Masai giraffe rifting apart: Loss of genetic connectivity across the Gregory Rift Valley in Tanzania

George G. Lohay[1], David Pearce[2], Lan Wu-Cavener[1], Xiaoyi Hou[1], Monica L. Bond[3,4], Derek E. Lee[1,3], Douglas R. Cavener [1*]

1. Biology department, Penn State University, 110 Life Sciences Building, University Park, PA, USA
2. Department of Ecosystem Science and Management, Penn State University, University Park, PA, USA
3. Wild Nature Institute, Concord, NH, USA
4. Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

[*]Corresponding author drc9@psu.edu

Abstract

The Masai giraffe has experienced a population decline from 70,000 to 35,000 in the past three decades and was declared an endangered subspecies by the IUCN in 2019. The remaining population is divided into smaller subpopulations dispersed west and east of the Gregory Rift Valley (GRV) in Tanzania, Kenya, and Zambia. The steep escarpments of the GRV are formidable barriers to migration and gene flow and the few remaining natural corridors are now occupied by human settlements. To assess the impact of the GRV on Masai giraffe gene flow, we examined nuclear and mitochondrial DNA (mtDNA) variation in subpopulations east and west of the Manyara and Eyasi escarpments of the Rift in northern Tanzania. Strikingly, evidence from mtDNA variation, which measures female-mediated gene flow, suggests that females have not migrated across the GRV between major subpopulations in the Serengeti and Tarangire

ecosystems since the approximate time that Masai giraffes emerged as a (sub)species 250kya. In contrast the analysis of nuclear DNA variation shows that male-mediated gene flow across the GVR has occurred over most of the history of Masai giraffes but may have ceased in the recent past as natural migratory routes across the escarpment have been blocked by human activities. These findings suggest that the Masai giraffe is now split into two separate metapopulations and in the absence of gene flow will eventually become separate species with populations of less than 20,000.

Keywords: Masai Giraffe, Genetic connectivity, Wildlife Corridors, Gregory Rift

Conservation Genetics, Genomics/Proteomics

Introduction

As a result of human activities wild animal populations have declined over the past 10,000 years and now account for less than 4% of mammals on the planet, with humans, pets, and livestock constituting ~96% (Bar-On, Phillips, & Milo, 2018; Ritchie & Roser, 2021). This downward trend has accelerated over the past 30 years for the charismatic megaherbivores on the African continent including giraffes (*Giraffa camelopardalis*), elephants (*Loxodonta cyclotis* and *L. Africana*), and rhinoceros (*Diceros bicornis* and *Ceratotherium simum*), leaving highly fragmented populations separated by land converted to agriculture and other human activities. Giraffe populations declined by 40% since 1990s to less than 100,000 individuals in parallel with a more than doubling of the human population in the sub-Saharan African countries that still harbor wild giraffe populations (Bolger et al., 2019). In this study we focused on the Masai giraffe (subspecies *G. c. tippelskirchi*: (Muller et al., 2018) species *G. tippelskirchi*: (Coimbra et al., 2021; Petzold & Hassanin, 2020),

which number only about 35,000 located in protected areas and national parks in Tanzania and southern Kenya and in a small population in Zambia. The Masai giraffe was listed an endangered subspecies (IUCN Red List) after their population declined by 50% in the preceding three decades (Bolger et al., 2019). The Gregory Rift Valley (GRV), which cuts through Kenya and Tanzania, bisects the Masai giraffe range into two regions: West GRV and East GRV populations with ca. 14,000 and ca. 21,000 individuals, respectively. The steep Manyara and Eyasi escarpments of the GRV system impose formidable barriers to animal movements (Baker, Mohr, & Williams, 1972; Scoon, 2018) but natural corridors traversing these escarpments were previously shown to act as wildlife movement routes (Lohay, Weathers, Estes, McGrath, & Cavener, 2020; Prins & Jong, 2022). However, these natural corridors are also ideal locations for roads, towns, and agricultural development, and during the past 30 years most of these corridors have rapidly been overcome by human activities and are unlikely to support wildlife migration. In regions without major geographic barriers Masai giraffes can roam over large areas (e.g., mean home range of 114 km^2 for females and 157 km^2 for males; (Knüsel, Lee, König, & Bond, 2019), and observational data have documented connectivity movements of individuals across the Tarangire-Manyara-Ecosystem (Lee & Bolger, 2017). Although males are more likely to mediate gene flow between populations, young adult female Masai giraffe have been shown to also disperse to adjacent populations, sometimes moving more than 26 km from their natal area (Bond, König, Ozgul, Farine, & Lee, 2021; Ferres et al., 2021).

A critical question germane to the long-term survival of Masai giraffes is whether populations located east and west of the GVR escarpments are reproductively isolated. Population genetic analysis provides a means to ascertain gene flow between populations, and female- versus male-

mediated gene flow can be assessed by comparing mitochondrial DNA (mtDNA) variation versus nuclear DNA variation because mtDNA is strictly maternally inherited whereas nuclear DNA is inherited from both parents (Abdul-Muneer, 2014; Allendorf, 2017; Allendorf, Hohenlohe, & Luikart, 2010; Miah et al., 2013). An earlier study of the population genetics of the major giraffe subspecies (Brown, Brenneman, Koepfli, Pollinger, Mila, et al., 2007) showed a significant differentiation of Masai giraffe mitochondrial haplotypes but not nuclear DNA variation among several populations in Tanzania and Kenya suggesting possible differences in male versus female mediated gene flow. The objective of this study was to determine if the escarpments of the GVR posed impediment to sex-dependent gene flow between West Rift Valley (WRV) and East Rift Valley (ERV) Masai giraffe populations, and to potentially date the time period when gene flow ceased to occur. To address these questions, we obtained WGS of mtDNA and nuclear microsatellite markers genotypic data for a large number of giraffes located east and west of the GRV in northern Tanzania where the major remaining populations of Masai giraffe still exist.

Materials and Methods

Study sites

The ca. 25,000 km^2 Tarangire Manyara Ecosystem and the 33,000 km^2 Serengeti Ecosystem are two of the most critical ecosystems in Tanzania for biodiversity conservation (Fig.1). Both ecosystems conserve biodiversity and large landscapes, and support Africa's most numerous migrations of large mammals including white-bearded wildebeests (*Connochaetes taurinus*) and plains zebras (*Equus quagga*) (Estes, 2014). Tarangire-Manyara and Serengeti are two of only three long-distance migrations of wildebeests remaining in Africa (Estes, 2014; Morrison, 2014).

Both systems also host important populations of Masai giraffes (Strauss et al., 2015; Lee & Bolger, 2017; Lee et al., 2022).

The Tarangire Manyara Ecosystem consists of several protected areas including Tarangire National Park (TNP), Manyara Ranch Conservancy (MRC), Lake Manyara National Park (LMNP), and Burunge Wildlife Management Area (BWMA). Since its establishment in 1950s, changes in the landscape due to human activities have significantly reduced wildlife corridors (Borner, 1985; Lamprey, 1964; Morrison, Link, Newmark, Foley, & Bolger, 2016; Mwalyosi, 1991). Manyara Ranch Conservancy (MRC) is a wildlife area unique in Tanzania, as it does not fall in the categories of formal protected areas. Rather, MRC is an open area supported by African Wildlife Foundation for wildlife conservation and livestock keeping, and functions as part of a wildlife corridor between TNP and LMNP and the Lake Natron area (Bond, Kiffner, & Lee, 2022). BWMA, located between TNP and LMNP, is a community-based conservation initiative started about 20 years ago by several villages (Bluwstein, 2018). This WMA is used for promoting eco-tourism and provides habitat for several wildlife species. The WMA is also part of the corridor connecting TNP and MRC and Lake Natron (Kiffner et al., 2020; Lee, 2018). Serengeti National Park (SNP) and Ngorongoro Conservation Area (NCA) form a major part of the Serengeti Ecosystem. While SNP is reserved for photo tourism and wildlife management, the NCA allows multiple land uses, although the only economic activities allowed in the NCA are tourism and pastoralism. Over the past few years, the number of livestock and humans has increased within the NCA (Masao, Makoba, & Sosovele, 2015).

Migratory routes

Giraffes are not considered to be long-distance migrants, but seasonality drives their resource use which may result in short-distance seasonal migrations between wet and dry season ranges (Pellew 1984; Brown & Bolger 2020). However, over the longer term, individual giraffes may perform longer-distance natal or breeding dispersal movements into new populations, which here we define as 'migration'. The giraffe populations west of the Rift in the SNP and NCA are contiguous with no impediments to migration. Populations east of Rift including LMNP, MRC, BWMA, and TNP are in proximity (Fig. 1), but migration between these populations is constrained by the presence of Lake Manyara and intensive agriculture and human habitation. LMNP, comprised of narrow strip of land between the escarpment to the west and Lake Manyara to the east, is particularly isolated and historical migration routes through the southwest and northeast around Lake Manyara are now largely blocked by agriculture and townships. Similarly, the Makuyuni wildlife corridor between MRC and TNP has seen a recent dramatic increase in agriculture and human populations, reducing possible wildlife migration (Kikoti & Griffin, 2009; Lohay, Riggio, Lobora, Kissui, & Morrison, 2022; Msofe, Sheng, & Lyimo, 2019). Among these four giraffe populations east of the Rift, the migratory routes between BWMA and TNP appear to be the most intact. By sharp contrast, migratory routes across the steep escarpments of the GRV have always been limited. Within a 50 km radius of Lake Manyara, which is the approximate center of the four East Rift populations, the Selela-Kitete wildlife corridor is the only known active wildlife corridor across the Rift that would connect the West and East Rift giraffe populations studied herein and entails a ~100 km trek across the Ngorongoro Highlands (Fig.1). The Selela-Kitete corridor was established in 1970 and the migration of elephants continues to occur to the present (PAMS personal communication, SIT paper). However, PAMS have stated that giraffe have not used this corridor

in recent years and may have not used it since it was established. Although there are a few other possible migratory routes through the escarpment in this area, intensive agriculture in the Karatu and Rotia townships just above the escarpment and leading up to the Ngorongoro Highlands precludes effective wildlife movement. Giraffe migration across the Rift is likely to have only occurred either southwest of Lake Manyara via a ~400 km trek around two escarpments and Lake Eyasi (Fig. 1) or north of MRC near Engaresero and Lake Natron (Fig. 1) which would involve ~200 km trek. The northern migratory route appears the most likely to be used by giraffes to cross the Manyara Escarpment separating the West and East Rift populations. Giraffes are continually sighted over the 85 km distance between MRC and Lake Natron, and animal trails and roads traverse the escarpment from Engaresero area to the top of the escarpment and down to the Salei Plains which leads to the NCA and SNP.

Field data collection

Fecal samples

We conducted fieldwork in the six protected areas in Tanzania (Fig. 1) between December 2019 and March 2021. Once giraffes were sighted, we observed and waited for them to defecate. We recorded sex and estimated the age of each giraffe (Strauss, Kilewo, Rentsch, & Packer, 2015). We collected giraffe fecal samples as soon as possible after defecation because giraffe pellets dry quickly. The epithelial cells adhering to the outside layer of pellets (2-4 pellets) were collected. A razor blade was used to scrape/peel the thin outer layer from each pellet (Austin et al. 2018) and placed it into a 50 ml tube. Queen's college buffer (Ahlering et al., 2012) was added immediately into the tube containing samples. We recorded GPS coordinates for each sample collected.

Tissue sample collection

Remote biopsy darts were used to obtain tissue samples from 100 wild giraffes. Darting was performed by a trained veterinarian from the Tanzania Wildlife Research Institute. Pneu darts were used with remote biopsy device type U (3cc) that is suitable for giraffes. The vet aimed at flat surfaces on thighs or shoulders. Darting was done from a distance between 10-20 m. Tissue samples were removed from the needles, placed in a 2ml microcentrifuge tube, and secured in a cool box with ice packs and frozen in -20°C after 6 hours.

DNA extraction, PCR amplification, and sequencing of mitochondrial DNA

Fecal DNA was extracted using the QIAamp PowerFecal DNA kit (QIAGEN) and tissue DNA was isolated using Monarch Nucleic Acid Purification Kits using manufacturer's protocol, but we increased incubation time to 12 hrs to ensure the whole tissue was completely lysed. DNA was extracted at the Nelson Mandela African Institution of Science and Technology.

We PCR amplified 1140nt long of cytochrome b gene (Bock et al., 2014). PCR amplification was performed using at least 10ng of DNA template, 0.5µl of 10 µM of both primers, 7.5 µl of 2x GoTaq master mix (Promega), and 3 µl of DNA template. The PCR reaction was performed with the initial polymerase activation step at 95°C for 3 min, denaturation at 95°C for 30 sec, annealing temperature at 58°C for 45 seconds, and extension at 72°C for 30 seconds for 35 cycles. We sequenced PCR products using both forward and reverse primers. We visually inspected sequence results in the trace file format using SnapGene® software 4.2.4 (from GSL Biotech; available at snapgene.com). Clean sequences were aligned with a previously published sequence of a giraffe from Tanzania (Brown et al. 2007; Bock et al. 2014; Coimbra et al. 2021). We trimmed sequences and collapsed haplotypes using FaBox (VILLESEN 2007). Haplotype diversity (H_d) and

nucleotide diversity (π) were calculated from DnaSP whereas pairwise genetic fixation (F_{ST}) was calculated using the Arlequin version 3.5 (EXCOFFIER & LISCHER, 2010). A median-joining network was constructed using PopArt 4.8.4 (Leigh & Bryant, 2015). Mantel test was conducted using IBD program (Bohonak, 2002) to test for correlations between Slatkin's linearized pairwise F_{ST} and geographical distance(François Rousset, 1997). DNA extracted from 101 dart biopsy tissue samples were subjected to Illumnia sequence to obtain the entire 16,430nt sequence of the mtDNA. A uniquely indexed library was made from the samples using Illumina DNA Prep Kit which uses PCR at the Pennsylvania State University Huck Institute's genomic core and sequenced at the Pennsylvania State University Hershey genomics core on an Illumnia Novaseq.

Amplification of microsatellite loci

Microsatellite analyses was performed using DNA extracted from tissue samples only. We selected 31 microsatellites that were previously shown to be highly polymorphic in one or more giraffe (sub)species (D. M. Brown, Brenneman, Koepfli, Pollinger, Mila, et al., 2007; Carter, Seddon, Carter, Goldizen, & Hereward, 2012; Crowhurst, Mullins, Mutayoba, & Epps, 2013; Huebinger et al., 2002). Refer to Table S2 for detailed information. We used Miseq Illumina sequencing technology which provides reads long enough to fully cover most microsatellite loci (Barbian et al. 2018). Illumina overhanging adapters were added to both forward and reverse primers at the 5' prime end: forward overhang TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and reverse overhang GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG. The first round of PCR was run to obtain products which have the Illumina adapter overhang. The PCR reagents and conditions were the same as for mtDNA amplification. Equal volume of each of the PCR amplicons was pooled in a single 96-well plate for each individual (Barbian et al., 2018).

Library preparations and MiSeq sequencing

The second round of PCR was performed at the Pennsylvania State University Huck Institute's genomic core facility, in which the indexes and the rest of the Illumina adapters were added. An approximately equimolar pool of libraries was made. The pool was sequenced using standard MiSeq 250 x 250 paired-end sequencing run and expected to get ~10 million pairs of reads in total. Once sequencing was completed, the sequence reads were demultiplexed into separate files for each sample based on unique indexes, and the adapter sequences were removed by default. Downstream analysis was performed using CHIIMP pipeline (Barbian et al., 2018).

MiSeq data analysis workflow

The raw sequence data were inspected using FastQC, which revealed the existence of residual adapter sequences. The adapter sequences were trimmed with cutadapt 3.5 with Python 3.8.12. Then, the paired-end Illumina reads were assembled using PANDAseq with the default algorithm (Masella, Bartram, Truszkowski, Brown, & Neufeld, 2012). The assembled sequences were fed into the *Computational High-throughput Individual Identification through Microsatellite Profiling* (CHIIMP) pipeline which was used to call alleles for each microsatellite loci following the GUIDE for CHIIMP R package.

Population genetic analyses

A Bonferroni correction for multiple comparisons was applied using a Holm-Bonferroni sequential correction for both Hardy-Weinberg equilibrium (HWE) (HOCHBERG 1988; Rice 1989) using a web based Genepop program (Raymond & Rousset, 1995; Francois Rousset, 2008). We then

checked for the presence of null alleles/allele dropout using MICRO-CHECKER (OOSTERHOUT, HUTCHINSON, WILLS, & SHIPLEY, 2004). Microsatellites that did not meet the assumptions of HWE or hard significant number of null alleles were omitted from the analysis. We retained loci that are polymorphic. Indices of genetic diversity such as observed heterozygosity (Ho), expected heterozygosity (He) and number of alleles per population were calculated using GenAIEX (Peakall & Smouse, 2012)and allelic richness (AR) were calculated using FSTAT (Goudet, 1995). To quantify the level of genetic differentiation between subpopulations we computed F_{ST} using Arlequin 3.5 (EXCOFFIER & LISCHER, 2010). To detect if there is isolation by distance, we used a Mantel test implemented in program R using the adegenet package (Jombart, 2008). We performed a statistical analysis of the F_{ST} normalized to geographic distance (F_{ST}/Km) within the Tarangire Manyara Ecosystem (East of the Rift) and between populations across the Rift (excluding LMNP microsatellite data due to insufficient sample size) to determine if the average F_{ST} normalized to geographic distance is significantly higher across the rift than within ecosystems west and east of the Rift.

Inference of genetic structure was made using Bayesian clustering algorithms implemented in STRUCTURE 2.3.4 (Falush, Stephens, & Pritchard, 2003; Pritchard, 1, & Donnelly, 2000). We sampled 100,000 steps following a 100,000-step burn-in, with 10 replicates for the maximum number of populations (K) between 1 and 10. The CLUMPAK webserver was used to obtain average results and to infer the optimum K based on posterior probability of K (Pritchard et al., 2000), and delta K(ΔK) generated from STRUCTURE harvester (Evanno, Regnaut, & Goudet, 2005). We then used CLUMPAK output for the most likely value of K (Francis, 2017; Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) and used it to create pie charts for each

individual and displayed on a map to help visualize data using ArcGIS map. Principal coordinate analysis implemented in GENAIEX (Peakall & Smouse, 2012) was used to detect the number of clusters using F_{ST} values.

We used Discriminant Analysis of Principal Components (DAPC), a multivariate method, designed to identify and describe clusters of genetically related individuals when group priors are lacking (Jombart, Devillard, & Balloux, 2010). DAPC implemented in the R package adegenet (Jombart, Devillard, Dufour, & Pontier, 2008), transforms multi-locus genotype data using principal component (PC) analysis to derive the uncorrected variables that serve as an input for discriminant analysis (Miller, Miller-Butterworth, Diefenbach, & Walter, 2020). We conducted DAPC analyses using two approaches. First, we used sampling locations as a priori groups. Then, we used the function find.cluster() to determine the number of clusters (K) de novo, with optimal K selected as that with the lowest Bayesian Information Criterion (BIC) value. In the a priori approach, the optimal number of PCs retained for the DAPC was determined using cross-validation and in the de novo approach a-score optimization was used. These methods were applied to carefully assess population structure and allow repeatability and transparency(Miller et al., 2020).

We used spatial principal component analysis (sPCA) that uses geographic information(GPS coordinates) to investigate the spatial pattern of genetic variability using allelic frequency data of individuals or populations (Jombart et al., 2008). We applied sPCA to understand how geographical and environmental aspect's structure genetic information (Jombart et al. 2008). The sPCA detects both global and local structures from strong genetic similarity or positive autocorrelation between neighbors. Individuals that are located beyond 30km were not considered

as neighbors. Global and local Monte Carlo tests was carried out with 10,000 permutations to evaluate the significance of detected global and local patterns (Jombart et al. 2008). We used 30km because it is an average distance known for the male giraffes to disperse between different groups (Ferres et al., 2021).

Results

We collected DNA samples from six subpopulations of Masai giraffes in the Serengeti and Tarangire Manyara ecosystems which lie either west or east of the Manyara and Eyasi escarpments of the GRV (**Fig. 2**). To assess mtDNA variation an 1140nt fragment was sequenced for 320 individuals and the entire 16,422bp mtDNA genome was sequenced for 101 individuals. From the 101 mtDNA WGS we identified 54 unique haplotypes in one or more of the six populations (Fig. 2) which were inclusive of all the 13 unique haplotypes found among the 320 individuals sequenced for only the 1140nt mtDNA fragment. We identified 14 haplotype clades among the 54 unique whole genome mtDNA haplotypes with subclade members of a group differing by no more than 3bp from each other. The 14 mtDNA haplotype clades exhibit an extreme geographic sorting with 13 of the groups found exclusively on one side or the other of the Manyara and Eyasi Escarpments of the GRV. WRV1 is the dominant clade in the Serengeti Ecosystem West Rift populations and is also found in several individuals in LMNP, but was not found in any other East Rift populations. The other three Tarangire Manyara Ecosystem populations shared no haplotypes with the Serengeti Ecosystem populations. This remarkable divergence has occurred despite the fact these populations across the GRV escarpments are relatively near each other (i.e. between 72-214 km). Even more surprising is the radical difference between the WRV and ERV mtDNA clades

with over 100nt differences separating them (Fig. 2). To ascertain the age and origin of these two major haplogroups, we compared them to WGS mtDNA of other giraffe (sub)species and to *Okapai johnstoni*. We found that the WRV and ERV haplogroups are equally related to other giraffe (sub)species and to okapi, suggesting that these two haplogroups diverged from a common ancestor after speciation. Moreover, WRV and ERV haplogroups are as similar to South African giraffe haplotypes as they are to each other and only marginally more similar to each other than they are to the other giraffe (sub)species. Masai and South African giraffes diverged from a common ancestor ~250 kya (Brown, Brenneman, Koepfli, Pollinger, Milá, et al., 2007; Coimbra et al., 2021). Together these data suggest that the WRV and ERV haplogroups diverged from a common ancestor shortly after Masai and South African giraffes diverged from a common ancestor 250 kya.

The WRV1 haplotype clade found in 70% of the the West Rift samples and appears to be the ancestral WRV haplogroup with the other four WRV clades showing substantial divergence from it (Fig. 2). WRV1 is also the lone clade found among individuals in the East Rift LMNP population. Interestingly, there are three different WRV1 subclades found in LMNP which differ 1-3nt from each other, and two of them are unique and not found in West Rift populations (Fig. 2; Table S1). This suggest that a few females, perhaps only one, bearing a WRV1 haplotype migrated to LMNP from a West Rift population a relatively long time ago and subsequently this WRV1 haplotype mutated to give rise to the unique WRV1 subclades found only in LMNP. The nine ERV haplotype clades exhibit a more star-like phylogenetic relationship with nearly equal genetic distances between them. Previous studies obtained a limited number of mtDNA sequences of Masai giraffes from Tanzania, Kenya, and Zambia (Agaba et al., 2016; Brown, Brenneman, Koepfli, Pollinger,

Mila, et al., 2007; Coimbra et al., 2021). To compare our results, we focused on a 652 fragment that had been sequenced in all studies (Fig. S1; Table S3). Five additional unique haplotypes, one in the Serengeti (Lobo) and four in East Rift populations in southern Kenya were found in the samples from Brown and coworkers (2007). Four of these unique haplotypes differed by only 1 or 2 nt from other haplotypes, but one haplotype—found only in the Athi River Ranch—exhibited 22 nt difference from next most similar haplotypes (Fig. S1). A small number of individuals in Selous Game Reserve in southeastern Tanzania and Luangwa Valley National Park in Zambia (Coimbra et al., 2021) exhibit haplotypes that are identical or nearly so to one of the 13 mtDNA haplotype clades described herein, and sort east and west of the GRV as expected. The first giraffe genome to be sequenced (Agaba et al., 2016) was denoted as MA1 originating from the Masai Mara in Kenya. The WGS mtDNA sequence of MA1 has one nucleotide difference with WRV1, the most frequent haplotypes in the west of GRV escarpments. Importantly, these additional samples demonstrate the same east-west segregation of the major haplotypes (Table S4). Except for LMNP no haplotypes were shared between populations east and west of the GRV.

To assess nuclear genetic variations, 23 microsatellites dispersed across the nuclear genome were sequenced for 95 individuals among five populations including SNP and NCA in the west and TNP, MRC, and BWMA in the east. An insufficient number of tissue samples were obtained from the LMNP to include in the analysis of microsatellites. Among the 23 loci 154 alleles were found to be present in in one or more of the five subpopulations, and of these 31 and 19 alleles were present in one and two populations, respectively (Table 2). Analysis of the rare alleles present only in two populations revealed a potential impact of the Manyara escarpment, with a significantly higher number of shared rare alleles between populations on the same side versus across the

Manyara escarpment (p=0.012). Moreover, the four cases of rare alleles shared between two populations across the escarpment involved MRC and either NCA or SNP. The mean number of alleles (Na), allelic richness (AR), observed heterozygosity, and the unbiased expected heterozygosity were very similar across all five populations and indicated a relatively high degree of genetic variation is being maintained in these populations (Table 2).

Population differentiation and structure

To examine population differentiation F_{ST} were estimated for mtDNA haplotypes and microsatellite data for all pairwise populations. F_{ST} values for mtDNA showed relatively low values between populations on the same side of the GRV compared to pairwise comparisons of populations across the GRV, as expected given the very large differences between the WRV and ERV haplotype clades. LMNP was found to be an exception to this general trend by exhibiting intermediate F_{ST} values in pairwise comparisons with west and east GRV populations, as was also expected given that it is the only population that contains both ERV and WRV haplotype clades. For microsatellites all pairwise F_{ST} values between populations west and east of the Rift showed significant genetic differentiation (Fig. 3a). Within ecosystems, low F_{ST} values were observed, and, in some cases, they were not statistically significant (e.g. MRC -TNP F_{ST} =0.0019).

F_{ST} values are known to generally increase as a function of geographic distance, and we found this to be the case for the mtDNA and STR using the Mantel isolation by distance test (Fig. S3). To determine if the apparent F_{ST} differences on each side of the GRV versus across the GRV were simply due to geographic distance, we normalized F_{ST} by geographic distance (F_{ST}/Km) and

compared same-side Rift populations with cross-Rift populations. We found the average normalized F_{ST}/Km for microsatellites was 1.8 higher for cross-Rift versus same-side Rift but did not reach statistical significance (p=0.0845). For mtDNA the F_{ST}/km was 9.3 higher cross-Rift than same-side Rift which was highly statistically significant (Table 1).

To assess the population genetic structure, we performed principal coordinate analysis (PCoA) on the mtDNA and microsatellite data. PCoA showed a clear separation of east and west GRV populations, except for the mtDNA of LMNP which clusters somewhat closer to the SNP and NCA. (Fig**.** 3c &d). Population structure based only on the microsatellites was further examined by three additional clustering methods: STRUCTURE, Discriminant Analysis of Principle Components (DAPC), and spatial PCA (sPCA). STRUCTURE revealed two distinct clusters: the Serengeti Ecosystem forming one cluster and the Tarangire Manyara Ecosystem forming the second cluster (Fig. 3a) with a high delta K value of 290. MRC and LMNP exhibited a small proportion of membership coefficient from the first cluster (orange) suggesting evidence of gene flow across the Rift. Similarly, Serengeti giraffes showed a small proportion of cluster 2 (blue) (Fig. 4). DAPC showed three significant genetic clusters: the SNP, NCA, and east GRV populations (Fig. 5). The division of the SNP and NCA is likely due to the relatively large distance between them compared to the East GRV populations (Fig. S3). The sPCA analysis, which takes into account geographic location, shows only two clusters (Fig. 6), further supporting this interpretation. Our results show two main genetic clusters, one in the Serengeti Ecosystem and the second cluster in the Tarangire Manyara Ecosystem (Fig. 6b**).**

Discussion

Masai giraffe populations have plummeted in the past 30 years as the result of human activities and are now dispersed into highly fragmented subpopulations with reduced opportunity for migration between them (Bolger et al., 2019). Overlying the impact of human activities, geographic barriers constrain migration with the escarpment faults of the Gregory Rift Valley system as the major impediment to migration of terrestrial animals in Tanzania and Kenya, where almost all the remaining Masai giraffe exist. Based on the whole genome mtDNA sequence data, a proxy for female-mediated gene flow, we show that female migration has likely not occurred across the Manyara escarpment of the GRV near the time that the Masai giraffe arose as a distinct (sub)species 250kya, except for LMNP. No mtDNA haplotypes are shared between the major populations of West Rift Valley (SNP and NCA), and the East Rift Valley (TNP, BWMA, and MRC). In addition, more than 100nt differences separate the WRV and ERV haplotypes, and they are no more closely related to each other than they are to South African giraffe mtDNA haplotypes. Our analysis of mtDNA haplotypes of Masai giraffes from four other studies (Agaba et al., 2016; Brown, Brenneman, Koepfli, Pollinger, Milá, et al., 2007; Coimbra et al., 2021) from Kenya, Zambia, and southern Tanzania confirm the radical separation of mtDNA haplotypes east and west of the escarpments of the GRV. Given that the Masai giraffe (*G. c. tippelskirchi*) and South African giraffe (*G. c. giraffa*) arose from a common ancestor of a southern African clade approximately 250 kya (Brown et al., 2007; Coimbra et al., 2021), we conclude that the WRV and ERV Masai giraffe mtDNA clades diverged near the time of the speciation event. The existence of mtDNA clades within a species that predate speciation is not uncommon. For example, a major mtDNA haplotype of the African savannah elephant (*Loxodonta africana*) apparently was introgressed from the forest elephant (*Loxodonta cyclotis*), and the forest elephant mtDNA is even more highly diverged from other savannah elephant mtDNA haplotypes (Ishida, Georgiadis, Hondo, & Roca,

414 2013; Ishida et al., 2011) than the Masai giraffe West Rift Valley and East Rift Valley haplotypes.
415 However, the forest elephant mtDNA haplotype is present in populations east and west of the GRV
416 (Ahlering et al., 2012; Lohay et al., 2020), albeit with higher frequencies west of the GRV nearer
417 the existing forest elephant populations in the West Africa and the Congo basin. That the forest
418 elephant mtDNA haplotype is present in major populations east of the GRV whereas the WRV
419 Masai giraffe haplotype is not, is likely due to the relative mobility of these two animals in
420 traversing mountainous terrain. Savannah elephants reportedly traverse the Manyara escarpment
421 through the Kitete-Selela corridor, which is the only known corridor directly across the Manyara
422 escarpment (Douglas-Hamilton, 1973; Prins & Jong, 2022). However, giraffe are not know to use
423 this corridor (Jones et al 2009; and personal communication). The combination of giraffe's high
424 anterior center of gravity and elevated forelegs and neck (Mitchell, 2021) makes climbing difficult
425 as can be seen in video recordings of giraffe attempting to climb modest inclines. In contrast to the
426 radical mtDNA differentiation across the GRV, evidence from nuclear genome microsatellite
427 genetic variation suggest that male migration has occurred throughout most of Masai giraffe
428 history. Male Masai giraffes range more widely than females (Knüsel et al., 2019; Lee & Bolger,
429 2017) and are thought to be more important in mediating gene flow between distant populations.
430
431 Population structure analysis, DAPC, and sPCA of nuclear genetic variation all show that WRV
432 and ERV populations are significantly differentiated. Presently these spatial differences in nuclear
433 genetic variation across the GRV are not large enough to conclude that male-mediated gene flow
434 across the GRV has ceased. Our analysis of potential migratory routes across the Manyara
435 Escarpment immediately separating these populations suggests that male giraffe migration has
436 stopped sometime in the past few decades. Circumnavigating the Manyara Escarpment may be

possible by first traveling north via the Manyara Ranch – Natron wildlife corridor (Caro, Jones, & Davenport, 2009) that runs parallel to the escarpment and then crossing over the escarpment near the village of Engare Sero and then down through Salei plains to the western NCA and SNP populations. This circuitous route (~175 km) appears to be the only potential migratory route that is not blocked by human activities. In support of this northern migratory route, we found that that the MRC, which is the most northern ERV population, was the only ERV population that shared rare alleles with WRV populations. The distance of this the northern route is well beyond the normal male home range of 157 km^2 (Knüsel et al., 2019), therefore we speculate that this route would be used rarely but could account for a small amount of gene flow. Migration via a more southerly route is highly unlikely inasmuch as an additional fault, known as the Eyasi escarpment, splits off the main fault system near the Ngorongoro Crater thus adding an additional barrier to potential southern migratory routes between the Tarangire Manyara Ecosystem and Serengeti Ecosystem.

A high degree of genetic connectivity occurs between the WRV populations in the SNP and NCA as shown by very low F_{ST} values. Comparatively these populations are far apart but are within highly protected contiguous regions with no geographic or human activities/settlements existing between them. The ERV populations, by contrast, are adjacent to high density human settlements, agricultural and pastoral lands, and tarmac roads that have expanded rapidly in the past two decades (Newmark, 1996; Newmark & McNeally, 2018). The MRC, TNP, and BWMA populations still show strong genetic connectivity despite the recent expansion of human activities. However, the Masai giraffe population in LMNP is very small, perhaps less than 100 individuals (Lee & Bond, 2022) , and is significantly differentiated from nearby populations in MRC, TNP,

and BWMA. LMNP is comprised of a narrow strip of land bounded by the steep Manyara escarpment and Lake Manyara on the west and east side, respectively and Mto wa Mbu and Magara village lands at the north and south ends, respectively. The previously identified wildlife corridors (Caro et al., 2009; Riggio & Caro, 2017; Riggio et al., 2022) now appear to be completely blocked by human activities. Unexpectedly, approximately half the sampled LMNP giraffe carry a common WRV mtDNA haplotype clade (WRV1) with the other half showing ERV haplotypes. Among the LMNP giraffe with the WRV1 haplotype, 3 distinct subclades are seen differing by only 1 or 2 nt and these unique WRV1 subclades are not seen in SNP and NCA populations. This suggests that a WRV1 female migrated from a WRV population to the LMNP some time ago and subsequently new mutations have occurred in her descendants giving rise to these unique subclades. A previous study using microsatellites showed that LMNP giraffes are equally closely related to the SNP and TNP (D. M. Brown, Brenneman, Koepfli, Pollinger, Mila, et al., 2007). Together, mtDNA and microsatellite data in LMNP show that LMNP is unique in displaying both ERV and WRV characteristics. Rather than being a source population for spreading genetic variation across Manyara escarpment, it appears that LMNP is a sink population for a genetic variation coming from the WRV and ERV populations. Gene flow from the WRV populations to LMNP likely ceased many years ago given that LMNP appears to be nearly isolated and that there is no evidence from our data that the LMNP has ever served as a conduit of gene flow from the WRV populations to the other ERV populations east of the Manyara escarpment.

Our study underscores the dire prospects of long-term survival of Masai giraffes. Previous studies had shown that the overall population of Masai giraffes had substantially declined and that populations were highly fragmented, leading the IUCN to declare the Masai giraffe an endangered

subspecies. We show that the escarpments of the Gregory Rift Valley system superimpose a further geographic subdivision of Masai giraffe into two metapopulations of ERV and WRV, located east and west of the GRV escarpment and which are comprised of <21,000 and <15,000 individuals, respectively. It is premature to claim that these two populations are distinct subspecies. However, they are very likely to be completely reproductively isolated and as a consequence their future persistence is independent of each other. These two major populations are further fragmented into smaller subpopulations, some that are small enough to experience inbreeding depression. A recent report was published by the Giraffe Conservation Foundation (GCF) in a non-refereed publication claiming that Masai giraffe populations had increased by 44% in just 5 years, and this report has been widely distributed in popular media outlets (M. B. Brown et al., 2022). However, this report does not provide the source of the primary data and sampling methods used, and an increase of this magnitude appears to be biologically implausible. Extensive demographic studies conducted by two of us (DEL and MLB) in northern Tanzanian populations of Masai giraffe also do not suggest a significant increase in population numbers during the past ten years (Lee, Lohay, Cavener, & Bond, 2022).

Acknowledgements

Funding for this project was provided by the Department of Biology, Eberly College of Science, and Huck Institutes of the Life Sciences of Penn State University, and the Wild Nature Institute. We thank the Tanzania Wildlife Research Institute, Tanzania National Park Authority, Ngorongoro Conservation Authority, Tanzania Wildlife Authority, and Commission of Science and Technology for providing research permits, the Nelson Mandela African Institutions of Science and Technology, Lab technician Emmanuel Lyimo, research assistants James Madeli,

506 Philip Ferdinand, and Emmanuel Kimaro. Our special gratitude to Rebecca Bourne a member of
507 Cavener lab for her technical support.

508 References

509
510

511 Abdul-Muneer, P. M. (2014). Application of microsatellite markers in conservation genetics and
512 fisheries management: recent advances in population structure analysis and conservation
513 strategies. *Genetics Research International*, *2014*(691759), 1--11. doi: 10.1155/2014/691759

514 Agaba, M., Ishengoma, E., Miller, W. C., McGrath, B. C., Hudson, C. N., Reina, O. C. B., …
515 Cavener, D. R. (2016). Giraffe genome sequence reveals clues to its unique morphology and
516 physiology. *Nature Communications*, *7*(1), 11519. doi: 10.1038/ncomms11519

517 Ahlering, M. A., Eggert, L. S., Western, D., Estes, A., Munishi, L., Fleischer, R., … Maldonado,
518 J. E. (2012). Identifying Source Populations and Genetic Structure for Savannah Elephants in
519 Human-Dominated Landscapes and Protected Areas in the Kenya-Tanzania Borderlands.
520 *PLoS ONE*, *7*(12), e52288. doi: 10.1371/journal.pone.0052288

521 Allendorf, F. W. (2017). Genetics and the conservation of natural populations: allozymes to
522 genomes. *Molecular Ecology*, *26*(2), 420--430. doi: 10.1111/mec.13948

523 Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of
524 conservation genetics. *Nature Publishing Group*, *11*(10), 697--709. doi: 10.1038/nrg2844

525 Baker, B. H., Mohr, P. A., & Williams, L. A. J. (1972). Geology of the Eastern Rift System of
526 Africa. In ["B. H. Baker", "P. A. Mohr", & "L. A. J. Williams"] (Eds.), *Geology of the*
527 *Eastern Rift System of Africa* (Vol. 136, p. 0). Geological Society of America. doi:
528 10.1130/spe136-p1

529 Barbian, H. J., Connell, A. J., Avitto, A. N., Russell, R. M., Smith, A. G., Gundlapally, M. S., …
530 Hahn, B. H. (2018). CHIIMP: An automated high-throughput microsatellite genotyping
531 platform reveals greater allelic diversity in wild chimpanzees. *International Journal of*
532 *Business Innovation and Research*, *17*(3), 7946--7963. doi: 10.1002/ece3.4302

533 Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings*
534 *of the National Academy of Sciences*, *115*(25), 6506–6511. doi: 10.1073/pnas.1711842115

535 Bluwstein, J. (2018). *Biopolitical Landscapes Governing People and Spaces through*
536 *Conservation in Tanzania*. (January).

537 Bock, F., Fennessy, J., Bidon, T., Tutchings, A., Marais, A., Deacon, F., & Janke, A. (2014).
538 Mitochondrial sequences reveal a clear separation between Angolan and South African giraffe
539 along a cryptic rift valley. *BMC Evolutionary Biology*, *14*(1), 219. doi: 10.1186/s12862-014-
540 0219-7

Bohonak, A. J. (2002). IBD (Isolation by Distance): A Program for Analyses of Isolation by Distance. *The Journal of Heredity*, (93), 153--154.

Bolger, D., Ogutu, J., Strauss, M., Lee, D., Muneza, A., Fennessy, J., & Brown, D. (2019). Giraffa camelopardalis ssp. tippelskirchi, Masai Giraffe Assessment. *The IUCN Red List of Threatened Species 2019*, *8235*. doi: 10.2305/iucn.uk.2019-1.rlts.t88421036a88421121.en

Bond, M. L., Kiffner, C., & Lee, D. E. (2022). Tarangire: Human-Wildlife Coexistence in a Fragmented Ecosystem. *Ecological Studies*, 163–188. doi: 10.1007/978-3-030-93604-4_8

Bond, M. L., König, B., Ozgul, A., Farine, D. R., & Lee, D. E. (2021). Socially Defined Subpopulations Reveal Demographic Variation in a Giraffe Metapopulation. *The Journal of Wildlife Management*. doi: 10.1002/jwmg.22044

Borner, M. (1985). The increasing isolation of Tarangire National Park. *Oryx*, *19*(2), 91–96. doi: 10.1017/s0030605300019797

Brown, D. M., Brenneman, R. A., Koepfli, K.-P., Pollinger, J. P., Mila, B., Georgiadis, N. J., … Wayne, R. K. (2007). Extensive population genetic structure in the giraffe. *BMC Biology*, *5*(1), 57. doi: 10.1186/1741-7007-5-57

Brown, D. M., Brenneman, R. A., Koepfli, K.-P., Pollinger, J. P., Milá, B., Georgiadis, N. J., … Wayne, R. K. (2007). Extensive population genetic structure in the giraffe. *BMC Biology*, *5*(1), 57. doi: 10.1186/1741-7007-5-57

Brown, M. B., Kulkarni, T., Ferguson, S., Fennessy, S., Muneza, A., Stabach, J. A., & Fennessy, J. (2022). *Imperiled: The Encyclopedia of Conservation*. 471–487. doi: 10.1016/b978-0-12-821139-7.00139-2

Caro, T., Jones, T., & Davenport, T. R. B. (2009). Realities of documenting wildlife corridors in tropical countries. *Biological Conservation*, *142*(11), 2807–2811. doi: 10.1016/j.biocon.2009.06.011

Carter, K. D., Seddon, J. M., Carter, J. K., Goldizen, A. W., & Hereward, J. P. (2012). Development of 11 microsatellite markers for Giraffa camelopardalis through 454 pyrosequencing, with primer options for an additional 458 microsatellites. *Conservation Genetics Resources*, *4*(4), 943–945. doi: 10.1007/s12686-012-9679-5

Coimbra, R. T. F., Winter, S., Kumar, V., Koepfli, K.-P., Gooley, R. M., Dobrynin, P., … Janke, A. (2021). Whole-genome analysis of giraffe supports four distinct species. *Current Biology*. doi: 10.1016/j.cub.2021.04.033

Crowhurst, R. S., Mullins, T. D., Mutayoba, B. M., & Epps, C. W. (2013). Characterization of eight polymorphic loci for Maasai giraffe (Giraffa camelopardalis tippelskirchi) using non-invasive genetic samples. *Conservation Genetics Resources*, *5*(1), 85--87. doi: 10.1007/s12686-012-9739-x

576 Douglas-Hamilton, I. (1973). Short contributions. *African Journal of Ecology*, *11*(3-4), 401–403.
577 doi: 10.1111/j.1365-2028.1973.tb00101.x

578 Estes, R. D. (2014). *The Gnu's World: Serengeti Wildebeest and Life History*. University of
579 California Press.

580 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals
581 using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611--
582 2620. doi: 10.1111/j.1365-294x.2005.02553.x

583 EXCOFFIER, L., & LISCHER, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs
584 to perform population genetics analyses under Linux and Windows. *Molecular Ecology*
585 *Resources*, *10*(3), 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

586 Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using
587 multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*(4),
588 1567--1587. doi: 10.1111/j.1471-8286.2007.01758.x

589 Ferres, J. M. L., Lee, D. E., Nasir, M., Chen, Y.-C., Bijral, A. S., Bercovitch, F. B., & Bond, M.
590 L. (2021). Social connectedness and movements among communities of giraffes vary by sex
591 and age class. *Animal Behaviour*, *180*, 315–328. doi: 10.1016/j.anbehav.2021.08.008

592 Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population
593 structure. *Molecular Ecology Resources*, *17*(1), 27--32. doi: 10.1111/1755-0998.12509

594 Goudet, J. (1995). FSTAT (version 1.2): A computer Program to calculate F-Statistics. *Journal*
595 *of Heredity*, *86*(6), 485--86.

596 Huebinger, R. M., Pierson, D. J., Maar, T. W. D., Borwn, D. M., Brenneman, R. A., & Louis, E.
597 E. (2002). Characterization of 16 microsatellite marker loci in the Maasai giraffe (Giraffa
598 camelopardalis tippelskirchi). *Molecular Ecology Notes*. doi: doi: 10.1046/j.1471-8278
599 .2002.00307.x

600 Ishida, Y., Georgiadis, N. J., Hondo, T., & Roca, A. L. (2013). Triangulating the provenance of
601 African elephants using mitochondrial DNA. *Evolutionary Applications*, *6*(2), 253--265. doi:
602 10.1111/j.1752-4571.2012.00286.x

603 Ishida, Y., Oleksyk, T. K., Georgiadis, N. J., David, V. A., Zhao, K., Stephens, R. M., … Roca,
604 A. L. (2011). Reconciling Apparent Conflicts between Mitochondrial and Nuclear
605 Phylogenies in African Elephants. *PLoS ONE*, *6*(6), e20642. doi:
606 10.1371/journal.pone.0020642

607 Jombart, T, Devillard, S., Dufour, A.-B., & Pontier, D. (2008). Revealing cryptic spatial patterns
608 in genetic variability by a new multivariate method. *Heredity*, *101*(1), 92–103. doi:
609 10.1038/hdy.2008.34

Jombart, Thibaut. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, Thibaut, Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94–94. doi: 10.1186/1471-2156-11-94

Kiffner, C., Thomas, S., Speaker, T., O'Connor, V., Schwarz, P., Kioko, J., & Kissui, B. (2020). Community-based wildlife management area supports similar mammal species richness and densities compared to a national park. *Ecology and Evolution*, *10*(1), 480–492. doi: 10.1002/ece3.5916

Kikoti, A. P., & Griffin, C. R. (2009). Seasonal home range sizes, transboundary movements and conservation of elephants in northern Tanzania. *Wildlife Fisheries Conservation*, *PhD*(Open Access Dissertations. Paper 108.). Retrieved from http://scholarworks.umass.edu/openaccessdissertations/108

Knüsel, M. A., Lee, D. E., König, B., & Bond, M. L. (2019). Correlates of home range sizes of giraffes, Giraffa camelopardalis. *Animal Behaviour*, *149*, 143–151. doi: 10.1016/j.anbehav.2019.01.017

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179--1191. doi: 10.1111/1755-0998.12387

Lamprey, H. F. (1964). Estimation of the Large Mammal Densities, Biomass and Energy Exchange in the Tarangire Game Reserve and the Masai Steppe in Tanganyika. *African Journal of Ecology*, *2*(1), 1--46. doi: 10.1111/j.1365-2028.1964.tb00194.x

Lee, D. E. (2018). Evaluating conservation effectiveness in a Tanzanian community wildlife management area. *The Journal of Wildlife Management*, *82*(8), 1767--1774. doi: 10.1002/jwmg.21549

Lee, D. E., & Bolger, D. T. (2017). Movements and source–sink dynamics of a Masai giraffe metapopulation. *Population Ecology*, *59*(2), 157–168. doi: 10.1007/s10144-017-0580-7

Lee, D. E., & Bond, M. L. (2022). Tarangire: Human-Wildlife Coexistence in a Fragmented Ecosystem. *Ecological Studies*, 189–207. doi: 10.1007/978-3-030-93604-4_9

Lee, D. E., Lohay, G. G., Cavener, D. R., & Bond, M. L. (2022). Using spot pattern recognition to examine population biology, evolutionary ecology, sociality, and movements of giraffes: a 70-year retrospective. *Mammalian Biology*, 1–17. doi: 10.1007/s42991-022-00261-3

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110--1116. doi: 10.1111/2041-210x.12410

Lohay, G. G., Riggio, J., Lobora, A. L., Kissui, B. M., & Morrison, T. A. (2022). Tarangire: Human-Wildlife Coexistence in a Fragmented Ecosystem. *Ecological Studies*, 255–276. doi: 10.1007/978-3-030-93604-4_12

Lohay, G. G., Weathers, T. C., Estes, A. B., McGrath, B. C., & Cavener, D. R. (2020). Genetic connectivity and population structure of African savanna elephants (Loxodonta africana) in Tanzania. *Ecology and Evolution*, *10*(20), 11069–11089. doi: 10.1002/ece3.6728

Masao, C. A., Makoba, R., & Sosovele, H. (2015). Will Ngorongoro Conservation Area remain a world heritage site amidst increasing human footprint? *International Journal of Biodiversity and Conservation*, *7*(9), 394--407. doi: 10.5897/ijbc2015.0837

Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G., & Neufeld, J. D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, *13*(1), 31–31. doi: 10.1186/1471-2105-13-31

Miah, G., Rafii, M. Y., Ismail, M. R., Puteh, A. B., Rahim, H. A., Islam, N. K., & Latif, M. A. (2013). A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. *International Journal of Molecular Sciences*, *14*(11), 22499--22528. doi: 10.3390/ijms141122499

Miller, W. L., Miller-Butterworth, C. M., Diefenbach, D. R., & Walter, W. D. (2020). Assessment of spatial genetic structure to identify populations at risk for infection of an emerging epizootic disease. *Ecology and Evolution*, *10*(9), 3977–3990. doi: 10.1002/ece3.6161

Mitchell, G. (2021). *How Giraffes Work*. N.Y: Oxford University Press,.

Morrison, D. B. T. (2014). Connectivity and bottlenecks in a migratory wildebeest Connochaetes taurinus population. *Fauna Flora International, Oryx*, *48*(4), 613--621, doi:10.1017/S0030605313000537. doi: 10.1017/s0030605313000537

Morrison, T. A., Link, W. A., Newmark, W. D., Foley, C. A. H., & Bolger, D. T. (2016). Tarangire revisited: Consequences of declining connectivity in a tropical ungulate population. *Biological Conservation*, *197*, 53--60. doi: 10.1016/j.biocon.2016.02.034

Msofe, N. K., Sheng, L., & Lyimo, J. (2019). Land Use Change Trends and Their Driving Forces in the Kilombero Valley Floodplain, Southeastern Tanzania. *Sustainability*, *11*(2), 505. doi: 10.3390/su11020505

Muller, Z., F, B., R, B., D, B., M, B., D, B., … T, W. (2018). *Giraffa camelopardalis (amended version of 2016 assessment). The IUCN Red List of Threatened Species 2018: e.T9194A136266699, Gland, Switzerland*. Retrieved from e.T9194A136266699

678 Mwalyosi, R. B. B. (1991). Ecological evaluation for wildlife corridors and buffer zones for
679 Lake Manyara National Park, Tanzania, and its immediate environment. *Biological*
680 *Conservation*, *57*(2), 171–186. doi: 10.1016/0006-3207(91)90137-x

681 Newmark, W. D. (1996). Insularization of Tanzanian Parks and the Local Extinction of Large
682 Mammals. *Conservation Biology*, *10*(6), 1549–1556. doi: 10.1046/j.1523-
683 1739.1996.10061549.x

684 Newmark, W. D., & McNeally, P. B. (2018). Impact of habitat fragmentation on the spatial
685 structure of the Eastern Arc forests in East Africa: implications for biodiversity conservation.
686 *Biodiversity and Conservation*, *27*(6), 1387–1402. doi: 10.1007/s10531-018-1498-x

687 OOSTERHOUT, C. V., HUTCHINSON, W. F., WILLS, D. P. M., & SHIPLEY, P. (2004).
688 micro-checker: software for identifying and correcting genotyping errors in microsatellite
689 data. *Molecular Ecology Notes*, *4*(3), 535--538. doi: 10.1111/j.1471-8286.2004.00684.x

690 Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population
691 genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537--2539.
692 doi: 10.1093/bioinformatics/bts460

693 Petzold, A., & Hassanin, A. (2020). A comparative approach for species delimitation based on
694 multiple methods of multi-locus DNA sequence analysis: A case study of the genus Giraffa
695 (Mammalia, Cetartiodactyla). *PLoS ONE*, *15*(2), e0217956. doi:
696 10.1371/journal.pone.0217956

697 Prins, H. H. T., & Jong, J. F. de. (2022). Tarangire: Human-Wildlife Coexistence in a
698 Fragmented Ecosystem. *Ecological Studies*, 129–161. doi: 10.1007/978-3-030-93604-4_7

699 Pritchard, J. K., 1, M. S. and P. D. G. J., & Donnelly, P. (2000). Inference of Population
700 Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959.

701 Raymond, M., & Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for
702 exact tests and ecumenicism. *Journal of Heredity*, *86*, 248--249.

703 Riggio, J., & Caro, T. (2017). Structural connectivity at a national scale: Wildlife corridors in
704 Tanzania. *PLOS ONE*, *12*(11), e0187407. doi: 10.1371/journal.pone.0187407

705 Riggio, J., Foreman, K., Freedman, E., Gottlieb, B., Hendler, D., Radomille, D., … Kiffner, C.
706 (2022). Predicting wildlife corridors for multiple species in an East African ungulate
707 community. *PLoS ONE*, *17*(4), e0265136. doi: 10.1371/journal.pone.0265136

708 Ritchie, H., & Roser, M. (2021). Biodiversity". Published online at OurWorldInData.org.
709 Retrieved from https://ourworldindata.org/mammals#wild-mammals-make-up-only-a-few-
710 percent-of-the-world-s-mammals

Rousset, François. (1997). Genetic Differentiation and Estimation of Gene Flow from F - Statistics Under Isolation by Distance. *Genetics*, *145*(4), 1219–1228. doi: 10.1093/genetics/145.4.1219

Rousset, Francois. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103--106. doi: 10.1111/j.1471-8286.2007.01931.x

Scoon, R. N. (2018). *Geology of National Parksof Central/Southern Kenyaand Northern Tanzania*. doi: 10.1007/978-3-319-73785-0_14

Strauss, M. K. L., Kilewo, M., Rentsch, D., & Packer, C. (2015). Food supply and poaching limit giraffe abundance in the Serengeti. *Population Ecology*, *57*(3), 505–516. doi: 10.1007/s10144-015-0499-9

Data Accessibility Statement

Whole genome sequence data for mitochondrial has been submitted to GenBank (GenBank Submissions grp 8715258) but we have not got the accession number yet . Multilocus genotype data from microsatellites and all parameter settings for DAPC and sPCA, R-scripts, CHIIMP pipeline and input files are included in Dryad, https://doi.org/10.5061/dryad.m905qfv4h

Benefit-Sharing statement: Benefits Generated: A research collaboration was developed with scientists from the Tanzania providing genetic samples, all collaborators are included as co-authors and/or acknowledged for their contributions, and the research addresses a priority concern, in this case the conservation of Masai giraffe in Tanzania and more broadly giraffe East Africa. Our group is committed to international scientific partnerships, as well as institutional capacity building.

Authors Contributions

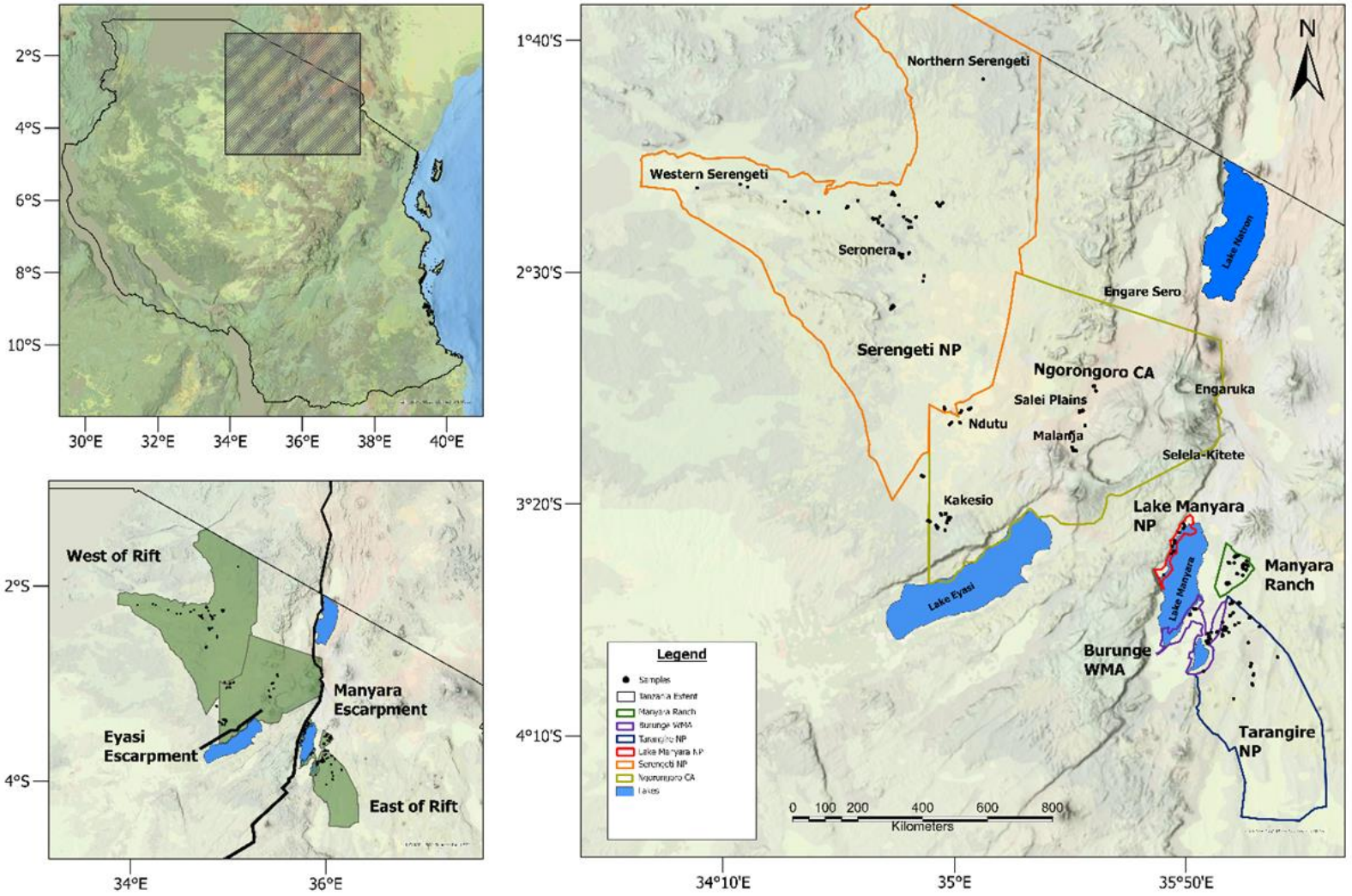
GGL and DRC conceived of the project. GGL collected the biological samples and obtain the permits for collecting and importing samples. GGL prepared genomic DNAs for sequencing. GGL, DRC, DP, and XH analyzed the data. LWC extracted and compiled the WGS mtDNA sequences.

737 XH assisted in the compiling the microsatellite genotypes. DRC, GGL, DEL, and MLB contributed
738 to writing the manuscript.

739 **Research Permits**

740 Permission to conduct this research was provided by the Commission of Science and Technology
741 #2020-185-NA-1990-172, Tanzania Wildlife Research Institute, Tanzania National Park
742 Authority, Ngorongoro Conservation Authority, Manyara Ranch Conservancy, and IACUC#
743 PROTO201901219 by Penn State University.

744 Tables and Figures



745
746 Fig. 1. A map showing areas where samples were collected in Serengeti and Tarangire Ecosystems in northern Tanzania (WMA= Wildlife Management Area,
747 NP=National Park, CA=Conservation Area).

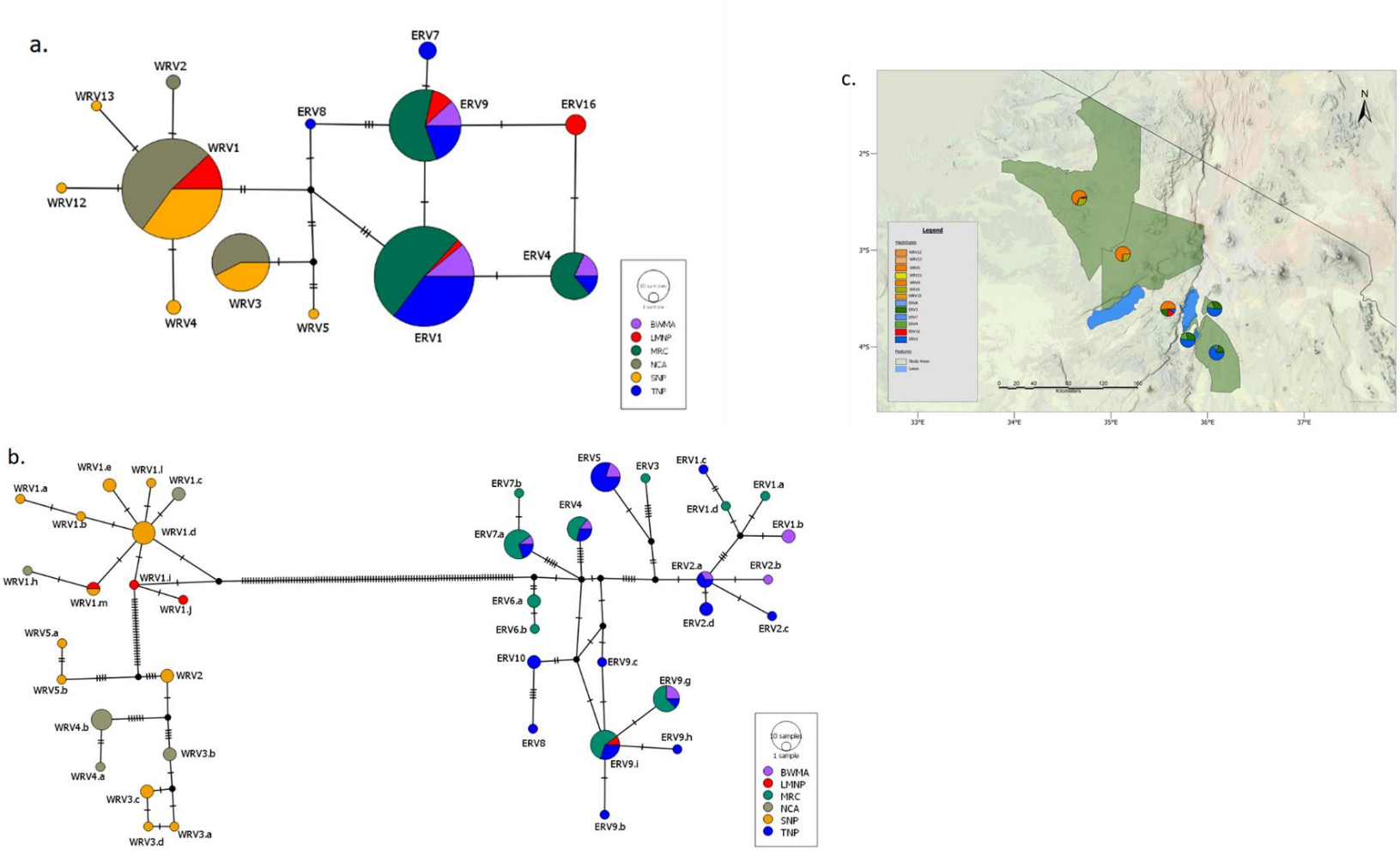


Fig. 2. Neighbor joining network showing genetic differentiation of the Masai giraffe east and west of the Rift using mtDNA data. Haplotype networks were created using (a) 1140bp mtDNA from 320 giraffes and (b) 16422bp from 101 giraffes. The size of circles corresponds to haplotype frequencies and hatch marks represent the number of mutations/nucleotide differences between haplotypes. c) a distribution of the 13 mtDNA haplotypes of Masai giraffes in northern Tanzania.

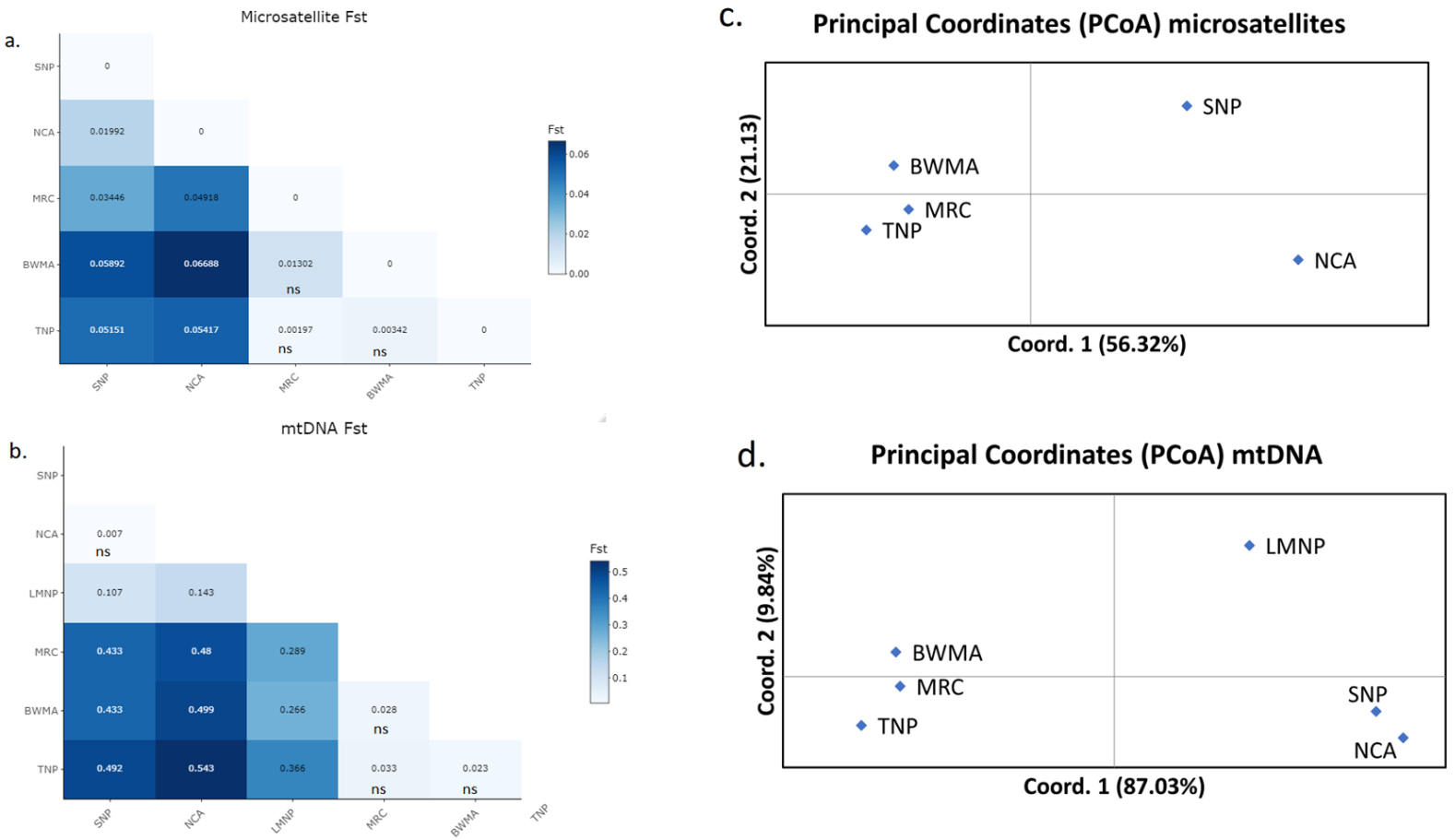


Fig. 3. Extent of genetic differentiation shown by pairwise F_{ST} and principal coordinate analysis (PCoA) based on the 23 microsatellites and mtDNA. (a &b) pairwise F_{ST} with heat map representing low and high values where the highest F_{ST} values were observed between giraffes west and east of the Rift. (c & d) principal coordinate analysis (PCoA) shows a genetic differentiation between giraffes found east and west of the Rift. For the mtDNA (d) LMNP was distinct from the rest of the TME (BWMA, MRC, TNP) because giraffes carried haplotypes found in both ecosystems.

762 Table 1. Normalized linearized F_{ST} to geographic distance (km). Average normalized F_{ST} /km was 1.8 higher across the Rift than within than across the Rift
763 (p=0.0845). Similarly, for the mtDNA the normalized F_{ST} /km was 9.3 higher across the Rift than within (p=0.0015).
764

Gregory Rift -point of reference	**Populations**	**Microsatellites Fst/km**	**mtDNA Fst/km**
Within	SNP-NCA	2.08E-04	7.27E-05
	MRC-BWMA	4.26E-04	9.29E-04
	MRC-TNP	4.11E-05	7.11E-04
	BWMA-TNP	1.32E-04	9.05E-04
	AVERAGE	2.02E-04	6.55E-04
Across	SNP-MRC	1.83E-04	3.92E-03
	SNP-BWMA	2.87E-04	3.52E-03
	SNP-TNP	2.25E-04	4.02E-03
	NCA-MRC	4.75E-04	8.47E-03
	NCA-BWMA	6.29E-04	8.74E-03
	NCA-TNP	3.87E-04	8.03E-03
	Average	3.64E-04	6.11E-03

765
766

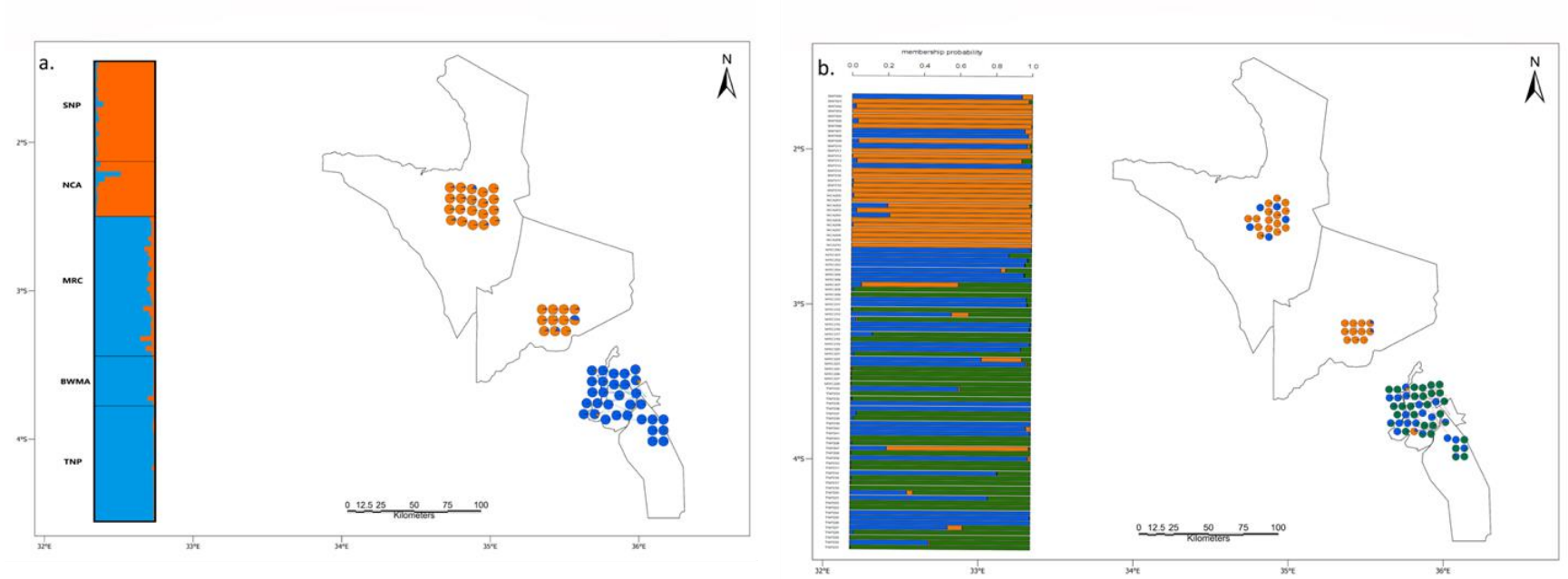


Fig. 4. Genetic subdivision of the Masai giraffe inferred from Bayesian clustering (STRUCTURE) and multivariate analysis (DAPC) based on microsatellite data. (a) Bar plot from STRUCTURE program, each vertical line represents an individual giraffe along with a map showing membership coefficients from STRUCTURE. The population is divided into two clusters (orange=Serengeti ecosystem, blue=Tarangire-Manyara ecosystem) according to the most informative K value using Evanno's method (K=2 Delta K= 150; Fig. S2). (b.) STRUCTURE-like results and membership coefficient inferred from DAPC show the presence of at least two genetic clusters. Both STRUCTURE and DAPC analysis reveal a clear genetic differentiation between giraffes found east and west of the Rift.

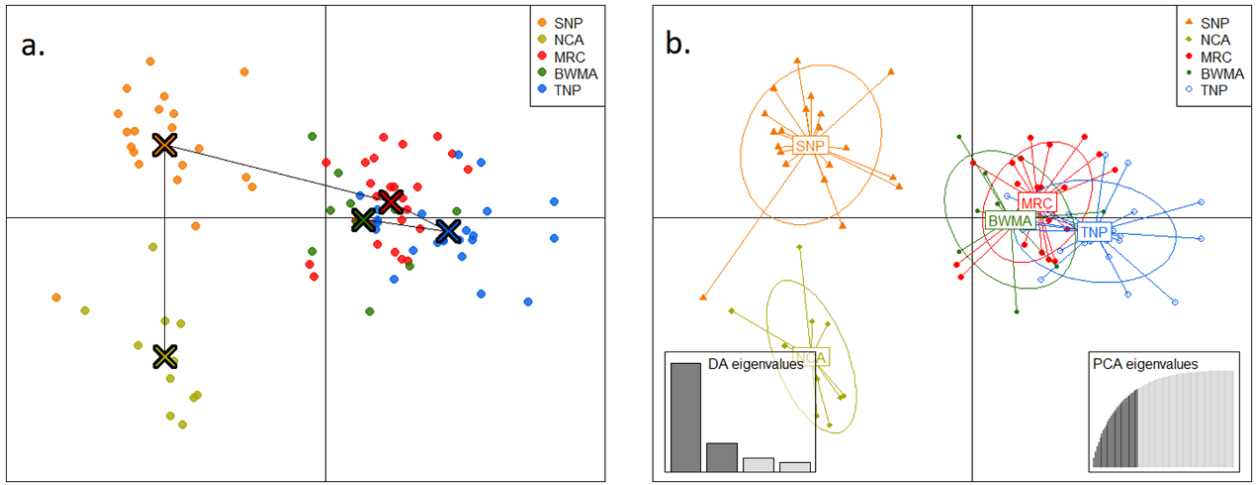


Fig. 5. Discriminant Analysis of Principal Components (DAPC) for microsatellite data of Masai Giraffe in Tanzania. (a) A minimum spanning tree based on the (squared) distances between populations which indicate the center of each group with crosses. (b) Scatterplots of DAPC of 23 microsatellite markers showing the first two principal components of the DAPC. A scatter plot from the DAPC using a priori approach (using sampling locations) suggested at least three clusters with SNP and NCA showing as different clusters. Within the TME, there was an overlap between population clusters. A minimum spanning tree based on the squared distance between populations indicate the NCA are most distantly related to the TME than SNP.

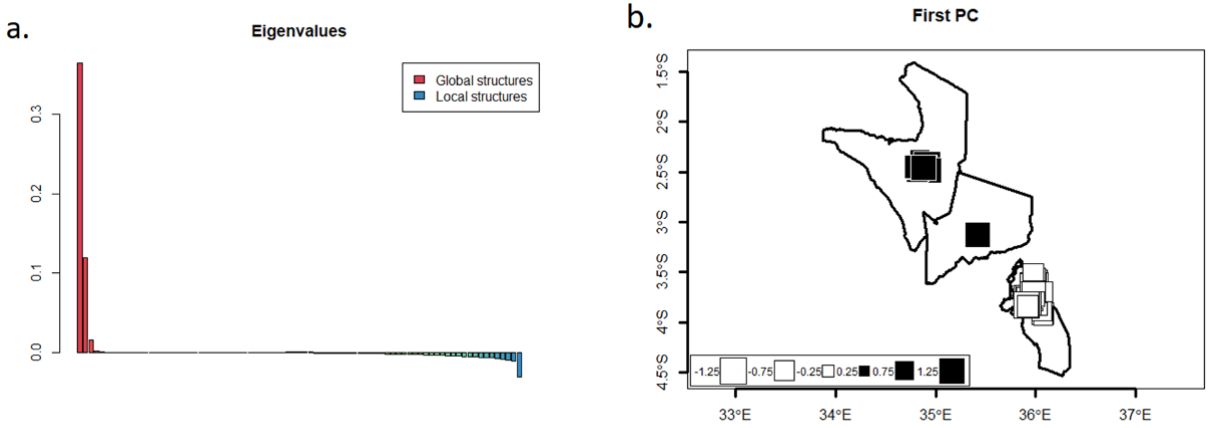


Fig. 6. Spatial Principal component analysis (sPCA) for Masai Giraffes using microsatellite markers . a) sPCA eigenvalues distribution confirming Global Structure. (b) scores of individuals along the first sPCA axis showing two subpopulations: black squares (Serengeti ecosystem) white squares (Tarangire ecosystem). The analysis of sPCA revealed cryptic spatial patterns using multi-locus genotype data. The first PCA was observed to have the highest eigen values meaning that the data could be best explained by the first PCA. We found the evidence for the presence of global Structure (p=0.001) and not local structure.

807 Table 2. Genetic parameters estimated at each locality for 23 microsatellites and mitochondrial DNA for the Masai giraffes

			Microsatellites						Mitochondrial DNA			
Location	Pop code	GPS coordinates	N	Na	A_R	P_A	Ho (SE)	µHE(SE)	N	Nh	Hd	π (%)
Burunge Wildlife Management Area	BWMA	3.959° S, 35.809° E	10	4.65 0.483	3.41 0.349	0.111 0.076	0.888 0.040	0.703 0.027	21	2	0.638	0.066
Lake Manyara National Park	LMNP	3.4459° S, 35.8093° E							23	4	0.654	0.292
Tarangire National Park	TNP	4.0057° S, 35.9788° E	23	5.55 0.633	3.32 0.365	0.611 0.244	0.820 0.055	0.657 0.036	51	5	0.495	0.066
Manyara Ranch Conservancy	MRC	3.5846° S, 36.0021° E	28	5.83 0.544	3.375 0.298	0.5 0.218	0.846 0.053	0.667 0.037	96	3	0.602	0.061
Ngorongoro Conservation Area	NCA	3.2279° S, 35.5075° E	11	4.55 0.364	3.10 0.279	0.166 0.090	0.818 0.067	0.645 0.049	74	4	0.464	0.157
Serengeti National Park	SNP	2.3333° S, 34.8333° E	20	5.61 0.506	3.61 0.363	0.277 0.109	0.842 0.055	0.687 0.035	55	7	0.537	0.107

808

809 N= Number of individuals with microsatellite data using tissue samples, N_A=Number of alleles, A_R=Allelic richness, P_A=Private
810 alleles, Ho= observed heterozygosity (SE=Standard Error), µHe=Unbiased expected heterozygosity, N= number of sequences from
811 fecal and tissue samples, N_H=Number of haplotypes, H_d=Haplotype diversity and π= haplotype diversity in %.
812

Supporting Information for Online Publication

Table S 1. Mitochondrial DNA haplotypes for Masai Giraffes recorded in this study. Haplotype names contains information of where the haplotype was collected whether West or East of Rift Valley (ERV= East of the Rift Valley, WRV= West of the Rift Valley). Nine haplotypes have been published (Brown, Brenneman, Koepfli, Pollinger, Mila, et al., 2007; Coimbra et al., 2021) and 11 new haplotypes found from this study.

Haplotypes	BWMA	TNP	MRC	LMNP	NCA	SNP	Total	GenBank Accession No.
ERV1	11	35	51	2			99	OL840825
ERV9	6	10	30	5			51	OL840826
ERV4	4	2	15				21	OL840827
WRV1				12	53	36	101	OL840828
WRV3					18	15	33	OL840829
ERV7		3					3	OL840830
ERV8		1					1	OL840831
WRV4						2	2	OL840832
WRV5						1	1	OL840833
WRV2					2		2	OL840834
WRV12						1	1	OL840835
WRV13						1	1	OL840836
ERV16				4				OL840839
Total	21	51	96	23	73	56	320	

823 *Table S 2. Microsatellite loci information for the 31 markers used in this study*

SSR names	*Abbreviation*	*Fragment size*	*Primer sequences Forward (5'-3')*	*Primer sequences Reverse (5'-3')*	*Repeat motif*	*Accession number*	*Reference*
Gica13905	A	230-236	CAGACAGATGGGGAAACTGAG	TTTGGCTAAATTTTTCATACACACA	(GT)7A(TG)17	JX424290	(Crowhurst et al. 2013)
Gica13619	B	257-261	CAG GTT TTC ATT GTA TTG CTC TG	ATGCAGAATGGGGGTTACAG	(TG)10	JX424289	
Gica9976	C	250-272	GGG AGG AGA CTG GAT TGT CA	AGT GGC TCT CCA AAG CAC AT	(GT)18	JX424293	
Gica10894	D	240-254	TGT TGT CAC TTA CCC GTT TTC C	AGA GTC TGG GAT GCA TTT GG	(GT)16	JX424294	
Gica9905	E	289-327	ATG ATA TTC AGC TGG GCC TCT	CCT GAT GGA CAC CAG GTT G	(TG)11 A(GT)2	JX424292	
Gica14170	F	261-269	GTG AGG TGC CAT CAC CTT CT	CAC TGG AGG CAA GTC AAC AA	(AT)3 (GT)18	JX424291	
Gica16160	G	133-167	TGC AGA GCA ATT GCA AAC AT	GTG GGC AAC TGT TCA TAG GG	(GT)24	JX424296	
Gica16120	H	134-140	AAA GTA ATT TGG GCA AAT GTG G	TTT GGC CAG TCT TCA GAT CA	(TG)6 A (GT)13	JX424295	
11HDZ484	I	~197	GCC TGG GGG AGC TAG AGT C	AAC TCA GAT TGC CTT GCC C	(CA)2(CG)2(CA)6	AY727871	(Brown et al. 2007)
11HDZ073	J		AGA CCT AAT GCC ACC AGA ATG	GAG GGT AGT GGA ACT GGG A	(GT)12	AY102406	(Huebinger et al. 2002)
11HDZ102	K	193-195	TGG AAT AGG AAA TGG CAA CC	GAT TGA AGG AAA CCA GAC ACG	(CA)10	AY102407	
11HDZ334	L	311-321	TTC ACT CAT TGT CCA TTT AGG G	TAG GCT GGC TTC TGC TGC	(TG)13	AY102408	
11HDZ443	M	129-147	CAT AAA ATT AAA AGG CAC TTG TTC C	ATG GGG GTC ACA AAG AGT CTG	(GT)17	AY102409	
11HDZ447	N	158-188	CTC AAC AGA CAG CTC AAT ACT AGA AC	AGT TCC TTC AAT AAG CCC ATA TC	(CA)7(AT)2CA(AT)5	AY102410	
11HDZ480	O	122-126	TGC TTT AGT AAA GTG TGT GAA ATG C	CAC AGA ATC TAC ACA CAT CAC ACA TC	(TG)15	AY102411	
11HDZ550	P	168-180	GGA CAG TGG ACT AGG AGA AAA GG	GCC TGG GAT TCC TGG TAA AC	GTCT(GT)14CT(GT)4	AY102412	
11HDZ561	Q	181-185	CAA CAA AGA CAA ACT GGA TAG C	TCT AAC ATC TGA GCC ACC G	(CA)3G(CA)18	AY102413	
11HDZ562	S	138-144	AAA GAG TTA GAT GCA ACT GAG TGA C	TCA GCA TCC TAT ATT TTC ACA CC	(CA)19	AY102414	
11HDZ567	T	182-212	GGT TTC AGA AGG TTT GTT GGC	TGC ATT ATC CCA AGT TCT TTA GC	(CA)13	AY102415	
11HDZ582	U	121-123	TTC CTA AGT TAC CCT CTC TGC C	TTA GCA CCA CCC CTC TCA AC	(CA)3(GT)6A(GT)6	AY102416	

824

11HDZ626	V	182-194	CAT TGG CAG GTG GAT TCT TTA C	AGC CCA ATT ATT CTT TTA CTT CCC	(GT)16	AY102417	
11HDZ665	W	194-228	GCC CCT TGC CTA GCT TAA C	CCG ACT GTA GAA ATG AAG CG	(CA)16	AY102418	
11HDZ748	X	241-245	TTT TGG AGA GGA TTG AAA TCT G	GAA TCA TCT GTG GCT AAG CAT C	(GT)14	AY102419	
11HDZ835	Y	201-220	CCC ACA CTG CAA CTA AAC CTG	AAG AAA CTC AAA AGC CTG CAA G	(CA)12	AY102420	
11HDZ1004	Z	142-158	CTC ATG TCT CTT GCA CTG GC	GTA ATG GCA TAT TTC ACT CTT TTT C	(CA)6(TA)2(CA)5TA(CA)6	AY102421	
Gca01	BA	85-89	GCATGCTACCAACACCTCTG	ACCAATCGAAGGACTGTTGC	(AC)13	JQ973777	(Carter et al. 2012)
Gca09	CA	126-132	GCATGCATCTTGAAAGGAAA GG	GGAGTCCCTTCCTGGTTCTG	(ACACC)9	JQ973778	
Gca14	EA	157-159	CAAGATGTGGAAGCAGCCTG	CCCTCTAGGTCCATTCGTATTG	(AC)8	JQ973780	
Gca16	FA	178-180	GCAACCTTCCCAGTTTCCAG	AAGACCCTGAGAGTGAGCAC	(AC)8	JQ973781	
Gca21	GA	237-257	GAGACACAGAACCAACAGG C	TCATACTTTGAGCATCCCAGC	(AGAT)11	JQ973782	
Gca25	KA	271-283	TGAAGTTGCCAGGGAGATCC	AGAGTCCACTGAAGCTGGTG	(AAT)10	JQ973785	

825
826

827 Table S3. Haplotype frequencies for the 652-bp sequences used in this study. Athi, Chyulu and Naivaisha are in Kenya (Brown et al.
828 2016), SGR=Selous Game Reserve in Tanzania and LVNP=Luangwa Valley National Park in Zambia (Coimbra et al. 2021)

Haplotype ID	ATHI	BWMA	CHYULU	LMNP	LVNP	MRC	NCA	NAIVASHA	SNP	TNP	SGR	Total
ERV1.C		1				1				8		10
ERV1.A						1						1
ERV2.B		2										2
ERV2.D										1		1
ERV4			2	2		2		2		2		10
ERV5		2								10		12
ERV9.G	5	5	12	6		24		4		17	5	78
WRV1.C				7			9		17			33
WRV1.M							2		2			4
WRV2									2			2
WRV3.C									4			4
WRV3.B									2			2
ERV8										3		3
WRV5.B									1			1
WRV5.A									1			1
WRV4.A							1					1
WRV4.B							8					8
LVNP8-14					6							6
AthiM18	2											2
AthiM8	10											10
AthiM25	1											1
ChyuluM25			1									1
Lobo2064									1			1
Total	18	10	15	15	6	28	20	6	30	41	5	194

829
830
831
832

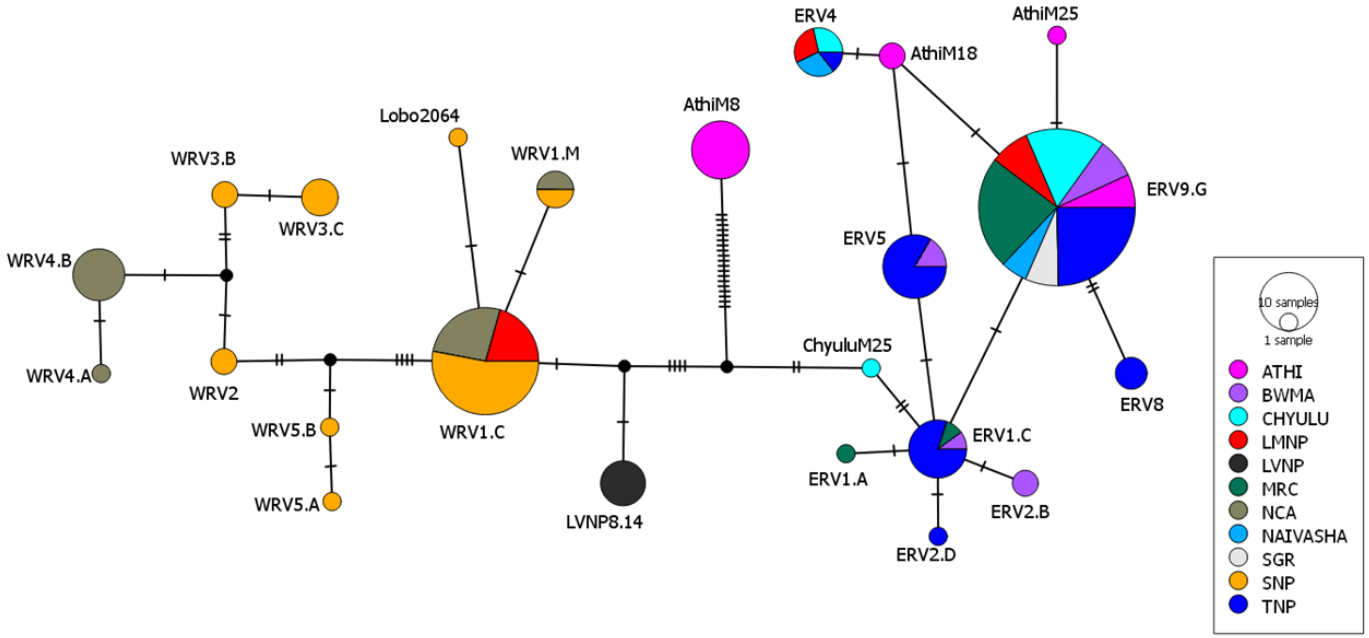
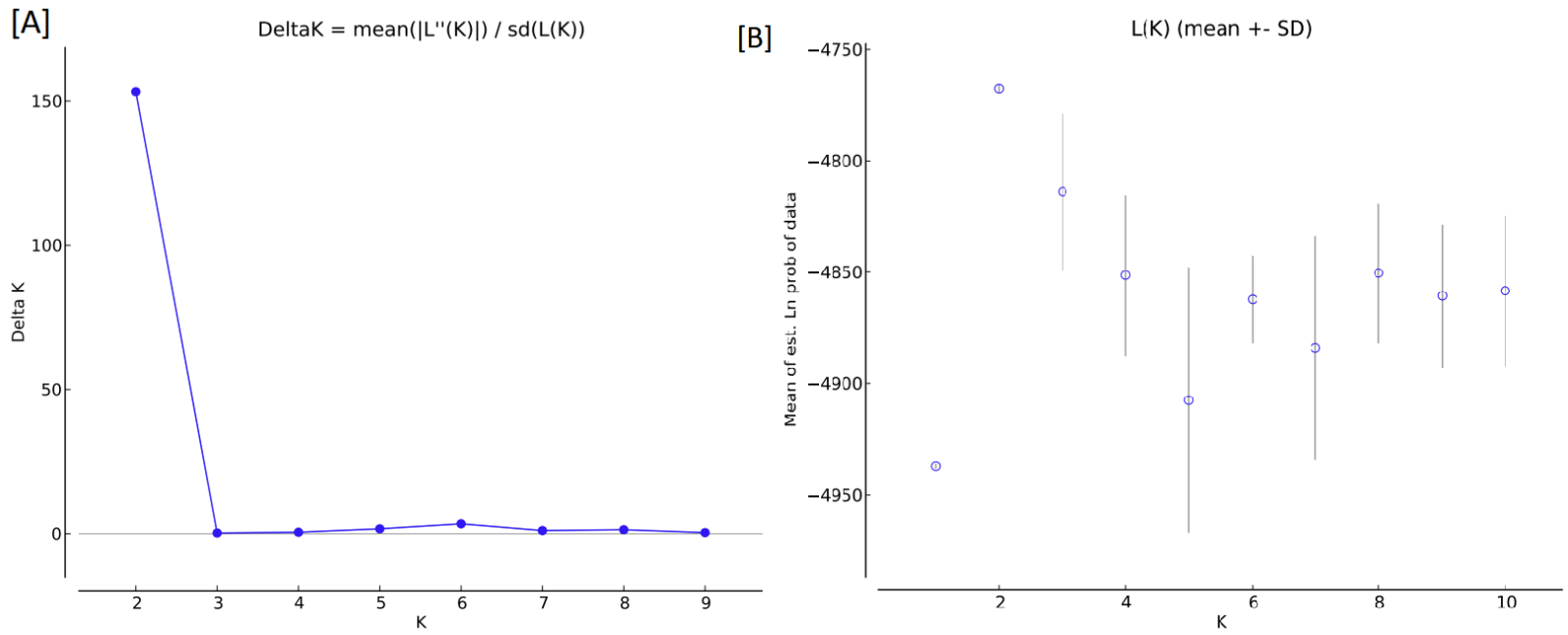


Fig. S1. A neighbor joining network for 652bp mtDNA haplotypes for 194 Masai giraffes. Giraffes from Athi, Chyulu and Naivasha in Kenya carry haplotypes that are clustered in the East Rift Valley subclade.

Table S4. Pairwise F_{ST} for 183 Masai giraffes based on 652 bp of the mitochondrial DNA. In this analysis we included published sequences (Brown et al. 2007). F_{ST} values with asterisk sign (*) indicate significance at $p \leq 0.05$

Pairwise FST	Athi	Chyulu	BWMA	LMNP	MRC	TNP	NCA	SNP
Athi	0							
Chyulu	0.349*	0						
BWMA	0.222*	0.083	0					
LMNP	0.279*	0.228*	0.145*	0				
MRC	0.446*	0.047	0.154*	0.354*	0			
TNP	0.217*	0.131*	-0.010	0.154*	0.185*	0		
NCA	0.354*	0.473*	0.327*	0.173*	0.564*	0.299*	0	
SNP	0.349*	0.453*	0.324*	0.108*	0.530*	0.299*	0.106*	0



852
853 Fig. S2. Delta K Δ*K* plots and ln Pr(X|*K*) used to estimate the Masai Giraffe population clusters inferred from STRUCTURE analysis
854

855

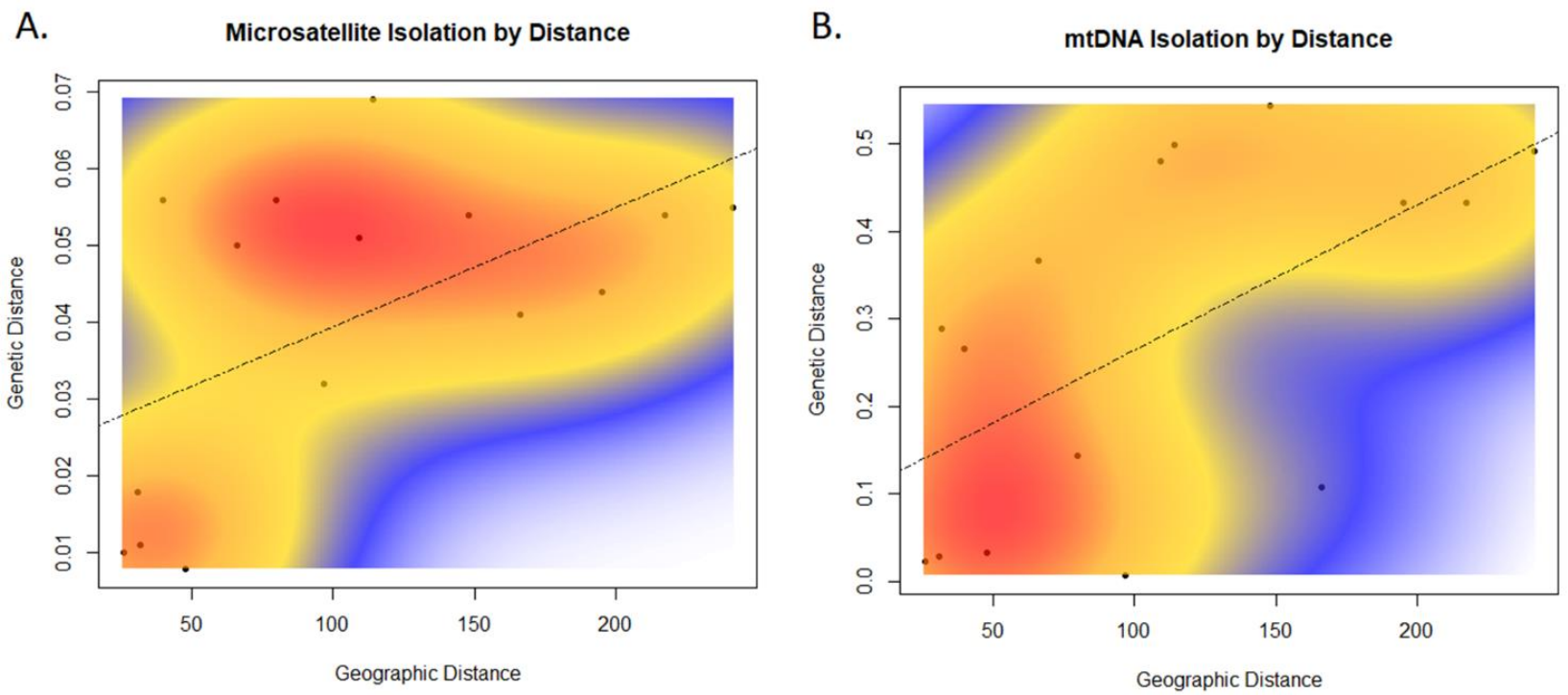


Fig. S3. Graphs showing a correlation between genetic distance (F_{ST}) and geographic distance (km). Isolation by distance was detected, populations that are geographically close to each other show higher genetic similarities that distant populations. Mantel test shows a positive correlation between genetic distance and geographic distance . Monte-Carlo test based on 10000 replicate using Euclidean distance for all pairwise genetic and geographic distance was significant with the observed p=0.3139, simulated p-value 9.999e-5. When IBD was tested using pairwise F_{ST} value against the pairwise geographic distance simulated p=0.0161 for microsatellites and p=0.0189 for mtDNA