Genetic Differentiation of *Glossina pallidipes* Tsetse Flies in Southern Kenya

Winnie A. Okeyo,^{1,2,3,*†} Norah P. Saarmaan,^{4†} Rosemary Bateta,² Kirstin Dion,⁴ Michael Mengual,⁴ Paul O. Mireji,^{2,3,5} Collins Ouma,¹ Sylvance Okoth,² Grace Murilla,^{2,3} Serap Aksoy,³ and Adalgisa Caccone^{3,4}

¹Department of Biomedical Sciences and Technology, School of Public Health and Community Development, Maseno University, Kisumu, Kenya;

²Biotechnology Research Institute, Kenya Agricultural and Livestock Research Organization, Nairobi, Kenya; ³Yale School of Public Health, Yale University, New Haven, Connecticut; ⁴Department of Ecology & Evolutionary Biology, Yale University, New Haven, Connecticut; ⁵Center for Geographic Medicine Research Coast, Kenya Medical Research Institute, Kilifi, Kenya

Abstract. The tsetse fly *Glossina pallidipes*, the major vector of the parasite that causes animal African trypanosomiasis in Kenya, has been subject to intense control measures with only limited success. The *G. pallidipes* population dynamics and dispersal patterns that underlie limited success in vector control campaigns remain unresolved, and knowledge on genetic connectivity can provide insights, and thereby improve control and monitoring efforts. We therefore investigated the population structure and estimated migration and demographic parameters in *G. pallidipes* using genotypic data from 11 microsatellite loci scored in 250 tsetse flies collected from eight localities in Kenya. Clustering analysis identified two genetically distinct eastern and western clusters (mean between-cluster $F_{ST} = 0.202$) separated by the Great Rift Valley. We also found evidence of admixture and migration between the eastern and western clusters, isolation by distance, and a widespread signal of inbreeding. We detected differences in population dynamics and dispersal patterns between the western and eastern clusters. These included lower genetic diversity (allelic richness; 7.48 versus 10.99), higher relatedness (percent related individuals; 21.4% versus 9.1%), and greater genetic differentiation (mean within-cluster F_{ST} ; 0.183 versus 0.018) in the western than the eastern cluster. Findings are consistent with the presence of smaller, less well-connected populations in Western relative to eastern Kenya. These data suggest that recent anthropogenic influences such as land use changes and vector control programs have influenced population dynamics in *G. pallidipes* in Kenya, and that vector control efforts should include some region-specific strategies to effectively control this disease vector.

INTRODUCTION

Animal African trypanosomiasis (AAT) and human African trypanosomiasis (HAT) are fatal diseases in sub-Saharan Africa that are transmitted primarily by tsetse flies.^{1–3} Human African trypanosomiasis has killed hundreds and thousands of people in the 20th century with the latest epidemic, which started in the 1980s and came to an end in the early 2000s,⁴ whereas AAT, also known as Nagana, causes high morbidity and death in livestock throughout the continent. Diseases caused by AAT inflict daily economic losses of an estimated \$1.24,⁵ affecting already disadvantaged populations living in rural communities. Unfortunately, disease prevention in the mammal by using a vaccine has not been possible because the parasite is covered with a thick surface coat glycoprotein, which undergoes antigenic variation that makes the design of an effective vaccine difficult.⁶ Drugs available for control of these diseases are also difficult to administer,^{7,8} particularly for late-stage disease, and an estimated 10,000 cases of drug resistance has been reported annually.^{9–11} Vector control remains an important component of HAT and AAT control in sub-Saharan Africa.^{12,13} Among the methods used for tsetse control are insecticide-treated traps and targets, aerial or ground spraying, pour-ons applied on cattle, sterile insect technique (SIT), and habitat interference,^{7,14–22} as well as newer ones, including the use of sticky small targets and possibilities of using microbial symbionts.^{23–26} However, their success has been limited by the large geographical scale of the problem and by the need to sustain long-term

monitoring efforts, which requires continued deployment of resources.^{20,21,27}

Kenya is plagued by tsetse flies and trypanosomes.²⁸ At present, the country is mainly affected by the animal disease (AAT), although the presence of human infectious parasites in animal reservoirs in the game parks remains a threat for HAT infections. Indeed, there was a case of the human disease reported in the Maasai Mara National Park in 2012 long after the eradication of the main HAT epidemic, indicating that human infection remains a real risk in Kenya.^{4,29} Eight species of the genus *Glossina* are found in Kenya, including *Glossina pallidipes*, *Glossina austeni*, *Glossina brevipalpis*, and *Glossina fuscipes*.²⁸ Among these, *G. pallidipes* is by far the most epidemiologically important species because of its extensive range.³⁰ *Glossina pallidipes* and other members of the *Morsitans* group of tsetse (including *Glossina morsitans*, *Glossina swynertoni*, and *G. austeni*) are patchily distributed across the landscape and are found locally in high density in low-temperature and high-humidity habitats, such as forest thickets in shrublands, savannah, and grassy woodlands.^{8,31–33} The highest densities of *G. pallidipes* occur along a belt of about 800 km bordering the Indian Ocean to the east, and along Lake Victoria and western Uganda to the west, and Tanzania to the south (Figure 1). Here, the presence of grazing fields and game reserves with large numbers of domesticated and wildlife herds provide abundant sources of blood meals and facilitate the maintenance of high-density tsetse fly populations.^{32–37} Given the high densities of flies³⁰ and the fact that Kenya shares game parks with Tanzania along its southern border where reservoir animals migrate between the two countries, there is a risk of disease emergence in Kenya, especially along the border.

Tsetse control efforts have been ongoing in many parts of Kenya since the 1980s.^{14–16,18,22} These control efforts have included collaboration with local farmers, and have been led by the State Department of Livestock's Tsetse Control

* Address correspondence to Winnie A. Okeyo, Department of Biomedical Sciences and Technology, School of Public Health and Community Development, Maseno University, Kisumu, Kenya. E-mail: okeyo.winnie@gmail.com

† These authors contributed equally to this work.

FIGURE 1. Map showing the location of the eight sampling sites in Kenya analyzed in this study. Black dots identify sampling sites together with their symbols (see Table 1). Thick black borders show the approximate extension of five registered protected areas (National Parks, Game Reserves, and forests). Their full names are included next to the sampling site symbols. Black broken line represents the Great Rift Valley. KAP = Kapesur; KIB = Kibwezi; KIN = Kinango; MRT = Mara Talek; NGU = Nguruman; RUM = Ruma; SHI = Shimba; TSW = Tsavo west.

Program, the Kenya Agricultural and Livestock Research Organization's Kenya Trypanosomiasis Research Institute, the Kenya Wildlife Service, and the Pan African Tsetse and Trypanosomiasis Eradication Campaign.^{14–16,18,22} The control methods have included trapping, insecticide-treated targets, aerial or ground spraying, insecticide-treated cattle, SIT, and habitat interference.^{7,14–22} In southwestern Kenya, tsetse densities declined by 99% after application of intense control efforts in the late 1980s.^{15,18,22} However, populations bounced back to nearly to pre-campaign levels in the past decades, after control measures were relaxed.^{30,34,38,39} Use of population genetics data can help identify the causes for these failures, providing spatially specific data that can help maximize the efficiency of future control and monitoring efforts, and evaluate the possibility of eradication.^{40,41}

Population genetics tools have been used to show macro- and micro-geographic differentiation of *G. pallidipes* populations across eastern Africa, including Kenya.^{30,34,38,39} One such study sampled tsetse flies in eastern Africa from Ethiopia to Zimbabwe (3,000 km).³⁰ Using microsatellite and mitochondrial DNA (mtDNA) markers, five distinct *G. pallidipes* genetic clusters were identified. Genetic differentiation was also noted among the Kenyan samples, which were approximately 600 km apart.³⁰ Using the same type of markers, genetic differentiation was also detected at a smaller spatial scale among *G. pallidipes* populations that were 100 km apart.³⁴ A third study showed evidence of geographic differentiation and sporadic gene flow among *G. pallidipes*

populations analyzed from Kenya and Uganda.³⁸ Previous work has also shown that *G. pallidipes* populations undergo temporal changes in allele frequencies because of both natural (such as seasonal environmental fluctuations), and human-induced factors (such as control measures).⁴² Nonetheless, there remain gaps in our knowledge of the population structure of this dangerous vector in Kenya, especially at a regional scale.

Our study focuses at a regional scale and includes the geographic localities with the highest animal and human disease risk in Kenya, the area spanning from the Lake Victoria shores in southwestern Kenya to the coastal area along the Indian Ocean in southeastern Kenya. Along this 800 km transect, we screened for variation at 13 microsatellite loci in *G. pallidipes* flies sampled from eight sites, including from national parks, farmlands, and ranches (Figure 1). The goal of this work was to quantify patterns and levels of genetic connectivity of *G. pallidipes* populations in the region, and to investigate the processes that have contributed to their origin and maintenance. We also discuss the relevance of these findings to ongoing and future vector control efforts.

METHODS

Study area and sampling. Samples were collected from eight sites that spanned over 800 km (Figure 1): Kapesur (KAP), Ruma National Park (RUM), Mara Talek Junction in Maasai Mara National Park (MRT), Nguruman Escarpment (NGU), Kibwezi Forest (KIB), Tsavo West National Reserve (TSW), Kinango Ranches (KIN), and Shimba Hills National

Reserve (SHI). The geographic distances separating each from their closest sampled neighbor range from 15 km (KIN–SHI) to 300 km (NGU–KIN). These eight sites host a large number of both wild and domestic animals with five locations (RUM, MRT, KIB, TSW, and SHI) falling within protected areas. The area climate is hot and dry. The habitat is characterized by a semi-arid mix of savannah, shrublands, and grassy woodlands that get thicker along the riverbanks, a favorite place for the *morsitans* group of tsetse flies, such as *G. pallidipes*.^{31,43}

Table 1 lists the number of samples per site and provides geographic coordinates for each location. We used 10 biconical and *Ngu* traps, mounted about 1 km apart, to collect at least 30 flies per site. We collected similar numbers of male and female flies per sampling site to avoid any potential sexual bias during analysis and to facilitate continuity and connectivity with previous and future studies. All samples were collected between 2015 and 2016 in the wet seasons, when the flies are naturally abundant. The flies were individually stored in 80% ethanol and kept at 4°C until DNA extraction.

DNA extraction and microsatellite genotyping. We screened for genetic variation within and between 250 flies from eight sampling sites (approximately 30 flies/site). Genomic DNA was extracted from two legs per fly using the Qiagen DNAeasy blood and tissue extraction kits (Qiagen, Hilden, Germany), following the manufacturers' instructions. The genome was screened for 13 microsatellite loci previously used in *G. pallidipes*⁴² and developed for *G. pallidipes*, *G. fuscipes*, and *G. morsitans*. Table S1 of Supplemental Appendix 1 lists the markers and their publication of origin. Polymerase chain reactions (PCRs) for each locus consisted of 1 µL DNA template (1–10 ng), 6 µL of distilled H₂O, 2.6 µL of 5X Promega PCR Buffer, 0.1 µL of 100X bovine serum albumin (molecular grade), 0.5 µL each of 10 mM fluorescently-labeled forward and reverse primers (Table S1 of Supplemental Appendix 1), 1.1 µL of 25 mM MgCl₂, 1.1 µL of 10 mM dinucleoside triphosphate mix, and 0.1 µL of 5 U/µL Promega GoTaq DNA polymerase (Promega Corporation, Madison, WI). Polymerase chain reaction products were multiplexed into a combination of either two or three loci, and fragment analysis was carried out on an ABI 3730xL automated sequencer (Thermo Fisher Scientific, Waltham, MA) at the DNA Analysis Facility located on Science Hill at Yale University (<http://dna-analysis.yale.edu/>) using GelCo Rox size standard. The software GeneMarker v 2.4.0 (Soft Genetics, State College, PA) was used to score our data. All alleles were called automatically, reviewed, and manually edited, if necessary.

Statistical analyses. Microsatellite validation. Null alleles were identified with Micro-Checker v2.2 (University of Hull, Hull, United Kingdom).⁴⁴ The frequency of null alleles were estimated with the Brookfields (1996) method⁴⁵ for each locus and sampling locality, and significance was assessed with the exact binomial test for heterozygote deficits. Loci that showed that a consistent signal of null alleles across four or more sampling locations were excluded from subsequent analyses. Genepop Web v4.2⁴⁶ was used to test for deviations from Hardy–Weinberg equilibrium (HWE) and for linkage disequilibrium (LD). For both tests, 10,000 dememorizations, 1,000 batches, and 10,000 iterations were performed. Our results were subsequently corrected for the effect of multiple comparisons using the Benjamini–Hochberg method.⁴⁷

Genetic diversity and relatedness. To assess levels of genetic diversity within sampling sites, we calculated allelic richness (AR) and the coefficient of inbreeding (F_{IS}), using the program FSTAT v2.9.3.2.⁴⁸ Observed heterozygosity (H_o) and expected heterozygosity (H_e) were estimated using the program Arlequin v3.5.⁴⁹ To evaluate how levels of genetic diversity were partitioned among different hierarchical levels (identified using clustering analyses, see following text), an analysis of molecular variance (AMOVA) was implemented in Arlequin v3.5.⁴⁹

Maximum likelihood relatedness and relationship between samples within a sampling site was estimated using the software ML-Relate (Steven Kalinowski, Montana State University, Bozeman, MT),⁵⁰ as having related individuals can bias downstream analyses. Pairwise relationships within individual samples were divided into the following four categories: unrelated, half-siblings, full siblings, and parents/offspring.

The possible influence of high relatedness among individuals on the sample-wide estimates of diversity and heterozygosity was assessed by repeating estimates of AR, F_{IS} , H_o , and H_e with a subset of the data excluding highly related individuals (one of each parent-offspring or sibling pair). Linear regression was used to compare estimates between the full dataset and this reduced one using JMP® version 11.0.0 (SAS Institute Inc., Cary, NC).

Geographic genetic structure. To investigate population structure and levels of genetic admixture among tsetse from different sites, the Bayesian clustering method was implemented in STRUCTURE v2.3.4 (Pritchard Lab, Stanford University, Stanford, CA).⁵¹ This program identifies unique genetic units and estimates a probability of assignment (q -value) into each unit for every individual. A model that infers that alpha and lambda was used, running 10 independent runs

TABLE 1

Summary statistics based on 11 microsatellite loci for eight *Glossina pallidipes* sampling sites from an 800 km transect spanning the Lake Victoria region in western Kenya to the coastal region in southeastern Kenya (symbols as in Figure 1)

Site name	ID	Lat	Long	Year	N	AR	H_o	H_e	F_{IS}	F_{IS} P value	Relatedness			
											U	HS	FS	PO
Kapesur	KAP	0.733	34.316	2016	30	3.87	0.417	0.507	0.179	0.001	70.6	13.6	4.8	11.0
Ruma	RUM	0.656	34.276	2016	30	3.45	0.345	0.381	0.096	0.017	80.5	11.3	3.0	5.3
Maasai Mara	MRT	1.435	35.065	2015	30	5.10	0.511	0.564	0.094	0.007	84.6	12.0	0.9	2.5
Nguruman	NGU	−1.832	36.089	2015	48	3.80	0.507	0.537	0.056	0.042	79.0	14.0	2.7	4.3
Kibwezi	KIB	2.416	37.954	2015	30	6.65	0.582	0.659	0.119	0.001	89.9	9.0	0.7	0.5
Tsavo west	TSW	3.027	38.218	2015	30	7.02	0.573	0.654	0.125	0.001	90.1	9.2	0.7	0.0
Kinango	KIN	4.108	38.874	2015	30	6.75	0.554	0.654	0.155	0.001	91.3	7.6	1.1	0.0
Shimba Hills	SHI	4.343	39.521	2015	22	6.57	0.520	0.665	0.222	0.001	92.2	6.9	0.4	0.4

For each locality, we report the sampling site name and ID, latitude (Lat), longitude (Long), year of collection (Year), number of samples analyzed (N), mean allelic richness (AR) across loci, observed heterozygosity (H_o) and expected heterozygosity (H_e), inbreeding coefficient (F_{IS}), and F_{IS} P values. The last four columns report relatedness measures for unrelated individuals (U), half-siblings (HS), full siblings (FS), and parents/offspring (PO) calculated using the software ML-Relate.⁵⁰

for K from $K = 1$ through 10. The most likely K was inferred by calculating the ad hoc statistic " ΔK "⁵¹ with the online program STRUCTUREHARVESTER v0.6.⁵² CLUMPAK main pipeline server⁵³ was subsequently used to align the results for replicate runs of K . To complement this analysis, a discriminant analysis of principal components⁵⁴ was conducted using the "adagenet" package v1.4-2⁵⁵ in R v3.0.2 (R core team, Vienna, Austria).⁵⁶ To identify the optimal number of clusters, the *find.clusters* function in *R* was adopted, using the Bayesian information criterion (BIC), where the optimal cluster should correspond to the lowest BIC.⁵⁷

Genetic differentiation between sampling sites and between groupings of sampling sites was quantified with pairwise F_{ST} values, using Arlequin v3.5⁴⁹ with Wright's statistics,⁵⁸ following the variance method developed by Weir and Cockerham.⁵⁹ Significance was assessed with 10,000 permutations. Isolation by distance (IBD) was investigated using the online program IBD,⁶⁰ which implements the Mantel test⁶¹ to compare genetic (F_{ST}) and geographic distance. These geographic distances were obtained using the Geographic Distance Matrix Generator v1.2.3.⁶²

Admixture and migration. Levels of genetic admixture among tsetse from different sites were investigated using the program STRUCTURE v2.3.4,⁵¹ which provides individual probability of assignment to a given cluster (q -value, ranging from 0 to 1).

To evaluate the impact of migration and its geographic extent, we used the programs GENECLASS v2.0 (INRA, Montpellier, France)⁶³ and FLOCK v3.1⁶⁴ to identify recent migrants and conservatively identified an individual as a migrant only if it was detected by both types of analysis. In GENECLASS, the "detection of first-generation migrants" function was used to compute the likelihood of an individual's probability of belonging to a locality in which it was sampled (L_h), the highest likelihood among all population samples plus the population where the individual was sampled (L_{max}) and their ratio (L_h/L_{max}) to detect migrants.⁶⁵ The Bayesian method of Rannala and Mountain⁶⁶ was used to identify true migrants and the Monte Carlo resampling algorithm described by Paetkau et al.⁶⁵ ($N = 1,000$) to compute the test statistics, L_h/L_{max} . Individuals were identified as migrants, when the probability of assignment to the reference population where they were caught was $P < 0.05$. In FLOCK, we started by assigning each fly to one "home" region. We then compared the likelihood of assignment of each fly to its "home" against the likelihood of assignment to an "away" region with 50 iterations. A log-likelihood difference threshold value of 1 was used to identify an individual as a likely migrant.

RESULTS

Microsatellite validation. Of the 13 microsatellite loci selected, two loci (GpB6b and Gmm8) showed consistent signals of null alleles across four or more of the samples included in the Micro-Checker analysis (Table S2 of Supplemental Appendix 1), and were excluded from subsequent analyses. After adjustment for multiple comparisons, using the Benjamini-Hochberg false discovery rate method, our results indicated 17 instances of significant deviation from HWE (Table S2 of Supplemental Appendix 1). However, none of the 11 remaining loci or the eight sampling sites showed a consistent pattern of deviation. Similarly, we did not find evidence of LD among the loci (Table S4 of Supplemental Appendix 1).

Thus, the final dataset included 11 loci and eight sampling sites.

Genetic diversity and relatedness. Estimates of AR in the eight sampling sites ranged from 3.4 in RUM to 7.02 in TSW (Table 1). Observed heterozygosity ranged from 0.345 to 0.582, and was consistently lower than H_e , which ranged from 0.381 to 0.665 (Table 1). F_{IS} estimates ranged from 0.056 in RUM to 0.222 in KAP, with all estimates being significantly greater than zero. These estimates of H_o and F_{IS} indicate a slight heterozygote deficit compared with that expected under HWE with random mating (Table 1). Estimates of relatedness suggest that samples in each site had 10–30% related individuals.

Estimates of AR after excluding closely related individuals mirrored results from the full dataset, and ranged from 2.86 in RUM to 4.77 in TSW (Table S5 of Supplemental Appendix 1). Observed heterozygosity in this subset ranged from 0.418 to 0.575, and as with the full dataset, was consistently lower than H_e , which ranged from 0.446 to 0.666 (Table S5 of Supplemental Appendix 1). Even with this subset of samples, F_{IS} estimates were greater than zero, ranging from 0.063 in RUM to 0.278 in KAP, and all estimates except for in RUM remained significantly greater than zero (Table S5 of Supplemental Appendix 1). Comparisons of diversity statistics between the full dataset and the one excluding related individuals indicated nearly perfect correlation (Figure S1 of Supplemental Appendix 2), whereas correlation was not as high when comparing F_{IS} estimates (R^2 of 0.602; Figure S2 of Supplemental Appendix 2).

Geographic genetic structure. Figure 2 shows the results of the STRUCTURE analyses for $K = 2$, which is the most likely number of genetically distinct population groups based on ΔK ⁴² (Figure S3b of Supplemental Appendix 2). The four western-most samples (KAP, RUM, MRT, and NGU) cluster together apart from the four eastern-most samples (KIB, TSW, KIN, and SHI; Figure 2). These two clusters will be referred to from now on as "western" and "eastern" clusters. Figure S3 of Supplemental Appendix 2 shows the STRUCTURE plots from $K = 1$ –10 (Figure S3a of Supplemental Appendix 2), and the ad hoc statistic ΔK (Figure S3b of Supplemental Appendix 2). Clustering at higher K values suggests sub-clustering within the western cluster, with KAP and RUM grouping together separate from both NGU and MRT (apparent at $K = 4$; Figure S3a of Supplemental Appendix 2), but no sub-clustering within the eastern cluster (Figure S3a of Supplemental Appendix 2).

Discriminant analysis of principal components supports the separation of samples into two major clusters (eastern and western clusters) and also corroborates presence of sub-clusters within the western cluster (Figure 3). The *find.clusters* function in the *adegenet* package in *R* did not clearly identify an optimal K value, but the BIC dropped quickly between $K = 1$ and $K = 4$ (Figure S3 of Supplemental Appendix 2), suggesting four clusters, with KAP and RUM together, MRT on its own, NGU on its own, and all of the eastern clusters together (Figure 3). This aligned with the results from the STRUCTURE analysis at $K = 4$ (Figure S3a of Supplemental Appendix 2).

Table 2 shows the pairwise F_{ST} values and geographic distances in km among sampling sites. Pairwise F_{ST} ranged from 0.309 (TSW versus RUM) to 0.006 (KIB versus SHI), and averaged 0.202 between pairs spanning the two

FIGURE 2. Bayesian clustering plot obtained using the program STRUCTURE. Results for $K = 2$ are shown. Each vertical bar represents the assignment based on q -values for each individual, or proportion of ancestry into each cluster. Bars with a single color identify individuals assigned to only one cluster. Bars with multiple colors identify individuals with assignment to multiple clusters, indicating admixed ancestry. Black lines demark sample sites. They are grouped geographically from west to east along the x axis, using the same symbols as in Figure 1 and Table 1. KAP = Kapesur; KIB = Kibwezi; KIN = Kinango; MRT = Mara Talek; NGU = Nguruman; RUM = Ruma; SHI = Shimba; TSW = Tsavo west. This figure appears in color at www.ajtmh.org.

STRUCTURE-defined clusters (Table 2). Average F_{ST} was higher within the western ($F_{ST} = 0.183$) than within the eastern ($F_{ST} = 0.018$) cluster (Table 2), and this difference was significant according to the t test ($P < 0.001$). This is consistent with the finding that all pairwise F_{ST} values across clusters were significantly different from zero (Table 2), as were all F_{ST} values from pairs within the western cluster (Table 2). However, we

also found two non-significant F_{ST} values between sites within the eastern cluster (SHI and KIN: $F_{ST} = 0.009$ and SHI and KIB: $F_{ST} = 0.006$). Average per-locus AR also indicated significant differences in levels of genetic diversity between the two clusters ($P = 0.024$), with lower genetic diversity in the western ($AR = 7.48$) than the eastern ($AR = 10.99$) cluster. The Mantel test for IBD using all sampling sites indicates a positive

FIGURE 3. Results of the Discriminant analysis of principal components (DAPC), using “ade4” v1.4-2 R package. Dots represent individuals connected by a line to the centroid. Ellipses encompass 95% of the variance within each sample. Different colors represent samples from different sampling sites. KAP = Kapesur; KIN = Kinango; MRT = Mara Talek; NGU = Nguruman; RUM = Ruma; SHI = Shimba. This figure appears in color at www.ajtmh.org.

TABLE 2

Pairwise geographic distances (km) among sampling localities above the diagonal, and pairwise estimates of genetic differentiation (F_{ST}) below the diagonal

		Western				Eastern			
		KAP	RUM	MRT	NGU	KIB	TSW	KIN	SHI
Western	KAP	—	154.7	255.3	347.1	535.6	603.1	740.0	809.0
	RUM	0.138	—	123.4	240.6	453.8	511.9	639.7	713.2
	MRT	0.140	0.147	—	122.3	339.6	393.0	517.6	591.8
	NGU	0.253	0.251	0.168	—	217.4	271.6	400.0	472.9
Eastern	KIB	0.236	0.304	0.137	0.175	—	74.0	214.3	276.2
	TSW	0.227	0.309	0.129	0.178	0.021	—	140.7	205.9
	KIN	0.207	0.277	0.108	0.184	0.016	0.028	—	76.5
	SHI	0.220	0.278	0.117	0.149	0.006	0.029	0.009	—

KAP = Kapesur; KIB = Kibwezi; KIN = Kinango; MRT = Mara Talek; NGU = Nguruman; RUM = Ruma; SHI = Shimba; TSW = Tsavo west. Bold F_{ST} values indicate significance after Benjamini–Hochberg correction for multiple testing. Geographic distances were generated by the Geographic Distance Matrix Generator v1.2.3,⁶² and F_{ST} values were estimated in Arlequin v.3.5⁵⁹ with Wright's statistics,⁵⁸ following the variance method developed by Weir and Cockerham.⁵⁹

correlation between genetic and geographic distances ($P = 0.013$; Figure S4 of Supplemental Appendix 2).

The AMOVA results suggest that all four hierarchical levels contribute significantly to the observed patterns of molecular variance ($P < 0.05$): 70.3% of the variance within individuals, 9.3% among individuals, 10.0% among sampling sites, and 10.3% among the eastern and western clusters (Table S6 of Supplemental Appendix 1). These results corroborate the existence of two distinct genetic units identified by previous analyses, but also suggests that differentiation among sites is an important component in describing the genetic variance in these samples (Table S6 of Supplemental Appendix 1).

Admixture and migration. Individual assignment probability (q) for each fly based on the STRUCTURE analysis indicates that most sampling sites include a homogenous group of individuals with assignment almost exclusively to one cluster ($q > 0.8$; Figure 2; Table S7 of Supplemental Appendix 1). One exception is MRT, where nine of 30 individuals were assigned with $q < 0.8$, suggesting that these samples are the result of mating between individuals from different clusters. Admixed individuals were also present in other sampling sites: six of 22 individuals in SHI, three of 30 individuals in KIN, one of 30 individuals in TSW, and two of 30 individuals in KIB. Two individuals in MMR (MMR_005 and MMR_006) had higher assignment to the opposite cluster than the other individuals from MMR, suggesting that they were recent migrants from the eastern cluster (Table S7 of Supplemental Appendix 1).

Figure 4 shows the results of the migrant detection analyses. Using GENECLASS, we detected three first-generation migrants from the eastern cluster to the western cluster (MRT_005, MRT_006, and MRT_018) with significant P values of 0.002, 0.006, and 0.002, respectively (Table S7 of Supplemental Appendix 1). FLOCK analyses identified the same three individuals detected by the previous analyses as migrants plus an additional one, MRT_038, and one migrant (TSW_076) in the reverse direction from the western to the eastern cluster (Table S7 of Supplemental Appendix 1).

DISCUSSION

Genetic diversity and relatedness. There was significantly lower genetic diversity and a trend of higher relatedness in the western compared with the eastern genetic cluster (Table 1). This finding suggests smaller population sizes in western as compared with eastern Kenya. Smaller effective population sizes are consistent with the history of greater vector control

efforts in Western Kenya because of the higher risk of human disease, which likely created population bottlenecks and reduced effective population sizes. This is also consistent with the large-scale conversion of natural habitat into farmland in Western Kenya that did not occur in eastern Kenya, which likely reduced habitat size and connectivity among populations.^{18,67–69}

We detected a widespread pattern of heterozygote deficiency and positive F_{IS} values significantly greater than zero (Table 1) across all samples, with the most extreme deficit detected in KAP in western, and SHI in eastern Kenya. The level of inbreeding detected in this study is consistent with

FIGURE 4. Migration patterns of *Glossina pallidipes*. Arrows represent movement of first-generation migrants and progeny of recent migrants evaluated using GENECLASS and FLOCK, respectively. Number of migrants are shown on the arrowheads, with a slash separating GENECLASS and FLOCK migrants, respectively. Thickness of the arrow represents the relative number of migrants. KAP = Kapesur; KIB = Kibwezi; KIN = Kinango; MRT = Mara Talek; NGU = Nguruman; RUM = Ruma; SHI = Shimba; TSW = Tsavo west.

findings from a previous study with different loci that found similar results for two sites that overlapped with this study (RUM and NGU in western Kenya).³² Furthermore, this pattern remained significant in the subset of the data that excluded related individuals (Table S5 of Supplemental Appendix 1). A signal of inbreeding could reflect technical artifacts, or ongoing biological processes. For instance, heterozygote deficits could be due to the presence of null alleles⁴⁵ or an artificial Wahlund effect caused by non-random sampling of multiple distinct genetic groups of individuals.⁷⁰ Although possible, it is unlikely that null alleles could be responsible for the observed heterozygote deficit because there is evidence of significant deviations from HWE in some populations in the overall (Table 1) or per-locus F_{IS} estimates (Table S3 of Supplemental Appendix 1) that had no indication of null alleles in the Micro-Checker analysis in this study (Table S2 of Supplemental Appendix 1), or in past studies.³⁸ Another possibility is that these population samples were non-random, and included a biased sample from multiple distinct groups (i.e., genetic clusters or families of closely related individuals). Such non-random sampling of a panmictic population can cause population-level underestimates in heterozygosity and overestimates of F_{IS} . However, there is no indication of multiple genetic clusters in the samples with the strongest signals of heterozygote deficit (Figures 2 and 3) and the F_{IS} estimates from analyses that included/exclude related individuals (Figure S1 of Supplemental Appendix 2) showed similar patterns. These results suggest that non-random sampling of multiple genetic units or families of highly related individuals were not a major driver of the observed patterns. If indeed, our finding reflects ongoing biological processes rather than artifacts, and this suggests that *G. pallidipes* has widespread low levels of heterozygosity and high levels of inbreeding. Reasons for this could include high site fidelity, small numbers of breeding individuals, high variance in male mating success,^{55,56} or the effects of endosymbionts such as *Wolbachia*, which are well known to cause non-random mating patterns in *Glossina* spp. under some circumstances.^{13,71}

Spatial genetic differentiation and migration. To understand the amount of genetic connectivity among *G. pallidipes* populations, the levels of spatial genetic differentiation and migration were evaluated within and between sampling sites (Figure 1). All analyses found evidence of a genetic break separating populations into two genetically distinct clusters (Figure 2; Figure S1 of Supplemental Appendix 2; Table 2); one in Western and the other in eastern Kenya. These results are consistent with previous analyses on *G. pallidipes* analyzed from within the same regions.^{30,34,72} The two genetic clusters are separated by the Great Rift Valley (Figure 1). This geological landmark has been known to act as a powerful biogeographic barrier in several groups of plants, animals, and disease vectors.^{73–78} In another important disease vector, the mosquitoes from the *Anopheles gambiae* complex, genetic analyses have also shown a clear, distinct break between populations in the west and east of the Great Rift Valley.^{79,80} However, despite the clear genetic break between the western and eastern clusters coinciding with this well-known biogeographic break, our findings also suggest that this barrier has been and still is a permeable one, as we also found evidence of few past and recent genetic exchanges among the two genetic clusters. Past genetic exchanges are the likely cause of the genetic admixture signals

found in several locations, with MRT being the one population displaying the highest level (Figure 2; Table S7 of Supplemental Appendix 1). Recent ongoing migration events are the likely reason for finding first-generation migrants between sampling sites from the two clusters (Figure 4).

Within the western cluster, significant levels of differentiation among all sites were detected (mean F_{ST} = 0.183, Table 2), whereas mean genetic differentiation among sites within the eastern cluster were lower and not all were significant (mean F_{ST} = 0.018, Table 2). This difference in the level of differentiation among sites may, in part, reflect the larger geographic distances among the western sites (range: 110–150 km) than the eastern sites (range: 76–276 km). However, differences in geographic distance cannot fully explain the lower F_{ST} found in the eastern cluster. For example, the two eastern sites SHI and KIB (Figure 1) are separated by more geographic distance than any pair of western sites (276.2 km), but do not show significant differentiation. Given that the ability for flies to disperse over their lifetime (~9.7 km at the most⁸¹) is far lower than the distance between these sites, the low differentiation among sites in the eastern cluster suggests higher habitat connectivity⁸² than that found in the western cluster. The more homogenous and less human-disturbed habitat in eastern Kenya,⁶⁷ together with a history of less intense vector control activities likely facilitated the persistence of larger, more stable, and more genetically connected tsetse populations in eastern than in western Kenya.

Conclusions and implications for vector control. The finding of different levels of genetic diversity within *G. pallidipes* samples from western and eastern Kenya, and the occurrence of a genetic break among sites for these two regions provide important insights for the implementation and monitoring of ongoing and future vector control strategies. First, they show that population genetics can identify population-level aspects of *G. pallidipes* biology, such as levels of genetic inbreeding, diversity, and connectivity with neighboring populations, that can inform vector control. Second, these data allow inferring the spatial scale over which control efforts are most effective for this species in this region. Although all analyses indicated a strong genetic break across the Great Rift Valley, which is likely to have been generated thousands of years ago, our analyses also identified recent long-range dispersal events between samples from the two clusters. From a vector control perspective, this study identified the geographic regions and their spatial extent that should be controlled simultaneously, as sampling sites within each cluster are genetically connected. This means that vector control has to be carried out over an area of about 500 km, the spatial extent of the two genetic clusters. Monitoring should be conducted more intensely in eastern than in western Kenya, as populations in the east are more genetically connected, a condition that increases the possibility of reinvasion from neighboring sites if they are not treated at the same time. Unfortunately, the fact that we detected recent migrants among clusters also suggests that, albeit rare, gene flow can occur between geographically distant regions. How much these rare events play a role in restoring tsetse population density and success after control is debatable, and is a topic to be pursued in future studies possibly via modeling approaches. In any case, the fact that we did find long-range migrants implies that monitoring efforts are a necessary strategy for detecting and halting re-establishment of tsetse from distant locations.

Received February 20, 2018. Accepted for publication June 15, 2018.

Published online August 13, 2018.

Note: Supplemental appendices, tables, and figures appear at www.ajtmh.org.

Acknowledgments: We would like to acknowledge Joanna Auma, Paul Thande, Patrick Obore, Lillian Mwendu Mwaniki, Rose Wanjiru Ndung'u, the Biotechnology Research Institute, and the Kenya Wildlife Services and their resourceful rangers for the immense amount of input and assistance in sample collection and DNA extraction. We also acknowledge Carol Mariani of the DNA Analysis Facility at Yale University for genotyping of the samples. The research was accomplished while WAO was a Research Fellow at Yale University.

Financial support: NIH Grant no. U01 AI115648; NIH-Fogarty Global Infectious Diseases Training Grant (D43TW007391).

Authors' addresses: Winnie A. Okeyo, Department of Biomedical Sciences and Technology, School of Public Health and Community Development, Maseno University, Kisumu, Kenya, Biotechnology Research Institute, Kenya Agricultural and Livestock Research Organization, Nairobi, Kenya, and Yale School of Public Health, Yale University, New Haven, CT, E-mail: okeyo.winnie@gmail.com. Norah P. Saarman, Kirstin Dion, Michael Mengual, and Adalgisa Caccone, Department of Ecology & Evolutionary Biology, Yale University, New Haven, CT, E-mails: norah.saarman@yale.edu, kirstin.dion@yale.edu, michael.mengual@yale.edu, and adalgisa.caccone@yale.edu. Rosemary Bateta, Sylvance Okoth, and Grace Murilla, Biotechnology Research Institute, Kenya Agricultural and Livestock Research Organization, Nairobi, Kenya, E-mails: batetarw@yahoo.com, sokotho@gmail.com, and gmurilla@yahoo.co.uk. Paul O. Mireji, Biotechnology Research Institute, Kenya Agricultural and Livestock Research Organization, Nairobi, Kenya, and Yale School of Public Health, Yale University, New Haven, CT, E-mail: mireji.paul@gmail.com. Collins Ouma, School of Public Health and Community Development, Maseno University, Kisumu, Kenya, E-mail: profcollinsouma@gmail.com. Serap Aksoy, Yale School of Public Health, Yale University, New Haven, CT, E-mail: serap.aksoy@yale.edu.

REFERENCES

- Buguet A, Mpanzou G, Bentivoglio M, 2014. Human african trypanosomiasis: a highly neglected neurological disease. Bentivoglio M, Cavalheiro E, Kristensson K, Patel N, eds. *Neglected Tropical Diseases and Conditions of the Nervous System*. New York, NY: Springer, 165–181.
- Courtin F et al., 2015. Reducing human-tsetse contact significantly enhances the efficacy of sleeping sickness active screening campaigns: a promising result in the context of elimination. *PLoS Negl Trop Dis* 9: e0003727.
- Alsan M, 2015. The effect of the tsetse fly on African development. *Am Econ Rev* 105: 382–410.
- Fèvre EM, Wissmann BV, Welburn SC, Lutumba P, 2008. The burden of human African trypanosomiasis. *PLoS Negl Trop Dis* 2: e333.
- Muhanguzi D, Picozzi K, Hattendorf J, Thrusfield M, Kabasa JD, Waiswa C, Welburn SC, 2014. The burden and spatial distribution of bovine African trypanosomes in small holder crop-livestock production systems in Tororo District, south-eastern Uganda. *Parasit Vectors* 7: 603.
- La Greca F, Magez S, 2011. Vaccination against trypanosomiasis. *Hum Vaccin* 7: 1225–1233.
- Aksoy S, 2003. Control of tsetse flies and trypanosomes using molecular genetics. *Vet Parasitol* 115: 125–145.
- CDC, 2015. *Parasites—Africa trypanosomiasis (Also Known as Sleeping Sickness)*. Centres for Disease Control and Prevention. Available at: <https://www.cdc.gov/parasites/sleepingsickness/treatment.html>. Accessed July 13, 2017.
- Anene BM, Onah DN, Nawa Y, 2001. Drug resistance in pathogenic African trypanosomes: what hopes for the future? *Vet Parasitol* 96: 83–100.
- Wilkinson SR, Kelly JM, 2009. Trypanocidal drugs: mechanisms, resistance and new targets. *Expert Rev Mol Med* 11: e31.
- Barrett MP, Vincent IM, Burchmore RJ, Kazibwe AJ, Matovu E, 2011. Drug resistance in human African trypanosomiasis. *Future Microbiol* 6: 1037–1047.
- Welburn SC, Molyneux DH, Maudlin I, 2016. Beyond tsetse—implications for research and control of human African trypanosomiasis epidemics. *Trends Parasitol* 32: 230–241.
- Aksoy S, Caccone A, Galvani AP, Okedi LM, 2013. *Glossina fuscipes* populations provide insights for human African trypanosomiasis transmission in Uganda. *Trends Parasitol* 29: 394–406.
- Brightwell R, Dransfield R, 1997. Odour attractants for tsetse: *Glossina austeni*, *G. brevipalpis* and *G. swynnertoni*. *Med Vet Entomol* 11: 297–299.
- Dransfield RD, Brightwell R, Kyorku C, Williams B, 1990. Control of tsetsefly (Diptera: Glossinidae) populations using traps at Nguruman, south-west Kenya. *Bull Entomol Res* 265–276.
- Echessah PN, Swallow BM, Kamara DW, Curry JJ, 1997. Willingness to contribute labor and money to tsetse control: application of contingent valuation in Busia District, Kenya. *World Dev* 25: 239–253.
- Maudlin I, 2006. African trypanosomiasis. *Ann Trop Med Parasitol* 100: 679–701.
- Muriuki GW, Njoka TJ, Reid RS, Nyariki DM, 2005. Tsetse control and land-use change in Lambwe valley, south-western Kenya. *Agric Ecosyst Environ* 106: 99–107.
- Schofield CJ, Kabayo JP, 2008. Trypanosomiasis vector control in Africa and Latin America. *Parasit Vectors* 1: 24.
- Torr SJ, Hargrove JW, Vale GA, 2005. Towards a rational policy for dealing with tsetse. *Trends Parasitol* 21: 537–541.
- Vreysen MJ, 2001. Principles of area-wide integrated tsetse fly control using the sterile insect technique. *Med Trop (Mars)* 61: 397–411.
- Wellde BT, Waema D, Chumo DA, Reardon MJ, Oloo F, Njogu AR, Opiyo EA, Mugutu S, 1989. Review of tsetse control measures taken in the Lambwe Valley in 1980–1984. *Ann Trop Med Parasitol* 83 (Suppl 1): 119–125.
- Mbewe NJ, Saini RK, Torto B, Irungu J, Yusuf AA, Pirk CWW, 2018. Sticky small target: an effective sampling tool for tsetse fly *Glossina fuscipes fuscipes* Newstead, 1910. *Parasit Vectors* 11: 268.
- Shaw AP, Tirados I, Mangwiro CT, Esterhuizen J, Lehane MJ, Torr SJ, Kovacic V, 2015. Costs of using “tiny targets” to control *Glossina fuscipes fuscipes*, a vector of gambiense sleeping sickness in Arua District of Uganda. *PLoS Negl Trop Dis* 9: e0003624.
- Lindh JM, Torr SJ, Vale GA, Lehane MJ, 2009. Improving the cost-effectiveness of artificial visual baits for controlling the tsetse fly *Glossina fuscipes fuscipes*. *PLoS Negl Trop Dis* 3: e474.
- Arora AK, Douglas AE, 2017. Hype or opportunity? Using microbial symbionts in novel strategies for insect pest control. *J Insect Physiol* 103: 10–17.
- WHO, 2017. *Human African trypanosomiasis*. WHO Factsheets. Available at: <http://www.who.int/mediacentre/factsheets/fs259/en/>. Accessed May 2, 2017.
- Messina JP, Moore NJ, DeVisser MH, McCord PF, Walker ED, 2012. Climate change and risk projection: dynamic spatial models of tsetse and African trypanosomiasis in Kenya. *Ann Assoc Am Geogr* 102: 1038–1048.
- Wolf T, Wichelhaus T, Göttig S, Kleine C, Brodt HR, Just-Nuebling G, 2012. *Trypanosoma brucei rhodesiense* infection in a German traveller returning from the Masai Mara area, Kenya, January 2012. *Euro Surveill* 17: pii:20114.
- Ouma JO, Marquez JG, Krafur ES, 2005. Macrogeographic population structure of the tsetse fly, *Glossina pallidipes* (Diptera: Glossinidae). *Bull Entomol Res* 95: 437–447.
- FAO, 2006. *Tsetse Fly Habitat and Land Cover: An Analysis at Continental Level*. Rome, Italy: GLCN, FAO, 1–16.
- Ouma JO, Krafur ES, 2010. The influence of temporal and seasonal changes on genetic diversity and population structure of the tsetse fly, *Glossina pallidipes* in Kenya. *East African Agric Forum J* 77: 59–68.
- Williams B, Dransfield R, Brightwell R, 1990. Monitoring tsetse fly populations. I. The intrinsic variability of trap catches of *Glossina pallidipes* at Nguruman, Kenya. *Med Vet Entomol* 4: 167–179.
- Ouma JO, Marquez JG, Krafur ES, 2006. Microgeographical breeding structure of the tsetse fly, *Glossina pallidipes* in south-western Kenya. *Med Vet Entomol* 20: 138–149.
- Swai ES, Kaaya JE, 2012. A parasitological survey for bovine trypanosomiasis in the livestock/wildlife ecozone of northern Tanzania. *Vet World* 5: 459–464.

36. Hyseni C, Kato AB, Okedi LM, Masembe C, Ouma JO, Aksoy S, Caccone A, 2012. The population structure of *Glossina fuscipes fuscipes* in the Lake Victoria basin in Uganda: implications for vector control. *Parasit Vectors* 5: 222.
37. Echodu R, Sistrom M, Hyseni C, Enyaru J, Okedi L, Aksoy S, Caccone A, 2013. Genetically distinct *Glossina fuscipes fuscipes* populations in the Lake Kyoga region of Uganda and its relevance for human African trypanosomiasis. *BioMed Res Int* 2013: 614721.
38. Ouma JO, Beadell JS, Hyseni C, Okedi LM, Krafur ES, Aksoy S, Caccone A, 2011. Genetic diversity and population structure of *Glossina pallidipes* in Uganda and western Kenya. *Parasit Vectors* 4: 122.
39. Krafur ES, 2002. Population structure of the tsetse fly *Glossina pallidipes* estimated by allozyme, microsatellite and mitochondrial gene diversities. *Insect Mol Biol* 11: 37–45.
40. Solano P, Ravel S, de Meeus T, 2010. How can tsetse population genetics contribute to African trypanosomiasis control? *Trends Parasitol* 26: 255–263.
41. Solano P et al., 2010. Population genetics as a tool to select tsetse control strategies: suppression or eradication of *Glossina palpalis gambiensis* in the Niayes of Senegal. *PLoS Negl Trop Dis* 4: e692.
42. Okeyo WA et al., 2017. Temporal genetic differentiation in *Glossina pallidipes* tsetse fly populations in Kenya. *Parasit Vectors* 10: 471.
43. Dransfield RD, Brightwell R, Kiilu J, Chaudhury MF, Abie DAAD, 1989. Size and mortality rates of *Glossina pallidipes* in the semi-arid zone of southwestern Kenya. *Med Vet Entomol* 3: 83–95.
44. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P, 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538.
45. Brookfield JF, 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* 5: 453–455.
46. Rousset F, 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 8: 103–106.
47. Benjamini Y, Hochberg Y, 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57: 289–300.
48. Goudet J, 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available at: <https://www2.unil.ch/popgen/softwares/fstat.htm>. Accessed February 22, 2017.
49. Excoffier L, Lischer HE, 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567.
50. Kalinowski ST, Wagner AP, Taper ML, 2006. ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* 6: 576–579.
51. Evanno G, Regnaut S, Goudet J, 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611–2620.
52. Earl DA, vonHoldt BM, 2011. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4: 359–361.
53. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I, 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour* 15: 1179–1191.
54. Jombart T, Devillard S, Balloux F, 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11: 94.
55. Jombart T, 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
56. R Core Team, 2011. *R: A Language And Environment For Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>. Accessed February 14, 2017.
57. Jombart T, Collins C, 2015. *A Tutorial for Discriminant Analysis of Principal Components (DAPC) Using Adegenet 2.0.0*. London, United Kingdom: Imperial College London, MRC Centre for Outbreak Analysis and Modelling, 1–43.
58. Wright S, 1951. The genetical structure of populations. *Ann Eugen* 15: 323–354.
59. Weir BS, Cockerham CC, 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
60. Jensen JL, Bohonak AJ, Kelley ST, 2005. Isolation by distance, web service. *BMC Genet* 6: 13.
61. Smouse PE, Long JC, Sokal RR, 1986. Multiple regression and correlation extensions of the mantel test of matrix correspondence. *Syst Zool* 35: 627–632.
62. Ersts PJ, 2011. *Geographic Distance Matrix Generator (Version 1.2.3)*. New York City, NY: American Museum of Natural History.
63. Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A, 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J Hered* 95: 536–539.
64. Duchesne P, Turgeon J, 2009. FLOCK: a method for quick mapping of admixture without source samples. *Mol Ecol Resour* 9: 1333–1344.
65. Paetkau D, Slade R, Burden M, Estoup A, 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* 13: 55–65.
66. Rannala B, Mountain JL, 1997. Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci USA* 94: 9197–9201.
67. DeVisser MH, Messina JP, 2009. Optimum land cover products for use in a *Glossina morsitans* habitat model of Kenya. *Int J Health Geogr* 8: 39.
68. Osano PM, Said MY, deLeeuw J, Moiko SS, OleKaelo D, Schomers S, Birner R, Ogutu JO, 2013. Pastoralism and ecosystem-based adaptation in Kenyan Masailand. *Int J Clim Chang Strateg Manag* 5: 198–214.
69. Wamwiri FN, Changasi RE, 2016. Tsetse flies (*Glossina*) as vectors of human African trypanosomiasis: a review. *Biomed Res Int* 2016: 8.
70. Zhivotovskiy LA, 2015. Relationships between Wright's FST and FIS statistics in a context of Wahlund effect. *J Hered* 106: 306–309.
71. Alam U et al., 2012. Implications of microfauna-host interactions for trypanosome transmission dynamics in *Glossina fuscipes fuscipes* in Uganda. *Appl Environ Microbiol* 78: 4627–4637.
72. Krafur ES, Wohlford DL, 1999. Breeding structure of *Glossina pallidipes* populations evaluated by mitochondrial variation. *J Hered* 90: 635–642.
73. Ruiz Guajardo JC, Schnabel A, Ennos R, Preuss S, Otero-Arnaiz A, Stone G, 2010. Landscape genetics of the key African acacia species *Senegalia mellifera* (Vahl)—the importance of the Kenyan Rift Valley. *Mol Ecol* 19: 5126–5139.
74. Bertola LD et al., 2011. Genetic diversity, evolutionary history and implications for conservation of the lion (*Panthera leo*) in west and central Africa. *J Biogeography* 38: 1356–1367.
75. Brown DM, Brenneman RA, Koepfli KP, Pollinger JP, Milá B, Georgiadis NJ, Louis EE Jr., Grether GF, Jacobs DK, Wayne RK, 2007. Extensive population genetic structure in the giraffe. *BMC Biol* 5: 57.
76. Dubach J, Patterson BD, Briggs MB, Venzke K, Flamand J, Stander P, Scheepers L, Kays RW, 2005. Molecular genetic variation across the southern and eastern geographic ranges of the African lion, *Panthera leo*. *Conserv Genet Resour* 6: 15–24.
77. Evans BJ, Bliss SM, Mendel SA, Tinsley RC, 2011. The Rift Valley is a major barrier to dispersal of African clawed frogs (*Xenopus*) in Ethiopia. *Mol Ecol* 20: 4216–4230.
78. Franco JR, Simarro PP, Diarra A, Jannin JG, 2005. Epidemiology of human African trypanosomiasis. *Clin Epidemiol* 6: 257–275.
79. Lehmann T, Hawley WA, Grebert H, Danga M, Atieli F, Collins FH, 1999. The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya. *J Hered* 90: 613–621.
80. Lehmann T, Licht M, Elissa N, Maega BTA, Chimumbwa JM, Watsenga FT, Wondji CS, Simard F, Hawley WA, 2003. Population structure of *Anopheles gambiae* in Africa. *J Hered* 94: 133–147.
81. Cano J, Descalzo MA, Ndong-mabale N, Ndong-asumu P, Bobuakasi L, Nzambo-ondo S, Benito A, Roche J, 2007. Predicted distribution and movement of *Glossina palpalis palpalis* (Diptera: Glossinidae) in the wet and dry seasons in the Kogo trypanosomiasis focus (Equatorial Guinea). *J Vector Ecol* 32: 218–225.
82. Wint W, Rogers D, 2000. *Predicted Distributions of Tsetse in Africa*. Oxford, United Kingdom: FAO Consultants' Report, ERGO Ltd and the TALA Research Group.