Copyright © 2005 by the Genetics Society of America
DOI: 10.1534/genetics.104.036145

An Active Transposable Element, *Herves*, From the African Malaria Mosquito *Anopheles gambiae*

Peter Arensburger,* Yu-Jung Kim,* Jamison Orsetti,† Channa Aluvihare,† David A. O'Brochta† and Peter W. Atkinson*,[1]

**Department of Entomology, University of California, Riverside, California 92521-0314 and †Center for Biosystems Research, University of Maryland Biotechnology Institute, College Park, Maryland 20742*

Manuscript received August 28, 2004
Accepted for publication October 26, 2004

ABSTRACT

Transposable elements have proven to be invaluable tools for genetically manipulating a wide variety of plants, animals, and microbes. Some have suggested that they could be used to spread desirable genes, such as refractoriness to Plasmodium infection, through target populations of *Anopheles gambiae*, thereby disabling the mosquito's ability to transmit malaria. To achieve this, a transposon must remain mobile and intact after the initial introduction into the genome. Endogenous, active class II transposable elements from *An. gambiae* have not been exploited as gene vectors/drivers because none have been isolated. We report the discovery of an active class II transposable element, *Herves*, from the mosquito *An. gambiae*. *Herves* is a member of a distinct subfamily of *hAT* elements that includes the *hopper-we* element from *Bactrocera dorsalis and B. cucurbitae. Herves* was transpositionally active in mobility assays performed in *Drosophila melanogaster* S2 cells and developing embryos and was used as a germ-line transformation vector in *D. melanogaster. Herves* displays an altered target-site preference from the distantly related *hAT* elements, *Hermes* and *hobo. Herves* is also present in *An. arabiensis* and *An. merus* with copy numbers similar to that found in *An. gambiae.* Preliminary data from an East African population are consistent with the element being transpositionally active in mosquitoes.

DESPITE breakthroughs in the generation of transgenic mosquitoes that, in laboratory studies, are refractory to the transmission of rodent malaria (Ito *et al.* 2002; Moreira *et al.* 2002; Kim *et al.* 2004), a key remaining obstacle to the extension of this technology to the field remains the absence of gene vectors that can efficiently drive genes through mosquito populations. Thus, while the distribution and spread of *P* and *hobo* elements through field populations of *Drosophila melanogaster* is well documented (Anxolabehere *et al.* 1988; Daniels *et al.* 1990a,b), and while the distribution of *mariner* elements in arthropods suggests that these elements were once capable of spreading both within and between genomes (Robertson and MacLeod 1993), no evidence for the ability of transposable elements to move within mosquito populations exists. This is in contrast to recent developments in gene transfer technologies in mosquitoes in which several transposable elements have been used to genetically transform mosquito species. These include the *Hermes* element from *Musca domestica* (Jasinskiene *et al.* 1998; Allen *et al.* 2001), the *piggyBac* element from *Trichoplusia ni* (Grossman *et al.* 2001; Kokoza *et al.* 2001; Moreira *et al.* 2002; Perera *et al.* 2002), the *Minos* element from *Drosophila hydei* (Catteruccia *et al.* 2000), and the *Mos1* element from *Drosophila mauritiana* (Coates *et al.* 1998). The remobilization properties of *Hermes*, *piggyBac*, and *Mos1* in somatic and germ-line nuclei were examined in transgenic lines of the yellow fever mosquito, *Aedes aegypti* (O'Brochta *et al.* 2003; Wilson *et al.* 2003). Little or no evidence of germ-line remobilization was found for these transposable elements in these transgenic lines of mosquito, indicating that these elements may not be suitable agents for driving transgenes through insect populations. The basis for this low frequency of remobilization is not known; however, it is consistent with the behavior of *mariner* elements in transgenic lines of *D. melanogaster* in which remobilization of these modified elements is also low (Lohe and Hartl 1996; Lozovsky *et al.* 2002).

We were interested in determining if active transposable elements were present in the genome of the malaria vector *Anopheles gambiae.* This mosquito species is the principal vector of the pathogen of human malaria, *Plasmodium falciparum*, in Africa and so is the target for genetic control and population replacement strategies aimed at preventing the transmission of Plasmodium through the female mosquito. An active transposable element from *An. gambiae* might serve as a platform for constructing gene vectors from the target mosquito

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession no. AY462096.

[1]*Corresponding author:* Department of Entomology, University of California, 900 University Ave., Riverside, CA 92521-0001.
E-mail: peter.atkinson@ucr.edu

species itself and would also facilitate the study of transposable element movement and spread through *An. gambiae* populations. We report here the isolation of *Herves*, an active transposable element from *An. gambiae*. *Herves* is a member of the *hAT* superfamily of transposable elements but is only distantly related to the *Hermes* and *hobo* elements that also are found in insects and are active. This element was discovered *in silico* through the use of a unique algorithm designed to identify active *hAT* elements from the output of genome-sequencing projects on the basis of conserved structural features of these elements in addition to conserved amino acid sequences. We describe the successful use of this element in interplasmid transposition assays in a Drosophila cell line and embryos as well as its use as a genetic transformation vector in this species. We show that *Herves* is present in one east African population of *An. gambiae sensu stricto*. Initial observations of copy number and site-occupancy data from this population are consistent with this element being transpositionally active. *Herves* is perhaps an illustrative example of transposable element invasion and spreading through *An. gambiae*, which could serve as an example of how class II transposable elements will behave in this species when they are used as part of a genetic drive strategy aimed at introducing, and then spreading, beneficial genes through mosquito populations that vector human disease.

MATERIALS AND METHODS

Computer searches for active *hAT* elements in *An. gambiae*: All 8987 scaffolds of the *An. gambiae* genome project (build 2, version 1; available at ftp.ncbi.nih.gov/genbank/genomes/Anopheles_gambiae/) were searched for nucleotide stretches with characteristics of recently active *hAT* elements (characteristics described below). All searches were performed using custom-written programming scripts in PERL v.5.8.0 (scripts available as online supplemental material at http://www.entomology.ucr.edu/people/atkinson.html). Nucleotide stretches were identified as those of potentially active *hAT* elements and saved for further analysis if they had the following characteristics: (1) they had to be <10,000 bp in length; (2) they had to have an 8-bp duplication at the beginning and end of the stretch (duplications >8 bp were not saved to avoid duplicated regions of the genome); (3) inverted terminal repeat sequences (ITRs) had to be adjacent to the 8-bp duplications and be at least 11 bp long with up to three mismatches; and (4) they had to have an ORF ≥ 200 aa. This list of potential elements was sorted further by isolating stretches in which the same ITR was found adjacent to at least two different 8-bp duplications. From the resulting list of nucleotide stretches those containing ORFs ≥200 aa were identified and the ORF compared to the "nr" peptide database in GenBank (http://www.ncbi.nih.gov) using the program BLASTP with expected value $E \leq 10^{-6}$ (Altschul *et al.* 1990) for similarities to known proteins. Finally, annotations from the BLASTP output were screened manually for references to known transposable elements.

Isolation of the *Herves* element from whole-genome extractions: Intact copies of *Herves* were amplified using PCR primer pairs designed for sequences surrounding three apparently intact copies of the element identified by the computer searches (see results). We attempted to isolate *Herves* from the RSP-1 strain because it is closely related to the PEST strain from which the *An. gambiae* genome sequence was obtained and because the PEST strain is no longer available. Genomic DNA from a dozen pupae from the RSP strain was extracted using the Wizard Genomic DNA purification kit (Promega, Madison, WI). Primer pairs were 5′-TAA GTG TGG TGA CTC GGG AT-3′ (lgam8859) and 5′-CCC TGG ACA GTT GTA GTT GA-3′ (rgam8859), 5′-TGA GGG ACG CTA ATT GAT TC-3′ (lgam8964) and 5′-GCT TTG CGT TCC TTG TAT CT-3′ (rgam8964), and 5′-TCC TAC ACA GGG CTG CTT CTC-3′ (lgam8978) and 5′-CAG CGA GAG GAC TAA TTT GG-3′ (rgam8978). A first round of PCR was performed using Vent_r DNA polymerase (New England Biolabs, Beverly, MA) under the following conditions: 94° for 3 min and 30 cycles of 94° for 30 sec, 65° for 30 sec, and 72° for 5 min. A second round of PCR, using the products of the first reaction as a DNA template, was performed to amplify internal sections of the element.

Transposition assays and Drosophila transformations: The amplified *Herves* element was cloned into plasmids for use in an interplasmid transposition assay similar to the one described in Sarkar *et al.* (1997). Donor plasmids were created by cloning the entire *Herves* sequence into a pBluescriptSK+ vector (Stratagene, La Jolla, CA) and replacing the ORF with a gene conferring resistance to kanamycin, the ColE1 origin of replication, and a *lacZ* reporter gene (all three were obtained from the pKSacO∝ plasmid; Sarkar *et al.* 1997). The *Herves* ORF isolated from the RSP strain was cloned into a pK19 vector containing a heat-shock protein 70 promoter (pKhsp70) to create the helper plasmid pKhsp70Herves-RSP. The ORF from RSP differed from that found in the PEST genome in that the amino acid at position 505 was P instead of L. This variable amino acid was mutated so that the ORF from RSP was identical to that found in the PEST strain. This altered ORF was cloned into pKhsp70 to create pKhsp70-Herves-PEST. The target plasmid was pGDV1.

Helper, donor, and target plasmids were introduced into *D. melanogaster* Schneider2 (S2) cells (Schneider 1972) using Cellfectin (Invitrogen, San Diego) following the manufacturer's suggested procedure. In each experiment 5 million cells were transfected with 2.5 μg each of donor and helper plasmids and 5.0 μg of target plasmid. Twenty-four hours after transfection cells were heat-shocked at 41° for 2 hr and allowed to recover at 23° for an additional 24 hr. Cells were then collected and DNA was extracted using the Wizard Genomic purification kit (Promega).

Helper, donor, and target plasmids were introduced into *D. melanogaster* Canton-S,*w* embryos by direct micro-injection. Injection plasmid mixtures were at 0.5, 0.5, and 1.0 mg/ml of donor, helper, and target plasmids, respectively. Drosophila preblastoderm embryos were injected between 30 and 60 min postoviposition. Eggs were allowed to develop at 26° for 16 hr at which time they were heat-shocked for 1 hr at 37°. After heat shock the embryos were allowed to recover at 26° at which time plasmids were recovered. Plasmid recovery was as described (Sarkar *et al.* 1997).

Plasmids recovered from both transfected cell lines and injected embryos were electroporated into competent *Escherichia coli* cells (DH10β, Invitrogen) and plated on Luria-Bertani (LB) plates containing kanamycin and chloramphenicol (25 and 10 μg/ml, respectively). Colonies from these plates were grown overnight in LB media containing kanamycin (25 μg/ml) and DNA was extracted using the Wizard Plus miniprep kit (Promega). DNA was digested with *Pst*I restriction enzyme (New England Biolabs) to check for transposition. Because the *Herves* donor element and the target plasmid contain *Pst*I restriction sites, recombinant plasmids resulting

from the transposition of *Herves* into pGDV1 produce a characteristic pattern of three fragments following digestion with this enzyme. A 622-bp fragment arising from two *Pst*I sites inside the *lacZ* gene in the *Herves* donor element is common to all recombinant plasmids. Two fragments totaling 6449 bp but varying in size as a function of the position of the integration site within the target plasmid permit the location of the integrated element to be determined. Transposition events were confirmed by DNA sequencing at the University of California Riverside Institute for Integrative Genome Biology.

Drosophila transformations were performed essentially as described (Rubin and Spradling 1982). A 1.3-kb *Bst*XI-*Bgl*II fragment containing the enhanced green fluorescence (EFGP) gene placed under the control of the 3xP3 promoter from *D. melanogaster* was removed from plasmid pBac[3xP3-EGFP,afm] and the 5′ and 3′ overhangs were removed with the large fragment of Klenow DNA polymerase and cloned into the blunt-ended *Pst*I site, also generated by the Klenow holoenzyme, in plasmid pBS*Herves*, creating plasmid p*Herves*[3xP3-EGFP], which contained 1.4 kb of *Herves* L end, 302 bp of *Herves* R-end sequence, and the 8-bp target-site duplication 5′-GTAGCAAC-3′ from *An. gambiae*. Plasmid pHerves [3xP3-EGFP] (250 mg/ml) was coinjected with the *Herves* helper plasmid pKhsp70*Herves* (300 mg/ml) into preblastoderm Drosophila Canton-S,*w* embryos. Surviving G_0 adults were backcrossed to Canton-S,*w* and G_1 progeny were examined for the expression of EGFP genetic marker. Homozygous lines were established by repeated backcrossing of transgenic individuals.

Phylogenetic analysis: Transposase protein sequences from 11 known *hAT* elements and *Herves* were aligned using CLUSTALW v.1.83 and the PAM250 matrix (Thompson *et al.* 1994). Phylogenetic analyses aimed at establishing the placement of *Herves* on a tree of *hAT* element transposases were performed using programs based on maximum-likelihood optimality criteria [MRBAYES v.3.0b4 (Huelsenbeck and Ronquist 2001) and TREE-PUZZLE v.5.0 (Schmidt *et al.* 2002)] and on maximum parsimony (PAUP*4.0b10; Swofford 1998). Amino acid sequences of *hAT* element transposases were selected from prior publications to represent the diversity of the *hAT* superfamily with particular attention to sequences likely to be closely related to *Herves* (Calvi *et al.* 1991; Rubin *et al.* 2001; Robertson 2002; Handler 2003). Amino acid substitution rates were modeled using the JTT (Jones *et al.* 1992) and WAG (Whelan and Goldman 2001) substitution matrixes. Among-site rate variation was modeled by a discrete gamma parameter estimated from the data by MRBAYES and TREE-PUZZLE.

Mosquito samples: Mosquito strains examined in this study were the G3 strain maintained at the University of Maryland Biotechnology Institute; Suakoko (SUA) from A. Crisanti, Imperial College, London, and A. Richman, Department of Veterans Affairs, Washington, DC.; RSP-2 from F. H. Collins, University of Notre Dame, Notre Dame, Indiana; KIL from P. J. Eggleston, Keele University, Keele, UK, and C. Curtis, London School of Hygiene and Tropical Medicine, London; and isofemale lines from recently caught field samples of *An. gambiae* from Mali from G. C. Lanzaro, University of California, Davis, California. Insects collected from Furvela, Mozambique, were provided by D. Charlwood, Danish Bilharziasis Institute, Charlottenlund, Denmark.

***Herves* transposon display analysis (TEDA):** The procedure used for transposon display has previously been described (Guimond *et al.* 2003) and was modified for use with the *Herves* element. Genomic DNA was isolated as described and processed for transposable element display after digesting with *Mse*I (Guimond *et al.* 2003). For these experiments, the *Herves*-specific primers HervTEDAL1 5′-AAT TCG ACG GGT TCC TAC C-3′ (preselective PCR) and HervTEDAL2 Cy5/5′-GTT GAT TAG ATG AAC GTA GG-3′ (selective PCR) were used in addition to the adapter-specific primer *Mse*Ia GAC GAT GAG TCC TGA G previously described (Guimond *et al.* 2003). Preselective PCR conditions were 95° for 3 min; 25 cycles of 95° for 15 sec, 60° for 30 sec, and 72° for 1 min; and 72° for 5 min. Selective PCR conditions were 95° for 3 min; 5 cycles of 95° for 15 sec; 64°–60° at 1° per cycle for 30 sec; 72° for 1 min; 25 cycles of 95° for 15 sec; 60° for 30 sec; 72° for 1 min; and 72° for 5 min. Selective PCR products were separated on a 6% denaturing polyacrylamide gel. The gel was dried onto 3MM paper and scanned with a Typhoon phosphorimager (Amersham, Buckinghamshire, UK). Individual bands were excised from the dried gels, eluted, reamplified, and sequenced as previously described (Guimond *et al.* 2003).

RESULTS

***hAT* elements in the PEST genome:** The search for nucleotide stretches with characteristics of active *hAT* elements yielded 3553 nucleotide stretches that satisfied the structural criteria. Of these, 1536 also contained ORFs ≥200 aa with significant similarities ($E \leq 10^{-6}$) to a number of class I and II elements, some of which have previously been described, including *P* elements (Sarkar *et al.* 2003a) and *Tc*1-*mariner* elements (Holt *et al.* 2002). Three different ORF sequences had significant similarities to known *hAT* elements: *hopper*, *Hermes*, *hobo*, *Ac*, and *Tam3*. The first ORF was 603 aa long and was present three times within the genome (Figure 1). The second ORF was almost identical to the first and differed from it only by the insertion of a 328-bp section of a *Topi* transposable element (Grossman *et al.* 1999) at aa 231, causing a frameshift mutation. The third ORF was 297 aa long and was present only once within the genome sequence.

The presence of three identical ORFs within stretches identified as possible active *hAT* elements and a fourth almost identical sequence was encouraging and these sequences were investigated further. The three 603-aa ORFs were almost identical and were located between identical ITR sequences forming three putative transposable elements of 3699, 3702, and 3707 bp. These elements compose a new subfamily of transposable elements referred to as *Herves*. One copy of *Herves* was located on chromosome 2 and two copies on chromosome 3. The fourth copy, containing the *Topi* insertion, was located in an unmapped area of the genome. The *Herves* element was characterized by 11-bp ITRs containing two mismatched bases, 8-bp target-site duplications showing sequence similarity to the *hAT* element consensus sequence 5′-GTNNNNAC-3′, and a transposase with amino acid sequence similarity to *hAT* element transposases. Sequences identical to *Herves* have been reported by Jurka (2000).

Phylogenetic analyses: All phylogenetic analyses using either maximum-likelihood- or maximum-parsimony-based methods resulted in topologies consistent with the tree shown in Figure 2. Maximum-likelihood methods, using more sophisticated models of evolution, had gen-

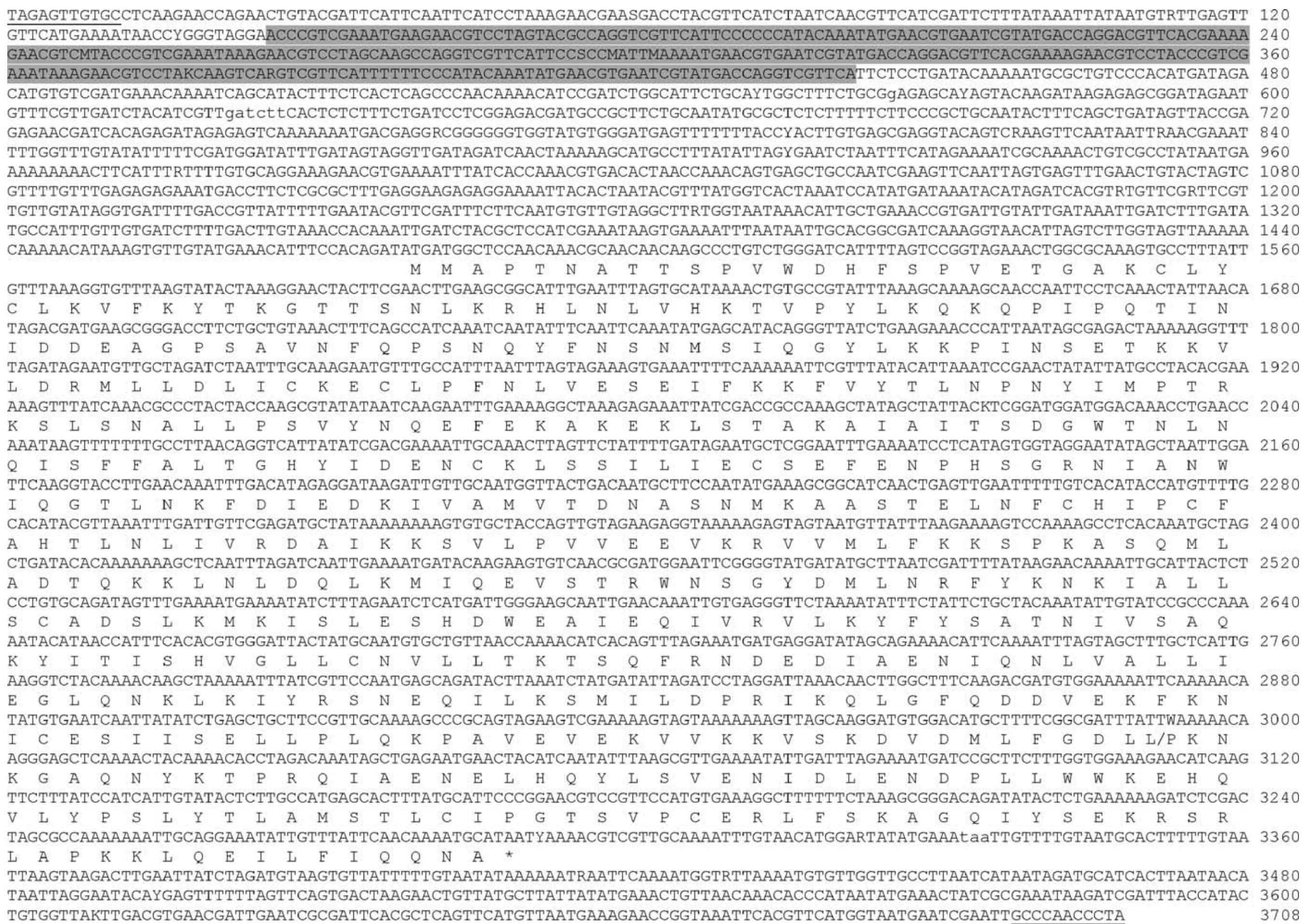

FIGURE 1.—Consensus sequence of three *Herves* elements from the *An. gambiae* genome (see text) and amino acid translation of the transposase. IUPAC ambiguity codes indicate sites where the three copies differ; lowercase letters indicate a deletion in at least one copy. Inverted terminal repeats are underlined. A region containing three almost-perfect tandem repeats is shaded.

erally higher nodal support for deep branches than did maximum parsimony. Among-site rate variation was low (the lowest estimated α-shape parameter was 2.64).

The *Herves* transposase sequence was most closely related to the *hopper-we* sequence (HANDLER 2003) and did not form a monophyletic group with the other insect elements *hopper*, *Hermes*, *Homer*, and *Hermit*. Corrected distance estimates indicated that *Herves* was as distantly related to these insect elements as to the human *Tramp* element (ESPOSITO *et al.* 1999; Table 1). The close relationship of *Herves* and *hopper-we* was confirmed by high levels of transposase sequence similarity (28% amino acid identity, 53% amino acid similarity), overall nucleotide sequence similarity (40% nucleotide identity when aligned using ClustalW), and ITR sequence similarity.

***Herves* is a functional element:** Three almost identical copies of *Herves* were identified in the published scaffolds of the *An. gambiae* genome-sequencing project and a fourth copy was also identified that differed by the insertion of 238 bp of the *Topi* transposable element inside its ORF. Circumstantial evidence of recent movement by *Herves* was discovered in the course of isolating *Herves* from the *An. gambiae* RSP strain. Three chromosomal sites within the genome of RSP that contained full-length *Herves* sequences in the PEST strain were amplified using PCR and sequenced. The genome of RSP differed from PEST at these three sites (Figure 3). The first site did not contain *Herves* but instead the observed sequence was consistent with what would be expected prior to transposition into that site, *i.e.*, a single copy of the 8-bp host duplication. The second site also did not contain the *Herves* element, but instead the *Herves* sequence was replaced with a single A/T base pair. This was consistent with the predicted excision footprint of the *Hermes hAT* element following a recently proposed mechanism of *Hermes* excision and transposition (ZHOU *et al.* 2004). The third site appeared to be heterozygous and PCR amplifications yielded two types of fragments: (1) a fragment with the expected nucleotide sequence prior to *Herves* insertion and (2) a fragment containing the full-length *Herves* element. These results suggest that (1) either some *Herves* sequences

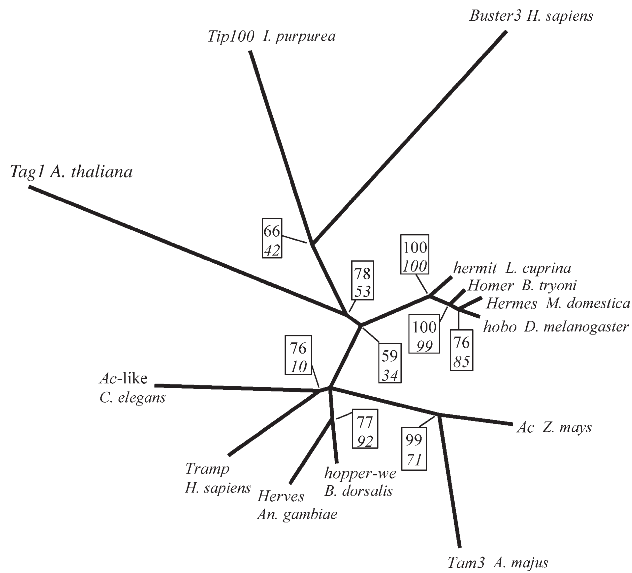

Figure 2.—Phylogenetic relationships of transposase amino acid sequences from selected *hAT* elements (see text). Boxed numbers next to each node represent measures of nodal support. The top number is the quartet puzzling reliability percentage from TREE-PUZZLE using the WAG matrix substitution model and among-site rate variation estimated using a discrete gamma shape parameter. The bottom number in italics is the percentage of bootstrap support from unweighted maximum-parsimony analysis using PAUP*. Tree topology and branch lengths correspond to the tree obtained with TREE-PUZZLE from 100 maximum-likelihood quartets. All transposase sequences except that of *Herves* were obtained from GenBank (http://www.ncbi.nih.gov/): *Tam3* (accession no. CAA38906), *Ac* (accession no. CAA29005), *hobo* (accession no. A39652), *Hermes* (accession no. AAC37217), *Homer* (accession no. AAD03082), *hermit* (accession no. AAA64851), *Buster3* (accession no. NP_071373), *Tip100* (accession no. BAA36-225), *Tag1* (accession no. T52187), *Ac*-like (accession no. NP 509255), *hopper* (accession no. AAL93203), and *Tramp* (accession no. CAA76545).

have mobilized since the PEST and RSP strains have become isolated or, alternatively, there has been differential assortment of an ancestral polymorphism of *Herves* insertions prior to the creation of these two strains; (2) *Herves* is capable of persisting in a single genome location for at least several generations; and (3) the RSP strain is not homozygous with respect to *Herves* integrations. It is unclear if this heterogeneity stems from the mating of individuals with and without the insertion, from movement of the element, or from both.

To test if the *Herves* element present in RSP was functional, interplasmid transposition assays were performed using two *Herves* ORF sequences as sources of transposases: first, *Herves*-PEST, with the amino acid sequence shown (Figure 1) and second, *Herves*-RSP, which has the same amino acid sequence but with L505 replaced with phenylalanine (see above). Assays performed in Drosophila S2 cells and Drosophila embryos demonstrated the transpositional activity of *Herves* (Table 2). Twenty-seven perfect transpositions and 13 imperfect transpositions of *Herves* were recorded (Table 2). All were dependent on the presence of *Herves* transposase since transposition events were not recovered in the absence of transposase. Perfect transpositions involved transposition of only the *Herves* element with the creation of an 8-bp target-site duplication at the point of insertion. The target-site selection of *Herves*, as reflected in the distribution of integrations in the pGDV1 target plasmid, differed from that of both the *Hermes* and *hobo* elements. From transposition assays performed in Drosophila Canton-S,*w* embryos, both *Hermes* and *hobo* show a very similar pattern of site selection within this plasmid, with insertion hotspots at nucleotides 736, 2154, 2271, and 2303 (Sarkar *et al.* 1997; Y.-J. Kim, D. A. O'Brochta and P. W. Atkinson, unpublished results). This was not the case for *Herves*. None of the 21 transpositions recovered from Drosophila embryos inserted at hotspots 736, 2154, and 2271 and only one insertion was seen at 2203. A new hotspot (8/21 insertions) was observed at nucleotide 476, at which no previous insertions of either *hobo* or *Hermes* have been recorded (Sarkar *et al.* 1997; Y.-J. Kim, D. A. O'Brochta and P. W. Atkinson, unpublished results).

To test whether *Herves* could be used to genetically transform insects, the plasmid p*Herves*[3xP3-EGFP] was co-injected with helper *Herves* plasmid into preblastoderm embryos of *D. melanogaster*. From two injection experiments, four independent transgenic lines were obtained with an overall transformation frequency of 30.1% (Table 3). This frequency of transformation is comparable to transformation frequencies obtained when other class II elements such as *P*, *hobo*, *Hermes*, and *piggyBac* are used to transform this species and is higher than predicted from the low frequency of transposition of *Herves* observed in Drosophila (Table 2). Three of these transgenic lines were selected for further analysis. Cut-and-paste integration of solitary *Herves* elements was confirmed by TEDA in the three lines examined (Figure 4). For each line, a single but different fragment was amplified. Purification and sequencing of each of these revealed that, for lines 2 and 3, at both ends, the *Herves* element insertion was delimited by the first nucleotide of *Herves* with flanking sequence from *D. melanogaster*. An 8-bp sequence consistent with the consensus insertion sites of *hAT* elements was present adjacent to each *Herves* element (Figure 4). For line 1, only the left-end integration site has been characterized and the insertion is delimited by the terminal nucleotide of *Herves*. The 8-bp site flanking *Herves* conforms to the target-site consensus of *hAT* elements. A BLAST search of the flanking sequences indicated that in line 1 *Herves* had inserted into chromosome 3L at 66A, in line 2 *Herves* had inserted into chromosome 3R at 95E, and in line 3 *Herves* had inserted into a sequence found in five locations throughout the genome.

Distribution of *Herves*: Seven laboratory lines of *An. gambiae* were tested and all contained *Herves* (Figure

TABLE 1

Measurements of amino acid similarity and distance of 12 *hAT* superfamily elements

Transposable element	*hAT* superfamily element											
	Herves	*hopper-we*	*Tramp*	*Ac-like*	*Activator*	*Hermes*	*hobo*	*Homer*	*Hermit*	*Buster3*	*Tip100*	*Tag1*
Herves	—	1.90	3.09	4.23	4.31	5.29	4.64	5.02	4.99	8.15	7.93	8.54
hopper-we	0.30 (0.49)	—	5.57	3.52	4.65	4.75	4.73	4.90	4.67	7.12	8.05	9.00
Tramp	0.25 (0.42)	0.25 (0.43)	—	4.32	6.16	5.65	4.99	5.44	5.57	6.99	9.00	9.00
Ac-like	0.21 (0.39)	0.22 (0.44)	0.21 (0.37)	—	5.27	5.68	5.59	6.16	6.15	8.47	9.00	9.00
Activator	0.24 (0.43)	0.23 (0.4)	0.20 (0.38)	0.21 (0.37)	—	6.24	7.16	6.75	6.70	8.88	9.00	9.00
Hermes	0.21 (0.37)	0.22 (0.43)	0.20 (0.38)	0.20 (0.37)	0.21 (0.41)	—	0.75	0.85	1.26	7.11	8.19	8.43
hobo	0.22 (0.38)	0.25 (0.43)	0.21 (0.4)	0.20 (0.4)	0.21 (0.4)	0.53 (0.70)	—	0.80	1.20	7.25	9.00	7.88
Homer	0.21 (0.37)	0.20 (0.40)	0.20 (0.36)	0.19 (0.35)	0.19 (0.37)	0.51 (0.69)	0.52 (0.69)	—	1.09	6.66	8.36	8.12
hermit	0.22 (0.38)	0.23 (0.44)	0.22 (0.36)	0.20 (0.38)	0.23 (0.39)	0.38 (0.58)	0.40 (0.59)	0.41 (0.59)	—	7.24	8.63	7.17
Buster3	NA	NA	NA	NA	NA	NA	NA	NA	NA	—	7.40	9.00
Tip100	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	—	9.00
Tag1	0.20 (0.37)	0.19 (0.38)	NA	NA	NA	0.22 (0.37)	0.23 (0.38)	0.33 (0.61)	0.36 (0.58)	NA	NA	—

Numbers above the diagonal are maximum-likelihood distances calculated by TREE-PUZZLE using the WAG amino acid substitution matrix and the among-site variation estimated by a discrete gamma parameter ($\alpha = 2.64$). Sequences were aligned using ClustalW. Numbers below the diagonal are the proportion of amino acid identity and, in parentheses, similarity (using the Blosum62 matrix). Amino acid identity and similarity were estimated from pairwise alignments generated by the BLASTP program ($E \leq 0.01$). NA indicates that no significant similarity between sequences could be identified by BLASTP.

```
                              SCAFFOLD AAAB01008959
PEST  ATCCATTTCGCGCCTTCGTACtagagttgtgcctcaagaacc(...)gaatcgaattgcccaacccctaCTTCGTACACCTTCAACTAC
RSP   ATCCATTTCGCGCCTTCGTAC-------------------------------------------------------ACCTTCAACTAC

                              SCAFFOLD AAAB01008964
PEST  CCACCCGGGGATGGTCCCCATtagagttgtgcctcaagaacc(...)gaatcgaattgcccaacccctaGTCCCCATCCTCCCCGGTTG
RSP   CCACCCGGGGATGGTCCCCA------------------------A-------------------------------CCTCCCCCGGGG

                              SCAFFOLD AAAB01008978
PEST  CTGCCTCAATGGGGTAGCAACtagagttgtgcctcaagaacc(...)gaatcgaattgcccaacccctaGTAGCAACTACATCATTACC
RSP   CTGCCTCAATGGGGTAGCAACtagagttgtgcctcaagaacc(...)gaatcgaattgcccaacccctaGTAGCAACTACATCATTACC
RSP   CTGCCTCAATGGGGTAGCAAC-------------------------------------------------------TACATCATTACC
```

Figure 3.—Evidence of past mobility of *Herves* in *An. gambiae* PEST and RSP strains. Sequences are grouped by locus under the heading of the scaffold from the *An. gambiae* sequencing project. PEST sequences were obtained from GenBank while RSP sequences were from PCR amplifications; the two RSP lines for scaffold AAAB01008978 represent different PCR amplification products. *Herves* ITRs are underlined, omitted *Herves* nucleotides are represented by (. . .), and dashes indicate insertions/deletions.

5). Likewise, separate colonies of the laboratory line Suakoko also showed notable differences in the number and position of elements. The G3 and KIL (no. 2) lines showed intraline variation in element position and copy number. Twenty-five bands from Figure 5 were excised from the gel, reamplified, and sequenced. Ten of the elements were inserted in unique, single-copy DNA; 9 were inserted in sequences repeatedly found throughout the genome, precluding localization; and 6 were integrated in sequences that were not present in the *An. gambiae* sequence database. Seven of the localized elements were on the second chromosome, two were on the X chromosome, and one was inserted in a sequence that had not yet been placed within the existing scaffold structure of the *An. gambiae* genome database.

Preliminary analysis of *An. gambiae sensu lato* collected from a village (Furvela) located along the southern coast of Mozambique revealed the presence of *Herves* not only in *An. gambiae s.s.* but also in *Anopheles arabiensis* and *Anopheles merus* (Figure 6). All insects analyzed contained at least one element ($n = 35$, *An. gambiae s.s.*; $n = 45$, *An. arabiensis*; $n = 6$, *An. merus*; Figure 7). Bands arising from transposable element display are dominant markers and copy numbers of elements were estimated to be $4x/3$ where $x =$ number of occupied sites ("bands") and Hardy-Weinberg equilibrium was assumed. In the three species examined, copy numbers were 4.5 elements/genome in *An. merus* (range 1–7), 4.6 elements/genome in *An. gambiae s.s.* (range 1–8), and 7.3 elements/genome in *An. arabiensis* (range 1–11). Elements were observed in a variety of different positions (sites) within the genome. Twenty-five sites were detected in *An. gambiae s.s.*, 31 sites in *An. arabiensis*, and 14 sites in *An. merus*.

DISCUSSION

We have identified a functional class II element of the *hAT* element superfamily, called *Herves*, that displays variation in both copy number and site distribution in both laboratory and field populations of *An. gambiae*, consistent with it being active, or recently active, in these populations.

Originally, Holt *et al.* (2002) identified 15 *hAT* sequences in the genome of *An. gambiae*. Of these 15 sequences, 4 matched the three active and the one presumably inactive copy of *Herves* (with a *Topi* element insertion) described here. The remaining 11 *hAT* sequences had high amino acid identity and similarity to the *Herves* ORF (identity ranged from 24 to 51%, similarity ranged from 41 to 69%, and for sequences aligned using BLASTP, $E \leq 0.01$). This suggests that these sequences were closely related to *Herves* and may be from inactive forms of this element.

Several features of the *Herves* transposable element placed it in the *hAT* superfamily. First, *Herves* contained several amino acid motifs shared among *hAT* elements. For example, *Herves* contains the well-conserved Ww*xxx*-*xxxx*P*x*L*xxx*A*xxx*L motif described by Calvi *et al.* (1991)

TABLE 2

Herves* transposition in cell lines and embryos of *D. melanogaster

Cell line name/ strain	Transposase source	No. of experiments	No. of plasmids screened	No. of cut-and-paste transpositions	No. of transposase-mediated events[a]	Events per 10^4 plasmids screened[b]
Cell line/S2	hsp70*Herves*-PEST	3	551,400	5	7	0.09
Cell line/S2	hsp70*Herves*-RSP	2	306,200	1	4	0.03
Cell line/S2	0	3	887,800	0	0	0
Embryos/Canton-S,*w*	hsp70*Herves*-PEST	4	649,000	21	2	0.32
Embryos/Canton-S,*w*	0	2	615,000	0	0	0

[a] "Transposase-mediated events" were defined as the insertion of at least one ITR into new DNA with the breakpoint immediately outside the ITR.

[b] Rates based on the number of "cut-and-paste" transpositions only.

TABLE 3

Transformation of *D. melanogaster* with p*Herves*[3xP3-EGFP] and pKhsp70*Herves*

Experiment no.	No. of embryos injected	No. of G_0 adults	No. of fertile matings	No. of G_0 matings producing transgenic adults	No. of transgenic G_1's	Transformation frequency (%)
1	120	38	9	3	15 males, 25 females 7 males, 17 females 2 males, 1 female	33
2	60	7	4	1	3 total	25
Total	180	45	13	4		30.1

with 12 of the 17 amino acids in this region being identical to those found in *Ac*, *hobo*, and *Tam3*. Also, of the six conserved *hAT* protein blocks described by Rubin *et al.* (2001), *Herves*, like *hobo*, contained five (blocks A–E). Second, *Herves* ITRs were consistent with a proposed consensus sequence of *hAT* ITRs (Warren *et al.* 1995). Third, the three active *Herves* copies identified in the *An. gambiae* genome and the seven observed "cut-and-paste" transposition events (Table 2) inserted by forming 8-bp target-site duplications, characteristic of *hAT* elements. These duplications had a NWNNNNAY 85% consensus sequence, similar to consensus sequences proposed for other *hAT* elements (*e.g.*, NTNNNNAC proposed for *hobo* and *Hermes*; O'Brochta *et al.* 1996).

Within the *hAT* superfamily the *Herves* transposase sequence was most closely related to the *hopper-we* element from *Bactrocera dorsalis* and *Bactrocera cucurbitae* (Handler 2003) but was distantly related to transposase sequences from other insect elements (*e.g.*, *Hermes* from *M. domestica*). Beyond their transposase sequences, *Herves* and *hopper-we* also shared similar ITR sequences. This suggests that *Herves* and *hopper-we* might form a new group of insect *hAT* transposable elements that diverged prior to the Brachycera-Nematocera divergence. This hypothesis would be supported by the discovery of closely related elements in other species within these suborders. BAC-end clones from the *Ae. aegypti* sequencing project (http://www.tigr.org/tdb/e2k1/aabe/) were searched using the TBLASTN program ($E \leq 10^{-20}$; Altschul *et al.* 1990) and the *Herves* transposase amino acid sequence. This revealed the presence of a 213-aa sequence in *Ae. aegypti* with high similarity to *Herves* and *hopper-we* (distance to *Herves*: 0.96; to *hopper-we*, 1.4; to all other ORFs, 2.1–2.9; distance calculated using the program PROTDIST; Felsenstein 1993). While more research will be necessary to confirm that this sequence is indeed from a transposable element, these results suggest the possibility of a new, unexplored group of insect transposable elements within the *hAT* superfamily.

Herves can transpose in Drosophila cell lines and in Drosophila embryos and can genetically transform Drosophila. Transposition rates are highest in Drosophila embryos but are ~30-fold less than those seen with the related *Hermes* element in similar assays performed in the Canton-S,*w* strain (Sarkar *et al.* 1997). Nevertheless, *Herves* transforms Drosophila at frequencies of the same order of magnitude as the *P*, *hobo*, *Hermes*, and *piggyBac* elements, indicating that it will be an efficient gene vector in at least this species. As such, *Herves* is the first active class II element isolated from any mosquito species and is the first to be used as a gene vector in an insect species. The *An. gambiae* genome project has facilitated the identification of potentially active *P* and *piggyBac* elements (Sarkar *et al.* 2003a,b), while there has been, previous to the publication of this genome sequence, circumstantial evidence for activity of the class II element *Ikirara* (Romans *et al.* 1998) and the class I elements *Moose* (Biessmann *et al.* 1999) and *mtanga* (Rohr *et al.* 2002). Donor *Herves* plasmids used in both transposition assays and fly transformations contain identical amounts of *Herves* left (1.4 kb) and right (302 bp) sequences, indicating that these sequences are sufficient for transposition of the element. Why *Herves* should perform relatively poorly in interplasmid transposition assays conducted in embryos and cell culture, but not in transformation of Drosophila, is unknown. The value of transposition assays may be primarily qualitative rather than quantitative; alternatively, *Herves* may be more efficient at transposition in germ-line nuclei than in somatic nuclei, the latter being measured in transposition assays performed in both cell cultures and developing embryos. The DNA target is also different between the transposition assays and genetic transformations, being a plasmid in the former and chromatin in the later. This difference might lead to a difference in transposition frequency but it is difficult, at this stage, to identify the precise structural or functional basis of such a difference between the *Herves* and *Hermes* elements.

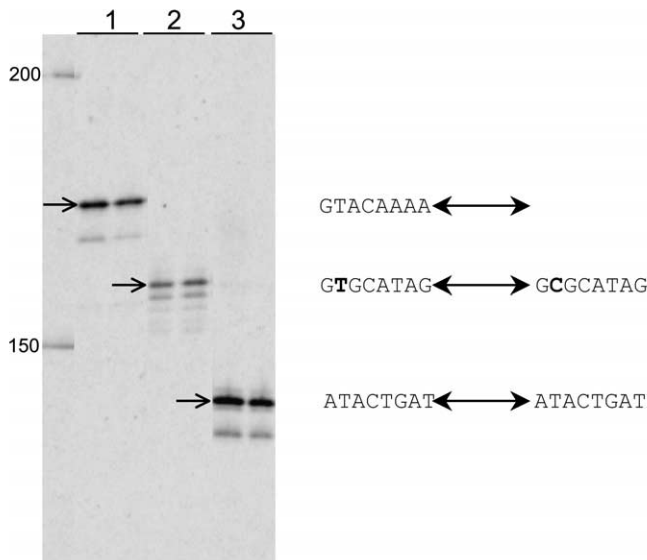

FIGURE 4.—Transposable element display of the left end of *Herves* in transgenic *D. melanogaster* lines 1, 2, and 3. The results from two individuals from each line are shown along with the position of the molecular weight makers. Bands were excised, eluted, reamplified, and sequenced. Chromosomal position was determined by BLAST searching the *D. melanogaster* sequence database. Bands below the main bands (arrows) are identical in sequence to the main bands and are often associated with abundant PCR products <250 bp in length. The sequences adjacent to the right inverted terminal repeat were determined by direct amplification using primers specific for the *Herves* element and flanking genomic DNA. The 8-bp target-site duplications flanking each of the three independent transpositions are shown to the right and each corresponds to the arrowed band at the corresponding migration distance on the gel. For line 1, only sequence flanking the left end of *Herves* was obtained. For line 2, the 8-bp target-site duplication was imperfect with the single-base-pair difference (T *vs.* G) shown in boldface type.

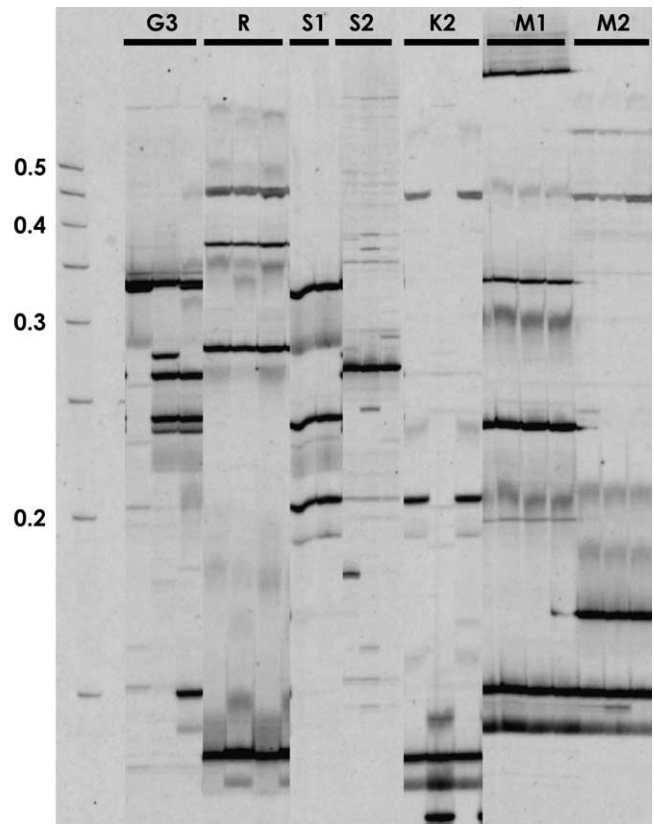

FIGURE 5.—Comparison of the *Herves* content of laboratory lines of *An. gambiae s.s.* Transposable element display was performed on individuals from the laboratory lines G3, RSP (R), Suakoko obtained from two sources (S1 and S2), KIL obtained from two sources (K1 and K2), and two recently established isofemale lines from individuals collected from Mali (M1 and M2). Molecular weight markers are indicated.

The difference in target-site specificity between *Herves* and both *hobo* and *Hermes* is intriguing. Within the constraints of generating an 8-bp target-site duplication conforming to the consensus sequence of *hAT* element insertions, our data show that *Herves* does not favor insertion at the major *hobo/Hermes* hotspots in pGDV1, but rather prefers its own unique hotspot for integration. The existence of at least three functional insect *hAT* elements, *hobo*, *Hermes*, and *Herves*, provides a unique opportunity to examine the role that element sequences within a transposable element family play in determining insertion site specificity. Our data show that *Herves*, while clearly being a *hAT* element, is both structurally and functionally distinct from the *hobo* and *Hermes* elements. We have preliminary evidence that *Herves* may be active in *An. gambiae*. This, combined with its mobility in *D. melanogaster*, opens up the possibility that *Herves* might be used to improve the efficiency of mosquito transformation, particularly of *An. gambiae* and perhaps of other species as well. *Herves* transposition into the anopheline chromosomes might occur at high frequencies if this element has already adapted to being able to move in this mosquito genome. A converse argument is that mechanisms to suppress *Herves* transposition may have already evolved in *An. gambiae* but, as shown above and discussed below, we have preliminary evidence that suggests that *Herves* is currently mobile in this species. We have shown that *Herves* is functionally different from *Hermes* and *hobo*, and it is of interest to determine if these differences will manifest themselves in the behavior of engineered *Herves* elements in transgenic mosquitoes. Genetic transformation of *An. gambiae* remains problematic with only three successful transformations reported since 1987 (MILLER *et al.* 1987; GROSSMAN *et al.* 2001; KIM *et al.* 2004). Furthermore, the difficult husbandry of *An. gambiae* demands that transformation rates be high so as to ensure that transformed genotypes can be propagated from adults surviving the micro-injection procedure used to introduce gene vectors into developing germ-line cells.

Herves is likely to be actively transposing in natural populations of *An. gambiae*, *An. arabiensis*, and *An. merus* on the basis of the abundance of chromosomal sites that are occupied by *Herves* elements at low frequencies

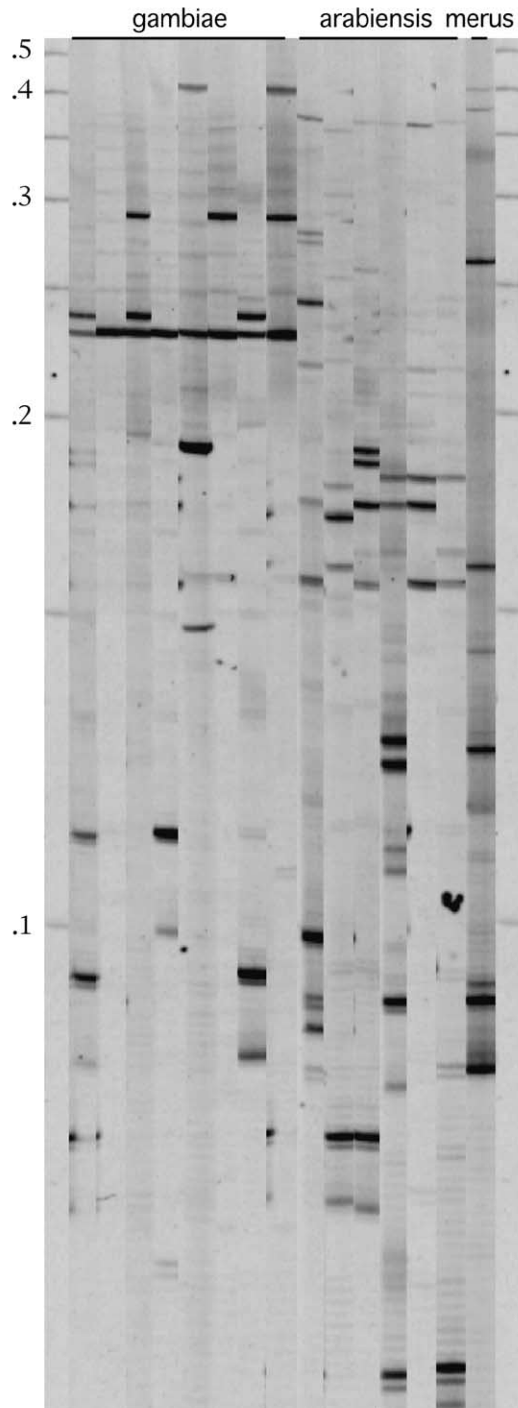

Figure 6.—Transposable element display of individuals collected from a local population in Mozambique. Three species are represented: *An. arabiensis*, *An. gambiae s.s.*, and *An. merus.* Molecular weight markers (in kilobase pairs) are shown.

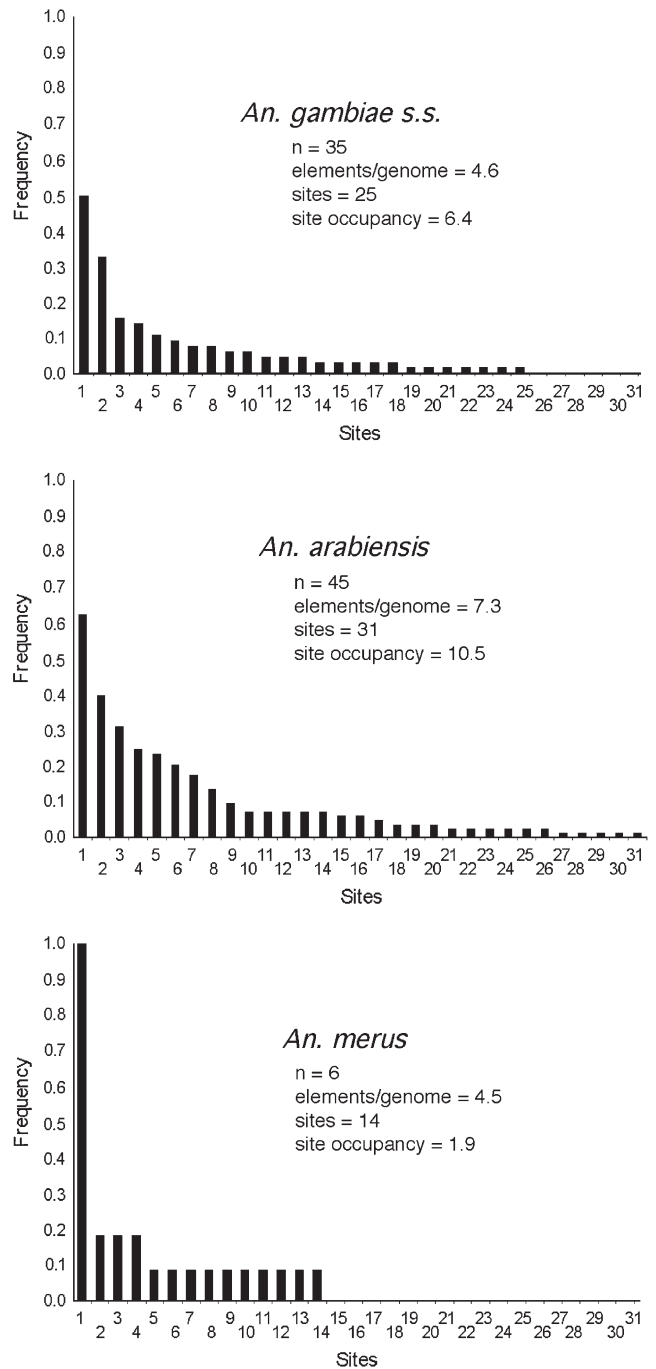

Figure 7.—*Herves* in three species of Anopheles in Mozambique. Data were derived from transposable element displays on the number of individuals indicated (n). "Site" refers to a genomic position occupied by at least one *Herves* element in the sample and corresponds to a unique PCR product on a transposable element display. Sites are arbitrarily numbered and are not similar between species. "Site occupancy" refers to the total number of elements within a sample divided by the number of sites.

within populations. Earlier studies of transposable elements in natural populations of *D. melanogaster* described similar distributions and indicated that element movement was probably responsible for producing the observed frequency spectrum (Kaplan and Brookfield 1983; Langley *et al.* 1983; Montgomery and Langley 1983). Given the generally small effective population sizes for *An. gambiae* and *An. arabiensis* (on the order of 10^3 individuals), the observed frequency spectrums of *Herves* in *An. gambiae s.l.* in Furvela, Mozambique, are consistent with the hypothesis that the element is active

(Taylor *et al.* 1993; Lehmann *et al.* 1998; Taylor and Manoukis 2003). The question that arises is whether *Herves* is at equilibrium in these populations, equilibrium being when the number of *Herves* elements in the genomes of members of a population is constant through a balance between replicative transposition and forces, such as genetic drift and natural selection, that might eliminate these elements. If the element is at equilibrium, then the frequency spectrum reflects the ongoing transposition and deletion activity of *Herves*. If the element is not at equilibrium, then the observed frequency spectrum may reflect a recent invasion of *An. gambiae s.l.* by *Herves*. These alternative hypotheses cannot be resolved using the data presented here. There are two interesting implications of these observations. First, current natural populations of *An. gambiae* should be able to support transpositional activity of *Herves*. Therefore, gene vectors or genetic drive agents constructed from this element will be functional in *An. gambiae*. Second, the active transposition of *Herves* in *An. gambiae* provides an opportunity to examine the dynamics of transposable element movement in a species being considered as a target for genetic control using recombinant DNA technologies. Using transposable elements to carry and spread genes of interest through natural populations of *An. gambiae* will require an understanding of how such elements are likely to behave. *Herves* may provide a critical bridge between laboratory-based gene vector technology development and the natural history of *An. gambiae*.

We thank G. C. Lanzaro (University of California, Davis, CA) and J. D. Charlwood (Danish Bilharziasis Institute) for supplying field material, P. Eggleston (Keele University), and F. H. Collins (University of Notre Dame) for supplying laboratory material used in these studies. We thank J. F. Y. Brookfield (University of Nottingham) for helpful and stimulating discussions and A. Arensburger (University of Maryland) for help with algorithm design. This research was supported by U.S. Public Health Service National Institutes of Health grants GM48102 to D.A.O. and AI45741 to P.W.A.

LITERATURE CITED

Allen, M. L., D. A. O'Brochta, P. W. Atkinson and C. S. LeVesque, 2001 Stable germ-line transformation of *Culex quinquefasciatus* (Diptera: Culicidae). J. Med. Entomol. **38:** 701–710.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990 Basic local alignment search tool. J. Mol. Biol. **215:** 403–410.

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

Biessmann, H., M. F. Walter, D. Le, S. Chuan and J. G. Yao, 1999 *Moose*, a new family of LTR-retrotransposons in the mosquito, *Anopheles gambiae*. Insect Mol. Biol. **8:** 201–212.

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

Catteruccia, F., T. Nolan, T. G. Loukeris, C. Blass, C. Savakis *et al.*, 2000 Stable germline transformation of the malaria mosquito *Anopheles stephensi*. Nature **405:** 959–962.

Coates, C. J., N. Jasinskiene, L. Miyashiro and A. A. James, 1998 *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. Proc. Natl. Acad. Sci. USA **95:** 3748–3751.

Daniels, S. B., A. Chovnick and I. A. Boussy, 1990a Distribution of *hobo* transposable elements in the genus *Drosophila*. Mol. Biol. Evol. **7:** 589–606.

Daniels, S. B., K. R. Peterson, L. D. Strausbaugh, M. G. Kidwell and A. Chovnick, 1990b Evidence for horizontal transmission of the *P* transposable element between Drosophila species. Genetics **124:** 339–355.

Esposito, T., F. Gianfrancesco, A. Ciccodicola, L. Montanini, S. Mumm *et al.*, 1999 A novel pseudoautosomal human gene encodes a putative protein similar to *Ac*-like transposases. Hum. Mol. Genet. **8:** 61–67.

Felsenstein, J., 1993 *PHYLIP version 3.5c*. Department of Genetics, University of Washington, Seattle.

Grossman, G. L., A. J. Cornel, C. S. Rafferty, H. M. Roberston and F. H. Collins, 1999 *Tsessebe*, *Topi* and *Tiang*: three distinct *Tc*-1 like transposable elements in the malaria vector, *Anopheles gambiae*. Genetica **105:** 69–80.

Grossman, G. L., C. S. Rafferty, J. R. Clayton, T. K. Stevens, O. Mukabayire *et al.*, 2001 Germline transformation of the malaria vector, *Anopheles gambiae*, with the *piggyBac* transposable element. Insect Mol. Biol. **10:** 597–604.

Guimond, N. D., D. K. Bideshi, A. C. Pinkerton, P. W. Atkinson and D. A. O'Brochta, 2003 Patterns of *Hermes* element transposition in *Drosophila melanogaster*. Mol. Genet. Genomics **268:** 779–790.

Handler, A. M., 2003 Isolation and analysis of a new *hopper hAT* transposon from the *Bactrocera dorsalis* white eye strain. Genetica **118:** 17–24.

Holt, R. A., G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab *et al.*, 2002 The genome sequence of the malaria mosquito Anopheles gambiae. Science **298:** 129–149.

Huelsenbeck, J. P., and F. Ronquist, 2001 MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:** 754–755.

Ito, J., A. Ghosh, L. A. Moreira, E. A. Wimmer and M. Jacobs-Lorena, 2002 Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. Nature **417:** 387–388.

Jasinskiene, N., C. J. Coates, M. Q. Benedict, A. J. Cornel, C. Salazar-Rafferty *et al.*, 1998 Stable, transposon-mediated transformation of the yellow fever mosquito, *Aedes aegypti*, using the *Hermes* element from the house fly. Proc. Natl. Acad. Sci. USA **95:** 3743–3747.

Jones, D. T., W. R. Taylor and J. M. Thornton, 1992 The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8:** 275–282.

Jurka, J., 2000 Repbase update: a database and an electronic journal of repetitive elements. Trends Genet. **16:** 418–420.

Kaplan, N., and J. F. Y. Brookfield, 1983 Transposable elements in Mendelian populations. III. Statistical results. Genetics **104:** 485–495.

Kim, W., H. Koo, A. M. Richman, D. Seeley, J. Vizioli *et al.*, 2004 Ectopic expression of a cecropin transgene in the human malaria vector mosquito *Anopheles gambiae* (Diptera: Culicidae): effects on susceptibility to Plasmodium. J. Med. Entomol. **41:** 447–455.

Kokoza, V., A. Ahmed, E. A. Wimmer and A. S. Raikhel, 2001 Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the *piggyBac* transposable element vector pBac [3xP3-EGFPafm]. Insect Biochem. Mol. Biol. **31:** 1137–1143.

Langley, C. H., J. F. Y. Brookfield and N. Kaplan, 1983 Transposable elements in Mendelian populations. I. A theory. Genetics **104:** 457–471.

Lehmann, T., W. A. Hawley, H. Grebert and F. H. Collins, 1998 The effective population size of *Anopheles gambiae* in Kenya: implications for population structure. Mol. Biol. Evol. **15:** 264–276.

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

Lozovsky, E. R., D. I. Nurminsky, E. A. Wimmer and D. L. Hartl, 2002 Unexpected stability of *mariner* transgenes in Drosophila. Genetics **160:** 527–535.

Miller, L. H., R. K. Sakai, P. Romans, R. W. Gwadz, P. Kantoff *et al.*, 1987 Stable integration and expression of a bacterial gene in the mosquito, *Anopheles gambiae*. Science **237:** 779–781.

Montgomery, E. A., and C. H. Langley, 1983 Transposable ele-

ments in Mendelian populations. II. Distribution of three *copia*-like elements in a natural population. Genetics **104:** 473–483.

Moreira, L. A., J. Ito, A. Ghosh, M. Devenport, H. Zieler *et al.*, 2002 Bee venom phosopholipase inhibits malaria parasite development in transgenic mosquitoes. J. Biol. Chem. **277:** 40839–40843.

O'Brochta, D. A., W. D. Warren, K. J. Saville and P. W. Atkinson, 1996 *Hermes*, a functional non-drosophilid insect gene vector from *Musca domestica*. Genetics **142:** 907–914.

O'Brochta, D. A., N. Sethuramuran, R. Wilson, R. H. Hice, A. C. Pinkerton *et al.*, 2003 Gene vector and transposable element behavior in mosquitoes. J. Exp. Biol. **203:** 3823–3834.

Perera, O. P., R. A. I. Harrell and A. M. Handler, 2002 Germ-line transformation of the South American malaria vector, *Anopheles albimanus*, with a *piggyBac*/EGFP transposon vector is routine and highly efficient. Insect Mol. Biol. **11:** 291–297.

Robertson, H. M., 2002 Evolution of DNA transposons in eukaryotes, pp. 1093–1110 in *Mobile DNA II*, edited by N. L. Craig, R. Craigie, M. Gellert and A. M. Lambowitz. ASM Press, Washington, DC.

Robertson, H. M., and E. G. MacLeod, 1993 Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. Insect Mol. Biol. **2:** 125–139.

Rohr, C. J. B., H. A. Ranson, X. Wang and N. J. Besansky, 2002 Structure and evolution of *mtange*, a retrotransposon actively expressed on the Y chromosome of the African malaria vector *Anopheles gambiae*. Mol. Biol. Evol. **19:** 149–162.

Romans, P., R. K. Bhattacharyya and A. Colavita, 1998 *Ikirara*, a novel transposon family from the malaria vector mosquito, *Anopheles gambiae*. Insect Mol. Biol. **7:** 1–10.

Rubin, E., G. Lithwick and A. A. Levy, 2001 Structure and evolution of the hAT transposon superfamily. Genetics **158:** 949–957.

Rubin, G. M., and A. C. Spradling, 1982 Genetic transformation of *Drosophila* with transposable element vectors. Science **218:** 348–353.

Sarkar, A., C. J. Coates, S. Whyard, U. Willhoeft, P. W. Atkinson *et al.*, 1997 The *Hermes* element from *Musca domestica* can transpose in four families of cylorrhaphan flies. Genetica **99:** 15–29.

Sarkar, A., R. Sengupta, J. Krzywinski, X. Wang, C. Roth *et al.*, 2003a *P* elements are found in the genomes of nematoceran insects of the genus *Anopheles*. Insect Biochem. Mol. Biol. **33:** 381–387.

Sarkar, A., C. Sim, Y. S. Hong, J. R. Hogan, M. J. Fraser *et al.*, 2003b Molecular evolutionary analysis of the widespread *piggyBac* transposon family and related "domesticated" sequences. Mol. Genet. Genomics **270:** 173–180.

Schmidt, H. A., K. Strimmer, M. Vingron and A. von Haeseler, 2002 TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18:** 502–504.

Schneider, I., 1972 Cell lines derived from late embryonic stages of *Drosophila melanogaster*. J. Embryol. Exp. Morphol. **27:** 363–365.

Swofford, D. L., 1998 *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, MA.

Taylor, C. E., and N. C. Manoukis, 2003 Effective population size in relation to genetic modification of *Anopheles gambiae sensu stricto*, pp. 133–146 in *Ecological Aspects for Application of Genetically Modified Mosquitoes*, edited by W. Takken and T. W. Scott. Kluver Academic, Dordrecht, The Netherlands.

Taylor, C. E., Y. T. Toure, M. Coluzzi and V. Petrarca, 1993 Effective population size and persitance of *Anopheles arabiensis* during the dry season in west Africa. Med. Vet. Entomol. **7:** 351–357.

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.

Warren, W. D., P. W. Atkinson and D. A. O'Brochta, 1995 The Australian bushfly *Musca vetustissima* contains a sequence related to transposons of the *hobo*, *Ac* and *Tam3* family. Gene **154:** 133–134.

Whelan, S., and N. Goldman, 2001 A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. **18:** 691–699.

Wilson, R., J. Orsetti, A. K. Klocko, C. Aluvihare, E. Peckham *et al.*, 2003 Post-integration behavior of a *Mos1 mariner* gene vector in *Aedes aegypti*. Insect Biochem. Mol. Biol. **33:** 853–863.

Zhou, L., R. Mitra, P. W. Atkinson, A. B. Hickman, F. Dyda *et al.*, 2004 Transposition of *hAT* elements links transposable elements and V(D)J recombination. Nature **432:** 995–1001.

Communicating editor: T. Eickbush