The Plant Cell, Vol. 4, 1283–1294, October 1992 © 1992 American Society of Plant Physiologists

Tourist: A Large Family of Small Inverted Repeat Elements Frequently Associated with Maize Genes

Thomas E. Bureau and Susan R. Wessler[1]

Botany Department, University of Georgia, Athens, Georgia 30602

The *wx-B2* mutation results from a 128-bp transposable element-like insertion in exon 11 of the maize *Waxy* gene. Surprisingly, 11 maize genes and one barley gene in the GenBank and EMBL data bases were found to contain similar elements in flanking or intron sequences. Members of this previously undescribed family of elements, designated *Tourist*, are short (133 bp on average), have conserved terminal inverted repeats, are flanked by a 3-bp direct repeat, and display target site specificity. Based on estimates of repetitiveness of three *Tourist* elements in maize genomic DNA, the copy number of the *Tourist* element family may exceed that of all previously reported eukaryotic inverted repeat elements. Taken together, our data suggest that *Tourist* may be the maize equivalent of the human *Alu* family of elements with respect to copy number, genomic dispersion, and the high frequency of association with genes.

INTRODUCTION

Two general classes of transposable elements have been described in plants (Finnegan, 1989; Gierl and Saedler, 1992; Grandbastien, 1992). One class, exemplified by the retroposon *Cin4* of maize and the retrotransposon *Tnt1* of tobacco (Shepherd et al., 1984; Grandbastien et al., 1989), transposes via an RNA intermediate. The second class of transposable elements is characterized by terminal inverted repeats and transposes via a DNA intermediate (Gierl and Saedler, 1992). For elements in this second class, such as *Ac/Ds*, *En/Spm*, and *Mu* of maize (reviewed in Fedoroff, 1989; Chandler and Hardeman, 1992) and *Tam3* of snapdragon (reviewed in Coen et al., 1989), activity has been demonstrated directly by their ability to excise and insert. In contrast, several DNA sequences with the structural features of inverted repeat transposable elements have been suggested to be active solely on the basis of allelic polymorphism. For instance, *Tpc1* was found in one of two chalcone synthase alleles of parsley (Herrman et al., 1988) and *Tst1* in only one member of the patatin class II gene family of potato (Koster-Topfer et al., 1990).

The *Waxy* (*Wx*) gene of maize has served as an excellent molecular trap for a potpourri of transposable elements. *Wx* encodes a starch granule-bound ADP glucose glucosyltransferase and is expressed in pollen, endosperm, and embryo sac. The viable and easily visualized *Wx* phenotype has facilitated the isolation of more than 50 mutant *wx* alleles. Most of these have been mapped at both the genetic and molecular levels. Wessler and Varagona (1985) found that of 16 spontaneous *wx* mutants, seven were due to insertions. Six of these were due to large insertions (>4.5 kb), which were subsequently identified as retrotransposons (Varagona et al., 1992). The seventh, *wx-B2*, is due to an insertion of approximately 150 bp in the 3′ half of the *Wx* transcription unit.

In this study, we have characterized the *wx-B2* insertion and found that it corresponds to an unusual inverted repeat transposable element that we have designated *Tourist-Zm1*. Computer-assisted data base searches revealed that this insertion is a member of a very large family of previously undescribed elements. We have named this family of elements *Tourist*. Surprisingly, 11 maize genes reported to the GenBank and EMBL data bases were found to be flanked by or to harbor a previously unrecognized *Tourist* element. *Tourist* elements are small (133 bp on average), have conserved terminal inverted repeats, and display target site specificity. The shared terminal inverted repeat sequence, 3-bp target site, and other structural characteristics suggest that all *Tourist* elements are mobilized by the same transposase and thus represent a new family of plant inverted repeat transposable elements. Furthermore, based on the frequent occurrence of *Tourist* elements in sequenced maize genes and estimates of the repetitiveness of *Tourist* elements in maize genomic DNA, we suggest that the copy number of the *Tourist* element family may exceed that of all previously reported eukaryotic inverted repeat transposable elements.

RESULTS

Structural Characteristics of the Insertion Located in Exon 11 of the Maize *wx-B2* Allele

The *wx-B2* allele is a spontaneous null mutant derived from the maize inbred line W23 (Brink and Nilan, 1952; Nelson, 1968). Its somatic and germinal stability allowed Nelson (1968, 1976) to place *wx-B2* on a fine structure genetic map of the *Wx* locus. The accuracy of Nelson's map has been confirmed

---

[1] To whom correspondence should be addressed.

by its excellent correlation with the *Wx* physical map (Wessler and Varagona, 1985). Restriction site mapping showed that the *wx-B2* mutation corresponded to an approximately 150-bp insertion within an 800-bp SalI fragment in the 3′ half of the *Wx* transcription unit. To isolate this insertion, a DNA fragment was amplified by polymerase chain reaction from genomic DNA of a *wx-B2*–containing line using primers located in exons 10 and 12 of the maize *Wx* gene, as shown in Figure 1A. A single band of approximately 800 bp was observed on an ethidium bromide–stained agarose gel, corresponding to the predicted length of the insertion and its associated *Wx* flanking sequences (data not shown).

The DNA sequence of the amplified fragment indicates that the *wx-B2* insertion is 128 bp in length and is located within exon 11 at position 2650 with respect to the start of *Wx* transcription (Figure 1B; Klosgen et al., 1986). This insertion, designated *Tourist-Zm1*, has the structural features of an inverted repeat transposable element. The 14 bp at its termini form a perfect terminal inverted repeat (TIR). The TIR is flanked by a 3-bp duplication of the *Wx* target site 5′-GCA-3′. The subterminal region of *Tourist-Zm1* contains seven copies of the pentamer 5′-GGATT-3′ (Figure 1B, underlined sequence). In addition, the subterminal sequences form an imperfect inverted repeat (subterminal inverted repeat or SIR) separated from the TIR by 9 bp, as shown in Figure 2.

The TIR and SIR may allow the *Tourist-Zm1* sequence to form a hairpin structure (Figure 2). Other foldback transposable elements have been described previously including *FB* (foldback) of Drosophila (Smith and Corces, 1991), *Tc6* of *Caenorhabditis elegans* (Dreyfus and Emmons, 1991), and *TU* of sea urchin (Hoffman-Liebermann et al., 1989). The TIRs of *FB* and *TU* are composed of short tandem repeats. *Tourist-Zm1*

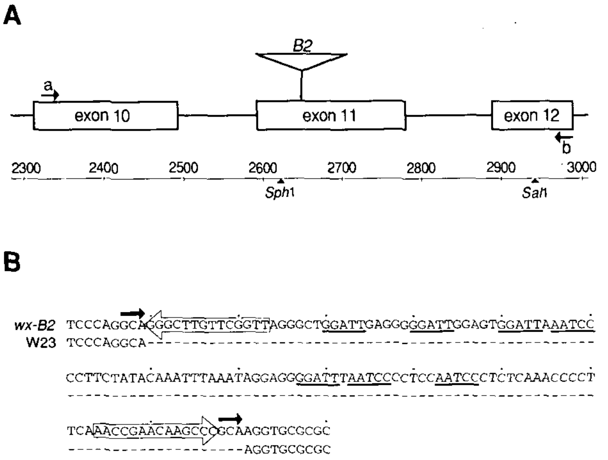

Figure 1. Position and DNA Sequence of the *wx-B2* Insertion (*Tourist-Zm1*).

(A) Site of the *wx-B2* insertion with respect to the 3′ half of the *Wx* gene. The solid arrows (a and b) indicate the position of the primers used to amplify the *wx-B2* insertion and flanking sequences.

(B) Sequence comparison between *wx-B2* and the progenitor *Wx* gene in the W23 inbred line (positions 2642 to 2660; Klosgen et al., 1986). Terminal inverted repeats are enclosed by open arrows and solid arrows are drawn over the 3-bp direct repeats. The seven subterminal repeats, 5′-GGATT-3′/5′-AATCC-3′, are underlined. Dashed line indicates nucleotides missing in the *Wx* gene of the W23 inbred line relative to *wx-B2*.

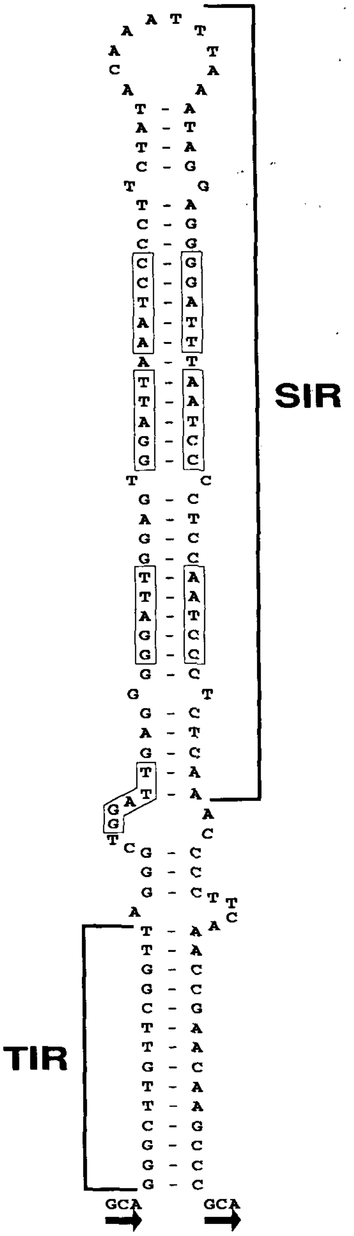

Figure 2. Potential DNA Secondary Structure of *Tourist-Zm1*.

The stem-loop structure was generated using the FOLD program (UWGCG). Boxed areas indicate the position of the seven 5′-GGATT-3′/5′-AATCC-3′ repeats. The direct repeat 5′-GCA-3′ is noted by solid arrows.

lacks any significant sequence homology to these elements, and its TIR does not appear to be composed of tandem repeats.

Identification of Sequences Similar to *Tourist-Zm1* in Maize and Barley

Computer-based sequence similarity searches revealed that *Tourist-Zm1* has no significant sequence identity to any known transposable element. To our surprise, however, sequence similarity was found among several previously sequenced maize genes (and pseudogenes) and with one barley gene. Closer examination revealed that these genes in fact contain small, previously unreported elements with structural and sequence similarities to *Tourist-Zm1*. These elements have been designated as *Tourist* family members, as shown in Table 1. With one exception, *Tourist* elements were retrieved from the GenBank (release 72.0) and EMBL (release 31.0) nucleic acid data bases using the query sequences shown in Table 1. *Tourist-Zm4* was identified by comparing the *Tourist-Zm1* sequence to the sequence of the maize *R* gene, *Lc* (S. Ludwig, L. Habera, and S. Wessler, unpublished data).

The location and structural characteristics of the *Tourist* elements are listed in Table 1. All *Tourist* elements are approximately the same size (133 ± 9 bp), have one to seven copies of the pentamer 5′-GGATT-3′ subterminal to their TIRs, and, with the exception of *Tourist-Zm1* (the *wx-B2* insertion), are located in noncoding regions. The calculated stabilization of free energy value (ΔG°) for each element suggests that *Tourist* elements have the potential to form stable intrastrand DNA secondary structures.

Percent sequence similarity between individual *Tourist* elements is given in Table 2. In each case, the orientation that consistently showed the highest sequence similarity to other *Tourist* elements was used in a multiple sequence alignment, as shown in Figure 3. *Tourist* family members are most similar at the element TIRs, which have the consensus sequence 5′-GGCCTTGTTCGGTT-3′. The similarity between TIRs of individual *Tourist* elements (5′ TIR versus 3′ TIR) and between each TIR sequence and the consensus is presented in Table 3. Internal sequences appear to be conserved primarily within subsets of *Tourist* elements (Figure 3). For instance, *Tourist-Zm2*, *Tourist-Zm4*, and *Tourist-Zm5* are more similar to each other than they are to other *Tourist* elements.

Alignment of the sequences flanking the *Tourist* elements reveals an apparent sequence preference for insertion sites, as shown in Table 4. *Tourist-Zm7*, *Tourist-Zm9*, *Tourist-Zm10*, and *Tourist-Zm12* were omitted from this alignment because

Table 1. *Tourist* Comparison

Element Designation[a]	Query Sequence[b]	Element Location[c]	Element position[d]	Size (bp)	No. of GGATTs[e]	ΔG° kcal/mol[f]
Tourist-Zm1	—	*wx-B2* (9S)	Exon 11 (+2650)	128	7	−69.4
Tourist-Zm2	*Zm1*	pseudo-*Gpa2*	5′ flanking (−2439†)	137	3	−40.2
Tourist-Zm3	*Zm1*	*Adh1-C^{m}* (1L)	3′ flanking (+4025*)	126	3	−39.2
Tourist-Zm4	*Zm1*	*Lc* (10L)	Intron 2 (+864*)	137	1	−31.8
Tourist-Zm5	*Zm1*	KD18	3′ flanking (+1548*)	142	5	−32.0
Tourist-Zm6	*Zm1*	*Aux311*	Intron 4 (+2749†)	131	6	−28.8
Tourist-Zm7	*Zm6*	*Bz-McC* (9S)	3′ flanking (+1937†)	134	3	−20.1
Tourist-Zm8	*Zm1*	*Adh1-1S* (1L)	3′ flanking (+4498*)	137	4	−34.0
Tourist-Zm9	*Zm8*	pseudo-*Gpa1*	5′ flanking (−572†)	130	3	−31.9
Tourist-Zm10	*Zm1*	*Zc2*	5′ flanking (−649†)	132	1	−44.2
Tourist-Zm11	TIR	*Aux311*	Putative promoter (−299†)	125	3	−26.0
Tourist-Zm12	*Zm8*	NBP1 (7L)	Intron 2 (+1545†)	137	3	−60.6
Tourist-Hv1	*Zm1*	*Acl1* (7)	3′ flanking (+2718*)	128	6	−35.7

[a] *Zm*, *Zea mays*; *Hv*, *Hordeum vulgare*.

[b] Computer-based sequence similarity searches were performed using the FASTDB and FASTA programs in the Intelligenetics and UWGCG computer program suites, respectively. TIR, *Zm1* terminal inverted repeat sequence.

[c] Gene loci encode the following: *wx*, ADP glucose glucosyl-transferase (Klosgen et al., 1986; this paper); pseudo-*Gpa2*, pseudogene of glyceraldehyde-3-phosphate dehydrogenase (Quigley et al., 1989); *Adh1-C^{m}*, alcohol dehydrogenase C^{m} allele (Osterman and Dennis, 1989); *Lc*, member of *R* gene family (Ludwig et al., 1989; S. Ludwig, L. Habera, and S. Wessler, unpublished data); KD18, oleosin (Qu and Huang, 1990); *Aux311*, auxin-binding protein (Yu and Lazarus, 1991); *Bz-McC* (also *-W22* and *-R* alleles), UDP glucose flavonoid glucosyl-transferase (Ralston et al., 1988); *Adh1-1S*, alcohol dehydrogenase *1S* allele (Sachs et al., 1986); pseudo-*Gpa1*, pseudogene of glyceraldehyde-3-phosphate dehydrogenase (Quigley et al., 1989); *Zc2* (also 27kD alleles), zein (Reina et al., 1990; Das et al., 1991); NBP1, nucleic acid–binding protein (Cook and Walker, 1992); *Acl1*, acyl carrier protein (Hansen and von Wettstein-Knowles, 1991). Chromosome position, if known, is indicated in parentheses.

[d] Numbers in parentheses give the *Tourist* element position proximal to the start of transcription (*) or translation (†).

[e] The number of 5′-GGATT-3′ repeats includes the occurrence of the reverse complement sequence 5′-AATCC-3′.

[f] Optimal folding of DNA sequences was performed using the FOLD program (UWGCG). Free energy values were determined for conditions at 1 M NaCl at 37°C. ΔG°, stabilization free energy.

Table 2. Percent Nucleotide Similarity between *Tourist* Elements[a]

	Tourist												
Tourist	*Zm1*	*Zm2*	*Zm3*	*Zm4*	*Zm5*	*Zm6*	*Zm7*	*Zm8*	*Zm9*	*Zm10*	*Zm11*	*Zm12*	*Hv1*
Zm1	100	66	64	67	59*	66*	58	64	57	73	58	46	73*
Zm2		100	66	79	88*	70*	65	63	66	60	66	56	69*
Zm3			100	59	68*	56*	58	66	61	65	69	54	54*
Zm4				100	78*	67*	68	59	60	60	69	64	57*
Zm5					100	74	63*	69*	71*	58*	64*	69*	68
Zm6						100	76*	57*	60*	58*	56*	66*	69
Zm7							100	56	53	62	67	61	62*
Zm8								100	81	57	58	78	65*
Zm9									100	55	57	64	68*
Zm10										100	59	51	61*
Zm11											100	63	68*
Zm12												100	54*
Hv1													100

[a] *Tourist* sequences were aligned using the GAP program (UWGCG) using a gap penalty of 1.00 and a gap length penalty of 0.30. An asterisk (*) indicates that the reverse complement of one of the *Tourist* elements was used to obtain an optimal alignment.

they are not flanked by an obvious target site duplication (Figure 3). The sequences of the nine remaining elements have a target site consensus sequence of 5′-MNCNNTAARKNK-3′ (Table 4). The central core of this sequence, the trimer 5′-TAA-3′, has presumably been duplicated upon element insertion in seven of 13 cases.

Tourist Elements in Maize *Adh1*, *Bz*, and *Zein* Alleles

The locations of *Tourist-Zm3* and *Tourist-Zm8* correspond to two different insertion polymorphisms previously noted in the maize *Adh1* alleles *1S* and *C*^{m}, as shown in Figure 4 (Sachs et al., 1986; Osterman and Dennis, 1989). There is no previous report, however, of either insertion polymorphism having the structure of a transposable element-like sequence. *Tourist-Zm3* corresponds to the insertion polymorphism located in the 3′ flanking sequences of the *Adh1-C*^{m} allele (Figure 4A; Osterman and Dennis, 1989). Like other members of the *Tourist* family, this element is flanked by the direct repeat 5′-TAA-3′. The presence of only a single copy of 5′-TAA-3′ in the *Adh1-1S* allele suggests that *Adh1-1S* is the progenitor condition and that this sequence was duplicated upon insertion of the *Tourist-Zm3* element. The second polymorphism is a 270-bp insertion in the 3′ flanking sequences of the *Adh1-1S* allele (Figure 4B; Sachs et al., 1986). This insertion is composed of two identical *Tourist* elements (designated *Tourist-Zm8*) arranged in tandem and each flanked by the direct repeat 5′-TAA-3′. At the same position in the *Adh1-C*^{m} allele is the sequence 5′-GCAA-3′ flanked by the direct repeat 5′-TAA-3′. Although this sequence is reminiscent of a transposable element footprint arising from *Tourist-Zm8* excision, there is no additional evidence indicating that *Tourist* elements excise or that excision produces transposable element footprints.

Identical *Tourist* elements are found at the same position in the 3′ flanking region of the maize *Bronze* (*Bz*) alleles *McC*, *R*, and *W22* (Ralston et al., 1988). Interestingly, this element, designated *Tourist-Zm7*, has sustained an insertion of a *Ds* element in a fourth allele, *bz-m4 D6856* (Klein et al., 1988). *Tourist* elements are also located at the same position in the 5′ flanking region of four different genes that encode the 27-kD class of zeins (Reina et al., 1990; Das et al., 1991). These elements, collectively designated *Tourist-Zm10*, share approximately 90%

Table 3. Nucleotide Similarity between *Tourist* TIRs

	Percent Similarity between	
Tourist	5′ and 3′ TIR[a]	TIR Consensus and 5′TIR/3′ TIR[b]
Zm1	100	100/100
Zm2	100	100/100
Zm3	100	100/100
Zm4	93	100/93
Zm5	86	100/93
Zm6	86	93/100
Zm7	86	86/86
Zm8	86	76/86
Zm9	86	76/71
Zm10	71	86/64
Zm11	64	43/100
Zm12	93	73/80
Hv1	86	93/100

[a] Orientation as indicated in references mentioned in Table 1.
[b] TIR consensus, 5′-GGCCTTGTTCGGTT-3′.

Figure 3. Multiple Sequence Alignment of the *Tourist* Family of Elements.

Tourist sequences were aligned using the PILEUP program (UWGCG) with a gap penalty of 1.00 and a gap length penalty of 0.30. The positions of the direct repeats are indicated by solid arrows and the TIRs by open arrows. Conserved nucleotides are indicated by white letters on a black background.

sequence identity. The finding of *Tourist* elements in multiple *Bronze* and *Zein* alleles indicates that the elements were probably present in the common progenitors of these alleles.

Tourist-Zm1, *Tourist-Zm4*, and *Tourist-Hv1* Copy Number

As mentioned above, *Tourist* elements were detected in one of six maize genes sequenced to date. This frequent occurrence and the lack of identity between any two *Tourist* elements suggest that this is a very large element family. To obtain independent evidence for high copy number, the approximate numbers of *Tourist-Zm1*, *Tourist-Zm4*, and *Tourist-Hv1* in maize genomic DNA were determined by slot blot hybridization, as shown in Figure 5. Hybridization stringencies that allowed for approximately 20% (low stringency) and 2% (high stringency) nucleotide mismatch were selected (Bolton and McCarthy, 1962; see Methods). At low stringency, all three elements appear to be highly repetitive (*Tourist-Zm1*, ~5.0 × 10^4; *Tourist-Zm4*, ~10^3; *Tourist-Hv1*, ~10^3). At high stringency, however, only *Tourist-Zm1* appears to be repetitive (~10^4). The copy numbers of the three *Tourist* elements analyzed were essentially the same in the HY and W23 inbreds with the exception of *Tourist-Zm4*, which was approximately one-half as abundant in W23 (data not shown).

Tourist-Zm1, *Tourist-Zm4*, and *Tourist-Hv1* Species Distribution

Tourist-Zm1, *Tourist-Zm4*, and *Tourist-Hv1* were also used as probes for DNA gel blots to determine the distribution of these elements among representative species of the subtribe Tripsacinae and in barley. At low-stringency conditions, *Tourist-Zm1* and *Tourist-Zm4* appear to be repeated in the genomes of *Zea mays* (HY and W23), teosinte (*Z. luxurians*), and *Tripsacum dactyloides*, as shown in Figures 6A and 6B. At high stringency, *Tourist-Zm1* still appears to be repetitive in these grasses. In contrast, at high stringency, *Tourist-Zm4* is present in what appears to be a single copy in the HY genome and is absent from the W23 genome. The identification of a sequence in one inbred line of maize but not in another is highly unusual. The *Tourist-Zm4* element was originally identified within intron 2 of the *R* gene *Lc* (Table 1). Because this region of other *R* alleles has not been sequenced, we do not know if *Tourist-Zm4* is in the same position of other *R* genes as well. If *Tourist-Zm4* insertion into *R* occurred recently, then it would not be surprising

Table 4. *Tourist* Target Site Preference

	Nucleotide Position[a]																
Tourist	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10
Zm1	T	T	**C**	C	**C**	A	G	G	C	**A**	**A**	**G**	G	**T**	G	C	G
Zm2	A	G	**C**	A	T	A	T	**T**	**A**	**A**	**A**	C	A	**G**	T	A	G
Zm3	G	T	**A**	T	**C**	G	G	**T**	**A**	**A**	**G**	**G**	C	**T**	T	C	G
Zm4	T	T	**C**	T	**C**	T	C	**T**	**A**	C	**G**	**T**	G	**G**	T	C	A
Zm5	A	A	**A**	T	**C**	C	T	**T**	**A**	**A**	**G**	**T**	G	**G**	G	T	T
Zm6	C	A	**A**	C	**C**	C	T	**T**	**A**	**A**	**A**	**G**	G	**G**	C	A	A
Zm8	G	C	**C**	C	**C**	G	C	**T**	**A**	**A**	**G**	**T**	G	**T**	C	T	A
Zm11	T	T	T	T	**C**	T	C	**T**	**A**	**A**	**A**	**T**	A	**T**	A	T	G
Hv1	T	C	**C**	A	**C**	C	A	**T**	**A**	**A**	**A**	**T**	T	**T**	T	C	C
Consensus[b]	N	N	**M**	N	**C**	N	N	**T**	**A**	**A**	**R**	**K**	N	**K**	N	N	N

[a] The target sites (positions 1 through 3; underlined) and 5′ (positions −6 through 0) and 3′ (positions 4 through 10) flanking sequences of each *Tourist* element were obtained from references mentioned in Table 1. *Tourist-Zm7*, *Tourist-Zm9*, *Tourist-Zm10*, and *Tourist-Zm12* lack a 3-bp direct repeat and were omitted from the alignment.

[b] N = G, A, T, or C; M = A or C; R = G or A; K = G or T.

to find this element in a subset of the naturally occurring *R* alleles. Thus, the W23 genome may contain an *R* gene that lacks *Tourist-Zm4*, whereas the *R* gene in the HY genome contains it. Finally, *Tourist-Hv1* is more widely distributed than either *Tourist-Zm1* or *Tourist-Zm4* (Figure 6C). At high stringency, however, hybridization only occurs with barley genomic DNA and at a single copy level. This single band may correspond to the *Acl1* gene from which the *Tourist-Hv1* element was originally identified (Table 1).

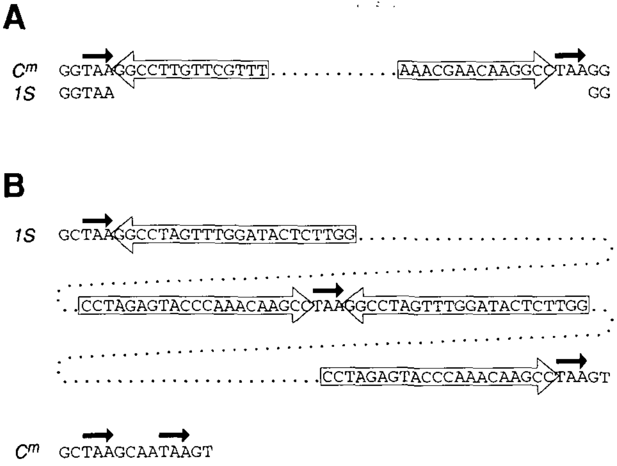

Figure 4. Insertion Polymorphisms in Maize *Adh1* Alleles Are *Tourist* Elements.

(A) Sequence comparison between 3′ flanking regions of the *Adh1* alleles *C^m* (positions 4020 to 4155; Osterman and Dennis, 1989) and *1S* (positions 3815 to 3821; Sachs et al., 1985). Only the TIR (open arrows) of the 126-bp *Tourist-Zm3* insertion and the 3-bp direct repeat (solid arrows) generated upon insertion are shown.

(B) A tandem repeat of two identical *Tourist-Zm8* elements corresponds to the 270-bp insertion polymorphism revealed when the 3′ flanking sequences of the *Adh1-1S* (positions 4493 to 4779) and *Adh1-C^m* (positions 4824 to 4836) alleles are compared. As in **(A)**, only the TIRs of each 137-bp *Tourist-Zm8* element and the direct repeats flanking each element are shown.

DISCUSSION

The maize *wx-B2* mutation is caused by the insertion of a previously unreported inverted repeat insertion element into exon 11 of the *Wx* transcription unit. This element has been named *Tourist-Zm1* and is characterized by its small size (128 bp), a 3-bp target site duplication, a 14-bp TIR, a subterminal pentamer repeat (5′-GGATT-3′), and the potential to form a hairpin structure. Eleven previously unreported *Tourist*-like elements were identified following computer-based sequence similarity searches. These elements have been designated as members of the *Tourist* family of elements because they share many of the structural characteristics of *Tourist-Zm1* (Tables 1 to 4 and Figure 3). In addition, *Tourist* elements appear to have a target site preference (Table 4); seven of 13 *Tourist* elements have inserted into the trinucleotide 5′-TAA-3′.

Tourist Forms a Unique Family of Plant Inverted Repeat Transposable Elements

Previous investigators have grouped plant inverted repeat transposable elements into families based on TIR sequence, TIR size, and the number of base pairs duplicated upon insertion (families I, II, and III of Table 5). A comparison of the structural characteristics of these plant inverted repeat transposable elements with *Tourist* suggests that *Tourist* represents

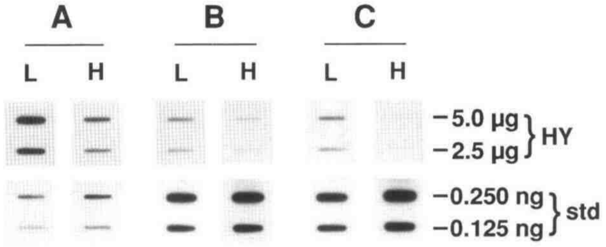

Figure 5. Copy Number of *Tourist-Zm1*, *Tourist-Zm4*, and *Tourist-Hv1*.

(A) *Tourist-Zm1*.
(B) *Tourist-Zm4*.
(C) *Tourist-Hv1*.
Genomic DNA from the maize inbred line HY and standard DNA (std) (pB2119, pLcH3, or pgLH1) were serially diluted, transferred to a nitrocellulose membrane, probed with the *Tourist* elements given above, and washed at low (L) and then high (H) stringency. Only two dilutions of each DNA sample are shown.

a unique family. Although both *Tourist* and members of the CACTA family are flanked by a 3-bp direct repeat, the TIR sequence of *Tourist* shares no significant homology to the terminal sequence 5′-CACTA-3′. Similarly, despite the fact that *Mu* and *Tourist* elements display internal sequence heterogeneity among family members, *Mu* and *Tourist* elements do not belong in the same family because their TIRs and the size of the target site duplication differ.

Two other characteristics distinguish the *Tourist* element family from other plant inverted repeat transposable elements. First, whereas other plant inverted repeat elements lack a target sequence preference, *Tourist* elements appear to prefer the target sequence 5′-TAA-3′. Second, *Tourist* elements are on average 133 bp in length, which is less than half the length of the smallest reported plant inverted repeat transposable element, *dTph1* of petunia (283 bp; Gerats et al., 1990).

Common Structural Features Suggest *cis* Requirements for *Tourist* Transposition

The *wx-B2* mutation was first identified approximately 40 years ago by Nilan and Brink (1952). Recent insertion of *Tourist-Zm1* into the maize *Wx* gene and the polymorphisms caused by insertion of *Tourist-Zm3* and *Tourist-Zm8* into the 3′ flanking regions of the maize *Adh1* gene clearly indicate a history of *Tourist* mobility. In contrast, the stability of the *wx-B2* mutant phenotype in a wide range of genetic backgrounds indicates that excision of *Tourist-Zm1* and probably other *Tourist* elements, if it occurs, is rare.

Due to their limited coding capacity, the *Tourist* elements identified in this study almost certainly do not encode a transposase. Rather, *Tourist* elements at *Wx* and *Adh1* were probably mobilized by transposase encoded at another chromosomal locus. Despite our inability to identify the gene(s) responsible for *Tourist* mobilization, we believe that the numerous conserved

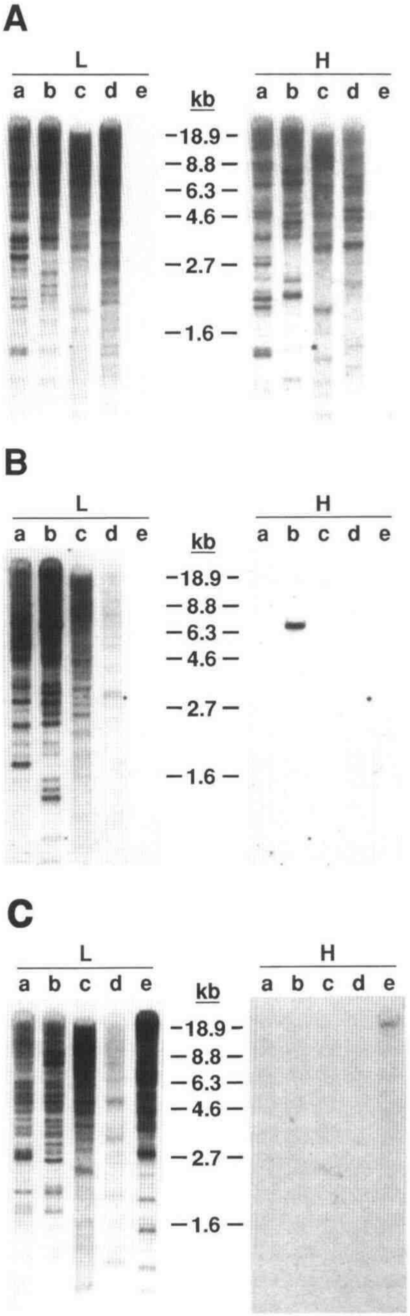

Figure 6. Species Distribution of *Tourist* Elements.

(A) *Tourist-Zm1*.
(B) *Tourist-Zm4*.
(C) *Tourist-Hv1*.
Five micrograms of EcoRI-digested genomic DNA from two maize inbred lines (W23 and HY, a and b, respectively), *Zea luxurians* (teosinte, c), *Tripsacum dactyloides* (d), and barley (e) was electrophoresed through a 0.8% agarose gel, transferred onto a nylon membrane, probed with the *Tourist* elements given above, and washed at low (L) and then high (H) stringency.

Table 5. Plant Inverted Repeat Transposable Element Families

Family Name	Element Name (Host Organism)[a]	Terminal Sequence Consensus[b]	TIR (bp)[c]	Target Site (bp)	Target Site Consensus[d]
I. *Ac/Ds*		YAGGG	11–15	8	None
	Ac/Ds, Bg/rbg, ruq, rDt (maize)				
	Tam3 (snapdragon)				
	Tdph1 (petunia)				
	Tpc1 (parsley)				
	Ips-r (pea)				
	Tst1 (potato)				
	Tat1 (Arabidopsis)				
	Gulliver (Chlamydomonas)				
II. CACTA		CACTA	13–14	3	None
	En/Spm, Mpi1 (maize)				
	Tam1, 2, 4, 7, 8, and *9* (snapdragon)				
	Tgm1 (soybean)				
	Pis1 (pea)				
III. *Mu*		GAGATAATTGCCA...	185–220	9	None
	Mu1–9 (maize)				
IV. *Tourist*		GGGCTTGTTCGGTT	15–23	3	TAA
	Tourist-Zm1–12 (maize)				
	Tourist-HV1 (barley)				

[a] Structural characteristics of plant inverted repeat transposable element families were obtained from the following references: *Ac/Ds* family (Federoff, 1989; Ferris, 1989; Gerats et al., 1990; Peleman et al., 1991; Gierl and Saedler, 1992); CACTA family (Coen et al., 1989; Fedoroff, 1989; Nacken et al., 1991; Gierl and Saedler, 1992); *Mu* family (Chandler, 1992).

[b] Y = T or C.

[c] TIR, terminal inverted repeat.

[d] None, no target site specificity reported.

structural features of the *Tourist* family argue for mobilization by a single transposase or family of related transposases.

In general, *Tourist* elements share TIR sequence similarity, generate a 3-bp direct repeat upon insertion, and contain one or more copies of a subterminal pentamer repeat (Tables 1 and 3 and Figure 3). Previous studies of other elements have highlighted the importance of each of these features in transposase-element interactions. For example, TIRs are believed to be the sites of enzymatic cleavage by transposase. Elements that share similar TIRs (e.g., *Ac, Tam3,* and the Drosophila *hobo* or *En/Spm, Tam1,* and *Tgm1*) also display amino acid similarity among their putative transposases (Calvi et al., 1991; Nacken et al., 1991). Similarly, the length of the target site duplication is almost always characteristic of an element family (Table 5) and reflects the size of the transposase-mediated staggered endonuclease cut required for element insertion. This duplication is 3 bp for all *Tourist* elements that are flanked by a direct repeat (Figure 3). In addition, the finding that seven of 13 *Tourist* element target sites have the same sequence, 5′-TAA-3′, argues strongly for the involvement of a single transposase in *Tourist* element mobility. Finally, many elements have repeated sequences internal to their TIRs. In a few instances, these subterminal repeats have been implicated as sites for transposase binding. For example, Kunze and Starlinger (1989) determined that the product of the *Ac*-encoded open reading frame "a" binds specifically to the *Ac* subterminal repeat 5′-AAACGG-3′. Moreover, *Ac* derivatives involving the deletion of some of these repeats have reduced transposition frequencies in transgenic tobacco (Coupland et al., 1988). Similarly, the *Spm*-encoded *TnpA* protein binds specifically to *Spm* subterminal repeats (5′-ACCGACACTCTTA-3′), and deletions of these repeats reduce the frequency of *Spm* excision (Gierl et al., 1988). Although all *Tourist* elements have between one and seven copies of the pentamer repeat 5′-GGATT-3′, the role of this pentamer and of any other sequence as a *cis* requirement for element mobility remain to be tested.

In addition to similarities in TIRs, direct repeats, and subterminal repeats, the *Tourist* elements identified in this study are remarkably homogeneous in length; all are 133 ± 9 bp. Such length homogeneity is extremely unusual among nonautonomous elements; most are deletion derivatives of the autonomous element and thus form a heterogeneous collection with respect to their length. One notable exception is the *Ds1* group of elements, which are about 90% conserved and approximately 400 bp long. It has been suggested that both size and sequence of *Ds1* elements are important for transposition (MacRae and Clegg, 1992). The finding of length homogeneity despite sequence heterogeneity among *Tourist* elements suggests that length may be an additional *cis* requirement for transposition of this element family.

Tourist Elements Are Highly Repetitive in Maize

The detection of several *Tourist* elements within maize nuclear genomic sequences submitted to the GenBank (release 72.0) and EMBL (release 31.0) nucleic acid data bases strongly suggests that the *Tourist* family of elements is highly repetitive. In fact, one *Tourist* element has been found for every 30 kb of maize nuclear genomic sequence submitted to these data bases. Alternatively, one of six reported maize nuclear genes either harbor or are adjacent to a *Tourist* element. No plant element described to date is as frequently associated with genes. In addition, *Tourist* elements are dispersed throughout the genome, as evidenced by their presence in genes located on at least four different chromosome arms (Table 1).

The ubiquitous nature of *Tourist* elements is reminiscent of the human *Alu* family of retroposons. Like *Tourist*, members of the *Alu* family are often found in introns and flanking sequences of genes. When members of an element family are frequently encountered in genes, it follows that the genomic copy number of that family should be very high. This is certainly borne out for the *Alu* family: there are approximately 500,000 *Alu* elements in the human genome. In other words, *Alu* comprises 5% of the human genome and is found in one copy per 5000 bp on average.

Could the copy number of the *Tourist* family in the maize genome approach the number of *Alu* elements in the human genome? To assess the copy number of *Tourist* in the maize genome, we utilized both slot and DNA gel blot analyses. One problem in accurately determining *Tourist* element copy number results from the high degree of sequence heterogeneity among family members (Table 2 and Figure 3). In fact, in pairwise combinations, most elements differ by greater than 30% (Table 2). If the stringency of hybridization is lowered to a level that would allow all *Tourist* elements in the genome to hybridize with a particular *Tourist* element probe, the subsequent nonspecific hybridization would yield a gross overestimate of copy number. Due to this sequence heterogeneity, we chose to estimate the magnitude of the *Tourist* family by approximating the copy number of individual *Tourist* elements under conditions determined to be stringent enough to preclude both nonspecific hybridization and cross hybridization between most of the sequenced *Tourist* elements.

The three elements, *Tourist-Zm1*, *Tourist-Zm4*, and *Tourist-Hv1*, tested by this method appear to be highly repetitive in the maize genome; elements with greater than 80% sequence similarity with *Tourist-Zm4* and *Tourist-Hv1* are present in about 10^3 copies, whereas elements with greater than 98% sequence similarity to *Tourist-Zm1* are present at about 10^4 copies. To the best of our knowledge, the copy number of the *Tourist-Zm1* element alone exceeds that reported for any other eukaryotic inverted repeat transposable element. For comparison, the copy number of *Tourist-Zm1* is more than 100 times that reported for *Ac/Ds*, *P*, or *Mu* (Fedoroff et al., 1983; Smith and Corces, 1991; Chandler and Hardeman, 1992). This finding, coupled with the recent insertion of a *Tourist-Zm1* element into the *Wx* gene, suggest a more recent history of activity for this subfamily.

In estimating *Tourist* family copy number, it is important to keep in mind that we have assayed only three of 13 known *Tourist* elements and that no two elements were encountered more than once in our computer-assisted data base searches. Thus, the actual copy number of the entire *Tourist* family is most likely substantially higher and could certainly account for the preponderance of *Tourist* elements detected in maize genes.

Taken together, our data indicate that *Tourist* may be the maize equivalent of the human *Alu* element with respect to copy number, dispersion, and the high frequency of association with genes. In this regard, it is relevant to note the important role *Alu* elements have played in restructuring the human genome. *Alu* elements have been implicated in altering gene expression, providing sites for homologous and illegitimate recombination, limiting gene conversion, and leading to the formation of macrotransposons (reviewed in Deininger, 1989). Now that the *Tourist* family has been detected and elements can be easily identified, it will be interesting to see if these elements are responsible for similar phenomena in maize.

METHODS

Plant Material and Genomic DNA Isolation

Maize lines containing the *wx-B2* allele were obtained from O. Nelson (University of Wisconsin, Madison). Maize genomic DNA was isolated as reported previously (Wessler and Varagona, 1985). *Zea luxurians* (teosinte) and *Tripsacum dactyloides* genomic DNA were obtained from S. White (University of Georgia, Athens), and *Hordeum vulgare* (Betzes barley) from L. Hansen (Carlsberg Laboratory, Copenhagen).

Tourist-Zm1 Isolation

The region containing the *Tourist-Zm1* element was amplified from a *wx-B2*–containing maize line by polymerase chain reaction (PCR) using synthetic oligonucleotides corresponding to sites in exon 10 (a: 5′-AGGCGTTGCAGGCGGAGGCG-3′ [nucleotides 2325 to 2344]) and exon 12 (b: 5′-CGCTGAGACGGCCCATGTGG-3′ [nucleotides 2974 to 2955]) of the maize *Wx* gene (Figure 1A; Klosgen et al., 1986).

Genomic DNA (300 ng) was added to 100 μL of a PCR cocktail consisting of 1 × PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, and 0.01% gelatin), 200 μM of each deoxynucleotide triphosphate (Pharmacia), 70 ng each primer (a and b), 10% glycerol, and 2.5 units of *Thermus aquaticus* DNA polymerase (Ampli*Taq*; Perkin-Elmer-Cetus). Using a programmable thermal cycler (Perkin-Elmer-Cetus), 40 cycles of amplication were carried out using a step program (95°C, 1 min; 65°C, 2 min; 72°C, 3 min), followed by a 10-min final extension at 72°C. Fifty microliters of the reaction mixture was digested with 10 units each of the restriction endonucleases SphI and

SalI (see Figure 1 for the position of these sites). The SphI-SalI PCR fragments were isolated (Weil and Bureau, 1992), cloned into pUC119 (construct designated as pB2119; Vieira and Messing, 1987), and sequenced by the dideoxy chain termination method (Sanger et al., 1977) using a Sequenase (United States Biochemical Corp.) kit as directed by the manufacturer.

DNA Sequence Analysis

Data base searches and sequence analyses were performed using the computer program suites of Intelligenetics (version 5.4; Intelligenetics, Inc., Mountain View, CA) and the University of Wisconsin Genetics Computer Group, Madison (UWGCG, version 7.0; Devereaux et al., 1984) accessed through the BioScience Computing Facility at the University of Georgia.

Optimal folding of DNA sequences was performed using the recursive algorithm of Zuker and Steigler (1989) as part of the FOLD program (UWGCG). Base pair and stacking energies for DNA were obtained from Breslauer et al. (1986) and the loop destabilizing energies from Freier et al. (1986).

Slot and DNA Gel Blot Hybridization

Slot blots were performed using a Minifold II apparatus (Schleicher & Schuell), as previously described by Zhao et al. (1989). The concentration of genomic (maize inbred lines HY and W23) and plasmid (pB2119, pLcH3 [containing the 5′ half of the *R* gene *Lc*], pgLH1 [containing the 3′ barley *Acl1* gene and 3′ flanking regions; obtained from L. Hansen], and pUC119) DNA was measured on a Hoefer TKO100 mini-fluorometer using the fluorescent dye bis-benzimidole (Hoechst 33258). Genomic DNA and plasmid standards were serially diluted ranging from 5 to 0.625 μg and 1 to 0.125 ng, respectively. Each slot blot run included the diluted genomic DNA, plasmid standard, and pUC119 control (to establish the degree of nonspecific hybridization).

For DNA gel blots, total genomic DNA (5 μg) was digested with 20 units of the restriction endonuclease EcoRI and electrophoresed through a 0.8% agarose gel (Bethesda Research Laboratories) at 100 V. DNA transfer onto GeneScreen Plus (New England Biolabs) was performed according to the procedure of Sambrook et al. (1989).

Tourist-Zm4 and *Tourist-Hv1* elements were isolated by PCR for use as probes of slot and DNA gel blots. A degenerate oligonucleotide corresponding to the TIR consensus sequence between *Tourist-Zm1*, *Tourist-Zm4*, and *Tourist-Hv1* was synthesized (*Tourist*A: 5′-GG[C/A]C-[T/A]T[G/A]TT[C/T][G/A]GTT-3′) and used in a PCR protocol with pLcH3 and pgLH1 constructs as target sequences. PCR conditions were as before except that an annealing temperature of 45°C for 1 min and an extension time of 2 min was used. PCR products were subsequently cloned and sequenced, and were found to be identical with the published sequences (see Table 1 for references).

Tourist-Zm1 could not be amplified by PCR as described above, presumably due to its high degree of DNA secondary structure (Table 1 and Figure 2). In order to use *Tourist-Zm1* as a labeled probe for slot and DNA gel blots, a SphI-SalI insert isolated from the pB2119 construct was digested with the restriction endonucleases BstNI and BssHI, which recognize sites immediately flanking the TIRs of *Tourist-Zm1*. To prevent foldback of single-stranded DNA during primer extension, this fragment was further digested with the restriction endonuclease DraI, which recognizes a site midway between the termini.

The *Tourist-Zm1* fragments and the *Tourist-Zm4* and *Tourist-Hv1* PCR products (50 ng) were radiolabeled by incorporation of ^{32}P-dATP and ^{32}P-dGTP (New England Biolabs) during *Tourist*A primer extension using a random primer kit (Bethesda Research Laboratories; random primers were not used). Hybridization conditions were as mentioned by Zhao and Kochert (1992). Filters were washed at low (2 × SSC [1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate], 0.5% SDS) and then high (0.1 × SSC, 0.5% SDS) stringency for 60 min at 65°C. Autoradiography was performed using Kodak X-Omat film and a Du Pont intensifying screen according to Sambrook et al. (1989). Scanning densitometry of slot blot autoradiograms was performed using a Beckman DU-7 spectrophotometer with a gel scan attachment and analyzed according to Rivin et al. (1986). For DNA gel blots, band intensity does not reflect copy number because the film exposure times (12 hr to 3 days) were varied to obtain optimum resolution.

ACKNOWLEDGMENTS

The authors would like to thank Shawn White and Cliff Weil for critically reviewing this manuscript. We would also like to thank Michael Weiss and Jonathan Arnold for their valuable advice in the use of the Intelligenetics and UWGCG computer program suites. This work was supported by the National Institutes of Health and the Rockefeller Foundation.

Received July 27, 1992; accepted August 12, 1992.

REFERENCES

Bolton, E.T., and McCarthy, B.J. (1962). A general method for the isolation of RNA complementary to DNA. Proc. Natl. Acad. Sci. USA **48,** 1390–1397.

Breslauer, K.J., Frank, R., Blocker, H., and Marky, L.A. (1986). Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. USA **83,** 3746–3750.

Brink, R., and Nilan, R. (1952). The relation between light variegated and medium variegated pericarp in maize. Genetics **37,** 519–544.

Calvi, B.R., Hong, T.J., Findley, S.D., and Gelbart, W.M. (1991). Evidence for a common evolutionary origin of inverted repeat transposons in *Drosophila* and plants: *hobo*, *Activator*, and *Tam3*. Cell **66,** 465–471.

Chandler, V.L., and Hardeman, K.J. (1992). The *Mu* elements of *Zea mays*. Adv. Genet. (in press).

Coen, E.S., Robbins, T.P., Almeida, J., Hudson, A., and Carpenter, R. (1989). Consequences and mechanisms of transposition in *Antirrhinum majus*. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, DC: American Society of Microbiology), pp. 413–436.

Cook, W.B., and Walker, J.C. (1992). Identification of a maize nucleic acid-binding protein (NBP) belonging to a family of nuclear-encoded chloroplast proteins. Nucl. Acids Res. **20,** 359–364.

Coupland, G., Baker, B., Schell, J., and Starlinger, P. (1988). Characterization of the maize transposable element *Ac* by internal deletions. EMBO J. **7,** 3653–3659.

Das, P.O., Poliak, E., Ward, K., and Messing, J. (1991). A new allele of the duplicated 27kD zein locus of maize generated by homologous recombination. Nucl. Acids Res. **19,** 3325–3330.

Deininger, P. (1989). SINEs: Short interspersed repeated DNA elements in higher eucaryotes. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, DC: American Society of Microbiology), pp. 619–636.

Devereaux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. **12**, 387–395.

Dreyfus, D.H., and Emmons, S.C. (1991). A transposon-related palindromic repetitive sequence from *C. elegans.* Nucl. Acids Res. **19**, 1871–1877.

Fedoroff, N. (1989). About maize transposable elements. Cell **56**, 181–191.

Fedoroff, N., Wessler, S., and Shure, M. (1983). Isolation of the transposable maize controlling elements *Ac* and *Ds.* Cell **35**, 235–242.

Ferris, P.J. (1989). Characterization of a *Chlamydomonas* transposon, *Gulliver,* resembling those in higher plants. Genetics **122**, 363–377.

Finnegan, D.J. (1989). Eukaryotic transposable elements and genome evolution. Trends Genet. **5**, 103–107.

Freier, S., Kierzek, R., Jaeger, J., Sugimoto, N., Caruthers, M., Neilson, T., and Turner, D. (1986). Improved free-energy parameters for prediction of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**, 9373–9377.

Gerats, A.G.M., Huits, H., Vrijlandt, E., Maraña, C., Souer, E., and Beld, M. (1990). Molecular characterization of a nonautonomous transposable element (*dTph1*) of petunia. Plant Cell **2**, 1121–1128.

Gierl, A., and Saedler, H. (1992). Plant-transposable elements and gene tagging. Plant Mol. Biol. **19**, 39–49.

Gierl, S.R., Lutticke, S., and Saedler, H. (1988). *TnpA* product encoded by the transposable element *En-1* of *Zea mays* is a binding protein. EMBO J. **7**, 4045–4053.

Grandbastien, M. (1992). Retroelements in higher plants. Trends Genet. **8**, 103–108.

Grandbastien, M., Spielman, A., and Caboche, M. (1989). *Tnt1*, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature **337**, 376–380.

Hansen, L., and Von Wettstein-Knowles, P. (1991). The barley genes *Acl1* and *Acl3* encoding the acyl carrier proteins I and III are located on different chromosomes. Mol. Gen. Genet. **229**, 467–478.

Herrmann, A., Schulz, W., and Hahlbrock, K. (1988). Two alleles of the single-copy chalcone synthase gene in parsley differ by a transposon-like element. Mol. Gen. Genet. **212**, 93–98.

Hoffman-Liebermann, B., Liebermann, D., and Cohen, S. (1989). *TU* elements and *Puppy* sequences. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, DC: American Society of Microbiology), pp. 575–592.

Klein, A.S., Clancy, M., Paje-Manalo, L., Furtek, D.B., Hannah, L.C., and Nelson, O.E. (1988). The mutation *bronze-mutable 4 Derivative 6856* in maize is caused by the insertion of a novel 6.7-kilobase pair transposon in the untranslated leader region of the *bronze-1* gene. Genetics **120**, 779–790.

Klosgen, R., Gierl, A., Schwarz-Sommer, Z., and Saedler, H. (1986). Molecular analysis of the *waxy* locus of *Zea mays.* Mol. Gen. Genet. **203**, 237–244.

Koster-Topfer, M., Frommer, W.B., Rocha-Sosa, M., and Willmitzer, L. (1990). Presence of a transposon-like element in the promoter region of an inactive patatin gene in *Solanum tuberosum* L. Plant Mol. Biol. **14**, 239–247.

Kunze, R., and Starlinger, P. (1989). The putative transposase of transposable element *Ac* from *Zea mays* L. interacts with subterminal sequences of *Ac.* EMBO J. **8**, 3177–3185.

Ludwig, S.R., Habera, L.F., Dellaporta, S.L., and Wessler, S.R. (1989). *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region. Proc. Natl. Acad. Sci. **86**, 7092–7096.

MacRae, A.F., and Clegg, M.T. (1992). Evolution of *Ac* and *Ds1* elements in selected grasses (Poaceae). Genetica (in press).

Nacken, W.K.F., Piotrowiak, R., Saedler, H., and Sommer, H. (1991). The transposable element *Tam1* from *Antirrhinum majus* shows structural homology to the maize transposon *En/Spm* and has no sequence specificity of insertion. Mol. Gen. Genet. **228**, 201–208.

Nelson, O.E. (1968). The *waxy* locus in maize. II. The location of the controlling element alleles. Genetics **60**, 507–524.

Nelson, O.E. (1976). Previously unreported *wx* heteroalleles. Maize Genet. Coop. Newslett. **50**, 109–113.

Osterman, J.C., and Dennis, E.S. (1989). Molecular analysis of the *ADH1-C^{m}* allele of maize. Plant Mol. Biol. **13**, 203–212.

Pelemen, J., Cottyn, B., Van Camp, W., Van Montagu, M., and Inze, D. (1991). Transient occurrence of extrachromosomal DNA of an *Arabidopsis thaliana* transposon-like element, *Tat1*. Proc. Natl. Acad. Sci. USA **88**, 3618–3622.

Qu, R., and Huang, A. (1990). Oleosin KD 18 on the surface of oil bodies in maize. J. Biol. Chem. **265**, 2238–2243.

Quigley, F., Brinkmann, H., Martin, W.F., and Cerff, R. (1989). Strong functional GC pressure in a light regulated maize gene encoding subunit GAPA of chloroplast glyceraldehyde-3-phosphate dehydrogenase: Implications for the evolution of GAPA pseudogenes. J. Mol. Evol. **29**, 412–421.

Ralston, E.J., English, J.J., and Dooner, H.K. (1988). Sequence of three *bronze* alleles of maize and correlation with the genetic fine structure. Genetics **119**, 185–197.

Reina, M., Ponta, I., Guillen, P., Boronet, A., and Palau, J. (1990). Sequence analysis of a genomic clone encoding a Zc2 protein from *Zea mays* W64A. Nucl. Acids Res. **18**, 6426.

Rivin, C.J., Cullis, C.A., and Walbot, V. (1986). Evaluating variation in the genome of *Zea mays.* Genetics **113**, 1009–1019.

Sachs, M.M., Dennis, E.S., Gerlach, W.L., and Peacock, W.J. (1986). Two alleles of maize *alcohol dehydrogenase 1* have 3′ structural and poly(A) addition polymorphisms. Genetics **113**, 449–467.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Sanger, F., Nicklen, S., and Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**, 5463–5467.

Shepherd, N.S., Schwarz-Sommer, Z., Blumberg vel Spalve, J., Gupta, M., Weinand, U., and Saedler, H. (1984). Similarity of the *Cin1* repetitive family of *Zea mays* to eukaryotic transposable elements. Nature **307**, 185–187.

Smith, P., and Corces, V. (1991). Drosophila transposable elements: Mechanisms of mutagenesis and interactions with the host genome. Adv. Genet. **29**, 229–300.

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.

Vieira, J., and Messing, J. (1987). Production of single-stranded plasmid DNA. Methods Enzymol. **153,** 3–11.

Weil, C.F., and Bureau, T.E. (1992). Preparation of genomic libraries in lambda phage. In The Maize Handbook, M. Freeling and V. Walbot, eds (New York: Springer-Verlag Press) (in press).

Wessler, S.R., and Varagona, M.J. (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.

Yu, L., and Lazarus, C.M. (1991). Structure and sequence of an auxin-binding protein gene from maize (*Zea mays* L.). Plant Mol. Biol. **16,** 925–930.

Zhao, X., and Kochert, G. (1992). Characterization and genetic mapping of a short, highly repeated, interspersed DNA sequence from rice (*Oryza sativa* L.). Mol. Gen. Genet. **231,** 353–359.

Zhao, X., Wu, T., Xie, Y., and Wu, R. (1989). Genome-specific repetitive sequences in the genus *Oryza*. Theor. Appl. Genet. **78,** 201–209.

Zuker, M., and Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxillary information. Nucl. Acids Res. **9,** 133–148.