

Retrotransposon Insertion into the Maize *waxy* Gene Results in Tissue-Specific RNA Processing

Sylvestre Marillonnet¹ and Susan R. Wessler²

Departments of Botany and Genetics, University of Georgia, Athens, Georgia 30602

We previously reported that three alleles of the maize *waxy* (*wx*) gene were alternatively spliced as a result of the insertion of retrotransposons into intronic sequences. In addition, inefficient splicing of element sequences with the surrounding intron produced wild-type transcripts that presumably were responsible for the observed residual gene expression in the endosperm. In this study, we report that one of these alleles, *wxG*, has a tissue-specific phenotype with 30-fold more *WX* enzymatic activity in pollen than in the endosperm. Quantification of *wxG*-encoded transcripts in pollen and the endosperm demonstrates that this difference can be accounted for by tissue-specific differences in RNA processing. Specifically, there is ~30-fold more correctly spliced RNA in pollen than in the endosperm. Based on an analogy to similar examples of tissue-specific alternative splicing in animal systems, we hypothesize that the tissue-specific phenotype of the *wxG* allele may reflect differences in the concentration of splicing factors in these tissues.

INTRODUCTION

Most plant genomes are populated with enormous quantities of transposable elements (TEs) or sequences derived from TEs. In large genomes such as maize, most of these sequences reside in the vast expanses between genes (Springer et al., 1994; SanMiguel et al., 1996). In addition, element-derived sequences have recently been identified in hundreds of normal genes from both monocot and dicot species (reviewed in Wessler et al., 1995). In several instances, sequences involved in gene regulation, including the start and stop sites of transcription and the binding sites for regulatory proteins, have been derived from TEs (Bureau and Wessler, 1992, 1994; White et al., 1994). Despite these examples, the contribution of TEs to normal gene expression remains uncertain because it is not known whether the insertion of any single element altered the preexisting pattern of gene expression.

One way to determine whether TEs have the potential to alter gene regulation is by studying the expression of TE-induced mutant alleles. There are numerous examples of the effect of the insertion of class 2 or DNA elements on gene regulation (reviewed in Weil and Wessler, 1990). Class 2 elements in maize include members of the *Activator/Dissociation* (*Ac/Ds*), *Suppressor-mutator/Enhancer* (*Spm/En*), and *Mutator* (*Mu*) families. Insertion of class 2 elements into 5' flanking sequences can put the adjacent gene under element control

by replacing endogenous regulatory sequences with element-encoded promoters and/or regulatory sequences (Masson et al., 1987; Barkan and Martienssen, 1991). Excision of most class 2 elements usually changes the insertion site by leaving behind a few nucleotides derived from the direct repeat of host sequences (called transposon footprints). Excision can also be accompanied by inversions or deletions of the flanking gene sequences. Such changes may be neutral, or they can have profound effects on the transcriptional regulation of the downstream gene (Coen et al., 1986; Kloeckener-Gruissem and Freeling, 1995). Similarly, TE insertion into and excision from coding sequences can alter both the quality of the protein and its regulation. For example, insertion of *Ds* and *defective Spm* elements can dramatically change splicing patterns (Kim et al., 1987; Wessler et al., 1987) or lead to the synthesis of truncated proteins with altered cellular localization (Liu et al., 1996).

Much less is known about the ability of plant retrotransposons to alter gene expression. Retrotransposons are class 1 or retroelements that transpose via an RNA intermediate. Recent studies suggest that retrotransposons are ubiquitous in plants and may be the most abundant TE class in most plant genomes (Joseph et al., 1990; Flavell et al., 1992; SanMiguel et al., 1996). Not only do they comprise a large fraction of repetitive, intergenic sequences, but they have been shown to be a causative agent for spontaneous mutations in maize (Johns et al., 1985; Wessler and Varagona, 1985; Varagona et al., 1992; Purugganan and Wessler, 1994; White et al., 1994; Vignols et al., 1995) and for mutations attributed to somaclonal variation in rice (Hirochika et al., 1996).

¹Current address: The Sainsbury Laboratory, Norwich Research Park, Colney, Norwich NR4 7UH, UK.

²To whom correspondence should be addressed at Department of Genetics, University of Georgia, Athens, GA 30602. E-mail sue@dogwood.botany.uga.edu; fax 706-542-3910.

The ability of retrotransposons to alter the regulation of genes has been well documented in *Drosophila*, where these elements are responsible for the vast majority of spontaneous mutations that result in a discernible phenotype in this genus (Green, 1988). In *Drosophila*, retrotransposon sequences have been shown to be spliced from pre-mRNA (Fridell et al., 1990) and to be influenced by *trans*-acting factors that can either enhance or suppress element-induced mutant phenotypes (Parkhurst and Corces, 1985; Kubli, 1986). Some of these *trans*-acting factors encode proteins that facilitate the transcription or splicing of gene-promoted mRNAs containing element sequences (reviewed in Arkhipova et al., 1995).

The *waxy* (*wx*) gene of maize provides an excellent system for studying the effect of retrotransposon insertion on gene expression. *Wx* encodes a starch granule-bound ADP-glucosyl transferase that is required for amylose biosynthesis in endosperm and pollen. The viable *wx* phenotype is easily scored by staining either endosperm or pollen with I/KI, which turns black in the presence of amylose.

Retrotransposons are responsible for five of six characterized stable spontaneous insertion mutations of the maize *wx* gene (Wessler and Varagona, 1985; Varagona et al., 1992; Purugganan and Wessler, 1994; White et al., 1994). Surprisingly, three of these five alleles (*wxStonor*, *wxG*, and *wxB5*) encode some *Wx* enzymatic activity, despite containing elements of 4.5, 5.6, and 6.1 kb in three different *wx* introns (Varagona et al., 1992). For all three alleles, leaky expression in the endosperm resulted from the synthesis of wild-type transcripts after the splicing of element sequences with the surrounding intron. Although each allele generated wild-type *Wx* mRNA, all three were also shown to encode alternatively spliced transcripts in which exons around the element insertion sites were skipped (Varagona et al., 1992). In addition, tissue-specific differences in *wx* expression in endosperm and pollen were suggested.

In this study, we have followed up this latter observation and show that the *wxG* allele contains 30-fold more activity in pollen than in the endosperm. Quantification of enzymatic activity and *wx* transcription in both tissues indicates that this difference can be accounted for by tissue-specific differences in splicing and 3' end formation. As such, the *wxG* allele represents one of only a very few examples of tissue-specific splicing in plants and is a novel example of a tissue-specific phenotype caused by retrotransposon insertion into the coding region of a gene.

RESULTS

Tissue-Specific Expression of the *wxG* Allele

Nelson (1968) found that pollen grains from plants homozygous for the *wxG* allele were indistinguishable from wild-type grains when stained for *wx* expression with I/KI. This

result is surprising because the *wxG* endosperm was found to encode <5% of wild-type *WX* enzymatic activity (Varagona et al., 1992). To determine the extent of *wxG* tissue-specific expression, *WX* enzymatic activity was quantified from starch granules isolated from both endosperm and pollen of strains containing *Wx*, *wxG*, and *wxC34* (*wxΔ*) alleles. In addition, the enzymatic activity of two other *wx* alleles that also contain retrotransposon insertions in *wx* introns (*wxStonor* and *wxB5*) was measured. As summarized in Table 1, *wxG* pollen has almost 30-fold more *WX* enzymatic activity than does the *wxG* endosperm (27 versus 0.9% that of wild-type levels, respectively). In contrast, *wxStonor* and *wxB5* show no significant difference in activity in the two tissues.

wxG Transcripts in Endosperm and Pollen

The presence of a retrotransposon (called G) in *wx* intron 8 results in alternative splicing of the *wx* pre-mRNA (Figure 1) (Varagona et al., 1992). One splicing event generates a wild-type transcript, presumed to account for *WX* activity in this organ. This result led us to hypothesize that the higher level of enzymatic activity in pollen results from the production of more wild-type transcripts in pollen versus the endosperm. This possibility was initially addressed by RNA gel blot analysis of endosperm and pollen RNA isolated from *Wx* and *wxG* strains by using probes located upstream and downstream of the *wxG* insertion (P5 and P3, respectively; see Figure 1). Transcripts of wild-type size (2.2 kb) and shorter were detected in both tissues when probed with P5 (Figures 2A and 2F). Most of the shorter transcripts were not detected when the same blot was reprobed with P3 (Figures 2B and 2G), suggesting that these transcripts terminate prematurely within or near retrotransposon sequences.

Although the gross patterns of *wx*-encoded transcripts in endosperm and pollen appear similar, a closer examination suggests that the amount of all *wxG* transcripts is higher in pollen than in the endosperm (relative to wild-type levels in these tissues). This was confirmed by quantifying the mRNA signal in each sample by using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA) and adjusting for the amount of RNA loaded by normalizing to hybridization of a maize actin

Table 1. *Wx* Enzymatic Activity^a

Allele	Endosperm	Pollen
<i>wxG</i>	0.9 ± 0.2	26.6 ± 4.0
<i>wxStonor</i>	2.1 ± 0.8	0.8 ± 0.1
<i>wxB5</i>	1.8 ± 0.7	3.0 ± 0.6
<i>wxC34(Δwx)</i> ^b	2.0 ± 0.2	1.0 ± 0.1

^a Percentage of wild-type activity.

^b The percentages for *wxC34* have been subtracted from the other values shown.

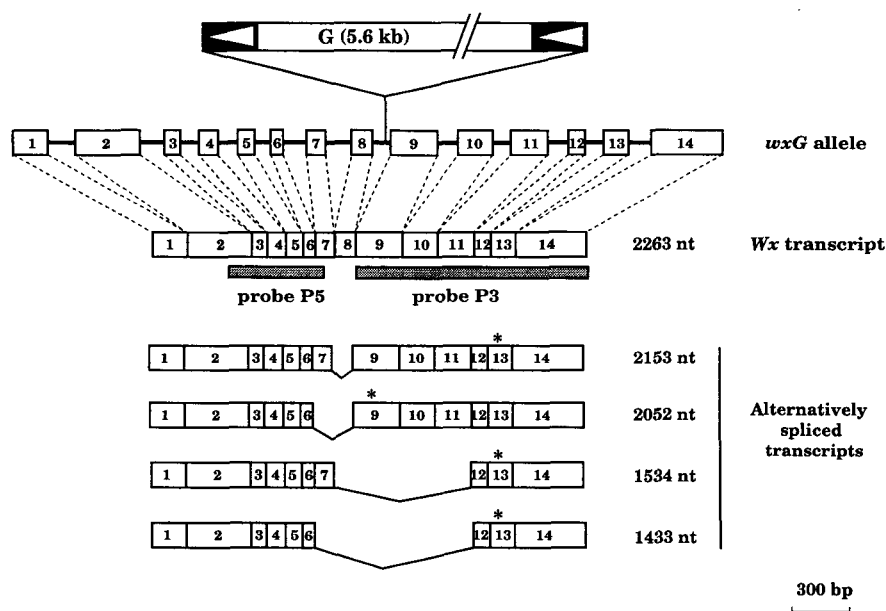


Figure 1. Structure of the *wxG* Allele and Transcripts.

The long terminal repeats of the G retrotransposon are represented by black squares containing a white triangle. Open boxes with numbers represent exons of the *wx* gene, and the horizontal lines connecting them represent the introns. Diagonal lines connect exons in alternatively spliced transcripts. Shaded boxes represent the probes used for RNA gel blot analysis. Asterisks denote the positions of premature stop codons. nt, nucleotides.

probe (Figures 2E and 2J). The amount of *wxG*-encoded transcripts was 13% that of the wild type in the endosperm and 54% that of the wild type in pollen (determined using probe P5). The proportion of wild-type size transcript is very similar in both tissues, representing 53% of all *wx* transcripts in endosperm and 64% of all *wx* transcripts in pollen. From these data, the amount of wild-type size transcript in the endosperm is ~7% that of wild-type levels (13% of 53%) and ~35% that of wild-type levels in pollen (54% of 64%).

Tissue-Specific Alternative Splicing in *wxG* Endosperm and Pollen

The pollen value of 35% correlates well with the enzymatic activity in this tissue. The same cannot be said for the endosperm value of 7%, which is almost 10-fold too high. RNA blots indicate on a gross level that there are tissue-specific differences in the size and amount of *wxG* transcripts. However, blots of this sort cannot be used to determine the precise splicing or termination events that occur in these tissues. Neither can they resolve wild-type transcripts from a previously identified alternatively spliced transcript (exons 7 to 9, 110 bp smaller; Figure 1). Inability to discriminate these RNAs would lead to an overestimate of the amount of wild-type size transcript in each sample.

To assess alternative splicing events in the two tissues, a reverse transcriptase–polymerase chain reaction (RT-PCR) assay was used with primers located in exons 5 and 12. This combination of primers was shown previously to detect all alternative splicing events among *wxG* endosperm RNA (Figure 1; Varagona et al., 1992). By repeating the assay with endosperm mRNA, four of the five splicing events identified previously were easily detected (8 to 9 [wild type], 7 to 9, 7 to 12, and 6 to 12; Figure 3A); the fifth event (6 to 9) was not seen, despite repeated attempts. In contrast, only two of the five splicing events were detected in pollen: 8 to 9 (wild type) and 7 to 9. Furthermore, these splicing differences were not a function of the developmental time that the RNA was isolated. Only two RNA products were detected among pollen RNA isolated from tassels just emerging from shoots (the stage when *wx* transcripts begin to accumulate during pollen development; G. Baran and S.R. Wessler, unpublished data; Figure 3B, lane 2) or from tassels just before shedding pollen 5 days later (the time of maximum accumulation of *wx* transcripts; Figure 3B, lane 1). Similarly, four splicing events were detected in endosperm isolated 14, 18, or 22 days after pollination (DAP) (Figure 3B, lanes 3 to 5).

Although not strictly quantitative, the RT-PCR data suggest that wild-type transcripts are relatively more abundant in pollen than in endosperm. This can be seen by comparing the ratio of splicing events 8 to 9 with 7 to 9 in the two tissues

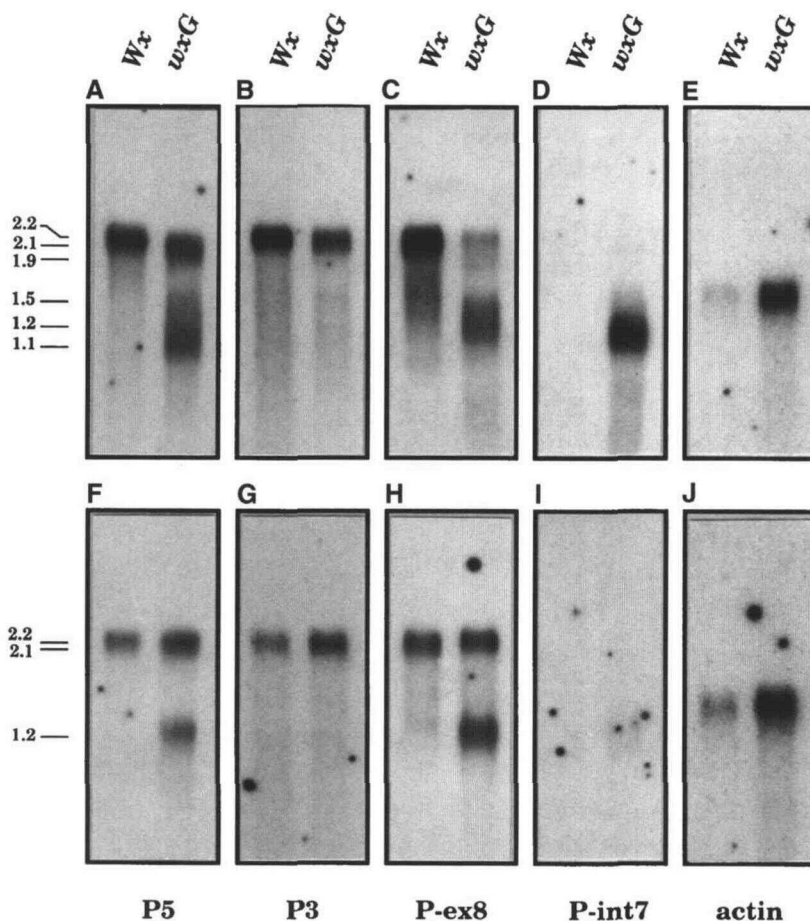


Figure 2. RNA Gel Blot Analysis of *Wx*- and *wxG*-Encoded Transcripts.

Transcript lengths are indicated at left in kilobases.

(A) to (E) Gel blots of endosperm RNA. *Wx* and *wxG* lanes contain 0.6 and 5.0 μg of poly(A)⁺ RNA, respectively.

(F) to (J) Gel blots of pollen RNA. *Wx* and *wxG* lanes contain 1.5 and 5.0 μg of poly(A)⁺ RNA, respectively.

(Figure 3A). However, it is not possible from this experiment to determine the exact relative amount of each splicing event, because small transcripts were amplified more efficiently than large ones. For example, comparison of the RNA gel blot (Figure 2B) with RT-PCR results clearly indicates that splicing events 7 to 9 and 6 to 12 were more efficiently amplified than the larger products. That is, these shorter transcripts are either barely visible (7 to 12) or not visible at all (6 to 12) on RNA gel blots (Figure 2B), but they account for a large fraction of the PCR products.

Quantifying *wxG* Transcripts

An RNase protection assay was used to quantify wild-type transcripts in *wxG* endosperm and pollen (Figure 4). By using

an RNA probe containing antisense *Wx* cDNA sequences, wild-type transcripts were detected as a 454-nucleotide protected fragment in *Wx* endosperm and pollen. This fragment was also prominent in *wxG* pollen but was greatly reduced in *wxG* endosperm. Additional fragments protected by *wxG* RNAs may correspond to the splicing and termination events that are depicted in Figure 4 (see below). Quantification of each band by using a PhosphorImager permitted us to measure the relative amount of each product in the sample. In this way, we found that wild-type transcripts account for 4.5 and 40% of all *wxG*-encoded transcripts in the endosperm and pollen, respectively. Recall that the steady state level of all *wxG* transcripts (as determined by RNA gel blots) relative to *Wx* transcripts was 13 and 54% in endosperm and pollen, respectively. When taken together, the amount of wild-type transcripts in *wxG* endosperms represents

0.6% of that in *Wx* endosperms (4.5% of 13%) and 22% of that in *Wx* pollen (40% of 54%). These values correlate well with the levels of enzymatic activity of 0.9% in endosperm and 27% in pollen.

These results suggest that the wild-type size bands visualized on the RNA gel blot in *wxG* endosperm and pollen (Figures 2A, 2B, 2F, and 2G) differ with respect to the relative contribution of wild-type and exon 7-to-9 splicing events. That is, in endosperm, most of this band would be expected to represent exon 7-to-9 events, whereas wild-type transcripts should predominate in pollen. This issue was addressed directly by reprobating the RNA gel blot with

an exon 8 probe that can hybridize only with wild-type but not exon 7-to-9 transcripts. As shown in Figures 2C and 2H, this expectation was borne out, providing direct evidence for tissue-specific alternative splicing.

Of the remaining transcripts detected in the RNase protection assay, alternative splicing can only account for the 291-nucleotide protected fragment resulting from the splicing of exon 7 to a downstream exon. Based on PCR results, the downstream exon could be either exon 9 or 12 in the endosperm but only exon 9 in pollen (Figure 3). In the endosperm, in which both products are possible, the vast majority are probably exon 7-to-9 events because the

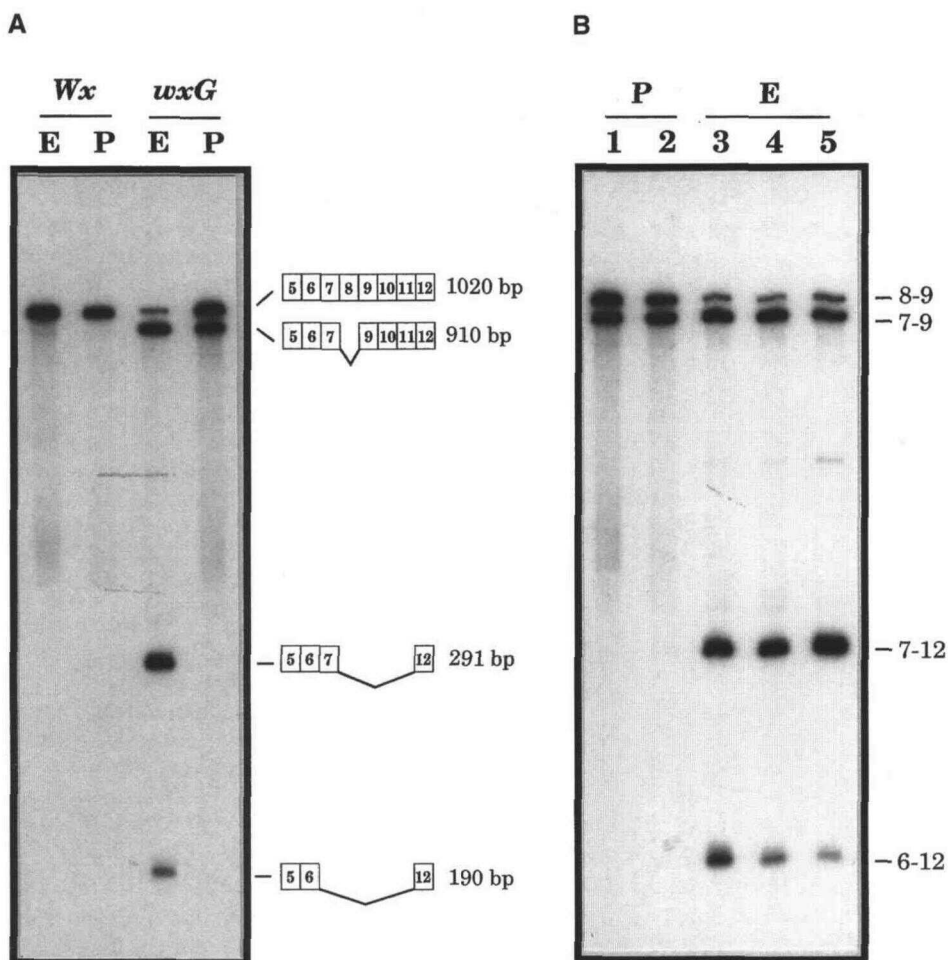


Figure 3. RT-PCR Analysis of *wxG*-Encoded Transcripts.

(A) Autoradiography of a denaturing acrylamide gel of end-labeled RT-PCR products from *Wx* and *wxG* endosperm (E) and pollen (P) RNAs. The presumed structure of each product is depicted at right. Sequences corresponding to exons 5 to 12 are shown as open boxes with numbers. Oblique lines represent alternative splicing events.

(B) As given in **(A)**, except that RNA was extracted from tassels just starting to shed pollen (lane 1) and just emerging from the shoot (lane 2) and from endosperms harvested 14 (lane 3), 18 (lane 4), and 22 DAP (lane 5). Numbers at right indicate the nature of the splicing event corresponding to each amplified product.

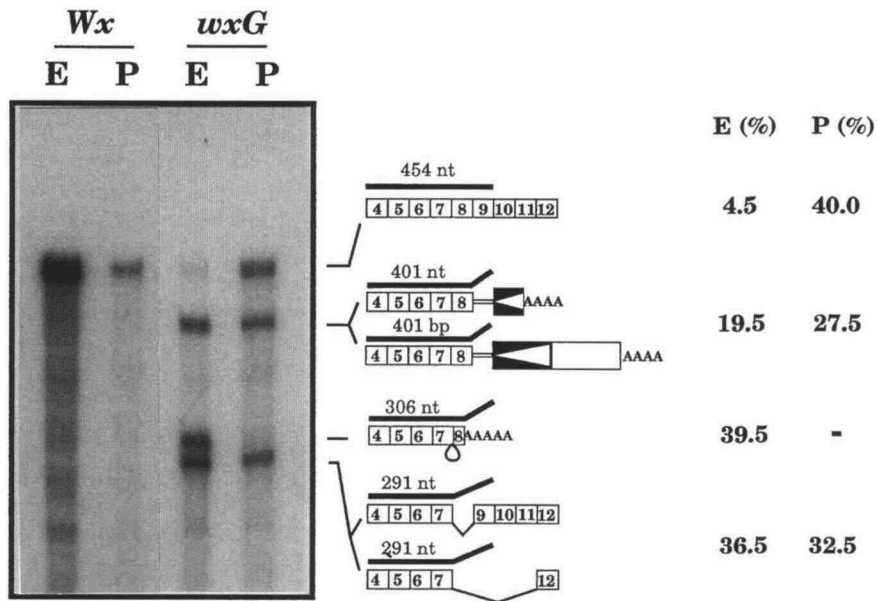


Figure 4. RNase Protection Assay of *wxG*-Encoded Transcripts.

Autoradiography of a denaturing acrylamide gel of protected products. Assays were performed using 1 and 15 μg of total *Wx* RNA isolated from endosperm (E) and pollen (P), respectively, and 30 and 90 μg of total *wxG* RNA isolated from endosperm and pollen, respectively. The deduced structure of each product is depicted next to the autoradiograph. *wx* exon sequences are represented as open boxes with numbers. Retrotransposon long terminal repeat sequences are shown as black boxes with a triangle, and retrotransposon internal sequences are shown as an open box. A thick black line represents the antisense *Wx* probe annealing to each transcript. Prematurely terminated transcripts are represented with a stretch of A residues at their 3' end. A protected fragment of 306 nucleotides (nt) in endosperm can be explained by hybridization of the probe to termination product T1. Intron 7 (which would be eliminated by RNase protection) is shown to indicate transcript structure. We suspect that the presence of this protected fragment in some but not all experiments indicates incomplete RNase digestion, even though the plasmid sequences contained in the probe were digested to completion (as shown by the absence of a 524-nucleotide full-length probe fragment in all lanes). The relative amount of each alternatively processed transcript in endosperm [E(%)] and pollen [P(%)] is shown at right. The absence of the 306-nucleotide protected fragment in pollen is noted as (-).

smaller transcript produced by splicing exon 7 to 12 (1534 nucleotides) was not prominent when the RNA gel blot was probed with the downstream P3 sequence (Figure 2B).

Truncated *wxG* Transcripts

Two RNase protection products of 401 and 306 nucleotides, one of which is only detected in endosperm, cannot be explained by alternative splicing. These products may be derived from transcripts that terminate prematurely in *wx* or element sequences. To test this notion, 3' rapid amplification of cDNA ends (RACE) was used to determine the sequences at the ends of truncated transcripts in *wxG* endosperm and pollen. Three products were obtained from endosperms (T1, T2, and T3), only one of which, T2, was also present at significant levels in pollen (Figure 5). The T2 transcript, found in both tissues, is polyadenylated in the 5' long terminal repeat of the *wxG* retrotransposon and protects a fragment of 401 nucleotides (Figure 4). The T3 transcript, which is also poly-

adenylated in element sequences but found only in endosperm, also protects a 401-nucleotide fragment.

The only remaining product, T1, is not expected to produce the prominent 306-nucleotide product seen only in the endosperm sample of the RNase protection assay. However, this product may be explained by postulating the occurrence of incomplete ribonuclease digestions of the probe in the RNA-RNA hybrids like that shown in Figure 4. Although this explanation seems unreasonable, it is consistent with the variable appearance of this band in other RNase protection experiments (data not shown). T1 is an interesting product because it terminates at a site 15 nucleotides within exon 8, upstream of the inserted G element. In addition, T1 retains intron 7; all other upstream introns are spliced correctly. When a DNA blot of 3' RACE products from endosperm and pollen was probed with an intron 7 fragment, the endosperm product was found to be much more abundant, thus providing additional evidence that the 306-nucleotide RNase fragment is T1 (data not shown). More convincing evidence was obtained when the same intron 7

fragment was used to probe blots containing RNA isolated from *Wx* and *wxG* endosperm and pollen (Figures 2D and 2I). Consistent with the RNase protection and 3' RACE results is the finding that intron 7-containing poly(A)⁺ transcripts are abundant in endosperm but rare in pollen. Furthermore, the hybridization signal corresponds in length to the truncated T1 transcript (1.1 kb).

DISCUSSION

In this study, we have shown that the retrotransposon-induced *wxG* allele has a tissue-specific phenotype with ~30-fold higher levels of WX enzymatic activity in pollen than in endosperm. Analysis of steady state RNA products by RNA gel blotting, RT-PCR, and RNase protection demonstrates that two factors contribute to this difference. First, as determined from RNA gel blots, there is relatively more steady state *wx* mRNA in *wxG* pollen than in endosperm. Pollen contains 54% of wild-type levels of all *wx* transcripts

compared with 13% in endosperm. Second, as determined from RNase protection assays, wild-type *Wx* transcripts account for 40% of all *wxG*-encoded transcripts in pollen but for only 4.5% of those in endosperm. Taken together, the higher level of wild-type *Wx* transcripts in pollen than in endosperm accounts for virtually all of the difference in enzymatic activity.

Why is there more of the wild-type *Wx* transcripts in pollen than in endosperms? Formally, at least three reasons can explain this difference. Perhaps pre-mRNA processing of *wxG* transcripts is the same in both tissues, but wild-type *Wx* transcripts are more unstable in the endosperm. Although this scenario would also explain the lower levels of *wxG*-encoded mRNA in the endosperm, it is very unlikely because wild-type *Wx* mRNA is quite stable in this organ (Figure 2). Tissue-specific differences in the rate of transcription could explain the lower levels of mRNA in the endosperm. However, differential rates of synthesis cannot explain the preferentially lower levels of wild-type transcripts in *wxG* endosperms. That is, although there is fourfold more *wxG*-encoded transcripts in pollen than in endosperm (54 and 13%, respectively), there is >30-fold more wild-type size transcripts in pollen (22% versus 0.6% in endosperm).

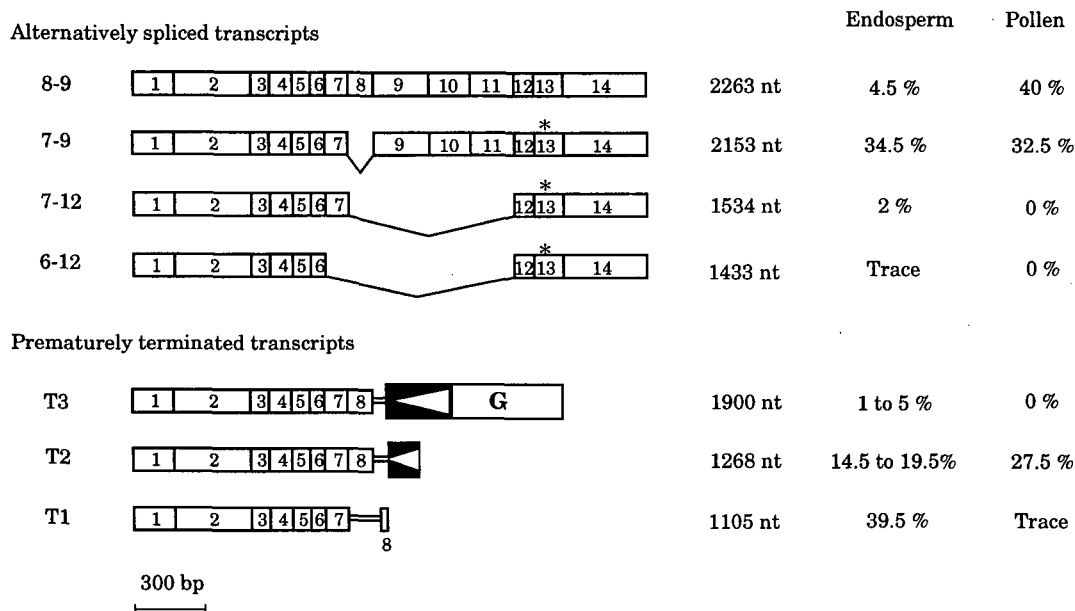


Figure 5. Summary of *wxG*-Encoded Transcripts in Endosperm and Pollen.

wx exon sequences are represented by open boxes with numbers. Filled boxes with open triangles represent LTR sequences from the G element. The open box labeled G represents internal element sequences. Asterisks denote the position of premature stop codons. The relative amount of the transcripts in each tissue is shown at right. These values were obtained from RNase protection analysis (transcripts 8 to 9 and T1 in endosperm; transcripts 8 to 9, 7 to 9, and T2 in pollen), from RNA gel blot analysis, or 3' RACE products (transcripts 7 to 12 and 6 to 12 in endosperm; transcript T1 in pollen), by estimating the relative intensity of the 3' RACE products from *wxG* tissues after agarose gel electrophoresis (transcript T3 in endosperm; transcripts 7 to 12, 6 to 12, and T3 in pollen), by subtracting 2% (the amount for alternatively spliced transcripts 7 to 12) from 36.5% (the sum of transcripts 7 to 9 and 7 to 12, obtained by RNase protection analysis; transcript 7 to 9 in endosperm), and by subtracting 1 to 5% (amount of transcript T3 in endosperm) from 19.5% (the sum of transcripts T2 and T3 as estimated from RNase protection analysis; transcript T2 in endosperm).

Furthermore, it is unlikely that the transcription rate would be affected by an insertion in the middle of the *wx* gene several kilobases downstream of the transcription start site. The third reason, which we favor, is that there are both qualitative and quantitative differences in pre-mRNA splicing and 3' end formation in these two tissues. Most importantly, there is more aberrant processing in the *wxG* endosperm.

This model also provides an explanation for the fourfold lower steady state level of *wxG*-encoded RNA in the endosperm. Aberrant splicing can generate transcripts containing premature stop codons resulting from the introduction of frameshifts. In yeast, *Caenorhabditis elegans*, and mammals, stop codons encountered prematurely in the 5' half of a transcript induce rapid mRNA degradation (Pulak and Anderson, 1993; Belgrader and Maquat, 1994; Jacobson, 1996). Premature stop codons in the 3' half have less of a destabilizing effect. The positions of premature stop codons in *wxG* transcripts found in the endosperm and/or pollen are shown in Figure 1. Three transcripts with stop codons in exon 13 (241 nucleotides upstream of the normal stop codon) were easily detected (Figure 5). In contrast, we were unable to detect the 6-to-9 event that is predicted to contain a premature stop codon in exon 9 sequences 995 nucleotides upstream of the normal stop codon (Figure 1). Although this splicing event was shown to occur in a previous study through the isolation of a rare cDNA (Varagona et al., 1992), it cannot be seen in the RT-PCR experiment shown in Figure 3.

Aside from the 6-to-9 event, many other alternatively spliced transcripts were not detected in this study. Of particular interest is an apparent bypass of the intron 9 and 10 splice acceptor sites in favor of the acceptor site from intron 11. That is, transcripts containing splicing events 6 to 10, 6 to 11, 7 to 10, or 7 to 11 (Figure 5) were not detected, even though these do not contain premature stop codons. It was noted previously (Varagona et al., 1992) that this may be due to the extremely high G + C content of introns 9 and 10 (60 and 58%, respectively) and the simultaneous availability of splice sites from introns 9, 10, and 11 in *wxG* but not in wild-type pre-mRNA.

Although tissue-specific alternative splicing is common in some animals, few examples have been reported in plants (reviewed in Luehrsen et al., 1994; Filipowicz et al., 1995). Recently, tissue-specific alternative splicing has been reported for the H protein gene (*gdcsH*) of *Flaveria trinervia* (Kopriva et al., 1995) and for the chorismate synthase gene (*LeCS2*) of tomato (Gorlach et al., 1995). For *gdcsH* and *LeCS2*, splicing involves the alternative use of two closely spaced 3' or 5' splice sites, respectively. In neither case is it known whether alternative splicing has physiological or regulatory consequences.

Unlike these examples of the splicing of wild-type genes, *wxG* is a mutant allele of a gene that is not normally alternatively spliced. However, closer examination of *wxG* pre-mRNA processing reveals underlying similarities between plants and animals and suggests a mechanism for the tissue-specific effect that is observed.

Tissue-Specific Alternative Splicing and 3' End Formation

Despite insertion of the G element at a considerable distance from the donor and acceptor splice sites of intron 8 (73 and 29 bp, respectively), the existence of alternatively spliced *wxG* transcripts with skipped exons indicates that splice site recognition was altered. This may be because plants do not efficiently process large introns (intron 8 is almost 6 kb) or because the *wxG* element has inserted into an intronic sequence required for correct splicing. That internal sequences in plant introns, especially UA-rich sequences, strongly influence splice site recognition has been documented and discussed (Luehrsen et al., 1994; Filipowicz et al., 1995).

Whatever the reason for inefficient splicing may be, our data suggest that the correct splice sites are recognized more efficiently in pollen than in endosperm. To understand the possible mechanism(s) underlying this cell type-specific splicing difference, it is necessary to extrapolate from animal studies because this question has not been pursued in plant systems. In animals, cell type-specific alternative splicing can be mediated by differences in the level of accessory or general splicing factors (reviewed in Valcarcel et al., 1995). For example, varying the concentration of a specific group of mRNA-associated proteins, the heterogeneous nuclear ribonucleoproteins, can promote exon skipping in *Drosophila* (Shen et al., 1995). Similarly, the concentration of the general splicing factor SF2/ASF can influence the choice between competing splice sites (Mayeda and Krainer, 1992) or can determine whether exons are included or excluded (Valcarcel et al., 1995). Recently, a protein related to SF2/ASF has been isolated from *Arabidopsis* and shown to function in a mammalian splicing assay (Lazar et al., 1995). It is not known whether there is variation in the concentration of splicing factors in different plant cell types. However, our data suggest that such differences exist.

Tissue-specific differences are also evident in the relative amounts of the truncated *wxG* transcripts. T1, the most abundant transcript in the endosperm, is virtually absent from pollen. T1 is an unusual transcript that terminates in *wx* but not in element sequences and that retains intron 7. That all transcripts terminating in exon 8 (T1) retain intron 7 whereas all transcripts terminating in element sequences (T2 and T3) have successfully processed this intron suggests a connection between 3' end formation and splicing. The simplest explanation for the tissue-specific splicing and polyadenylation events is that they are related. That is, there may be two consequences of the inefficient splicing of intron 7 observed preferentially in endosperm—one leading to exon skipping and the other to 3' end formation in exon 8 sequences. Such a connection has been established in vertebrates, where growing evidence suggests that 3' end formation and splicing may be coupled (Niwa et al., 1990; Niwa and Berget, 1991). Specifically, splicing of the 3' terminal intron has been shown to be influenced by and to influence a downstream polyadenylation site. Evidence for the coupling

of splicing and polyadenylation in maize has also been reported (Luehrsen and Walbot, 1994).

TE Insertions and Gene Regulation

Taken together, our data indicate that tissue-specific alternative processing of *wxG* pre-mRNA accounts for the tissue-specific phenotype of this allele. Previous studies demonstrated that TE insertion into or excision from 5' flanking or untranslated leader sequences can alter spatial and/or temporal parameters of transcription (reviewed in Weil and Wessler, 1990). For example, excision of the Antirrhinum Tam3 element from its position ~70 bp upstream from the *pallida* transcription start site generated new alleles with altered patterns of floral pigmentation (Coen et al., 1986). Similarly, new alleles produced by the excision of the *Ds1* element from the untranslated leader of the maize alcohol dehydrogenase *Adh1* gene displayed higher levels of *Adh1* mRNA in pollen than in the scutellum and anaerobic root (Dawe et al., 1993). Also, the insertion of a *Mu1* element into intron 1 of the maize *Adh1* gene was shown to result in tissue-specific alternative splicing in roots subjected to anaerobic stress and in kernels (Ortiz and Strommer, 1990). Recently, it was shown that retroviral integration in an intron of the murine *dilute* gene produced a tissue-specific phenotype as a result of splicing differences (Seperack et al., 1995). *wxG* represents a novel example of a retrotransposon insertion that has resulted in tissue-specific alternative splicing. As such, it expands the repertoire of ways that TEs can diversify gene regulation and potentially contribute to the evolution of gene expression.

METHODS

Strains

Strains of maize with *wxG*, *wxB5*, and *wxStonor* were obtained from O. Nelson (University of Wisconsin, Madison) and crossed into different inbred backgrounds, including W23, Mo20, and Kent21. *waxy* (*wx*) transcripts were not affected by the genetic backgrounds tested. *wxC34* was obtained from the Maize Genetics Cooperation Stock Center (Urbana, IL). All strains used in this study were homozygous for the designated *wx* allele.

Starch Granule Isolation and the WX Enzyme Assay

Starch granules were extracted from endosperm tissue collected 19 days after pollination (DAP), as described by Shure et al. (1983), and were isolated from mature pollen by using a modification of this protocol. Approximately 3 g of pollen was added to 45 mL of sucrose buffer (30% [w/v] sucrose, 50 mM Tricine, pH 8, 0.1 M KCl, 5 mM MgCl₂, 1 mM EDTA, and 1 mM DTT) and homogenized with a tissue grinder (Wheaton Scientific, Millville, NJ). Cell walls were eliminated by filtration through a nylon mesh (pore size of 41 μm), and the filtrate

was centrifuged at 2000g for 10 to 15 min. The pellet was washed twice with sucrose buffer before the purification of starch granules, as was described for endosperms.

WX enzyme activity was assayed by measuring the rate of incorporation of ¹⁴C-UDP glucose (Amersham) into starch granules. The assay was performed as previously described (Nelson, 1968), with the exception of the last step (consisting of trapping the starch granules on a glass filter fiber for the final wash), which was omitted because starch granules from pollen were lost at this stage as a result of their small size. The values of enzymatic activity for *wxG* and wild-type plants were corrected by subtracting the rate of incorporation measured from *wxC34* (*wxΔ*) (Wessler and Varagona, 1985). Assays were done in triplicate for starch granules isolated from endosperms and in duplicate for granules isolated from pollen.

RNA Gel Blot Analysis

Total RNA was isolated as described previously (Fedoroff et al., 1983) from endosperm collected 19 DAP and from tassels before shedding (when the middle of the main spike has just begun to extrude anthers). These stages represent times of maximum *wx* transcript accumulation in both tissues (G. Baran and S.R. Wessler, unpublished data). Tassels were used to analyze *Wx* expression in pollen, because microspores are the only cells in the tassel that express the *wx* gene (G. Baran and S.R. Wessler, unpublished results). When this protocol was used without modification, starch granules from *wxG* endosperms dissolved in the aqueous phase after lysis and homogenization and resulted in a gel that prevented the subsequent isolation of RNA. To prevent gelation, the tissue was homogenized with a polytron for only 5 sec. This amount of time was not long enough for the starch granules to dissolve but was sufficient for the endosperm cells to be broken. Homogenized material was then quickly spun at 16,000g for 3 min. Poly(A)⁺ RNA was isolated using the Poly-A-Tract mRNA isolation system IV (Promega).

DNA fragments used to prepare probes for RNA gel blot analysis were obtained by polymerase chain reaction (PCR) amplification of cloned *wx* cDNA sequences (probes P3, P5, and P-ex8) or *wx* genomic sequences (probe P-int7) and gel purified before labeling. Probe P3 extends from nucleotide position 1110 (exon 9) to the 3' end of the cDNA, P5 extends from position 418 (exon 2) to position 909 (exon 7), P-ex8 contains sequences from the entire exon 8, and P-int7 contains the entire intron 7. The actin probe was prepared by using a 485-bp HindIII-SstI fragment isolated from the plasmid pMAC1 (containing genomic sequences of a maize actin gene) provided by R.B. Meagher (University of Georgia). Probes used for RNA gel blot analysis were labeled by using the random primers DNA labeling kit (Bethesda Research Laboratories).

Reverse Transcriptase-PCR Amplification of *Wx* and *wxG* Transcripts

First-strand cDNAs were synthesized from ~50 ng of poly(A)⁺ RNA by using a reverse transcription system (Promega) and diluted 1:5 with distilled water. Primers *wx212* and *wx18* used for subsequent PCR amplification were first end labeled with γ-³²P-ATP by incubating 350 ng of each primer with 10 units of T4 polynucleotide kinase, 2 μL of 5 × reaction buffer, and 2 μL of γ-³²P-ATP (10 μCi/μL and 3000 Ci/mmol) in a 10-μL reaction for 10 min at 37°C. PCR amplification was

performed in a 40- μ L reaction containing 14 ng of each labeled primer, 4 μ L of diluted first-strand cDNA solution, 200 μ M of each deoxynucleotide triphosphate, and 1.2 μ L of the Elongase enzyme (Bethesda Research Laboratories) in PCR buffer (60 mM Tris-SO₄, pH 9.1, 18 mM [NH₄]₂SO₄, and 1.8 mM MgSO₄). Amplification was performed for 21 to 27 cycles (21 and 26 cycles for reactions containing wild-type cDNAs from endosperm and pollen, respectively, and 27 cycles for reactions containing wxG cDNAs from endosperm and pollen), with each cycle consisting of 30 sec at 94°C, 30 sec at 65°C, and 2 min at 68°C. Amplified DNA was then ethanol precipitated, and the pellets were resuspended in 15 μ L of loading buffer (80% formamide, 0.1% xylene cyanol, 0.1% bromophenol blue, and 2 mM EDTA). Resuspended samples were denatured by heating for 2 min at 95°C, and 1 to 10 μ L of each reaction was loaded on a 5% denaturing acrylamide gel.

Plasmid Construction

Plasmid pW2, which was used to prepare an antisense probe for RNase protection assays, contains a 454-bp insert from the wx cDNA (from exon 4 to exon 9). This plasmid was constructed in three steps: (1) a fragment obtained by PCR amplification from cloned wx cDNA with primers wx9 and wxexon72 was subcloned in the TA cloning vector (Invitrogen, San Diego, CA); (2) an EcoRI fragment containing the entire insert was subcloned in pGEM-7Zf(+) to generate the plasmid pWTS; and (3) a 303-bp BamHI fragment from the wx cDNA was then subcloned into pWTS to replace a 110-bp fragment containing part of the cDNA insert previously cloned and part of the pGEM-7Z polylinker.

RNase Protection Assay

RNase protection was performed using the ribonuclease protection assay kit (Ambion Inc., Austin, TX), according to the manufacturer's protocol. Plasmid pW2, linearized with XbaI, was used as a template for the preparation of an antisense RNA probe (containing 454 nucleotides of antisense wx sequences and 70 nucleotides of plasmid-derived sequences), using the Riboprobe Gemini II core system (Promega). The probe was labeled with α -³²P-UTP and gel purified before use. Complete RNase digestion of the unhybridized probe was monitored in all assays by the absence of a 524-nucleotide (454 plus 70 nucleotides) fragment resulting from digestion of the plasmid-derived 70 nucleotides.

3' Rapid Amplification of cDNA Ends

3' Rapid amplification of cDNA ends (RACE) was performed using the Bethesda Research Laboratories kit, according to the manufacturer's protocol, with ~50 ng of poly(A)⁺ RNA. A first round of PCR amplification was performed with the AUAP primer provided in the 3' RACE kit (anneals to the end of the cDNAs) and a primer located in exon 5 (wx212) for 18 cycles (30 sec at 94°C, 30 sec at 65°C, and 2 min at 68°C). A second PCR amplification was performed (using a diluted aliquot of the previous reaction) with the AUAP primer and a second nested primer located in exon 5 (wx21) for 16 cycles. Amplified fragments were cloned in the TA cloning vector and sequenced.

Primer Sequences

The primers used in this study are as follows: wx9 (5'-CGCGTGTTCGTTGACCACCC-3'; exon 4, 5' to 3'), wxexon72 (5'-TTGAGGTAGCACGAGAGAGG-3'; exon 7, 3' to 5'), wx21 (5'-CGGACTACAGGGACAACCAG-3'; exon 5, 5' to 3'), wx212 (5'-CGCTGGAACGACTACAGGG-3'; exon 5, 5' to 3'), and wx18 (5'-CGCTGAGGCGGCCATGTGG-3'; exon 12, 3' to 5').

ACKNOWLEDGMENTS

We thank Rich Meagher for advice and for the actin plasmid and Lane Arthur for critical reading of the manuscript. This work was supported by a grant from the National Institutes of Health to S.R.W.

Received January 3, 1997; accepted March 26, 1997.

REFERENCES

- Arkhipova, I.R., Lyubomirskaya, N.V., and Ilyin, Y.V. (1995). *Drosophila* Retrotransposons. (Austin, TX: R.G. Landes Company).
- Barkan, A., and Martienssen, R.A. (1991). Inactivation of maize transposon-*Mu* suppresses a mutant phenotype by activating an outward-reading promoter near the end of *Mu1*. *Proc. Natl. Acad. Sci. USA* **88**, 3502-3506.
- Belgrader, P., and Maquat, L.E. (1994). Nonsense but not missense mutations can decrease the abundance of nuclear mRNA for the mouse major urinary protein, while both types of mutations can facilitate exon skipping. *Mol. Cell. Biol.* **14**, 6326-6336.
- Bureau, T.E., and Wessler, S.R. (1992). *Tourist*: A large family of small inverted repeat elements frequently associated with maize genes. *Plant Cell* **4**, 1283-1294.
- Bureau, T.E., and Wessler, S.R. (1994). *Stowaway*: A new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. *Plant Cell* **6**, 907-916.
- Coen, E.S., Carpenter, R., and Martin, C. (1986). Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. *Cell* **47**, 285-296.
- Dawe, R.K., Lachmansingh, A.R., and Freeling, M. (1993). Transposon-mediated mutations in the untranslated leader of maize *Adh1* that increase and decrease pollen-specific gene expression. *Plant Cell* **5**, 311-319.
- Fedoroff, N.V., Mauvais, J., and Chaleff, D. (1983). Molecular studies on mutations at the *Shrunken* locus in maize caused by the controlling element *Ds*. *J. Mol. Appl. Genet.* **2**, 11-29.
- Filipowicz, W., Gniadkowski, M., Klahre, U., and Liu, H.X. (1995). Pre-mRNA splicing in plants. In *Pre-mRNA Processing*, A.I. Lamond, ed (Austin, TX: R.G. Landes Company), pp. 65-77.

- Flavell, A.J., Dunbar, E., Anderson, R., Pearce, S.R., Hartley, R., and Kumar, A.** (1992). Ty1-copia group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Res.* **20**, 3639–3644.
- Fridell, R.A., Pret, A.M., and Searles, L.L.** (1990). A retrotransposon-412 insertion within an exon of the *Drosophila melanogaster vermilion* gene is spliced from the precursor RNA. *Genes Dev.* **4**, 559–566.
- Gorlach, J., Raesecke, H.R., Abel, G., Wehrli, R., Amrhein, N., and Schmid, J.** (1995). Organ-specific differences in the ratio of alternatively spliced chorismate synthase (*LeCS2*) transcripts in tomato. *Plant J.* **8**, 451–456.
- Green, M.M.** (1988). Mobile DNA elements and spontaneous gene mutation. In *Eukaryotic Transposable Elements as Mutagenic Agents*, M.E. Lambert, J.F. McDonald, and I.B. Weinstein, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 41–50.
- Hirochika, H., Sugimoto, K., Otsuki, Y., and Kanda, M.** (1996). Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl. Acad. Sci. USA* **93**, 7783–7788.
- Jacobson, A.** (1996). Poly(A) metabolism and translation: The closed-loop model. In *Translational Control*, J.W.B. Hershey, M.B. Mathews, and N. Sonenberg, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 451–480.
- Johns, M.A., Mottinger, J., and Freeling, M.** (1985). A low copy number, *copia*-like transposon in maize. *EMBO J.* **4**, 1093–1102.
- Joseph, J.L., Sentry, J.W., and Smyth, D.R.** (1990). Interspecies distribution of abundant DNA sequences in *Lilium*. *J. Mol. Evol.* **30**, 146–154.
- Kim, H.Y., Schiefelbein, J.W., Raboy, V., Furtek, D., and Nelson, O.** (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.
- Kloeckener-Gruissem, B., and Freeling, M.** (1995). Transposon-induced promoter scrambling: A mechanism for the evolution of new alleles. *Proc. Natl. Acad. Sci. USA* **92**, 1836–1840.
- Kopriva, S., Cossu, R., and Bauwe, H.** (1995). Alternative splicing results in two different transcripts for H-protein of the glycine cleavage system in the C_4 species *Flaveria trinervia*. *Plant J.* **8**, 435–441.
- Kubli, E.** (1986). Molecular mechanisms of suppression in *Drosophila*. *Trends Genet.* **2**, 204–208.
- Lazar, G., Schaal, T., Maniatis, T., and Goodman, H.M.** (1995). Identification of a plant serine-arginine-rich-protein similar to the mammalian splicing factor SF2/ASF. *Proc. Natl. Acad. Sci. USA* **92**, 7672–7676.
- Liu, Y., Alleman, M., and Wessler, S.R.** (1996). A *Ds* insertion alters the nuclear localization of the maize transcriptional activator *R*. *Proc. Natl. Acad. Sci. USA* **93**, 7816–7820.
- 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. Nucleic Acid Res.* **47**, 149–193.
- Masson, P., Surosky, R., Kingsbury, J.A., and Fedoroff, N.V.** (1987). Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. *Genetics* **177**, 117–137.
- Mayeda, A., and Krainer, A.R.** (1992). Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68**, 365–375.
- Nelson, O.E.** (1968). The *waxy* locus in maize. II. Location of the controlling element alleles. *Genetics* **99**, 275–284.
- Niwa, M., and Berget, S.M.** (1991). Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev.* **5**, 2086–2095.
- Niwa, M., Rose, S.D., and Berget, S.M.** (1990). In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev.* **4**, 1552–1559.
- Ortiz, D.F., and Strommer, J.N.** (1990). The *Mu1* maize transposable element induces tissue-specific aberrant splicing and polyadenylation in two *Adh1* mutants. *Mol. Cell. Biol.* **10**, 2090–2095.
- Parkhurst, S., and Corces, V.** (1985). Forked, gypsy and suppressors in *Drosophila*. *Cell* **41**, 429–437.
- Pulak, R., and Anderson, P.** (1993). mRNA surveillance by the *Caenorhabditis elegans smg* genes. *Genes Dev.* **7**, 1885–1897.
- Purugganan, M.D., and Wessler, S.R.** (1994). Molecular evolution of *magellan*, a maize *Ty3/gypsy*-like retrotransposon. *Proc. Natl. Acad. Sci. USA* **91**, 11674–11678.
- SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., and Bennetzen, J.L.** (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**, 765–767.
- Seperack, P.K., Mercer, J.A., Strobel, M.C., Copeland, N.G., and Jenkins, N.A.** (1995). Retroviral sequences located within an intron of the *dilute* gene alter *dilute* expression in a tissue-specific manner. *EMBO J.* **14**, 2326–2332.
- Shen, J., Zu, K., Cass, C.L., Beyer, A.L., and Hirsh, J.** (1995). Exon skipping by overexpression of a *Drosophila* heterogeneous nuclear ribonucleoprotein *in vivo*. *Proc. Natl. Acad. Sci. USA* **92**, 1822–1825.
- Shure, M., Wessler, S., and Fedoroff, N.** (1983). Molecular identification and isolation of the *waxy* locus in maize. *Cell* **35**, 225–233.
- Springer, P.S., Edwards, K.J., and Bennetzen, J.L.** (1994). DNA class organization on maize *adh1* yeast artificial chromosomes. *Proc. Natl. Acad. Sci. USA* **91**, 863–867.
- Valcarcel, J., Singh, R., and Green, M.R.** (1995). Mechanisms of regulated pre-mRNA splicing. In *Pre-mRNA Processing*, A.I. Lamond, ed (Austin, TX: R.G. Landes Company), pp. 97–112.
- Varagona, M.J., Purugganan, M., and Wessler, S.R.** (1992). Alternative splicing induced by insertion of retrotransposons into the maize *waxy* gene. *Plant Cell* **4**, 811–820.
- Vignols, F., Rigau, J., Torres, M.A., Capellades, M., and Puigdomènech, P.** (1995). The *brown midrib3 (bm3)* mutation in maize occurs in the gene encoding caffeic acid *O*-methyltransferase. *Plant Cell* **7**, 407–416.

- Weil, C.F., and Wessler, S.R.** (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., and Varagona, M.** (1985). Molecular basis of mutations at the *waxy* locus of maize: Correlation with the fine structure genetic map. *Proc. Natl. Acad. Sci. USA* **82**, 4177–4181.
- Wessler, S.R., Baran, G., and Varagona, M.** (1987). The maize transposable element *Ds* is spliced from RNA. *Science* **237**, 916–918.
- Wessler, S.R., Bureau, T.E., and White, S.E.** (1995). LTR-retrotransposons and MITEs: Important players in the evolution of plant genomes. *Curr. Opin. Genet. Dev.* **5**, 814–821.
- White, S.E., Habera, L.F., and Wessler, S.R.** (1994). Retrotransposons in the flanking regions of normal plant genes: A role for *copia*-like elements in the evolution of gene structure and expression. *Proc. Natl. Acad. Sci. USA* **91**, 11792–11796.