On the physiological significance of alternative splicing events in higher plants

Raquel F. Carvalho, Carolina V. Feijão, and Paula Duque

Instituto Gulbenkian de Ciência, Rua da Quinta Grande, 6, 2780-156 Oeiras, Portugal

Address correspondence to: Paula Duque, Email: duquep@igc.gulbenkian.pt; Phone: +351 214464630; Fax: +351 21 4407970.

Keywords Alternative splicing, Gene expression, Plant physiology and development, Precursor mRNA, Splice isoforms
Abstract

Alternative splicing, which generates multiple transcripts from the same gene and potentially different protein isoforms, is a key posttranscriptional regulatory mechanism for expanding proteomic diversity and functional complexity in higher eukaryotes. The most recent estimates, based on whole transcriptome sequencing, indicate that about 95% of human and 60% of Arabidopsis multi-exon genes undergo alternative splicing, suggesting important roles for this mechanism in biological processes. However, while the misregulation of alternative splicing has been associated with many human diseases, its biological relevance in plant systems is just beginning to unfold. We review here the few plant genes for which the production of multiple splice isoforms has been reported to have a clear in vivo functional impact. These case studies implicate alternative splicing in the control of a wide range of physiological and developmental processes, including photosynthetic and starch metabolism, hormone signaling, seed germination, root growth and flowering, as well as in biotic and abiotic stress responses. Future functional characterization of alternative splicing events and identification of the transcripts targeted by major regulators of this versatile means of modulating gene expression should uncover the breadth of its physiological significance in higher plants.
Introduction

The majority of eukaryotic genes have their coding exonic regions interrupted by stretches of non-coding intronic sequences. RNA splicing removes these intervening sequences from the precursor-mRNA (pre-mRNA) and joins the exonic sequences to obtain mature functional mRNAs. The pre-mRNA splicing reaction, occurring cotranscriptionally, is carried out by the spliceosome, a multi-megadalton RNA-protein complex containing small nuclear ribonucleoproteins (snRNPs) and a large number of accessory non-snRNP proteins. Whereas the composition of the human, yeast and Drosophila spliceosome has been analyzed in great detail (reviewed in Wahl et al., 2009), plant spliceosomes have yet to be isolated. Nevertheless, searches for homologs of the human spliceosome proteome revealed that most splicing factors are encoded by the Arabidopsis genome (reviewed in Ru et al., 2008; Koncz et al., 2012).

The differential selection of splice sites through alternative splicing allows for different pre-mRNA fragments to be joined together, thus giving rise to multiple mature mRNAs from a single gene. Remarkably, the number of mRNA isoforms encoded by a single gene can vary from two to several thousand, as in the extreme case of the Drosophila Dscam gene that can potentially generate more than 38,000 splice forms (Schmucker et al., 2000). Some alternative splicing events are constitutive, with multiple isoforms being produced at the same ratios in all or most cell types, whereas others are regulated events. Such regulation can be tissue-specific, controlled by developmental cues or modulated in response to external stimuli, with the underlying biochemical mechanisms controlling splice site usage ranging from spliceosome composition and concentration of splicing factors to chromatin structure and transcription rates (reviewed in Nilsen and Graveley, 2010). Most alternative splicing events can be classified into four main types: exon skipping (or cassette exons), alternative 5' splice sites, alternative 3' splice sites and intron retention (Fig. 1A). In humans, the most frequent alternative splicing event is exon skipping, followed by alternative 3' or 5' splice sites, while intron retention is the least common (Fig. 1B) (Kim et al., 2007; Kim et al., 2008). By contrast, intron retention accounts for a surprising large proportion (over 30%) of the recorded alternative splicing events in Arabidopsis and rice, while cassette exons are extremely rare (less than 3%) (Fig. 1B) (Ner-Gaon et al., 2004; Ner-Gaon and Fluhr, 2006; Wang and Brendel, 2006; Barbazuk et al., 2008; Filichkin et al., 2010; Marquez et al., 2012).
Among the different pre-mRNA processing possibilities, alternative splicing is the most prevalent mechanism to generate transcriptome complexity and protein diversity in metazoans. The effects of alternative splicing on proteome expansion include the production of protein isoforms displaying loss or gain of function and/or altered subcellular localization, stability, enzymatic activity or posttranslational modifications (reviewed in Stamm et al., 2005). In addition, alternative splicing within the untranslated regions (UTRs) may affect the stability or translational efficiency of mRNA transcripts, providing a rapid means of altering protein expression (reviewed in Hughes, 2006). Finally, alternative splicing also often generates nonfunctional mRNAs containing premature termination codons (PTCs), which can be targeted for degradation by nonsense-mediated decay (NMD), an mRNA surveillance mechanism that prevents accumulation of truncated and potentially harmful proteins (reviewed in Nicholson et al., 2010). Coupling of alternative splicing to NMD can function as a negative feedback loop, effectively downregulating physiological transcripts to control the amounts of functional protein. Such a mechanism appears to be of significance in plant gene expression, with 13-18% of Arabidopsis intron-containing genes estimated to be potentially regulated by alternative splicing/NMD (Kalyna et al., 2012).

Recent studies based on next-generation sequencing (NGS) have estimated alternative splicing to occur in up to 95% of human multi-exon genes (Pan et al., 2008; Wang et al., 2008). Until 2001, alternative splicing studies in plants were mostly limited to the analysis of individual genes and, since less than 40 Arabidopsis genes had been described to undergo alternative splicing, this posttranscriptional regulatory mechanism was considered to occur infrequently in the plant kingdom (Reddy, 2001). Strikingly, the estimated number of alternatively-spliced plant genes has been rising steadily in the last decade (Brett et al., 2002; Haas et al., 2003; Kikuchi et al., 2003; Iida et al., 2004; Ner-Gaon et al., 2004; Nagasaki et al., 2005; Campbell et al., 2006; Wang and Brendel, 2006; Filichkin et al., 2010; Marquez et al., 2012), owing first to an increase in the number of available expressed sequence tags (ESTs) and full-length cDNA sequences and then to the recent “boom” of RNA sequence data generated with NGS technologies. In fact, the most recent transcriptome-wide analyses of alternative splicing in plants using RNA-seq indicate that at least 61% of Arabidopsis intron-containing genes (Marquez et al., 2012) and about 48% of rice genes (Lu et al., 2010) undergo alternative splicing. It is also clear that these estimate figures are bound to increase as more RNA-seq data becomes
available, particularly from various tissues, cell types, developmental stages or from plants grown under different biotic or abiotic stresses.

The prevalence of alternative splicing in many genomes raises the question of what fraction of these genes undergoes regulated splicing with biological consequences. In humans the misregulation of alternative splicing has been associated with many diseases, such as breast cancer, cystic fibrosis, Alzheimer’s or spinal muscular atrophy (reviewed in Blencowe, 2006; Ward and Cooper, 2010), but surprisingly little is known about the physiological relevance of this pre-mRNA processing mechanism in plant systems. Here, we review reported case studies for which the production of at least two splice variants from the same plant gene has clear in vivo functional impact.

**Plant Metabolism**

In 1989, Ogren and coworkers reported the first case of alternative splicing in plants. Their study showed that, in both spinach and *Arabidopsis*, the gene encoding the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) activase (RCA) produces two proteins, a shorter (RCA$_S$) and a longer (RCA$_L$) isoform, which differ at their C termini due to the usage of an alternative 5' splice site in an intron near the 3’ end of the primary transcript (Werneke et al., 1989). Rubisco, which catalyses the first step of both the photosynthetic CO$_2$ assimilation and the photorespiratory pathways, is activated by RCA, whose ATP hydrolysis activity promotes the dissociation of inhibitory sugar phosphates, removing them from the Rubisco’s active site. When the ADP/ATP ratio is high, the ATP hydrolysis activity of RCA is inhibited, thus preventing Rubisco activation (reviewed in Portis, 2003). In vitro experiments by Shen et al. (1991) revealed that although both isoforms of RCA are capable of activating Rubisco, the larger isoform is more sensitive to ADP inhibition. Zhang and Portis (1999) later showed that this is due to the structural differences in the C-terminal regions of the two isoforms — unlike RCA$_S$, the longer RCA$_L$ contains two cysteine residues with disulfide bonds, which allow for its redox regulation. Whereas reduction of RCA$_L$ via thiorredoxin-f leads to its activation, oxidation of its disulfide bonds reduces RCA activity. Accordingly, Zhang et al. (2002) observed that light modulation of Rubisco in transgenic *Arabidopsis* plants expressing exclusively RCA$_L$ is similar to the wild type, but replacement of the C-terminal Cys residues with Ala no longer allows for Rubisco downregulation in response to low light. On the other hand, plants expressing only the small RCA$_S$ isoform, which lacks the redox...
regulatory region, show unaltered Rubisco activity under saturating light, but are unable to downregulate the enzyme under limiting light conditions (Zhang et al., 2002). These findings clearly show that RCA_L is regulated by stromal redox changes in response to varying light intensity and suggest that the two RCA splice isoforms provide for optimization of photosynthesis under fluctuating light environments — while the large isoform may be produced as a means of controlling the energy expense of high RCA activity rates, the short isoform might allow for a rapid response to changes in light intensity. Intriguingly, this regulatory mechanism does not appear to be universal among plants, as tobacco, for instance, does not produce the larger RCA form (Salvucci et al., 1987).

Alternative splicing of the RCA gene has also been reported to play a role in the thermostolerance of photosynthesis. In fact, early reports document Rubisco activation as the primary target of photosynthesis inhibition by high temperatures (Weis, 1981; Kobza and Edwards, 1987; Feller et al., 1998), which may be overcome, to a certain extent, through acclimation processes (Law and Crafts-Brandner, 1999; Yamori et al., 2006). In vitro studies have shown that the spinach RCA_L is more thermostable than RCA_S, with the former protecting the shorter isoform from high temperature inactivation (Crafts-Brandner et al., 1997). However, despite these differences in thermostability, Salvucci et al. (2006) found that transgenic Arabidopsis lines individually expressing suboptimal levels of either the redox-regulated RCA_L or the RCA_S isoforms exhibit similar increased sensitivity to heat inhibition of photosynthesis and Rubisco activation, indicating that the heat stress response is determined by the amount of the isoform rather than the isoform itself. In this context, it is interesting to note that, under normal temperature conditions, while spinach and Arabidopsis contain equal amounts of each RCA isoform, in rice the shorter isoform is more abundant than the longer one (Salvucci et al., 1987; To et al., 1999; Salvucci et al., 2006; Wang et al., 2009). In addition, Wang et al. (2010) later reported that heat stress dramatically increases the levels of rice RCA_L. Importantly, wild-type rice plants overexpressing the RCA_S isoform under control temperature conditions show higher net photosynthetic rates and increased biomass, whereas under heat stress this is only accomplished upon overexpression of the RCA_L isoform (Wang et al., 2010).

These results provide further evidence for the functional significance of each RCA isoform, with the long isoform fulfilling a determinant role in the photosynthetic acclimation of rice plants to heat stress and the short isoform maintaining Rubisco activity
under normal conditions. Interestingly, DeRidder et al. (2012) have recently revealed the existence of a third Arabidopsis RCA splice variant, $RCA_{S2}$, which is barely detectable in unstressed plants but whose expression is induced three fold by heat. This $RCA_{S2}$ transcript encodes an isoform five amino acids shorter than RCA$_S$ and contains a much shorter 3'UTR as well as increased stability when compared to the other two RCA mRNAs (Deridder et al., 2012). Future functional analysis of the $RCA_{S2}$ splice form should elucidate its significance in the thermal regulation of photosynthesis.

A biological role for alternative splicing in the regulation of starch metabolism has also been recently uncovered. Plants store carbohydrates in the form of starch, with different forms of environmental stress, including cold, generally leading to starch breakdown and accumulation of soluble sugars in leaves (reviewed in Krasensky and Jonak, 2012). Seo et al. (2011) identified an Arabidopsis gene encoding the Indeterminate Domain 14 (IDD14) transcription factor, which regulates the expression of a key gene in the induction of starch degradation, Qua-Quine Starch (QQS). The authors also found that cold promotes retention of the first intron in the $IDD14$ pre-mRNA, leading to the production, in addition to the full-length $IDD14\alpha$, of a splice variant, $IDD14\beta$, which encodes a truncated protein lacking a functional DNA-binding domain. Strikingly, independent overexpression of each alternative $IDD14$ transcript leads to opposing phenotypes, with $IDD14\alpha$ overexpression causing retarded growth, while transgenic $IDD14\beta$ plants display accelerated growth and early flowering similarly to the idd14-1 knockout mutant (Seo et al., 2011). These results suggest negative activity for the $\beta$ form, whereas $IDD14\alpha$, shown to bind directly to the QQS promoter, acts as a positive regulator of this gene. Yeast-two-hybrid assays showed that $IDD14\beta$ and $IDD14\alpha$ form heterodimers that are likely to act as competitive inhibitors, preventing the $\alpha$ isoform from binding the QSS gene promoter (Seo et al., 2011). This leads to a scenario in which the alternative $IDD14\beta$ splice form could serve as a cold adaptation strategy, helping the plant maintain during the dark period the amount of starch required to endure cold conditions during the day.

Peroxisomes are eukaryotic multipurpose organelles with important roles in various metabolic pathways, such as fatty acid $\beta$-oxidation, the glyoxylate cycle or photorespiration (reviewed in Hu et al., 2012). As peroxisomes lack DNA, an important component of their function is the ability to import proteins into the peroxisomal matrix.
Such proteins contain peroxisome-targeting signals (PTSs) that are recognized by specific Peroxin (PEX) receptor proteins. Two types of PTS have been described — a highly conserved C-terminal tripeptide sequence, PTS1, and a less conserved cleavable N-terminal nonapeptide, PTS2 (reviewed in Brown and Baker, 2008), which are recognized by PEX5 and PEX7 receptors, respectively. Furthermore, in both plants and mammals, PEX7 interaction with PEX5 is necessary to deliver PTS2 cargo into peroxisomes (Braverman et al., 1998; Woodward and Bartel, 2005). Interestingly, alternative splicing of exon 8 in the mammalian PEX5 pre-mRNA generates two isoforms, a longer PEX5L and the short PEX5S. Although both PEX5 isoforms can bind PTS1, only PEX5L binds PEX7 and functions in PTS2 import (Braverman et al., 1998). While Arabidopsis appears to generate only the longer PEX5L isoform, Lee et al. (2006) have reported that, similarly to mammals, rice generates two PEX5 splice variants, OsPEX5L, encoding the full-length protein, and OsPEX5S, which lacks the 7th exon as a result of an exon skipping event. In the same study, the functional role of OsPEX5L and OsPEX5S was addressed by expressing each isoform independently in a null Atpex5 mutant background. This knockout insertion mutant is unable to germinate in the absence of sucrose due to a severe block in peroxisomal β-oxidation and displays insensitivity to Indole-3-Butyric Acid (IBA), which is converted in peroxisomes into the active auxin Indole-3-Acetic Acid (IAA) (Lee et al., 2006). Remarkably, full rescue of these phenotypes was only observed upon complementation with OsPEX5L, which yeast two-hybrid assays also showed to be the sole isoform capable of interacting with PEX7. These results suggested that the long OsPEX5 isoform is involved in both PTS1 and PTS2 import, whereas the short isoform is only related to PTS1 import. Indeed, transient expression and targeting analysis of GFP-PTS1 and PTS2-GFP reporter constructs in protoplasts isolated from OsPEX5S- or OsPEX5L-expressing Atpex5 mutants demonstrated that while the long isoform is able to rescue targeting of both PTS1- and PTS2-tagged proteins to the peroxisome, the short isoform only complements PTS1 protein transport (Lee et al., 2006). Thus, by determining distinct roles for two isoforms, alternative splicing of the OsPEX5 gene may play a key role in controlling peroxisomal protein import in rice.

Plant development, hormone signaling and abiotic stress response
The transition from the vegetative to the reproductive stage is a crucial developmental switch in the plant’s life cycle controlled by a complex network of signaling pathways and environmental cues. The characterization of late-flowering mutants has led to the identification of the autonomous pathway where flowering is promoted through FCA, an RNA binding protein that blocks the expression of Flowering Locus C (FLC), an important flowering repressor (reviewed in Simpson, 2004). Macknight et al. (1997) showed that alternative splicing and polyadenylation of the FCA pre-mRNA generates four different transcripts, FCA-α, FCA-β, FCA-δ and FCA-γ, which respectively account for <1, 55, 10 and 35% of total FCA expression. Although the ratio of these alternative mRNAs remains constant throughout all developmental stages, in different tissues and under different growing conditions (Macknight et al., 1997), their functional relevance in promoting the transition to the reproductive stage differs considerably (Macknight et al., 2002). The FCA-γ transcript encodes the functional full-length protein, containing two RRM-type RNA binding motifs and a WW protein-protein interaction module, and the main repercussion of FCA alternative pre-mRNA processing appears to be the generation of transcripts lacking these important functional domains. In the case of FCA-δ, the use of both alternative 5' and 3' splice sites in intron 13 results in an early in frame translation termination codon, which yields a protein lacking the WW domain and the 63 downstream amino acids. For the FCA-α transcript, retention of intron 3 also introduces a PTC and additionally includes a second downstream open reading frame (ORF). Finally, while for FCA-γ the polyadenylation site is distal translating into the functional protein, for FCA-β premature cleavage and polyadenylation at a proximal site within intron 3 leads to the production of a truncated transcript encoding a protein that lacks both the RNA binding and WW domains. As expected, overexpression of the FCA-γ transcript is able to rescue the late-flowering defect of fca-1 mutants and leads to an early flowering phenotype in the wild-type background (Macknight et al., 2002). On the other hand, similarly to what was observed for FCA-β (Macknight et al., 1997), FCA-δ overexpression does not affect the mutant phenotype and neither transgene alters flowering time in wild-type plants (Macknight et al., 2002). Despite the lack of functionality of the encoded protein, the FCA-β mRNA is the most abundant, and indeed the production of alternative FCA transcripts may play a key role in limiting the amounts of functional FCA protein to control the flowering process. In agreement with this notion, Quesada et al. (2003) have shown that the active FCA form, FCA-γ, negatively regulates its own expression in a
temporal and spatial dependent manner by favoring polyadenylation at the third intron site (and hence production of the nonfunctional FCA-β transcript), via interaction of its WW domain with FY, a protein similar to the yeast polyadenylation factor Pfs2p (Simpson et al., 2003). Alternative pre-mRNA processing of the FCA gene therefore prevents overproduction of the active isoform, which could lead to deleterious effects on Arabidopsis viability and precocious flowering.

Another Arabidopsis alternatively-spliced RNA binding protein functioning in plant development is related to members of the serine/arginine-rich (SR) protein family, which play key roles in both constitutive and alternative splicing (reviewed in Long and Caceres, 2009; Shepard and Hertel, 2009). The SR45 plant-specific splicing factor contains two arginine/serine-rich (RS) domains flanking a single RNA-Recognition Motif (RRM) and has been well studied in Arabidopsis (Golovkin and Reddy, 1999; Ali et al., 2003; Ali and Reddy, 2006; Ali et al., 2008). A knockout mutant for this gene, sr45-1, exhibits pleiotropic phenotypes during normal plant development, such as late flowering, delayed root growth and altered leaf and flower morphology (Ali et al., 2007), as well as defects in glucose and abscisic acid (ABA) signaling (Carvalho et al., 2010). Alternative splicing generates two SR45 transcripts, with SR45.1 containing a 21-nucleotide sequence that is absent from SR45.2 due to selection of an alternative 3’ splice site in intron 6 (Palusa et al., 2007; Zhang and Mount, 2009). These splice variants encode very similar proteins, differing by only eight amino acids that include putative phosphorylation sites. Remarkably, and despite ectopic expression of the two transgenes, complementation studies clearly showed that SR45.1 rescues the floral but not the root sr45-1 phenotype, while SR45.2 complements exclusively the root growth defect (Zhang and Mount, 2009). By contrast, SR45 alternative splicing does not appear to play a role in sugar signaling, as both splice forms are able to restore the glucose hypersensitivity of the sr45-1 mutant (Carvalho et al., 2010). Elegant mutational analyses conducted by Zhang and Mount (2009) indicate that a potential threonine phosphorylation site within the alternatively-spliced segment is important in distinguishing the flower and root functions of the two SR45 isoforms. The SR45 gene represents a compelling example of alternative splicing dictating a clear dual role in plant development. Another case of a plant gene encoding splice variants that fulfill functions in strikingly different biological processes has been recently uncovered by Remy et al. (in review). Indeed, alternative splicing controls the tissue and subcellular distribution of Zinc-Induced Facilitator 1 (ZIFL1), an Arabidopsis
membrane transporter belonging to the Major Facilitator Superfamily (MFS), with the root tonoplast-localized full-length protein being involved in polar auxin transport, while a truncated isoform targeted to the plasma membrane of stomatal guard cells regulates drought tolerance.

Phytochrome Interacting Factors (PIFs) have been shown to play an extensive role in plant development by triggering responses to light (reviewed in Castillon et al., 2007) and temperature (Koini et al., 2009; Franklin et al., 2011). Germination is a classical example where the response to these environmental signals is integrated, and indeed PIF6 has been shown to control seed dormancy in Arabidopsis (Penfield et al., 2010). This transcription factor is expressed during seed development and its expression is severely reduced upon imbibition. The PIF6 pre-mRNA yields two alternative transcripts resulting from an exon-skipping event, with PIF6-α encoding the full-length protein and PIF6-β, lacking exon 3, giving rise to a truncated protein that lacks a basic Helix-Loop-Helix (bHLH) domain. Although the truncated form can no longer interact with DNA or with other bHLH family transcription factors, both isoforms retain phytochrome binding ability (Penfield et al., 2010). The structural difference between the two PIF6 isoforms leads to different phenotypical outcomes, as shown by analyses of transgenic Arabidopsis plants independently overexpressing each splice variant. While elevated levels of PIF6-α have no effect on germination, lines overexpressing PIF6-β show higher germination levels when compared to the wild type, indicating a role for this isoform in the seed’s germination potential. However, both forms are able to inhibit hypocotyl elongation to a similar extent under red light, showing that phytochrome binding is important for this process, whereas DNA binding activity does not seem to be required (Penfield et al., 2010). These results provide the first evidence of a PIF transcription factor splice form with a role in plant development, extending the accumulated evidence on the importance of PIF factors in the signal transduction of environmental cues and the functional significance of alternative splicing.

The functional characterization of a novel Arabidopsis E3 ubiquitin ligase, XBAT35, has been recently reported by Carvalho et al. (2012). E3 ligases are responsible for recruiting target proteins for ubiquitination and hence confer specificity to the 26S proteasome degradation pathway (reviewed in Smalle and Vierstra, 2004). XBAT35 belongs to a large class of E3 ligases characterized by the presence of a Really Interesting New Gene (RING) domain and contains in addition two N-terminal ankyrin
repeats, a protein-protein interaction motif believed to act as the E3 ligase substrate-binding domain (Stone et al., 2005). Interestingly, skipping of exon 8 excludes an in-frame alternatively-spliced segment in the XBAT35 pre-mRNA harboring an entire nuclear localization signal (NLS), thus determining dual targeting of this E3 ligase to the cytoplasm and nucleus of plant cells (Carvalho et al., 2012). Both the NLS-containing XBAT35.1 and the shorter XBAT35.2 splice variants are ubiquitously expressed in Arabidopsis and have been shown to encode E3 ubiquitin ligase activity in vitro (Stone et al., 2005; Carvalho et al., 2012). Loss of XBAT35 function confers hypersensitivity to the ethylene precursor 1-aminocyclopropane-1-carboxylate (ACC), resulting in a significant increase in apical hook curvature of etiolated seedlings. When ectopically expressed in the xbat35 mutant background, each splice form is able to rescue this mutant phenotype, but the cytoplasmic protein is clearly more effective in conferring an overexpression phenotype, i.e., reduced ethylene-mediated tightening of the apical hook (Carvalho et al., 2012). These observations suggest that the XBAT35 E3 ligase is required in two distinct cell compartments to mediate broad ethylene control of apical hook formation, through the degradation of positive signaling components present in both the nucleus and the cytoplasm. The fact that the XBAT35 splice forms exhibit different biological activity in this specific ethylene-mediated response may be due to the context in which they operate, reflecting either distinct sets and/or levels of target substrates or different in vivo ubiquitination activity of the two isoforms in each cellular compartment.

Jasmonates, including methyl jasmonate (MeJA) and its free-acid jasmonic acid (JA), also constitute important hormone signaling molecules implicated in a wide range of plant developmental processes and stress responses (reviewed in Avanci et al., 2010). Upon perception by receptor proteins, JA signals elicit a transduction pathway involving the degradation of downstream repressors capable of activating or repressing JA-regulated genes (reviewed in Cheong and Choi, 2003; Kazan and Manners, 2008). One family of such repressors is that of Jasmonate ZIM-domain-containing (JAZ) proteins that downregulate JA signaling by binding to MYC2, a bHLH transcription factor and key regulator of JA responses. The presence of JA is detected by the Coronatine-Insensitive 1 (COI1) F-box protein, a component of the SCF E3 ubiquitin ligase that interacts with JAZ proteins and targets them to degradation, thus releasing MYC2 repression and allowing JA-responsive gene activation (reviewed in Robert-Seilaniantz et al., 2011). In addition to a central ZIM domain involved in JAZ-JAZ protein interactions (Chung and...
Howe, 2009), JAZ proteins harbor a C-terminal Jas domain crucial for JAZ-COI1 and JAZ-MYC2 interactions (reviewed in Chung et al., 2009). In most JAZ genes, the Jas-motif coding region consists of two exons separated by an alternatively-spliced intron, which when partially retained generates truncated proteins with incomplete Jas domains, thus implying an important biological function for this so-called Jas intron (Chung et al., 2010). In the JAZ10 gene, proper excision of the Jas intron produces the JAZ10.1 transcript coding for the full-length protein, while usage of an alternative 5′ splice site results in retention of a small portion of this intron and introduces a PTC in the JAZ10.3 splice variant, which encodes a truncated isoform lacking the C-terminal region of the Jas domain. Additionally, selection of an alternative 5′ splice site in exon 3 to produce JAZ10.4 causes a frame-shift mutation that deletes the entire Jas domain and adds 20 C-terminal amino acids (Yan et al., 2007; Chung and Howe, 2009). Importantly, transgenic Arabidopsis lines expressing these individual JAZ10 mRNAs display distinguishable phenotypes. Overexpression of JAZ10.4, but not of JAZ10.1 or JAZ10.3, affects flower fertility due to defects in filament elongation and anther dehiscence, a characteristic feature of strong JA synthesis and perception mutant alleles. On the other hand, upon exogenous application of the strong plant growth repressor MeJA, the JAZ10.1 overexpressor is unaffected in MeJA inhibition of root growth, whereas lines overexpressing JAZ10.3 or JAZ10.4 show partial or severe MeJA insensitivity to this response, respectively (Yan et al., 2007; Chung and Howe, 2009). This is consistent with the prompt degradation of JAZ10.1 by the 26S proteasome, while alterations in the Jas motif impair binding to the COI1 protein. In this way, JAZ10.3 carrying a partially truncated Jas domain can interact weakly with COI1, but the absence of this domain in JAZ10.4 renders it unable to bind COI1 and hence highly resistant to hormone-induced degradation. Nevertheless, as JAZ10.4 does not bind MYC2 directly but through the formation of complexes with other JAZ proteins (Chung and Howe, 2009), the partial or complete absence of the Jas domain in truncated JAZ10 versions is unlikely to impair their interaction with MYC2, explaining the repressed JA-responses in the corresponding JAZ10 transgenic plants. It is interesting to note that overexpression of a splice form of the JAZ2 protein with an incomplete Jas domain, JAZ2.2, but not of the full-length JAZ2.1, also leads to mild MeJA insensitivity (Chung et al., 2010). Therefore, alternative splicing of the C-terminal Jas domain of JAZ proteins plays an important role in attenuating JA responses, thus preventing potential negative consequences on plant growth and development.
A conserved alternative splicing event with implications in plant abiotic stress tolerance was reported by (Fu et al., 2009). Transcription Factor for polymerase III A (TFIIIA) binds both 5S rDNA and 5S rRNA and mediates RNA-polymerase-III-based transcription of 5S rRNA, which is crucial in the assembly of the large ribosomal subunit (reviewed in Szymanski et al., 2003). Comparative sequence analysis between A. thaliana and O. sativa allowed the identification of five conserved cassette exon alternative splicing events between the two species, all associated with genes potentially encoding RNA-binding activity, including TFIIIA (Fu et al., 2009). Whereas in animals there is no evidence of this gene being alternatively spliced, in plants two TFIIIA transcripts are produced, an exon-skipping (ES) and an exon-including (EI) splice variant. While the former encodes the full-length protein, the latter contains a PTC and is targeted to degradation by NMD. Accordingly, Arabidopsis lines overexpressing the EI transcript display no evident phenotype (Fu et al., 2009). On the other hand, transgenic ES overexpressor lines display stress phenotypes, such as low tolerance to high salt and mannitol concentrations, and show an increased accumulation of the endogenous EI isoform, supporting the existence of a post-transcriptional negative feedback regulatory mechanism for TFIIIA. Indeed, the evolutionary conserved cassette exon in TFIIIA contains a 5S-rRNA-like element (Fu et al., 2009) that could provide a means of this transcription factor binding its own pre-mRNA. In this scenario, high TFIIIA abundance would promote inclusion of the third exon in TFIIIA, thereby introducing a PTC and targeting this transcript to degradation. By contrast, low TFIIIA levels would favor skipping of exon 3, leading to the production of the active TFIIIA protein. Such a mechanism would allow maintaining TFIIIA homeostasis in the plant cell, which may be of particular importance under osmotic stress conditions.

**Plant resistance to biotic stress**

Plant disease Resistance (R) genes, important components in plant defense responses, mediate the specific recognition of invading pathogens carrying cognate avirulence (avr) determinants, thereby triggering responses that prevent pathogen colonization and multiplication in the plant (reviewed in Gassmann, 2008). Such responses need to be tightly regulated in order not to cause detrimental effects in the plant. The Arabidopsis genome includes approximately 150 putative R genes, which encode proteins containing a nucleotide binding site (NBS) and a C-terminal extracellular, surface-exposed leucine-
rich repeat (LRR) domain. The majority of these proteins also possess an N-terminal toll/interleukin-1 receptor (TIR) domain, making up the TIR-NBS-LRR (TNL) class of \( R \) genes, many of which have been reported to produce alternative splice products that generally encode truncated TIR-NBS proteins (reviewed in Jordan et al., 2002; Gassmann, 2008). Evidence that these alternative transcripts fulfill crucial in vivo functions has been provided for both the tobacco \( N \) gene and the \textit{Arabidopsis RPS4}.

The tobacco \( N \) gene conferring resistance to the tobacco mosaic virus (TMV) was first characterized by Whitham et al. (1994), who reported the production of two splice variants, one predicted to encode the full-length protein (\( N_S \)) and the other a truncated protein (\( N_L \)). Dinesh-Kumar and Baker (2000) later showed that the \( N_L \) splice variant arises from retention of an alternative exon within intron 3, causing a change in the ORF that includes a PTC and leads to exclusion of most of the LRR domain. Expression of both \( N_S \) and \( N_L \) transcripts, either independently or in combination, in TMV-susceptible tobacco plants indicated not only that both isoforms are required to yield fully resistant plants, but also that the ratio between the two is fundamental in this defense response. Indeed, the relative ratio of the two \( N \) transcripts is regulated by TMV signals in such a way that the \( N_S \) prevalence relative to \( N_L \) is reverted upon 4-8 h of infection, returning to the original splicing pattern after 9 h (Dinesh-Kumar and Baker, 2000).

The \textit{Arabidopsis RPS4} gene, which confers resistance towards strains of \textit{Pseudomonas syringae} pv. \textit{tomato} expressing \textit{avrRps4} (Hinsch and Staskawicz, 1996), generates several splice variants that retain introns 2 and/or 3 and/or splice out or not a cryptic intron within exon 3, thus encoding multiple truncated RPS4 proteins that differ in the number of LRR domains (Gassmann et al., 1999; Zhang and Gassmann, 2003, 2007). Importantly, functional analysis of different combinations of intron-deprived \( RPS4 \) transgenes revealed these introns to be biologically active in disease resistance (Zhang and Gassmann, 2003), providing compelling evidence for the relevance of the encoded truncated proteins. As with the \( N \) gene, the expression ratio of each splice variant was subsequently shown to be crucial in \textit{RPS4}-mediated responses (Zhang and Gassmann, 2007). This suggests that the dynamic regulation of the relative levels of splice isoforms produced by TIR-NBS-LRR \( R \) genes may be an important conserved strategy to modulate plant disease resistance responses.

These studies have established the production of alternative transcripts via alternative splicing as fundamental in achieving full plant disease resistance, but the underlying
mechanisms by which this is accomplished remain unknown. According to the literature reviewed by Gassmann (2008), one hypothesis is that the full-length protein is self-inhibited through interactions between its N and C-terminal domains and that, upon pathogen perception, truncated splice isoforms of the protein would bind its full-length version, thus relieving it from self-inhibition and triggering disease resistance responses. Alternatively, instead of interacting with the full-length R protein, the truncated versions could function as adaptor proteins for downstream signaling events (Gassmann, 2008).

Conclusions and Future Perspectives

With recent global transcriptome analyses indicating that the majority of genes encoded by higher plants undergo alternative splicing, a future major challenge will be to uncover the biological significance of this key posttranscriptional regulatory mechanism in plant systems. In the present review, we have specifically focused on studies in which, through the in vivo functional analysis of different splice variants, alternative splicing of a given gene was clearly shown to confer an advantage to the plant. Nevertheless, several other reports in the literature have provided evidence in support of a physiological meaning for alternative splicing.

A first indication for functional relevance stemmed from early indications that alternative splicing events occurred preferentially in the pre-mRNAs of Arabidopsis genes with regulatory functions, such as those involved in signal transduction or encoding enzymes, receptors and transcription factors (Kazan, 2003). In addition, plant stress-associated genes are particularly prone to alternative splicing (Ner-Gaon et al., 2004), which is also markedly affected by abiotic stress (Iida et al., 2004; Ali and Reddy, 2008; Filichkin et al., 2010), suggesting a role for this mechanism in the response to environmental signals. Consistent with this notion, the expression of SR proteins is stress-regulated at multiple levels, and the functional characterization of these key modulators of alternative splicing has revealed important roles in various aspects of plant development and stress responses (reviewed in Reddy, 2007; Barta et al., 2008; Duque, 2011).

Plant circadian components have also been shown to be regulated by alternative splicing, with a recent analysis of the Arabidopsis transcriptome revealing the existence of clock-regulated intron retention events, thus providing a link between circadian and alternative splicing networks (Hazen et al., 2009). Moreover, the clock-regulated
Arabidopsis Protein Arginine Methyl Transferase 5 (PRMT5), which methylates arginine residues in histones and Sm spliceosomal proteins, has been shown to play a key role in the regulation of pre-mRNA splicing events (Sanchez et al., 2010). This interplay between the circadian clock and alternative splicing control may allow the synchronization of physiological processes with periodic environmental changes (reviewed in Sanchez et al., 2011).

In addition to roles established in planta for the tobacco N and Arabidopsis RPS4 genes, alternative splicing may be of more general importance in plant responses to biotic stress. In fact, although the in vivo functional impact of their multiple splice forms remains elusive, many other alternatively-spliced R genes have been characterized, including RPS6 (Kim et al., 2009), the Arabidopsis RLM3 encoding a TIR-NBS protein required for resistance to the fungal pathogen Leptosphaeria maculans (Staal and Dixelius, 2008; Staal et al., 2008) and the rice Pi-ta gene that confers resistance to races of Magnaporthe oryzae (Costanzo and Jia, 2009). On the other hand, expression of several Arabidopsis splice variants in yeast or bacterial cells, such as of from the MGT7 (Mao et al., 2008) and UPS5 (Schmidt et al., 2006) transporters in Salmonella and yeast, respectively, has also proved their functionality in heterologous systems.

The importance of alternative splicing in plant physiological and developmental processes has therefore recently begun to unfold. Further in planta analyses of the biological significance of individual isoforms generated by alternative splicing as well as the identification of the endogenous transcripts targeted by key regulators of this highly versatile gene regulation mechanism should help elucidate the full extent of its functional relevance in the plant kingdom.

Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia (Grant PTDC/AGR-PRO/119058/2010 as well as Postdoctoral Fellowship SFRH/BPD/80073/2011 awarded to R.F.C.). P.D. is supported by Programa Ciência (QREN/MCTES).

References


Carvalho SD, Saraiva R, Maia TM, Abreu IA, Duque P (2012) XBAT35, a Novel Arabidopsis RING E3 Ligase Exhibiting Dual Targeting of Its Splice Isoforms,
Is Involved in Ethylene-Mediated Regulation of Apical Hook Curvature. Mol Plant. doi: 10.1093/mp/sss048


Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C (2003) FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell 113: 777-787


Figure Legends

Figure 1. Basic types of alternative splicing events and their observed frequencies in the *Arabidopsis* and human transcriptomes.

(A) Schematic diagrams of the main types of alternative splicing. Exons and introns are represented by boxes and horizontal lines, respectively, and diagonal lines indicate alternative splicing events.

(B) Comparison of the relative prevalence of alternative splicing events between *Arabidopsis* and humans (data from Kim et al., 2008; Marquez et al., 2012). More complex alternative splicing events are grouped under “Others”.

Exon S, exon skipping; Alt 5’ ss, alternative 5’ splice site; Alt 3’ ss, alternative 3’ splice site; Intron R, intron retention.
Figure 1. Basic types of alternative splicing events and their observed frequencies in the Arabidopsis and human transcriptomes.

(A) Schematic diagrams of the main types of alternative splicing. Exons and introns are represented by boxes and horizontal lines, respectively, and diagonal lines indicate alternative splicing events.

(B) Comparison of the relative prevalence of alternative splicing events between Arabidopsis and humans (data from Kim et al., 2008; Marquez et al., 2012). More complex alternative splicing events are grouped under “Others”.

Exon S, exon skipping; Alt 5’ ss, alternative 5’ splice site; Alt 3’ ss, alternative 3’ splice site; Intron R, intron retention.
Table 1. Plant genes reported to undergo physiologically relevant alternative splicing events

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Product / Function</th>
<th>AS Function</th>
<th>Species</th>
<th>Type of AS</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCA</td>
<td>Rubisco Activase / Activation of Rubisco</td>
<td>Optimization of photosynthesis — light and thermal regulation of Rubisco activity</td>
<td>Arabidopsis thaliana, Spinacia oleracea, Oryza sativa</td>
<td>Alt 5’ ss (last intron)</td>
<td>Zhang and Portis, 1999, Zhang et al., 2002, Wang et al., 2010</td>
</tr>
<tr>
<td>IDD14</td>
<td>Indeterminate Domain 14 / Transcription factor</td>
<td>Cold regulation of starch metabolism</td>
<td>Arabidopsis thaliana</td>
<td>Intron R (intron 1)</td>
<td>Seo et al., 2011</td>
</tr>
<tr>
<td>PEX5</td>
<td>Peroxin 5 / Peroxin receptor</td>
<td>Control of peroxisomal protein import</td>
<td>Oryza sativa</td>
<td>Exon S (exon 7)</td>
<td>Lee et al., 2006</td>
</tr>
<tr>
<td>FCA</td>
<td>RNA-binding protein / Negative regulator of the FLC flowering repressor</td>
<td>Autoregulation of pre-mRNA processing; Control of flowering</td>
<td>Arabidopsis thaliana</td>
<td>Alt 5’ + 3’ ss (intron 13) Intron R (intron 3) Alt PolyA (intron 3)</td>
<td>Macknight et al., 2002, Quesada et al., 2003</td>
</tr>
<tr>
<td>SR45</td>
<td>SR-related protein / Splicing regulator</td>
<td>Flower and root development</td>
<td>Arabidopsis thaliana</td>
<td>Alt 3’ss (intron 6)</td>
<td>Zhang and Mount, 2009</td>
</tr>
<tr>
<td>PIF6</td>
<td>Phytochrome Interacting Factor 6 / Transcription Factor</td>
<td>Seed germination and hypocotyl elongation</td>
<td>Arabidopsis thaliana</td>
<td>Exon S (exon 3)</td>
<td>Penfield et al., 2010</td>
</tr>
<tr>
<td>XBAT35</td>
<td>E3 Ubiquitin Ligase / Protein recruitment to degradation via the 26S proteasome</td>
<td>Dual targeting to the nucleus and cytoplasm; Regulation of ethylene control of apical hook curvature</td>
<td>Arabidopsis thaliana</td>
<td>Exon S (exon 8)</td>
<td>Carvalho et al., 2012</td>
</tr>
<tr>
<td>JAZ10</td>
<td>Jasmonate Zim domain protein / Jasmonate signaling repressor</td>
<td>Attenuation of jasmonate responses; Control of plant growth and development</td>
<td>Arabidopsis thaliana</td>
<td>Alt 5’ss (Jas intron 4 or intron 3)</td>
<td>Yan et al., 2007, Chung and Howe, 2009, Chung et al., 2010</td>
</tr>
<tr>
<td>TFIIIA</td>
<td>Transcription Factor for Polymerase III A / Transcriptional regulation of the 5S RNA gene</td>
<td>Coupling with NMD; Regulation of osmotic and salt stress tolerance</td>
<td>Arabidopsis thaliana, Oryza sativa</td>
<td>Exon S (Arabidopsis: cryptic exon in intron 2; Rice: cryptic exons in intron 3)</td>
<td>Fu et al., 2009</td>
</tr>
<tr>
<td>N-gene</td>
<td>TNL resistance (R) protein / Defense responses</td>
<td>Regulation of resistance to the Tobacco Mosaic Virus (TMV)</td>
<td>Nicotiana glutinosa</td>
<td>Exon S (cryptic exon in intron 3)</td>
<td>Dinesh-Kumar and Baker, 2000</td>
</tr>
</tbody>
</table>

AS, alternative splicing; NMD, nonsense-mediated decay; SR, serine/arginine-rich; TNL, TIR(toll/interleukin-1 receptor)-NBS(nucleotide binding site)-LRR(leucine-rich repeat).
Alt 3’ ss, alternative 3’ splice site; Alt 5’ ss, alternative 5’ splice site; Alt 3’ + 5’ ss, alternative 5’ and 3’ splice sites; Alt PolyA, alternative polyadenylation; Exon S, exon skipping; Intron R, intron retention.