Insect Molecular Biology (2004) **13**(3), 251–258

Population structure of the malaria vector *Anopheles funestus* in Senegal based on microsatellite and cytogenetic data

A. Cohuet*, I. Dia†, F. Simard‡, M. Raymond§ and D. Fontenille*

**Laboratoire de Lutte Contre les Insectes Nuisibles, Institut de Recherche pour le Développement (LIN-IRD), Montpellier, France;* †*Laboratoire d'Entomologie Médicale, Institut Pasteur, Dakar, Senegal;* ‡*Laboratoire d'Entomologie, antenne IRD, Organisation de Coordination pour la lutte contre les Endémies en Afrique Centrale (OCEAC), Yaoundé, Cameroon; and* §*Institut des Sciences de l'Evolution, Laboratoire de Génétique et Environnement, CC065, UMR CNRS 5554, Université de Montpellier II, France*

Abstract

The study of chromosomal inversions distribution within natural *Anopheles funestus* populations from West Africa revealed high levels of genetic structuring. In Burkina Faso, this was interpreted as evidence for incipient speciation, and two chromosomal forms were described, namely 'Folonzo' and 'Kiribina'. Assignation of field collected specimens to one chromosomal form depends upon application of an algorithm based on chromosomal inversions. We assessed relevance and applicability of this algorithm on *An. funestus* populations from Senegal, where both forms occur. Furthermore, we estimated the level of genetic differentiation between populations using microsatellite loci spread over the whole genome. Significant genetic differentiation was revealed between geographical populations of *An. funestus*, and the pattern observed suggested isolation by distance. Chromosomal heterogeneity was not detected by microsatellite markers. Thus, although incipient speciation could not be ruled out by our data, our results suggest that differential environmental selection pressure acting on inversions should be considered a major factor in shaping their distribution in wild *An. funestus* populations.

Keywords: *Anopheles funestus*, population structure, microsatellites, chromosomal inversions, speciation.

Received 23 October 2003; accepted 5 January 2004. Correspondence: Dr Didier Fontenille, Laboratoire de Lutte Contre les Insectes Nuisibles, Institut de Recherche pour le Developpement (LIN-IRD), 911 Avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France. Tel.: +33 4 67 04 32 22; fax: +33 4 67 54 20 44; e-mail: didier.fontenille@mpl.ird.fr

Introduction

Anopheles funestus Giles is a major vector of human malaria in Africa. Together with *An. gambiae* Giles and *An. arabiensis* Patton, these species complement each other and sustain perennial, high intensity malaria transmission all over sub-Saharan Africa. *An. funestus* often follows in peak abundance *An. gambiae* after the end of the rainy season, thereby extending malaria transmission into the dry season (Gillies & De Meillon, 1968). Moreover, in some areas of Africa, *An. funestus* is the main vector of malaria, with infection rates as high or higher than its counterparts from the *An. gambiae* complex (Fontenille & Lochouarn, 1999; Fontenille *et al*., 1997; Manga *et al*., 1997; Mendis *et al*., 2000). As such, any vector control measure to be implemented in the field will need to target at least *An. funestus* as well as members of the *An. gambiae* complex. However, despite such paramount epidemiological importance in malaria transmission, *An. funestus* has received far less attention than members of the *An. gambiae* complex and basic data on the biology, role as a malaria vector and population genetic structuring within this species are crucially lacking. The importance of detailed understanding of vector population structure has been emphasized in the context of genetic control (Aultman *et al*., 2001; Collins & Besansky, 1994; Curtis *et al*., 1999; Green, 1981). Comprehensive knowledge on gene flow between populations is also pivotal to the successful use and management of classical control techniques based on insecticides, as it makes it possible to foresee the spread of genes of interest such as those involved in insecticide resistance, refractoriness/susceptibility to *Plasmodium* infection, or feeding behaviour.

An. funestus belongs to a group of ten species that are morphologically very similar and can only be distinguished at specific stages of their development (Gillies & Coetzee,

1987; Gillies & De Meillon, 1968). The members of this group are *An. funestus*, *An. vaneedeni* Gillies and Coetzee, *An. parensis* Gillies, *An. aruni* Sobti, *An. confusus* Evans and Leeson, *An. rivulorum* Leeson, *An. fuscivenosus* Leeson, *An. leesoni* Evans, *An. brucei* Service and the recently identified *An. rivulorum*-like (Cohuet *et al.*, 2003; Hackett *et al.*, 2000). Of the ten species, *An. funestus* has the widest distribution, extending throughout the whole of sub-Saharan Africa, and is also highly anthropophilic and endophilic. This is the only member of the group to play a significant role in human malaria transmission continent-wise, although other species of the group have been found naturally infected with *Plasmodium falciparum* (Wilkes *et al.*, 1996).

Few studies have focused on the genetic structure of *An. funestus* populations. However, several observations led to the prediction of population subdivision. Evidence for population heterogeneity within and between *An. funestus* populations comes from the distribution of paracentric inversions on chromosomal arms 2R, 3R and 3L (Green & Hunt, 1980).

In Burkina Faso, *An. funestus* populations are polymorphic for at least four paracentric chromosomal inversions (Boccolini *et al.*, 1994). Wide variations in inversion frequencies were observed among samples without consistent geographical or temporal clines (Costantini *et al.*, 1999). Highly significant linkage disequilibrium and departures from Hardy–Weinberg expectations were recorded for inversions 3Ra, 3Rb and 2Rs in most samples, suggesting a Walhund effect (pooling of subpopulations). Furthermore, significantly higher frequencies of the inverted arrangements 3Ra and 3Rb were found in indoor human-fed samples vs. corresponding outdoor animal-fed samples.

Based on these findings, two chromosomal forms, named 'Kiribina' and 'Folonzo' were defined by Costantini *et al.* (1999). Specimens of *An. funestus* can be assigned to a chromosomal form using an algorithm based on chromosomal inversions. The Kiribina form is characterized by standard 2R and 3R chromosomes, or presence of the 2Rs inversion. The Folonzo form is nearly fixed for inversions 3Ra and 3Rb, polymorphic on arm 2R, and seems to have a higher vectorial capacity. Such level of structuring led Costantini *et al.* (1999) to hypothesize that chromosomal forms represent reproductive units and thus that incipient speciation occurs in *An. funestus*.

To further assess the taxonomic status of these chromosomal forms, we studied genetic variability at nine microsatellite loci in four *An. funestus* populations collected along a west-east transect in Senegal (West Africa), where previous cytogenetic studies demonstrated high levels of chromosomal heterogeneities, both within and between populations (Dia *et al.*, 2000; Lochouarn *et al.*, 1998). We particularly focused on the village of Kouvar where chromosomal data showed strong heterokaryotype deficits, suggesting that both chromosomal forms are sympatric and synchronous in this area.

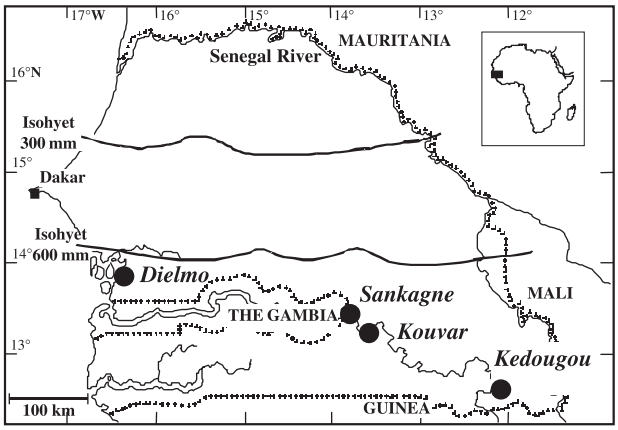

Figure 1. Map of Senegal showing the 4 study sites (in italics).

Results

Microsatellite data

Genetic diversity. Genotypes at nine microsatellite loci were determined from 247 *An. funestus* specimens collected in our four study sites (Fig. 1). All loci were highly polymorphic showing between six and twenty-five alleles (Table 1). Across the four populations, mean number of alleles per locus ranged from 9 to 11.4 and mean observed heterozygosity across all loci ranged from 0.70 to 0.73.

Hardy–Weinberg and linkage disequilibrium. When considering all samples as belonging to a single population, Hardy–Weinberg expectations were significantly rejected ($P < 0.05$, single test level) for AFND3, AFND20, FunG, FunD and FunF, with positive values of F_{is} (Table 1, 'overall'). Five out of nine loci (or 56%, $P < 10^{-4}$, binomial test with 0.05% 'success' rate) therefore showed significant deficit of heterozygotes, suggesting pooling of different gene pools (Wahlund effect). Alternatively, using the Bonferroni procedure to detect locus-specific deviations, two F_{is} estimates remained statistically significant at the 5% threshold, for loci FunD and FunG. At the population level, slight deviations from Hardy–Weinberg expectations occurred. In Kedougou, the deficit of heterozygotes across all loci was significant ($P < 0.05$), suggesting a subdivision in the population (Table 1). Exact tests for linkage disequilibrium within each population and in the pooled population resulted in two significant values after the sequential Bonferroni procedure was applied (36 pair-wise comparisons within each population), between AFND19 and FunL and between FunO and FunD in the population of Kedougou. Together with significant heterozygote deficit observed over all loci, linkage disequilibrium in Kedougou strengthens our findings of population heterogeneity in this sample.

No pair of loci appeared in linkage disequilibrium in more than one population, suggesting the absence of statistical

Table 1. Genetic variability and significance level for goodness of fit tests to Hardy-Weinberg equilibrium within *Anopheles funestus* populations from 4 localities in Senegal

Locus (chromosomal location)		Populations				
		Dielmo	Sankagne	Kouvar	Kedougou	Overall[a]
		$2n = 100$	$2n = 120$	$2n = 144$	$2n = 130$	$2n = 494$
AFND3	Nall	3	5	5	4	6
(unknown)	Hobs	0.56	0.53	0.37	0.39	0.46
	F_{is}	−0.064	−0.028	**0.113**	−0.146	**0.024**
AFND5	Nall	7	6	6	5	7
(2R: 15 A)	Hobs	0.59	0.61	0.74	0.63	0.65
	F_{is}	−0.007	0.023	−0.040	−0.121	−0.020
FunO	Nall	8	11	12	11	15
(2R: 18 A)	Hobs	0.79	0.75	0.77	0.69	0.75
	F_{is}	−0.003	−0.016	0.007	0.086	0.032
AFND20	Nall	8	11	9	8	11
(3R: 32D)	Hobs	0.75	0.79	0.65	0.67	0.72
	F_{is}	0.042	0.046	0.125	**0.138**	**0.091**
AFND19	Nall	8	12	12	11	15
(3R: 34 A)	Hobs	0.74	0.76	0.74	0.76	0.75
	F_{is}	−0.044	0.040	0.033	**0.062**	0.029
FunG	Nall	11	14	11	8	17
(3R: 35 A)	Hobs	0.84	0.85	0.77	0.86	0.83
	F_{is}	−0.085	0.030	0.081	−0.041	**0.028**
FunD	*Nall*	15	20	21	21	25
(3R: 35B)	Hobs	0.71	0.75	0.87	0.92	0.81
	F_{is}	0.121	0.117	0.038	**0.026**	**0.092**
FunL	Nall	14	16	15	9	21
(2L: 24C)	Hobs	0.86	0.80	0.75	0.72	0.78
	F_{is}	−0.061	0.037	0.066	0.142	0.053
FunF	Nall	7	8	8	6	8
(3L: 43 A)	Hobs	0.70	0.70	0.70	0.64	0.69
	F_{is}	**0.060**	−0.064	0.063	0.136	**0.055**
Mean	Nall	9	11.4	11	9.2	14
across	Hobs	0.72	0.73	0.70	0.70	0.71
all loci[b]	F_{is}	−0.002	0.027	0.049	**0.044**	**0.045**

$2n$, number of chromosomes scored.
Hobs, observed heterozygosity (direct count).
Nall, number of alleles.
F_{is} was calculated according to Weir & Cockerham (1984).
Bold values: $P < 0.05$; Bold underlined values: $P < 0.05$ after Bonferonni correction (see text).
[(a)] all specimens were considered as belonging to one single panmictic population.
[(b)] multilocus F_{is} estimates were computed according to Weir & Cockerham (1984) using GENEPOP 3.2. (Raymond & Rousset, 1995); statistical significance of F_{is} across loci was assessed through Fisher exact test.

linkage between loci despite their proximity in the genome (AFND19, AFND20, FunD and FunG on arm 3R) (Fig. 2).

Genetic differentiation. Estimates of pair-wise F_{st} between populations and across all loci are shown in Table 2, together with corresponding geographical distance between sampling sites.

Average F_{st} based on the whole data set (9 loci and 4 populations) was low (0.0214) but highly significant ($P < 10^{-4}$), suggesting restricted gene flow between *An. funestus* populations at this geographical scale. Single locus estimates varied among loci from 0.0057 to 0.0510; all of them were statistically significant ($P < 0.05$), thus showing genome-wide differentiation. The highest single locus F_{st} estimates were observed for loci AFND3, FunG and FunD, a result that is consistent with previous evidence for strong and significant heterozygote deficits observed at these loci when populations are pooled (see above). Mean pair-wise F_{st} estimates over all loci between the different locations ranged from 0.0080 ($P < 0.05$) to 0.0424 ($P < 10^{-4}$) (Table 2). Global and pair-wise F_{st} estimates were computed separately for loci located within (4 loci) and outside (4 loci) chromosomal inversions. There was no significant difference (Mann–Whitney test, $P = 0.48$), suggesting homogeneity between these groups of loci and negligible (if any) influence of chromosomal inversions on our estimates.

Isolation by distance. Isolation by distance was tested and showed a correlation between genetic (pair-wise F_{st}) and geographical distance ($P < 0.05$), when considering the whole data set, despite a low number of populations (Fig. 3). These results suggest that the level of the genetic differentiation between populations of *An. funestus* in Senegal would be mainly due to a restriction of gene flow by geographical distance. The correlation became nonsignificant when considering only loci located within inversions

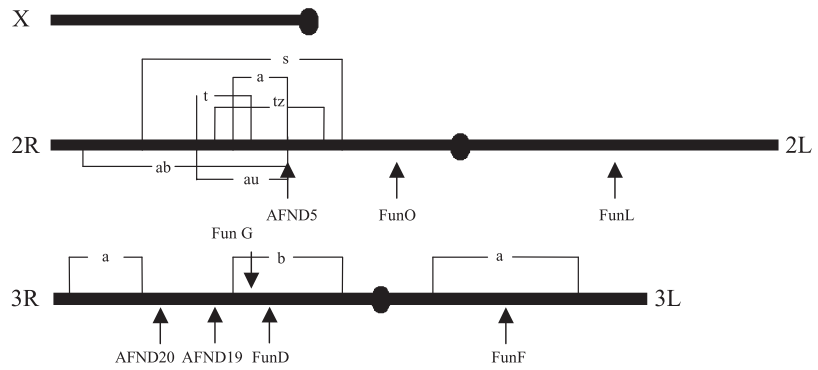

Figure 2. Chromosomal inversions of *An. funestus* in Senegal, and location of microsatellite loci, AFND3 is not located yet (Sharakhov *et al.*, 2003).

Table 2. Genetic differentiation (F_{st}) between *Anopheles funestus* populations from Senegal

	Sankagne vs. Kouvar (25 km)	Kouvar vs. Kedougou (170 km)	Sankagne vs. Kedougou (195 km)	Dielmo vs. Sankagne (305 km)	Dielmo vs. Kouvar (330 km)	Dielmo vs. Kedougou (500 km)	All
Loci							
AFND3	0.0083	–0.0021	0.0332	**0.0325**	**0.0938**	**0.1404**	**0.0510**
AFND5	0.0065	**0.0427**	0.0125	–0.0002	**0.0282**	–0.0045	**0.0142**
FunO	**0.0083**	0.0067	–0.0095	**0.0604**	**0.0178**	**0.0597**	**0.0239**
AFND20	0.0104	–0.0023	–0.0019	**0.0052**	**0.0287**	**0.0127**	**0.0088**
AFND19	–0.0003	0.0077	0.0007	**0.0157**	**0.0059**	**0.0274**	**0.0095**
FunG	**0.0124**	–0.0042	**0.0139**	**0.0811**	**0.0783**	**0.0675**	**0.0415**
FunD	0.0060	0.0121	**0.0401**	**0.0394**	**0.0282**	**0.0756**	**0.0336**
FunL	0.0179	0.0059	0.0119	**0.0143**	–0.0043	**0.0096**	**0.0092**
FunF	0.0008	0.0024	**0.0208**	0.0034	–0.0029	**0.0096**	**0.0057**
Mean across all loci	**0.0080**	**0.0085**	**0.0131**	**0.0292**	**0.0274**	**0.0424**	**0.0214**
Loci within inversions[a]	**0.0064**	**0.0135**	**0.0228**	**0.0342**	**0.0305**	**0.0398**	**0.0245**
Loci outside inversions[b]	**0.0095**	0.0048	0.0004	**0.0237**	**0.0124**	**0.0280**	**0.0131**

Geographical distance between sites is given in parentheses.
Bold values: $P < 0.05$; bold underlined values; $P < 0.05$ after Bonferroni correction.
[a]Loci within polymorphic inversions are AFND5, FunG, FunD and FunF.
[b]Loci outside polymorphic inversions are AFND19, AFND20, FunO and FunL.

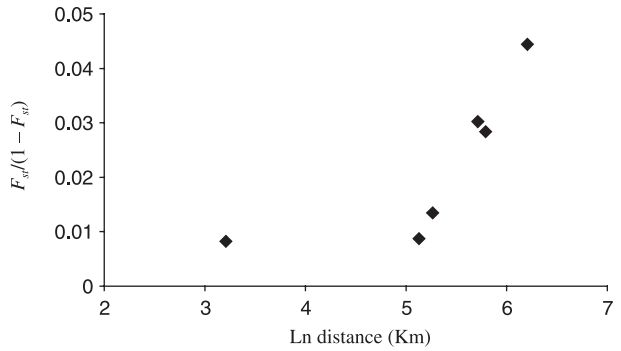

Figure 3. Correlation between $F_{st}/(1 - F_{st})$ and logarithm of distance (in km) for pair-wise comparisons of 4 *Anopheles funestus* populations from Senegal genotyped at 9 microsatellite loci.

($P = 0.08$) or outside inversions ($P = 0.19$), probably due to lack of statistical power.

Population of Kouvar. In the Kouvar sample, 42 individuals were analysed both for chromosomal inversions and microsatellite genotypes. All of them were collected between July and October 2000.

No inversion was found on chromosomal arm 2L or on the X heterosome in accordance with other studies on this species (Boccolini *et al.*, 2002; Boccolini *et al.*, 1994; Boccolini *et al.*, 1998; Costantini *et al.*, 1999; Dia *et al.*, 2000; Green & Hunt, 1980; Kamau *et al.*, 2003; Lochouarn *et al.*, 1998; Sharakhov *et al.*, 2001). One inversion was detected on each of the three other arms: 2Rs, 3Rb and 3La. Significant departure from Hardy–Weinberg equilibrium was observed for all inversions, showing a deficit of heterokaryotypes (Table 3). All pairs of inversions appeared in linkage disequilibrium ($P < 0.05$). Linkage disequilibrium was tested between chromosomal inversions and microsatellite alleles. No pair-wise test was significant ($P > 0.05$). When individuals were assigned to chromosomal forms following the algorithm of Costantini *et al.* (1999), thirty-eight were assigned without ambiguity: seven belong to the Folonzo form and thirty-one to the Kiribina form. Hardy–Weinberg equilibrium was re-established in the Folonzo form, but deficit of heteorokaryotypes remained significant for inversions

Table 3. Chromosomal arrangements and inversions frequencies (f) in 42 *Anopheles funestus* individuals from Kouvar, analysed both for chromosomal inversions and microsatellite genotypes

Chromosomal arm 2R	2*n*	f(s/s)	f(s/+)	f(+/+)	f(s)	f(+)	F_{is}
	80	0.625	0.1	0.275	0.675	0.325	**0.797**
Chromosomal arm 3R	2*n*	f(b/b)	f(b/+)	f(+/+)	f(b)	f(+)	F_{is}
	80	0.225	0	0.775	0.225	0.775	**1**
Chromosomal arm 3L	2*n*	f(a/a)	f(a/+)	f(+/+)	f(a)	f(+)	F_{is}
	84	0.262	0.238	0.5	0.381	0.619	**0.589**

2*n*, number of chromatids scored.
F_{is} calculated according to Weir & Cockerham (1984).
Bold values $P < 0.05$; bold underlined values: $P < 0.01$.

2Rs ($P < 0.05$) and 3La ($P < 0.05$) in the Kiribina form, as opposed to what was found in Burkina Faso.

Discussion

Population structure of *An. funestus* from Senegal was previously studied at the chromosomal level (Dia *et al*., 2000; Lochouarn *et al*., 1998) and showed a high degree of structuring. The population of Dielmo was fixed for 2R+ and inversions 3Ra and 3Rb were observed floating at low frequency (< 18%). Population of Sankagne was nearly fixed for 2Rs and 3R+. When applying the algorithm proposed by Costantini *et al*. (1999), most of specimens from Dielmo and all specimens from Sankagne were assigned to the Kiribina form. However, Dielmo and Sankagne populations were almost fixed for alternative arrangements on arm 2R, suggesting that the Kiribina form is not homogenous in Senegal. The population of Kedougou was typical to the chromosomal form Folonzo with the chromosomal arm 2R polymorphic (without 2Rs inversion) and 3Ra and 3Rb nearly fixed. The population of Kouvar showed both chromosomal forms with 2Rs, 3Ra, 3Rb and 3La polymorphic and significant departure from Hardy–Weinberg equilibrium and linkage disequilibrium between each pair of inversions (Lochouarn *et al*., 1998). Our results also showed strong heterokaryotype deficit and linkage disequilibrium between chromosomal inversions in Kouvar. In this locality, most of the specimens were assigned to the Kiribina form. However, significant deficit in heterokaryotypes persisted for inversions 2Rs and 3La after assignment of specimens to their chromosomal form. The fact that Hardy–Weinberg equilibrium was not re-established in the Kouvar population by separating specimens following Costantini's algorithm, confirmed that chromosomal structuring differs between Burkina Faso and Senegal. If the algorithm is well adapted in Burkina Faso it seems unlikely that it can be extended to populations distant from Burkina Faso as previously shown with samples from Angola (Boccolini *et al*., 2002).

It was shown that chromosomal inversions reduce or even suppress recombination in certain regions of the genome when found at the heterozygous state (Caceres *et al*., 1999; Navarro *et al*., 1997; Rieseberg, 2001), and therefore can protect well adapted gene associations from recombination (Alvarez & Zapata, 1997; Caceres *et al*., 1999; Coluzzi, 1982; Coluzzi *et al*., 1979). With this respect, the most likely explanations for heterozygote deficiencies observed for chromosomal arrangements in our *An. funestus* samples are: (i) reproductive isolation between chromosomal forms (i.e. the 'speciation' hypothesis) or (ii) natural selection against certain heterokaryotypes (i.e. the 'selection' hypothesis), inversion being commonly viewed as underdominant mutations decreasing fitness of hybrids compared to homozygotes (Rieseberg, 2001). Neutral markers such as microsatellites could help address this issue. Selection on inversions is a locus-specific evolutionary force, and in this case, genetic differences should be observable only at these loci directly exposed to natural selection or at tightly linked loci, whereas genetic heterogeneity following speciation should be reflected throughout the genome. In *An. gambiae*, microsatellite loci located within polymorphic chromosomal inversions, revealed higher levels of genetic structuring than loci located outside inversions (Onyabe & Conn, 2001). This was observed even between sympatric populations belonging to alternative chromosomal forms of *An. gambiae* in Mali (Lanzaro *et al*., 1998; Taylor *et al*., 2001). Furthermore, significant linkage disequilibrium was detected between loci within chromosomal inversions (Lanzaro *et al*., 1998; Zheng *et al*., 1996), suggesting that recombination rate is lower within polymorphic chromosomal inversions in *An. gambiae*. In our study however, similar levels of genetic differentiation between *An. funestus* populations were detected for loci located within and outside polymorphic chromosomal inversions. Specimens from Kouvar, that were both karyotyped and genotyped, revealed no linkage disequilibrium between microsatellite loci and chromosomal inversions as well as among loci located within inversions. However, significant linkage disequilibrium was found between chromosomal inversions even though these were observed on different chromosomal arms.

In Dielmo and Sankagne as well, no evidence for departure from panmixia was observed using our set of microsatellite markers. The opposite was observed in Kedougou where microsatellite markers revealed population substructure, in spite of the fact that all specimens from this area were identified as belonging to the same chromosomal form, Folonzo (Dia *et al*., 2000; Lochouarn *et al*., 1998). This may reflect bias in our sampling design in this area, where overall very low human densities triggered several human settlements to be sampled in order to reach adequate sample sizes, representing a pool of potentially differentiated mosquitoes. Preference of *An. funestus* for breeding in patchily isolated permanent bodies of water, and the generally uneven distribution of this vector in Africa, indeed suggest discontinuous populations (Charlwood *et al*., 2000; Gillies

& Coetzee, 1987; Gillies & De Meillon, 1968; Molineaux & Gramiccia, 1980). Moreover, the isolation by distance found in the present study provides support for the usefulness of microsatellite markers in detecting population substructure in *An. funestus*.

In conclusion, our results suggest that microsatellite markers did not sustain the genetic structuring revealed by chromosomal markers, both within and among *An. funestus* populations in Senegal. These results are thus consistent with the selection hypothesis. It is noticeable however, that fine-scale analysis of the evolutionary dynamics of different parts of the *An. funestus* genome would require dense linkage maps while only few microsatellite markers are available to date. Boosted by recent advances in molecular entomology and the genomics of *An. gambiae* (Holt *et al.*, 2002), renewed interest of the scientific community for this 'neglected' malaria vector will undoubtedly result in the development and implementation of new molecular tools that will help shade light on the complex genetic structure of anopheline malaria vectors in Africa.

Experimental procedures

Study sites

The study was carried out in four villages along river Gambia from west to east Senegal (Fig. 1). Dielmo (13°45′N, 16°25′W) is located 280 km south-east of Dakar and approximately 15 km north of the Gambian border. The vegetation is a Sudan-type savannah. The rainy season extends from June to mid-October and the average annual rainfall is approximately 700 mm. The village is situated on the marshy bank of a small permanent stream which permits the persistence of anopheline larval development sites throughout the year. A population of 250 inhabitants lives in the village. Malaria transmission dynamics in Dielmo was previously described in detail (Fontenille *et al.*, 1997). The village of Sankagne (13°24′N, 13°45′W), is east of The Gambia, 2 km from river Gambia, in the Sudanian domain. The rainy season extends from June to October The population of Sankagne is 1600 inhabitants. Kouvar (13°23′N, 13°37′W), with 1000 inhabitants, is located 5 km east of Sankagne, 3 km from the river. A permanent pool is located 500 m from the village. In the region of Sankagne and Kouvar, the mean annual rainfall is approximately 900 mm. The Kedougou area (12°33′N, 12°11′W) is situated in the extreme south-east of Senegal in a Sudan-Guinean phytogeographical domain. The rainy season extends from June to November with an average annual rainfall of 1250 mm. Collections were carried out in small villages around Kedougou within a radius of 30 km. The human population density is fairly low, with 2.5 inhabitants per km^2 living in small dispersed agricultural villages.

Mosquito collection and field processing of specimens

Mosquitoes were collected by indoor pyrethrum spraying between November 1997 and October 2001. Females from the *An. funestus* group were visually sorted from other anophelines according to morphological identification keys (Gillies & Coetzee, 1987; Gillies & De Meillon, 1968). All specimens were stored individually in tubes containing desiccant. In Kouvar, half-gravid females of the *An. funestus* group were dissected and ovaries were immediately fixed in Carnoy's fixative (one part of glacial acetic acid and three parts of pure ethanol). The corresponding carcasses were stored in tubes with desiccant. Back in the laboratory, all tubes were stored at −20 °C until processing.

DNA extraction and species identification

Genomic DNA was extracted from wings and legs of each individual mosquito following a slightly modified version of the protocol of Cornel & Collins (1996) and resuspended in sterile water. *An. funestus s.s.* females were included in the analysis after species identification was carried out by the recently described diagnostic PCR assay (Cohuet *et al.*, 2003; Koekemoer *et al.*, 2002). No other member of the *An. funestus* group was observed in our samples.

Microsatellite amplification

Nine microsatellite loci were selected from published *An. funestus* sequence data (Cohuet *et al.*, 2002; Sinkins *et al.*, 2000), based on high polymorphism, no evidence for null alleles, mean allele sizes ranging from 110 to 220 bp, and cytological location allowing whole-genome scan (Fig. 2). We used AFND5 and FunO located on chromosomal arm 2R; FunL on chromosomal arm 2L; FunG, FunD, AFND19 and AFND20 on chromosomal arm 3R; FunF on chromosomal arm 3L; and AFND3, the precise cytological location of which is still undefined (Sharakhov *et al.*, 2003).

PCR amplification was carried out in 25 μl reaction volume, from 5 to 10 ng of template DNA. The reaction mixture contained 1 × PCR buffer containing 1.5 mM $MgCl_2$ (Qiagen, France), 200 μM of each dNTP, 10 pmol of each primer, and 0.5–1 U Taq polymerase (Qiagen, France). The forward primer was labelled in 5′ with either TET, HEX or FAM fluorescent markers (Eurogentec, Belgium) to allow multiplex electrophoresis. Amplification was performed under the following conditions: an initial denaturation step at 94 °C for 2 min followed by 36 cycles of 30 s at 94 °C, 30 s at 54 °C, 30 s at 72 °C and a final elongation step of 10 min at 72 °C. Fragment analyses were conducted with an ABI PRISM 377. Alleles were sized relatively to an internal size standard using GENESCAN version 3.1 (Applied Biosystems, France).

Polytene chromosomes preparation

Ovaries of half-gravid females were used to obtain squash preparations of the polytene chromosomes according to Hunt (1973). The preparations were examined under phase contrast microscopy and paracentric inversion karyotypes were scored according to the chromosomal map and nomenclature of Sharakhov *et al.*, 2003).

Data analysis

For each microsatellite locus, goodness-of-fit to Hardy–Weinberg expectations was tested in each location and overall. Cytogenetic data from Kouvar were analysed in the same way, considering alternative chromosomal arrangements as different alleles at one locus as previously defined (Costantini *et al.*, 1999; Lochouarn *et al.*, 1998). Statistical significance was assessed by the exact probability test available in GENEPOP 3.2 (Raymond & Rousset, 1995).

Linkage disequilibrium between pairs of microsatellite loci in the pooled population and within each population was assessed using

exact tests available in GENEPOP 3.2. In Kouvar, cytogenetic data (i.e. karyotypes) were included in the analysis.

Genetic differentiation between geographical populations was examined by *F*-statistics (Wright, 1951), calculated according to Weir & Cockerham (1984). Significance of F_{st} was assessed using the G-based exact test of genotypic differentiation (Goudet *et al*., 1996). The correlation between genetic and geographical distances was assessed by the regression of $F_{st}/(1 - F_{st})$ on the logarithm (ln) of geographical distance (Rousset, 1997), and tested using the Mantel test available in GENEPOP.

Global tests were employed to evaluate significance of multiple tests. Fisher's exact test was used to assess statistical significance of mean F_{is} and F_{st} estimates (across loci or across samples). The sequential Bonferroni procedure (Holm, 1979) can detect a single test-specific departure when multiple tests are performed: it was applied to evaluate significance of locus-specific departures within populations. Finally, the binomial test, which estimates the probability of obtaining the observed number of significant tests at the 0.05 level given the total number of tests, was used to explore genome-wide signatures.

Acknowledgements

This study was funded by the French ministry of research throughout the PAL + project. We thank F. Rousset for helpful comments. We would like to acknowledge the villagers for their cooperation throughout the survey.

References

Alvarez, G. and Zapata, C. (1997) Conditions for protected inversion polymorphism under supergene selection. *Genetics* **146**: 717–722.

Aultman, K.S., Beaty, B.J. and Walker, E.D. (2001) Genetically manipulated vectors of human disease: a practical overview. *Trends Parasitol* **17**: 507–509.

Boccolini, D., Carrara, G.C., Cani, P.J. and Costantini, C. (2002) Preliminary data on *Anopheles funestus* chromosomal polymorphism in a peri-urban site of Western Angola. *Parassitologia* **44**: 22.

Boccolini, D., Sabatini, A., Sanogo, E., Sagnon, N., Coluzzi, M. and Costantini, C. (1994) Chromosomal and vectorial heterogeneities in *Anopheles funestus* from Burkina Faso, West Africa. *Parassitologia* **36**: 20.

Boccolini, D., Sagnon, N. and Touré, Y.T. (1998) Chromosomal polymorphism in *Anopheles funestus* and description of new inversions in Burkina Faso and Mali. *Parassitologia* **40**: 14.

Caceres, M., Barbadilla, A. and Ruiz, A. (1999) Recombination rate predicts inversion size in Diptera. *Genetics* **153**: 251–259.

Charlwood, J.D., Vij, R. and Billingsley, P.F. (2000) Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of east Africa. *Am J Trop Med Hyg* **62**: 726–732.

Cohuet, A., Simard, F., Berthomieu, A., Raymond, M., Fontenille, D. and Weill, M. (2002) Isolation and characterization of microsatellite DNA markers in the malaria vector *Anopheles funestus*. *Mol Ecol Notes* **2**: 498–500.

Cohuet, A., Simard, F., Toto, J.C., Kengne, P., Coetzee, M. and Fontenille, D. (2003) Species identification within the *Anopheles funestus* group of malaria vectors in Cameroon and evidence for a new species. *Am J Trop Med Hyg* **69**: 200–205.

Collins, F.H. and Besansky, N.J. (1994) Vector biology and the control of malaria in Africa. *Science* **264**: 1874–1875.

Coluzzi, M. (1982) Spatial distribution of chromosomal inversions and speciation in Anopheline mosquitoes. In *Mechanisms of Speciation* (Barigozzi, C., ed.), pp. 143–153. Alan Liss Inc, New York.

Coluzzi, M., Sabatini, A., Petrarca, V. and Di Deco, M.A. (1979) Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans R Soc Trop Med Hyg* **73**: 483–497.

Cornel, A.J. and Collins, F.H. (1996) PCR of the ribosomal DNA intergenic spacer regions as a method for identifying mosquitoes in the *Anopheles gambiae* complex. *Methods Mol Biol* **50**: 321–332.

Costantini, C., Sagnon, N., Ilboudo-Sanogo, E., Coluzzi, M. and Boccolini, D. (1999) Chromosomal and bionomic heterogeneities suggest incipient speciation in *Anopheles funestus* from Burkina Faso. *Parassitologia* **41**: 595–611.

Curtis, C.F., Pates, H.V., Takken, W., Maxwell, C.A., Myamba, J., Priestman, A., Akinpelu, O., Yayo, A.M. and Hu, J.T. (1999) Biological problems with the replacement of a vector population by *Plasmodium*-refractory mosquitoes. *Parassitologia* **41**: 479–481.

Dia, I., Lochouarn, L., Boccolini, D., Costantini, C. and Fontenille, D. (2000) Spatial and temporal variations of the chromosomal inversion polymorphism of *Anopheles funestus*. Senegal. *Parasite* **7**: 179–184.

Fontenille, D. and Lochouarn, L. (1999) The complexity of the malaria vectorial system in Africa. *Parassitologia* **41**: 267–271.

Fontenille, D., Lochouarn, L., Diagne, N., Sokhna, C., Lemasson, J.J., Diatta, M., Konate, L., Faye, F., Rogier, C. and Trape, J.F. (1997) High annual and seasonal variations in malaria transmission by anophelines and vector species composition in Dielmo, a holoendemic area in Senegal. *Am J Trop Med Hyg* **56**: 247–253.

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

Gillies, M. and De Meillon, B. (1968) *The Anophelinae of Africa South of the Sahara (Ethiopian zoogeographical region)* 2nd edn. The South African Institute for Medical Research, Johannesburg.

Goudet, J., Raymond, M., de Meeus, T. and Rousset, F. (1996) Testing differentiation in diploid populations. *Genetics* **144**: 1933–1940.

Green, C.A. (1981) Malaria epidemiology and anophelines cytogenetics. In *Cytogenetics and Genetics of Vectors*, (Pal, R., Kitzmiller, J.B. and Kanda, T., eds), pp. 1–20. Elsevier, Amsterdam.

Green, C. and Hunt, R. (1980) Interpretation of variation in ovarian polytene chromosomes of *Anopheles funestus* Giles, *Anopheles parensis* Gillies and *Anopheles aruni? Genetica* **51**: 87–195.

Hackett, B.J., Gimnig, J., Guelbeogo, W., Costantini, C., Koekemoer, L.L., Coetzee, M., Collins, F.H. and Besansky, N.J. (2000) Ribosomal DNA internal transcribed spacer (ITS2) sequences differentiate *Anopheles funestus* and *An. rivulorum*, and uncover a cryptic taxon. *Insect Mol Biol* **9**: 369–374.

Holm, S. (1979) A simple sequential rejective multiple test procedure. *Scand J Stat* **6**: 65–70.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M. and Wides, R. (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**: 129–149.

Hunt, R.H. (1973) A cytological technique for the study of *Anopheles gambiae* complex. *Parassitologia* **15**: 137–139.

Kamau, L., Munyekenye, G.O., Koekemoer, L.L., Hunt, R.H. and Coetzee, M. (2003) A survey of the *Anopheles funestus* (Diptera: Culicidae) group from 10 sites in Kenya with special emphasis on population genetic structure based on chromosomal inversion karyotypes. *J Med Entomol* **40**: 664–671.

Koekemoer, L.L., Kamau, L., Hunt, R.H. and Coetzee, M. (2002) A cocktail polymerase chain reaction assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *Am J Trop Med Hyg* **66**: 804–811.

Lanzaro, G.C., Toure, Y.T., Carnahan, J., Zheng, L., Dolo, G., Traore, S., Petrarca, V., Vernick, K.D. and Taylor, C.E. (1998) Complexities in the genetic structure of *Anopheles gambiae* populations in west Africa as revealed by microsatellite DNA analysis. *Proc Natl Acad Sci USA* **95**: 14260–14265.

Lochouarn, L., Dia, I., Boccolini, D., Coluzzi, M. and Fontenille, D. (1998) Bionomical and cytogenetic heterogeneities of *Anopheles funestus* in Senegal. *Trans R Soc Trop Med Hyg* **92**: 607–612.

Manga, L., Toto, J.C., Le Goff, G. and Brunhes, J. (1997) The bionomics of *Anopheles funestus* and its role in malaria transmission in a forested area of southern Cameroon. *Trans R Soc Trop Med Hyg* **91**: 387–388.

Mendis, C., Jacobsen, J.L., Gamage-Mendis, A., Bule, E., Dgedge, M., Thompson, R., Cuamba, N., Barreto, J., Begtrup, K., Sinden, R.E. and Hogh, B. (2000) *Anopheles arabiensis* and *An. funestus* are equally important vectors of malaria in Matola coastal suburb of Maputo, southern Mozambique. *Med Vet Entomol* **14**: 171–180.

Molineaux, L. and Gramiccia, G. (1980) *The Garki Project. Research on Epidemiology and Control of Malaria in the Sudan Savannah of West Africa.* World Health Org, Geneva.

Navarro, A., Betran, E., Barbadilla, A. and Ruiz, A. (1997) Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**: 695–709.

Onyabe, D.Y. and Conn, J.E. (2001) Genetic differentiation of the malaria vector *Anopheles gambiae* across Nigeria suggests that selection limits gene flow. *Heredity* **87**: 647–658.

Raymond, M. and Rousset, F. (1995) GENEPOP, Version 1.2. A population genetics sofware for exact tests and ecumenicism. *J Hered* **26**: 248–249.

Rieseberg, L.H. (2001) Chromosomal rearrangements and speciation. *Trends Ecol*, **16**: 351–358.

Rousset, F. (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**: 1219–1228.

Sharakhov, I.V., Braginets, O., Grushko, O.G., Cohuet, A., Guelbeogo, W., Boccolini, D., Weill, M., Sagnon, N., Fontenille, D., Yan, G. and Besansky, N.J. (2003) A microsatellite physical map of the African human malaria vector *Anopheles funestus*. *J Hered* **95**: 29–34.

Sharakhov, I.V., Sharakhova, M.V., Mbogo, C.M., Koekemoer, L.L. and Yan, G. (2001) Linear and spatial organization of polytene chromosomes of the African malaria mosquito *Anopheles funestus*. *Genetics* **159**: 211–218.

Sinkins, S.P., Hackett, B.J., Costantini, C., Vulule, J., Ling, Y.Y., Collins, F.H. and Besansky, N.J. (2000) Isolation of polymorphic microsatellite loci from the malaria vector *Anopheles funestus*. *Mol Ecol* **9**: 490–492.

Taylor, C., Toure, Y.T., Carnahan, J., Norris, D.E., Dolo, G., Traore, S.F., Edillo, F.E. and Lanzaro, G.C. (2001) Gene flow among populations of the malaria vector, *Anopheles gambiae*, in Mali, West Africa. *Genetics* **157**: 743–750.

Weir, B. and Cockerham, C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Wilkes, T.J., Matola, Y.G. and Charlwood, J.D. (1996) *Anopheles rivulorum*, a vector of human malaria in Africa. *Med Vet Entomol* **10**: 108–110.

Wright, S. (1951) The genetical structure of populations. *Ann Eugen* **15**: 323–354.

Zheng, L., Benedict, M.Q., Cornel, A.J., Collins, F.H. and Kafatos, F.C. (1996) An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* **143**: 941–952.