Plant Molecular Biology **41:** 637–644, 1999.

Test of the combinatorial model of intron recognition in a native maize gene

Maita J. Latijnhouwers[1,3], Claudio F. Pairoba[1], Volker Brendel[2,4], Virginia Walbot[1] and Jose C. Carle-Urioste[1,5,*]

[1]*Department of Biological Sciences and* [2]*Department of Mathematics, Stanford University, Stanford, CA 94305-5020, USA; present addresses:* [3]*Stadhouderslaan 47, 3583 JD Utrecht, Netherlands;* [4]*Department of Zoology and Genetics, 2112 Molecular Biology Building, Iowa State University, Ames, IA 50011-3260, USA;* [5]*Hyseq Inc., 670 Almanor Avenue, Sunnyvale, CA 94086, USA (*author for correspondence)*

Received 7 June 1999; accepted in revised form 8 September 1999

Key words: Bronze2, exon, G+C content, RNA processing, splicing

Abstract

Previous studies have established that splice site selection and splicing efficiency in plants depend strongly on local compositional contrast consisting of high exon G+C content relative to high intron U content. The combinatorial model of plant intron recognition posits that splice site sequences as well as local intron and exon sequences contribute to splice site selection and splicing efficiency. Most of the previous studies used synthetic or chimeric constructs, often tested in heterologous hosts. To perform a more critical test of the combinatorial model in a native context, the single intron of the maize *Bronze2* gene and its flanking exons were modified by site-directed mutagenesis. Splicing efficiency was tested in maize protoplasts. Results show that a higher U content in the flanking 5′ exon, whether close to or distant from the 5′ splice site, did not modify splicing efficiency. Decreasing exon G+C content dramatically impaired splicing. Increasing intron G+C content or decreasing intron U content adversely impacted splicing. In all constructs splicing occurred exclusively at the original 5′ and 3′ splice sites. These results are consistent with the hypothesis that exon G+C content and intron U content contribute separate but complementary aspects of intron definition in the native *Bz2* transcript.

Introduction

Efficient and accurate removal of non-coding sequences from pre-mRNA transcripts is a critical step in message maturation. Required *cis*-acting factors include transcript motifs (such as the semi-conserved 5′ and 3′ splice sites) recognized by *trans*-acting factors (proteins, small nuclear RNAs and ribonucleoproteins). Interaction between *cis*- and *trans*-acting factors leads to intron removal after two transesterification reactions (Simpson and Filipowicz, 1996; Brown and Simpson, 1998).

Plants appear to have some unique requirements for intron recognition. Plant introns are usually spliced in mammals, indicating that they typically contain all of the *cis* factors for recognition, but mammalian introns are often not or incorrectly spliced in plant cells (Luehrsen *et al.*, 1994; Simpson and Filipowicz, 1996). Mammalian genes generally do not display a consistent difference in base composition between exons and introns, but this is characteristic of plant genes (Filipowicz *et al.*, 1994; Luehrsen *et al.*, 1994). Many experiments with native, site-mutated and synthetic introns established that introns must be AU-rich for efficient splicing in dicots (Goodall and Filipowicz, 1989, 1991). In monocots, G+C-rich introns are accurately and efficiently spliced, suggesting that AU-richness is not an absolute requirement (Goodall and Filipowicz, 1991).

More recent statistical analysis of *Arabidopsis* and maize established that individual exons are typically 15% more G+C-rich than their flanking intron and conversely that introns are typically 15% more U-

rich than the flanking exons; there is no difference in A content (Brendel *et al.*, 1998a). Experimental evidence also indicates that it is the relative U-richness of introns compared to the flanking exons that allows G+C-rich introns to be properly spliced (Carle-Urioste *et al.*, 1994; Ko *et al.*, 1998).

Carle-Urioste *et al.* (1994, 1997) demonstrated that exon sequences have a specific role in splicing in maize and that high G+C content *per se* cannot account for differences in splicing efficiency. It has been proposed that a combination of local features is required for plant intron definition (McCullough *et al.*, 1993; Filipowicz *et al.*, 1994; Luehrsen *et al.*, 1994). It was shown that for about 90% of native splice sites the local site context of about 50 bases upstream and downstream is sufficient for highly accurate computer-aided splice site prediction (Brendel *et al.*, 1998b).

Studies of splicing in plants have faced two main problems. First, there is no *in vitro* splicing assay. Second, virtually all data derive from artificial constructs, often expressed in heterologous hosts. For example, the combinatorial model hypothesis was based on analysis of the maize *Bz2* intron in chimeric reporter gene constructs (Carle-Urioste *et al.*, 1997; Brendel *et al.*, 1998a). Given the likely participation of both intronic and exonic elements in intron recognition and splice site selection in plants, it is important to evaluate the relative contributions of elements within the context of a specific intron and its exons. The aim of the present work was to test the contribution of compositional contrast to intron recognition with the native maize *Bz2* gene expressed in maize cells.

Materials and methods

Plasmid constructs

Constructs as shown in Figure 1 were derived from pCa*Bz2*i (Luehren and Walbot, 1994). pCa*Bz2*i contains the complete *Bz2* coding region (including the intron) with a poly(A) sequence under control of the cauliflower mosaic virus (CaMV) 35S promoter in pBluescript (pBS KS+). pCa*Bz2*i was digested with *Stu*I, blunt-ended with Mung Bean Nuclease (Pharmacia) and subsequently digested with *Sst*I. The released 481 bp fragment containing exon 1, the intron and 77 bases of exon 2 was cloned into pBS KS+ cut with *Sst*I and *Sal*I. The resulting subclone of pCa*Bz2*i was designated pMA2.

Base changes were made in pMA2 by site-directed mutagenesis and were verified by sequencing. To clone the mutated fragment back into pCa*Bz2*i, both the mutated plasmid pMA2 and pCa*Bz2*i were digested with *Sst*I and *Sst*II. The *Sst*I/*Sst*II fragment from pMA2 was cloned into the *Sst*I and *Sst*II sites of pCa*Bz2*i, resulting in the piGC, ppA2T, puA2T, ppGC2T, puGC2T, ppGC2A and puGC2A series of vectors (Table 1).

In puA2T, the 9 adenosines present between positions 926 and 976 in exon 1 (see Figure 1) were replaced by thymidines, thereby increasing the U content of this window without changing the G+C content. In ppA2T, the 9 adenosines present between positions 992 and 1042 of exon 1, were replaced by thymidines, which again increases the %U in this window, leaving the %G+C untouched. In both constructs all codon changes were non-synonymous. In piGC, the 8 adenosines present between positions 1052 and 1073 within the intron were replaced by 4 Gs and 4 Cs, a change that increases the G+C content of the intron without changing the U content. The same base changes as in ppA2T and puA2T were made in piGC, resulting in piGCuA2T and piGCpA2T, respectively.

In p3′, the 3′ splice site of the *Bz2* intron was changed from CAG/TTT to CAT/TTT. In pNU, the T-stretch in the *Bz2* intron (TGTCTTTTCTT) was replaced by a G+C-rich sequence (CGCCGCGCCGG). In pU2A, the stretch was disrupted by replacing four Ts with As (TGACTAATCAT). In ppGC2A, 11 Cs of Gs present in third-codon positions between positions 992 and 1042 were changed to As. These substitutions resulted in only three, conservative amino acid replacements (D to E). In puGC2T, 10 Gs or Cs present in third-codon positions between positions 926 and 976 were changed to Ts. These substitutions resulted in no amino acid changes. In puGC2A, 10 Cs or Gs in third-codon positions were changed to As between positions 926 and 976. The amino acid sequence remained unchanged. Each of these substitutions was also made in piGC, resulting in piGCpGC2A, piGCpGC2T, piGCuGC2A and piGCuGC2T.

DNA sequencing

DNA sequencing was performed using the chain-termination method with T7 DNA polymerase (United States Biochemicals). The polymerase reaction was carried out at room temperature for 5 min. The chain termination reaction was carried out at 37 °C for

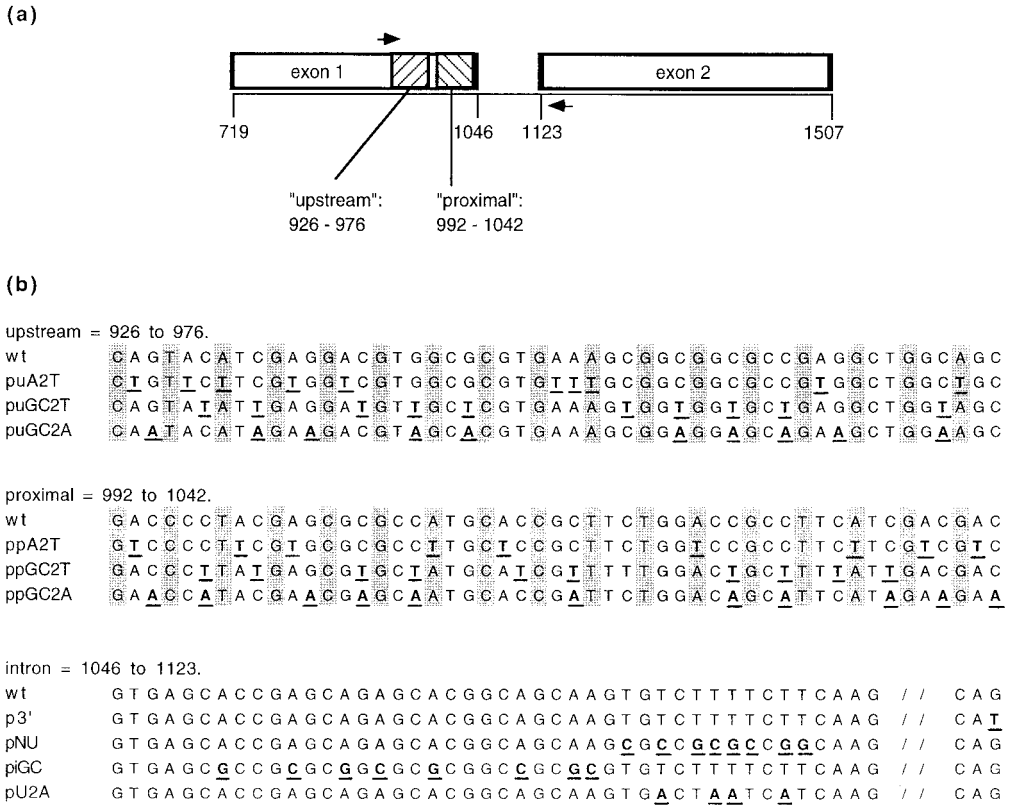

Figure 1. Schematic display of the *Bronze2* gene and the mutated regions. Numbering is according to GenBank ZMU14599 (accession number U14599). The coding parts of exons 1 and 2 are bold-face-boxed. Mutations within exon 1 were made in two regions of 17 codons each. The 'upstream' region extends from 926 to 976, and the 'proximal' region extends from 992 to 1042. The mutant constructs are displayed in the lower part of the figure. The first position of a codon is displayed on shaded background. Mutations relative to the wild type (wt) are indicated in underlined bold face. An unmutated region in the intron is displayed by //. The mutations are identified by the labels given in the first column. Double mutants are labeled by double identifiers; e.g. iGCpA2T refers to the construct consisting of pA2T in the proximal region and iGC in the intron.

5 min. The fragments were separated on 5% polyacrylamide, 7 M urea gel. X-Omat film (Kodak) was exposed overnight to obtain an autoradiograph.

Transient assays

Black Mexican Sweet (BMS) suspension-cultured cells were used for protoplast preparation and electroporation as described (Carle-Urioste *et al.*, 1994). Assays were performed in triplicate within an experiment, and experiments were repeated three times.

RNA extraction and analysis

Total RNA was extracted using the RNeasy kit (Qiagen) and treated twice with DNaseI (Ambion) at 37 °C for 1 h. For RT-PCR analysis first-strand cDNA was synthesized with 100 ng of primer 5′-CGGCGGCAGGGAGACGGCA-3′ complementary to exon 2 between positions 1141 and 1159. Reverse transcription and PCR amplification were done as previously described (Carle-Urioste *et al.*, 1994). The PCR products were analyzed on 2.5% agarose gels (FMC) and transferred onto nylon membrane (Hybond N^+). Hybridization and washing of the membranes were performed as previously described (Carle-Urioste *et al.*, 1994). The splicing efficiency was quantified on an autoradiograph using a BioRad Phosphorimager.

A search for alternative 5′ and 3′ splicing sites was performed using primers 1250 (5′-CCCTGCCGTTG-CTGCGGTCCT-3′) and 26 (5′-CAGAAGCTGATC-AGCGAGGAGGACC-3′), which hybridize at 1230–1250 and 728–750, respectively.

For RNase protection analysis, a probe was prepared from pMA3, which contains the last 58 bases of the 78 bp *Bz2* intron and the first 236 bp of exon 2 (see

Table 1. Compositional analysis, quality of 5′ donor site context and splicing efficiency of constructs.

Construct	GC content (%)			U content (%)			Donor	Splicing efficiency (%)
	upstream	proximal	intron	upstream	proximal	intron	*P* value[a]	RT-PCR ± SD
Experiment 1								
wt	71	67	54	10	16	22	0.82	82.4 ± 8.5
puA2T[b]	71	67	54	**29**	16	22	0.82	80.6 ± 6.1
ppA2T	71	67	54	10	**33**	22	0.54	81.6 ± 11
piGC	71	67	**64**	10	16	22	0.47	49.2 ± 4.0
piGCuA2T	71	67	**64**	**29**	16	22	0.47	52.1 ± 3.4
piGCpA2T	71	67	**64**	10	**33**	22	0.19	58.9 ± 4.0
pNU	71	67	**64**	10	16	**12**	0.19	32.5 ± 13
p3′	71	67	53	10	16	23	0	1.7 ± 2.9
Experiment 2								
wt	71	67	54	10	16	22	0.82	82.7 ± 2.0
ppGC2A	71	**45**	54	10	16	22	0.37	20.0 ± 8.0
puGC2T	**51**	67	54	**29**	16	22	0.82	50.5 ± 7.0
puGC2A	**51**	67	54	10	16	22	0.82	59.3 ± 6.0
ppT2A	71	67	54	10	16	**17**	0.70	58.3 ± 3.0
piGCpGC2T	71	**45**	**64**	10	**37**	22	0.02	19.0 ± 4.0
piGCpGC2A	71	**45**	**64**	10	16	22	0.10	22.2 ± 5.0
piGCuGC2T	**51**	67	**64**	**29**	16	22	0.47	35.3 ± 5.0
piGCuGC2A	**51**	67	**64**	10	16	22	0.47	40.5 ± 2.0

SD: standard deviation.
Numbers in bold characters: G+C or U content after changes described in the text.
[a] *P* value is calculated by the SplicePredictor program (Kleffe *et al.*, 1996; http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
[b] Code for plasmid names: puA2T, upsteam A to T; ppA2T, proximal A to T; piGC, intron GC; pNU, G for U in intron U-rich motif; ppGC2A, proximal GC to A; puGC2T, upstream GC to T; puGC2A, upstream GC to A; ppT2A, proximal T to A; ppGC2T, proximal GC to T.

Figure 1). A second probe was prepared from plasmid Ubi95, which is complementary to a 95 bp fragment of maize ubiquitin mRNA (Christensen *et al.*, 1992). Probes pMA3 and Ubi95 were prepared by *in vitro* transcription using T3 polymerase and T7 polymerase (Gibco-BRL), respectively. The *in vitro* transcription was carried out in a buffer containing 40 mM Tris-HCl pH 8.0, 8 mM $MgCl_2$, 2 mM spermidine, 25 mM NaCl, 4 U RNase inhibitor, 0.5 mM each of ATP plus UTP and GTP, 25 μM CTP and 60 μCi ^{32}P-CTP as the label. The probe was separated from unincorporated nucleotides on a gel filtration column (Stratagene) in STE buffer.

About 10 μg of RNA and 1×10^5 cpm ^{32}P-labeled probe were co-precipitated in 70% ethanol. The pellet was resuspended in 10 μl HybSpeed buffer (Ambion). Hybridizations, RNase digestion, separation and detection of protected fragments were performed as recommended (HybSpeed RPA manual, Ambion).

Donor site scoring

Predicted donor site strength was calculated as a *P* value as in Kleffe *et al.* (1996). The *P* value was shown to correlate positively with experimentally measured splicing efficiencies for chimeric *Bz2* intron constructs (Brendel *et al.*, 1998a).

Results

Combinatorial model of intron recognition

We define base compositional contrast as [(% exon G+C) − (% intron G+C)] + [(% intron U) − (% exon U)] in the 50 bases of exon and 50 bases of intron flanking splice sites; the absolute base composition is defined as the total G+C content in exons or as the total U content in introns. In the combinatorial model of intron recognition we propose that the compositional contrast is one of several features crucial to intron recognition. From this model, we would predict

the following results when changes are introduced in exon or intron sequences:

1 Increasing the U content in the exons while keeping the exon G+C content constant will decrease splicing efficiency.

2 The impact of exon changes on splicing will be significantly higher for alterations within 50 bases of the splice sites than for alterations further away.
3 Increasing the G+C content in the intron but keeping the intron U content constant will decrease splicing efficiency.
4 Decreasing the G+C content in the exons will decrease splicing efficiency.
5 Decreasing the U content in the intron will decrease splicing efficiency.

We tested these predictions in the following set of experiments.

Effect of A-to-U changes in the exon and A-to-G+C changes in the intron

To test predictions 1 and 2 we constructed ppA2T and puA2T. In ppA2T, nine As in a 51 base exon 1 fragment proximal to the 5′ splice site were replaced by Ts. In puA2T, ten As were replaced by Ts in the 51 base exon 1 fragment located between 70 and 120 bases upstream of the 5′ splice site (Figure 1 and Materials and methods). The U content of transcripts of the 51 base fragment proximal or distal to the 5′ splice site was increased from 16% to 33% in ppA2T and from 10% to 29% in puA2T. The absolute G+C content remained constant in both cases. The results obtained by either RT-PCR or RNase protection (Tables 1 and 2) showed that splicing efficiency was not affected by these changes. This is not consistent with our prediction of a decrease in splicing efficiency for the proximal construct (with a decreased donor splice site score, Table 1).

Three constructs were made to test predictions 2 and 3: piGC, piGCuA2T and piGCpA2T. In piGC eight As within the intron were changed to Gs or Cs, increasing the G+C content of the intron from 54% to 64%. piGCuA2T and piGCpA2T were generated by combining the changes between piGC and puA2T (see Materials and methods). As expected, splicing efficiency decreased significantly (from about 80% to 50–60%, Table 1) when the G+C content of the intron increased from 54% to 64%. Again, no changes were detected in the splicing efficiency when As in the exon were changed to Us, in either the proximal or the upstream substitution.

Table 2. Comparison of splicing efficiency measured by two methods.

Construct	Splicing efficiency	
	RT-PCR	RNase protection assay
wt	82.4 ± 8.5	84.8 ± 1.0
puA2T	80.6 ± 6.1	81.0 ± 5.8
ppA2T	81.6 ± 11.0	80.1 ± 3.0
piGC	49.2 ± 4.0	50.6 ± 3.8
piGCuA2T	52.1 ± 3.4	51.1 ± 6.7
piGCpA2T	58.9 ± 4.0	55.4 ± 7.5
pNU	32.5 ± 13.0	35.7 ± 1.0

Effect of G+C-to-U and G+C-to-A changes in the exon

In a second set of experiments we tested predictions 2 and 4 by changing Gs and Cs to As or Gs and Cs to Us in pCa*Bz2*i and in piGC (Figure 1 and Materials and methods) to generate ppGC2A, puGC2A, puGC2T, piGCpGC2A, piGCuGC2A, piGCpU2A and piGCuU2A. In these seven constructs, the bases were changed in exon 1 in either the proximal or distal region. In the upstream constructs the G+C content of the 51 base exon 1 fragment descreased from 71% to 51%. In the proximal constructs the G+C content of the 51 base exon 1 fragment decreased from 67% to 45%. For these constructs, the nucleotide changes preserved the predicted translation product apart from conservative aspartate-to-glutamate substitutions. The results (Figure 2B, Table 1) show that splicing efficiency is affected when Gs and Cs are exchanged for As or Us, fulfilling prediction 4. Furthermore, the effect depends on where the changes were made, supporting prediction 2: the deleterious effect of changing base composition is significantly stronger in the proximal than in the upstream constructs.

Effect of U-to-A and U-to-G+C changes within the intron and base changes in the 3′ splice site

To test prediction 5 we made two constructs. In one of them, all eight Us of the U-rich tract within the intron were changed to Gs or Cs (pNU), and in the other construct four Us were changed to As (pU2A). Splicing effiency decreased in both cases (Table 1), and the

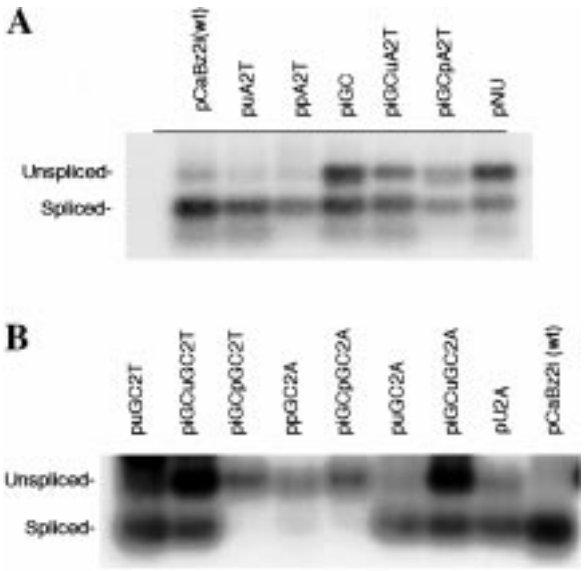

Figure 2. RT-PCR analysis of splicing of constructs with intron and exon alterations. After electroporation with the indicated constructs, total RNA was isolated from maize protoplasts and analyzed by RT-PCR. The products were blotted onto a nylon membrane and probed with an oligonucleotide complementary to exon sequences as described in Materials and methods. A. Splicing profile of constructs with changes in the intron and/or A-to-T changes in the exon. B. Splicing profile of constructs with changes in the intron and/or GC-to-A(T) changes in the exon.

effect was stronger for pNU. These results corroborate prediction 5.

The AG dinucleotide is a nearly absolute requirement for 3′ splice sites in plants (Simpson and Filipowicz, 1996). The question is what makes a particular AG the chosen 3′ splice site when proximal or distal AGs could, in theory, be selected. In the *Bz2* gene there are three proximal AG dinucleotides at 45, 55 and 65 bases downstream of the native *Bz2* 3′ splice site that have never been observed as alternative splice sites. We reasoned that by eliminating the real 3′ splice site, one of those three AG (or more than one) could be used. To test this idea we changed the G of the 3′ splice site to U. Neither RNase protection experiments nor RT-PCR detected alternative splicing events. These results indicate that the context is critical to define a 3′ splice site. In this case the three AGs are embedded in GC-rich sequences and no branchpoint-like motif is found nearby.

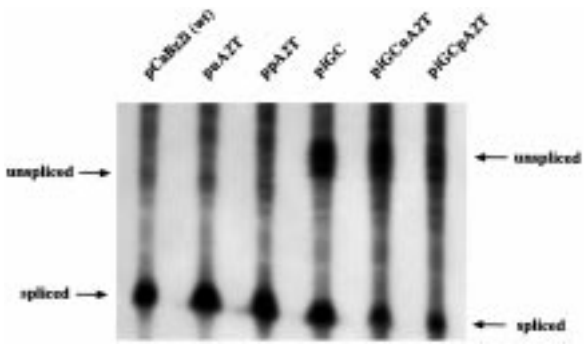

Figure 3. RNase protection analysis of constructs with changes in the intron and/or the exon. After electroporation with the indicated constructs, total RNA was isolated from maize protoplasts and analyzed by RNase protection as described in Materials and methods.

Comparison of RNase protection and RT-PCR assays for unspliced transcripts

Because no alternative splice sites have ever been detected in *Bz2* transcripts, we wished to determine if RNase protection and RT-PCR could detect similar contributions of spliced and unspliced transcripts. For a subset of vectors (Figure 3; Table 2) the two methods were compared. The two methods yield highly similar results for vectors with highly efficient or poor splicing. Although the standard deviation is lower for RNase protection assays, both methods are reproducible and appropriate to test transcripts with a simple splicing pattern.

Discussion

The principles that govern intron splicing in plant nuclei have been studied for a decade, and the available information suggests that intron and exon features as well as splice sites and a branchpoint motif are crucial for intron definition in plants. These general rules for splicing are based mainly on experiments done using artificial constructs, usually expressed in heterologous hosts. Either authentic plant introns were placed in constructs containing reporter genes or artificial introns were inserted in plant sequences. The aim of the present work was to test two principles of intron splicing in the native environment of both the gene (maize *Bz2* gene in maize nuclei) and the intron (*Bz2* intron flanked by *Bz2* exons). One of the principles is that introns are defined by U-rich sequences flanked by G+C-rich exons. The second principle derives from the first and states that splice site definition is deter-

mined mostly by local contrast or, in other words, that splice site locations coincide with 'transition points' from G+C-rich to U-rich sequences. The questions we asked were, first, is it the contrast between the intron and its flanking exons that affects splicing decisions or is it the absolute base composition of intron and exons? Second, what is the impact of local sequences vs. 'distant' sequences in splice site definition? These questions were formulated into five predictions that we tested experimentally.

In ppA2T and puA2T we changed As to Us in the first exon, maintaining the absolute G+C content and splicing efficiency was identical to the native gene. In piGC, for which As were changed to Gs and Cs within the intron maintaining the absolute U content in the intron, splicing efficiency was about 50% of wild type. Combining these individual alterations in piGCpA2T and piGCuA2T showed no additive effect: splicing remained about 50% of wild type. These results contradict prediction 1 of the model in that increasing U content in the exon did not alter splicing. At the same time the results are in agreement with prediction 3 of the model: increasing the G+C content of the intron is deleterious to splicing. Similar conclusions were obtained with artificial constructs in our previous work (Carle-Urioste *et al.*, 1997).

To address the question of the G+C content in the first exon and the importance of local context in splicing more specifically, we designed a second set of experiments. We changed Gs and Cs for Us in construct ppuGC2T and Gs and Cs for As in constructs ppGC2A and puGC2A. In both cases the changes were made either close to the 5′ splice site (pp constructs) or 70 bases upstream of the 5′ splice site (pu constructs). The idea was to change the absolute G+C base composition in the first exon and separately analyze the impact of increasing the U content or the A content farther upstream compared to changes near the 5′ splice site. The results show that changes of Cs and Gs for either Us or As in the first exon affected splicing, as predicted. This effect is stronger when the modifications take place in the proximal region (Table 1). These experiments verify predictions 2 and 4 and are consistent with previous observations that G+C sequences in the exon contribute to efficient splicing. They also provide clear evidence that local sequences have a greater impact than upstream sequences on efficient splicing.

We also tested prediction 5 by investigating the impact of U-to-A changes and U-to-G or U-to-C changes in the U-rich tract within the intron. In one construct, four out of eight Us in the U-rich intron motif were changed to As (pU2A), and in another construct eight Us were changed to Gs or Cs (pNU). The results showed that splicing efficiency decreases significantly relative to wild type in both cases, consistent with previous work that established that the U tract within the *Bz2* intron constitutes a specific and important motif for intron definition (Ko *et al.*, 1998).

The combinatorial model of intron recognition is generally confirmed by these results examining the splicing efficiency of a native maize gene. The present results highlight, however, the importance of G+C content (rather than the absence of Us) in the exon and the importance of a U-rich motif (rather than just the absence of G+Cs) in the intron. Because the G+C exon and U intron motifs contribute directly it seems highly likely that they are recognized by *trans*-acting factors. The SR proteins interact with G+C- and GA-rich exon splicing enhancers in animals; the enhancer motifs contain one or more SR binding sites (Manley and Tacke, 1996) and serve to activate splicing by recruiting the splicing machinery to the adjacent intron through protein-protein interactions (Hertel and Maniatis, 1998). It is plausible that plant SR proteins recognize G+C-rich exon motifs to perform a parallel role in splicing.

We propose that compositional contrast reflects natural selection to insure that G+C-rich motifs are present in exons to bind exon-recognition proteins and that U-rich motifs are present in maize introns to bind intron-recognition proteins. Thus, the compositional contrast reflects the mechanism of intron and exon definition and may be a guide to defining the range of exon and intron motifs that are used *in vivo*.

Acknowledgements

Research was financed by a grant from the National Science Foundation IBN96-03927 to V.W. V.B. was supported in part by the National Institutes of Health grant 2R01HG00335-O9.

References

Brendel, V., Carle-Urioste, J.C. and Walbot, V. 1998a. Intron recognition in plants. In: J. Bailey-Serres and D.R. Gallie (Eds.), A Look beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants, American Society of Plant Physiologists, Rockville, MD. pp. 20–28.

Brendel, V., Kleffe, J., Carle-Urioste, J.C. and Walbot, V. 1998b. Prediction of splice sites in plant pre-mRNA from sequence properties. J. Mol. Biol. 276: 85–104.

Brown, J.W.S. and Simpson, C.G. 1998. Splice site selection in plant pre-mRNA splicing. Annu. Rev. Plant Plysiol. Plant Mol. Biol. 49: 77–95.

Carle-Urioste, J.C., Ko, C., Benito, M.-I. and Walbot, V. 1994. *In vivo* analysis of intron processing using splicing-dependent reporter gene assays. Plant Mol. Biol. 26: 1785–1795.

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

Christensen, A.H., Sharrock, R.A. and Quail, P.H. 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol. Biol. 18: 675–689.

Filipowicz, W., Gniadkowski, M., Klahre, U. and Hong-Xiang, L. 1994. Pre-mRNA splicing in plants. In: A.I. Lamond (Ed.), Pre-mRNA Processing, R.G. Landes Company, Georgetown, TX, pp. 65–77.

Goodall, G.J. and 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, G.J. and 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.

Hertel, K.J. and Maniatis, T. 1998. The function of multisite splicing enhancers. Mol. Cell 1: 449–455.

Kleffe, J., Hermann, K., Vahrson, W., Wittig, B. and Brendel, V. 1996. Logitlinear models for the prediction of splice sites in plant pre-mRNA sequences. Nucl. Acids Res. 24: 4709–1718.

Ko, C.H., Brendel, V., Taylor, R.D. and Walbot, V. 1998. U-richness is a defining feature of plant introns and may function as an intron recognition signal in maize. Plant Mol. Biol. 36: 573–583.

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

Luehrsen, K.R., Taha, S. and Walbot, V. 1994. Nuclear pre-mRNA processing in higher plants. Prog. Nucl. Acid Res. Mol. Biol. 47: 149–193.

Manley, J.L. and Tacke, R. 1996. SR proteins and splicing control. Genes Dev. 10: 1569–1579.

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

Simpson, G.G. and Filipowicz, W. 1996. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organisation of the spliceosomal machinery. Plant Mol. Biol. 32: 1–41.

Simpson, C.G., McQuade, C., Lyon, J. and Brown, J.W.S. 1998. Characterization of exon skipping mutants of the *COP1* gene from *Arabidopsis*. Plant J. 15: 125–131.