REPRODUCTIVE ISOLATION AND LOCAL ADAPTATION QUANTIFIED FOR A CHROMOSOME INVERSION IN A MALARIA MOSQUITO

Diego Ayala,[1,2,3] Rafael F. Guerrero,[1,3] and Mark Kirkpatrick[1,3,4]

[1] Section of Integrative Biology C-0930, University of Texas, Austin, Texas 78712

[2] UMR 224 MIVEGEC/BEES, Institut de Recherche pour le Développement, P.O. Box 64501, 34394 Montpellier, France

[3] These authors contributed equally to this work

[4] E-mail: kirkp@mail.utexas.edu

Chromosome inversions have long been thought to be involved in speciation and local adaptation. We have little quantitative information, however, about the effects that inversion polymorphisms have on reproductive isolation and viability. Here we provide the first estimates from any organism for the total amount of reproductive isolation associated with an inversion segregating in natural populations. We sampled chromosomes from 751 mosquitoes of the malaria vector *Anopheles funestus* along a 1421 km transect in Cameroon that traverses savannah, highland, and rainforest ecological zones. We then developed a series of population genetic models that account for selection, migration, and assortative mating, and fit the models to the data using likelihood. Results from the best-fit models suggest there is strong local adaptation, with relative viabilities of homozygotes ranging from 25% to 130% compared to heterozygotes. Viabilities vary qualitatively between regions: the inversion is underdominant in the savannah, whereas in the highlands it is overdominant. The inversion is also implicated in strong assortative mating. In the savannah, the two homozygote forms show 92% reproductive isolation, suggesting that this one inversion can generate most of the genetic barriers needed for speciation.

KEY WORDS: *Anopheles*, assortative mating, selection, postzygotic, prezygotic, speciation.

Chromosome inversions have a rich history in evolutionary genetics, much of it focused on the roles they may play in speciation (White 1973; King 1993; Hoffmann and Rieserberg 2008; Kirkpatrick 2010). Two genetic properties of inversions make them particularly favorable for the evolution of reproductive isolation (Butlin 2005). First, large inversions cover hundred or even thousands of loci, and so could easily span genes involved both prezygotic and postzygotic isolations. Second, inversions dramatically reduce recombination when heterozygous, increasing linkage between any genes present that contribute to the two types of isolation. Consider the situation when prezygotic isolation results from assortative mating of inversion homozygotes and postzygotic isolation results from selection against inversion heterozygotes. Then pre- and postzygotic isolations will reinforce each other, a situation that greatly facilitates speciation (Rice 1987; Kirkpatrick and Ravigné 2002; Gavrilets 2004). Inversions have been implicated in prezygotic and postzygotic isolations in *Rhagoletis* flies (Feder et al. 2003; Michel et al. 2010) and monkeyflowers (Lowry and Willis 2010). Quantitative measures for the overall strength of reproductive isolation generated by an inversion, however, are as yet still lacking.

The mosquito *Anopheles funestus* offers an unusually compelling opportunity to study inversions and speciation. Widely distributed across the sub-Saharan Africa, *A. funestus* is one of the most proficient vectors of human malaria, and in some places its transmission rate surpasses *A. gambiae*, its better studied congener (Coetzee and Fontenille 2004). The species is highly polymorphic for inversions, but a clear picture of their evolutionary

significance remains elusive because of conflicting results from different geographical regions and from different genetic markers (Costantini et al. 1999; Kamau et al. 2003; Boccolini et al. 2005; Cohuet et al. 2005; Michel et al. 2006).

Inversion 3Ra in *A. funestus* is a prime candidate for detailed investigation. It is large, spanning roughly 30% of the right arm of Chromosome 3, and 7% of the entire genome (which has 3 pairs of chromosomes). Its frequencies in Africa are correlated with humidity (Costantini et al. 1999; Guelbeogo et al. 2005; Michel et al. 2006; Ayala et al. 2011b), suggesting that this inversion is involved in local adaptation. Other phenotypes that have been associated with 3Ra are resting behavior (Costantini et al. 1999), host preference (Costantini et al. 1999), and wing shape (Ayala et al. 2011a). Previous studies have also suggested that this inversion contributes to the genetic isolation between populations in Burkina Faso (Costantini et al. 1999; Guelbeogo et al. 2005).

Chromosome inversions in *Anopheles* mosquitoes have important implications for human health (Sharakhov et al. 2002; Besansky et al. 2003; White et al. 2007). Malaria is a severe problem in Africa, responsible for over one million deaths per year and more than 250 million people infected (Murray et al. 2012). Inversion polymorphisms in *Anopheles* are critical to the epidemiology of malaria: they affect habitat preferences (Coluzzi et al. 1979; Costantini et al. 1999), feeding behavior (Petrarca and Beier 1992; Lochouarn et al. 1998), and they may have enabled large range expansions (Toure et al. 1994; Besansky et al. 2003; Cohuet et al. 2005; Manoukis et al. 2008; Ayala et al. 2011b). Furthermore, inversions affect mosquito control because they limit the potential for insecticide resistance alleles to introgress between populations and impact the possibility of using genetically modified mosquitoes for disease control (Alphey et al. 2002; Boete and Koella 2003; Tripet et al. 2007; Enayati and Hemingway 2010).

In this study, we estimate the effects of inversion 3Ra in *A. funestus* on viability, local adaptation, and both pre- and postzygotic reproductive isolations. We sampled mosquitoes along a transect in Cameroon that traverses diverse habitats. We next constructed a series of spatially explicit population genetic models, then fit the models to the data. The best-fit models suggest that the inversion is subject to strong viability selection that varies between the zones, and to strong assortative mating. The results have implications for how inversions evolve, and provide hypotheses for further experimental studies of their genetics, physiology, and ecology.

Methods
SAMPLING
We sampled 751 mosquitoes from 105 villages that lie along a 1421 km transect in Cameroon that follows the only highway that runs the length of the country. The transect, shown
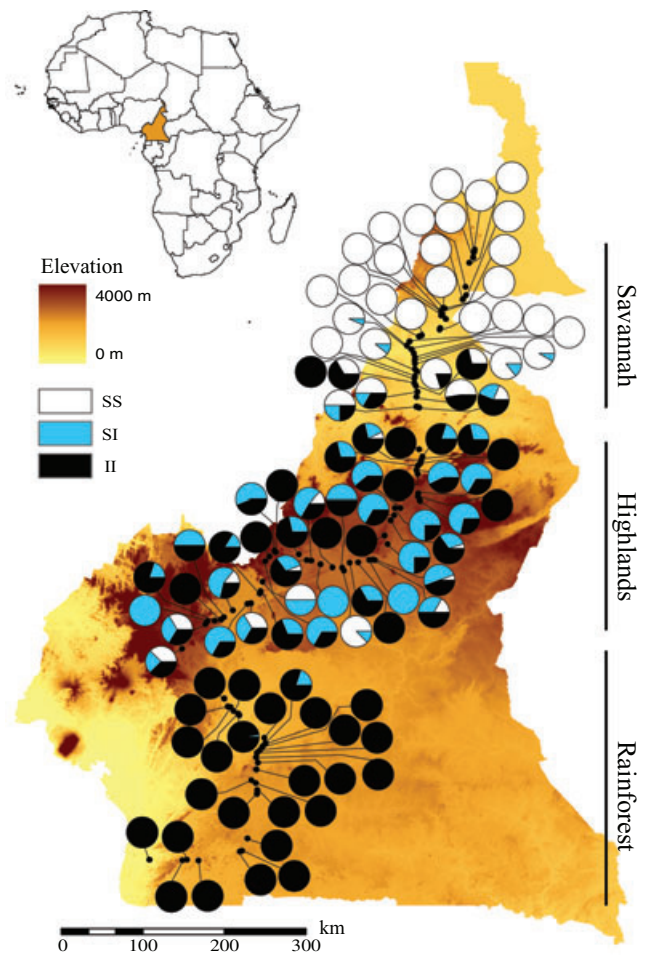

Figure 1. Karyotype frequencies for inversion 3Ra in Cameroon. Solid points show villages sampled for *Anopheles funestus*. Pie diagrams show the frequencies of the standard homozygote (SS, white), heterozygote (SI, blue), and inverted homozygote (II, black).

in Figure 1, traverses three markedly distinct ecological zones. It begins in the sub-Saharan savannah, crosses a highland region with mountains that reach 4000 m, then descends into the lowland tropical rainforest. *A. funestus* is largely commensal on humans. Most of its population is thought to be strongly associated with human presence, and our transect follows the region of suitable habitat that runs north–south in Cameroon (Ayala et al. 2009).

Adult mosquitoes were collected between August and December (the wet season) in 2005 and 2006 from inside homes (Ayala et al. 2009). Sampling dates for each zone were timed so that they were in comparable phases of the wet season. Species identification was confirmed using polymerase chain reaction (Cohuet et al. 2003). Karyotypes were determined with a phase contrast microscope following the *A. funestus* cytological map (Sharakhov et al. 2004). Sixty percent of the individuals were independently karyotyped a second time by different technician,

and no discrepancies were found. Coordinates, sample sizes, and karyotype count for all of the localities are given in Table S1.

THE MODELS

We fit a series of models to the data to estimate the strengths of evolutionary forces acting on the inversion. The models differ in their assumptions about selection, mating, and migration. We implemented the models in simulations, then used likelihood to determine the model and parameter values that best fit the data.

The life cycle, based on the natural history of *Anopheles*, assumes that viability selection occurs at the larval stage (Clements 1999). Viability selection is followed by mating, which may be assortative. Mated females disperse and then lay their eggs. Generations are nonoverlapping.

The viabilities of the three genotypes depend on the ecological zone in which a deme lies. We allowed there to be one, two, or three zones, and assumed that selection is uniform within a zone. With two and three zones, we assumed the zone boundaries run east–west, and allowed the locations of the boundaries to be variables in the model. Viability selection is soft, and so the number of surviving adults in a deme is independent of the fitnesses and genotype frequencies. Heterozygotes are assigned a relative viability of 1, and we denote the viabilities of standard and inverted homozygotes as W_{SS} and W_{II}, respectively. These parameters can be smaller or greater than 1, which allows for directional, overdominant, or underdominant selection.

Other inversions in *Anopheles* are thought to be involved in assortative mating (Lehmann and Diabate 2008; Perevozkin et al. 2012). We therefore allowed for this possibility by considering four models for mating. The first assumes mating is random. In the second, assortative mating occurs when a fraction F of individuals to mate only with others sharing the same karyotype, whereas the remainder mate at random. In the third model for mating, the strength of assortative mating takes on a different value in each ecological zone. In this case, there are as many assortment parameters as there are ecological zones (F_1, F_2, etc.). Finally, we allowed the karyotypes to differ in the strength of assortment. This model includes, for example, the case in which only homozygotes have a mating preference. In this case there are three assortment parameters (F_{SS}, F_{SI}, and F_{II}).

We have limited information about how mosquitoes are distributed in space and how they move. We therefore considered two very different models for population structure and migration that hopefully bracket the actual situation. The first is motivated by an analysis of habitat suitability (Ayala et al. 2009) and the first author's experience while collecting mosquitoes over 2 years. The species appears to be largely commensal on humans and strongly concentrated in the villages that we sampled along the transect. Mosquitoes are rarely found in the surrounding habitat, even where nonhuman animals are abundant. Our first model is

therefore a one-dimensional stepping-stone array with 105 demes of equal size that represent the villages sampled. To account for the variation in distances between demes, the probability of dispersal to the neighboring deme that is at a linear distance d is calculated as the area under the tail of a Gaussian distribution with expectation 0 and variance σ_m^2 evaluated at $d/2$.

The second model of migration takes an strongly contrasting view of how mosquitoes are distributed by assuming that the mosquitoes live in a continuous two-dimensional habitat with constant densities. Migration is uniform in all directions with a variance σ_m^2. Without any loss of accuracy, we modeled this situation using a continuous one-dimensional habitat running north–south (i.e., orthogonal to the boundaries between the ecological zones). The simulation model then predicts the genotype frequencies for a given village based on its latitude, that is, its position along the ecological gradient.

Table 1 summarizes the combinations of these assumptions that we simulated. The simulations were written in C++. Runs were initialized with karyotype frequencies at the values observed in each population (except if an observed frequency was zero, where it was adjusted to 10^{-3}). Simulations were stopped when all karyotype frequencies changed less than 10^{-8} per generation. The equilibrium frequencies were evaluated at the point in the life cycle following dispersal, corresponding to when individuals in our field samples were collected. The number of generations needed to reach equilibrium depends on the parameter values and initial conditions. With our maximum likelihood parameter estimates, the system evolves roughly halfway to the equilibrium in 200 generations, corresponding to about 15 years for this species.

Fitting the Models

We estimated the parameters for each model by finding the values that give equilibrium frequencies for the inversion that best fit the data. The fit was evaluated using likelihood. Write $f_{i,j}$ for the number of individuals sampled from deme (village) i with karyotype j ($= 0, 1, 2$). Denote the vector of parameter values as $\mathbf{\Pi}$, and the equilibrium frequency in deme i of karyotype j from the simulation under those parameter values as $p_{i,j}(\mathbf{\Pi})$. Then the log likelihood for the vector $\mathbf{f}$ of all the data, given the parameter values $\mathbf{\Pi}$, is found from the multinomial distribution:

$$lnL(\mathbf{f}|\Pi) = C \sum_{i=1}^{n_p} \sum_{j=1}^{3} f_{i,j} \ln(p_{i,j}(\Pi)), \qquad (1)$$

where C is a constant that is independent of the parameters $\mathbf{\Pi}$. The maximum likelihood estimate for the parameters is found by maximizing equation (1) with respect to $\mathbf{\Pi}$.

We found the parameter values that maximize the likelihood using a simulated annealing method (Press et al. 2007).

Table 1. Models that were fit to the data. The four best-fit models are shown at top, with the best-fit model listed first. The first three columns give the assumptions. Under Migration, "SS" is the stepping stone model and "C" is the continuous space model. The following columns give the parameter estimates, the negative log likelihood, and the Akaike's information criterion (AIC) values for each model. The likelihood and AIC values are scaled relative to the best-fit model.

Number of zones	Assortative mating	Migration	Savannah		Highlands		Rainforest		F	σ_m^2	$-\ln(L)$	AIC
			W_{SS}	W_{II}	W_{SS}	W_{II}	W_{SS}	W_{II}				
3	Yes	SS	1.29	1.30	0.25	0.50	0.73	0.86	0.82	7.34	–	–
		C	1.44	1.44	0.27	0.51	0.83	0.90	0.82	7.12	−1.4	2.8
	By genotype	SS	1.21	1.21	0.31	0.57	0.95	1.05	SS: 0.95 SI: 0.68 II: 0.002	80.8	1.3	1.4
	By zone	SS	1.22	1.23	0.21	0.44	1.65	1.28	Sa: 0.83 Hi: 0.96 Ra: 0.3	5.8	1.1	1.8
1	Yes	SS	1.00	1.00					0.5	3.0	−102.9	195.8
		C	1.00	1.00					0.5	7.12	−117.7	224.8
2	Yes	SS	1.28	1.27	0.33	0.60			0.80	49.4	−80.8	157.6
		C	3.22	3.16	0.81	0.95			0.39	34.0	−96.0	188.0
3	No	SS	1.04	1.04	0.56	0.84	1.08	1.08		37.0	−35.8	69.6
		C	1.33	1.25	0.56	0.84	1.07	1.1		19.0	−36.5	71.0

This heuristic algorithm samples the parameter space sequentially, starting from an arbitrary set of parameters. Random changes are then made to the parameter set. The new values are always accepted if they increase the likelihood of the data. A change that decreases the likelihood is accepted with probability proportional to a control parameter T. These steps are repeated with gradually decreasing values of T, with the goal of reaching the global maximum of the likelihood function.

We obtained confidence regions for our maximum-likelihood estimates by calculating likelihood profiles of parameters for the best-fit model. Profiles for migration, mating, and selection parameters were estimated by evaluating each parameter on a lattice of points, then maximizing the likelihood with respect to the remaining parameters. This process is computationally costly, so we fixed the region boundaries at their maximum-likelihood estimates to analyze profiles of other parameters. (Our boundary estimates correspond closely to those previously defined on ecological grounds (Olivry 1986), and are the parameters of least interest.) Approximate 95% and 99% confidence regions were identified as the parameter space for which $\ln(L)$ was within 2 and 3 units, respectively, of the maximum. These likelihood profiles were also useful in verifying that our annealing runs did in fact reach optima.

To evaluate the fit of different models we used Akaike's information criterion (AIC; Akaike 1974). This measure takes into account both the likelihood and the number of parameters in the model, penalizing those models that have more parameters. We judged a model to have better fit than another if it has a lower AIC score. The fit of two models whose AIC scores differ by less than 6 units is judged to be not significantly different ($P > 0.05$), whereas the fit of models differing by more than 14 is highly significantly different ($P < 0.001$; Burnham and Anderson

2004). Some of our models are nested (i.e., special cases of other models). In these cases, we also compared models although χ^2-tests of their likelihood ratios (Whitlock and Schluter (2009), ch. 20).

We estimated total reproductive isolation (RI_{total}) in each ecological zone by combining the contributions of prezygotic (RI_{pre}) and postzygotic isolation (RI_{post}):

$$RI_{total} = RI_{pre} + (1 - RI_{pre})RI_{post} \qquad (2)$$

(Coyne and Orr 1989, 1997). Prezygotic isolation is the proportional reduction in heterokaryotypic matings caused by assortment in a population consisting of 50% standard and 50% inverted individuals, and is equal to

$$RI_{pre} = 2F/(1+F), \qquad (3)$$

where F is the fraction of individuals that mate with their own karyotype (with a fraction $1 - F$ mating at random). Postzygotic isolation is the reduction in survival of F_1 hybrids relative to the average of their parents:

$$RI_{post} = 1 - 1/[(W_{SS} + W_{II})/2]. \qquad (4)$$

where W_{SS} and W_{II} are the viabilities of standard and inverted homozygotes, respectively, relative to the heterozygotes.

Results
THE INVERSION CLINE

Inversion 3Ra in *A. funestus* shows a dramatic cline in frequency along the transect. As shown in Figure 1, the inversion is near fixation in the rainforest, at intermediate frequencies in the highlands, and nearly absent in the savannah. Shifts in frequencies

at the boundaries between the zones are abrupt, which is consistent with how rapidly the environment changes at these transition points. Within each of the three zones, inversion frequencies are generally similar. The data also suggest qualitatively that there are heterozygote deficits in the highlands, as can result from selection against inversion heterozygotes and/or assortative mating.

VIABILITY, MATING, AND REPRODUCTIVE ISOLATION

We fit 10 models that differ in their combinations of assumptions about selection, mating, and migration. Four models gave very similar fits to the data. However, they also give very similar qualitative (and even quantitative) results regarding selection and mating, suggesting that those conclusions are robust. We will first describe the model that best fits the data, then turn to the alternatives. They include three models whose fit is statistically indistinguishable from the best-fit model.

Under the best-fit model, viability varies in space and there are three ecological zones corresponding to the savannah, highlands, and rainforest. The estimated locations for the boundaries between the zones lie very close to boundaries that were previously defined on ecological grounds (Table S1; Olivry 1986). The inversion is involved with assortative mating, but its strength does not vary in space or with the karyotype. Migration follows a stepping stone model in which the demes correspond to villages along the highway where the mosquitoes were sampled.

The fit of the model to the data is shown in Figure 2. Qualitatively, predictions for the average karyotype frequencies within each zone are good. Much of the variation around the prediction (visible particularly in the highlands) results from the small sample sizes in individual villages. Only three of the 105 villages have frequencies that differ significantly from the model (Table S1), fewer than is expected by chance.

The estimated strength of viability selection is intense. Figure 3 and Table 1 show estimates for the viabilities of the karyotypes in the three ecological zones. The viabilities of homozygotes range from 25% to 130% relative to those of heterozygotes. Furthermore, the pattern of selection differs strikingly between the three zones.

Heterozygotes have low fitness in the savannah, reduced by about 23% compared to the two homozygotes, which in turn have similar fitness. The evidence for underdominance in the savannah is strong, as the 99% confidence region excludes values for the homozygote fitnesses that are equal to or lower than the heterozygote fitness.

In the highlands and rainforest, by contrast, standard homozygotes are strongly selected against, whereas heterozygotes have highest fitness. This pattern of overdominance is statistically significant in the highlands, where the viabilities of standard and inverted homozygotes are respectively 25% and 50% that of het-
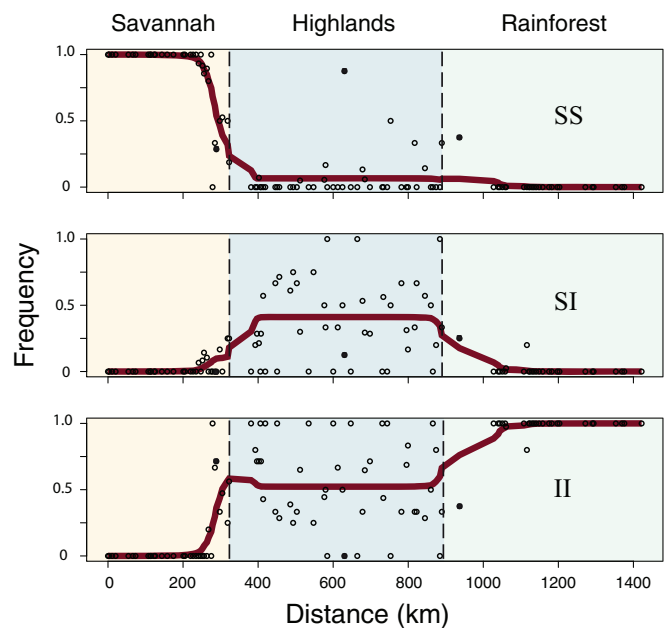

Figure 2. Comparison of the observed frequencies (circles) and predicted frequencies from the best-fit model (curves) for the three karyotypes. SS = standard homozygotes, SI = heterozygotes, II = inverted homozygotes. Populations that differ significantly from the model are shown as filled circles. Further details are given in Table S1.
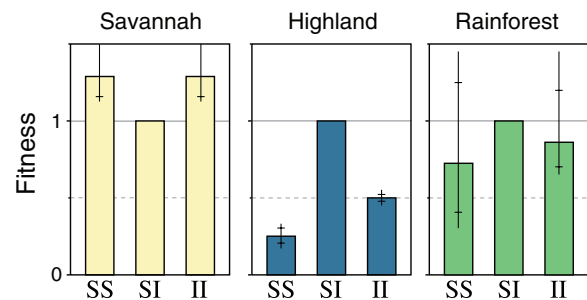

Figure 3. Estimated viabilities of standard homozygotes (W_{SS}) and inverted homozygotes (W_{II}) relative to heterozygotes (W_{SI}) in the three ecological regions. Bars show maximum likelihood estimates, and the error bars show 95% and 99% confidence regions (identified respectively by a change of 2 and 3 log-likelihood units). In the savannah, the upper 99% limits are 1.94 and 1.95 for standard and inverted homozygotes, respectively.

erozygotes (Fig. 3). Results from the rainforest are consistent with overdominance. The pattern is not significant, however, because the inversion is near fixation there and selection parameters are therefore poorly estimated.

The data suggest that the inversion is involved with strong assortative mating. The maximum likelihood estimate for F is 0.82 (Fig. 4). The lower bound of the 99% confidence interval is $F = 0.77$, which implies that even the weakest assortment consistent with the data is still strong. To our knowledge, this is the
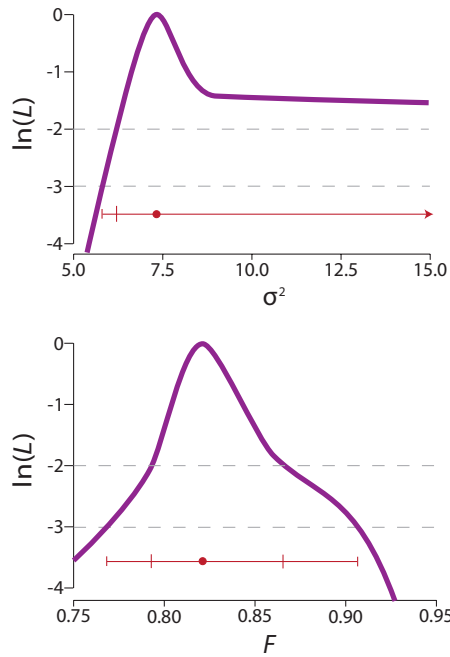
Figure 4. Likelihood profiles for the dispersal variance (σ^2, above) and assortative mating (*F*, below). The vertical axis shows the decrease (in ln(*L*) units) from the maximum likelihood value. The maximum likelihood estimate is shown as the closed circle, and the confidence regions are indicated by crosshairs on the horizontal line (95% for the inner and 99% for the outer crosshairs). The upper limit for the confidence region of *F* is unbounded.

first estimate in any natural population for the overall strength of assortment associated with a chromosomal rearrangement. Our analysis provides no suggestion, however, regarding the mechanism of assortment. It might result from active mate choice or from passive mechanisms such as habitat preference and the pattern of diurnal activity (Diabate et al. 2009; Pennetier et al. 2009).

Our inferences about viability and assortment are based on the assumption that the inversion (or genes that it carries) is causal for these effects. We are unable to test this assumption formally, for example by a classic genetic analysis using recombinants, because we (as well as others) have been unable to establish a breeding colony in Cameroon despite several years of effort. In the Discussion we return to the possibility that the patterns in our data are caused by genetic factors outside of 3Ra.

We can combine our inferences about natural selection and assortative mating to arrive at an overall estimate of the strength of reproductive isolation conferred by this inversion. The most interesting context is the savannah, where the premating isolation that results from assortment is reinforced by the postmating isolation that results from underdominance. Using our estimates for viability and assortment with equation (2), we estimate that the total isolation between standard and inverted homozygotes is 92%. Genetic introgression between the two homozygous forms in the

savannah continues as the result of both the incomplete isolation and gene flow from neighboring populations in the highlands.

We can also estimate the total reproductive isolation in the rainforest and highlands habitats. Heterozygotes in those two habitats have highest viability, however, which produces negative values for postzygotic isolation. If one is willing to allow for that possibility, the values for total isolation in the highlands and rainforest are 0.74 and 0.88, respectively.

The last parameter estimated in our analysis is the dispersal variance. The maximum likelihood estimate is 7.3 km^2, corresponding to an average movement of 2.2 km/generation (Fig. 4). This is the parameter that is estimated with least precision (the 95% confidence region is [6.2, ∞], where the upper bound corresponds to free movement between adjacent villages). This lack of precision presumably results because migration only has strong effects on the predictions for the frequencies in villages near to boundaries between ecological zones. It is, however, the first estimate of the migration rate for this malaria vector, and for any organism using an inversion. In the best-fit model, migration follows a stepping stone model in which the demes correspond to villages. The alternative model for migration, in which space is continuous, fits the data virtually as well and gives very similar parameter estimates (Table 1).

Alternative Models

These qualitative conclusions about the patterns of evolutionary forces and the quantitative parameter estimates depend on the model that is assumed. Table 1 summarizes results from the 10 models that we analyzed. The two right-hand columns give the negative of the log likelihood (where a larger value corresponds to better agreement between model and data) and the AIC (where a smaller value implies a better fit after accounting for the number of parameters in the model). Both of these scores are measured relative to the model that fit best, as judged by AIC.

Six models can be rejected because they give a significantly worse fit to the data (based on an AIC score and/or a likelihood ratio test). For the remaining four models, the fits to the data are very similar. Importantly, the patterns and strength of selection and assortative mating that they estimate are very similar, suggesting that our central conclusions may be robust.

The four models that fit best share several key features. First, they show strong evidence that the fitness of the inversion varies between three ecological zones. Second, they suggest the inversion is overdominant in some habitats but underdominant in others. All of these models give similar maximum likelihood estimates for the viabilities of the inversion genotypes in the savannah and highlands, where all the estimates fall within the 95% confidence intervals shown for viability parameters in Figure 3.

Finally, all four models implicate the inversion in strong assortative mating.

There are, however, important differences between the four best-fit models regarding mating and dispersal. We cannot distinguish between a model in which the strength of assortment is constant, one in which it varies by genotype, and one in which it varies by ecological zone. The latter two models show slightly improved likelihood over the best-fit model, but the two extra parameters that they each require lead to slightly worse AIC scores. In all cases, however, there is support for relatively strong assortment in the savannah and highlands. (Estimates in the rainforest are imprecise because there is little polymorphism there.)

We also cannot distinguish between two very different sets of assumptions about spatial distribution and dispersal, a linear stepping-stone versus a continuous two-dimensional habitat. Figure S1 plots the data and model predictions as a function of latitude (rather than position along the transect, as in Fig. 2). It shows that the two models give very similar fits to the data. Inferences about dispersal and population distribution are the weakest point of our analysis. It is encouraging, however, that the qualitative results for viability and assortment are the same under these very different models for population structure.

We can clearly reject several alternative models for how viability selection varies in space. In one alternative there is only a single ecological zone, and fitnesses for the karyotypes do not vary along the cline. This model includes the situation in which the cline is maintained in a "tension zone" where a combination of underdominance and assortative mating causes fixation of standard homozygotes at one end of the cline and inversion homozygotes at the other (Barton and Hewitt 1985; Barton and Gale 1993; The estimated viability of the heterozygote is very slightly less than that of the homozygotes, but the difference is too small to appear in Table 1.) In another alternative model of selection, there are two ecological zones. Both the one- and two-zone models are decisively rejected ($P \ll 10^{-6}$, χ^2-test of the likelihood ratio).

We also considered several alternative assumptions about mating. The simplest is random mating. Here departures from Hardy–Weinberg result only from viability selection and migration. This model fits the data poorly ($P \ll 10^{-6}$, χ^2-test of the likelihood ratio). As we discussed earlier, other models make different assumptions about how assortative mating works. In one, the strength of assortment is allowed to vary between the three ecological zones. In the second, the strength of assortment depends on the karyotype but does not vary in space. The fit to the data are marginally improved in both models by the addition of two more parameters, but the increase in likelihood is not significant ($P > 0.05$, χ^2-tests of the likelihood ratios).

Finally, we considered the possibility that our results are driven by aberrant samples. Karyotype frequencies depart significantly from the model predictions at the $P < 0.05$ level in three of the 105 villages we sampled (Table S1). Although at least that many are expected by chance alone, we removed those villages from the dataset and reran the analysis for the best-fit model. The parameter estimates changed little (results not shown), and so the analyses reported above are based on the full data set.

Discussion

This study estimates the viability effects and nonrandom mating associated with an inversion. Three key conclusions are suggested by the four models that fit best. First, viability selection appears to be intense. The maximum likelihood estimates from the best-fit model suggest that survival of homozygotes ranges from 25% to 130% of the heterozygote fitness. Second, the pattern of selection appears to differ strikingly between habitats. In the savannah, inversion 3Ra is strongly underdominant, whereas in the highlands it is strongly overdominant. Third, assortative mating seems to be strong. The likelihood analysis suggests that between 77% and 91% of matings are assortative based on karyotype. These estimates are based on the assumptions of the models. Independent estimates, based for example on experiments in the field, would be very valuable. The following conclusions can be viewed as hypotheses that might be tested using other approaches.

In the savannah, where selection is underdominant, the inversion segregates as a single genetic element that confers both prezygotic and postzygotic isolation. Natural selection depresses the frequency of heterozygotes. This enhances the power of assortative mating to inhibit matings between carriers of standard and inverted chromosomes. In short, natural selection reinforces premating isolation because the inversion links viability and mating phenotypes. A trait with these properties is particularly favorable to speciation (Kirkpatrick and Ravigné 2002), and so is sometimes referred to as a "magic trait" (Gavrilets 2004). The total reproductive isolation that results in the savannah between chromosome heterozygotes is 92%, which is perhaps the strongest that ever been reported from sympatric forms of a single species. This degree of isolation is similar to the figure of 94% reported for sympatric pairs of *Drosophila* species when measured in the lab (Coyne and Orr 1997). It is difficult to compare these numbers directly, however, as the *Drosophila* estimates do not take into account the effect of factors such as habitat preferences that are excluded from the lab environment.

It may seem remarkable that an inversion, which segregates like a single locus with two alleles, could at once confer both prezygotic and postzygotic isolations. Evidence from two other species suggests this situation may not be uncommon, however, at least for large inversions that capture many loci. In *Rhagoletis* flies, inversions influence diapause, which in turn adapts races of the flies to alternative hosts and contributes to assortative mating (Feder et al. 2003; Michel et al. 2010). In the yellow

monkeyflower, *Mimulus guttatus*, a recently discovered inversion affects several morphological and phenological traits (Lowry and Willis 2010). These changes adapt the plant to two very different habitats and incidentally contribute to both prezygotic and postzygotic isolations between the two chromosomal forms. There is also abundant evidence from other species of *Anopheles* that inversions have diverse phenotypic effects, including aridity tolerance (White et al. 2007; Gray et al. 2009; Cassone et al. 2011; Fouet et al. 2012), temperature tolerance (Rocca et al. 2009), mating preferences (Perevozkin et al. 2012), host preference (Coluzzi et al. 1979; Petrarca and Beier 1992; Costantini et al. 1999), resting behavior (Coluzzi et al. 1977; Rishikesh et al. 1985; Bryan et al. 1987; Costantini et al. 1999), susceptibility to parasites (Toure et al. 1996), breeding site preference (Manoukis et al. 2008), insecticide resistance (Brooke et al. 2002), and environmental tolerance (Coluzzi et al. 1979; Toure et al. 1998; Petrarca et al. 2000; Coluzzi et al. 2002; Cohuet et al. 2004; Cohuet et al. 2005; Simard et al. 2009; Ayala et al. 2011b).

What evolutionary events led to a polymorphism for inversion 3Ra that mediates both local adaptation and reproductive isolation? When two or more loci experience local adaptation to environmental conditions that change in space, an inversion that spans the loci will spread under quite general conditions (Kirkpatrick and Barton 2006). The result predicted by theory is a cline that looks very much like that seen in Figure 1. One could imagine that when inversion 3Ra first appeared, it captured alleles that allow its carriers to survive well in more humid environments and also to mate in a different time or place than savannah-adapted mosquitoes do. Our data are consistent with this hypothesis for how inversions become established, and it could explain why inversion 3Ra causes both prezygotic and postzygotic isolations.

The data are, however, also consistent with at least two other hypotheses for how the inversion was established (Kirkpatrick 2010). One of the inversion breakpoints could be responsible for a mutation at a single locus that has pleiotropic effects on both viability and mating. Alternatively, the inversion's effects could result from genetic divergence at one or more loci that accumulated after the inversion was established by some other mechanism (Navarro and Barton 2003). Distinguishing between these hypotheses will require other kinds of data, for example molecular markers for the inversion breakpoints and candidate loci for adaptation inside the region it spans. Developing these resources will require substantial effort, however, because the genome of this species has not been sequenced. Regardless of the inversion's history and the genetic basis of its phenotypic effects, our results suggest the inversion experiences very strong local adaptation and nonrandom mating.

Three previous studies have estimated the viability effects of polymorphic inversions. Lewontin and White (1960) studied two inversions segregating on different chromosomes in the grasshop-

per *Moraba scurra*. They estimated the genotype viabilities using the deviation between their observed frequencies and those expected under random mating. They found very strong selection, with statistical support for some genotypes differing by 40% in viability. They constructed adaptive topographies (the surface of mean fitness as a function of the inversion frequencies) for two populations and reported that both lie at saddle points. This result is surprising, because under many conditions a saddle point is evolutionarily unstable. Lewontin and White offered several hypotheses, for example that fitnesses are frequency dependent. They did not, however, consider two hypotheses suggested by our results. First, fitnesses may vary in space. If the localities that they sampled are at midpoints along a cline, genotype frequencies could resemble the pattern of disruptive selection they describe. Second, mating might be assortative. Although Lewontin and White assumed random mating, assortment could produce the appearance of disruptive selection. Unfortunately, their data do not seem able to test either of these two possibilities.

The most direct evidence for selection acting on a common inversion comes from a 900 kb inversion on human chromosome 8 (Stefansson et al. 2005). Data on fertilities of the karyotypes in Iceland lead to the estimate that inversion heterozygotes have about 3% greater fitness than standard chromosome homozygotes. An intriguing observation is that the inversion seems to be spreading in Europe but not in Africa, suggesting that its fitness effects vary in space.

Schaeffer (2008) analyzed the famous inversion polymorphisms in *Drosophila pseudoobscura* studied by Dobzansky et al. (Dobzhansky 1944; Anderson et al. 1991). He used an approach similar to ours that fit a population genetic model to karyotype frequencies sampled from populations sampled across the western United States. Two of Schaeffer's major conclusions are qualitatively consistent with ours. First, selection can be strong: the fitnesses of common karyotypes often differ on the order of 20%, and some karyotypes in some environments are nearly lethal. Second, there is substantial geographical variation in fitnesses. There are, however, differences between the Schaeffer's analysis and ours that make it difficult to compare the results quantitatively. In Schaeffer's model, mating is assumed to be random, migration rates are assumed rather than estimated from the data, and fitnesses are estimated by minimizing the summed differences between observed and predicted frequencies rather than by likelihood. Finally, because the *D. pseudoobscura* system involves five different segregating arrangements, it is not possible to describe patterns of selection in the simple terms of overdominance and underdominance as is done here for inversion 3Ra in *A. funestus*.

A large number of studies from the mid-twentieth century quantified the contribution that inversions make to postzygotic isolation by measuring the fitnesses of hybrids in crosses between species that have karyotypic differences (White 1973; King 1993).

It is difficult to compare those data to our results for two reasons. We do not know what evolutionary forces established those inversions. Second, it is not known whether the fitness effects that we see now were present when the inversion first appeared, or whether they accumulated after the two species were fixed for alternative arrangements.

Speciation and Inversion 3Ra

Do our results imply that *A. funestus* is speciating, or perhaps already has? Microsatellite loci on chromosome 3 both inside and outside of the inversion show signs of differentiation between standard and inverted chromosomes in some geographical regions but not others (Cohuet et al. 2005; Michel et al. 2006; Ayala et al. 2011b). This suggests that there is some ongoing recombination and genetic introgression between the standard and inverted arrangements, at least in some populations. In Cameroon, the opportunity for recombination is increased in the highlands, where heterozygotes are common. At the genomic level, we expect the amount of genetic exchange between standard and inverted chromosomes to vary along the inversion. The regions surrounding the breakpoints and genes within the inversion that experience local adaptation should show higher divergence than other regions (Guerrero et al. 2012). Early in the process of speciation, variation across the genome in the degree of differentiation may be quite common (Wu 2001; Diabate et al. 2009; Nosil et al. 2009; Butlin 2010; Michel et al. 2010; White et al. 2010). The current situation in *A. funestus* could therefore represent an intermediate step on the path to speciation that will be completed in the near future, perhaps with contributions from other inversions.

It is also possible that we are observing a long-term equilibrium in which genetic isolation will never be completed. The genus *Anopheles* is replete with examples of inversions that are shared across species boundaries (Coluzzi 1982). These result from both shared ancestral polymorphisms and introgression between species (Besansky et al. 2003; White et al. 2009). One consequence is a mosaic pattern of differentiation within and between species (Turner et al. 2005; Turner and Hahn 2007; Wang-Sattler et al. 2007; White et al. 2009; Turner and Hahn 2010; Cheng et al. 2012).

It may ultimately be possible to test hypotheses regarding the role that this inversion plays in reproductive isolation and local adaptation by analyzing patterns of neutral molecular variation with coalescent models of inversions (Guerrero et al. 2012). In an important recent advance, Cheng et al. (2012) resequenced inversions from *A. gambiae* along a cline in Cameroon that shares many qualitative features with the one studied here. Their results are consistent with the pattern expected for an inversion polymorphism maintained by local adaptation of loci carried by the inversion, but quantitative tests of that hypothesis have not yet been done.

Our results imply there is strong assortative mating, but they give no information about the mechanism responsible. In principle, it could involve differentiation between standard and inverted chromosomes in both a signaling system and a receiving system. A genetically simpler possibility is that the chromosome forms differ in a single trait that leads to assortative mating, for example habitat preference. This is what happens with the *M* and *S* molecular forms of *A. gambiae*, which form mating swarms in different habitats (Diabate et al. 2009). Another mechanism that could contribute to the assortment we observe is suggested by the recent discoveries that mated pairs tend to have matching wing beat frequencies (Gibson and Russell 2006; Pennetier et al. 2009), and that inversion 3Ra affects wing shape in *A. funestus* (Ayala et al. 2011a). Because wing shape could plausibly affect wing beat frequency, this could be one pathway that contributes to the prezygotic isolation between inversion homozygotes (Sanford et al. 2011).

Crossing an Adaptive Valley

An enduring conundrum in evolutionary genetics is to explain how underdominant chromosome rearrangements are established (White 1973; King 1993). Some inversions are unconditionally underdominant, causing decreased heterozygote fitness regardless of the environment. Many pericentric inversions have this property because recombination in inversion heterozygotes generates aneuploid gametes and so decreases fertility (Hoffmann and Rieseberg 2008). Paracentric inversions such as 3Ra, however, may often be free of these effects. They can then evolve as the result of the alleles that they carry, which may have fitness effects that depend on the environmental context.

Our results suggest how a paracentric inversion that is underdominant can become established when its fitness varies in space. Heterozygotes for 3Ra have lowest relative fitness in the savannah, but they have highest fitness in the highlands. Assuming that the savannah-adapted standard chromosome is ancestral (Green and Hunt 1980), an inverted chromosome that appeared by mutation could have spread in the highlands (and perhaps enabled a range expansion there and into the rainforest). If highland populations were to evolve reproductive isolation from the savannah populations, the stage would be set for inverted homozygotes to reinvade the savannah if suitable habitat becomes available there. The result would be a pair of sympatric sister species separated by an adaptive valley of low heterozygote fitness. This scenario is hypothetical in the case of *A. funestus*. Nevertheless, the potential for this mechanism is supported by the fact that previous studies (Stefansson et al. 2005; Schaeffer 2008; Lowry and Willis 2010)

as well as our results show that the fitness of inversions can vary in space.

Caveats and Extensions

Our strategy here uses statistical analyses of genotype frequencies to make inferences about selection and mating. A fundamental strength of this approach is that the data come from natural populations. Model-based estimation is the basis for the vast majority of inferences about how selection acts on genetic variation in other species, including humans.

This approach, however, also has weaknesses. We noted earlier that we do not have a formal genetic analysis proving that inversion 3Ra is the causal agent of the patterns seen in Figure 1. Controlled crosses to test the effects of the inversion are not possible because no karyotyped breeding colony of *A. funestus* is available. (We and others have tried several times without success to establish a colony.) Another technical limitation of this system is that the data are based on classical cytogenetic methods that can be applied only to one sex at one point in the life cycle (half-gravid females). Molecular markers are not available for the inversion breakpoints of *A. funestus*. Those would be valuable to develop in the future because they would allow further tests of the model, for example by sampling inversion frequencies at different life stages and by correlating the genotypes of females and the sperm of the males they mated.

Progress toward identifying the loci involved in local adaptation and isolation is likely to be slow, even in those species where additional genetic resources are available. Because recombination is almost completely blocked in inversion heterozygotes, it is generally not possible to map phenotypic effects using recombinants. For now, inversion 3Ra is perhaps best regarded as a single (albeit very large) quantitative trait locus segregating for two alleles that have major effects on multiple phenotypes.

It is conceivable that the cline in inversion 3Ra shown in Figure 1 is in fact caused by selection on loci outside of the inverted region. If so, then those loci must be in very strong linkage disequilibrium with 3Ra, which would require both strong epistatic selection and tight linkage between the causal locus and the inversion. That would change our quantitative results about selection, but likely would not change the qualitative conclusion that there is strong assortative mating and strong natural selection that varies in space. Previous studies using molecular markers have highlighted the importance of gene flow between natural populations of this mosquito in Cameroon (Cohuet et al. 2005; Ayala et al. 2011b). These results are consistent with our assumptions of very strong linkage disequilibrium between those loci and the inversion.

Our inferences depend on the assumptions made in our model. We found that the four models that best-fit the data have significantly better statistical support than several alternatives: models that assume fitnesses do not change in space, that mating is random, and that there are fewer than three ecological zones. There are, however, additional alternatives that might be considered. Our model assumes that selection acts on viability but not fertility. That assumption is justified by experimental results showing that karyotypes affect larval survival in *Anopheles* (Rocca et al. 2009) and that many paracentric inversions have little if any effect on fertility in dipterans (Hoffmann and Rieseberg 2008). We have run simulations that assume selection acts on fertility rather than viability and found that the results are not much changed (results not shown). We assumed densities are equal everywhere. The position and shape of clines are affected by variation in density (Barton and Gale 1993), which could bias our parameter estimates. We have no way to estimate density, however, so it does not seem possible to make progress on this issue with the available data.

Our inferences are partly limited by the sampling protocol. Because our data come only from the wet season, we are not able to include seasonal variation in the model. The frequencies of 3Ra in *A. funestus* vary between the wet and dry seasons in the dry savannah of West Africa (Guelbeogo et al. 2009). These changes are relatively small, however, and certainly less dramatic than the frequency differences seen between the three zones in our transect. Consequently we expect that there are seasonal changes in the parameters of our model, but that the qualitative patterns from the analyses are robust. A second issue involves the geography of the samples. Our samples are limited to the one-dimensional transect along a highway. Our analyses are not able to distinguish between a one-dimensional stepping-stone and an alternative model in which mosquitoes are distributed uniformly in two-dimensional space. Consequently, we have little confidence about any inference concerning spatial distribution and migration patterns from these data.

Inversion 3Ra is only one of several common inversions in *A. funestus*. Intriguing observations are that their frequencies vary geographically, and that there is strong linkage disequilibrium between some pairs of inversions even when they occur on different chromosomes (Costantini et al. 1999; Dia et al. 2000; Ayala et al. 2011b). Disequilibria could result from epistatic selection between inversions and/or from spatial variation in selection and migration. These possibilities could be tested by extending our modeling framework to multiple inversions.

Our results can be viewed from two perspectives. On the one hand, they provide estimates of the strengths of evolutionary forces and the amount of reproductive isolation that results. An alternative view is that this study generates hypotheses about the phenotypic effects of the inversion. An important next step will be to test these hypotheses using other approaches. They might include experimental studies of the effects of the inversion on

mating behavior and viability, for example, and analyses of patterns of neutral molecular variation inside and outside of the inversion. Although the constraints of *A. funestus* discussed above make some approaches infeasible, it should be possible to pursue them in other species (such as *A. gambiae*). Together, these complementary research programs offer the prospect of quantifying the evolutionary forces that act on inversions and ultimately settling some of the oldest debates in evolutionary genetics.

ACKNOWLEDGMENTS

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19:716–723.

Alphey, L., C. B. Beard, P. Billingsley, M. Coetzee, and A. Crisanti. 2002. Malaria control with genetically manipulated insect vectors. Science 298:119–121.

Anderson, W. W., J. Arnold, D. G. Baldwin, A. T. Beckenbach, C. J. Brown, S. H. Bryant, J. A. Coyne, L. G. Harshman, W. B. Heed, D. E. Jeffrey, et al. 1991. Four decades of inversion polymorphism in *Drosophila pseudoobscura*. Proc. Natl. Acad. Sci. USA 88:10367–10371.

Ayala, D., H. Caro-Riaño, J.-P. Dujardin, N. Rahola, D. Fontenille, and F. Simard. 2011a. Chromosomal and environmental determinants of morphometric variation in natural populations of the malaria vector *Anopheles funestus* in Cameroon. Infect., Genet. Evol. 11:940–947.

Ayala, D., C. Costantini, K. Ose, G. C. Kamdem, C. Antonio-Nkondjio, J. P. Agbor, P. Awono-Ambene, D. Fontenille, and F. Simard. 2009. Habitat suitability and ecological niche profile of major malaria vectors in Cameroon. Malaria J. 8:307.

Ayala, D., M. C. Fontaine, A. Cohuet, D. Fontenille, R. Vitalis, and F. D. R. Simard. 2011b. Chromosomal inversions, natural selection and adaptation in the malaria vector *Anopheles funestus*. Mol. Biol. Evol. 28:745–758.

Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13–45 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, Oxford, U. K.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Ann. Rev. Ecol. Syst. 16:113–148.

Besansky, N. J., J. Krzywinski, T. Lehmann, F. Simard, M. Kern, O. Mukabayire, D. Fontenille, Y. Toure, and N. F. Sagnon. 2003. Semipermeable species boundaries between Anopheles gambiae and Anopheles arabiensis: evidence from multilocus DNA sequence variation. Proc. Natl. Acad. Sci. USA 100:10818–10823.

Boccolini, D., G. C. Carrara, I. Dia, F. Fortes, P. J. Cani, and C. Costantini. 2005. Chromosomal differentiation of *Anopheles funestus* from Luanda and Huambo Provinces, western and central Angola. Am. J. Trop. Med. Hygene 73:1071–1076.

Boete, C., and J. C. Koella. 2003. Evolutionary ideas about genetically manipulated mosquitoes and malaria control. Trends Parasitol. 19:32–38.

Brooke, B. D., R. H. Hunt, F. Chandre, P. Carnevale, and M. Coetzee. 2002. Stable chromosomal inversion polymorphisms and insecticide resistance in the malaria vector mosquito Anopheles gambiae (Diptera : Culicidae). J. Med. Entomol. 39:568–573.

Bryan, J. H., V. Petrarca, M. A. di Deco, and M. Coluzzi. 1987. Adult behaviour of members of the *Anopheles gambiae* complex in the Gambia with special reference to An. melas and its chromosomal variants. Parassitologia (Rome) 29:221–249.

Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociol. Methods Res. 33:261–304.

Butlin, R. K. 2005. Recombination and speciation. Mol. Ecol. 14:2621–2635.

———. 2010. Population genomics and speciation. Genetica 138:409–418.

Cassone, B. J., M. J. Molloy, C. D. Cheng, J. C. Tan, M. W. Hahn, and N. J. Besansky. 2011. Divergent transcriptional response to thermal stress by Anopheles gambiae larvae carrying alternative arrangements of inversion 2La. Mol. Ecol. 20:2567–2580.

Cheng, C., B. J. White, C. Kamdem, K. Mockaitis, C. Costantini, M. W. Hahn, and N. J. Besansky. 2012. Ecological genomics of Anopheles gambiae along a latitudinal cline: a population resequencing approach. Genetics 190:1417–1432.

Clements, A. N. 1999. The biology of mosquitoes: Development, nutrition and reproduction v. 1. Chapman & Hall, London.

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, and D. Fontenille. 2004. Population structure of the malaria vector *Anopheles funestus* in Senegal based on microsatellite and cytogenetic data. Insect Mol. Biol. 13:251–258.

Cohuet, A., I. Dia, F. Simard, M. Raymond, F. Rousset, C. Antonio-Nkondjio, P. H. Awono-Ambene, C. S. Wondji, and D. Fontenille. 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.

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

Coluzzi, M. 1982. Spatial distribution of chromosomal inversions and speciation in Anopheline mosquitoes. Pp. 143–153 *in* C. Barigozzi, ed. Mechanisms of speciation. Alan R. Liss, New York.

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

Coluzzi, M., A. Sabatini, V. Petrarca, and M. A. Dideco. 1977. Behavioral divergences between mosquitos with different inversion karyotypes in polymorphic populations of Anopheles gambiae complex. Nature 266:832–833.

———. 1979. Chromosomal differentiation and adaptation to human environments in the Anopheles gambiae complex. Trans. Roy. Soc. Trop. Med. Hygiene 73:483–497.

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

Coyne, J. A., and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. Evolution 43:362–381.

———. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution 51:295–303.

Dia, I., L. Lochouarn, D. Boccolini, C. Costantini, and D. Fontenille. 2000. Spatial and temporal variations of the chromosomal inversion

polymorphism of *Anopheles funestus* in Senegal. Parasite-Journal de la Societe Francaise de Parasitologie 7:179–184.

Diabate, A., A. Dao, A. S. Yaro, A. Adamou, R. Gonzalez, N. C. Manoukis, S. F. Traore, R. W. Gwadz, and T. Lehmann. 2009. Spatial swarm segregation and reproductive isolation between the molecular forms of Anopheles gambiae. Proc. Roy. Soc. B-Biol. Sci. 276:4215–4222.

Dobzhansky, T. 1944. Chromosomal races in *Drosophila pseudoobscura* and *Drosophila persimilis*. Carnegie Inst. Washington Publ. 554:47–144.

Enayati, A., and J. Hemingway. 2010. Malaria management: past, present, and future. Ann. Rev. Entomol. 55:569–591.

Feder, J. L., F. B. Roethele, K. Filchak, J. Niedbalski, and J. Romero-Severson. 2003. Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. Genetics 163:939–953.

Fouet, C., E. Gray, N. J. Besansky, and C. Costantini. 2012. Adaptation to aridity in the malaria mosquito Anopheles gambiae: chromosomal inversion polymorphism and body size influence resistance to desiccation. PLoS One 7:e34841.

Gavrilets, S. 2004. Fitness landscapes and the origin of species. Princeton Univ. Press, Princeton NJ.

Gibson, G., and I. Russell. 2006. Flying in tune: sexual recognition in mosquitoes. Curr. Biol. 16:1311–1316.

Gray, E. M., K. A. C. Rocca, C. Costantini, and N. J. Besansky. 2009. Inversion 2La is associated with enhanced desiccation resistance in *Anopheles gambiae*. *Malaria J.* 8:215.

Green, C. A., and R. H. Hunt. 1980. Interpretation of variation in ovarian polytene chromosomes of *Anopheles funestus* Giles, *A. parensis* Gillies, and *A. Aruni*? Genetica 51:187–195.

Guelbeogo, W. M., O. Grushko, D. Boccolini, P. A. Ouedraogo, N. J. Besansky, N. F. Sagnon, and C. Costantini. 2005. Chromosomal evidence of incipient speciation in the Afrotropical malaria mosquito *Anopheles funestus*. Med. Veterinary Entomol. 19:458–469.

Guelbeogo, W. M., N. Sagnon, O. Grushko, M. A. Yameogo, D. Boccolini, N. J. Besansky, and C. Costantini. 2009. Seasonal distribution of *Anopheles funestus* chromosomal forms from Burkina Faso. *Malaria J.* 8:239.

Guerrero, R. F., F. Rousset, and M. Kirkpatrick. 2012. Coalescence patterns for chromosomal inversions in divergent populations. Philos. Trans. Roy. Soc. B-Biol. Sci. 367:430–438.

Hoffmann, A. A., and L. H. Rieseberg. 2008. Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? Ann. Rev. Ecol. Evol. Syst. 39:21–42.

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

King, M. 1993. Species evolution: the role of chromosomal change. Cambridge Univ. Press, Cambridge, U.K.

Kirkpatrick, M. 2010. How and why chromosome inversions evolve. PLoS Biol. 8:e1000501.

Kirkpatrick, M., and N. Barton. 2006. Chromosome inversions, local adaptation, and speciation. Genetics 173:419–434.

Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. Am. Nat. 159:S22–S35.

Lehmann, T., and A. Diabate. 2008. The molecular forms of *Anopheles gambiae*: a phenotypic perspective. Infect., Genet. Evol. 8:737–746.

Lewontin, R. C., and M. J. D. White. 1960. Interaction between inversion polymorphisms of two chromosome pairs in the grasshopper. *Moraba scurra*. Evolution 14:116–129.

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

Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PloS Biol. 8:e1000500.

Manoukis, N. C., J. R. Powell, M. B. Toure, A. Sacko, F. E. Edillo, M. B. Coulibaly, S. F. Traore, C. E. Taylor, and N. J. Besansky. 2008. A test of the chromosomal theory of ecotypic speciation in Anopheles gambiae. Proc. Natl. Acad. Sci. USA 105:2940–2945.

Michel, A. P., O. Grushko, W. M. Guelbeogo, N. F. Lobo, N. Sagnon, C. Costantini, and N. J. Besansky. 2006. Divergence with gene flow in *Anopheles funestus* from the Sudan Savanna of Burkina Faso, West Africa. Genetics 173:1389–1395.

Michel, A. P., S. Sim, T. H. Q. Powell, M. S. Taylor, P. Nosil, and J. L. Feder. 2010. Widespread genomic divergence during sympatric speciation. Proc. Natl. Acad. Sci. USA 107:9724–9729.

Murray, C. J. L., L. C. Rosenfeld, S. S. Lim, K. G. Andrews, K. J. Foreman, D. Haring, N. Fullman, M. Naghavi, R. Lozano, and A. D. Lopez. 2012. Global malaria mortality between 1980 and 2010: a systematic analysis. Lancet 379:413–431.

Navarro, A., and N. H. Barton. 2003. Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation. Evolution 57:447–459.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

Olivry, J. C. 1986. Fleuves et Rivières du Cameroun. ORSTOM, Paris.

Pennetier, C., B. Warren, K. R. Dabire, I. J. Russell, and G. Gibson. 2009. "Singing on the wing" as a mechanism for species recognition in the malarial mosquito Anopheles gambiae. Curr. Biol. 20:131–136.

Perevozkin, V. P., A. A. Printseva, P. V. Maslennikov, and S. S. Bondarchuk. 2012. Genetic aspects of sexual behavior in malaria mosquitoes on the basis of specific acoustic signals at mating. Russian J. Genet. 48:587–591.

Petrarca, V., and J. C. Beier. 1992. Intraspecific chromosomal polymorphism in the Anopheles gambiae complex as a factor affecting malaria transmission in the Kisumu area of Kenya. Am. J. Trop. Med. Hygiene 46:229–237.

Petrarca, V., A. D. Nugud, M. A. E. Ahmed, A. M. Haridi, M. A. Di Deco, and M. Coluzzi. 2000. Cytogenetics of the Anopheles gambiae complex in Sudan, with special reference to An. arabiensis: relationships with East and West African populations. Med. Vet. Entomol. 14:149–164.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 2007. Numerical recipes: The art of scientific computation. Cambridge Univ. Press, Cambridge, U.K.

Rice, W. R. 1987. Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. Evol. Ecol. 1:301–314.

Rishikesh, N., M. A. Dideco, V. Petrarca, and M. Coluzzi. 1985. Seasonal variations in indoor resting Anopheles gambiae and Anopheles arabiensis in Kaduna, Nigeria. Acta Trop. 42:165–170.

Rocca, K. A. C., E. M. Gray, C. Costantini, and N. J. Besansky. 2009. 2La chromosomal inversion enhances thermal tolerance of Anopheles gambiae larvae. Malaria J. 8:147.

Sanford, M. R., B. Demirci, C. D. Marsden, Y. Lee, A. J. Cornel, and G. C. Lanzaro. 2011. Morphological differentiation may mediate mate choice between incipient species of Anopheles gambiae s.s. PLoS One 6:e27920.

Schaeffer, S. W. 2008. Selection in heterogeneous environments maintains the gene arrangement polymorphism of *Drosophila pseudoobscura*. Evolution 62:3082–3099.

Sharakhov, I., O. Braginets, O. Grushko, A. Cohuet, W. M. Guelbeogo, D. Boccolini, M. Weill, C. Costantini, N. Sagnon, D. Fontenille, et al. 2004. A microsatellite map of the African human malaria vector *Anopheles funestus*. J. Heredity 95:29–34.

Sharakhov, I. V., A. C. Serazin, O. G. Grushko, A. Dana, N. Lobo, M. E. Hillenmeyer, R. Westerman, J. Romero-Severson, C. Costantini, N. Sagnon, et al. 2002. Inversions and gene order shuffling in *Anopheles gambiae* and *A. funestus*. Science 298:182–185.

Simard, F., D. Ayala, G. C. Kamdem, M. Pombi, J. Etouna, K. Ose, J.-M. Fotsing, D. Fontenille, N. J. Besansky, and C. Costantini. 2009. Ecological niche partitioning between Anopheles gambiae molecular forms in Cameroon: the ecological side of speciation. BMC Ecol. 9: 17.

Stefansson, H., A. Helgason, G. Thorleifsson, V. Steinthorsdottir, G. Masson, J. Barnard, A. Baker, A. Jonasdottir, A. Ingason, V. G. Gudnadottir, et al. 2005. A common inversion under selection in Europeans. Nat. Genet. 37:129–137.

Toure, Y. T., V. Petrarca, S. F. Traore, A. Coulibaly, H. M. Maiga, O. Sankare, M. Sow, M. A. Di Deco, and M. Coluzzi. 1998. The distribution and inversion polymorphism of chromosomally recognized taxa of the Anopheles gambiae complex in Mali, West Africa. Parassitologia (Rome) 40:477–511.

Toure, Y. T., V. Petrarca, S. F. Traore, A. Coulibaly, H. M. Maiga, O. Sankare, M. Sow, M. A. Dideco, and M. Coluzzi. 1994. Ecological genetic studies in the chromosomal form Mopti of Anopheles gambiae s.str. in Mali, West Africa. Genetica 94:213–223.

Toure, Y. T., S. F. Traore, O. Sankare, M. Y. Sow, A. Coulibaly, F. Esposito, and V. Petrarca. 1996. Perennial transmission of malaria by the Anopheles gambiae complex in a North Sudan Savanna area of Mali. Med. Vet. Entomol. 10:197–199.

Tripet, F., J. Wright, A. Cornel, A. Fofana, R. McAbee, C. Meneses, L. Reimer, M. Slotman, T. Thiemann, G. Dolo, et al. 2007. Longitudinal survey of knockdown resistance to pyrethroid (kdr) in Mali, West Africa, and evidence of its emergence in the Bamako form of *Anopheles gambiae* s.s. Am. J. Trop. Med. Hygiene 76:81–87.

Turner, T. L., and M. W. Hahn. 2007. Locus- and population-specific selection and differentiation between incipient species of Anopheles gambiae. Molecular Biology and Evolution 24:2132–2138.

———. 2010. Genomic islands of speciation or genomic islands and speciation? Mol. Ecol. 19:848–850.

Turner, T. L., M. W. Hahn, and S. V. Nuzhdin. 2005. Genomic islands of speciation in Anopheles gambiae. PLoS Biol. 3:1572–1578.

Wang-Sattler, R., S. Blandin, Y. Ning, C. Blass, G. Dolo, Y. T. Toure, A. della Torre, G. C. Lanzaro, L. M. Steinmetz, F. C. Kafatos, et al. 2007. Mosaic genome architecture of the Anopheles gambiae species complex. PLoS One 2:e1249.

White, B. J., C. Cheng, F. Simard, C. Costantini, and N. J. Besansky. 2010. Genetic association of physically unlinked islands of genomic divergence in incipient species of Anopheles gambiae. Mol. Ecol. 19:925–939.

White, B. J., C. D. Cheng, D. Sangare, N. F. Lobo, F. H. Collins, and N. J. Besansky. 2009. The population genomics of trans-specific inversion polymorphisms in Anopheles gambiae. Genetics 183:275–288.

White, B. J., M. W. Hahn, M. Pombi, B. J. Cassone, N. F. Lobo, F. Simard, and N. J. Besansky. 2007. Localization of candidate regions maintaining a common polymorphic inversion (2La) in Anopheles gambiae. PLoS Genet. 3:2404–2414.

White, M. J. D. 1973. Animal cytology and evolution. Cambridge Univ. Press, Cambridge, MA.

Whitlock, M. C., and D. Schluter. 2009. The analysis of biological data. Roberts and Company Publishers, Greenwood Village.

Wu, C. I. 2001. The genic view of the process of speciation. J. Evol. Biol. 14:851–865.

Associate Editor: T. Lenormand

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Data from natural populations of *Anopheles funestus* used to fit the genetic model.

Figure S1. Comparison of the observed frequencies (circles) and predicted frequencies from two models for the three karyotypes.