Review Article

Shaping the Arabidopsis Transcriptome through Alternative Splicing

Dorothee Staiger

Molecular Cell Physiology, Faculty of Biology, Bielefeld University, Universitaetsstrasse 25, 33615 Bielefeld, Germany

Correspondence should be addressed to Dorothee Staiger; dorothee.staiger@uni-bielefeld.de

Received 16 November 2014; Accepted 17 January 2015

Academic Editor: Tomotsugu Koyama

Copyright © 2015 Dorothee Staiger. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alternative splicing is a molecular tool of the cell to generate more than one messenger RNA from the same gene. Through variable combinations of exons blueprints for different proteins are assembled from one and the same pre-messenger RNA, thus increasing the complexity of the proteome. Moreover, through alternative splicing different transcript variants with different stabilities and different regulatory motifs can be generated, leading to variation in the transcriptome. The importance of alternative splicing in plants has been increasingly recognized in the last decade. Alternative splicing has been found during abiotic and biotic stress and during development. Here, recent advancements in the understanding of alternative splicing in higher plants are presented. Mechanistic details and functional consequences of alternative splicing are discussed with a focus on the model plant Arabidopsis thaliana.

1. Introduction

The realization of the information encoded in the eukaryotic genome can be fine-tuned at multiple steps once transcription has been initiated in the nucleus. This layer of regulation includes capping of the mRNA 5’ end, splicing, processing at the 3’ end including addition of the poly(A) tail, RNA editing, covalent modification of bases, export from the nucleus to the cytoplasm, mRNA degradation, and localization of mRNAs, as well as the control of translation initiation [1]. These regulatory steps are collectively referred to as posttranscriptional control. The importance of this posttranscriptional regulation for all aspects in the life of a plant is increasingly being recognized [2–8].

The original view that each eukaryotic protein is encoded by a dedicated gene was compromised by the discovery of intronic sequences in an adenovirus gene that are no longer present in the mRNA [9, 10]. Subsequently, it was found that one gene can have many such introns that disrupt the open reading frame and need to be removed from the pre-mRNA through the splicing process [11, 12]. Short consensus sequences, the approx. nine nucleotide-long 5’ splice site with a GU dinucleotide at the beginning of the intron, and the 3’ splice site with an AG dinucleotide at the end of the intron and a region enriched in pyrimidines located further upstream delimit the introns (Figure 1). Additional sequence motifs located in the mRNAs serve as docking sites for a plethora of RNA-binding proteins and accessory factors involved in splicing [13].

The removal of introns from one and the same pre-mRNA does not always follow the default pathway of removing the intron and joining the flanking exons, designated as constitutive splicing [14, 15]. Rather, through the variable use of splice sites exonic sequences can be lost or intronic sequences can remain in the mRNA, thus changing the composition of the mRNA and affecting the open reading frames of the encoded proteins (Figure 2(a)). This variation in the pre-mRNA splicing patterns is designated as alternative splicing [15, 16].

Here, recent advancements in our understanding of alternative splicing in higher plants are reviewed. Mechanistic details and functional consequences of alternative splicing are discussed, focusing on the model plant Arabidopsis thaliana. For comparative discussions of the splicing process in other
model plants including rice, Brachypodium, or the moss Physcomitrella patens, readers are referred to recent excellent summaries [17–20].

2. Alternative Splicing

In the process of alternative splicing, not every splice site is used during pre-mRNA maturation. Rather, exons can be removed together with flanking introns, designated exon skipping (Figure 2(a)), or introns can stay in the pre-mRNA, designated intron retention. Through the use of alternative 5' splice sites or alternative 3' splice sites variable portions of introns can be removed and variable portions of the exons remain in the mRNA. This variation in the splicing patterns entails major consequences for the resulting mRNA isoforms. The encoded proteins can be composed of distinct domains and thus have different functions [16]. At the RNA level, variation in the sequence of alternative splice isoforms can lead to differences in subcellular localization, stability, translatability, or regulation by microRNAs (miRNAs). Finally, alternative splice isoforms can be recognized as “aberrant” and channeled into RNA decay pathways. For example, if alternative splicing leads to retention of an intron, the open reading frame may terminate at a stop codon in the intron (Figure 2(b)). Furthermore, if alternative splicing removes a sequence whose length cannot be divided by three, the open reading frame is shifted and may also include a premature termination codon (PTC). Such PTCs can be recognized by a surveillance mechanism designated as nonsense-mediated decay (NMD) that rapidly removes PTC-containing transcripts from the cellular transcriptome [21–25]. Thus, through linkage with NMD, alternative splicing can lead to quantitative changes in overall transcript levels [23, 26].

In the early days of expressed sequence tag (est) sequencing occasionally cDNAs were found that retained either an entire intron or a part of an intron [27, 28]. Around a decade ago, 1.2% of all Arabidopsis genes were estimated to be alternatively spliced [29]. Whereas standard microarrays based on exonic hybridization probes did not allow discriminating between spliced and unspliced transcript variants, the use of tiling arrays enabled the detection of intronic sequences and the differential incorporation of exons in transcript variants [30, 31]. Furthermore, a high resolution Arabidopsis alternative splicing panel was developed based on reverse transcription followed by PCR-amplification using primers flanking known alternative splice sites [32]. This system proved exquisitely sensitive to detect changes in the ratios of known alternatively spliced transcript variants [33–38]. At the same time, however, previously unknown alternative splice isoforms were detected for certain splicing events. Importantly, the most recent data based on whole transcriptome sequencing that allow de novo detection of unknown transcript variants point to 61% of all Arabidopsis genes to be alternatively spliced [39–42]. For comparison, in mammals around 95% of all multixen genes undergo alternative splicing [43]. It is assumed that also in plants new splice variants will continue to be discovered, as alternative splicing varies with developmental state, organ, and tissue and cell type and upon exposure to biotic and abiotic stress [39, 44–48].

More recently, it was found that not only precursor transcripts of protein coding genes but also precursors of miRNAs can undergo alternative splicing. MiRNAs are 20–24 nucleotide-long RNAs that regulate the expression of cognate target mRNAs. They are generated from transcripts with internal stem-loop structures by RNA polymerase II. In plants, these pri- (primary-) miRNAs are converted to stem-loop pre- (precursor-) miRNAs and further processed...
to miRNA/miRNA* duplexes by DICER-LIKE1 (DCL1) [49]. Thus, the pri-miRNAs undergo both splicing and processing. Consequently, both processes can interact or interfere with each other.

3. The Players

In mammalian cells, major insights into the splicing process have been obtained through biochemical purification of splicing factors and their extensive characterization. The recognition of the splice sites, cutting, and religation are handled by the spliceosome, a high molecular weight molecular machine in the cell [50–52]. The spliceosome is assembled in its functional form in a highly ordered process at each intron. Key components are five small nuclear ribonucleoproteins (snRNPs) consisting of short uridine-rich RNAs and designated small nuclear RNAs (snRNAs) U1, U2, U4, U5, and U6, as well as a distinct set of proteins. The core particles of the U1, U2, U4, and U5 snRNPs are formed by Sm proteins. The U6 snRNPs contains the related Sm-like proteins Lsm2 (like Sm 2) to Lsm8 [53]. Some Sm proteins (SmB, SmB’, SmD1, and SmD3) contain arginine-glycine (RG) rich motifs that can be posttranslationally modified by arginine dimethylation through PROTEIN METHYLTRANSFERASE 5 [54].

The first step in the splicing process is the interaction of the U1 snRNP with the 5’ splice site. This is accomplished via base pairing of the U1 snRNA to complementary sequences across the exon/intron border. At the end of the intron, U2 auxiliary factor (U2AF) interacts with the 3’ splice site. U2AF is a heterodimer of a 35kDa and a 65kDa subunit. U2AF35 binds to the intron/exon border, whereas U2AF65 binds to the so-called poly(pyrimidine) tract, a stretch of pyrimidine nucleotides located in close vicinity upstream of the intron/exon border. Subsequently, the U2 snRNP binds to the so-called branch point in the intron, again via base pairing of U2 snRNA. An assembly of the U4, U5, and U6 snRNPs is then docked onto U2 snRNP. After major rearrangements and release of the U1 and U4 snRNPs the splicing reaction takes place.

The splicing process in plants has not been studied extensively at the biochemical level, not least because it has not yet been possible to establish an in vitro system faithfully recapitulating the splicing reaction. Whereas some plant introns have been shown to be accurately spliced in nuclear extracts from HeLa cells, a mammalian intron was not efficiently spliced in plants, pointing to differences in intron recognition [55–57]. This may partly be due to the vastly differing size with introns from animals being on average 5kb in length compared to plant introns of 160bp [58]. Furthermore, the sequence composition is thought to be important, with a high UA content of plants introns favouring efficient splicing [57, 59]. More recently, a higher UA content relative to the exons was also found for animal introns [60].

The identification of counterparts of most of the splicing factors and the U snRNAs from mammals and yeast in the Arabidopsis genome suggests that central principles of the splice reaction are conserved [61, 62]. Notably, several paralogs have often been identified, pointing to some redundancy within the splicing factor families.

4. Regulation of Alternative Splicing

Whereas the conserved splice sites determine where to splice, pre-mRNAs contain additional sequence motifs that, through interaction with cognate RNA-binding proteins, dictate where splice sites are to be used (Figure 1). These motifs are collectively referred to as splicing enhancers and splicing silencers and can be located in intronic or exonic sequences [63]. The interplay between such cis-active motifs and the cognate trans-acting factors that dictate the splicing outcome has been named the splice code.

A number of Arabidopsis proteins have been identified that can regulate alternative splicing in trans. They mostly belong to the two main classes of splicing regulators known from metazoa, the serine-arginine-rich (SR) proteins and the heterogeneous nuclear ribonucleoproteins.

4.1. SR Proteins. The SR proteins contain one or two RNA recognition motifs (RRMs) that bind to the cis-acting motifs in the mRNA and an RS domain of at least 50 amino acids that mediates protein-protein interaction [64, 65]. Arabidopsis has eighteen SR proteins, some of them being orthologs of metazoan SR proteins whereas others are plant-specific [64, 66, 67]. Ectopic expression of several SR proteins produces a range of morphological and physiological phenotypes and impacts alternative splicing of a suite of other transcripts [68, 69]. Recently, alternative splicing of the SR protein RS31 in different light regimes has been shown to be governed by the reduced pool of plastoquinone in photosynthetic electron transport in the chloroplasts [70]. This suggests that retrograde signaling from the chloroplast to coordinate nuclear gene expression programs can also rely on alternative splicing.

In particular, several SR proteins undergo autoregulation through alternative splicing to PTC-containing isoforms encoding truncated proteins devoid of one or more domains [69, 71–73]. Half of the PTC-containing isoforms are degraded via NMD and thus lead to alterations in transcript abundance [74]. Several alternative splicing events in SR genes are evolutionarily conserved, underscoring their physiological relevance [75, 76]. Recently, regulatory elements within an intron of SCL33 have been identified that mediate autoregulation of SCL33 alternative splicing [77]. SCL33 binds to a conserved sequence motif in this intron with four GAAG motifs that are necessary for correct splicing. This represents an important step towards understanding the plant splice code.

Initially, SR45 and SR45a that harbor two RS domains, one of each side of the RRM, were also classified as SR proteins but have recently been renamed to SR-like proteins [64, 78]. Nevertheless, SR45 has been extensively characterized. It interacts with the U1 snRNP protein U1-70K involved in the recognition of 5’ splice sites and with U2AF35 involved in the recognition of 3’ splice sites. SR45 and U2AF35 bind to intron 10 of the SR30 pre-mRNA, whose splicing is altered in sr45 mutant plants [79]. The binding site of SR45 localizes to the 5’ end of the intron and U2AF35 binds to the 3’ end of the intron, suggesting that SR45 helps to recruit U1 snRNP to 5’ splice site and U2AF to the 3’ splice sites.
The functional relevance of individual splice isoforms of SR proteins was nicely demonstrated for the sr45 mutant that exhibits developmental abnormalities, including narrow leaves and petals, altered number of petals and stamens, and short roots [80]. The use of alternative 3′ splice sites generates two alternative splice isoforms which differ by only eight amino acids and yet perform different functions: expression of the individual isoforms differentially complement the defects of the sr45 mutant in petal development or root growth [81]. Additionally, the sr45 mutant displays hypersensitivity to glucose and abscisic acid, which is complemented by both splice isoforms [82].

4.2. Heterogeneous Nuclear Ribonucleoproteins. The second class of splicing regulators corresponds to the heterogeneous nuclear ribonucleoproteins (hnRNPs) [83, 84]. Chief among those are the polypyrimidine tract binding proteins (PTBs) that interact with motifs rich in pyrimidines in introns. PTBs have been comprehensively characterized in mammals where they either activate or repress the usage of specific splice sites in the vicinity of their binding sites. PTBs can act in a combinatorial manner with other hnRNPs or SR proteins [85, 86].

The Arabidopsis PTB1 and PTB2 proteins show negative autoregulation by alternative splicing of their own pre-mRNAs where inclusion of a PTC-containing cassette exon creates an NMD substrate [87]. Furthermore, PTB1 and PTB2 reciprocally cross-regulate through the same mechanism. A genome-wide RNA-seq analysis of transgenic Arabidopsis plants demonstrated changes in the splicing pattern of a plethora of transcripts upon overexpression or knock-down of PTB1 or PTB2 [38, 41]. The more distantly related PTB3 did not exhibit a significant impact on alternative splicing [41].

To delineate the PTB-responsive sequence motifs within its splicing targets, a splice reporter was used that allows monitoring of skipping of an exon in response to increased PTB1 levels in a Nicotiana plumbaginifolia protoplast transient expression assay [38]. The negative effect of PTB1 on inclusion of this exon is mediated by motifs in the intron upstream of the alternative exon, and a pyrimidine rich motif is located between the branch point and the 3′ splice site. Notably, U2AF65 acts antagonistically to PTB1, promoting inclusion of the exon, similar to what has been found in mammals [38].

Furthermore, AtGRP7 (Arabidopsis thaliana glycine-rich RNA binding protein 7) and AtGRP8 are simplified versions of mammalian hnRNPs proteins. A single N-terminal RRM mediates RNA binding and chaperone activity of AtGRP7 [88–92]. The C-terminal region is enriched in glycine residues and shows similarity to M9 domains implicated in nuclear trafficking [93–95]. Indeed, AtGRP7 is not only imported from the cytoplasm into the nucleus but also exported from the nucleus to the cytoplasm and thus is a shuttling protein [96, 97]. Recently it was shown that the glycine stretch has an accessory role in RNA binding [98].

Through reverse genetics, AtGRP7 and AtGRP8 have been shown to play a regulatory role in the circadian timing system [27, 99, 100]. Both AtGRP7 and AtGRP8 negatively autoregulate the circadian oscillations of their own transcripts through alternative splicing-NMD [35, 88, 101]. Furthermore, both AtGRP7 and AtGRP8 reciprocally cross-regulate, similar to PTB1 and PTB2. AtGRP7 regulates a range of downstream targets, a third of which are circadianally regulated themselves [102, 103]. This suggests that the clock-regulated AtGRP7 feedback loop transduces timing information from the core clockwork [104].

Using the high resolution RT-PCR alternative splicing panel developed by the Brown laboratory it was shown that AtGRP7 has a more global impact on alternative splicing [105]. In RNA immunoprecipitation experiments, a suite of transcripts coprecipitated with an AtGRP7-GREEN FLUORESCENT PROTEIN (GFP) fusion protein but not with AtGRP7-R49Q-GFP carrying a mutation of the conserved arginine [97] of the RRM, showing for the first time that an Arabidopsis splicing regulator affects alternative splicing by in vivo binding to its targets [105, 106]. This very arginine is ADP-ribosylated by a Pseudomonas syringae type III effector protein upon bacterial infection, suggesting that virulent bacteria interfere with plant immune responses by disabling AtGRP7 binding to its targets [107–110].

4.3. Other Splicing Factors

4.3.1. snRNP and Related Proteins. Splicing defects were detected in mutants of spliceosomal Lsm proteins. The lsm5 mutant was originally identified because of its enhanced sensitivity to abscisic acid (ABA) and thus is also known as sad1 (supersensitive to ABA and drought 1). It shows reduced levels of U6 snRNA and accumulation of unspliced pre-mRNAs, suggesting that it has a role in pre-mRNA splicing by contributing to U6 stability [111]. Recently, a widespread effect of Lsm5 on splicing has been described [112]. Mutants defective in Lsm4 also show splicing defects [113].

STA1 (STABILIZED 1) shows homology to the human U5 snRNP-associated I02 kDa protein [114]. STA1 is required for correct splicing of the cold-induced COR15A transcript and the mutant defective in STA1 is cold-sensitive.

Another protein complex that is essential for the catalytic activity of the spliceosome in mammals and yeast is the NineTeen complex, named for its core component Prp19 [115]. Several proteins with homology to components of the NineTeen complex have been identified in Arabidopsis. Among those, the MOS (MODIFIER OF sncl) proteins have been identified in a suppressor screen for mutants that revert the constitutive pathogen resistance of the sncl (SUPPRESSOR OF NPR1-1, CONSTITUTIVE) mutant with the constitutively active R (resistance) protein SNC1 [116]. MOS4 shows homology to human Breast Cancer-Amplified Sequence 2 (BCAS2) and interacts with AtCDC5 (CELL DIVISION CYCLE 5) Myb-transcription factor, the WD-40 repeat PRL1 (PLEIOTROPIC REGULATORY LOCUS 1), and the pair of related MAC3A (MOS4 associated complex 3A) and MAC3B proteins with sequence homology to Prp19 [117, 118]. In the mos4, cdc5, and mac3a mac3b mutants the SNC1 splicing pattern is altered, providing evidence that the MOS complex, similar to the NineTeen complex, is involved in alternative splicing [119]. Another protein associated with the MOS complex is...
MAC5A with similarity to human RBM22 (RNA Binding Motif Protein 22) [120]. As RBM22 interacts with the U6 snRNA and pre-mRNA and participates in splicing, MAC5A may also be involved in splicing in Arabidopsis [62, 121].

ROOT INITIATION DEFECTIVE1 (RID1) encodes a DEAH-box RNA helicase similar to the splicing factors Prp2 in yeast and DEAH box polypeptide8 (DHX8) in humans [122]. Indeed, in the rid1-1 mutant several transcripts were aberrantly spliced.

SPLICOSOMAL TIMEKEEPER LOCUS1 (STIPL1) is a homolog of TFPII in humans and Ntr1p in yeast involved in spliceosome disassembly [123]. In the stipl1 mutant splicing of many introns is reduced. At the physiological level, stipl1 shows a long period of the circadian clock. In accordance with this, the accumulation of circadian transcripts is altered. In particular, retention of intron 3 in the core clock gene PSEUDORESPONSE REGULATOR 9 (PRR9) is increased.

Another long period mutant is defective in the homologue of mammalian SNW/Ski-interacting protein domain protein (SKIP) [124]. Arabidopsis SKIP interacts with the splicing factor SR45 and has a global effect on splicing. In particular, the core clock genes PRR7 and PRR9 are aberrantly spliced.

4.3.2. Factors for mRNP Formation. The two subunits of the CAP binding complex that interacts with the cap structure at the mRNA 5' end have been shown to affect splicing, CBP80, also known as ABCISIC ACID HYPERSENSITIVE1, and CBP20 contribute to the regulation of alternative splicing and preferentially affect alternative splicing of the first intron, particularly at the 5’ splice site [33, 125]. Both CBP20 and CBP80 interact with the zinc finger protein SERRATE (SE), a component of miRNA biogenesis pathway. Like CBC20 and CBC80, SE affects alternative splicing, and it also preferentially targets the 5’ splice site of first introns [125, 126].

4.3.3. Modifiers of Splicing Factors. An important feature modulating the activity of splicing factors is posttranslational modification. Thus, alterations in the corresponding enzymes can also affect alternative splicing. Mutants deficient in PROTEIN ARGININE METHYLTRANSFERASE 5 (PRMT5) have a defect in circadian timekeeping, showing a longer period of leaf movement and gene expression rhythms [34, 127]. Substrates for PRMT5 have been identified and include spliceosomal U snRNP proteins [128]. Furthermore, PRMT5 affects splicing including splicing of PRR9, suggesting that the circadian defect in prmt5 can partly be attributed to changes in PRR9 splice patterns [34, 128].

A large proportion of the SR proteins can undergo phosphorylation [129, 130]. Phosphorylation and dephosphorylation have been shown to impact the movement of SR proteins within the nucleus [66]. In particular, ectopic expression of a LAMMER protein kinase, named for a conserved EHLAMMERILG motif in their catalytic subdomain, leads to aberrant splicing of SR30 and SR34 as well as other transcripts [131]. Notably, the AFC2 transcript encoding the Arabidopsis LAMMER kinase undergoes multiple alternative splicing events itself, pointing to the possibility of a feedback regulation.

5. Intron Retention and Alternative Splicing-NMD

Early on, intron retention has been considered as the most prevalent type of alternative splicing events in plants, whereas intron retention represents only a minor fraction of alternative splicing events in metazoan [39]. In a recent Arabidopsis transcriptome study it was found that intron-retained transcripts have a very low read coverage [40].

Alternative splicing often generates transcript isoforms with characteristic NMD features. These include PTCs located more than around 50 nucleotides upstream of exon junctions or long 3' UTRs. Moreover, splicing in the 5' UTR can affect upstream open reading frames with uORFs of more than amino acids triggering NMD [35]. In this way, alternative splicing linked to NMD makes a substantial contribution to transcript levels [26].

In Arabidopsis 13–18% of the genes undergo alternative splicing and NMD [35, 132]. Notably, many factors involved in alternative splicing themselves undergo regulation by alternative splicing and NMD, as shown for SR proteins, PTBs and GRPs [83]. The resulting changes in the level of the splicing factors have consequences on their splicing substrates. Thus, AS-NMD of splicing factors appears a mechanism to coordinate alternative splicing of large sets of downstream targets. The functional importance of this type of autoregulation is underscored by the observation in mammals that alternative splicing of SR genes associated with NMD of the alternative splice variants is frequently mediated by regulatory elements that show a very high conservation in the genome [133, 134].

Intron-retained transcripts show features that would allow degradation via the NMD pathway such as a PTC in a defined distance upstream of an exon junction, long 3' UTRs, or uORFs in the 5' UTR [35]. It was found that the majority of intron retention transcripts were not NMD sensitive despite displaying NMD features, explaining their prevalence. For example, in the case of AtGRP7 and AtGRP8 the pre-mRNAs retaining the entire intron including a PTC were not NMD sensitive whereas the alternative splice variants retaining only the first part of the intron underwent NMD [35, 89]. This suggested that the pre-mRNAs may escape detection by the NMD pathway which is thought to occur during a pioneer round of translation in the cytosol. Indeed, hybridization with molecular beacons and confocal microscopy allowed the detection of transcripts for the SR protein RSZZ33 retaining entire introns exclusively in the nucleus [135].

6. Functional Outcome of Alternative Splicing

Alternative splice isoforms of one pre-mRNA can encode different proteins, thus greatly enlarging the coding capacity of the genome (Figure 2(a)). For example, the encoded proteins can be composed of distinct domains and thus possess different functions, differentially engage in protein-protein interactions, and contribute to alternative splicing.
interaction, or localize to different subcellular compartments. We are beginning to understand the physiological relevance of distinct protein isoforms resulting from regulated alternative splice events.

6.1. Protein Variants Encoded by Alternative Splice Isoforms. The RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE ACTIVASE (RCA) transcript undergoes circadian oscillations in steady-state abundance as well as alternative splicing [34]. A long transcript isoform codes for a protein whose activity is regulated by light intensity, whereas the short transcript isoform codes for a protein whose activity is light independent [136]. Alternative splicing of the mRNA isoform that encodes the light-regulated protein increases during the day [34].

The Arabidopsis heat shock transcription factor HsfA2 undergoes temperature dependent alternative splicing. In plants exposed to 37°C, a 31 bp mini-exon within the conserved intron in the DNA-binding domain is spliced into the transcript. This exon introduces a PTC and targets the alternative splice isoform HsfA2 II to NMD, thus providing a mechanism to adjust the level of active HsfA2 protein [137]. At 42°C, the alternative splice variant HsfA2 III is generated encoding a shorter protein, while HsfA2 II decreases [138]. This truncated protein retains the DNA-binding domain, localizes to the nucleus, and binds to the HsfA2 promoter, pointing to a positive autoregulatory loop of HsfA2 expression through alternative splicing.

Alternative splicing also impacts the activity of the JAZ (JASMONATE ZIM-domain) proteins that inhibit jasmonic acid (JA) responsive gene expression by sequestering MYC2, the key transcription factor involved in JA signaling. For several JAZ proteins, intron retention leads to truncated protein variants that show reduced interaction with the JA receptor CORONATINE INSSENSITIVE 1 in the presence of the active JA-Ile conjugate and thus are resistant to proteosomal degradation [139, 140]. These small interfering peptides are examples of so-called micropeptides that have been named in analogy to miRNAs. The JAZ micropeptides may serve to limit defence against herbivores and necrotrocytes.

Another type of variation of the protein sequence results from the repeated 3′ splice site dinucleotides AG within three nucleotides, leading to the generation of NAGNAG tandem splice acceptor sites [141, 142]. The resultant protein variants differ by only one amino acid. In Arabidopsis around 7000 introns with such NAGNAG acceptor sites have been identified mostly in DNA-binding proteins and splicing factors. A survey of the Arabidopsis SR protein coding genes predicted a suite of NAGNAG acceptor sites and eight of them were experimentally validated [143]. The ratios of splice site usage changed in different organs and in response to heat shock and cold shock with a coregulation of all analyzed genes, suggesting that the differential effects on NAGNAG alternative splicing in SR and SR-related genes are organ- and condition-specific rather than gene-specific.

6.2. Altered Localization of Protein Variants. Altered subcellular localization of splice variants was found for several proteins. A ubiquitously expressed YUCCA4 transcript isoform encodes an auxin biosynthetic enzyme that localizes to the cytoplasm, whereas a second flower-specific transcript isoform codes for a protein localized at the endoplasmic reticulum (ER) [144].

In the case of bZIP60, a transcription factor associated with the ER, a 23-nt intron is removed in response to ER stress elicited by misfolded proteins, leading to a frame shift, introduction of a PTC, and loss of the membrane anchoring domain [145]. The resulting smaller protein variant relocates to the nucleus and activates ER stress-inducible genes.

Alternative splicing of a Major Facilitator Superfamily transporter, Zinc-Induced Facilitator-Like 1 (ZIFL1), leads to two protein isoforms, both of which modulate \( H^{+} \)-coupled \( K^{+} \) transport, but differ in tissue distribution and subcellular localization [146]. The protein encoded by full length ZIFL1.1 transcript is targeted to the tonoplast of root cells. The ZIFL1.3 transcript, generated through an alternative 3′ splice site, two nucleotides downstream of the authentic 3′ splice site, causing a frame-shift and introduction of a PTC, leads to the loss of the two last C-terminal membrane-spanning domains and localisation of the protein to the plasma membrane of leaf stomatal guard cells. Differential complementation of the zif1.1 drought sensitivity and auxin-related defects shows that the full length ZIFL1.1 protein influences cellular auxin efflux and polar auxin transport in roots, whereas the truncated ZIFL1.3 isoform regulates stomatal movement [146].

The INDOLE-3-BUTYRIC ACID RESPONSE5 (IBR5) gene encodes a dual specificity phosphatase associated with MAP kinase signaling that is involved in the response to auxin. Alternative splicing gives rise to a the IBR5.1 transcript isoform encoding a protein with phosphatase activity that localizes to the cytosol and the nucleus, and the IBR5.3 isoform encoding a protein lacking phosphatase activity that exclusively localizes to the nucleus [147]. Complementation of ibr5 mutants revealed that the encoded IBR5.1 and IBR5.3 have both distinct and overlapping functions. It remains to be shown how this correlates with the differential subcellular localization.

ZIF2 transporter localizes at the tonoplast of root cortex cells and mediates \( Zn^{2+} \) efflux [148]. Two splice variants encoding the same protein have been detected that differ by an intron in the 5′ UTR. Under high \( Zn^{2+} \) concentrations the long variant ZIF2.2 retaining the intron is made. When expressed in transgenic plants, this construct confers a higher tolerance against excess \( Zn^{2+} \). This is due to an increased translation of the intron containing splice variant in response to \( Zn^{2+} \). The enhanced translation is caused by a stem loop structure upstream of the translational start codon that is present in the intron containing ZIF2.2 but not in the intron less ZIF2.1 alternative splice variant [148].

7. Regulation of Alternative Splicing by RNA Secondary Structure

Single-stranded RNA undergoes a high degree of secondary and tertiary structure formation through intramolecular base pairing. The dynamics of partially double-stranded regions in stem-loops can mask or unmask cis-active motifs within
the mRNA. Alternative splicing can be heavily influenced by RNA secondary structures if these structures alter the access to the splice sites or the splicing silencers or splicing enhancers. For example, a heterologous stem-loop interfered with the splicing reaction in a position-dependent manner [149]. Furthermore, control of alternative splicing by an mRNA secondary structure in response to a small metabolite has also been shown for Arabidopsis. Such structures are known as riboswitches. They undergo conformational changes upon metabolite binding with a concomitant impact on gene expression.

A plant thiamine pyrophosphate (TPP) riboswitch located in the 3’ UTR of the Arabidopsis thiamine biosynthetic THIAMINE C SYNTHASE gene mediates alternative splicing in response to elevated TPP levels [150, 151]. During this splicing event an mRNA 3’ processing site is removed, resulting in a longer transcript variant with reduced stability and thus ultimately limiting THIAMINE C SYNTHASE expression [150, 151].

In the case of TFIIIA, a secondary structure within the mRNA mimicking part of the 5S ribosomal RNA serves as a binding site for an alternative splicing regulator [152, 153]. The ribosomal L5 protein binds to the 5S rRNA in the ribosomes. Free L5 protein indicates an excess of L5 over 5S rRNA. L5 then binds to the 5S rRNA structural mimic in the TFIIIA mRNA, causing a shift from splicing to the PTC-containing transcript isoform to an mRNA variant that can be translated into TFIIIA protein, a transcription factor for RNA polymerase III transcription of ribosomal RNA precursors.

A long noncoding RNA regulates the activity of the alternative splicing regulators AtNSR (Arabidopsis thaliana nuclear speckle RNA-binding protein) [154]. This highly structured long noncoding RNA, AS competitor long noncoding RNA (ASCO-lncRNA), binds to AtNSR and prevents it from interaction with their splicing targets, thus modulating the splicing profile of AtNSR downstream targets.

Apart from these specialized structures, further information on how structure impacts alternative splicing patterns will come from determination of secondary structure at the genome level [155].

8. Crosstalk between Alternative Splicing and miRNA Mediated Regulation

Alternative splicing and miRNA-based regulation are connected at several levels.

Firstly, alternative splicing of pre-mRNAs can affect the miRNA target site within these pre-mRNAs and thus generate miRNA-susceptible or resistant transcripts. Computational simulation indeed suggests that the frequency of alternative splicing at miRNA binding sites of mRNAs is significantly higher than at other regions [156]. Such a differential regulation of transcript isoforms by miRNA was observed for a family member of the SQUAMOSA-PROMOTER BINDING PROTEIN LIKE (SPL) transcription factors that regulate developmental phase transitions in plants [157]. The SPLs are targets of miR156 and a decrease in miR156 levels during development allows accumulation of the SPLs. A differential regulation of SPL4 transcript isoforms by miR156 was observed [156, 158]. Only SPL4-1 contains a miR156 target site and the other three alternative splice variants are not subject to inhibition by miR156.

Recently, miRNA binding sites have also been detected in intronic sequences of mRNAs [159]. Thus, alternative splice isoforms with or without the corresponding introns can be differentially controlled.

Secondly, the primary transcripts of miRNAs also can undergo alternative splicing, thus affecting mature miRNA levels. Currently, around 30 Arabidopsis pri-miRNAs are known that harbour introns [160, 161]. Alternative splicing of pri-miR62 generates transcripts lacking part of the stem-loop and thus cannot give rise to mature pri-miR62 [161, 162]. For the intronic miR400 a temperature-induced alternative splicing event affects miRNA processing and causes the miRNA to be retained in the transcript [163]. The altered miR400 level in turn feeds back on the level of its host transcript.

Thirdly, pri-miRNAs also contain introns and undergo alternative splicing. A reciprocal interaction between splicing of introns in pri-miRNAs that are located downstream of the stem-loop and accumulation of the corresponding mature miRNA has been described for pri-miR72 and pri-miR63 [160, 164, 165]. Several splicing regulators that affect splicing of pre-miRNAs have been shown to affect splicing of pri-miRNAs as well, including the zinc finger protein SE, the CBP20 and CBP80 subunits of the CAP BINDING COMPLEX [125], the hnRNP-like protein AtGRP7 [166], and STAI [167].

9. Conclusion

Alternative splicing has emerged as a versatile strategy to increase the functionality and regulatory potential of the Arabidopsis genome in the last decade. The ongoing identification of a wealth of novel splice variants will lead to more refined gene models not present in the current version of the Arabidopsis genome (TAIR10) and will have to be incorporated in future databases [168]. A comprehensive description of the interplay between cis-active motifs and trans-acting factors will be crucial to decipher the splice code [169]. Furthermore, more in-depth characterization of individual alternative splicing events will be required to fully appreciate the regulatory potential of alternative splicing to bring about transcriptome changes during development and as a way to adjust to changing environmental conditions. For example, global differences in alternative splicing patterns appear to contribute to natural variation between ecotypes [170]. Another emerging area into Arabidopsis alternative splicing research addresses its connection to epigenetic regulation. Despite being historically categorized as a posttranscriptional layer of control, we now know that alternative splicing is tightly intertwined with transcription. The carboxyterminal domain of RNA polymerase II represents a binding platform for splicing factors that are then translocated along with the transcriptional machinery during generation of the nascent pre-mRNA [171]. The rate of transcription elongation dynamically varies in the cell and can determine whether or not particular splice sites are used under the given conditions.
So far, this has been extensively studied in mammals and yeast [172]. Furthermore, alternative splicing patterns can be determined at the chromatin level: epigenetic marks such as DNA methylation and histone modification ultimately impact the outcome of the splicing process [173–175], adding another level of versatility to the alternative splicing programs in the cell that remains to be fully explored in plants.

Conflict of Interests

The author declares no competing interests.

Acknowledgments

The author wishes to thank Katja Meyer for critical reading of the paper. Work in the author’s laboratory is supported by the DFG (STA653 and priority program 1530).

References

Advances in Botany 9


Advances in Botany


