

Divergence With Gene Flow in *Anopheles funestus* From the Sudan Savanna of Burkina Faso, West Africa

Andrew P. Michel,* Olga Grushko,* Wamdaogo M. Guelbeogo,[†] Neil F. Lobo,* N'Fale Sagnon,[‡] Carlo Costantini[‡] and Nora J. Besansky*¹

*Center for Tropical Disease Research and Training, Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, 46556,

[†]Centre National de Recherche et de Formation sur le Paludisme, Ouagadougou 01, Burkina Faso and

[‡]Institut de Recherche pour le Développement, Ouagadougou 01, Burkina Faso

Manuscript received April 19, 2006

Accepted for publication April 26, 2006

ABSTRACT

Anopheles funestus is a major vector of malaria across Africa. Understanding its complex and non-equilibrium population genetic structure is an important challenge that must be overcome before vector populations can be successfully perturbed for malaria control. Here we examine the role of chromosomal inversions in structuring genetic variation and facilitating divergence in Burkina Faso, West Africa, where two incipient species (chromosomal forms) of *A. funestus*, defined principally by rearrangements of chromosome 3R, have been hypothesized. Sampling across an ~300-km east–west transect largely contained within the Sudan–Savanna ecoclimatic zone, we analyzed chromosomal inversions, 16 microsatellite loci distributed genomewide, and 834 bp of the mtDNA *ND5* gene. Both molecular markers revealed high genetic diversity, nearly all of which was accounted for by within-population differences among individuals, owing to recent population expansion. Across the study area there was no correlation between genetic and geographic distance. Significant genetic differentiation found between chromosomal forms on the basis of microsatellites was not genomewide but could be explained by chromosome 3R alone on the basis of loci inside and near inversions. These data are not compatible with complete reproductive isolation but are consistent with differential introgression and sympatric divergence between the chromosomal forms, facilitated by chromosome 3R inversions.

MALARIA continues to claim over 1 million lives annually and remains the leading pathogen-specific cause of morbidity and mortality in children from tropical Africa (WORLD HEALTH ORGANIZATION 2003). Among its principal African vectors are two highly anthropophilic mosquito species, *Anopheles gambiae* and *A. funestus*, whose population genetic structure shows remarkable parallels (LEHMANN *et al.* 2003; MICHEL *et al.* 2005b). The tendency of both species to specialize on humans not only maintains a high malaria transmission intensity but also has shaped their population history. This is evidenced by very shallow population structure across Africa and signs of recent population expansion in response to human population growth triggered by the rise and spread of agriculture. The shallow and nonequilibrium population genetic structure of these mosquitoes represents a major challenge to understanding contemporary rates and patterns of gene flow but one that must be overcome if vector populations are to be successfully perturbed for malaria control (COETZEE and FONTENILLE 2004).

Against the backdrop of weak population genetic structure inferred from neutral molecular markers, nonrandom patterns of chromosomal inversion polymorphism in West Africa suggest an ongoing speciation process in both vectors, associated with environmental changes, including irrigation schemes for rice cultivation (DELLA TORRE *et al.* 2002). Although this phenomenon has been most extensively studied in *A. gambiae* (COLUZZI *et al.* 2002), karyotypic analysis of >5,000 *A. funestus* specimens also has revealed significant and stable departures from Hardy–Weinberg equilibrium due to heterokaryotype deficit and linkage disequilibrium between arrangements on independently assorting chromosome arms (COSTANTINI *et al.* 1999; GUELBEOGO *et al.* 2005). Equilibrium was restored under the hypothesis of two sympatric and morphologically indistinguishable but assortatively mating chromosomal forms designated Folonzo and Kiribina. Corresponding ecological and behavioral differences suggest that Kiribina is associated with rice cultivation and is less likely than Folonzo to rest indoors, feed on humans, or carry malaria parasites (COSTANTINI *et al.* 1999).

Owing to their role in suppressing recombination between alternative arrangements, chromosomal inversions may facilitate divergence in the face of gene flow between incompletely isolated populations (NOOR *et al.*

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. DQ126687–127169.

¹Corresponding author: 317 Life Sciences Bldg., P.O. Box 369, University of Notre Dame, Notre Dame, IN, 46556. E-mail: nbesansk@nd.edu

2001; RIESEBERG 2001; AYALA and COLUZZI 2005). Whereas gene flow would be expected to homogenize homosequential regions of the genome, gene exchange is restricted between rearranged regions, allowing differences to accumulate and leading to the prediction of greater divergence in such regions. To date, molecular evidence supporting this prediction for *A. funestus* has been contradictory between recent studies in Burkina Faso and Cameroon (COHUET *et al.* 2005; MICHEL *et al.* 2005a). Here, we examine this problem on a wider geographic scale in Burkina Faso, where Folonzo and Kiribina forms are sympatric. Using chromosomal inversions, 16 microsatellite markers, and mitochondrial DNA sequence, we assessed the relative impact of inversions and geographic distance on the partitioning of genetic variation.

MATERIALS AND METHODS

Mosquito sampling and processing: Villages, the spatial sampling unit used in this study, are 1–3 km in radius consisting of 20- to 50-m² compounds, each containing 2–13 small, closely spaced huts. Collections were made inside multiple huts and compounds per village by pyrethrum spray catches. Eight villages were sampled along an east–west transect in Burkina Faso during November and December 2002 (Figure 1). The villages of Vi (11°45'N, 3°8'W), Siby (11°51'N, 2°58'W), Ividie (11°49'N, 2°38'W), La (12°4'N, 2°18'W), Pehele (12°15'N, 1°11'W), and Sabtenga (11°48'N, 0°30'W) are all located within the Sudan–Savanna ecoclimatic zone; Bagre (11°32'N, 0°28'W) and Dirze (11°24'N, 0°36'W) lie on or just inside the border with the Guinean–Savanna zone. The area sampled is characterized by annual rainfall amounts of 550–900 mm and large flat expanses of tall grasses or shrubs with scattered trees.

Mosquitoes were sorted morphologically to the *Funestus* Group (GILLIES and COETZEE 1987) in the field under a dissecting scope. Ovaries from females at the appropriate gonotrophic stage were dissected in the field into individually labeled tubes containing modified Carnoy's solution (ethanol:glacial acetic acid, 3:1) and held on ice until they could be stored at –20° for later polytene chromosome analysis. Mosquito carcasses were stored individually at room temperature in correspondingly labeled tubes with desiccant. Chromosome preparation followed SHARAKHOV *et al.* (2001). Assignment of individual females to the Folonzo or Kiribina form was performed on the basis of karyotype, according to GUELBEOGO *et al.* (2005). Only those specimens that could be unambiguously assigned were analyzed further. After removal of abdomens for later blood meal analysis (results will be presented in a future article), genomic DNA was extracted from single mosquito carcasses in 96-well plates (Wizard SV-96 genomic DNA purification system, Promega) with the aid of a Biomek FX workstation (Beckman Coulter). Prior to analysis by PCR, genomic DNA was diluted 1:10 in H₂O (~5 ng/μl). Identification of *A. funestus sensu stricto* apart from closely related species with similar or identical adult morphology was achieved by an rDNA-based PCR assay (KOEKEMOER *et al.* 2002) as modified by MICHEL *et al.* (2005a). Sporozoite infection was assayed by nested PCR targeting the *MspI* gene (RANFORD-CARTWRIGHT *et al.* 1993; MICHEL *et al.* 2005a).

mtDNA sequencing and microsatellite genotyping: A segment of the mtDNA *ND5* gene was amplified by PCR and sequenced directly on an ABI-3700 (Applied Biosystems) as

described previously (MICHEL *et al.* 2005a). Sequences were aligned, trimmed to a common 834-bp segment, and edited using Lasergene software (DNASTAR). Sequences have been deposited under GenBank accession nos. DQ126687–127169. Sixteen physically mapped reference microsatellite loci (SHARAKHOV *et al.* 2004) were amplified and products were diluted, pool-plexed (two groups of eight loci each), and genotyped on a Beckman Coulter CEQ8000 as described previously (MICHEL *et al.* 2005a).

Genetic data analysis: For each village and chromosomal form, summary statistics of mtDNA and microsatellite polymorphism were generated using DnaSP 4.0 (ROZAS *et al.* 2003) and Fstat2.9.3.2 (GOUDET 2001), respectively. Fstat2.9.3.2 was also used to assess deviations from Hardy–Weinberg equilibrium per locus and per village/chromosomal form (by estimating the inbreeding coefficient F_{is}) and linkage disequilibrium between pairs of loci within each village/chromosomal form. Significance was tested using the randomization approach implemented in Fstat, which applies Bonferroni corrections. Within each village the frequency of suspected null alleles was calculated using the Brookfield 2 estimate (BROOKFIELD 1996). Allele and genotype frequencies were modified accordingly in MICRO-CHECKER (VAN OOSTERHOUT *et al.* 2004) to create a null allele-adjusted data set that was compared to the original data set to explore the impact of null alleles on estimates of genetic differentiation.

Genetic differentiation between pairs of population samples within and between chromosomal forms was estimated using F_{ST} . For mtDNA, F_{ST} based on pairwise nucleotide difference was calculated in Arlequin 2.0 (SCHNEIDER *et al.* 2000). For microsatellites, F_{ST} was computed in Microsatellite Analyzer (DIERINGER and SCHLOTTERER 2003). Global estimates of F_{ST} for both markers were generated by analysis of molecular variance (AMOVA) implemented in Arlequin 2.0. For all calculations, significance was assessed by 10,000 permutations and reported *P*-values were Bonferroni adjusted.

The isolation-by-distance model was tested in GENEPOP 3.3 (RAYMOND and ROUSSET 1995) using a semi-matrix of F_{ST} values based on mtDNA or microsatellite data sets. Significance of the regression of genetic distance on geographic distance was tested using the Mantel procedure (MANTEL 1967), permuting columns of the semi-matrix 10,000 times. For microsatellites we also used the program SPAIDA (PALSSON 2004), which calculates two estimates of spatial autocorrelation on the basis of frequencies of different alleles assuming an infinite allele model (Moran's *I*) or differences in the numbers of repeats (Geary's *C*) assuming a stepwise mutation model of microsatellite evolution. Significance of the correlation for each distance class was evaluated by 1000 permutations of individuals among locations.

Three approaches were followed to investigate whether populations cluster by geographic proximity or chromosomal form. In the first, we reconstructed relationships on the basis of microsatellite or mtDNA genetic distances. For microsatellites, NEI *et al.*'s (1983) D_A distance was used to construct a neighbor-joining tree with Populations 1.2.28 software (<http://www.cnrs-gif.fr/pge/bioinfo/populations/index.php>). Bootstrapping over individuals 100 times was used to assess robustness of individual branches. For mtDNA, a haplotype network was constructed using the statistical parsimony method implemented in TCS 1.13 (CLEMENT *et al.* 2000). The second approach, applied only with microsatellites, used the model-based clustering method implemented in STRUCTURE 2.0 (PRITCHARD *et al.* 2000), which estimates the most likely number of distinct populations on the basis of multilocus genotypes with *no a priori* assumptions of population structure. Each simulation consisted of a burn-in period of 500,000 chains and a run length of 1,000,000 chains for each of $K = 1–8$, replicated

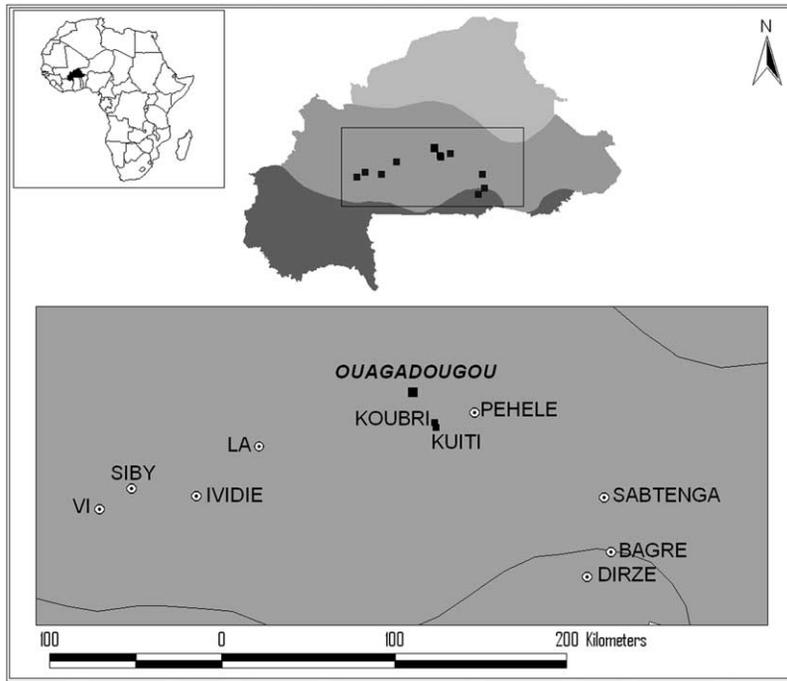


FIGURE 1.—Map of Burkina Faso showing the capital Ouagadougou, the villages of Koubri and Kuiti sampled in a previous study (MICHEL *et al.* 2005a), and the villages where mosquito collections were made for the present study. Three climatic regions indicated by different shading (from north to south: Sahel, Sudan-Savanna, and Sudan-Guinea) are delimited by the 600- and 900-mm isohyets.

three times to assess consistency. Additionally, assignment tests using microsatellite data were performed with the program SPASSIGN (PALSSON 2004) using a frequency-based method developed by PAETKAU *et al.* (1995) to calculate the probability of assignment to the Kiribina or Folonzo chromosomal forms.

Tests of non-neutral evolution were performed for microsatellite and mtDNA data sets. For mtDNA, Tajima's D (TAJIMA 1989), R_2 (RAMOS-ONSINS and ROZAS 2002), and Fu's F_s (FU 1997) were computed using DnaSP. The first two statistics are based on the distribution of mutation frequencies, whereas the latter is based on haplotype distribution; all have low values where low-frequency mutations are in excess, as expected following a selective sweep or a population expansion. Statistical significance was evaluated by coalescent simulations (10,000 replicates) in DnaSP. For microsatellites, KIMMEL *et al.*'s (1998) β -imbalance index was computed using the β_1 estimator of $\beta(t)$ (KING *et al.* 2000). This statistic is based on the imbalance between allele size variance and heterozygosity at a locus, and it assumes values <1 following a population expansion. For the imbalance index, 95% confidence intervals were estimated by bootstrapping over loci using an SAS program kindly provided by T. Lehmann (DONNELLY *et al.* 2001).

RESULTS

During November and December 2002 we collected indoor-resting *A. funestus* from eight villages spanning an ~300-km east-west transect across Burkina Faso (Figure 1). Although both Folonzo and Kiribina forms were collected from each village, their relative proportions were unbalanced such that in only two locales (La and Pehele villages) were adequate sample sizes (*i.e.*, >30) obtained for both forms in the same collection.

mtDNA and microsatellite polymorphism: Levels of polymorphism for both marker classes were high (supplemental Table S1 at <http://www.genetics.org/supplemental/>). In the set of 483 mtDNA *ND5* sequen-

ces (302 Folonzo, 181 Kiribina), 180 of 834 sites were variable. Nearly half (80) of the variable sites were singletons, resulting in a large total number of haplotypes (280) and a high haplotype diversity (0.98). Average nucleotide diversity overall was only 0.006 per site, or 4.6 nucleotide differences between sequences. Thus, the total mtDNA data set is composed of closely related yet often unique sequences, as reflected in the greater estimates of $N_e\mu$ based on θ *vs.* π (supplemental Table S1, <http://www.genetics.org/supplemental/>).

Microsatellite polymorphism levels were high at all 16 loci genotyped from 502 *A. funestus* (315 Folonzo, 187 Kiribina). Overall allelic richness averaged 9.0 per locus with a range of 3–20. Values of observed heterozygosity ranged from 0.26–0.89. Deviations from Hardy-Weinberg equilibrium within some samples of Kiribina and Folonzo (19 of 160 tests) indicated heterozygote deficit (supplemental Table S1, <http://www.genetics.org/supplemental/>), most likely due to null alleles as they were not genomewide, consistent with the presence of null alleles in microsatellite studies of this and other species (LEHMANN *et al.* 1996; DAKIN and AVISE 2004; MICHEL *et al.* 2005a; STUMP *et al.* 2005). Notably, all population samples showed heterozygote deficits at locus FunD, and three more showed significant deficits at AFND12 and AFUB12, together accounting for 16 of 19 significant tests. Under the assumption of random mating, we estimated null allele frequencies for each locus and population and explored their impact by repeating all population genetic analyses under two different treatments: (1) with a data set adjusted for estimated null allele frequencies, or (2) by excluding FunD, AFND12, and AFUB12. The effect of either treatment was minimal and not statistically significant.

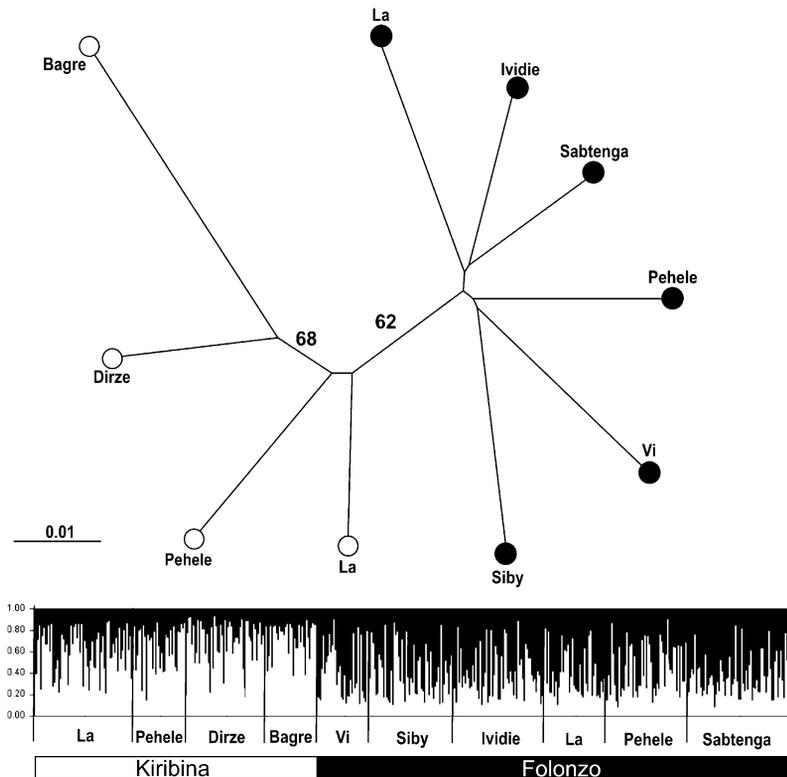


FIGURE 2.—Clustering of *A. funestus* study populations from Burkina Faso, inferred from microsatellites. Top: Neighbor-joining tree based on NEI *et al.*'s (1983) D_A distance. Bootstrap values $>50\%$ are shown. Branches are labeled by sampling location (village) and chromosomal form assignment (solid circles, Folonzo; open circles, Kiribina). Bottom: Plot from an unsupervised STRUCTURE run of highest estimated probability at $k = 2$ (two clusters). Each individual is represented by a vertical line in which shading (solid or open) indicates membership coefficients (scale bar at left) for alternative clusters corresponding to Folonzo and Kiribina. Plot is partitioned into 10 segments representing samples of Kiribina or Folonzo from different villages. Village names are indicated below the plot, and chromosomal form membership (determined independently by karyotyping) is indicated at the bottom by a bar.

Genetic clustering: Neighbor-joining trees were reconstructed to explore whether *A. funestus* would cluster on the basis of physical proximity of sampling sites or chromosomal form membership. Mitochondrial sequences were uninformative, resulting in a star-like phylogeny (data not shown). However, microsatellite D_A distances clustered by chromosomal form (Figure 2). Notably, samples of both chromosomal forms collected in sympatry from the villages of La and Pehele were more closely related to samples of the same form separated by ~ 120 km than to sympatric samples of the other form.

The same pattern was revealed by unsupervised Bayesian clustering analysis of microsatellite data (Figure 2). Two clusters were inferred as most likely corresponding to the chromosomal forms. Predefined Kiribina and Folonzo samples were consistently different in their proportion of membership to cluster 1 (mean, 0.67 ± 0.05 and 0.40 ± 0.02 for Kiribina and Folonzo, respectively) or cluster 2 (0.33 ± 0.05 and 0.60 ± 0.02 for Kiribina and Folonzo). Assignment tests based on microsatellites showed similar results, correctly allocating 81% of Kiribina specimens and 74% of Folonzo specimens. The proportion of specimens correctly assigned to chromosomal form was significantly better than random ($X^2 = 143.7$, $P < 0.001$). By contrast, the proportion of Kiribina or Folonzo specimens collected from La and Pehele villages that were correctly assigned to those villages was not significantly different from random ($X^2 = 1.48$, $P > 0.1$ and $X^2 = 2.64$, $P > 0.1$, respectively).

Population structure: The genetic structure of these *A. funestus* populations was investigated using a hierar-

chical AMOVA approach based on mtDNA and microsatellites. We defined three hierarchical levels: within populations, among populations within chromosomal forms, and between forms. Results for both types of markers were virtually identical and in accord with the clustering analyses. Almost all genetic diversity (99%) was partitioned within populations, indicating high genetic similarity overall. Global F_{ST} estimates revealed slight but significant overall genetic structure (mtDNA: 0.009, $P = 0.002$; microsatellites: 0.010, $P = 0.000$). However, most AMOVA tests of population structure within forms (*i.e.*, two hierarchical levels: within populations and among populations within forms) were not significant (Kiribina mtDNA $F_{ST} = 0.002$, $P = 0.32$ and microsatellite $F_{ST} = 0.003$, $P = 0.06$; Folonzo mtDNA $F_{ST} = 0.003$, $P = 0.11$ and microsatellite $F_{ST} = 0.004$, $P = 0.003$).

Within chromosomal forms, pairwise F_{ST} comparisons among samples gave a range of values from 0 to 0.006 for microsatellites and 0 to 0.020 for mtDNA (supplemental Table S2 at <http://www.genetics.org/supplemental/>). Of 40 comparisons in total, only 7 revealed significant values after correction for multiple tests. Two of the largest such values, corresponding to mtDNA contrasts between Siby and either La or Ividie, involved a trio of nearby villages, whereas comparisons across much larger distances (*e.g.*, Ividie *vs.* Sabtenga) were not statistically different from zero. These results are inconsistent with positively correlated geographic and genetic distances and instead may reflect heterogeneous sampling of rare alleles, which collectively compose a substantial fraction

of allelic variation. An excess of rare mtDNA alleles was signaled by significantly negative Tajima's D (see Table 2); singletons alone accounted for 44% of mtDNA variable sites. For microsatellites, rare alleles (defined here as having frequencies <2.5%) accounted for 43% of alleles at a locus in the Kiribina form (range across 16 loci, 20–62%) and 52% of alleles at a locus in Folonzo (33–65%). The proportion of rare microsatellite alleles in the earlier study from Koubri/Kuiti (MICHEL *et al.* 2005a) was similarly high in both forms (54% of alleles from ~380 Kiribina individuals, 41% of alleles from ~180 Folonzo individuals).

Formal tests of isolation by distance were conducted separately for each chromosomal form and marker class. For Folonzo, tests included the villages of Vi, Siby, Ividie, La, Pehele, and Sabtenga, spanning pairwise distances of 20–280 km. For Kiribina, the tests included La, Pehele, Dirze, and Bagre (19–200 km; see Figure 1). In no case did the data fit the isolation-by-distance model based on Mantel tests (mtDNA: Kiribina $P > 0.08$, Folonzo $P > 0.95$; microsatellites: Kiribina $P > 0.08$; Folonzo $P > 0.54$). In addition, using the SPAIDA program to calculate two estimates of spatial autocorrelation extended for microsatellites (PALSSON 2004), we found no significant correlations with either Moran's I under the infinite allele model or Geary's C under the stepwise mutation model. Therefore, for distances <300 km the data do not support an isolation-by-distance model for *A. funestus*. It is instructive to note that Mantel tests of isolation by distance using microsatellite data from each village without regard to chromosomal form membership of the samples produced a potentially misleading, marginally significant correlation ($P < 0.053$). This emphasizes the need for karyotype information on all specimens and the importance of separate analyses of Folonzo and Kiribina forms of *A. funestus*.

For comparisons between chromosomal forms, the analysis was limited to samples in which sufficient numbers of both forms were captured simultaneously from the same locale (La and Pehele villages). The results are given in Table 1, which also provides data from a previous study in Koubri/Kuiti (see Figure 1). The microsatellite data were consistent across space (among villages) and time (among years) and gave evidence of

TABLE 1
Microsatellite and mtDNA differentiation (F_{ST}) between *A. funestus* chromosomal forms in Burkina Faso

Data set	Sampling location		
	La	Pehele	Koubri/Kuiti ^a
Microsatellites			
All loci (16)	0.004*	0.006**	0.004***
3R loci (5)	0.013***	0.015***	0.010***
Non-3R loci (11)	0	0	0.001 ^{NS}
mtDNA	0.002 ^{NS}	0.004 ^{NS}	0.023**

NS, $P > 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

^aFrom MICHEL *et al.* (2005a).

significant differentiation between forms, although the magnitude of differentiation was slight. Notably, microsatellite differentiation was not genomewide, as shown by independent analyses of the five loci on chromosome 3R and the 11 loci outside of this chromosome arm. The right arm of chromosome 3 contains inversions 3R*a* and 3R*b*, whose arrangement is crucial in chromosomal form assignment. All significant differentiation between forms (and none within forms) could be explained by the five 3R loci alone: three outside of inversions, two within inversion 3R*b*. The lack of microsatellite differentiation outside of 3R is consistent with slight or nonexistent mtDNA differentiation.

Demographic inference: Typical of villages across the transect, an excess of low-frequency mtDNA alleles in Folonzo and Kiribina samples from La and Pehele villages was indicated by each of three tests on the basis of the frequencies of segregating sites or the haplotype distribution (Table 2). These results are consistent with alternative scenarios—a recent selective sweep or population expansion—both of which are characterized by an elevated proportion of recent mutations occurring on the external branches of the genealogy. Supporting evidence from 16 microsatellite loci, on the basis of a β -imbalance index significantly <1 in both Kiribina and Folonzo from La and Pehele, favors population expansion as the more likely explanation (KIMMEL *et al.* 1998; KING *et al.* 2000) because the effect is genomewide. This index was significantly <1 in virtually all individual

TABLE 2
Summary of neutrality tests against population growth

	Mitochondrial DNA			Microsatellites
	Tajima's D	R_2	Fu's F_s	β_1 (95% CI) ^a
Kiribina	-2.23**	0.029***	-59.104***	0.63 (0.76–0.49)**
Folonzo	-2.39**	0.024***	-109.837***	0.70 (0.98–0.48)*

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

^a95% confidence interval based on bootstrapped β_1 values.

samples (only Folonzo from Pehele was marginally significant; data not shown).

DISCUSSION

As in Burkina Faso, chromosomal polymorphism in Cameroon appears consistent with the presence of Kiribina and Folonzo forms, given heterokaryotype deficits and linkage disequilibria between inversions in central localities of that country (COHUET *et al.* 2005). However, at odds with the present study, molecular data from 10 microsatellite loci—of which 7 were shared between studies—were not consistent with chromosomal discontinuities (COHUET *et al.* 2005). Instead, the presence of significant population structure among localities was attributed not to genetic isolation between the forms but to an isolation-by-distance effect that encompassed all loci both inside and outside of inversions. The discrepancy between studies may reflect a true difference between West and Central African populations of *A. funestus*. However, as only 25% of the ~600 Cameroon specimens analyzed molecularly also were karyotyped, individuals were not identified as Folonzo or Kiribina and the data were not partitioned by chromosomal form prior to analysis. In light of the nonrandom distribution of chromosomal inversions along a north–south cline of decreasing aridity in Cameroon, the significant correlation of genetic and geographic distance may have been driven by hidden population structure due to displacement of the chromosomal forms along the cline. Moreover, the seeming contradiction regarding the influence of chromosomal inversions on estimates of genetic differentiation also may be a function of different methodology rather than biology. Whereas COHUET *et al.* (2005) partitioned their loci strictly according to whether they fell inside or outside inversions, we found that loci on chromosome 3R (whether within or near inversions) accounted for all significant differentiation. Viewed in this way, the data from Cameroon are more congruent with the present study; their mean F_{ST} estimate for all 4 loci from chromosome 3R (FunG, FunD, AFND19, and AFND20) was 0.016, twofold higher than the estimate across the 4 remaining loci whose genome position has been mapped (0.008).

Under a model of allopatric speciation (or complete reproductive isolation), it is expected that genetic divergence should be found genomewide and in proportion to time since isolation. The present data are not compatible with this model. Yet the absence of genomewide divergence does not contradict an ongoing process of lineage splitting. Divergence with gene flow models in which incipient species experience residual gene flow predict a mosaic pattern of undifferentiated genome segments homogenized by recombination interspersed with islands of differentiation due to strong selection against introgression (MACHADO *et al.* 2002; WU and TING 2004; FEDER *et al.* 2005). These islands are likely to

coincide with low-recombination regions of the genome, including polymorphic chromosomal inversions and pericentromeric areas, and are likely to contain many of the genes underlying differential ecological adaptations and premating isolation (NOOR *et al.* 2001; ORTIZ-BARRIENTOS *et al.* 2002). The African malaria vector *A. gambiae* appears to represent this situation very well, with nascent species M and S in sympatry across West and Central Africa (STUMP *et al.* 2005; TURNER *et al.* 2005). In the case of Folonzo and Kiribina forms of *A. funestus*, alternative arrangements on chromosome 3R are core components of a deterministic algorithm that makes taxonomic assignments on the basis of karyotype (GUELBEOGO *et al.* 2005); no independent molecular markers have been discovered. Such an algorithm has limitations. It should be emphasized that no inversions or inversion combinations are strictly diagnostic for different forms, neither for *A. funestus* nor for *A. gambiae*. Inversions are known to be shared across forms, albeit at significantly different frequencies. Nevertheless, their importance in partitioning genetic variation in *A. funestus* is underscored by the finding that all significant nuclear differentiation derives from loci within or near inversions on chromosome 3R. Neither candidate genes nor candidate phenotypes exist that can be mapped to this chromosome arm at present, although tantalizing evidence suggests that the inversions may be linked to alternative utilization of anthropogenic (ricefield) *vs.* natural (*e.g.*, marsh) habitats (COSTANTINI *et al.* 1999), to which they may be differentially adapted. The present data do not support a link between karyotype and rate of malaria parasite infection, as sporozoite rates were uniformly high in both forms (8.8% and 9.1% for Kiribina and Folonzo, respectively).

As is the case for *A. gambiae* and our own species, most genetic diversity in *A. funestus* (99%) is accounted for by within-population differences among individuals, owing to recent population expansion. The implications are twofold. First, the assumption of mutation-drift equilibrium so common to analyses of population genetic structure and especially to indirect estimates of contemporary gene flow is not valid for either vector. Migration rates among demes within a chromosomal form, or gene flow between chromosomal forms, might be much smaller than what a simple transformation of F_{ST} to N_m would suggest (WHITLOCK and McCAULEY 1999). Second, it suggests a shared history that follows from the interplay between human population expansion, associated anthropogenic modifications to the environment, and the correlated rise of pest species with the infectious diseases they transmit. The parallel instances of incipient speciation in West Africa by these major malaria vectors, both associated with rice cultivation, may offer mechanistic insights into the rapid ecological adaptation and lineage splitting that seem to occur in response to common environmental cues.

We thank the entomological technicians at the Centre National de Recherche et de Formation sur le Paludisme for assistance with field collections. This project was funded by National Institute of Health grant R01-AI48842 to N.J.B. A.P.M was supported by an Arthur J. Schmidt Graduate Fellowship from the University of Notre Dame.

LITERATURE CITED

- AYALA, F. J., and M. COLUZZI, 2005 Chromosome speciation: humans, *Drosophila*, and mosquitoes. *Proc. Natl. Acad. Sci. USA* **102**(Suppl. 1): 6535–6542.
- BROOKFIELD, J. F., 1996 A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* **5**: 453–455.
- CLEMENT, M., D. POSADA and K. A. CRANDALL, 2000 TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**: 1657–1659.
- COETZEE, M., and D. FONTENILLE, 2004 Advances in the study of *Anopheles funestus*, a major vector of malaria in Africa. *Insect Biochem. Mol. Biol.* **34**: 599–605.
- COHUET, A., I. DIA, F. SIMARD, M. RAYMOND, F. ROUSSET *et al.*, 2005 Gene flow between chromosomal forms of the malaria vector *Anopheles funestus* in Cameroon, Central Africa and its relevance in malaria fighting. *Genetics* **169**: 301–311.
- COLUZZI, M., A. SABATINI, A. DELLA TORRE, M. A. DI DECO and V. PETRARCA, 2002 A polyploid chromosome analysis of the *Anopheles gambiae* species complex. *Science* **298**: 1415–1418.
- COSTANTINI, C., N. F. SAGNON, E. ILBOUDO-SANOGO, M. COLUZZI and D. BOCCOLINI, 1999 Chromosomal and bionomic heterogeneities suggest incipient speciation in *Anopheles funestus* from Burkina Faso. *Parassitologia* **41**: 595–611.
- DAKIN, E. E., and J. C. AVISE, 2004 Microsatellite null alleles in parentage analysis. *Heredity* **95**: 504–509.
- DELLA TORRE, A., C. COSTANTINI, N. J. BESANSKY, A. CACCONE, V. PETRARCA *et al.*, 2002 Speciation within *Anopheles gambiae*—the glass is half full. *Science* **298**: 115–117.
- DIERINGER, D., and C. SCHLOTTERER, 2003 Microsatellite analyzer (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* **3**: 167–169.
- DONNELLY, M. J., M. C. LIGHT and T. LEHMANN, 2001 Evidence for recent population expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae*. *Mol. Biol. Evol.* **18**: 1353–1364.
- FEDER, J. L., X. XIE, J. RULL, S. VELEZ, A. FORBES *et al.*, 2005 Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. *Proc. Natl. Acad. Sci. USA* **102**(Suppl. 1): 6573–6580.
- FU, Y. X., 1997 Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.
- GILLIES, M. T., and M. COETZEE, 1987 *A Supplement to the Anophelinae of Africa South of the Sahara*. The South African Institute for Medical Research, Johannesburg.
- GOUDET, J., 2001 *FSTAT: A Program To Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3)* (<http://www.unil.ch/izea/software/fstat.html>).
- GUELBEOGO, W. M., O. GRUSHKO, D. BOCCOLINI, P. A. OUEDRAOGO, N. J. BESANSKY *et al.*, 2005 Chromosomal evidence of incipient speciation in the Afrotropical malaria mosquito *Anopheles funestus*. *Med. Vet. Entomol.* **19**: 458–469.
- KIMMEL, M., R. CHAKRABORTY, J. P. KING, M. BAMSHAD, W. S. WATKINS *et al.*, 1998 Signatures of population expansion in microsatellite repeat data. *Genetics* **148**: 1921–1930.
- KING, J. P., M. KIMMEL and R. CHAKRABORTY, 2000 A power analysis of microsatellite-based statistics for inferring past population growth. *Mol. Biol. Evol.* **17**: 1859–1868.
- KOEKEMOER, L., M. M. WEETO, L. KAMAU, R. H. HUNT and M. COETZEE, 2002 A cocktail polymerase chain reaction (PCR) assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *Am. J. Trop. Med. Hyg.* **66**: 804–811.
- LEHMANN, T., W. A. HAWLEY and F. H. COLLINS, 1996 An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics* **144**: 1155–1163.
- LEHMANN, T., M. LIGHT, N. ELISSA, B. T. MAEGA, J. M. CHIMUMBWA *et al.*, 2003 Population structure of *Anopheles gambiae* in Africa. *J. Hered.* **94**: 133–147.
- MACHADO, C. A., R. M. KLIMAN, J. A. MARKERT and J. HEY, 2002 Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* **19**: 472–488.
- MANTEL, N., 1967 The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**: 209–220.
- MICHEL, A. P., W. M. GUELBEOGO, O. GRUSHKO, B. J. SCHEMERHORN, M. KERN *et al.*, 2005a Molecular differentiation between chromosomally defined incipient species of *Anopheles funestus*. *Insect Mol. Biol.* **14**: 375–387.
- MICHEL, A. P., M. J. INGRASCI, B. J. SCHEMERHORN, M. KERN, G. LE GOFF *et al.*, 2005b Rangewide population genetic structure of the African malaria vector *Anopheles funestus*. *Mol. Ecol.* **14**: 4235–4248.
- NEI, M., F. TAJIMA and Y. TATENO, 1983 Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J. Mol. Evol.* **19**: 153–170.
- NOOR, M. A., K. L. GRAMS, L. A. BERTUCCI and J. REILAND, 2001 Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* **98**: 12084–12088.
- ORTIZ-BARRIENTOS, D., J. REILAND, J. HEY and M. A. NOOR, 2002 Recombination and the divergence of hybridizing species. *Genetica* **116**: 167–178.
- PAETKAU, D., W. CALVERT, I. STIRLING and C. STROBECK, 1995 Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* **4**: 347–354.
- PALSSON, S., 2004 Isolation by distance, based on microsatellite data, tested with spatial autocorrelation (SPAIDA) and assignment test (SPASSIGN). *Mol. Ecol. Notes* **4**: 143–145.
- PRITCHARD, J. K., M. STEPHENS and P. DONNELLY, 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- RAMOS-ONSINS, S. E., and J. ROZAS, 2002 Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* **19**: 2092–2100.
- RANFORD-CARTWRIGHT, L. C., P. BALFE, R. CARTER and D. WALLIKER, 1993 Frequency of cross-fertilization in the human malaria parasite *Plasmodium falciparum*. *Parasitology* **107**(Pt 1): 11–18.
- RAYMOND, M., and F. ROUSSET, 1995 GENEPOP Version 1.2. A population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248–249.
- RIESEBERG, L. H., 2001 Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16**: 351–358.
- ROZAS, J., J. C. SANCHEZ-DELBARRIO, X. MESSEGUER and R. ROZAS, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- SCHNEIDER, S., D. ROESSLI and L. EXCOFFIER, 2000 Arlequin: a software for population genetics data analysis, Ver. 2.000. University of Geneva, Geneva.
- SHARAKHOV, I., O. BRAGINETS, O. GRUSHKO, A. COHUET, W. M. GUELBEOGO *et al.*, 2004 A microsatellite map of the African human malaria vector *Anopheles funestus*. *J. Hered.* **95**: 29–34.
- SHARAKHOV, I. V., M. V. SHARAKHOVA, C. M. MBOGO, L. L. KOEKEMOER and G. YAN, 2001 Linear and spatial organization of polytene chromosomes of the African malaria mosquito *Anopheles funestus*. *Genetics* **159**: 211–218.
- STUMP, A. D., J. A. SHOENER, C. COSTANTINI, N. SAGNON and N. J. BESANSKY, 2005 Sex-linked differentiation between incipient species of *Anopheles gambiae*. *Genetics* **169**: 1509–1519.
- TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- TURNER, T. L., M. W. HAHN and S. V. NUZHIDIN, 2005 Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol.* **3**: e285.
- VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS and P. F. SHIPLEY, 2004 Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**: 535–538.
- WHITLOCK, M. C., and D. E. MCCAULEY, 1999 Indirect measures of gene flow and migration: FST not equal to 1/(4Nm + 1). *Heredity* **82**(Pt 2): 117–125.
- WORLD HEALTH ORGANIZATION, 2003 World Health Report 2003. 210 World Health Organization, Geneva.
- WU, C.-J., and C. T. TING, 2004 Genes and speciation. *Nat Rev Genet* **5**: 114–122.