

Review

# First Come, First Served: Sui Generis Features of the First Intron

David Zalabák and Yoshihisa Ikeda \* 

Department of Molecular Biology, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, 78371 Olomouc, Czech Republic; david.zalabak@upol.cz

\* Correspondence: yoshihisa.ikeda@upol.cz

Received: 31 May 2020; Accepted: 17 July 2020; Published: 19 July 2020



**Abstract:** Most of the transcribed genes in eukaryotic cells are interrupted by intervening sequences called introns that are co-transcriptionally removed from nascent messenger RNA through the process of splicing. In *Arabidopsis*, 79% of genes contain introns and more than 60% of intron-containing genes undergo alternative splicing (AS), which ostensibly is considered to increase protein diversity as one of the intrinsic mechanisms for fitness to the varying environment or the internal developmental program. In addition, recent findings have prevailed in terms of overlooked intron functions. Here, we review recent progress in the underlying mechanisms of intron function, in particular by focusing on unique features of the first intron that is located in close proximity to the transcription start site. The distinct deposition of epigenetic marks and nucleosome density on the first intronic DNA sequence, the impact of the first intron on determining the transcription start site and elongation of its own expression (called intron-mediated enhancement, IME), translation control in 5'-UTR, and the new mechanism of the *trans*-acting function of the first intron in regulating gene expression at the post-transcriptional level are summarized.

**Keywords:** first intron; alternative splicing; intron-mediated enhancement; upstream ORF

## 1. Introduction

More than four decades ago, we underwent a paradigm shift in gene organization in higher organisms and caught a glimpse of the molecular processes in manufacturing messenger RNA from heterogeneous nuclear RNA (hnRNA). In 1977, electron microscopy images visualizing DNA:RNA hybrids annealed between adenovirus 2 hexon mRNA and restriction endonuclease cleavage fragments of viral DNA were the smoking gun proof demonstrating that cytosolic mRNA is matured through the removal of intervening sequences from nascent mRNA, called splicing [1,2]. In the following year, 1978, the terms “introns” and “exons” were proposed for the first time to describe a mosaic feature of eukaryotic genes, comprising coding sequences (exons) interrupted by intragenic regions (introns), by Gilbert, who wondered “Why genes in pieces?” [3]. He also coined the statement “Evolution can seek new solutions without destroying the old” [3]. Instead of duplicating the genetic materials, higher organisms take advantage of alternative splicing (AS) to increase protein diversification. Despite the fact that more than 40 years have passed since its discovery, very simple questions such as the origin and evolution of spliceosomal introns remain unsolved [4] and have been contentious in the intensive, long-standing “introns-early/introns-late” debate [5,6]. Either way, although intron densities among eukaryotes are variable, introns and spliceosome machinery are ubiquitous in eukaryotes [7,8]. Regulatory molecular mechanisms of pre-mRNA processing [9], benefits of splicing that contribute to the diversification of proteins by AS [10], epigenetic regulation of AS [11], the roles of introns in modulating expression in diverse regulatory layers [9], how introns influence and enhance eukaryotic gene expression. [9,12], and AS in plant development and stress responses [13–16] have

been intensively discussed and summarized elsewhere. Along with providing concise and general background knowledge of AS, we will focus particularly on unique signatures of the first intron sequence on DNA/nucleosome and on pre-mRNA in its own gene regulation. Additionally, we will highlight the novel functions of yeast first introns modulating the translation of other genes for fitness to nutrient deprivation.

## 2. First Intronic DNA Sequence

In humans, the first intronic DNA sequence is the longest and most conserved, and exhibits the highest density of regulatory chromatin marks when compared to the remaining intronic sequences [17,18]. DNA and histone methylation levels are also related to nucleosomes [14–16]. In general, the DNA methylation level correlates with that of H3K4me1 but is exclusive with H3K4me2 and H3K4me3 in transcribed genes [19]. Similar to animal chromosome organization [20], genome-wide nucleosome positioning analysis in *Arabidopsis* revealed that both 5'- and 3'-splice sites (5'-SSs and 3'-SSs) are preferentially enriched in nucleosomes, nucleosomal DNA is more highly methylated, and nucleosomes are enriched in exons [21]. Interestingly, nucleosome deposition at the first intron 3'-SS is more clearly defined than the other intron 3'-SSs [22]. Repressive epigenetic marks such as DNA methylation and histone di-methylations are deposited on some introns, especially on long introns (>2 kb) [23]. In human cells, histone modifications such as H3K4me3 and H3K9ac are enriched in the 5'-SS of the first intron [24]. Recent findings suggest an effect of DNA methylation on splicing [25,26] In humans, DNA methylation is known to affect 22% of AS, presumably through (1) modulation of the elongation rate of RNA polymerase II, or (2) the formation of protein bridges of HP1 that recruit splicing factors onto exons [26,27] IR in plants is highly enriched in DNase I hypersensitive sites (DHSs) [28]. In rice, approximately 7% of AS events in *met1-2* null allele, whose methyl-CG level is globally lost, are affected, suggesting the involvement of DNA methylation in AS, but the contribution is relatively minor [29].

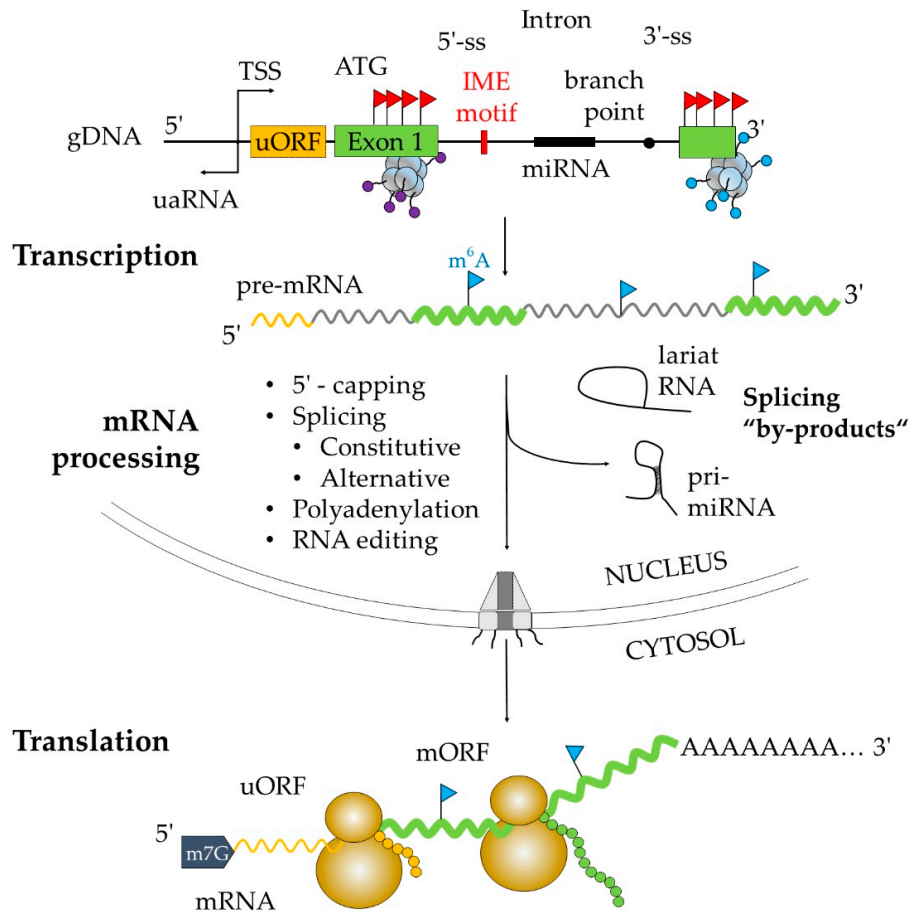
When intronic regions are larger, regulatory elements are likely to be embedded within them to regulate their gene expression [30,31] A well-documented example of this is intronic transposable elements (TEs) [32]. Around 3% of TEs are located within gene bodies, with the majority at intronic regions [33] often triggering the formation of transcriptionally repressive heterochromatin [32]. The second example of an intron-encoded element is microRNA (miRNA). Similar to canonical intron-containing gene expression, intron-embedded miRNA is co-transcribed with its host gene, and its subsequent maturation is controlled by AS. This regulation is unique and intriguing, in that environmental cues such as heat stress affect AS events to control the level of mature miRNA, which, in turn, regulates the miRNA-target gene expression *in trans* [34]. On the contrary, some intronic regions contain a sequence complementary to miRNA [35] or are targeted by noncoding RNA (ncRNA). HIDDEN TREASURE 1 (*HID1*) encodes ncRNA that forms a protein complex in the nucleus to bind to the *PIF3* first intronic sequence to repress *PIF3* expression at the transcriptional level [36]. This was probably the first case in plants demonstrating the regulation of ncRNA binding to chromatin for regulation of gene expression.

The FLOWERING LOCUS C (*FLC*) first intron contains two RY (R, purine; Y, pyrimidine) cis-elements (TGCATG) that are recognized by the B3 DNA binding domain of VAL1 (VIVIPAROUS1/ABI3-LIKE factor 1), a transcriptional repressor capable of repressing *FLC* expression by modifying the H3K27me3 level in the *FLC* intronic region and decreasing H3 acetylation in the first exon by recruiting plant homeodomain-polycomb repressive complex 2 (PHD-PRC2) [37].

## 3. Splicing is Predominantly Co-Transcriptional

As described above, exonic and intronic DNA exhibit distinct nucleosome occupancy/chromatin structure, as well as DNA and histone depositions, all of which could define RNA polymerase II (Pol II) elongation speed. Due to high GC content and concentrated nucleosomes in the exonic region, elongation speed is slower than that in the intronic region (Figure 1, top) [38,39] As a result, AS is

regulated by the speed of transcriptional machinery [11,40–43]. A myriad of articles demonstrate that splicing in many cases takes place co-transcriptionally [39,44–46] and the speed of transcription elongation driven by Pol II affects splicing [40,41,47]. In the case of plants, the rate of splicing appears to be much slower than that of yeast. Jia et al. showed that by developing a nanopore-based method, more than half of the introns remain unspliced after Pol-II transcribed 1 kb past the 3'-splice site, and the authors called such introns post-transcriptionally spliced introns (pts introns) [48].

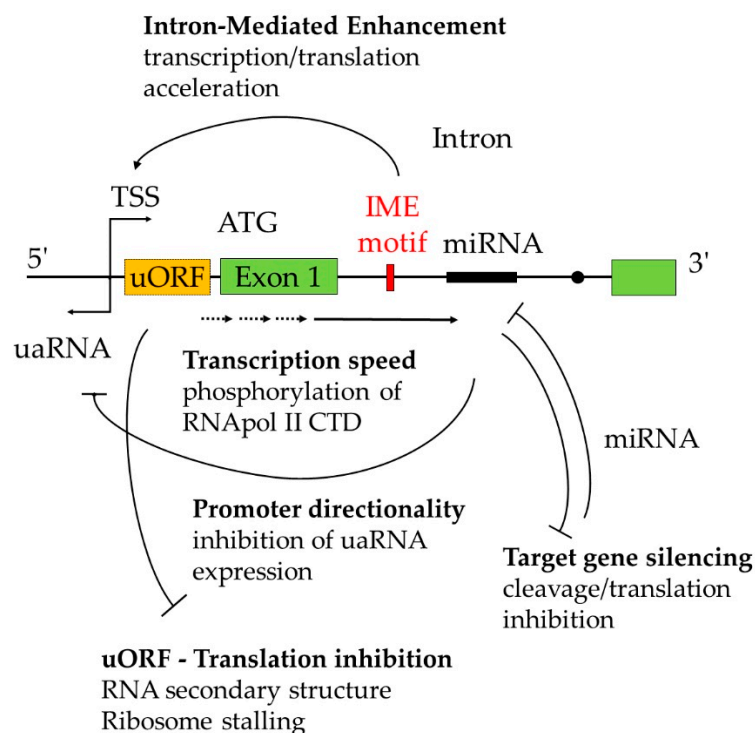


**Figure 1.** Organization of eukaryotic gene and key steps in gene expression. Structure of eukaryotic gene encoded in nuclear genomic DNA (gDNA) is shown on top. Orange box represents an upstream open reading frame (uORF); green boxes are exons of main open reading frame (mORF) carrying initiation codon (ATG). The first intron is flanked by a 5'-splice site (5'-SS) and a 3'-splice site (3'-SS), with branch point (black circle) located close to 3'-SS. The first intron might carry an intron-derived motif (IDM), gene encoding micro RNA (miRNA) or miRNA-binding site. Transcription is initiated from transcription start site (TSS) in sense orientation, producing premature messenger RNA (pre-mRNA) and, to a lesser extent in antisense orientation, upstream antisense RNA (uaRNA). Gene transcription is further regulated by specific epigenetic marks, e.g., CpG methylation (red flags) enriched in exonic regions of gDNA and adenine methylation (blue flags) within RNA molecule. Another layer of epigenetic control is mediated by histone (grey octamer structure) methylation as H3K4me2 together with H3K4me3 pattern (purple dots) are enriched in the 5'-end and H3K4me1 (blue dots) in the 3'-end of gene body. pre-mRNA is further processed to mature messenger RNA (mRNA) and exported to cytosol through nuclear pores. The processing involves 5'-capping (m<sup>7</sup>G), splicing, polyadenylation, and occasionally RNA editing. During the splicing process, lariat RNA is produced. Primary miRNA (pri-miRNA) is released from spliced intron-containing miRNAs and requires additional processing to become functional. Translation of mRNA into protein occurs on ribosomes composed of small and large subunits (yellow structure).

Throughout the whole processes of transcription, Pol II interacts with various proteins (such as transcription factors, chromatin modifiers, and RNA processing enzymes) via its carboxy-terminal domain (Pol II-CTD) in post-translational modification, especially in a phosphorylation-dependent manner [49]. It is worth noting that Pol II interacts with active spliceosome complex (both snRNA and proteins) in a Ser-5 phosphorylation-dependent manner [46]. The distribution of Ser-5P Pol II has a particular peak at the 5'-SS of exon–intron boundaries, and Ser-5P Pol II accumulation persists over exon sequences; however, such accumulation is diminished at the 5'-SS of exon excluded by exon skipping. This finding suggests that Ser-5P Pol II pauses at 5'-SSs to allow time for the U2 snRNP splicing complex to cleave 5'-SS [33].

Nuclear cap-binding complex (CBC) consists of two subunits, CBP20 and CBP80, which bind to the cap structure of transcripts and affect splicing [50]. From the 5'-end of pre-mRNA, the CBC interacts with U1 snRNP at the 5'-SS of the first intron [50]. In *Arabidopsis*, cap-binding proteins, CBP20 and CBP80, and its interacting partner SERRATE (*SE*) modulate the first intron AS by affecting the 5'-SS of the first intron [51–53]

Non-protein-coding RNAs are shown to modulate AS, especially IR in plants. AS COMPETITOR LONG NONCODING RNA (*ASCO-lncRNA*) hijacks AS regulators by interacting with NUCLEAR SPECKLE RNA-BINDING PROTEIN (*NSR*), resulting in IR of *NSR*-target transcripts [54]. In many cases, micro RNA (*miRNA*) molecules are complementary to mRNA sequences; however, some *miRNAs* found in *Arabidopsis* and rice genomes target intronic sequences [35], raising the possibility that *miRNA* regulates pre-mRNA (Figure 2). Meng et al. [35] proposed a novel regulatory cascade called “*miRNA*-intron-phased secondary *sRNAs* targets” by which such overlooked annotated intron-targeting *miRNAs* are processed by ARGONATURE 1 silencing complex to form mature *miRNAs* that bind to and cleave target pre-mRNAs in the nucleus. As a consequence, cleaved remnants are subjected to the RNA-dependent RNA polymerase template to give rise to the production of *sRNA* by DICER to degrade target mRNA in the cytoplasm or modify target intronic DNA sequences (such as methylation) as AGO4-associated *sRNA* in the nucleus (experimental validation is required) [35]. *miRNAs* complementary to the first intronic DNA are listed in Table 1.



**Figure 2.** The first intron plays a key role in several mechanisms regulating gene expression.

The presence of intron-mediated enhancement (IME) motif within the first intron determines the TSS and increases gene expression. Phosphorylation status of polymerase II (Pol II) carboxy-terminal domain (CTD) varies during transcription. Ser-5P Pol II poses at the exon–intron junction. The first intron inhibits transcription of uaRNA by recruiting termination factors that confer promoter directionality. The presence of uORF within the mRNA transcript might negatively affect the translation efficiency of main ORF by forming the secondary structure or by ribosome stalling. Another layer of gene expression control is mediated through miRNA/TE encoded within intronic DNA. Alternatively, the intronic sequence is complementary to miRNA.

#### 4. Intron-Mediated Enhancement (IME), a Mysterious Phenomenon

It is widely recognized that the presence of intron stimulates mRNA accumulation [12]. Back in 1979, two years after the discovery of intervening DNA in the adenovirus 2 genome [1,2] it was reported that intron was required for the accumulation of stable mRNA [12]. In 1987, the same outcome was confirmed by studying ALCOHOL DEHYDROGENASE 1 (*ADH1*) in maize [55], and the authors revealed that only the first intron is sufficient to retain expression similar to that of the genomic *ADH1* construct [55,56]. This result suggests a unique ability of the first intron to enhance its expression, yet the precise molecular mechanism underlying intron-mediated enhancement (IME) remains vague (Figure 2).

##### 4.1. Splicing-Dependent IME

The majority of mammalian promoters initiate transcription on both sides in opposite directions, called divergent transcription [57–59] whereas plant promoters appear to act unidirectionally [44,60]. One way to increase mRNA accumulation is to orient transcription initiation toward one direction. The splicing signal has been shown to play a predominant role in enhancing gene transcription [61]. This was corroborated in yeast cells, showing that splicing-competent introns affect promoter directionality, favored toward transcribing downstream coding region producing mRNA rather than the upstream antisense RNA [62]. The mechanism behind to account for the splicing-competent IME appears to be gene looping. Gene loops juxtapose promoters and terminators in yeast [63,64]. O’Sullivan et al. showed that a looped configuration of genes (where promoter and terminator regions are in close proximity by physical interaction) is associated with early transcriptional activation [63]. It appears that the looped configuration is indispensable for splicing-competent intronic DNA to enhance mRNA accumulation and that the gene looping is crucial for the recruitment of termination factors in the promoter proximal region of an intron-containing gene [62]. Due to the unidirectional nature of plant promoters, it is not entirely clear whether or not IME in plants is achieved in a splicing-competent intron-dependent manner.

##### 4.2. Splicing-Independent IME in Plants

Investigations into plant IME have been extensively carried out by Rose’s group [65,66]. For instance, first introns of maize *ADH1* [55], *SHRUNKEN-1* [67,68], *GAP1* [69], *UBI1* [70], rice *SODCC2* (cytosolic superoxide dismutase) [71], *Arabidopsis PAT1* (phosphoribosylanthranilate transferase 1) [72], *UBIQUITIN (UBQ)* genes [73], and *MHX* [74,75] have been documented to boost mRNA accumulation. Hereafter, intron that can significantly increase gene expression is termed “stimulating intron.” Based on the assumption that introns in close proximity to TSS would be more frequently enriched in IME signals than distal introns, genome-wide in silico surveys of *Arabidopsis* and rice were carried out, and a word-based discriminator called the IMEter was developed [76]. Consequently, the most common motifs found using NestedMICA were CGATT and TTNGATYTG in *Arabidopsis* and TCGATC in rice [76–78]. Hereafter, the TTNGATYTG sequence is termed “IME motif.” A similar consensus sequence responsible for IME in rice was found too: GATCTG [71]. The ability of IME motif enriched in promoter-proximal introns to boost mRNA accumulation was confirmed by a gain-of-function approach of converting non-stimulating intron into stimulating intron by integrating IME motif [77]. The unique features of the IME motif found in *Arabidopsis* have been further revealed. Contrary



to IME observed in yeast cells [61,62], splicing is not required for plant-stimulating introns [79,80]. The *UBQ10* first intron not only increases mRNA accumulation, but also affects tissue specificity when the reporter gene containing the *UBQ10* first intron is under the control of different promoters [81], suggesting that *UBQ10* intron renders tissue specificity over promoter sequence. Likewise, the first intron of *PROFILIN2 (PRF2)*, encoding a small actin-binding protein expressed in vegetative organs, is capable of boosting mRNA accumulation and confers vegetative tissue expression when driven by reproductive tissue-specific *PRF5* promoter [82].

Intriguingly, the *UBQ10* first intron determines the transcription start site (TSS) from where it locates (Figure 2) [83]. In order for an intron to accumulate mRNA, its position has to be within 1000 bp downstream from the TSS [83,84]. Contrary to its name (intron-mediated enhancement), IME takes place in a splicing-independent manner because its ability to boost mRNA accumulation persists even when IME motif is inserted in 5'-UTR (IMEter score is high in 5'-UTR) or present in the coding region (as long as it is positioned within 1000 bp downstream from TSS) [85,86]. Although IME motif is capable of increasing translational efficiency [72,75], it is likely that it is more relevant to IME when located in the DNA rather than the strand-specific nascent RNA (prior to splicing) [87], which is in agreement with its ability to determine TSS [83].

### 5. Post-Transcriptional Regulation by Intron Retention (IR) within 5'-UTR

Intron retention (IR) is the major AS in plants [88,89], but the idea that IR contributes to protein diversity is controversial [90–92]. The outcome of IR differs depending on the position of the retained first intron, or most AS events are the consequence of stochastic noise in the process machinery [93]. In plants, around 15% of all AS events take place within the 5'-UTR, which can affect the transport and stability of mRNAs that are kept under constant surveillance by the nonsense-mediated decay (NMD) pathway [10,94,95].

When AS of the first intron within 5'-UTR takes place, the impact on gene expression will be quite the opposite. A positive effect of the first intron retention on boosting mRNA accumulation was shown by Laxa et al.: the first intron within 5'-UTR of Glu:glyoxylate aminotransferase 1 (*GGT1*) contains a TGTGATTTG sequence that is highly similar to IME motif and enhances leaf-specific expression at the transcriptional level by enriching Pol II abundance [96]. The topological secondary structure built by base-pairing also accounts for the accessibility of translational machinery. First intron retention within 5'-UTR of *ZIF2* (encoding ZINC-INDUCED FACILITATOR 2) is shown to provide a stable stem-loop structure that enables more efficient translation when compared to the spliced variant, in which the stable structure is lost by splicing [97]. On the contrary, base-pairing interaction between retained intron and 5'-UTR of yeast *HAC1* represses translation [98].

The vast number of eukaryotic mRNAs contain a short open-reading frame within their 5'-UTR, called upstream ORF (uORF). In *Arabidopsis*, about 35% of genes transcribing mRNA contain at least one uORF by prediction [99], and high-resolution ribosome profiling verified 187 uORFs being translated [100]. The presence of uORF within 5'-UTR negatively affects the translation of main ORF by ribosome stalling [101–103]. Although it is a process whose molecular mechanism is still being investigated, the repression of main ORF by uORF acts in both a uORF peptide translation-dependent [104] and -independent manner [105,106].

Riboswitches are regulatory elements in mRNA, predominantly found in the noncoding region, that act as metabolite sensors by selectively binding to ligands and regulate transcription or translation without the need for protein factors [107]. It is prevalent in prokaryotes, and the most widespread riboswitch known in bacteria responds to the coenzyme thiamine pyrophosphate (TPP), which is also conserved in fungi and plants [108–110]. Álvarez et al. proposed alluring retrograde signaling as a feedback regulatory mechanism of carotenogenesis, where *ASV1*, one of the splicing variants containing longer 5'-UTR of PHYTOENE SYNTHASE (*PSY*) mRNA, acting as a riboswitch-like module to repress translation and AS of the first intron is the key to modulating retrograde signaling [111]. Similarly, a phytohormone riboswitch has been proposed. Grojean and Downes hypothesized the involvement

of riboswitch in cytokinin perception (beyond membrane-bound histidine kinase cytokinin receptors AHK2, AHK3, and AHK4) in *Arabidopsis* [112]. Having gone through biochemistry combined with bioinformatics data of the cytokinin-responsive gene, one candidate encoding CYSTEINE-RICH RECEPTOR-LIKE KINASE 10 (*CRK10*), whose expression is downregulated by cytokinin application, is likely to contain an adenine aptamer-related sequence in its first intronic sequence [112]. In both cases, further experimental results to corroborate their hypothesis are anticipated.

## 6. Regulation of First Intron Retention Located Downstream of ATG

IR that takes place in introns positioned after the ATG affects protein in its sequence, function, subcellular localization, or IR mRNAs is subjected to degradation by the NMD [113,114]. IR, in many cases, causes disruption of original ORF by premature stop codons that subsequently activate the NMD pathway. However, not all IR transcripts are sensitive to NMD [95]. The splicing of retained introns is facilitated by stress signals. Tomato WRKY transcriptional factor *SlyWRKY75* contains the first intron downstream of ATG, and its expression is regulated by *miR1127-3p*, which is complementary to the first intron sequence of *SlyWRKY75*. Under noninfected conditions, the first intron is retained, whereas upon *Botrytis cinerea* pathogen challenge, the overall *SlyWRKY75* transcript level is elevated, which is in sharp contrast to compromised *miR1127-3p* expression. Interestingly, pathogen challenges facilitate splicing of the retained first intron that remained unspliced under noninfected conditions [115]. These results imply that intron-targeting *miR1127-3p* binds to *SlyWRKY75* pre-mRNA in the nucleus to inhibit the splicing of target intronic region under noninfected conditions.

## 7. AS in Response to Stresses and Developmental Cues

Pathogen infection [115] is not the only example to show the effect of splicing a retained intron. Sessile plants have adopted sophisticated mechanisms to adjust to the varying environment, and the regulation of AS appears to play a more pivotal role than previously assumed. Splicing efficiency is regulated by stresses such as abiotic stress [116], drought [117], salt [118], cold [119,120], high temperature [34,121] phytochrome B signaling [122,123] thermo-priming [124], temperature-dependent regulation of circadian clock genes [125], and phosphate starvation [126]. Vernalization is, in part, regulated by AS of *FLC*, a central repressor in the flowering transition [37]. Tsugeki et al. suggested the crucial role of CLUMSY VEIN (*CUV*) encoding DEAD-Box RNA-dependent ATPase *PRP16* in facilitating the splicing of genes involved in auxin biosynthesis, transport, perception, and signal transduction [127]. AS of the splicing factor *At-RS31* has been shown to be regulated under retrograde signaling control [128]. Light/dark transition affects AS via the plastoquinone redox state [128], suggesting close communication between two organelles for fine-tuning of gene expression through AS. Such close communication is further supported by a study of LEFKOTHEA (*LEF*) encoding PORR domain-containing RNA-binding protein [129]. Nuclear-encoded LEF proteins exhibit dual targeting to the nucleus and chloroplasts to regulate AS of nuclear-encoded genes as well as group II introns in chloroplast [129]. According to the function and position of the first intron (relative to the main ORF ATG translation initiation site), genes whose expression is under the influence of AS of the first intron are summarized in Table 1. The biggest obstacle to account for the modulation of AS by environmental cues is the lack of knowledge of their perception mechanism (i.e., stress receptors). Besides, it should be noted that the spliceosome is among the largest molecular complexes in the cell, comprising five snRNAs and more than 150 proteins [130], which also impede our comprehension of stress-induced AS changes.

**Table 1.** Plant first intron with functions and stress responses.

Category	Gene Name (Gene ID)	Position Relative to ATG	Function	Organism	Reference
IME	<i>UBQ10</i> (At4g05320)	downstream	IME	<i>Arabidopsis</i>	[80]
	<i>PAT1/TRP1</i> (At5g17990)	downstream	IME	<i>Arabidopsis</i>	[72]
	<i>ADH1</i> (GRMZM2G442658)	downstream	IME	maize	[55]
	<i>Sh1</i> (GRMZM2G089713)	downstream	IME, splicing-dependent	maize	[56,68]
	<i>PRF1</i> (At2g19760)	downstream	IME	<i>Arabidopsis</i>	[82]
	<i>PRF2</i> (At4g29350)	downstream	IME	<i>Arabidopsis</i>	[82]
miRNA/ lncRNA	<i>OsCLV1</i> (Os1g07060)	upstream	target of miR2123abc	rice	[35]
	<i>YSL3</i> (At5g53550)	upstream	miRNA-binding site	<i>Arabidopsis</i>	[35]
	<i>SLK2</i> (At5g62090)	upstream	miRNA-binding site	<i>Arabidopsis</i>	[35]
	<i>SlyWRKY75</i> (Solyco05g015850)	downstream	target of sly-miR1127	tomato	[115]
	n.a. (At1g32583)	upstream	source of miR400	<i>Arabidopsis</i>	[34]
	n.a. (At5g08185)	downstream	source of miR162a	<i>Arabidopsis</i>	[131]
	<i>NPF2.9/NRT1.9</i> (At1g18880)	downstream	source of miR837	<i>Arabidopsis</i>	[131]
	n.a. (At2g23348)	upstream	source of miR844	<i>Arabidopsis</i>	[131]
<i>PIF3</i> (At1g09530)	upstream	target of HID1	<i>Arabidopsis</i>	[36]	
IR secondarystructure riboswitch	<i>ZIF2</i> (At2g48020)	upstream	translation enhancement	<i>Arabidopsis</i>	[97]
	<i>PSY</i> (At5g17230)	upstream	translation inhibition	<i>Arabidopsis</i>	[111]
riboswitch	<i>CRK10</i> (At4g23180)	downstream	putative cytokinin-binding riboswitch	<i>Arabidopsis</i>	[112]
uORF	<i>PIF3</i> (At1g09530)	upstream	translation inhibition	<i>Arabidopsis</i>	[123]
uORF	<i>ARF3</i> (At2g24765)	upstream	translation inhibition	<i>Arabidopsis</i>	[132]
uORF	<i>OsNLA1</i> (Os07g0673200)	upstream	confers phosphate-dependent induction	Rice	[126]
IR	<i>ATGGT-IB</i> (At2g39550)	downstream	enhanced splicing upon cold	<i>Arabidopsis</i>	[119]
	n.a. (At2g43160)	downstream	enhanced splicing upon cold	<i>Arabidopsis</i>	[119]
	n.a. (At3g47630)	downstream	enhanced splicing upon cold	<i>Arabidopsis</i>	[119]
	<i>CDKG1</i> (At5g63370)	downstream	altered subcellular localization	<i>Arabidopsis</i>	[133]
	n.a. (At3g48070)	downstream	target of the LEF	<i>Arabidopsis</i>	[129]
	n.a. (At2g37510)	downstream	target of the LEF	<i>Arabidopsis</i>	[129]
	n.a. (At3g19920)	downstream	target of the LEF	<i>Arabidopsis</i>	[129]
	n.a. (At4g30993)	downstream	target of SME1	<i>Arabidopsis</i>	[120]
	<i>RD28/PIP2C</i> (At2g37180)	downstream	target of SME1	<i>Arabidopsis</i>	[120]
	n.a. (At3g26360)	downstream	target of SME1	<i>Arabidopsis</i>	[120]
	<i>TTA2/OBE4</i> (At3g63500)	downstream	target of SME1	<i>Arabidopsis</i>	[120]
	<i>IPGAM2</i> (At3g08590)	downstream	target of SME1	<i>Arabidopsis</i>	[120]
	<i>SPL2</i> (At5g43270)	downstream	target of SE	<i>Arabidopsis</i>	[53]
	<i>APE2</i> (At5g46110)	downstream	target of SE	<i>Arabidopsis</i>	[53]
	<i>HsfA2</i> (At2g26150)	downstream	target of SE	<i>Arabidopsis</i>	[53]
	<i>SnrK2.8</i> (At1g78290)	downstream	target of SE	<i>Arabidopsis</i>	[53]
	<i>YUC3</i> (At1g04610)	downstream	target of CUV	<i>Arabidopsis</i>	[127]
	<i>SHY2/IAA3</i> (At1g04240)	downstream	target of CUV	<i>Arabidopsis</i>	[127]
	<i>AXR3/IAA17</i> (At1g04250)	downstream	target of CUV	<i>Arabidopsis</i>	[127]
	<i>NPH4/ARF7</i> (At5g20730)	downstream	target of CUV	<i>Arabidopsis</i>	[127]
	<i>SLR/IAA14</i> (At4g14550)	downstream	target of CUV	<i>Arabidopsis</i>	[127]
	n.a. (At4G27050)	downstream	target of AtNSRa/b	<i>Arabidopsis</i>	[54]
	<i>AtDRM2, DAP2</i> (At2G33830)	downstream	target of AtNSRa/b	<i>Arabidopsis</i>	[54]

\* non-annotated (n.a.). Abbreviations: *HID1*, HIDDEN TREASURE 1; *LEF*, LEFKOTHEA; *SME1*, Sm protein E1; *SE*, SERRATE; *CUV*, CLUMSY VEIN; *AtNSR1a/b*, NUCLEAR SPECKLE RNA-BINDING PROTEIN a/b.

## 8. Intron Regulation beyond the Host Gene

In eukaryote cells, transcribing intronic regions and correctly finding junctions to cut and subsequently combine both ends are energy-consuming and laborious. Indeed, introns are junk on a good day but can be treasures on a bad day in the case of budding yeast [134]. In yeast, global splicing efficiency is affected by the expression level of ribosomal protein genes (RPGs) because the splicing apparatus is a limiting factor for non-RPG pre-mRNAs [135]. Parenteau et al. elegantly concluded that introns, as part of the unspliced pre-mRNA, promote resistance to nutrient starvation by enhancing the repression of RPGs [136]. In addition, a small subset of spliced introns with short distances between the lariat branch point and 3'-SS (lariat RNA, a byproduct of spliced introns) could also promote cell survival under starvation conditions [137]. Lariat RNAs control miRNA biogenesis by interacting with the dicing complex. A similar mechanism is likely to be conserved in humans [138] and plants [139]. *Arabidopsis* weak mutant allele, *db1-2*, defective in RNA debranching enzyme 1 (*DBR1*) accumulates lariat RNAs that act as a decoy by competing with the dicing complex to inhibit genome-wide miRNA processing [139]. Collectively, these results strongly suggest in trans regulation of excised intron beyond host gene expression.



## 9. Conclusions and Future Aspects

Information on defined epigenetic marks deposited on the first intronic DNA sequence and the higher sequence conservation within the longest first introns in all domains of life suggests that the first intronic DNA sequence is more distinguishable than the other introns. The underlying mechanisms of promoter-proximal introns to boost mRNA accumulation and determine TSS from downstream, a phenomenon called intron-mediated enhancement (IME), appear different from those of the initiation of DNA sequence-specific transcription by transcription factors or conventional enhancers. The biggest breakthrough for a better understanding of IME in plants was the discovery of the consensus IME motif enriched in introns capable of boosting mRNA accumulation [76]. However, it is not entirely clear whether the IME motif is unique in the plant kingdom.

When pondering the underlying mechanism for stimulating introns to determine TSS from the gene body in a relatively short range (up to about 1000 bp) [83], IME motif is likely to fulfill its potential when located in the DNA rather than when transcribed in pre-mRNA [87]. Is stimulating intron marked with specific epigenetic marks on DNA? Analogous to yeast, does stimulating intron facilitate a looped configuration of genes? What kind of underlying mechanisms are there for the enhancing intron to determine TSS? Our exploration of the unique features of plant first intron is ongoing.

**Author Contributions:** D.Z. and Y.I. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Czech Science Foundation grants; GACR17-23702S and GACR 18-23972Y.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Berget, S.M.; Moore, C.; Sharp, P.A. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 3171–3175. [[CrossRef](#)] [[PubMed](#)]
- Chow, L.T.; Gelinis, R.E.; Broker, T.R.; Roberts, R.J. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **1977**, *12*, 1–8. [[CrossRef](#)]
- Gilbert, W. Why genes in pieces? *Nature* **1978**, *271*, 501. [[CrossRef](#)]
- Catania, F.; Lynch, M. Where do introns come from? *PLoS Biol.* **2008**, *6*, 2354–2361. [[CrossRef](#)]
- Rogozin, I.B.; Carmel, L.; Csuros, M.; Koonin, E.V. Origin and evolution of spliceosomal introns. *Biol. Direct* **2012**, *7*, 1–28. [[CrossRef](#)] [[PubMed](#)]
- Irimia, M.; Roy, S.W. Origin of spliceosomal introns and alternative splicing. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a016071. [[CrossRef](#)]
- Koonin, E.V.; Csuros, M.; Rogozin, I.B. Whence genes in pieces: Reconstruction of the exon-intron gene structures of the last eukaryotic common ancestor and other ancestral eukaryotes. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 93–105. [[CrossRef](#)]
- Csuros, M.; Rogozin, I.B.; Koonin, E.V. A detailed history of intron-rich eukaryotic ancestors inferred from a global survey of 100 complete genomes. *PLoS Comput. Biol.* **2011**, *7*, e1002150. [[CrossRef](#)]
- Braunschweig, U.; Gueroussov, S.; Plocik, A.M.; Graveley, B.R.; Blencowe, B.J. Dynamic integration of splicing within gene regulatory pathways. *Cell* **2013**, *152*, 1252–1269. [[CrossRef](#)]
- Barbazuk, W.B.; Fu, Y.; McGinnis, K.M. Genome-wide analyses of alternative splicing in plants: Opportunities and challenges. *Genome Res.* **2008**, *18*, 1382–1391. [[CrossRef](#)]
- Alló, M.; Schor, I.E.; Muñoz, M.J.; De La Mata, M.; Agirre, E.; Valcárcel, J.; Eyra, E.; Kornblihtt, A.R. Chromatin and alternative splicing. *Cold Spring Harb. Symp. Quant. Biol.* **2010**, *75*, 103–111. [[CrossRef](#)] [[PubMed](#)]
- Le Hir, H.; Nott, A.; Moore, M.J. How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* **2003**, *28*, 215–220. [[CrossRef](#)]
- Szakonyi, D.; Duque, P. Alternative splicing as a regulator of early plant development. *Front. Plant Sci.* **2018**, *9*, 1–9. [[CrossRef](#)] [[PubMed](#)]
- Shang, X.; Cao, Y.; Ma, L. Alternative splicing in plant genes: A means of regulating the environmental fitness of plants. *Int. J. Mol. Sci.* **2017**, *18*, 432. [[CrossRef](#)] [[PubMed](#)]

15. Mastrangelo, A.M.; Marone, D.; Laidò, G.; De Leonardis, A.M.; De Vita, P. Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. *Plant Sci.* **2012**, *185–186*, 40–49. [[CrossRef](#)]
16. Jabre, I.; Reddy, A.S.N.; Kalyna, M.; Chaudhary, S.; Khokhar, W.; Byrne, L.J.; Wilson, C.M.; Syed, N.H. Does co-transcriptional regulation of alternative splicing mediate plant stress responses? *Nucl. Acids Res.* **2019**, *47*, 2716–2726. [[CrossRef](#)]
17. Park, S.G.; Hannenhalli, S.; Choi, S.S. Conservation in first introns is positively associated with the number of exons within genes and the presence of regulatory epigenetic signals. *BMC Genomics* **2014**, *15*, 526. [[CrossRef](#)]
18. Jo, S.S.; Choi, S.S.; Hurst, L. Analysis of the functional relevance of epigenetic chromatin marks in the first intron associated with specific gene expression patterns. *Genome Biol. Evol.* **2019**, *11*, 786–797. [[CrossRef](#)]
19. Zhang, X.; Bernatavichute, Y.V.; Cokus, S.; Pellegrini, M.; Jacobsen, S.E. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* **2009**, *10*, 1–14. [[CrossRef](#)]
20. Schwartz, S.; Meshorer, E.; Ast, G. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* **2009**, *16*, 990–995. [[CrossRef](#)]
21. Chodavarapu, R.K.; Feng, S.; Bernatavichute, Y.V.; Chen, P.Y.; Stroud, H.; Yu, Y.; Hetzel, J.A.; Kuo, F.; Kim, J.; Cokus, S.J.; et al. Relationship between nucleosome positioning and DNA methylation. *Nature* **2010**, *466*, 388–392. [[CrossRef](#)] [[PubMed](#)]
22. Pass, D.A.; Sornay, E.; Marchbank, A.; Crawford, M.R.; Paszkiewicz, K.; Kent, N.A.; Murray, J.A.H. Genome-wide chromatin mapping with size resolution reveals a dynamic sub-nucleosomal landscape in *Arabidopsis*. *PLoS Genet.* **2017**, *13*, e1006988. [[CrossRef](#)] [[PubMed](#)]
23. Deremetz, A.; Le Roux, C.; Idir, Y.; Brousse, C.; Agorio, A.; Gy, I.; Parker, J.E.; Bouché, N. Antagonistic actions of FPA and IBM2 regulate transcript processing from genes containing heterochromatin. *Plant Physiol.* **2019**, *180*, 392–403. [[CrossRef](#)] [[PubMed](#)]
24. Bieberstein, N.I.; Oesterreich, F.C.; Straube, K.; Neugebauer, K.M. First exon length controls active chromatin signatures and transcription. *Cell Rep.* **2012**, *2*, 62–68. [[CrossRef](#)]
25. Gelfman, S.; Cohen, N.; Yearim, A.; Ast, G. DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon-intron structure. *Genome Res.* **2013**, *23*, 789–799. [[CrossRef](#)]
26. Lev Maor, G.; Yearim, A.; Ast, G. The alternative role of DNA methylation in splicing regulation. *Trends Genet.* **2015**, *31*, 274–280. [[CrossRef](#)]
27. Wong, J.J.L.; Gao, D.; Nguyen, T.V.; Kwok, C.T.; Van Geldermalsen, M.; Middleton, R.; Pinello, N.; Thoeng, A.; Nagarajah, R.; Holst, J.; et al. Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. *Nat. Commun.* **2017**, *8*, 1–13. [[CrossRef](#)]
28. Ullah, F.; Hamilton, M.; Reddy, A.S.N.; Ben-Hur, A. Exploring the relationship between intron retention and chromatin accessibility in plants. *BMC Genomics* **2018**, *19*, 21. [[CrossRef](#)]
29. Wang, X.; Hu, L.; Wang, X.; Li, N.; Xu, C.; Gong, L.; Liu, B. DNA methylation affects gene alternative splicing in plants: An example from rice. *Mol. Plant* **2016**, *9*, 305–307. [[CrossRef](#)]
30. Chorev, M.; Joseph Bekker, A.; Goldberger, J.; Carmel, L. Identification of introns harboring functional sequence elements through positional conservation. *Sci. Rep.* **2017**, *7*, 1–13. [[CrossRef](#)]
31. Hou, J.; Lu, D.; Mason, A.S.; Li, B.; Xiao, M.; An, S.; Fu, D. Non-coding RNAs and transposable elements in plant genomes: Emergence, regulatory mechanisms and roles in plant development and stress responses. *Planta* **2019**, *250*, 23–40. [[CrossRef](#)] [[PubMed](#)]
32. Espinas, N.A.; Tu, L.N.; Furci, L.; Shimajiri, Y.; Harukawa, Y.; Miura, S.; Takuno, S.; Saze, H. Transcriptional regulation of genes bearing intronic heterochromatin in the rice genome. *PLoS Genet.* **2020**, *16*, e1008637. [[CrossRef](#)] [[PubMed](#)]
33. Le, T.N.; Miyazaki, Y.; Takuno, S.; Saze, H. Epigenetic regulation of intragenic transposable elements impacts gene transcription in *Arabidopsis thaliana*. *Nucl. Acids Res.* **2015**, *43*, 3911–3921. [[CrossRef](#)] [[PubMed](#)]
34. Yan, K.; Liu, P.; Wu, C.A.; Yang, G.D.; Xu, R.; Guo, Q.H.; Huang, J.G.; Zheng, C.C. Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in *Arabidopsis thaliana*. *Mol. Cell* **2012**, *48*, 521–531. [[CrossRef](#)] [[PubMed](#)]
35. Meng, Y.; Shao, C.; Ma, X.; Wang, H. Introns targeted by plant microRNAs: A possible novel mechanism of gene regulation. *Rice* **2013**, *6*, 1–10. [[CrossRef](#)] [[PubMed](#)]
36. Wang, Y.; Fan, X.; Lin, F.; He, G.; Terzaghi, W.; Zhu, D.; Deng, X.W. *Arabidopsis* noncoding RNA mediates control of photomorphogenesis by red light. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10359–10364. [[CrossRef](#)] [[PubMed](#)]

37. Qüesta, J.I.; Song, J.; Geraldo, N.; An, H.; Dean, C. Arabidopsis transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization. *Science* **2016**, *353*, 485–488. [[CrossRef](#)]
38. Churchman, L.S.; Weissman, J.S. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **2011**, *469*, 368–373. [[CrossRef](#)]
39. Nojima, T.; Gomes, T.; Grosso, A.R.F.; Kimura, H.; Dye, M.J.; Dhir, S.; Carmo-Fonseca, M.; Proudfoot, N.J. Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* **2015**, *161*, 526–540. [[CrossRef](#)]
40. Leng, X.; Ivanov, M.; Kindgren, P.; Malik, I.; Thieffry, A.; Brodersen, P.; Sandelin, A.; Kaplan, C.D.; Marquardt, S. Organismal benefits of transcription speed control at gene boundaries. *EMBO Rep.* **2020**, *21*, 1–16. [[CrossRef](#)]
41. Godoy Herz, M.A.; Kornblihtt, A.R. Alternative splicing and transcription elongation in plants. *Front. Plant Sci.* **2019**, *10*, 1–5. [[CrossRef](#)] [[PubMed](#)]
42. Fong, N.; Kim, H.; Zhou, Y.; Ji, X.; Qiu, J.; Saldi, T.; Diener, K.; Jones, K.; Fu, X.D.; Bentley, D.L. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev.* **2014**, *28*, 2663–2676. [[CrossRef](#)] [[PubMed](#)]
43. Dujardin, G.; Lafaille, C.; de la Mata, M.; Marasco, L.E.; Muñoz, M.J.; Le Jossic-Corcós, C.; Corcos, L.; Kornblihtt, A.R. How slow RNA polymerase II elongation favors alternative exon skipping. *Mol. Cell* **2014**, *54*, 683–690. [[CrossRef](#)] [[PubMed](#)]
44. Zhu, J.; Liu, M.; Liu, X.; Dong, Z. RNA polymerase II activity revealed by GRO-seq and pNET-seq in Arabidopsis. *Nat. Plants* **2018**, *4*, 1112–1123. [[CrossRef](#)]
45. Wallace, E.W.J.; Beggs, J.D. Extremely fast and incredibly close: Cotranscriptional splicing in budding yeast. *RNA* **2017**, *23*, 601–610. [[CrossRef](#)] [[PubMed](#)]
46. Nojima, T.; Rebelo, K.; Gomes, T.; Grosso, A.R.; Proudfoot, N.J.; Carmo-Fonseca, M. RNA Polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. *Mol. Cell* **2018**, *72*, 369–379.e4. [[CrossRef](#)]
47. Fong, N.; Saldi, T.; Sheridan, R.M.; Cortazar, M.A.; Bentley, D.L. RNA pol II dynamics modulate co-transcriptional chromatin modification, CTD phosphorylation, and transcriptional direction. *Mol. Cell* **2017**, *66*, 546–557.e3. [[CrossRef](#)]
48. Jia, J.; Long, Y.; Zhang, H.; Li, Z.; Liu, Z.; Zhao, Y.; Lu, D.; Jin, X.; Deng, X.; Xia, R.; et al. Post-transcriptional splicing of nascent RNA contributes to widespread intron retention in plants. *Nat. Plants* **2020**, *6*, 780–788. [[CrossRef](#)]
49. Harlen, K.M.; Churchman, L.S. The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 263–273. [[CrossRef](#)]
50. Izaurralde, E.; Lewis, J.; McGuigan, C.; Jankowska, M.; Darzynkiewicz, E.; Mattaj, I.W. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **1994**, *78*, 657–668. [[CrossRef](#)]
51. Raczynska, K.D.; Simpson, C.G.; Ciesiolka, A.; Szewc, L.; Lewandowska, D.; McNicol, J.; Szweykowska-Kulinska, Z.; Brown, J.W.S.; Jarmolowski, A. Involvement of the nuclear cap-binding protein complex in alternative splicing in Arabidopsis thaliana. *Nucl. Acids Res.* **2009**, *38*, 265–278. [[CrossRef](#)]
52. Laubinger, S.; Sachsenberg, T.; Zeller, G.; Busch, W.; Lohmann, J.U.; Rättsch, G.; Weigel, D. Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 8795–8800. [[CrossRef](#)] [[PubMed](#)]
53. Raczynska, K.D.; Stepien, A.; Kierzkowski, D.; Kalak, M.; Bajczyk, M.; McNicol, J.; Simpson, C.G.; Szweykowska-Kulinska, Z.; Brown, J.W.S.; Jarmolowski, A. The SERRATE protein is involved in alternative splicing in Arabidopsis thaliana. *Nucl. Acids Res.* **2014**, *42*, 1224–1244. [[CrossRef](#)]
54. Bardou, F.; Ariel, F.; Simpson, C.G.; Romero-Barrios, N.; Laporte, P.; Balzergue, S.; Brown, J.W.S.; Crespi, M. Long noncoding RNA modulates alternative splicing regulators in Arabidopsis. *Dev. Cell* **2014**, *30*, 166–176. [[CrossRef](#)] [[PubMed](#)]
55. Callis, J.; Fromm, M.; Walbot, V. Introns increase gene expression in cultured maize cells. *Genes Dev.* **1987**, *1*, 1183–1200. [[CrossRef](#)] [[PubMed](#)]
56. Clancy, M.; Curtis Hannah, L. Splicing of the maize Sh1 first intron is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing. *Plant Physiol.* **2002**, *130*, 918–929. [[CrossRef](#)] [[PubMed](#)]
57. Core, L.J.; Waterfall, J.J.; Lis, J.T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **2008**, *322*, 1845–1848. [[CrossRef](#)] [[PubMed](#)]

58. Wu, X.; Sharp, P.A. Divergent transcription: A driving force for new gene origination? *Cell* **2013**, *155*, 990–996. [[CrossRef](#)] [[PubMed](#)]
59. Mayer, A.; Di Iulio, J.; Maleri, S.; Eser, U.; Vierstra, J.; Reynolds, A.; Sandstrom, R.; Stamatoyannopoulos, J.A.; Churchman, L.S. Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **2015**, *161*, 541–554. [[CrossRef](#)]
60. Hetzel, J.; Duttke, S.H.; Benner, C.; Chory, J. Nascent RNA sequencing reveals distinct features in plant transcription. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12316–12321. [[CrossRef](#)]
61. Furger, A.; O’Sullivan, J.M.; Binnie, A.; Lee, B.A.; Proudfoot, N.J. Promoter proximal splice sites enhance transcription. *Genes Dev.* **2002**, *16*, 2792–2799. [[CrossRef](#)] [[PubMed](#)]
62. Agarwal, N.; Ansari, A. Enhancement of transcription by a splicing-competent intron is dependent on promoter directionality. *PLoS Genet.* **2016**, *12*, 1006047. [[CrossRef](#)] [[PubMed](#)]
63. O’Sullivan, J.M.; Tan-Wong, S.M.; Morillon, A.; Lee, B.; Coles, J.; Mellor, J.; Proudfoot, N.J. Gene loops juxtapose promoters and terminators in yeast. *Nat. Genet.* **2004**, *36*, 1014–1018. [[CrossRef](#)]
64. Moabbi, A.M.; Agarwal, N.; El Kaderi, B.; Ansari, A. Role for gene looping in intron-mediated enhancement of transcription. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8505–8510. [[CrossRef](#)] [[PubMed](#)]
65. Gallegos, J.E.; Rose, A.B. The enduring mystery of intron-mediated enhancement. *Plant Sci.* **2015**, *237*, 8–15. [[CrossRef](#)] [[PubMed](#)]
66. Rose, A.B. Introns as gene regulators: A brick on the accelerator. *Front. Genet.* **2019**, *10*, 1–6. [[CrossRef](#)]
67. Vasil, V.; Clancy, M.; Ferl, R.J.; Vasil, I.K.; Hannah, L.C. Increased gene expression by the first intron of maize shrunken-1 locus in grass species. *Plant. Physiol.* **1989**, *91*, 1575–1579. [[CrossRef](#)]
68. Clancy, M.; Vasil, V.; Curtis Hannah, L.; Vasil, I.K. Maize Shrunken-1 intron and exon regions increase gene expression in maize protoplasts. *Plant Sci.* **1994**, *98*, 151–161. [[CrossRef](#)]
69. Donath, M.; Mendel, R.; Cerff, R.; Martin, W. Intron-dependent transient expression of the maize GapA1 gene. *Plant Mol. Biol.* **1995**, *28*, 667–676. [[CrossRef](#)]
70. Vain, P.; Finer, K.R.; Engler, D.E.; Pratt, R.C.; Finer, J.J. Intron-mediated enhancement of gene expression in maize (*Zea mays L.*) and bluegrass (*Poa pratensis L.*). *Plant. Cell Rep.* **1996**, *15*, 489–494. [[CrossRef](#)]
71. Morita, S.; Tsukamoto, S.; Sakamoto, A.; Makino, H.; Nakauji, E.; Kaminaka, H.; Masumura, T.; Ogiyama, Y.; Satoh, S.; Tanaka, K. Differences in intron-mediated enhancement of gene expression by the first intron of cytosolic superoxide dismutase gene from rice in monocot and dicot plants. *Plant Biotechnol.* **2012**, *29*, 115–119. [[CrossRef](#)]
72. Rose, A.B.; Last, R.L. Introns act post-transcriptionally to increase expression of the Arabidopsis thaliana tryptophan pathway gene PAT1. *Plant J.* **1997**, *11*, 455–464. [[CrossRef](#)] [[PubMed](#)]
73. Norris, S.R.; Meyer, S.E.; Callis, J. The intron of Arabidopsis thaliana polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. *Plant Mol. Biol.* **1993**, *21*, 895–906. [[CrossRef](#)] [[PubMed](#)]
74. David-Assael, O.; Berezin, I.; Shoshani-Knaani, N.; Saul, H.; Mizrachy-Dagri, T.; Chen, J.; Brook, E.; Shaul, O. AtMHX is an auxin and ABA-regulated transporter whose expression pattern suggests a role in metal homeostasis in tissues with photosynthetic potential. *Funct. Plant. Biol.* **2006**, *33*, 661–672. [[CrossRef](#)]
75. Akua, T.; Shaul, O. The Arabidopsis thaliana MHX gene includes an intronic element that boosts translation when localized in a 5’ UTR intron. *J. Exp. Bot.* **2013**, *64*, 4255–4270. [[CrossRef](#)] [[PubMed](#)]
76. Rose, A.B.; Elfersi, T.; Parra, G.; Korf, I. Promoter-proximal introns in Arabidopsis thaliana are enriched in dispersed signals that elevate gene expression. *Plant Cell* **2008**, *20*, 543–551. [[CrossRef](#)]
77. Rose, A.B.; Carter, A.; Korf, I.; Kojima, N. Intron sequences that stimulate gene expression in Arabidopsis. *Plant Mol. Biol.* **2016**, *92*, 337–346. [[CrossRef](#)]
78. Bradnam, K.R.; Korf, I. Longer first introns are a general property of eukaryotic gene structure. *PLoS ONE* **2008**, *3*. [[CrossRef](#)]
79. Rose, A.B.; Beliakoff, J.A. Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing. *Plant Physiol.* **2000**, *122*, 535–542. [[CrossRef](#)]
80. Rose, A.B. Requirements for intron-mediated enhancement of gene expression in Arabidopsis. *RNA* **2002**, *8*, 1444–1453. [[CrossRef](#)]
81. Emami, S.; Arumainayagam, D.; Korf, I.; Rose, A.B. The effects of a stimulating intron on the expression of heterologous genes in Arabidopsis thaliana. *Plant Biotechnol. J.* **2013**, *11*, 555–563. [[CrossRef](#)] [[PubMed](#)]



82. Jeong, Y.M.; Mun, J.H.; Lee, I.; Woo, J.C.; Hong, C.B.; Kim, S.G. Distinct roles of the first introns on the expression of Arabidopsis profilin gene family members. *Plant Physiol.* **2006**, *140*, 196–209. [[CrossRef](#)]
83. Gallegos, J.E.; Rose, A.B. Intron DNA sequences can be more important than the proximal promoter in determining the site of transcript initiation. *Plant Cell* **2017**, *29*, 843–853. [[CrossRef](#)] [[PubMed](#)]
84. Rose, A.B. The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. *Plant. J.* **2004**, *40*, 744–751. [[CrossRef](#)] [[PubMed](#)]
85. Gallegos, J.E.; Rose, A.B. An intron-derived motif strongly increases gene expression from transcribed sequences through a splicing independent mechanism in *Arabidopsis thaliana*. *Sci. Rep.* **2019**, *9*, 1–9. [[CrossRef](#)]
86. Parra, G.; Bradnam, K.; Rose, A.B.; Korf, I. Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants. *Nucl. Acids Res.* **2011**, *39*, 5328–5337. [[CrossRef](#)]
87. Rose, A.B.; Emami, S.; Bradnam, K.; Korf, I. Evidence for A DNA-based mechanism of intron-mediated enhancement. *Front. Plant Sci.* **2011**, *2*, 1–9. [[CrossRef](#)]
88. Ner-Gaon, H.; Halachmi, R.; Savaldi-Goldstein, S.; Rubin, E.; Ophir, R.; Fluhr, R. Intron retention is a major phenomenon in alternative splicing in Arabidopsis. *Plant J.* **2004**, *39*, 877–885. [[CrossRef](#)]
89. Filichkin, S.A.; Priest, H.D.; Givan, S.A.; Shen, R.; Bryant, D.W.; Fox, S.E.; Wong, W.K.; Mockler, T.C. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res.* **2010**, *20*, 45–58. [[CrossRef](#)]
90. Fesenko, I.; Khazigaleeva, R.; Kirov, I.; Kniazev, A.; Glushenko, O.; Babalyan, K.; Arapidi, G.; Shashkova, T.; Butenko, I.; Zgoda, V.; et al. Alternative splicing shapes transcriptome but not proteome diversity in *Physcomitrella patens*. *Sci. Rep.* **2017**, *7*, 1–14. [[CrossRef](#)]
91. Chaudhary, S.; Khokhar, W.; Jabre, I.; Reddy, A.S.N.; Byrne, L.J.; Wilson, C.M.; Syed, N.H. Alternative splicing and protein diversity: Plants versus animals. *Front. Plant Sci.* **2019**, *10*, 1–14. [[CrossRef](#)] [[PubMed](#)]
92. Yu, H.; Tian, C.; Yu, Y.; Jiao, Y. Transcriptome Survey of the Contribution of Alternative Splicing to Proteome Diversity in *Arabidopsis thaliana*. *Mol. Plant.* **2016**, *9*, 749–752. [[CrossRef](#)] [[PubMed](#)]
93. Melamud, E.; Moulton, J. Stochastic noise in splicing machinery. *Nucl. Acids Res.* **2009**, *37*, 4873–4886. [[CrossRef](#)] [[PubMed](#)]
94. Srivastava, A.K.; Lu, Y.; Zinta, G.; Lang, Z.; Zhu, J.K. UTR-dependent control of gene expression in plants. *Trends Plant Sci.* **2018**, *23*, 248–259. [[CrossRef](#)] [[PubMed](#)]
95. Kalyna, M.; Simpson, C.G.; Syed, N.H.; Lewandowska, D.; Marquez, Y.; Kusenda, B.; Marshall, J.; Fuller, J.; Cardle, L.; McNicol, J.; et al. Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucl. Acids Res.* **2012**, *40*, 2454–2469. [[CrossRef](#)] [[PubMed](#)]
96. Laxa, M.; Müller, K.; Lange, N.; Doering, L.; Pruscha, J.T.; Peterhänsel, C. The 5'UTR intron of arabidopsis GGT1 aminotransferase enhances promoter activity by recruiting RNA polymerase II. *Plant. Physiol.* **2016**, *172*, 313–327. [[CrossRef](#)]
97. Remy, E.; Cabrito, T.R.; Batista, R.A.; Hussein, M.A.M.; Teixeira, M.C.; Athanasiadis, A.; Sá-Correia, I.; Duque, P. Intron retention in the 5'UTR of the novel ZIF2 transporter enhances translation to promote zinc tolerance in Arabidopsis. *PLoS Genet.* **2014**, *10*, e1004375. [[CrossRef](#)]
98. Sathe, L.; Bolinger, C.; Amin-Ul Mannan, M.; Dever, T.E.; Dey, M. Evidence that base-pairing interaction between intron and mRNA leader sequences inhibits initiation of HAC1 mRNA translation in yeast. *J. Biol. Chem.* **2015**, *290*, 21821–21832. [[CrossRef](#)]
99. Von Arnim, A.G.; Jia, Q.; Vaughn, J.N. Regulation of plant translation by upstream open reading frames. *Plant. Sci.* **2014**, *214*, 1–12. [[CrossRef](#)]
100. Hsu, P.Y.; Calviello, L.; Wu, H.Y.L.; Li, F.W.; Rothfels, C.J.; Ohler, U.; Benfey, P.N. Super-resolution ribosome profiling reveals unannotated translation events in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E7126–E7135. [[CrossRef](#)]
101. Hinnebusch, A.G.; Ivanov, I.P.; Sonenberg, N. Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* **2016**, *352*, 1413–1416. [[CrossRef](#)] [[PubMed](#)]
102. Lim, C.S.; Wardell, S.J.T.; Kleffmann, T.; Brown, C.M. The exon-intron gene structure upstream of the initiation codon predicts translation efficiency. *Nucleic Acids Res.* **2018**, *46*, 4575–4591. [[CrossRef](#)] [[PubMed](#)]
103. Johnstone, T.G.; Bazzini, A.A.; Giraldez, A.J. Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J.* **2016**, *35*, 706–723. [[CrossRef](#)] [[PubMed](#)]



104. Ebina, I.; Takemoto-Tsutsumi, M.; Watanabe, S.; Koyama, H.; Endo, Y.; Kimata, K.; Igarashi, T.; Murakami, K.; Kudo, R.; Ohsumi, A.; et al. Identification of novel *Arabidopsis thaliana* upstream open reading frames that control expression of the main coding sequences in a peptide sequence-dependent manner. *Nucl. Acids Res.* **2015**, *43*, 1562–1576. [[CrossRef](#)]
105. Calvo, S.E.; Pagliarini, D.J.; Mootha, V.K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7507–7512. [[CrossRef](#)]
106. Noh, A.L.; Watanabe, S.; Takahashi, H.; Naito, S.; Onouchi, H. An upstream open reading frame represses expression of a tomato homologue of arabidopsis ANAC096, a NAC domain transcription factor gene, in a peptide sequence-dependent manner. *Plant Biotechnol.* **2015**, *32*, 157–163. [[CrossRef](#)]
107. Henkin, T.M. Riboswitch RNAs: Using RNA to sense cellular metabolism. *Genes Dev.* **2008**, *22*, 3383–3390. [[CrossRef](#)]
108. Sudarsan, N.; Barrick, J.E.; Breaker, R.R. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **2003**, *9*, 644–647. [[CrossRef](#)]
109. Bocobza, S.; Adato, A.; Mandel, T.; Shapira, M.; Nudler, E.; Aharoni, A. Riboswitch-dependent gene regulation and its evolution in the plant kingdom. *Genes Dev.* **2007**, *21*, 2874–2879. [[CrossRef](#)]
110. Cheah, M.T.; Wachter, A.; Sudarsan, N.; Breaker, R.R. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* **2007**, *447*, 497–500. [[CrossRef](#)]
111. Álvarez, D.; Voß, B.; Maass, D.; Wüst, F.; Schaub, P.; Beyer, P.; Welsch, R. Carotenogenesis is regulated by 5'UTR-mediated translation of phytoene synthase splice variants. *Plant. Physiol.* **2016**, *172*, 2314–2326. [[CrossRef](#)] [[PubMed](#)]
112. Grojean, J.; Downes, B. Riboswitches as hormone receptors: Hypothetical cytokinin-binding riboswitches in *Arabidopsis thaliana*. *Biol. Direct* **2010**, *5*, 60. [[CrossRef](#)] [[PubMed](#)]
113. Reddy, A.S.N. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant Biol.* **2007**, *58*, 267–294. [[CrossRef](#)] [[PubMed](#)]
114. Reddy, A.S.N.; Marquez, Y.; Kalyna, M.; Barta, A. Complexity of the alternative splicing landscape in plants. *Plant Cell* **2013**, *25*, 3657–3683. [[CrossRef](#)]
115. Crespo-Salvador, Ó.; Sánchez-Giménez, L.; López-Galiano, M.J.; Fernández-Crespo, E.; Schalschi, L.; García-Robles, I.; Rausell, C.; Real, M.D.; González-Bosch, C. The histone marks signature in exonic and intronic regions is relevant in early response of tomato genes to *Botrytis cinerea* and in miRNA regulation. *Plants* **2020**, *9*, 300. [[CrossRef](#)]
116. Egawa, C.; Kobayashi, F.; Ishibashi, M.; Nakamura, T.; Nakamura, C.; Takumi, S. Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat. *Genes Genet. Syst.* **2006**, *81*, 77–91. [[CrossRef](#)]
117. Chong, G.L.; Foo, M.H.; Lin, W.D.; Wong, M.M.; Verslues, P.E. Highly ABA-induced 1 (HAI1)-interacting protein HIN1 and drought acclimation-enhanced splicing efficiency at intron retention sites. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 22376–22385. [[CrossRef](#)]
118. Feng, J.; Li, J.; Gao, Z.; Lu, Y.; Yu, J.; Zheng, Q.; Yan, S.; Zhang, W.; He, H.; Ma, L.; et al. SKIP confers osmotic tolerance during salt stress by controlling alternative gene splicing in *Arabidopsis*. *Mol. Plant* **2015**, *8*, 1038–1052. [[CrossRef](#)]
119. Leviatan, N.; Alkan, N.; Leshkowitz, D.; Fluhr, R. Genome-wide survey of cold stress regulated alternative splicing in *Arabidopsis thaliana* with tiling microarray. *PLoS ONE* **2013**, *8*, e66511. [[CrossRef](#)]
120. Huertas, R.; Catalá, R.; Jiménez-Gómez, J.M.; Castellano, M.M.; Crevillén, P.; Piñeiro, M.; Jarillo, J.A.; Salinas, J. *Arabidopsis* SME1 regulates plant development and response to abiotic stress by determining spliceosome activity Specificity. *Plant. Cell* **2019**, *31*, 537–554. [[CrossRef](#)]
121. Kim, J.Y.; Ryu, J.Y.; Baek, K.; Park, C.M. High temperature attenuates the gravitropism of inflorescence stems by inducing SHOOT GRAVITROPISM 5 alternative splicing in *Arabidopsis*. *New Phytol.* **2016**, *209*, 265–279. [[CrossRef](#)] [[PubMed](#)]
122. Xin, R.; Kathare, P.K.; Huq, E. Coordinated regulation of pre-mRNA splicing by the SFPS-RRC1 complex to promote photomorphogenesis. *Plant Cell* **2019**, *31*, 2052–2069. [[CrossRef](#)] [[PubMed](#)]
123. Dong, J.; Chen, H.; Deng, X.W.; Irish, V.F.; Wei, N. Phytochrome B induces intron retention and translational inhibition of PHYTOCHROME-INTERACTING FACTOR3. *Plant. Physiol.* **2020**, *182*, 159–166. [[CrossRef](#)]

124. Ling, Y.; Serrano, N.; Gao, G.; Atia, M.; Mokhtar, M.; Woo, Y.H.; Bazin, J.; Veluchamy, A.; Benhamed, M.; Crespi, M.; et al. Thermoprimering triggers splicing memory in Arabidopsis. *J. Exp. Bot.* **2018**, *69*, 2659–2675. [[CrossRef](#)] [[PubMed](#)]
125. James, A.B.; Syed, N.H.; Bordage, S.; Marshall, J.; Nimmo, G.A.; Jenkins, G.I.; Herzyk, P.; Brown, J.W.S.; Nimmo, H.G. Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. *Plant. Cell* **2012**, *24*, 961–981. [[CrossRef](#)]
126. Yang, S.Y.; Lu, W.C.; Ko, S.S.; Sun, C.M.; Hung, J.C.; Chiou, T.J. Upstream open reading frame and phosphate-regulated expression of rice OsNLA1 controls phosphate transport and reproduction. *Plant Physiol.* **2020**, *182*, 393–407. [[CrossRef](#)]
127. Tsugeki, R.; Tanaka-Sato, N.; Maruyama, N.; Terada, S.; Kojima, M.; Sakakibara, H.; Okada, K. CLUMSY VEIN, the Arabidopsis DEAH-box Prp16 ortholog, is required for auxin-mediated development. *Plant J.* **2015**, *81*, 183–197. [[CrossRef](#)]
128. Petrillo, E.; Godoy Herz, M.A.; Fuchs, A.; Reifer, D.; Fuller, J.; Yanovsky, M.J.; Simpson, C.; Brown, J.W.S.; Barta, A.; Kalyna, M.; et al. A chloroplast retrograde signal regulates nuclear alternative splicing. *Science* **2014**, *344*, 427–430. [[CrossRef](#)]
129. Daras, G.; Rigas, S.; Alatzas, A.; Samiotaki, M.; Chatzopoulos, D.; Tsitsekian, D.; Papadaki, V.; Templalexis, D.; Banilas, G.; Athanasiadou, A.M.; et al. LEFKOTHEA Regulates Nuclear and Chloroplast mRNA Splicing in Plants. *Dev. Cell* **2019**, *50*, 767–779.e7. [[CrossRef](#)]
130. Wahl, M.C.; Will, C.L.; Lührmann, R. The spliceosome: Design principles of a dynamic RNP machine. *Cell* **2009**, *136*, 701–718. [[CrossRef](#)]
131. Brown, J.W.S.; Marshall, D.F.; Echeverria, M. Intronic noncoding RNAs and splicing. *Trends Plant Sci.* **2008**, *13*, 335–342. [[CrossRef](#)]
132. Rosado, A.; Li, R.; Van De Ven, W.; Hsu, E.; Raikhel, N.V. Arabidopsis ribosomal proteins control developmental programs through translational regulation of auxin response factors. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19537–19544. [[CrossRef](#)] [[PubMed](#)]
133. Cavallari, N.; Nibau, C.; Fuchs, A.; Dadarou, D.; Barta, A.; Doonan, J.H. The cyclin-dependent kinase G group defines a thermo-sensitive alternative splicing circuit modulating the expression of Arabidopsis ATU2AF65A. *Plant J.* **2018**, *94*, 1010–1022. [[CrossRef](#)] [[PubMed](#)]
134. Parenteau, J.; Abou Elela, S. Introns: Good Day Junk Is Bad Day Treasure. *Trends Genet.* **2019**, *35*, 923–934. [[CrossRef](#)] [[PubMed](#)]
135. Munding, E.M.; Shiue, L.; Katzman, S.; Donohue, J.P.; Ares, M. Competition between Pre-mRNAs for the splicing machinery drives global regulation of splicing. *Mol. Cell* **2013**, *51*, 338–348. [[CrossRef](#)] [[PubMed](#)]
136. Parenteau, J.; Maignon, L.; Berthoumieux, M.; Catala, M.; Gagnon, V.; Abou Elela, S. Introns are mediators of cell response to starvation. *Nature* **2019**, *565*, 612–617. [[CrossRef](#)] [[PubMed](#)]
137. Morgan, J.T.; Fink, G.R.; Bartel, D.P. Excised linear introns regulate growth in yeast. *Nature* **2019**, *565*, 606–611. [[CrossRef](#)]
138. Zhang, Y.; Zhang, X.O.; Chen, T.; Xiang, J.F.; Yin, Q.F.; Xing, Y.H.; Zhu, S.; Yang, L.; Chen, L.L. Circular intronic long noncoding RNAs. *Mol. Cell* **2013**, *51*, 792–806. [[CrossRef](#)] [[PubMed](#)]
139. Li, Z.; Wang, S.; Cheng, J.; Su, C.; Zhong, S.; Liu, Q.; Fang, Y.; Yu, Y.; Lv, H.; Zheng, Y.; et al. Intron lariat RNA inhibits microRNA biogenesis by sequestering the dicing complex in Arabidopsis. *PLoS Genet.* **2016**, *12*, e1006422. [[CrossRef](#)]

