glucose	2.25 ml 40% glucose
	10 mM EDTA, pH 8	2 ml 0.5 M EDTA, pH 8
	10 mM Tris, pH 8	1 ml 1 M Tris, pH 8
	(store at 4 °C)	94.75 ml sterile distilled water
Solution 2	0.2 M NaOH	5 ml of 2 M NaOH
	1% SDS	5 ml of 10% SDS
	(store at room temperature)	40 ml sterile distilled water (add this first)
Solution 3	3 M KOAc, pH 4.8	29.45 g KOAc (crystal)
	(store at 4 °C)	11.5 ml glacial acetic acid
		Add sterile distilled water to 100 ml

1) The PCR products are cloned into the pENTR TOPO vector and transformed into chemically competent TOP10 cells as described in Protocol 2 (Section 23.2.2.3) and spread onto LB plates with an appropriate selection agent.

2) Selected colonies are grown in 96-well format deep-well plates overnight. The plates are centrifuged at 1300 × *g* for 3 min, and the broth removed by inverting the plate.
3) 80 µl of cold Solution 1 is added to each well and vortexed for 1 min, followed by the addition of 80 µl of Solution 2. This is mixed together by vortexing for 1 min and left at room temperature for 2 min (*Note*: Do not leave for more than 5 min). 80 µl of Solution 3 is added and mixed by vortexing for 1 min.
4) A multiscreen binding plate (Millipore, MAFBNOB) is placed onto the base of the vacuum manifold, and 160 µl of binding solution (8 M guanidine hydrochloride) is added. The upper section of the vacuum manifold is placed above this, and a clearing plate (Millipore, MANANLY) then placed on top.
5) 140 µl of the bacterial lysate is added to the clearing plate and drawn through the clearing plate by vacuum for 3 min. The lysate and binding solution are mixed by multichannel pipetting and drawn through the column by vacuum for 1 min.
6) The bound lysate is washed twice with 180 µl of 70% ethanol by drawing through the columns by vacuum for 1 min. After the second wash, the vacuum is maintained for a further 2 min, and the plate then centrifuged at 1300 × *g* for 10 min to dry the columns.
7) 50 µl sterile distilled water is added to each column, left at room temperature for 10 min, and placed on a microtiter PCR plate using the alignment frames and centrifuged at 1300 × *g* for 10 min.

Sequencing and Sequence Analysis

Sequencing was carried out as described in Protocol 2 (Section 23.2.2.4), and the sequences were analyzed using Sequencher v3.8 (Gene Codes Corp, Inc., Ann Arbor, MI, USA) to generate contigs and aligned using ClustalW [12].

23.3 Example Experiments

SR proteins have been identified as positive regulators of splice site selection and have a key role in AS. By using overexpressing plant lines of plant SR proteins, the AS RT-PCR panels were used to identify changes in AS as a consequence of overexpressing three different *Arabidopsis* SR proteins, namely RSp31, RSZ33, and SRp30.

RNA was extracted from three biological repeats of a wild-type plant and the three overexpressing SR lines. AS RT-PCR analysis revealed that about one-quarter of the AS events tested showed significant changes. Two examples occurred in two repair endonuclease genes, where the ratio of alternatively spliced products was changed significantly between the wild-type plant and the over-expressing lines (see Figure 23.1).

Novel RT-PCR products are detected in many of the AS RT-PCR reactions (Figure 23.2a). To identify these products, RT-PCR reactions were carried out on RNA from mutants to the NMD pathway and cycloheximide-treated plants, and cloned. The products of the ligation reactions were amplified by PCR to indicate cloning of multiple alternatively spliced products, followed by PCR of individual colonies to identify those containing different RT-PCR products for sequencing.

Colony PCR of the ligation reaction for At5g13730 (Figure 23.2b) identified plasmids for sequencing which resulted in three different sequences representing alternatively spliced forms (shown schematically in Figure 23.2c). One of the forms showed an accurate splicing of the second intron, while the second isoform utilized an alternative 3′ splice site (3′SS) 14 nt upstream of the normal 3′SS. A novel AS event showed the presence of the 170 nt of intron 2 as a result of intron retention (Figure 23.2c).

Full-length cDNA sequencing of transcripts from selected genes of interest has also been used to successfully identify new AS events (Figure 23.2d). In At3g09150, three alternatively spliced transcripts are annotated on the Arabidopsis Information Resource (TAIR; http://www.arabidopsis.org/) database, and the present analysis

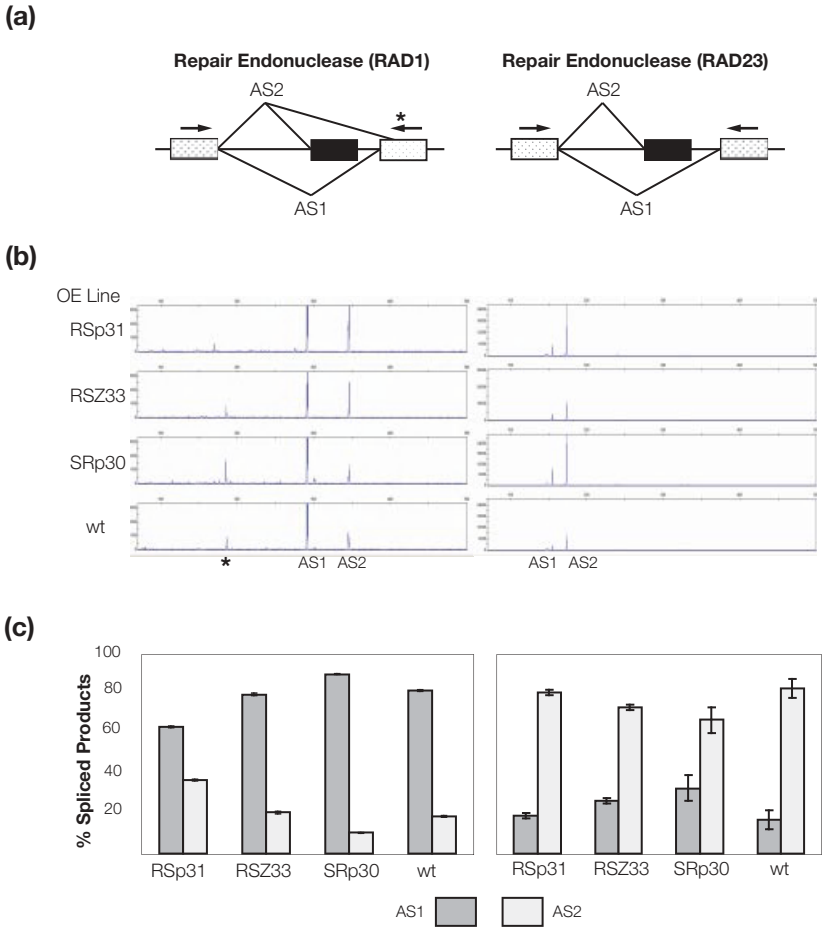


Fig. 23.1 Examples of changing alternative splicing (AS) patterns in overexpressing lines of SR proteins. (a) Schematic diagram showing AS pattern schemes for two repair endonuclease genes (RAD1 and RAD23). The open boxes refer to exon sequences, and black bars to intron sequence. The smaller open box shows the alternatively selected sequence, as indicated by lines labeled with AS1 and AS2, to indicate the events giving rise to different AS isoforms; (b) The * refers to a novel previously unannotated alternative splice site, while the black arrows indicate primers used for RT-PCR to generate the AS profiles from three different SR protein overexpressing (OE) lines, RSp31, RSZ33, and SRp30. wt = wild-type plant, and AS1, AS2 and * correspond to the alternatively spliced products indicated in the schematic diagram in panel (a); (c) Plots showing the percentage of spliced products for AS1 and AS2.

discovered a new AS event that results in the skipping of two exons (Figure 23.2d). By systematically analyzing the genes of interest, all AS events can be identified and included on AS RT-PCR panels.

23.4 Troubleshooting

Problem	Reason + solution
Overlap between LIZ and 6-FAM	Poor or inaccurate base-pair calling may occur as a result of peaks with high relative fluorescence; this leads to an overlap with the LIZ-labeled size markers. The peaks are present but may not be correctly assigned to the expected product sizes. The relative fluorescence units may still be used
High levels of short DNA products	DNA cleaning methods can be highly variable, and lead to varying amounts of low-molecular-weight products that inhibit the ligation reaction, leading to severely reduced numbers of colonies on plates. Agencourt AMPure beads (Beckman Coulter Genomics) were found efficiently and consistently to remove low-molecular-weight nucleic acids below 150 bp.

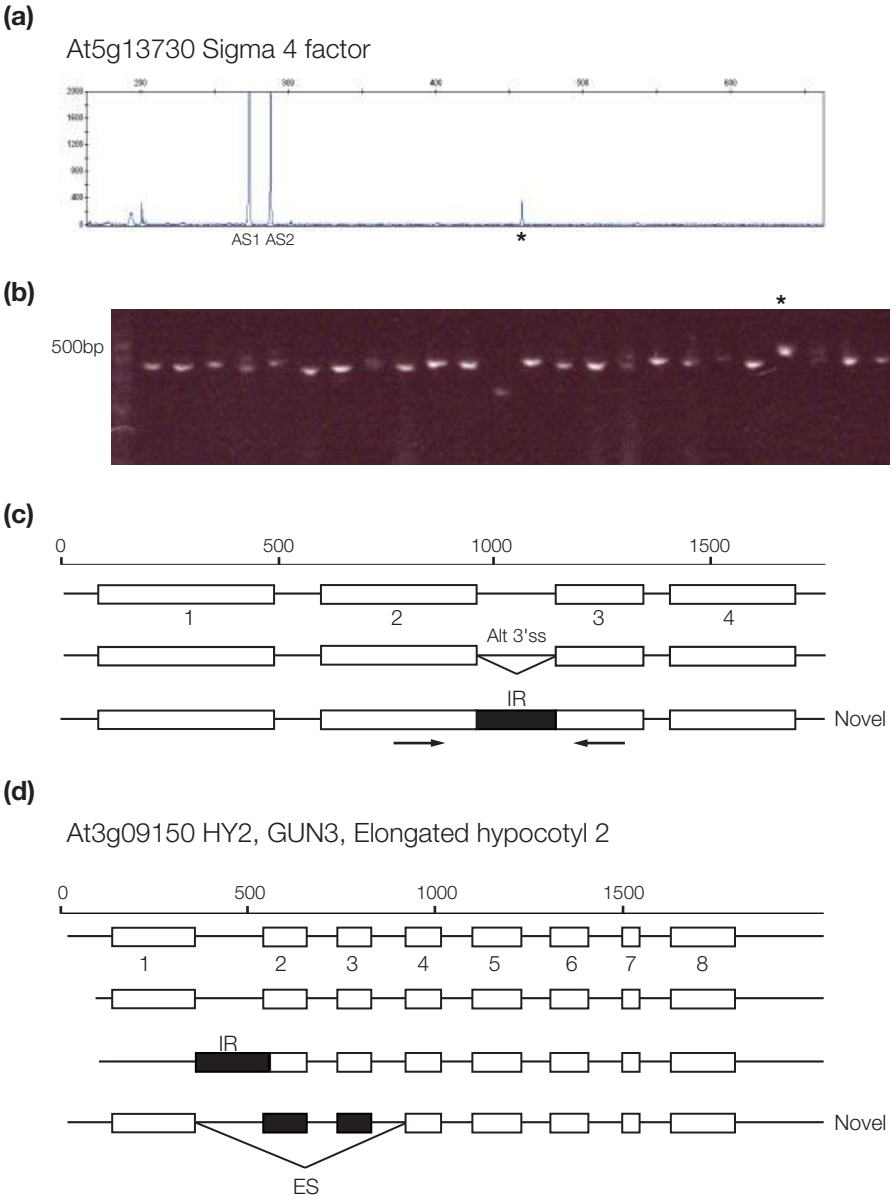


Fig. 23.2 Characterizing novel alternatively spliced events. (a) Genemapper profile of region of At5g13730, showing expected alternatively spliced transcripts AS1 and AS2 and an unknown product (*); (b) Colony PCR amplification identified clones that contain the novel alternatively spliced product (*). A 500 bp marker band is shown; (c) Gene transcript models derived from sequence analysis and alignment of cloned products. The alternative use of 3′ splice sites is indicated (Alt 3′ss), and the novel intron retained product is shown (IR). Gene length is indicated by the bar at the top of the gene model in steps of 500 bp. Open boxes refer to coding exon sequences, and are numbered. Black lines between the boxes show the intron sequence. The black bars indicate 5′ and 3′UTR sequences, and the lines below the splice type label (Alt 3′ss) indicate the splicing event. Newly discovered AS events are labeled as Novel; (d) Full-length cDNA sequencing of transcripts from the *HY2* gene (At3g09150). The multiple exon skipping event is indicated (ES), and the intron retained transcript is shown (IR). Gene length is indicated by the bar at the top of the gene model in steps of 500 bp. Open boxes refer to coding exon sequences and are numbered. Black lines between the boxes show intron sequence. The black bars indicate 5′ and 3′UTR sequences, and the lines above the splice type label (ES) indicate the splicing event. Newly discovered alternative splicing events are labeled as Novel.

Acknowledgments

The authors acknowledge funding for this research from the Biotechnology and Biological Sciences Research Council (BBSRC) [BB/G000212/1 and BB/G024979/1]; from the Scottish Government Rural and Environment Research and Analysis Directorate (RERAD) [WP114]; from the EU FP6 Programme Network of Excellence on Alternative Splicing (EURASNET) [LSHG-CT-2005-518238]; ERA-NET Plant Genomics Programme; and from the Austrian Science Foundation (FWF: SFB-F017/10/11) and the Austrian GEN-AU program.

References

1 Xing, Y. and Lee, C. (2006) Alternative splicing and RNA selection pressure-evolutionary consequences for eukaryotic genomes. *Nat. Rev. Genet.*, **7**, 499–509.
2 Reddy, A.S. (2007) Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant. Biol.*, **58**, 267–294.
3 Simpson, C.G., Lewandowska, D., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Raczynska, D., Jarmolowski, A., Barta, A., and Brown, J.W.S. (2008) Alternative splicing in plants. *Biochem. Soc. Trans.*, **36**, 508–510.
4 Barbazuk, W.B., Fu, Y., and McGinnis, K.M. (2008) Genome-wide analyses of alternative splicing in plants: opportunities and challenges. *Genome Res.*, **18**, 1381–1392.

5 Barta, A., Kalyna, M., and Lorković, Z.J. (2009) Plant SR proteins and their functions. *Curr. Top. Microbiol. Immunol.*, **326**, 83–102.
6 Ner-Gaon, H. and Fluhr, R. (2006) Whole-genome microarray in *Arabidopsis* facilitates global analysis of retained introns. *DNA Res.*, **13**, 111–121.
7 Lister, R., Gregory, B.D., and Ecker, J.R. (2009) Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Curr. Opin. Plant. Biol.*, **12**, 107–118.
8 Simpson, C.G., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Barta, A., and Brown, J.W.S. (2008) Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts. *Plant J.*, **53**, 1035–1048.
9 Kim, S.H., Koroleva, O.A., Lewandowska, D., Pendle, A.F., Clark, G.P., Simpson, C.G., Shaw, P.J., and Brown, J.W.S. (2009) Aberrant mRNA transcripts and the nonsense-mediated decay proteins UPF2 and UPF3 are enriched in the *Arabidopsis* nucleolus. *Plant Cell.*, **21**, 2045–2057.
10 Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**, 1413–1415.
11 Schlueter, S.D., Dong, Q., and Brendel, V. (2003) GeneSeqer@PlantGDB: Gene structure prediction in plant genomes. *Nucleic Acids Res.*, **31**, 3597–3600.
12 Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

2.3. Publication

Alternative splicing in plants

Craig G. Simpson, Dominika Lewandowska, John Fuller, Monika Maronova, Maria Kalyna, Diane Davidson, Jim McNicol, Dorota Raczynska, Artur Jarmolowski, Andrea Barta, John W.S. Brown

Biochem. Soc. Trans. (2008) 36, 508–510.

REVIEW

Alternative splicing in plants

Craig G. Simpson*, Dominika Lewandowska*, John Fuller*, Monika Maronova†, Maria Kalyna†, Diane Davidson*, Jim McNicol‡, Dorota Raczynska§, Artur Jarmolowski§, Andrea Barta† and John W.S. Brown*‖[1]

*Genetics Programme, SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K., †Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria, ‡Biomathematics and Statistics Scotland, SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K., §Adam Mickiewicz University, Institute of Molecular Biology and Biotechnology, Department of Gene Expression, Miedzychodzka 5, 60-371 Poznan, Poland and ‖Plant Sciences Division, University of Dundee at SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K.

Abstract

The impact of AS (alternative splicing) is well-recognized in animal systems as a key regulator of gene expression and proteome complexity. In plants, AS is of growing importance as more genes are found to undergo AS, but relatively little is known about the factors regulating AS or the consequences of AS on mRNA levels and protein function. We have established an accurate and reproducible RT (reverse transcription)–PCR system to analyse AS in multiple genes. Initial studies have identified new AS events confirming that current values for the frequency of AS in plants are likely to be underestimates.

Regulation of gene expression by AS (alternative splicing)

Pre-mRNA (precursor mRNA) splicing is the excision of intron sequences from pre-mRNAs mediated by the spliceosome. AS produces more than one mRNA from a single gene through the selection and utilization of alternative splice sites in the pre-mRNA [1]. AS is widespread in higher eukaryotes, with up to 74% of human genes containing one or more AS events [2]. The main consequences of AS are changes in protein structure/function, mRNA stability and regulation of gene expression. The inclusion or exclusion of whole or parts of exons or introns can alter protein domains, activity, localization, interactions with other proteins or substrates, and post-translational modification [3]. As such, AS increases the protein complexity of an organism (the ~30000 human genes produce 150000–200000 proteins). Gene expression can be affected directly by AS of TFs (transcription factors) and SFs (splicing factors), which modulate DNA recognition and binding, transcriptional complex assembly and splicing. In addition, AS can affect mRNA stability and turnover because many alternatively spliced transcripts contain premature termination codons and are therefore substrates for NMD (nonsense-mediated decay) [4,5].

The selection of alternative splice sites involves recognition of classical intron splicing signals or exonic/intronic enhancers or suppressors by RNA-binding factors that enhance or suppress use of particular splice sites [1]. Splice site choice is determined by virtue of their position relative to competing splice sites and/or through interactions or interference with other proteins or complexes. Major factors involved in alternative splice site regulation are families of SR (serine/arginine-rich) and hnRNP (heterogeneous ribonucleoprotein particle) proteins [1]. In addition, AS of specific gene transcripts or sets of transcripts is regulated by specific regulatory proteins. SR and hnRNP proteins have multiple roles associated with splicing, mRNA transport and translation, and often act antagonistically [6]. The relative amounts of alternatively spliced transcripts in different cells and tissues under different conditions reflect the relative levels and activities of the interacting factors, leading to the hypothesis of a 'cellular code' of *trans*-acting factors [6] (Figure 1). Further fine-tuning of expression is achieved by the co-ordinated AS of genes involved in the same biological pathway or process and the superimposition of such AS networks on transcriptional regulatory networks [7].

AS in plants

Although AS appears to occur less frequently in plants than in animal systems, it is clearly significant, with over 35% of genes in *Arabidopsis* and rice showing AS [8–10]. Alternatively spliced genes are involved in a range of plant functions, including growth and development, signal transduction, disease resistance, biotic and abiotic stress responses, flowering time and the circadian clock. For the vast majority of the reported alternatively spliced plant genes, there is little or no information on functional differences among the proteins produced or any mechanistic understanding of how alternative splice sites are selected. In addition, nothing is known about higher-order combinatorial control or co-ordination of AS of sets of genes.

Extensive microarray studies have shown how transcriptional profiles change during development and in response to environmental factors and stresses. Similarly, patterns of AS across the genome change under different conditions [11] and the essential role of AS is demonstrated by the many developmental and growth defects in plants

Key words: alternative splicing, *Arabidopsis*, precursor mRNA, splicing factor, transcription factor.
Abbreviations used: AS, alternative splicing; hnRNP, heterogeneous ribonucleoprotein particle; pre-mRNA, precursor mRNA; RT, reverse transcription; SF, splicing factor; SR, serine/arginine-rich; TF, transcription factor.
[1]**To whom correspondence should be addressed (email j.w.s.brown@dundee.ac.uk or john.brown@scri.ac.uk).**

Figure 1 | **Regulation by AS**

External stimuli and developmental cues induce changes in transcription, splicing, protein modification and localization. The resulting changes in the levels and activities of different SFs (histograms) induce changes in expression and AS of TFs and SFs (which feedback to further regulate expression) and other downstream genes to produce and modulate the cell's response.

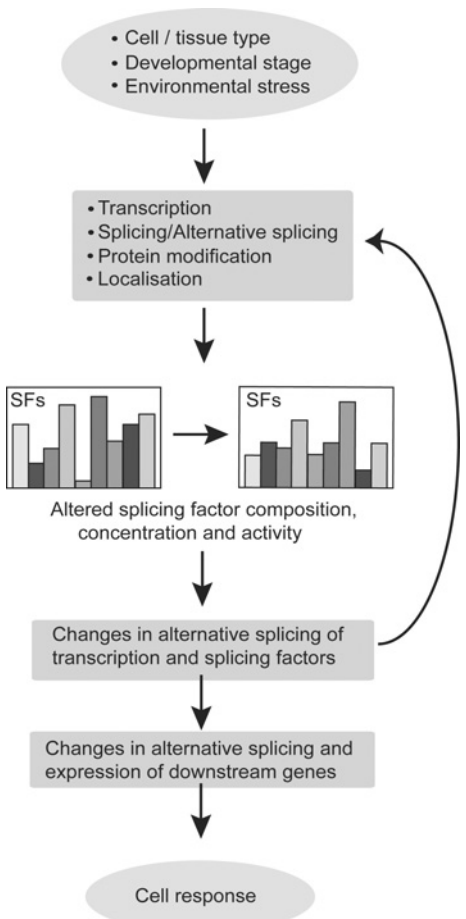
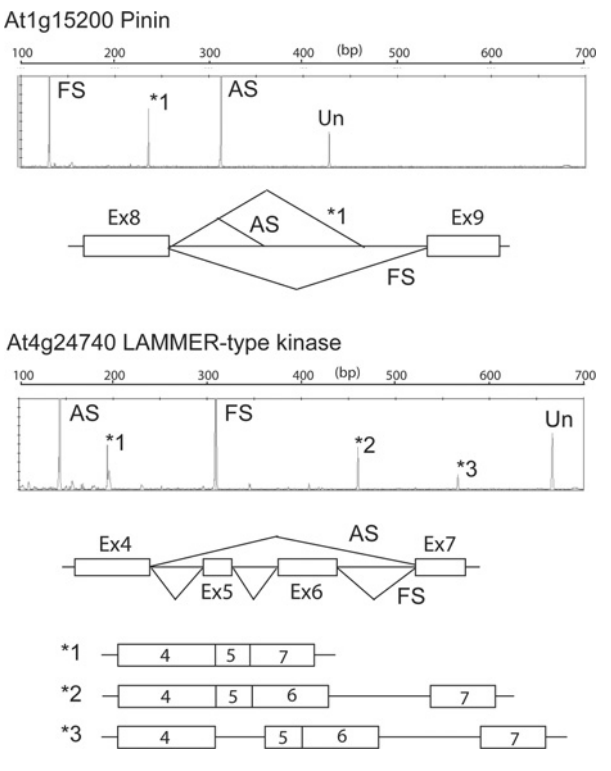


Figure 2 | **Detection of new AS events**

Scans (ABI3730) of RT–PCR products across known AS events detect unspliced (Un), fully spliced (FS) and the expected alternatively spliced (AS) transcripts shown for two splicing-related genes. Novel transcripts (*1–*3) represent new AS events and are indicated on the gene structures or as transcripts. Boxes, exons; horizontal lines, introns; lines above and below gene structures, splicing events.



overexpressing splicing regulators such as SR proteins [12,13]. Thus environmental and developmental cues affect the transcriptional and AS regulation of these factors which in turn modulate AS of downstream targets to control metabolic and developmental processes. In particular, AS of genes encoding RNA- or DNA-interacting proteins (e.g. SFs and TFs) can generate major changes in expression profiles, thereby amplifying signals and responses (Figure 1).

Monitoring AS events in plants detects novel AS transcripts

To begin to address the functions of RNA-binding proteins in AS and the co-ordination of AS in genes involved in the same biological process, we have established an AS RT (reverse transcription)–PCR panel to monitor multiple AS events from multiple plant genes simultaneously. By measuring the relative amounts of alternatively spliced transcripts, we are able to reproducibly detect statistically significant changes in the ratios of AS products [14]. To date, we have detected changes in AS of numerous genes in seedlings grown in the light and dark, in different organs (flower, leaf and root) of seedlings grown under short and long days, in transgenic lines overexpressing SFs [14] and in mutants of proteins involved in mRNA biogenesis. Significantly, many new AS events were detected in over 25% of the genes studied. Characterization of some of these events confirmed that they were previously unidentified AS transcripts (Figure 2), demonstrating the much wider occurrence of AS than currently expected in plants.

Given the current level of knowledge of AS in plants and the number of potentially undescribed AS events, RT–PCR-based approaches are the most accurate method of directly quantifying and monitoring the relative amounts of different alternatively spliced transcripts. We are currently increasing the number of AS events analysed on the RT–PCR panels and are developing distinct sets of AS events to measure changes in AS in genes involved in developmental and biological processes. Nevertheless, to obtain a true picture of AS in plants, a concerted effort to discover and

characterize all AS events will be needed. High-throughput sequencing systems will undoubtedly help in the discovery phase. This information will then allow the application of high-throughput technologies to measure changes in plant AS, as routinely as microarrays are currently used to measure transcript levels. Alternatively, splice junction, exon and tiled genome arrays have been used in animals to investigate global AS [2,7]. These approaches have advantages and disadvantages in terms of *de novo* discovery compared with reliance on annotated events, detection of different types of event, quantification of changes in AS, expense and computational requirements [15], but ultimately will benefit our understanding of the key role of AS in plants.

We acknowledge funding for this research from the European Alternative Splicing Network of Excellence (EURASNET), the Scottish Government, the Polish Government (grant number PBZ-KBN-110/P04/2004) and the Austrian Science Foundation (FWF).

References

1 Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. **72**, 291–336
2 Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science **302**, 2141–2144
3 Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. and Soreq, H. (2005) Function of alternative splicing. Gene **344**, 1–20
4 Maquat, L E. (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat. Rev. Mol. Cell Biol. **5**, 89–99
5 Lewis, B.P., Green, R.E. and Brenner, S.E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. U.S.A. **7**, 189–192
6 Smith, C.W. and Valcárcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem. Sci. **25**, 381–388
7 Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. Cell **126**, 37–47
8 Wang, B.B. and Brendel, V. (2006) Genome-wide comparative analysis of alternative splicing in plants. Proc. Natl. Acad. Sci. U.S.A. **103**, 7175–7180
9 Chen, F.-C., Wang, S.-S., Chaw, S.-M., Huang, Y.-T. and Chuang, T.-J. (2007) Plant gene and alternatively spliced variant *annotator*: a plant genome annotation pipeline from rice gene and alternatively spliced variant identification with cross-species expressed sequence tag conservation from seven plant species. Plant Phys. **143**, 1086–1095
10 Xiao, Y.-L., Smith, S.R., Ishmael, N., Redman, J.C., Kumar, N., Monaghan, E.L., Ayele, M., Haas, B.J., Wu, H.C. and Town, C.D. (2005) Analysis of cDNAs of hypothetical genes on *Arabidopsis* chromosome 2 reveals numerous transcript variants. Plant Physiol. **139**, 1323–1337
11 Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, K., Toyoda, T., Konagaya, A. and Shinozaki, K. (2004) Genome-wide analysis of alternative splicing in *Arabidopsis thaliana* based on full-length cDNA sequences. Nucleic Acids Res. **32**, 5096–5103
12 Lopato, S., Kalyna, M., Dorner, S., Kobayashi, R., Krainer, A.R. and Barta, A. (1999) atSRp30, one of two SF2/ASF-like proteins from *Arabidopsis thaliana*, regulates splicing of specific plant genes. Genes Dev. **13**, 987–1001
13 Kalyna, M., Lopato, S. and Barta, A. (2003) Ectopic expression of atRSZ33 reveals its function in splicing and causes pleiotropic changes in development. Mol. Biol. Cell **14**, 3565–3577
14 Simpson, C.G., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Barta, A. and Brown, J.W.S. (2008) Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts. Plant J. **53**, 1035–1048
15 Moore, M.J. and Silver, P.A. (2008) Global analysis of mRNA splicing. RNA **14**, 197–203

Received 19 February 2008
doi:10.1042/BST0360508

2.4. Publication

Regulation of plant gene expression by alternative splicing

Craig G. Simpson, Sujatha Manthri, Katarzyna Dorota Raczynska, Maria Kalyna, Dominika Lewandowska, Branislav Kusenda, Monika Maronova, Zofia Szweykowska-Kulinska, Artur Jarmolowski, Andrea Barta and John W.S. Brown

Biochem. Soc. Trans. (2010) 38, 667–671.

REVIEW

Regulation of plant gene expression by alternative splicing

Craig G. Simpson*, Sujatha Manthri†, Katarzyna Dorota Raczynska‡, Maria Kalyna§, Dominika Lewandowska*, Branislav Kusenda§, Monika Maronova§, Zofia Szweykowska-Kulinska‡, Artur Jarmolowski‡, Andrea Barta§ and John W.S. Brown*†[1]

*Genetics Programme, SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K., †Plant Sciences Division, University of Dundee at SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K., ‡Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Miedzychodzka 5, 60-371 Poznan, Poland, and §Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria

Abstract

AS (alternative splicing) is a post-transcriptional process which regulates gene expression through increasing protein complexity and modulating mRNA transcript levels. Regulation of AS depends on interactions between *trans*-acting protein factors and *cis*-acting signals in the pre-mRNA (precursor mRNA) transcripts, termed 'combinatorial' control. Dynamic changes in AS patterns reflect changes in abundance, composition and activity of splicing factors in different cell types and in response to cellular or environmental cues. Whereas the SR protein family of splicing factors is well-studied in plants, relatively little is known about other factors influencing the regulation of AS or the consequences of AS on mRNA levels and protein function. To address fundamental questions on AS in plants, we are exploiting a high-resolution RT (reverse transcription)–PCR system to analyse multiple AS events simultaneously. In the present paper, we describe the current applications and development of the AS RT–PCR panel in investigating the roles of splicing factors, cap-binding proteins and nonsense-mediated decay proteins on AS, and examining the extent of AS in genes involved in the same developmental pathway or process.

Introduction

AS (alternative splicing) generates more than one spliced mRNA isoform from the same gene through the selection of alternative splice sites. Different types of AS event include alternative 5′ and 3′ splice site selection, intron retention, exon skipping and mutually exclusive exon splicing, resulting in the inclusion or exclusion of intronic or exonic sequences [1–3]. The consequences of such changes in transcript sequence are altered protein sequence and functionality (activity, protein–protein or protein–substrate interactions, localization, protein modification, etc.) or the introduction of premature termination codons giving truncated proteins or leading to degradation of AS isoforms by NMD (nonsense-mediated decay) [1–5]. In humans, 95 % of multi-exon genes undergo AS [6], whereas, in plants, current estimates are that approx. 35 % of genes (in *Arabidopsis* and rice) undergo AS [7–11]. This level of AS is likely to be underestimated owing to the relatively lower level of ESTs (expressed sequence tags) available in plants when compared with, for example, humans or mice, and because many AS events are not represented or are under-represented in EST collections as they occur only in specific cells and tissues, growth conditions or stages of development [10–13]. AS is found in genes involved in a wide range of cellular processes, including transcription, splicing, development, signal transduction, and responses to biotic and abiotic stress contributing to, for example, seed development and quality, germination, disease resistance, flowering time and the circadian clock [10,11,14–20].

Although more and more examples of AS in plant genes are being discovered, very little is known about (i) the dynamic changes in AS of the majority of annotated alternatively spliced genes, (ii) how such changes are regulated, or (iii) the functions of protein isoforms produced from the AS transcripts. In order to begin to address the aspects of dynamic modulation and regulation of AS, we developed a high-resolution RT (reverse transcription)–PCR system to monitor changes in AS isoform abundance of approx. 300 *Arabidopsis* AS events [21]. The AS events represent different types of AS and are mainly in transcription factor, RNA-binding protein or stress-related genes. By using overexpression lines and knockout mutants of factors involved in splicing regulation [e.g. SR (serine/arginine-rich) proteins [21] and PTB (polypyrimidine tract-binding protein)-like proteins] and mRNA biogenesis [e.g. CBPs (cap-binding proteins)] [22], the AS RT–PCR system detects whether the factors influence AS and which events and types of AS event are affected. Similarly, by analysing mutants impaired in NMD, it is possible to identify AS isoforms which are subject to degradation by NMD. Although the AS RT–PCR system does not report on

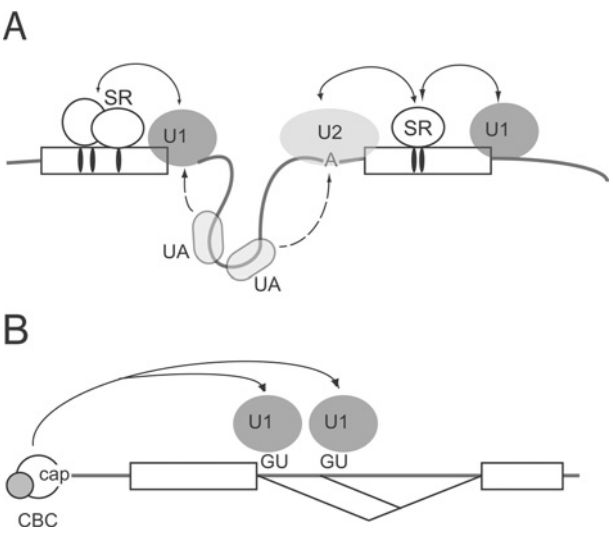
Key words: alternative splicing, *Arabidopsis*, precursor mRNA, splicing factor.

Abbreviations used: AS, alternative splicing; AtGRP, *Arabidopsis thaliana* glycine-rich protein; CBC, cap-binding complex; CBP, cap-binding protein; EST, expressed sequence tag; NMD, nonsense-mediated decay; PTC, premature termination codon; RT, reverse transcription; snRNP, small nuclear ribonucleoprotein; SR, serine/arginine-rich; UPF, up-frameshift; UTR, untranslated region.

[1]**To whom correspondence should be addressed (email j.w.s.brown@dundee.ac.uk or john.brown@scri.ac.uk).**

Figure 1 | **Regulation by AS**

(**A**) Splice site selection depends on multiple interactions between splicing factors and components and splicing signals in the transcript. SR proteins interact with other splicing factors and U1 and U2 snRNPs early in spliceosome assembly to select splice sites. In plants, this process may further involve UA-binding proteins interacting with the UA-rich introns. Differential interactions of different SR proteins with exon splicing enhancers promote alternative splice site selection. (**B**) In plants, the CBC can influence AS and does this preferentially at the 5′ splice site of the first intron, but can promote splicing at either the proximal or distal 5′ splice site [22]. Boxes represent exons; lines represent introns; vertical lines in exons represent exon splicing enhancer sequences; U1, U1 snRNP at the 5′ splice site; U2, U2 snRNP at the branchpoint (A); UA, UA-binding protein.



AS across the whole genome, the number of different events (approx. 300) is large enough to indicate the extent to which particular proteins affect AS. Finally, in addition to exploiting the genetic resources of *Arabidopsis* (in terms of mutant lines and transgenic overexpression lines), the AS RT–PCR system is being used to assess splicing efficiency and AS in different plant organs, developmental stages and under different growth and stress conditions (e.g. temperature). For example, a subset of intron-retention events have been used to examine splicing efficiency in plants under anoxic conditions which cause a core exon junction complex protein to relocalize to the nucleolus [23].

Regulation of AS

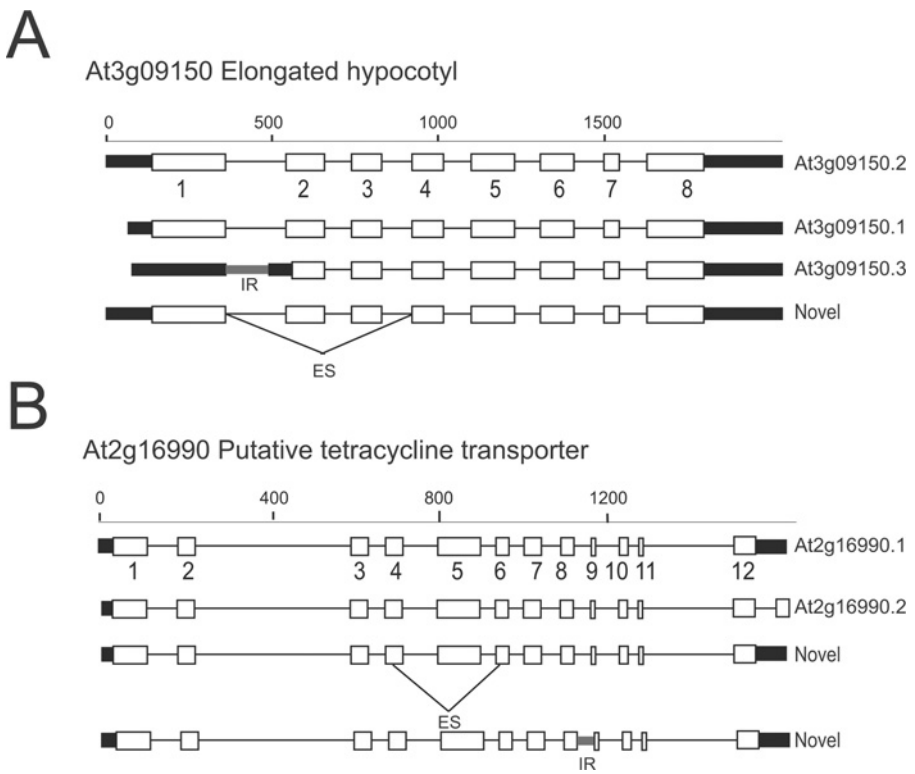
Regulation of AS is determined by the interactions of multiple *trans*-acting factors with intronic and exonic signals in the transcript giving rise to the concept of combinatorial control and a splicing code [24,25] (Figure 1A). Most progress in terms of *trans*-acting factors involved in plant splicing regulation has been made in the characterization of *Arabidopsis* SR proteins [26,27]. SR proteins are involved in both constitutive splicing and AS and generally promote splice site selection by binding to splicing enhancer sequences in the pre-mRNA and through protein–protein interactions (Figure 1A). In *Arabidopsis*, the essential role of these proteins is reflected in the range of developmental changes in plants where SR proteins are overexpressed or mutated [27–30]. SR protein genes are also of particular interest because the majority are themselves alternatively spliced and some of the AS events are evolutionarily conserved, suggesting that their regulation by AS is important for plant development [31]. Moreover, protein isoforms, produced from different AS transcripts of the *Arabidopsis* SR protein gene, *SR45*, are able to complement different developmental phenotypes of an *sr45* mutant demonstrating functional differences for the AS products [32]. Changes in expression of SR proteins are, not surprisingly, associated with changes in AS of many genes, including other SR protein genes themselves [20,28,29], and the AS RT–PCR panel is currently being used to analyse the subsets of AS events which are affected by particular SR proteins.

In addition to splicing factors, other proteins influence splicing. For example, the CBC (cap-binding complex) has a number of functions in mRNA biogenesis, including splicing, mRNA export, protection of the transcripts from nuclease degradation and 3′-end formation by stabilizing the interaction of the 3′-end processing machinery. In mammals and yeast, the CBC at the 5′ end of the pre-mRNA transcript promotes the initial interaction between U1 snRNP (small nuclear ribonucleoprotein) and the 5′ splice site of the first intron in the transcript and enhances the formation of spliced mRNAs [33]. However, the CBC has not been shown to affect AS to date. The single-knockout mutants of CBP20 and CBP80 and the double mutant are viable, but have slow growth, have serrated leaf margins, are late-flowering, are hypersensitive to abscisic acid and have enhanced drought tolerance [34–37]. Effects on splicing by knocking out CBC proteins have been observed in flowering time genes and in the increased occurrence of unspliced introns detected on microarrays [36–38]. To investigate effects on AS, the AS RT–PCR panel has been used in conjunction with single and double mutants of CBP20 and CBP80 [22]. Of ~250 AS events, 101 showed significant changes in AS in at least one mutant, and 41 of these showed changes in all three mutant lines. In the latter group, the number of events involving the first intron in a transcript and alternative 5′ splice site selection was significantly increased, suggesting that the *Arabidopsis* CBC preferentially (but not exclusively) affected AS of the first intron and particularly at the 5′ splice site [22] (Figure 1B). Moreover, we observed that CBP80 exerts a stronger influence on splicing than CBP20.

The NMD pathway targets mRNAs containing PTCs (premature termination codons) for degradation. In both mammals and plants, mRNAs with an increased distance between the PTC and the 3′ end of the mRNA [long 3′ UTRs (untranslated regions)] and/or with an exon junction complex downstream of a PTC are targeted for degradation [4,39,40]. NMD is thought to reduce the consequences of mutations or mistakes in expression, to reduce genomic noise

Figure 2 | Identification of novel AS events for process-specific RT–PCR panels

Cloning and sequencing of RT–PCRs covering the gene sequences has identified novel AS events of flowering time genes (an example is shown in **A**) and cold-stress induced genes (an example is shown in **B**). AS events in TAIR (The *Arabidopsis* Information Resource) are shown below the exon–intron structure of the gene, and novel events are labelled. White boxes represent exons; black boxes represent 5′ and 3′ UTRs; horizontal lines represent introns; lines below gene structures represent splicing events. IR, intron retention; ES, exon skip.



and prevent production of potentially detrimental truncated proteins. In addition, NMD modulates the levels of AS isoforms which contain PTCs. Approx. 10% of the human and yeast transcriptome is turned over by NMD [41,42] and 30% of human AS transcripts contain PTCs and are degraded by NMD [5]. In *Arabidopsis*, examples of regulation by AS and NMD include the genes encoding the circadian clock proteins, AtGRP (*Arabidopsis thaliana* glycine-rich protein) 7 and AtGRP8 [43,44]. By analysing mutants of the NMD proteins, UPF (up-frameshift) 1 and UPF3 (*upf3-1* and *upf1-5*), the AS RT–PCR system readily detects increases in the abundance of specific AS isoforms in a significant number of AS events. These potentially represent transcripts which are turned over by the NMD degradation pathway and are currently being characterized.

Process-specific AS

Ultimately, we wish to understand how transcriptional and post-transcriptional control systems are integrated in the regulation of gene expression. Current models suggest that post-transcriptional control networks such as AS are superimposed on transcriptional networks and that regulation of AS must be co-ordinated with regulation of transcription [45]. For example, the fine-tuning of developmental or metabolic pathways will reflect the integrated outputs of both transcription and AS networks such that changes in expression of one gene at the transcriptional or post-transcriptional levels could feed back on its own expression and the expression of other genes in the pathway. An excellent example of this type of control is the auto- and cross-regulation of the clock proteins AtGRP7 and AtGRP8 mentioned above.

Our current knowledge of AS events in plants is incomplete, making global approaches to studying AS (such as the use of exon arrays or splice junction arrays) presently unfeasible. Indeed, one of the outcomes of the use of the AS RT–PCR panel has been the unexpected number of novel RT–PCR products, some of which have been characterized and shown to be novel unannotated AS events. The number of novel RT–PCR products which we detect in our amplicons suggests that the level of AS is much higher in plants than currently estimated. Therefore, to be able to address co-ordination of post-transcriptional control by AS in genes in the same pathway, it is necessary to identify all the possible AS events in these genes. To this end, we are systematically cloning and sequencing AS events for all genes in particular pathways (e.g. the flowering time pathway) (Figure 2). AS

events are assembled from existing transcript information available in databases and from cloning and sequencing of RT–PCR products covering the coding sequence and UTRs of each gene. Novel AS events have been observed and characterized (Figure 2) and, combining this information, the feasibility of generating process- or pathway-specific AS RT–PCR panels is currently being tested. In the future, next-generation sequencing is expected to provide a platform for the global identification and quantification of AS isoforms in different plant species, but robust methods of data analysis for this particular function are still under development. The detailed characterization of AS events by the AS RT–PCR system will be invaluable both upstream in developing the next-generation sequencing data-analysis systems and downstream in validating changes in AS.

Funding

We acknowledge funding for this research from the Biotechnology and Biological Sciences Research Council (BBSRC) [grant numbers BB/G000212/1 and BB/G024979/1]; the Scottish Government Rural and Environment Research and Analysis Directorate (RERAD) [grant number WP114]; the European Union Framework Programme 6 Network of Excellence on Alternative Splicing (EURASNET) [grant number LSHG-CT-2005-518238]; European Research Area Network (ERA-NET) Plant Genomics; the Polish Government [grant numbers PBZ-KBN-110/P04/2004; PBZ/2/3/2006; PZ-19/2006/4/ZSK; 312/N-COST/2008/0]; the Austrian Science Foundation [grant number SFB-F017/10/11] and the Genome Research in Austria (GEN-AU) programme.

References

1 Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. **72**, 291–336
2 Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science **302**, 2141–2144
3 Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. and Soreq, H. (2005) Function of alternative splicing. Gene **344**, 1–20
4 Maquat, L.E. (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat. Rev. Mol. Cell Biol. **5**, 89–99
5 Lewis, B.P., Green, R.E. and Brenner, S.E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. U.S.A. **7**, 189–192
6 Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. **40**, 1413–1415
7 Xiao, Y.-L., Smith, S.R., Ishmael, N., Redman, J.C., Kumar, N., Monaghan, E.L., Ayele, M., Haas, B.J., Wu, H.C. and Town, C.D. (2005) Analysis of cDNAs of hypothetical genes on *Arabidopsis* chromosome 2 reveals numerous transcript variants. Plant Physiol. **139**, 1323–1337
8 Wang, B.B. and Brendel, V. (2006) Genome-wide comparative analysis of alternative splicing in plants. Proc. Natl. Acad. Sci. U.S.A. **103**, 7175–7180
9 Campbell, M.A., Haas, B.J., Hamilton, J.H., Mount, S.M. and Buell, C.R. (2006) Comprehensive analysis of alternative splicing in rice and comparative analysis with *Arabidopsis*. BMC Genomics **7**, 327–343
10 Barbazuk, W.B., Fu, Y. and McGinnis, K.M. (2008) Genome-wide analyses of alternative splicing in plants: opportunities and challenges. Genome Res. **18**, 1381–1392
11 Reddy, A.S.N. (2007) Alternative splicing of pre-messenger RNAs in plants in the genomic era. Annu. Rev. Plant Biol. **58**, 267–294
12 Hirose, T., Sugita, M. and Sugiura, M. (1993) cDNA structure, expression and nucleic acid-binding properties of three RNA-binding proteins in tobacco: occurrence of tissue-specific alternative splicing. Nucleic Acids Res. **21**, 3981–3987
13 Yoshimura, K., Yabuta, Y., Ishikawa, T. and Shigeoka, S. (2002) Identification of a *cis* element for tissue-specific alternative splicing of chloroplast ascorbate peroxidase pre-mRNA in higher plants. J. Biol. Chem. **277**, 40623–40632
14 Schöning, J.C., Streitner, C., Meyer, I.M., Gao, Y. and Staiger, D. (2008) Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in *Arabidopsis*. Nucleic Acids Res. **36**, 6977–6987
15 Larkin, P.D. and Park, W.D. (1999) Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. Plant Mol. Biol. **40**, 719–727
16 Egawa, C., Kobayashi, F., Ishibashi, M., Nakamura, T., Nakamura, C. and Takumi, S (2006) Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat. Genes Genet. Syst. **81**, 77–91
17 Balasubramanian, S., Sureshkumar, S., Lempe, J. and Wiegel, D. (2006) Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. PLoS Genet. **2**, e106
18 Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, K., Toyoda, T., Konagaya, A. and and Shinozaki, K. (2004) Genome-wide analysis of alternative splicing in *Arabidopsis thaliana* based on full-length cDNA sequences. Nucleic Acids Res. **32**, 5096–5103
19 Bove, J., Kim, C.Y., Gibson, C.A. and Assmann, S.M. (2008) Characterization of wound-responsive RNA-binding proteins and their splice variants in *Arabidopsis*. Plant Mol. Biol. **67**, 71–88
20 Dinesh-Kumar, S.P and Baker, B.J. (2000) Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. Proc. Natl. Acad. Sci. U.S.A. **97**, 1908–1913
21 Simpson, C.G., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Barta, A. and Brown, J.W.S. (2008) Monitoring changes in alternative precursor messenger RNA splicing in multiple gene transcripts. Plant J. **53**, 1035–1048
22 Raczynska, K.D., Simpson, C.G., Ciesiolka, A., Szewc, L., Lewandowska, D., McNicol, J., Szweykowska-Kulinska, Z., Brown, J.W.S. and Jarmolowski, A. (2009) Involvement of the nuclear cap-binding protein complex in alternative splicing in *Arabidopsis thaliana*. Nucleic Acids Res. **38**, 265–278
23 Koroleva, O.A., Calder, G., Pendle, A.F., Kim, S.H., Lewandowska, D., Simpson, C.G., Jones, I.M., Brown, J.W.S. and Shaw, P.J. (2009) Dynamic behaviour of the eIF4A-III putative core protein of the exon junction complex: fast relocation to nucleolus and speckles under hypoxia. Plant Cell **21**, 1592–1606
24 Smith, C.W. and Valcárcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem. Sci. **25**, 381–388
25 Matlin, A.J., Clark, F. and Smith, C.W.J. (2005) Understanding alternative splicing: towards a cellular code. Nat. Rev. Mol. Cell Biol. **5**, 89–99
26 Kalyna, M. and Barta, A. (2004) A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel gene functions. Biochem. Soc. Trans. **32**, 561–564
27 Barta, A., Kalyna, M. and Lorković, Z.J. (2009) Plant SR proteins and their functions. Curr. Top. Microbiol. Immunol. **326**, 83–102
28 Kalyna, M., Lopato, S. and Barta, A. (2003) Ectopic expression of atRSZ33 reveals its function in splicing and causes pleiotropic changes in development. Mol. Biol. Cell **14**, 3565–3577
29 Lopato, S., Kalyna, M., Dorner, S., Kobayashi, R., Krainer, A.R. and Barta, A. (1999) atSRp30, one of two SF2/ASF-like proteins from *Arabidopsis thaliana*, regulates splicing of specific plant genes. Genes Dev. **13**, 987–1001
30 Ali, G.S., Palusa, S.G., Golovkin, M., Prasad, J., Manley, J.L. and Reddy, A.S. (2007) Regulation of plant developmental processes by a novel splicing factor, PLoS One, e471
31 Kalyna, M., Lopato, S., Voronin, V. and Barta, A. (2006) Evolutionary conservation and regulation of particular alternative splicing events in plant SR proteins. Nucleic Acids Res. **34**, 4395–4405
32 Zhang, X.-N. and Mount, S.M. (2009) Two alternatively spliced isoforms of the *Arabidopsis* SR45 protein have distinct roles during normal plant development. Plant Physiol. **150**, 1450–1458
33 Lewis, J.D. and Izaurralde, E. (1997) The role of the cap structure in RNA processing and nuclear export. Eur. J. Biochem. **247**, 461–469

34 Hugouvieux, V., Kwak, J.M. and Schroeder, J.I. (2001) An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in *Arabidopsis*. Cell **106**, 477–487
35 Papp, I., Mur, L.A., Dalmadi, A., Dulai, S. and Koncz, C. (2004) A mutation in the cap binding protein 20 gene confers drought tolerance to *Arabidopsis*. Plant Mol. Biol. **55**, 679–686
36 Bezerra, I.C., Michaels, S.D., Schomburg, F.M. and Amasino, R.M. (2004) Lesions in the mRNA cap-binding gene *ABA HYPERSENSITIVE 1* suppress FRIGIDA-mediated delayed flowering in *Arabidopsis*. Plant J. **40**, 112–119
37 Kuhn, J.M., Hugouvieux, V. and Schroeder, J.I. (2008) mRNA cap binding proteins: effects on abscisic acid signal transduction, mRNA processing, and microarray analyses. Curr. Top. Microbiol. Immunol. **326**, 139–150
38 Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Rätsch, G. and Weigel, D. (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U.S.A. **105**, 8795–8800
39 Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A., Gyula, P., Lacomme, C., Barta, E., Nagy, F. and Silhavy, D. (2008) Inter-kingdom conservation of mechanism of nonsense-mediated decay. EMBO J. **27**, 1585–1595
40 Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Palfy, T., Barta, E. and Silhavy, D. (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
41 Mendall, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F. and Dietz, H.C. (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. **36**, 1073–1078
42 He, F., Li, X., Spatrick, P., Casillo, R., Dong, S. and Jacobson, A. (2003) Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. Mol. Cell **12**, 1439–1452
43 Schöning, J.C., Streitner, C., Page, D.R., Hennig, S., Uchida, K., Wolf, E., Furuya, M. and Staiger, D. (2007) Auto-regulation of the circadian slave oscillator component AtGRP7 and regulation of its targets is impaired by a single RNA recognition motif point mutation. Plant J. **52**, 1119–1130
44 Schöning, J.C., Streitner, C., Meyer, I.M., Gao, Y. and Staiger, D. (2008) Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in *Arabidopsis*. Nucleic Acids Res. **36**, 6977–6987
45 Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. Cell **126**, 37–47

Received 29 September 2009
doi:10.1042/BST0380667

Chapter 3

Plant-specific SR protein At-RS31 modulates alternative splicing and plant stress response

3.1. Introduction

The above analysis has shown that overexpression of two SR proteins At-SR30 and At-RS2Z33, caused significant changes in the patterns of several AS events on the HR RT-PCR screen (Simpson et al., 2008a). We are particularly interested to study functions of plant-specific SR proteins as they might have evolved to carry out plant-specific tasks (Barta et al., 2008). Gene *At-RS31* (At3g61860) has also been extensively studied in our laboratory and it became the focus of my research work on Arabidopsis.

At-RS31 has been identified as a member of plant-specific subfamily of SR proteins. It is able to restore splicing in deficient human S100 extract, which lacks SR proteins but contains all other factors necessary for splicing (Lopato et al., 1996). At-RS31 interacts with several proteins involved in splicing: an SR protein At-RSZ21, a homolog of human 9G8 (SFRS7) (Lopato et al., 2002), with U11–35K protein, a component of U11 snRNP (Lorkovic et al., 2005), with At-CYP59, a multidomain cyclophilin that interacts with C-terminal domain of RNA Polymerase II (Gullerova et al., 2006). In addition, At-RS31 is phosphorylated *in vitro* by SRPK4, SR protein-specific kinase 4 (de la Fuente van Bentem et al., 2006) and its subnuclear distribution changes upon phosphorylation/de-phosphorylation inhibition (Tillemans et al., 2005, 2006). Nuclear import of At-RS31 depends on transportin-SR MOS14, which is a nuclear import receptor for SR proteins (Xu et al., 2011). *At-RS31* itself is alternatively spliced (Figure 3.1), whereby only one of the splice variants produces the protein, and its splicing pattern is regulated by another plant-specific SR protein At-RS2Z33 (Kalyna et al., 2006; Simpson et al., 2008a). Alternative splicing pattern of *At-RS31* was shown to be influenced by light in a way that the light shifts alternative splicing to production of fully–spliced protein-coding isoform (Petrillo et al., 2014b). Alternative splicing of the orthologues of *At-RS31* is evolutionary highly conserved from green algae to angiosperms (Kalyna et al., 2006) and represents one of the most ancestral alternative splicing described to date (Irimia et al., 2011) implicating an

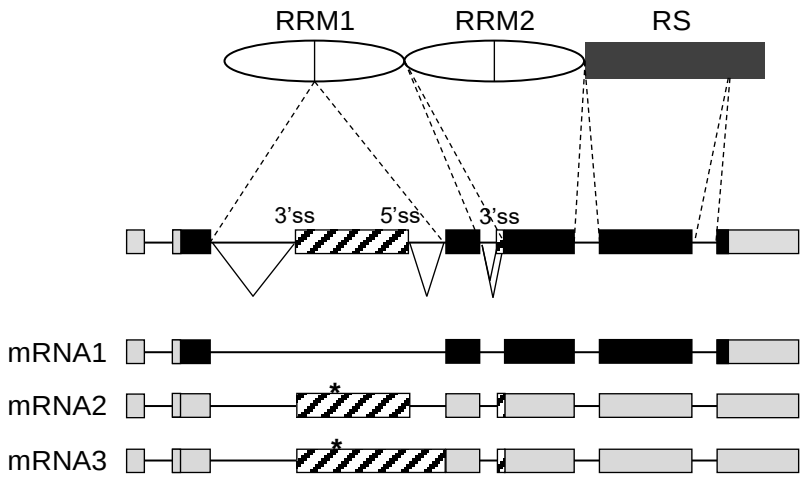


Figure 3.1. Schematic representation of *At-RS31* and the alternatively spliced transcripts. Depicted is protein domain structure of At-RS31, comprising two RNA recognition motifs (RRM) and the arginine/serine-rich region (RS). Dashed lines connecting the protein domains with the gene structure specify positions of introns. Gene structure denotes the alternative splice sites present in the long intron and upstream of the fourth exon. mRNA1 is the fully spliced protein-coding isoform, mRNA2 and mRNA3 are AS variants with premature termination codon (PTC) introduced by AS.
Exons are illustrated as boxes, introns as lines. Black and grey boxes depict coding and non-coding exons, respectively. Striped boxes represent intronic regions included due to alternative splicing, and asterisk indicates the first downstream PTC. Positions of start codons are shown by vertical lines. 5'ss/3'ss indicate alternative 5' and 3' splice sites. The schemes are drawn to scale.

ancient function of this regulation. However, neither effect of *At-RS31* on plant growth and development, nor downstream target genes of *At-RS31* have been identified.

3.2. Materials and methods

3.2.1. *At-RS31* overexpressing line 35S:At-RS31

To study impact of *At-RS31* on selected set of 90 AS events on the HR RT-PCR screen (Simpson et al., 2008a) we used transgenic line 35S:At-RS31, stably overexpressing At-RS31. In this line, expression of cDNA of protein-coding splice isoform (mRNA1, Figure 3.1.) is driven by the 35S CaMV (cauliflower mosaic virus) promoter and 35S

CaMV terminator. Overexpression of fully-spliced mRNA1 was verified by RT-PCR (Figure 3.2) and Northern blot (not shown).

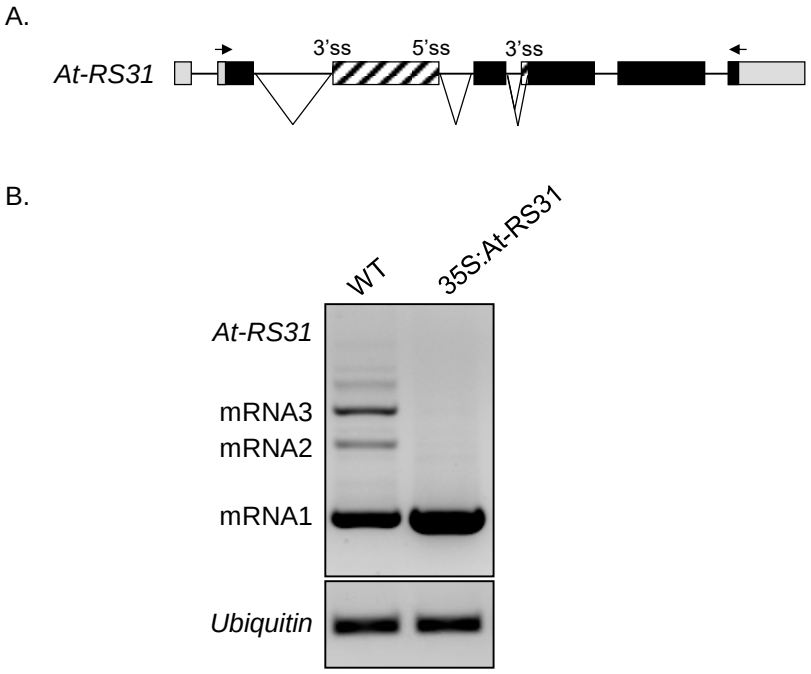


Figure 3.2. Expression of *At-RS31* in transgenic line 35S:At-RS31 compared to wild type (WT).
A. Schematic representation of the gene *At-RS31* with positions of primers used in RT-PCR analysis. Exons are illustrated as black and grey boxes, indicating the coding and non-coding regions, respectively. Striped boxes represent intronic regions included due to alternative splicing. Lines represent the introns. 5'ss/3'ss indicate alternative 5' and 3' splice sites. Arrows define positions of primers used in B. Scheme is drawn to scale.
B. RT-PCR analysis of *At-RS31*. RNA1-3 mark the AS isoforms (see Figure 3.1.). *Ubiquitin* was used as a loading control.

3.2.2. T-DNA insertion mutant lines of *At-RS31*

To complement the study of *At-RS31*, we identified two T-DNA insertion lines. PCR analysis and sequencing of the insertion sites localized the T-DNA into the following positions (Figure 3.3A). In line SALK_021332 (*at-rs31mut1*), the T-DNA insertion was detected in the fifth exon, 81 nt from exon beginning. In SALK_077769 (*at-rs31mut2*), insert was detected in the second intron, 131 nt downstream of the 5' splice site and 154 nt upstream of alternative 3' splice site.

RT-PCR analysis confirmed no detectable production of the *At-RS31* transcripts in the mutant lines compared to wild type control plants (Figure 3.3B).

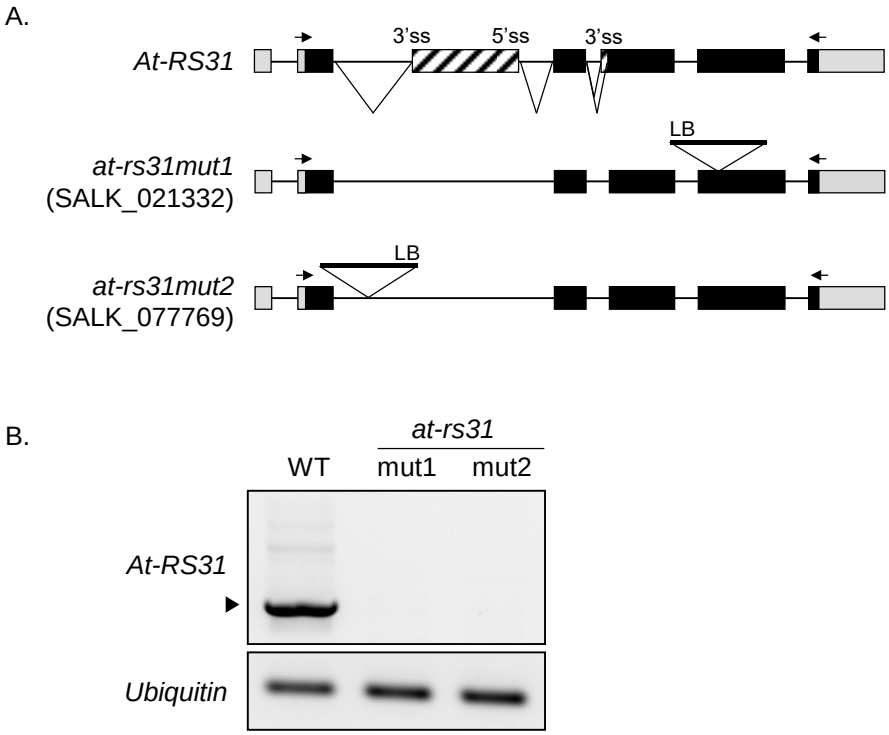


Figure 3.3. RT-PCR verification of *at-rs31mut1* and *at-rs31mut2* T-DNA insertion lines.
A. Schemes denote sites of two distinct T-DNA insertions in the *At-RS31* gene. Exons are depicted as black and grey boxes, indicating the coding and non-coding regions, respectively. Striped boxes represent intronic regions included due to alternative splicing. Lines represent the introns. 5'ss/3'ss indicate alternative 5' and 3' splice sites. LB – T-DNA left border. Arrows indicate positions of primers used in B. Schemes are drawn to scale.
B. RT-PCR detection of the *At-RS31* transcripts in wild type (WT) compared to mutant lines. Arrowhead indicates fully spliced protein-coding transcript (see Figure 3.1.). *Ubiquitin* was used as a loading control.

3.2.3. Preparation of plant material for HR RT-PCR screen

35S:At-RS31 and wild type (ecotype Col-0) seeds were stratified for three days in the dark at 4 °C and then germinated on agar plates containing GM media (Valvekens et al., 1988). Seedlings were grown in the growth chamber under light regime 16h day/8h night at 23°C. Whole seedlings were harvested 10 days after germination. For each line, three biological replicates were performed.

3.2.4. RNA isolation

Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen), according to manufacturer's instructions.

3.2.5. HR RT-PCR screen

HR RT-PCR screen was conducted as described in (Simpson et al., 2008a, 2012), Chapter 2.

3.2.6. Transcript diagrams

Schemes of AS transcripts were obtained by using a drawing tool accessible at http://rtd2.petervenhuizen.nl/ (Venhuizen, unpublished). The tool is based on the models annotated in AtRTD2 database (Zhang et al., 2017). Input data (AGI codes and sequences of primers) are listed in Table 3.1.

3.2.7. Conventional RT-PCR

Total RNA was reversely transcribed with oligo$(dT)_{15}$ primers using the Reverse Transcription System (Promega) according to manufacturer's protocol; 1 µg of RNA was used per RT reaction and RT reaction was incubated for 60 min.

PCRs were performed in 50 µL reaction volume and the PCR mixture consisted of 1 × Optimized DyNAzyme Buffer (Finnzymes), 1 mM $MgCl_2$, 0.25 mM of each dATP, dCTP, dGTP, dTTP, 1 µM of each primer, 2 µL of cDNA and 0.5 U DyNAzyme DNA Polymerase (Finnzymes). PCR products were by ethidium bromide stained 2% agarose gel-electrophoresis based on TAE. PCR programmes and primer sequences are listed in Table 3.2.

Table 3.1. AGI codes and primers used to obtain transcript schemes with the tool accessible at http://rtd2.petervenhuizen.nl/ (Venhuizen, unpublished)

Primer Pair	**AGI Code**	**Gene Name**	**Primer Sequences from 5' to 3'**
3	At1g09140	*At-SR30*	CCTGCTAGATCCATTTCCCC
			GATCTTGATCTTGATTTTG
42	At5g04430	*BTR1*	CCTTCGAGGAGCAGATGCGGGC
			CCTCCAGAACCGTTGGGTGC
67	At1g20693	*HMGB2*	GGCTAAAGGAGCAGGGCGTGG
			GCTACAGACTTGTTCTTGGGG
75	At2g36000	*mTERF25/EMB3114*	CTCGTTTAGTTTGGAGAATC
			CTTCATCATCATTCATTACAAA
58	At5g65050	*AGL31/MAF2*	CAAAGATCATTGATCGTTAC
			CTTCACTTCCCCCATCATTAGTTC
49	At5g41150	*UVH1/ATRAD1*	CCATCCTGACATGGGTTTTGTC
			CCAGTTCCTTTCTTCCGCCTGC
13	At1g79650	*RAD23B*	GGAGAACAAAGTTACGGAGG
			GGAACAGGTGAAGACTGGG
6	At1g52500	*AtFPG/AtMMH*	CCCGTTGCAGACTGCTTCTAGCC
			CCATCAACAAAGGCCTTTCC
19	At2g32320	tRNAHis guanylyltransferase	CTTTTGGAATGAGCACTCC
			CCGTAGGCAAAGGCAATATCC
71	At1g78810	expressed protein	GGACAACAATGAGGCTGAGG
			GAGTACACTAATACATATTG

Table 3.2. Primer sequences and PCR programmes used for confirmation of the mutant lines and for validation of the HR RT-PCR screen.

Gene (AGI)	Primer pair	Sequence from 5' to 3'	PCR programme					
			ID*	D*	A*	E*	FE*	C
At-RS31 (At3g61860)	RSp31SacI	AATGAGCTCGAATTGCGAATTAAGATAA	95/1m	95/30s	50/1m	72/1m	72/5m	35
	G47B73	CTTTGAGTAGCTTCAAGGG						
At-SR30 (AT1G09140)	SRp30-F1721pan	CCTGCTAGATCCATTTCCCC	95/1m	95/30s	50/1m	72/1m	72/5m	30
	SRp30-R2766pan	GATCTTGATCTTGATTTTG						
HMGB2 (AT1G20693)	HMGB2-F665	GGCTAAAGGAGCAGGGCGTGG	95/1m	95/30s	50/1m	72/1m	72/5m	35
	HMGB2-R1077	GCTACAGACTTGTTCTTGGGG						
UVH1/ATRAD1 (AT5G41150)	RAD1forw	CCATCCTGACATGGGTTTTGTC	95/1m	95/30s	50/1m	72/1m	72/5m	35
	RAD1rev	CCAGTTCCTTTCTTCCGCCTGC						
AtFPG/AtMMH (AT1G52500)	FPG-F1716pan	CCCGTTGCAGACTGCTTCTAGCC	95/1m	95/30s	55/1m	72/1m	72/5m	35
	FPG-R2173pan	CCATCAACAAAGGCCTTTCC						
Ubiquitin (AT3G62250)	NUBQ	GGTGCTAAGAAGAGGAAGAAG	95/1m	95/30s	50/1m	72/1m	72/5m	20
	CUBQ	CTCCTTCTTTCTGGTAAACGT						

* Temperature [°C]/Time

ID—initial denaturation, D—denaturation, A—annealing, E—extension, FE—final extension, C—number of cycles.

3.2.8. Phenotype analyses

Seeds of *A. thaliana* wild type (Col-0), 35S:AtRS31 overexpression line, *at-rs31mut1*, *at-rs31mut2* and *uvh1-1* (Liu et al., 2000) mutant lines were germinated and grown on ½ GM medium (Valvekens et al., 1988) at 16h light/8h dark, 23ºC, with stress conditions applied as described below.

Effect of abscisic acid (ABA) was tested by germination and growth of wild type, 35S:AtRS31 and *at-rs31mut1* plants on ½ GM plates supplemented with 0, 1 µM of ABA for 14 days after germination (DAG).

Sensitivity to NaCl was assayed by germination and growth of wild type, 35S:AtRS31 and *at-rs31mut1* plants on ½ GM plates supplemented with 0, 0.1 M NaCl for 14 DAG.

Senescence phenotype was tested in the seedlings of wild type, 35S:AtRS31 and *at-rs31mut1* grown on ½ GM for up to six weeks. Effect of sugars was checked by adding 0.12 M glucose to the ½ GM media.

Sensitivity to methyl viologen (MV) (Sigma-Aldrich) was examined by germination and growth of wild type, 35S:AtRS31 (two distinct lines N1 and N4), *at-rs31mut1* and at-rs31mut2 plants on ½ GM plates supplemented with 0, 0.1 and 0.5 µM of MV. Plates were incubated for 20 days in the vertical position.

Sensitivity to DNA damage was tested by exposure of 10 DAG old seedlings of wild type, 35S:AtRS31, *at-rs31mut1* and *at-rs31mut2* and *uvh1-1* (Liu et al., 2000) mutant lines to UV-C (254 nm) irradiation of 2000 J/m^2 generated by UV crosslinker (Stratagene). After the UV-C treatment plates with seedlings were incubated either in light (standard growth conditions) allowing photoreactivation, or were kept in dark for two days followed by transfer to the standard growth conditions. Phenotype analysis was performed one week after irradiation.

3.3. Results

3.3.1. At-RS31 modulates alternative splicing

Overexpression of *At-RS31* lead to significant changes in AS patterns of ten genes out of 90 monitored alternative splicing events in 89 genes (Table 3.3).

Among them, we detected changes in the alternative splicing of two genes involved in RNA processing: *At-SR30* (At1g09140) and *BTR1* (At5g04430) (Figure 3.4, Table 3.3).

At-SR30 is homolog of the human splicing factor, canonical SR protein ASF/SF2 (Lopato et al., 1999). Major alternative splicing events occur in the longest 10th intron of this gene. The fully spliced (FS) isoform encodes the full-length reference protein. Other isoforms are generated either by usage of alternative 3' splice site resulting in addition of 338 nt (AS) and PTC (Lopato et al., 1999), or by other AS events (mainly intron retentions) (Zhang et al., 2017) that are not monitored by this primer pair (Figure 3.4 A). At-RS31 overexpression inhibits usage of the alternative 3' splice site in the 10th intron and promotes complete removal of this intron thus potentially leading to production of the FS isoform coding for full-length protein (Table 3.3, Figure 3.4A). Regions upstream of the used primer pair have to be tested to make conclusions about the functional impact of At-RS31 overexpression on *At-SR30*.

BTR1 encodes a protein with K homology (KH) RNA binding domain which binds to tomato mosaic virus RNA and inhibits its multiplication (Fujisaki and Ishikawa, 2008). BTR1 is a homolog of human neuron-specific RNA-binding protein NOVA that regulate alternative splicing in brain and, in Arabidopsis, it interacts with the spliceosomal protein U1A (Arabidopsis Interactome Mapping Consortium, 2011; de la Fuente van Bentem et al., 2006; Jensen et al., 2000). *BTR1* gene possesses an alternative 3' splice site in the 5th intron that removes 63 nt from the fully-spliced form (FS) of the transcript, termed *BTR1L*. Selection of the alternative splice site gives rise to a shorter protein isoform BTR1S, which lacks 21 amino acids from the large spacer region between the second and the third KH-1 domains and is considered to be the active protein isoform (Fujisaki and Ishikawa, 2008). Interestingly, the longer variant BTR1L was shown to be phosphorylated by ABA-activated protein kinase SnRK2.6/OST1 (Wang et al., 2013). At-RS31 overexpression inhibits selection of the 3' alternative splice site in the 5th intron, thus potentially downregulating the BTR1S and promoting the BTR1L (Table 3.3, Figure 3.4 B).

Table 3.3. Genes from a panel of 90 alternative splicing events (in 89 genes) that showed significant alternative splicing changes induced by overexpression of At-RS31 (t-test significance at 5% or stronger)[a].

Primer Pair	Gene ID	Gene Name	Alternative Splicing Events		Wild Type		35S:RS31	
			Event	**Consequence**	**FS**	**AS**	**FS**	**AS**
Splicing Factors								
3	At1g09140	*At-SR30*	E11 Alt 3'ss (+338)	PTC	57	43	75	25
42	At5g04430[b]	*BTR1*	E6 Alt3'ss (-63)	-21 aa	33	67	38	62
Transcription Factors								
67	At1g20693	*HMGB2*	E3 Alt 5'ss (+104)	PTC	97	3	90	10
75	At2g36000	*mTERF25/EMB3114*	E1 Alt 5'ss (-104)	C-terminal change	63	37	73	27
58	At5g65050	*AGL31/MAF2*	E4 Alt 3'ss (-33)	-11 aa	84	16	74	26
DNA-repair genes								
49	At5g41150	*UVH1/ATRAD1*	E6 Alt 3'ss (+53)	PTC	82	18	64	36
13	At1g79650	*RAD23B*	E4 Alt 3'ss (-18)	- 6 aa	77	23	82	18
6	At1g52500	*AtFPG/AtMMH*	E8 Alt 3'ss (+158)	PTC	85	15	90	10
Other								
19	At2g32320	tRNAHis guanylyltransferase	E9 Alt 3'ss (-71)	PTC	63	37	69	31
71	At1g78810	expressed protein	E3 Alt 5'ss (+57)	PTC	80	20	73	27

[a]The results are the means from three different experiments using biological replicates and are represented in the percentage splicing value for a splicing event. FS corresponds to TAIR reference model. AS corresponds to alternative splicing event described. [b]FS corresponds to BTR1L (AT5G04430.2, TAIR reference protein); AS - BTR1S (AT5G04430.1) produces active protein.

Abbreviations: En Alt5'ss and En Alt3'ss - alternative 5' splice site and alternative 3' splice site in exon, where n is the exon number. The number in brackets indicates the number of nucleotides added (+) or removed (-) from the reference exonic sequence. PTC, premature termination codon; aa, amino acids.

Three transcription factors showed changes in alternative splicing, namely *HMGB2*, *mTERF/EMB3114* and *AGL31/MAF2* (Table 3.3, Figure 3.5).

HMGB2 belongs to a family of chromatin-associated proteins, that possess the high-mobility group (HMG) box domain (responsible for DNA-binding) and that are involved in regulation of transcription (Pedersen et al., 2010). The expression of *HMGB2* transcript was shown to be modulated by abiotic stresses. It was substantially increased by cold stress and down-regulated by either drought or salt stress, suggesting its role in expression of stress-responsive genes (Kwak et al., 2007). We have monitored alternative splicing in the 3rd intron, where selection of alternative 5' splice site leads to addition of 104 nt to the 3rd exon and introduces a PTC (Figure 3.5 A). Other potentially protein-coding transcripts contain additional AS events outside of the monitored region. At-RS31 overexpression decreased the production of the transcripts with the fully-spliced 3rd intron (FS), relatively to the alternatively spliced PTC+ variant. However, to evaluate the biological significance of this shift, it is needed to monitor other AS events (Table 3.3, Figure 3.5 A).

mTERF25/EMB3114 is a protein of mitochondrial transcription termination factor-related family. While vertebrate homologs of the mTERF family are responsible for regulation of mitochondrial DNA transcription, Arabidopsis mTERF25/EMB3114 protein was shown to be localized in plastids (Babiychuk et al., 2011). Alternative splicing of the sole intron produces three transcripts. Two expected transcripts generated by alternative 5' splice site selection were detected. They differ by 104 nt and possibly code for two proteins variable in C-terminal part. Transcript containing cassette exon (At2g36000.JC1) was not detected. Overexpression of At-RS31 increases the production of the longer splice variant (FS) coding for the longer protein isoform (Table 3.3, Figure 3.5 B).

AGL31/MAF2 is a transcription factor that functions as a temperature-dependent regulator of flowering time due to alternative splicing sensitive to the ambient temperature (Airoldi et al., 2015). Overexpression of At-RS31 restrains production of the longer splice variant encoding the reference protein and increases shorter isoform that is translated into 11 aa shorter protein (Table 3.3). An unexpected product appeared as a 324 nt peak. It is generated by retention of the 3rd intron together with selection of alternative 3' splice site that removes 4 nt from the 5th exon and is potentially non-coding (Figure 3.5 C).

Three DNA repair genes were detected, *UVH1/ATRAD1*, *RAD23B* and *AtFPG/AtMMH*, all of them involved in some kind of DNA excision repair (Table 3.3, Figure 3.6).

DNA endonuclease UVH1/ATRAD1 is a homolog of *S. cerevisiae* Rad1, involved in nucleotide excision repair (Fidantsef et al., 2000; Liu et al., 2000). Selection of alternative 3' splice site in the 5th intron adds 53 nt to the 6th exon. Overexpression of At-RS31 considerably decreases the production of the fully-spliced variant (FS) (Table 3.3, Figure 3.6 A), while the longer AS isoform is increased. The latter contains NMD signals (PTC >50 nt upstream of splice junctions) and was shown to be up-regulated upon cycloheximide treatment, therefore it is presumably degraded via NMD pathway (Kalyna et al., 2012). An intron retention event, also generating PTC+ transcript (At5g41150.P4) (Figure 3.6 A), was later confirmed by sequencing, but its contribution to the AS ratio was not determined in this experiment.

Arabidopsis RAD23B belongs to a family of proteins playing an essential role in proteolytic pathway (ubiquitin/26S proteasome system) as well as in base excision repair of the DNA (Farmer et al., 2010). A region between the 2nd and 5th exons was monitored, where selection of the alternative 3' splice site in the the 3rd intron causes removal of 18 nt in frame, from the 4th exon. At-RS31 overexpression shifts the relative abundance in favor of transcripts without the deletion. The abundance of alternatively spliced isoforms, lacking 18 nucleotides, is diminished (Table 3.3, Figure 3.6 B). The 18 nt removal introduced by selection of the alternative 3' splice site results in the deletion of 6 aa between the ubiquitin-like and UBA1 domains. Alternative in-frame exons are present in the region downstream of the used primer pair (At1g79650.P1, At1g79650.P4 and At1g79650.1), and impact of At-RS31 on their inclusion needs to be tested with specific primer pairs (Figure 3.6 B).

Formamidopyrimidine DNA glycosylase (AtFPG/AtMMH) is DNA repair enzyme, that recognizes and excises oxidized purines in the damaged DNA (Gao and Murphy, 2001). Alternative splicing in *AtFPG/AtMMH* is complex and generates 9 transcript isoforms (Figure 3.6 C) (Gao and Murphy, 2001; Marquez et al., 2012; Zhang et al., 2017). Primer pair was designed to monitor usage of alternative 3' splice site that adds 158 nt to the 8th exon and introduces a PTC (Table 3.3). Overexpression of At-RS31 down-regulates usage of this alternative splice site, thus potentially down-regulating production of the full-length protein (Table 3.3, Figure 3.6 C). Due to the complexity of AS in this gene, functional consequences of At-RS31 overexpression have to be tested with additional primer pairs.

Two additional genes, not falling into the above categories, were affected by overexpression of At-RS31, tRNAHis guanylyltransferase and At1g78810 (Table 3.3, Figure 3.7).

tRNAHis guanylyltransferase is involved in the processing of tRNA by adding an essential extra GMP to the 5**'** end of tRNA(His) (Placido et al., 2010). We monitored usage of alternative 3' splice site in the 9th intron, which removes 71 nucleotides from the 9th exon. At-RS31 overexpression inhibits this event and enhances relative abundance of the variants with the fully-spliced intron (Table 3.3, Figure 3.7 A). Down-regulated alternatively spliced isoform is an NMD target due to presence of NMD signals and its up-regulation in the NMD-impaired mutant *upf3-1* (Kalyna et al., 2012). Additional, mostly intron retentions, events were later annotated (Marquez et al., 2012), but their abundance is probably low due to generally low abundance of IR-containing transcripts. Inclusion of cassette exon (At2g32320.ID19 transcript) was not observed both in wild type and in At-RS31 overexpressing plants (Figure 3.7 A).

For the gene of yet unknown function (At1g78810), At-RS31 overexpression down-regulates protein-coding transcript and enhances production of the longer PTC+ AS variant generated by usage of alternative 5' splice site in the 4th intron (Table 3.3, Figure 3.7 B). Its sensitivity to NMD is not known. Intron retention event was annotated later (Marquez et al., 2012), and impact of At-RS31 overexpression on its abundance was not evaluated.

Figures 3.4-3.7. Genes with significantly changed alternative splicing ratio induced by overexpression of the SR protein At-RS31.
Top: Structures of the alternatively spliced transcripts and primers used (red arrowheads). Transcript models are derived from AtRTD2 database (Zhang et al., 2017). Number in parenthesis indicates the primer pair number (see Tables 3.1 and 3.3). Black and grey boxes represent coding and non-coding exons, respectively. Lines depict introns. Red numbers indicate sizes of the PCR products.
Bottom: Electropherograms of RT-PCR products in wild type and line overexpressing At-RS31, generated by GeneMapper. Numbers on the x-axis represent the size markers in bp; y-axis scale indicates relative fluorescence, reflecting transcript abundance. The AS event related to each product is denoted by the corresponding peak. FS corresponds to TAIR reference model. AS corresponds to alternative splicing event described in the Table 3.3. IR - intron retention. Unk – unknown.

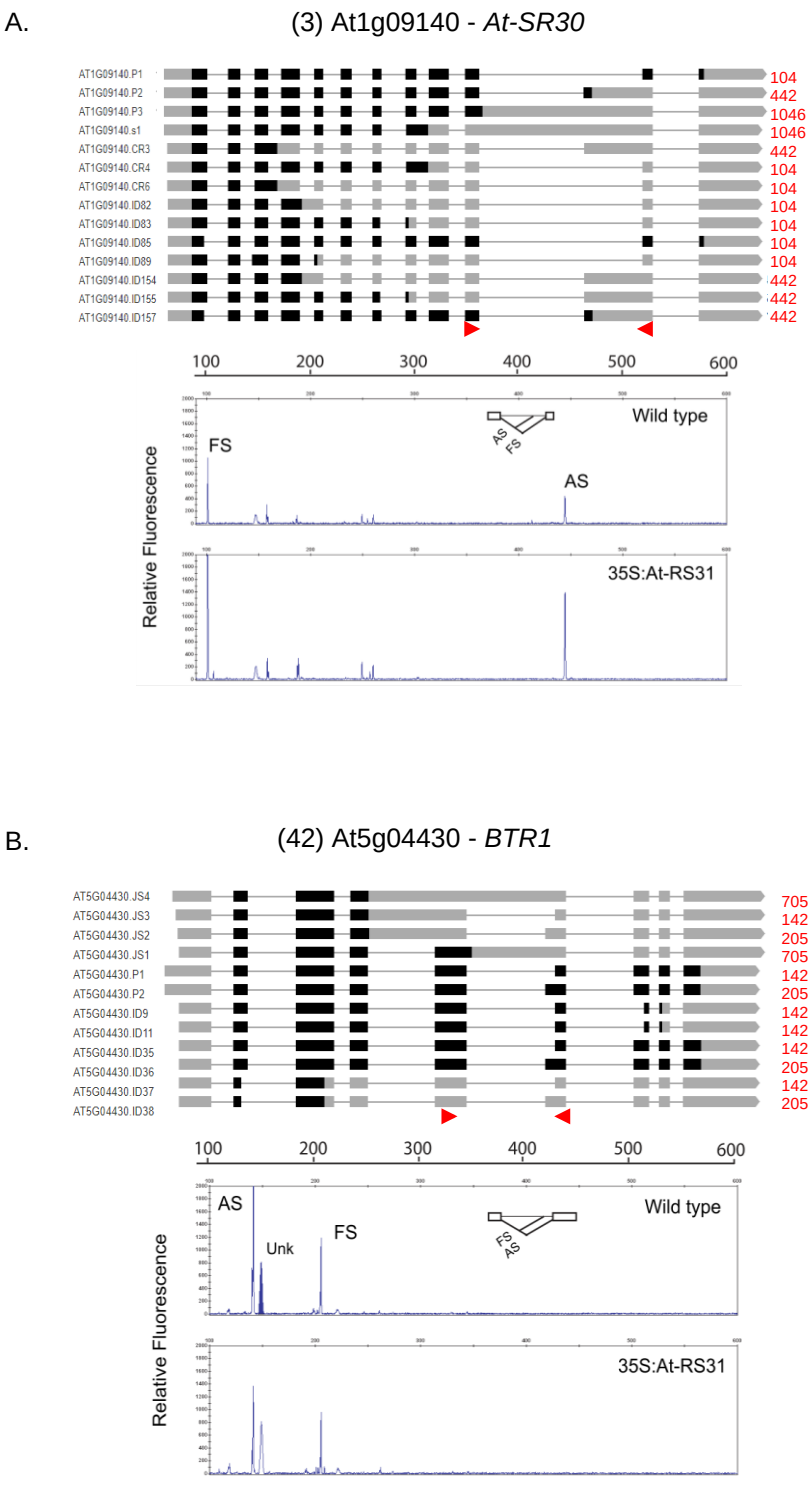


Figure 3.4. RNA processing genes with significantly changed alternative splicing ratio by At-RS31 overexpression.
A. At1g09140—SF2/ASF-like splicing factor, *At-SR30*. B. At5g04430—RNA-binding protein, *BTR1*. For the description see the legend, page 81.

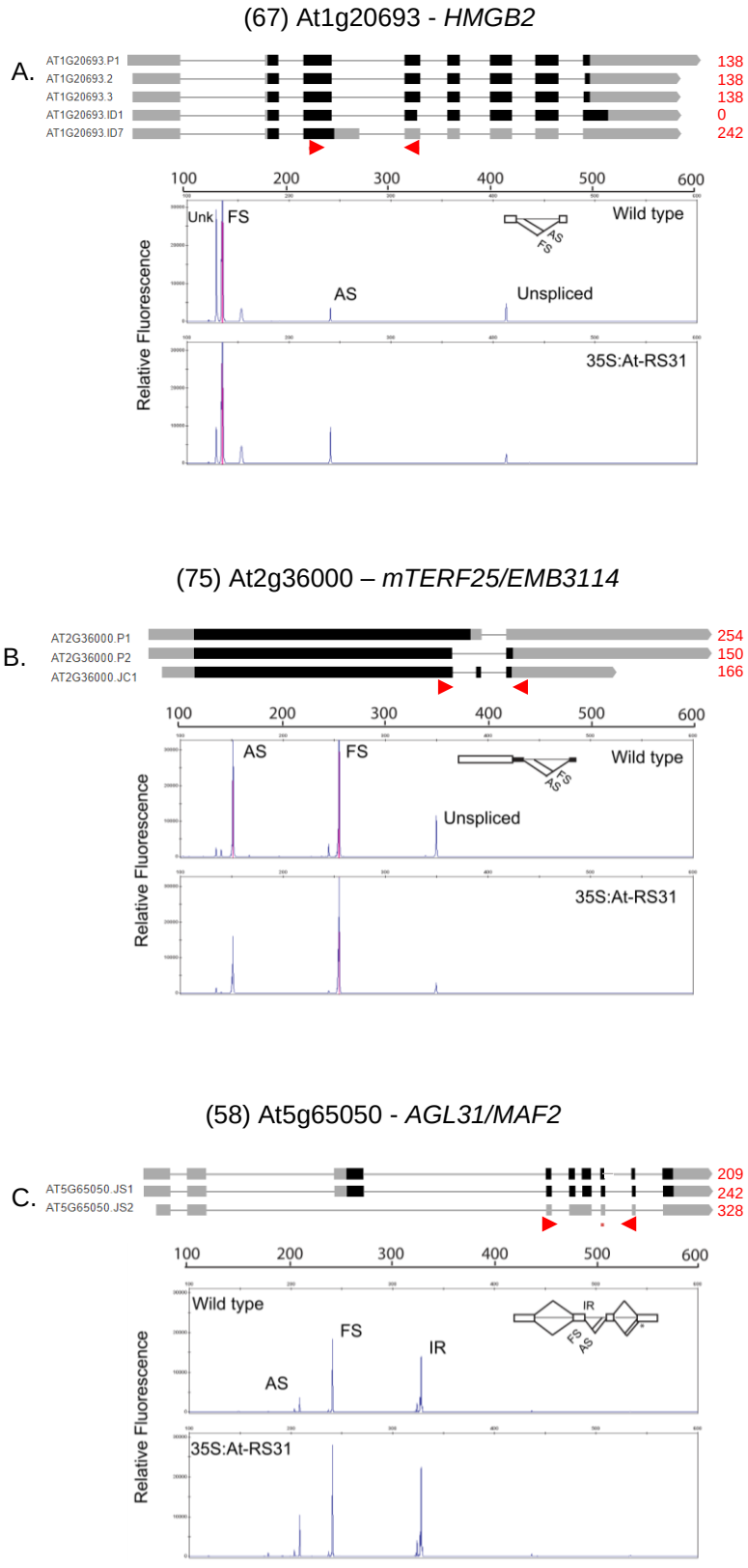


Figure 3.5. Transcription factors with significantly changed alternative splicing ratio due to At-RS31 overexpression.
A. At1g20693—HMG-box domain containing B2 protein, *HMGB2*. B. At2g36000—mitochondrial transcription termination factor-related protein, *mTERF25/EMB3114*. C. At5g65050—MADS box protein, *AGL31/MAF2*.
For the description see the legend, page 81

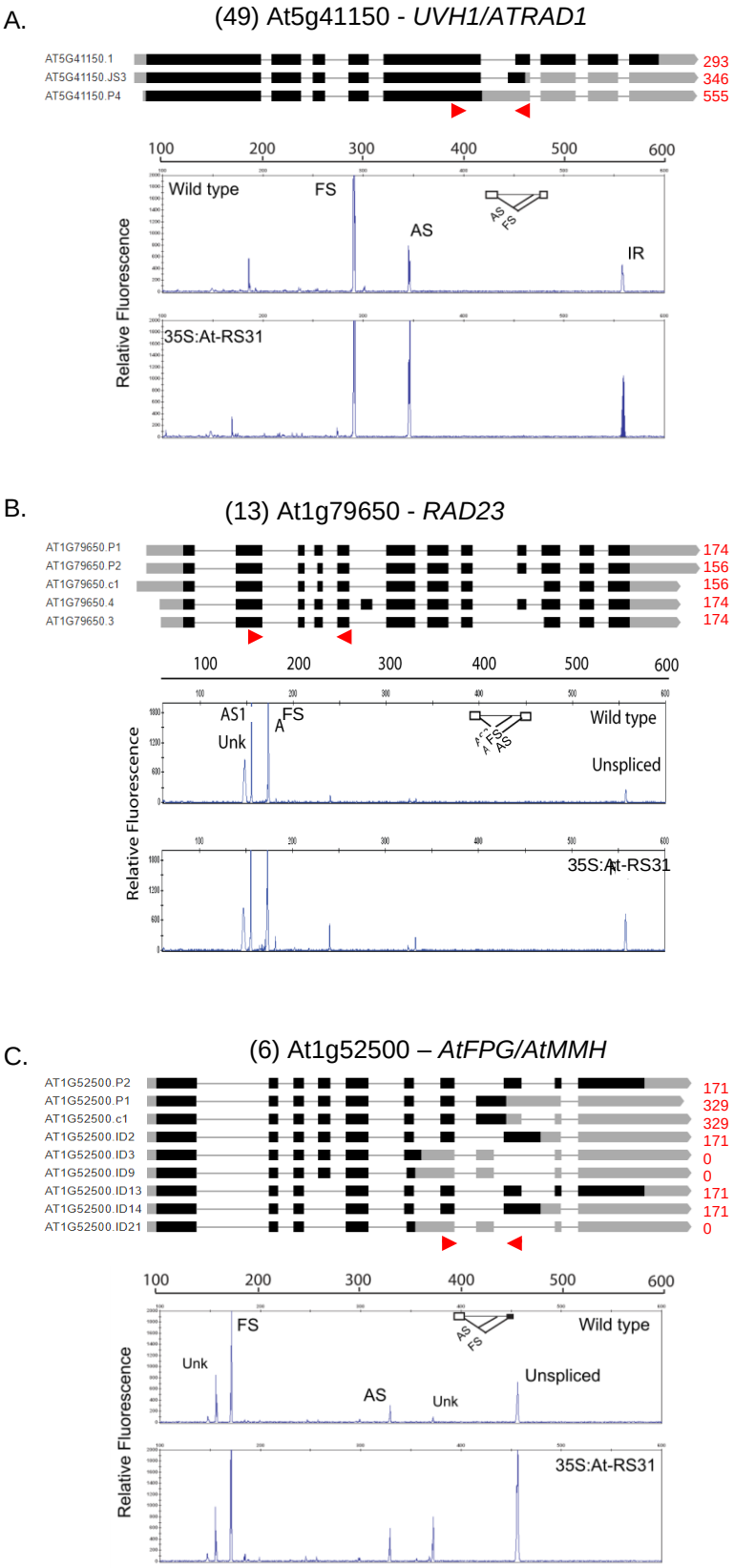
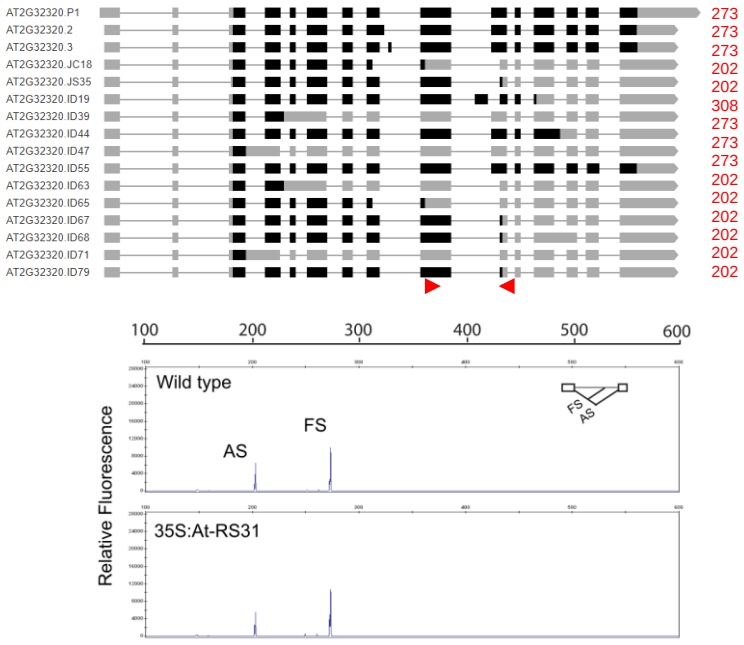


Figure 3.6. DNA repair genes with significantly changed alternative splicing ratio by At-RS31 overexpression.
A. At5g41150—DNA repair endonuclease, *UVH1/ATRAD1*. B. At1g79650—radiation sensitive 23B protein, *RAD23B*. C. At1g52500—Formamidopyrimidine DNA glycosylase, *AtFPG/AtMMH*. For the description see the legend, page 81

A. (19) At2g32320 - tRNAHis guanylyltransferase



B. (71) At1g78810 – expressed protein

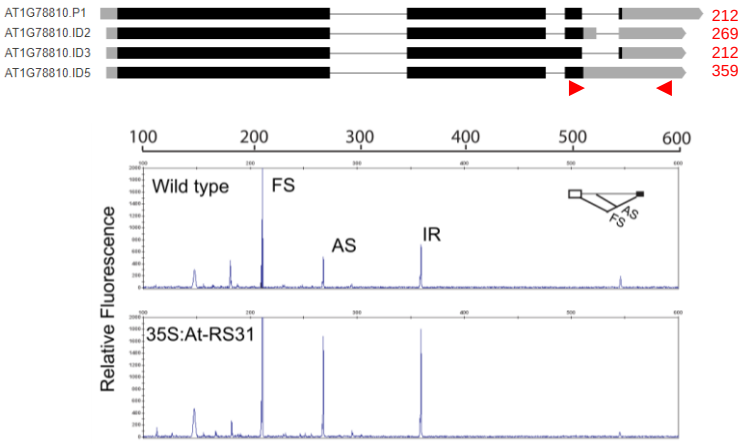


Figure 3.7. Other genes with significantly changed alternative splicing ratio by At-RS31 overexpression.
A. At2g32320—tRNAHis guanylyltransferase. B. At1g78810—expressed protein.
For the description see the legend, page 81.

3.3.2. Validation of the HT RT-PCR screen results, including the *At-RS31* mutant line

We used standard RT-PCR to further analyze *At-RS31* targets pointed out by HR RT-PCR screen, and we included meanwhile confirmed T-DNA insertion knock-out line *at-rs31mut1* (Figure 3.3.) to the analysis (Figure 3.8.).

Splicing factor *At-SR30* displayed the most prominent change in the HR RT-PCR screen, a shift from 57% to 75% towards the distal 3' splice site in the 10th intron used in the fully-spliced (FS) protein-coding isoform. The shift was observed with the RT-PCR, however the mutant line did not display any obvious change in the splicing pattern (Figure 3.8.).

In the HR RT-PCR screen, transcription factor *HMGB2* showed increase in usage of alternative 5' splice site in the 4th intron and therefore, potentially, decrease in production of FS isoform, and the same tendency was observed with the RT-PCR (Figure 3.8.). However, the effect seems to be stronger and in addition, overall down-regulation of the gene in overexpression line compared to wild type can be observed. Knock-out line appeared to have similar splicing pattern as the WT.

DNA repair endonuclease *UVH1/ATRAD1* was another gene strongly affected by At-RS31 overexpression and indeed, the RT-PCR confirmed the same direction of AS variant shift, an increase of the AS isoform generated by usage of alternative 3' splice site in the 5th intron (Figure 3.8.). *at-rs31mut1* shows much stronger expression of the fully spliced protein-coding isoform in comparison to the wild type. In addition, overall expression of both splice variants is enhanced in the *at-rs31mut1* line.

RT-PCR of the DNA repair enzyme *FPG* is difficult to assess. A shift towards the FS form is not convincing from the RT-PCR, and additional bands around the size of the expected AS product are present on the gel. Later sequencing of one of the additional bands, which is strongly upregulated in *At-RS31* overexpression line, appeared to be a transcript from the overlapping gene (At1g52510) on the opposite DNA strand (Fuchs et al., unpublished).

From the above RT-PCR validation we conclude that these results are in line with the data obtained from the HR RT-PCR screen. Changes in AS ratio in the At-RS31 overexpressing line are well detected in AS events, where the measured change in ratio was sufficiently high (at least 10%) and the size difference of the amplified splice variants can be well discriminated after separation on the agarose gel.

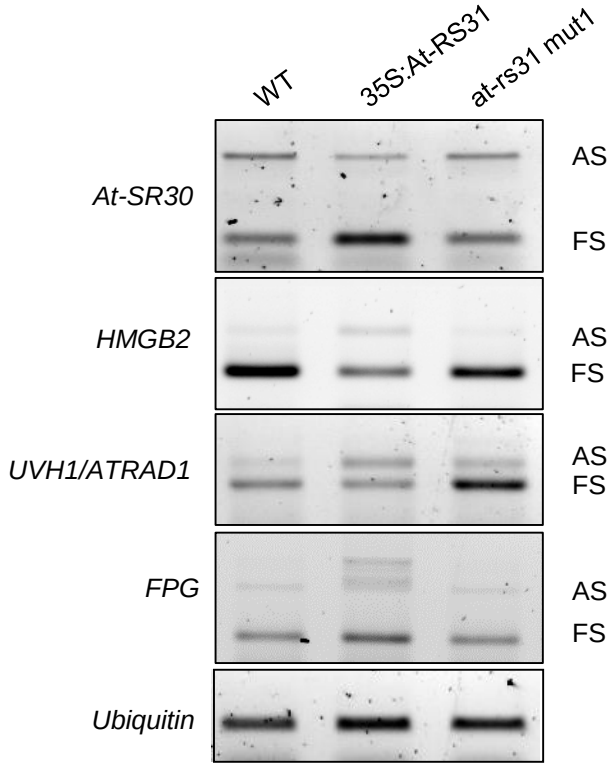


Figure 3.8. Conventional RT-PCR analysis of selected genes, with significantly changed AS ratio upon At-RS31 overexpression.
Splicing patterns of several genes were tested in At-RS31 overexpression and mutant lines, compared to wild type control (WT). *Ubiquitin* was used as a loading control. FS marks TAIR reference model, and AS corresponds to alternative splicing event depicted and described in Table 3.3 and Figures 3.4-3.6.

3.3.3. Expression level of *At-RS31* has impact on abiotic stress responses and senescence in Arabidopsis seedlings

It is known that SR genes play essential role in environmental and developmental signalling pathways, by modulating AS of gene expression regulators, as well as directly affect AS of downstream target genes (Filichkin et al., 2015; Staiger and Brown, 2013; Reddy, 2007; Duque, 2011).

Our HR RT-PCR analysis revealed that At-RS31 overexpression led to significant changes in AS of splicing and transcription factors as well as genes of DNA repair

pathway, suggesting a possible role for At-RS31 in gene expression regulation and stress response.

We were interested to examine phenotype of Arabidopsis plants with changed levels of At-RS31 upon continuous application of different abiotic stresses and DNA damaging treatments.

3.3.3.1. Overexpression of At-RS31 increases plant sensitivity to abscisic acid and high salt stress

Plant hormone abscisic acid (ABA) represents an important regulator of development and stress response in plants. Levels of endogenous ABA increase upon exposure to stress, such as drought, high salinity, cold and heat, indicating engagement of ABA in stress signal transduction (Cruz et al., 2014; Fernando and Schroeder, 2016). High salinity severely affects plant growth and development. Exposure to high salt environment causes osmotic stress to the plant cells and promotes increased levels of endogenous ABA (Cruz et al., 2014; Tuteja, 2007).

HR RT-PCR analysis suggested connection of At-RS31 to ABA through detection of a RNA-binding protein BTR1 as a potential AS target of At-RS31. Longer AS variant of *BTR1* (BTR1L) was identified as a substrate of ABA-activated protein kinase SnRK2.6/OST1 (Wang et al., 2013) and abundance of this mRNA variant was elevated upon overexpression of At-RS31.

We grew Arabidopsis seedlings overexpressing and lacking At-RS31, as well as wild type plants on medium supplemented with 1 µM ABA to mimic the stress conditions. Expecting the similar effect, due to cooperation of high salt stress and ABA signalling, we exposed seedlings of the same lines to high salt stress by adding 0.1 M NaCl to the growth media.

In our experimental setup, altered levels of At-RS31 have not influenced the plant phenotype in the absence of stress conditions. However, plants overexpressing At-RS31 were hypersensitive to both ABA and NaCl compared to wild type plants. *at-rs31mut1* mutant line showed slightly enhanced resistance to ABA than the wild type, but in high salt environment, the phenotype of *at-rs31mut1* resembled wild type line (Figure 3.9). This suggests that At-RS31 negatively regulates response to abscisic acid and high salt stress.

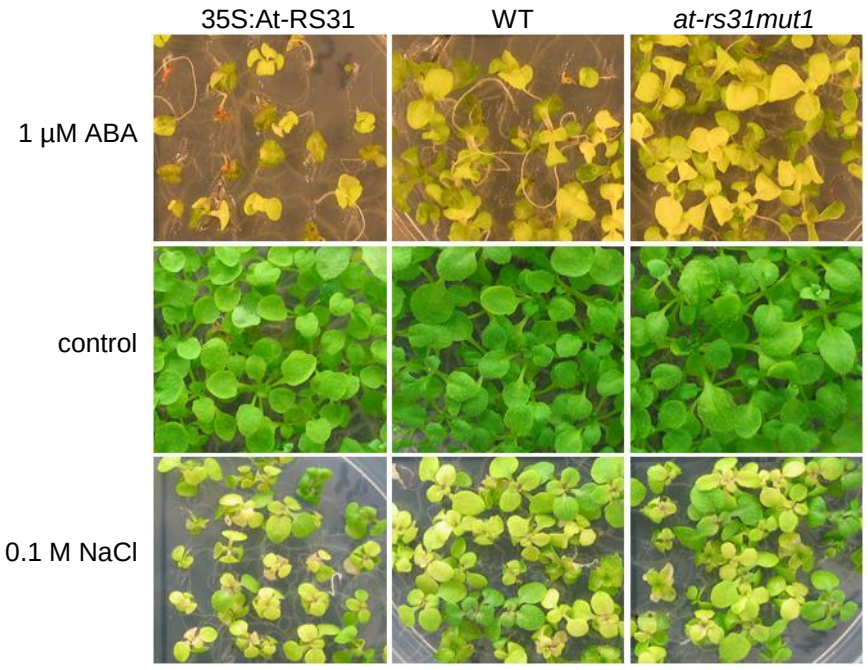


Figure 3.9. Increased levels of At-RS31 induce hypersensitivity to ABA and NaCl. Seedlings of wild type, At-RS31overexpressing (35S:AT-RS31) and mutant (*at-rs31mut1*) lines were germinated and grown on ½ GM plates, supplemented with 1 µM of abscisic acid (ABA) or 0.1 M NaCl for 2 weeks (23°C, 16h day/8h night).

3.3.3.2. At-RS31 promotes early senescence that is enhanced by glucose

Leaf senescence occurs naturally in plant development, but it is also induced by stress. The process of ageing is regulated by plant hormones, among others also ABA that acts as a promoter of senescence (Schippers et al., 2015). Sugars are also involved in regulation of leaf senescence, and exogenous sugars are able to induce an early senescence (Wingler et al., 2009; Wingler and Roitsch, 2008).

Under normal growth conditions, we have not observed any major changes in the phenotype of young At-RS31 overexpression and mutant plants in comparison to wild type (see Figure 3.9). We explored whether different expression levels of At-RS31 affect phenotype of plants at the later stages of development. After 5 weeks of growth under normal growth conditions, all three lines looked very similar, so the level of At-RS31 does not influence phenotype at this stage. After one more week, elevated levels of At-RS31 caused onset of early senescence in overexpressing line compared to wild type and mutant

plants. Interestingly, adding exogenous glucose into the media promoted leaf senescence in At-RS31 overexpressing plants already at 5 weeks after germination and in 6 weeks this effect is even more pronounced (Figure 3.10.). Excess of At-RS31 induces premature ageing that is augmented by exogenous glucose.

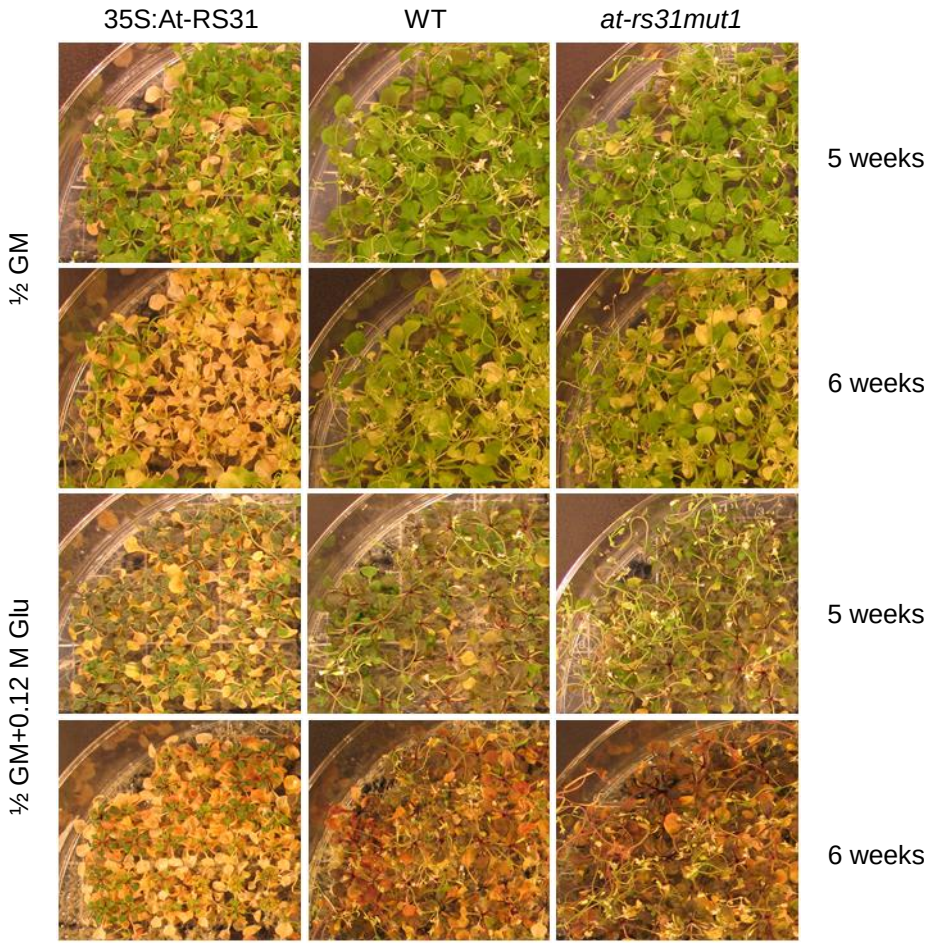


Figure 3.9. Increased levels of At-RS31 induce hypersensitivity to ABA and NaCl. Seedlings of wild type, At-RS31overexpressing (35S:AT-RS31) and mutant (*at-rs31mut1*) lines were germinated and grown on ½ GM plates, supplemented with 1 µM of abscisic acid (ABA) or 0.1 M NaCl for 2 weeks (23°C, 16h day/8h night).

3.3.3.3. Levels of At-RS31 influence plant sensitivity to oxidative stress

Abiotic stresses, such as drought, high temperature, salinity and heavy metals, as well as biotic stress initiate the production of high levels of reactive oxygen species (ROS) that generate reactive oxygen species (ROS) that are in high concentration toxic to the

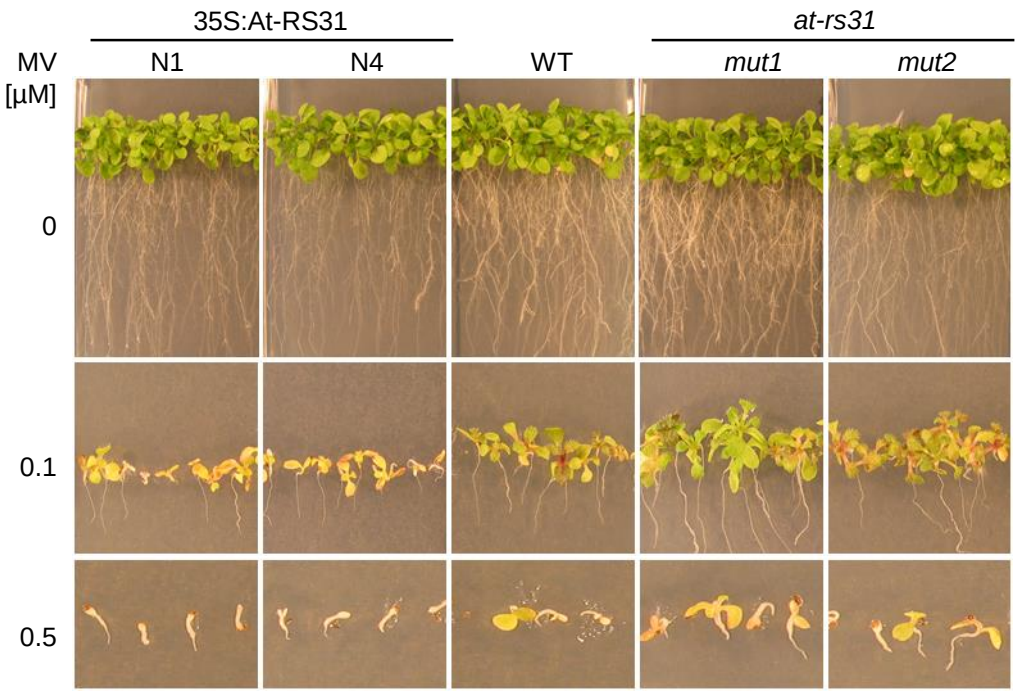


Figure 3.11. *At-RS31* overexpressing and mutant lines exhibit inverse phenotypic reactions to oxidative stress induced by methyl viologen (MV).
Seeds of wild type, two independent At-RS31 overexpression lines (35S:AT-RS31, N1 and N4) and two mutant (*at-rs31mut1*, *at-rs31mut2*) lines were germinated on ½ GM or ½ GM plates supplemented with 0, 0.1 and 0.5 µM methyl viologen (MV) and grown vertically at 23°C, 16h day/8h night for 20 days.

cells and cause oxidative damage to protein, lipids and DNA. ROS are produced also as a by-product of the normal oxygen metabolism. In moderate amounts function as signalling molecules and coordinate homeostasis and adaptation to stress. This regulated functions are executed in crosstalk with plant hormone networks, for instance auxin (Hirt, 2000; Krishnamurthy and Rathinasabapathi, 2013; Tognetti et al., 2012).

Methyl viologen (MV) is a redox-active compound that catalyses the formation of reactive oxygen species (ROS) by reduction of molecular oxygen (O^2) to the superoxide free radical, the precursor of most other ROS (Bus and Gibson, 1984), therefore it is used experimentally to induce oxidative stress (Slade, 1966).

We tested how Arabidopsis plants with altered levels of At-RS31 cope with the exogenously induced ROS production by growing the seedlings on media supplemented with different concentrations of MV (0, 0.1 and 0.5 µM). As mentioned previously, altered levels of At-RS31 had no visible effect on the phenotype of young seedlings under normal growth conditions (Figure 3.9), however application of MV induced phenotypic

differences dependent on the genetic background. Mutant lines showed enhanced resistance to oxidative stress in comparison to wild type, whereby cotyledons and leaves are more expanded and greener and the roots are longer. The effect was stronger in *at-rs31mut1* (disruption in the 5th exon) than in *at-rs31mut2* (disruption in the 2nd intron).

Conversely, elevated levels of At-RS31 aggravated sensitivity to oxidative stress compared to wild type, causing poorly developed rosettes and short roots in two independent At-RS31 overexpressing lines at the MV concentration 0.1 µM. Similar effects of At-RS31 are observed at the higher concentration of MV (0.5 µM) (Figure 3.11).

3.3.3.4. At-RS31 modulates sensitivity of Arabidopsis seedlings to UV-C irradiation

Our analysis of At-RS31 overexpression line using the HR RT-PCR screen of AS patterns in 89 genes with various functions revealed three DNA repair genes significantly affected in alternative splicing of their transcripts. The most prominent change was observed in the gene *UVH1/ATRAD1*. The gene is alternatively spliced in the fifth intron, and the usage of alternative 3' splice site introduces a PTC upstream of a splice junction (Vonarx et al., 2002). AS isoform was shown to be up-regulated upon cycloheximide treatment, which suggests the AS variant to be NMD sensitive and the RNA is presumably degraded via NMD pathway (Kalyna et al., 2012).

Plants depleted in At-RAD1/UVH1 are hypersensitive to DNA damaging agents, such as ultraviolet (UV-B and UV-C), ionizing radiation, and mitomycin C (Gallego et al., 2000; Harlow et al., 1994; Liu et al., 2000). Molecular characterization *uvh1-1* mutant showed that the point mutation affects the fifth intron and creates an alternative 3' splice site which is favored by the splicing machinery leading to higher production of an aberrant transcript (Liu et al., 2000). Interestingly, the latter is almost identical to the wild type *At-RAD1/UVH1* alternatively spliced transcript, which is upregulated by overexpression of At-RS31.

Elevated levels of At-RS31 shifted the AS ratio towards the alternative transcript isoform, therefore both, the overexpression and *uvh1-1* lines produce increased levels of almost identical mRNAs.

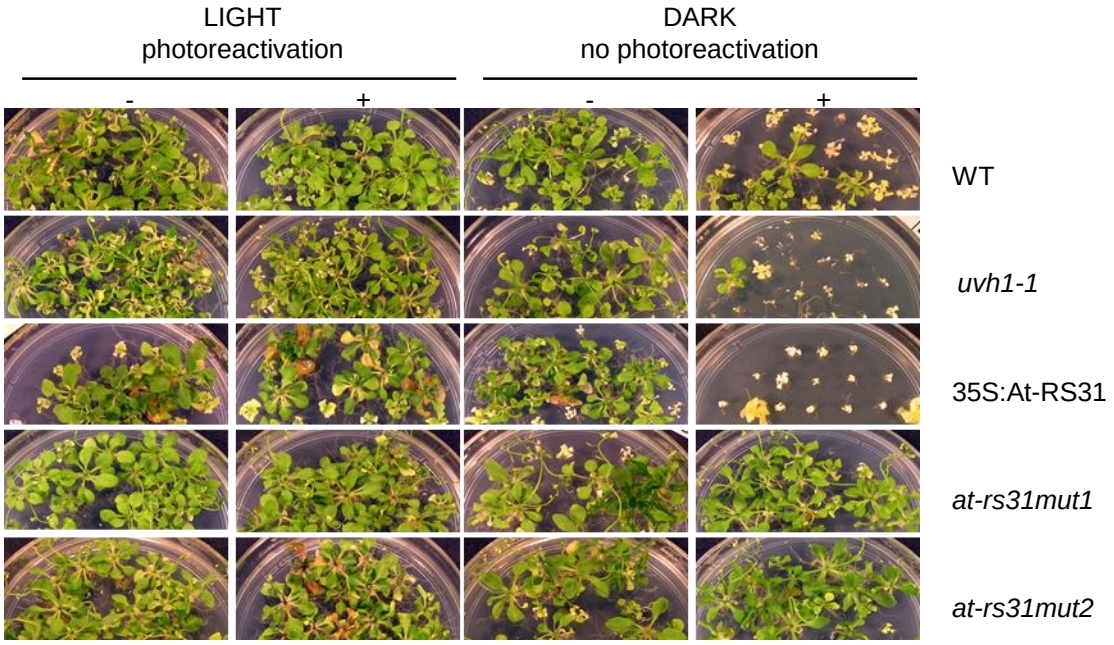


Figure 3.12. At-RS31 affects the UV-C sensitivity of the Arabidopsis seedlings.
Young seedlings of the At-RS31 overexpression line, two distinct *At-RS31* knock-out lines, *uvh1-1* mutant and wild type control were irradiated by UV-C (2000 J/m^2). "Light"-labelled plates were afterwards incubated under standard growth conditions (23°C, 16h day/8h night), "dark"-labelled plates were kept in the dark for two days, transferred to normal growth conditions for 5 days and photographed. At-RS31 overexpression and the *uvh1-1* mutant lines both show hypersensitivity to the UV-C, unlike *at-rs31* mutants that are more resistant to UV-C irradiation.

To investigate the links between the At-RS31 and the genes important in the DNA repair pathway we applied DNA damaging UV-C irradiation to young seedlings of the At-RS31 overexpressing and mutant lines, as well as to *uvh1-1* mutant. UV-C treatment was followed by incubation either in the light, to promote photoreactivation repair of DNA photoproducts, or by growth in the dark to test the performance of dark repair mechanisms.

Examination of UV sensitivity and phenotype showed that photoreactivation is not affected in any of the lines. However, in the experiment setup 35S:At-RS31 overexpression line and *uvh1-1* mutant both show hypersensitivity to the UV-C, whereas two independent knock-out mutants of *At-RS31* compensate the defect and are more resistant to UV-C irradiation than WT (Figure 3.12). This confirms the connection of *At-RS31* to the *UVH1/ATRAD1* pathway.

3.4. Discussion

To elucidate the function of SR protein At-RS31 in plant development and stress response, we used the transgenic overexpressing At-RS31 as well as a knock-out line. We used HR RT-PCR screen to detect genes that responded to overexpression of At-RS31 by change in alternative splicing pattern, and we analysed phenotype of the overexpression and knock-out lines upon application of various abiotic stresses.

At the standard growth conditions, younger plants, overexpressing or lacking the At-RS31, showed phenotype comparable to wild type plants. Altered phenotypic responses were induced by application of abiotic stresses or were detected in ageing plants. Notably, under some conditions (ABA, oxidative stress, DNA damage) both genetic backgrounds developed changes in the phenotypes, however other factors (high salinity, glucose) and prolonged growth affected only the At-RS31 overexpression line, while mutant exhibited no change or only mild phenotype. These results indicate that At-RS31 overexpression and knock-out lines have conditional phenotypes. At-RS31 overexpression has broader phenotype probably due to stable permanent overexpression of the protein. On contrary, mutation in At-RS31 probably reveals changes only at the specific time and conditions when the gene acts. This can be considered for further phenotype analyses that could include testing plants at different periods of development and growth conditions, and more exact methods, such as quantification of the leaf chlorophyll or germination rates, and treatments with increased intensity, or combination of stress-factors.

In addition, conditional mutant phenotype could be explained by potential functional redundancy of the four paralogues in the Arabidopsis RS subfamily. Their functions might have partially diverged and could act differently or overlap in dependence of the conditions and tissues. Indeed, at least three out of four paralogues have both overlapping and tissue-specific patterns of expression (Kalyna and Barta, 2004).

Using HR RT-PCR screen we have found ten genes with changed alternative splicing, induced by overexpression of At-RS31. Two splicing and three transcription factors are among them, suggesting that the effect of At-RS31 overexpression might be caused by multiple factors in combination, shaping the transcriptome in dependence of changing conditions and affecting the resulting phenotype (Figure 1.6) (Simpson et al., 2008b; Syed et al., 2012). Direct binding of At-RS31 to pre-mRNAs of the identified genes could be tested by electrophoretic mobility shift assay. Additional RNA targets of

At-RS31 could be identified by iCLIP (individual-nucleotide-resolution UV cross-linking and immunoprecipitation) (König et al., 2010) or RIP (RNA immunoprecipitation) sequencing (Zhao et al., 2010).

To gain detailed insight into the role of At-RS31 in the complex regulation of plant development and stress responses, future work might employ the extended version of HR RT-PCR screen, with the possibility to design specific primers to the genes of affected pathways and processes and analyses of plants treated by abiotic stimuli. High-throughput transcriptome sequencing (RNA-seq) of treated At-RS31 overexpression and mutant plants under various conditions would provide detailed view on extent and quantitative measure of alternative splicing changes exerted by At-RS31 under specific conditions.

HR RT-PCR screen allowed monitoring shorter regions comprising one AS event and not the whole transcript. Recent RNA-seq analyses revealed many more splice variants per gene (Marquez ez al., 2012, Zhang et al., 2017, see Figure 3.4-3.7) that were known when primers for HR RT-PCR were designed. Therefore, additional and larger gene regions should be tested by new primer pairs to draw conclusions on functional consequences. Together with this, high sensitivity of RNA-seq analyses (Marquez et al., 2012; Zhang et al 2017) resulted in detection of many novel transcripts generated via intron retention (e.g. *At-SR30*, *AtFPG/AtMMH*, tRNAHis guanylyltransferase, Figures 3.4, 3.6, 3.7). It has been shown that IR transcripts are usually very low abundant and probably represent partially processed transcripts (Marquez et al., 2012) and thus stay in the nucleus (Göhring et al., 2014; Kim et al., 2009). Therefore, their potential contribution to the peaks in the HR RT-PCR screen is probably low.

Despite the high frequency of intron retention events in Arabidopsis transcriptome (40%), 51% of intron-containing genes are alternatively spliced using other types of AS events and only about 24% of alternative transcripts contain one or more retained introns (Marquez et al., 2012; Syed et al., 2012). Indeed, no intron retention was detected by HR RT-PCR screen among ten significantly changed AS events. On the other hand, a prevalence of affected alternative 3' splice sites (7 of 10 events) suggests specific role of At-RS31 in the recognition of 3' splice site that is corroborated by interaction of At-RS31 with U2AF65, detected in yeast two- hybrid screen (Lopato et al., unpublished).

Six AS events led to incorporation of a PTC and for some of them NMD sensitivity was experimentally proved (Kalyna et al., 2012) suggesting that shift in AS mediated by At-RS31 modulates protein levels of these genes.

In four cases, the selection of alternative spice site occurs in frame and is predicted to generate an alternative protein isoform. Production of two functionally different proteins from *BTR1* was experimentally shown (Fujisaki and Ishikawa, 2008). This suggests contribution of At-RS31 in generating proteome diversity.

Three of ten genes with changed AS pattern in the At-RS31 overexpression background were also affected upon At-SR30 and At-RS2Z33 overexpression in our previous HR RT-PCR study (Simpson et al., 2008a). *UVH1/AtRAD1* has shifted AS ratio in the opposite directions in At-RS31 and At-SR30 overexpression lines. tRNAHis guanylyltransferase showed significant differences in all three transgenic lines. AS changed towards variant with fully spliced intron in At-RS31 and At-RS2Z33 transgenes, whereas AS was promoted in At-SR30 overexpressing plants. AS change of *BTR1* is identical in At-RS31 and At-SR30 overexpressors. This suggests that some effects are specific to distinct SR proteins, while other might be indirect due to cross-regulation of SR proteins and/or other splicing regulators.

At-RS31 link to the ABA signalling is suggested by several results. At-RS31 overexpression and disruption both induced phenotypic response upon exogenous application of ABA. Further, alternatively spliced variant of *BTR1* splicing factor, BTR1L transcript was promoted by At-RS31 overexpression, and the corresponding protein was shown to be phosphorylated by ABA-activated protein kinase SnRK2.6/OST1 (Wang et al., 2013).

Our phenotypical analyses uncovered hypersensitivity of the At-RS31 overexpression line to oxidative stress (caused by methyl viologen) and UVC-induced DNA damage, whereas mutation enhances the resistance to these factors. Since ROS act destructively also on the DNA and may invoke genome instability, we hypothesize that the negative response to methyl viologen in At-RS31 overexpression line is partially caused by perturbed ability to cope with the DNA damage. Indeed, three DNA-repair genes significantly changed their AS profile in At-RS31 overexpression background suggesting link of At-RS31 to DNA-damage response. In addition, premature ageing of the At-RS31 overexpression line can be possibly explained by accumulation of various DNA lesions. Out results can corroborate the rising evidence of the importance of pre-mRNA processing and alternative splicing in DNA damage response (Giono et al., 2016; Shkreta and Chabot, 2015).

Our findings indicate that At-RS31 is mediating response to the environmental cues by influencing the alternative splicing of other genes, however this dynamic regulation and exact function of At-RS31 in the stress response need to be addressed in further research.

Chapter 4

Study of *Physcomitrella patens* serine/arginine-rich (SR) genes reveals extensive and highly conserved regulation of alternative splicing

4.1. Introduction

RS subfamily of SR proteins is specific to plants. In *A. thaliana,* it consists of four paralogous genes, *At-RS31, At-RS31a, At-RS40* and *At-RS41*, which have arisen by gene duplication, and the former and the latter pairs are evolutionary closer to each other (Figure 1.6). Genes are composed of six exons and five introns, encoding proteins with two RRM domains followed by RS region (Barta et al., 2010). We have studied At-RS31 by stable overexpression and also by using two distinct T-DNA insertion mutants that do not show phenotype under the standard growth conditions or display mild phenotypic variations when exposed to abiotic stresses (Chapter 3). In contrast to all other analyzed plant species, *P. patens* genome contains only one member of the RS protein subfamily, Pp-RS27, with conserved gene and protein architecture (Figure 4.3A) (Kalyna et al., 2006), thus making *P. patens* a useful model to study functions of this subfamily. Interestingly, analysis of Physcomitrella EST collections revealed long alternatively spliced intron at the conserved position of the first RRM, but no conserved alternative splice sites as present in Arabidopsis, rice and maize (Kalyna et al., 2006), raising a question about the evolution of alternative splicing regulation in this subfamily.

RS2Z proteins constitute another plant-specific subfamily. In Arabidopsis, RS2Z subfamily includes two paralogous genes, *At-RS2Z32* and *At-RS2Z33*, which consist of seven exons and six introns and code for proteins with one RRM, two Zn knuckles domains, and SR and SP (serine/proline) regions (Figure 1.6) (Barta et al., 2010). In *P. patens* genome, we have detected two members of RS2Z family: *Pp-RS2Z38* (Kalyna et al., 2006) and, later with new EST data, a second gene *Pp-RS2Z37*. Physcomitrella RS2Z subfamily members share exon–intron structures, protein domain features, and alternatively spliced long introns at the conserved positions with Arabidopsis orthologues (Kalyna et al., 2006). In contrast to the

Physcomitrella RS subfamily member, alignment of the *P. patens*, Arabidopsis, rice and maize long intron sequences revealed highly conserved alternative 3' splice site in *Pp-RS2Z38*, though no supporting ESTs were detected (Kalyna et al., 2006), implying existence of a conserved mechanism regulating alternative splicing in the RS2Z subfamily genes.

To complement our bioinformatics study of highly conserved plant-specific SR genes, we aimed to experimentally explore alternative splicing and its regulation in Physcomitrella orthologues of plant-specific SR genes. We identified and characterized alternative splicing events in Physcomitrella RS and RZ2Z orthologues, and investigated their evolutionary conservation. We examined alternative splicing patterns of RS and RS2Z genes under different growth conditions and stress treatments and identified the highly conserved alternative splicing responses to particular stimuli. Further, we disrupted *Pp-RS27* and *Pp-RS2Z37* genes to monitor the effects of depletion of these proteins on alternative splicing and used the knock-out lines in phenotypic studies to elucidate functions of these proteins in Physcomitrella development and stress response. We expect that our findings would contribute to understanding the evolutionary history of the molecular mechanisms by which environmental cues modulate the activity of the splicing factors and thus the splicing pattern of target pre-mRNAs.

4.2. Materials and methods

4.2.1. Accession numbers

Sequences of the Physcomitrella genes studied here can be accessed at http://plants.ensembl.org by the following IDs: *Pp-RS27* – PP1S144_89V6; *Pp-RS2Z37* – PP1S69_23V6; *Pp-RS2Z38* – PP1S65_286V6; *LEA2* – PP1S442_22V6; *Actin* – PP1S381_18V6.

Arabidopsis and rice sequences can be accessed at https://www.ncbi.nlm.nih.gov under the following IDs: *At-RS2Z32* (At3g53500), *At-RS2Z33* (At2g37340), *Os-RS2Z37a* (Os01g06290), *Os-RS2Z37b* (Os03g17710).

4.2.2. Plant material and growth conditions

We used wild type strain of *P. patens* (Hedw.), collected in Gransden Wood, Huntingdonshire, U.K. (Schaefer et al., 1991).

For routine subculture, the plants were maintained as protonemal tissue, vegetatively propagated by homogenization, subcultured every 7 days and grown on agar plates with BCD-AT medium, overlaid with cellophane. For all described experiments and maintenance of the moss cultures, the plates were incubated at 23°C, 16-h light/8-h dark regime, and 120 μmol/m^2/s. For detailed description of materials and methods to culture, maintain and collect Physcomitrella tissues and plants refer to Section 4.2.16 (Maronova and Kalyna, 2016).

4.2.3. Tissue collection

To collect protonemal tissue grown on cellophane overlay, the tissue was scraped off of the cellophane with a sterile spatula, gently pressed against the wall of the Petri dish and subsequently pressed softly between two layers of filter paper to remove excess water.

Tissue from regenerants after protoplast transformation was obtained from colonies, which were briefly dried between two layers of filter paper.

Gametophores (leafy shoots) were collected from 5 week old colonies growing on BCD-AT media without cellophane. Individual gametophores were collected and cut in half with a scalpel. Upper (younger) parts (UG) and lower parts (LG) were collected separately into the microcentrifuge tubes.

In each instance, the tissue was harvested into a microcentrifuge tube in amount of 100 mg and flash frozen in liquid nitrogen. Grinding of the tissue was carried out in liquid nitrogen with teflon pestle attached to a rotor at a speed of 2000 rpm (Maronova and Kalyna, 2016).

4.2.4. Generation of *pp-rs27* and *pp-rs2z37* knock-out lines

To disrupt the Physcomitrella SR genes *Pp-RS27* and *Pp-RS2Z37*, we first constructed plasmids comprising NPTII cassette flanked by gene homologous sequences of about 0.8 kb each (Figure 4.1 A, 4.1 C). The NPTII cassette confers the kanamycin and G418 resistance and consists of the CaMV (cauliflower mosaic virus) 35S RNA promoter, the neomycin phosphotransferase II (*NPTII*) gene and the CaMV 35S terminator and was obtained from

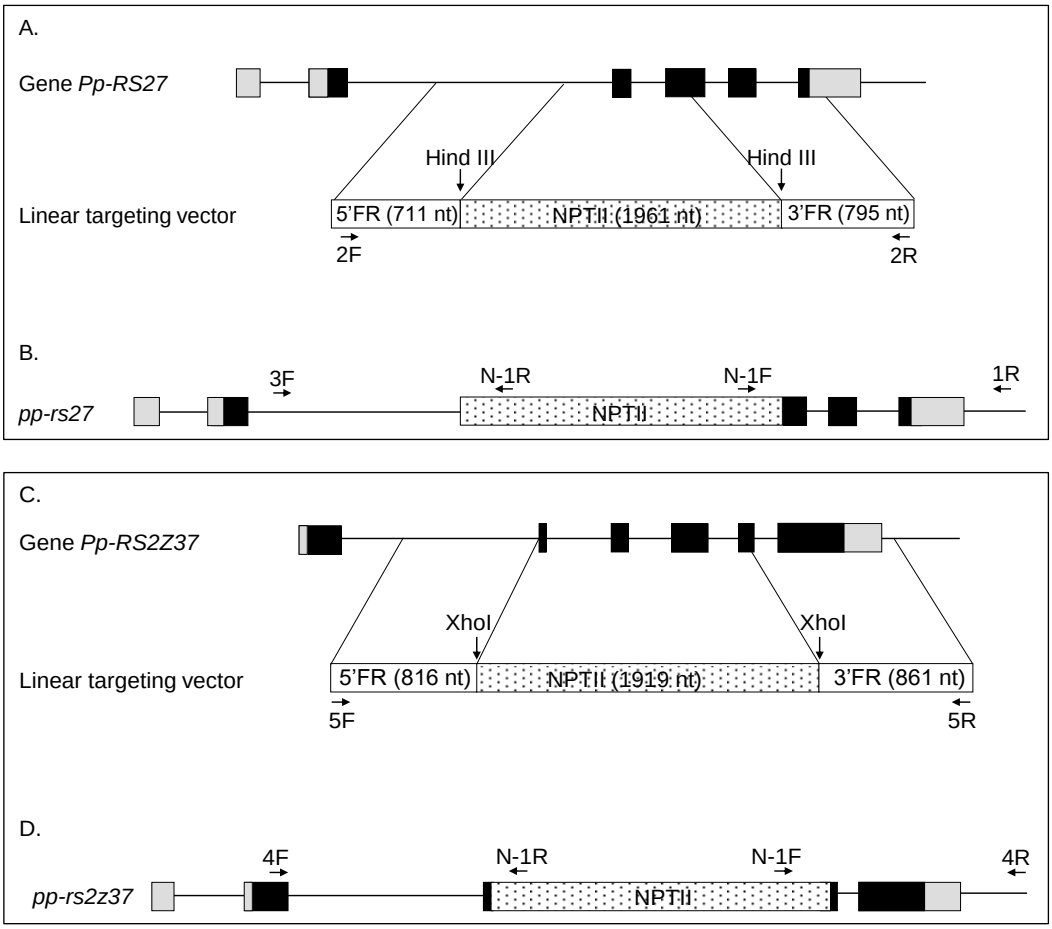


Figure 4.1. Generation of the *pp-rs27* and *pp-rs2z37* knock-out mutants

A. Scheme depicting targeted replacement of the internal part of the gene *Pp-RS27* with NPTII selectable marker by homologous recombination (HR) of the linear targeting vector with the intact gene. Vector consists of NPTII selection cassette, inserted into HindIII sites, bordered by 5' and 3' flanking regions (FR) that are identical to the gene sequences marked by slanted lines. Primers 2F and 2R used to amplify the fragment from the plasmid template (not shown) are depicted by arrows.

B. Altered *Pp-RS27* gene disrupted by NPTII. Primers used to verify the insertion site by PCR are shown by arrows.

C. Scheme illustrates targeted replacement of the internal part of the gene *Pp-RS2Z37* with NPTII selectable marker by HR of the linear targeting vector with the intact gene. NPTII selection cassette is inserted into XhoI sites of the vector and is bordered by 5' and 3' flanking regions (FR) that are identical to the gene sequences marked by slanted lines. Primers 4F and 4R, used to amplify the fragment from the plasmid template (not shown) are depicted by arrows.

D. Altered *Pp-RS2Z37* gene disrupted by NPTII cassette. Primers used to verify the insertion site by PCR are shown by arrows.

Grey and black boxes represent non-coding and coding exons, respectively. Introns are shown as lines. Slanted lines associate regions, shown as white boxes, flanking the NPTII cassette from the 5' (5' FR) and 3' (3' FR). Sizes of particular elements are shown in nucleotides (nt).

the plasmid pMBL6 (GenBank Accession DQ228131) (Knight et al., 2002). Plasmids were used for PCR amplification of DNA targeting vectors, bearing the sequence of the gene, interrupted by NPTII cassette (Figure 4.1 B, 4.1 D), which were introduced into the Physcomitrella protoplasts. Disruption of the target genes occurred by homologous recombination (HR) of the transformed vectors with the corresponding endogenous sequences (Kamisugi et al., 2005; Knight et al., 2002; Schaefer et al., 1991).

4.2.4.1. DNA extraction

Physcomitrella total DNA was extracted from 100 mg of protonemal tissue using DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's protocol "Purification of Total DNA from Plant Tissue (Mini Protocol)".

4.2.4.2. Construction of the *Pp-RS27* and *Pp-RS2Z37* gene targeting vectors

For gene *Pp-RS27*, a fragment of 3862 nt was PCR-amplified from Physcomitrella total DNA with the primers 3F and 1R and, for *Pp-RS2Z37*, primers 4F and 4R were used to amplify genomic fragment of 3444 nt. All primers are shown in the Table 4.1. PCR mixture (50 µL) consisted of 1 × Optimized DyNAzyme EXT Buffer (Finnzymes), 1 mM $MgCl_2$, 0.25 mM of each dATP, dCTP, dGTP, dTTP, 2 µM of each primer, 250 ng DNA and 1 U DyNAzyme EXT DNA Polymerase (Finnzymes). PCR programmes are described in the Table 4.2. PCR products were resolved by ethidium bromide stained 1.2% (w/v) agarose gel-electrophoresis based on TAE (Tris-Acetate-EDTA) buffer.

Purified PCR products were inserted by TA-cloning into the pGEM®-T Easy plasmid vector (pGEM®-T Easy Vector System I, Promega) according to the manufacturer's instructions and introduced into competent *E. coli* cells (strain XL1-Blue) by heat-shock transformation. Plasmid DNA (pDNA) from positive colonies was screened by PCR and by restriction enzyme (RE) analysis.

To introduce NPTII cassette, plasmids were linearized by restriction enzymes (HindIII in *Pp-RS27*- and XhoI for *Pp-RS2Z37*-containing plasmids). 5´-ends were dephosphorylated using Calf Intestinal Alkaline Phosphatase following the manufacturer's instructions (New England Biolabs, Inc.). Restriction enzymes HindIII and XhoI, respectively, were used to excise the selection cassette NPTII from the plasmid pMBL6 (GenBank Accession DQ228131) (Knight et al., 2002). NPTII cassettes were ligated with the dephosphorylated

linear plasmids using T4 DNA ligase according to the manufacturer's instruction (New England Biolabs, Inc.). Ligation products were transformed into the competent *E.coli* cells (strain XL1-Blue). Colonies were screened by RE analysis of the extracted pDNA, and insert of the positive plasmid was confirmed by sequencing.

Table 4.1. Primers used for generation and validation of *pp-rs27* and *pp-rs2z37* knock-out lines

Gene	Name	Sequence from 5' to 3'	Internal Name*
Pp-RS27	1R	ATTAAGGACTGTGACTCGTATCTCG	RSp27-R4110
Pp-RS27	2F	TTTCCGTGGTGTATTCCTTGGTTC	PpRSp27-F1288:
Pp-RS27	2R	GCAGTCCAACTCCTAAATCCCTC	PpRSp27-R3498
Pp-RS27	3F	GGAGGTTGTAGTCGGTTTATTTCTC	ppRSp27-F968
Pp-RS2Z37	4F	ATCTGCGCTAGAAATCCTCCTTGTG	PpRSZ37-F3176
Pp-RS2Z37	4R	ATTGTAGTAATGTTGAGGCCCACTTG	RSZ37-R6619
Pp-RS2Z37	6R	TGTCAACTTTGTGGAATCAATGC	PpRSZ37-R6053
NptII	N-1F	ATTGGTATCAGAGCCATGAATAGGT	nptII-g6termF
NptII	N-1R	AGATAGCTGGGCAATGGAATCCGA	nptII-35Spro2R
NptII	N-2F	ATCTCCTGTCATCTCACCTTGC	nptIIfw1 probe
NptII	N-2R	CTCTTCAGCAATATCACGGGTAG	nptIIrev1 probe

* Internal names as utilized in the laboratory, not mentioned in this thesis

Table 4.2. Primer pairs and PCR programmes used for generation and validation of *pp-rs27* and *pp-rs2z37* knock-out lines

Gene	Primer pair	Thermocycler programme					
		ID*	D*	A*	E*	FE*	C
Pp-RS27	2F + 2R	95/2m	95/30s	60/1m	72/4m	72/10m	35
pp-rs27	3F + N-1R	95/1m	95/30s	60/1m	72/90s	72/5m	35
pp-rs27	N-1F + 1R	95/1m	95/30s	60/1m	72/90s	72/5m	35
Pp-RS2Z37	4F + 4R	95/2m	95/30s	65/1m	72/4m	72/10m	35
pp-rs2z37	5F + N-1R	95/1m	95/30s	60/1m	72/90s	72/5m	35
pp-rs2z37	N-1F + 6R	95/1m	95/30s	60/1m	72/90s	72/5m	35
NPTII	N-2F + N-2R	95/1m	95/30s	60/1m	72/90s	72/5m	35

* Temperature [°C]/Time
ID - initial denaturation, D - denaturation, A - annealing, E - extension, FE - final extension, C - number of cycles.

Targeting vectors for *P. patens* protoplast transformation were amplified from plasmid constructs by PCR with the primers 2F and 2R (for *PpRS27*) and 4F and 4R (for *PpRS2Z37*) (Figure 4.1 B, 4.1 D, Table 4.1). PCR mixture of 50 µL was composed as previous PCR reaction, except that 25 ng of pDNA was used as a template. PCR programme is described in Table 4.2.

PCR products were precipitated by adding 1/10 of the reaction volume 3 M sodium acetate (pH 5.2), 2.5 of the reaction volume of ethanol, mixing and incubating at -20°C overnight, followed by centrifugation at 14,000 rpm for 15 min. Pellet was washed with cold 80% ethanol, dried on the heating block at 37°C and dissolved in max. 30 µL dH_2O to the concentration of 2 µg/µL.

4.2.4.3. *P. patens* protoplast transformation

P. patens protoplast transformation, regeneration and selection of stable genomic transformants are described to the detail in section 4.2.16 (Maronova and Kalyna, 2016).

4.2.5. Identification of the stable *P. patens* transformants using PCR screening

DNA (for DNA extraction see Section 4.2.4.1) of stable regenerants obtained from spot inocula (Maronova and Kalyna, 2016) was analysed by PCR to detect the disruption of the gene by NPTII cassette. Combinations of *NPTII*-specific with the gene-specific primers, designed to the region adjacent, but not present in the linear targeting vector, were used to investigate the integration of NPTII cassette into the endogenous sequence of the corresponding gene (Figure 4.1; Table 4.1).

PCR mixture (50 µL) consisted of 1 × Optimized DyNAzyme Buffer (Finnzymes), 1 mM $MgCl_2$, 0.25 mM of each dATP, dCTP, dGTP, dTTP, 1 µM of each primer, 10 ng DNA and 0.5 U DyNAzyme DNA Polymerase (Finnzymes). PCR programme, primer combinations and sequences are described in Tables 4.1 and 4.2. Insertion sites of the identified mutants were sequenced.

4.2.6. Southern blot analysis of stable regenerants

DNA in amount of 1 µg extracted from independent *Pp-RS27* regenerated colonies was digested with 20 U of restriction enzyme AflII. For *Pp-RS2Z37*, DNA samples were

digested simultaneously with the restriction enzymes NdeI (20 U) and XbaI (20 U). Reactions of 60 µL were incubated for 8 hours at 37°C, according to the manufacturer's instructions (New England Bioloabs). Digested DNA samples were precipitated as described in the section 4.2.4.2 and dissolved in 18 µL dH_2O. DNA samples, stained with ethidium bromide, were separated on 0.6% agarose gel in TAE buffer for about 6 hours. Separated DNA was transferred from the gel to the nylon membrane Hybond-N+ (Amersham Biosciences) and cross-linked according to Draper (1988).

DNA probe (392 nt), specific to the sequence of NPTII cassette, was obtained by PCR from the plasmid pMBL6 with the primers N-2F and N-2R (Table 4.1 and 4.2). PCR mixture in 50 µL volume contained 1 × Optimized DyNAzyme Buffer (Finnzymes), 1 mM $MgCl_2$, 0.25 mM of each dATP, dCTP, dGTP, dTTP, 1 µM of each primer, 10 ng pDNA and 0.5 U DyNAzyme DNA Polymerase (Finnzymes). PCR programme is described in Table 4.2. Purified probe was fluorescein-labelled using the Gene Images ™ Random-Prime Labelling Module (Amersham Biosciences) according to the instructions, and incubated with the membrane overnight. Detection procedure was performed using the Gene Images ™ CDP-Star ™ Detection Module (Amersham Biosciences) following the manufacturer's protocol.

4.2.7. RNA extraction

Total RNA was extracted from 100 mg of frozen ground tissue using RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's protocol "Isolation of Total RNA from Plant Cells and Tissues and Filamentous Fungi". Buffer RLT was used as lysis buffer. Optional DNA removal by DNase I treatment using the RNase-Free DNase Set (Qiagen) was carried out according to the manufacturer's instructions.

4.2.8. Reverse-transcription (RT)-PCR

RT-PCR was used for detection of gene expression and changes in AS pattern in various moss tissues, in wild type and mutant genetic background and under various treatment conditions.

To synthesize the first strand cDNA, total RNA was reversely transcribed using oligo$(dT)_{15}$ primers and the Reverse Transcription System (Promega) according to the manufacturer's protocol, with following modifications. We used 2 µg of RNA per reaction

(due to high amount of chloroplast RNA in the total RNA from moss), and the incubation time at 42°C was extended to 2 h.

RT-PCRs of SR and *Actin* (PP1S381_18V6) transcripts were performed in Piko Thermal Cycler 24-well system (Finnzymes) in a reaction volume of 20 µL with the reaction composition as follows: 5 × Phusion HF Buffer, 0.2 mM of each dATP, dCTP, dGTP, dTTP, 1 µM of each primer, 2 µL of cDNA, 0.4 U of Phusion High-Fidelity DNA Polymerase (Finnzymes).

Table 4.3. Primers used for PCRs to detect gene expression and changes in AS pattern

Gene	Exon	Sequence from 5' to 3'	Internal name*
Pp-RS27	E1-F	AGGAGAATAGAGGAATCCCTTGC	PpRSp27-F1
Pp-RS27	E2-F	CCGGTGTATTGTGGGAACTTTG	ppRSp27-F574
Pp-RS27	E2-R	TCTTCATGTCAACACGCTCCAC	ppRSp27-R670
Pp-RS27	E3-F	CTACATGGAAGATGAGCGTGATG	ppRSp27-F2300
Pp-RS27	E3-R	CATCACGCTCATCTTCCATGTAG	PpRSp27-R2322
Pp-RS27	E4-F	AGACCTGGAACGACACTTCGAG	ppRSp27-F2672
Pp-RS27	E4-R	AGTGTCGTTCCAGGTCTCGTACTC	PpRSp27-R2688
Pp-RS27	E5-F	ATCTAGTGGTCGGGGTGGTAGT	ppRSp27-F2997
Pp-RS27	E5-R	ACTACCACCCCGACCACTAGAT	PpRSp27-R3018
Pp-RS27	E6-R	AAACTGCCTGCAGAACATTGC	PpRSp27-R3741
Pp-RS2Z37	E1-F	CTCTGTTCATTCCAGTGAGTGTCTC	ppRSZ37-F1243
Pp-RS2Z37	E2-F	ACAAGAAGCCGAGATTTGGAAGAC	ppRSZ37-F2726
Pp-RS2Z37	E2-R	GTCTTCCAAATCTCGGCTTCTTG	ppRSZ37-F2749
Pp-RS2Z37	E3-R	CAAAGTCGTGCTTCACATCCAC	ppRSZ-Rex3
Pp-RS2Z37	E7-R	AAGTATTTTCATTCACAGGTCATGG	ppRSZ37-R5784
Pp-RS2Z38	E1-F	GCCAAGACCGTTGCGTAG	ppRSZ38-F194,
Pp-RS2Z38	E7-R	AGGACTCCGATCACCCTTGACTG	ppRSZ38-R4373
Actin	ActF	GTCGAGCATGAAGATCAAAGTGG	ppF-Act
Actin	ActR	ACCAGCCGTTAGAATTGAGCCCAG	ppR-Act
LEA-2	LeaF	AGAAGTACCAGGAGCATCGAGAC	ppLEA2-F
LEA-2	LeaR	GGATTCCAAGCTCGTAGGTTC	ppLEA2-R

* Internal names as utilized in the laboratory, not mentioned in this thesis

Table 4.4. Primer pairs and PCR programmes used to detect changes in gene expression and AS pattern

Gene (AGI)	Primer pair*	PCR programme					
		ID**	D**	A**	E**	FE**	C
Pp-RS27	E1-F + E2-R	98/3m	98/10s.	63/10s	72/30s	72/10m	30
Pp-RS27	E2-F + E3-R	98/3m	98/10s	63/10s	72/30s	72/10m	30
Pp-RS27	E3-F + E4-R	98/3m	98/10s	63/10s	72/30s	72/10m	30
Pp-RS27	E4-F + E5-R	98/3m	98/10s	63/10s	72/30s	72/10m	30
Pp-RS27	E5-F + E6-R	98/3m	98/10s	63/10s	72/30s	72/10m	30
Pp-RS27	E1-F + E6-R	98/3m	98/10s	63/10s	72/75s	72/10m	30
Pp-RS2Z37	E1-F + E2-R	98/3m	98/10s	63/10s	72/30s	72/10m	30
Pp-RS2Z37	E1-F + E7-R	98/3m	98/10s	63/10s	72/75s	72/10m	30
Pp-RS2Z37	E2-F + E3-R	98/3m	98/10s	63/10s	72/30s	72/10m	30
Pp-RS2Z37	E2-F + E7-R	98/3m	98/10s	63/10s	72/75s	72/10m	30
Pp-RS2Z38	E1-F + E7-R	98/3m	98/10s	63/10s	72/75s	72/10m	30
Actin	ActF + ActR	98/3m	98/10s	63/10s	72/15s	72/3m	20
Lea-2	LeaF + LeaF	95/2m	95/30s	50/45s	72/30s	72/5m	30

* Primers positions are shown by arrows in the gene schemes drawn above each image of the gel (Section 4.3.). E*n* specifies the exon number, F and R means forward and reverse direction, respectively.
** Temperature [°C]/Time
ID - initial denaturation, D - denaturation, A - annealing, E - extension, FE - final extension, C - number of cycles.

Amplification of *LEA-2* (PP1S442_22V6) was conducted on Biometra thermal cycler in a 25 µl reaction volume that contained 10 × Optimized DyNAzyme Buffer, 1 mM $MgCl_2$, 0.25 mM of each dATP, dCTP, dGTP, dTTP, 1.2 µM of each primer, 3 µl of cDNA and 0.3 µl of DyNAzyme DNA Polymerase (Finnzymes). Primer pairs and PCR programmes are specified in the Tables 4.3 and 4.4. PCR products were resolved by ethidium bromide stained 1.2% (w/v) agarose gel-electrophoresis based on TAE buffer.

4.2.9. Cloning and sequencing of alternatively spliced isoforms

PCR products representing putative AS isoforms were extracted and purified from agarose gel using NucleoSpin® Extract II kit (Macherey-Nagel) according to the manufacturer's protocol. Taq DNA polymerase was used to add adenine (A)-overhangs at

the 3´-ends of extracted DNA. Reaction (10 µL) consisted of 200-300 ng of DNA, 10 × reaction buffer (Promega), 0.2 mM dATP and 1 U GoTaq DNA Polymerase (Promega) and was incubated at 72°C for 10 min.

Processed DNA fragments were ligated into pGEM T-Easy Vector system (Promega). Ligation was performed as described by the manufacturer's protocol, and the ligation products were transformed into the competent *E.coli* cells (strain GT116). Plasmid DNA (appx. 20-50 ng) extracted from the colonies was screened by PCR using the same primers and PCR reaction mixture, as utilized for the RT-PCR, and the positive clones were sequenced.

Sequence analyses and AS transcript models were done by alignments of the sequences of the respective Physcomitrella SR genes with the sequences of transcripts or their fragments using GeneSeqer splice alignment tool (http://www.plantgdb.org/) (Usuka et al., 2000).

4.2.10. Nucleotide sequence alignment

Alignments were done by ClustalW tool at http://www.genome.jp/tools-bin/clustalw. Shading of alignments was done with BoxShade tool at http://embnet.vital-it.ch/software/BOX_form.html.

4.2.11. Scoring of the splice sites based on position weight matrices (PWM)

To determine the splice site strength, we generated position weight matrices (PWMs) based on the collection of splice sites of 160,800 introns obtained from the *P. patens* genome sequence annotation, Version 3.3 (http://www.cosmoss.org/fgb2/gbrowse/V3.3/).

PWMs for 5' splice sites were calculated from the nucleotide sequences of the last 3 nt of the upstream exon and the first 10 nt of the intron. For PWMs of 3' splice sites, the last 14 nt of the intron and the first 3 nt of the downstream exon were used.

4.2.12. Phenotype analyses of Physcomitrella mutant lines

For phenotypic analyses, control wild type, two independent *pp-rs27* mutant lines (#36 and #38) and two independent *pp-rs2z37* mutant lines (#36 and #47) were inoculated

together onto an agar plate without cellophane. Each *P. patens* line was represented by two colonies per plate (see Figure 4.16).

Auxin (2,4-D, 2,4-Dichlorophenoxyacetic acid), cytokinin (6-BAP, 6-Benzylaminopurine), abscisic acid (ABA) and methyl viologen (MV) were dissolved in 10% DMSO (Sigma-Aldrich) as 10,000 × stock solutions, filter sterilized and added to the autoclaved BCD-AT agar media, cooled to appx. 60ºC. Control BCD-AT medium contained 0.001% DMSO. Glucose, mannitol and NaCl were added to the BCD-AT agar media before autoclaving. Final concentrations of supplements are listed in the Table 4.5. Agar media were poured into sterile Petri dishes (9 cm diameter) in amount of 20 mL per plate under the laminar flow hood by using a sterile pipette. Plates were not overlaid with cellophane.

Preceding inoculation, protonemal tissue was cultivated on cellophane-covered BCD-AT solid medium for 5 days at 23°C, 16-h light/8-h dark regime. Tissue was scraped from the cellophane using sterile spatula, gently pressed against the wall of the Petri dish and then softly pressed between two pieces of sterile filter paper to remove excess water. The tissue was weighed under sterile conditions, and 150 mg were blended and resuspended in 6 mL of 0.1% agar in water (sterilized by autoclaving). Inoculation onto supplemented and control plates was carried out by dropping 5 µL of the tissue suspension on the plate according to

Table 4.5. List of the BCD-AT supplements used for stress treatments

Supplement	Cat. Nr., Producer	Molar mass [g/mol]	Final concentrations
2,4-Dichlorophenoxyacetic acid (2,4-D, Auxin)	D70724, Sigma-Aldrich	221.04	0.1, 1, 10 µM*
6-Benzylaminopurine (6-BAP, Cytokinin)	B3408, Sigma-Aldrich	225.25	0.1, 1, 10 µM *
Abscisic acid (ABA)	A1049, Sigma-Aldrich	264.32	0.1, 1, 50, 100 µM*
Methyl viologen dichloride hydrate (MV)	856177, Sigma-Aldrich	257.16	0.1, 1, 10 µM*
Glucose	Sigma-Aldrich	180.16	1, 100, 300 mM**
Mannitol	Sigma-Aldrich	182.17	1, 100, 300 mM**
NaCl	Sigma-Aldrich	58.44	1, 100, 150 mM**

* in 10% DMSO (v/v)
** in dH_2O

the scheme illustrated on Figure 4.16. Tissue regenerated into filaments and formed colonies. Plates were incubated in the growth chamber at 23°C, 16-h light/8-h dark regime for 5 weeks from the inoculation and then photographed.

4.2.13. Stress treatments for analyses of splicing patterns

Drought, ABA and MV were applied on 6 days old protonemal tissue of wild type, grown at 23°C, 16-h light/8-h dark regime on BCD-AT plates covered with cellophane discs.

To induce drought stress, empty Petri dishes (9 cm diameter) were lined with a single filter paper soaked with 1 mL of BCD-AT medium. Cellophane discs with the tissue were transferred onto the filter paper and kept open under the laminar flow hood for 40 min. In the control, Petri dishes were closed with the lids during the incubation time.

For stress treatments with ABA or MV, three sterile filter papers were placed into empty Petri dishes and soaked with 6 ml of BCD-AT medium containing 10 µM ABA or 10 µM MV in 0.005% DMSO. For the control, the medium contained 0.005% DMSO. Cellophane discs with protonemal tissue were transferred into prepared Petri dishes with soaked filter papers and incubated for 2 or 6 h in the tissue culture room at 23°C, 16-h light/8-h dark regime.

To test the effect of light, plates were incubated for 3 days at standard conditions and collected. For dark, the plates were wrapped into aluminium foil for two days, cultivated together with "light" samples, returned back to standard conditions for one day and collected.

Irradiation by UV-C (254 nm) was performed using UV crosslinker (Stratagene) in the fluence of 0.5, 1, 2, 3 kJ/m^2, with lids of the Petri dishes removed. Untreated tissue was used for the control. Irradiation was followed by 4 hours incubation in the dark (wrapped in aluminium foil) or in the light, at the standard conditions.

Tissue was collected by gentle scraping from the cellophane with the spatula and softly pressing it to the wall of the Petri dish. The tissue was briefly dried on filter paper, flash frozen in liquid nitrogen and stored at -80°C.

4.2.14. Cycloheximide treatment

Six days old *P. patens* wild type protonemal tissue was grown on BCD-AT plate with cellophane at 23°C, 16-h light/8-h dark. To analyse *Pp-RS27*, the tissue was transferred from the plate to a 70 µm cell strainer (Falcon) (tissue from one plate per one strainer). Strainers

were placed into 6-well plate (Falcon) containing 7.5 mL of 20 µM cycloheximide (Sigma-Aldrich) or 7.5 mL of water for control. Tissue samples were infiltrated by 10 min. vacuum application. After 4 hours the tissue was collected, dried with filter paper and flash frozen. Experiment was conducted in two biological replicates.

To analyse *Pp-RS2Z37* and *Pp-RS2Z38*, the tissue on the plates was covered with 20 µL of 20 µM cycloheximide or with the 20 µL of water for control. No vacuum infiltration was performed. The tissue was collected after 4 hours, briefly dried with the filter paper and flash frozen. Experiment was done in two biological replicates.

4.2.15. UV-C irradiation treatments for phenotypic analysis

Six days old Physcomitrella protonemal tissue of wild type, *pp-rs27#36* and *pp-rs2z37#47* mutant lines, grown on BCD-AT plates with cellophane discs at 23°C, 16-h light/8-h dark (standard conditions), was used.

Irradiation by UV-C (254 nm) was conducted using UV crosslinker (Stratagene) in fluence of 1 and 3 kJ/m^2, with lids of the Petri dishes removed. Untreated tissue was used for the control. Half of the plates was afterwards cultivated for three days in the standard conditions (light) and the other half was cultivated wrapped in aluminium foil for two days and one day unwrapped (dark), at the standard conditions.

4.2.16. Publication

Generating Targeted Gene Knockout Lines in *Physcomitrella patens* to Study Evolution of Stress-Responsive Mechanisms

Monika Maronova and Maria Kalyna

In: P. Duque (ed.), Environmental Responses in Plants: Methods and Protocols. Methods in Molecular Biology 1398: 221–234.

Author's contribution: Establishment and adaptation of the method, writing the manuscript.

Chapter 18

Generating Targeted Gene Knockout Lines in *Physcomitrella patens* to Study Evolution of Stress-Responsive Mechanisms

Monika Maronova and Maria Kalyna

Abstract

The moss *Physcomitrella patens* possesses highly efficient homologous recombination allowing targeted gene manipulations and displays many features of the early land plants including high tolerance to abiotic stresses. It is therefore an invaluable model organism for studies of gene functions and comparative studies of evolution of stress responses in plants. Here, we describe a method for generating targeted gene knockout lines in *P. patens* using a polyethylene glycol-mediated transformation of protoplasts including basic in vitro growth, propagation, and maintenance techniques.

Key words Homologous recombination, Gene targeting, Knockout, *Physcomitrella*, Plant, Stress, Evolution, Protoplast

1 Introduction

Living organisms have to cope with various stresses, both of biotic and abiotic nature. Studies of stress response not only are essential for a better understanding of life and its diversity but also have significant medical and economic implications. Different species respond and adapt very differently from one another. Fight-or-flight strategy used by animals as the first stage of the response to a harmful event is not applicable for plants that are sessile organisms. Plants have their own panoply of stress responses that was shaped by migration of plant ancestors from water habitats to land. At the terrestrial surroundings, they had to develop strategies to respond to much more substantial fluctuations in temperature, light intensity, and quality, including a stronger exposure to the DNA-damaging ultraviolet compound of the light, and now to water availability ranging from drought to flooding. Bryophytes are considered to be the closest to the common ancestor of plants. Among them, the moss *Physcomitrella patens* provides us with powerful

Paula Duque (ed.), *Environmental Responses in Plants: Methods and Protocols*, Methods in Molecular Biology, vol. 1398,
DOI 10.1007/978-1-4939-3356-3_18,

tools to study functions of genes and together with other model plants allows comparative analyses of evolution of molecular mechanisms of stress response and adaptation [1]. *Physcomitrella* displays many features characteristic to the early land plants including a high tolerance to drought, salinity, osmotic, and other abiotic stresses [2–5].

Physcomitrella is haploid through almost all life cycle. The diploid stage, sporophyte, is short and results in haploid spores. Spores germinate and produce filamentous structures, protonemata. The protonemal filaments, the juvenile stage of gametophyte, can initiate buds that further develop into gametophores, the adult stage of gametophyte. The gametophores have more complex organization, with leaflike structures, rhizoids, female and male sex organs producing gametes, and subsequently diploid zygote that develops into sporophyte.

Physcomitrella is a remarkable model organism. Minimalist anatomy and simple developmental stages of *Physcomitrella* can be easily monitored. This is further facilitated by efficient methods of in vitro growth and propagation [6, 7] that also provide ample source of material for biochemical and molecular analyses. Moreover, *Physcomitrella* is highly amenable to genetic manipulations [8]. Dominant haploid generation of *Physcomitrella* allows fast and simpler recognition of mutant phenotypes [9, 10]. Development of approaches for genetic transformation of *Physcomitrella* yielded not only routine and efficient procedures, such as polyethylene glycol (PEG)-mediated DNA delivery into protoplasts [11], but also led to the realization that it undergoes high frequency of homologous recombination, comparable to that in yeast [12]. Gene targeting through homologous recombination is one of the most efficient means to study gene functions. This technology is far less efficient when applied in plants that have low rates of homologous recombination. Gene targeting is not limited to generation of knockout mutants by gene disruption. It allows also knock-in manipulations, such as introduction of various reporters and tags for visualization of expression of endogenous genes and isolation of molecular complexes. Moreover, endogenous genes or their portions can be replaced by their modified versions, thus providing means for site-directed mutagenesis. Importantly, the *Physcomitrella* genome is sequenced and fully assembled, therefore enabling reverse genetic approaches [1].

Here, we provide a protocol for generating targeted gene knockout lines in *P. patens* using a PEG-mediated transformation of protoplasts including basic techniques for in vitro growth, propagation, and maintenance of *P. patens*. This protocol is routinely used in our lab and is based on the protocols developed earlier [6, 11, 13]. Rich information resources providing further *Physcomitrella* protocols and techniques are currently available online (http://moss.nibb.ac.jp/ and http://biology4.wustl.edu/moss/methods.html).

2 Materials

2.1 Materials and Equipment

1. *Physcomitrella patens*, the "Gransden" strain, freshly grown protonemal tissue.
2. Gene targeting linear DNA vector (Fig. 1), 10–15 µg of purified DNA dissolved in a maximum of 30 µL dH_2O.
3. Laminar airflow hood.
4. Blender suitable for sterilization (*see* **Note 1**).
5. Water bath 45 and 25 °C (room temperature).
6. Hemocytometer.
7. Lighttight box (*see* **Note 2**).
8. Petri dishes, 9 cm in diameter with one vent.
9. Glass Petri dish.
10. Filter paper discs, 5–8 cm in diameter.
11. Sterile cellophane discs, 8 cm in diameter (*see* **Note 3**). To sterilize the cellophane discs separate each other by a disc of filter paper. Assemble ca. 30 cellophane discs in this sandwich way into a glass Petri dish, wrap into an aluminum foil, and autoclave.
12. Sterile wide-bored (or cut tip) blue micropipette tips.
13. Sterile 70 µm nylon mesh attached to sterile 50 mL centrifuge tube.

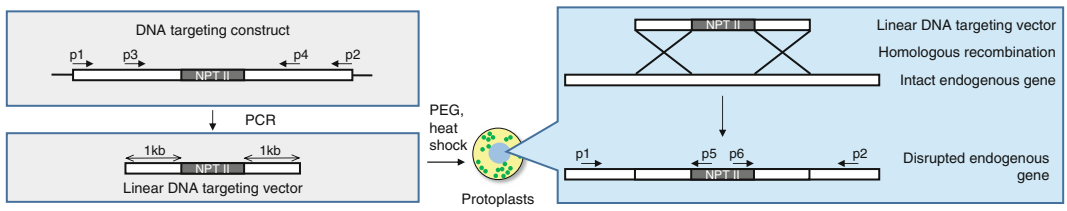


Fig. 1 Strategy of targeted gene disruption by homologous recombination in *Physcomitrella patens* protoplasts. To create the DNA targeting construct, a sufficiently long region of the target gene is amplified by PCR using primers p1 and p2 (outer primers) and integrated into the plasmid vector. The antibiotic resistance selection cassette (in this case neomycin phosphotransferase gene, NPT II) is then ligated into convenient restriction sites within the cloned target gene to replace the coding sequence (*upper left scheme*). The DNA targeting construct is used as a template in a subsequent PCR with primers p3 and p4 (inner primers) to obtain a linear targeting vector. The selection cassette is flanked by about 1 kb regions of homology to the target gene (*lower left scheme*). The linear DNA targeting vector is used for PEG-mediated transformation of *Physcomitrella* protoplasts, where it homologously recombines with the genomic DNA (*upper right scheme*). As a result, the endogenous gene is disrupted by the selection cassette. Stable transformants after the selection process are verified by two PCRs using outer primers in combination with primers to the selection cassette (*lower right scheme*; primer combinations: p1 and p5, p6 and p2). *White rectangles* represent the endogenous gene or the homology regions of the construct and the vector; *gray rectangles* represent the selection cassette

14. Centrifuge 10–15 mL tubes with round bottom and a cap.
15. Surgical tape.

2.2 Media and Solutions

1. Stock solution B (100×): 101 mM $MgSO_4$. Add 25 g of $MgSO_4 \cdot 7H_2O$ (or 12 g of anhydrous $MgSO_4$) to 500 mL of water in a beaker. Dissolve using a magnetic stirrer, transfer to the 1 L graded cylinder, and make up to 1 L with water. Autoclave or filter sterilize. Store at 4 °C.
2. Stock solution C (100×): 184 mM KH_2PO_4, pH 6.5. Weigh 25 g of KH_2PO_4 and dissolve in 500 mL of water as in previous step. Adjust pH to 6.5 with 4 M KOH. Make up to 1 L. Autoclave or filter sterilize. Store at 4 °C.
3. Stock solution D (100×): 1 M KNO_3, 4.5 mM $FeSO_4$. Weigh 101 g of KNO_3 and dissolve in 200 mL of water. Weigh 1.25 g of $FeSO_4 \cdot 7H_2O$ and dissolve separately in 200 mL of water. Adding of H_2SO_4 may help dissolve $FeSO_4 \cdot 7H_2O$. Mix both solutions and make up to 1 L. Autoclave. Store at 4 °C for 2–3 month and discard when iron precipitate forms.
4. Trace element solution (TES, 1000×): 614 mg H_3BO_3, 110 mg $AlK(SO_4)_2 \cdot 12H_2O$, 55 mg $CuSO_4 \cdot 5H_2O$, 28 mg KBr, 28 mg LiCl, 25 mg $Na_2MoO_4 \cdot 2H_2O$, 389 mg $MnCl_2 \cdot 4H_2O$, 55 mg $CoCl_2 \cdot 6H_2O$, 55 mg $ZnSO_4 \cdot 7H_2O$, 28 mg of KI, 28 mg $SnCl_2 \cdot 2H_2O$, 59 mg $NiCl_2 \cdot 6H_2O$. Weigh and separately dissolve all the components in ca. 50 mL of water. Combine the solutions and make up to 1 L. Store at 4 °C.
5. Ammonium tartrate (AT) stock solution (100×): 0.5 M Diammonium (+) tartrate. Weigh 9.21 g of diammonium (+) tartrate and dissolve in 80 mL of water as in previous steps. Make up to 100 mL. Autoclave or filter sterilize. Store at 4 °C.
6. $CaCl_2$ stock solution: 1 M $CaCl_2$. Weigh 11.1 g of $CaCl_2$, dissolve in 80 mL of water, and make up to 100 mL as in previous steps. Autoclave or filter sterilize. Store at room temperature (*see* **Note 4**).
7. Plant agar (*see* **Note 5**).
8. BCD medium (*see* Table 1): Basal minimum medium. Pour ca. 20–25 mL of a medium per a 9 cm Petri dish. Store plates at 4–10 °C.
9. BCD-AT medium (*see* Table 1): BCD medium supplemented with NH_4 as a nitrogen source. Pour ca. 20–25 mL of a medium per a 9 cm Petri dish. Store plates at 4–10 °C.
10. LB broth.
11. Mannitol 8 % (w/v): Weigh 8 g of mannitol and dissolve in 100 mL of water. If sterile solution is needed, autoclave or filter sterilize and store at room temperature.

Table 1
***Physcomitrella patens* media**

	BCD (1 L)	BCD-AT (1 L)	PRM-L (20 mL)	PRM-B (500 mL)	PRM-T (100 mL)	Final concentration
Stock solution B	10 mL	10 mL	200 µL	5 mL	1 mL	1 mM $MgSO_4$
Stock solution C	10 mL	10 mL	200 µL	5 mL	1 mL	1.84 mM KH_2PO_4
Stock solution D	10 mL	10 mL	200 µL	5 mL	1 mL	10 mM KNO_3, 45 µM $FeSO_4$
TES	1 mL	1 mL	20 µL	0.5 mL	0.1 mL	1× TES
AT	–	10 mL[a]	200 µL	5 mL[a]	1 mL	5 mM
Mannitol	–	–	1.2 g	30 g	6 g	6 % (w/v)
Plant agar	5.5 g	5.5 g	–	2.75 g	0.4 g	0.55 % (0.4 % for PRM-T)
H_2O	up to 1 L	up to 1 L	up to 20 mL	up to 495 mL	up to 99 mL	N/A
$CaCl_2$ 1 M	1 mL[b]	1 mL[b]	200 µL	5 mL[b]	1 mL[b]	1 or 10 mM $CaCl_2$
Sterilization	Autoclave	Autoclave	Filter	Autoclave	Autoclave	N/A

[a]Or add appropriate amount of the diammonium (+) tartrate
[b]Add after autoclaving

12. Driselase solution: 1 % (w/v) Driselase (*see* **Note 6**) in 8 % (w/v) mannitol. Add 50 mg Driselase into 5 mL of 8 % (w/v) mannitol solution (not sterile). Dissolve the powder in the centrifuge tube by inverting the tube occasionally and gently for 15 min. Centrifuge at 2500 rpm for 5 min. Filter sterilize the supernatant through a 0.22 µm filter. Prepare Driselase solution in amount of 5 mL per one *Physcomitrella* protonemal homogenate plate.
13. MaMg solution: 0.15 M $MgCl_2$, 8 % (w/v) mannitol, 0.1 % (w/v) MES. Weigh 3.05 g of $MgCl_2 \cdot 6H_2O$, 8 g of mannitol, and 0.1 g of MES. Dissolve all ingredients consecutively in 90 mL of dH_2O. Adjust the pH to 5.6 with 1 M KOH. Make up to 100 mL. Autoclave. Aliquot the solution in 10 mL plastic tubes and store at –20 °C (*see* **Note 7**). Before use thaw by placing the tube in a warm water and vortex until no precipitate is visible. Spin down.
14. PEG-CMS: 0.1 M $Ca(NO_3)_2$, 20 mM HEPES, 7.28 % (w/v) mannitol, 40 % PEG 6000. Weigh 0.236 g of $Ca(NO_3)_2.4H_2O$, 0.0476 g HEPES, and 0.728 g of mannitol and consecutively

dissolve in 5 mL of water. Make up to the volume of 6 mL. Measure the pH. A pH between 7 and 8 is acceptable (ideally pH 7.5). To adjust the pH, use 1 M KOH (usually 30–40 μL). Add PEG, dissolve the solution by incubation at 37 °C in water bath, and shake occasionally. Make up to 10 mL and mix the viscous solution thoroughly (*see* **Note 8**). Leave it for several hours to stabilize the pH. Filter sterilize and dispense in 1 mL aliquots in 1.5 mL microtubes. Store at –20 °C. Before use thaw PEG-CMS by placing tubes in warm water, vortex, and pulse spin (*see* **Note 9**). Dispense 300 μL aliquots into 10 mL centrifuge tubes and centrifuge the tubes briefly. Prepare one tube per transformation.

15. PRM-L (liquid) (*see* Table 1 and **Note 10**): Prepare 2 mL per transformation on the day of transformation.
16. PRM-B (bottom layer) (*see* Table 1 and **Note 10**): Plates with solid medium overlaid with cellophane, prepared 1–2 days before transformation, three plates per transformation. Pour maximum 20 mL of PRM-B into 9 cm Petri dish.
17. PRM-T (top layer) (*see* Table 1 and **Note 10**): Low agarose medium, prepare 10 mL per transformation on the day of embedding the protoplasts.
18. Antibiotics (*see* **Note 11**).

3 Methods

The tissue culture and transformation procedures require sterile conditions and are carried out in the laminar airflow hood.

3.1 Physcomitrella patens Protonemal Tissue Growth, Propagation, and Collection

1. Prepare BCD-AT growth plates and overlay them with sterile cellophane discs (*see* **Note 12**).
2. Take a well-grown (5–7 days old) protonemal homogenate BCD-AT plate (Fig. 2a). Using a sterile spatula, scrape the protonemal tissue off of the cellophane and transfer it to the sterile glass tube of a blender assembly filled with 10 mL of sterile water.
3. Blend the tissue in water until it is cut completely into ca. 0.5 mm small pieces.
4. Using a sterile pipette, inoculate cut tissue suspension onto five cellophane overlaid BCD-AT plates 2 mL each. Make quick circular moves with the plate to evenly distribute the inoculum.
5. To check for a bacterial contamination, add a few drops of the protonemal suspension to the 10 mL tube with 2 mL LB broth. Let it stay at the room temperature for 1–2 days and shake once a day (*see* **Note 13**).

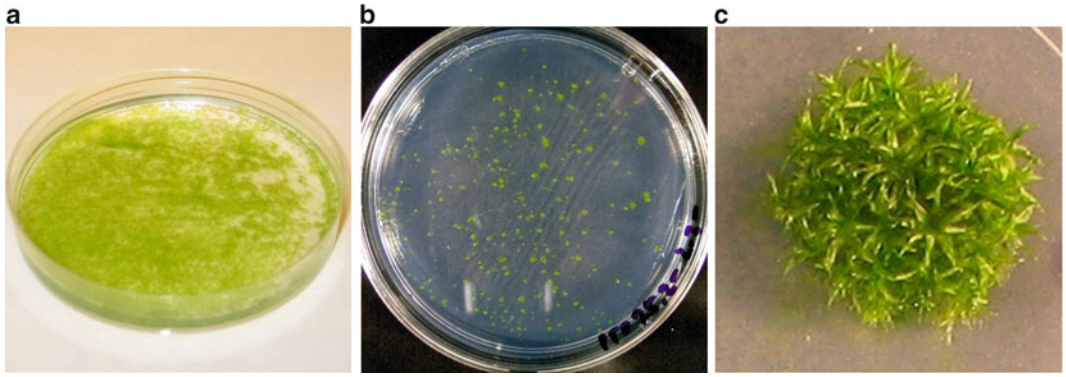


Fig. 2 The moss *Physcomitrella patens* grown in vitro. (**a**) Protonemal homogenate plate, 6 days after incubation on BCD-AT plate. (**b**) Regenerating protoplasts after transformation, 5 days after incubation on PRM. (**c**) Colony with gametophores, incubated for 3 weeks on BCD-AT plate

6. Cultivate the plates for 5–7 days (*see* **Note 14**) at 25 ± 2 °C, discontinuous white light (16-h light/8-h dark, 120 μmol/m^2/s).

3.2 Physcomitrella patens Protoplast Isolation, PEG-Mediated Transformation, and Regeneration

1. Collect the tissue from two 5–7 day old, well-grown BCD-AT cellophane overlaid plates (*see* **Note 15**) with a sterile spatula into a sterile container (e.g., Petri dish) containing 10 mL Driselase solution.
2. Let the tissue digest at room temperature with occasional gentle stirring until the tissue loses its filamentous character (45–90 min) (*see* **Note 16**).
3. Using a sterile pipette transfer the digested protoplast suspension onto a 70 μm filter placed in the corresponding container (*see* **Note 17**). Transfer the collected filtered suspension into the centrifuge tube.
4. Centrifuge at 800 rpm for 4 min, no brakes (*see* **Note 18**).
5. Carefully aspirate the supernatant at once using the pipette and leave ca. 0.1 mL of the liquid over the pellet. Resuspend the protoplasts in residuum very gently by rocking, tilting, and turning the tube. Slowly add 10 mL of 8 % mannitol (pour liquid slowly over the tube wall) and gently mix by rocking the tube. The cells should be evenly resuspended. Repeat **steps 4** and **5**.
6. Set aside 0.2–0.3 mL of protoplast suspension to measure the protoplast density using the counting chamber (hemocytometer).
7. Recover the remaining protoplasts by centrifugation at 800 rpm for 4 min, no brakes (*see* **Note 18**). Meanwhile calculate the protoplast density.

8. To calculate the protoplast density pipette small amount of protoplast suspension to both chambers of the hemocytometer. Count protoplasts in eight 1 mm^2 large squares and calculate the mean value *n*. This gives the density of the protoplasts in 0.1 μL and $n \times 10^5$ corresponds to the total amount of protoplasts in 10 mL of 8 % mannitol. Calculate the amount of MaMg solution to be added to give a final protoplast density of 1.6×10^6/mL.
9. Carefully add calculated amount of MaMg solution to the pellet by slowly pouring it over the tube wall.
10. (Optional) Transfer the resuspended protoplasts from the 10 mL tube to a wider and shallower container (e.g., Petri dish) in order to make further manipulations easier and reduce the risk of contamination.
11. Pipette 300 μL of the protoplast suspension using a wide-bored blue tip into the microtube containing 30 μL DNA solution.
12. Using a Pasteur pipette transfer the mixture of protoplasts and DNA to a 10 mL tube containing 300 μL PEG-CMS. Stir the mixture gently but thoroughly with the pipette tip. Gently suck up and expel the whole mixture once.
13. Heat-shock the protoplasts by placing the tubes in the water bath at 45 °C for 5 min.
14. Take the tubes out from the water bath and cool them in water at room temperature for 10 min.
15. Reduce the concentration of PEG by diluting the mixture with 8 % mannitol. Diluting has to be done slowly over the next 30–60 min in 6–7 steps. Add stepwise twice 300 μL, twice 600 μL, and twice 1 mL aliquots of mannitol and finally top up to the volume of 8 mL. The protoplasts must be always mixed very gently but thoroughly by tilting and rolling the tube. Wait for 3 min between each addition to give the protoplasts enough time to recover.
16. Centrifuge the protoplasts at 800 rpm for 4 min, no brakes (*see* **Notes 18** and **19**).
17. Carefully remove the supernatant (*see* **Note 20**) and gently add 2 mL of PRM-L to each tube. Mix gently; do not shake.
18. Incubate the tubes at 25 °C in the dark overnight (*see* **Note 21**).
19. Next morning, incubate the regenerating protoplasts under the normal growth room light conditions until the PRM-T medium is ready to embed the protoplasts.
20. Prepare PRM-T but do not add $CaCl_2$ yet. Divide it into 10 mL aliquots into the tubes, one tube per transformation. Cool down the PRM-T to 45 °C in the preheated water bath.

21. Centrifuge the protoplasts at 800 rpm for 3 min, no brakes (*see* **Note 18**).
22. Take one of the PRM-T aliquots from the water bath and transfer it to the hood. Add $CaCl_2$ in amount 100 µL per 10 mL of PRM-T.
23. Carefully aspirate the supernatant from the tube with protoplasts and resuspend the protoplasts in residuum very gently by rocking, tilting, and turning the tube. Add 10 mL of PRM-T and pour this on three PRM-B plates (*see* **Note 22**).

3.3 Selection of the Genomic Transformants

An outline of the selection process is shown in Fig. 3.

1. Cultivate the embedded protoplasts 5 days under standard conditions in the growth chamber (Fig. 2b) (*see* **Note 23**).
2. Initiate the first selection by transferring the cellophane with the top agar containing regenerated protoplasts on BCD-AT plate supplemented with appropriate antibiotics. Cultivate for 2–3 weeks under the standard conditions in the growth chamber (*see* **Note 24**).
3. Release the selection. With sterile forceps transfer colonies that show growth on the selective medium to the BCD-AT plates with no antibiotics and no cellophane. Cultivate for 2 weeks under the standard conditions in the growth chamber (*see* **Note 25**).
4. Repeat the selection. Using sterile forceps transfer small part of the protonemal tissue from the edge of each colony (spot inoculum) to BCD-AT medium supplemented with antibiotics and no cellophane. Cultivate for 2 weeks under the standard conditions in the growth chamber.

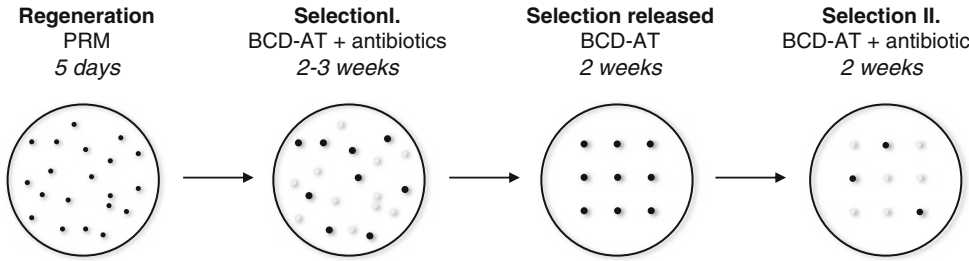


Fig. 3 Outline of the selection of *Physcomitrella* stable transformants. Protoplasts regenerate 5 days on PRM plate. To select the transformed colonies, a cellophane disc with colonies is transferred onto selective BCD-AT plate with antibiotics and incubated for 2–3 weeks. Surviving colonies are transferred on the BCD-AT plate and grown for 2 weeks without antibiotics, allowing unstable transformants (with non-integrated vector) to lose their antibiotic resistance. Spot inocula of all the colonies are then transferred to the selective BCD-AT plate with antibiotics in order to select stable transformants (with targeting DNA integrated into the genomic DNA). *Black dots* represent surviving colonies; *grey dots* represent dying colonies

5. Select stable transformants that survive on the selection medium (*see* **Note 26**).

3.4 Preparation of the Material for Verification and Molecular Characterization of the Targeted Gene Disruption

1. Transfer spot inoculum (small part of the protonemal tissue from the edge of the colony) of the wild type and the lines to be analyzed on the BCD-AT plates with cellophane. Seal the dish with the surgical tape to reduce evaporation. Cultivate under the standard conditions in the growth chamber for 3–4 weeks or longer until the colonies provide sufficient tissue (ca. 0.5–1 cm in diameter) (*see* **Note 27**).
2. To obtain material for molecular analyses from colonies, pick the colony from the cellophane with forceps, dry it briefly with a filter paper, transfer into a 1.5 mL microcentrifuge tube, and freeze in liquid nitrogen.
3. Grind the tissue to a very fine powder using a plastic pestle (cooled in liquid nitrogen) attached to a homogenizer. Keep the tissue frozen by scooping some liquid nitrogen into the tube.
4. To establish a protonemal homogenate subculture from the confirmed transformants transfer the colony (Fig. 2c) on a 1/6 sector of the BCD-AT plate with cellophane, dissect the protonemal part of the colony to as small pieces as possible with forceps, and distribute over the sector of the plate (*see* **Note 28**). Cultivate under standard conditions in the growth chamber for about 2 weeks, until the tissue is dense and has fresh green color. Subculture the protonemal tissue several times (*see* Subheading 3.1, **steps 2–4**) (*see* **Note 29**). The aim is to achieve a yield 500 mg of the tissue per protonemal homogenate plate (*see* **Note 30**).
5. To obtain material for molecular analyses from a protonemal homogenate tissue (Fig. 2a), scrape the protonemal tissue with a sterile spatula off of the cellophane and dry briefly between two layers of filter paper. Repeat drying approx. three times until almost no water comes out of the tissue. Weigh the tissue, transfer a maximum of 100 mg into the 1.5 mL microcentrifuge tube, and freeze in liquid nitrogen. Perform **step 3** to grind the tissue.
6. Extract DNA and RNA using a plant extraction kit (*see* **Note 31**).

3.5 Preparation of the Material for Phenotypic Analyses of the Protonemal Tissue and Gametophores

1. To obtain protonemal tissue suitable for observation, subculture the protonemal tissue on the BCD-AT plates with cellophane (*see* Subheading 3.1, **steps 2–4**). Cultivate for a maximum of 7 days under standard conditions in the growth chamber (*see* **Note 32**).
2. To induce gametophore production, take a little piece (about 1 mm) of the protonemal tissue from freshly grown protonemal homogenate plate (Fig. 2a) with the forceps (*see* **Note**

32). Transfer the spot inoculum to the BCD-AT plate overlaid with cellophane. Cultivate under the standard conditions in the growth chamber for 1 week or until the colonies are about 1 cm in diameter. Transfer the colonies onto the BCD plates, with no cellophane. Cultivate under standard conditions in the growth chamber until the colonies develop gametophores (leafy shoots) with approximately ten leafs (*see* **Note 33**).

3. To compare colonies of different lines, inoculate protonemal spot inoculum (*see* **step 2**) of maximum six lines (including wild type) on one plate containing desired solid media. Arrange the colonies to the circle with approximately the same distance from the edge of the Petri dish and from each other. Cultivate under standard conditions in the growth chamber until the colonies are about 0.5–1 cm in diameter (*see* **Note 34**).

3.6 Short-Term Storage and Regeneration

1. To store *P. patens* as a protonemal homogenate tissue, seal well-grown (5–7 days old) protonemal homogenate plate (Fig. 2a) with parafilm. Store the plate at 4–10 °C and illuminate for 2 h per day with a white light (*see* **Note 35**). To fully regenerate the tissue, subculture the tissue (*see* Subheading 3.1, **steps 2–4**) at least twice. Protonemal tissue plates can be stored for up to 3 months.
2. To store *P. patens* as a colony, take a small piece of the protonemal tissue with sterile forceps or a small piece of the tissue from the edge of the colony (spot inoculum) and transfer to the BCD plate, no cellophane. Cultivate under standard conditions in the growth chamber for 2–3 weeks. Seal the dish with parafilm, store at 4–10 °C and illuminate the plates for 2 h per day with white light (*see* **Note 35**). Plates with colonies can be stored for 2–3 months. Regenerate the tissue by making spot inoculum and transfer to the BCD-AT plates (*see* **Note 36**).

4 Notes

1. For example IKA Ultra Turrax. Previously we used a custom-made blender assembly, consisting of a glass tube and a metallic part with rotating blades. The blender assembly was attached to a drill-like rotor.
2. A cardboard box without holes that closes very tightly is sufficient.
3. Discs can be purchased from a packaging company.
4. $CaCl_2$ cannot be autoclaved added to the medium. Add corresponding amount of the $CaCl_2$ solution to the cooled medium right before pouring it.

5. We use Plant Agar P1001, Duchefa. Use of another agar may require optimization of the final concentration in the medium.
6. We use Driselase from *Basidiomycetes* sp., SIGMA, D8037.
7. Since adjusting pH in small volume might be difficult, making 100 mL of the MaMg solution and freezing is recommended.
8. Proper dissolving of PEG requires repeated thorough mixing and may take time.
9. Make sure that no precipitate is visible.
10. Protoplast regeneration medium (PRM) is BCD-AT medium supplemented with 6 % (w/v) mannitol and higher concentration (10 mM) of $CaCl_2$.
11. According to the antibiotic resistance gene incorporated in the DNA targeting vector we use G418 at the concentration of 40 mg/L.
12. It may take a while to practice this. The media must not be too wet (condensed water) to prevent wrinkling of the cellophane.
13. Clear LB medium means no contamination. If the medium turns cloudy or opaque in 1–2 days, this is a sign of contamination.
14. Cultivate until the plate is homogenously and densely grown by protonemal tissue of intense green color (Fig. 2a).
15. Two plates are optimal to obtain sufficient amount of the tissue (about 1 g) and protoplasts for one transformation.
16. The extent of the digestion can be monitored under the microscope.
17. Filtration can be repeated for better removal of undigested tissue.
18. Since the protoplasts are very fragile, it is recommended to turn the brake off completely. If it takes too long until the centrifuge stops, it is possible to turn the brake on when speed is down to ca. 80 rpm.
19. Alternatively let the protoplasts settle down by leaving the tube standing upright on the bench for 30–60 min.
20. The pellet is very loose.
21. Place the tubes in a lighttight box and incubate in the growth chamber.
22. PRM-T solidifies quickly at room temperature; therefore this step must be done fast.
23. The protoplasts recovery may take 3–6 days. Check under the stereomicroscope if the cells proliferate and the filaments are formed.

24. Colonies originated from untransformed cells turn light brown and die.
25. This step allows non-genomic transformants to lose the non-integrated construct.
26. Do not miss positive transformants showing reduced growth of the colony as an effect of the gene disruption.
27. One colony gives enough tissue for PCR analysis of the gene targeting events, but not for all the subsequent analyses. It is also possible to establish a protonemal homogenate subculture (*see* Subheading 3.4, **step 4**) at this stage, which gives more tissue in shorter time, but is more laborious, especially in case of many positive colonies after the selection process.
28. If an increased amount of the starting tissue is needed, gametophores (leafy shoots) can also be dissected and used as a source of the protonemal tissue.
29. Reduce the amount of water for blending and the number of plates for inoculation during the first round of subculture.
30. Another option is to homogenize the colony in a small amount of water in the homogenizer and inoculate on a BCD-AT plate with cellophane.
31. To confirm targeted insertion event in the genome, PCR amplify and sequence both regions of insertion outward from the selection cassette. Use the outer primers in combination with primers to the selection cassette (Fig. 1). Southern blotting allows detection of multiple insertions in case they occur to the nonhomologous sites of the moss genome. To monitor the transcript production of the targeted gene, conduct the RT-PCR and Northern blotting.
32. Use well-grown tissue that was subcultured at least twice consecutively and grown for 5–7 days under standard conditions in the growth chamber.
33. While growth of gametophores is enhanced by cultivation on BCD medium, the overall growth of the colony is slower than on BCD-AT plates, hence the transfer from BCD-AT to BCD.
34. PEG-mediated transformation can generate polyploids due to protoplast fusion. Polyploid colonies have different shape, more caulonemal tissue, and less gametophores than the wild-type haploid colonies.
35. The most important is regular daily illumination rather than the quality and intensity of the light.
36. Additional options for storage of the *P. patens* lines can be found in the literature.

Acknowledgements

We thank Dr. Cuming and Dr. Kamisugi for expert advice and training in *Physcomitrella* transformation, as well as for providing us with the "Gransden" strain of *P. patens* and advice on its handling and for the vectors used for preparing constructs. This work was supported by the Austrian Science Fund (FWF [P26333]) to M.K.

References

1. Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin IT, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu SH, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science 319(5859):64–69. doi:10.1126/science.1150646, 1150646 [pii]
2. Wang XQ, Yang PF, Liu Z, Liu WZ, Hu Y, Chen H, Kuang TY, Pei ZM, Shen SH, He YK (2009) Exploring the mechanism of *Physcomitrella patens* desiccation tolerance through a proteomic strategy. Plant Physiol 149(4):1739–1750. doi:10.1104/pp. 108.131714
3. Frank W, Ratnadewi D, Reski R (2005) *Physcomitrella patens* is highly tolerant against drought, salt and osmotic stress. Planta 220(3): 384–394. doi:10.1007/s00425-004-1351-1
4. Charron AJ, Quatrano RS (2009) Between a rock and a dry place: the water-stressed moss. Mol Plant 2(3):478–486. doi:10.1093/mp/ssp018
5. Saavedra L, Svensson J, Carballo V, Izmendi D, Welin B, Vidal S (2006) A dehydrin gene in *Physcomitrella patens* is required for salt and osmotic stress tolerance. Plant J 45(2):237–249. doi:10.1111/j.1365-313X.2005.02603.x
6. Knight CD, Cove DJ, Cuming AC, Quatrano RS (2002) Moss gene technology. In: Gilmartin PM, Bowler C (eds) Molecular plant biology, vol 2. Oxford University Press, Oxford, pp 285–301
7. Cove D (2005) The moss *Physcomitrella patens*. Annu Rev Genet 39:339–358. doi:10.1146/annurev.genet.39.073003.110214
8. Quatrano RS, McDaniel SF, Khandelwal A, Perroud PF, Cove DJ (2007) *Physcomitrella patens:* mosses enter the genomic age. Curr Opin Plant Biol 10(2):182–189. doi:10.1016/j.pbi.2007.01.005
9. Ashton NW, Cove DJ (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss, *Physcomitrella patens*. Mol Gen Genet 154(1):87–95. doi:10.1007/bf00265581
10. Ashton NW, Grimsley NH, Cove DJ (1979) Analysis of gametophytic development in the moss, *Physcomitrella patens*, using auxin and cytokinin resistant mutants. Planta 144(5):427–435. doi:10.1007/BF00380118
11. Schaefer D, Zryd JP, Knight CD, Cove DJ (1991) Stable transformation of the moss *Physcomitrella patens*. Mol Gen Genet 226(3):418–424
12. Schaefer DG, Zryd JP (1997) Efficient gene targeting in the moss *Physcomitrella patens*. Plant J 11(6):1195–1206
13. Kamisugi Y, Cuming AC, Cove DJ (2005) Parameters determining the efficiency of gene targeting in the moss *Physcomitrella patens*. Nucleic Acids Res 33(19):e173. doi:10.1093/nar/gni172

4.3. Results

4.3.1. Extensive alternative splicing in pre-mRNA of Physcomitrella SR gene *Pp-RS27*

Pp-RS27 is the sole orthologue of RS protein subfamily in *P. patens*. The *Pp-RS27* (PP1S144_89V6) gene consists of 6 exons and 5 introns with characteristic long second intron (1613 nt). Information about alternative splicing and expression of this gene is scarce, based mainly on ESTs collections. Though alternative splicing of the long intron was found in dicots and monocots and in a more distant single celled green alga *Chlamydomonas reinhardtii* orthologue *crRSp35* (Kalyna et al., 2006), no experimental data supported alternative splicing in *Pp-RS27*.

To investigate alternative splicing profile of *Pp-RS27*, we performed RT-PCR using the RNA from the upper parts of gametophytes. We used primers designed to amplify transcripts from the first to the last exon (Figure 4.2A) or to the adjacent exons, to monitor alternative splicing exclusively within each intron (Figure 4.2B, C).

RT-PCR amplification using primers to the first and last exons revealed about ten potential AS isoforms (Figure 4.2A). Cloning and sequencing was successful for the three shortest and the most abundant variants. Two alternative 3' splice sites and one alternative 5' splice site were identified in the long second intron, giving rise to AS isoforms that include cassette exons in the length of 246 and 446 nt, both introducing premature termination codons (PTCs) (Figure 4.2A).

Analyses of the splicing patterns for each pair of adjacent exons showed extensive alternative splicing in the region between E2-E3, in the long second intron, whereas in all other regions only complete intron removal is detected. The additional very weak band present in the PCR using E5-E6 primers is probably unspecific, because its size is higher than the potential intron retention event (Figure 4.2B).

At least 16 distinct bands were detected using E2-E3 primers, and sixteen different AS variants were identified by sequencing. Selection of the alternative splice sites in all AS variants resulted into PTC. We detected eight alternative 5' and five alternative 3' splice sites (Figures 4.2C and D).

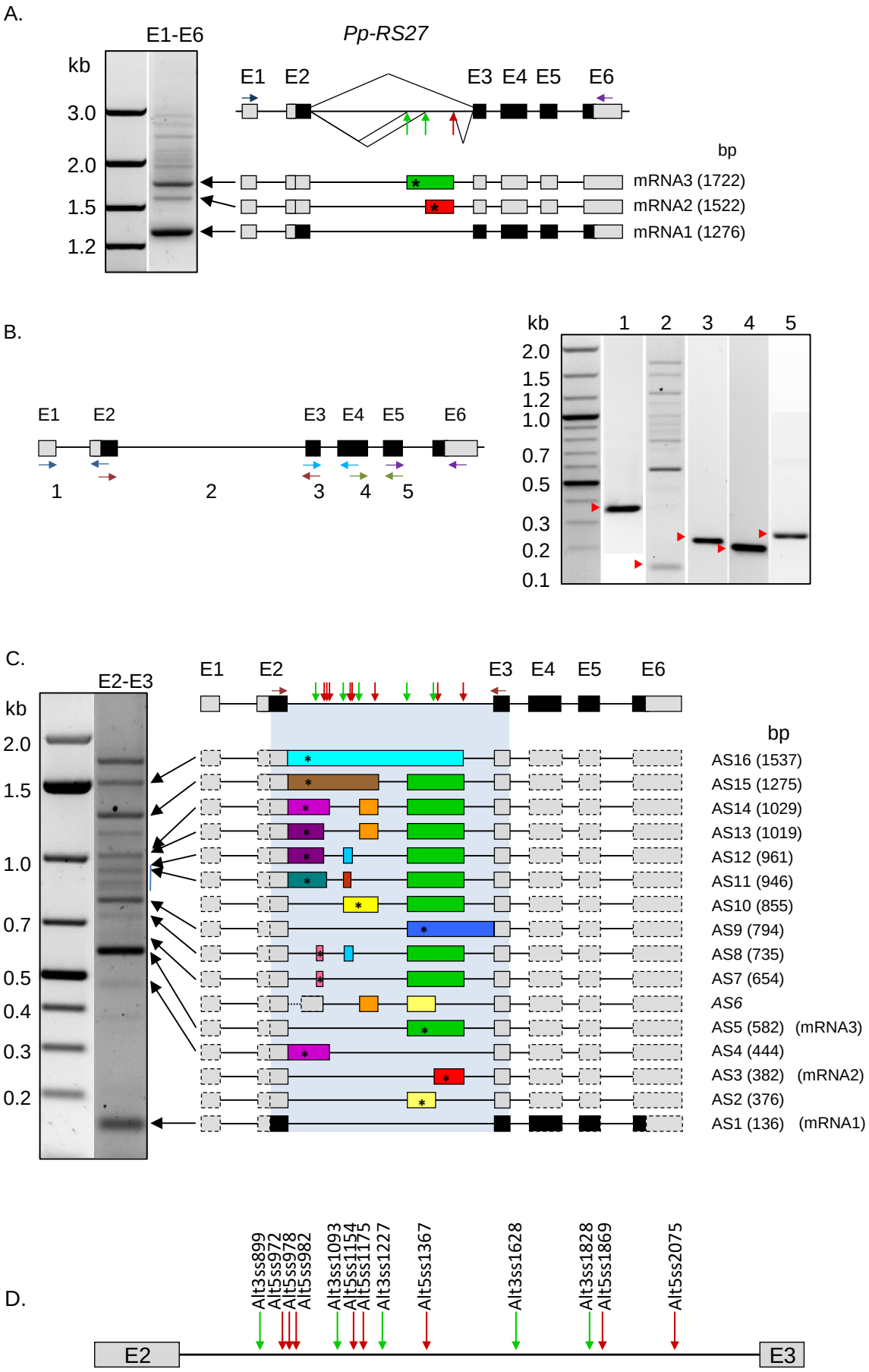

A.
E1-E6
Pp-RS27
kb
3.0
2.0
1.5
1.2
E1
E2
E3
E4
E5
E6
bp
mRNA3 (1722)
mRNA2 (1522)
mRNA1 (1276)
B.
E1
E2
E3
E4
E5
E6
1
2
3
4
5
kb
1
2
3
4
5
2.0
1.5
1.2
1.0
0.7
0.5
0.3
0.2
0.1
C.
E2-E3
kb
2.0
1.5
1.0
0.7
0.5
0.4
0.3
0.2
E1
E2
E3
E4
E5
E6
bp
AS16 (1537)
AS15 (1275)
AS14 (1029)
AS13 (1019)
AS12 (961)
AS11 (946)
AS10 (855)
AS9 (794)
AS8 (735)
AS7 (654)
AS6
AS5 (582) (mRNA3)
AS4 (444)
AS3 (382) (mRNA2)
AS2 (376)
AS1 (136) (mRNA1)
D.
Alt3ss899
Alt5ss972
Alt5ss978
Alt5ss982
Alt3ss1093
Alt5ss1154
Alt5ss1175
Alt3ss1227
Alt5ss1367
Alt3ss1628
Alt3ss1828
Alt5ss1869
Alt5ss2075
E2
E3

Figure 4.2. Pre-mRNA of *Physcomitrella* SR protein Pp-RS27 undergoes extensive alternative splicing.
RNA from the tissue of upper part of the gametophore was analysed. AS events were detected using RT-PCR with primers either to the first and last exons of *Pp-RS27* (A) or to adjacent exons (B and C).
A. RT-PCR analysis of the entire transcripts. Fully spliced and two most abundant and shortest alternatively spliced forms were sequenced. Identified mRNA isoforms are numbered and the length of the sequenced region in base pairs (bp) is indicated in brackets.
B. Each intron was analysed separately. Colored arrows and numbers define the amplified region of the gene. The long second intron shows multiple potential AS events. Red arrowheads mark bands corresponding to the events of complete intron removal.
C. Sequencing of the multiple bands amplified from E2 to E3 region revealed 16 different alternatively spliced variants (AS) that are numbered and the length (bp) of the amplified region is shown in brackets. Dashed lines depict regions that were not sequenced.
D. Scheme depicts relative positions of alternative splice sites detected in the second intron of *Pp-RS27*. Red and green arrows indicate 5' and 3' splice sites, respectively. The numbers refer to the positions of splice sites in the gene sequence. Boxes are exons (E), line is intron. Not to scale.
For A and C. Schemes summarize the identified splice sites. Red and green arrows mark alternative 5' and 3' splice sites, respectively. Black arrows connect DNA bands with the corresponding mRNA isoforms. Small colored arrows show the positions of primers.
For A-C. Exons (E) are illustrated as boxes, introns as lines. Black and grey boxes depict coding and non-coding exons, respectively. Positions of start codons and the first downstream premature termination codons are shown by vertical lines and asterisks within boxes, respectively. Colored boxes represent transcript regions included due to alternative splicing. The schemes are drawn to scale. DNA ladders are shown and indicate the size in kilobases (kb).

4.3.2. *Pp-RS27* possesses plenitude of alternative splice sites with diverse features

Intriguingly, sequence alignment of the detected alternative splice sites revealed that neither of the *Pp-RS27* thirteen alternative splice sites shares an evolutionary conservation observed among *Arabidopsis*, rice and maize orthologues, supporting previous analyses of EST sequences (Kalyna et al., 2006). Moreover, splice site sequences do not show similarity amid each other (data not shown).

Interestingly, some of the *Pp-RS27* alternative splice sites are located very closely to each other, separated only by 4 or 6 nucleotides (e.g. Alt5ss972, Alt5ss978, Alt5ss982) (Figure 4.2D) that may have influence on their recognition by snRNAs.

There are two 5’ splice sites with GC dinucleotide instead of almost invariant GT, i.e conventional 5’ and Alt5ss1869 splice sites. It is unusual to find two GC-type 5’ splice sites in one gene, because they are generally rare (<1%). On the other hand, GC-type 5’ splice sites are enriched in alternatively spliced genes in plants and animals (Campbell et al., 2006; Farrer et al., 2002; Thanaraj and Clark, 2001).

To estimate the potential of a particular splice site identified in the second intron to be recognized and selected for splicing, we determined strength of splice sites by scoring each splice site according to the position weight matrices (PWM) generated from the collection of splice sites of 160,800 introns in *P. patens* genome annotation (Venhuizen, unpublished). Splice site scores for 13 alternative splice sites of the second intron and two conventional splice sites (5’ and 3’), situated at the 5’ and 3’ end of the second intron are shown in the Table 4.6.

Alternative splice sites in *Pp-RS27* are generally weaker than splice sites responsible for complete intron removal, despite that they are selected relatively frequently (Table 4.6, Figure 4.2 A and C). Some alternative splice sites are used much more frequently than the others and the rate of their selection does not correlate with their splice site score. For instance, Alt5ss2075 with 69.45 score is used to generate 11 different alternatively spliced transcripts, and Alt3ss1628, though it does not have the highest score, is used for the highest number of alternatively spliced transcripts (12). In addition, Alt3ss1628 is used to generate the second most abundant transcript isoform (mRNA3) after the fully spliced transcript (mRNA1). Alt3ss1093 with the highest score among alternative 3’ splice sites (86.95) is used for four alternatively spliced variants only (AS8, AS10-AS12). Interestingly, though both GC-type 5’ splice sites have the highest scores (94.40 and 89.11, Table 4.6), Alt5ss1869 is used to generate only two very lowly abundant splice isoforms (AS2 and AS6, Figure 4.2 C).

Table 4.6. Scores of splice sites detected in the second intron of *Pp-RS27* (Figure 4.5) Scores were calculated using PWMs generated from the *P. patens* splice sites obtained from the *P. patens* genome annotation. Splice sites leading to splicing of the entire second intron are designated 5' and 3'. In alternative 5' and 3' splice site (Alt5ss and Alt3ss) names, the numbers refer to positions of splice sites in the gene sequence. Upper and lower case in splice site sequences represent the exonic and intronic sequences, respectively. Splice sites scores above 65 threshold are in bold. Last column lists AS events according to the numbering in Figure 4.3 C and the number in brackets is the total count of the AS events utilizing the particular splice site.

Splice site type	Sequence	Splice site scores				Used in AS (n)
		AT-AC	GC-AG	GT-AG	GT-AG (U12)	
5'	CAGgcaagtttga	9.96	**94.40**	44.61	29.70	1- 3, 5, 7- 10 (8)
Alt3ss899	gaagctcgtgtcagGAG	8.71	63.58	61.39	**67.31**	7, 8 (2)
Alt5ss972	GAGgttgtagtcg	0.00	1.66	60.45	18.81	6, 7, 8, 12, 13 (5)
Alt5ss978	GTAgtcggtttat	11.69	0.00	56.19	26.96	11 (1)
Alt5ss982	TCGgtttatttct	20.85	3.45	**65.28**	35.52	4, 14 (2)
Alt3ss1093	gttttgtttctcagATT	23.73	**84.13**	**81.13**	**86.95**	8, 10-12 (4)
Alt5ss1154	GAGgtagtcacag	25.02	10.63	**66.64**	37.50	11 (1)
Alt5ss1175	AATgttaagtaat	10.47	4.51	60.86	21.91	8, 12 (2)
Alt3ss1227	tactccttttcagAGT	29.69	**75.73**	**72.67**	**82.33**	6, 13, 14 (3)
Alt5ss1367	ATTgtgagttctg	10.83	8.52	**70.65**	27.98	6, 10, 13- 15 (5)
Alt3ss1628	gtgtgttaccgcagGCA	1.17	**78.29**	**73.84**	**73.43**	2, 5- 15 (12)
Alt3ss1828	gcaatgggtggcagCTC	3.62	63.43	62.04	**68.43**	3 (1)
Alt5ss1869	AAGgcatgtgcct	11.43	**89.11**	43.00	36.67	2, 6 (2)
Alt5ss2075	CAGgtggctatcg	16.47	8.11	**69.45**	33.73	3, 5, 7, 8, 10- 16 (11)
3'	gtttggttgcacagGAT	3.46	**83.17**	**78.13**	**72.06**	1-8, 10-16 (15)

4.3.3. PTC-containing AS transcripts of *Pp-RS27* are not affected by cycloheximide

All of the fifteen AS isoforms, generated due to alternative splicing in the second intron of *Pp-RS27*, have incorporated a PTC into their reading frame. To shed more light on the fate of these AS isoforms, we examined the AS pattern of *Pp-RS27* gene upon exposure of wild type Physcomitrella protonemal tissue to cycloheximide, an inhibitor of translation. Transcripts accumulating after cycloheximide treatment are potential targets for NMD, since it is a translation-dependent process, inhibited by cycloheximide.

Surprisingly, as shown in Figure 4.3, the abundance of AS isoforms changes slightly compared to mRNA1 after cycloheximide application, suggesting insensitivity of *Pp-RS27* AS transcripts to regulation by NMD pathway. Since NMD operates in the cytoplasm, PTC-containing AS transcripts of *Pp-RS27* possibly do not leave the nucleus. In contrast, mRNA1 accumulates upon cycloheximide treatment, compared to mock-treated sample.

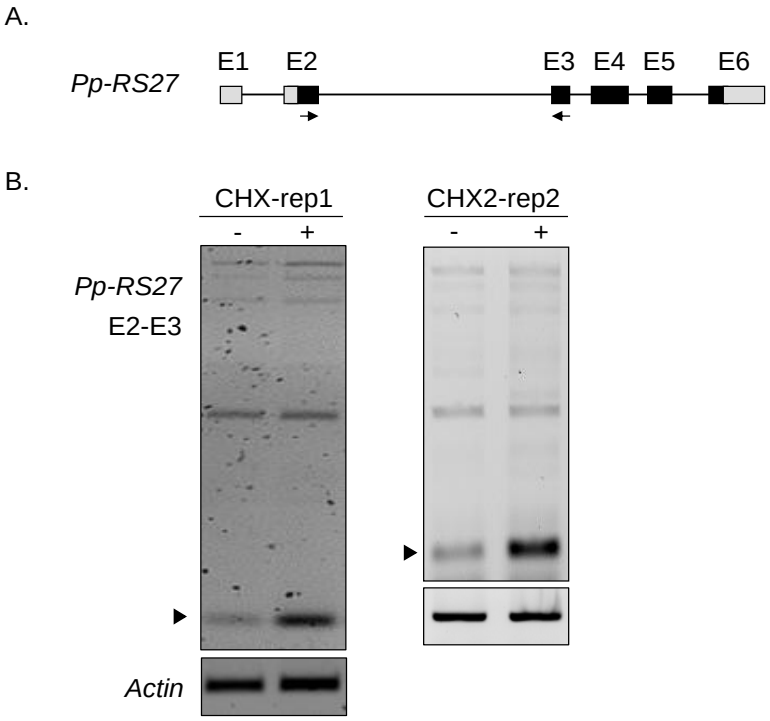


Figure 4.3. PTC-containing AS isoforms of *Pp-RS27* do not increase in abundance upon cycloheximide treatment.
A. Scheme of the gene *Pp-RS27* with positions of primers used in RT-PCR analysis. Exons (E) are illustrated as black and grey boxes, indicating the coding and non-coding regions, respectively. Lines represent the introns. Scheme is drawn to scale.
Wild type protonemal tissue was vacuum infiltrated with 20 µM cycloheximide (CHX) for 4 h (+) or with water (-) for control. Two biological replicates are shown.
Analysed region is defined by the pair of exons (E*n*-E*n*) to which primers align. Arrowheads mark AS1 (mRNA1) with fully spliced second intron. Bands above are the PTC+ AS variants (See Figure 4.3). *Actin* (PP1S381_18V6) was used as a loading control.

4.3.4. Physcomitrella SR genes *Pp-RS2Z37* and *Pp-RS2Z38* display multiple alternative splicing events

Pp-RS2Z37 (PP1S69_23V6) and *Pp-RS2Z38* (PP1S65_286V6) belong to the *P. patens* RS2Z subfamily of SR proteins, which is orthologous to *RS2Z* subfamily in *Arabidopsis*. Previous evolutionary studies predicted alternative 3' splice site in the long second intron of *Pp-RS2Z38,* based on the sequence conservation with rice and *Arabidopsis*, however EST data to support this AS event were not available (Kalyna et al., 2006). We used RT-PCR to examine AS events in *P. patens* RS2Z orthologues and to verify the occurrence of predicted AS splice site.

To investigate AS in *Pp-RS2Z37,* total RNA of the wild type protonemal tissue was subjected to three different RT-PCRs designed to monitor splicing of the region between E2 and E7, and in the first and the second intron separately (Figure 4.4).

Multiple PCR products were detected in the region of E2–E7, and five different AS transcripts were successfully sequenced. The fragments over 1.5 kb failed to be cloned and sequenced, probably due to the length and overrepresentation of the shorter fragments. One alternative 5' splice site and one alternative 3' splice site were identified in the second intron of *Pp-RS2Z37* (Figure 4.4 A). Selection of both alternative splice sites leads to inclusion of alternative exon (119 nt) and usage of 3' splice site alone results into addition of 257 nt sequence to the 3rd exon. Both events lead to the PTC. Two intron retention events affect the 4th and the 5th intron, respectively. Detected intron retention events are always combined with alternative splicing in E2–E3 region (Figure 4.4 A). The first intron analysed with the primers to the E1 and E2 was fully spliced and analysis of the second intron revealed multiple PCR products. Two most abundant PCR products correspond by size to the alternative 3' splice site event detected in the second intron (band sized between 0.3-0.4 kb) and to the potential intron retention of the second intron (band about 1.2 kb), however these possibilities were not confirmed by sequencing (Figure 4.4 B).

Although both genes of Pp-RS2Z subfamily share the same exon-intron structure and highly homologous gene sequences, RT-PCR analysis of *Pp-RS2Z38* showed differences in AS pattern detected in wild type protonemal tissue.

We analysed *Pp-RS2Z38* transcripts by RT-PCR with primers to 1st exon and to the beginning of the 7th exon, spanning almost the entire transcript sequence (Figure 4.5). Analysis uncovered multiple PCR products and six distinct mRNAs were identified by sequencing. In the region between E1-E2, we have found one alternative 3' splice site and

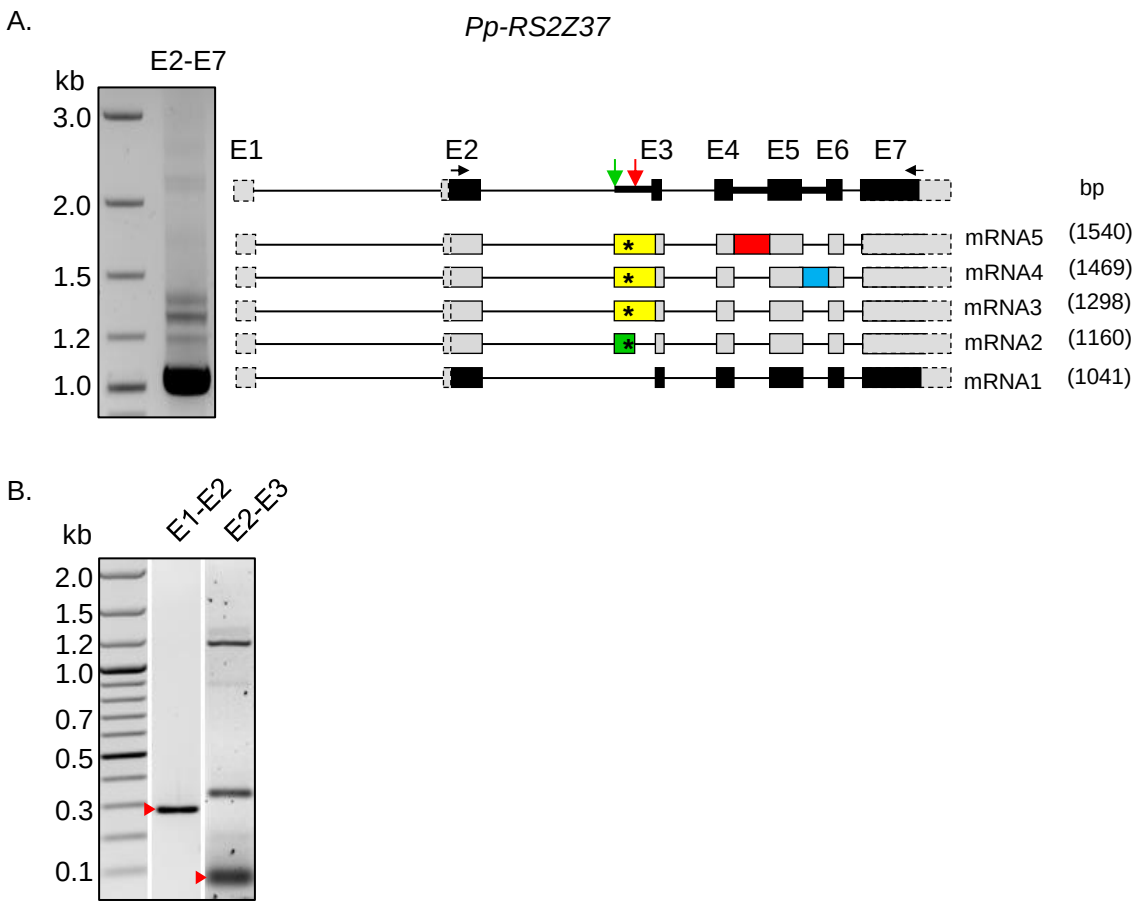


Figure 4.4. *Pp-RS2Z37* gene produces at least five splice variants.
RNA from the protonemal tissue was analysed by RT-PCR with primers to the second and last exons (A) and to the exons adjacent to the first and to the second intron (B).
A. Multiple bands were amplified in the region between the E2 to E7, and five alternatively spliced mRNA isoforms were identified by sequencing. They are denoted as mRNA1-5 and the length of the amplified region in base pairs (bp) is indicated in brackets.
Scheme above the mRNAs shows the positions of primers (small black arrows) and summarizes the identified alternative splice sites and intron retention events. Exons (E) are shown as boxes, introns as lines. Red and green arrows mark alternative 5' and 3' splice sites, respectively. Black and grey boxes depict coding and non-coding exons. Positions of start codons and the first downstream premature termination codons are shown by vertical lines and asterisks within boxes, respectively. Colored boxes represent transcript regions included due to alternative splicing. Dashed lines depict regions that were not sequenced. Schemes are drawn to scale.
B. Two introns were analysed individually. The first intron (corresponding to region E1-E2) was shown to be fully spliced, and the long second intron (region E2-E3) reveals multiple putative AS events. Red arrows mark the products with the fully spliced introns.
DNA ladders indicate the sizes in kilobases (kb).

one alternative 5’ splice site that are selected simultaneously, resulting into incorporation of 51 nt into the 5’UTR. In the region of the long second intron, one alternative 3’ splice site and two alternative 5’ splice sites are present. Their various utilization gives rise to three different mRNAs, two with distinct alternative exons (77 nt and 120 nt) and the third one with 258 nt extension to the 3rd exon. All identified AS events in this region cause incorporation of PTCs into the transcripts. Another alternative 3’ splice site is located between E4–E5 and leads to the in frame inclusion of 15 nt. Therefore, in contrast to *Pp-RS27* and *Pp-RS2Z37*, *Pp-RS2Z38* is the only gene where alternative splicing has capacity to generate two protein isoforms. These isoforms differ by five amino acids (RMMDQ) that are added to the C-terminal end of the RRM motif.

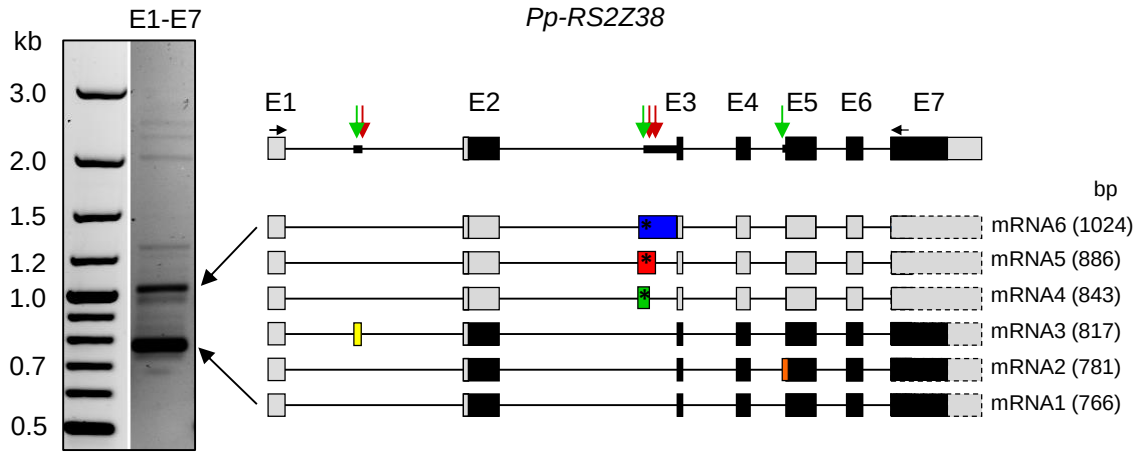


Figure 4.5. *Pp-RS2Z38* generates at least six alternatively spliced mRNAs.
Entire transcripts (E1-E7) of the *Pp-RS2Z38* gene were analysed by RT-PCR of RNA extracted from the protonemal tissue. Multiple bands are visible on the gel, and six sequenced mRNAs are numbered and depicted in schemes. Sizes of the amplified regions in base pairs (bp) are indicated in the brackets. Black arrows connect schemes with the corresponding PCR bands. DNA ladder indicates the size in kilobases (kb). The top gene scheme summarizes the identified AS events. Intronic regions included due to alternative splicing are shown as thick lines. Red and green arrows mark alternative 5’ and 3’ splice sites, respectively. Exons (E) are shown as boxes, introns as lines. Black and grey boxes depict coding and non-coding exons. Colored boxes represent transcript regions included due to alternative splicing. Positions of start codons and the first downstream premature termination codons are shown by vertical lines and asterisks within boxes, respectively. Black arrows show the positions of primers. Dashed lines depict regions that were not sequenced. Schemes are drawn to scale.

4.3.5. Physcomitrella RS2Z homologues possess unique as well as evolutionary conserved alternative splice sites

Our analyses confirmed alternative splicing in two *P. patens* orthologues of RS2Z subfamily and identified several alternative splice sites. To address the question of evolutionary conservation of the splice sites, we aligned identified splice sites with the corresponding introns of rice and Arabidopsis orthologues (Figure 4.6). Further, we determined strength of all splice sites in both genes, to evaluate splicing potential for each particular site (Table 4.7, 4.8).

Alternative 3' splice site in the long intron of both genes is highly conserved among dicots (Arabidopsis) and monocots (rice) (Figure 4.6), confirming experimentally the prediction of a splice site in *P. patens* orthologues by sequence analysis (Kalyna et al., 2006).

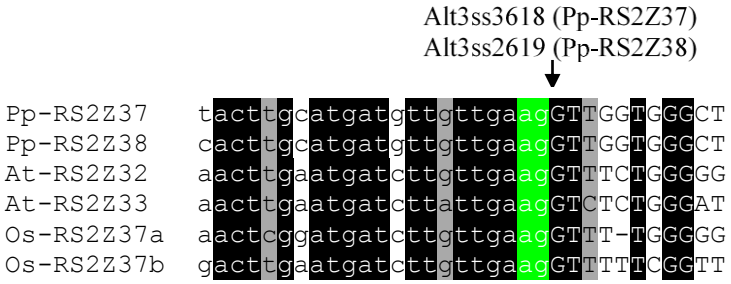


Figure 4.6. Conservation of sequences around the alternative 3' splice site in the long intron of RS2Z orthologous genes in *P. patens*, Arabidopsis and rice. Nucleotides identical in all genes are shaded. Intronic and exonic sequences are in small and capital letters, respectively. Terminal dinucleotide of the 3' splice site is highlighted green.

Two alternative 5' splice sites, identified in the long intron of *Pp-RS2Z38* are both present in the long intron of *Pp-RS2Z37*, but only one of them was detected by RT-PCR in *Pp-RS2Z37* (Figure 4.4, 4.5, 4.7). In orthologous sequences of angiosperms (Arabidopsis and rice), these sites are not present. Interestingly, both splice sites are situated in the region towards the 3' end of the second intron, which is characterized by highly similar sequence in both genes, suggesting its importance and stronger selective pressure, in contrast to the central and the 5' parts of this intron that are more dissimilar (Figure 4.7).

In the first intron, located in the 5' UTR of *Pp-RS2Z38*, two AS events are creating a 51 nt alternative exon in mRNA3. This AS event is specific to *Pp-RS2Z38* and no identical or similar splice sites are found in its paralogue *Pp-RS2Z37*. Nucleotide sequence analysis

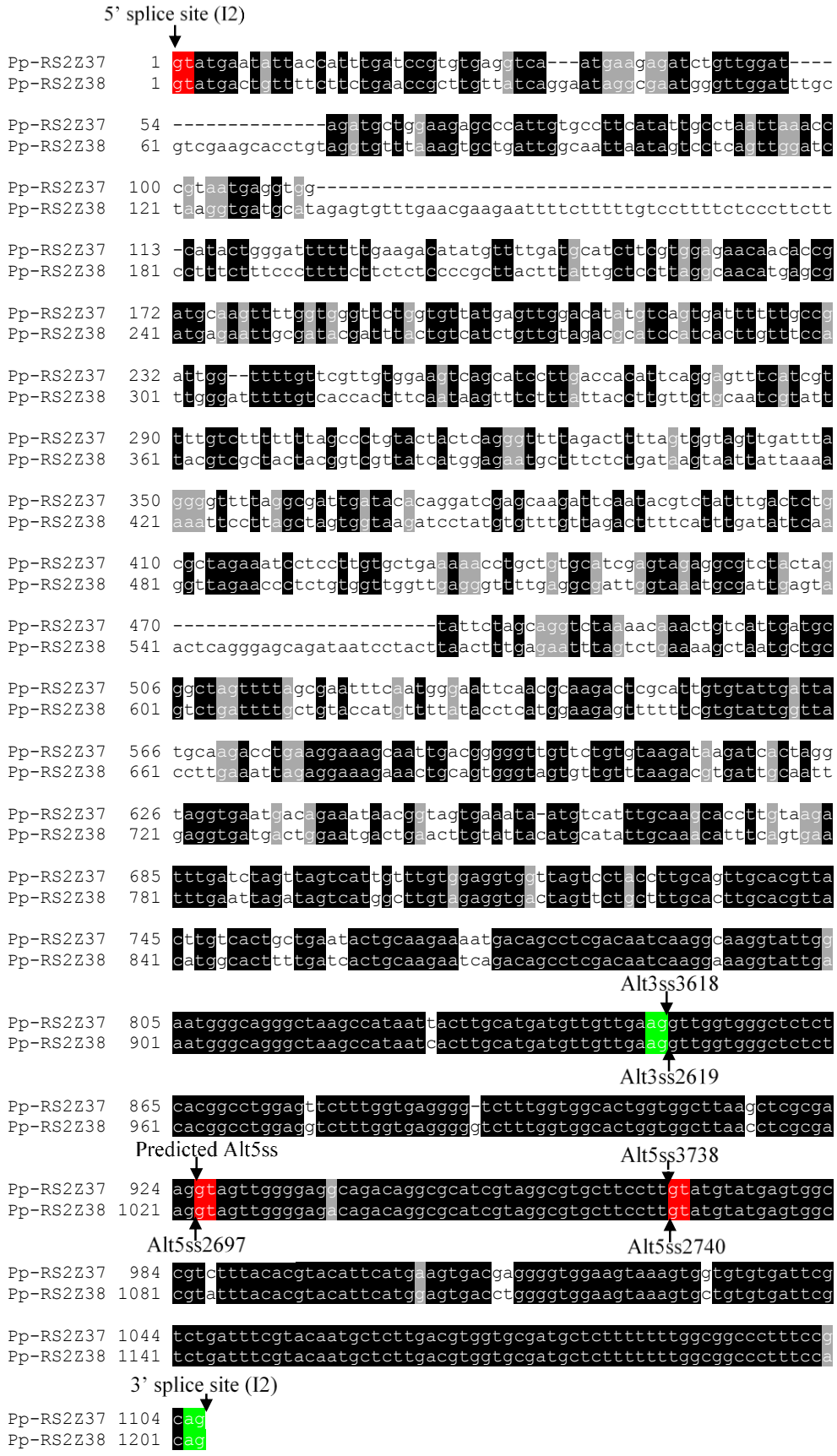


Figure 4.7. Alignments of the long introns of *Pp-RS2Z37* and *Pp-RS2Z38*. Nucleotides identical in both genes are shaded. Terminal dinucleotides of the 5' and 3' splice sites are highlighted red and green, respectively.

of the first introns of Arabidopsis and rice orthologues revealed no sequence similarities that would predict such an event.

Selection of alternative 3' splice site in the 4th intron, detected in mRNA2 of *Pp-RS2Z38*, adds 15 nt in frame to the 5th exon, corresponding to additional five amino acids (RMMDQ) in the potential protein isoform. Identical sequence is found in *Pp-RS2Z37* (Figure 4.8), but not in rice and Arabidopsis. We searched for a similar AS event in the protein sequence databases by blasting the translated *Pp-RS2Z38* mRNA2. Few hits showed variability in this region (e.g. in *Marchantia polymorpha*) but none of the sequences proved to contain an AS event.

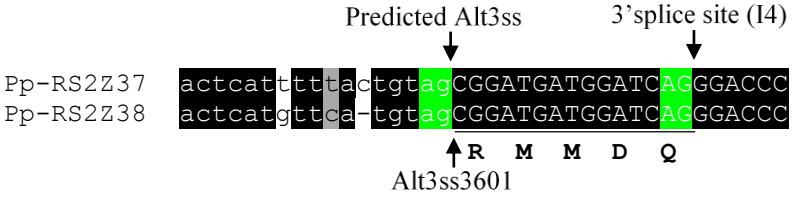


Figure 4.8. Conservation of the alternative 3' splice site in the fourth intron of *Pp-RS2Z37* and *Pp-RS2Z38* orthologues. Line below the alignment indicates added sequence due to alternative splicing and translation into amino acid sequence.
Identical nucleotides are shaded. Intronic and exonic sequences are in small and capital letters, respectively. Terminal dinucleotide of the 3' splice site is highlighted green.

In Pp-RS2Z37, we detected two intron retention events. While in the 4th intron, 3' splice site is as expected weak, just above the threshold, the 5th intron contains the strongest splice sites in this gene (except 3' splice site in the 1st intron) (Table 4.7). Splice sites of corresponding introns in *Pp-RS2Z38* are weaker than in *Pp-RS2Z37*. Especially both (5' and 3') splice sites of the 4th intron are very weak, however no intron retention was detected in *Pp-RS2Z38*, suggesting that intron retention occurs independently of the splice site strength (Table 4.8). Alternative splice site (Alt3ss3601) in the same intron is also very weak (65.41) (Table 4.8), however it is conserved in *Pp-RS2Z37* (Figure 4.8) and potentially generates a protein-coding isoform. Highly conserved alternative 3' splice sites in the long intron of RS2Z genes are quite strong (78.36) with scores comparable to other 3' splice sites in RS2Z genes, except to the competing conventional 3' splice sites of the same introns in both RS2Z genes that are just above threshold.

Table 4.7. Scores of splice sites detected in *Pp-RS2Z37*

Scores were calculated using PWMs generated from the *P. patens* splice sites obtained from the *P. patens* genome annotation. Splice sites leading to splicing of the entire intron (I*n*) are designated 5' and 3'. In Alternative 5' and 3' splice site (Alt5ss and Alt3ss) the numbers refer to the positions of splice sites in the gene sequence. Upper and lower case in splice site sequences represent the exonic and intronic sequences, respectively. Splice sites scores above 65 threshold are in bold. Last column lists AS events according to the numbering in Figure 4.4 A and the number in brackets is the total count of the AS events utilizing the particular splice site.

Intron	Splice site type	Sequence	Splice site scores				Usage in mRNA (n)
			AT-AC	GC-AG	GT-AG	GT-AG (U12)	
I1	5'	TCGgtacgtggca	6.93	3.20	**68.86**	30.94	1-5 (5)
	3'	gaatatttttgcagCTA	21.54	**83.07**	**78.66**	**87.01**	1-5 (5)
I2	5'	GAGgtatgaatat	27.12	24.62	**84.06**	38.07	1-5 (5)
	Alt3ss3618	tgatgttgttgaagGTT	11.72	**78.36**	**76.01**	**65.95**	2-5 (4)
	Alt5ss3738	CTTgtatgtatga	25.53	8.07	**69.80**	43.03	2 (1)
	3'	ggccctttccgcagAGT	22.99	**69.12**	**65.51**	**73.81**	1, 2 (2)
I3	5'	GTTgtgagtatac	6.85	0.00	**68.68**	25.94	1-5 (5)
	3'	attcaattcttcagGAA	23.16	**78.27**	**72.67**	**81.47**	1-5 (5)
I 4	5'	AGGgtgagtgtta	0.00	11.89	**78.82**	23.79	1-4 (4)
	3'	ggatgatggatcagGGA	4.89	**69.26**	**65.42**	**69.35**	1-4 (4)
I5	5'	CAGgtgagcttct	19.62	24.25	**89.26**	30.62	1-3, 5 (4)
	3'	tgtttctggtgcagTCC	13.07	**76.85**	**75.56**	**68.15**	1-3, 5 (4)
I6	5'	CAGgtttgttctc	19.72	22.88	**84.06**	35.15	1-5 (5)
	3'	tttgatttattcagTCG	18.60	**77.43**	**74.35**	**73.03**	1-5 (5)

Table 4.8. Scores of splice sites detected in *Pp-RS2Z38*

Scores were calculated using PWMs generated from the *P. patens* splice sites obtained from the *P. patens* genome annotation. Splice sites leading to excision of the entire intron (I*n*) are named 5' and 3'. In alternative 5' and 3' splice site (Alt5ss and Alt3ss) names, the numbers refer to the positions of splice sites in the gene sequence. Upper and lower case in splice site sequences represent the exonic and intronic sequences, respectively. Splice sites scores above 65 threshold are in bold. Last column lists AS events according to the numbering in Figure 4.5 and the number in brackets is the total count of the AS events utilizing the particular splice site.

Intron	Splice site type	Sequence	Splice site scores				Usage in mRNA (n)
			AT-AC	GC-AG	GT-AG	GT-AG (U12)	
I1	5'	ACGgtacgtctcg	14.06	11.68	**76.61**	33.03	1-6 (6)
	Alt3ss687	gcgtgtggttgtagGAA	0.00	**75.22**	72.67	71.12	3 (1)
	Alt5ss739	TTGgtagggcttg	16.03	0.87	**66.95**	29.00	3 (1)
	3'	ttttgtttttgcagTTT	21.06	**95.42**	91.72	82.85	1-6 (6)
I2	5'	GAGgtatgactgt	24.78	20.09	**80.74**	35.92	1-6 (6)
	Alt3ss2619	tgatgttgttgaagGTT	11.72	**78.36**	**76.01**	**65.95**	4-6 (3)
	Alt5ss2697	AAGgtagttgggg	5.48	4.90	**67.81**	30.07	4 (1)
	Alt5ss2740	CTTgtatgtatga	25.53	8.07	**69.80**	43.03	5 (1)
	3'	ggccctttccacagAGT	22.14	**65.27**	61.73	**76.33**	1-5 (5)
I3	5'	GTTgtgagtgttc	1.06	0.48	**70.15**	25.19	1-6 (6)
	3'	gttcacttttcagGAA	21.24	**82.46**	**77.69**	**83.53**	1-6 (6)
I4	5'	AGGgtgagcacaa	14.63	3.23	**70.17**	28.44	1-6 (6)
	Alt3ss3601	tcatgttcatgtagCGG	3.98	**65.41**	63.66	51.38	2 (1)
	3'	ggatgatggatcagGGA	4.89	**69.26**	**65.42**	**69.35**	1, 3-6 (5)
I5	5'	CAGgtttgagtgc	8.39	9.88	**73.28**	26.85	1-6 (6)
	3'	tggttttgatgcagTCC	12.81	**76.83**	**75.38**	**65.80**	1-6 (6)
I6	5'	CAGgtgcacttcc	19.34	2.84	**75.49**	30.62	1-6 (6)
	3'	tttatttttacagCCG	24.47	**84.27**	**80.69**	**78.89**	1-6 (6)

4.3.6. PTC-containing AS transcripts of *Pp-RS2Z37* and *Pp-RS2Z38* do not accumulate upon cycloheximide treatment

Majority of *Pp-RS2Z37* and *Pp-RS2Z38* transcripts contain PTCs followed by splice junctions (Figures 4.4, 4.5), thus suggesting their sensitivity to NMD. To analyse the fate of the *Pp-RS2Z37* and *Pp-RS2Z38* transcripts, we treated the wild type tissue with cycloheximide by flooding for four hours. We did not observe significant accumulation of any of the splice variants (Figure 4.9). Previous studies in Arabidopsis have shown that only

a fraction of transcripts, which possess predicted NMD features, are subjected to NMD. Especially, transcripts with retained introns were shown to escape NMD (Kalyna et al., 2012). However, the treatment with cycloheximide without vacuum infiltration might be less efficient and should be repeated using conditions described for *Pp-RS27*.

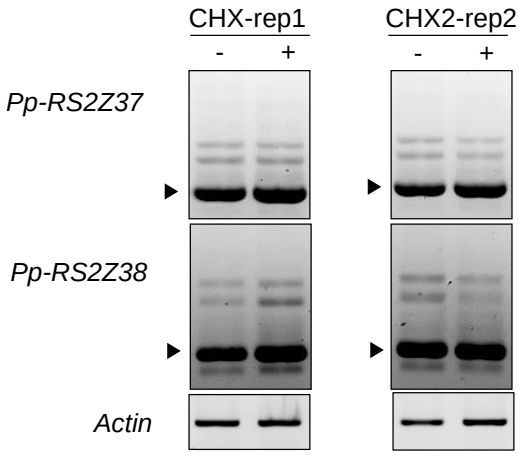


Figure 4.9. Alternatively spliced isoforms of *Pp-RS2Z37* and *Pp-RS2Z38* do not accumulate upon cycloheximide treatment.
A. Scheme of *Pp-RS2Z37* and *Pp-RS2Z38* with positions of primers used in RT-PCR analysis. Exons (E) are illustrated as black and grey boxes, indicating the coding and non-coding regions, respectively. Introns are represented as lines. Schemes are drawn to scale.
B. Wild type protonemal tissue was treated with 20 µM cycloheximide (CHX) for 4 h (+) or with water (-) for control. Two biological replicates are shown.
Analysed region is defined by the pair of exons (E*n*-E*n*) to which primers align. Arrowheads mark the fully spliced mRNA1. Bands above are the AS variants (See Figure 4.6, 4.7). *Actin* was used as a loading control.

4.3.7. Alternative splicing of *Pp-RS27, Pp-RS2Z37* and *Pp-RS2Z38* is regulated in a tissue-specific manner

Identification of multiple alternative splicing events in *Pp-RS27*, *Pp-RS2Z37* and *Pp-RS2Z38* raised the question, if they are regulated in different tissues.

Transcriptional activity and AS pattern of SR genes was demonstrated to be tightly regulated in tissue-specific manner both in animals (Barbosa-Morais et al., 2012; Merkin et al., 2012) and plant species, e.g. Arabidopsis, rice, maize and sorghum (Ali and Reddy, 2008; Kalyna et al., 2006, 2003, Lopato et al., 2002, 1999; Palusa et al., 2007; Simpson et al., 2008; Tillemans et al., 2005).

We were interested to see if the AS pattern and expression of the three SR genes in Physcomitrella would exhibit differences in tissues that differ greatly in cell morphology and physiology. We collected four tissue samples (Figure 4.10 A, 4.11 A). Protonemal sample (PRO) was collected from homogenate protonemal tissue grown for 7 days on the solid medium. The filamentous protonema tissue consisted of chloronema and caulonema cells. Gametophores, which are represented by leafy shoots, partly with rhizoids, were collected from 5 week old moss colonies growing on BCD-AT medium. Individual gametophores were collected and cut in half, and upper (younger) parts (UG) were analysed separately from the lower parts (LG) that contained also rhizoids. Colonies (COL) composed predominantly of gametophores with small proportion of rhizoids, but also of chloronema and caulonema, were also used in the analysis. Plants were grown at 23°C, 16-h light/8-h dark regime. RNA of the four tissue samples was subjected to RT-PCR analysis of *Pp-RS27*, *Pp-RS2Z37* and *Pp-RS2Z38*.

For *Pp-RS27*, two sets of primers were utilized, amplifying the region between E2 and E3 and the entire transcript (E1-E6). In both cases, we detected increased abundance of AS isoforms relatively to the fully spliced isoform in the gametophores, particularly in their upper parts, when compared to filamentous protonemal tissue (Figure 4.10 C, D). These AS isoforms are generated by alternative splicing in the second intron and contain PTCs (see Figure 4.2).

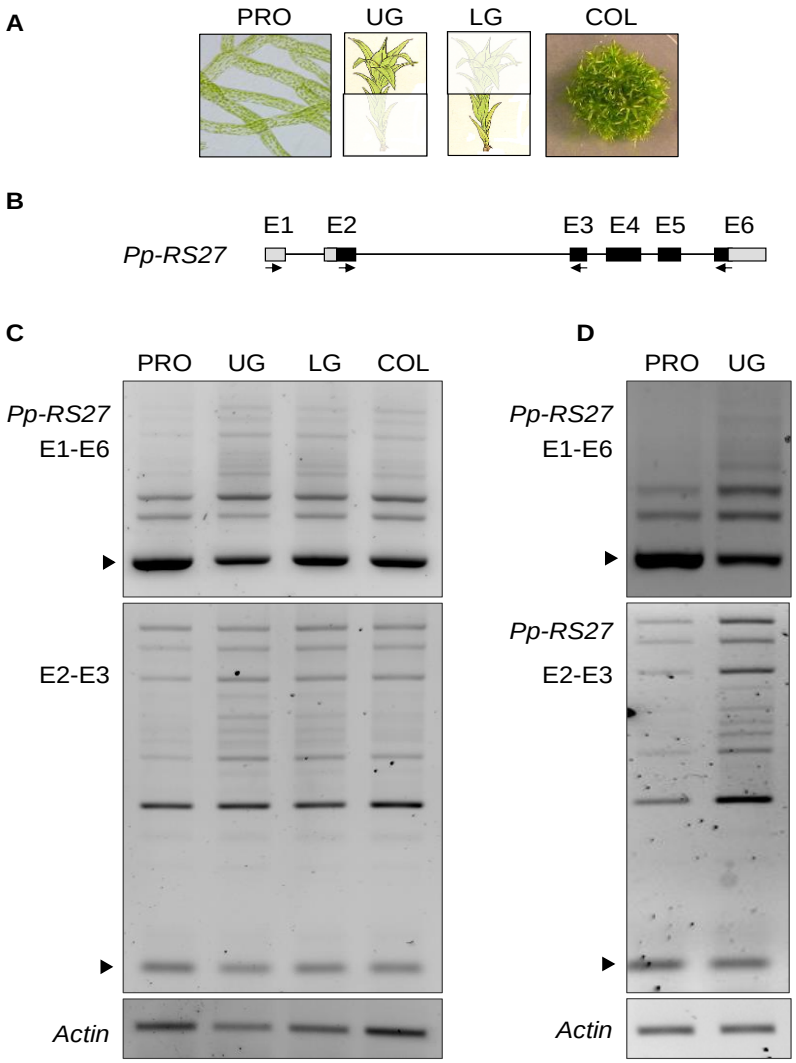


Figure 4.10 Alternative splicing pattern of *Pp-RS27* is regulated in a tissue-specific manner**.**

Pp-RS27 shows a shift towards the alternatively spliced variants in upper parts of gametophores.

A. Four tissue samples were analysed: protonemal tissue (PRO), upper (UG) and lower (LG) half of the leafy gametophore (removed parts are shaded) and a colony (COL). Protonemal tissue image adopted from Anja Martin (https://commons.wikimedia.org) and gametophore drawing adopted from W. Fitch (source - www.biolib.de).

B. *Pp-RS27* gene structure and position of the primers used for RT-PCRs. Exons (E) are depicted as black and grey boxes, indicating the coding and non-coding regions, respectively. Lines represent the introns. Scheme is drawn to scale.

C. and D. RT-PCR analyses of *Pp-RS27* in two biological replicates. AS variants in protonemal tissue, upper and lower half of the leafy gametophore and colony (C) and in protonemal tissue and upper leafy gametophore (D).

Analysed regions are defined by the pairs of exons (E*n*-E*n*) to which primers were designed. Black arrow marks the fully spliced transcript encoding Pp-RS27 protein. Bands above the fully spliced transcript are the alternatively spliced variants (see Figure 4.2).

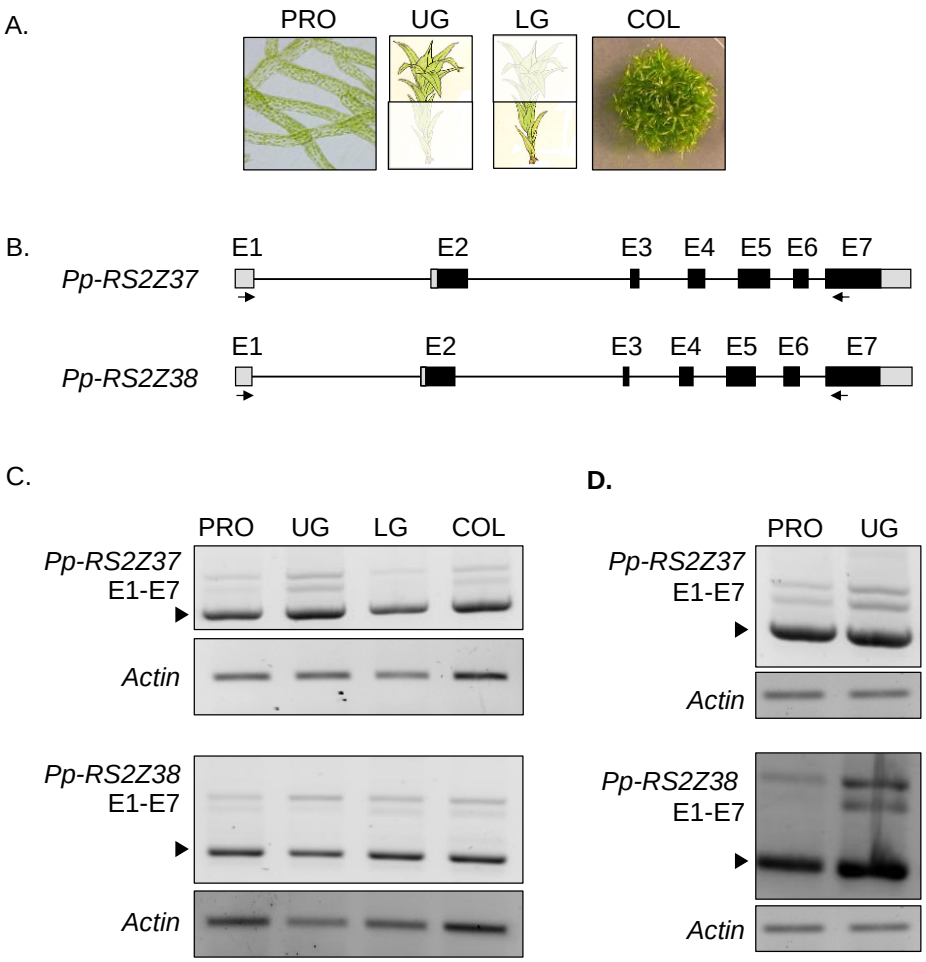


Figure 4.11. RT-PCRs analyses of AS pattern of *Pp-RS2Z37* and *Pp-RS2Z38* in different tissues. Enhanced expression of the fully as well as alternatively spliced mRNAs variants is observed for both genes in upper parts of gametophores.

A. Four tissue samples were analysed: protonemal tissue (PRO), upper (UG) and lower (LG) half of the leafy gametophore (removed parts are shaded) and colony (COL). Protonemal tissue image adopted from Anja Martin (https://commons.wikimedia.org) and gametophore drawing adopted from W. Fitch (source - www.biolib.de).

B. Schemes of the genes *Pp-RS2Z37* and *Pp-RS2Z38* with indicated primer positions. Exons (E) are depicted as black and grey boxes, indicating the coding and non-coding regions, respectively. Lines represent the introns. Schemes are drawn to scale.

C. and D. RT-PCR analyses of *Pp-RS2Z37* and *Pp-RS2Z38* in two biological replicates. AS variants in protonemal tissue, upper and lower half of the leafy gametophore and colony (C) and in protonemal tissue and upper leafy gametophore (D).

Analysed regions are defined by the pair of exons (E*n*-E*n*) to which primers align. Black arrow marks the fully spliced transcript that encode Pp-RS2Z37 or Pp-RS2Z38 protein, respectively. Bands above the fully spliced transcript are the alternatively spliced variants (see Figures 4.4, 4.5). *Actin* was used as a loading control.

RT-PCR for *Pp-RS2Z37* revealed changes in overall expression and in the alternative splicing pattern in different tissues, with an increased abundance of AS transcripts, observed in upper parts of gametophores (Figure 4.11 C and D), similarly to *Pp-RS27*. For *Pp-RS2Z38*, we did not obtain consistent results (Figure 4.11 C and D), though the same tendency of accumulation of AS transcripts is observed in the upper gametophores (Figure 4.11 D).

4.3.8. Generation of *pp-rs27* and *pp-rs2z37* knock-out lines

To elucidate the functions of RS subfamily we decided to knock-out *Pp-RS27,* which is the only gene of the RS subfamily in *P. patens,* in contrast to the multigene RS subfamily in Arabidopsis with four paralogs.

Physcomitrella RS2Z subfamily contains at least two genes, *Pp-RS2Z37* and *Pp-RS2Z38,* and we disrupted the higher expressed *Pp-RS2Z37*.

We constructed two linear DNA targeting vectors, each composed of the NPTII cassette flanked from 5' and 3' ends with gene homology regions of appx. 800 bp. Each targeting vector was transformed into Physcomitrella protoplasts, and the genes were disrupted by homologous recombination of the vectors with the corresponding intact genes. Precise integration of the vector lead to replacing part of the gene by NPTII cassette, which served also as a selection marker (G418 resistance).

We conducted two PEG-mediated transformations with each linear vector. Selection of the regenerating protoplasts yielded 15 and 8 stable G418-resistant regenerants for the *Pp-RS27* and *Pp-RS2Z37,* respectively. Regenerants were screened by PCRs to monitor the 5' and 3' integration of the NPTII and the integrations sites were sequenced in selected lines.

Genomic DNA of regenerants and wild type control was analysed by Southern blot to discern the homologous single insertion events from non-homologous and multiple ones. Multiple insertion occurs due to concatenation of the transforming DNA preceding the integration (Kamisugi et al., 2006). Restriction enzyme that excises the whole insert, including the homology regions (5'FR and 3'FR, see Figure 4.1) was used to fragment the DNA, and a probe specific to NPTII cassette was utilized to detect the desired disruption site (Figure 4.12 A and 4.13 A).

For *Pp-RS27,* six lines showed multiple insertions, two lines possible ectopic insertion(s) and five lines showed no detectable insert. Two lines were successfully

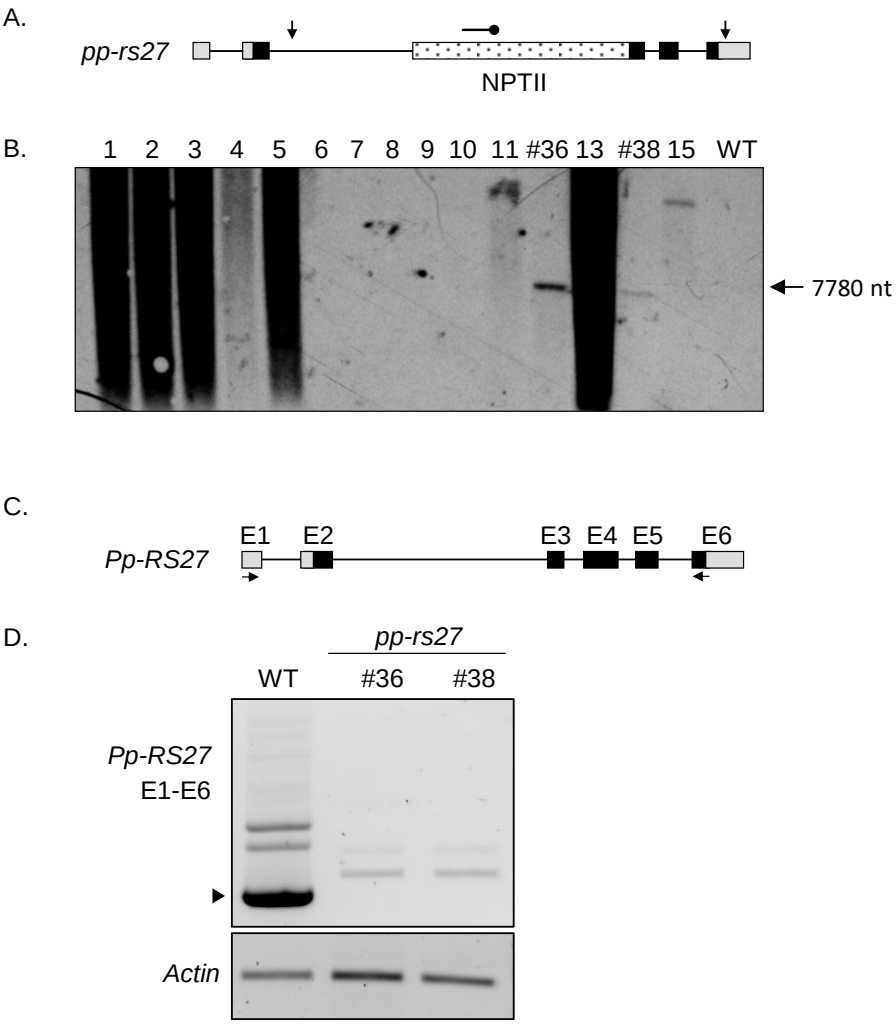


Figure 4.12. Detection and verification of *pp-rs27* mutant lines by Southern blot and RT-PCR analyses

A. Schematic representation of targeted disruption of *Pp-RS27*. Dotted box denotes NPTII cassette. Grey and black boxes represent non-coding and coding exons, respectively. Introns are shown as lines. Arrows mark the restriction enzyme cuts outside the insertion. Line with black circle represents the hybridization probe.

B Southern blot analysis of the *Pp-RS27* stable regenerants. Lines with multiple insertions (1-5, 13), lines with no detectable insert (6-10), lines with possible ectopic insert(s) (11, 15), lines with correct single insert (#36, #38), and wild type control (WT). Arrow marks the expected band of 7780 nt.

C. Scheme of the intact gene *Pp-RS27* and the primers (shown by arrows) used for RT-PCR. Grey and black boxes illustrate non-coding and coding exons, respectively. Introns are shown as lines. E*n* indicates the number of exon.

D. RNA of two potential knock-out lines and wild type control was subjected to RT-PCR to investigate the production of *Pp-RS27* transcripts. No specific transcript was detected, and two weak bands are considered unspecific. Arrowhead indicates the fully spliced transcript that codes for the reference protein. Upper bands in wild type control represent AS variants (see Figure 4.2).

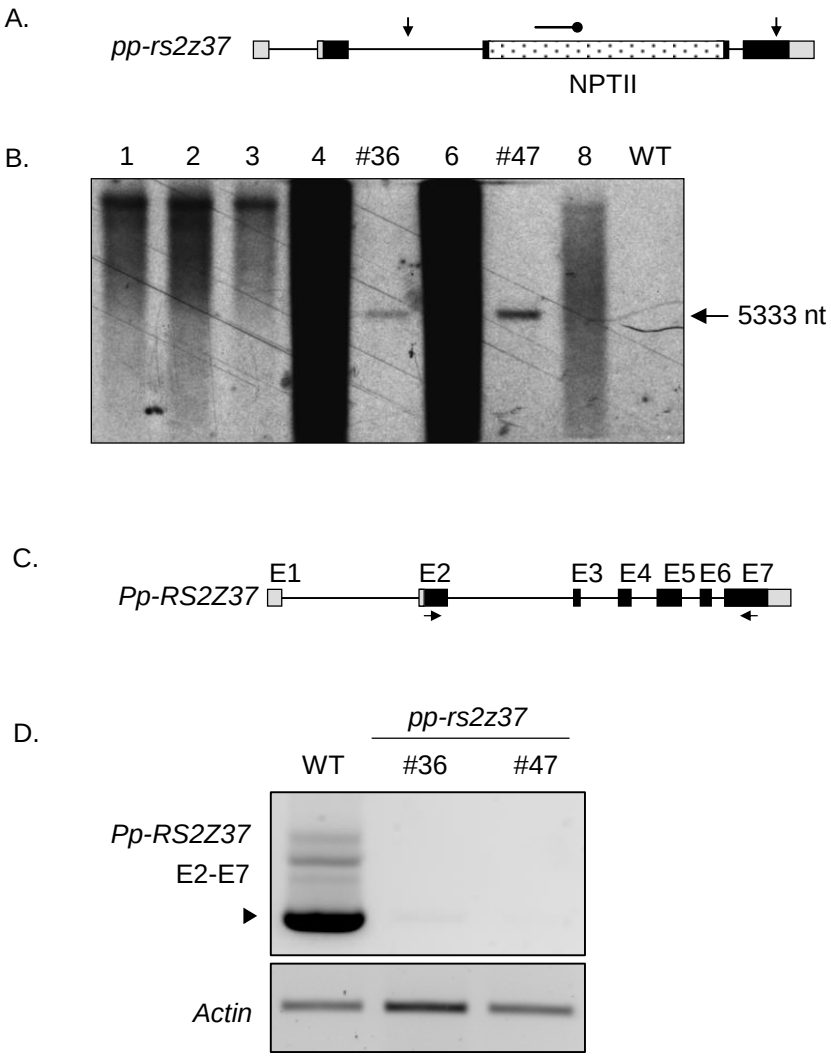


Figure 4.13. Detection and verification of *pp-rs2z37* mutant lines by Southern blot and RT-PCR analyses

A. Schematic illustration of *Pp-RS2Z37* disruption by NPTII. Grey and black boxes represent non-coding and coding exon, respectively. Dotted box indicates NPTII cassette. Introns are drawn as lines. Arrows mark the RE cuts outside the insertion. Line with black circle denotes the probe.

B. Southern blot of the *Pp-RS2Z37* stable regenerants. Lines with multiple insertions (4, 6), lines with likely ectopic insert(s) (1-3, 8), targeted lines with correct single insert (#36, #47), wild type control (WT). Arrow marks the expected band of 5333 nt.

C. Scheme of the intact gene *Pp-RS2Z37* and the primers (arrows) utilized for RT-PCR. Grey and black boxes show non-coding and coding exons, respectively. Introns are represented by lines. E*n* indicates the number of exon.

D. RNA of two potential knock-out lines and wild type control was subjected to RT-PCR to investigate the production of *Pp-RS2Z37* transcripts. RT-PCR detected no transcript of *Pp-RS2Z37* produced in the two putative *pp-rs2z37* lines. Arrowhead indicates the mRNA1 that codes for the reference protein. Upper bands are AS variants (see Figure 4.4).

confirmed as *Pp-RS27* mutants with a single and correct NPTII insertion into the *Pp-RS27* gene (Figure 4.12 A).

For *Pp-RS2Z37,* two transformants contained multiple insertions and four lines displayed ectopic insertion(s). Two lines proved to be correct single copy knock-outs of *Pp-RS2Z37* (Figure 4.13 A).

RNA of the confirmed lines was subjected to RT-PCRs that ruled out any production of the wild type transcripts in both of the *pp-rs27* (from now labelled #36 and #38) as well as in both *pp-rs2z37* (from now labelled #36 and #47) analysed lines, compared to wild type. (Figure 4.12 B and 4.13 B).

4.3.9. *Pp-RS27* and *Pp-RS2Z37* disruptions do not affect alternative splicing of other SR genes

We observed cross-regulation of Arabidopsis SR homologues when overexpression of *At-RS2Z33* dramatically affected AS of *At-RS31* (Kalyna et al., 2006; Simpson et al., 2008) and overexpression of *At-RS31* changed AS splicing pattern of *At-SR30* (Chapter 3). We were interested to see whether disruption of *Pp-RS27* and *Pp-RS2Z37* in Physcomitrella will have an effect on the expression and AS of other SR genes.

We used RNA from the protonemal tissue of the mutant lines *pp-rs27#36, pp-rs27#38, pp-rs2z37#36* and *pp-rs2z37#47* in the RT-PCR to check the expression of *Pp-RS27, Pp-RS2Z37* and *Pp-RS2Z38* (Figure 4.14, 4.16). At the tested conditions, we did not detect any significant effect of *Pp-RS27* and *Pp-RS2Z37* knock-outs on alternative splicing or expression of the tested SR genes.

4.3.10. Investigating the regulation and functions of *Pp-RS27* and *Pp-RS2Z37* in responses to developmental and stress signals

P. patens, as an ancient lineage of land plants, one of the first that moved from aquatic to terrestrial environment, was inevitably forced to acquire mechanisms for coping with abiotic stresses, such as drought, extreme temperature and increased radiation. Stress tolerance, as well as signalling pathways of important growth and developmental regulators, such as ABA, auxin and cytokinin, have developed together with appearance of *Bryophytes*

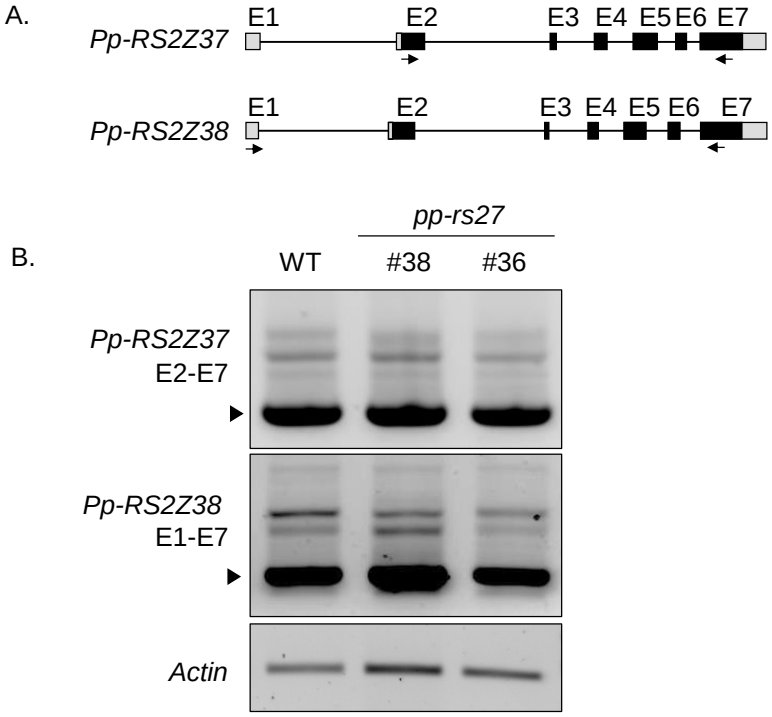


Figure 4.14. Mutation of *Pp-RS27* does not significantly affect the expression and AS pattern of *Pp-RS2Z37* and *Pp-RS2Z38*

A. Schemes of the SR genes with positions of primers used in RT-PCRs. Black and grey boxes indicate coding and non-coding exons (E), respectively. Lines represent the introns. Gene schemes are to scale.

B. RT-PCR of protonemal RNA was used to analyse transcripts of the *Pp-RS2Z37* and *Pp-RS2Z38* genes in mutant lines *pp-rs27#36* and *pp-rs27#38*.

Amplified region is defined by the pair of exons (E*n*-E*n*) to which primers align. Black arrowhead marks the fully spliced transcript that encodes a reference protein. Upper bands are the alternatively spliced variants (see Figures 4.4, 4.5). *Actin* was used as the loading control.

and are considered highly conserved up to the vascular plants (Prigge and Bezanilla, 2010; Rensing et al., 2008a).

Importance of alternative splicing in managing growth, development and stress responses in plants was established by many experimental studies (Filichkin et al., 2015; Staiger and Brown, 2013; Carvalho et al., 2013; Duque, 2011). Effects of At-RS31 overexpression and mutation, described in this thesis, also suggest that regulation of alternative splicing is engaged in control of plant growth and development as well as in mechanisms of stress tolerance and DNA damage response. Moreover, alternative splicing

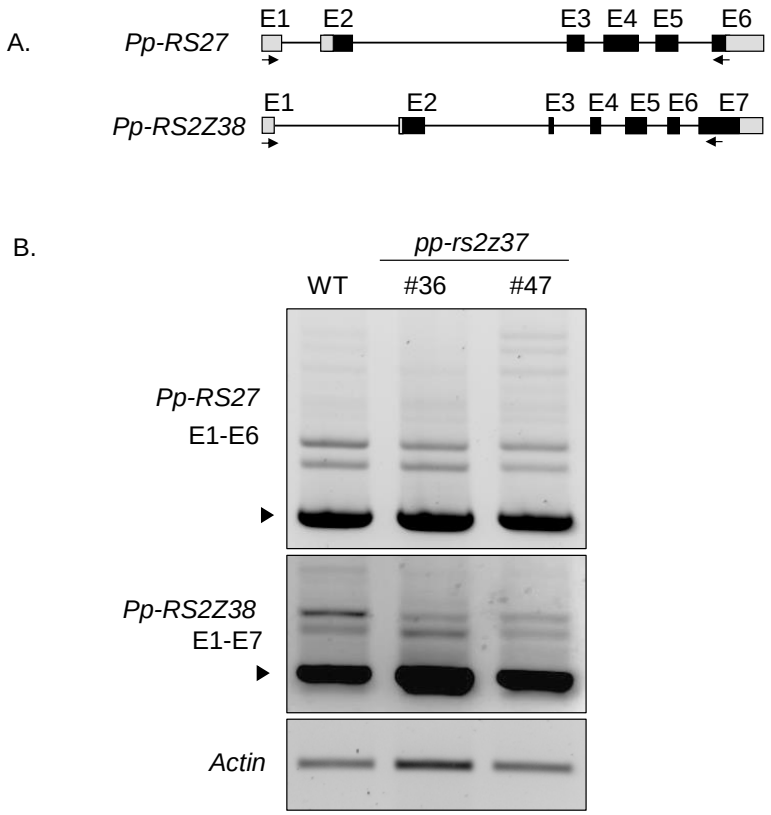


Figure 4.15. Disruption in *pp-rs2z37* does not significantly affect the expression and alternative splicing of *Pp-RS27* and *Pp-RS2Z38*.
A. Schemes of the SR genes with positions of primers used in RT-PCRs. Black and grey boxes indicate coding and non-coding exons (E), respectively. Lines represent the introns. Gene schemes are to scale.
B. RT-PCR of protonemal RNA was performed to analyse transcripts of the genes *Pp-RS27* and *Pp-RS2Z38* in mutant lines *pp-rs2z37#36* and *pp-rs2z37#47*.
Amplified region is defined by the pair of exons (E*n*-E*n*) to which primers align. Black arrowhead marks the fully spliced transcript that encodes a reference protein. Upper bands are the alternatively spliced variants (see Figures 4.2, 4.5). *Actin* was used as the loading control.

of *At-RS31* was modulated by different light regimes (Petrillo et al., 2014). However, interplay of SR proteins and alternative splicing with development and abiotic stress adaptation in *Bryophytes* is poorly investigated.

Due to haploid dominant phase of the moss life cycle, recessive phenotypes can be easily detected by changed colony morphology. Physcomitrella colony forms after few weeks long cultivation of small pieces of freshly grown protonemal tissue, inoculated onto BCD-AT solid medium. Tissue regenerates and differentiates into both types of protonemal

filaments (chloronema and caulonema), eventually forming a colony with gametophores and rhizoids (Cove, 2005; Cove et al., 2006).

To test the possible roles for Physcomitrella SR genes in stress response and developmental pathways, we examined phenotypes of wild type, two distinct *pp-rs27* and two distinct *pp-rs2z37* mutant lines for changes of the colony morphology and growth as a reaction to phytohormones and stress treatments (Figure 4.16), and to DNA damage induced by UV-C irradiation.

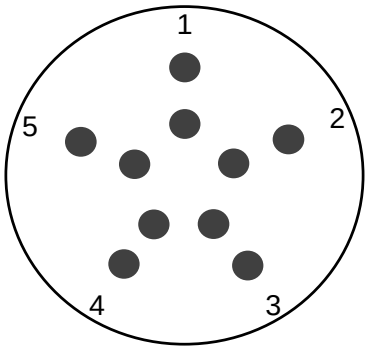


Figure 4.16. Arrangements of the spot inocula on the agar plate to test the possible roles for Physcomitrella SR genes in stress response.
Each line is represented by two colonies. 1 - Wild type; 2 - pp-rs27#36; 3 - pp-rs27#38; 4 - pp-rs2z37#36; 5 - pp-rs2z37#47. Numbers after hash mark indicate the number of the independent mutant line.

We also performed RT-PCR analysis of wild type protonemal tissue subjected to transient abiotic stress exposure, light/dark transitions and DNA damaging UV-C irradiation to detect changes in the gene expression and alternative splicing.

Knock-out lines *pp-rs27#36, pp-rs27#38, pp-rs2z37#36* and *pp-rs2z37#47* did not show any obvious changes of protonemata growth or appearance, shape and growth of the colony when compared to wild type at the standard conditions (on BCD-AT agar plates as presented in the control experiments in this section, and on BCD agar plates lacking ammonium tartrate, not shown).

4.3.10.1. Auxin

Auxins are plant hormones regulating numerous important steps of plant morphogenesis, development and growth, as well as stress response. At the cellular level, auxin regulates morphogenesis by governing cell division, elongation and differentiation (Bargmann et al., 2013; Decker et al., 2006). Auxins play important role in response to osmotic stress, such as water shortage, high salinity and in ROS metabolism (Naser and Shani, 2016; Tognetti et al., 2012). In Physcomitrella, auxin was studied mostly for its importance in development and promotion of the cell differentiation from chloronema to caulonema, rhizoid formation and elongation of gametophore leaflets (Ashton et al., 1979; Cove et al., 2006; Decker et al., 2006; Prigge and Bezanilla, 2010).

Interestingly, overexpression of *At-RS2Z33*, one of the Arabidopsis orthologues of *Pp-RS2Z37* and *Pp-RS2Z38* was changing the expression activity of an auxin-responsive promoter in the tissue-dependent manner, and plants showed auxin-dependent phenotypes (Kalyna et al., 2003). Therefore, we aimed to test a link between auxin and Physcomitrella SR genes.

We cultivated colonies of Physcomitrella *pp-rs27* and *pp-rs2z37* mutant lines and wild type control on media with various concentrations of auxin, 2,4-dichlorophenoxyacetic acid (Figure 4.17). Colonies showed auxin-induced changes of protonemal filaments at the concentration of 1 µM auxin, and the effect was enhanced with increased concentration of auxin to 10 µM. However, mutation of any of the SR genes did not affect the colony phenotype in response to the applied auxin (Figure 4.17).

Figure 4.17 Colonies of wild type, *pp-rs27* and *pp-rs2z37* mutants were grown on media with auxin.
Wild type (WT) and mutant protonemal tissue forming colonies, cultivated for 5 weeks on BCD-AT plates, supplemented either with various concentrations of auxin 2,4-dichlorophenoxyacetic acid (2,4-D) or with 0.005% DMSO as a control. Left column shows whole plate with all tested lines (see Figure 4.16 for the reference). Three other columns represent close-up views of the representative colonies. For each concentration: upper row - images with light background, lower row - images with dark background.

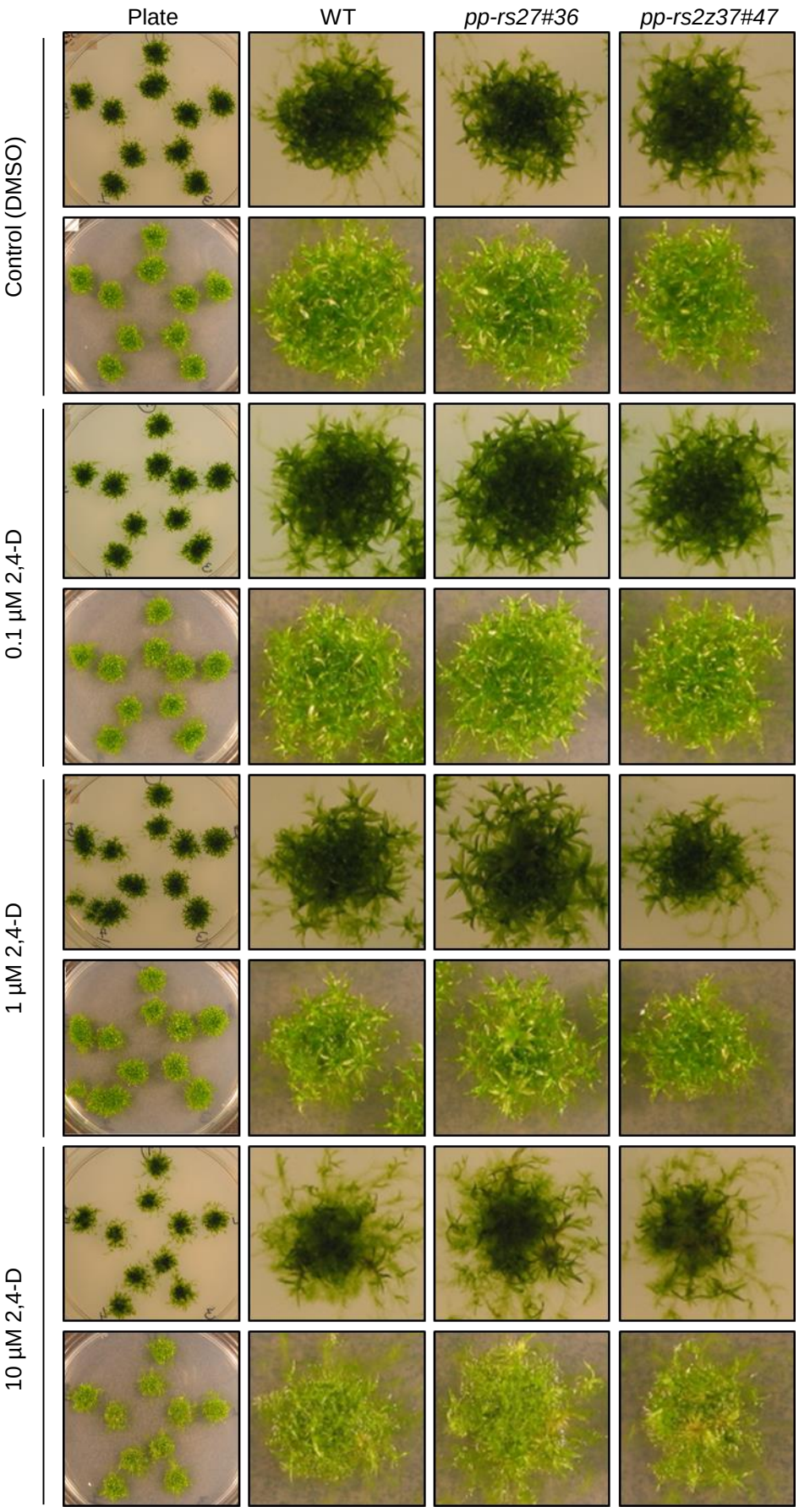


Plate
WT
pp-rs27#36
pp-rs2z37#47
Control (DMSO)
0.1 µM 2,4-D
1 µM 2,4-D
10 µM 2,4-D

4.3.10.2. Cytokinin

Cytokinins are phytohormones with plethora of biological functions, mostly known for stimulation of cell proliferation and differentiation, in which they act in relation with auxins, in a ratio-dependent manner. In association with ABA, they were shown to modulate abiotic stress response to high salinity, desiccation and cold, and also evince activity in resistance to pathogens. Cytokinins are able to delay leaf senescence in plants by controlling the cell cycle, and this process is carried out in an interplay with sugars (Hutchison and Kieber, 2002; Hwang et al., 2012; Wingler et al., 1998).

Components of the cytokinin signal transduction pathway are present in Physcomitrella genome confirming an ancient origin of cytokinin regulation mechanisms and suggesting high similarity in biological functions of cytokinin (Rensing et al., 2008b; Schwartzenberg, 2006). As in vascular plants, cytokinin in Physcomitrella promotes bud formation from protonemal filaments that leads eventually to gametophore development (Ashton et al., 1979; Cove et al., 2006; Decker et al., 2006).

In our experimental setup, colonies of *pp-rs27* and *pp-rs2z37* mutant lines and wild type control were cultivated for five weeks on media with different concentrations of cytokinin 6-benzylaminopurine (Figure 4.18). We observed increased bud production at the lowest cytokinin concentration (0.1 µM) and callus formation at the levels of 1 and 10 µM which is a sign of high concentration of cytokinins (Decker et al., 2006). All the observed changes were in the similar fashion for all tested lines, indicating that mutations of *Pp-RS27* or *Pp-RS2S37* do not affect the colony morphology and growth in the used experimental arrangement and do not interfere with cytokinin responses (Figure 4.18).

Figure 4.18. Wild type, *pp-rs27#36, pp-rs27#38, pp-rs2z37#36* and *pp-rs2z37#47* mutant colonies grown in the presence of cytokinin.
Wild type (WT) and mutant colonies were cultivated 5 weeks on BCD-AT plate with various concentrations of cytokinin 6-Benzylaminopurine (6-BAP) or with 0.005% DMSO as a control. Left column shows whole plate with all tested lines (see Figure 4.16 for the reference). Three other columns show close-up views of the representative colonies. For 0.1 and 1 µM 6-BAP and DMSO: upper row - images with light background, lower row - images with dark background. For 10 µM 6-BAP, only images with dark background are shown.

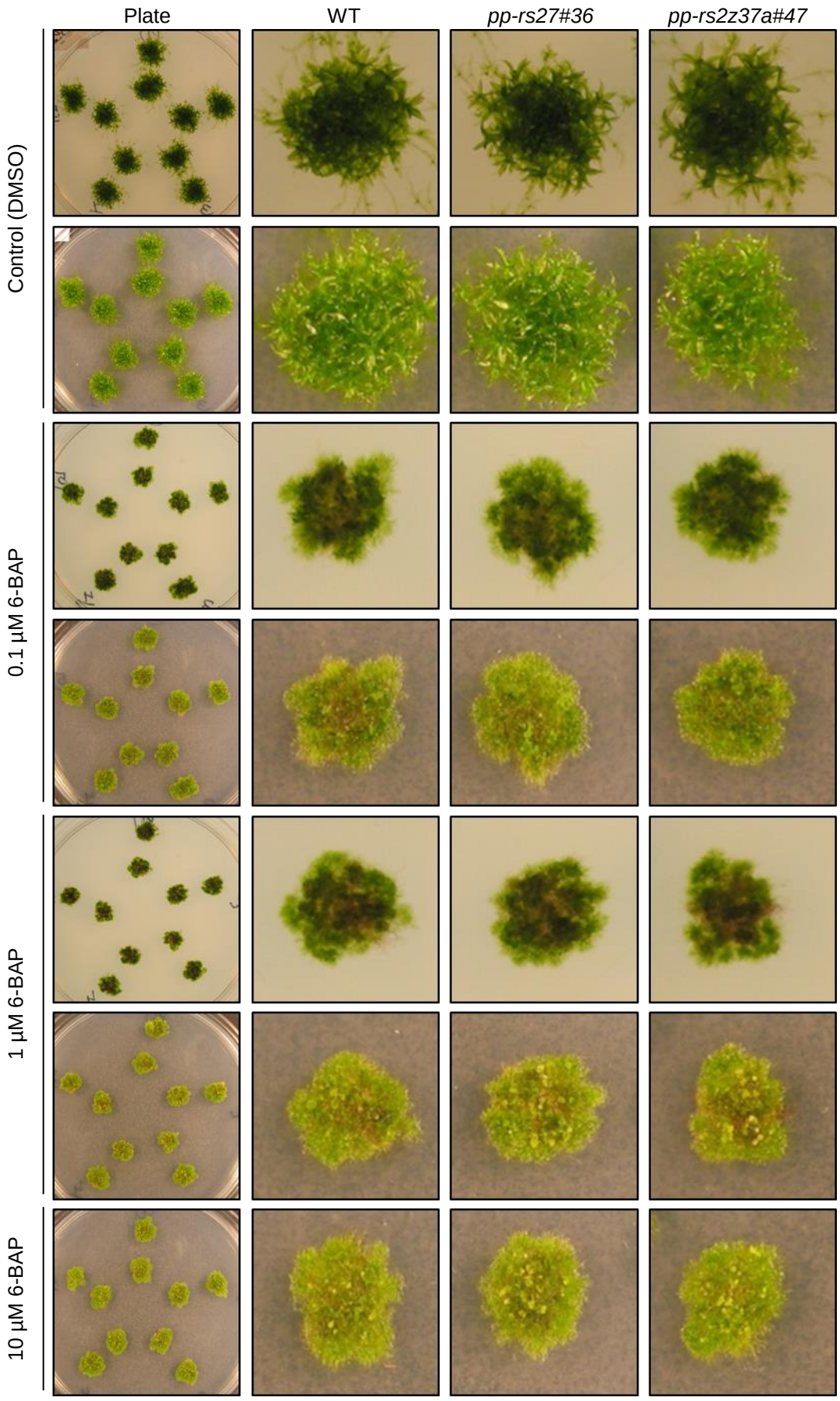


Plate
WT
pp-rs27#36
pp-rs2z37a#47
Control (DMSO)
0.1 µM 6-BAP
1 µM 6-BAP
10 µM 6-BAP

4.3.10.3. Methyl viologen

Methyl viologen (MV) is a redox-active compound that is able to generate ROS (Bus and Gibson, 1984). We used it to induce oxidative stress in Physcomitrella.

We detected changes in the phenotype of Arabidopsis plants with modified expression levels of At-RS31 in the presence of methyl viologen (see Chapter 3). We tested Physcomitrella sensitivity to the different concentrations of methyl viologen on the colonies of *pp-rs27*, *pp-rs2z37* and wild type control lines. Although methyl viologen treatment clearly affected colony growth, wild type and four mutants showed identical phenotype under the applied conditions (Figure 4.19).

Figure 4.19. Wild type, *pp-rs27#36*, *pp-rs27#38*, *pp-rs2z37#36* and *pp-rs2z37#47* mutant colonies grown in presence of methyl viologen.
Wild type (WT) and mutant protonemal tissue was cultivated for 5 weeks (23°C, 16-h light/8-h dark) on BCD-AT plates supplemented with indicated concentrations of methyl viologen (MV) or with 0.005% DMSO as a control. Left column shows plates with all tested lines (see Figure 4.16 for the reference). Three columns to the right represent close-up views of the representative colonies. For concentrations 0.1 and 1 µM and DMSO: upper row - images with light background, lower row - images with dark background. For 10 µM, only images with dark background are shown.

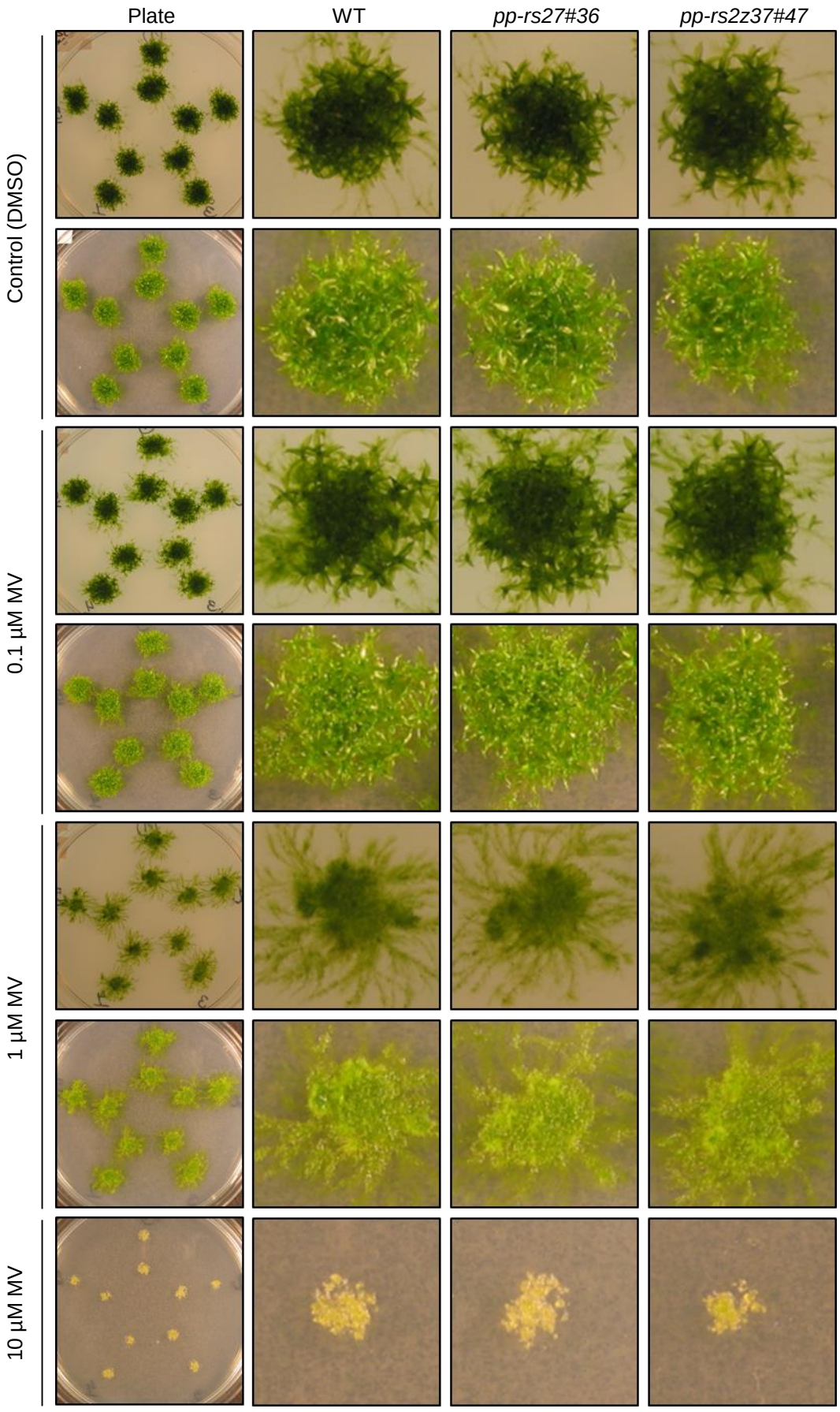


Plate
WT
pp-rs27#36
pp-rs2z37#47
Control (DMSO)
0.1 µM MV
1 µM MV
10 µM MV

4.3.10.4. Abscisic acid

Increasing evidence links plant splicing factors and other RNA-binding proteins to ABA-mediated stress responses (Carvalho et al., 2010; Cruz et al., 2014; Duque, 2011; Wong et al., 2017).

Our experiments with Arabidopsis suggest that splicing factor At-RS31 is involved in ABA triggered regulation of stress response (Chapter 3). Moreover, in At-RS31 overexpression background, we detected changes in alternative splicing of *BTR1* gene (Chapter 3), which encodes protein phosphorylated by ABA-activated protein kinase SnRK2.6 (Wang et al., 2013).

Plant hormone ABA is present also in early land plants, such as *P. patens*, and is a part of conserved stress-sensing and signalling pathways, likewise in angiosperms (Takezawa et al., 2011).

We tested growth of moss mutants *pp-rs27* and *pp-rs2z37* and wild type for five weeks on BCD-AT media supplemented with increasing amounts of ABA. We observed similar ABA-dependent changes to the colony shape regardless of the genotype (Figure 4.20), suggesting that shutting down the expression of *Pp-RS27* and *Pp-RS2Z37* does not interfere with ABA responses, at least under tested conditions.

Figure 4.20. Colonies of wild type, *pp-rs27#36, pp-rs27#38*, *pp-rs2z37#36* and *pp-rs2z37#47* mutants upon ABA treatment.
Wild type (WT) and mutant protonemal tissue was cultivated for 5 weeks *(*23°C, 16-h light/8-h dark) on BCD-AT plates, supplemented with indicated concentrations of abscisic acid (ABA) or with 0.005% DMSO as a control. Left column shows plates with all tested lines (see Figure 4.16 for the reference). Three columns to the right represent close-up views of the representative colonies. For each concentration: upper row - images with light background, lower row - images with dark background.

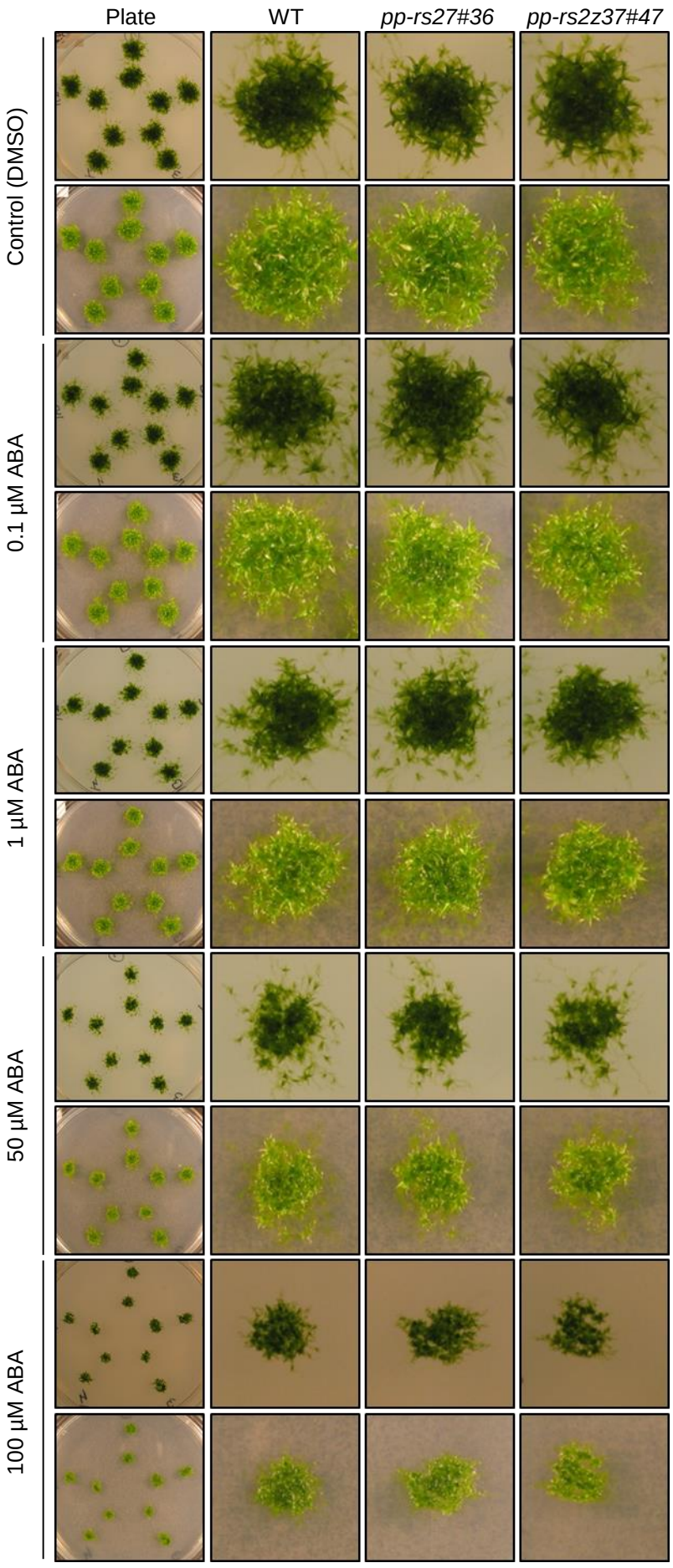


Plate
WT
pp-rs27#36
pp-rs2z37#47
Control (DMSO)
0.1 µM ABA
1 µM ABA
50 µM ABA
100 µM ABA

4.3.10.5. Sodium chloride

In Arabidopsis, we observed that overexpression of At-RS31 negatively affected growth of the seedlings in high salt environment and At-RS31 mutant plants did not differ from the wild type. We therefore checked whether *P. patens* SR mutant lines display any phenotypic changes compared to wild type in response to increased salinity. We cultivated Physcomitrella colonies for five weeks in the presence of NaCl (0, 1, 100 and 150 mM) at the standard conditions. We observed that high concentration of NaCl imposed similar changes to the colony phenotype, independently from the genotype (Figure 4.21).

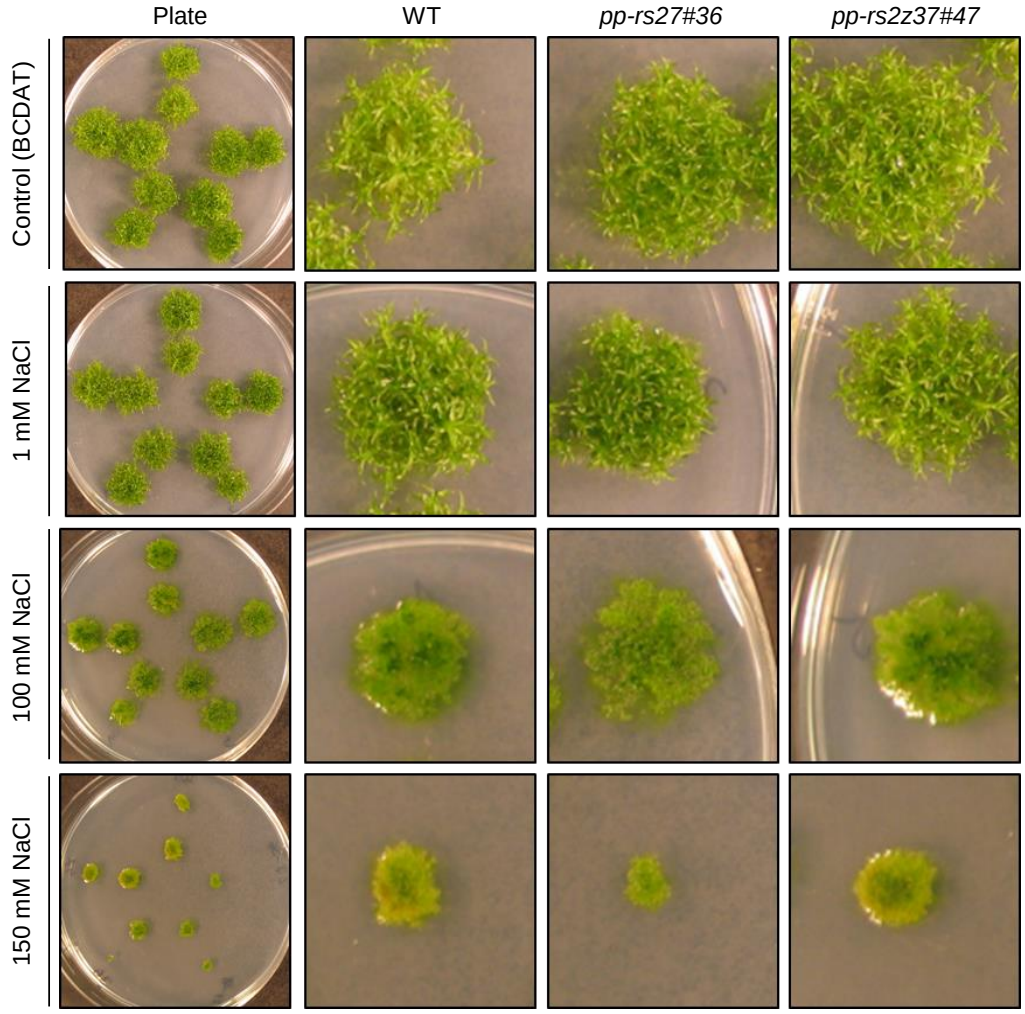


Figure 4.21. Colonies of wild type and *pp-rs27#36*, *pp-rs27#38*, *pp-rs2z37#36* and *pp-rs2z37#47* mutants upon salinity stress induced by NaCl.
Wild type (WT) and mutant protonemal tissue was cultivated for 5 weeks (23°C, 16-h light/8-h dark) on BCD-AT plates, supplemented with indicated concentrations of NaCl. Left column shows whole plates with all tested lines (see Figure 4.16 for the reference). Three columns to the right represent close-up views of the representative colonies.

4.3.10.6. Analyses of alternative splicing changes upon drought, ABA and oxidative stress treatments

We applied drought (for 40 min), ABA, MV, and DMSO as a control (for 2 and 6h) in a form of short-term stress on Physcomitrella protonemal tissue grown for 5 days at 23°C, 16-h light/8-h dark. We performed RT-PCR analyses of *Pp-RS27, Pp-RS2Z37* and *Pp-RS2Z38* to monitor abundance of AS transcripts upon stress. Late embryogenesis abundant protein 2 (*LEA-2*, PP1S442_22V6) was used as a positive control for ABA and drought treatments. Drought exposure should increase endogenous levels of ABA, thereby promoting expression of abiotic stress related gene *LEA-2* (Cuming et al., 2007; Shinde et al., 2012).

LEA-2 did not exhibit changes in expression upon drought treatment when compared to the control, which might indicate insufficient desiccation and a need to reconsider the experimental approach (Figure 4.22, 4.23). Expression of *LEA-2* upon exogenous ABA application however, was increasing with the time of exposure (2 vs. 6 hours).

AS pattern or overall expression of *Pp-RS27* was not influenced with applied stress treatments (Figure 4.22). Interestingly, amount of the second most abundant *Pp-RS27* splice variant, mRNA3, decreases in majority of 6 hours samples when compared to those collected after 2 hours, while abundance of mRNA1 increases or remains the same (Figure 4.22), suggesting diurnal regulation of *Pp-RS27* splicing pattern. This possibility could be tested in the extended time-course by quantitative RT-PCR of these isoforms.

RT-PCR of the *Pp-RS2Z37* showed no changes in alternative splicing or gene expression upon any of the treatments. In *Pp-RS2Z38*, an overall up-regulation in later collected control (6h) and MV samples (6h) is present, which might be indicative of a diurnal regulation of this gene (Figure 4.23). As in case of *Pp-RS27*, extended time-course and quantitative RT-PCR can be performed to test this.

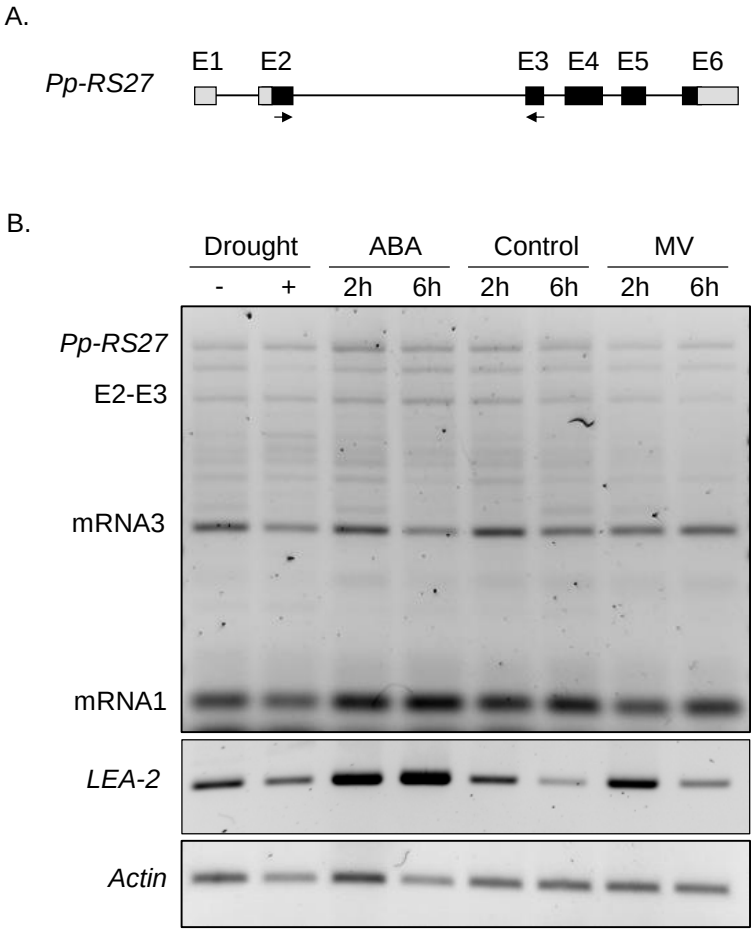


Figure 4.22. Analysis of abiotic stress impact on the alternative splicing pattern and expression level of *Pp-RS27* in wild type protonemal tissue.
A. Schematic representation of the gene *Pp-RS27* with positions of primers used in RT-PCR analysis. Exons (E) are illustrated as black and grey boxes, indicating the coding and non-coding regions, respectively. Lines represent the introns. Scheme is drawn to scale.
B. RT-PCR analysis of *Pp-RS27* splice variants and of the expression of abiotic stress related gene *LEA-2*. Wild type protonemal tissue was subjected to drought for 40 minutes, or treated with 10 µM abscisic acid (ABA) or with 10 µM methyl viologen (MV) for 2 and 6 hours. Untreated (-) or mock-treated samples were used as controls for drought and ABA/MV, respectively. Region analysed in *Pp-RS27* is defined by the pair of exons (E*n*-E*n*) to which primers align. Fully spliced transcript, encoding the reference protein (mRNA1), and alternatively spliced mRNA3, are indicated. Other bands above the mRNA1 are the alternatively spliced variants (see Figure 4.2). *Actin* was used as a loading control.

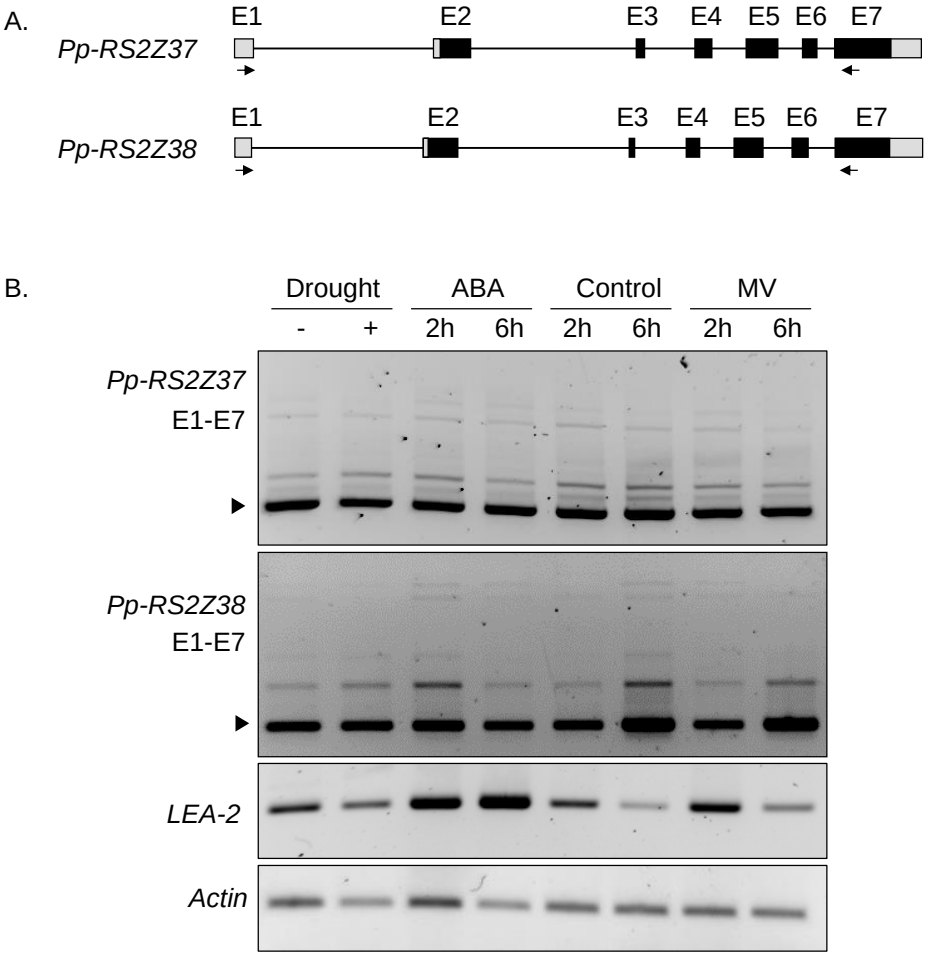


Figure 4.23. Analysis of abiotic stress treatment impact on the AS pattern and expression levels of *Pp-RS2Z37*, *Pp-RS2Z38* in wild type protonemal tissue.

A. Schemes of the genes *Pp-RS2Z37* and *Pp-RS2Z38* with indicated primer positions used in RT-PCR analysis. Exons (E) are depicted as black and grey boxes, indicating the coding and non-coding regions, respectively. Lines represent the introns. Schemes are drawn to scale.

B. RT-PCR analysis of *Pp-RS2Z37* and *Pp-RS2Z38* splice variants and the expression of *LEA-2*. Wild type protonemal tissue (23°C, 16-h light/8-h dark) was subjected to drought for 40 min, or treated with 10 µM abscisic acid (ABA) or with 10 µM methyl viologen (MV) for 2 and 6 hours. Untreated (-) or mock-treated samples were used as controls for drought and ABA/MV, respectively. Analysed region is defined by the pair of exons (E*n*-E*n*) to which primers align. Black arrowheads mark fully spliced transcripts. Bands above the fully spliced transcripts are the alternatively spliced variants (See Figure 4.4, 4.5). *Actin* was used as a loading control.

4.3.10.7. Disruption of *Pp-RS27* increases sensitivity to glucose excess

In Arabidopsis, we observed that an early senescence was induced by elevated levels of At-RS31 by overexpression, in contrast to the mutant that showed phenotype similar to wild type (Chapter 3). By growth in the presence of glucose, Arabidopsis plants of all three genetic backgrounds were negatively affected by earlier onset of ageing.

High-energy conditions, generated e.g. by exogenous glucose, promote growth of Physcomitrella juvenile tissues (caulonema), and the transition to sexual reproduction is diminished (Thelander et al., 2005). We tested growth of the moss *Pp-RS27* and *Pp-RS2Z37* mutants on glucose-supplemented medium to investigate a possible connection of these SR genes to the glucose metabolism or ageing. To this end, wild type and mutant lines were grown on media supplemented with 1, 100 and 300 mM glucose.

Lowest concentration of glucose (1 mM) did not affect the phenotype of the colonies (Figure 4.24). Addition of 100 mM glucose stimulated growth of the tissue with less chlorophyll and reduced the development of gametophores. This effect was enhanced in both independent *Pp-RS27* knock-out lines *pp-rs27#36 and pp-rs27#38*, in comparison to the wild type.

To exclude the possibility that effect of glucose on the *pp-rs27* phenotype was caused by osmotic stress, the colonies of wild type and all four mutant lines were grown in the presence of mannitol (1, 100 and 300 mM) (Figure 4.25). We did not detect differences among the different backgrounds with any concentration (1, 100, 300 mM) of mannitol. Higher concentrations of mannitol induced, similar morphological changes in all genotypes, but different from those induced by glucose. We assume that the effect of glucose on phenotype was not a result of an osmotic stress.

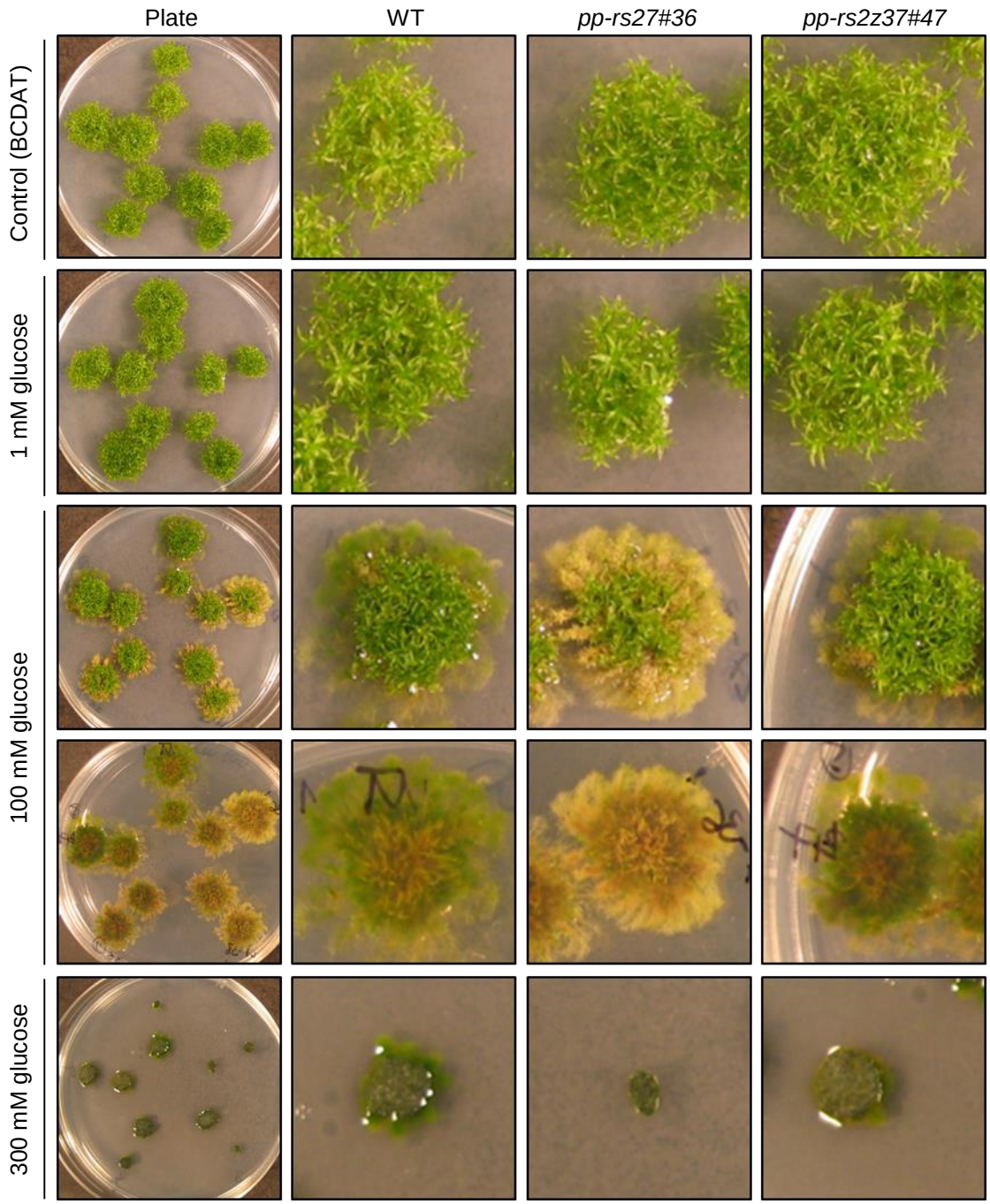


Figure 4.24. Disruption of *Pp-RS27* influences the colony phenotype at high concentration of glucose.

Colonies of wild type (WT), *pp-rs27#36*, *pp-rs27#38*, *pp-rs2z37#36* and *pp-rs2z37#47* mutants were cultivated for 5 weeks (23°C, 16-h light/8-h dark) on BCD-AT plates, supplemented with indicated concentrations of glucose. Left column shows whole plates with all tested lines (see Figure 4.2 for the reference). Three columns to the right represent close-up views of the representative colonies. For 100 mM glucose treatment, colonies are additionally shown from the bottom (lower row of images).

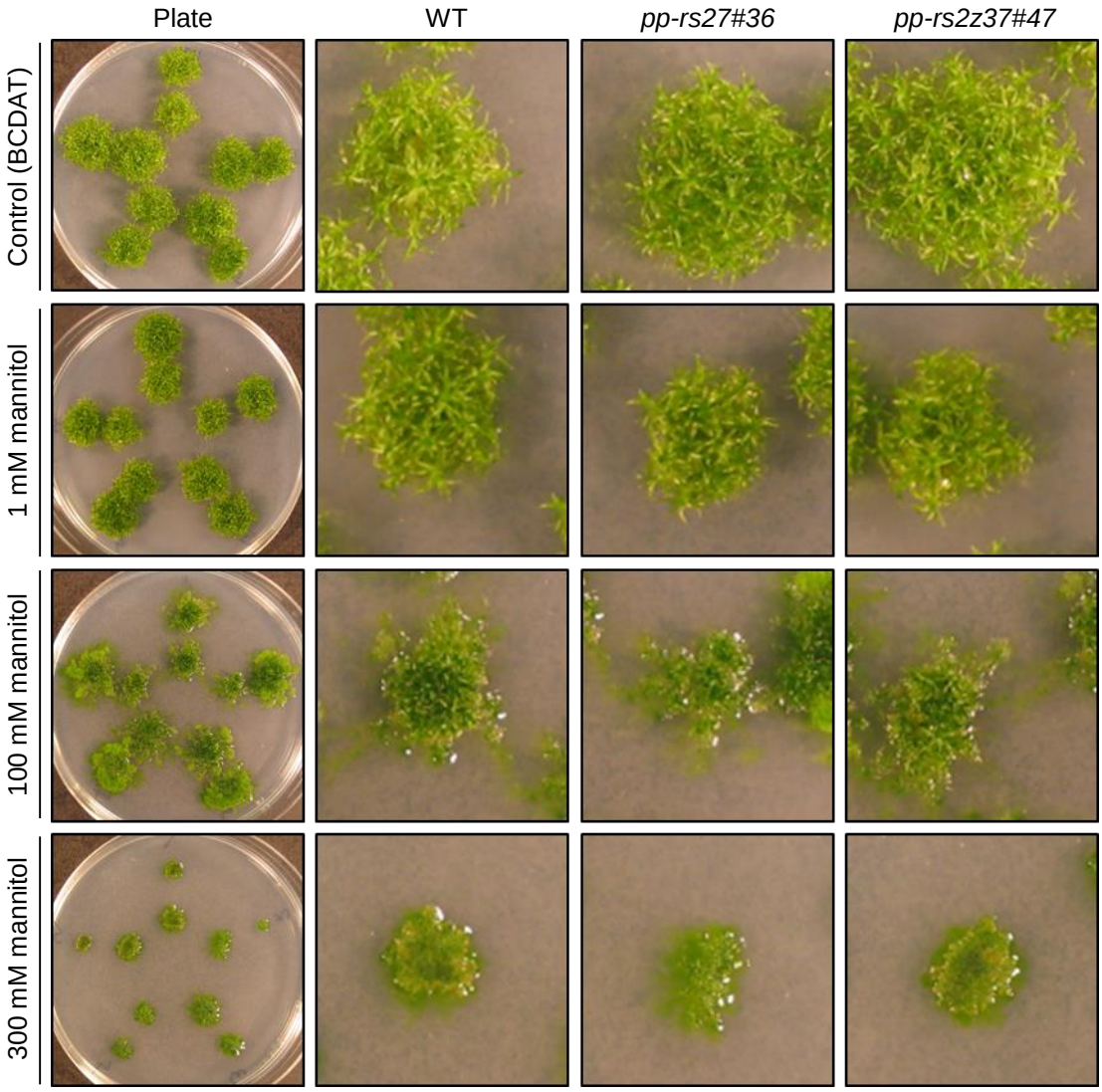


Figure 4.25. Wild type, *pp-rs27#36, pp-rs27#38, pp-rs2z37#36* and *pp-rs2z37#47* mutant colonies grown in the presence of mannitol.
Colonies of wild type (WT) and mutants were cultivated for 5 weeks (23°C, 16-h light/8-h dark) on BCD-AT plates, supplemented with indicated concentrations of mannitol. Left column shows whole plates with all tested lines (see Figure 4.2 for the reference). Three columns to the right represent close-up views of the representative colonies.

4.3.10.8. Evolutionary conserved regulation of *Pp-RS27* alternative splicing in response to light and dark

Light is one of the most important environmental factors that trigger plenty of signalling pathways and biological processes during plant development and growth.

Our study using the HR RT-PCR screen identified 9 genes with differential AS in light- and dark-grown seedlings, among them RS2Z subfamily gene *At-RS2Z33* showed shift to the protein-coding splice isoform in the light, while PTC-containing isoform was downregulated (see Chapter 2) (Simpson et al., 2008). Interestingly, light was also able to modulate AS of RS subfamily gene *At-RS31* and increase abundance of protein-coding mRNA1, while in the dark AS transcripts were more pronounced (Petrillo et al., 2014).

To examine whether Physcomitrella gene *Pp-RS27*, similarly to its Arabidopsis orthologue *At-RS31*, displays any change in alternative splicing in response to light/dark transitions, we kept wild type protonemal tissue for four hours in dark and checked the expression of *Pp-RS27* by RT-PCR (Figure 4.26). The abundance of mRNA3, the most prominent AS transcript, was enhanced upon dark treatment in contrast to incubation in light. In contrast, the protein-coding mRNA1 was the most abundant transcript in light. Since all alternative isoforms introduce PTCs (Figure 4.2 C), we can assume that mRNA1 and hence a functional protein are down-regulated in the absence of light. This alternative splicing control of abundance of RS subfamily proteins, Pp-RS27 and At-RS31, in response to light and dark is highly conserved between Physcomitrella and Arabidopsis, two species diverged 450 million years ago.

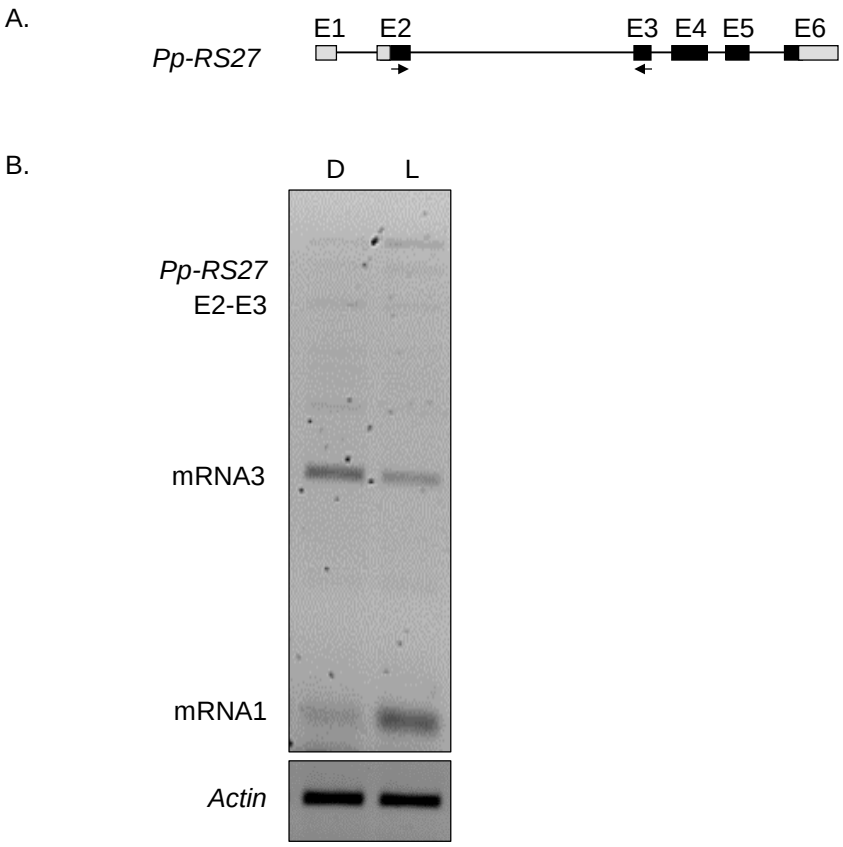


Figure 4.26. Alternative splicing of *Pp-RS27* is modulated by light
A. Schematic illustration of the gene *Pp-RS27* and primers specific to exon 2 and exon 3, used in RT-PCR analysis. Exons (E) are shown as black and grey boxes, indicating the coding and non-coding regions, respectively. Scheme is drawn to scale.
B. Wild type protonemal tissue, cultivated under standard conditions (23°C, 16-h light/8-h dark) for six days, was transferred to dark (D) for 4 h or kept in light (L). Fully spliced transcript encoding the reference protein (mRNA1) and alternatively spliced mRNA3 are indicated. Other bands above the mRNA1 are the alternatively spliced variants (see Figure 4.2). *Actin* was used as a loading control.

4.3.10.9. *Pp-RS27* disruption enhances tolerance to ultraviolet-C irradiation through photoreactivation

When ancestors of vascular plants left the water for land, they became much more exposed to ultraviolet radiation that imposes cytotoxic as well as genotoxic effects. Plants were constrained to develop adaptations to the new light conditions. Apart from the generation of protective molecular ultraviolet (UV) "sunscreens", UV-absorbing and reflecting compounds such as flavonoids, hydroxycinnamate or other antioxidants (Tohge et al., 2016), plants also had to improve DNA repair mechanisms to maintain the genome integrity. The development of an efficient DNA repair by precise homologous recombination in mosses, a feature unique among plants, might serve as a proof of this idea (Rensing et al., 2008).

Using HR RT-PCR screen we detected three DNA damage repair genes (*FPG*, *RAD23*, and *UVH1/ATRAD1*) with alternative splicing modulated by At-RS31 overexpression that suggests a role for At-RS31 in DNA damage response (Chapter 3). Our further research revealed that At-RS31 overexpression increased sensitivity of Arabidopsis seedlings to DNA damaging UV-C irradiation, suggesting At-RS31 as a negative regulator of DNA dark repair (Chapter 3).

Intrigued by our findings in Arabidopsis, we tested phenotypes of Physcomitrella in a similar experimental set-up to detect signs of an evolutionary conserved response to DNA damage upon the UV-C irradiation.

Since UV-C is the most suitable wavelength spectrum to inflict DNA lesions (Stapleton, 1992), we irradiated five days old protonemal tissue of the wild type and two mutant lines *pp-rs27#36* and *pp-rs2z37#47* with the increasing dose of UV-C light. Afterwards, half of the plates were grown in the light to allow photoreactivation, and the other half was grown in the absence of light to test the condition of the dark repair mechanisms.

UV-C untreated dark-grown protonema showed slower growth and less intense green color than the tissue grown in the light and comparable effect was observed when lower intensity of the UV-C was applied (1 kJ/m^2) (Figure 4.27). Significant differences in growth were observed when lines were treated with 3 kJ/m^2. Firstly, higher UV-C dose was detrimental to all three genotypes if the UV-C treatment was followed by dark cultivation. Secondly, whereas wild type and *pp-rs2z37#47* mutant line displayed high sensitivity to UV-C also after light exposure, mutation in *pp-rs27#36* improved the resistance to UV-C

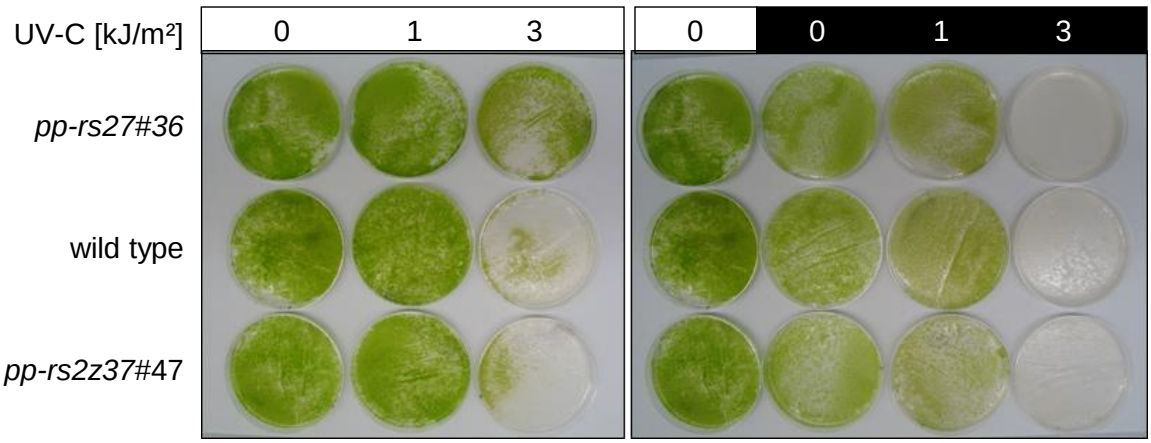


Figure 4.27. Mutant *pp-rs27#36* shows enhanced tolerance to UV-C irradiation upon photoreactivation.
UV-C irradiation in two intensities, 1 and 3 kJ/m^2, was applied on Physcomitrella protonemal tissue of wild type, *pp-rs27#36* and *pp-rs2z37#47* mutant lines grown for five days at 23°C, 16-h light/8-h dark. Half of the plates was afterwards cultivated for three days in the normal light conditions (white labels), and the other half was kept for two days in the dark and one day in the normal light conditions (black labels) and photographed. Experiment was conducted in 2 replicates.

substantially in the standard conditions with the light (Figure 4.27). Intriguingly, similar effect of the mutation in Arabidopsis orthologue *At-RS31* was observed in plants upon dark treatment, while the photoreactivation was not affected by *at-rs31* allele (Chapter 3).

4.3.10.10. Ultraviolet-C irradiation influences the splicing patterns of Physcomitrella SR genes

To monitor impact of UV-C irradiation on expression and alternative splicing of *Pp-RS27*, *Pp-RS2Z37* and *Pp-RS2Z38 SR* genes, we applied increasing dose of UV-C (0.5, 1, 2, 3 kJ/m^2) on the wild type protonemal tissue. Irradiation was followed by incubation of the tissue either under the light or dark conditions for four hours. RT-PCR analysis was performed.

In the control samples (0 kJ/m², light vs. 4 hours dark), confirming the previous results (Figure 4.26), we detected decrease in abundance of *Pp-RS27* fully spliced mRNA1 and shift towards AS isoforms, especially mRNA3, in response to dark (Figure 4.28).

In the samples incubated in the dark after UV-C irradiation, there is noticeable rise in the abundance of AS transcripts, which is not observed in mRNA1 (0.5 kJ/m², dark). With increased UV-C dose, the overall expression of *Pp-RS27* is shut down, similarly to the decreased expression of *Actin* control, suggesting general transcription shut down in response to UV-C in dark. In the tissue allowed to photoreactivate the UV-photoproducts by the growth in the light, shift to AS variants including mRNA3 is visible at 2 and 3 kJ/m² exposure.

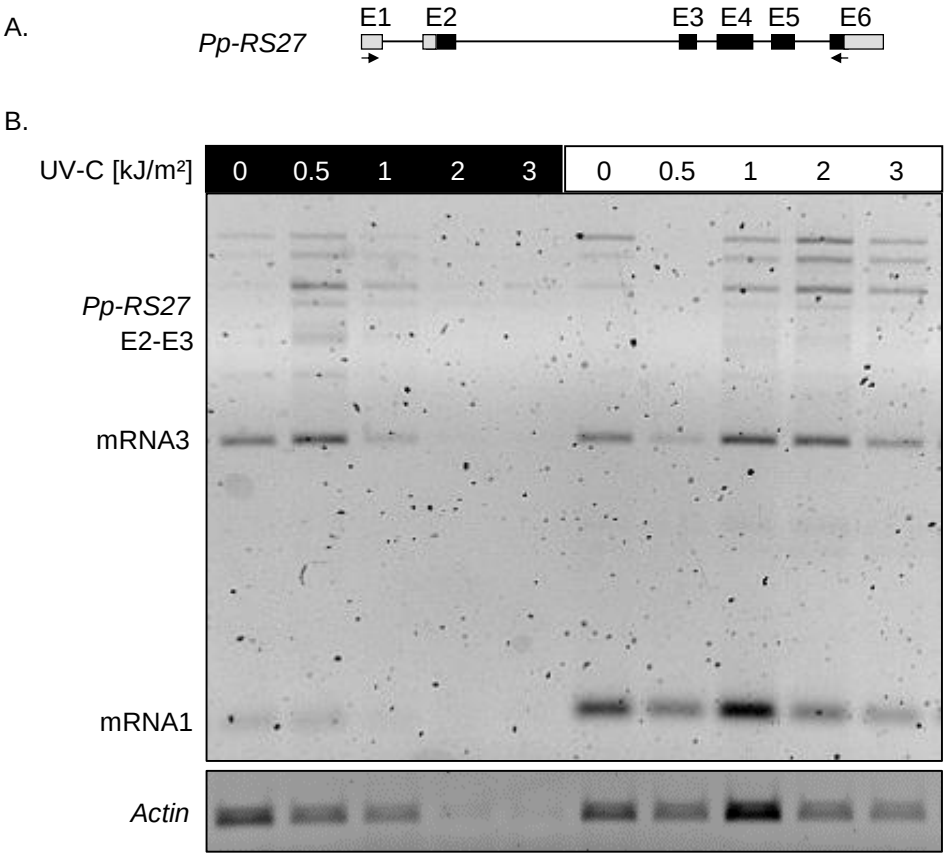


Figure 4.28. Alternative splicing of *Pp-RS27* is affected by UV-C.
A. Scheme of *Pp-RS27* with primer positions. Black and grey boxes indicate coding and non-coding exons (E), respectively. Lines represent the introns. Schemes are drawn to scale. B. RT-PCR analysis of the *Pp-RS27* alternative splicing and expression pattern. Wild type protonemal tissue was irradiated with increasing dose of UV-C (0.5, 1, 2, 3 kJ/m²) followed by dark (black labels) or light (white labels) incubation for 4 h. Analysed region is defined by the pair of exons (En-En) to which primers align. Fully spliced transcript encoding the reference protein (mRNA1) and alternatively spliced mRNA3 are indicated. Other bands above the mRNA1 are the alternatively spliced variants (see Figure 4.2). *Actin* was used as a loading control.

RS2Z genes of Physcomitrella display similar effect of UV-C and light on their AS profile. In *Pp-RS2Z37*, UV-C exposure of 0.5 kJ/m² followed by 4 hours in dark led to decreased abundance of protein-coding splice variant (Figure 4.29). Higher UV-C exposures, similarly to *Pp-RS27* and *Actin*, caused substantial decrease in *Pp-RS2Z37*expression. For plants kept in the light after irradiation, RT-PCR failed for 0.5 kJ/m² dose, however 1 kJ/m² dose did not affect AS ratio. We observed changes in the AS ratio after the stronger exposure (2 and 3 kJ/m²) resulting in enrichment of the PTC-containing AS isoforms and decrease in protein-coding mRNA1 (Figure 4.29).

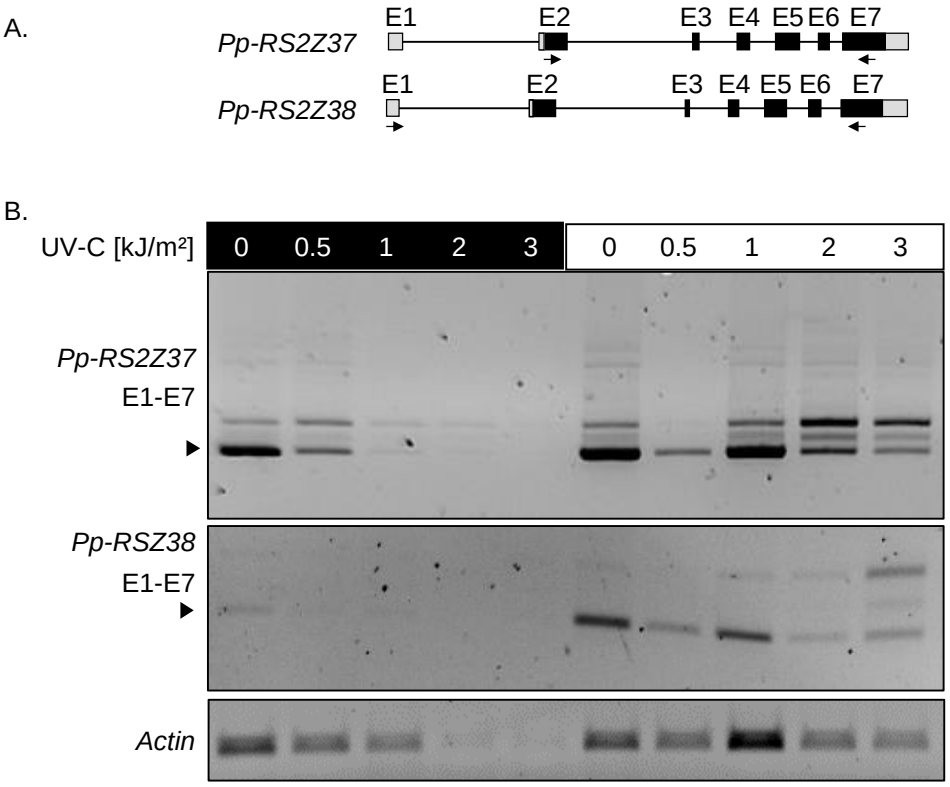


Figure 4.29. UV-C irradiation has impact on alternative splicing of Physcomitrella SR genes *Pp-RS2Z37* and *Pp-RS2Z38*.
A. Schemes of the SR genes with primer positions. Black and grey boxes indicate coding and non-coding exons (E), respectively. Lines represent the introns. Schemes are drawn to scale.
B. RT-PCR analysis of the *Pp-RS2Z37* and *Pp-RS2Z38* alternative splicing and expression pattern. WT protonemal tissue was irradiated with increasing strength of UV-C (0.5, 1, 2, 3 kJ/m²) followed by 4 hours incubation in the dark (black labels) or in the light (white labels). Analysed region is defined by the pair of exons (E*n*-E*n*) to which primers align. Black arrowheads mark fully spliced transcripts and the upper bands represent the alternatively spliced variants (See Figure 4.4, 4.5). *Actin* was used as a loading control.

Expression of *Pp-RS2Z38* was greatly down-regulated in dark even in the absence of UV-C exposure (Figure 4.29), therefore only experiments followed by photoreactivation period can be evaluated. Likewise to the previous observations, lower dose (1 kJ/m²) of the UV-C did not affect the AS pattern of *Pp-RS2Z38*, however higher doses (2 and 3 kJ/m²) led to relative changes in alternative splicing ratios, enhancing the AS variants and decreasing the fraction of the protein-coding mRNA1 (Figure 4.29).

4.4. Discussion

We previously uncovered alternative splice sites and their surrounding regions in plant-specific SR genes that are evolutionary highly conserved from moss to dicots (Kalyna et al., 2006). Here we detected extensive alternative splicing in three SR genes of *P. patens* and identified conserved AS responses to environmental cues that are in *P. patens* triggered similarly to Arabidopsis.

Although the potential of SR proteins to generate multiple AS variants is well documented both in plants and animals (Gupta et al., 2005; Iida and Go, 2006; Isshiki et al., 2006; Kalyna et al., 2003, 2006; Lazar and Goodman, 2000; Lejeune et al., 2001; Long and Caceres, 2009; Lopato et al., 1996, 1999, 2002; Palusa et al., 2007; Wang and Brendel, 2006), our analysis of three *P. patens* plant-specific SR genes revealed surprisingly high number of alternative splicing events, producing multitude of alternatively spliced isoforms. In *Pp-RS27*, we detected 16 splice variants generated by alternative splicing, occurring only in the long second intron, which is situated at the evolutionary conserved position in all plant RS orthologues (Iida and Go, 2006; Kalyna et al., 2006). In Arabidopsis, from 7 to 9 splice variants are generated in RS orthologues (Palusa et al., 2007). Importantly and in contrast to Physcomitrella *Pp-RS27*, majority of Arabidopsis RS subfamily splice variants are generated by intron retention events and only one 5' and one 3' alternative splice sites are usually present in the conserved long intron. However, thirteen alternative splice sites were identified in *Pp-RS27* long intron, raising a question on the mechanisms of their selection to generate the plethora of *Pp-RS27* transcripts. The *Pp-RS27* alternative splice sites have very variable features. The nucleotide sequences around the splice sites and their positions did not show any homology to the conserved alternative 5' and 3' spliced sites detected in the monocot and dicot RS genes. Three weak alternative 5' splice sites are situated only 4 and 6

nt to each other, which may pose sterical hindrances during their recognition by snRNAs. Selection of such closely spaced 5' splice sites suggests a competition during base pairing with U1 snRNA or with other components of the spliceosome at the later stages of splicing.

Two alternative 5' splice sites in *Pp-RS27* comprise a rare (<1%) GC instead of GU dinucleotide at the terminal position. One of the splice sites is a conventional 5' splice site, and the second one is alternative splice site Alt5ss1869, and both are rated with the highest splice site scores (94.40 and 89.11, respectively). However, Alt5ss1869 is used in only two AS isoforms (AS2 and AS6), generated in low amounts. In humans, it was observed that constitutively spliced GC-AG introns compensate the mispairing at the second intron position (C vs U) by generally stronger consensus sequence of the 5' splice site to improve base pairing with U1 and U5/U6 snRNAs (Thanaraj and Clark, 2001). This might be true for the identified conventional GC 5' splice site, it has the highest score of all splice sites in *Pp-RS27*. In humans, if GC-type 5' splice site is used in alternative splicing and has a weak consensus sequence, its weakness tends to be compensated by a strong 3' splice site (Thanaraj and Clark, 2001). Indeed, Alt5ss1869 is used only with conventional 3' splice site with rather high score and rate of selection. On the other hand, Alt5ss1869 has a strong consensus sequence with very high splice site score questioning the 3' splice site compensatory mechanism suggested for human GC-type 5' alternative splice sites. Our RT-PCR analyses of *Pp-RS27* transcripts show that the rate of a splice site usage in the second intron of *Pp-RS27* is not dependent on its strength only. Some strong splice sites are selected seldom and some weak ones are used frequently, suggesting that additional determinants are involved in the splice site recognition, such as splicing enhancers and silencers or RNA secondary structures.

Long introns of both RS2Z genes comprise three alternative splice sites. We experimentally confirmed predicted 3' alternative splice site that is conserved from moss to monocots and dicots. Such a strong preservation indicates importance of this splicing event in AS regulation of RS2Z genes and raises a question on its mechanism. Additional two 5' splice sites present in the second intron of RS2Z genes are specific to *P. patens* and absent in both dicots and monocots. All three alternative splice sites are situated in a region towards the 3' end of the second intron that is highly conserved between two Physcomitrella RS2Z paralogues, whereas the rest of the intron is much more variable. This implies an evolutionary selective pressure on the alternatively spliced intronic region, suggesting presence of important regulatory features.

Cassette exon was identified in the 5'UTR in the first intron of *Pp-RS2Z38*. Interestingly, this AS event is specific to the *Pp-RS2Z38,* because no significant sequence similarities to the cassette exon were found in the first intron of *Pp-RS2Z37.*

Alternative 3' splice site found in the fourth intron of *Pp-RS2Z38* extends the third exon by 15 nt and represents the only detected alternative transcript, potentially coding for an alternative protein isoform. Five amino acids would be added to the C-terminus of the RRM motif, however effect of this change on protein structure and function needs to be elucidated.

Selection of identified alternative splice sites in Physcomitrella RS and RS2Z genes leads predominantly to the production of PTC-containing (PTC+) transcripts. What is the fate of these products is unclear. They may code for proteins, although extremely truncated. Another option is that PTC followed by splice junctions renders them sensitive to NMD pathway and degradation. Inhibition of NMD by cycloheximide did not show increase in the abundance of PTC+ transcripts suggesting that they are not targets of NMD. Nevertheless, it does not indicate that these PTC+ transcripts code for truncated proteins. It is possible that they are retained in the nucleus and are not transported into the cytoplasm, thereby escaping the NMD machinery and translation.

A study of the fern *Marsilea vestita* demonstrated that regulated intron retention occurs in the microspores during their maturation by drying. Partially spliced transcripts are stored in the nucleus and upon initiation of spermatogenesis the retained introns are removed in a temporally regulated pattern (Boothby et al., 2013). This interesting example of control through nuclear storage and on demand transcript maturation could explain the presence of numerous AS isoforms in *Pp-RS27*. Wide range of *Pp-RS27* alternative mRNA isoforms could possibly represent partially spliced intermediate products. This hypothesis can be supported by different strategies of post-transcriptional and multi-step splicing described in animals (Hatton et al., 1998; Kameyama et al., 2012; Parra et al., 2008).

Out of 24 alternative transcripts generated in Physcomitrella RS and RS2Z genes, only one, mRNA2 in *Pp-RS2Z38,* codes for a protein isoform, suggesting that identified PTC+ alternative transcripts are important in regulation of protein abundance, rather than for protein isoform diversity.

Alternative splicing of SR protein genes in both plant and animal kingdoms was shown to be controlled in tissue-specific manner (Kalyna et al., 2003; Palusa et al., 2007; Lopato et al., 1999, 2002; Kalyna et al., 2006; Iida et al., 2004; Tillemans et al., 2005; Lopato et al., 1999; Ali and Reddy, 2008). In moss RS and RS2Z paralogs, we uncovered relative

enrichment of PTC+ alternatively spliced variants to protein-coding transcript in gametophores. Since the PTC+ alternative transcripts of *Pp-RS27*, *Pp-RS2Z37* and *Pp-RS2Z38* genes are not protein-coding a shift to the production of PTC+ variants might represent a mechanism to regulate the SR protein levels in different tissues.

In Arabidopsis, At-RS2Z33 affects alternative splicing and expression of *At-RS31* (Kalyna et al., 2006). We did not detect any changes in *Pp-RS27* expression in *pp-rs2z37* mutant line. It can be explained by redundancy in Physcomitrella RS2Z gene functions and can be tested in the double knock-out line, or we did not match the conditions triggering this type of control in Physcomitrella, because the cross-regulation in Arabidopsis was dependent on the age of the plant (unpublished results). Alternatively, this regulation might evolved later. Indeed, identified alternative splice sites used in *Pp-RS27* and *At-RS31* are not conserved. Moreover, in Arabidopsis, two other paralogues in RS subfamily (*At-RS40* and *At-RS41*) do not possess these conserved alternative splice sites (Kalyna et al., 2006) and are not regulated by At-RS2Z33 (Simpson et al., 2008).

To get insights into roles of Physcomitrella RS and RS2Z proteins in regulation of plant development, growth and response to environmental cues, we tested a range of various triggers, including factors already known from Arabidopsis (Chapter 3). We examined effects of plant hormones (auxin, cytokinin, and abscisic acid), drought, salinity, osmotic and oxidative stresses, UV-C irradiation, high energy (glucose) supply, and light and dark growth conditions on the phenotypes of *Pp-RS27* and *Pp-RS2Z37* mutants and on AS pattern of Physcomitrella RS and RS2Z genes. We observed that some factors did not trigger any response or the response was different from the one observed in Arabidopsis, while other exerted a highly conserved response at the phenotype and/or alternative splicing levels.

One of the examples of the divergent responses is sensitivity to glucose. In Arabidopsis, high amount of glucose inhibits growth and promotes premature senescence in At-RS31 overexpressing plants. In Physcomitrella, glucose also triggered response in phenotype, but, to the contrary, hypersensitivity was observed in mutants of *Pp-RS27*, where glucose inhibited growth of gametophores and promoted growth of caulonemata, filaments with fewer and less developed chloroplasts and reduced photosynthetic activity. We suppose that different number of paralogs in the RS subfamilies, a single gene in moss and four genes Arabidopsis, could be one of the reasons of the inconsistency. Arabidopsis paralogs might have diverged in function. Overexpression of At-RS31 leads to changes in splicing pattern and, as a consequence, to down-regulation of other three paralogs, At-RS31a, At-RS40 and At-RS41 (Kalyna et al., unpublished). Therefore the response to glucose in Arabidopsis

might be a secondary effect of At-RS31 overexpression mediated by one or more of its paralogs.

Enhanced caulonema formation and inhibition of chloronemata, cells that have many chloroplasts and are photosynthetically active, are promoted not only by excess of glucose but also by high light (Thelander et al., 2005).

Interestingly, chloroplast retrograde signaling modulates alternative splicing of At-RS31, and protein-coding mRNA1 is upregulated by light while non-productive splice variants are relatively upregulated in dark (Petrillo et al., 2014). Physcomitrella displayed the same effect: upregulation of PTC-containing transcripts in response to dark and increase of protein-coding splice variant in light, likely modulating Pp-RS27 protein levels in response to changing environmental conditions. This suggests strong conservation and importance of alternative splicing in regulating light response mediated by orthologues of RS subfamily. Indeed, recent genome-wide studies in Arabidopsis and Physcomitrella using RNA-seq revealed that light modulates thousands of genes by alteration of their alternative splicing, among them multiple splicing related proteins, including SR proteins and hnRNPs (Shikata et al., 2014; Wu et al., 2014). Moreover, photoreceptor pathways regulate alternative splicing by repression of IR events in splicing factors, suggesting increase in the level of functional proteins encoded by fully spliced transcripts (Wu et al., 2014).

Phenotype analysis revealed that depletion of Pp-RS27 improved the resistance of the mutant to the UV-induced DNA damage, compared to both wild type and mutant in Pp-RS2Z37. Interestingly, Arabidopsis At-RS31 mutant responded similarly, by enhanced resistance, while an overexpressing line showed hypersensitivity to UV-C. Response to DNA damage is therefore evolutionary conserved, suggesting importance of alternative splicing in the DNA damage response. Moreover, all three tested moss SR genes changed their AS patterns upon UV-C irradiation, producing PTC+ transcripts in relatively larger abundance, whereas protein-coding isoforms diminished. Participation of alternative splicing and other pre-mRNA processing mechanisms in the response to DNA damage has been studied in animals with increasing interest for the last years (Giono et al., 2016, Shkreta and Chabot, 2015), however the mechanisms and regulation of their cooperation have to yet be clarified.

Our comparative study of SR proteins in Physcomitrella and Arabidopsis, two species separated by 450 million years of evolution, uncovered specific as well as highly conserved AS events and alternative splicing regulation in response to light/dark transitions and DNA

damage, suggesting an important function of alternative splicing in the processes essential for colonizing the land by plants.

References

Airoldi, C.A., McKay, M., Davies, B., 2015. MAF2 Is Regulated by Temperature-Dependent Splicing and Represses Flowering at Low Temperatures in Parallel with FLM. PLOS ONE 10, e0126516. doi:10.1371/journal.pone.0126516

Ali, G.S., Reddy, A.S.N., 2008. Spatiotemporal organization of pre-mRNA splicing proteins in plants. Curr. Top. Microbiol. Immunol. 326, 103–118.

Arabidopsis Interactome Mapping Consortium, 2011. Evidence for network evolution in an Arabidopsis interactome map. Science 333, 601–607. doi:10.1126/science.1203877

Ashton, N.W., Grimsley, N.H., Cove, D.J., 1979. Analysis of gametophytic development in the moss, Physcomitrella patens, using auxin and cytokinin resistant mutants. Planta 144, 427–435. doi:10.1007/BF00380118

Babiychuk, E., Vandepoele, K., Wissing, J., Garcia-Diaz, M., Rycke, R.D., Akbari, H., Joubès, J., Beeckman, T., Jänsch, L., Frentzen, M., Montagu, M.C.E.V., Kushnir, S., 2011. Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family. Proc. Natl. Acad. Sci. 108, 6674–6679. doi:10.1073/pnas.1103442108

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., Frey, B.J., 2010. Deciphering the splicing code. Nature 465, 53–59. doi:10.1038/nature09000

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., Kim, T., Misquitta-Ali, C.M., Wilson, M.D., Kim, P.M., Odom, D.T., Frey, B.J., Blencowe, B.J., 2012. The evolutionary landscape of alternative splicing in vertebrate species. Science 338, 1587–1593. doi:10.1126/science.1230612

Bardou, F., Ariel, F., Simpson, C.G., Romero-Barrios, N., Laporte, P., Balzergue, S., Brown, J.W.S., Crespi, M., 2014. Long Noncoding RNA Modulates Alternative Splicing Regulators in Arabidopsis. Dev. Cell 30, 166–176. doi:10.1016/j.devcel.2014.06.017

Bargmann, B.O.R., Vanneste, S., Krouk, G., Nawy, T., Efroni, I., Shani, E., Choe, G., Friml, J., Bergmann, D.C., Estelle, M., Birnbaum, K.D., 2013. A map of cell type-specific auxin responses. Mol. Syst. Biol. 9, 688. doi:10.1038/msb.2013.40

Barta, A., Kalyna, M., Reddy, A.S.N., 2010. Implementing a Rational and Consistent Nomenclature for Serine/Arginine-Rich Protein Splicing Factors (SR Proteins) in Plants. Plant Cell 22, 2926–2929. doi:10.1105/tpc.110.078352

Barta, A., Marquez, Y., Brown, J.W.S., 2012. Challenges in Plant Alternative Splicing, in: Stamm, S., Smith, C.W.J., Lührmann, R. (Eds.), Alternative Pre-mRNA Splicing. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 79–91. doi:10.1002/9783527636778.ch7

Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M.A., Matzke, A.J., 1986. The expression of a nopaline synthase - human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. Plant Mol. Biol. 6, 347–357. doi:10.1007/BF00034942

Behzadnia, N., Golas, M.M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., Dube, P., Will, C.L., Urlaub, H., Stark, H., Lührmann, R., 2007. Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. EMBO J. 26, 1737–1748. doi:10.1038/sj.emboj.7601631

Berget, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Rev. Med. Virol. 10, 356-362; discussion 355–356.

Black, D.L., 2003. Mechanisms of Alternative Pre-Messenger RNA Splicing. Annu. Rev. Biochem. 72, 291–336. doi:10.1146/annurev.biochem.72.121801.161720

Boothby, T.C., Zipper, R.S., van der Weele, C.M., Wolniak, S.M., 2013. Removal of Retained Introns Regulates Translation in the Rapidly Developing Gametophyte of Marsilea vestita. Dev. Cell 24, 517–529. doi:10.1016/j.devcel.2013.01.015

Brown, J.W., Feix, G., Frendewey, D., 1986. Accurate in vitro splicing of two pre-mRNA plant introns in a HeLa cell nuclear extract. EMBO J. 5, 2749–2758.

Bus, J.S., Gibson, J.E., 1984. Paraquat: model for oxidant-initiated toxicity. Environ. Health Perspect. 55, 37–46.

Calixto, C.P.G., Simpson, C.G., Waugh, R., Brown, J.W.S., 2016. Alternative Splicing of Barley Clock Genes in Response to Low Temperature. PLOS ONE 11, e0168028. doi:10.1371/journal.pone.0168028

Campbell, M.A., Haas, B.J., Hamilton, J.P., Mount, S.M., Buell, C.R., 2006. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. BMC Genomics 7, 327. doi:10.1186/1471-2164-7-327

Cartegni, L., Chew, S.L., Krainer, A.R., 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat. Rev. Genet. 3, 285–298. doi:10.1038/nrg775

Carvalho, R.F., Carvalho, S.D., Duque, P., 2010. The Plant-Specific SR45 Protein Negatively Regulates Glucose and ABA Signaling during Early Seedling Development in Arabidopsis1[W]. Plant Physiol. 154, 772–783. doi:10.1104/pp.110.155523

Carvalho, R.F., Feijão, C.V., Duque, P., 2013. On the physiological significance of alternative splicing events in higher plants. Protoplasma 250, 639–650. doi:10.1007/s00709-012-0448-9

Chamala, S., Feng, G., Chavarro, C., Barbazuk, W.B., 2015. Genome-wide identification of evolutionarily conserved alternative splicing events in flowering plants. Front. Bioeng. Biotechnol. 3, 33. doi:10.3389/fbioe.2015.00033

Chang, C.-Y., Lin, W.-D., Tu, S.-L., 2014. Genome-Wide Analysis of Heat-Sensitive Alternative Splicing in Physcomitrella patens. Plant Physiol. 165, 826–840. doi:10.1104/pp.113.230540

Chow, L.C., Gelinas, R.E., Broker, T.R., Roberts, R.J., 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Rev. Med. Virol. 10, 362-371; discussion 355–356.

Cove, D., 2005. The Moss Physcomitrella patens. Annu. Rev. Genet. 39, 339–358. doi:10.1146/annurev.genet.39.073003.110214

Cove, D., Bezanilla, M., Harries, P., Quatrano, R., 2006. Mosses as Model Systems for the Study of Metabolism and Development. Annu. Rev. Plant Biol. 57, 497–520. doi:10.1146/annurev.arplant.57.032905.105338

Cruz, T.M.D., Carvalho, R.F., Richardson, D.N., Duque, P., 2014. Abscisic Acid (ABA) Regulation of Arabidopsis SR Protein Gene Expression. Int. J. Mol. Sci. 15, 17541–17564. doi:10.3390/ijms151017541

Cuming, A.C., Cho, S.H., Kamisugi, Y., Graham, H., Quatrano, R.S., 2007. Microarray analysis of transcriptional responses to abscisic acid and osmotic, salt, and drought stress in the moss, Physcomitrella patens. New Phytol. 176, 275–287. doi:10.1111/j.1469-8137.2007.02187.x

de la Fuente van Bentem, S., Anrather, D., Djamei, A., Hufnagl, T., Barta, A., Csaszar, E., Dohnal, I., Lecourieux, D., Hirt, H., 2006. Phosphoproteomics reveals extensive in vivo phosphorylation of Arabidopsis proteins involved in RNA metabolism. Nucleic Acids Res. 34, 3267–3278. doi:10.1093/nar/gkl429

Decker, E.L., Frank, W., Sarnighausen, E., Reski, R., 2006. Aux Cyt ABA Moss Systems Biology en Route: Phytohormones in Physcomitrella Development. Plant Biol. 8, 397–406. doi:10.1055/s-2006-923952

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H., Lührmann, R., 2006. Protein Composition and Electron Microscopy Structure of Affinity-Purified Human Spliceosomal B Complexes Isolated under Physiological Conditions. Mol. Cell. Biol. 26, 5528–5543. doi:10.1128/MCB.00582-06

Draper, J., Scott, R., Armitage, P., 1988. Plant Genetic Transformation and Gene Expression: A Laboratory Manual. Blackwell Science Inc, Oxford ; Boston.

Drechsel, G., Kahles, A., Kesarwani, A.K., Stauffer, E., Behr, J., Drewe, P., Rätsch, G., Wachter, A., 2013. Nonsense-Mediated Decay of Alternative Precursor mRNA Splicing Variants Is a Major Determinant of the *Arabidopsis* Steady State Transcriptome. Plant Cell Online tpc.113.115485. doi:10.1105/tpc.113.115485

Duclos, S., Aller, P., Jaruga, P., Dizdaroglu, M., Wallace, S.S., Doublié, S., 2012. Structural and biochemical studies of a plant formamidopyrimidine-DNA glycosylase reveal why eukaryotic Fpg glycosylases do not excise 8-oxoguanine. DNA Repair 11, 714–725. doi:10.1016/j.dnarep.2012.06.004

Duque, P., 2011. A role for SR proteins in plant stress responses. Plant Signal. Behav. 6, 49–54. doi:10.4161/psb.6.1.14063

Farmer, L.M., Book, A.J., Lee, K.-H., Lin, Y.-L., Fu, H., Vierstra, R.D., 2010. The RAD23 Family Provides an Essential Connection between the 26S Proteasome and Ubiquitylated Proteins in Arabidopsis. Plant Cell 22, 124–142. doi:10.1105/tpc.109.072660

Farrer, T., Roller, A.B., Kent, W.J., Zahler, A.M., 2002. Analysis of the role of Caenorhabditis elegans GC-AG introns in regulated splicing. Nucleic Acids Res. 30, 3360–3367. doi:10.1093/nar/gkf465

Fernando, V.C.D., Schroeder, D.F., 2016. Role of ABA in Arabidopsis Salt, Drought, and Desiccation Tolerance. doi:10.5772/61957

Fidantsef, A.L., Mitchell, D.L., Britt, A.B., 2000. The Arabidopsis UVH1 Gene Is a Homolog of the Yeast Repair Endonuclease RAD1. Plant Physiol. 124, 579–586. doi:10.1104/pp.124.2.579

Filichkin, S., Priest, H.D., Megraw, M., Mockler, T.C., 2015. Alternative splicing in plants: directing traffic at the crossroads of adaptation and environmental stress. Curr. Opin. Plant Biol. 24, 125–135. doi:10.1016/j.pbi.2015.02.008

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.-K., Mockler, T.C., 2010. Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome Res. 20, 45–58. doi:10.1101/gr.093302.109

Fujisaki, K., Ishikawa, M., 2008. Identification of an Arabidopsis thaliana protein that binds to tomato mosaic virus genomic RNA and inhibits its multiplication. Virology 380, 402–411. doi:10.1016/j.virol.2008.07.033

Gallego, F., Fleck, O., Li, A., Wyrzykowska, J., Tinland, B., 2000. AtRAD1, a plant homologue of human and yeast nucleotide excision repair endonucleases, is involved in dark repair of UV damages and recombination. Plant J. Cell Mol. Biol. 21, 507–518.

Gao, M.-J., Murphy, T.M., 2001. Alternative Forms of Formamidopyrimidine-DNA Glycosylase from Arabidopsis thaliana¶. Photochem. Photobiol. 73, 128–134. doi:10.1562/0031-8655(2001)0730128AFOFDG2.0.CO2

Giono, L.E., Nieto Moreno, N., Cambindo Botto, A.E., Dujardin, G., Muñoz, M.J., Kornblihtt, A.R., 2016. The RNA Response to DNA Damage. J. Mol. Biol., Macromolecular Complexes in Transcription and Co-transcriptional RNA Processing 428, 2636–2651. doi:10.1016/j.jmb.2016.03.004

Göhring, J., Jacak, J., Barta, A., 2014. Imaging of Endogenous Messenger RNA Splice Variants in Living Cells Reveals Nuclear Retention of Transcripts Inaccessible to Nonsense-Mediated Decay in Arabidopsis. Plant Cell 26, 754–764. doi:10.1105/tpc.113.118075

Graveley, B.R., Hertel, K.J., Maniatis, T., 2001. The role of U2AF35 and U2AF65 in enhancer-dependent splicing. RNA N. Y. N 7, 806–818.

Gullerova, M., BARTA, A., LORKOVIĆ, Z.J., 2006. AtCyp59 is a multidomain cyclophilin from Arabidopsis thaliana that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. RNA 12, 631–643. doi:10.1261/rna.2226106

Gupta, S., Wang, B.-B., Stryker, G.A., Zanetti, M.E., Lal, S.K., 2005. Two novel arginine/serine (SR) proteins in maize are differentially spliced and utilize non-canonical splice sites. Biochim. Biophys. Acta BBA - Gene Struct. Expr. 1728, 105–114. doi:10.1016/j.bbaexp.2005.01.004

Han, S.P., Tang, Y.H., Smith, R., 2010. Functional diversity of the hnRNPs: past, present and perspectives. Biochem. J. 430, 379–392. doi:10.1042/BJ20100396

Hancock, R.D., Morris, W.L., Ducreux, L.J.M., Morris, J.A., Usman, M., Verrall, S.R., Fuller, J., Simpson, C.G., Zhang, R., Hedley, P.E., Taylor, M.A., 2014. Physiological, biochemical and molecular responses of the potato (Solanum tuberosum L.) plant to moderately elevated temperature. Plant Cell Environ. 37, 439–450. doi:10.1111/pce.12168

Harlow, G.R., Jenkins, M.E., Pittalwala, T.S., Mount, D.W., 1994. Isolation of uvh1, an Arabidopsis mutant hypersensitive to ultraviolet light and ionizing radiation. Plant Cell 6, 227–235. doi:10.1105/tpc.6.2.227

Hartmuth, K., Barta, A., 1986. In vitro processing of a plant pre-mRNA in a HeLa cell nuclear extract. Nucleic Acids Res. 14, 7513–7528.

Hatton, A.R., Subramaniam, V., Lopez, A.J., 1998. Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon-exon junctions. Mol. Cell 2, 787–796.

Hirt, H., 2000. Connecting oxidative stress, auxin, and cell cycle regulation through a plant mitogen-activated protein kinase pathway. Proc. Natl. Acad. Sci. U. S. A. 97, 2405–2407.

Hutchison, C.E., Kieber, J.J., 2002. Cytokinin Signaling in Arabidopsis. Plant Cell 14, s47–s59. doi:10.1105/tpc.010444

Hwang, I., Sheen, J., Müller, B., 2012. Cytokinin signaling networks. Annu. Rev. Plant Biol. 63, 353–380. doi:10.1146/annurev-arplant-042811-105503

Ibrahim, E.C., Schaal, T.D., Hertel, K.J., Reed, R., Maniatis, T., 2005. Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers. Proc. Natl. Acad. Sci. U. S. A. 102, 5002–5007. doi:10.1073/pnas.0500543102

Iida, K., Go, M., 2006. Survey of Conserved Alternative Splicing Events of mRNAs Encoding SR Proteins in Land Plants. Mol. Biol. Evol. 23, 1085–1094. doi:10.1093/molbev/msj118

Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, K., Toyoda, T., Konagaya, A., Shinozaki, K., 2004. Genome-wide analysis of alternative pre-mRNA splicing in Arabidopsis thaliana based on full-length cDNA sequences. Nucleic Acids Res. 32, 5096–5103. doi:10.1093/nar/gkh845

Isshiki, M., Tsumoto, A., Shimamoto, K., 2006. The serine/arginine-rich protein family in rice plays important roles in constitutive and alternative splicing of pre-mRNA. Plant Cell 18, 146–158. doi:10.1105/tpc.105.037069

Iwata, H., Gotoh, O., 2011. Comparative analysis of information contents relevant to recognition of introns in many species. BMC Genomics 12, 45. doi:10.1186/1471-2164-12-45

James, A.B., Syed, N.H., Bordage, S., Marshall, J., Nimmo, G.A., Jenkins, G.I., Herzyk, P., Brown, J.W.S., Nimmo, H.G., 2012. Alternative Splicing Mediates Responses of the Arabidopsis Circadian Clock to Temperature Changes[W]. Plant Cell 24, 961–981. doi:10.1105/tpc.111.093948

Jensen, K.B., Dredge, B.K., Stefani, G., Zhong, R., Buckanovich, R.J., Okano, H.J., Yang, Y.Y.L., Darnell, R.B., 2000. Nova-1 Regulates Neuron-Specific Alternative Splicing and Is Essential for Neuronal Viability. Neuron 25, 359–371. doi:10.1016/S0896-6273(00)80900-9

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., Shoemaker, D.D., 2003. Genome-Wide Survey of Human Alternative Pre-mRNA Splicing with Exon Junction Microarrays. Science 302, 2141–2144. doi:10.1126/science.1090100

Jurica, M.S., Moore, M.J., 2003. Pre-mRNA Splicing. Mol. Cell 12, 5–14. doi:10.1016/S1097-2765(03)00270-3

Kalyna, M., Barta, A., 2004. A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel gene functions? Biochem. Soc. Trans. 32, 561–564. doi:10.1042/BST0320561

Kalyna, M., Lopato, S., Barta, A., 2003. Ectopic Expression of atRSZ33 Reveals Its Function in Splicing and Causes Pleiotropic Changes in Development. Mol. Biol. Cell 14, 3565–3577. doi:10.1091/mbc.E03-02-0109

Kalyna, M., Lopato, S., Voronin, V., Barta, A., 2006. Evolutionary conservation and regulation of particular alternative splicing events in plant SR proteins. Nucleic Acids Res. 34, 4395–4405. doi:10.1093/nar/gkl570

Kalyna, M., Simpson, C.G., Syed, N.H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H.Q., Barta, A., Brown, J.W.S., 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

Kameyama, T., Suzuki, H., Mayeda, A., 2012. Re-splicing of mature mRNA in cancer cells promotes activation of distant weak alternative splice sites. Nucleic Acids Res. 40, 7896–7906. doi:10.1093/nar/gks520

Kamisugi, Y., Cuming, A.C., Cove, D.J., 2005. Parameters determining the efficiency of gene targeting in the moss Physcomitrella patens. Nucleic Acids Res. 33, e173–e173. doi:10.1093/nar/gni172

Kamisugi, Y., Schlink, K., Rensing, S.A., Schween, G., von Stackelberg, M., Cuming, A.C., Reski, R., Cove, D.J., 2006. The mechanism of gene targeting in Physcomitrella patens: homologous recombination, concatenation and multiple integration. Nucleic Acids Res. 34, 6205–6214. doi:10.1093/nar/gkl832

Kazan, K., 2003. Alternative splicing and proteome diversity in plants: the tip of the iceberg has just emerged. Trends Plant Sci. 8, 468–471. doi:10.1016/j.tplants.2003.09.001

Kim, S.H., Koroleva, O.A., Lewandowska, D., Pendle, A.F., Clark, G.P., Simpson, C.G., Shaw, P.J., Brown, J.W.S., 2009. Aberrant mRNA Transcripts and the Nonsense-Mediated Decay

Proteins UPF2 and UPF3 Are Enriched in the Arabidopsis Nucleolus. Plant Cell Online 21, 2045–2057. doi:10.1105/tpc.109.067736

Knight, C.D., Cove, D.J., Cuming, A.C., Quatrano, R.S., 2002. Moss Gene Technology, in: Gilmartin, P.M., Bowler, C. (Eds.), Plant Molecular Biology – a Practical Approach.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., Ule, J., 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat. Struct. Mol. Biol. 17, 909–915. doi:10.1038/nsmb.1838

Kornblihtt, A.R., Schor, I.E., Alló, M., Dujardin, G., Petrillo, E., Muñoz, M.J., 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell Biol. 14, 153–165. doi:10.1038/nrm3525

Krishnamurthy, A., Rathinasabapathi, B., 2013. Oxidative stress tolerance in plants. Plant Signal. Behav. 8. doi:10.4161/psb.25761

Kwak, K.J., Kim, J.Y., Kim, Y.O., Kang, H., 2007. Characterization of Transgenic Arabidopsis Plants Overexpressing High Mobility Group B Proteins under High Salinity, Drought or Cold Stress. Plant Cell Physiol. 48, 221–231. doi:10.1093/pcp/pcl057

Lazar, G., Goodman, H.M., 2000. The Arabidopsis splicing factor SR1 is regulated by alternative splicing. Plant Mol. Biol. 42, 571–581.

Lejeune, F., Cavaloc, Y., Stevenin, J., 2001. Alternative splicing of intron 3 of the serine/arginine-rich protein 9G8 gene. Identification of flanking exonic splicing enhancers and involvement of 9G8 as a trans-acting factor. J. Biol. Chem. 276, 7850–7858. doi:10.1074/jbc.M009510200

Leviatan, N., Alkan, N., Leshkowitz, D., Fluhr, R., 2013. Genome-Wide Survey of Cold Stress Regulated Alternative Splicing in Arabidopsis thaliana with Tiling Microarray. PLOS ONE 8, e66511. doi:10.1371/journal.pone.0066511

Lim, L.P., Burge, C.B., 2001. A computational analysis of sequence features involved in recognition of short introns. Proc. Natl. Acad. Sci. U. S. A. 98, 11193–11198. doi:10.1073/pnas.201407298

Liu, Z., Hossain, G.S., Islas-Osuna, M.A., Mitchell, D.L., Mount, D.W., 2000. Repair of UV damage in plants by nucleotide excision repair: Arabidopsis UVH1 DNA repair gene is a homolog of Saccharomyces cerevisiae Rad1. Plant J. 21, 519–528. doi:10.1046/j.1365-313x.2000.00707.x

Llorian, M., Smith, C.W.J., 2012. Quantification of Alternative Splice Variants, in: Stamm, S., Smith, C.W.J., Lührmann, R. (Eds.), Alternative Pre-mRNA Splicing. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 227–236. doi:10.1002/9783527636778.ch21

Long, J.C., Caceres, J.F., 2009. The SR protein family of splicing factors: master regulators of gene expression. Biochem. J. 417, 15–27. doi:10.1042/BJ20081501

Lopato, S., Forstner, C., Kalyna, M., Hilscher, J., Langhammer, U., Indrapichate, K., Lorković, Z.J., Barta, A., 2002. Network of Interactions of a Novel Plant-specific Arg/Ser-rich Protein,

atRSZ33, with atSC35-like Splicing Factors. J. Biol. Chem. 277, 39989–39998. doi:10.1074/jbc.M206455200

Lopato, S., Kalyna, M., Dorner, S., Kobayashi, R., Krainer, A.R., Barta, A., 1999. atSRp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes. Genes Dev. 13, 987–1001.

Lopato, S., Waigmann, E., Barta, A., 1996. Characterization of a novel arginine/serine-rich splicing factor in Arabidopsis. Plant Cell 8, 2255–2264. doi:10.1105/tpc.8.12.2255

Lorković, Z.J., Barta, A., 2002. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. Nucleic Acids Res. 30, 623–635.

Lorkovic, Z.J., Lehner, R., Forstner, C., Barta, A., 2005. Evolutionary conservation of minor U12-type spliceosome between plants and humans. RNA 11, 1095–1107. doi:10.1261/rna.2440305

Manley, J.L., Krainer, A.R., 2010. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev. 24, 1073–1074. doi:10.1101/gad.1934910

Maronova, M., Kalyna, M., 2016. Generating Targeted Gene Knockout Lines in Physcomitrella patens to Study Evolution of Stress-Responsive Mechanisms. Methods Mol. Biol. Clifton NJ 1398, 221–234. doi:10.1007/978-1-4939-3356-3_18

Marquez, Y., Brown, J.W.S., Simpson, C., Barta, A., Kalyna, M., 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome Res. 22, 1184–1195. doi:10.1101/gr.134106.111

Marquez, Y., Höpfler, M., Ayatollahi, Z., Barta, A., Kalyna, M., 2015. Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. Genome Res. 25, 995–1007. doi:10.1101/gr.186585.114

Martinez-Contreras, R., Cloutier, P., Shkreta, L., Fisette, J.-F., Revil, T., Chabot, B., 2007. hnRNP proteins and splicing control. Adv. Exp. Med. Biol. 623, 123–147.

Merkin, J., Russell, C., Chen, P., Burge, C.B., 2012. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science 338, 1593–1599. doi:10.1126/science.1228186

Naser, V., Shani, E., 2016. Auxin response under osmotic stress. Plant Mol. Biol. 91, 661–672. doi:10.1007/s11103-016-0476-5

Palusa, S.G., Ali, G.S., Reddy, A.S.N., 2007. Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. Plant J. 49, 1091–1107. doi:10.1111/j.1365-313X.2006.03020.x

Palusa, S.G., Reddy, A.S.N., 2010. Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. New Phytol. 185, 83–89. doi:10.1111/j.1469-8137.2009.03065.x

Pan, Q., Saltzman, A.L., Kim, Y.K., Misquitta, C., Shai, O., Maquat, L.E., Frey, B.J., Blencowe, B.J., 2006. Quantitative microarray profiling provides evidence against widespread

coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. Genes Dev. 20, 153–158. doi:10.1101/gad.1382806

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. 40, 1413–1415. doi:10.1038/ng.259

Pan, Q., Shai, O., Misquitta, C., Zhang, W., Saltzman, A.L., Mohammad, N., Babak, T., Siu, H., Hughes, T.R., Morris, Q.D., Frey, B.J., Blencowe, B.J., 2004. Revealing Global Regulatory Features of Mammalian Alternative Splicing Using a Quantitative Microarray Platform. Mol. Cell 16, 929–941. doi:10.1016/j.molcel.2004.12.004

Parra, M.K., Tan, J.S., Mohandas, N., Conboy, J.G., 2008. Intrasplicing coordinates alternative first exons with alternative splicing in the protein 4.1R gene. EMBO J. 27, 122–131. doi:10.1038/sj.emboj.7601957

Pedersen, D.S., Merkle, T., Marktl, B., Lildballe, D.L., Antosch, M., Bergmann, T., Tönsing, K., Anselmetti, D., Grasser, K.D., 2010. Nucleocytoplasmic Distribution of the Arabidopsis Chromatin-Associated HMGB2/3 and HMGB4 Proteins. Plant Physiol. 154, 1831–1841. doi:10.1104/pp.110.163055

Petrillo, E., Godoy Herz, M.A., Barta, A., Kalyna, M., Kornblihtt, A.R., 2014a. Let there be light: Regulation of gene expression in plants. RNA Biol. 11, 1215–1220. doi:10.4161/15476286.2014.972852

Petrillo, E., Herz, M.A.G., Fuchs, A., Reifer, D., Fuller, J., Yanovsky, M.J., Simpson, C., Brown, J.W.S., Barta, A., Kalyna, M., Kornblihtt, A.R., 2014b. A chloroplast retrograde signal regulates nuclear alternative splicing. Science 344, 427–430. doi:10.1126/science.1250322

Placido, A., Sieber, F., Gobert, A., Gallerani, R., Giegé, P., Maréchal-Drouard, L., 2010. Plant mitochondria use two pathways for the biogenesis of tRNA His. Nucleic Acids Res. 38, 7711–7717. doi:10.1093/nar/gkq646

Prigge, M.J., Bezanilla, M., 2010. Evolutionary crossroads in developmental biology: *Physcomitrella patens*. Development 137, 3535–3543. doi:10.1242/dev.049023

Raczynska, K.D., Simpson, C.G., Ciesiolka, A., Szewc, L., Lewandowska, D., McNicol, J., Szweykowska-Kulinska, Z., Brown, J.W.S., Jarmolowski, A., 2010. Involvement of the nuclear cap-binding protein complex in alternative splicing in Arabidopsis thaliana. Nucleic Acids Res. 38, 265–278. doi:10.1093/nar/gkp869

Rauch, H.B., Patrick, T.L., Klusman, K.M., Battistuzzi, F.U., Mei, W., Brendel, V.P., Lal, S.K., 2014. Discovery and Expression Analysis of Alternative Splicing Events Conserved among Plant SR Proteins. Mol. Biol. Evol. 31, 605–613. doi:10.1093/molbev/mst238

Reddy, A.S.N., 2007. Alternative Splicing of Pre-Messenger RNAs in Plants in the Genomic Era. Annu. Rev. Plant Biol. 58, 267–294. doi:10.1146/annurev.arplant.58.032806.103754

Reddy, A.S.N., Marquez, Y., Kalyna, M., Barta, A., 2013. Complexity of the Alternative Splicing Landscape in Plants. Plant Cell 25, 3657–3683. doi:10.1105/tpc.113.117523

Reddy, A.S.N., Shad Ali, G., 2011. Plant serine/arginine-rich proteins: roles in precursor messenger RNA splicing, plant development, and stress responses. Wiley Interdiscip. Rev. RNA 2, 875–889. doi:10.1002/wrna.98

Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.-F., Lindquist, E.A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W.B., Barker, E., Bennetzen, J.L., Blankenship, R., Cho, S.H., Dutcher, S.K., Estelle, M., Fawcett, J.A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K.A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D.R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P.J., Sanderfoot, A., Schween, G., Shiu, S.-H., Stueber, K., Theodoulou, F.L., Tu, H., Peer, Y.V. de, Verrier, P.J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A.C., Hasebe, M., Lucas, S., Mishler, B.D., Reski, R., Grigoriev, I.V., Quatrano, R.S., Boore, J.L., 2008a. The Physcomitrella Genome Reveals Evolutionary Insights into the Conquest of Land by Plants. Science 319, 64–69. doi:10.1126/science.1150646

Richardson, D.N., Rogers, M.F., Labadorf, A., Ben-Hur, A., Guo, H., Paterson, A.H., Reddy, A.S.N., 2011. Comparative Analysis of Serine/Arginine-Rich Proteins across 27 Eukaryotes: Insights into Sub-Family Classification and Extent of Alternative Splicing. PLoS ONE 6. doi:10.1371/journal.pone.0024542

Rukov, J.L., Irimia, M., Mørk, S., Lund, V.K., Vinther, J., Arctander, P., 2007. High Qualitative and Quantitative Conservation of Alternative Splicing in Caenorhabditis elegans and Caenorhabditis briggsae. Mol. Biol. Evol. 24, 909–917. doi:10.1093/molbev/msm023

Sakharkar, M.K., Chow, V.T.K., Kangueane, P., 2004. Distributions of Exons and Introns in the Human Genome. In Silico Biol. 4, 387–393.

Sanchez, S.E., Petrillo, E., Beckwith, E.J., Zhang, X., Rugnone, M.L., Hernando, C.E., Cuevas, J.C., Godoy Herz, M.A., Depetris-Chauvin, A., Simpson, C.G., Brown, J.W.S., Cerdán, P.D., Borevitz, J.O., Mas, P., Ceriani, M.F., Kornblihtt, A.R., Yanovsky, M.J., 2010. A methyl transferase links the circadian clock to the regulation of alternative splicing. Nature 468, 112–116. doi:10.1038/nature09470

Schaefer, D., Zryd, J.P., Knight, C.D., Cove, D.J., 1991. Stable transformation of the moss Physcomitrella patens. Mol. Gen. Genet. MGG 226, 418–424.

Schippers, J.H.M., Schmidt, R., Wagstaff, C., Jing, H.-C., 2015. Living to Die and Dying to Live: The Survival Strategy behind Leaf Senescence. Plant Physiol. 169, 914–930. doi:10.1104/pp.15.00498

Schwartzenberg, K. von, 2006. Moss Biology and Phytohormones - Cytokinins in Physcomitrella. Plant Biol. 8, 382–388. doi:10.1055/s-2006-923962

Shalgi, R., Hurt, J.A., Lindquist, S., Burge, C.B., 2014. Widespread Inhibition of Posttranscriptional Splicing Shapes the Cellular Transcriptome following Heat Shock. Cell Rep. 7, 1362–1370. doi:10.1016/j.celrep.2014.04.044

Sharp, P.A., Burge, C.B., 1997. Classification of introns: U2-type or U12-type. Cell 91, 875–879.

Shikata, H., Hanada, K., Ushijima, T., Nakashima, M., Suzuki, Y., Matsushita, T., 2014. Phytochrome controls alternative splicing to mediate light responses in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 111, 18781–18786. doi:10.1073/pnas.1407147112

Shinde, S., Nurul Islam, M., Ng, C.K.-Y., 2012. Dehydration stress-induced oscillations in LEA protein transcripts involves abscisic acid in the moss, Physcomitrella patens. New Phytol. 195, 321–328. doi:10.1111/j.1469-8137.2012.04193.x

Shkreta, L., Chabot, B., 2015. The RNA Splicing Response to DNA Damage. Biomolecules 5, 2935–2977. doi:10.3390/biom5042935

Simpson, C.G., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Barta, A., Brown, J.W.S., 2008a. 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

Simpson, C.G., Lewandowska, D., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Raczynska, D., Jarmolowski, A., Barta, A., Brown, J.W.S., 2008b. Alternative splicing in plants. Biochem. Soc. Trans. 36, 508–510. doi:10.1042/BST0360508

Simpson, C.G., Manthri, S., Raczynska, K.D., Kalyna, M., Lewandowska, D., Kusenda, B., Maronova, M., Szweykowska-Kulinska, Z., Jarmolowski, A., Barta, A., Brown, J.W.S., 2010. Regulation of plant gene expression by alternative splicing. Biochem. Soc. Trans. 38, 667–671. doi:10.1042/BST0380667

Simpson, C.G., Syed, N.H., Manthri, S., Fuller, J.D., Maronova, M., Kusenda, B., Kalyna, M., Barta, A., Brown, J.W.S., 2012. Monitoring Changes in Plant Alternative Splicing Events, in: Stamm, S., Smith, C.W.J., Lührmann, R. (Eds.), Alternative Pre-mRNA Splicing. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 248–259. doi:10.1002/9783527636778.ch23

Simpson, C.G., Thow, G., Clark, G.P., Jennings, S.N., Watters, J.A., Brown, J.W.S., 2002. Mutational analysis of a plant branchpoint and polypyrimidine tract required for constitutive splicing of a mini-exon. RNA 8, 47–56.

Singh, R., Valcárcel, J., 2005. Building specificity with nonspecific RNA-binding proteins. Nat. Struct. Mol. Biol. 12, 645–653. doi:10.1038/nsmb961

Slade, P., 1966. The Fate of Paraquat Applied to Plants. Weed Res. 6, 158–167. doi:10.1111/j.1365-3180.1966.tb00876.x

Staiger, D., 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

Stapleton, A., 1992. Ultraviolet Radiation and Plants: Burning Questions. Plant Cell 4, 1353–1358.

Streitner, C., Köster, T., Simpson, C.G., Shaw, P., Danisman, S., Brown, J.W.S., Staiger, D., 2012. An hnRNP-like RNA-binding protein affects alternative splicing by in vivo interaction with transcripts in Arabidopsis thaliana. Nucleic Acids Res. 40, 11240–11255. doi:10.1093/nar/gks873

Sugnet, C.W., Srinivasan, K., Clark, T.A., O'Brien, G., Cline, M.S., Wang, H., Williams, A., Kulp, D., Blume, J.E., Haussler, D., Ares, M., 2006. Unusual Intron Conservation near Tissue-

Regulated Exons Found by Splicing Microarrays. PLoS Comput. Biol. 2. doi:10.1371/journal.pcbi.0020004

Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., Brown, J.W.S., 2012. Alternative splicing in plants – coming of age. Trends Plant Sci. 17, 616–623. doi:10.1016/j.tplants.2012.06.001

Syed, N.H., Prince, S.J., Mutava, R.N., Patil, G., Li, S., Chen, W., Babu, V., Joshi, T., Khan, S., Nguyen, H.T., 2015. Core clock, SUB1, and ABAR genes mediate flooding and drought responses via alternative splicing in soybean. J. Exp. Bot. 66, 7129–7149. doi:10.1093/jxb/erv407

Takezawa, D., Komatsu, K., Sakata, Y., 2011. ABA in bryophytes: how a universal growth regulator in life became a plant hormone? J. Plant Res. 124, 437–453. doi:10.1007/s10265-011-0410-5

Tarn, W.-Y., Steitz, J.A., 1996. A Novel Spliceosome Containing U11, U12, and U5 snRNPs Excises a Minor Class (AT–AC) Intron In Vitro. Cell 84, 801–811. doi:10.1016/S0092-8674(00)81057-0

Thanaraj, T.A., Clark, F., 2001. Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. Nucleic Acids Res. 29, 2581–2593.

Thelander, M., Olsson, T., Ronne, H., 2005. Effect of the energy supply on filamentous growth and development in Physcomitrella patens. J. Exp. Bot. 56, 653–662. doi:10.1093/jxb/eri040

Tillemans, V., Dispa, L., Remacle, C., Collinge, M., Motte, P., 2005. Functional distribution and dynamics of Arabidopsis SR splicing factors in living plant cells. Plant J. 41, 567–582. doi:10.1111/j.1365-313X.2004.02321.x

Tillemans, V., Leponce, I., Rausin, G., Dispa, L., Motte, P., 2006. Insights into Nuclear Organization in Plants as Revealed by the Dynamic Distribution of Arabidopsis SR Splicing Factors. Plant Cell 18, 3218–3234. doi:10.1105/tpc.106.044529

Tognetti, V.B., Mühlenbock, P., Van Breusegem, F., 2012. Stress homeostasis – the redox and auxin perspective. Plant Cell Environ. 35, 321–333. doi:10.1111/j.1365-3040.2011.02324.x

Tohge, T., Wendenburg, R., Ishihara, H., Nakabayashi, R., Watanabe, M., Sulpice, R., Hoefgen, R., Takayama, H., Saito, K., Stitt, M., Fernie, A.R., 2016. Characterization of a recently evolved flavonol-phenylacyltransferase gene provides signatures of natural light selection in Brassicaceae. Nat. Commun. 7. doi:10.1038/ncomms12399

Tuteja, N., 2007. Abscisic Acid and Abiotic Stress Signaling. Plant Signal. Behav. 2, 135–138.

Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.-S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., Zeeberg, B.R., Kane, D., Weinstein, J.N., Blume, J., Darnell, R.B., 2005. Nova regulates brain-specific splicing to shape the synapse. Nat. Genet. 37, 844–852. doi:10.1038/ng1610

Usuka, J., Zhu, W., Brendel, V., 2000. Optimal spliced alignment of homologous cDNA to a genomic DNA template. Bioinforma. Oxf. Engl. 16, 203–211.

Valvekens, D., Van Montagu, M., Van Lijsebettens, M., 1988. Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection. Proc. Natl. Acad. Sci. U. S. A. 85, 5536–5540.

Vandenbroucke, I.I., Vandesompele, J., Paepe, A.D., Messiaen, L., 2001. Quantification of splice variants using real-time PCR. Nucleic Acids Res. 29, e68.

Vonarx, E.J., Howlett, N.G., Schiestl, R.H., Kunz, B.A., 2002. Detection of Arabidopsis thaliana AtRAD1 cDNA variants and assessment of function by expression in a yeast rad1 mutant. Gene 296, 1–9. doi:10.1016/S0378-1119(02)00869-7

Wahl, M.C., Will, C.L., Lührmann, R., 2009. The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell 136, 701–718. doi:10.1016/j.cell.2009.02.009

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

Wang, P., Xue, L., Batelli, G., Lee, S., Hou, Y.-J., Oosten, M.J.V., Zhang, H., Tao, W.A., Zhu, J.-K., 2013. Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. Proc. Natl. Acad. Sci. 110, 11205–11210. doi:10.1073/pnas.1308974110

Warf, M.B., Berglund, J.A., 2010. The role of RNA structure in regulating pre-mRNA splicing. Trends Biochem. Sci. 35, 169–178. doi:10.1016/j.tibs.2009.10.004

Will, C.L., Lührmann, R., 2011. Spliceosome Structure and Function. Cold Spring Harb. Perspect. Biol. 3. doi:10.1101/cshperspect.a003707

Wingler, A., Masclaux-Daubresse, C., Fischer, A.M., 2009. Sugars, senescence, and ageing in plants and heterotrophic organisms. J. Exp. Bot. 60, 1063–1066. doi:10.1093/jxb/erp067

Wingler, A., Roitsch, T., 2008. Metabolic regulation of leaf senescence: interactions of sugar signalling with biotic and abiotic stress responses. Plant Biol. 10, 50–62. doi:10.1111/j.1438-8677.2008.00086.x

Wingler, A., von Schaewen, A., Leegood, R.C., Lea, P.J., Paul Quick, W., 1998. Regulation of Leaf Senescence by Cytokinin, Sugars, and Light. Plant Physiol. 116, 329–335.

Wong, M.M., Chong, G.L., Verslues, P.E., 2017. Epigenetics and RNA Processing: Connections to Drought, Salt, and ABA? Methods Mol. Biol. Clifton NJ 1631, 3–21. doi:10.1007/978-1-4939-7136-7_1

Wu, H.-P., Su, Y., Chen, H.-C., Chen, Y.-R., Wu, C.-C., Lin, W.-D., Tu, S.-L., 2014. Genome-wide analysis of light-regulated alternative splicing mediated by photoreceptors in Physcomitrella patens. Genome Biol. 15, R10. doi:10.1186/gb-2014-15-1-r10

Xing, Y., Lee, C.J., 2005. Protein Modularity of Alternatively Spliced Exons Is Associated with Tissue-Specific Regulation of Alternative Splicing. PLoS Genet. 1. doi:10.1371/journal.pgen.0010034

Zhang, H., Lin, C., Gu, L., 2017. Light Regulation of Alternative Pre-mRNA Splicing in Plants. Photochem. Photobiol. 93, 159–165. doi:10.1111/php.12680

Zhang, R., Calixto, C.P.G., Marquez, Y., Venhuizen, P., Tzioutziou, N.A., Guo, W., Spensley, M., Entizne, J.C., Lewandowska, D., ten Have, S., Frei dit Frey, N., Hirt, H., James, A.B., Nimmo, H.G., Barta, A., Kalyna, M., Brown, J.W.S., 2017. A high quality Arabidopsis transcriptome for accurate transcript-level analysis of alternative splicing. Nucleic Acids Res. 45, 5061–5073. doi:10.1093/nar/gkx267

Zhang, R., Calixto, C.P.G., Tzioutziou, N.A., James, A.B., Simpson, C.G., Guo, W., Marquez, Y., Kalyna, M., Patro, R., Eyras, E., Barta, A., Nimmo, H.G., Brown, J.W.S., 2015. AtRTD – a comprehensive reference transcript dataset resource for accurate quantification of transcript-specific expression in Arabidopsis thaliana. New Phytol. 208, 96–101. doi:10.1111/nph.13545

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., Lee, J.T., 2010. Genome-wide identification of Polycomb-associated RNAs by RIP-seq. Mol. Cell 40, 939–953. doi:10.1016/j.molcel.2010.12.011

Zhou, Z., Fu, X.-D., 2013. Regulation of splicing by SR proteins and SR protein-specific kinases. Chromosoma 122, 191–207. doi:10.1007/s00412-013-0407-z

Zhu, W., Schlueter, S.D., Brendel, V., 2003. Refined annotation of the Arabidopsis genome by complete expressed sequence tag mapping. Plant Physiol. 132, 469–484. doi:10.1104/pp.102.018101