Molecular Ecology (2006) doi: 10.1111/j.1365-294X.2006.03150.x

Insertion polymorphism of transposable elements and population structure of *Anopheles gambiae* M and S molecular forms in Cameroon

M. BOULESTEIX,* F. SIMARD,†‡ C. ANTONIO-NKONDJIO,†‡ H. P. AWONO-AMBENE,†§ D. FONTENILLE‡ and C. BIÉMONT*

Laboratoire de Biométrie et Biologie Evolutive, UMR 5558, CNRS, Université Claude Bernard Lyon1, 69622 Villeurbanne cedex, France, †Laboratoire de Recherche sur le Paludisme, Organization de Coordination pour la lutte Contre les Endémies en Afrique Centrale (OCEAC), BP 288, Yaoundé, Cameroon, ‡Unité de Recherche 016, Institut de Recherche pour le Développement (IRD), 911, Avenue Agropolis, BP 64501, 34394 Montpellier cedex 5, France, §Centre de Recherches Médicales, Institut de Recherches Médicales et des Plantes Médicinales (IMPM), BP 6163, Yaoundé, Cameroon

Abstract

The insertion polymorphism of five transposable element (TE) families was studied by Southern blots in several populations of the M and S molecular forms of the mosquito *Anopheles gambiae sensu stricto* from southern Cameroon. We showed that the mean TE insertion site number and the within-population insertion site polymorphism globally differed between the M and S molecular forms. The comparison of the TE insertion profiles of the populations revealed a significant differentiation between these two molecular forms ($0.163 < \Phi_{ST} < 0.371$). We cloned several insertions of a non-LTR retrotransposon (Aara8) that were fixed in one form and absent in the other one. The only insertion that could be clearly located on a chromosome arm mapped to cytological division 6 of chromosome X, confirming the importance of this region in the ongoing speciation between the M and S molecular forms.

Keywords: *Anopheles gambiae*, insertion polymorphism, molecular forms, population genetics, speciation, transposable elements

Received 5 May 2006; revision accepted 1 September 2006

Introduction

Parasites from the genus *Plasmodium* are transmitted to humans by anopheline mosquitoes, causing malaria, which is responsible for the death of more than 1 million people annually (WHO/UNICEF World Malaria Report 2005). In Africa, where 80% of the total mortality associated with this disease occurs, different species of anopheline are able to transmit *Plasmodium* very efficiently, which in part explains why this continent bears the bulk of the malaria burden. Among these vectors, *Anopheles gambiae sensu stricto* plays a major role in the transmission of the disease, because of its ubiquity throughout sub-Saharan Africa (della Torre *et al*. 2005) and its extreme adaptation to human environments (Coluzzi *et al*. 2002).

The wide distribution of *A. gambiae* throughout Africa implies adaptation of the populations to a large range of local climatic conditions. It has been shown that several chromosomal inversions play a role in these local adaptations (Coluzzi *et al*. 2002). Some of these inversions are locally fixed, and the quasi-absence of hybrids in nature has led to the definition of several 'chromosomal forms'. In addition, polymorphism studies of rDNA regions located on the pericentromeric region of the X chromosome have distinguished two major haplotypes that segregate in nature (della Torre *et al*. 2001; Gentile *et al*. 2001). These haplotypes, which have been called the M and S molecular forms, can segregate within some of the chromosomal forms. Again, hybrids between these two forms are very rarely met in nature even in areas of sympatry. The M and S molecular forms are distributed from Senegal to Angola

Correspondence: Biémont Christian, Laboratoire «Biométrie et Biologie Evolutive», UMR 5558, CNRS, Université Claude Bernard Lyon1, 69622 Villeurbanne cedex, France. Fax: (33) 4 72 43 13 88; E-mail: biemont@biomserv.univ-lyon1.fr

but only a few M specimens from East Africa have been reported to date (Masendu *et al.* 2004). Recent evidence suggests that the current gene flow between these forms is considerably reduced (della Torre *et al*. 2002; Wondji *et al.* 2002). For example the *kdr* mutation, which confers resistance to pyrethroid insecticides on mosquitoes, is almost exclusively restricted to the S molecular form (Chandre *et al*. 1999; Awolola *et al*. 2003; Gentile *et al*. 2004; Yawson *et al*. 2004). Moreover, it is possible that the M and S forms exploit slightly different breeding sites or have different vectorial capacities (della Torre *et al*. 2005; Wondji *et al.* 2005). Microsatellite data provide evidence of a significant but low overall genetic differentiation between these forms, with however, high F_{ST} values (e.g. 0.468 in Stump *et al*. 2005a) for microsatellite loci of the pericentromeric region of the X chromosome (Wang *et al*. 2001; Lehmann *et al*. 2003; Stump *et al*. 2005a). This leads to the idea of a mosaic genome with restricted gene flow for the X chromosome while gene flow is possible for the autosomes. These results have been recently confirmed with DNA sequence polymorphism data (Stump *et al*. 2005b) and SINE insertion polymorphism (Barnes *et al*. 2005). Turner *et al*. (2005), using a microarray approach with no a priori assumption about the differentiated genomic regions between M and S, confirm the differentiation in the proximal region of the X chromosome but also suggest that two regions on chromosome 2 differ significantly between M and S populations. All the probes that had multiple exact matches with the *A. gambiae* sequenced genome, thus including the transposable elements (TEs), were excluded from this analysis although these sequences may provide considerable information on the differentiation of the M and S molecular forms.

TEs are mobile genetic element dispersed throughout the genomes of most eukaryotic organisms, including mosquitoes (Tu & Coates 2004; Boulesteix & Biémont 2005). They are classified according to their transposition mechanism, some moving from place to place via a DNA intermediate by a cut and paste mechanism (class II elements or transposons), others replicating themselves by transcription and reverse transcription of their RNA intermediate by a copy and paste mechanism (class I elements or retrotransposons). Retrotransposons are further subdivided into long terminal repeat (LTR) and non-LTR retrotransposons. These subclasses differ in their structure and modality of retrotransposition. TE dynamics is dependent upon host genes that regulate their transpositional activity, and upon the population genetic parameters of the host such as effective population size and population structure (Deceliere *et al*. 2005). In this respect, TEs have proved to be powerful markers for population genetic studies in different organisms (Lepetit *et al*. 2002; Zampicinini *et al*. 2004; Barnes *et al*. 2005). In addition, TEs have been used extensively to introduce genes of interest into insects, and over the past few years several species of anopheline mosquitoes have been genetically modified using TEs as gene vectors (Catteruccia *et al*. 2000; Grossman *et al*. 2001; Perera *et al*. 2002). In particular, *Anopheles stephensi* has been modified in a way that drastically reduced its ability to transmit *Plasmodium berghei* (Ito *et al*. 2002). Refractory genetically modified mosquitoes could thus be introduced in nature and mate with the native parasite-transmitting populations, hence transmitting the transgene that would eventually be fixed. One way to quickly drive the transgene to fixation is to use the ability of TEs to be transmitted, on average, to more than half of the progeny because of their ability to replicate faster than the host genome (Ribeiro & Kidwell 1994; Boëte & Koella 2003). This genetic drive will depend, however, on the dynamics of the TE concerned among the natural populations. It is thus clear that we need a better understanding of the dynamics of TEs among natural populations of *Anopheles gambiae*, particularly between the M and S molecular forms.

In this paper, the insertion polymorphism of five TE families (three LTR retrotransposons, one non-LTR retrotransposon and one transposon) was studied in various populations of the M and S molecular forms of *A. gambiae ss* from southern Cameroon. In this area, all the specimens belong to the same chromosomal form, which makes it possible to dissociate the molecular form dynamics from the chromosomal form dynamics (Wondji *et al*. 2002). We show that the mean insertion number and the within-population insertion polymorphism vary for some TE families between the M and S molecular forms. Used as dominant multilocus markers, each TE family indicates a significant differentiation between these two molecular forms despite low sample sizes. We cloned several insertions of Aara8, a non-LTR retrotransposon, that are fixed in one form and absent from the other one. The only insertion that could be located without ambiguity mapped to the division 6 of chromosome X, confirming the importance of this region in the ongoing speciation between the M and S molecular forms.

Materials and methods

Natural populations

Adult females of *Anopheles gambiae* were collected in South Cameroon at Simbock (3°51′N, 11°30′E) and Ipono (2°30′N, 9°50′E) in July 2003, and at Mfou (3°41′N, 11°32′E) in July 2004. They were individually stored in absolute ethanol and kept at –20 °C upon arrival at the laboratory. *Anopheles gambiae sensu lato* specimens were distinguished from other anopheline species by visual inspection using the identification keys of Gillies & Coetzee (1987). For the Simbock and Ipono samples, the species identification was

performed according to Scott *et al*. (1993), and the molecular form was determined following the method of Favia *et al*. (2001). For the Mfou samples, the species and molecular form were assessed using the protocol of Fanello *et al*. (2002). Sampling in 2004 was much more successful than in 2003. All the specimens from Simbock and Ipono belonged to the M molecular form. In Mfou, M and S specimens were found in roughly equal proportions. Individuals from the Kisumu laboratory strain (S form) were used as controls for the Southern blot experiments. Individuals from the strains Vallée du Kou (M form) and Youndé (M form) were used in preliminary experiments to select appropriate probes (see below).

Transposable elements and probes selection

From a dozen of the TE families that we cloned, the elements Crusoe, ozymamdias, Amer6, Agam10 and Aara8 were selected on the basis of their restriction map and previous hybridization trials on the laboratory stocks of Kisumu, Vallée du Kou and Yaounde. Crusoe and ozymamdias were originally described by Hill *et al*. (2001). Crusoe is a Tc1-mariner family class II element related to Tiang (Grossman *et al*. 1999). Ozymandias is an LTR retrotransposon, distantly related to 412 of *Drosophila melanogaster*. Fragments of the reverse transcriptases of Amer6, Agam10 and Aara8 were described by Cook *et al*. (2000). Amer6 and Agam10 are a copia-like LTR retrotransposon and a Pao-like LTR retrotransposon, respectively. Aara8 is a non-LTR retrotransposon. Primers were designed to amplify the reverse transcriptase for the class I elements and the transposase for the class II elements. The resulting polymerase chain reaction (PCR) fragments were cloned into a pGEM-T easy vector (Promega), and the inserts were sequenced to confirm the identity of the clones.

Southern blot experiments

Total genomic DNA was extracted from single females by a standard phenol–chloroform method after proteinase K digestion. DNA was digested with *Bam*HI (to hybridize with Amer6 and Agam10) or *Hpa*I (to hybridize with Aara8, ozymandias, and Crusoe for the Mfou samples only as the Simbock and Ipono samples were not large enough to perform this experiment), which cut only once within the sequences of the TEs. With the combination of restriction enzymes and probes, each TE insertion was represented by a single band on the autoradiographs. Digested DNA of 14 individuals from natural populations, plus two individuals from the Kisumu strain used as internal controls, was loaded onto a 0.8% agarose gel, and electrophoresis was performed at 50 V for 17 h. DNA was then transferred onto nylon N + membranes (Amersham). Prehybridization was performed overnight at 42 °C in 50% formamide, 5× SSC, 0.5% SDS, 5× Denhardt solution and denatured herring sperm DNA (100 µg/mL), and hybridized in a fresh buffer of the same composition at the same temperature for 18 h with the TE probes. Membranes were washed at 42 °C once in 5× SSC and 0.5% SDS (30 min), then in 1× SSC and 0.1% SDS (2 × 30 min). Probes were random prime labelled with a Megaprime kit (Amersham). The X-ray films were exposed at –80 °C for 4–10 days with intensifying screens. The autoradiographs were scanned and the patterns were assessed with a DNA molecular weight ladder. Membranes were dehybridized in 0.4 N NaOH (2 × 20 min). Preliminary experiments indicated that this treatment did not affect subsequent hybridizations. Lanes that provided a weak signal were excluded in the subsequent analyses.

Data analysis

Profile determination. Automated detection of the bands in each individual lane on the Southern blots was performed using the BIO-PROFIL image analysis software (Vilber Lourmat). This software ascribed a molecular weight to each band on the autoradiographs according to the DNA ladder, allowing different gels to be compared. Assuming a 5% error in the molecular weight of the bands, the gels were divided into classes according to their molecular weight, with size intervals proportional to their central value. Only the region of the gel between 1.8 and 18.5 kb, which could be analysed adequately, was considered. By assuming that DNA bands within the same molecular weight class shared the same TE insertion, all the individual profiles on the blots were scored as discrete characters using a binary matrix of 0 and 1, where 0 and 1 corresponded to the absence or presence of a given band, respectively. The individual insertion site number thus corresponded to the number of '1's present in a single lane. This is a slight underestimation of the true number of insertion sites because very large and very small bands were not taken into account.

Insertion polymorphism. In order to compare insertion polymorphism between populations, we computed Jaccard distances between all the possible pairs of individuals within a population. This distance allowed us to compare insertion polymorphism between populations whose insertion site numbers per individual were different. Because any pairwise distances which share an individual are correlated, we used the procedure described in Gilbert *et al*. (2005) to compare the mean within-population distances. These authors propose a method that accommodates such correlations, allowing the calculation of correct estimates of the mean within-population distances and their variance, and a comparison of these mean distances between populations.

Suppose that N individuals are sampled from a population, and that all pairwise distances D_{ij} are computed from their Southern profiles. If μ is the true mean pairwise distance, then an estimate of μ is given by $\hat{\mu} = \{N(N-1)/2\}^{-1}\Sigma_{i<j}\ D_{ij}$. Gilbert *et al*. (2005) show that the variance of $\hat{\mu}$ is $\mathrm{Var}(\hat{\mu}) = \{N(N-1)/2\}^{-1}\{2(N-2)\sigma_1^2 + \sigma_2^2\}$, where $\sigma_1^2 = \mathrm{Cov}(D_{ij}, D_{ik})$ is the covariance of two pairwise distances that share one sequence *i*, and $\sigma_2^2 = \mathrm{Var}(D_{ij})$. Empirical estimates of σ_1^2 and σ_2^2 are given by equations (3) and (4) in Gilbert *et al*. (2005). To compare two populations with mean μ_1 and μ_2, given sample sizes of N_1 and N_2 individuals, Gilbert *et al*. (2005) propose an asymptotically standard normal test statistic for evaluating H_0: $\mu_1 = \mu_2$ as:

$$T = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\left[\sum_{p=1}^{2}\{N_p(N_p-1)/2\}^{-1}\{2(N_p-2)\hat{\sigma}_{p1}^2 + \hat{\sigma}_{p2}^2\}\right]^{1/2}}$$

where *p* denotes the populations.

Population structure. Genetic distances between all the possible pairs of individual Southern profiles were computed according to Huff *et al*. (1993) for each TE family. This distance, which is the sum of the differences between two profiles, is frequently used in population genetics studies (Maguire *et al*. 2002; Gaudeul *et al*. 2004). Principal coordinate analyses (PCoA) were then computed from these distance matrices using the R package ADE4 (Chessel *et al*. 2004), which give a visual representation of the genetic proximities between individuals.

The data from the original 0/1 matrix were analysed by molecular variance (AMOVA) (Excoffier *et al*. 1992) based on the distances between all pairs of profiles calculated above, using GENALEX 6.0 (Peakall & Smouse 2006). The total variation was subdivided into within-population and between-population levels. The AMOVA variance components were tested for significance by nonparametric randomization tests using 9999 permutations with the null hypothesis of no population structure. To allow for the dominant nature of our TE markers, we calculated the Φ_{ST}, which is an analogue of the fixation index of interpopulation differentiation, F_{ST} (Excoffier *et al*. 1992; Peakall *et al*. 1995).

Fixed Aara8 insertions

To isolate fixed Aara8 insertions, two complementary methods were chosen. First, we designed locus specific PCR for every Aara8 insertion found in the sequenced genome of *A. gambiae*. Second, because the *A. gambiae* strain sequenced was a mixture between M and S forms, some insertions fixed in one or in the other form are not present in the genome sequence. To isolate such insertions we thus cloned several Aara8 insertions present in our M and S populations from Mfou, using TAIL PCR (see below).

Amplification of Aara8 insertions present in the genome sequence. Among the 26 Aara8 insertions annotated by Biedler & Tu (2003) (denominated as Ifam3 or AgI3 in their nomenclature) only 12, which had reverse transcriptase fragments, were retained because only those insertions were detected in our Southern blotting experiments (see Table 1 for insertion names and location). For each of these 12 insertions, one primer was designed within the element and another in the flanking DNA (Table 2). Whenever the insertion was too short, primers were designed in flanking regions on both sides of the element. PCR amplification for each of the 12 insertions was then performed on two M and two S individuals from Mfou. When an amplicon of the expected size was observed for the two individuals of a given form and absent in the other form, six additional M individuals and six S individuals were further screened for the presence of the particular insertion. An insertion was considered differentially fixed between M and S when it was amplified in each individual of a given form and never amplified in the other one. In every case, some amplicons were randomly chosen, cloned and sequenced in order to confirm whether they were identical to the expected product.

Amplification of Aara8 insertions from natural populations using TAIL PCR. Thermal asymmetric interlaced (TAIL) PCR was initially developed to sequence unknown DNA regions flanking known sequences. In this way this technique has been used for the recovery of insert end sequences from YAC and P1 clones (Liu & Whittier 1995) and sequences flanking T-DNA insertions in *Arabidopsis thaliana* (Liu *et al*. 1995; Sessions *et al*. 2002) or Mu transposable elements in maize (Settles *et al*. 2004). As this

Table 1 Name and location of the 12 Aara8 insertions that are present in the genome sequence of *Anopheles gambiae* and that contain a reverse transcriptase fragment

Insertion	Chromosome	Cytological division
Ifam3-1	UNK	—
Ifam3-2	3R	32A
Ifam3-3	X	5C
Ifam3-4	X	6
Ifam3-5	UNK	—
Ifam3-6	2R	19D
Ifam3-11	UNK	—
Ifam3-13	UNK	—
Ifam3-15	UNK	—
Ifam3-21	UNK	—
Ifam3-24	3R	37D
Ifam3-26	UNK	—

Table 2 Primers used for the Aara8 insertion/locus specific PCR. Two specific primers were designed for each Aara8 insertion present in the genome sequence of *Anopheles gambiae* (Ifam3) except for Ifam3-21 for which the amplification was performed as for the Aara8 insertions recovered by TAIL PCR

Aara8 insertion	Primers
Ifam3-1	CTGTTGTGCGTATGTGGTGC
	ACTTTTCGGCTGTGTTTTCG
Ifam3-2	ACATCATCCTATCACCCGAC
	CGACAGGTTCTTTGAGGTGA
Ifam3-3	CTAAGGGGGGAGGGTCTAAC
	CGTGAGTGAGCGAGTGTGTC
Ifam3-4	TTCGTTTCAGTTCGTGTAGC
	GGAACCAACGAGTGCTATCA
Ifam3-5	ATGATTTTTCGGGGTTGGGA
	TACAAGGGGAGCATACAGAA
Ifam3-6	AGATACCCTTGAACACAGCG
	TCAAACCATCAACAACAGCC
Ifam3-11	TGAGAAAGAATGGAATGGCG
	TGGTGCGGTGAAACAACAAC
Ifam3-13	AAGTGTAGGAGAAAGGATGG
	ACCTGTTACTCCCACGCTTC
Ifam3-15	TGCTTTGGATGAGACTGGTA
	CGTTGGCAGCAGCAGTGTAA
Ifam3-24	CCAGTCTAAATCCAGGGCGA
	TAACAAACACACGCACACCC
Ifam3-26	TGTCCCTCTGTTCTGTTTTG
	ATGTATGCCGCCGTCTGCTC
Ifam3-21	TTTGAACTCTGGAACTTGAC
S6220-16	TGGTAATGATGAAGACACGA
C58A3-2	ATGATTATCCGTCGAATCTGC
M44-3-5	GAAGCTTTTTACCCCTTGGGCT
M46-10-3	TAGCGTCCTGCTGCAAAGCGA
M44-10-5	CAAACGCCGATTTCATCACGC
M44-10-1	GAATGACGATCAAGAATGAGGGA
Kiss-W4-M13-5	TTCGGCAGATACTAAGTGGTC
Kiss-CA2-M13-1	GCACGGGATTGTACGTTCACGA
Kiss-A6-M13-2	AGGATTTTGTTGGGCTGGGAGA

Table 3 Name and sequence of the primers used for the TAIL PCR experiments

Aara8 specific primers	
RA1	CGGATTAAGGTTTACACAGGGAG
RA2	GAAATGCCCTGGGTAAGAAGATTG
RA3	TTAGGTTTAAGTAKTTTAMGAARAGG
Arbitrary Degenerate primers	
AD1†	NTCGASTWTSGWGTT
AD2†	NGTCGASWGANAWGAA
AD3†	WGTGNAGWANCANAGA
AD10‡	TTGIAGNACIANAGG
AD20‡	TCTTICGNACITNGGA
W4‡	AGWGNAGWANCANAGA
geeky1§	GKYKGCKGCNGC
CG2§	GCNGNWCGWCGWG
GAG3§	GWSIDRAMSCTGCTC

†Liu *et al*. (1995); ‡Smith *et al*. (1996); §Settles *et al*. (2004).

method is entirely PCR-based, it requires less manipulation (no restriction digest, no ligation) than other methods, such as inverse PCR or adapter PCR methods. Moreover, many adapter-based methods use restriction enzymes with a 4-bp recognition site, resulting in very short amplified flanking regions (Settles *et al*. 2004). TAIL PCR is based on the use of nested primers internal to the known sequence (here Aara8) with a high melting temperature (Tm) in combination with an arbitrary degenerate (AD) primer with a low Tm. A first PCR is performed on genomic DNA with the most internal Aara8 primer (RA1, Table 3) and the AD primer. Because of its short size and degeneracy, the AD primer anneals at several sites, sometimes in the neighbourhood of an Aara8 insertion. In this case, the amplification product corresponds to the extremity of the element and its flanking sequence. In order to limit the amplification of nonspecific products due to annealing of the AD primer at two close sites on each of the DNA strands, high annealing temperature cycles alternate with low temperature cycles. During the high temperature cycles, only the Aara8 primer can anneal, whereas both the Aara8 and AD primers can anneal during the low temperature cycles. This first PCR product is then diluted and used as a template for a second PCR with a more external Aara8 primer (RA2) and the same AD primer. The resulting product is then diluted and used as the template in a third PCR with the most external Aara8 primer (RA3) and the same arbitrary degenerate primer. These nested amplifications lead to highly specific products. The third amplicon is then cloned and sequenced. As a single AD primer will not anneal close to all the Aara8 insertions, several AD primers were used in different experiments to amplify the maximum number of Aara8 insertions (see Table 3). Because non-LTR retrotransposons are often 5′-truncated upon insertion, the Aara8 nested primers were designed at the 3′ extremity of the element, at the end of the reverse transcriptase region. As our aim was to amplify the maximum number of Aara8 insertions, degeneracy was included in the most external primer (RA3) because the alignment of Aara8 insertions present in the sequenced *A. gambiae* genome revealed two nucleotide polymorphisms in this part of the element. The TAIL PCRs were performed on two M and two S individuals from Mfou plus one individual from the Kisumu S strain. When Aara8 insertions isolated by this method had no counterpart in the genome sequence, a primer was designed in their flanking region (Table 2). This latter primer, used in combination with the RA2 primer, allowed for an insertion/locus specific PCR. Each insertion/locus specific PCR was

then performed following the same scheme as for the insertions present in the sequenced genome (see above) in order to determine which insertions were specific to one molecular form.

Results

TE individual insertion site numbers

Table 4 summarizes the mean number of bands in individual profiles of Southern filters for the elements Aara8, ozymandias, Crusoe, Amer6 and Agam10 in the M and S populations analysed. The mean number of insertion sites varied according to the TE family and the population. In particular, the S population of Mfou had more insertion sites for Crusoe than the M population (Mann–Whitney *U*-test, U = 2, $P < 0.05$), and more Amer6 sites than all the M populations (Kruskal–Wallis *H*-test, H = 15.58, $P < 0.01$). The least significant difference (LSD) method performed on ranks showed that the difference between the S population of Mfou and the M populations were statistically significant at the 1% level. The numbers of insertion sites of Amer6 were not statistically different between the M populations. Agam10 seemed to follow the same trend but the differences between the S and M molecular forms were not as pronounced as they were for Amer6 (Kruskal–Wallis *H*-test, H = 15.36, $P < 0.01$). The LSD method performed on ranks showed that the difference between the S population of Mfou and the M populations of Simbock and Ipono were significant at the 1% level, as was the difference between the M populations of Mfou and Simbock. On the other hand, the S population of Mfou had fewer ozymandias insertion sites than its M counterpart (Mann–Whitney *U*-test, $U = 1$, $P < 0.05$). Finally, the individual number of Aara8 insertion sites did not differ statistically between the M and S populations of Mfou. Hence, the mean of the TE insertion site numbers appeared to be similar between populations of the M molecular form, despite large geographic distances between the populations, but it showed a general tendency to differ between the populations of the S and M molecular forms sampled in the same village.

TE insertion polymorphism

The measures of within-population diversity for the different TE families, i.e. the mean within-population Jaccard distances, are reported in Table 5. To compare these distances between populations for each TE family, we calculated the *T* statistic of Gilbert *et al*. (2005), and tested whether it was statistically different from zero. Crusoe and Aara8 displayed a very similar level of polymorphism in Mfou S and Mfou M populations. This contrasted with the other elements. Ozymandias insertion

Table 4 Mean insertion site numbers of the five TE families studied in the M and S populations. Standard deviations are in brackets. N, sample size

	Transposable element									
	Aara8		ozymandias		Crusoe		Amer6		Agam10	
	Mean	*N*	Mean	*N*	Mean	*N*	Mean	*N*	Mean	*N*
Mfou S	10.0 (1.2)	5	15.4 (1.7)	5	20.6 (2.9)	5	19.1 (3.4)	7	24.0 (3.0)	6
Mfou M	11.0 (1.4)	6	20.8 (3.7)	6	16.5 (0.8)	6	12.6 (2.7)	5	20.0 (2.8)	5
Ipono M	—		—		—		11.5 (2.3)	13	17.2 (2.1)	11
Simbock M	—		—		—		10.9 (1.9)	8	16.1 (3.1)	7

Table 5 Mean within-population Jaccard distances for the five TE families studied in the M and S populations. N, sample size

	Transposable element									
	Aara8		ozymandias		Crusoe		Amer6		Agam10	
	Mean	*N*	Mean	*N*	Mean	*N*	Mean	*N*	Mean	*N*
Mfou S	0.7812	5	0.7208	5	0.7689	5	0.7739	7	0.7057	6
Mfou M	0.7370	6	0.8157	6	0.7816	6	0.8443	5	0.7378	5
Ipono M	—		—		—		0.8422	13	0.8063	11
Simbock M	—		—		—		0.8595	8	0.7388	7

polymorphism was slightly lower in Mfou S than in Mfou M ($T = 2.40$, $P < 0.05$). All the M populations displayed similar estimates of diversity for the element Amer6, while the S population of Mfou had a lower insertion polymorphism for this element than all the M populations (Mfou S vs. Mfou M: $T = 3.43$, $P < 0.01$; Mfou S vs. Ipono M: $T = 3.18$, $P < 0.01$; Mfou S vs. Simbock M: $T = 5.92$, $P < 0.001$; P values corrected for multiple comparisons). Insertion polymorphism for Agam10 was also lower in the S population of Mfou than in the M populations, although only the difference between Mfou S and the M population of Ipono was significant ($T = 5.24$, $P < 0.01$ after correction for multiple comparisons).

Population structure

Figure 1 shows the scatter diagrams of the PCoA computed for each of the TE families. These diagrams provide visual representations of the Southern profile proximities between individuals within- and between-populations. As data for the five TE families were available only for the S and M samples from Mfou, the populations from Ipono and Simbock were first excluded from the analysis of Amer6 and Agam10 distance matrices to facilitate the comparisons with the other TE families (Fig. 1a). The first two axes accounted for 41.7% to 52.6% of the total variation, depending on the TE family. Whichever TE family was considered, the analysis distinguished clearly between the M and S molecular forms, indicating that individuals of the same molecular form tended to share more insertion bands than individuals from the two different forms. Whenever an individual displayed an ambiguous pattern for one TE family, it was properly clustered with individuals of the same molecular form for the other TE families. When the data of Agam10 and Amer6 of the Ipono and Simbock populations were included in the analysis (Fig. 1b), the S population from Mfou was again well isolated. While the Mfou M population clustered with the other M populations into a diffuse group with the Amer6 data, the different M populations were discriminated against with the Agam10 data.

The significance of the population structure detected by PCoA was tested by AMOVA. The AMOVA detected relatively high between-population variation from the data of the Mfou S and M populations, from 16% of the total variation for Amer6 to 37% for Aara8. All the Φ_{ST} values calculated between the M and S populations of Mfou were high and significant at the 0.005 level, whatever the TE (Table 6). When the populations of Simbock and Ipono were included in the analyses, Amer6 failed to discriminate between the different M populations, in agreement with the PCoA analysis, but all the pairwise Φ_{ST} values were significant with the element Agam10. Overall, the differentiation between the M populations was, however, lower than the differentiation between the S population of Mfou and all the other populations. The geographic distance between Ipono and the other villages was much greater than the distance between Simbock and Mfou, and this was reflected by the pairwise Φ_{ST} estimates between the M populations.

Diagnostic Aara8 insertions

Southern profiles showed that several Aara8 insertions were fixed in one form and absent in the other (Fig. 2). We attempted to localize these insertions using the dual approach of direct locus specific PCR on insertions of the sequenced genome, and TAIL PCR followed by locus specific PCR, on Mfou M and S individuals. Among the 12 Aara8 insertions retrieved from the sequenced genome that contained a reverse transcriptase (RT) fragment, only three (Ifam3-4, Ifam3-6 and Ifam3-21) were redundant with those recovered by TAIL PCR. Among 21 screened insertions (the 12 from the sequenced genome plus nine isolated by TAIL PCR that were not redundant with those from the sequenced genome), one (Ifam3-21) was found in all the M and S individuals screened, including individuals from the Kisumu strain, two (Ifam3-4 and Ifam3-5) were fixed in the M molecular form and absent in the S form, and three (Ifam3-15, C58A3-2 and S6220-16) were fixed in the S form and absent in the M form. Among these 'diagnostic' insertions, only Ifam3–4 could be mapped without ambiguity to cytogenetic division 6 on the X chromosome. Ifam3-5 and Ifam3-15 were located on the UNK chromosome (see Table 1), which groups unplaced scaffolds in ENSEMBL, as was the flanking region of S6220-16. C58A3-2

Table 6 Pairwise Φ_{ST} values for the five TE families studied. Significance was established by 9999 permutations (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$)

	Transposable element				
	Aara8	ozymandias	Crusoe	Amer6	Agam10
Mfou S vs. Mfou M	0.371**	0.231**	0.259**	0.163**	0.233**
Mfou S vs. Ipono M	—	—	—	0.242***	0.275***
Mfou S vs. Simbock M	—	—	—	0.221***	0.183**
Mfou M vs. Ipono M	—	—	—	0.061	0.199***
Mfou M vs. Simbock M	—	—	—	0.019	0.139*
Ipono M vs. Simbock M	—	—	—	0.030	0.178***

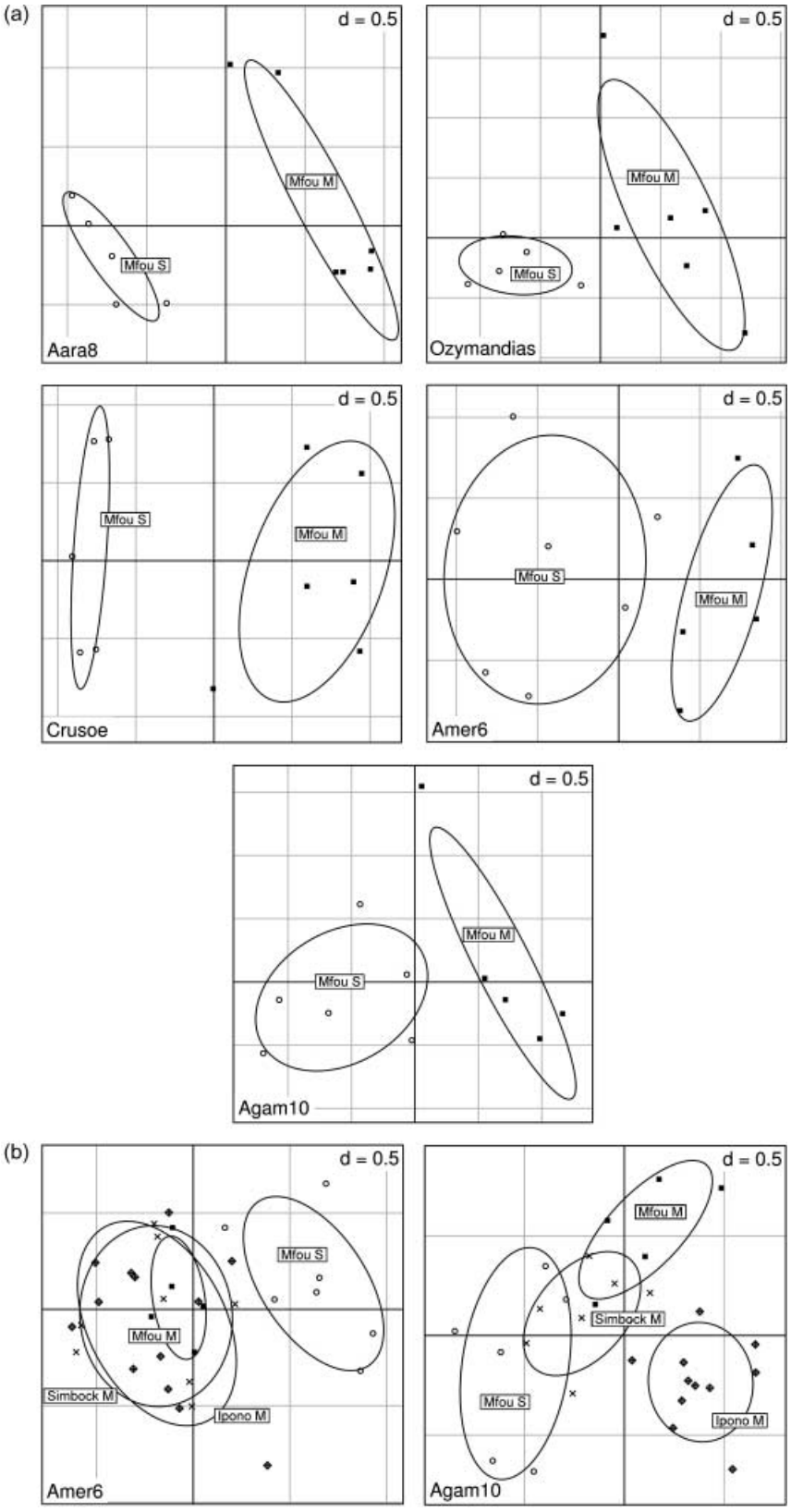

Fig. 1 (a) Principal coordinate analyses computed for the sympatric M and S populations of Mfou, using the five TE families. Each symbol represents a single individual. White circles, S individuals; black squares, M individuals. (b) Principal coordinate analyses computed for all the M and S populations with the Amer6 and Agam10 families. White circles, S individuals from Mfou; black squares, M individuals from Mfou; crosses, M individuals from Simbock; diamonds, M individuals from Ipono. Note: ellipses are centred on the gravity centre of each scatterplot, with axes size equal to 1.5 times the standard deviation of the coordinates of the projections on the axes.

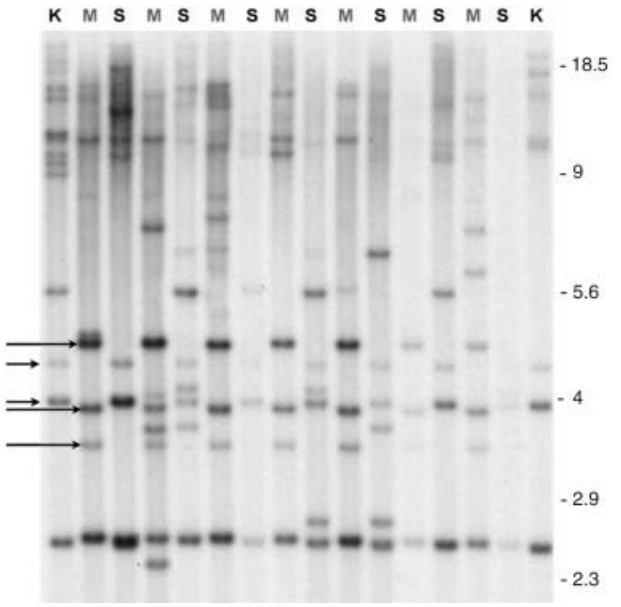

Fig. 2 Southern blot of genomic DNA of *Anopheles gambiae* individuals from Mfou with the probe Aara8. K, control individuals from the Kisumu strain; M, M individual; S, S individual. The arrows show diagnostic insertions. Note that the two S diagnostic insertions observed are present in the Kisumu strain, which originates from Kenya.

could not be assigned to a precise location because its immediate flanking sequence did not match with a high percentage of identity to any sequence present in the genome sequence of *Anopheles gambiae*.

Discussion

The analysis of the distribution of five TE families in M and S populations of *Anopheles gambiae* from Cameroon shows that the number of TE insertion sites per individual varies according to the M and S molecular forms, even in an area of sympatry, whereas the values for distant populations of the M form are more homogeneous. Such a pattern is to be expected if gene flow is severely reduced between the M and S molecular forms, while there is more intense migration between M populations. This difference between M and S forms may result from genetic drift and independent transposition events that removed and added sites independently in the M and S molecular forms. Differences in insertion site numbers could, however, also be the consequence of real differences in TE dynamics in the populations of these two molecular forms.

TE dynamics is the rate at which new TE insertion sites are produced and removed within a genome. It depends on three fundamental parameters: transposition rate, which continuously adds new sites, and excision rate and selection intensity against TE insertions, which remove insertion sites. Many factors control the intensity of these parameters, and any variation of these factors can either have consequences on the dynamics of all the TE families or act more specifically on the parameters of a particular TE family. For example, effective population size, which governs the relative importance of drift and selection against TE insertions (Biémont *et al*. 1997; Charlesworth *et al*. 1997), is a factor which should affect the dynamics of all TE families in the same way. This is also the case for generalist defense pathways of the host against TE insertions, such as methylation in plants (Zilberman & Henikoff 2005) or primates (Meunier *et al*. 2005), which reduce the transposition rate. On the other hand, we know of mechanisms that specifically regulate the transposition of particular families. For example the activity of the endogenous retrovirus gypsy is regulated by the heterochromatic locus *flamenco* (Sarot *et al*. 2004), ZAM and Idefix transposition by the locus COM (Desset *et al*. 2003), and P element transposition is repressed by the presence of defective P copies (Castro & Carareto 2004). As TE dynamics is reflected by the level of insertion polymorphism, we estimated and compared the within-population diversity for different TE families in M and S populations. We showed that the within-population diversity differed between M and S for a given TE family. What are the factors involved in such differences? The fact that Amer6 and, in a less pronounced way, Agam10 displayed a lower diversity (and thus higher insertion frequencies) in the S form than in the M form, in agreement with the higher insertion site number found for these elements in the S form, suggests that the S form may have a lower effective population size than the M form. However, this hypothesis is not strengthened by the results obtained with the other TE families. Indeed, while comparable results concerning within-population diversity were obtained with ozymandias, its insertion site number was higher in the M than in the S molecular form. There is no evidence for a different dynamics between the molecular forms for Crusoe despite significant differences in the insertion site numbers. Moreover, Aara8 exhibited the same insertion site number and within-population diversity in M and S forms. Further data on other TE families, more individuals, and more populations are thus needed to see if a general pattern emerges.

Genetic divergence between the M and S molecular forms is further supported by the PCoA and AMOVA performed on these data, which consistently showed strong genetic differentiation between the sympatric populations of the M and S molecular forms from Mfou, regardless of the TE family. Estimates of Φ_{ST} between the two molecular forms are remarkably close to those reported in a preliminary study by della Torre *et al*. (2005) with the SINE200 TE family. Overall, pairwise Φ_{ST} values calculated between the M populations were lower, when significant. This is in agreement with microsatellite data for other villages in South Cameroon (Wondji *et al*. 2002).

To date, with the exception of the *kdr* mutation and the report of Turner *et al*. (2005), who show that two regions on chromosome 2 exhibited fixed differences between M and S, the two molecular forms appeared to be differentiated essentially for the proximal region of the X chromosome (Barnes *et al*. 2005; Stump *et al*. 2005a, b). In our study, several insertions were differentially fixed between the M and S molecular forms. In particular, the element Aara8 had two insertions fixed in the M form and three in the S form. Over these five insertions, one was located on the X chromosome in the pericentromeric cytological division 6. This confirms the divergence of the M and S molecular forms in this region, which could host genes implicated in the assortative mating of molecular forms (Wang *et al*. 2001; Lehmann *et al*. 2003; Stump *et al*. 2005b). However, the four remaining insertions could not be mapped on the chromosomes. One possible reason of our inability to locate these insertions is that they are embedded within the heterochromatin, which is known to accumulate TEs and is difficult to sequence.

Our results show that TEs can be very efficient markers for population genetics investigations in *A. gambiae*. Because transposition rate has been repeatedly found to be much higher than nucleotide mutation rate, the presence of insertions specific to one lineage (private insertions) after the rupture of gene flow is likely to arise more quickly than specific alleles of other kinds of markers, such as amplified fragment length polymorphism (AFLP). Conversely, very high transposition or excision rates would obscure population structure because they would prevent the sharing of insertions by several individuals. For example, in our study Amer6, which displayed higher within-population diversities than Agam10, did not provide significant Φ_{ST} values between M populations.

Given the current research effort toward the release of transgenic mosquitoes, and the possible use of TEs as a genetic drive to help malaria refractoriness to reach fixation in natural populations, it is an absolute necessity to continue to investigate the dynamics of TEs in the natural populations of *A. gambiae ss* and other malaria vectors. In particular it will be essential to carry out studies of this kind for the endogenous families close to that of the TE chosen to achieve the genetic drive, first because these specific TEs could be used as a model for the diffusion of the transgene; second, because they could interact with the transgenic TE, as has already been reported (Sundararajan *et al*. 1999).

Acknowledgements

We thank the people from the Mfou, Ipono and Simbock villages for their help with mosquito sampling, Isabelle Morlais and Catherine Loevenbruck for technical assistance, Anne-Béatrice Dufour for advice about the statistical procedures, Valérie James for her help with the English text, and the reviewers for their useful comments.

References

Awolola TS, Brooke BD, Koekemoer LL, Coetzee M (2003) Absence of the kdr mutation in the molecular 'M' form suggests different pyrethroid resistance mechanisms in the malaria vector mosquito *Anopheles gambiae* s.s. *Tropical Medicine and International Health*, **8**, 420–422.

Barnes MJ, Lobo NF, Coulibaly MB *et al*. (2005) SINE insertion polymorphism on the X chromosome differentiates *Anopheles gambiae* molecular forms. *Insect Molecular Biology*, **14**, 353–363.

Biedler J, Tu Z (2003) Non-LTR retrotransposons in the African malaria mosquito, *Anopheles gambiae*: unprecedented diversity and evidence of recent activity. *Molecular Biology and Evolution*, **20**, 1811–1825.

Biémont C, Tsitrone A, Vieira C, Hoogland C (1997) Transposable element distribution in *Drosophila*. *Genetics*, **147**, 1997–1999.

Boëte C, Koella JC (2003) Evolutionary ideas about genetically manipulated mosquitoes and malaria control. *Trends in Parasitology*, **19**, 32–38.

Boulesteix M, Biémont C (2005) Transposable elements in mosquitoes. *Cytogenetic and Genome Research*, **110**, 500–509.

Castro JP, Carareto CMA (2004) *Drosophila melanogaster* P transposable elements: mechanisms of transposition and regulation. *Genetica*, **121**, 107–118.

Catteruccia F, Nolan T, Loukeris TG *et al*. (2000) Stable germline transformation of the malaria mosquito *Anopheles stephensi*. *Nature*, **405**, 959–962.

Chandre F, Manguin S, Brengues C *et al*. (1999) Current distribution of a pyrethroid resistance gene (kdr) in *Anopheles gambiae* complex from west Africa and further evidence for reproductive isolation of the Mopti form. *Parassitologia*, **41**, 319–322.

Charlesworth B, Langley CH, Sniegowski PD (1997) Transposable element distributions in *Drosophila*. *Genetics*, **147**, 1993–1995.

Chessel D, Dufour AB, Thioulouse J (2004) The ADE 4 package — I: one-table methods. *R News*, **4**, 5–10.

Coluzzi M, Sabatini A, della Torre A, Di Deco MA, Petrarca V (2002) A polytene chromosome analysis of the *Anopheles gambiae* species complex. *Science*, **298**, 1415–1418.

Cook JM, Martin J, Lewin A, Sinden RE, Tristem M (2000) Systematic screening of *Anopheles* mosquito genomes yields evidence for a major clade of Pao-like retrotransposons. *Insect Molecular Biology*, **9**, 109–117.

Deceliere G, Charles S, Biémont C (2005) The dynamics of transposable elements in structured populations. *Genetics*, **169**, 467–474.

Desset S, Meignin C, Dastugue B, Vaury C (2003) COM, a heterochromatic locus governing the control of independent endogenous retroviruses from *Drosophila melanogaster*. *Genetics*, **164**, 501–509.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fanello C, Santolamazza F, della Torre A (2002) Simultaneous identification of species and molecular forms of the *Anopheles gambiae* complex by PCR-RFLP. *Medical and Veterinary Entomology*, **16**, 461–464.

Favia G, Lanfrancotti A, Spanos L, Siden-Kiamos I, Louis C (2001) Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. *Insect Molecular Biology*, **10**, 19–23.

Gaudeul M, Till-Bottraud I, Barjon F, Manel S (2004) Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae):

comparison of AFLP and microsatellite markers. *Heredity*, **92**, 508–518.

Gentile G, Slotman M, Ketmaier V, Powell JR, Caccone A (2001) Attempts to molecularly distinguish cryptic taxa in *Anopheles gambiae* s.s. *Insect Molecular Biology*, **10**, 25–32.

Gentile G, Santolamazza F, Fanello C *et al.* (2004) Variation in an intron sequence of the voltage-gated sodium channel gene correlates with genetic differentiation between *Anopheles gambiae* s.s. molecular forms. *Insect Molecular Biology*, **13**, 371–377.

Gilbert PB, Rossini AJ, Shankarappa R (2005) Two-sample tests for comparing intra-individual genetic sequence diversity between populations. *Biometrics*, **61**, 106–117.

Gillies MT, Coetzee M (1987) *A Supplement to the Anophelinae of Africa South of the Sahara*. The South African Institute For Medical Research, Johannesburg, South Africa.

Grossman GL, Cornel AJ, Rafferty CS, Robertson HM, Collins FH (1999) Tsessebe, Topi and Tiang: three distinct Tc1-like transposable elements in the malaria vector, *Anopheles gambiae*. *Genetica*, **105**, 69–80.

Grossman GL, Rafferty CS, Clayton JR *et al.* (2001) Germline transformation of the malaria vector, *Anopheles gambiae*, with the piggyBac transposable element. *Insect Molecular Biology*, **10**, 597–604.

Hill SR, Leung SS, Quercia NL *et al.* (2001) Ikirara insertions reveal five new *Anopheles gambiae* transposable elements in islands of repetitious sequence. *Journal of Molecular Evolution*, **52**, 215–231.

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloë dactyloides* (Nutt.) Engelm.]. *Theoretical and Applied Genetics*, **86**, 927–934.

Ito J, Ghosh A, Moreira LA, Wimmer EA, Jacobs-Lorena M (2002) Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature*, **417**, 452–455.

Lehmann T, Licht M, Elissa N *et al.* (2003) Population structure of *Anopheles gambiae* in Africa. *Journal of Heredity*, **94**, 133–147.

Lepetit D, Brehm A, Fouillet P, Biémont C (2002) Insertion polymorphism of retrotransposable elements in populations of the insular, endemic species *Drosophila madeirensis*. *Molecular Ecology*, **11**, 347–354.

Liu YG, Whittier RF (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, **25**, 674–681.

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant Journal*, **8**, 457–463.

Maguire T, Peakall R, Saenger P (2002) Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. *Theoretical and Applied Genetics*, **104**, 388–398.

Masendu HT, Hunt RH, Govere J *et al.* (2004) The sympatric occurrence of two molecular forms of the malaria vector *Anopheles gambiae* Giles *sensu stricto* in Kanyemba, in the Zambezi Valley, Zimbabwe. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **98**, 393–396.

Meunier J, Khelifi A, Navratil V, Duret L (2005) Homology-dependent methylation in primate repetitive DNA. *Proceedings of the National Academy of Sciences, USA*, **102**, 5471–5476.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peakall R, Smouse PE, Huff DR (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloë dactyloides*. *Molecular Ecology*, **4**, 135–147.

Perera OP, Harrell IIRA, Handler AM (2002) Germ-line transformation of the South American malaria vector, *Anopheles albimanus*, with a piggyBac/EGFP transposon vector is routine and highly efficient. *Insect Molecular Biology*, **11**, 291–297.

Ribeiro JM, Kidwell MG (1994) Transposable elements as population drive mechanisms: specification of critical parameter values. *Journal of Medical Entomology*, **31**, 10–16.

Sarot E, Payen-Groschene G, Bucheton A, Pélisson A (2004) Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the *Drosophila melanogaster* flamenco gene. *Genetics*, **166**, 1313–1321.

Scott JA, Brogdon WG, Collins FH (1993) Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. *American Journal of Tropical Medicine and Hygiene*, **49**, 520–529.

Sessions A, Burke E, Presting G *et al.* (2002) A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell*, **14**, 2985–2994.

Settles AM, Latshaw S, McCarty DR (2004) Molecular analysis of high-copy insertion sites in maize. *Nucleic Acids Research*, **32**, e54.

Smith D, Yanai Y, Liu YG *et al.* (1996) Characterization and mapping of Ds-GUS-T-DNA lines for targeted insertional mutagenesis. *Plant Journal*, **10**, 721–732.

Stump AD, Shoener JA, Costantini C, Sagnon N, Besansky NJ (2005a) Sex-linked differentiation between incipient species of *Anopheles gambiae*. *Genetics*, **169**, 1509–1519.

Stump AD, Fitzpatrick MC, Lobo NF *et al.* (2005b) Centromere-proximal differentiation and speciation in *Anopheles gambiae*. *Proceedings of the National Academy of Sciences, USA*, **102**, 15930–15935.

Sundararajan P, Atkinson PW, O'Brochta DA (1999) Transposable element interactions in insects: crossmobilization of hobo and Hermes. *Insect Molecular Biology*, **8**, 359–368.

della Torre A, Fanello C, Akogbeto M *et al.* (2001) Molecular evidence of incipient speciation within *Anopheles gambiae* s.s. in West Africa. *Insect Molecular Biology*, **10**, 9–18.

della Torre A, Costantini C, Besansky NJ *et al.* (2002) Speciation within *Anopheles gambiae* — the glass is half full. *Science*, **298**, 115–117.

della Torre A, Tu Z, Petrarca V (2005) On the distribution and genetic differentiation of *Anopheles gambiae* s.s. molecular forms. *Insect Biochemistry and Molecular Biology*, **35**, 755–769.

Tu Z, Coates C (2004) Mosquito transposable elements. *Insect Biochemistry and Molecular Biology*, **34**, 631–644.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, **3**, e285.

Wang R, Zheng L, Toure YT, Dandekar T, Kafatos FC (2001) When genetic distance matters: measuring genetic differentiation at microsatellite loci in whole-genome scans of recent and incipient mosquito species. *Proceedings of the National Academy of Sciences, USA*, **98**, 10769–10774.

Wondji C, Simard F, Fontenille D (2002) Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. *Insect Molecular Biology*, **11**, 11–19.

Wondji C, Simard F, Petrarca V *et al.* (2005) Species and populations of the *Anopheles gambiae* complex in Cameroon with special emphasis on chromosomal and molecular forms of *Anopheles gambiae s.s*. *Journal of Medical Entomology*, **42**, 998–1005.

Yawson AE, McCall PJ, Wilson MD, Donnelly MJ (2004) Species abundance and insecticide resistance of *Anopheles gambiae* in

selected areas of Ghana and Burkina Faso. *Medical and Veterinary Entomology*, **18**, 372–377.

Zampicinini G, Blinov A, Cervella P, Guryev V, Sella G (2004) Insertional polymorphism of a non-LTR mobile element (NLRCth1) in European populations of *Chironomus riparius* (Diptera, Chironomidae) as detected by transposon insertion display. *Genome*, **47**, 1154–1163.

Zilberman D, Henikoff S (2005) Epigenetic inheritance in *Arabidopsis*: selective silence. *Current Opinion in Genetics and Development*, **15**, 557–562.

This research is part of a larger project being conducted under the supervision of Christian Biémont, which is intended to understand the dynamics of transposable ekements in natural populations of Anopheles. Christophe Antonio-Nkondjio, Herman P. Awono-Ambene, Didier Fontenille, Frédéric Simard, are IRD, OCEAC, and IMPM scientists specializing in natural populations of mosquitoes. Matthieu Boulesteix is a PhD student specializing in population genetics and genomics.