Plant Physiology, May 1999, Vol. 120, pp. 65–72, www.plantphysiol.org © 1999 American Society of Plant Physiologists

The AG Dinucleotide Terminating Introns Is Important but Not Always Required for Pre-mRNA Splicing in the Maize Endosperm[1]

Shailesh Lal[2], Jae-Hyuk Choi[2], and L. Curtis Hannah*

Program in Plant Molecular and Cellular Biology and Horticultural Sciences, 1143 Fifield Hall, P.O. Box 110690, University of Florida, Gainesville, Florida 32611–0690

Previous RNA analysis of lesions within the 15 intron-containing *Sh2* (*shrunken2*) gene of maize (*Zea mays*) revealed that the majority of these mutants affect RNA splicing. Here we decipher further two of these mutants, *sh2-i* (*shrunken2 intermediate* phenotype) and *sh2-7460*. Each harbors a G-to-A transition in the terminal nucleotide of an intron, hence destroying the invariant AG found at the terminus of virtually all nuclear introns. Consequences of the mutations, however, differ dramatically. In *sh2-i* the mutant site is recognized as an authentic splice site in approximately 10% of the primary transcripts processed in the maize endosperm. The other transcripts exhibited exon skipping and lacked exon 3. A G-to-A transition in the terminus of an intron was also found in the mutant *sh2-7460*, in this case intron 12. The lesion activates a cryptic acceptor site downstream 22 bp within exon 13. In addition, approximately 50% of *sh2-7460* transcripts contain intron 2 and 3 sequences.

Although introns in nuclear genes are ubiquitous in nature, the signals required to define precisely and to recognize exon-intron borders are not fully understood. Studies from all eukaryotes suggest that splicing is essentially a two-step cleavage-ligation reaction. The first step involves the cleavage at the 5′-splice site that leads to the formation of an intron lariat with the adenosine residue of the branch point sequence located upstream of the 3′-splice site. This step is followed by the ligation of the exon and by the release of the intron lariat (Moore and Sharp, 1993; Brown, 1996; Simpson and Filipowicz, 1996). This complex set of events is carried out by pre-mRNA association with a conglomeration of small nRNAs and nuclear proteins that forms a dynamic, large ribonucleosome protein complex termed a spliceosome (for review, see Moore et al., 1993; Sharp, 1994). This fundamental process, common to all eukaryotic gene expression, can have a diverse impact on the regulation of gene expression. For example, imprecise or inaccurate pre-mRNA splicing often imparts a mutant phenotype (for review, see Weil and Wessler, 1991), whereas alternative splicing is sometimes important in the regulation of gene expression (Gorlach et al., 1995; Golovkin and Reddy, 1996; Nishihama et al., 1997). Certain introns dramatically enhance gene expression in transient and stably transformed callus tissue (Callis et al., 1987; Vasil et al., 1989; Clancy et al., 1994). Finally, intron splicing is required for some plant viruses to be pathogenic (Kiss-Laszlo et al., 1992).

Unlike yeast and vertebrates, the lack of a plant in vitro system capable of efficiently splicing introns has hindered our understanding of the mechanism of splicing in plants. Despite the universal nature of the splicing pathway, primary transcripts of animal origin are not efficiently or accurately spliced in plant cells. Conversely, the majority of plant pre-mRNAs are not faithfully spliced in animal cells (Barta et al., 1986; van Santen and Spritz, 1987; Wiebauer et al., 1988; Pautot et al., 1989; Waigmann and Barta, 1992). This species barrier between the heterologous splicing of pre-mRNA is also observed between monocots and dicots. Some monocot introns are not spliced in dicots (Keith and Chua, 1986; Goodall and Filipowicz, 1991). In contrast, introns of dicot origin are efficiently spliced in monocots, suggesting that monocot-splicing machinery is more flexible or complex. There are also fundamental structural/sequence differences that differentiate plant introns from those of vertebrate and yeast introns (Goodall and Filipowicz, 1991; for review, see Simpson and Filipowicz, 1996). Vertebrate introns possess a polypyrimidine track that is not found in plant introns (Simpson et al., 1996).

Branch points have recently been mapped in plants, the consensus of which is similar to that of the vertebrates (Liu and Filipowicz, 1996; Simpson et al., 1996; for review, see Brown, 1998). A feature distinguishing plant introns from those of other organisms is their AU richness. This has been indicated to be essential for intron processing and for definition of the intron/exon junction (Goodall and Filipowicz, 1989; Lou et al., 1993; McCullough et al., 1993; Carle-Urioste et al., 1994; Luehrsen and Walbot, 1994; Gniadkowski et al., 1996). It is interesting that the requirements of the AU-rich region are more stringent in dicots

[1] This research was supported in part by the National Science Foundation (grant nos. IBN-9316887 and MCB-9420422) and by the U.S. Department of Agriculture Competitive Grants Program (grant nos. 94-37300-453, 97-36306-4461, 95-37301-2080, and 98-01006). This is Florida Agricultural Experiment Station journal series no. R-06349.

[2] These authors contributed equally to the paper.

* Corresponding author; e-mail hannah@gnv.ifas.ufl; fax 1–352–392–5653.

Abbreviations: AGP, ADP-Glc pyrophosphorylase; RT-PCR, reverse transcription-PCR.

than in monocots (Goodall and Filipowicz, 1991), and some monocot introns are GC rich.

Here we describe two mutants of the *sh2* (*shrunken2*) gene of maize (*Zea mays*) that affect RNA splicing. *Sh2* encodes the large subunit of a heterotetrameric enzyme, AGP, a key regulatory enzyme in the starch biosynthetic pathway. Null *sh2* mutants cause the dramatic reduction of the starch content in maize kernels that leads to a collapsed, brittle, or shrunken phenotype (Tsai and Nelson, 1966). Previous analysis of the various *sh2* mutants revealed that each of these mutants produced multiple transcripts or transcripts of a non-wild-type size. Because *Sh2* contains 15 introns, northern analyses suggest that the primary lesion of many mutants occurs at RNA processing (Giroux and Hannah, 1994). To characterize further the molecular events underlying the mutational lesion of two of these mutants, *sh2-i* (*shrunken2 intermediate* phenotype) and *sh2-7460*, mutant transcripts and genomic DNA were isolated, cloned, sequenced, and expressed in *Escherichia coli* and in maize tissue culture cells.

To our knowledge, the results of these studies reveal a number of features of pre-RNA processing not yet reported in plants. Namely, we show that, although the AG dinucleotide terminating a nuclear intron is important for splicing, it is not always essential. We note exon skipping, a phenomenon commonly found in vertebrates. We also report a case of improper splicing of two adjacent introns. These latter two phenomena are hallmarks of the exon definition of pre-RNA splicing.

MATERIALS AND METHODS

Plant Materials

The maize (*Zea mays*) mutants *sh2-i* and *sh2-7460* were kindly provided by Dr. M.G. Nueffer and Dr. Oliver Nelson and have been described elsewhere (Hannah et al., 1980; Giroux and Hannah, 1994). The wild-type *Sh2* allele used here was isolated from McClintock's *a1-m3* stock (Giroux et al., 1996). Plants were maintained and grown in the greenhouse or in the field at the University of Florida, Gainesville.

RNA and Genomic DNA Isolation

Total RNA from the 20- to 22-d postpollination kernels and leaf genomic DNA were isolated as described previously (McCarty, 1986; Giroux et al., 1994). RNA from suspension cells was extracted using TRIZOL reagent (Life Technologies) according to the protocol of the manufacturer. Northern analysis was performed as described by Maniatis et al. (1993). Band intensities were measured by a digital imaging system (model IS-1000, Alpha Innotech Corp., San Leandro, CA).

PCR Amplification and Cloning

RT-PCR was used to synthesize the full-length cDNA clones of the *sh2*-i and *sh2*-7460 alleles by using a reverse transcriptase kit (Superscript, Life Technologies). First-strand cDNA synthesis was primed with oligo(dT) primer from developing maize endosperm total RNA, and full-length clones were isolated by use of the primers SH2FO.1 (5′-CAAGATCACGTCGACAGGCAAGTG-3′) and SH2.3 (5′-GGTTTGCTGCAGC TTCTAGGGC-3′). These are complementary to the 5′- and 3′-nontranslated regions of *Sh2* cDNA, respectively. Underlined sequences contain modified bases to incorporate restriction sites for *Sal*I and *Pst*I. The restriction sites were used to subsequently clone the amplified fragment into the corresponding restriction site of vector pBluescript KS+.

To amplify cDNA-spanning exons 1 to 4 from *sh2-i*, primers SH2F0.1 and SH2R0.1 (5′-GCCTGTAACATCCTC-CTGCAGGT-3′) were used. The resulting products were separated on a 1% agarose gel and alkaline transferred onto a membrane (Hybond H^+, Amersham), according to the protocol provided by the manufacturer. This blot was first probed with an exon 3-specific probe, according to the procedure described by Church and Gilbert (1984). The exon 3-specific probe was generated by PCR amplification of *Sh2* cDNA using primers SH2LHS1 (5′-CATTCT-CAAACACAGTCGACTAG-3′) and SH2LHSR (5′-AG-CAGGCGCAGCTCTAG-3′). The blot was then washed twice with 2× SSPE/0.1%SDS and 0.2× SSPE/0.1%SDS at 65°C for 20 min and then subjected to overnight exposure to x-ray film to monitor the efficacy of probe removal. It was then probed with full-length *Sh2* cDNA. Resulting autoradiograms were quantified by a phosphor imager (Molecular Dynamics, Sunnyvale, CA) or with a digital imaging system (model IS-1000).

Genomic sequences from exons 1 to 4 were amplified using primers SH2FO.1 and SH2RO.1 (5′-GCCTGTAA-CATCCTCCTGCAGGT-3′) using leaf DNA as the template. Similarly, exons 7 to 14 of mutant *sh2-7460* were amplified using primers, SHLH872 (5′-ACATGTCGAC-GATGCTGCTG CTATC-3′) and SHLH871R (5′-GAGT-TCACCTGCAGA GCTGAC-3′). Resulting fragments were cloned into pBluescript KS+ or pUC19. DNA sequencing was done at the University of Florida DNA Sequencing Core Laboratory (Gainesville), using the ABI Prism Dye Terminator sequencing protocol (Applied Biosystems). DNA sequences derived from the mutants were compared with the *Sh2* sequence (Shaw and Hannah, 1992) using DNA analysis software (Lasergene, DNASTAR, Madison, WI).

Bacterial Expression of AGP

The *Escherichia coli* expression (Iglesias et al., 1993) was used to monitor *sh2-i* and *sh2-7460* transcripts for functional AGP activity. The full-length *Sh2* transcript from wild type, the smaller transcript of mutant *sh2-i*, and the wild-type-size transcript of *sh2-7460* were RT-PCR amplified using primers LH377 (5′-GGGGCCATGGCCCAG TT-TGCACTTGCATTGGACGACA CG-3′) and LH396 (5′- CCC-CGAGCTCACTATATGACAGACCCATCGTTGATGG) as described earlier and cloned into the *Nco*I and *Sst*I restriction sites of bacterial expression vector pMSH (Iglesias et al., 1993). Resulting clones are termed pSHW, pMSHi, and pMSH7460, respectively. The low abundance of the larger

sh2-i transcript and its close proximity after electrophoresis to the more abundant, smaller transcript made it difficult to effectively resolve the DNA bands. Hence, the region from exon 1 to exon 7 of *sh2-i* transcripts was amplified using primers LH377 and SH796R (5′-CTCTCATCAACA GGAGCA C-3′). This allowed for a definitive resolution of fragments of approximately 600 and 700 bp on the 1% agarose gel. The larger fragment of approximately 700 bp of mutant *sh2-i* was eluted from the gel, digested with *Nco*I and *Xho*I, and cloned into the corresponding site of pMSHi replacing the corresponding sequence in the smaller transcript, giving rise to pMSHWi. Constructs were separately transformed into an *E. coli* strain AC70R-504, which lacks endogenous AGP activity but harbors the *Bt2* gene on the compatible vector described by Giroux et al. (1996). Transformed cells were grown for 16 h on Luria-Bertani-medium plates containing 1% Glc and then iodine stained to monitor AGP activity as described earlier (Iglesias et al., 1993; Greene et al., 1996).

Particle Bombardment and Expression Vectors

Maize cell line PC5, established from the mesocotyl tissue of germinating seed (Chourey and Zurawski, 1981), was cultured in a liquid Murashige and Skoog medium supplemented with 2 mg/L 2,4-D. Cells were grown in the dark at 27°C on a shaker at 150 rpm and were routinely subcultured at 7-d intervals. After 3 d of subculture, approximately 2 mL of cells was transferred onto a filter disc (Whatman) for particle bombardment. The disc was placed on a Petri plate containing Murashige and Skoog-agarose medium and used immediately as a target for the particle bombardment. Preparation of the DNA/gold mixture and the parameters for bombardment were previously described (Taylor and Vasil, 1991). The Biolistic Particle Delivery System (model PDS-1000/HE, Bio-Rad) was used for bombardment. Cells were harvested 22 h postbombardment, frozen in liquid N_2, and stored at −70°C for further analysis.

Exons 2 to 4 were isolated from 1 μg of leaf genomic DNA from the wild type and *sh2-i* by 30 cycles of PCR amplification using primers SH2F0.1 and SH2R0.1. Resulting 1.6-kb fragments were blunt ended with *Pfu* polymerase (New England Biolabs) and ligated into the solitary *Eco*RV site present in the luciferase-coding region of the plant expression vector pAHC18, kindly provided by Dr. Peter Quail (Christensen and Quail, 1996). Resulting constructs were expressed in maize suspension cells as described earlier. Total RNA extracted from the cells was treated with amplification grade DNase I (BRL) and subjected to RT-PCR using primers LucLo.Pr2 (5′-CCCGGTTTTAATGAATACGT-3′) and LucUp.Pr2 (5′-CCGTGCTCCAAAACAACAA-3′), which flanked the insertion. The resulting PCR fragments were blotted and probed with the luciferase-coding region of pAHC18 as described earlier. The fragments were eluted from the gel and directly sequenced.

RESULTS

Origin of AGP Activity in the Mutant *sh2-i*

The mutant *sh2-i*, which was generously supplied by Dr. M.G. Neuffer (University of Missouri, Columbia), was generated by ethyl methanesulfonate mutagenesis and conditions an intermediate or leaky phenotype in comparison to virtually all other *sh2* mutant alleles. Giroux and Hannah (1994) showed that this mutant produced a major transcript and protein smaller than those found in the wild type. To determine whether the detected diminutive SH2 protein conditioned AGP activity in *sh2-i*, the small transcript of *sh2-i* was cloned. Sequencing revealed that this transcript lacked exon 3. Exon 3 is a multiple of 3 (123 bp) in length; therefore, deletion of this exon maintains translational continuity. Conservation of the reading frame combined with the fact that initiation of translation begins in exon 2 (Shaw and Hannah, 1992) caused us to consider the possibility that this mutant SH2 protein conditioned the partial AGP activity of *sh2-i*.

Accordingly, this transcript was cloned and coexpressed with the wild-type *Bt2* (*Brittle2*) gene in an *E. coli* mutant lacking the endogenous bacterial AGP, glg-C gene (Igleasias et al., 1993). Expression of wild-type *Sh2* and *Bt2* genes (Fig. 1) complemented the *E. coli* mutant giving rise to glycogen, which can be visualized easily by exposure to iodine vapors. In contrast, expression of the abbreviated *sh2-i* transcript (Fig. 1, pMSH-i/pMBT-W) did not complement the glg-C mutant. Because mutants with less than 3% wild-type activity give rise to some staining (S. Lal, J.-H. Choi, and L.C. Hannah, unpublished data), we have concluded that this *sh2-i* transcript probably does not underlie the leaky kernel phenotype. Furthermore, the lack of enzymatic activity encoded by the small transcript has been predicted from the recent findings that the completely

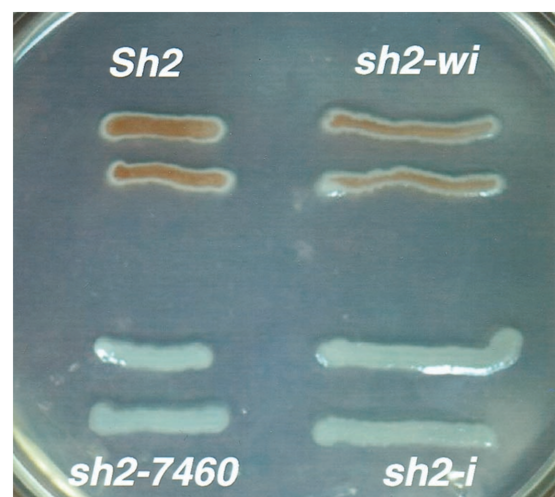

Figure 1. Expression of maize endosperm AGP in *E. coli*. The full-length wild-type transcript *Sh2*, the wild-type-size transcript of mutants *sh2-i* (*sh2-iw*) and *sh2-7460*, and the smaller transcript of *sh2-i* were expressed with *Bt2* in an *E. coli* strain deficient in the endogenous AGP gene. Complementation of the glg-C *E. coli* mutant resulting in glycogen synthesis was visualized by iodine staining.

conserved, exon 3-encoded motif PAV is critical for enzymatic activity in potato (Greene et al., 1996) and in *E. coli* (Meyer et al., 1993). We have concluded from these experiments that the truncated *sh2-i* transcript does not encode the AGP activity encountered in this mutant and that an additional source of AGP must condition the leaky phenotype of this mutant.

Therefore, we analyzed whether the *sh2-i* primary transcript undergoes multiple splicing events possibly giving rise to other less-abundant, mature transcripts containing exon 3. RT-PCR analysis was performed on oligo(dT)-primed first-strand cDNA from wild-type and *sh2-i* developing endosperm poly(A^+) RNA using primers flanking exons 1 to 4. A less-abundant, wild-type-size PCR product from *sh2-i* (Fig. 2A, left, lane 2) was observed. This fragment hybridized to a full-length *Sh2* cDNA probe (Fig. 2A, center, lane 2) and to a probe specific to exon 3 (Fig. 2A, right, lane 2) As judged by digital imaging, this fragment makes up approximately 10% of the total *Sh2* transcript. Furthermore, quantitative analysis of RNA blots probed with exon 3 suggests that this transcript makes up 5% to 10% of the total transcript.

Full-length clones making up the larger transcript of *sh2-i* were isolated by PCR, sequenced, and expressed in *E. coli.* Sequencing showed that this transcript contains a completely wild-type exon 3 that perfectly abuts exons 2 and 4

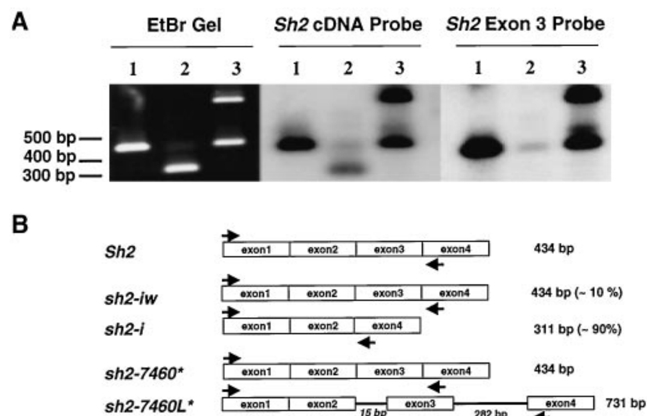

Figure 2. RT-PCR analysis of mutant *sh2-i* and *sh2-7460* transcripts. A, Total endosperm RNA from the wild type and that from the mutants *sh2-i* and *sh2-7460* were subjected to RT-PCR using primers spanning exons 1 to 4 of the *Sh2* transcript. The left panel depicts the resultant RT-PCR products resolved on a 1% agarose gel and stained with ethidium bromide. Lanes 1, 2, and 3 represent the RT-PCR product derived from *Sh2* and the mutants *sh2-i* and *sh2-7460*, respectively. The center and right panels are DNA blots of the agarose gel probed with radiolabeled full-length *Sh2* cDNA and *Sh2* exon 3-specific probes, respectively. B, Schematic structure of the RT-PCR products derived from the mutants *sh2-i* and *sh2-7460* and the wild-type *Sh2* transcripts, as shown in A. Transcript sources are labeled at the left side of the panel; the numbers on the right give the size and the relative proportion (in percentages) of each transcript in the mutant endosperm. Arrows represent primers used for PCR amplification. Exons are represented by boxes, and the intronic sequences that remained in the mutant *sh2-7460* transcript are marked by a thick line. Asterisks indicate that these transcripts contain a 22-bp deletion of a downstream exon 13 sequence.

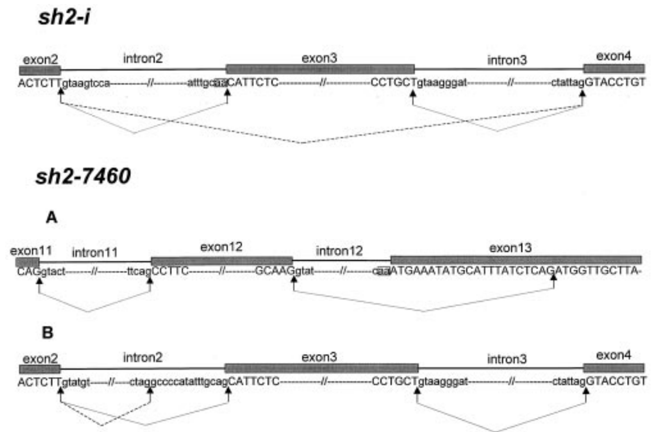

Figure 3. Schematic representation of the genomic sequence bearing the splice-site alterations of mutants *sh2-i* (exons 2–4) and *sh2-7460* (exon 11–13 [A]) and (exon 2–4 [B]). The point mutations that altered the 3′-splice site A*G* to A*A* of intron 2 in mutant *sh2-i* and intron 12 of mutant *sh2-7460* are boxed. Arrows joined by lines mark the donor and acceptor sites used during RNA splicing to generate the mutant transcripts.

(Fig. 2B). Expression of this transcript leads to a functional AGP (Fig. 1, top right). We have concluded that the low levels of AGP activity found in *sh2-i* arise from this less-abundant transcript.

Because two transcripts arise from one mutant gene, the *sh2-i* pre-RNA must undergo multiple splicing events. Primers spanning exons 1 to 4 were used to isolate *sh2-i* genomic DNA. Sequencing revealed that *sh2-i* harbors a G-to-A transition at the terminus of intron 2 (Fig. 3). Hence, loss of a wild-type splice site at the terminus of intron 2 leads to the skipping of exon 3 in approximately 90% of the processed transcripts. In these cases, splicing occurs between the donor site of intron 2 and the acceptor site of intron 3. In approximately 10% of the transcripts, however, the intron 2 acceptor site functions in splicing, although it lacks the invariant AG terminus.

Whereas the mutated 3′-splice site of intron 2 functions as an acceptor site in the maize endosperm, this variant splice site lacks function when this portion of *Sh2* is expressed in cultured maize cells. Exons 2 to 4 of the wild type and *sh2-i* were cloned into the lone *Eco*RV site within the luciferase-coding region of the plant expression vector AHC18 (Christensen and Quail, 1996). These recombinant constructs and AHC18 were separately introduced into the maize suspension cell line PC5 via particle bombardment. Total RNA extracted from these cells was analyzed by RT-PCR using primers flanking the adjacent luciferase-coding region of AHC18. Resulting DNA was electrophoresed, blotted, and probed with the luciferase-coding region (Fig. 4). A single fragment of 981 bp was amplified from cells expressing the wild-type *Sh2* gene. Sequencing revealed that this fragment lacked introns 2 and 3, which was expected if *Sh2* pre-RNA splice signals are recognized in the cultured cells as they are in the maize endosperm. In contrast, the *sh2-i*-containing construct produced a single, highly abundant transcript of 858 bp. From sequencing, this transcript lacked exon 3. Therefore, whereas the AA

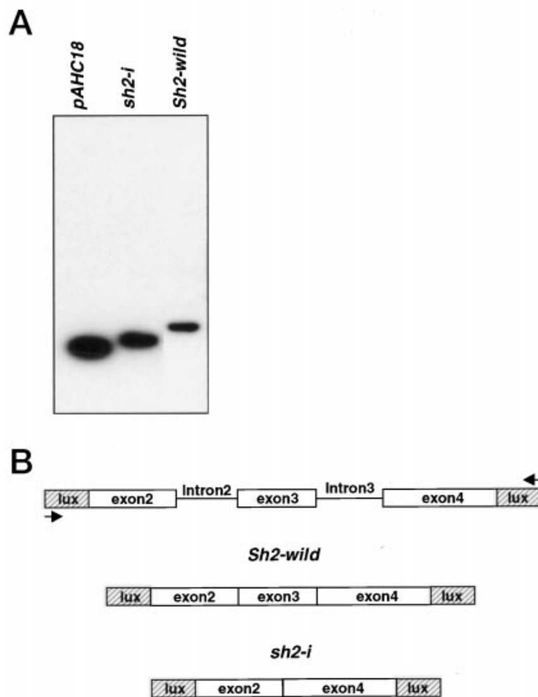

Figure 4. Expression of *Sh2* transcripts in maize suspension cells. A, Southern blot of RT-PCR products amplified using primers specific to the luciferase-coding region. These products were detected with the luciferase-coding region after expression of exons 2 to 4 from *Sh2* and the mutant *sh2-i* cloned into the luciferase-coding region of the expression vector AHC18 and were transiently expressed in maize suspension cells using particle bombardment. B, Top, Represents the genomic construct cloned into the luciferase (lux)-coding region of the vector AHC18. *Sh2* exons are marked by blank boxes, whereas introns are depicted by a line. The hatched boxes represent the flanking luciferase-coding sequences. Arrows represent the primers used in the RT-PCR reaction. The construct source is labeled at the top of each panel. Structures of the various constructs were determined by cloning and sequencing.

terminus of intron 2 is recognized in the maize endosperm, we find no evidence that this can also function in this genic context in maize tissue culture cells.

The Mutant *sh2-7460* also Contains a G-to-A Mutation at an Intron Acceptor Site

Giroux and Hannah (1994) showed that the mutation *sh2-7460* produces two transcripts, pointing to an alteration in RNA splicing. RT-PCR was used to isolate full-length clones from the two transcripts. Sequencing of these transcripts identified three alterations in RNA splicing. Both transcripts lacked the first 22 bp of exon 13. In addition, the larger transcript contained 15 bp of the distal portion of intron 2 and all of intron 3 (Fig. 2A, lanes 3), which lead to a shift in the reading frame and a proximal chain-termination codon. From sequencing, we concluded that a premature chain-termination mutation was created in exon 13 of the wild-type-size transcript and, as expected, that this transcript does not condition AGP activity (Fig. 1, lower left).

To precisely define the mutation underlying the complex splicing pattern exhibited by *sh2-7460*, genomic DNA corresponding to the altered spliced sites was amplified and cloned. A G-to-A transition at the terminus of intron 12 was detected, which provided a ready explanation for the loss of function of the intron 12 acceptor site. Here an AG in the AU-rich proximal region of exon 13 was activated and used as an acceptor site. Unlike the situation with *sh2-i,* we found no evidence for splicing of the mutant A-terminating intron.

DISCUSSION

Wild-Type Transcripts Arise from Splicing of an AA-Terminating Intron Giving Rise to the Leaky or Non-Null Phenotype of the Maize Mutant *sh2-i*

Here we elucidated the mechanism underlying the residual activity in the mutant *sh2-i*. The phenotype of mutant *sh2-i* is quite leaky, and Giroux and Hannah (1994) speculated that a predominant but truncated protein observed in this mutant might have partial enzymatic activity giving rise to the leaky phenotype. Here we cloned the cognate transcript and showed that the truncated protein is nonfunctional in an *E. coli* expression system. In an attempt to elucidate the cause of the residual activity, we found that, although this mutant has suffered a G-to-A transition in the terminus of intron 2, the mutant splice site can still be recognized and properly spliced, albeit in only approximately 10% of the primary transcripts. The low level of this wild-type transcript causes the leaky phenotype of *sh2-i*. Giroux et al. (1994) noted that AGP activity, the product of the *Sh2* locus, from *sh2-i* is approximately 20% of the wild type, a value comparable with that found for the wild-type transcript from this mutant.

The presence of the wild-type *Sh2* transcript in the mutant *sh2-i* indicates that at low, physiologically significant levels, the 3′-splice site of intron 2 is used although it lacks the virtually invariant AG-terminating dinucleotide. To our knowledge, this is the first report of a conventional nuclear intron now terminating in AA being spliced in a plant system. This splicing does not reflect a unique feature of *Sh2* RNA splicing, because an identical change in a latter intron in the *sh2-7460* mutation abolishes splicing at that position. Hence, this splicing is context dependent. Context dependency or tissue dependency is also evident, because we cannot detect intron 2 splicing of *sh2-i* when exons 2 to 4 of this mutant are expressed in a maize tissue culture system.

This observation is intriguing because the terminal AG of nuclear introns is considered invariant in many diverse species, including plants. This dinucleotide is likely associated with the binding of U5 small nuclear ribonucleoprotein particles, one of the major subunits of spliceosomes, and represents an essential step in RNA processing (Brown, 1996). Furthermore, the terminal G is important for the second step of splicing in yeast (Parker and Siliciano, 1993). Mutations altering the terminal AG abolish splicing and sometimes activate nearby, downstream cryptic acceptor sites (Kiss-Laszlo et al., 1992; Lou et al., 1993; Parker and

Siliciano, 1993; Carle-Urioste et al., 1994; Simpson and Filipowicz, 1996; Simpson et al., 1996). Several cases of a G-to-A transition at the 3′-splice site have been reported in Arabidopsis floral homeotic genes *ag-1*, *ag-4*, and *ap1-1* (Yanofsky et al., 1990; Mandel et al., 1992; Sieburth et al., 1995); the *spy-2* gene associated with GA signal transduction (Jacobsen et al., 1996); the *tt4*(*2YY6*) gene encoding chalcone synthase (Burbulis et al., 1996); the *stm-3* homeotic gene essential for meristem formation during embryogenesis (Long et al., 1996); and the *cop1-1* essential regulatory gene for the photomorphogenic development of the plant (McNellis et al., 1994). These mutations completely abolish the recognition of the acceptor site and result in inclusion of the intron in the processed transcript (McNellis et al., 1994), in exon skipping (Sieburth et al., 1995; Jacobsen et al., 1996), or in activation of a cryptic acceptor splice site in the adjacent exon (Sieberth et al., 1995; Burbulis et al., 1996).

Alterations in splicing caused by the intron terminal G-to-A transition have been extensively documented in vertebrates (Krawczak et al., 1992). These alterations are the basis for Citrullinemia (Dunn et al., 1989), Tay Sachs disease (Arpaia et al., 1988), and Fabry disease (Okumiya et al., 1995).

Whereas *sh2-i* splicing is the first report, to our knowledge, of this type of splicing in a plant system, the use of the dinucleotide AA in the acceptor site has been noted in other systems. Two cases were reported in *Caenorhabditis elegans* (Aroian et al., 1993) and one was reported in swine (Brown et al., 1994). The 3′-splice-site mutation AG to AA functions at 2% of the wild-type frequency in yeast (Parker and Siliciano, 1993). Recently, McCullough et al. (1996) expressed a β-conglicinin intron construct in tobacco nuclei and noted the use of a 3′ AA splice site. However, this splicing event also involved use of the nonconical 5′-splice site, AU.

The G-to-A Transitions of the Terminal Nucleotides of Introns 2 and 12 Have Drastically Different Effects on Pre-mRNA Splicing

Mutant *sh2-7460* contains an AG-to-AA alteration of the 3′ terminus of intron 12. Therefore, this lesion is identical to that of *sh2-i*, except that it resides in a different intron. In contrast to *sh2-i*, this mutant site is not functional. Rather, a cryptic 3′-splice site 22 bp downstream within exon 13 is used as an acceptor. This divergence in consequence on splicing must reflect a difference in sequence, context, or position within the gene. These types of mutations in which the splicing apparatus uses a site found in an exon as an acceptor site have been amply documented in plant literature. We note that the 22 bp of exon 13 now included in the intron of *sh2-7460* are 73% AU rich. Contrast between GC and AT richness has been indicated in splice-site selection (Goodall and Filipowicz, 1989; Lou et al., 1993; McCullough et al., 1993; Simpson and Brown, 1993; Luehrsen and Walbot, 1994) that may aid in the selection of the cryptic splice site.

Mutant *sh2-7460* Depicts a Complex Pattern of Pre-mRNA Splicing

The larger transcript of *sh2-7460* exhibits abnormalities in the splicing of exons 2, 3, and 13. The former change involves enclosure of the entire 282-bp intron 3 and 15 bp of the distal portion of intron 2. Hence, the authentic 3′-splice site of intron 2 is skipped in approximately 50% of the mature transcripts. Coupled with this, the splicing of intron 3 does not occur.

Plant genes are rich in introns and their accurate recognition and removal from the primary transcript is a complex process that may involve sequences distant from the splice site, as might be the case here. Mutations acting from a distance to affect splicing have been found in mammalian systems and in Arabidopsis (McNellis et al., 1994). An insertion in intron 11 of the maize mutant *sh2-7527* results in the inclusion of intron 9 in approximately 40% of the processed transcripts. This mutant is identical in sequence to wild-type *Sh2* from exons 6 to 11 (S. Lal and L.C. Hannah, unpublished data; Giroux and Hannah, 1994). There may be a global, three-dimensional nature of splicing, in which the splicing of various introns in context to a particular gene is interrelated. Whether the downstream mutation in intron 12 in *sh2-7460* affects the splicing of upstream exons is being investigated.

These Mutations Provide Evidence for Both the Intron Definition and the Exon Definition of Pre-RNA Splicing

Two theories account for the initial recognition of introns. In exon definition of splicing, recognition of splice sites on each end of a single exon occurs initially and is coordinated (Berget, 1995). Mutations that alter recognition at one end of the exon also alter recognition at the other end of the same exon. In other words, alteration in binding at the 5′-donor site also alters use of the upstream 3′-acceptor site abutting the same exon. Recently, the evidence for exon definition in plants was reported (Yi and Jack, 1998). In contrast, in the intron definition of splicing, the initial recognition is in the intron, probably at the branch point for lariat formation. This is followed by scanning in the 3′ direction for detection of the first AG dinucleotide. This AG then defines the acceptor site (for summary, see Luehrsen and Walbot [1994]).

The exon definition of splicing is attractive in animal systems where introns can be quite large compared with the adjoining exons. This model circumvents the need to scan the very large animal introns for splice sites. This model predicts that mutations abolishing splicing at one terminus of an exon also block splicing at the other terminus of the same exon. The exon consequently is not recognized and is removed during splicing, a phenomenon termed exon skipping. In support of this model, 51% of more than 100 splice mutations in humans exhibited exon skipping (Berget, 1995). Although plant exon sequences influence pre-mRNA splicing (Carle-Urioste et al., 1994, 1997), little evidence favors the exon definition of splicing in plants (for review, see Simpson and Filipowicz, 1996).

The vast majority of the evidence from splice-site mutants in plants favors the intron definition of splicing. In this regard, the splicing pattern of exon 13 in *sh2-7460* is easily explained by the intron definition of splicing. The use of the first AG nucleotide downstream of the mutated site in *sh2-7460* is consistent with the scanning model, beginning in the intron and proceeding 3′ until the first AG is encountered.

In contrast, the alteration found in *sh2-i* favors the exon definition of splicing. Exon skipping, the hallmark of the exon definition of splicing, occurs in 90% of the *sh2-i* endosperm-splicing events and all of those produced in maize tissue culture cells. Blockage of splicing at the 3′ end of intron 2 apparently blocks splicing at the 5′ end of intron 3 and results in total exclusion of exon 3 from the processed transcript. Furthermore, the aberrations in the splicing of the two adjacent introns (introns 2 and 3) in *sh2-7460* endosperms are in accord with the exon definition of splicing.

In summary, splicing of introns 2 and 3 of *Sh2* transcripts seemingly follows the exon definition of splicing, whereas splicing of intron 12 of the same transcript is apparently in accord with the intron definition of splicing. One possible explanation for this paradox is that the critical factor in determining the mode of recognition is the distance between adjacent splice sites. Those sequences, whether exon or intron, separated by the shortest distance between splice sites determine whether the exon or the intron definition, respectively, is followed in splicing. We note that exon 3 of *Sh2* is short (123 nucleotides) relative to the flanking introns (476 and 284 nucleotides), whereas intron 12 is short (70 nucleotides) compared with the adjacent exons (87 and 105 nucleotides). Berget (1995) similarly noted this possibility in reference to observations that increases in exon length in genes with short exons and long introns or expanding introns in genes with short introns and long exons led to aberrant splicing.

ACKNOWLEDGMENTS

We thank Dr. Gerry Neuffer for *sh2-i*, Dr. Oliver Nelson for *sh2-7460*, Dr. Peter Quail for the plant expression vector, and Dr. Prem Chourey for the maize tissue culture cells. Synthesis of DNA primers and DNA sequencing were done in the facilities of the Interdisciplinary Center for Biotechnology Research at the University of Florida, Gainesville.

Received November 16, 1998; accepted January 25, 1999.

LITERATURE CITED

Aroian RV, Levy AD, Kago M, Oshima Y, Kramer JM, Sternberg PW (1993) Splicing in *Caenorhabditis elegans* does not require an AG at the 3′ splice acceptor site. Mol Cell Biol **13:** 626–637

Arpaia E, Dumbrille-Ross A, Maler T, Neote K, Tropak M, Stirling JL, Pitts JS, Bapat B, Lamhonwah AM (1988) Identification of an altered splice site in Ashkenazi Tay-Sachs disease. Nature **333:** 85–86

Barta A, Sommergruber K, Thompson D, Harmuth K, Matzke ME, Matzke AJM (1986) The expression of nopaline synthase-human growth hormone chimeric gene in transformed tobacco and sunflower callus tissue. Plant Mol Biol **6:** 347–357

Berget SM (1995) Exon recognition in vertebrate splicing. J Biol Chem **270:** 2411–2414

Brown JWS (1996) Arabidopsis intron mutations and pre-mRNA splicing. Plant J **10:** 771–780

Brown JWS (1998) Splice site selection in plant pre-mRNA splicing. Annu Rev Plant Physiol Plant Mol Biol **49:** 77–95

Brown WR, Kacskovics I, Amendt BA, Blackmore NB, Rothschild M, Shinde R, Butler JE (1994) The hinge deletion allelic variant of porcine IgA results from a mutation at the splice acceptor site in the first Cα intron. J Immunol **154:** 3836–3842

Burbulis IA, Lacobucci M, Shirley BW (1996) A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in Arabidopsis. Plant Cell **8:** 1013–1025

Callis J, Fromm M, Walbot V (1987) Introns increase gene expression in cultured maize cells. Genes Dev **1:** 1183–1200

Carle-Urioste JC, Brendel V, Walbot V (1997) A combinational role for exon, intron and splice site sequences in splicing in maize. Plant J **11:** 1253–1263

Carle-Urioste JC, Ko CH, Bonito M, Walbot V (1994) *In vivo* analysis of intron processing using splicing-dependent reporter gene assays. Plant Mol Biol **26:** 1785–1795

Chourey PS, Zurawski D (1981) Callus formation from protoplasts of a maize cell culture. Theor Appl Genet **59:** 341–344

Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res **5:** 213–218

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA **81:** 1991–1995

Clancy M, Vasil V, Hannah LC, Vasil IK (1994) Maize shrunken1 intron and exon regions increase gene expression in maize protoplasts. Plant Sci **98:** 151–161

Dunn JM, Phillips RA, Zhu X, Becker A, Gallie BL (1989) Mutations in the RB1 gene and their affects on transcription. Mol Cell Biol **9:** 4596–4604

Giroux MJ, Boyer C, Feix G, Hannah LC (1994) Coordinated transcriptional regulation of storage product genes in the maize endosperm. Plant Physiol **106:** 713–722

Giroux MJ, Hannah LC (1994) ADP-glucose pyrophosphorylase in *shrunken2* and *brittle2* mutants of maize. Mol Gen Genet **243:** 400–408

Giroux MJ, Shaw GB, Cobb BG, Greene T, Okita T, Hannah LC (1996) A single gene mutation increases maize seed weight. Proc Natl Acad Sci USA **93:** 5824–5829

Gniadkowski M, Hammings-Mieszczak M, Klahre U, Liu H, Filipowicz W (1996) Characterization of intronic uridine-rich sequence elements acting as possible targets for nuclear proteins during pre-mRNA splicing in *Nicotiana plumbaginifolia*. Nucleic Acids Res **24:** 619–627

Golovikin M, Reddy AS (1996) Structure and expression of a plant U1 snRNP 70K gene: alternative splicing of U1 snRNP 70K pre-mRNAs produces two different transcripts. Plant Cell **8:** 1421–1435

Goodall GJ, Filipowicz W (1989) The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58:** 473–483

Goodall GJ, Filipowicz W (1991) Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants. EMBO J **10:** 2635–2644

Gorlach J, Raesecke HR, Abel G, Wehrli R, Amrhein N, Schmid J (1995) Organ-specific differences in the ratio of alternatively spliced chorismate synthase (LeCS2) transcripts in tomato. Plant J **8:** 451–456

Greene T, Chantler SE, Kahn B, Preiss J, Okita TW (1996) Mutagenesis of the potato ADP-glucose pyrophosphorylase and characterization of an allosteric mutant defective in 3-phosphoglycerate activation. Proc Natl Acad Sci USA **93:** 1509–1513

Hannah LC, Tuschall DM, Man RJ (1980) Multiple forms of maize endosperm ADP-glucose pyrophosphorylase and their control by *Shrunken-2* and *Brittle-2*. Genetics **95:** 961–970

Iglesias AA, Barry GF, Myer C, Bloksberg L, Nakata PA, Greene T, Laughlin MJ, Okita TW, Kishore GM, Preiss J (1993) Expression of the potato tuber ADP-glucose pyrophosphorylase in *Escherichia coli.* J Biol Chem **268:** 1081–1086

Jacobsen SE, Binkowski K, Olszewski NE (1996) *SPINDLY*: a tetratricopeptide repeat protein involved in gibberellin signal transduction in Arabidopsis. Proc Natl Acad Sci USA **93:** 9292–9296

Keith B, Chua N-H (1986) Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. EMBO J **5:** 2419–2425

Kiss-Laszlo Z, Blanc S, Hohn T (1992) Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. EMBO J **14:** 3552–3562

Krawczak M, Reiss J, Cooper DN (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet **90:** 41–54

Liu HX, Filipowicz W (1996) Mapping of branchpoint nucleotides in mutant pre-mRNAs expressed in plant cells. Plant J **9:** 381–389

Long JA, Moan EI, Medford JI, Barton MK (1996) A member of *KNOTTED* class of homeodomain proteins encoded by the STM gene of *Arabidopsis*. Nature **379:** 66–69

Lou H, McCullough AJ, Schuler MA (1993) 3′ Splice site selection in dicot plant nuclei is position dependent. Mol Cell Biol **13:** 4485–4493

Luehrsen KR, Walbot V (1994) Intron creation and polyadenylation in maize are directed by AU-rich RNA. Genes Dev **8:** 1117–1130

Mandel MA, Gustafson-Brown C, Savidge B, Yanofsky MF (1992) Molecular characterization of the *Arabidopsis* floral homeotic gene *APTALA1*. Nature **360:** 273–277

Maniatis T, Fritsch EF, Sambrook KJ (1993) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

McCarty DR (1986) A simple method for extraction of RNA from maize tissue. Maize Genet Coop Newslett **60:** 61

McCullough AJ, Bayton CE, Schuler MA (1996) Interaction across exons can influence splice site recognition in plant nuclei. Plant Cell **12:** 2295–2307

McCullough AJ, Lou H, Schuler MA (1993) Factors affecting authentic 5′ splice site selection in plant nuclei. Mol Cell Biol **13:** 1323–1331

McNellis TW, von Arnim AG, Araki T, Komeda Y, Misera S, Deng X-W (1994) Genetic and molecular analysis of an allelic series of cop1 mutants suggests functional roles for multiple protein domains. Plant Cell **6:** 487–500

Meyer CR, Ghosh P, Nadler S, Preiss J (1993) Cloning, expression, and sequence of an allosteric mutant ADPglucose pyrophosphorylase from *Escherichia coli B*. Arch Biochem Biophys **302:** 64–71

Moore MJ, Query CC, Sharp PA (1993) Splicing of precursors to messenger RNAs by the spliceosome. *In* R Gestland, J Atkins, eds, The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 303–308

Moore MJ, Sharp PA (1993) Evidence of two active sites in the spliceosome provided by stereochemistry of pre-mRNA. Nature **365:** 364–368

Nishihama R, Banno H, Kawahara E, Irie K, Machida Y (1997) Possible involvement of differential splicing in regulation of the activity of Arabidopsis ANP1 that is related to mitogen-activated protein kinase kinase kinase (MAPKKKs). Plant J **12:** 39–48

Okumiya T, Ishii S, Kase R, Kamei S, Sakuraba H, Yoshiyki S (1995) α-Galactosidase gene mutations in Fabry disease: heterogeneous expressions of mutant enzyme proteins. Hum Genet **95:** 557–556

Parker R, Guthrie C (1985) A point mutation in the conserved hexa-nucleotide at a yeast 5′ splice site junction uncouples recognition, cleavage and ligation. Cell **41:** 107–118

Parker R, Siliciano PG (1993) Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. Nature **361:** 660–662

Pautot V, Brzezinski R, Tepfer M (1989) Expression of mouse metallothionein gene in transgenic plant tissue. Gene **77:** 133–140

Sharp PJ (1994) Split genes and RNA splicing. Cell **77:** 805–815

Shaw JR, Hannah LC (1992) Genomic sequence of the wild type *Shrunken-2* allele of *Zea mays*. Plant Physiol **98:** 1214–1216

Sieburth LE, Running MP, Meyerowitz EM (1995) Genetic separation of the third and fourth whorl functions of *AGAMOUS*. Plant Cell **7:** 1249–1258

Simpson CG, Brown JWS (1993) Efficient splicing of AU-rich antisense intron sequence. Plant Mol Biol **21:** 205–211

Simpson CG, Clark G, Davidson D, Smith P, Brown JWS (1996) Mutation of putative branchpoint consensus sequences in plant introns reduces splicing efficiency. Plant J **9:** 369–380

Simpson CG, Filipowicz W (1996) Splicing of precursors to mRNA in higher plants: regulation and sub-nuclear organization of the spliceosomal machinery. Plant Mol Biol **32:** 1–41

Smith CJW, Chu TT, Nadal-Ginard B (1993) Scanning and competition between AGs are involved in 3′ splice site selection in mammalian introns. Mol Cell Biol **13:** 4939–4952

Smith CJW, Patton JG, Nadal-Ginard B (1989) Alternative splicing in the control of gene expression. Annu Rev Genet **23:** 527–577

Taylor MG, Vasil IK (1991) Histology of, and physical factors affecting, transient GUS expression in plant millet (*Pennisetum glaucum* (Li) R. Rv.) embryos following microprojectile bombardment. Plant Cell Rep **10:** 120–125

Tsai CY, Nelson OE (1966) Starch deficient maize mutant lacking ADP-glucose pyrophosphorylase activity. Science **151:** 341–343

van Santen VL, Spritz RA (1987) Splicing of plant pre-mRNAs in animal system and vice versa. Gene **56:** 253–265

Vasil V, Clancy M, Ferl RJ, Vasil IK, Hannah LC (1989) Increased gene expression by the first intron of maize *shrunken-1* locus in grass species. Plant Physiol **91:** 1575–1579

Waigmann E, Barta A (1992) Processing of chimeric introns in dicot plants: evidence for close cooperation between 5′ and 3′ splice sites. Nucleic Acids Res **20:** 75–81

Weil CF, Wessler SR (1991) The effects of plant transposable element insertion on transcription initiation and RNA processing. Annu Rev Plant Physiol **42:** 527–552

Wiebauer K, Herrero J-J, Filipowicz W (1988) Nuclear pre-mRNA processing in plants: distinct modes of 3′-splice-site selection in plants and animals. Mol Cell Biol **8:** 2042–2051

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM (1990) The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. Nature **346:** 35–39

Yi Y, Jack T (1998) An intragenic suppressor of the Arabidopsis floral organ identity mutant *apetala3-1* functions by suppressing defects in splicing. Plant Cell **10:** 1465–1477