HHS Public Access

Author manuscript

Mol Ecol. Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:

Mol Ecol. 2017 October ; 26(20): 5552–5566. doi:10.1111/mec.14335.

Chromosome arm specific patterns of polymorphism associated with chromosomal inversions in the major African malaria vector, Anopheles *funestus*

Colince Kamdem[1,*], **Caroline Fouet**[1], and **Bradley J. White**[1,#]

[1]Department of Entomology, University of California, Riverside, CA 92521

Abstract

Chromosomal inversions facilitate local adaptation of beneficial mutations and modulate genetic polymorphism, but the extent of their effects within the genome is still insufficiently understood. The genome of *Anopheles funestus*, a malaria mosquito endemic to sub-Saharan Africa, contains an impressive number of paracentric polymorphic inversions, which are unevenly distributed among chromosomes and provide an excellent framework for investigating the genomic impacts of chromosomal rearrangements. Here we present results of a fine-scale analysis of genetic variation within the genome of two weakly differentiated populations of *Anopheles funestus* inhabiting contrasting moisture conditions in Cameroon. Using population genomic analyses, we found that genetic divergence between the two populations is centered on regions of the genome corresponding to three inversions, which are characterized by high values of F_{ST}, absolute sequence divergence and fixed differences. Importantly, in contrast to the 2L chromosome arm, which is collinear, nucleotide diversity is significantly reduced along the entire length of three autosome arms bearing multiple overlapping chromosomal rearrangements. These findings support the idea that interactions between reduced recombination and natural selection within inversions contribute to sculpt nucleotide polymorphism across chromosomes in *An. funestus*.

Keywords

Anopheles funestus; chromosomal inversion; recombination; genetic divergence; nucleotide diversity

Introduction

Although much progress has been made in analyzing genetic variation among populations, less is known about the complex interactions between the different driving forces — including mutation, selection, recombination, gene flow, demography or incomplete lineage

[*]Corresponding author: colincek@ucr.edu.

[#]Present address: Verily Life Sciences, South San Francisco, CA 94080, USA

DR. CAROLINE FOUET (Orcid ID : 0000-0001-8347-4786)

Author contributions

CK, CF and BJW conceived, designed and performed the experiments. CK, CF analysed the data. CK wrote the paper.

Data Archiving Statement

Raw data (fastq files) for 132 *Anopheles funestus* individuals are available at: https://doi.org/10.5061/dryad.qp4kb

sorting — which shape nucleotide polymorphism across genomes (Begun & Aquadro 1992; Betancourt & Presgraves 2002; Nordborg & Tavar 2002; Comeron *et al.* 2008; Cutter & Payseur 2013; Gosset & Bierne 2013; Huber *et al.* 2014; Cruickshank & Hahn 2014). Most genetic variability is neutral or nearly so (Kimura 1983), but patterns of polymorphism within species are also influenced by several types of natural selection whose signatures are heterogeneous in nature and across the genome (Hill & Robertson 1966; Maynard Smith & Haigh 1974; Hudson & Kaplan 1988; Charlesworth *et al.* 1993; Gillespie 1997). One fundamental, unresolved question is how the adaptation of different gene pools to different fitness conditions imposed by spatially varying natural selection in heterogeneous environments (local adaptation) affects their genomic variation (Williams 1966; Kawecki & Dieter 2004; Savolainen *et al.* 2013).

Population genetics theory and empirical models suggest that candidate genomic regions, which contain genes underlying responses to spatially heterogeneous selection pressures exhibit high divergence among populations, skewed polymorphism or an extended correlation among alleles from different loci (linkage disequilibrium) (Lewontin & Krakauer 1973; Schlötterer 2002; Beaumont 2005; Nielsen 2005; Storz 2005; Nosil *et al.* 2009). Because migration, gene flow and recombination among diverging populations can break down locally adapted allelic combinations, genetic targets of local adaptation are often protected within genomic regions of reduced recombination – including chromosomal inversions (Noor *et al.* 2001; Rieseberg 2001; Ortiz-Barrientos *et al.* 2002; Butlin 2005; Kirkpatrick & Barton 2006; Joron *et al.* 2011; Roesti *et al.* 2012, 2013; Yeaman 2013; Nishikawa *et al.* 2015).

Structural rearrangements known as chromosomal inversions, which occur when a piece of DNA within a single chromosome breaks and rotates 180° before being reinserted in the reversed orientation (Sturtevant 1921), are widespread in plants and animals (Dobzhansky 1970). A large body of literature supports a major implication of paracentric inversions (which do not encompass the centromere) in evolutionary adaptation in a wide variety of species (see Krimbas & Powell 1992; Powell 1997; Hoffmann *et al.* 2004; Hoffmann & Rieseberg 2008; Kirkpatrick 2010 for a review). This includes many examples of polymorphic inversions whose frequencies are correlated with environmental variables and temporal changes consistent with natural selection in dipteran species (Dobzhansky 1943; Mettler *et al.* 1977; Coluzzi *et al.* 1979; Knibb 1982; Krimbas & Powell 1992; De Jong & Bochdanovits 2003; Schaeffer *et al.* 2003; Anderson *et al.* 2005; Schaeffer 2008). The adaptive potential of some rearrangements is also highlighted by experimental evolution and phenotypic studies showing a remarkable association between the inversion and several fitness-related traits (Wright & Dobzhansky 1946; Dobzhansky 1948; Rako *et al.* 2006; Hoffmann & Weeks 2007; Kennington *et al.* 2007; Lowry & Willis 2010; Lee *et al.* 2011; Fouet *et al.* 2012; Kapun *et al.* 2014, 2016).

Another important property of inversions concerns the significant role they play in genome evolution as a whole. Recombination is strongly reduced between two paired chromosomes that differ by an inversion through several mechanisms, which impede crossing over (reviewed in Roberts 1976). Local recombination rates affect a myriad of processes including selection, gene conversion, diversity and divergence throughout the genome

(Begun & Aquadro 1992; Kliman & Hey 1993; Andolfatto *et al.* 2001; Nordborg *et al.* 2005; Haddrill *et al.* 2007; Kulathinal *et al.* 2008; Comeron *et al.* 2012; Nachman & Payseur 2012; Roesti *et al.* 2012, 2013; Campos *et al.* 2014; Ortiz-barrientos *et al.* 2016). Since the pioneering work of Begun and Aquadro (1992), which demonstrated a positive correlation between recombination rate and nucleotide diversity within the genome of *Drosophila melanogaster*, multiple studies have provided empirical evidence that reduced crossing over alters genetic variation along chromosomes in a wide variety of organisms (Hellmann *et al.* 2003, 2005; Tenaillon *et al.* 2004; Begun *et al.* 2007; Kulathinal *et al.* 2008, 2009; Pegueroles *et al.* 2010; Corbett-Detig & Hartl 2012; McGaugh *et al.* 2012; Pool *et al.* 2012). Consequently, inversions, which are coldspots of recombination, have the potential to modulate patterns of genetic variation across relatively significant regions of the genome. More specifically, the presence of one or several rearrangements may reduce crossing over and diversity along an entire chromosome or chromosome arm, which in turn evolves differently from the rest of the genome. This scenario is best illustrated by the nonrecombining part of the Y chromosome in mammals, which is maintained by multiple inversions that have suppressed recombination over the entire length of the chromosome (Lahn & Page 1999). The potential of inversions to reduce recombination and diversity in large regions of the genome has long been exploited to engineer chromosomes or chromosomal arms with multiple inversions that block recombination, allowing lethal or sterile mutations to be stably maintained in a heterozygous state (balancer chromosomes) (Muller 1918; Sturtevant 1921; Hentges & Justice 2004; Edgley *et al.* 2006). Despite these clear examples, our understanding of how inversions affect polymorphism over the entire length of a chromosome remains limited in part due to the lack of empirical support across taxa. Targeted studies based on fine-scale examinations of genomic variation in species with well-characterized patterns of naturally occurring inversions are important for assessing both the genomic extent and the ubiquity of diversity reduction associated with chromosomal rearrangements.

An excellent opportunity to investigate the role of inversions in genome evolution exists in *An. funestus*, the second most important vector of *Plasmodium* parasites in the Afrotropical region behind the best known species, *An. gambiae* (Gillies & De Meillon 1968; Gillies & Coetzee 1987; Sinka *et al.* 2010; Coetzee & Koekemoer 2013). The two taxa share several characteristics including a near continental distribution, a marked preference for human hosts and intradomicillary host-seeking and resting behavior, which reflect their efficiency as vector of malaria parasites (Dia *et al.* 2013; Lanzaro & Lee 2013). Numerous studies have suggested that ongoing adaptive divergence contributes in part to the significant environmental flexibility, which underlies the widespread distribution of *An. funestus* populations. This species has been described as an amalgam of at least two relatively differentiated ecotypes whose ecology, phenotypic divergence, distribution range and role in malaria transmission remain obscure (Green & Hunt 1980; Costantini *et al.* 1999; Dia *et al.* 2000; Kamau *et al.* 2002; Boccolini *et al.* 2005; Cohuet *et al.* 2005; Michel *et al.* 2005; Ayala *et al.* 2011; Barnes *et al.* 2017). A very high number of paracentric polymorphic chromosomal inversions have been identified in polytene chromosomes of *An. funestus* via cytogenetic studies (Green & Hunt 1980; Sharakhov *et al.* 2004). Some of these rearrangements are spread along stable geographic clines and exhibit a significant deficit in

heterokaryotypes, suggesting that they may play a role in environmental adaptation (Costantini *et al.* 1999; Dia *et al.* 2000; Kamau *et al.* 2002; Boccolini *et al.* 2005; Ayala *et al.* 2011). Importantly, dozens of large rearrangements have been detected on all autosomes except for 2L (Sharakhov *et al.* 2004). This uneven distribution of inversions between chromosome arms provides an ideal framework for testing hypotheses about the impacts of rearrangements at the chromosome level via cross-chromosomal comparisons.

In this study, we have used Restriction site Associated DNA Sequencing (RAD-Seq) (Baird *et al.* 2008) to address the effects of inversions on the genetic architecture of selection and diversity across the genome of *An. funestus* populations collected from Cameroon. Although the genome-wide level of differentiation is weak, signals of divergence are apparent within three large chromosomal rearrangements whose frequencies vary along moisture gradients in this region (3R*a*, 3R*b* and 2R*h*). Genome scans show that, in contrast to the collinear 2L arm, the proportion of polymorphic sites is drastically reduced along the entire length of three autosomal arms bearing multiple polymorphic inversions. These findings suggest that the evolution of chromosomes is impacted by the combined effects of suppressed recombination and selection within paracentric polymorphic inversions in *An. funestus*.

Materials and methods

Mosquito samples

Populations of *An. funestus* that occur in Cameroon have been assimilated to two weakly differentiated ecotypes distributed along a moisture gradient spanning the whole country (Costantini *et al.* 1999; Cohuet *et al.* 2005; Ayala *et al.* 2011). We sequenced 132 mosquitoes collected from two locations separated by ~500km, representative of the savannah and the forest ecogeographic domains (Fig. 1A and Table S1. In Mfou, a small city of the forest region, *An. funestus* mosquitoes breed in an artificial lake and maintain abundant populations throughout the year. In Tibati, which lies within the forest-savannah transition zone, several artificial lakes also provide abundant breeding sites for dense and perennial populations of *An. funestus*. Between August and November 2013, we used several sampling methods (Service 1993) to collect *An. funestus* larvae in breeding sites and adults seeking human hosts in and around human dwellings at night, or resting indoors during daytime. (Fig. 1A and Table S1). With this diversified sampling, we aimed to maximize the chances that our sample represents a good approximation of the genetic diversity of *An. funestus* populations, as genetic differentiation within malaria vectors species sometimes overlap with subtle microgeographic and temporal segregations (e.g. Riehle *et al.* 2011). *An. funestus* belong to a large taxonomic unit comprising at least seven taxa that are identified by slight morphological variations and a diagnostic PCR based on mutations of the ribosomal DNA (Gillies & De Meillon 1968; Gillies & Coetzee 1987; Cohuet *et al.* 2003). We verified, using morphology and PCR, that all samples included in this study belonged to the nominal species and only important malaria vector of the group: *An. funestus*.

Library preparation, sequencing and SNP identification

We extracted genomic DNA of larvae and adult specimens with the DNeasy Blood and Tissue kit (Qiagen) and the Zymo Research MinPrep kit, respectively. Double-digest

Restriction-site Associated DNA (ddRAD) libraries were prepared as described in Kamdem *et al.* (2017) following a modified protocol of Peterson *et al.* (2012) and single-end sequenced to 100 base reads on Illumina HiSeq2000.

Illumina sequences were sorted according to barcode tag and filtered using the *process_radtags* program of the Stacks v 1.35 software (Catchen *et al.* 2013). Reads with ambiguous barcode, inappropriate restriction site or low sequencing quality score were removed. We checked the depth of sequencing coverage of the final data set after all filtering steps using VCFtools (Danecek *et al.* 2011). To call and genotype SNPs, we first aligned the remaining high-quality reads to the *An. funestus* reference sequence with GSNAP (Wu & Nacu 2010), with a maximum of five mismatches allowed and terminal alignments prevented. We then used Stacks to build consensus RAD loci and to identify SNPs within each locus. We set the minimum number of reads required to form a "stack" to three and allowed two mismatches during catalogue creation. Following assembly and genotyping, the polymorphism data was further filtered to maximize data quality. To do this, we retained only RAD loci scored in every population and in at least 60% of individuals within each population and we randomly selected one SNP per locus for further analyses. SNP files in different formats used for downstream analyses were created with the *populations* program in Stacks, PLINK v 1.09 and PGDSpider v 2.0.8.2 (Purcell *et al.* 2007; Lischer & Excoffier 2012; Catchen *et al.* 2013).

Population structure and genetic divergence

We examined the genetic relatedness among individuals with a Principal Component Analysis (PCA), a neighbor-joining (NJ) tree and the STUCTURE v 2.3.4 software (Pritchard *et al.* 2000) using SNPs identified in Stacks. We utilized the R package *adegenet* to implement the PCA and *ape* to compute a genotype-based Euclidian distance matrix between individuals and to infer individual-based NJ networks (Paradis *et al.* 2004; Jombart 2008; R Development Core Team 2016). In STRUCTURE, we ran five replicates of 1 to 10 assumed clusters (k). Each run consisted of 200000 iterations, and the first 50000 iterations were discarded as burn-in. CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007) was used to aggregate results across multiple STRUCTURE runs and the clustering results were visualized graphically using DISTRUCT v1.1 (Rosenberg 2004). To find the optimal number of genetically distinct clusters, we used both the Discriminant Analysis of Principal Component (DAPC) implemented in *adegenet* and the ad hoc statistic DeltaK of Evanno *et al.* (2005) (Evanno *et al.* 2005; Earl & VonHoldt 2012).

To assess the level of genetic differentiation between savannah and forest populations, we estimated the overall F_{ST} (Weir & Cockerham 1984) in Genodive v1.06 (Meirmans & Van Tienderen 2004) using a subset of 2000 randomly selected SNPs. To examine to what extent the geographic region contributes to the genetic variance among samples, we conducted a hierarchical analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) in GenoDive. We used 10000 permutations to assess the statistical significance of F_{ST} and AMOVA.

Genomic targets of selection

The current *An. funestus* draft genome assembly consists of 12243 scaffolds ranging from 2334 to 3334433bp in length (Giraldo-Calderon *et al.* 2015; Neafsey *et al.* 2015). Therefore, to perform genome scans and inspect footprints of selection throughout the genome, we ordered and concatenated 104 long scaffolds whose positions on the physical map have been inferred via alignment and orthology (Neafsey *et al.* 2015) and we created "pseudo-chromosomes" corresponding to the five chromosome arms of *An. funestus*. These mapped scaffolds accounted for 33% in length of the approximate size of the *An. funestus* genome (Fig. S1 and Table S2).

To delineate genomic signatures of selection, we performed an outlier analysis in order to detect genomic regions exhibiting exceptional differentiation or diversity that are putative targets of selection (Lewontin & Krakauer 1973; Storz 2005). Locus-specific F_{ST} values between savannah and forest populations were estimated with the *populations* program in *Stacks* and visualized graphically using the subset of SNPs located on mapped scaffolds. SNPs with F_{ST} values above the top 1% of the empirical distribution were considered as outliers of genetic differentiation. Loci with unusually low or high F_{ST} values relative to neutral expectations were also detected using the coalescence-based method FDIST2 (Beaumont & Nichols 1996) as implemented in LOSITAN (Antao *et al.* 2008). The mean neutral F_{ST} across all SNPs was approximated by choosing the "neutral mean F_{ST} option" with 99% confidence interval in LOSITAN. We ran LOSITAN with 100000 simulations and assumed a false discovery rate of 0.1 to minimize the number of false positives. F_{ST} values are dependent on within-population genetic diversity, which may bias estimates of the level of divergence among populations (Noor & Bennett 2009; Cruickshank & Hahn 2014). To alleviate this effect in our analyses, we used two complementary statistics to assess the degree of genetic divergence among populations across the genome — the absolute sequence divergence (d_{xy}) and the proportion of fixed differences between populations (d_f). Both statistics were estimated in ngsTools using genotype likelihood without SNP calling (Fumagalli *et al.* 2014), and kernel smoothed values were visualized in non-overlapping 90-kb windows along pseudo-chromosomes in R.

To inspect genomic patterns of genetic diversity and allele frequency spectra, we calculated pairwise nucleotide diversity (θ_π), Watterson's estimator of segregating sites (θ_w) and Tajima's *D* across RAD loci located on mapped scaffolds in ANGSD v 0.612 (Korneliussen *et al.* 2014). This program derives diversity and allele frequency spectrum statistics using genotype likelihoods without SNP calling, thereby alleviating some of the uncertainties and biases associated with SNP calling in low coverage Next Generation Sequencing (Korneliussen *et al.* 2013). To gain a genome-wide view and identify genomic regions of exceptional diversity and/or skewed allele frequency spectra, average values of θ_π, θ_w and Tajima's *D* were determined in non-overlapping 90-kb windows and plotted along pseudo-chromosomes.

Results

Genetic differentiation within *An. funestus*

Using alignments to the draft reference genome, we assembled 490871 unique RAD loci. We identified a total of 10687 high-quality biallelic markers by randomly choosing one SNP across loci that were present in all populations and in at least 60% of individuals within every population. A NJ tree and the first three PCA axes based on these variants clearly distinguished two genetic clusters likely corresponding to the two ecotypes previously described in Cameroon and in several other countries with a diversity of genetic markers (Costantini *et al.* 1999; Dia *et al.* 2000; Kamau *et al.* 2002; Boccolini *et al.* 2005; Cohuet *et al.* 2005; Michel *et al.* 2005; Ayala *et al.* 2011; Barnes *et al.* 2017) (Fig. 1B and 1C). The method of Evanno *et al.* (2005) and DAPC confirmed the occurrence of two to three distinct gene pools reflecting the ecological divergence known between forest and savannah populations in Cameroon (Fig. 1D and 1E). However, despite this apparent geographic segregation, the overall genetic differentiation is low between the two putative subgroups (F_{ST} = 0.033, $p < 0.005$). Consistent with weak genetic divergence, STRUCTURE analyses revealed a single cluster of individuals with admixed ancestry at k = 2 (Fig. 1F). The moderate genetic differentiation between savannah and forest populations is also illustrated by the results of an AMOVA, which show that the greatest proportion of the genetic variance (98.2%, $p < 0.005$) among our samples is explained by within-individual variations. The geographic origin of individuals accounts for less than 2% of the genetic variance. Overall, the genetic differentiation of *An. funestus* in Cameroon suggests a low level of genomic divergence, consistent with extensive gene flow and/or recent split between the two putative ecotypes.

Divergence within chromosomal inversions

We scanned the genome of savannah and forest populations to identify the few regions of the genome that diverge within a largely homogeneous background. We found that, in our panel of 10687 SNPs, highly differentiated loci are non-randomly distributed, the highest F_{ST} values being clustered in genomic regions bearing known polymorphic chromosomal inversions that have been previously verified in specimen collected from this area (Cohuet *et al.* 2005; Ayala *et al.* 2011) (Fig. 2 and Fig. S1). We identified a total of 107 outliers that fell above the 99th percentile of the empirical distribution of F_{ST}, including 31 SNPs that were successfully mapped to concatenated scaffolds and were confirmed as statistical outliers in LOSITAN. F_{ST} outliers were located exclusively on two chromosome arms containing nearly 20 polymorphic inversions (3R and 2R) (Fig. 2 and Fig. S1) (Sharakhov *et al.* 2004). Moreover, in contrast to 2L, 3L and X, which have no discriminatory power, SNPs that mapped to 3R or 2R reproduce the segregation observed between the savannah and the forest at the genome level (Fig. 3). The number of SNPs with F_{ST} values above the 1% threshold also varies between the 2R and 3R (9 against 22) implying that mutations or structural variants of the 3R arm are more strongly correlated with the genetic differentiation between the two populations. Overall, the genomic distribution of F_{ST} outliers suggests that divergent genomic regions between ecotypes are located within polymorphic inversion on two chromosome arms.

To further explore the role of inversions in genetic divergence, we examined the population structure separately for three large chromosomal rearrangements present on 3R (3R*a*, 3R*b* and 3R*d*) and for the 2R*h* that occurs on 2R. Although several inversions coexist and overlap along 2R, six out of nine F_{ST} outliers identified on this chromosome mapped to scaffolds specific to portions of the 2R*h* inversion, making it an interesting candidate locus. SNPs located within the 3R*a* assign more that 90% of individuals to their respective ecotypes, and the 3R*b* also separates a significant proportion of individuals between both sampling sites (Fig. 3). In addition, consistent with the genomic pattern of differentiation, which indicates that 83% of F_{ST} outliers map to these two inversions (Fig. 2), the genome-wide distribution of d_{xy} and fixed differences shows clear peaks within these two large inversions (Fig. 4). The population structure of 2R*h* is comparable with that of 3R*b*, suggesting that this inversion also contributes to the savannah-forest segregation of *An. funestus* populations in Cameroon. Conversely, SNPs identified outside the 3R*a* and 3R*b*, or within the 3R*d* have no discriminatory power (Fig. 3). In summary, the population structure of inversions on 2R and 3R reflects the genomic distribution of differentiated loci, which suggests that divergence centered on three chromosomal rearrangements (3R*a*, 3R*b* and 2R*h*) maintains some level of genomic integrity despite low genome-wide differentiation between ecotypes of *An. funestus* in this geographic area. This divergence translates into F_{ST} values, which increase from 0.033 at the genome level to 0.053, 0.08 and 0.22 in 2R*h*, 3R*b* and 3R*a*, respectively. Similarly, the proportion of the genetic variance explained by the geographic origin of samples rises from 1.8% for the genome-wide SNPs to 3.2%, 4.7% and 12.9% when only variants present within the 2R*h*, 3R*b* and 3R*a* rearrangements, respectively, are included.

Chromosome arm-specific diversity associated with inversions

We investigated patterns of polymorphism across the genome using scans based on estimates of nucleotide diversity (θ_π and θ_w) and Tajima's *D* (Fig. 5). The proportion of polymorphic sites (θ_w) varies substantially between chromosome arms, the highest values being found on the only collinear chromosome arm (2L) (Fig. 5 and Fig. 6). Precisely, the average value of θ_w differs significantly between 2L and each of the three other autosomes (Wilcoxon rank sum test, $p < 0.001$) (Fig. 6). θ_w is reduced by 34.1% on 3R, 24.0% on 2R and 13.2% on 3L relative to the 2L arm, regardless of the sampling location. Estimates of the number of polymorphic sites along each of the five chromosome arms are also significantly different between Mfou (forest) and Tibati (savannah) populations (Wilcoxon rank sum test, $p < 0.001$). Additionally, the drop of the average θ_w on inverted autosomes relative to 2L is 34.7% on 3R, 29.2% on 2R and 13.2% on 3L in Mfou compared with 33.5% on 3R, 18.8% on 2R and 13.2% on 3L in Tibati (Fig. 5 and Table 1). A different demographic history between ecotypes or diverse other factors including the variation in sample size between the two locations can account for this difference in the level of polymorphism between populations. Nevertheless, since most of the polymorphic chromosomal inversions (notably 3R*a* and 3R*b*) are fixed in forest populations and fluctuate in Tibati (Cohuet *et al.* 2005; Ayala *et al.* 2011), it is likely that variations in karyotype frequencies of inversions contribute at least in part to the stronger reduction of nucleotide diversity on inverted autosomes observed in Mfou compared with Tibati. The X chromosome bears no known chromosomal rearrangement, but likely due to its particular divergence and its distorted effective population size, its average θ_w is also weak compared with 2L (21.1% reduction

relative to 2L) regardless of the sampling site. A significant reduction in pairwise nucleotide diversity (θ_π) relative to the 2L is found on chromosomes 3R, X and 2R to a lesser extent, but there are no clear patterns among chromosomes and sampling locations as is the case with θ_w (Fig. 5, Fig. 6 and Table 1).

Local recombination rates are strongly correlated with genetic diversity and the effective population size (Begun & Aquadro 1992). Therefore, it is possible and parsimonious that the strong variation in polymorphism observed among autosomes in *An. funestus* is due to the presence of multiple overlapping inversions that are under selection in different geographic contexts throughout the species' range. These inversions reduce meiotic crossover and genetic variation along the 2R, 3R and 3L autosomes, which contain multiple rearrangements. For instance, three inversions (3R*a*, 3R*b* and 3R*d*) account for at least 75% of the length of the 3R arm (Fig. S1). Similarly, nearly 80% of the length of 2R and 3L is affected by chromosomal rearrangements. The idea that the combined effects of low recombination and selection within multiple inversions contributes to reduce polymorphism throughout the three inverted autosomal arms is supported by the pattern of diversity observed on the collinear chromosome arm 2L, which acts as a negative control. Conceivably, polymorphism is highest on this chromosome arm due to the absence of inversions, which preserves more substantial levels of recombination.

The genome-wide diversity falls in the range described in other highly polymorphic *Anopheles* populations with RADseq markers (θ_w = 0.0083 and θ_π = 0.0042 in Mfou; θ_w = 0.0097 and θ_π = 0.0043 in Tibati) (O'Loughlin *et al.* 2014; Fouet *et al.* 2017; Kamdem *et al.* 2017). The range of values of Tajima's *D* (from −2.47 to −1.58 among Tibati samples and from −2.20 to −1.19 in Mfou) was shifted towards negative values suggestive of recent population expansion leading to an accumulation of low-frequency variants, although the effects of population structure, selective sweeps, background selection or changes in sample size can contribute to skewed site frequency spectra across the genome (Donnelly *et al.* 2001; Thornton 2005; Gattepaille *et al.* 2013) (Fig. 5). Unsurprisingly, Tajima's *D* is more negative on the 2L chromosome, which exhibits the greatest difference between the number of polymorphic sites and the pairwise nucleotide diversity in both ecotypes.

Discussion

Recombination, inversions and genetic divergence

Our analysis based on more than 10000 genome-wide SNPs revealed that *An. funestus* populations collected from two different moisture conditions in Cameroon are very weakly differentiated. These results corroborate two previous studies based on microsatellites and polymorphic chromosomal inversions, which have found similar levels of divergence between *An. funestus* populations in this region (Cohuet *et al.* 2005; Ayala *et al.* 2011). The population genetic structure of this mosquito throughout Africa remains largely unresolved, but based on many previous studies in different parts of the continent, it has been proposed that at least two ecotypes segregate within this species (Green & Hunt 1980; Costantini *et al.* 1999; Dia *et al.* 2000; Kamau *et al.* 2002; Boccolini *et al.* 2005; Cohuet *et al.* 2005; Michel *et al.* 2005; Ayala *et al.* 2011; Barnes *et al.* 2017). These studies have also indicated that polymorphic chromosomal inversions play a key role in ecological divergence in *An.*

funestus (Green & Hunt 1980; Costantini *et al.* 1999; Dia *et al.* 2000; Cohuet *et al.* 2005; Ayala *et al.* 2011). The 3R*a* and 3R*b* inversions in particular have long been suspected to be strongly associated with ecological divergence due to the clinal distribution of alternative karyotypes and the significant deficit in heterokaryotypes observed in several countries (Costantini *et al.* 1999; Dia *et al.* 2000; Ayala *et al.* 2011). Here we provide the first genomic evidence that genetic divergence within this species is limited to a few loci bearing large chromosomal rearrangements. Our conclusions come with some caveats due to the potential limitations of the sampling scheme, the reference genome and the RAD sequencing approach we used. For instance, it is likely that some genomic signatures of selection are not detected because of the limited genomic coverage of RAD tags and pseudo-chromosomes (Arnold *et al.* 2013; Tiffin & Ross-Ibarra 2014). Bearing in mind these caveats, the clustering of genetic differentiation within the 3R*a*, 3R*b* and 2R*h* inversions appears to be consistent with the hypothesis that these rearrangements are the major genetic targets of divergent selection in Cameroon.

The phenotypic, behavioral or ecological variations associated with inversions in *An. funestus* remain unknown. In principle, the segregation between the arid zone and the humid forest may potentially involve fitness traits and phenotypes contributing to thermal tolerance. Indeed, examples of inversions whose alternative karyotypes provide selective advantages in different moisture conditions are well known in dipteran species (e.g. the 2L*a* inversion and to a lesser extent the 2R*b* in *An. gambiae*) (Coluzzi *et al.* 1979; Gray *et al.* 2009; Lee *et al.* 2009; Fouet *et al.* 2012; Cheng *et al.* 2012). Comparative cytogenetic studies also showed a high proportion of conserved genes between the 3R*b* in *An. funestus* and the 2L*a* in *An. gambiae* and between the 2R*h* in *An. funestus* and the 2R*b* in *An. gambiae* (Sharakhova *et al.* 2011). Although this hypothesis awaits thorough investigation, the two pairs of inversions may have inherited similar phenotypes via a common ancestor or via convergent evolution between *An. funestus* and *An. gambiae*.

Theoretical models and empirical data support the idea that, at early stages of divergence as in the case of *An. funestus* ecotypes, genetic differentiation is restricted to a few genomic regions because the effects of natural selection at this step are very localized (Feder *et al.* 2012; Nosil & Feder 2012; Andrew & Rieseberg 2013; Seehausen *et al.* 2014). Also, selection at a locus affects the level of genetic differentiation among populations and reciprocally the degree of subdivision within a population impacts patterns of variation at selected loci (Lewontin & Krakauer 1973; Charlesworth *et al.* 1997; Slatkin & Wiehe 1998; Majewski & Cohan 1999; Kim & Maruki 2011; Schneider & Kim 2013). In fact, it has been proposed that selection modulates levels of divergence locally and globally across the genome through two types of genetic hitchhiking — divergence hitchhiking and genomic hitchhiking (Feder *et al.* 2012). Divergent hitchhiking (the spread of a mutation and linked neutral sequences within a population) (Maynard Smith & Haigh 1974) causes the diffusion of favorable alleles at loci important for local adaptation and ecological differentiation. Limited gene flow that occurs between divergent lineages can naturally lead to genomic hitchhiking, which reflects the global accumulation of differences at the genome level through different selective or demographic mechanisms (Feder *et al.* 2012). In *An. funestus*, it is plausible that divergent hitchhiking within the three important inversions, 3R*a*, 3R*b* and

2R*h*, is so far strongly counterbalanced by extensive gene flow and the weak overall number of selected loci across the genome, leading to minimal genomic divergence among ecotypes.

Although some controversy persists (see McGaugh *et al.* 2012), the idea that strongly differentiated genomic regions between species or between populations within the same species accumulate disproportionately in regions of low genetic recombination including chromosome centers and chromosomal rearrangements has been strongly supported in many species notably humans, *Drosophila* flies, stickleback fish and maize (Hellmann *et al.* 2003, 2005; Tenaillon *et al.* 2004; Kulathinal *et al.* 2008; Cai *et al.* 2009; Keinan & Reich 2010; Nachman & Payseur 2012; McGaugh & Noor 2012; Roesti *et al.* 2012, 2013). This coincidence can be explained by a mechanical effect of recombination, which reduces genetic polymorphism within populations and thereby generates sequence divergence between populations as a simple byproduct of diversity reduction (Begun & Aquadro 1992; Roesti *et al.* 2012). Concordance between limited rate of crossing over and increased divergence may also be directly correlated with more pronounced effects of divergent selection in genomic regions in which recombination is less frequent (Charlesworth *et al.* 1997; Charlesworth 1998; Nachman 2002; Cutter & Payseur 2013; Cruickshank & Hahn 2014). Within inversions in particular, reduced recombination enhances divergent selection acting on locally adapted gene complexes, which co-segregate with or without epistatic interactions (Dobzhansky 1950; Kirkpatrick & Barton 2006).

Recombination, inversions and genetic polymorphism

We have found chromosome arm specific patterns of polymorphism in *An. funestus* characterized by a significant difference between collinear and inverted chromosomes. We first conducted several tests to insure that uncertainties associated with our sequencing and analytical approach cannot account for the observed disparities in genetic polymorphism between chromosomes. We compared the distribution and the length of mapped scaffolds and confirmed that each chromosome arm is represented by a substantial number of long scaffolds (Table S2 and Fig. S1). The length of pseudo-chromosomes is also proportional to the length of chromosome arms, ranging from 6.8 Mb on X (the smallest) to 27.4 Mb on 2R (the largest). Therefore, we can reasonably rule out the possibility that the chromosome-bias diversity is due to a systematic error associated with our reference sequence. Additionally, mapped scaffolds are evenly distributed along the length of chromosomes, suggesting that such important variations in the amount of polymorphism between chromosome arms are not due to centromere- and/or telomere-proximal effects (Fig. S1) (Aguade *et al.* 1989; Stephan & Langley 1989). Finally, we noted that the mean depth of sequencing coverage per mapped scaffolds was consistent across the five chromosome arms (Fig. S2), implying that chromosome-specific diversity is not simply a covariate of sequencing biases.

As estimates of the number of segregating sites (θ_w) also known as the population-scaled mutation rate (Watterson 1975) are the most drastically affected by between-chromosome variations observed in *An. funestus*, the stark contrast in the amount of polymorphism among inverted and collinear autosomes may in theory be due to lower mutation rates on inverted chromosomes. However, as shown in other insect species, such a variability in mutation rates between chromosomes or between large segments of the genome is unlikely

(Begun *et al.* 2007; Keightley *et al.* 2015). Instead, the occurrence of chromosome-specific patterns of diversity in *An. funestus* is consistent with the presence of weakly recombining autosomes bearing multiple overlapping chromosomal rearrangements. These same causes produce similar effects along the non-recombining portion of the Y chromosome in mammals (Lahn & Page 1999) or along balancer chromosomes described in *Drosophila*, *Caenorhabditis* and rodent species (Muller 1918; Hentges & Justice 2004; Edgley *et al.* 2006). Indeed, population genetic analyses across genomes of diverse taxa have found a positive correlation between the rate of recombination and genetic variation (Aguade *et al.* 1989; Stephan & Langley 1989; Begun & Aquadro 1992; Hellmann *et al.* 2003, 2005; Tenaillon *et al.* 2004; Begun *et al.* 2007; Kulathinal *et al.* 2008, 2009). This correlation translates into specific patterns of diversity that can be observed at the species, genomic or chromosomal levels. At the species level, plants and some animal species, which reproduce by self-fertilization have reduced overall genomic diversity due to low genetic recombination (Nordborg *et al.* 1996; Akhunov *et al.* 2010; Cutter & Choi 2010; Andersen *et al.* 2012; Thomas *et al.* 2015). The local effects of recombination on genetic polymorphism are also particularly evident across centromere and telomeres of chromosomes and within genomic regions bearing structural rearrangements (Aguade *et al.* 1989; Stephan & Langley 1989; Andolfatto *et al.* 2001; Corbett-Detig & Hartl 2012; Pool *et al.* 2012).

The relationship between diversity and crossing over has been ascribed to two processes: selection and mutagenesis. Recombination may influence polymorphism because of new mutations created during crossing over, which increase diversity (the mutagenic effect of recombination) (Magni & Von Borstel 1962). However, multiple studies that have analyzed fine-scale recombination in humans, yeast, *Arabidopsis*, *Anopheles*, *Drosophila*, and *Caenorhabditis* have undermined the notion that recombination, or some correlate, generally exerts mutagenic effects on genomes (Betancourt & Presgraves 2002; Stump *et al.* 2005; Spencer *et al.* 2006; Wright *et al.* 2006; Begun *et al.* 2007; Noor 2008; Denver *et al.* 2009; Cutter & Choi 2010; Mcgaugh *et al.* 2012). Alternatively, recombination may affect diversity indirectly by modulating the effects of positive or negative selection across the genome. Indeed, the hitchhiking of favorable alleles (selective sweep) or the removal of recurrent deleterious mutations (background selection) affect linked neutral sequences more strongly in low-recombination genomic regions, which gives rise to a positive correlation between diversity and recombination rate (Maynard Smith & Haigh 1974; Kaplan *et al.* 1989; Begun & Aquadro 1992; Charlesworth *et al.* 1993, 1997; Nordborg *et al.* 1996; Andolfatto 2001; Nachman 2002). Another indirect effect of recombination on genetic diversity may be driven by the Hill-Roberston interference, which occurs when two sites under weak selection are in physical linkage and as a result, selection at one site interferes with selection at another site (Hill & Robertson 1966; Felsenstein 1974). Hill-Roberston interference can also be thought of as a reduction in the effective population size (N_e) caused by selection at linked loci (Comeron *et al.* 2008; Cutter & Payseur 2013; Castellano *et al.* 2016). Recombination, by alleviating interference between linked sites, alleviates this reduction in N_e leading to a positive correlation between recombination rate and levels of neutral polymorphism. Because of the relationship between recombination and polymorphism, measures of the skew in the allele frequency spectrum, such as Tajima's D values (Tajima 1989), are expected to positively correlate with the rate of recombination (Braverman *et al.* 1995). The

An. funestus genome shows a strong relationship between Tajima's *D* values and recombination, and as expected, the chromosome 2L, whose recombination is not limited by inversions, exhibits the sharpest skew in the AFS.

Conclusions and perspectives

We found evidence that, among the many chromosomal rearrangements identified in *An. funestus*, genomic footprints of divergence are centered on three inversions that are potential targets of selection among ecotypes in Cameroon. The other rearrangements are likely either cosmopolitan or endemic inversions under selection at various geographic extents that have yet to be resolved. Our data support the idea that interactions between recombination and selection — which amplify the effects of selective sweeps or background selection within genomic regions where recombination is blocked by multiple chromosomal inversions — account for the strong disparity in nucleotide diversity observed between autosomes in this mosquito.

To deepen our understanding of the adaptive role of inversions and their contribution to chromosome specific patterns of diversity, a complete reference genome assembly and extensive sampling across the species range are needed. These resources will make it possible to design more sensitive tests including the functional characterization of footprints of selection and their detailed signatures among individuals and populations. The presence of hallmarks of both reduced recombination and linked selection across large genomic sequences in *An. funestus* highlights the important contribution of multiple aspects of linkage in genome evolution. These aspects have significant implications for the detection of genomic signatures of adaptation in species whose genome contains multiple polymorphic chromosomal inversions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the associate editor and six reviewers for their constructive comments. This work was supported by NIH grant 1R01AI113248 to BJW.

References

Aguade M, Miyashita N, Langley CH. Reduced variation in the yellow-achaete-scute region in natural populations of Drosophila melanogaster. Genetics. 1989; 122:607–15. [PubMed: 17246506]

Akhunov ED, Akhunova AR, Anderson OD, et al. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. BMC genomics. 2010; 11:702. [PubMed: 21156062]

Andersen EC, Gerke JP, Shapiro Ja, et al. Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. Nature genetics. 2012; 44:285–90. [PubMed: 22286215]

Anderson AR, Hoffmann AA, McKechnie SW, Umina PA, Weeks AR. The latitudinal cline in the In (3R) Payne inversion polymorphism has shifted in the last 20 years in Australian Drosophila melanogaster populations. Molecular Biology and Evolution. 2005; 14:851–858.

Andolfatto P. Adaptive hitchhiking effects on genome variability. Current Opinion in Genetics and Development. 2001; 11:635–641. [PubMed: 11682306]

Andolfatto P, Depaulis F, Navarro A. Inversion polymorphisms and nucleotide variability in Drosophila. Genetical research. 2001; 77:1–8. [PubMed: 11279826]

Andrew RL, Rieseberg LH. Divergence is focused on few genomic regions early in speciation: incipient speciation of sunflower ecotypes. Evolution; international journal of organic evolution. 2013; 67:2468–82. [PubMed: 24033161]

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC bioinformatics. 2008; 9:323. [PubMed: 18662398]

Arnold B, Corbett-Detig RB, Hartl D, Bomblies K. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Molecular Ecology. 2013; 22:3179–3190. [PubMed: 23551379]

Ayala D, Fontaine MC, Cohuet A, et al. Chromosomal inversions, natural selection and adaptation in the malaria vector Anopheles funestus. Molecular biology and evolution. 2011; 28:745–758. [PubMed: 20837604]

Baird NA, Etter PD, Atwood TS, et al. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. PLoS ONE. 2008; 3:e3376. [PubMed: 18852878]

Barnes KG, Weedall GD, Ndula M, et al. Genomic Footprints of Selective Sweeps from Metabolic Resistance to Pyrethroids in African Malaria Vectors Are Driven by Scale up of Insecticide-Based Vector Control. PLoS Genetics. 2017; 13(2):e10, 1–22.

Beaumont MA. Adaptation and speciation : what can F st tell us? Trends in Ecology & Evolution. 2005; 20:435–440. [PubMed: 16701414]

Beaumont MA, Nichols RA. Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society B: Biological Sciences. 1996; 263:1619–1626.

Begun D, Aquadro C. Levels of naturally occurring DNA polymorphism correlate with recombination rates in D. melanogaster. Nature. 1992; 356:519–520. [PubMed: 1560824]

Begun DJ, Holloway AK, Stevens K, et al. Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans. PLoS biology. 2007; 5:e310. [PubMed: 17988176]

Betancourt AJ, Presgraves DC. Linkage limits the power of natural selection in Drosophila. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:13616–20. [PubMed: 12370444]

Boccolini D, Carrara GC, Dia I, et al. Chromosomal differentiation of Anopheles funestus from Luanda and Huambo Provinces, western and central Angola. American Journal of Tropical Medicine and Hygiene. 2005; 73:1071–1076. [PubMed: 16354814]

Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics. 1995; 140:783–796. [PubMed: 7498754]

Butlin RK. Recombination and speciation. Molecular Ecology. 2005; 14:2621–2635. [PubMed: 16029465]

Cai JJ, Macpherson JM, Sella G, Petrov DA. Pervasive hitchhiking at coding and regulatory sites in humans. PLoS Genetics. 2009; 5

Campos JL, Halligan DL, Haddrill PR, Charlesworth B. The relation between recombination rate and patterns of molecular evolution and variation in drosophila melanogaster. Molecular Biology and Evolution. 2014; 31:1010–1028. [PubMed: 24489114]

Castellano D, Coronado-Zamora M, Campos JL, Barbadilla A, Eyre-Walker A. Adaptive evolution is substantially impeded by hill-Robertson interference in drosophila. Molecular Biology and Evolution. 2016; 33:442–455. [PubMed: 26494843]

Catchen J, Hohenlohe Pa, Bassham S, Amores A, Cresko W. Stacks: An analysis tool set for population genomics. Molecular Ecology. 2013; 22:3124–3140. [PubMed: 23701397]

Charlesworth B. Measures of divergence between populations and the effect of forces that reduce variability. Molecular Biology and Evolution. 1998; 15:538–543. [PubMed: 9580982]

Charlesworth B, Morgan MT, Charlesworth D. The effect of deleterious mutations on neutral molecular variation. Genetics. 1993; 134:1289–1303. [PubMed: 8375663]

Charlesworth B, Nordborg M, Charlesworth D. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res., Camb. 1997; 70:155–174.

Cheng C, White BJ, Kamdem C, et al. Ecological genomics of Anopheles gambiae along a latitudinal cline: a population-resequencing approach. Genetics. 2012; 190:1417–32. [PubMed: 22209907]

Coetzee M, Koekemoer LL. Molecular systematics and insecticide resistance in the major African malaria vector Anopheles funestus. Annual Review of Entomology. 2013; 58:393–412.

Cohuet A, Dia I, Simard F, et al. Gene flow between chromosomal forms of the malaria vector Anopheles funestus in Cameroon, Central Africa, and its relevance in malaria fighting. Genetics. 2005; 169:301–311. [PubMed: 15677749]

Cohuet A, Simard F, Toto J-C, et al. Species identification within the Anopheles funestus group of malaria vectors in Cameroon and evidence for a new species. The American journal of tropical medicine and hygiene. 2003; 69:200–205. [PubMed: 13677376]

Coluzzi M, Sabatini A, Petrarca V, Di Deco M. Chromosomal differentiation and adaptation to human environments in the Anopheles gambiae complex. Trans R Soc Trop Med Hyg. 1979; 73:483–97. [PubMed: 394408]

Comeron JM, Ratnappan R, Bailin S. The Many Landscapes of Recombination in Drosophila melanogaster. PLoS Genetics. 2012; 8:33–35.

Comeron JM, Williford a, Kliman RM. The Hill-Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. Heredity. 2008; 100:19–31. [PubMed: 17878920]

Corbett-Detig RB, Hartl DL. Population Genomics of Inversion Polymorphisms in Drosophila melanogaster. PLoS Genetics. 2012; 8

Costantini C, Sagnon N, Ilboudo-Sanogo E, Coluzzi M, Boccolini D. Chromosomal and bionomic heterogeneities suggest incipient speciation in Anopheles funestus from Burkina Faso. Parassitologia. 1999; 41:595–611. [PubMed: 10870569]

Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Molecular ecology. 2014; 23:3133–57. [PubMed: 24845075]

Cutter AD, Choi JY. Natural selection shapes nucleotide polymorphism across the genome of the nematode Caenorhabditis briggsae Natural selection shapes nucleotide polymorphism across the genome of the nematode Caenorhabditis briggsae. Genome Research. 2010:1103–1111. [PubMed: 20508143]

Cutter AD, Payseur BA. Genomic signatures of selection at linked sites: unifying the disparity among species. Nature reviews. Genetics. 2013; 14:262–74. [PubMed: 23478346]

Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. Bioinformatics. 2011; 27:2156–2158. [PubMed: 21653522]

Denver DR, Dolan PC, Wilhelm LJ, et al. A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:16310–4. [PubMed: 19805298]

Dia, I., Guelbeogo, M., Ayala, D. Advances and Perspectives in the Study of the Malaria Mosquito Anopheles funestus. In: Manguin, S., editor. Anopheles mosquitoes - New insights into malaria vectors. 2013.

Dia I, Lochouarn L, Boccolini D, Costantini C, Fontenille D. Spatial and temporal variations of the chromosomal inversion polymorphism of Anopheles funestus in Senegal. Parasite. 2000; 7:179–184. [PubMed: 11031753]

Dobzhansky T. Genetics of natural populations. IX. Temporal changes in the composition of populations of Drosophila pseudoobscura. Genetics. 1943; 28:162–186. [PubMed: 17247077]

Dobzhansky T. Genetics of natural populations. XVIII. Experiments on chromosomes of Drosophila pseudoobscura from different geographic regions. Genetics. 1948; 33:588–602. [PubMed: 18100290]

Dobzhansky T. Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of Drosophila pseudoobscura. Genetics. 1950; 35:288–302. [PubMed: 15414931]

Dobzhansky, T. Genetics of the evolutionary process. New York: 1970.

Donnelly MJ, Licht MC, Lehmann T. Evidence for recent population expansion in the evolutionary history of the malaria vectors Anopheles arabiensis and Anopheles gambiae. Molecular biology and evolution. 2001; 18:1353–1364. [PubMed: 11420373]

Earl DA, VonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 2012:4.

Edgley, ML., Baillie, DL., Riddle, DL., Rose, AM. Genetic balancers. Wormbook, editor. 2006.

Evanno G, Goudet J, Regnaut S. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology. 2005; 14:2611–2620. [PubMed: 15969739]

Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics. 1992; 131:479–491. [PubMed: 1644282]

Feder JL, Egan SP, Nosil P. The genomics of speciation-with-gene-flow. Trends in Genetics. 2012; 28:342–350. [PubMed: 22520730]

Felsenstein J. The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics. 1974; 78:737–756. [PubMed: 4448362]

Fouet C, Gray E, Besansky NJ, Costantini C. Adaptation to aridity in the malaria mosquito Anopheles gambiae: chromosomal inversion polymorphism and body size influence resistance to desiccation. PloS one. 2012; 7:e34841. [PubMed: 22514674]

Fouet C, Kamdem C, Gamez S, White BJ. Extensive genetic diversity among populations of the malaria mosquito Anopheles moucheti revealed by population genomics. Infection, Genetics and Evolution. 2017; 48:27–33.

Fumagalli M, Vieira FG, Linderoth T. ngsTools : methods for population genetics analyses from Next-Generation Sequencing data. Bioinformatics. 2014

Gattepaille LM, Jakobsson M, Blum MGB. Inferring population size changes with sequence and SNP data: lessons from human bottlenecks. Heredity. 2013; 110:409–19. [PubMed: 23423148]

Gillespie JH. Junk ain't what junk does: Neutral alleles in a selected context. Gene. 1997; 205:291–299. [PubMed: 9461403]

Gillies MT, Coetzee M. A supplement to the Anophelinae of Africa south of the Sahara. The South African Institute for Medical Research, Johannesburg. 1987

Gillies MT, De Meillon B. The Anophelinae of Africa South of the Sahara. Publications of the South African Institute for Medical Research, Johannesburg. 1968

Giraldo-Calderon GI, Emrich SJ, MacCallum RM, et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Research. 2015; 43:D707–D713. [PubMed: 25510499]

Gosset CC, Bierne N. Differential introgression from a sister species explains high FST outlier loci within a mussel species. Journal of Evolutionary Biology. 2013; 26:14–26. [PubMed: 23199184]

Gray EM, Rocca KaC, Costantini C, Besansky NJ. Inversion 2La is associated with enhanced desiccation resistance in Anopheles gambiae. Malaria journal. 2009; 8:215. [PubMed: 19772577]

Green CA, Hunt RH. Interpretation of variation in ovarian polytene chromosomes of Anopheles funestus Giles, A. parensis Gillies, and A. aruni? Genetica. 1980; 51:187–195.

Haddrill PR, Halligan DL, Tomaras D, Charlesworth B. Reduced efficacy of selection in regions of the Drosophila genome that lack crossing over. Genome biology. 2007; 8:R18. [PubMed: 17284312]

Hellmann I, Ebersberger I, Ptak SE, Pääbo S, Przeworski M. A Neutral Explanation for the Correlation of Diversity with Recombination Rates in Humans. The American Journal of Human Genetics. 2003; 72:1527–1535. [PubMed: 12740762]

Hellmann I, Prüfer K, Ji H, et al. Why do human diversity levels vary at a megabase scale? Genome Research. 2005; 15:1222–1231. [PubMed: 16140990]

Hentges KE, Justice MJ. Checks and balancers : balancer chromosomes to facilitate genome annotation. Trends in Genetics. 2004:20.

Hill WG, Robertson A. The effect of linkage on the limits of artificial selection. Genetical Research. 1966; 8:269–294. [PubMed: 5980116]

Hoffmann, Aa, Rieseberg, LH. Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? Annual review of ecology, evolution, and systematics. 2008; 39:21–42.

Hoffmann AA, Sgro CM, Weeks AR. Chromosomal inversion polymorphisms and adaptation. Trends in Ecology & Evolution. 2004; 19:482–488. [PubMed: 16701311]

Hoffmann AA, Weeks AR. Climatic selection on genes and traits after a 100 year-old invasion : a critical look at the temperate-tropical clines in Drosophila melanogaster from eastern Australia. Genetica. 2007; 129:133–147. [PubMed: 16955331]

Huber CD, Nordborg M, Hermisson J, Hellmann I. Keeping It Local: Evidence for Positive Selection in Swedish Arabidopsis thaliana. Molecular Biology and Evolution. 2014; 31:3026–3039. [PubMed: 25158800]

Hudson RR, Kaplan NL. The coalescent process in models with selection and recombination. Genetics. 1988; 120:819–829. [PubMed: 3066685]

Jakobsson M, Rosenberg N. CLUMPP: a cluster matching and permutation program for dealing with multimodality in analysis of population structure. Bioinformatics. 2007; 23:1801–1806. [PubMed: 17485429]

Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008; 24:1403–1405. [PubMed: 18397895]

De Jong G, Bochdanovits Z. Latitudinal clines in Drosophila melanogaster : body size, allozyme frequencies, inversion frequencies, and the insulin-signalling pathway. Journal of Genetics. 2003; 82:207–223. [PubMed: 15133196]

Joron M, Frezal L, Jones RT, et al. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature. 2011; 477:203–206. [PubMed: 21841803]

Kamau L, Hunt RH, Coetzee M. Analysis of the population structure of Anopheles funestus (Diptera: Culicidae) from western and coastal Kenya using paracentric chromosomal inversion frequencies. Journal of Medical Entomology. 2002; 39:78–83. [PubMed: 11931275]

Kamdem C, Fouet C, Gamez S, White BJ. Pollutants and insecticides drive local adaptation in African malaria mosquitoes. Molecular Biology and Evolution. 2017; 34:1261–1275. [PubMed: 28204524]

Kaplan NL, Hudson RR, Langley CH. The “hitchhiking effect” revisited. Genetics. 1989; 123:887–899. [PubMed: 2612899]

Kapun M, Van Schalkwyk H, McAllister B, Flatt T, Schlötterer C. Inference of chromosomal inversion dynamics from Pool-Seq data in natural and laboratory populations of Drosophila melanogaster. Molecular Ecology. 2014; 23:1813–1827. [PubMed: 24372777]

Kapun M, Schmidt C, Durmaz E, Schmidt PS, Flatt T. Parallel effects of the inversion In (3R) Payne on body size across the North American and Australian clines in Drosophila melanogaster. Journal of Evolutionary Biology. 2016; 29:1059–1072. [PubMed: 26881839]

Kawecki TJ, Dieter E. Conceptual issues in local adaptation. Ecology Letters. 2004; 7:1225–1241.

Keightley PD, Pinharanda A, Ness RW, et al. Estimation of the Spontaneous Mutation Rate in Heliconius melpomene. Molecular Biology and Evolution. 2015; 32:239–243. [PubMed: 25371432]

Keinan A, Reich D. Human population differentiation is strongly correlated with local recombination rate. PLoS Genetics. 2010:6.

Kennington WJ, Hoffmann AA, Partridge L. Mapping Regions Within Cosmopolitan Inversion In(3R)Payne Associated With Natural Variation in Body Size in Drosophila melanogaster. Genetics. 2007; 177:549–556. [PubMed: 17603103]

Kim Y, Maruki T. Hitchhiking effect of a beneficial mutation spreading in a subdivided population. Genetics. 2011; 189:213–226. [PubMed: 21705748]

Kimura, MK. The Neutral Theory of Molecular Evolution. Cambridge, UK: 1983.

Kirkpatrick M. How and why chromosome inversions evolve. PLoS Biology. 2010:8.

Kirkpatrick M, Barton N. Chromosome inversions, local adaptation and speciation. Genetics. 2006; 173:419–434. [PubMed: 16204214]

Kliman R, Hey J. Reduced natural selection associated with low recombination in Drosophila melanogaster. Molecular biology and evolution. 1993; 10:1239–1258. [PubMed: 8277853]

Knibb WR. Chromosome inversion polymorphisms in Drosophila melanogaster II. Geographic clines and climatic associations in Australasia, North America and Asia. Genetica. 1982; 58:213–221.

Korneliussen T, Albrechtsen A, Nielsen R. ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics. 2014; 15:356. [PubMed: 25420514]

Korneliussen TS, Moltke I, Albrechtsen A, Nielsen R. Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC bioinformatics. 2013; 14:289. [PubMed: 24088262]

Krimbas, C., Powell, JR. Introduction. In: Krimbas, C., Powell, JR., editors. Drosophila inversion polymorphism. London: 1992. p..

Kulathinal RJ, Bennett SM, Fitzpatrick CL, Noor MAF. Fine-scale mapping of recombination rate in Drosophila refines its correlation to diversity and divergence. Proceedings of the National Academy of Sciences. 2008; 105:10051–10056.

Kulathinal RJ, Stevison LS, Noor MAF. The genomics of speciation in Drosophila: Diversity, divergence, and introgression estimated using low-coverage genome sequencing. PLoS Genetics. 2009:5.

Lahn BT, Page DC. Chromosome Four Evolutionary Strata on the Human X Chromosome. Science. 1999; 286:964–967. [PubMed: 10542153]

Lanzaro, GC., Lee, Y. Speciation in Anopheles gambiae — The Distribution of Genetic Polymorphism and Patterns of Reproductive Isolation Among Natural Populations. In: Manguin, S., editor. Anopheles mosquitoes - New insights into malaria vectors. 2013.

Lee SF, Chen Y, Varan AK, et al. Molecular Basis of Adaptive Shift in Body Size in Drosophila melanogaster : Functional and Sequence Analyses of the Dca Gene Research article. Molecular Biology and Evolution. 2011; 28:2393–2402. [PubMed: 21393605]

Lee Y, Meneses CR, Fofana A, Lanzaro GC. Desiccation Resistance Among Subpopulations of Anopheles gambiae s.s. From Selinkenyi, Mali. Journal of medical entomology. 2009; 46:316–320. [PubMed: 19351082]

Lewontin RC, Krakauer J. DISTRIBUTION OF GENE FREQUENCY AS A TEST OF THE THEORY OF THE SELECTIVE NEUTRALITY OF and LEWONTIN. Genetics. 1973; 74:175–195. [PubMed: 4711903]

Lischer HEL, Excoffier L. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 2012; 28:298–299. [PubMed: 22110245]

Lowry DB, Willis JH. A Widespread Chromosomal Inversion Polymorphism Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. PLoS Biology. 2010:8.

Magni GE, Von Borstel RC. Different Rates of Spontaneous Mutation during Mitosis and Meiosis in Yeast. Genetics. 1962; 47:1097–1108. [PubMed: 17248123]

Majewski J, Cohan FM. Adapt globally, act locally: The effect of selective sweeps on bacterial sequence diversity. Genetics. 1999; 152:1459–1474. [PubMed: 10430576]

Maynard Smith J, Haigh J. The hitch-hiking effect of a favourable gene. Genetical research. 1974; 23:23–35. [PubMed: 4407212]

Mcgaugh SE, Heil CSS, Manzano-winkler B, et al. Recombination Modulates How Selection Affects Linked Sites in Drosophila. PLoS Biology. 2012:10.

McGaugh SE, Noor MaF. Genomic impacts of chromosomal inversions in parapatric Drosophila species. Philosophical transactions of the Royal Society of London. Series B, Biological sciences. 2012; 367:422–9.

Meirmans P, Van Tienderen P. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes. 2004; 4:792–794.

Mettler LE, Voelker RA, Mukai T. Inversion clines in populations of Drosophila melanogaster. Genetics. 1977; 87:169–176. [PubMed: 17248756]

Michel, aP, Ingrasci, MJ., Schemerhorn, BJ., et al. Rangewide population genetic structure of the African malaria vector Anopheles funestus. Molecular Ecology. 2005; 14:4235–4248. [PubMed: 16313589]

Muller HJ. Genetic variability, twin hybrids, and constant hybrids, in a case of balanced lethal factors. Genetics. 1918; 3:422–499. [PubMed: 17245914]

Nachman MW. Variation in recombination rate across the genome: Evidence and implications. Current Opinion in Genetics and Development. 2002; 12:657–663. [PubMed: 12433578]

Nachman MW, Payseur BA. Recombination rate variation and speciation : theoretical predictions and empirical results from rabbits and mice. Philosophical Transactions of the Royal Society B: Biological Sciences. 2012; 367:409–421.

Neafsey DE, Waterhouse RM, Abai MR, et al. Highly evolvable malaria vectors: The genomes of 16 Anopheles mosquitoes. Science. 2015; 347:1258522–1258522. [PubMed: 25554792]

Nielsen R. Molecular Signatures of Natural Selection. Annual review of genetics. 2005; 39:197–218.

Nishikawa H, Iijima T, Kajitani R, et al. A genetic mechanism for female-limited Batesian mimicry in Papilio butterfly. Nature genetics. 2015; 47:405–9. [PubMed: 25751626]

Noor MAF. Mutagenesis from meiotic recombination is not a primary driver of sequence divergence between Saccharomyces species. Molecular Biology and Evolution. 2008; 25:2439–2444. [PubMed: 18723832]

Noor, MaF, Bennett, SM. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. Heredity. 2009; 103:439–444. [PubMed: 19920849]

Noor MAF, Grams KL, Bertucci LA, Reiland J. Chromosomal inversions and the reproductive isolation of species. Proceedings of the National Academy of Sciences. 2001; 98:12084–12088.

Nordborg M, Charlesworth B, Charlesworth B. Increased levels of polymorphism surrounding selectively maintained sites in highly selfing species. Proceedings of the Royal Society B: Biological Sciences. 1996; 263:1033–1039.

Nordborg M, Hu TT, Ishino Y, et al. The pattern of polymorphism in Arabidopsis thaliana. PLoS Biology. 2005; 3:1289–1299.

Nordborg M, Tavar S. Linkage disequilibrium : what history has to tell us. Trends in Genetics. 2002; 18:83–90. [PubMed: 11818140]

Nosil P, Feder JL. Widespread yet heterogeneous genomic divergence. Molecular Ecology. 2012; 21:2829–2832. [PubMed: 22676072]

Nosil P, Funk DJ, Ortiz-Barrientos D. Divergent selection and heterogeneous genomic divergence. Molecular Ecology. 2009; 18:375–402. [PubMed: 19143936]

O'Loughlin SM, Magesa S, Mbogo C, et al. Genomic Analyses of Three Malaria Vectors Reveals Extensive Shared Polymorphism but Contrasting Population Histories. Molecular biology and evolution. 2014:1–14.

Ortiz-barrientos D, Engelstädter J, Rieseberg LH. Recombination Rate Evolution and the Origin of Species. Trends in Ecology & Evolution. 2016; 31:226–236. [PubMed: 26831635]

Ortiz-Barrientos D, Reiland J, Hey J, Noor MAF. Recombination and the divergence of hybridizing species. Genetica. 2002; 116:167–178. [PubMed: 12555775]

Paradis E, Claude J, Strimmer K. Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 2004; 20:289–290. [PubMed: 14734327]

Pegueroles C, Ordóñez V, Mestres F, Pascual M. Recombination and selection in the maintenance of the adaptive value of inversions. Journal of Evolutionary Biology. 2010; 23:2709–2717. [PubMed: 20964762]

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLoS ONE. 2012; 7:e37135. [PubMed: 22675423]

Pool JE, Corbett-Detig RB, Sugino RP, et al. Population Genomics of Sub-Saharan Drosophila melanogaster: African Diversity and Non-African Admixture. PLoS Genetics. 2012:8.

Powell, JR. Progress and prospects in evolutionary biology: the Drosophila model. New York: 1997.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. [PubMed: 10835412]

Purcell S, Neale B, Todd-Brown K, et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics. 2007:81.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2014.

Rako L, Anderson AR, Sgro CM, Stocker AJ, Hoffmann AA. The association between inversion In (3R) Payne and clinally varying traits in Drosophila melanogaster. Genetica. 2006; 128:373–384. [PubMed: 17028965]

Riehle MM, Guelbeogo WM, Gneme A, et al. A cryptic subgroup of Anopheles gambiae is highly susceptible to human malaria parasites. Science (New York, N.Y.). 2011; 331:596–8.

Rieseberg LH. Chromosomal rearrangements and speciation. Trends in Ecology and Evolution. 2001; 16:351–358. [PubMed: 11403867]

Roberts, P. The genetics of chromosomal aberration. In: Ashburner, M., Novitski, E., editors. The genetics and biology of Drosophila. London: 1976. p. 67-184.

Roesti M, Hendry AP, Salzburger W, Berner D. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. Molecular ecology. 2012; 21:2852–62. [PubMed: 22384978]

Roesti M, Moser D, Berner D. Recombination in the threespine stickleback genome - Patterns and consequences. Molecular Ecology. 2013; 22:3014–3027. [PubMed: 23601112]

Rosenberg N. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Resources. 2004; 4:137–138.

Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation. 2013:14.

Schaeffer SW. Selection in heterogeneous environments maintains the gene arrangement polymorphism of drosophila pseudoobscura. Evolution. 2008; 62:3082–3099. [PubMed: 18764919]

Schaeffer SW, Goetting-minesky MP, Kovacevic M, et al. Evolutionary genomics of inversions in Drosophila pseudoobscura : Evidence for epistasis. Proceedings of the National Academy of Sciences. 2003; 100:8319–8324.

Schlötterer C. Towards a molecular characterization of adaptation in local populations. Current Opinion in Genetics and Development. 2002; 12:683–687. [PubMed: 12433582]

Schneider KA, Kim Y. Genetic Hitchhiking under Heterogeneous Spatial Selection Pressures. PLoS ONE. 2013:8.

Seehausen O, Butlin RK, Keller I, et al. Genomics and the origin of species. Nature reviews. Genetics. 2014; 15:176–92. [PubMed: 24535286]

Service, MW. Mosquito ecology: field sampling methods. UK Elsevier Applied Science; London: 1993.

Sharakhova MV, Xia A, Leman SC, Sharakhov IV. Arm-specific dynamics of chromosome evolution in malaria mosquitoes. BMC evolutionary biology. 2011; 11:91. [PubMed: 21473772]

Sharakhov I, Braginets O, Grushko O, et al. A Microsatellite Map of the African Human Malaria Vector Anopheles funestus. Journal of Heredity. 2004; 95:29–34. [PubMed: 14757727]

Sinka ME, Bangs MJ, Manguin S, et al. The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis. Parasites & vectors. 2010; 3:117. [PubMed: 21129198]

Slatkin M, Wiehe T. Genetic hitch-hiking in a subdivided population. Genetical Research. 1998; 71:155–160. [PubMed: 9717437]

Spencer CCA, Deloukas P, Hunt S, et al. The influence of recombination on human genetic diversity. PLoS Genetics. 2006; 2:1375–1385.

Stephan W, Langley CH. Molecular genetic variation in the centromeric region of the X chromosome in three Drosophila ananassae populations. I. Contrasts between the vermilion and forked loci. Genetics. 1989; 121:89–99. [PubMed: 2563714]

Storz JF. Using genome scans of DNA polymorphism to infer adaptive population divergence. Molecular Ecology. 2005; 14:671–688. [PubMed: 15723660]

Stump AD, Fitzpatrick MC, Lobo NF, et al. Centromere-proximal differentiation and speciation in Anopheles gambiae. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:15930–5. [PubMed: 16247019]

Sturtevant A. A case of rearrangement of genes in Drosophila. Proceedings of the National Academy of Sciences. 1921; 7:235–237.

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989; 123:585–595. [PubMed: 2513255]

Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS. Selection versus demography: A multilocus investigation of the domestication process in maize. Molecular Biology and Evolution. 2004; 21:1214–1225. [PubMed: 15014173]

Thomas CG, Wang W, Jovelin R, et al. Full-genome evolutionary histories of selfing, splitting, and selection in Caenorhabditis. 2015:667–678.

Thornton K. Recombination and the properties of Tajima's D in the context of approximate-likelihood calculation. Genetics. 2005; 171:2143–8. [PubMed: 15998723]

Tiffin P, Ross-Ibarra J. Advances and limits of using population genetics to understand local adaptation. Trends in Ecology & Evolution. 2014; 29:673–680. [PubMed: 25454508]

Watterson GA. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology. 1975; 7:256–276. [PubMed: 1145509]

Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984; 38:1358–1370. [PubMed: 28563791]

Williams, GC. Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought. Princeton, NJ: 1966.

Wright S, Dobzhansky T. Genetics of natural populations. XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of Drosophila pseudoobscura. Genetics. 1946; 31:125–156. [PubMed: 21021044]

Wright SI, Foxe JP, DeRose-Wilson L, et al. Testing for effects of recombination rate on nucleotide diversity in natural populations of Arabidopsis lyrata. Genetics. 2006; 174:1421–1430. [PubMed: 16951057]

Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics. 2010; 26:873–881. [PubMed: 20147302]

Yeaman S. Genomic rearrangements and the evolution of clusters of locally adaptive loci. Proceedings of the National Academy of Sciences. 2013; 110:E1743–51.

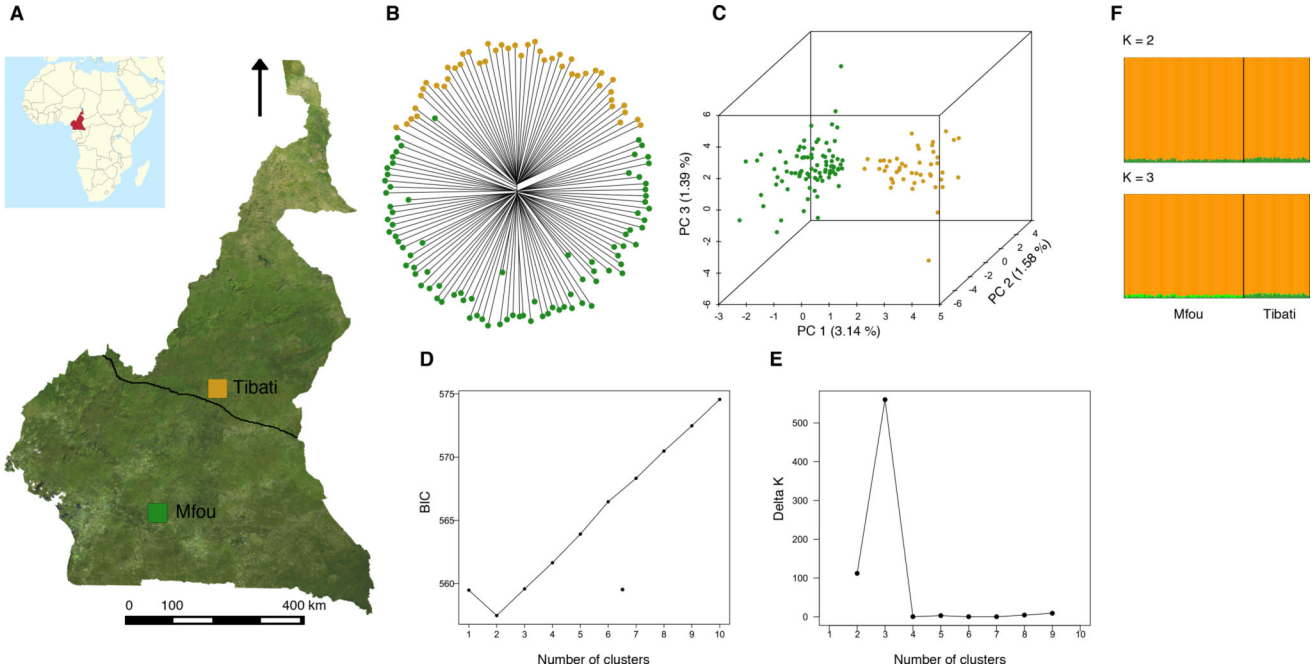

Figure 1.

Geographic origin and genetic relatedness of *An. funestus* populations in Cameroon. (A) Map showing Mfou and Tibati, the two locations where mosquitoes were collected. A delimitation of the approximate distribution ranges of the two ecotypes that occur in Cameroon is shown (continuous line in the middle of the map). (B) and (C) Population genetic structure as revealed respectively by a neighbor-joining tree and a PCA. The percentage of variance explained is indicated on each PCA axis. (D) and (E) Confirmation of the presence of two *An. funestus* clusters with DAPC and the delta k method of Evanno *et al.* (2005). The lowest Bayesian Information Criterion (BIC) and cross-validation error and the highest delta k indicate the most probable number of clusters (2–3). (F) Bayesian clustering in STRUCTURE.

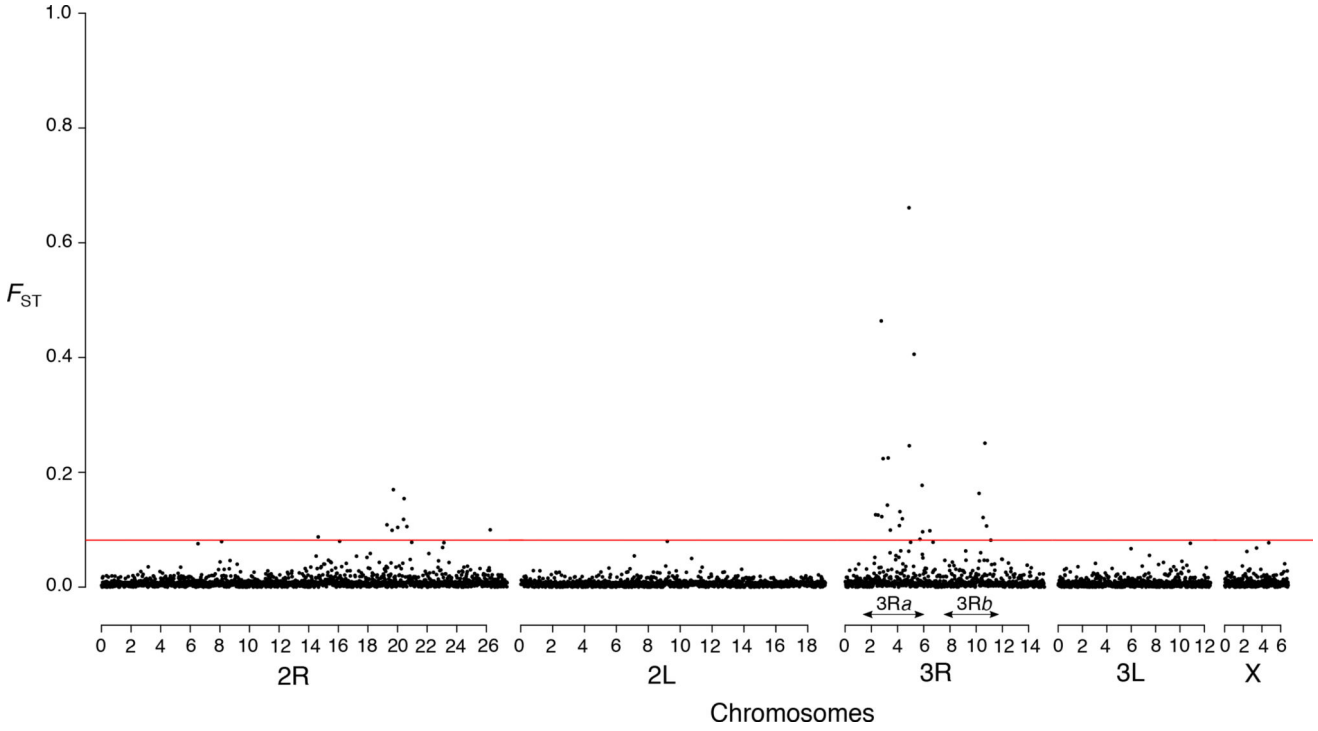

Figure 2. aaaa

Estimates of pairwise population differentiation (F_{ST}) based on SNPs ordered by position along pseudo-chromosomes representing the five chromosome arms of *An. funestus* F_{ST} values on top of the red line are above the 99th percent of the empirical distribution. Arrows indicate the genomic coordinates of the 3R*a* and 3R*b* inversions.

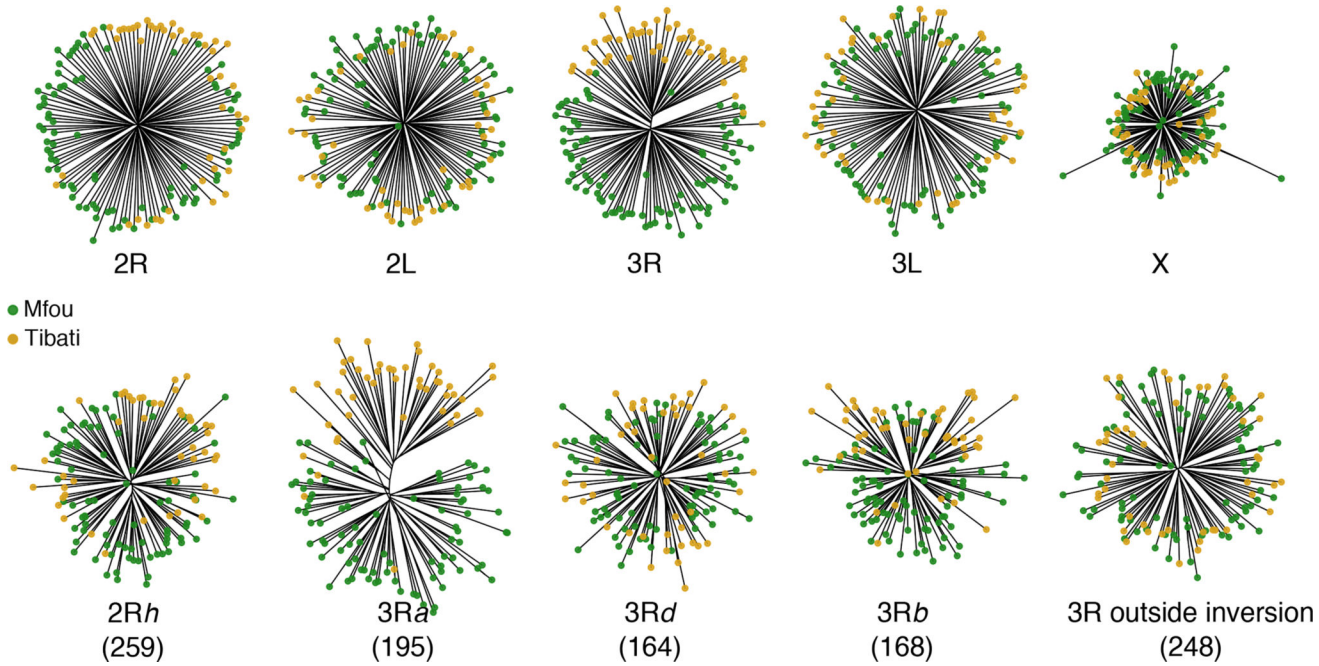

Figure 3.

Population genetic structure revealed by each chromosome arm and by four polymorphic chromosomal inversions (2R*h*, 3R*a*, 3R*b* and 3R*d*). The number of SNPs is indicated in parenthesis.

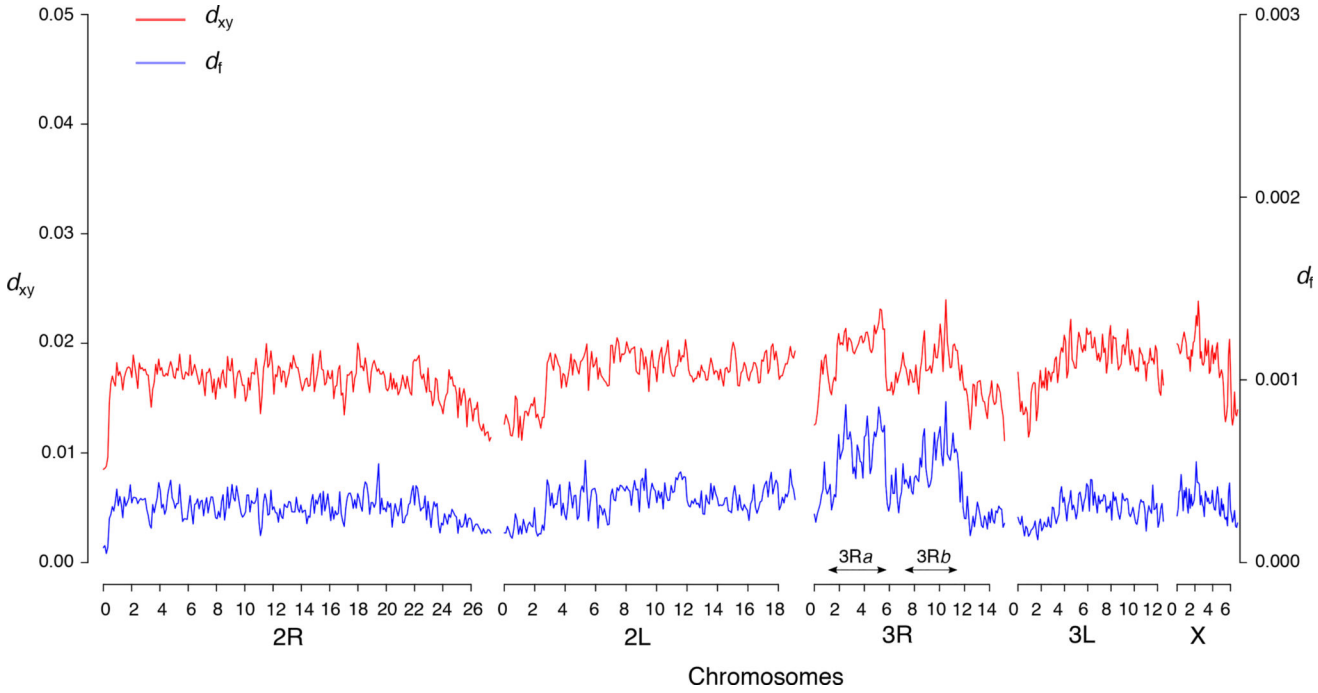

Figure 4.

Genome-wide distribution of d_{xy} and fixed differences (d_f) across 90-kb non-overlapping windows along the five pseudo-chromosomes in Mfou and Tibati populations. Strong sequence divergence within the 3R*a* and 3R*b* inversions is characterized by the presence of peaks of d_{xy} and d_f at these loci.

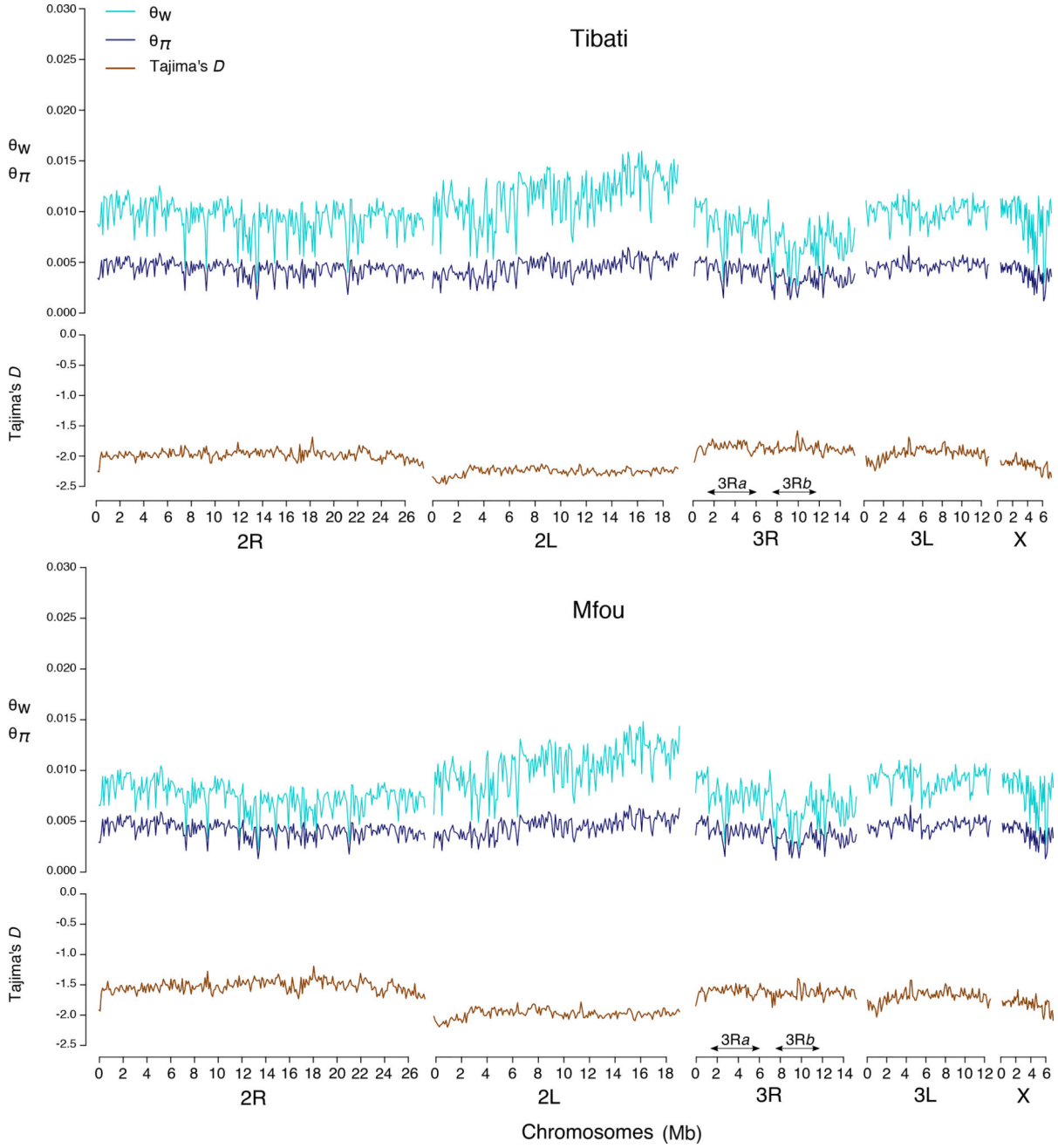

Figure 5.

Estimates of nucleotide diversity (θ_π and θ_w) and allele frequency spectrum (Tajima's *D*) across 90-kb non-overlapping windows illustrating the uneven distribution of genetic diversity along the five chromosome arms. The collinear autosome (2L) is the most polymorphic in both Mfou and Tibati populations. In contrast, autosomes bearing multiple inversions exhibit a significant reduction in the amount of polymorphic sites.

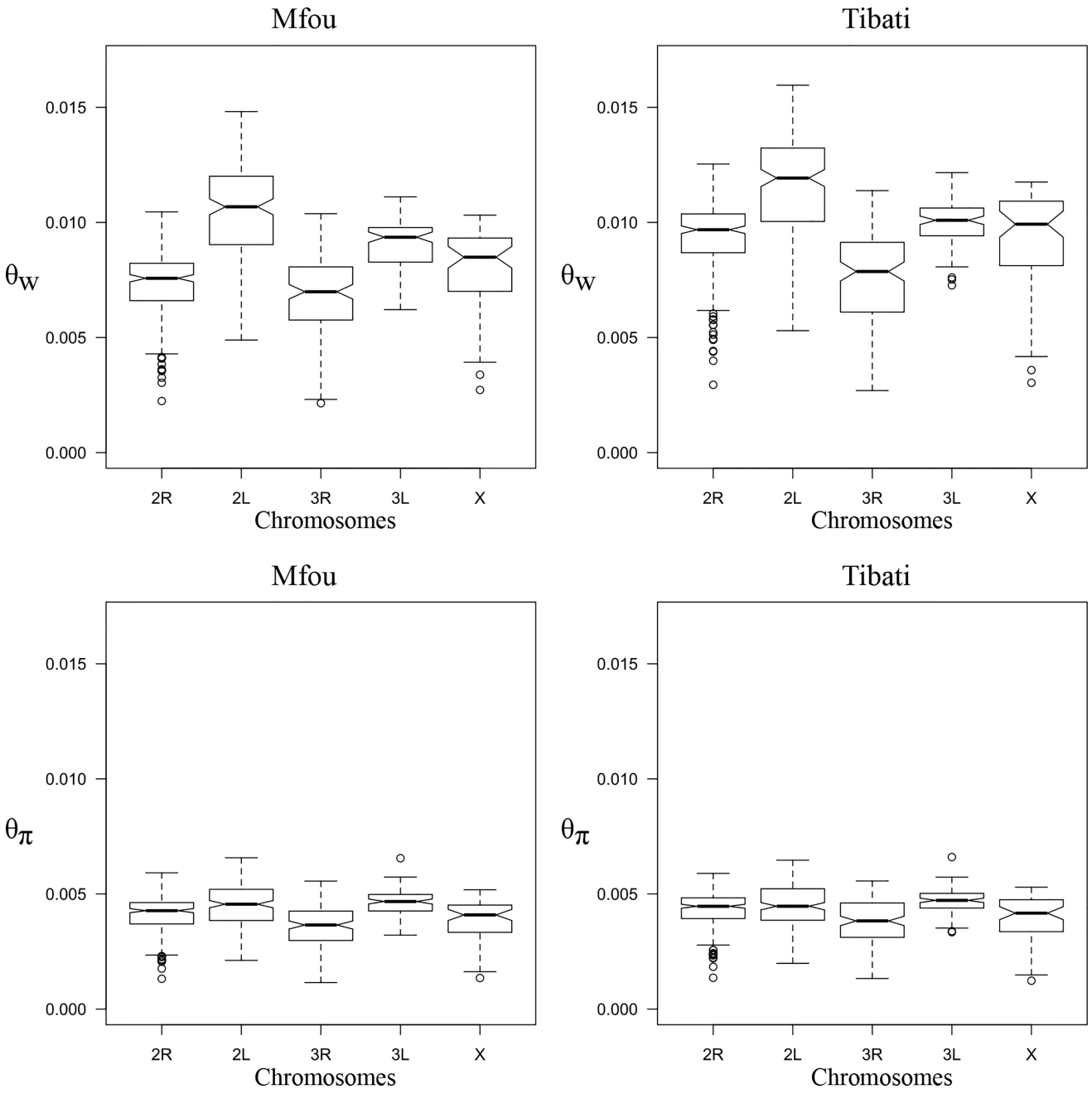

Figure 6.
Box plot depicting the distribution of θ_{π} and θ_{w} among chromosomes in Mfou and Tibati.

Table 1

Reduction in nucleotide diversity relative to the 2L chromosome arm (%).

		2R	3L	3R	X
Mfou	θ_w	29.22 *	13.24 *	34.69 *	23.35 *
	θ_π	7.92 *	−3.23 *	19.23 *	13.92 *
Tibati	θ_w	18.79 *	13.24 *	33.54 *	18.94 *
	θ_π	3.23 *	−4.84 *	15.61 *	11.86 *

* Statistically significant (Wilcoxon rank sum test, $p < 0.001$)