Copyright © 1998 by the Genetics Society of America

Factors Affecting Transposition of the *Himar1 mariner* Transposon *in Vitro*

David J. Lampe, Theresa E. Grant and Hugh M. Robertson

Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Manuscript received September 10, 1997
Accepted for publication January 15, 1998

ABSTRACT

Mariner family transposable elements are widespread in animals, but their regulation is poorly understood, partly because only two are known to be functional. These are particular copies of the *Dmmar1* element from *Drosophila mauritiana*, for example, *Mos1*, and the consensus sequence of the *Himar1* element from the horn fly, *Haematobia irritans*. An *in vitro* transposition system was refined to investigate several parameters that influence the transposition of *Himar1*. Transposition products accumulated linearly over a period of 6 hr. Transposition frequency increased with temperature and was dependent on Mg^{2+} concentration. Transposition frequency peaked over a narrow range of transposase concentration. The decline at higher concentrations, a phenomenon observed *in vivo* with *Mos1*, supports the suggestion that *mariners* may be regulated in part by "overproduction inhibition." Transposition frequency decreased exponentially with increasing transposon size and was affected by the sequence of the flanking DNA of the donor site. A noticeable bias in target site usage suggests a preference for insertion into bent or bendable DNA sequences rather than any specific nucleotide sequences beyond the TA target site.

HIMAR1 is an irritans subfamily member of the *mariner* family of class II, DNA-mediated, or short-inverted, terminal repeat-type transposable elements, and it is one of only two known active *mariner* elements. It was isolated from the horn fly, *Haematobia irritans*. The active copy is a reconstructed consensus sequence based on a series of genomic clones, each of which differ from the consensus at several positions (Robertson and Lampe 1995; Lampe *et al.* 1996). Closely related *mariners* (>95% encoded amino acid identity) were found in the genomes of a mosquito, *Anopheles gambiae*; a green lacewing, *Chrysoperla plorabunda*; and a drosophilid fly, *Drosophila ananassae*, and they are apparently the result of recent horizontal transfers from unknown sources (Robertson and Lampe 1995). The presence of very similar elements in divergent species indicated that *mariners* are capable of functioning in many genomic environments.

Activity of *mariners* in diverse hosts can be attributed to their very simple requirements for transposition. *Himar1* can complete transposition *in vitro* using only its purified transposase (Lampe *et al.* 1996), and it is active as an autonomous element in *D. melanogaster* (D. J. Lampe, unpublished results). The mauritiana subfamily *mariner* element, *Dmmar1* (particular copy = *Mos1* and referred to as such hereafter), first isolated from *D. mauritiana*, has been artificially transferred into *D. melanogaster* and *D. virilis* (Garza *et al.* 1991; Lohe and Hartl 1996a), and it is even functional in the protist *Leishmania major* (Gueiros-Filho and Beverley 1997), dramatically demonstrating a lack of species-specific host factor requirements.[1]

The ubiquity of *mariners* and their apparent ease of transfer between species does not mean that the activity of these elements is unregulated. Indeed, their regulation is complex and most likely has both inherent and stochastic components (Hartl *et al.* 1997). Each time a *mariner* finds its way into a new genome via horizontal transfer, it must create copies to ensure its persistence; however, the damage that creating those copies inevitably inflicts on the host must be limited. We have used our *in vitro* assay for *Himar1* activity (Lampe *et al.* 1996) to investigate parameters affecting its transposition, and our results suggest control points for activity. We confirm the finding for *Mos1* that elevated concentrations of transposase decrease transposition frequency (Lohe and Hartl 1996b; Hartl *et al.* 1997), a feature that may have regulatory consequences. Further results show a propensity for *Himar1* to insert into target TA dinucleotides that lie in regions of bent or bendable DNA and an effect of flanking DNA sequence at the donor site. These findings suggest that *mariners* have regulatory features that allow both their spread and their coexistence with their hosts.

Corresponding author: David Lampe, Department of Entomology, University of Illinois at Urbana-Champaign, 505 S. Goodwin, Urbana, IL 61801. E-mail: d-lampe@uiuc.edu

[1]Because of the great number of *mariner* family sequences published to date, we follow the formalized naming system for *mariner* family elements suggested in Robertson and Asplund (1996). By this convention, the original element found in *D. mauritiana* is named *Dmmar1*. The *Mos1* copy of *Dmmar1* was named before this convention came into use and we will refer to it by this name.

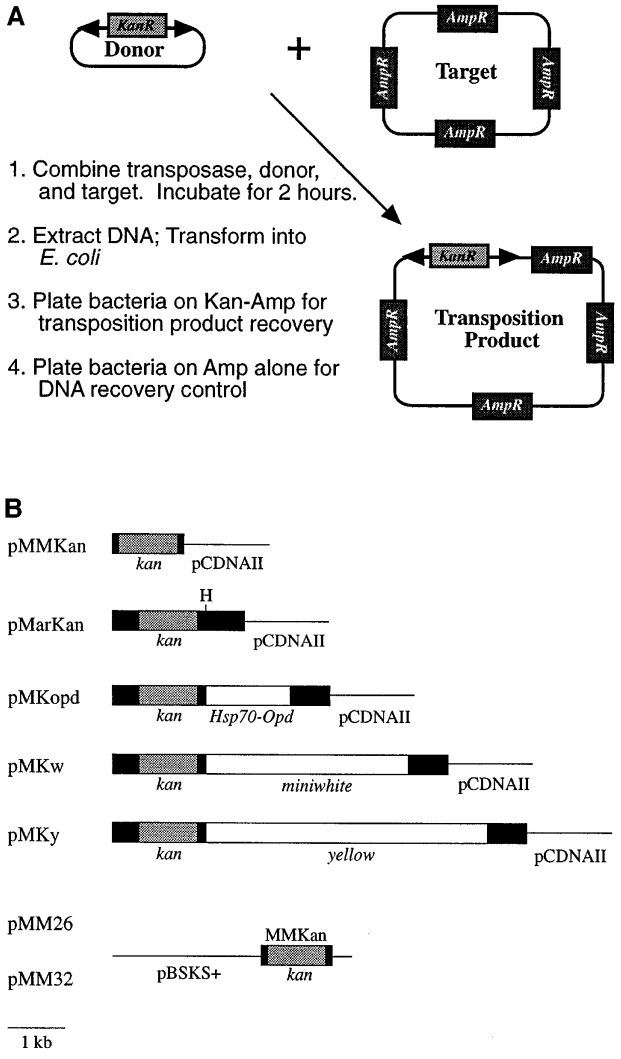

Figure 1.—(A) An outline of the *in vitro* reaction for *Himar1* transposition (Lampe *et al.* 1996). (B) The structure of transposon donor plasmids used in this study. Black boxes are *Himar1* sequences. Gray boxes are the *Kan*^{r} gene. White boxes are exogenous DNAs used to increase the size of the transposon and are labeled beneath each. Lines are the plasmid backbone. "H" in donor pMarKan indicates the *Hpa*I site where exogenous DNAs were inserted. Donors pMM26 and pMM32 have the MMKan transposon inserted at positions 2611 and 2622 in pBSKS+, respectively.

MATERIALS AND METHODS

Recombinant DNA: Donor constructs carrying different size *Himar1* transposons used in the *in vitro* transposition reaction (outlined in Figure 1A) were created by adding various DNA fragments to the *Hpa*I site of pMarKan (Figure 1B; Lampe *et al.* 1996). The DNAs inserted into pMarKan were a 5-kb *Sal*I fragment containing the intronless *D. melanogaster yellow* gene from pCy4Y (Patton *et al.* 1992; gift from P. Geyer), a 3.6-kb *Bam*HI fragment from pBS-walL (gift from G. Gloor) containing a *D. melanogaster white* minigene (Keeler *et al.* 1996), and a 1.5-kb *Xho*I/*Bam*HI fragment from pM[(opd)A] (gift from M. Benedict) containing the *hsp70-opd* gene (Benedict *et al.* 1995).

A transposon smaller than pMarKan was constructed by first creating a "minimariner" consisting of the first and last 100 bp of *Himar1* using a PCR-ligation-PCR technique exactly as described by Ali and Steinkasserer (1995). The primer pairs were T7 (TAATACGACTCACTATAGGG) and 76rSma (TAC CCGGGAATCATTTGAAGGTTGGTAC) and SP6 (CGATTT AGGTGACACTATAG) and 1218f (TCGCTCTTGAAGGGAA CTATG). The final PCR of the procedure used primers SP6 and T7 only. Cloning of the final PCR product produced the clone pMinimariner. The minimariner-based donor (pMMKan) for the *in vitro* transposition reaction was created essentially as was pMarKan (Lampe *et al.* 1996), except that the *Kan*^{r} gene was inserted into the *Sma*I site of pMinimariner.

Donor constructs containing different flanking DNAs were created from *in vitro* transposition products. MMKan (the *Himar1* transposon portion of the plasmid pMMKan) insertions into target TA dinucleotides at positions 2611 and 2622 of the target plasmid were cleaved with *Xho*I, and the products were separated on a 0.5% 1× TAE agarose gel. Insertions at these positions were known to be in the *Amp*^{r} gene, so the monomer of the target containing the insertion should be Kan^r (kanamycin resistant) but ampicillin sensitive (Amp^s). The 4.3-kb band containing one target monomer and the MMKan insertion was isolated, diluted to <10 ng/μl with TE, and 1 μl was used in a 10-μl ligation to recircularize the plasmid. The ligation products were transformed into bacteria, and the cells were plated on LB-kan plates. Subsequent plating onto LB-amp plates confirmed that the bacteria were Amp^s. Because these insertions were in the *Amp*^{r} gene of the plasmid, no further manipulations were necessary to derive new donor constructs. These donors are pMM26 (insert at 2611) and pMM32 (insert at 2622).

***In vitro* transposition reactions:** *In vitro* transposition reactions were performed as described in Lampe *et al.* (1996) unless otherwise specified. The donor DNA for these reactions is a *Himar1* transposon into which a *Kan*^{r} gene has been placed. The target is a tetramer of pBSKS+ (a pUC-like plasmid). This kind of target is used so that insertions into the *Amp*^{r} gene or an origin of replication can still be recovered. Transposase concentration was determined by measuring its absorbance at 280 nm and using a theoretically derived molar extinction coefficient (E = 53,510; Gill and von Hippel 1989). The basic transposition reactions contained 10% glycerol, 25 mm Hepes, pH 7.9, 250 μg acetylated BSA, 2 mm dithiothreitol, 100 mm NaCl, 5 mm $MgCl_2$, 30 nm purified transposase, 100 ng target plasmid, and an equimolar amount of donor plasmid in a 20-μl volume. Reactions were allowed to proceed for 2 hr at 25° before 80 μl of stop solution (50 mm Tris-HCl, pH 7.6, 0.5 mg/ml proteinase K, 10 mm EDTA, 250 μg/ml yeast tRNA) was added, and the reactions were incubated at 37° for an additional 30 min. The reactions were then phenol-chloroform extracted and ethanol precipitated using standard techniques. The precipitated reaction products were resuspended in 10 μl TE (10 mm Tris-HCl, pH 8.0, 1 mm EDTA), and 1 μl was used to transform competent *Escherichia coli* by electroporation. The cells were grown at 37° for only 30 min to ensure that each event was unique. Suitable dilutions of each reaction were plated on ampicillin-LB agar plates (to score for DNA recovery) and also on ampicillin-kanamycin LB agar plates (to score for transposition products).

Conditions were varied from the descriptions above as follows:

1. Time: The time at which the transposition reactions were stopped was varied between 30 and 360 min.
2. Temperature: The temperature at which the reactions were performed was varied between 15° and 40°.
3. $MgCl_2$ concentration: The $MgCl_2$ concentration was varied between 0 and 15 mm.
4. Transposase concentration: Transposase concentration was

varied between 0.6 and 600 nm.

5. Transposon size: The size of the donor construct was varied between 1.3 and 7.3 kb, as described above. Each donor was used in an equimolar concentration to that of the target DNA.
6. Donor flanking DNA: The effect of flanking DNA was tested by using the donors pMM26 and pMM32 (see above).

Transposition frequencies were previously reported as the number of Kan^r-Amp^r colonies/(number of Amp^r colonies $\times 10^{-3}$) (Lampe *et al.* 1996). Data were tested for significance where appropriate, either by regression analysis or pairwise *t*-tests, using StatView 4.01 (Abacus Concepts, Inc., Berkeley, CA).

Himar1 insertion sites into the target plasmid were ascertained by sequencing outward from the transposon across the *Himar1*/target junction and comparing the flanking DNA sequence to that of the known sequence of the target. The products were picked randomly from plates of transposition products produced under the standard conditions. Products from reactions performed on different days were pooled. We attempted to ensure that the colonies picked were independent by allowing the growth of bacteria transformed with transposition products to grow no more than 30 min. The number of the insertion site position is that of the T nucleotide in the TA dinucleotide target sequence using the forward orientation of the pBSKS+ sequence from GenBank (accession number X52331).

A "sequence logo" for the aligned sites of insertion was generated over the internet by the WebLogo program (Schneider and Stephens 1990; http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi) using the default parameters. A sequence logo is a way of presenting a group of aligned sequences that shows not only the consensus sequence, but also the functional importance of each nucleotide position, if any. The nucleotides at each position are shown in a "stack," with the most frequent nucleotides on top of the stack and the height of the nucleotide in the stack being proportional to its frequency at a given position. The height of the overall stack is measured in "bits," and it is a measure of the information content of that nucleotide position. Tall stacks represent important functional positions in the DNA.

RESULTS

Results obtained by varying reaction conditions are presented in raw and transformed formats in Table 1 and graphically in Figures 2–5. We previously reported an assay to detect the transposition of the *Himar1* transposon *in vitro* (Lampe *et al.* 1996). No effort was made to rigorously quantify the assay at that time, but a quantitative assay is desirable to investigate various aspects of the biochemistry of *Himar1* transposase. Three separate treatments were examined to determine if the *in vitro* assay was quantitative based on reasonable expectations for *Himar1* transposase activity. Transposition products increased over a period between 30 and 360 min, indicating that both the transposase and the transposition products were stable for the period of the assay (Figure 2A). Transposition frequency increased with increasing temperature from 15° to 28°, but it declined at higher temperatures (Figure 2B). $MgCl_2$ was absolutely required for transposition, but 5 mm was sufficient for maximal activity (Figure 2C). These results

TABLE 1

The effect of changing different reaction parameters on the *in vitro* transposition frequency of *Himar1*

Reaction parameter[a]	Absolute transposition frequency[b]	Relative transposition frequency[c]
Time (min)		
30 (5)	0.35 (±0.07)	0.02 (±0.004)
60 (4)	2.02 (±0.76)	0.12 (±0.05)
180 (3)	5.61 (±0.35)	0.34 (±0.02)
360 (5)	16.69 (±3.40)	1.0 (±0.20)
Temperature		
15.5° (4)	0.94 (±0.13)	0.11 (±0.02)
23° (4)	3.83 (±0.48)	0.45 (±0.06)
28° (4)	8.52 (±0.49)	1.0 (±0.06)
35° (3)	6.02 (±0.42)	0.71 (±0.05)
40° (4)	3.63 (±0.21)	0.42 (±0.02)
$MgCl_2$ (mm)		
0 (4)	0.00	0.00
5 (3)	5.45 (±0.90)	0.82 (±0.14)
10 (4)	5.64 (±0.89)	1.00 (±0.13)
15 (4)	5.68 (±0.96)	0.86 (±0.14)
Transposase (nm)		
0.6 (3)	0.19 (±0.06)	0.03 (±0.01)
1.2 (3)	0.87 (±0.05)	0.14 (±0.01)
6.0 (3)	6.39 (±0.57)	1.0 (±0.09)
15 (2)	5.85 (±0.67)	0.92 (±0.12)
60 (3)	1.0 (±0.06)	0.16 (±0.01)
150 (3)	0.64 (±0.09)	0.10 (±0.01)
300 (3)	0.20 (±0.06)	0.03 (±0.01)
600 (3)	0.05 (±0.01)	0.01 (±0.001)
Transposon size (kb)		
1.3 (4)	17.97 (±4.27)	1.0 (±0.10)
2.3 (12)	14.19 (±2.40)	0.82 (±0.14)
3.8 (5)	3.10 (±0.42)	0.18 (±0.02)
5.7 (3)	1.53 (±0.14)	0.09 (±0.01)
7.3 (3)	1.17 (±0.50)	0.07 (±0.03)
Donor flanking DNA		
pMMKan (4)	4.34 (±0.58)	1.0 (±0.13)
pMM26 (4)	10.47 (±0.60)	2.40 (±0.14)
pMM32 (4)	5.93 (±0.26)	1.36 (±0.06)

[a] Numbers in parentheses indicate the number of replicates for that treatment.

[b] Average absolute transposition frequency. Absolute frequencies for each parameter are calculated by dividing the number of Kan^r-Amp^r colonies produced in the *in vitro* transposition reaction by the number of Amp^r colonies and multiplying by 1000. The average is the sum of the individual frequencies for each parameter divided by the number of replicates. Each error is the standard error of the mean, which is calculated as the parameter standard deviation divided by the square root of the number of replicates.

[c] Relative frequencies are calculated by dividing each of the absolute frequencies in a given parameter group by the average absolute maximum transposition frequency within that group. In the case of the donor flanking DNA data set, the absolute transposition frequency for the donor containing the flanking DNA used for all other reactions, pMMKan, was used for normalization. Relative errors are computed by dividing the absolute errors within a group by the maximum average absolute frequency within that group.

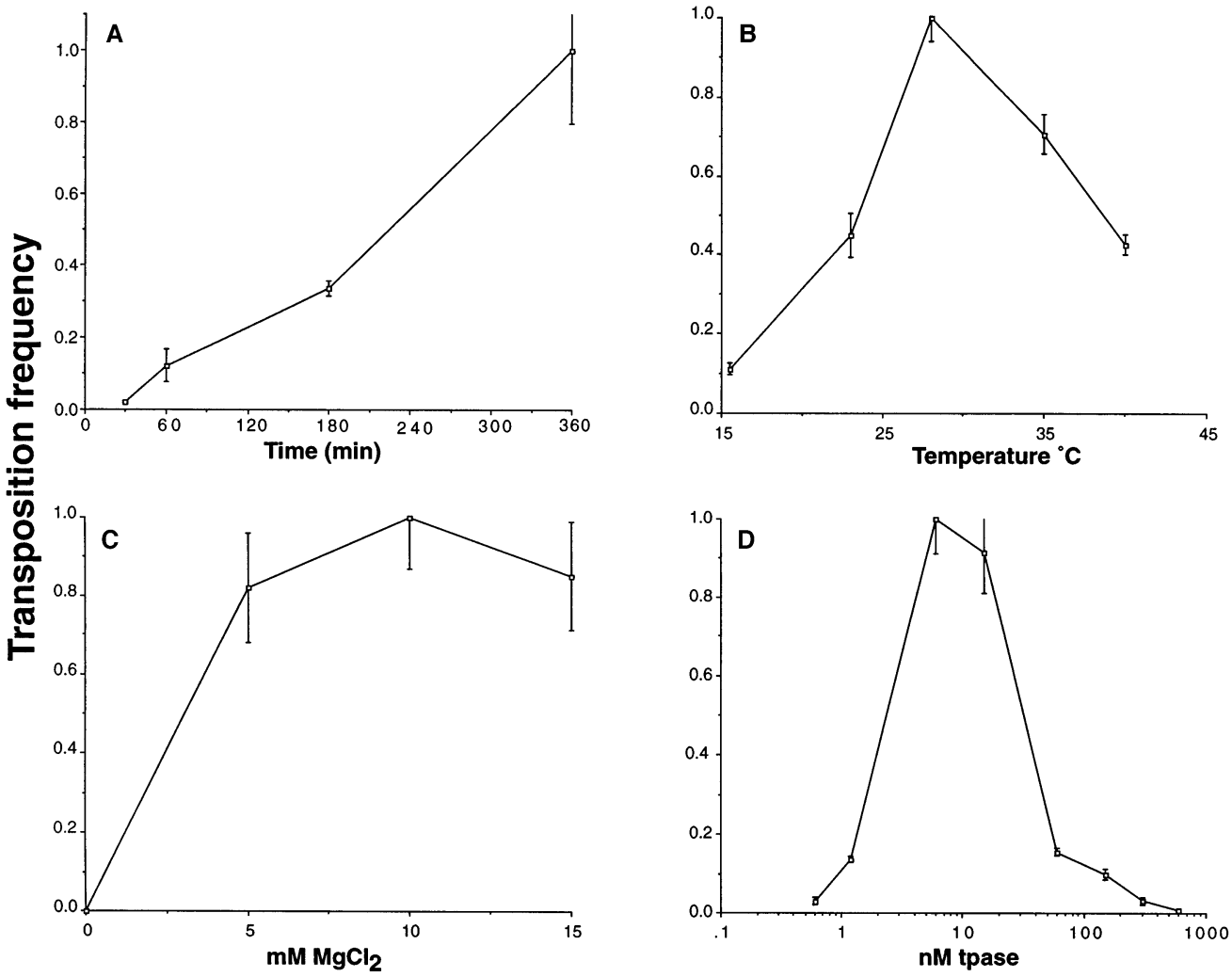

Figure 2.—Effect of changing several conditions on the transposition frequency of *Himar1 in vitro*. Transposition frequency is plotted on the *y*-axis. Each specific treatment is plotted on the *x*-axis. The value of 1.0 was used for the maximum transposition frequency for a particular treatment in each experiment, and the other treatments were normalized to it (see Table 1). Data points are averages ($N \geq 3$). Error bars are SEM.

are similar to those found for many eukaryotic and prokaryotic transposons (Berg and Howe 1989).

Transposase was required for transposition and was most active (>50% maximal activity) between 2.5 and 35 nm transposase, peaking at ~10 nm (Figure 2D). This result is similar to that found for purified Tn*10* transposase, which is also inhibited at high transposase concentrations (Chalmers and Kleckner 1994), although not to this degree. Recent findings using *Mos1* indicate that these *in vitro* results may reflect a property inherent in *mariner* transposases (Lohe and Hartl 1996b; Hartl *et al.* 1997).

Transposon size dramatically affected the frequency of transposition (Figure 3). Smaller elements transposed much more frequently than larger ones. We fitted an exponential curve to these data based on the behavior of other transposons (Way and Kleckner 1985). The curve is described by the function $y = 1.8121 \times 10^{(-0.2128x)}$ and has an $r^2 = 0.93$. The fitted curve suggests that transposition frequency decreases ~38% for each 1-kb increase in transposon size.

Random transposition products ($N = 65$) were selected and sequenced to determine the site and orientation of *Himar1* insertion into the target plasmid. The sequence, position, orientation, and frequency of use of each insertion site are shown in Figure 4A. A graphical representation of the data in the form of a sequence logo (Schneider and Stephens 1990) is shown in Figure 4B. *Himar1* used a variety of sites throughout the target plasmid, invariably inserting into a TA dinucleotide. Some sites were used preferentially (*e.g.*, position 2611). Other sites were used preferentially and exclusively in one orientation (*e.g.*, position 1659), and others were used in both orientations (*e.g.*, position 1947). It is unclear why a given site should be used in only one orientation. As shown in the sequence logo, there is

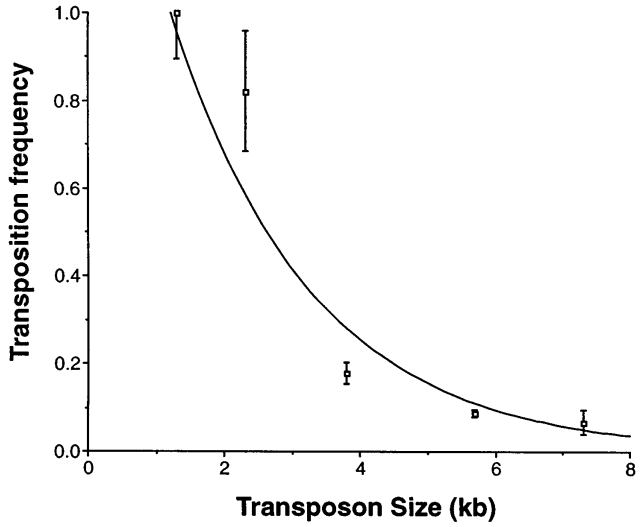

Figure 3.—Effect of transposon size on transposition frequency. Average ($N \geq 3$) of transposition reactions using variously sized donor constructs (see Figure 1B). Error bars are SEM. Data were normalized to the frequency shown by the pMMKan construct, which was most active and whose frequency was given a value of 1.0 (see Table 1).

no significant sequence similarity in the sites used for insertion beyond the TA dinucleotide sequence (Figure 4B), and this was the case even when only the preferred sites are incorporated into the data set (data not shown).

The effect of flanking DNA sequence on transposition frequency was tested by creating two new donor constructs. These were derived from transposition products in which the *Himar1* transposon was inserted into the *Amp*^{r} gene of the target plasmid, rendering the transposon-containing monomer Kan^r but Amp^s (see materials and methods). pMM26, derived from the insertion at site 2611, showed a 2.4-fold increase in transposition frequency, whereas pMM32, derived from the insertion at site 2622, was only slightly more active (1.36-fold higher) than the original pMMKan construct (Figure 5).

DISCUSSION

Mariner family transposable elements are extremely widespread in animals, occurring in several phyla (Robertson 1993, 1997; Garcia-Fernandez *et al.* 1995; Robertson *et al.* 1997). They are unusual in that they seem to undergo horizontal transfer between species with comparative ease. Artificial transfers of these elements have been accomplished across taxonomic groups as divergent as kingdoms (Gueiros-Filho and Beverley 1997). We have used an *in vitro* assay to examine several properties of one *mariner*, the *Himar1* element found in the horn fly *H. irritans.* The results confirm some properties found *in vivo* with the only other active *mariner*, the *Mos1* element of *D. mauritiana. Mos1* and *Himar1* represent highly divergent subfamilies of *mariners*, and thus these results may be generalizable to the entire *mariner* family of transposable elements.

Transposition of *Himar1 in vitro* is markedly affected by the concentration of the transposase protein, and it is most active over a narrow range of transposase concentration, peaking at ~10 nm. Significantly, these data mirror what occurs *in vivo* for *Mos1*. *Mos1* can mobilize a nonautonomous *mariner* element (w^{peach}) from the *white* locus in *D. mauritiana,* leading to a mosaic eye phenotype (Medhora *et al.* 1988). This assay system has been transferred into the more genetically tractable species *D. melanogaster* (Garza *et al.* 1991), where increasing the amount of *Mos1* transposase by either increasing the copy number of the element or by inducing a transposase source carrying a *hsp70* promoter led to a decrease in reversion frequency of w^{peach} (Lohe and Hartl 1996b). Although the precise concentration of transposase was not measured, these data imply that *Mos1 mariner* transposase also functions at an optimal protein concentration, with elevated concentrations lowering transposition rate. This phenomenon has been termed "overexpression inhibition," and a regulatory role has been suggested for this property (Lohe and Hartl 1996b; Hartl *et al.* 1997). These authors further suggested that this might be a unique property for *mariner* transposases, but, in fact, decreased transposition at elevated transposase concentration is known for other transposons (*e.g.*, Wiegand and Reznikoff 1992; Kunze *et al.* 1993). In the case of Tn*5*, the formation of inactive multimers of transposase seems to underlie the phenomenon (Weinreich *et al.* 1994).

The important question is whether the property of reduced transposition at elevated transposase concentration is regulatory and under selection, or if it is merely coincidental. All transposable elements have some kind of restrictions on activity, and it would be strange if *mariners* lacked these. Isolation of mutant *mariners* that eliminated concentration-dependent inhibition would help resolve this problem. We would predict that such mutants would be hyperactive at high transposase concentrations because more transposase would be available to participate in a transposition reaction rather than being bound up in a nonactive form, whether in inactive multimers or some unstructured aggregation.

Increasing transposon size decreases transposition frequency: The transposition frequency of *Himar1* decreased exponentially with increasing transposon size. This effect has been demonstrated with other transposable elements, but the degree to which it occurs depends on the element and, at least in bacteria, on whether the donor site is on a chromosome or a plasmid. For example, the bacterial transposon Tn*10* shows a decrease in transposition frequency of 40% per kb when mobilized from the bacterial chromosome but only 16% when mobilized from a plasmid (Way and Kleckner 1985). *Himar1* shows a 38% decrease in transposition frequency from a small plasmid *in vitro* for every 1-kb increase

A

Sequence	Position	N=	+/-
GGAACAAGAGTCCAC**TA**TTAAAGAACGTGGAC	165	1	-
ACAAGAGTCCACTAT**TA**AAGAACGTGGACTCC	168	1	+
GTGCTGCAAGGCGAT**TA**AGTTGGGTAACGCCA	564	1	-
CCAGGCTTTACACTT**TA**TGCTTCCGGCTCGTA	879	1	-
ACTCAAAGGCGGTAA**TA**CGGTTATCCACAGAA	1113	1	+
GAAACCCGACAGGAC**TA**TAAAGATACCAGGCG	1290	1	+
TCATAGCTCACGCTG**TA**GGTATCTCAGTTCGG	1420	1	+
TTGGAGCGAACGACC**TA**CACCGAACTGAGATA	1439	1	-
GCCTTATCCGGTAAC**TA**TCGTCTTGAGTCCAA	1514	2	+
TTGGACTCAAGACGA**TA**GTTACCGGATAAGGC			-
TAGTGTAGCCGTAGT**TA**GGCCACCACTTCAAG	1633	1	-
ACTGTCCTTCTAGTG**TA**GCCGTAGTTAGGCCA	1643	1	-
ACACTAGAAGGACAG**TA**TTTGGTATCTGCGCT	1659	4	+
TCTGCTGAAGCCAGT**TA**CCTTCGGAAAAAGAG	1690	1	+
TTTTGGTCATGAGAT**TA**TCAAAAAGGATCTTC	1881	1	+
AAAAGGATCTTCACC**TA**GATCCTTTTAAATTA	1901	1	+
AAGTTTACTCATATA**TA**CTTTAGATTGATTTA	1945	1	-
AATCAATCTAAAGTA**TA**TATGAGTAAACTTGG	1947	1	+
CCAAGTTTACTCATA**TA**TACTTTAGATTGATT		3	-
TCAATCTAAAGTATA**TA**TGAGTAAACTTGGTC	1949	1	+
GACCAAGTTTACTCA**TA**TATACTTTAGATTGA		3	-
CTGTCAGACCAAGTT**TA**CTCATATATACTTTA	1955	1	-
AATCAGTGAGGCACC**TA**TCTCAGCGATCTGTC	1999	1	+
GACAGATCGCTGAGA**TA**GGTGCCTCACTGATT		2	-
TCCCCGTCGTGTAGA**TA**ACTACGATACGGGAG	2058	1	+
CTCCCGTATCGTAGT**TA**TCTACACGACGGGGA		1	-
AGCCCTCCCGTATCG**TA**GTTATCTACACGACG	2062	1	-
CCCGGCAACAATTAA**TA**GACTGGATGGAGGCG	2215	1	-
CTCCATCCAGTCTAT**TA**ATTGTTGCCGGGAAG	2218	1	+
CTTCCCGGCAACAAT**TA**ATAGACTGGATGGAG		1	-
GCGAACTACTTACTC**TA**GCTTCCCGGCAACAA	2236	1	-
GCCGGGAAGCTAGAG**TA**AGTAGTTCGCCAGTT	2241	1	+
AACTGGCGAACTACT**TA**CTCTAGCTTCCCGGC		1	-
CAACGTTGCGCAAAC**TA**TTAACTGGCGAACTA	2260	1	-
ACGATCAAGGCGAGT**TA**CATGATCCCCCATGT	2362	1	+
GATCGTTGTCAGAAG**TA**AGTTGGCCGCAGTGT	2428	1	+
CATAACCATGAGTGA**TA**ACACTGCGGCCAACT	2445	1	-
CACAGAAAAGCATCT**TA**CGGATGGCATGACAG	2499	2	-
ATTCTGAGAATAGTG**TA**TGCGGCGACCGAGTT	2554	1	+
AACTCGGTCGCCGCA**TA**CACTATTCTCAGAAT		1	-
TATGTGGCGCGGTAT**TA**TCCCGTATTGACGCC	2596	1	-
TAATACCGCGCCACA**TA**GCAGAACTTTAAAAG	2611	1	+
CTTTTAAAGTTCTGC**TA**TGTGGCGCGGTATTA		3	-
AATGATGAGCACTTT**TA**AAGTTCTGCTATGTG	2622	3	-
AAACGCTGGTGAAAG**TA**AAAGATGCTGAAGAT	2740	1	-
AAGGAAGAGTATGAG**TA**TTCAACATTTCCGTG	2826	1	-
TTGAAAAAGGAAGAG**TA**TGAGTATTCAACATT	2832	4	-
TGATAAATGCTTCAA**TA**ATATTGAAAAAGGAA	2852	1	-
ATACATTCAAATATG**TA**TCCGCTCATGAGACA	2891	1	-

B

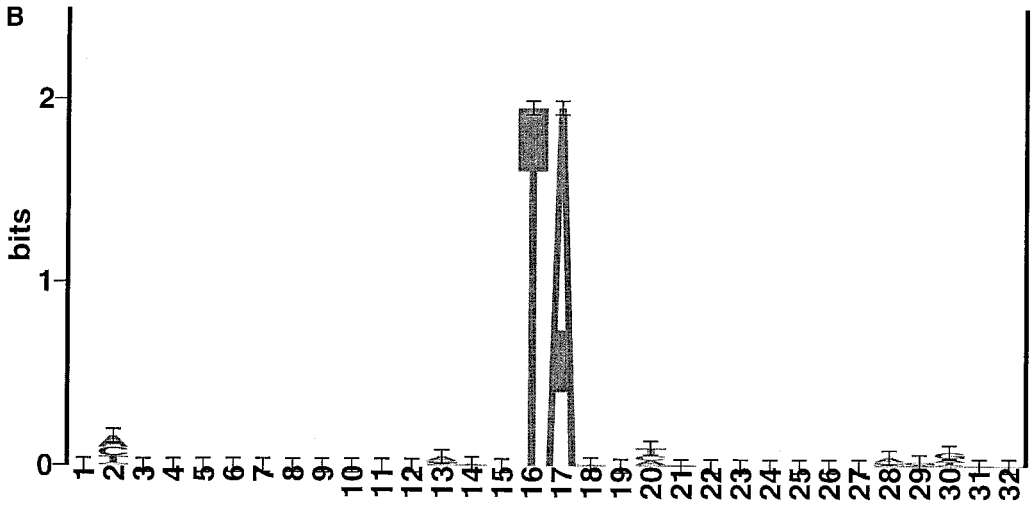

Figure 4.—Sequence, position, frequency, and orientation of insertions into a target plasmid by *Himar1*. (A) Numbers correspond to the position of the T of the TA target dinucleotide in the plasmid pBSKS+. N, frequency of use of the site; + and −, the orientation of the insertion relative to the 5′ end of the *Himar1* transposon. The TA target sequence is highlighted in bold. Sixty-five transposition products were analyzed and sequenced. (B) Sequence logo of insertion sites used by *Himar1 in vitro*. The program "Web Logo" (http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi) was used to analyze the information content of the insertion site sequences used by *Himar1* (Schneider and Stephens 1990). Each position was used once in the analysis except where both orientations were used at a particular site. In these cases, both orientations were entered into the data set. *y*-axis, bits of information for each position; *x*-axis, nucleotide position of the site of *Himar1* insertion. Positions 16 and 17 correspond to the TA target dinucleotide. The height and order of each letter in the stack of letters at each position corresponds to its relative frequency at that position. Nucleotides used most frequently are on top. Error bars indicate the variability of a comparable number of random sequences. The total height of the stack in bits is an indication of the information content of the particular site.

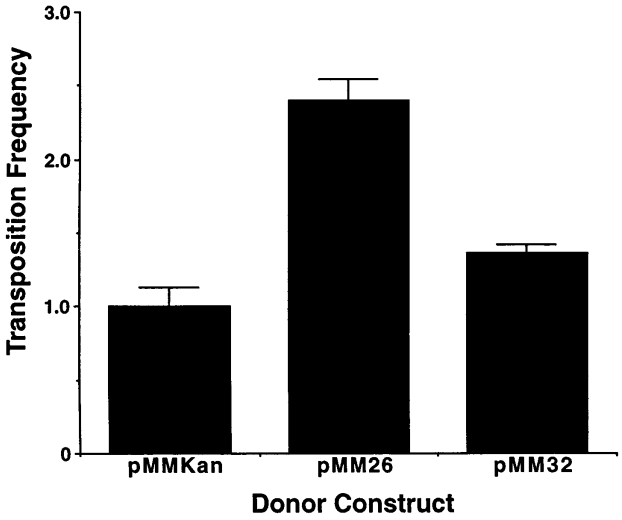

Figure 5.—Effect of donor flanking DNA sequence on transposition frequency. Values are the mean and error bars the SEM. A value of 1.0 was assigned to the pMMKan construct whose flanking DNA was identical to that of all other donors used herein and in Lampe *et al.* (1996). The means are significantly different (Student's *t*-test, $P \leq 0.05$).

in transposon length (Figure 3). If the disparities in transposition frequency for donor form hold for *Himar1* like they do for Tn*10*, then we could expect *Himar1* constructs or other *mariners* to transpose at even lower frequencies when mobilized from chromosomal sites. This phenomenon might explain the extreme stability of *Mos1* transposons carrying a *white* minigene (Lohe and Hartl 1996c). These elements were originally transferred into the *D. melanogaster* genome by transposition from a plasmid to a chromosome by injection into preblastoderm embryos. Subsequent attempts at remobilization off the chromosome to other genomic sites were unsuccessful. A greatly reduced level of transposition resulting from the form of the donor, such as that seen in bacteria, might account for this apparent stability.

DNA bending may underlie target site choice and flanking DNA effects at the donor site: *Himar1* showed a pronounced preference for insertion at some sites in the target plasmid used in the *in vitro* assay (Figure 4A). Sequence logo analysis of these sites shows that there is very little underlying sequence similarity in these sites (Figure 4B). These data are consistent with a lack of insertion site specificity (besides TA) found *in vivo* (Robertson and Lampe 1995; but see Bigot *et al.* 1994 for a possible exception) and are in contrast to other transposons, such as the *P* element and *Hermes*, which have more extensive target site preferences (Engels 1997; Sarkar *et al.* 1997). Why, then, does *Himar1* insert so frequently at certain positions? It may be that *Himar1* transposase does not recognize target DNA in any sequence-specific manner (beyond the TA dinucleotide), but that certain sequences surrounding the target TA may allow local *structures* to form that are conducive to *Himar1* insertion. Far from being a generalized flexible rod shape over its entire length, any given DNA molecule will have a certain local structure that deviates from the average, depending on the underlying sequence (Hagerman 1986; Dickerson 1992). Particular TA target sites that are significantly more "bendable" than others may be preferred for *Himar1* insertion because they might allow easier identification of the TA dinucleotide by the transposase in the first place, or, once located, they might allow the DNA to bend so that the excised transposon might be inserted more easily. There is precedence for this phenomenon in the nonrandom use of target sequences by retroviral integrases, proteins very distantly related to *Himar1* transposase (Doak *et al.* 1994). These proteins insert their cognate DNAs into sequences that show a pronounced propensity for bending (*e.g.*, Milot *et al.* 1994; Müller and Varmus 1994).

Fortuitously, the DNA bending propensity of the exact target region most preferred by *Himar1* for insertion was experimentally measured elsewhere by Brukner *et al.* (1995). These workers used fragments of the plasmid pUC18 to compare the actual degree of DNA bending of DNA sequences by DNaseI to that of a theoretical model. The sequence of our target plasmid centered around position 1945–1955, the short stretch of the target plasmid used most often by *Himar1* for insertion (10/65 insertions; see Figure 4A), showed a high degree of DNA bending in the DNaseI study (Brukner *et al.* 1995). This coincidence strongly suggests that DNA bending of the target sequence plays a significant role in the ability of *Himar1* to use a given target TA.

Similar arguments can explain the effects of flanking DNA sequence at the donor site on overall transposition frequency. Several transposases, including the related Tc*3* transposase, bend the DNA at the transposon termini upon binding (Arciszewska and Craig 1991; Derbyshire and Grindley 1992; von Pouderoyen *et al.* 1997; York and Reznikoff 1997). This bending is presumably a requirement for the formation of a DNA structure that can be cleaved by the transposase. Because the bending occurs so close to the end of the element, flanking sequences are likely to be involved, at least indirectly, in the formation of this structure. Sequences that would interfere with that bending would tend to decrease or eliminate excision and, hence, may decrease the overall frequency of transposition. Alternatively, flanking sequences could contribute to the recognition of the inverted repeat sequences by the transposase. The width of the minor groove at the end of the DNA recognition sequence for Tc*3* transposase appears to be important in its ability to make base-specific contacts, and the width of that groove is dependent on the primary DNA sequence (von Pouderoyen *et al.* 1997). Sequences that would change minor groove width could alter the ability of transposase to bind to its terminal repeat sequences.

Practical implications: The ability of *mariners* to func-

tion in diverse hosts makes them attractive candidates for development into generalized animal transformation vectors. The success of elements used in this fashion, however, has been mixed. The *Mos1* element has been used to transform *D. melanogaster* at low frequency, although the construct integrated was >13 kb in length (Lidholm *et al.* 1993). Repeated attempts to use *Himar1* as a germline transformation vector for *D. melanogaster* DNA have failed, although it is active in that species as an autonomous element (D. J. Lampe, unpublished results).

In vitro and *in vivo* studies are revealing the likely foci for the control of *mariner* activity. One feature of *mariner* transposases that may make them difficult to use as genetic tools and that may be a property under selection is the inhibitory effect of transposase at elevated concentrations. One of the most common ways to increase the activity of a heterologous genetic system is to increase its level of expression, a strategy clearly counterproductive for *mariners*. Suitable screens for mutant transposases lacking inherent control mechanisms should reveal forms of *mariners* more suitable for use as genetic tools; such strategies have worked for the bacterial transposon Tn*5* (Wiegand and Reznikoff 1992). It seems likely that only after mutants of *mariners* are found that can overcome their concentration-dependent phenotype will these elements fulfill their potential as generalized genetic tools.

We thank K. Walden for constructing the minimariner; M. Benedict, P. Geyer, and G. Gloor for the gift of plasmids; and S. Berlocher, I. Boussy, M. Churchill, P. Jambeck, D. O'Brochta, M. Simmons, and an anonymous reviewer for helpful comments. This work was supported by Public Health Service grant AI33586-01.

LITERATURE CITED

Ali, S. A., and A. Steinkasserer, 1995 PCR-ligation-PCR mutagenesis: a protocol for creating gene fusions and mutations. Biotechniques **18:** 746–750.

Arciszewska, L. K., and N. L. Craig, 1991 Interaction of the Tn*7*-encoded transposition protein TnsB with the ends of the transposon. Nucleic Acids Res. **19:** 5021–5029.

Benedict, M. Q., C. E. Salazar and F. H. Collins, 1995 A new dominant selectable marker for genetic transformation, Hsp70-opd. Insect Biochem. Mol. Biol. **25:** 1061–1065.

Berg, D. E., and M. M. Howe (Editors), 1989 *Mobile DNA*. American Society for Microbiology, Washington, DC.

Bigot, Y., M. H. Hamelin, P. Capy and G. Periquet, 1994 *Mariner*-like elements in hymenopteran species: insertion site and distribution. Proc. Natl. Acad. Sci. USA **91:** 3408–3412.

Brukner, I., R. Sánchez, D. Suck and S. Pongor, 1995 Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. EMBO J. **14:** 1812–1818.

Chalmers, R. M., and N. Kleckner, 1994 Tn*10*/IS*10* transposase purification, activation, and *in vitro* reaction. J. Biol. Chem. **269:** 8029–8035.

Derbyshire, K. M., and N. D. Grindley, 1992 Binding of the IS903 transposase to its inverted repeat *in vitro*. EMBO J. **11:** 3499–3455.

Dickerson, R. E., 1992 DNA structure from A to Z. Methods Enzymol. **211:** 3449–3455.

Doak, T. G., F. P. Doerder, C. L. Jahn and G. Herrick, 1994 A proposed superfamily of transposase-related genes: new members in transposon-like elements of ciliated protozoa and a common "D35E" motif. Proc. Natl. Acad. Sci. USA **91:** 942–946.

Engels, W. R., 1997 *P-elements in Drosophila*. Online monograph found at http://www.wisc.edu/genetics/CATG/engels/Pelements/index.html.

Garcia-Fernàndez, J., J. R. Batascas-Ramírez, G. Marfany, A. M. Muñoz-Mármol, A. Casali *et al.*, 1995 High copy number of highly similar *mariner*-like transposons in planarian (platyhelminthe): evidence for a trans-phyla horizontal transfer. Mol. Biol. Evol. **12:** 421–431.

Garza, D., M. Medhora, A. Koga and D. L. Hartl, 1991 Introduction of the transposable element *mariner* into the germline of *Drosophila melanogaster*. Genetics **128:** 303–310.

Gill, S. G., and P. H. von Hippel, 1989 Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182:** 319–326.

Gueiros-Filho, F. J., and S. M. Beverley, 1997 Trans-kingdom transposition of the Drosophila element *mariner* within the protozoan *Leishmania*. Science **276:** 1716–1719.

Hagerman, P. J., 1986 Sequence-directed curvature of DNA. Nature **321:** 449–450.

Hartl, D. L., E. R. Lozovskaya, D. I. Nurminsky and A. R. Lohe, 1997 What restricts the activity of *mariner*-like transposable elements? Trends Genet. **13:** 197–201.

Keeler, K. J., T. Dray, J. E. Penney and G. B. Gloor, 1996 Gene targeting of a plasmid-borne sequence to a double-strand DNA break in *Drosophila melanogaster*. Mol. Cell. Biol. **16:** 522–528.

Kunze, R., U. Behrens, U. Courage-Franzkowiak, S. Feldmar, S. Kühn *et al.*, 1993 Dominant transposition-deficient mutants of maize Activator (Ac) transposase, Proc. Natl. Acad. Sci. USA **90:** 7094–7098.

Lampe, D. J., M. E. A. Churchill and H. M. Robertson, 1996 A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. EMBO J. **15:** 5470–5479.

Lidholm, D. A., A. R. Lohe and D. L. Hartl, 1993 The transposable element *mariner* mediates germline transformation in *Drosophila melanogaster*. Genetics **134:** 859–868.

Lohe, A. R., and D. L. Hartl, 1996a Germline transformation of *Drosophila virilis* with the transposable element *mariner*. Genetics **143:** 365–374.

Lohe, A. R., and D. L. Hartl, 1996b Autoregulation of *mariner* transposase activity by overproduction and dominant-negative complementation. Mol. Biol. Evol. **13:** 549–555.

Lohe, A. R., and D. L. Hartl, 1996c Reduced germline mobility of a *mariner* vector containing exongenous DNA: effect of size or site? Genetics **143:** 1299–1306.

Medhora, M. M., A. H. MacPeek and D. L. Hartl, 1988 Excision of the Drosophila transposable element *mariner*: identification and characterization of the *Mos* factor. EMBO J. **7:** 2185–2189.

Milot, E., A. Belmaaxa, E. Rassart and P. Chartrand, 1994 Association of a host structure with retroviral integration sites in chromosomal DNA. Virology **201:** 408–412.

Müller, H. P., and H. E. Varmus, 1994 DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. EMBO J. **13:** 4704–4714.

Patton, J. S., X. V. Gomes and P. K. Geyer, 1992 Position-independent germline transformation in *Drosophila* using a cuticle pigmentation gene as a selectable marker. Nucleic Acids Res. **20:** 5859–5860.

Robertson, H. M., 1993 The *mariner* transposable element is widespread in insects. Nature **362:** 241–245.

Robertson, H. M., 1997 Multiple *mariner* transposons in flatworms and hydras are related to those of insects. J. Heredity **88:** 195–201.

Robertson, H. M., and D. J. Lampe, 1995 Recent horizontal transfer of a *mariner* transposable element among and between diptera and neuroptera. Mol. Biol. Evol. **12:** 850–862.

Robertson, H. M., and M. L. Asplund, 1996 *Bmmar1*: a basal lineage of the *mariner* family of transposable elements in the silkworm moth, *Bombyx mori*. Insect Biochem. Mol. Biol. **26:** 945–954.

Robertson, H. M., F. N. Soto-Adames, K. K. O. Walden, R. M. P. Avancini and D. J. Lampe, 1997 The *mariner* transposons of animals: horizontally jumping genes, pp. 268–284 in *Horizontal Gene Transfer*, edited by M. Syvanen and C. Kado. Chapman & Hall, London.

Sarkar, A., K. Yardley, P. W. Atkinson, A. A. James and D. A. O'Brochta, 1997 Transposition of the *Hermes* element in em-

bryos of the vector mosquito, *Aedes aegypti*. Insect Biochem. Mol. Biol. **27:** 359–363.

Schneider, T. D., and R. M. Stephens, 1990 Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. **18:** 6097–6100.

van Pouderoyen, G., R. F. Ketting, A. Perrakis, R. H. A. Plasterk and T. K. Sixma, 1997 Crystal structure of the specific DNA-binding domain of Tc*3* transposase of *C. elegans* in complex with transposon DNA. EMBO J. **16:** 6044–6054.

Way, J. C., and N. Kleckner, 1985 Transposition of plasmid-borne Tn*10* elements does not exhibit simple length dependence. Genetics **111:** 705–713.

Weinreich, M. D., A. Gasch and W. S. Reznikoff, 1994 Evidence that the *cis* preference of the Tn*5* transposon is caused by nonproductive multimerization. Genes Dev. **8:** 2363–2374.

Wiegand, T. W., and W. S. Reznikoff, 1992 Characterization of two hypertransposing Tn*5* mutants. J. Bacteriol. **174:** 1229–1239.

York, D., and W. S. Reznikoff, 1997 DNA binding and phasing analyses of Tn*5* transposase and a monomeric variant. Nucleic Acids Res. **25:** 2153–2160.

Communicating editor: M. J. Simmons