Intron creation and polyadenylation in maize are directed by AU-rich RNA

Kenneth R. Luehrsen and Virginia Walbot
Department of Biological Sciences, Stanford University, Stanford, California 94305-5020 USA

Intron recognition in Angiosperms is hypothesized to require AU-rich motifs within introns. In this report we examined the role of AU-rich motifs in pre-mRNA processing. AU-rich segments of maize introns inserted near the single intron of the maize Bronze-2 (Bz2) gene result in alternative splicing. Other insertions of AU-rich sequence in the Bz2 cDNA resulted in de novo intron creation using splice junctions at the edges of the AU-rich region. Surprisingly, the five AU-rich inserts that we tested also caused polyadenylation, even though none had been selected for that function in plants. Insertions of GC-rich sequence into Bz2 did not cause either splicing or polyadenylation. We propose that AU-rich motifs are a general signal for RNA processing in maize and that in the absence of a 5′ splice site, polyadenylation is the default pathway.

[Key Words: Splicing; polyadenylation; maize; Bronze-2]

Received September 27, 1993; revised version accepted March 24, 1994.

The primary transcripts (pre-mRNAs) of eukaryotic genes undergo a series of post-transcriptional modifications, including the additions of a cap structure to the 5′ end of the message and of a poly[A] tail to the 3′ end of the transcript. Many genes also have introns, which must be excised or spliced from the pre-mRNA before its translation. In yeast and mammals, the biochemistry of both splicing (Hernandez and Keller 1983; Krainer et al. 1984) and polyadenylation (Takagaki et al. 1988) have been elucidated through the use of in vitro assays. To date, such assays have not been developed for plants, where the biochemistry of these processes is not yet known.

Although plant and animal introns do share some structural similarities, there are differences. Animal introns have well-conserved 5′ and 3′ splice junction sequences and a branchpoint/polypyrimidine stretch located 20-50 nucleotides upstream of the 3′ splice junction. In vivo and in vitro studies show that these sequence motifs are necessary and probably sufficient for intron recognition, spliceosome assembly, and efficient splicing (Lührmann et al. 1990; Green 1991). Plant introns have the same consensus 5′ splice junction sequence (AG GUAGU) found in yeast and vertebrates (Lührsen et al. 1994). However, plant 5′ splice junctions are more variable: On the basis of a survey of positions -1 to -6, only 8.4% of plant 5′ splice junctions are expected to match the consensus versus values of 14.4% and 40.6% for vertebrates and yeast, respectively (Lührsen et al. 1994). Plant introns have the typical 3′ [YAG G] splice junction sequence but appear to lack both a poly(pyrimidine tract and a conserved branchpoint sequence (Hanley and Schuler 1988; Goodall and Filipowicz 1989, 1991). On the other hand, novel features of plant introns have been discovered by analysis of intron composition (Wiebauer et al. 1988; Csank et al. 1990; White et al. 1992). Unlike animal introns, plant introns are usually AU-rich; Filipowicz and co-workers have shown that this feature is required for efficient splicing in dicots and monocots (Wiebauer et al. 1988; Goodall and Filipowicz 1989, 1991). In Figure 1A, a base content survey of genes from the monocot maize shows that exons are GC-rich (58%), whereas introns are AU-rich (60%), with a distinct bias for U (34%). For maize, both optimal 5′ splice site complementarity with U1 small nuclear RNA (snRNA) and intron AU-richness contribute to efficient splicing (Goodall and Filipowicz 1991). A few maize introns are GC-rich, but these are flanked by even more GC-rich exons and still contain AU-rich motifs that could serve as signals for intron recognition. The functional significance of AU-rich sequence has not been determined. It may act to minimize secondary structure within introns, alternatively, the transition from AU-rich introns to GC-rich exons might be involved in specifying intron borders.

To test the hypothesis that AU-rich motifs are a primary signal for plant intron recognition, we inserted several AU-rich fragments into the maize Bronze-2 (Bz2) gene and assessed their impact on RNA processing in transient gene expression assays. Fragments of the Alcohol dehydrogenase-1 (Adh1) intron 1 inserted downstream of the single Bz2 intron resulted in alternative splicing events in which the native Bz2 3′ splice site was skipped in favor of either a native Adh1 3′ splice site or...
The 3' gene region was divided into three parts: the 100 nucleotides polyadenylation sites were analyzed for base composition. The exons range from 35% to 78% GC and introns range from 23% to 66% GC. (B) The 3'-end regions of 29 genes with defined polyadenylation sites were analyzed for base composition. The 3' gene region was divided into three parts: the 100 nucleotides of coding region preceding the 3' untranslated region (UTR), the UTR (to the major polyadenylation site), and the 50 nucleotides after the major polyadenylation site.

Figure 1. Base content survey of maize exons, introns and 3'-end regions. (A) The intron/exon structure of 26 genes was completely defined and analyzed for base composition using 184 exons (47,078 nucleotides) and 152 introns (42,134 nucleotides). Exons range from 35% to 78% GC and introns range from 23% to 66% GC. (B) The 3'-end regions of 29 genes with defined polyadenylation sites were analyzed for base composition. The 3' gene region was divided into three parts: the 100 nucleotides of coding region preceding the 3' untranslated region (UTR), the UTR (to the major polyadenylation site), and the 50 nucleotides after the major polyadenylation site.

a cryptic 3' splice site at the insertion ends. The insertion of random AU-rich sequence resulted in de novo intron creation in which the AU-rich regions were spliced from the surrounding GC-rich sequence. Surprisingly, when splicing was not observed, the AU-rich fragments directed efficient polyadenylation, even though none had been selected for that function in plants. Insertions of GC-rich sequence did not cause RNA processing. We present evidence consistent with the hypothesis that AU-rich RNA is a general signal for RNA processing and that the presence of a usable 5' splice site proximal to the AU-rich RNA causes splicing, in the absence of a 5' splice site, the AU-rich RNA directs polyadenylation.

Results

The cDNA and intron-containing Bz2 genes were cloned downstream of the cauliflower mosaic virus (CaMV) 35S promoter resulting in expression plasmids pCaBz2c and pCaBz2i, respectively (Fig. 2A). The Bz2 pre-mRNA is 0.87 kb in length and has two exons [69% and 62% GC] separated by a single 78-nucleotide intron (Fig. 2B). Although the Bz2 intron is 54% GC, it has one central U-rich motif (8 of 11 residues are U) that could be involved in intron recognition. Plasmid constructs were electroporated into maize protoplasts, and after a 20- to 24-hr incubation, total RNA was isolated. RNA processing was assayed by Northern blots, RNase protection, and PCR amplification of first-strand cDNA [reverse transcriptase polymerase chain reaction (RT–PCR)], the probes and primers used are shown in Figure 2C.

Northern blot analysis showed that the expression of pCaBz2c and pCaBz2i each resulted in transcripts of ~0.9 kb (Fig. 3, band 1), the length expected for wild-type-sized Bz2 mRNA. No Bz2 mRNA was observed in the untransformed control cells, indicating that the genomic copies of Bz2 are not expressed. Quantification of transcript levels for pCaBz2c and pCaBz2i showed that inclusion of the Bz2 intron consistently increased expression ~1.5-fold [data not shown] over the cDNA construct, in accordance with previous observations that introns can enhance transcript accumulation (Callis et al. 1987; Buchman and Berg 1988; Luehrsen and Walbot 1991). RNase protection [data not shown] indicated that the pre-mRNA from pCaBz2i is efficiently spliced (>90%), conforming to previous results for the endogenous Bz2 gene in Black Mexican Sweet (BMS) corn cells (Nash and Walbot 1992).

Insertions of AU-rich sequence into the Bz2 gene cause alternative splicing

We tested whether the addition of AU-rich sequence to the GC-rich Bz2 exons is a sufficient signal to cause alternative RNA processing. AU-rich fragments (Fig. 2D) derived from the Adh1 intron 1, actin intron 3, and bacteriophage λ were inserted at several restriction sites in the Bz2 gene. After transient expression of the Bz2 chimeric constructs, we deduced the structure of each major transcript from Northern blot, RNase protection, and cDNA sequence data. The alternative RNA processing events that we observed are summarized in Figures 4 and 5. For ease of identification, the correspondence codes in Figures 4 and 5 refer to the like-numbered bands observed in the Northern blots and in the Southern blots of RT–PCR products.

First, we tested the splicing potential of a construct containing a 113-nucleotide, AU-rich fragment (54% AU) from the 3' end of Adh1 intron 1, the fragment ends with the native 3' splice site (CAG). The 113-nucleotide fragment was placed in the Smal site of the cDNA construct pCaBz2c, resulting in c-s113S. Northern blot analysis of c-s113S (Fig. 3) showed that the transcript length was the sum of the Bz2-coding region and the 113-nucleotide insert; this result indicates that the inserted fragment did not initiate novel RNA processing. A different result was obtained when the 113-nucleotide fragment was inserted at the Smal site, 41 nucleotides downstream of the intron in pCaBz2i (i-s113S). By Northern
Splicing and polyadenylation in maize

**Figure 2.** Expression plasmids, probes, primers, and insertion sequences. (A) The structures of pCaBz2c and pCaBz2i are shown along with the restriction sites used for the AU-rich insertions. The nucleotide lengths of each exon and the single intron are noted. (B) The sequence of the 78-nucleotide Bz2 intron. The region with 8 of 11 U nucleotides is underlined. (C) Primers used in RT–PCR and segments of the gene used to make hybridization probes. First-strand cDNA for insertions at the SmaI site was primed using the antisense oligonucleotide 1365. Oligonucleotide 951 was used to prime at poly(A) tails. The positions of the PA139 and P300 segments used to make RNA probes are shown. Oligonucleotide 972 was end-labeled and used as a hybridization probe. (D) AU-rich insertion sequences. The broken line above each restriction map denotes the region used as an insertion sequence. The base composition of each insertion sequence is reported in Materials and methods.

---

blot analysis, RNA from i-s346S was ~0.85 kb [Figs. 3 and 5B, band 2], indicating that the pre-mRNA was processed from the expected full length of 0.98 kb. Using a Bz2 probe [P300] from exon 2, the same 0.85-kb transcript was detected [Fig. 5B]; however, an Adh1 intron 1 probe did not detect the 0.85-kb transcript [data not shown]. Taken together, the results suggest that the intron insert was spliced from the transcript.

Using RT–PCR [Fig. 5E], only two types of cDNAs for i-s113S were observed. These amplification products were cloned and sequenced, and they correspond to the mRNA structures depicted in Figure 4A. One cDNA (2a) had the Bz2 5’ splice site [AG GUGAGC; underline indicates a consensus match] spliced to the Adh1 3’ splice site; the 78-nucleotide Bz2 intron plus 41 nucleotides of exon sequence and the 113-nucleotide Adh1 insertion were excised as a 232-nucleotide intron. The other cDNA (2b) indicated that an alternative [cryptic] 5’ splice site [AG GUAACG] within the Bz2 intron was spliced to the Adh1 3’ splice site. Although the quantification of RT–PCR products suggests that the cryptic 5’ splice site is used preferentially over the native Bz2 5’ splice site, RNase protection showed that the native Bz2 5’ splice site was used about fourfold more often than the cryptic 5’ splice site [data not shown]. In the wild-type gene, we never detected use of the cryptic 5’ splice site [data not shown; Nash et al. 1990; Nash and Walbot 1992]. We found no cDNAs in which the Bz2 3’ splice site was chosen, indicating that it was always skipped in favor of the Adh1 3’ splice site. The evidence suggests that the AU-rich, 113-nucleotide sequence was recognized as an intron, provided that a usable 5’ splice site was supplied by the Bz2 intron. Because the Bz2 3’ splice site was always skipped, we hypothesize that the splicing machinery recognizes extended regions of AU-rich sequence and excises these and any inclusive GC-rich sequence [such as the intervening 41 nucleotides of Bz2 exon sequence]. The 5’ and 3’ splice sites are selected at the proximal and distal borders, respectively, of the AU-rich region.

We then prepared a 346-nucleotide fragment from the center of Adh1 intron 1; this fragment is 58% AU and lacks both native 5’ and 3’ splice junction sequences. The sense orientation of the 346-nucleotide fragment was inserted at the SmaI site downstream of the Bz2 intron [i-s346S]. Northern analysis showed that i-s346S yielded RNAs of 1.3 and 0.9 kb [Fig. 5C, bands u and 3]; the 1.3-kb species corresponds in size to the unprocessed transcript. The P300 exon 2 probe detected the same
The probe is PA139 RNA. The numbers to the right of each band refer to the correspondence code described in Fig. 4. Band 1 for pCaBz2i is a mixture of spliced and unspliced RNAs, which are not resolved on this Northern blot. Bands A1 and A2 represent transcripts polyadenylated at multiple, closely spaced sites (see Fig. 6).

Some AU-rich insertions cause polyadenylation

Some AU-rich insertions, however, did not result in intron creation. A construct containing the sense orientation of the 346-nucleotide Adh1 intron 1 fragment inserted into the Bz2 cDNA Smal site (c-s346S) yielded a major transcript of ~0.7 kb (Fig. 5C, band A4), significantly shorter than the expected primary transcript of 1.3 kb. The 0.7-kb band also hybridized with Adh1 intron sequence (data not shown), indicating that the intron sequence was not completely removed from the tran-
Splicing and polyadenylation in maize

| Plasmid | Insert | Correspondence code |
|---------|--------|---------------------|
| A       | i-s113S 113 bp end of Adh1 intron 1 | 2a, 2b             |
|         | i-s346S 346 bp middle of Adh1 intron 1 | 3a, 3b             |
|         | x-e346S | 4                  |
|         | c-a346S | 5a                 |
|         | i-s346S | 5b                 |
|         | i-a346S | 6a                 |
|         | c-a302S 302 bp middle of actin intron 3 | 6b                 |
|         | i-s346B | 6c                 |
|         |       | 6d                 |
| B       | c-s113B 113 bp 3' end of Adh1 intron 1 | A1                |
|         | i-s113B | A2                |
|         | c-a346A 346 bp middle of Adh1 intron 1 | A3                |
|         | c-a346S | A4                |
|         | c-a346B | A5                |
|         | i-s346S | A6                |
|         | i-s346B | A7                |
|         | c-a302B 302 bp middle of actin intron 3 | A8                |
|         | c-a346S | A9                |
|         | i-a346B | A10               |
| C       | c-s113S 113 bp 3' end of Adh1 intron 1 | -                 |
|         | c-Aa441S 441 bp phage | -                 |
|         | c-Aa441S | -                 |

Figure 4. Expression plasmids and RNA processing outcomes. Only the expression casettes are shown; the vector is pUC8. Refer to Fig. 2D for the sources of the insertion sequences; the solid or broken lines represent the sense or antisense orientation of the insert, respectively. The positions of spliced introns are shown with broken lines below each construct, and the alternative polyadenylation sites are denoted by vertical arrows. Each processed transcript type is assigned a code corresponding to the cartoon illustrating its structure. The correspondence code refers each processed transcript to the corresponding band on the Northern and RT-PCR Southern blots. The correspondence code 1 represents spliced mRNA. [A] Constructs that are alternatively spliced. When different spliced transcripts were unresolved in Northern blots, they were denoted a, b, etc. [B] Summary of constructs that were alternatively polyadenylated. Each transcript type is assigned a code of A#. In most cases, this refers to polyadenylation over a broad region as shown by two vertical bold arrows in the cartoon. The precise nucleotide positions of the polyadenylation sites are shown in Fig. 6. [C] No alternative RNA processing was observed for the constructs c-s113S, c-λs441S, and c-λα441S.
Luehrsen and Walbot

Figure 5. Northern blots of RNA derived from several expression constructs. Total RNA from transfected BMS protoplasts was displayed in three blots as described in Materials and methods. The top row of blots is probed with PA139 (Fig. 2C), derived from exon 1 of the Bz2 gene. The bottom row is probed with P300 from exon 2 of the Bz2 gene. Most of the hybridization below -0.35 kb presumably represents RNA degradation products; these products were not characterized further. The autoradiographs were sliced apart and reorganized for ease of viewing. The numbers to the right of the major transcripts refer to the correspondence codes described in Fig. 4 for each mRNA structure. (A) Unprocessed RNA. (A) Transcript RNA derived from control cells and the plasmids pCaBz2c and pCaBz2i. Transcript RNA from the BMS Bz2 gene is not detectable in the control lane. (B) Transcript RNA from the i-s113S plasmid. (C) Transcript RNA from constructs having insertions of the 346-nucleotide Adh1 intron fragment at the Smal site in the Bz2 gene. (D) Transcript RNA from constructs in which polyadenylation occurred in the Adh1 intron 1, actin intron 3, or bacteriophage λ insertion sequences. Many of the transcripts that were polyadenylated within the AU-rich insertions do not hybridize to Bz2 exon 2; note especially the absence of bands A3, A8, and A9 with the P300 probe. (E) Southern blot of RT-PCR products amplified using primers 1364 and 1365. The PCR-amplified cDNAs from several expression constructs were probed with oligonucleotide 972 (Fig. 2C). The numbers to the right of each amplified product are the correspondence codes described in Fig. 4. (F) cDNA from unprocessed transcript. The arrow to the right of the blot denotes an amplification product representing unspliced Bz2 transcript, presumably derived from low-level expression of the BMS genomic copy in this experiment. The bands immediately above bands 2b and 3b are possibly PCR artifacts, direct sequencing showed that each has the same sequence in the construct region as bands 2b and 3b. (F) Southern blot of RT-PCR products amplified using the poly(dT) primer. First-strand cDNA was synthesized using oligonucleotide 951 (Fig. 2C) and PCR-amplified using primer 1364. The blot is probed with oligonucleotide 972 (Fig. 2C). The exact nucleotide positions of the sequenced poly(A) addition sites are shown in Fig. 6. No amplified products were observed for the plasmids pCaBz2c and pCaBz2i, possibly because of inefficient reverse transcriptase or AmpliTaq polymerase activity through the GC-rich exon 2 sequence of the Bz2 gene.
synthesized by the 3' RACE (rapid amplification of cDNA ends) method (Frohman et al. 1988) using a poly[dT] primer for first strand synthesis, one major and several minor bands were amplified (Fig. 5F). These amplified cDNAs were cloned, and from the sequence it is clear that polyadenylation occurred at several sites, clustered in two regions ~100 and 225 nucleotides into the Adhl insert (Fig. 4B, sequence shown in Fig. 6B). When the 346-nucleotide insert was placed near the 5' end of Bz2 (ApaI site; c-s346A) or near the end of exon 2 (BsaAI site; i-s346B), similar 3' end-processing was observed (Fig. 4B, bands A1 and A10). Quantification of truncated and full-length (unprocessed) transcripts from the Northern blot in Figure 5C showed that 96% of the transcript from c-s346S was polyadenylated in the 346-nucleotide intron insert.

Taken together, the results indicate that an efficient signal for polyadenylation is wholly contained in the Adhl intron 1 insert; the polyadenylation sites were not created in conjunction with specific Bz2 sequences. It is surprising that the Adhl intron 1 insert directs polyadenylation, because this sequence is in the first fifth of the native Adhl transcript.

We also inserted the 113-nucleotide Adhl intron fragment at the BsaAI site (near the 3' end of the gene) of pCaBz2c and pCaBz2i, resulting in c-s113B and i-s113B, respectively. c-s113B and i-s113B yielded identically sized transcript populations of ~0.85 and ~0.95-kb (Fig. 3). Using RT–PCR with primers flanking the insertion (not shown), no spliced cDNAs were recovered, even though a sequence (GG GUCCAC) with a weak resemblance to the 5' splice junction consensus sequence was immediately upstream of the insert. Using RT–PCR and a poly[dT] primer for first-strand cDNA synthesis, several unique fragments were amplified (not shown). The sequence of the amplified cDNAs showed that the insertion shifted the preferred polyadenylation sites to within the Adhl sequence, just upstream of the normal Bz2 polyadenylation region (Fig. 6, cf. E with A). Although the 113-nucleotide fragment inserted in the SmaI site of pCaBz2c did not cause polyadenylation (Fig. 5E), the same fragment inserted in the BsaAI site led to premature polyadenylation. Thus, the 113-nucleotide fragment can alter polyadenylation when followed by the AU-rich Bz2 3' region. We speculate that the 113-nucleotide AU-rich insertion in c-s113S (embedded in GC-rich exon sequence) was itself not long enough to support efficient polyadenylation.

We have shown that both orientations of the 346-nucleotide Adhl intron 1 fragment inserted in the SmaI site of the Bz2 gene can result in alternative splicing, and the same fragments can also direct polyadenylation. Transcripts from i-s346S, c-a346S, and i-a346S were amplified using RT–PCR and the poly[dT] primer, and the cDNAs are displayed on the Southern blot shown in Figure 5E. For all three constructs, intensely hybridizing bands are visible along with a smear in the 300- to 700-bp range. When several cDNAs were cloned and sequenced, we found that both the sense and antisense orientations of the 346-nucleotide fragment were capable of directing
polyadenylation when placed within the GC-rich Bz2 exon (Fig. 6B,F). Northern blot data show that transcripts from i-s346S are predominantly spliced and that transcripts polyadenylated within the insertion are rare. Because the alternatively spliced and polyadenylated transcripts were not completely resolved on the Northern blot, we could not quantify the splicing or polyadenylation efficiencies for c-a346S and i-a346S.

Two other AU-rich insertions (Fig. 2D) were tested for their ability to promote either alternative splicing or polyadenylation. The 302-nucleotide (66% AU) antisense orientation of an internal fragment of maize actin intron 3 and a 352-nucleotide (62% AU) fragment of bacteriophage λ DNA were installed at the SmaI site of pCaBz2c, to construct c-a302S and c-λ352S, respectively. Neither sequence has been selected to function as an intron or a polyadenylation region in maize. Northern blot analysis (Fig. 5D) showed a major mRNA species of ~0.45 kb for c-a302S (band A8) and ~0.53 kb for c-λ352S (band A9). Neither of the RNA samples hybridized to the P300 exon 2 probe (Fig. 5D). Only a single band corresponding to unprocessed transcript was amplified by RT–PCR using primers 1364 and 1365 (Fig. 5E), indicating that alternatively spliced products did not accumulate to a detectable degree. Using RT–PCR and the poly(dT) primer (Fig. 5F), a major cDNA product within a background smear was amplified from each transcript pool. Several cDNAs from both c-a302S and c-λ352S were cloned and sequenced. Both the antisense actin intron 3 and λ fragments directed polyadenylation of transcripts beginning ~100 to ~200 nucleotides from the proximal end of the insert (Fig. 6G,H). AU-richness is the only apparent shared feature of the five inserts that directed polyadenylation, suggesting that random AU-rich sequence can be a sufficient signal for polyadenylation.

GC-rich insertions do not cause RNA processing

We found that insertions of AU-rich sequence into Bz2 caused either cryptic splicing or premature polyadenylation. To test whether RNA processing can be caused by the insertion of any sequence into Bz2, we inserted a 441-nucleotide, GC-rich fragment of λ in both orientations into the SmaI site of pCaBz2c (c-λ441S and c-λ441S, Fig. 4C). The λ fragment is 62% GC, similar to the 65% GC Bz2 cDNA. Northern blot analysis showed that the transcript length for both c-λ441S and c-λ441S was the sum of the Bz2 cDNA and the λ insertion [data not shown]. RT–PCR analysis failed to detect either spliced or polyadenylated cDNAs caused by the insertion sequences [data not shown]. The results indicate that the GC-rich inserts we tested did not initiate novel RNA processing.

A 5' splice site mediates the choice between splicing and polyadenylation

We have shown that AU-rich regions of pre-mRNA can be either excised as an intron or direct polyadenylation. What signal specifies whether an AU-rich region will be spliced or polyadenylated? We hypothesize that a usable 5’ splice junction proximal to an AU-rich region will cause the AU-rich region to be spliced, whereas in its absence the AU-rich RNA will direct polyadenylation. To test this hypothesis, we constructed x-s346S by deleting all but the first 13 nucleotides of the Bz2 intron from i-s346S. In x-s346S, the first 13 nucleotides of the Bz2 intron contains the 5’ splice junction and is 42 nucleotides upstream of the Adh1 insertion. RNA processing of pre-mRNA from c-s346S, i-s346S, and x-s346S was then compared (Fig. 7). x-s346S RNA is predominantly spliced (85%), similar to the mRNA profile for i-s346S. The deleted Bz2 intron sequence, which includes the central U-rich motif, is not necessary for intron recognition; these functions are supplied by the internal parts of the Adh1 intron 1 insertion. In contrast, mRNA from c-s346S, in which no 5’ splice site appears upstream of the 346-nucleotide insert, is predominantly polyadenylated (96%). Interestingly, when the 346-nucleotide insert is far downstream of the Bz2 intron [i-s346B], the insert directs polyadenylation but is not spliced (Fig. 5D), probably because the 375-nucleotide of GC-rich exon 2 sequence separates the Bz2 intron and the AU-rich insert and precludes intron recognition. Considered together, the results suggest that AU-rich RNA is recognized and

Figure 7. Northern blot comparing spliced and polyadenylated transcripts from constructs containing the 346-nucleotide Adh1 intron 1 insertion. PA139 antisense RNA was used as a hybridization probe. The numbers to the right of each band refer to the correspondence code described in Fig. 4. Band 1 for pCaBz2c is a mixture of spliced and unspliced RNAs, which are not resolved on this Northern blot. Band A4 is RNA polyadenylated at several sites within the 346-nucleotide insertion [see Fig. 6B]. Bands 3 and 4 are spliced using the Bz2 5’ splice site and a cryptic 3’ splice site at the end of the 346-nucleotide insertion. [u] Unprocessed RNA. The intron deletion in x-s346S makes its unprocessed RNA slightly shorter than the unprocessed RNA from i-s346S.
Splicing and polyadenylation in maize

destined for processing, and the presence of the nearby Bz2 5' splice junction is sufficient to shunt the transcript to the splicing pathway. An early step in splicing may allow this process to take precedence over polyadenylation. For example, the binding of the 5' splice junction sequence with U1 snRNP is one of the first steps in yeast and animal splicing (Rosbash and Seraphin 1991; Bennett et al. 1992). In mammalian and viral genes in which alternative splicing involves removal of polyadenylation signals, splicing generally precedes the choice of polyadenylation site [Leff et al. 1987; Peterson and Perry 1989].

Discussion

The cis-acting RNA signals that lead to splicing and polyadenylation are not as well characterized in plants as they are in yeast or animals. Elegant experimental evidence implicating a role for AU-rich sequence in splicing in plants has been provided by Filipowicz and colleagues based on the splicing efficiency of artificial and native introns [Wiebauer et al. 1988; Goodall and Filipowicz 1989, 1991]. They have shown that dicot introns must be >60% AU-rich to be spliced efficiently in vivo. The same group has also shown that although monocot introns are sometimes GC-rich, AU-richness is beneficial to introns having suboptimal splice junction sequences. Other studies have corroborated the importance of AU-rich RNA for splicing. The insertion of GC-rich cDNA sequences into two maize introns interferes with splicing, whereas the insertion of AU-rich intron sequence does not [Luehrsen and Walbot 1992; Paszkowski et al. 1992]. We have shown that the addition of A- and U-rich oligomers can improve the splicing efficiency of a deletion-impaired variant of Adhl intron 1 [Luehrsen and Walbot 1994]. McCullough et al. (1993) demonstrated that AU-rich sequence can contribute to the use of suboptimal 5' splice junction in chimeric expression constructs. Lou et al. (1993a,b) have recently demonstrated that AU-rich to GC-rich transition points in pre-mRNAs can influence 3' splice site choice. Simpson and Brown (1993) showed that AU-rich introns can be spliced from the antisense sequence of a pea legumin intron inserted into a maize zein gene.

Angiosperm gene organization is quite compact compared with that of mammals. In mammals, there are often very long introns punctuated by short exons [Hawkins 1988; Smith 1988], and definition of exon borders is emphasized [Robberson et al. 1990]. In plants, exons are generally long [Hawkins 1988] and separated by shorter introns [80% of maize introns are <200 nucleotides [Luehrsen et al. 1994]]. Thus, in plants, introns must be recognized within the context of pre-mRNAs that are mainly exon. Available experimental data favor the hypothesis of internal intron definition in plants, that is, that recognition of the AU- or U-rich motifs within introns is a first step in processing. Consequently, a long GC-rich sequence within an intron might be interpreted as an exon and either interfere with intron recognition or divide one intron into two [Luehrsen and Walbot 1990, 1992].

In this report we further define the roles of AU-rich sequence in pre-mRNA processing in maize. We reasoned that the insertion of AU-rich fragments into the GC-rich exon sequence of the Bz2 gene might cause novel RNA processing events. We inserted two partial fragments of Adhl intron 1 at three restrictions sites of the Bz2 gene and determined how each chimeric transcript was processed. When the 113- or the 346-nucleotide Adhl intron 1 fragments (sense orientation) are placed at the Smal site 41 nucleotides downstream of the native Bz2 intron, new introns are created. One type of intron uses the Bz2 5' splice site spliced to a 3' splice site at the end of the insertion. The 113-nucleotide fragment contains the native Adhl 3' splice site, which is used exclusively over the now proximal Bz2 3' splice site. When the 346-nucleotide insert is placed downstream of the Bz2 intron, the Bz2 3' splice site is again skipped in favor of a new 3' splice site found at the junction between the AU-rich intron insert and the GC-rich Bz2 exon 2. In each case, the entire Bz2 intron, the intervening 41 nucleotides of Bz2 exon 2 and the entire insert, is excised as an intron. We propose that the 113- and 346-nucleotide intron insertions effectively extend the region identified as an intron. Taken together, the data show that extended regions of AU-rich sequence are recognized as "intron" and that base content information supersedes the splicing signal provided by the Bz2 3' splice junction motif. The results also add weight to the idea that 3' splice site choice is mediated, in part, by the transition from AU-rich intron to GC-rich exon sequence and suggest that any AG doublet can be used as a 3' splice site if it is present at a transition point [Lou et al. 1993a]. These results parallel the findings that maize transposable element insertions close to introns often change splicing efficiency and contribute alternate splice sites [Kim et al. 1987; Wessler et al. 1987; Luehrsen and Walbot 1990; Menssen et al. 1990; Ortiz and Strohmer 1990; Wessler 1991; Varagona et al. 1992].

The results with our constructs in which 41 nucleotides of exon sequence separated the Bz2 intron and the Adhl intron insertion showed that the 41-nucleotide region was excised from the pre-mRNA as intron sequence. Although the 41 nucleotides of exon sequence separating the Bz2 intron and the Smal insertion site is 73% GC, it is probably too short to be recognized as exon. Generally, maize internal exons are \(>47\) nucleotides [at 24 nucleotides, exon 2 of the maize glyceraldehyde dehydrogenase gene is an exception [Quigley et al. 1988]]. In contrast, when the 113- and 346-nucleotide AU-rich insertions were placed downstream of the Bz2 intron [at the BsaAI site], the intervening exon region was 375 nucleotides and 71% GC; this is probably sufficient to define it as exon and prevent the excision of the AU-rich sequence as part of a larger intron. This interpretation is consistent with the observation that blocks of GC-rich sequence interfere with splicing in dicots (Goodall and Filipowicz 1989), that the insertion of GC-rich exon sequence interferes with splicing in maize...
Luehrsen and Walbot (1992), and that GC-rich transposable elements can cause splicing failure in maize (Luehrsen and Walbot 1990; Ortiz and Strommer 1990).

We also observed that insertion of both the 113- and 346-nucleotide *Adh1* fragments downstream of the *Bz2* intron activated a cryptic 5' splice site present within the *Bz2* intron. The sequence of the cryptic 5' splice site is AG GUACC, this is a 6/8 match to the consensus sequence for plants, AG GUAGU. We found no evidence that the cryptic 5' splice site was used with the native *Bz2* 3' splice site, the potential intron would be only 34 nucleotides, below the minimum functional length of 70-73 nucleotides determined for dicot and monocot introns (Goodall and Filipowicz 1990). In the i-s1135 construct, the cryptic 5' splice site is used fourfold less often than the wild-type *Bz2* 5' splice site. In the i-s3465 construct, the same cryptic 5' splice site is also used with the cryptic 3' splice site at the end of the 346-nucleotide *Adh1* insert, but RNase protection data showed that this product was not abundant. It has been reported that individual 5' splice sites might not splice equally well to all potential 3' splice site sequences (Goodall and Filipowicz 1990). In the i-s1135 construct, the cryptic 5' splice site within the *Bz2* intron is a poor match with the cryptic 3' splice site within the 346-nucleotide insert.

**Introns can be created from AU-rich sequence**

We also showed that introns could be "created" from an insertion of AU-rich sequence. We inserted the antisense orientation of the 346-nucleotide *Adh1* intron fragment into the *Smal* site of *Bz2* cDNA [c-a346S]. PCR amplification and sequencing of cDNA sequences showed that two introns were created de novo from the AU-rich insertion. The introns were 92 and 96 nucleotides, typical lengths of maize introns. Both the 5' and 3' junction sequences were similar to the consensus derived for plants. As expected, each intron was more AU-rich than the surrounding exons, further implying that the transition from AU-rich to GC-rich sequence is information that specifies the location of intron borders. The experiments also show that AU-rich sequence is sufficient to specify an intron, provided that sequences that approximate 5' and 3' splice junction sequences occur fortuitously. Concerning the evolution of introns, these findings provide experimental justification for the view that in plants, introns can arise simply by the insertion of AT-rich DNA in exon regions [which are usually GC-rich]. Many maize transposable elements are AT-rich or contain regions of AT-richness. For example, the AT-rich *Ds1* element can be considered a mobile intron when inserted in the *Wx* gene (Wessler 1991). In addition, *dSpm* and *Mu* elements also have AT-rich regions capable of causing aberrant splicing (Weil and Wessler 1990, Luehrsen et al. 1994).

**Cis-acting sequences involved in the polyadenylation of plant mRNAs**

In animals, polyadenylation of pre-mRNA occurs at a precise site for each mRNA and is set by the highly conserved AAUAAA motif and by additional sequence motifs near the poly(A) addition site (Proudfoot 1991). The cis-acting RNA sequences and the enzymology of polyadenylation in plants have not been well characterized. Unlike animals, control of plant polyadenylation is relaxed and occurs at multiple sites, 100-400 nucleotides downstream of the stop codon (Hunt et al. 1987b; Hunt and MacDonald 1989; Sanfaçon and Hohn 1990; Guerineau et al. 1991, MacDonald et al. 1991). Deletion analysis of the 3' ends of several plant genes has implicated regions that are important for efficient polyadenylation (Hunt and MacDonald 1989; Mogen et al. 1990, 1992, Guerineau et al. 1991, MacDonald et al. 1991, Sanfaçon et al. 1991, Wu et al. 1993). The far upstream element [FUE] is generally located 50-150 nucleotides upstream of the cleavage/polyadenylation site (CS). The FUE is U-rich but has no well-conserved sequence motif common to all plant 3' ends. The sequence element UUGUA is present in some plant FUEs and has been suggested to be functionally important (Mogen et al. 1992), but this motif [and variants of it] is not present in all plant 3' ends and would probably occur fortuitously in U-rich RNA. The near upstream element [NUE] occurs in the 50 nucleotides preceding the CS site and is proposed to contain sequences that are related to the AAUAAA motif that is essential for efficient polyadenylation in animals. Alterations of the AAUAAA-like motifs in plant 3' ends have led to ambiguous results. Several reports have shown that deletion or mutagenesis of the AAUAAA-like motifs is deleterious to polyadenylation, but polyadenylation function was not completely eliminated (Mogen et al. 1990, 1992; Guerineau et al. 1991; MacDonald et al. 1991; Wu et al. 1993). In only one study was deletion of the AAUAAA motif found to inhibit polyadenylation completely (Sanfaçon and Hohn 1990). The AAUAAA sequence is not present in many plant 3' ends (Joshi 1987a). There is apparently no functional, conserved sequence downstream of the CS site (Sanfaçon et al. 1991; Wu et al. 1993).

We were surprised to find that when not spliced from the pre-mRNA, all five of the AU-rich insertions into *Bz2* could direct polyadenylation. In some cases, the polyadenylation was very efficient: For c-a346S, 96% of the transcript was polyadenylated in the AU-rich intron sequence. Also, the 346-nucleotide *Adh1* intron 1 insert directed efficient polyadenylation when placed at three positions in the *Bz2* gene, showing that the functional sequence was within the insertion. Most surprisingly, a 352-nucleotide, AU-rich fragment of bacteriophage λ also directed polyadenylation. In each instance, the polyadenylation sites were scattered from 84 to 253 nucleotides into the AU-rich insert, consistent with the observation that most native transcripts are polyadenylated 100-400 nucleotides downstream of the stop codon. The five sequences that directed polyadenylation had no long sequence motif in common, and none contained the AAUAAA motif that is essential for efficient polyadenylation for animals. The UUGUA motif was found in only one (346 nucleotides, sense orientation of *Adh1* intron 1).
Splicing and polyadenylation in maize

A 5’ splice junction sequence specifies the choice between splicing and polyadenylation of AU-rich RNA

The survey of maize genes in Figure 1 shows that both introns and 3’-end regions are AU-rich. Furthermore, there are no obvious highly conserved sequence features of internal intron regions or 3’-end regions, suggesting that AU-richness per se is functionally important. If AU-rich sequence is a determinant of RNA processing, how is the choice between splicing and polyadenylation mediated? We found that when the 346-nucleotide Adhl fragment was inserted in the Smal site of the Bz2 3’ cDNA, it directed efficient polyadenylation, however, when either the native Bz2 intron or just the Bz2 5’ splice junction was upstream of the insertion, the insertion was efficiently spliced from the transcript and polyadenylation occurred instead at the native Bz2 3’ end. We propose that a usable 5’ splice site proximal to an AU-rich region will cause that region to be spliced; in the absence of a 5’ splice site, the same AU-rich region will direct the default pathway of polyadenylation. This is similar to the processing of the rat calcitonin/CGRP gene, in which polyadenylation regions that occur in spliceable introns are rendered silent, suggesting that splicing occurs before polyadenylation [Leff et al. 1987]. Further work will focus on the discovery of trans-acting factors that bind AU-rich RNA and mediate the choice in RNA processing.

Materials and methods

Expression plasmids

Bz2 encodes the last genetically defined step in the synthesis of anthocyanin, the purple pigment of maize [McLaughlin and Walbot 1987; Nash et al. 1990]. The complete Bz2-coding region from either pBl2c or pBl2i [Nash et al. 1990, Nash and Walbot 1992] was excised with Bsal (end filled in with Klenow and dNTPs) and BglII and inserted into the KpnI (end trimmed with T4 DNA polymerase)–BglII sites of plasmid pD300 [Luehrsén et al. 1992], resulting in constructs pCaBz2c and pCaBz2i, respectively (Fig. 2A).

Insertions of AT-rich DNA into pCaBz2c and pCaBz2i were made in the Apal, Smal, or BsaAI sites [Fig. 2A]. The insertion sequences (Fig. 2D) were derived in the following way: Adhl intron 1 was digested with Stul and PvuII releasing a 113-bp, 54% AT fragment [positions 704–816 (Dennis et al. 1984)] ending with the 3’ splice site sequence CAG. The base composition of the 113-bp fragment is 20% A, 24% C, 22% G, and 34% T. Adhl intron 1 was also digested with Alul and Stul releasing a 346-bp, 58% AT fragment (positions 358–703) after the stop codon, suggesting that this base content change could be a signal for polyadenylation. The 50 nucleotides after the poly(A) site is 62% AU and even more U-rich (40%).

Previous reports have concluded that plant 3’ ends are structurally complex, but our experimental evidence suggests that they are really quite simple. The experimental results combined with the survey suggest that in maize, polyadenylation occurs in an AU- or U-rich sequence environment, without requiring strictly conserved sequence motifs. The insertions that directed polyadenylation were derived from sequences that had not been selected for polyadenylation function and, therefore, approximate “random” AU-rich sequence. Because plant 3’ ends are AU-rich, the occurrence of UUGUA- and AAUAAA-like motifs should occur by chance alone. We suggest that the previously characterized FUE elements contain no extended, conserved motifs and that their function simply depends on AU- or U-richness. This conclusion is similar to studies of yeast 3’ ends, which are also AU-rich and where no strictly conserved sequences have been identified [Heidmann et al. 1992; Irmer et al. 1992]. Also, a recent report shows that random U-rich sequence at the 3’ ends of vaccinia virus transcripts is sufficient to direct polyadenylation [Gershon and Moss 1993]. Further experiments in which the length and AU content of plant 3’ ends are systematically altered will be necessary to test rigorously our hypothesis that polyadenylation regions are composed of AU-rich RNA rather than one or more conserved motifs.
with T4 DNA polymerase) or the Smal site of pCaBz2c. c-a346S and t-a346S: The antisense orientation of the 346-bp fragment was inserted in the Smal site of pCaBz2c and pCaBz2i. c-a302S: The antisense orientation of the 302-bp actin intron 3 fragment was placed in the Smal site of pCaBz2c. c-a352S: The 352-bp fragment of bacteriophage λ DNA was placed in the Smal site of pCaBz2c. x-s346S: i-s346S was digested with BsmI and Psrl (ends trimmed with T4 DNA polymerase) and the blunt ends were religated; 63 bp of the 78-bp Bz2 intron was deleted.

Hybridization probes

The segments of the Bz2 gene used to make hybridization probe plasmids are shown in Figure 2C. A 139-bp PstI-Apal fragment [positions 685–823 of Bz2 (Nash et al. 1990)] was cloned into the same sites of pBSKS(-) resulting in pBSPA139. The 300-bp PstI fragment containing exon 2 sequences [positions 1118–1418 of Bz2 (Nash et al. 1990)] was cloned into the PstI site of pBSKS(−) to give pBS300. 32P-Labeled antisense RNA was synthesized from pBSPA139 and pBS300, and used as hybridization probes. In addition to RNA probes, an oligomer (972, positions 1021–1041 of Bz2 [Nash et al. 1990]) was synthesized and end-labeled using [γ-32P]ATP.

Gene transfer and RNA analysis

Transcription from the Bz2 promoter in BMS tissue culture cells requires the R and C1 regulatory gene products (Bodeau and Walbot 1992). To circumvent that requirement and to preclude expression of the endogenous BMS copy of Bz2, the cDNA and intron-containing Bz2-coding regions were placed under the control of the constitutive CaMV 35S promoter. BMS protoplasts were electroporated with 50 μg of plasmid DNA and incubated overnight at 25°C (Luehrsen et al. 1992). Total RNA was isolated and purified by a cesium chloride step gradient as described previously (Luehrsen and Walbot 1991, Luehrsen et al. 1992). RNA processing outcomes from each expression plasmid were determined by RT–PCR (PCR amplification of first-strand cDNA), Northern blots, and RNase protection.

For Northern blots, 15 μg of total RNA from each expression construct was separated by agarose gel electrophoresis in the presence of formaldehyde and transferred to Hybond-N paper. The size standards are the 0.24- to 9.5-kb and 0.15- to 1.77-kb RNA ladders (BRL). Standard hybridization conditions (Sambrook et al. 1989) were used. The blots were washed in 0.1× SSPE, 0.1% SDS, at 70°C–80°C and autoradiographed with one intensifying screen. The relevant bands were excised from the autoradiograms.

RT–PCR

Purified total RNA was treated with DNase before RT–PCR analysis. One hundred fifty nanograms of antisense oligonucleotide [Fig. 2C, 1365 or 951 poly(T)T] was used to prime cDNA synthesis using 1 μg of total RNA (in a volume of 20 μl) and cDNA was synthesized using either MMLV [BRL] or AMV (Molecular Genetic Resources) reverse transcriptase. The first-strand cDNA was amplified in a 100-μl PCR cocktail containing a second sense primer (1364) and 5 units of AmpliTaq DNA polymerase (Perkin-Elmer). All PCR reactions used 35 rounds of amplification in a Perkin-Elmer DNA Thermal Cycler 480 with the following cycles: 1 min at 95°C, 1 min at 50°C. The amplified cDNAs were cloned into pBSKS(−) and sequenced. Alternatively, the PCR products were displayed on a Southern blot and hybridized with the 972 oligomer probe. To ensure that the recovered cDNAs were genuine, we amplified each plasmid construct using a primer pair that spanned each insertion. In all cases, only one amplified band was observed, indicating that the cDNAs identified by RT–PCR were not the result of fortuitous alterations in Escherichia coli.

Acknowledgments

Research was supported by a grant from the National Institutes of Health (GM32422). We thank Hanya Chrisepels, Jane Hershberger, and Chris Ko for critical comments on the manuscript. The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bennett, M., S. Piñol-Roma, D. Staknis, G. Dreyfuss, and R. Reed. 1992. Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to splicosome assembly in vitro. Mol. Cell. Biol. 12: 3165–3175.

Bodeau, J.P. and V. Walbot. 1992. Regulated transcription of the maize Bronze-2 promoter in electroporated protoplasts requires the C1 and R gene products. Mol. & Gen. Genet. 233: 379–387.

Buchman, A.R. and P. Berg. 1988. Comparison of intron-dependent and intron-independent gene expression. Mol. Cell. Biol. 8: 4395–4405.

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

Csank, C., F.M. Taylor, and D.W. Martindale. 1990. Nuclear pre-mRNA introns: Analysis and comparison of intron sequences from Tetrahymena thermophila and other eukaryotes. Nucleic Acids Res. 18: 5133–5141.

Dennis, E.S., W.L. Gerlach, A.J. Pryor, J.L. Benetzen, A. Inglis, D. Llewellyn, M.M. Sachs, R.J. Ferl, and W.J. Peacock. 1984. Molecular analysis of the Alcohol dehydrogenase [Adh1] gene of maize. Nucleic Acids Res. 12: 3983–4000.

Frohman, M.A., M.K. Dush, and G.A. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. 85: 8998–9002.

Gershon, P.D. and B. Moss. 1993. Uridylate-containing RNA sequences determine specificity for binding and polyadenylation by the catalytic subunit of vaccinia virus poly(A) polymerase. EMBO J. 12: 4705–4714.

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

—. 1990. The minimum functional length of pre-mRNA introns in monocots and dicots. Plant Mol. Biol. 14: 727–733.

—. 1991. Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants. EMBO J. 10: 2635–2644.

Green, M.R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. 7: 559–599.

Guerrineau, F., L. Brooks, and P. Mullineaux. 1991. Effect of deletions in the cauliflower mosaic virus polyadenylation sequence on the choice of the polyadenylation sites in tobacco protoplasts. Mol. & Gen. Gen. 226: 141–144.

1128 GENES & DEVELOPMENT
Hanley, B.A. and M.A. Schuler. 1988. Plant intron sequences: Evidence for distinct groups of introns. *Nucleic Acids Res.* 16: 7159–7176.

Hawkins, J.D. 1988. A survey of intron and exon lengths. *Nucleic Acids Res.* 16: 9893–9908.

Heidmann, S., B. Obermaier, K. Vogel, and H. Domdey. 1992. Identification of pre-mRNA polyadenylation sites in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 12: 4215–4229.

Hernandez, N. and W. Keller. 1983. Splicing of in vitro synthesized messenger RNA precursors in HeLa cell extracts. *Cell* 35: 89–99.

Hunt, A.G. and M.H. MacDonald. 1989. Deletion analysis of the polyadenylation signal of a pea ribulose-1,5-biphosphate carboxylase small-subunit gene. *Plant Mol. Biol.* 13: 125–138.

Hunt, A.G., N.M. Chu, J.T. Odell, F. Nagy, and N.-H. Chua. 1987. Plant cells do not properly recognize animal gene polyadenylation signals. *Plant Mol. Biol.* 8: 23–35.

Iriiger, S., H. Sanfaçon, C.M. Egli, and G.H. Braus. 1992. Different sequence elements are required for function of the Cauliflower Mosaic Virus polyadenylation site in *Saccharomyces cerevisiae* compared with plants. *Mol. Cell. Biol.* 12: 2322–2330.

Joshi, C.P. 1987a. An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Res.* 15: 6643–6653.

——. 1987b. Putative polyadenylation signals in nuclear genes of higher plants: A compilation and analysis. *Nucleic Acids Res.* 15: 9627–9639.

Kim, H.-Y., J.W. Schiefelbein, V. Raboy, D.B. Furtick, and O.E. Nelson Jr. 1987. RNA splicing permits expression of a maize gene with a defective suppressor–mutator transposable element insertion in an exon. *Proc. Natl. Acad. Sci. USA* 84: 5863–5867.

Kraimer, A.R., T. Maniatis, B. Ruskin, and M.R. Green. 1984. Normal and mutant human β-globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell* 36: 993–1005.

Leff, S.E., R.M. Evans, and M.G. Rosenfeld. 1987. Splice commitment dictates neuron-specific alternative RNA processing in calcitonin/CGRP gene expression. *Cell* 48: 517–524.

Lou, H., A.J. McCullough, and M.A. Schuler. 1993a. 3’ splice selection in dicot plant nuclei is position dependent. *Mol. Cell. Biol.* 13: 4485–4493.

——. 1993b. Expression of maize *Adh1* intron mutants in tobacco. *Plant* J. 3: 393–403.

Luehrsren, K.R. and V. Walbot. 1990. Insertion of *Mug* elements in the first intron of the *Adh1* S gene of maize results in novel RNA processing events. *Plant Cell* 2: 1225–1238.

——. 1991. Intron enhancement of gene expression and the splicing efficiency of introns in maize cells. *Mol. & Gen. Genet.* 225: 81–93.

——. 1992. Insertion of non-intron sequence into maize introns interferes with splicing. *Mol. & Gen. Genet.* 225: 81–93.

——. 1994. Addition of *A*- and U-rich sequence increases the splicing efficiency of a deleted form of a maize intron. *Plant Mol. Biol.* 24: 449–463.

Luehrsren, K.R., J.R. de Wet, and V. Walbot. 1992. Transient expression analysis in plants using firefly luciferase reporter gene. *Methods Enzymol.* 216: 397–414.

Luehrsren, K.R., S. Taha, and V. Walbot. 1994. Nuclear pre-mRNA splicing in higher plants. *Prog. Nucleic Acids Res. Mol. Biol.* 47: 149–193.

Lührmann, R., B. Kastner, and M. Bach. 1990. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochem. Biophys. Acta* 1087: 265–292.

MacDonald, M.H., B.D. Mogen, and A.G. Hunt. 1991. Charac-

terization of the polyadenylation signal from the T-DNA-encoded octopine synthase gene. *Nucleic Acids Res.* 19: 5575–5581.

McCullough, A.J., H. Lou, and M.A. Schuler. 1993. Factors affecting authentic 5′ splice site selection in plant nuclei. *Mol. Cell. Biol.* 13: 1323–1331.

McLaughlin, M. and V. Walbot. 1987. Cloning of a mutable *bz2* allele of maize by transposon tagging and differential hybridization. *Genetics* 117: 771–776.

Menssen, A., S. Hohmann, W. Martin, P.S. Schnable, P.A. Petersen, H. Saedler, and A. Giefl. 1990. The *En/Spm* transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a *dSpm* element in the A2 gene. *EMBO J.* 9: 3051–3057.

Mogen, B.D., M.H. MacDonald, R. Graybosch, and A.G. Hunt. 1990. Upstream sequences other than AUAAA are required for efficient messenger RNA 3′-end formation in plants. *Plant Cell* 2: 1261–1272.

Mogen, B.D., M.H. MacDonald, G. Leggewie, and A.G. Hunt. 1992. Several distinct types of sequence elements are required for efficient mRNA 3′ end formation in a pea *rbcS* gene. *Mol. Cell. Biol.* 12: 5406–5414.

Nash, J. and V. Walbot. 1992. *Bronze-2* gene expression and intron splicing patterns in cells and tissues of *Zea mays* L. *Plant Physiol.* 100: 664–671.

Nash, J., K.R. Luehrsren, and V. Walbot. 1990. *Bronze-2* gene of maize: Reconstruction of a wild-type allele and analysis of transcription and splicing. *Plant Cell* 2: 1039–1049.

Ortiz, D.F. and J.N. Strommer. 1990. The *Mu*1 maize transposable element induces tissue-specific aberrant splicing and polyadenylation in two *Adh1* mutants. *Mol. Cell. Biol.* 10: 2090–2095.

Paszewski, J., A. Peterhans, R. Bilang, and W. Filipowicz. 1992. Expression in transgenic tobacco of the bacterial neomycin phosphotransferase gene modified by intron insertions of various sizes. *Plant Mol. Biol.* 19: 825–836.

Peterson, M.L. and R.P. Perry. 1989. The regulated production of μS and μL mRNA is dependent on the relative efficiencies of μS poly[A] site usage and the C4-12 to M1 splice. *Mol. Cell. Biol.* 9: 726–738.

Proudfoot, N. 1991. Poly[A] signals. *Cell* 64: 671–674.

Quigley, F., W. Martin, and R. Cerff. 1988. Intron conservation across the prokaryote-eukaryote boundary: Structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Proc. Natl. Acad. Sci. USA* 85: 2672–2676.

Robberson, B.L., G.J. Cote, and S.M. Berget. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* 10: 84–94.

Rosbash, M. and B. Seraphin. 1988. Intron conservation across the prokaryote-eukaryote boundary: Structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Plant Cell* 2: 1225–1238.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanfaçon, H. and T. Hohn. 1990. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. *Nature* 346: 81–84.

Sanfaçon, H., P. Brodmann, and T. Hohn. 1991. A dissection of the cauliflower mosaic virus polyadenylation signal. *Genes & Dev.* 5: 141–149.

Sanger, F., A.R. Coulson, G.F. Hong, D.F. Hill, and G.B. Petersen. 1987. Nucleotide sequence of bacteriophage λ DNA. *J. Mol. Biol.* 162: 729–773.

Shah, D.M., R.C. Hightower, and R.B. Meagher. 1983. Genes encoding actin in higher plants: Intron positions are highly

GENES & DEVELOPMENT 1129
conserved but the coding sequences are not. J. Mol. Appl. Genet. 2: 111–126.
Simpson, C.G. and J.W.S. Brown. 1993. Efficient splicing of an AU-rich antisense intron sequence. Plant Mol. Biol. 21: 205–211.
Smith, M.W. 1988. Structure of vertebrate genes: A statistical analysis implicating selection. J. Mol. Evol. 27: 45–55.
Takagaki, Y., L.C. Ryner, and J.L. Manley. 1988. Separation and characterization of a poly(A) polymerase and cleavage/specificity factor required for pre-mRNA polyadenylation. Cell 52: 731–742.
Varagona, M.J., M. Purugganan, and S.R. Wessler. 1992. Alternative splicing induced by insertion of retrotransposons in the maize waxy gene. Plant Cell 4: 811–820.
Waigmann, E. and A. Barta. 1992. Processing of chimeric introns in dicot plants: Evidence for a close cooperation between 5' and 3' splice sites. Nucleic Acids Res. 20: 75–81.
Weil, C.F. and S.R. Wessler. 1990. The effects of plant transposable element insertion on transcription initiation and RNA processing. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41: 527–552.
Wessler, S.R. 1991. The maize transposable Ds1 element is alternatively spliced from exon sequences. Mol. Cell. Biol. 11: 6192–6196.
Wessler, S.R., G. Baran, and M. Varagona. 1987. The maize transposable element Ds is spliced from RNA. Science 237: 916–918.
White, O., C. Soderlund, P. Shanmugan, and C. Fields. 1992. Information contents and dinucleotide compositions of plant intron sequences vary with evolutionary origin. Plant Mol. Biol. 19: 1057–1064.
Wiebauer, K., J.I. Herrero, and W. Filipowicz. 1988. Nuclear pre-mRNA processing in plants: Distinct modes of 3'-splice-site selection in plants and animals. Mol. Cell. Biol. 8: 2042–2051.
Wu, L., T. Ueda, and J. Messing. 1993. 3'-end processing of the maize 27 kDa zein mRNA. Plant J. 4: 535–544.
Intron creation and polyadenylation in maize are directed by AU-rich RNA.

K R Luehrsen and V Walbot

Genes Dev. 1994, 8:
Access the most recent version at doi:10.1101/gad.8.9.1117

References
This article cites 64 articles, 24 of which can be accessed free at: http://genesdev.cshlp.org/content/8/9/1117.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

Copyright © Cold Spring Harbor Laboratory Press