Resistance to Gap Repair of the Transposon Tam3 in *Antirrhinum majus*: A Role of the End Regions

Shiko Yamashita,* Toshiyuki Takano-Shimizu,† Ken Kitamura,* Tetsuo Mikami* and Yuji Kishima*

**Laboratory of Genetic Engineering, Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan and †Department of Population Genetics, National Institute of Genetics, Mishima 411-8540, Japan*

Manuscript received May 22, 1999
Accepted for publication August 3, 1999

ABSTRACT

The extremely homogeneous organization of the transposon family Tam3 in *Antirrhinum majus* is in sharp contrast to the heterogeneity of the copies constituting many other transposon families. To address the issue of the Tam3 structural uniformity, we examined two possibilities: (1) recent invasion of Tam3 and (2) failure of gap repair, which is involved in conversion from autonomous forms to defective forms. The phylogenetic analysis of 17 Tam3 copies suggested that the invasion of Tam3 into the Antirrhinum genome occurred at least 5 mya, which is sufficiently long ago to have produced many aberrant copies by gap repair. Thus, we investigated gap repair events at the *nivea*$^{\text{recurrens:Tam3}}$ (*niv*^{rec}*::Tam3*) allele, where Tam3 is actively excised. We show here that the gap repair of *de novo* somatic Tam3 excision was arrested immediately after initiation of the process. All of the identified gap repair products were short stretches, no longer than 150 bp from the ends. The Tam3 ends have hairpin structures with low free energies. We observed that the gap repair halted within the hairpin structure regions. Such small gap repair products appear to be distributed in the Antirrhinum genome, but are unlikely to be active. Our data strongly suggest that the structural homogeneity of Tam3 was caused by immunity to gap repair at the hairpins in both of the end regions. The frequency of extensive gap repair of *de novo* excision products in eukaryotic transposons was found to be correlated with the free energies of the secondary structures in the end regions. This fact suggests that the fates of transposon families might depend on the structures of their ends.

THE cut-and-paste-type transposons in eukaryotes generally comprise a very limited number of autonomous copies that produce transposase and a larger number of nonautonomous copies with internal deletions or replacements (McClintock 1950; Döring *et al.* 1984; Döring and Starlinger 1986). Generation of nonautonomous elements is attributed to gap repair at the breakpoints after excision of the element (Engels *et al.* 1990; Gloor *et al.* 1991). Gap repair is associated with synthesis-dependent strand annealing (SDSA), which is an error-prone repair process (Formosa and Alberts 1986; Nassif *et al.* 1994; Rubin and Levy 1997). Aberrant copying of a transposon due to gap repair requires its homozygous form (Gloor *et al.* 1991). Excision of one of the homozygous copies initiates gap repair via the SDSA pathway (Nassif *et al.* 1994); *i.e.*, free 3′ ends of the donor sites, which were generated by the double-stranded break after excision of the element, invade homologous sequences on a sister chromatid, and DNA synthesis from 5′ to 3′ proceeds through a local loop of DNA (Figure 1). A similar reaction was catalyzed *in vitro* using purified T4 phage proteins (Formosa and Alberts 1986). As a consequence of the gap repair, active transposons frequently produce nonautonomous elements, resulting in a decline in the proportion of autonomous copies (Engels *et al.* 1990). In particular, the human genome has many fossilized transposons due to the loss of autonomous copies (Smit and Riggs 1996). The degree of gap repair appears to differ among eukaryotic transposons: *P*, Tc1, and Mu elements tend to produce extensive gap repair products (Doseff *et al.* 1991; Gloor *et al.* 1991; Plasterk and Groenen 1992), while the Ac and Tam3 elements cause occasional reversion by near-end-joining at the breakpoints or by leaving short footprints (Coen *et al.* 1989; Rubin and Levy 1997; Yamashita *et al.* 1998).

We previously isolated 40 independent clones carrying Tam3-homologous sequences to study Tam3 structural organization in the genome of *Antirrhinum majus* (Kishima *et al.* 1997, 1999). We found structural alterations in only 5 out of 40 copies, and most copies had highly conserved structures of nearly the same size. The Antirrhinum genome contains at least 8 autonomous Tam3 copies (Martin *et al.* 1989; Kishima *et al.* 1999). This study was conducted to elucidate the cause of the uniform structure of the Tam3 copies. Our results demonstrate that Tam3 is immune to gap repair due to arrest

Corresponding author: Yuji Kishima, Laboratory of Genetic Engineering, Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan. E-mail: gelab@abs.agr.hokudai.ac.jp

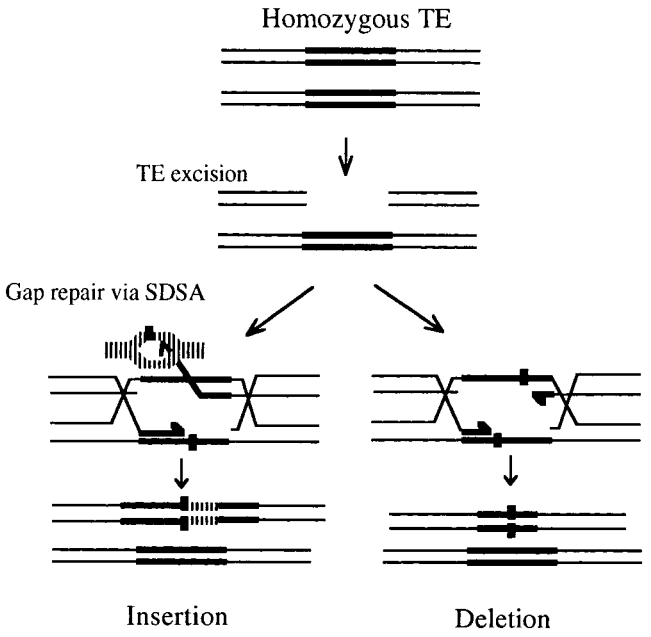

Figure 1.—Model for the processes causing the structural alterations that occur in general in transposons. Each thin line represents a single-stranded DNA. The figure shows the abortive gap repair processes following transposon (thick lines) excision from a sister chromatid in a homozygous allele. The gap repair utilizes the SDSA pathway. This ectopic repair process is initiated by invasion by the ends of the donor site into the sister chromatid. New synthesis of DNA occasionally gives rise to alterations of the template structure. Insertion results from strand invasion into a nonhomologous template (striped line). Deletion is due to annealing of incomplete repair products at short repeats (short vertical bar).

of the gap repair process at its end regions. Interestingly, Tam3 possesses 16 hairpin structures in the end regions. Comparison of the end regions of several known transposons suggests that there is an apparent correlation between the secondary structures in the end regions and the extent of gap repair after excision. We thus propose that these folded structures are involved in halting the gap repair process of *de novo* Tam3 excision.

MATERIALS AND METHODS

Plant materials and DNA isolation: The HAM5 line of *A. majus* was derived from the *niv*^{rec}*::Tam3*/*stabilizer*$^{-}$ line (Kishima *et al.* 1997). The plants were grown in an air-conditioned chamber; some were grown at ~25° for 2 mo and then moved to 15°, and others were grown at ~25° continuously. Leaf DNA was isolated from 10-mo-old plants according to the method of Martin *et al.* (1985).

Estimation of the divergence time: To compare the sequences corresponding to the Tam3 transposase gene, 17 Tam3 clones were chosen from 40 Tam3 clones isolated from the HAM5 genomic library (Kishima *et al.* 1999). These 17 clones carried the complete Tam3 transposase region. The rest of the clones were interrupted by the cloning sites of the vector, indicating partial clones of Tam3. The sequences were aligned using the CLUSTAL W program (Thompson *et al.* 1994) from the DNA Data Bank Japan (DDBJ). To evaluate the degree of nucleotide mutational bias, the numbers of synonymous and nonsynonymous changes and those of transitional and transversional changes were counted only for sites commonly shared by all the sequences. Furthermore, for the sake of simplicity, codons in which nucleotides differed at two or more positions were excluded from the calculation of synonymous and nonsynonymous substitutions. We found a transition/transversion bias, but the numbers of the synonymous and nonsynonymous substitutions did not significantly deviate from 1:3. We therefore calculated the pairwise nucleotide divergence (K) between the 17 independent Tam3 sequences on the basis of Kimura's two-parameter method (1980) without taking synonymous and nonsynonymous changes into account, and we constructed a neighbor-joining tree on the basis of these estimates (Saitou and Nei 1987). A tree was drawn using the DendroMaker program developed by Dr. T. Imanishi (National Institute of Genetics, Japan). The sequence data have been deposited in the DDBJ/GenBank/EMBL DNA databases (AB013982–AB013997 and AB012941). Among the 17 Tam3 copies, the highest K value (0.108) was scored between S-98 (AB012941) and S-CHS (AB013892), and the divergence time was calculated as follows: the K value/synonymous substitution rate in the plant genome (6×10^{-9}/site/yr) (Wolfe *et al.* 1989) / 2.

PCR amplification of excision products of Tam3 at *niv*^{rec}*::Tam3*: To amplify Tam3-excised sequences in *niv*^{rec}*::Tam3*, PCR was done using KOD DASH (TOYOBO), and 200–400 ng of genomic DNA was prepared from HAM5 plants carrying homozygous *niv*^{rec}*::Tam3*. The DNA templates and primers were denatured at 94° for 30 sec, annealed at a few degrees above the melting temperature calculated for each primer, and extended at 72° for 2 min. In the case of the PCR reaction to amplify the whole Tam3 sequence, 5% dimethyl sulfoxide was added. As illustrated in Figure 3A, primers A and B were the outermost primers, located in the two Tam3-flanking sequences (Sommer and Saedler 1986). Primers C and D were also located outside of the Tam3 sequence (Hehl *et al.* 1991; Kishima *et al.* 1999). E and F were located on the two borders between Tam3 and the *niv* promoter sequence, and G and H were located inside the Tam3 sequence. The sequences of these PCR primers were as follows: A (5′-TAGCTTCGGCGCCCGGCGGTAGAACTCCCG-3′); B (5′-TGGCCGGTCCCTCAGCCCTCTGAGCCCTAC-3′); C (5′-CCTATTGGGCAAAATTAGGTACC-3′); D (5′-GAACCTCCTCAACAGTCACCATTT-3′); E (5′-GCTAATAACCACGTATCTCAGCTAAAG-3′); F (5′-GGTACCAAGTATGGTAGCTGAGATTAAAG-3′); G (5′-AAGAATCGCGACATGGACGC-3′); H (5′-CTATATTGTTGGTCGAGCATGTCT-3′).

Construction of a plasmid carrying the *niv*^{rec}*::Tam3* promoter: We examined the PCR products to distinguish between two possibilities: the fragments produced reflect *de novo* gap repair at *niv*^{rec}*::Tam3* or artifacts that arose during the PCR reaction. To distinguish between these possibilities, we made a plasmid carrying the region between primers A and B in the *niv*^{rec}*::Tam3* allele in pBluescript SK vector (Stratagene, La Jolla, CA; Figure 3A). This construct contains a Tam3 copy, and it cannot be excised by the PCR reaction. When the plasmid was used as a template DNA in the same PCR reactions as used for the genomic DNA, if the PCR reactions gave rise to small fragments like the gap repair products, it would imply that the obtained fragments should be artifacts.

Southern hybridization and sequencing: The probes were PCR-amplified fragments prepared in the presence of DIG-11-dUTP. Southern hybridization procedures were carried out at 68° according to the protocol of DIG detection system (Boehringer Mannheim, Mannheim, Germany). Double-stranded DNA samples inserted into pBluescript vector were sequenced using a d-Rhodamine Terminator Cycle Sequenc-

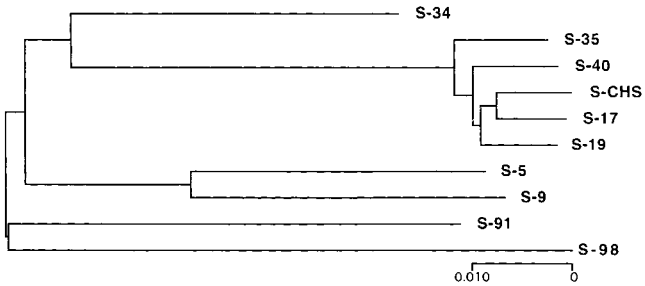

Figure 2.—Neighbor-joining tree for 10 different Tam3 copies based on the transposase-encoding gene sequences (Kishima *et al.* 1999). No sequence difference was found in 7 out of 17 copies studied, and the autonomous Tam3 element (S-CHS) represents these 7 copies in this figure. The root was placed at the midpoint of the path between the 2 most-distant copies. The sequence data have been deposited in the DDBJ/GenBank/EMBL DNA databases (AB013982–AB013997 and AB012941).

ing Ready Reaction sequencing kit (Applied Biosystems, Foster City, CA) and an ABI377 Automated DNA sequencer (Applied Biosystems).

Simulation of secondary structures of single-stranded DNA and calculation of free energy: Computer simulation and calculation of free energy (−kcal/mole) were performed using the Internet-supplied software DNA folding server: http://www.ibc.wustl.edu/~zuker/ from Professor M. Zuker (Washington University), which was established according to Santalucia (1998).

RESULTS

Divergence of the Tam3 transposase coding region and estimation of the invasion time: Previously, we investigated the organization of the transposon Tam3 family in *A. majus*. Characterization of 40 independent Tam3 clones isolated from an Antirrhinum plant revealed that the Tam3 family is highly conserved and that the copy sizes are uniform. We did not find any copy with a deleted internal sequence, unlike what is usually observed for other transposons (Kishima *et al.* 1999). To test whether the conservation of the Tam3 structure might be due to the recent invasion of the Antirrhinum genome by Tam3, as with the *P* element in *Drosophila melanogaster* (Kidwell 1979, 1983), the DNA sequences of the Tam3 transposase-encoding gene were determined for 17 clones carrying almost full-sized Tam3 sequences (Kishima *et al.* 1999). Figure 2 shows the neighbor-joining tree constructed from these data. The number of substitutions per site between the autonomous copy (S-CHS) and its most distantly related copy (S-98) among the 17 copies was estimated to be 0.108, which corresponds to about a 9-million-yr divergence, based on a calculation using the estimated synonymous substitution rate of 6×10^{-9}/site/yr in the plant genome (Wolfe *et al.* 1989). S-98 is no longer active and is immobilized in all lines in Antirrhinum (Kishima *et al.* 1999). The other stable copies with long terminal branches, such as S-5 and S-9, have ~5-million-yr-old lineages. Although we cannot rule out the possibility of multiple invasions by Tam3, the above-described copies in the tree do not appear to form obvious clusters whose presence would suggest multiple invasions (Anxolabehere *et al.* 1988; Figure 2). Even if we assume that Tam3 has invaded the Antirrhinum genome two or three times, the latest invasion should have occurred a few million years ago. This time interval should have allowed Tam3 to produce many aberrant copies resulting from gap repair. Nevertheless, we could not find any internally defective Tam3 elements, *i.e.*, any typical outcome of gap repair. This suggests that Tam3 has some mechanism to maintain its own structural stability against gap repair.

PCR detection of the Tam3 gap repair products at the *niv^{rec}::Tam3* allele: The above result raised the questions of whether gap repair was initiated by Tam3 excision and whether the processes involved in gap repair could proceed through the Tam3 structures. Therefore, we carried out PCR analysis, which is currently the only method available to detect gap repair products, to investigate the genomic structures of *de novo* somatic Tam3 excision sites. Template DNA was prepared from Antirrhinum plants carrying the homozygous *niv^{rec}::Tam3* allele (Sommer *et al.* 1985; Kishima *et al.* 1997). The *niv* locus is the sole chalcone synthase gene (CHS) in *A. majus* (Sommer and Saedler 1986). This allele has a Tam3 copy 64 bp upstream of the transcription start site of *niv*. Somatic excision of Tam3 is markedly activated by relatively low temperatures of ~15° (Carpenter *et al.* 1987). Therefore, DNA was prepared from plants grown for 8 mo at 15° to allow accumulation of somatic excision products. We also used two control DNAs to test whether the PCR products reflect *in vivo* reaction products or artifacts that arose during the PCR reaction; the first control was a genomic DNA isolated from the plants grown continuously at 25°, and the second was a plasmid DNA containing the PCR-targeted sequence in *niv^{rec}::Tam3*. We designed eight primers in the *niv^{rec}::Tam3* allele (Figure 3A). Primers A and B were used for primary PCR (Figure 3B, lane 1); the others were used for nested PCR. Among the primary PCR products, a 620-bp band in the sample from the plants grown at 15° would represent the *de novo* event of the Tam3 excision, but we failed to detect this band from the plants grown at 25° (Figure 3C, lane 1). When four combinations of the primers C-D, C-H, G-D, and G-H, were used to further amplify the primary PCR products from the 15° sample, we detected only the fragments expected from excised or nonexcised products and no other aberrant fragments (Figure 3B, lanes 2–5). With the primary PCR product from the 25° sample, we confirmed that no excised products arose using the above four primer sets (Figure 3C, lanes 2–5). These results suggested that the sites divided by Tam3 in the *niv* promoter were mostly reunified by end-joining after Tam3 excision and that abortive gap repair products, if there were any,

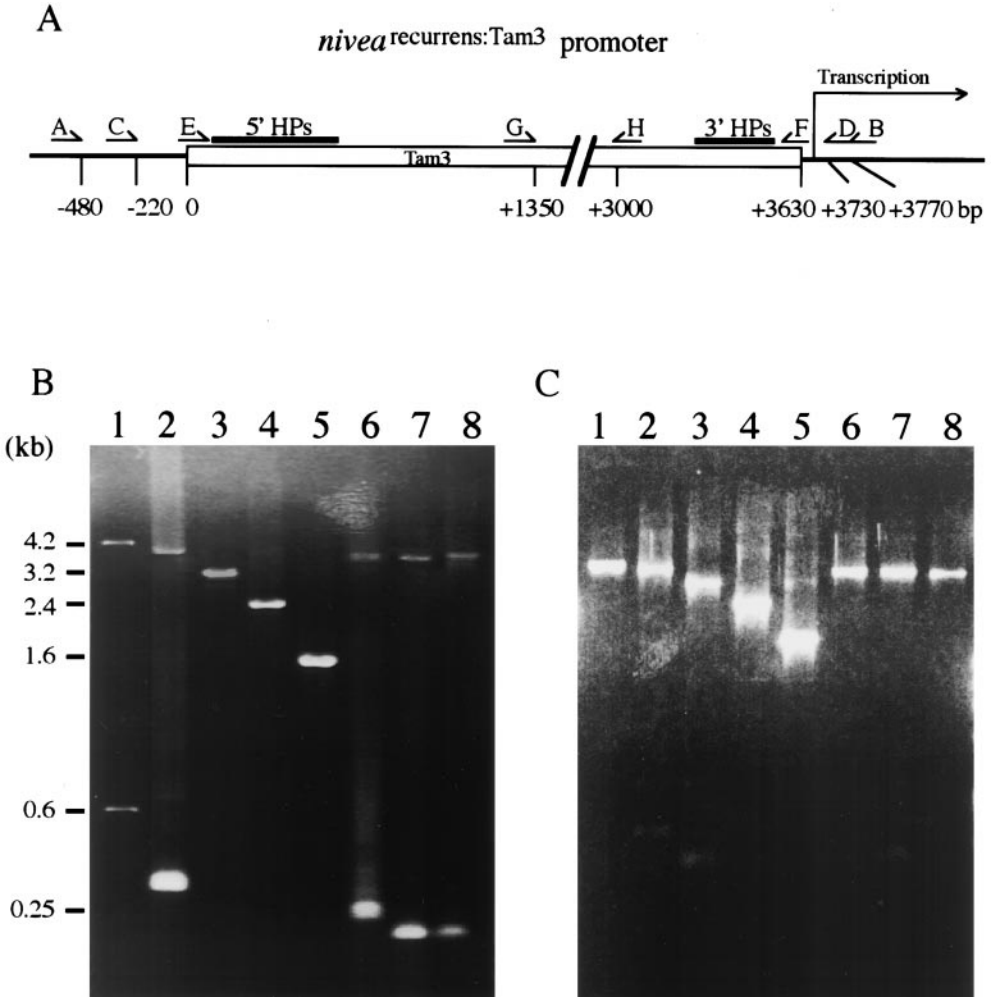

Figure 3.—PCR detection of the gap repair products at the *niv*^{rec}*::Tam3* allele, which has an actively excised Tam3 copy. (A) The map summarizes the locations of the eight primers (A–H) for the *niv*^{rec}*::Tam3* allele, as described in materials and methods. Numbers (base pair) under the map indicate distances from the Tam3 5′ end to each primer position. The two hairpin regions in the Tam3 sequence are labeled as 5′ HPs and 3′ HPs. (B) PCR amplifications using template DNA isolated from the plants grown at 15° were done with the following primers: lane 1, primers A–B; lane 2, C–D; lane 3, C–H; lane 4, G–D; lane 5, G–H; lane 6, E–D; lane 7, C–F; lane 8, E–F. (C) PCR amplifications using template DNA isolated from the plants grown at 25° were done with the same primers as those listed above. The same PCR patterns were also obtained when the plasmid containing the *niv*^{rec}*::Tam3* promoter region was used as a template.

were very short. To isolate short gap repair fragments carrying the Tam3 sequence, PCR amplifications were carried out with three combinations of primers across the junctions of Tam3 and the *niv* promoter region, E-D, C-F, and E-F. We found discrete, small fragments of length <300 bp along with the intact Tam3 (Figure 3B, lanes 6–8), whereas no such fragments were produced when the primary PCR product from the 25° sample was used (Figure 3C, lanes 6–8). Furthermore, when the plasmid DNA containing this region was used in the same PCR reactions as the template DNA, we could not detect such small-sized fragments (the patterns obtained were exactly the same as the one shown in Figure 3C). Together with the control experiments, these results implied that the small fragments were produced only when the template DNA from the plants grown at the low temperature was used. These fragments might be gap repair products of *de novo* Tam3 excision. However, the Tam3 excision at *niv*^{rec}*::Tam3* does not seem to have generated an abundance of such gap repair products, because the corresponding gap repair fragments were not detected with the primers C-D except for fragments as short as the end-joining products at the excision site, as shown in Figure 3B, lane 2.

Characterization of the gap repair products: Two hundred clones that were isolated from the three PCR products (Figure 3C, lanes 6–8) produced from the low-temperature sample could be classified into 13 types (Table 1). All the clones represented abortive gap repair products from the excision sites of Tam3 or from one break in Tam3 and one site in the flanking sequence (Figure 4 and Table 1). A stretch of nucleotides from the breakpoint was linked with either a stretch of Tam3 sequence from the opposite end or with Tam3-flanking sequences. No sequences were elongated by >150 bp from a single end of the excision site. Most of the gap repair products were ligated with complementary motifs in either the Tam3 sequence or its flanking sequences (Figure 4 and Table 1). This result agrees with the gap repair model prediction that single-stranded DNA ends would search for and anneal with complementary sequences (Nassif *et al.* 1994).

Arrest of DNA polymerase reaction and gap repair at hairpin structures in the Tam3 ends: When T7 DNA polymerase was used in DNA sequencing reactions under standard conditions, the reactions from both the Tam3-flanking primers were stacked or compressed at the positions 80 bp from the 5′ end and 28 bp from

TABLE 1

Structures of the gap repair products

Product no.	Combination of rejoined sites[a]	Overlapping nucleotide(s) at junction
1	1-1/13-1	tgggccg
2	1-2/13-5	c
3[b]	1-3/15-1	ggcat**t**gcc (t:insertion)
4	1-4/16-1	ccggcacggc
5	1-5/13-4	cggcacggccc
6	2-1/16-5	cggcacg
7	3-3/13-3	cggc
8	3-3/16-4	tcggc
9	CHS-1/16-2	cg
10	CHS-2/16-6	tgcac
11	CHS-3/16-3	g
12	3-1/CHS-4	gg
13	3-2/CHS-5	c

[a] The gap repair products are designated by the combinations of sites indicated in Figure 4.

[b] This gap repair product formed an exceptional palindromic motif with a "t" insertion at the junction.

the 3′ end (Figure 4). The sites where the polymerase reactions stopped corresponded to regions in firm hairpin structures. In total, 16 putative hairpins with relatively low free energy (−6.8 to −29.9 kcal/mole, with mostly GC-rich content) were found within 500 bp from each end (Figure 4). The remainder of the Tam3 sequence had no such strong hairpin structures and allowed sequencing reactions to proceed. It was also difficult to amplify the whole Tam3 sequence by PCR. The polymerization reactions were enhanced by addition of DMSO and/or high temperatures of annealing, but the addition of nucleotide analogues such as dITP or 7-deaza-dGTP had little effect. These findings suggested that the Tam3 template inhibited the polymerase reaction because of the strongly folded structures at the two ends. Considering this observation, we hypothesized that the gap repair process of *de novo* excision of Tam3 at *niv*^{rec}*::Tam3* could not proceed beyond the clustered low-free-energy hairpin regions. We previously reported that the two genomic hybridization patterns obtained by probing with the 600-bp sequences from both the end regions of Tam3 were identical to the pattern obtained using the internal Tam3 sequence as a probe (Kishima *et al.* 1999). These two findings suggest that no gap repair products that proceeded through the hairpins are present in the genome.

Detection of footprints of the gap repair in the genome: To verify the occurrence of the arrested gap repair at the Tam3 excision sites, we attempted to demonstrate the presence of gap repair footprints in the Antirrhinum genome using Southern blotting. The HAM5 genomic DNA was digested with *Eco*RI, of which recognition sites were absent from probes to be used (Figure 5). When the blot was probed by a full Tam3 sequence, a number of bands were generated in the hybridization pattern (Figure 5, left). Subsequently, we carried out the hybridization analyses probed with the gap repair products listed in Table 1. When the longest product (no. 7, Table 1), which consists of equal-sized sequences from both the ends, 130 bp from the 5′ end and 125 bp from the 3′ end, was used as a probe under high-stringency conditions, two bands were detected in the blot (Figure 5, right). The results suggested that the two homologous bands were specific footprints to the gap repair product; *i.e.*, the sequence homologous to the gap repair product should be present at two sites, independent of the Tam3 copies in the genome. The result led us to assume that the Antirrhinum genome contains various footprints of the arrested gap repair products. However, using the other gap repair products as probes, we were unable to detect specific bands under the same hybridization conditions. This is presumably due to extremely small sizes of the sequence derived from either Tam3 end in the gap repair products (data not shown).

DISCUSSION

Hairpins in the end regions of Tam3 could arrest the gap repair process and maintain the structural conservation: We can summarize the above results as follows: after the excision of Tam3, the SDSA process was arrested at the hairpin regions adjacent to the Tam3 ends, and the resultant short, single-stranded sequence searched for a complementary motif in the other strand and annealed with it (Figure 6). Most of these short stretches should lose the ability to transpose due to lacking the greater part of the Tam3 sequence, including the subterminal regions involved in putative *cis*-elements for transposition. They remain integrated in the genome as footprints of the transposed Tam3. At present, we do not have evidence that such short segments are capable of transposition. Consequently, Tam3 has maintained its structural homogeneity in the Antirrhinum genome by avoiding production of nonautonomous copies capable of transposition. In this sense, the end regions of the element have an important role in controlling the gap repair activity. The secondary structure with its strong hydrogen bonds is one of the possible causes of the arrest, because SDSA depends on DNA polymerase reactions (Holmes and Haber 1999) and is thought to occur independently of DNA topoisomerase I activity, which is related to formation of a large D loop (Formosa and Alberts 1986), so that the resolution of the DNA strands might be incomplete. The inhibition of the SDSA reaction may be analogous to the termination of the RNA polymerase reaction at hairpin structures in prokaryotes (Platt 1986). Recently, Moore *et al.* (1999) reported that hairpin structures containing CNG repeats were inefficiently repaired dur-

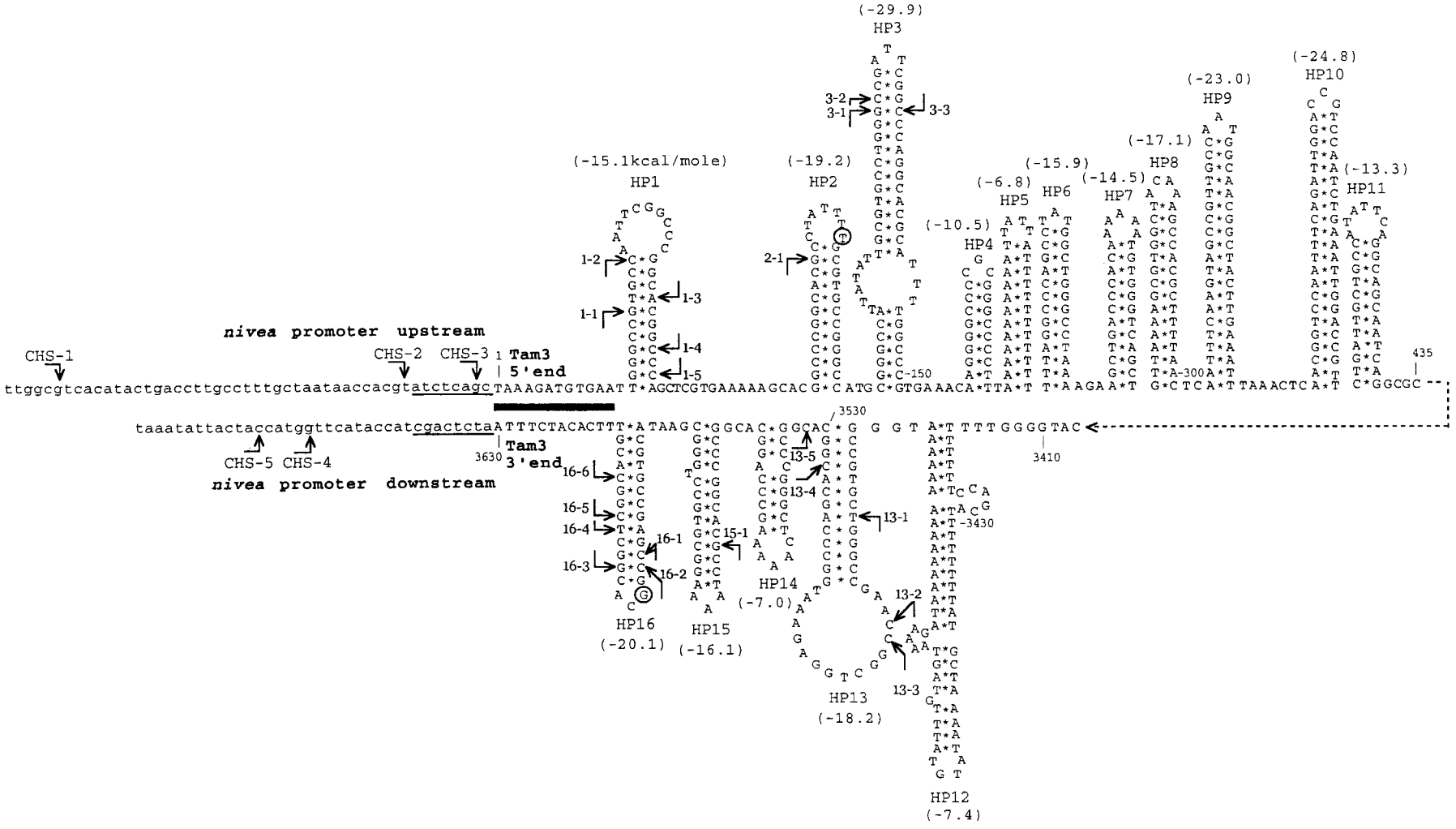

Figure 4.—Putative hairpin structures of single-stranded DNA in the Tam3 end regions that were simulated at 37°. Although Tam3 is active at 15° and not at 37°, the folded structures are basically the same at 15° and 37°. The 5′ end of Tam3 flanked by the *niv* sequence upstream of the insertion site in the promoter of the chalcone synthase gene and (bottom) the 3′ end of Tam3 flanked by the *niv* sequence downstream of the insertion site. These two parts, which are depicted on the same strand of DNA, contain a total of 16 hairpins, while no hairpins are found in the rest of the Tam3 sequence, which is represented by a broken line. We named each hairpin HP(*n*), with *n* indicating the order from the 5′ end. Asterisks show paired nucleotides that constitute the stems in the hairpins. The hairpin structures and the free energies (−kcal/mole) were simulated according to Santalucia (1998) using Zuker's home page. The encircled T 80 bp from the 5′ end and G 28 bp from the 3′ end of Tam3 indicate the arrested positions in the sequencing reaction using T7 polymerase. The bent arrows show the breakpoints of the gap repair products and the limits of the locations of the gap repair products elongated from the end region of Tam3 or the flanking *niv* sequence of *de novo* Tam3 excision. The uppercase letters represent Tam3 sequences, and the lowercase letters represent flanking sequences in the *niv* promoter. Each arrow's number corresponds with the breakpoint number in Table 1. The thick line shows the two 13-bp Tam3 terminal inverted repeats. Eight-base pair target-site duplications are underlined. Nucleotide positions in the Tam3 sequence are indicated at 1, 150, 300, 435, 3410, 3430, 3530, and 3630 bp from the 5′ end.

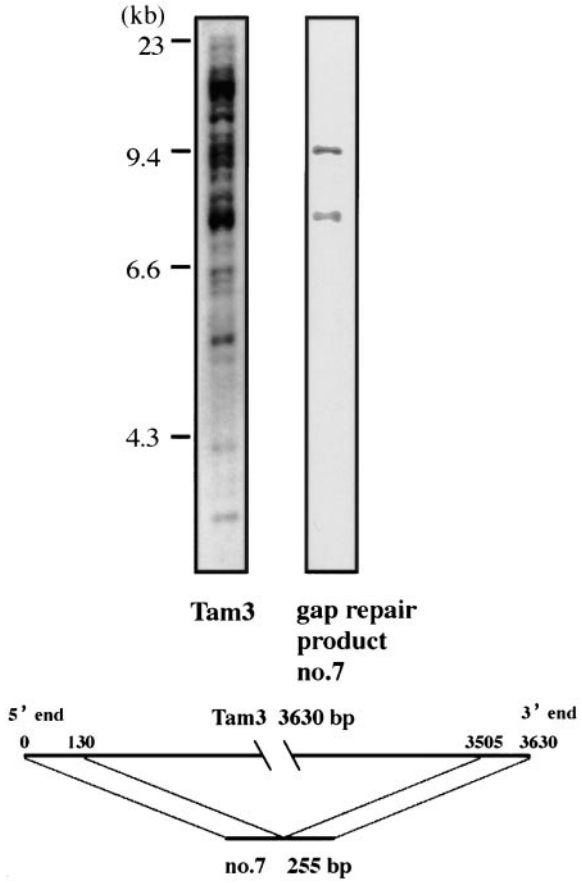

Figure 5.—Southern blots of genomic DNAs probed with the Tam3 intact sequence and one of the gap repair products, type 7, in Table 1. DNA prepared from HAM5 (5 μg) was digested with *Eco*RI in each lane. As illustrated below the blots, the probe used in the left lane corresponds to the whole Tam3 sequence, and the probe in the right lane contains both 130 bp from the 5′ end of Tam3 and 125 bp from the 3′ end of Tam3. Neither probe contains *Eco*RI-recognition sequences.

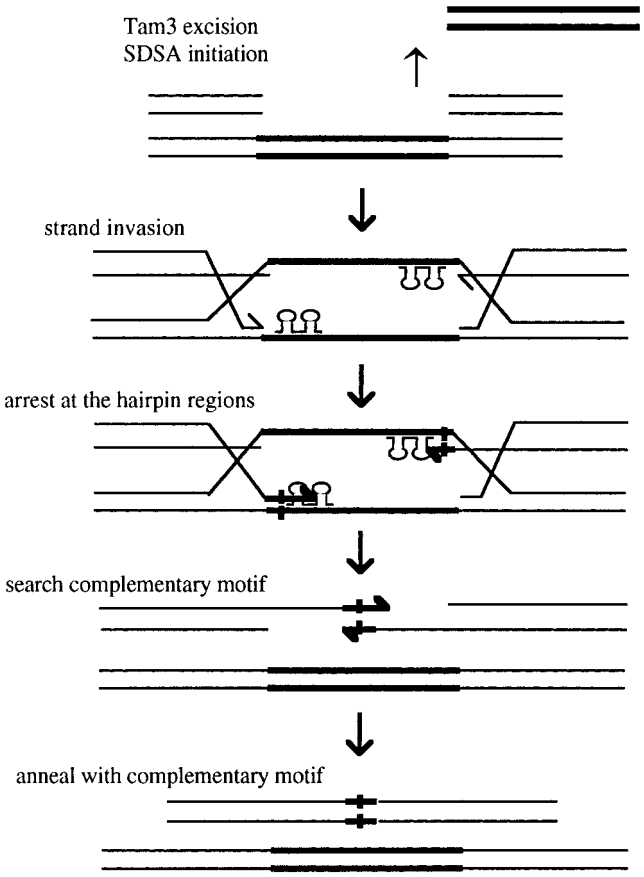

Figure 6.—Model for the processes of the arrested gap repair in Tam3. Each thin line represents a single-stranded DNA. The gap repair process (the SDSA pathway) is initiated by transposon (thick line) excision from a sister chromatid in a homozygous allele. The ends of the donor site then invade the sister chromatid. The DNA synthesis stops within the hairpin regions located in the terminals of Tam3. These short stretches search for complementary motifs (short vertical bar) and are annealed with these motifs.

ing meiosis in yeast. This finding is quite relevant to our hypothesis that firm hairpin structures block the repair process.

Correlation between the structure of the end region and extension of gap repair: Gap repair events have been well analyzed in active eukaryotic transposons such as *P* (Engels *et al.* 1990; Gloor *et al.* 1991) in *D. melanogaster*, Tc1 (Plasterk and Groenen 1992) in *Caenorhabditis elegans*, and Ac/Ds (Rubin and Levy 1997) and Mu (Doseff *et al.* 1991; Hsia and Schnable 1996) in maize. The frequency and size of gap repair products appears to vary among these transposons; the transpositions in *P* (Engels *et al.* 1990), Tc1 (Plasterk and Groenen 1992), and Mu (Doseff *et al.* 1991), which appear to use the replicative mode, often undertake extensive gap repair, but Ac (Rinehart *et al.* 1997; Rubin and Levy 1997) and En/Spm (Gierl 1996) show reversions from transposon-inserted mutations or usually leave only short footprints after excision. A particular example of the latter type is Tam3, for which copies deleting internal sequences have never been isolated (Coen *et al.* 1989; Yamashita *et al.* 1998), presumably because of the arrest of the gap repair at the hairpin structures of the end regions. To address these differences in degree of gap repair, we compared free energies in the end regions among several known transposons (Table 2). There were no other elements that had low free energies comparable to those of Tam3. This is very likely to be a reason for the difference between Tam3 and the other transposons with respect to structural conservation. The elements with relatively low free energy are postulated to be likely to transpose in the nonreplicative manner and to have only occasional reversions or short footprints, while we failed to find such low free energies in the elements showing the replicative mode of transposition, due to the high frequency of extensive gap repair (Table 2). Therefore, a close correlation exists between the secondary structures in the end regions and the extent of gap repair. It can also be predicted from these

TABLE 2

Free energies of the single-stranded DNA in the end regions of various transposons

Transposon	TP mode[a]	5′ 100 bp[b]	3′ 100 bp[c]	100 bp[d]	5′ 200 bp	3′ 200 bp	200 bp[d]	5′ 300 bp	3′ 300 bp	300 bp[d]
Tam3	Nonreplicative	−36.8	−44.5	−81.3	−77.7	−71.0	−148.7	−127.3	−82.8	−210.1
En/Spm	Nonreplicative	−15.6	−21.4	−36.6	−43.6	−41.2	−84.8	−47.0	−69.8	−116.8
hobo	Nonreplicative	−13.0	−7.9	−20.9	−28.8	−30.6	−59.4	−45.7	−55.1	−100.8
Ac	Nonreplicative	−18.2	−9.4	−27.6	−33.1	−34.4	−67.5	−36.6	−53.1	−89.7
Mu	Replicative	−12.1	−7.2	−19.3	−31.3	−24.5	−55.8	−47.4	−41.6	−89.0
P	Replicative	−8.9	−6.6	−15.5	−19.5	−13.3	−32.8	−29.0	−21.3	−50.3
Tcl	Replicative	−7.7	−9.7	−17.4	−15.0	−17.1	−32.1	−21.0	−25.1	−46.1

The free energy (kcal/mole) was calculated according to Santalucia (1998) in Zuker's home page.

[a] The preferred mode of transposition of each transposon is shown.

[b] Free energy in the 100 bp at the 5′ end of each transposon.

[c] Free energy in the 100 bp at the 3′ end of each transposon.

[d] Sum of the free energies in both of the end regions.

relationships that the gap repair (SDSA) process is sensitive to the secondary structure of the DNA. There is not a marked difference in the free energies of Ac and Mu. Actually, Ac has a number of nonautonomous Ds copies that originated from gap repair and have a wide range of lengths (Rubin and Levy 1997), but Mu is thought to give rise to extensive gap repair products more often than Ac, since the germinal excision frequencies for Mu elements are usually two orders of magnitude or more lower than transposition rates (Bennetzen *et al.* 1993). Computer simulation showed that the terminal 200-bp regions of Ac contain firmer hairpin structures than the Mu end regions (data not shown). The difference in the degrees of gap repair exhibited by these two transposons might be due to their differing secondary structures.

Comparison with other possible ideas: Engels *et al.* (1990) predicted that differences of the gap repair modes among transposons are attributable to species-specific repair enzyme activities. This might be partly true because the occurrence of the gap repair is closely associated with repair machineries (Lankenau and Gloor 1998). In this study, we have exclusively investigated the extent of the gap repair resulting from *de novo* Tam3 excision in the somatic cells. Our results led to the hypothesis that the mode of gap repair was strongly influenced by the extent of gap repair after excision, which is dependent on structure of the end regions in each transposon. Therefore, the activities of species-specific repair enzymes seem not to be directly related to the mode of gap repair. The frequency of the occurrence of gap repair still remains to be analyzed while taking account of repair activities.

Another possible cause of the arrest has been reported by Lankenau *et al.* (1996). They detected vector-mediated gap repair events (gene conversion) at certain frequencies after the excision of a *P* element, but when the binding site of the *suppressor of Hairy-Wing* [*su(HW)*] gene product was contained in a template for the *P*-element-induced gap repair, no gap repair products were obtained. One interpretation of this result was that the binding of the *su(HW)* product prevented the gap repair reaction and that the properties of a certain protein(s) bound to the subterminal repeats or the hairpin structures might include inhibition of the gap repair process. Antirrhinum may have such proteins, but it is doubtful, since this sort of protein must then have been present in all the cells and all the stages during evolution and development, otherwise the gap repair would have produced nonautonomous copies of Tam3. Although we cannot rule out some effects of transposon binding proteins, they do not seem to completely account for the inhibition of the gap repair.

A possible function of the end regions in determining the fates of transposons: The end regions in the transposons with preferential nonreplicative transposition, like Tam3, En/Spm, and Ac, tend to possess many short repeat motifs. Some of the repeats function as binding

sites for transposase or host factors that might be essential for transposition (Gierl 1996; Kunze 1996); the function of the remaining repeats remains to be explained. The subterminal repeats presumably form folded secondary structures that may inhibit the SDSA activity. In prokaryotic transposons, theoretically no gap repair can occur because of monoploidy; interestingly, repeat motifs are rarely present in the end regions of prokaryotic transposons. Hence, the acquisition of reiterated sequences in the subterminal regions might be a means of preventing overproduction of aberrant copies via gap repair.

Active transposons have been exposed to selective pressure by the gap repair mechanism of the host, and in most cases they have become fossilized due to loss of autonomous copies (Engels *et al.* 1990; Smit and Riggs 1996). Even currently active transposons may have the same fate sooner or later. However, Tam3-like transposons with strong hairpin structures in the ends should be immune from generation of defective copies via gap repair and might have a different fate in the genome.

We thank A. Wakatsuki and T. Yoshii for excellent technical assistance and Drs. T. Kubo and E. Nakajima for comments on the manuscript. We also thank Professors Y. Sano and C. Martin for providing facilities and the Antirrhinum lines, respectively. A part of this work was done at the Research Center for Molecular Genetics, Hokkaido University.

LITERATURE CITED

Anxolabehere, D., M. G. Kidwell and G. Periquet, 1988 Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of Drosophila melanogaster by mobile P elements. Mol. Biol. Evol. **5:** 252–269.

Bennetzen, J. L., P. S. Springer, A. D. Cresse and M. Hendrickx, 1993 Specificity and regulation of the Mutator transposable element system in maize. Crit. Rev. Plant Sci. **12:** 57–95.

Carpenter, R., C. Martin and E. S. Coen, 1987 Comparison of genetic behaviour of the transposable element Tam3 at two unlinked pigment loci in *Antirrhinum majus*. Mol. Gen. Genet. **207:** 82–89.

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

Döring, H. P., and P. Starlinger, 1986 Molecular genetics of transposable elements in plants. Annu. Rev. Genet. **20:** 175–200.

Döring, H. P., E. Tillmann and P. Starlinger, 1984 The origin and behavior of mutable loci in maize. Nature **307:** 127–130.

Doseff, A., R. Martinssen and V. Sundaresan, 1991 Somatic excision of the Mu1 transposable element of maize. Nucleic Acids Res. **19:** 579–584.

Engels, W. R., D. M. Johnson-Schlitz, W. B. Eggleston and J. Sved, 1990 High-frequency P element loss in *Drosophila* is homolog dependent. Cell **62:** 515–525.

Formosa, T., and B. M. Alberts, 1986 DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. Cell **47:** 793–806.

Gierl, A., 1996 The En/Spm transposable element of maize. Curr. Top. Microbiol. Immunol. **204:** 145–159.

Gloor, G. B., N. A. Nassif, D. M. Johnson-Schlitz, C. R. Preston and W. R. Engels, 1991 Targeted gene replacement in *Drosophila* via P element-induced gap repair. Science **253:** 1110–1117.

Hehl, R., W. Nacken, A. Krause, H. Saedler and H. Sommer, 1991 Structural analysis of Tam3, a transposable element from *Antirrhinum majus*, reveals homologies to the Ac element from maize. Plant Mol. Biol. **16:** 369–371.

Holmes, A. M., and J. E. Haber, 1999 Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. Cell **96:** 415–424.

Hsia, A., and P. S. Schnable, 1996 DNA sequence analyses support the role of interrupted gap repair in the origin of internal deletions of the maize transposon, *MuDR*. Genetics **142:** 603–618.

Kidwell, M. G., 1979 Hybrid dysgenesis in Drosophila melanogaster. The relationship between the P-M and I-R interaction systems. Genet. Res. **33:** 205–217.

Kidwell, M. G., 1983 Evolution of hybrid dysgenesis determinants in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA **88:** 1655–1659.

Kimura, M., 1980 A simple method for estimating evolutionary rates of base substitution comparative studies of nucleotide sequence. J. Mol. Evol. **16:** 111–120.

Kishima, Y., S. Yamashita and T. Mikami, 1997 Immobilized copies with nearly intact structure of the transposon Tam3 in *Antirrhinum majus*: implications for cis-element related to the transposition. Theor. Appl. Genet. **95:** 1246–1251.

Kishima, Y., S. Yamashita, C. Martin and T. Mikami, 1999 Structural conservation of the transposon Tam3 family in *Antirrhinum majus* and estimation of the number of copies able to transpose. Plant Mol. Biol. **39:** 299–308.

Kunze, R., 1996 The maize transposable element Activator (Ac). Curr. Top. Microbiol. Immunol. **204:** 161–194.

Lankenau, D. H., and G. B. Gloor, 1998 In vivo gap repair in Drosophila: a one-way street with many destinations. Bioessays **20:** 317–327.

Lankenau, D. H., V. G. Corces and W. R. Engels, 1996 Comparison of targeted-gene replacement frequencies in *Drosophila melanogaster* at the forked and white loci. Mol. Cell. Biol. **16:** 3535–3544.

Martin, C., R. Carpenter, H. Sommer, H. Saedler and E. S. Coen, 1985 Molecular analysis of instability in flower pigmentation of *Antirrhinum majus*, following isolation of the *pallida* locus by transposon tagging. EMBO J. **4:** 1625–1630.

Martin, C., A. Prescott, C. Lister and S. Mackay, 1989 Activity of the transposon Tam3 in *Antirrhinum* and tobacco: possible role of DNA methylation. EMBO J. **8:** 997–1004.

McClintock, B., 1950 The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. USA **36:** 344–355.

Moore, H., P. W. Greenwell, C.-P. Liu, N. Arnheim and T. D. Petes, 1999 Triplet repeats form secondary structures that escape DNA repair in yeast. Proc. Natl. Acad. Sci. USA **96:** 1504–1509.

Nassif, N., J. Penney, S. Pal, W. R. Engels and G. B. Gloor, 1994 Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol. Cell. Biol. **14:** 1613–1625.

Plasterk, R. H. A., and J. T. M. Groenen, 1992 Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. EMBO J. **11:** 287–290.

Platt, T., 1986 Transcription termination and the regulation of gene expression. Ann. Rev. Biochem. **55:** 339–372.

Rinehart, T. A., C. Dean and C. F. Weil, 1997 Comparative analysis of non-random DNA repair following *Ac* transposon excision in maize and *Arabidopsis*. Plant J. **12:** 1419–1427.

Rubin, E., and A. Levy, 1997 Abortive gap repair: underlying mechanism for *Ds* element formation. Mol. Cell. Biol. **17:** 6294–6302.

Saitou, N., and M. Nei, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Santalucia, J., Jr., 1998 A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc. Natl. Acad. Sci. USA **95:** 1460–1465.

Smit, A. F., and A. D. Riggs, 1996 *Tiggers* and other DNA transposon fossils in the human genome. Proc. Natl. Acad. Sci. USA **93:** 1443–1448.

Sommer, H., and H. Saedler, 1986 Structure of the chalcone synthase gene of *Antirrhinum majus*. Mol. Gen. Genet. **202:** 429–434.

Sommer, H., R. Carpenter, B. J. Harrison and H. Saedler, 1985 The transposable element Tam3 of *Antirrhinum majus* generates a novel type of sequence alterations upon excision. Mol. Gen. Genet. **199:** 225–231.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 CLUSTAL

W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

Wolfe, K. H., P. M. Sharp and W.-H. Li, 1989 Rates of synonymous substitution in plant nuclear genes. J. Mol. Evol. **29:** 208–211.

Yamashita, S., T. Mikami and Y. Kishima, 1998 Tam3 in *Antirrhinum majus* is an exceptional transposon in resistant to alteration by abortive gap repair: identification of nested transposons. Mol. Gen. Genet. **259:** 468–474.

Communicating editor: V. Sundaresan