Alternative splicing in plant stress response

JOANNA GRACZ
Institute of Bioorganic Chemistry, PAS, Poznań, Poland

Abstract
Alternative splicing is a cellular process in which multiple mRNA transcripts are created from a single gene. Comparative studies of splicing on plant and animal systems have revealed differences between those two kingdoms in splice-site recognition, distribution of various splice types, and overall frequency of splicing events. Moreover, global analysis of plant genomes showed an impact of alternative splicing on crop domestication and trait selection. In recent years, new functions of this process in plant development and stress responses were described. From known examples of spliced transcripts emerges a complicated picture of mutual relationship of splicing process with various molecular machineries taking part in regulation of gene expression. For now, we know that splicing is coupled with nonsense-mediated mRNA decay and dependent on chromatin modification. It also influences miRNA pathway. This review summarizes the role of alternative spliced transcripts in expression of plant genetic information especially in response to biotic and abiotic stress conditions.

Key words: alternative splicing, gene expression, stress conditions, plants

Introduction
In 1978, it has been discovered that in messenger RNAs protein-coding sequences (exons) are separated by noncoding protein sequences (introns) and that the latter are removed from primary transcripts thus forming mature ones. Soon after a mechanism in which different exons may be combined with each other, leading to the formation of number of mRNA transcripts from a single gene, was proposed. Alternative splicing (AS) is the phenomenon of selective choosing and combining different exons or introns. It is a common mechanism to increase the diversity of transcripts, which contributes significantly to enhancing the encoding capacity of plant and animal genomes and therefore the complexity of the proteome. Initially, however, the AS process was considered to be an isolated phenomenon occurring only in rare cases (first estimates reported that only 5% of human transcripts is subjected to this process (Sharp 1994)). However, in recent years, the use of high throughput research methods made it possible to re-estimate the scale of this phenomenon, and now it is believed that in about 95% of human (Pan et al. 2008) and in approx. 61% of the Arabidopsis thaliana genes’ (Marquez et al. 2012) AS events occur. In contrast to post-transcriptional control processes, AS modifies the structure of transcripts, which can affect virtually all aspects of the protein (structurally and functionally). Thus, AS appears to be one of the most important processes in the regulation of gene expression in eukaryotic cells. Many AS events occur cotranscriptionally and depend on DNA and chromatin modifications (Braunschweig et al. 2013). The resulting mRNA transcripts may differ in stability or location in cells. They can be translated into proteins that can have shorter or longer amino acid chains, altered activity (increased, decreased, or completely lost), and a different mode of regulation or durability. AS is also involved in regulating the level of transcripts in cells by coupling with the process of degradation of mRNA molecules – NMD (nonsense-mediated mRNA decay) – when a premature termination codon (PTC) appears (Kervestin and Jacobson 2012).

General mechanism of AS
Splicing reaction is catalyzed by a large ribonucleoprotein complex called spliceosome, which includes 5 snRNP (small nuclear ribonucleoprotein) particles –
RNA, from which they consist, is rich in uracil; thus, these molecules are referred as e.g. U1, U2 snRNP. A large part of spliceosome comprises variety of proteins (approximately 300), which belong to different classes, such as RNA-binding proteins (RBP), RNA-dependent ATPase, helicases and RNA kinases (Lorkovic et al. 2000). The spliceosome structure is created de novo for each intron allowing the adoption of a specific conformation of pre-mRNA and spatial approximation of exons located above and below the intron. Introns may be distinguished from exons in primary transcripts by the presence of specific and highly conserved sequences:

- 5’ splicing site with evolutionary conserved GU di-nucleotide, which is recognized by U1 snRNP,
- 3’ splicing site, with conserved AG sequence recognized by U2AF35 (U2 auxiliary factor 35),
- polypyrimidine tract on intron 3’ end,
- the branch point, with CURAY sequence (where R is a purine and Y pyrimidine), which is located 17 to 40 nucleotides upstream of the 3’ splice site; it is recognized by U2 snRNP (Reddy 2007).

Spliceosome assembly is a highly complex and dynamic reaction, which involves the hydrolysis of several ATP molecules and a number of structural rearrangements. After the formation of a spliceosome on an intron sequence, splicing takes place in two transesterification reactions. In the first step, the phosphate from guanosine from 5’ splice site is attacked by a 2’ hydroxyl group from adenosine from branch point sequence. This results in the formation of lariat structure by the intron and as a consequence the release of the exon. In the next step, the cleavage of mRNA by guanine residue from the sequence AG in 3’ splice site is performed, which is then followed by the connection of exons. The intron, released in the lariat form, is then degraded (Turner et al. 2004).

The general organization of genes and the process of pre-mRNA maturation in plants are similar to that occurring in vertebrates, but much less understood and described. The mechanism of splicing and the spliceosome has been described in an animal model, but so far, researchers have failed to isolate spliceosome from plants. However we know that mRNA splicing in plants, as well as in other organisms, consists of two transesterification reactions and a step of the lariat structure formation. Moreover, many analogs of animal spliceosome components have already been identified in plant genomes. Despite a number of similarities, plant cells are not able to perform splicing of heterologous pre-mRNA molecules, suggesting that AS in plants is carried out and controlled in a slightly different manner than in animals (Zhiguo et al. 2013). Significant differences at the stage of intron recognition have also been described. It was found that the sequences of both plant’s splice sites (3’ and 5’) have greater variation, and the location of branch point sequence is less defined compared with yeast or animals (so far, only two plant’s branch sequences were experimentally determined) (Severing et al. 2009). A characteristic feature of plant’s introns is high nucleotide content such as UA and U – introns in both mono- and dicots are approx. 15% richer in the AU sequences than exons. This bias in U content is important in identifying the 5’ and 3’ splice sites and the removal of introns (Ko et al. 1998). The choice of specific splice sites depends on the presence of additional cis-elements in the transcript sequence, which is referred to as splicing regulatory elements (SRES). They can be located both within introns and exons and perform the functions of splicing enhancers and silencers; hence, they are further divided into exonic splicing enhancer (ESE)/exonic splicing silencer (ESS) and intronic splicing enhancer (ISE)/intronic splicing silencer (ISS). Several proteins that specifically bind to these elements have been identified: for example, serine/arginine-rich proteins (SR proteins). These proteins are important factors involved in splicing and the spliceosome arrangement. They are mediating protein-protein and protein-mRNA interactions. SR proteins are linked through their RNA-binding domain with the corresponding motif sequence within the maturing transcript. Depending on the recognition sequence, joining the SR protein promotes or inhibits the inclusion of the sequence (intron or exon) to the maturing mRNA molecule (Reddy 2004).

AS events can be grouped into several major types: 1) exon skipping, in which the entire exon (called cassette exon) together with the adjacent introns is removed from the maturing mRNA; 2) selection of different 5’ splice sites (acceptor site) or 3’ splice site (donor site) – these two types of AS are associated with the recognition of two or more splice sites at the end of an exon, resulting in a displacement of the exon/intron boundary; and 3) intron retention, which results in leaving some or all of the intron sequences within the mature transcript (see Figure 1).
The analysis of the amount and types of splicing events showed differences in their distribution among species. The most frequent splicing event in vertebrates is exon skipping, which occurs in 40% of AS cases (Keren et al. 2010). The distribution of AS types on the eukaryotic evolutionary tree indicates an increasing contribution of this type of splicing and it is suggested that exon skipping affects the phenotypic variation to the greatest extent. In plants, however, whole genome duplication is the phenomenon of greater importance; therefore, exon-skipping events are less frequent (Kim et al. 2008). In *A. thaliana* and several other plant species, the dominant type of AS is intron retention. Initially, it was assumed that mature transcripts with intron sequences are mostly subjected to NMD. Nevertheless, recent studies have shown that such transcripts are not always degraded in the NMD process, which once again raised the question about their purpose and function (Kalyna et al. 2012). A clue to answer this question can be observed during the maturation of gametophyte of bryophyte *Marsilea vestia*. During the spermatogenesis, transcripts with retained intron are collected in spores to a moment when their final splicing and translation are induced (Boothby et al. 2013). Another hint allowing in determining the role of transcripts with retained introns came from the research on mammalian neurogenesis. It has been shown that retaining an intron sequence in transcripts specific to neurons inhibits their export into the cytoplasm and causes degradation within the nucleus independently of NMD (Yap et al. 2012). It appears possible that the AS process of transcripts withholding an intron can be a common mechanism for handling precursor RNA in the nucleus and adjusting the final stages of mRNA maturation in a manner dependent of the developmental stage and external factors.

**AS’ role in gene expression**

The increase in the proteome diversity with the limited coding capacity of the genome is a basic function of AS events. However, the extent of these events is still difficult to estimate. Some genes are alternatively spli-
ced constitutively, while others are tissue specific or alternatively spliced depending on the developmental phase or environmental conditions. The biosynthesis of protein isoforms also depends on the type of AS events. Some of them do not lead to a frame shift – analysis of enriched gene ontology showed that AS events with frame shifts are particularly common in transcripts encoding for RNA binding proteins (RBP), which are transcription factors and are involved in regulation of splicing (Severing et al. 2012). As a result of open reading frame (ORF) alteration by AS, proteins with shorter amino acid chains are often synthesized. In some cases, they may maintain their functional domains, while in others the sequences encoding these regions may be removed from the mature transcript, or sequences that carry information about new domains may be added. It has been shown that some alternatively included exons encode domains of undefined structures that, however, appear on the protein surface and enter into protein-protein interactions. In this way, the AS may modify the networks of protein interactions as well as signal transduction pathways (Ellis et al. 2012). In some cases, these truncated peptides retain the ability to dimerize and to form homo- or heteromers, but do not have any functional domains. It is suggested that such truncated proteins may compete during the dimerization process with full-length peptides, thereby acting as its negative regulators (Seo et al. 2013). Furthermore, these truncated peptides may alter post-translational modifications: for example, phosphorylation, which also can influence kinase-dependent signal flow in the cells (Merkin et al. 2012). It has been shown that in A. thaliana RS41, CypRS64, and SR45 proteins have specific phosphorylation sites for single splice variants (Bentem et al. 2006).

AS is involved in the regulation of the amount of transcripts in cells by linking with the decay of mRNA molecules – NMD. NMD is a cytoplasmic RNA degradation system, which is activated during the first round of translation, when the termination of translation is disturbed. Transcripts targeted by NMD have long 3′ UTR sequences with introns, premature transcription termination signal (PTC) situated about 50-55 nucleotides upstream of the splice site and above the upstream open reading frame (uORF). All these features can be introduced or modified by the AS events so that the splicing process can directly influence transcript stability. The analysis of transcriptome of A. thaliana using a collection of expressed sequence tags (ESTs) and RT-PCR estimated that from 11% to 18% of transcripts containing introns are targets for NMD (Kalyna et al. 2012; Drechsel et al. 2013). AS process is often coupled to the NMD and affects gene expression on a negative autoregulatory feedback loop (increased level of protein in cell leads to an increase in the amount of transcripts generated by AS with PTC that are then targeted by NMD). Such dependence between AS and NMD was observed, for instance, for splicing factors (Palusa and Reddy 2010), genes encoding for GRP7 and GRP8 that are part of the circadian clock (Steiger et al. 2003), or for the AFC2 gene encoding for a highly conserved LAMMER kinase (Marquez et al. 2012). The feedback loop process between splicing and NMD affects many genes, as demonstrated by the above-mentioned examples.

Another case illustrating AS’ influence on gene expression is its crosstalk with miRNA pathway. As the result of AS process, mRNA isoforms or non-coding RNAs can occur that may or may not contain the miRNA binding site. AS may also modify the process of excision of miRNAs’ from the pri-miRNA (miRNA precursor), thus influencing the amount of mature miRNAs in the cell, which in turn corresponds to changes in the amount of targeted mRNAs. Changes in the miRNAs’ path induced by splicing process have also an impact on the splicing pattern of pre-miRNA and mRNA of genes encoding enzymes involved in the biogenesis of miRNAs: for example, gene encoding a key endonuclease in miRNA pathway – DCL1 (Dicer-like 1). As a result of AS events associated with intron 12 removal, 4 transcript isoforms arise, but only one of them has the binding site for miRNA162 (Yang G.D. et al. 2012). In plants, miRNAs binding sites (MBS) may be located both in the coding sequences and in the UTR regions. Using bioinformatics analysis, it has been established that in A. thaliana, AS process affects the 12% MBS (44 out of 354 mRNA with specific MBS) (Yang X. et al. 2012). One of the most interesting examples of the relationship between the AS and miRNAs is the expression of miRNA400 in A. thaliana under heat stress conditions. This miRNA coding sequence is located within the intron of the transcript encoding an unknown protein (At1g32853), and therefore it is coexpressed with this transcript. In plants that overexpress miRNA400 under heat stress conditions, the germination and growth rate decrease. However, in the wild type under high temperature, an excision of the
Alternative splicing in plant stress response

Adverse geo-environmental conditions, which are known as abiotic stresses, require the ability to adapt cell metabolism so that plants can survive and thrive. Under abiotic stress condition, plants must precisely and tightly regulate gene expression at multiple levels. Moreover, to fight and defend against a wide range of pathogens (biotic stressors), plants devote substantial resources by reprogramming expression of many resistant and pathogen-related genes. Comparative analysis of several genomes showed greater pace of evolutionary change in genes undergoing AS, which in plants may be related to the acquisition of better adaptation properties and fitness to adverse environmental conditions, as well as to biotic stresses. Although the majority of AS events are poorly characterized, many genes with regulatory functions and related to stress response are overrepresented in databases of AS in plants. Furthermore, functional analysis of genes in A. thaliana, which produce transcripts with retained intron, showed that most of them are related to stress responses. An analysis of transcriptomes of A. thaliana under various stress conditions showed a significant increase in the amount of AS events compared to the control conditions (Kazan 2003). Transcriptome studies conducted on two maize lines differentially susceptible to herbicide Roundup also showed an increase in AS events after herbicide spraying, which in fact is another abiotic stress for growing plants. Genes with altered splicing pattern were mainly involved in stress response, photosynthesis, and secondary metabolism (Gracz et al; data unpublished). In addition to genome-wide analysis, there are numerous and rapidly growing reports of single genes, in which pre-mRNA splicing is altered under stress conditions; some examples are summarized in Table 1 and are described as follows.

The salt overly sensitive (SOS4) gene from A. thaliana is one of stress-induced genes. After AS event in intron 1, under high salt conditions, two transcripts are created, which then can be translated into a truncated protein with 34 amino acid differences in the N-terminus. Despite the differences, both isoforms were effective in rescuing salt-sensitive phenotype in sos4 mutant plants and both are overexpressed in salt stress (Shi et al. 2002). OsIIm – gene encoding oxidase that functions in the removal of stress-induced reactive oxygen species – is also overexpressed and alternatively spliced under salt stress conditions in rice. Two isoforms

intron encoding miRNA400 occurs, resulting in increased amount of the pri-miR400 transcripts and the decline of its mature, functional form. This dependence increases the chances of plant’s survival in stress conditions (Yan et al. 2012).

In recent years, the knowledge about AS general role in the gene expression process and the number of genes that undergo this process is constantly increasing. However, for the alternatively spliced transcripts, still there are only few proposed biological functions. Recently, the role of AS of a tobacco gene N in the acquisition of plant resistance to tobacco mosaic virus (TMV) infection was described. Plants synthesize two transcripts of a gene N – Ns and NL. NL is formed by including an alternative exon located within intron 3 and is overexpressed during the infection. It was also shown that the presence of both transcripts of the N gene is necessary to provide a resistance to the virus (Latijnhouwers et al. 1999). Another example of an important process depending on AS is the control of flowering in A. thaliana associated with the expression FCA gene (Flowering Time Control Protein). In transgenic plants with an overexpression of this gene, an acceleration in flowering time has been observed, which is associated with a small increase in the pool of FCA transcript that has all the introns removed (in the remaining 3 transcripts, a fragment of intron 3 is retained). It is suggested that the negative autoregulation of FCA gene transcripts maturation prevents the plant from premature flowering (Osesada et al. 2003). Other examples were provided by studies on maize genome, which revealed that part of the genes involved in processes such as germination or synthesis of anthocyanins are subjected to AS. During gene expression of Knox7 gene, which encodes a transcription factor regulating the germination, 7 alternative transcripts are formed (Morele-Le Paven et al. 2007). Other well-described example of AS events in maize is the expression of Bronze2 gene. Under stress factors such as the presence of cadmium, up to 50-fold increase in the level of an alternative transcript (from which the introns have been removed) is observed. This gene is responsible for the final step of the synthesis of anthocyanins (Marrs 1997). Also, genes encoding for SR proteins are subjected to AS in maize – 4 transcripts have been identified for zmRSp31A gene and 3 transcripts from zmRSp31B gene, which arise from the use of non-canonical splice sites (Gupta et al. 2005).
Table 1. Examples of alternatively spliced genes associated with plant stress response

<table>
<thead>
<tr>
<th>Species</th>
<th>Gene</th>
<th>Protein product</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biotic stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>RPS4</td>
<td>Resistance protein</td>
<td>Yang et al., 2014</td>
</tr>
<tr>
<td></td>
<td>RPP5</td>
<td></td>
<td>Parker et al., 1997</td>
</tr>
<tr>
<td></td>
<td>EDS1</td>
<td></td>
<td>Iida et al., 2004</td>
</tr>
<tr>
<td>Flax</td>
<td>L6</td>
<td></td>
<td>Ayliffe et al., 1999</td>
</tr>
<tr>
<td>Rice</td>
<td>OsMAPK5</td>
<td>MAP kinase</td>
<td>Xiong and Yang, 2003</td>
</tr>
<tr>
<td>Cold stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>SR34/SR1</td>
<td>SR protein</td>
<td>Lazar and Goodman, 2000</td>
</tr>
<tr>
<td>Grapevine</td>
<td>DHN1</td>
<td>Dehydrin</td>
<td>Xiao and Nassuth, 2006</td>
</tr>
<tr>
<td>Tomato</td>
<td>LeOX</td>
<td>Alternative oxidase</td>
<td>Fung et al., 2006</td>
</tr>
<tr>
<td>Heat stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>SR30, SR1, SR34b, RS40, RS31, SR33</td>
<td>SR protein</td>
<td>Palusa et al., 2007</td>
</tr>
<tr>
<td>Maize</td>
<td>HSP22</td>
<td>Heat shock protein</td>
<td>Lund et al., 2001</td>
</tr>
<tr>
<td>Salt stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>SOS4</td>
<td>Pyroxidal kinase</td>
<td>Shi et al., 2002</td>
</tr>
<tr>
<td>Rice</td>
<td>OsIM</td>
<td>Oxidase</td>
<td>Kong et al., 2003</td>
</tr>
<tr>
<td>Drought</td>
<td>DHN1</td>
<td>Dehydrin</td>
<td>Xiao and Nassuth, 2006</td>
</tr>
<tr>
<td>Heavy metal stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maize</td>
<td>Bronze2</td>
<td>Glutathione-S-transferase</td>
<td>Marrs, 1997</td>
</tr>
</tbody>
</table>

(OsIM1 and OsIM2) from this gene are produced, which differ in abundance between salt-tolerant and salt-sensitive rice varieties. After the exposure to salt treatment, high level of OsIM1 in salt-tolerant variety is maintained, whereas in salt-sensitive plant, the level of this transcript decreases rapidly (Kong et al. 2003).

Dehydrins are a group of hydrophilic proteins involved in response to diverse range of stress conditions – they are induced by cold, drought, and salt. Experiments with overexpression of dehydrins in Arabidopsis improved stress tolerance in these plants (Puhakainen et al. 2004). In Vitis, two genes encoding dehydrins were identified (DHN1a and DHN1b); they contain one intron producing two splice forms – fully spliced encoding for a full protein and its unspliced version (Xiao and Nassuth 2006). The unspliced form encodes for two short ORFs, each giving a truncated protein. After the exposure to cold in Vitis, raparia, the unspliced isoform production is induced. Drought, on the contrary, induces the transcription of both spliced and unspliced forms of DHN1, but with different kinetics. The amount of spliced transcript progressively increases from day 1 and unspliced form from day 4 of applied water deficit stress conditions in two varieties of Vitis. Unspliced form produces proteins with calcium binding domain – its elevated level in late stress response may be a mechanism of cytosolic calcium downregulation in stress conditions. Calcium is one of key components of signal transmission in plant cells, and under cold or drought treatment, its concentration is significantly increased (Puhakainen et al. 2004).

As mentioned before, plants produce reactive oxygen species (ROS) in response to most stress conditions. Because of their harmful effects, plants evolved numerous mechanisms to safely detoxify them, which includes synthesis of antioxidants such as glutathione,
ascorbate, tocopherol, and related enzymes such as superoxide dismutase, catalase, peroxidase, and other ROS scavengers (Tripathy and Oelmuller 2012). In tomato, one of alternative oxidase genes (LeAOX) produces three isoforms as a result of different splicing events of last intron, which in consequence is leading to the translation of three proteins with different activities (Fung et al. 2006). Expression of individual isoforms is tissue specific and changes upon cold treatment. During low-temperature conditions, the last intron is retained in maturing transcript, and as a result, a change in the ratio of different isoforms occurs. Moreover, those changes are correlated with splicing pattern alteration of the 9G8 SR splicing factor transcript. The level of a longer isoform containing additional introns was reduced under cold stress conditions, and it presumably encodes for a nonfunctional protein. The role of SR proteins in affecting splicing of other genes is well known, modification of 9G8 AS may be responsible for changing splicing patterns of LeAOX gene in tomato (Fung et al. 2006).

Different splicing pattern of transcripts from several SR genes under stress conditions was observed in Arabidopsis as well. Cold treatment alters splicing of SR34/SR1, SR34b, RS 40, RS31, and SR33, whereas heat changes the expression patterns of SR30, SR1, SR34b, RS31a, RS40, RSZ33, SR33, and SCL30a (Lazar and Goodman 2000; Palusa et al. 2007). SR genes with different AS under both stress conditions are sometimes affected in opposite manner. For example, isoform 2 of SR33 was reduced by cold and increased by heat conditions (Palusa et al. 2007). These results suggest that temperature stress simultaneously affects the expression of several specific splicing factor isoforms, which in turn can alter AS of downstream temperature-regulated genes.

Resistance genes (R genes) that provide plant disease resistance are also subjected to AS process. Protein products of R genes are involved in recognition of invading pathogens and triggering response, which inhibits pathogen colonization and multiplication. One of best described examples of fine-tuned AS role in plant immunity is expression regulation of RPS4 gene in A. thaliana. This gene confers resistance to Pseudomonas syringae pv. tomato strain DC3000 (DC3000). After intron 2 and/or 3 retention and usage of cryptic splice site in exon 3, six different isoforms of RPS4 are expressed. A precise adjustment of isoform ratios is key to maintaining RPS4 gene function. There is evidence that expressed isoforms encode truncated proteins, which regulate the activity of full-length RPS4 (Yang et al. 2014). AS events in N gene in tobacco mentioned before is yet another example of the role splicing plays in plant’s resistance to pathogens. Apart from R gene family, there are other genes such as EDS1 (Iida et al. 2004) and OsMAPK5 (Xiong and Yang 2003) that undergo AS in response to biotic stresses. The function of many isoforms arising during plant disease remains unknown but should definitely be one of priority areas in determining the role and importance of splicing process in plants, especially those that are agronomically important.

Conclusions

Recent years have brought many new methods to study the coding potential of living organisms. Thanks to high throughput, sensitive sequencing methods the scientific community manages to re-establish the level of AS events in plants and appreciate magnificent richness of different transcript isoforms expressed in various tissues or under diverse range of stress conditions. The function of specific transcripts in most cases is still a mystery awaiting for discovery, but instantly growing examples of AS studies shed light on the complicated role of splicing mechanism in gene expression and its relation with other transcriptional and posttranscriptional processes regulating plant stress responses. Functional analyzes indicate the involvement of splicing process in photosynthesis, defense response, flowering, or cereal grain quality (Barbazuk et al. 2008). Extending those analyzes on economically significant plants like Zea mays or other crops should help in faster and more efficient trait selection and future breeding programs.

Acknowledgment

This work was supported by grants from National Science Center Poland no. UMO-2011/01/N/NZ9/02900 and Polish Ministry of Science and Higher Education, under the KNOW program.

References


