Molecular Evolution of the *Ac/Ds* Transposable-Element Family in Pearl Millet and Other Grasses

Gavin A. Huttley,* Amy F. MacRae† and Michael T. Clegg*

** Department of Botany and Plant Sciences, University of California, Riverside, California 92521, and † Department of Biology, University of Missouri, St. Louis, Missouri 63121*

Manuscript received May 18, 1994
Accepted for publication November 15, 1994

ABSTRACT

We report an *Ac*-like sequence from pearl millet (*Pennisetum glaucum*) and deletion derivative *Ac*-like sequences from pearl millet and another grass species, *Bambusa multiplex*. Sequence relationships between the pearl millet and maize *Ac* elements suggest the *Ac/Ds* transposable-element family is ancient. Further, the sequence identity between the Bambusa *Ac*-like sequence and maize *Ac* implies that the *Ac/Ds* transposable-element family has been in the grass family since its inception. The *Ac*-like sequences reported from pearl millet and maize *Ac* are statistically heterogeneous in pair-wise distance comparisons to each other. Yet, we are unable to discriminate between differential selection or ectopic exchange (recombination and conversion) between nonidentical transposable element homologues, as the cause of the heterogeneity. However, the more extreme heterogeneity exhibited between the previously described pearl millet element and maize *Ac* seems likely to derive from ectopic exchange between elements with different levels of divergence.

TRANSPOSABLE elements occupy a significant proportion of virtually all eukaryotic genomes. These mobile genetic elements were first detected as a consequence of their ability to disrupt the expression of genes and produce unstable phenotypes (McCLINTOCK 1948). It was subsequently discovered that transposable elements are also capable of independent replication and insertion (for reviews see BERG and HOWE 1989). These observations have prompted considerable efforts to determine the principal forces responsible for the maintenance of transposable elements in natural populations.

It is currently believed that transposable elements are maintained in populations as a balance between transposition-related copy number increase and some opposing force(s) (CHARLESWORTH and LANGLEY 1989). A possible mechanism for containing transposable element copy number is ectopic recombination where, in some circumstances, recombination between transposable elements at nonhomologous genomic locations gives rise to deleterious chromosomal rearrangements (LANGLEY *et al.* 1988). The hypothesis assumes that the frequency of deleterious rearrangements increases with increasing copy number, resulting in selection against individuals with high copy number. This hypothesis predicts a specific genomic distribution of transposable elements and has been partially supported by observations in Drosophila (LANGLEY *et al.* 1988; CHARLESWORTH and LAPID 1989; MONTGOMERY *et al.* 1991) and in mice (BAKER and WICHMAN 1990). Importantly, if the elements involved in ectopic recombination events are nonidentical homologues, then the recombinant transposable elements will contain regions with different genealogies. Therefore, because not all ectopic recombination events will be deleterious, the occurrence of ectopic recombination can be assayed by sequence comparisons between transposable-element family members. The implication of detecting nondeleterious rearrangements is that they support the potential involvement of ectopic recombination in limiting the spread of the transposable element family under study. It is noteworthy, however, that gene conversion during ectopic pairing will also generate elements with genealogically distinct regions. Although the potential evolutionary consequences of gene conversion versus ectopic recombination are obviously different, a mechanism whereby gene conversion may also serve to limit the spread of transposable elements is outlined below. Events that could occur by either ectopic recombination or gene conversion will hereafter be referred to as ectopic exchange events.

Not all members of a transposable-element family are autonomous. For example, *Ds* elements in the maize genome are capable of transposition only in the presence of a putative transposase provided by an autonomous *Ac* element. *Ds* elements are therefore referred to as nonautonomous. Interestingly, almost all eukaryotic transposable-element families described have nonautonomous members (HARTL *et al.* 1992). Ectopic exchange is a potential mechanism by which nonautonomous elements may originate from autonomous ones. For example, ectopic exchange between autonomous and nonau-

Corresponding author: Gavin A. Huttley, Laboratory of Viral Carcinogenesis, National Cancer Institute, Bldg. 560, FCRDC, Frederick, MD 21702. E-mail: huttley@fcrfv1.ncifcrf.gov

tonomous elements might result in two nonautonomous elements, depending on the location of the exchange and the number of mutations defining the nonautonomous element. Eliminating autonomous elements presumably serves to decrease the ability of a transposable element family to increase in copy number.

As phylogenetic surveys have become available, observations of horizontal transfer in Drosophila and other taxa (MARUYAMA and HARTL 1991; PASCUAL and PERIQUET 1991; FLAVELL 1992) have become common. It has been suggested that the loss of autonomous elements is inevitable in a finite population, implying that horizontal transfer may be the only force preventing the ultimate death of a transposable-element family (HARTL *et al.* 1992). Consequently, transposable-element families are seen as transient members of eukaryotic genomes that are dependent on horizontal transfer for replenishment. The alternative view is that transposable elements are stable residents of eukaryotic genomes and are only transmitted vertically. The vertical transmission of a transposable-element family can be evaluated by sampling from a distantly related species and confirming the presence of the element family in that species. If the sequence relationships between the elements from the different taxa are consistent with the phylogenetic distance between those taxa, then this may be interpreted as support for the vertical transmission of the transposable-element family.

Most research aimed at elucidating the forces that influence eukaryotic transposable element evolution has been conducted in animal and fungal systems (see BERG and HOWE 1989). At present comparable studies of plant transposable elements are lacking, despite the fact that the first transposable-element family described was the *Ac/Ds* family of *Zea mays* (MCCLINTOCK 1948). Although the *Ac/Ds* family is perhaps the best characterized plant transposable-element family at the molecular level, evolutionary analysis of this family has received little attention (GERLACH *et al.* 1987; MACRAE and CLEGG 1992). Furthermore, few reports exist of *Ac*-like sequences in other plant taxa (CHERNYSHEV *et al.* 1988; MACRAE *et al.* 1994).

In this study we sample from a representative of the Bambusa genus, *Bambusa multiplex*, that is a very distant relative of Zea and Pennisetum (MORTON and CLEGG, 1994) and ask if there are *Ac*-like sequences in the genome of *B. multiplex*? Evidence for the *Ac/Ds* transposable-element family in Bambusa is described, implying a long evolutionary history of *Ac/Ds* transposable elements in the grass family. Furthermore, we sequence three additional pearl millet elements to address the following questions: (1) Is there evidence of ectopic exchange in comparisons between the different pearl millet elements? (2) What are the relationships between the pearl millet elements? Evidence from sequences of pearl millet elements consistent with the occurrence of ectopic exchange is presented.

MATERIALS AND METHODS

Plant materials, library construction and screening: The *B. multiplex* and *Pennisetum glaucum* genomic DNA samples prepared are described elsewhere (GEPTS and CLEGG 1989; DUVALL *et al.* 1993). The genomic library was made from *P. glaucum*, strain Tifton, GA DB. The construction and screening of this library is described in MACRAE and CLEGG (1992).

Nomenclature: To distinguish between the previously reported *Ac*-like sequence from pearl millet (MACRAE *et al.* 1994) and the additional *Ac*-like sequences characterized in this paper, we will use the abbreviations described in Table 1. In summary, the first letters of the genus and species names, *Pg* representing *P. glaucum*, and *Zm* representing *Z. mays*, will precede the element class designation. The *Ac*-L class refers to these sequences being *Ac*-like. All elements amplified by PCR using the ZmAcF and ZmAcR primers (described below) are referred to as *PgDs* elements. Although the *PgDs* elements are also considered *Ac*-like, the distinction between these two classes of elements is made to reflect the fact that *PgAc*-L1 (MACRAE *et al.* 1994), and possibly *PgAc*-L2, do not have inverted terminal repeats. The high stringency PCR conditions and primer design suggest that the *PgDs* elements do have these terminal repeats.

Cloning and sequencing of the *PgAc*-L2 element: Two primers designed for sequencing of the previously reported *Ac*-like element (MACRAE *et al.* 1994) were used in PCR experiments to determine whether related sequences existed in previously-isolated putative positive genomic clones. The reaction profile used is described below. A product from these reactions was cloned using the TA cloning kit from Invitrogen (San Diego, CA) according to manufacturers specifications. Additional regions of this element were obtained by inverse PCR (OCHMAN *et al.* 1990) from an *NdeI* digest of the original genomic clone. The single inverse PCR product was cloned as above. Both strands of these clones were sequenced only to the point at which they overlapped the central region of identity between *PgAc*-L1 and *ZmAc*. The sequence was obtained using both universal primers and primers designed from the element. All sequencing was performed on plasmid DNA using either standard procedures (SANGER *et al.* 1977) or the fmol kit from Promega (Madison, WI) according to manufacturers specifications.

PCR reactions, cloning and sequencing of *Ac*-like sequences: PCR reactions were performed using primers designed from the reported sequence of the *Ac* element of *Zea mays* (KUNZE *et al.* 1987). The 25-mer primers include the 11-bp inverted terminal repeats and an additional 14 bp of internal *ZmAc* sequence. These primers will be subsequently referred to as ZmAcF and ZmAcR. Reaction conditions for all amplifications (except for Bambusa) were as follows: 30 cycles of 94° for 1 min, 60° for 1 min, 72° for 5 min; 1 cycle of 72° for 15 min. The Bambusa products were obtained as above except the annealing temperature was 50°. The amount of genomic DNA used for the amplifications was 0.15–1 μg. Negative controls were run with each PCR reaction, and all reactions were set up using aerosol resistant tips. All PCR products were cloned as described above.

The *PgDs* elements whose sequences we report here were obtained from a wild-type individual from Sudan (85-2). Both strands of the clones were sequenced as described above, with the exception that elements from Bambusa were only partially characterized at the sequence level for direct confirmation that they were *Ac*-like sequences.

Sequence alignments and relationships between the pearl millet elements: Sequences were aligned using either the BESTFIT or GAP algorithms (DEVEREUX *et al.* 1984) or the CLUSTAL alignment procedure from the Genetic Data Environment package version 2.0 (SMITH 1993). The region se-

TABLE 1

Table of nomenclature and origins of elements used

Species	Element Class	Origin	Abbreviation	Reference
Z. mays	Autonomous	Genomic clone	*ZmAc*	KUNZE *et al.* (1987)
P. glaucum	*Ac*-like	Genomic clone	*PgAc*-L1	MACRAE *et al.* (1994)
	Ac-like	Genomic clone/PCR	*PgAc*-L2	This paper
	Ac-like	PCR	*PgDs1*	This paper
	Ac-like	PCR	*PgDs2*	This paper

The division of the pearl millet elements into two different classes (*PgAc*-L, *PgDs*) was based on the presence/absence of the 11-bp inverted terminal repeats (both repeats are probably present in the *PgDs* elements).

quenced from the *PgAc*-L2 element did not overlap with either of the *PgDs* elements. Consequently, alignments between the *PgDs1, PgDs2, PgAc*-L1 (MACRAE *et al.* 1994) and *ZmAc* elements (KUNZE *et al.* 1987) were generated. Given the constraints imposed by the degree to which these sequences have diverged, the alignments that involved all elements were limited to sequences that spanned 203 bp of the 5′-untranslated regions (*ZmAc* coordinates 35–232, see Appendix, Figure 2) and 130 bp of the 3′-untranslated regions (*ZmAc* coordinates 4326–4445, see Appendix, Figure 3). The sequences of *PgAc*-L1 and *ZmAc* were trimmed to the size of *PgAc*-L2 for comparisons involving these elements (*ZmAc* coordinates 1328–3708, see Appendix, Figure 4). The resulting alignments were used to determine relationships between the various elements and in the statistical analyses described below.

Evaluating the relationships between the different elements from pearl millet was based on pair-wise distances calculated between the elements from each class (*i.e., PgAc*-L and *PgDs*) and with the *ZmAc* element. Kimura's two-parameter distance measure was used for these calculations (KIMURA 1980). For determining distances between the *PgAc*-L elements and *ZmAc,* only the central region was used (*ZmAc* coordinates 1587–3438). Synonymous distances were calculated between the *PgAc*-L elements and *ZmAc.* The *ZmAc* sequence was used as the reading frame reference, and the region compared between the elements lay within the central region defined above (see Appendix, Figure 4). For comparisons to *ZmAc,* the second intron was deleted. A small section that contained frameshift mutations was deleted from the center of this region. Distances were calculated using MEGA version 1.0 (KUMAR *et al.* 1993). Divergence times were calculated using a substitution rate of 7.9×10^{-9} synonymous substitutions per site per year estimated from *ADH1* comparisons between maize and pearl millet (GAUT and CLEGG 1991).

Statistics: A log likelihood heterogeneity test was used to determine whether there are differences in substitution rates along the sequences (GAUT and CLEGG 1993). For the comparisons involving both the *PgAc*-L elements and *ZmAc,* the test was performed by dividing the sequences that could be unambiguously aligned into three unequally sized regions: a 305-bp block representing *ZmAc* coordinates 1328–1594 (referred to as region A), a block of 1830 bp representing *ZmAc* coordinates 1595–3408 (referred to as region B) and a 305-bp block representing *ZmAc* coordinates 3409–3708 (referred to as region C) (see Figure 1). All three possible pair-wise comparisons were performed.

The distribution of substitutions among synonymous and nonsynonymous sites between the *PgAc*-L elements was evaluated to test the hypothesis that their most recent common ancestor was nonautonomous. The region compared corresponds to that used to calculate synonymous distances described above (see Appendix, Figure 4). Stop codons that occurred were not considered in the calculations. Estimates of the numbers of synonymous and nonsynonymous sites were obtained, and substitutions were classified using the unweighted pathways method (NEI and GOJOBORI 1986).

RESULTS

Partial characterization of *PgAc*-L2: The *PgAc*-L2 sequence has been submitted to GenBank (accession number U17068). Although we present only a partial characterization of this element, it is apparent that it shares some striking similarities to the *PgAc*-L1 element. The sequenced region of *PgAc*-L2 is 2337 bp long. The relationships between *PgAc*-L2, *PgAc*-L1 and *ZmAc* are shown in Figure 1 and the Appendix (Figure 4). The regions of identity between *PgAc*-L2 and *ZmAc* designated in Figure 1 are based on a BESTFIT alignment. The region of identity shared between these sequences is almost equivalent to that shared between *PgAc*-L1 and *ZmAc,* with the exception that *PgAc*-L2 has an additional 26 bp of identity at the 5′-end and an additional 143 bp of identity at the 3′-end. The percent identity

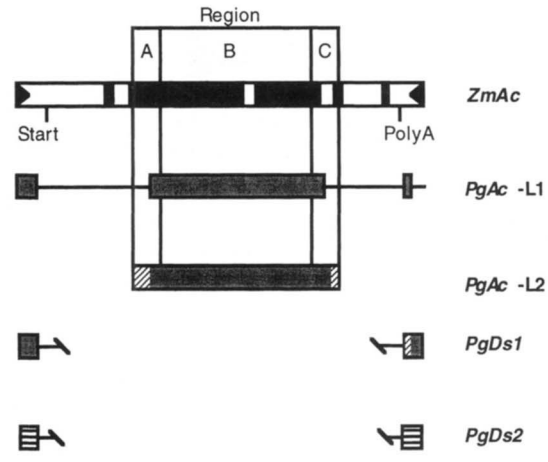

FIGURE 1.—Relationships between the *PgAc*-L, *PgDs* and *ZmAc* elements. The presence and size of the regions of identity were determined using the BESTFIT algorithm. Regions A, B and C indicate partitions used in the heterogeneity analyses. Start, major transcription start site; PolyA, poly(A) addition site; ■, *ZmAc* exons; □, *ZmAc* introns; ▒, >60% identity to *ZmAc;* ▨, >60% identity to *PgAc*-L1; ▩, >64% identity to *PgDs1*; —, <42% identity; ▶◀, inverted terminal repeats.

to *ZmAc* for this region is ~70%, an almost identical value to that between *PgAc*-L1 and *ZmAc*.

The identity between the two *PgAc*-L elements is ~78%, which spans the entire sequence of *PgAc*-L2. There are a total of seven insertion or deletion events, six of which occur within 150 bp of the limits of the *PgAc*-L2 sequence. The likelihood that the *PgAc*-L2 element might be functional is precluded by the occurrence of several stop codons and a frameshift mutation in the region that corresponds to the *ZmAc* open reading frame. In addition, a cluster of deletions relative to *ZmAc* occurs around the second intron. The deletions are similar to those reported for *PgAc*-L1 (MACRAE *et al.* 1994).

Features of the *PgDs* elements: The sequences for *PgDs1* and *PgDs2* have been submitted to GenBank (accession numbers U17069-U17070). The stringent amplification conditions strongly suggest that the *PgDs* elements have relatively intact inverted terminal repeats and thus may retain the potential for transposition when catalyzed by the transposase from an autonomous element. The *PgDs1* element is 868 bp in length, and *PgDs2* is 719 bp in length (these lengths do not include the primers). The relationships between these elements and *ZmAc* are illustrated in Figure 1. Although the length of the regions assumed to be homologous and the degree of sequence identity relative to *ZmAc* differs among *PgDs* elements, the terminal regions of both of these elements are similar to the flanking regions of *ZmAc*. The latter result is not surprising because the ZmAcF and ZmAcR primers would not amplify all possible types of elements equally, given the inclusion of internal *ZmAc* sequence in these primers. The central regions of the *PgDs* elements have very low identity to *ZmAc* or to each other. The presence of regions with no obvious relationship to *ZmAc* in nonautonomous elements has been previously reported (see FEDOROFF 1989; MACRAE *et al.* 1994). The relationships between the *PgDs* and *ZmAc* elements indicate that these elements are probably deletion-derivatives of *Ac*.

An *Ac*-like sequence from *B. multiplex*: Amplified products were obtained from *B. multiplex*. One of the products (data available on request) had regions of identity to *ZmAc* that were consistent with those observed for the *PgDs* elements. The total size of the latter product was ~760 bp. The region directly internal to the 5′-inverted terminal repeat had ~63.8% identity for 108 bp, whereas the region directly internal to the 3′-inverted terminal repeat had ~55.2% identity for 211 bp. The genetic distance between these two combined regions to the *ZmAc* element is ~0.7204. The size and pattern of sequence identity to *ZmAc* suggest that this element is a deletion-derivative of *Ac*.

***Ac*-like sequences from pearl millet are statistically heterogeneous:** Ectopic exchange between nonidentical transposable-element homologues may result in elements with genealogically distinct regions. The potential occurrence of ectopic exchange can therefore be evaluated by testing for distance heterogeneity among regions between elements.

TABLE 2

Log-likelihood ratio heterogeneity test for distance heterogeneity between three unequal regions in the *PgAc*-L1, *PgAc*-L2 and *ZmAc* comparisons

Pair-wise comparison	*p* per region			LR-statistic
	A	B	C	
PgAc-L1 *vs. PgAc*-L2	0.4127	0.2686	0.3989	20.80*
PgAc-L1 *vs. ZmAc*	0.5818	0.4202	0.7054	53.69*
PgAc-L2 *vs. ZmAc*	0.5254	0.4000	0.6791	44.86*

Regions A and C are the terminal 305 bp of the compared aligned sequences. Region A has *ZmAc* coordinates of 1328–1594, and region C has *ZmAc* ccordinates of 3409–3708. The central region B is 1830 bp (*ZmAc* coordinates 1595–3408). *p* is an estimate of the number of substitutions per site; LR-statistic is distributed chi square with 2 d.f. * Significant at the 1% level.

The aligned sequences in the comparisons involving *PgAc*-L1, *PgAc*-L2 and *ZmAc* were partitioned into three unequal regions: region A of 305 bp, region B of 1830 bp and region C of 305 bp. The results of the likelihood ratio heterogeneity tests (Table 2) indicate that all three pair-wise comparisons are significant. Interestingly, there is no significant difference between the distances for regions A and C in the *PgAc*-L1 to *PgAc*-L2 comparison (result not shown). The results from these tests indicate that the *PgAc*-L elements have a nonuniform distribution of substitutions relative both to each other and to *ZmAc*. Major heterogeneity is also obvious from the comparison between *PgAc*-L1 and *ZmAc*, as indicated by the regions of very low identity.

Relationships between the pearl millet elements: The alignments between the *PgAc*-L elements and between the *PgDs* elements indicate that the pairs of elements within these groups are most closely related to each other. The distances between these elements based on the 5′-region are presented in Table 3. Synonymous distances calculated from the central regions in the *PgAc*-L to *ZmAc* comparisons are shown in Table 4.

A test was performed to determine whether the most recent common ancestor of region B from the *PgAc*-L elements was nonautonomous. Because we expect

TABLE 3

DNA distance matrix between the aligned 5′ ends of *PgDs*, *PgAc*-L1 and *ZmAc* elements

	PgDs1	*PgDs2*	*PgAc*-L1	*ZmAc*
PgDs1	—	0.4660	0.7937	0.6357
PgDs2		—	0.6867	0.5104
PgAc-L1			—	0.7278
ZmAc				—

TABLE 4

Matrix of synonymous distance estimates (above the diagonal) and their standard errors (below the diagonal) between the *PgAc*-L and *ZmAc* elements

	PgAc-L1	*PgAc*-L2	*ZmAc*
PgAc-L1	—	0.88	1.77
PgAc-L2	0.08	—	1.56
ZmAc	0.26	0.20	—

The region tested corresponds roughly to the central region (see text and APPENDIX, Figure 4).

replacement substitutions to be unconstrained between nonautonomous elements, substitutions should be uniformly distributed between synonymous and nonsynonymous sites. Alternatively, an excess of synonymous substitutions implies that the coding regions of the most recent common ancestor were subject to selective constraint and that the ancestor may have been autonomous for a significant period of time before becoming nonautonomous. When a Chi-square test is performed on the occurrence of substitutions between synonymous and nonsynonymous sites for region B of the *PgAc*-L elements, there is a highly significant excess, roughly fivefold, of synonymous substitutions (result not shown).

DISCUSSION

The pattern of relationships revealed through the analyses of the pearl millet and maize elements show that there is significant distance heterogeneity within individual pearl millet elements. Heterogeneity among regions is consistent with the occurrence of ectopic exchange. However, the data do not exclude several alternative mechanisms that may also have contributed to the observed heterogeneity in some of these comparisons.

One alternative mechanism that could contribute to heterogeneity among regions of the *PgAc*-L/*ZmAc* elements is selection. For example, the large proportion of intron sites (≈60%) in region C may account for the higher divergence of this region in some comparisons, owing to weak selective constraints in intron regions. In contrast, region A is exclusively coding but has the same level of divergence as region C between the *PgAc*-L elements. Assuming that the *PgAc*-L elements diverged from an autonomous *Ac* element, as suggested by the constraint on missense substitutions, we expect regions A and B to have equivalent divergence. This expectation is not supported by the data (Table 2). This result could arise from either reduced selective constraint for region A in autonomous elements or ectopic exchange between elements with different levels of divergence. We are unable to eliminate either of these hypotheses as a partial explanation for the heterogeneity across regions.

There is a more extreme form of heterogeneity represented by the very low sequence identity in the intervals separating the 5′- and 3′-blocks from the central block in *PgAc*-L1 depicted in Figure 1. The remarkable feature of this pattern is the positional conservation of the 5′- and 3′-blocks relative to *ZmAc*. Such conservation requires either an insertion of precisely the correct length, or more likely, an ectopic exchange event. Ectopic exchange between elements of different levels of divergence appear to be the most likely cause of the extreme heterogeneity of *PgAc*-L1.

It is of interest to estimate the age of the relationships between the pearl millet and *ZmAc* elements. The estimation of divergence time is complicated by the heterogeneity discussed above, however, using the synonymous substitution rate estimated from *ADH1* comparisons (GAUT and CLEGG 1991) and the synonymous distances calculated for the *PgAc*-L, it appears that region B of the two *PgAc*-L elements diverged from each other ~55 mya (±5 my). This result dates their most recent common ancestor as occurring before the divergence of the Zea and Pennisetum lineages (about 25–30 mya). The divergence of the *PgAc*-L elements from *ZmAc* cannot be estimated with any confidence because all synonymous sites are saturated. In addition, divergence times between the *PgDs* and other elements are unreliable given the small regions of identity. Nevertheless, the data suggest that the *Ac/Ds* family is an ancient component of plant genomes.

The finding of an *Ac*-like element in Bambusa further supports the ancient origins of the *Ac/Ds* family. An estimate of the specific divergence time between this element and *ZmAc* would be unreliable because of the small regions of identity. Yet, it is clear that the Bambusa sequence is considerably diverged from *ZmAc* and to a greater extent than either of the *PgDs* elements.

The evidence from the *PgAc*-L1/*ZmAc* comparison, with equivocal support from the remaining *PgAc*-L/*ZmAc* comparisons, suggests that recombination or gene conversion can take place between nonidentical transposable-element homologues. This is consistent with the hypothesis that ectopic exchange can generate nonautonomous elements. The data also suggest that blocks of very low identity found in nonautonomous elements may be the "bones" of ancient ectopic exchange events.

We thank BRANDON GAUT and BRIAN MORTON for assistance in performing the statistical analyses and for their helpful comments, GERALD H. LEARN JR. for his computing assistance, and A. H. D. BROWN, A. PRYOR and two anonymous reviewers for their helpful comments. This work was supported by National Institutes of Health grant GM-45144.

LITERATURE CITED

BAKER, R. J., and H. A. WICHMAN, 1990 Retrotransposon *mys* is concentrated on the sex chromosomes: implications for copy number containment. Evolution **44:** 2083–2088.

BERG, D. E., and M. M. HOWE, 1989 *Mobile DNA.* ASM Press, Washington, DC.

CHARLESWORTH, B., and C. H. LANGLEY, 1989 The population genetics of *Drosophila* transposable elements. Annu. Rev. Genet. **23:** 251–287.

CHARLESWORTH, B., and A. LAPID, 1989 A study of ten families of transposable elements on X chromosomes from a population of *Drosophila melanogaster.* Genet. Res. Camb. **54:** 113–125.

CHERNEYSHEV, A. I., M. V. GOLOVKIN, N. V. MIL'SHINA, A. K. GAZUMYAN and E. V. ANAN'EV, 1988 Molecular genetic organization of mobile elements of the *Ac/Ds* family in cereal genomes: Identification of DNA sequences homologous to the *Ac* element of maize in Barley (*Hordeum vulgare* L.) genome. Genetika **11:** 1918–1927.

DEVEREUX, J., P. HAERBERLI and O. SMITHIES, 1984 A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:** 387–395.

DUVALL, M. R., G. H. LEARN, JR., L. E. EGUIARTE and M. T. CLEGG, 1993 Phylogenetic analysis of *rbcL* sequences identifies *Acorus calamus* as the primal extant monocotyledon. Proc. Natl. Acad. Sci. USA **90:** 4641–4644.

FEDOROFF, N. V., 1989 Maize Transposable Elements, pp. 375–411 in *Mobile DNA,* edited by M. HOWE and D. BERG. ASM Press, Washington, DC.

FLAVELL, A. J., 1992 *Ty1-copia* group retrotransposons and the evolution of retroelements in the eukaryotes. Genetica **86:** 203–214.

GAUT, B. S., and M. T. CLEGG, 1991 Molecular evolution of alcohol dehydrogenase 1 in members of the grass family. Proc. Natl. Acad. Sci. USA **88:** 2060–2064.

GAUT, B. S., and M. T. CLEGG, 1993 Molecular evolution of the *ADH1* locus in the genus *Zea.* Proc. Natl. Acad. Sci. USA **90:** 5095–5099.

GEPTS, P., and M. T. CLEGG, 1989 Genetic diversity in pearl millet (*Pennisetum glaucum* [L] R. Br.) at the DNA sequence level. J. Heredity **80:** 203–208.

GERLACH, W. L., E. S. DENNIS, W. J. PEACOCK and M. T. CLEGG, 1987 The *Ds1* controlling element family in maize and *Tripsicum.* J. Mol. Evol. **26:** 329–334.

HARTL, D. L., E. R. LOZOVSKAYA and J. G. LAWRENCE, 1992 Nonautonomous transposable elements in prokaryotes and eukaryotes. Genetica **86:** 47–48.

KIMURA, M., 1980 A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16:** 111–120.

KUMAR, S., K. TAMURA and M. NEI, 1993 *MEGA: Molecular Evolutionary Genetics Analysis, version 1.0.* The Pennsylvania State University, University Park, PA.

KUNZE, R., U. STOCHAJ, J. LAUFS and P. STARLINGER, 1987 Transcription of transposable element *Activator* (*Ac*) of *Zea mays.* L. EMBO J. **6:** 1555–1563.

LANGLEY, C. H., E. MONTGOMERY, R. HUDSON, N. KAPLAN and B. CHARLESWORTH, 1988 On the role of unequal exchange in the containment of transposable element copy number. Genet. Res. **52:** 223–225.

MACRAE, A. F., and M. T. CLEGG, 1992 Evolution of *Ac* and *Ds1* elements in select grasses (Poaceae). Genetica **86:** 55–66.

MACRAE, A. F., G. A. HUTTLEY and M. T. CLEGG, 1994 Molecular evolutionary characterization of an *Activator* (*Ac*)-like transposable element from pearl millet (*Pennisetum glaucum*) (Poaceae). Genetica **92:** 77–89.

MARUYAMA, K., and D. L. HARTL, 1991 Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus.* J. Mol. Evol. **33:** 514–524.

MCCLINTOCK, B., 1948 Mutable loci in maize. Carnegie Inst. Washington Year Book **47:** 155–169.

MONTGOMERY, E. A., B. CHARLESWORTH and C. H. LANGLEY, 1991 A test for the role of natural selection in the containment of transposable element copy number. Genet. Res. **52:** 223–235.

MORTON, B. M., and M. T. CLEGG, 1994 Substitutional dynamics and sequence composition of different regions of the chloroplast genome. Mol. Biol. Evol. (in press).

NEI, M., and T. GOJOBORI, 1986 Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions. Mol. Biol. Evol. **3:** 418–426.

OCHMAN, H., M. M. MADHORE, D. GARZA and D. L. HARTL, 1990 Amplification of flanking sequences by inverse PCR. pp. 219–227 in *PCR Protocols: A guide to methods and applications,* edited by M. A. INNIS, D. H. GELFAND, J. J. SNINSKY and T. J. WHITE. Academic Press, San Diego.

PASCUAL, L., and G. PERIQUET, 1991 Distribution of *hobo* transposable elements in natural populations of *Drosophila melanogaster.* Mol. Biol. Evol. **8:** 282–296.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SMITH, S., 1993 *Genetic Data Environment. Version 2.0.* Harvard Genome Laboratory. Cambridge, MA.

Communicating editor: A. H. D. BROWN

APPENDIX

Alignments Between *Ac*-like Elements

Name	Sequences				
ZmAc	GACCGTTATC	GTATAACCGA	TTTTGTTAGT	TTTATCCCGA	TCGATTTCGA ACCCGA-GGT
PgDs1	.....AC.C.	T....-..A-	----------	CA.....	..A.A.A.C. GATTTTC..G
PgDs2	.....GC.C.	.....-...-	----------	CG.....T.C	..A.A.A.CG TATT..C...
PgAc-L1	.....GC.C.	.C..C-...-	----------	.C..AA...T	C..TGAA.TG TTTT.TCT.G
ZmAc	AAAAAACGAA	AAC-GGAACG	GAAACGGGAT	ATACAAAACG	GTAAACGGAA ACGGAAAC-G
PgDs1	....T.T...	TT.A....AA	T...TACAGG	.A..CGG-..	..C.CG...T -..A..T.G.
PgDs2	.....CA...	.-.A....AA	TTCC..A...	TA..CGG-.A	....CA.... -.....C.A.
PgAc-L1	....CCA.G.	...T..G.AA	A..CT.-.GA	.A..CGG-..	.G....A.TG -..A..T...
ZmAc	GTAGAGCTAG	TTTCCCGACC	GTTTCACCGG	GATCCCGTTT	TTAATCGGGA TGATCCCGTT
PgDs1	.GGA..GCTT	-....A...T	.AA..CT...	-T..T..C..	..GTG..A.. .AT.....
PgDs2	..G.G.GC.A	A.....	A.A.TCA...	-.....	CC.GC..A.. .TC...A...
PgAc-L1	..T...G.GT	.....	.G..TT....	-T.A.T..A.	.AG.C..... AA....T.C-
ZmAc	--TCGTTACC	GTATTTTCTA	ATT		
PgDs1	CT.AT....G	.G.AA...CC	...		
PgDs2	TT.TC..G.G	.G.A....CC	GC.		
PgAc-L1	..AGCCAG..	.C.ACCA.A.	GCC		

FIGURE 2.—Alignment of 5′-regions of the *PgDs, PgAc*-L1 and *ZmAc* elements. *ZmAc* coordinates 35–232. Nucleotides that differ from *ZmAc* are given. A period indicates identity to *ZmAc,* a dash indicates a gap in the alignment.

```
Name                                    Sequences

ZmAc        GACCGTTATC GTATAACCGA TTTTGTTAGT TTTATCCCGA TCGATTTCGA ACCCGA-GGT
PgDs1       .....AC.C. T....-..A- ---------- CA........ ..A.A.A.C. GATTTTC..G
PgDs2       .....GC.C. .....-...- ---------- CG.....T.C ..A.A.A.CG TATT..C...
PgAc-L1     .....GC.C. .C..C-...- ---------- .C..AA...T C..TGAA.TG TTTT.TCT.G

ZmAc        AAAAAACGAA AAC-GGAACG GAAACGGGAT ATACAAAACG GTAAACGGAA ACGGAAAC-G
PgDs1       ....T.T... TT.A....AA T...TACAGG .A..CGG-.. ..C.CG...T -..A..T.G.
PgDs2       .....CA... .-.A....AA TTCC..A... TA..CGG-.A ....CA.... -.....C.A.
PgAc-L1     ....CCA.G. ...T..G.AA A..CT.-.GA .A..CGG-.. .G....A.TG -..A..T...

ZmAc        GTAGAGCTAG TTTCCCGACC GTTTCACCGG GATCCCGTTT TTAATCGGGA TGATCCCGTT
PgDs1       .GGA..GCTT -....A...T .AA..CT... -T..T..C.. ..GTG..A.. .AT......
PgDs2       ..G.G.GC.A A......... A.A.TCA... -......... CC.GC..A.. .TC...A...
PgAc-L1     ..T...G.GT .......... .G..TT.... -T.A.T..A. .AG.C..... AA....T.C-

ZmAc        --TCGTTACC GTATTTTCTA ATT
PgDs1       CT.AT....G .G.AA...CC ...
PgDs2       TT.TC..G.G .G.A....CC GC.
PgAc-L1     ..AGCCAG.. .C.ACCA.A. GCC
```

FIGURE 3.—Alignment of 3′-regions of the *PgDs*, *PgAc*-L1 and *ZmAc* elements. *ZmAc* coordinates 4326–4445. Nucleotides that differ from *ZmAc* are given. A period indicates identity to *ZmAc;* a dash indicates a gap in the alignment.

```
Name                                    Sequences
                  ↓
ZmAc          GAGAGGCATT TTATTCAGAG TGTAAGCAGT AGTAATG--- ------CAAA TGGTACAGCT
PgAc-L1       .CTG.TGG.G AC.G.GGA.. ..GTG.TG.. G..GG..GTG CTGCTG.TG. CA..GGT...
PgAc-L2       .CTG.TGG.. .G.A.GGC.. ..G...TG.. G...G..GTG AAGAAA..-. CA..CA.--.

ZmAc          AC---AGATC C--GA-GTCA AGATG-ATAT G--GCTATTG TTCATGAACC ACAACCACAA
PgAc-L1       G.TGG....G GTG..A..GG ..C..C.GC. .CT.GGGA.. .GGG..T.G. ..C-GGTGT.
PgAc-L2       C.TGG.-..T GTT..T..-- ---..T.GG. .CT.---A.. ..GG..G.A. .A..GGTGG.

ZmAc          CC-ACAACCA CAACCAGA-A CCACAA---C CA-CAGCCAC -AACCTGA-- -ACC--C---
PgAc-L1       AGG..TG.A. GG.A.T..T. .T..T.ATG. ..T..TTATT G.GGG...TA CG..TG-TGA
PgAc-L2       .AG-..G.A. GGGA.T..T. .TG...ATG. ..T..TTGTG G..GG...TG A...TG.TGA

ZmAc          --GAAGAAGA AGCACC-ACA G--AAGAGGG CAAAGAAGTG CACATCGGAT GTATGGCAGC
PgAc-L1       TT..T..T.. T.TG..TTGT .CA.....C AG..A..... T..T..A... ..G.....T
PgAc-L2       AG.T...-.. T.T...T..T .GG.....C ....A.TA.. ......AA.G ..C...G.T.

ZmAc          A-TTTCACCA AGAAGGAAAT TGAAGTGGAG GTCGATGGAA AGAAATACGT TCAGGTATGG
PgAc-L1       ........T. .....AC.G. GAC.AAA... .AGA....G. ..GT...TCA G..AA.G...
PgAc-L2       .C...G..A. GC.CCAG..- GATCA.T..A .ATA....G. ...CC..TC. G...A.T...
                  ↓⇓
ZmAc          GGACATTGCA ACTTTCCTAA TTGCAAGGCT AAGTATAGGG CTGAGGGTCA TCATGGAACA
PgAc-L1       .C........ .T.A.GA... A.....ACAC ...GGA..AT G..G.A.CA. .T....T...
PgAc-L2       .CT....... .ACA.GA... A..T.GATAC ...GCC..AT G......CA. .T........

ZmAc          AGCGGATTTC GAAATCACTT GAGAACATCA CATAGTTTAG TTAAAGGTCA GTTGTGTCTA
PgAc-L1       .CT..G...T .G.CA..TC. ....GTTG.. ......G.T. .G..G..... .CA.CAG..C
PgAc-L2       .CT..T...T .G.CG...C. A..G...A.. ...G..G.G. .G.....C.. .CAACAA..C

ZmAc          AAAAGTGAAA AGGATCATGG CAAAGACATA AATCTCATTG AGCCTTATAA GTACGATGAA
PgAc-L1       ..GGTA.... ....CAG.TC A..G.....T .C.ACT..A. CA.....C.G ...T......
PgAc-L2       ..GGCA..G. .....A...A A....GG..T .C.G.TG.G. TA.....C.G ...T......

ZmAc          GTGGTTAGCC TAAAGAAGCT TCATTTGGCA ATAATCATGC ATGAATATCC TTTCAATATT
PgAc-L1       .AA..C..TT ...G...AT. .T.C..... ...G.T.... ....G..C.. A........A
PgAc-L2       .A..CC.ATT .G.G....T. .T.C..A... ...G...... ....G..... A.....C..A

ZmAc          GTAGAACATG AGTACTTTGT TGAGTTTGTT AAGTCTCTGC GCCCTCACTT TCCAATAAAG
PgAc-L1       .CT....... ....T..... G.....CA.. .....CT... .T...AGT.. C.....C...
PgAc-L2       TCT..G.... .A........ .........A ..A..CT... .T...AG... C..C....G.

ZmAc          TCCCGTGTCA CTGCTAGAAA ATATATCATG GATTTGTATT TGGAAGAAAA AGAAAAGTTG
PgAc-L1       ..T...A.T. .CAT...G.. TG.A..TT.. .GCA...T.G ..C..... G.........
PgAc-L2       ..T...TGT. ..AT...G.. .G.A..T... ..CA...TC. ..C..... ...G..AC..

ZmAc          TATGGAAAAC TAAAAGATGT TCAGTCTCGC TTCAGTACAA CTATGGATAT GTGGACATCT
PgAc-L1       ....C.T.TT .C...AC... CA.A.....T ..T...G.C. .A.....C.. ......T..A
PgAc-L2       ....C...TT .C...ACCT. G....G.... ..T...G... .......... ......T..A

ZmAc          TGTCAAAATA AGTCATACAT GTGTGTCACC ATCCATTGGA TTGATGATGA TTGGTGTCTC
PgAc-L1       AA........ .......... .........T T.G....... .A......A. ......CA.T
PgAc-L2       AA...G..C. .A.....TT. .......GTT G.......... .CA....CA. C......A.G
```

FIGURE 4.—Alignment of *PgAc*-L1, *PgAc*-L2 and *ZmAc* elements. ↓, partition boundaries; ⇓, region selected for reading frame comparison. The partitions correspond to regions A, B and C with *ZmAc* coordinates, respectively: 1328–1594, 1595–3408, 3409–3708. Nucleotides that differ from *ZmAc* are given. A period indicates identity to *ZmAc;* a dash indicates a gap in the alignment.

```
ZmAc        CAAAAAAGAA TTGTTGGCTT TTTTCATGTT GAAGGGCGCC ACACTGGCCA AAGGTTATCA
PgAc-L1     .......... ....CAAT.. .G.......A A.G..C.... .T.....TAT ..AAC.G..T
PgAc-L2     ..G....... ....CAA... .G.G.....A ........T. ....A..AGC T.A...G...

ZmAc        CAAACCTTCA CTGCAATCAT GGTTAAGTGG AACATTGAGA AAAAATTGTT TGCCTTGTCT
PgAc-L1     G....A.... ..AA.C...G ...C..A... T.T.....C. .G..G..... .T.TC.CA.A
PgAc-L2     G.G....... ...A.C.... .......... T.......C. .G.GGA.... .T.T...A..

ZmAc        TTGGATAATG CTAGTGCAAA TGAAGTAGCT GTGCACGATA TAATTGAGGA TTTGCAGGAC
PgAc-L1     .......... ..TCA..C.. ......G..A ..CA.A..C. .C....CT.. .C..TTATC.
PgAc-L2     .....C.... .AGCG..... ...G....A ..CA.A..C. .C..CTCA.T CC..A...TT

ZmAc        ACTGA---TT CAAATCTAGT TTGTGATGGT GCTTTCTTTC ATGTGAGGTG TGCTTGTCAC
PgAc-L1     .G.ACTAG.. .TGCAT...C .......... ..C.....C. .......... .........T
PgAc-L2     ...AGCAG.G .TTCAT.G.. A......... ATA..T..C. .......... ..........

ZmAc        ATACTGAACT TGGTTGCAAA GGATGGCTTG GCTGTAATTG CAGGAACAAT TGAGAAAATC
PgAc-L1     ..T..T...C .A.......G ......T... AG........ ..CC..... ......T..T
PgAc-L2     ..T......C .T.....T.G A.....GC.. AGC....... GT..C..... ......T..T

ZmAc        AAAGCGATTG TTCTTGCTGT AAAATCTTCT CCTTTGCAGT GGGAAGAACT AATGAAGTGT
PgAc-L1     .G.CAAC... .A........ C..GGGA..C T.GC....A. T...T..GT. G....G.C..
PgAc-L2     .G...CC... .......... T...GGA... ..A.....A. .......G.. C......C..
                                                      ⇓         ⇓
ZmAc        GCTAGTGAAT GTGACTTGGA TAAATCTAAA GGGATCTCA- TATGATGTCT CAACTAGATG
PgAc-L1     ..G.C...G. ...GA..... ..CT..A..G ..AC.T..TC .TG-....TG .......G..
PgAc-L2     ....C...G. ...GGC.T.. C.C.AG...G ..TC.T..TC .CG.....G. .C........

ZmAc        GAATTCAACC TATTTGATGT TGAGGGATGC CTTATATTAT AAGCCTGCAC TAATAAGGCT
PgAc-L1     ...C...... ...C.....C .CC.A..... TC.G..C... ...G.....T .C..G.....
PgAc-L2     ...C..T... ..CA.....C .TC.A..... A..G..C..C ...G....TT .T..G...A.

ZmAc        TAAAACAAGT GATCCTCGCA GGTATGTTTG TCTCAATTGT TGTACATGTC ATCATTATAA
PgAc-L1     G...T..TCG ....GC..T. T.....A--- ---------- -.A.A..AA. .C...-----
PgAc-L2     G...T..TCA T...A...T. ......A--- ---------- -.A.A..C.. TC...-----

ZmAc        ATTCTCAATT AATCAAATGT CAATTATTGT AGGTACGATG CAATTTGTCC TAAAGCCGAG
PgAc-L1     ---...---- ---.T.G..- -...------ -..------- ---------- ----...---
PgAc-L2     ---...---- ---.T....- -...------ -..------- ---------- ----...---

ZmAc        GAGTGGAAGA TGGCATTAAC TCTTTTTAAG TGTTTGAAGA AGTTTTTTGA TCTCACTGAA
PgAc-L1     ----------. ....T..C.A A.AG..CC.A ..C..A.... .A....A... .........G
PgAc-L2     ----------. ....T.AT.G A..C..CC.A ..CC.A.... .A..C.A... ....G.....

ZmAc        CTCCTATCTG GTACTCAATA TTCCACTGCA AATTTATTTT ACAAAGGTTT CTGTGAGATA
PgAc-L1     ..TT.T.... .A...TT... .C.A...... .....T..C. .T.G...... .........T
PgAc-L2     ..TT.G...T .A...TT... .C.A..AA.. ...C.G..C. .T.G...... T.........

ZmAc        AAGGATTTGA TTGACCAATG GTGTGTTCAT GAAAAATTTG TCATTAGGAG AATGGCCGTT
PgAc-L1     ...AT....T .A..TG.T.. ....C.AAG. A..G.TGCAA .......TG. ......TAC.
PgAc-L2     ...AT....C ....TGGT.. .......AG. ..TG.TAACA C......T.C ......T.C.

ZmAc        GCAATGAGTG AAAAGTTTGA GAAATATTGG AAAGTGTCTA ATATTGCACT AGCTGTAGCA
PgAc-L1     AGT....A.A .G........ A..G..C... ..GAA...AT C..CCA.... TT.......T
PgAc-L2     T.T......A GG........ A.......... ..GAA...AT CC.C.A.... T.T...T..C

ZmAc        TGCTTCCTTG ACCCTAGGTA CAAGAAAATA TTGATTGAGT TCTATATGAA AAAATTTCAT
PgAc-L1     .....T..A. .T........ ........GG C.T..A.... ....C....G G.....C...
PgAc-L2     .....TT.A. .T........ ........GG C.T..A.... .T..C...GG ......C..C

ZmAc        GGTGATTCAT ACAAAGTTCA TGTAGATGAC TTTGTTAGGG TCATTAGAAA ATTGTATCAA
PgAc-L1     ..CAG.G... CTC....C.. ...T....CA C....AGATT ....C.AG.. .C.T..CT.G
PgAc-L2     ..C.G.G... CT...AA... ...T.....G C....A.AT. ....C.AGC. .C.......G

ZmAc        TTCTATTCTA GTTGTAGTCC TTCAGCTCCA AAGACAAAGA CAACTACTAA TGATAGTATG
PgAc-L1     .CT...GT.. ..GC.GAA.. A..TT..T.. ..AC.T.... GT.AAG.AG. ...ACA..G.
PgAc-L2     ..T...GT.. ..GC..CA.. ....CTG..T .....T.... GT.GAT.AG. ..CACC....

ZmAc        GA---TGATA C--C------ -TTGATGGAA AATGAAGATG ATGAATTTCA AAACTATTTG
PgAc-L1     ..TCT..... .AG.AGATCT A..AG.T..C .G.TG..... .....C.AG. G.....C...
PgAc-L2     C.ACT..... .AG.AGAATT G..A...... G...G..... .....C.AG. G.....CC.A

ZmAc        CATGAGTTGA AGGATTATGA TCAAGTAGAG TCAAATGAAT TGGATAAATA TATGTCTGAA
PgAc-L1     T......CT. GT.GGCC... .GGG.AT..T ATG.....G. ....C..G.. C...G.A..T
PgAc-L2     T....A.CT. GT.GAC...G .GCT.AT.CT ATG.....G. ....G..... ....G.A..C

ZmAc        CCCCTTTTGA AGCATAGTGG TCAGTTTGAT ATTTTATCAT GGTGGAGGGG AAGGGTTGCA
PgAc-L1     ..A.C..... G.TT...... C.......... .....GG... ......A.AA CCAATC..AT
PgAc-L2     ..A.C..... G.AT...... .......... ..C..GG... ......A.AA TCA.A...AT

ZmAc        GAATATCCTA TTCTCACCCA AATTGCAAGG GATGTGCTAG CAATACAAGT GTCAACTGTT
PgAc-L1     ..G....... ....GG.AA. ...A..CC.T ...C.A..G. .TG.G..... C..T.....C
PgAc-L2     ..G....... ....TG.AA. ......TC.T ...C...... TTG....... ......G..G
```

FIGURE 4.—*Continued*

```
ZmAc        GCTTCTGAGT CTGCGTTCAG TGCTGGTGGT CGTGTTGTTG ATCCTTACCG CAATCGTCTT
PgAc-L1     ..C.....A. ....C..T.. .........A .......... ....A.TT.. ..GC..C..A
PgAc-L2     ........A. .A..C..T.. .........G .....AA... ......T... A.G...C...

ZmAc        GGTTCGGAGA TTGTTGAAGC TTTGATATGC ACAAAAGATT GGGTAGCAGC ATCTAGAAAA
PgAc-L1     .A.C.T..A. .G...C.... .......... .TG....... ....T..T.. GGGA....GG
PgAc-L2     .A.C.T..A. .G.CA..... ...A..T... .TG..G.T.. ....T..T.. .GGAC...G.

ZmAc        GGTG---AAT GCATATATGT TATAATGAAG TTCCAATTTA TAGTTATTCA ACAATTATTT
PgAc-L1     ...AATA..A CT.AC.TA.. .CC....TTT G....GC..T .TACC..AAC .ATC..GGC.
PgAc-L2     ...TATA..A ..TCC.CA.. .T..G.TCTT G.A......T ATA.C..ATC G---..TGC.

ZmAc        TACTTATATT GATGCATATT TGTGTCATTC AAGGTGCTAC ATATTTTCCA ACAATGATTG
PgAc-L1     G.T......G TC.TGT..AC .T.A.TT... .-.A.AAG.A GA.GG..GGT T....TG...
PgAc-L2     A.T..G--.G T..T------ -T.CATTG.. .-.A.AAG.- --.GG..GGT T....TG...

ZmAc        GTGATCTCGA GGTGCTAGAC TCTGTTATTG CTGCTGCAAC AAATCATGAG AATCATATGG
PgAc-L1     ....C..T.. A..TA....G GTCC..G.C- ...AAAAGCT G..GTG.... GTA....A--
PgAc-L2     ....C..T.. A..TA....A G.CC..G--- ....AAA.CT GG.C.G.... GTATG...--

ZmAc        ATGAGGTATT TAAAGATTAT TATTTACTTC GTGCATGGGC TATTAATTTG CTATTAT-TC
PgAc-L1     .A..TA-..A CC...C..TA ..A.CTG..T T.-T..TCTT C.C.TGC..T A..A.C.T.G
PgAc-L2     ..C.AC-.A. AC..C---.A .....TGC.A T.AT..CCTA ...GTTCC.T ..CA...G.G
                                                  ↓
ZmAc        ACTACT-GTT TTGATGCATG GGCTGTTTGC TGTCGCCTTG
PgAc-L1     TT...AATC. C.C.ACTT.C A--CA.A.A. -.AATTAC.T
PgAc-L2     TA-..AA.-. G.T..TGT.A A.TCAGCAA. ..AAAT.A..
```

FIGURE 4.—*Continued*