Molecular Phylogenetics and Evolution 201 (2024) 108198

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Population genomics of the southern giraffe

David Prochotta [a,b,*], Sven Winter [b,c], Julian Fennessy [d,e], Axel Janke [a,b,f]

[a] Institute for Ecology, Evolution and Diversity, Goethe University, Max-von-Laue-Strasse. 9, Frankfurt am Main, Germany
[b] Senckenberg Biodiversity and Climate Research Centre (BiK-F), Georg-Voigt-Strasse 14-16, Frankfurt am Main, Germany
[c] Research Institute of Wildlife Ecology, Vetmeduni Vienna, Vienna, Austria
[d] Giraffe Conservation Foundation, PO Box 86099, Eros, Windhoek, Namibia
[e] School of Biology and Environmental Science, University College Dublin, Ireland
[f] LOEWE Centre for Translational Biodiversity Genomics, Frankfurt, Germany

A B S T R A C T

Studying wildlife taxonomic diversity and identifying distinct populations has traditionally been largely based on morphology and geographic origin. More recently, this method has been supplemented by genetic data from the mitochondrial genome. However, this is limited as only maternally inherited and may not reflect the true nature of a population's genetics. Within the giraffe (*Giraffa* spp.), subspecies and unique populations were successfully characterized using both mitochondrial and genomic DNA studies, which led to new insights and, in some cases, unexpected results that required further verification. Here, we sequenced the genomes of 85 southern giraffe (*G. giraffa*) individuals from ten populations across southern Africa for a detailed investigation into the genetic diversity and history of its two subspecies, the Angolan (*G. g. angolensis*) and the South African (*G. g. giraffa*) giraffe. While the overall genotypes show low levels of runs of homozygosity compared to other mammals, the degree of heterozygosity is limited despite the large population size of South African giraffe. The nuclear genotype is largely congruent with the mitochondrial genotype. However, we have identified that the distribution of the Angolan giraffe is not as far east as indicated in an earlier mitochondrial DNA study. Botswana's Central Kalahari Game Reserve giraffe are unique, with a clear admixture of Angolan and South African giraffe populations. However, the enigmatic desert-dwelling giraffe of northwest Namibia is locally distinct from other Angolan giraffe yet exhibits intra-subspecies signs of admixture resulting from a recent introduction of individuals from Namibia's Etosha National Park. Whole genome sequencing is an invaluable and nearly indispensable tool for wildlife management to uncover genetic diversity that is undetectable through mitogenomic, geographical, and morphological means.

1. Introduction

The southern giraffe (*Giraffa giraffa*) is one of four distinct giraffe species delimited based on whole genome analysis (Coimbra et al., 2023, 2021), which led to a reassessment of their taxonomic status (Bánki et al., 2024; ITIS, 2024; Mammal Diversity Database, 2024). Today, an estimated 48,000 southern giraffe inhabit southern African countries, with populations expanding due to concerted conservation efforts by the countries' respective governments, the private sector, and local communities (Brown et al., 2022).

Historically, giraffe were considered a single species with nine recognized subspecies distinguished primarily based on their pelage pattern, morphometrics, and assumed geographic distribution (Dagg, 1971; Lydekker, 1904). Other studies have suggested a three-species concept, although based on limited nuclear data (Petzold and Hassanin, 2020), or the historic one-species concept basing their findings on mitochondrial data (Hassanin et al., 2007).

Currently, two subspecies of the southern giraffe are accepted: the Angolan giraffe (*G. g. angolensis*) and the South African giraffe (*G. g. giraffa*). Five Southern African giraffe subspecies have historically been described in a convoluted and confusing nomenclature, most of which are no longer recognized: *G. g. giraffa* (first described as *G. camelopardalis giraffa*; Boddaert, 1785) and *G. g. angolensis* (first described as *G. c. angolensis;* Lydekker, 1903), *G. c. wardi* (Lydekker, 1904), *G. c. capensis* or Cape giraffe (Lesson, 1842), and *G. (c.) infumata* (Noack, 1908).

The traditional identification and distribution of the two southern giraffe subspecies highlight minimal differences in pelage pattern and large intrapopulation variation (Dagg and Foster, 1982; Fennessy, 2004; Hausen, 2017). Inhabiting large areas of Botswana and Namibia, with small populations in Angola, the Angolan giraffe range differs from the South African giraffe, which occurs in northern Botswana, Malawi, Mozambique, northeast Namibia, South Africa, Zambia, and Zimbabwe (Fig. 1). Due to a similarity in pelage pattern, skull morphology, and geographical range between the two southern giraffe subspecies, coupled with the feasibility of crossing theorized natural barriers, their

* Corresponding author.
E-mail addresses: david.prochotta@senckenberg.de (D. Prochotta), svenwinter.research@gmail.com (S. Winter), julian@giraffeconservation.org (J. Fennessy), axel.janke@senckenberg.de (A. Janke).

https://doi.org/10.1016/j.ympev.2024.108198
Received 15 November 2023; Received in revised form 30 July 2024; Accepted 11 September 2024
Available online 12 September 2024

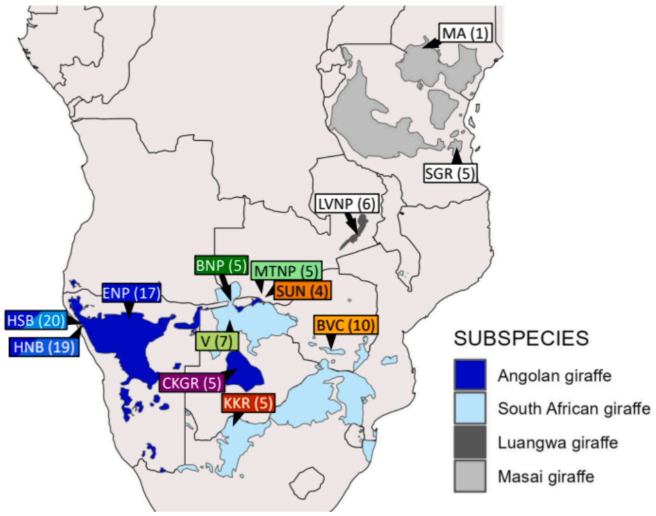

Fig. 1. Southern giraffe skin biopsy sampling sites and distribution patterns across southern Africa. The abbreviations (Table 1) in boxes identify the different sampling localities. Dark blue represents the Angolan giraffe and light blue represents the South African giraffe. Dark grey represents the Luangwa giraffe and grey represents the Masai giraffe. Updated geographical ranges by courtesy of Giraffe Conservation Foundation (2024) based on morphological and genetic data. HSB: Hoarusib River Catchment, HNB: Hoanib River Catchment, ENP: Etosha National Park, BNP: Bwabwata National Park, V: Vumbura Concession, MTNP: Mosi-oa-tunya National Park, SUN: Zambezi Sun, BVC: Bubye Valley Conservancy, KKR: Khamab Kalahari Reserve, CKGR: Central Kalahari Reserve, MA: Masai Mara National Reserve, LVNP: Luangwa National Park, and SGR: Selous Game Reserve. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

subspecies assignment has never been clear (Bock et al., 2014; Fennessy, 2004; Seymour, 2001).

Population-level genetic analyses of giraffe have predominantly been based on mitochondrial DNA (mtDNA) and initially aided the identification of distinct giraffe (sub)species and local populations (Bock et al., 2014; Fennessy et al., 2016). A complex distribution of the two southern giraffe subspecies was revealed, with the Angolan giraffe maternal haplotype occurring much further east in the Bubye Valley Conservancy (BVC), Zimbabwe, than previously assumed (Winter et al., 2018). However, based on pelage pattern and geography, this population has historically been considered South African giraffe. Coimbra et al. (2021) indicated admixture – sometimes human-facilitated through translocations – between the two subspecies, suggesting recent gene flow. However, a detailed genomic analysis of their taxonomic status and distribution was lacking.

Botswana's Central Kalahari Game Reserve (CKGR) population lies between traditionally recognized Angolan giraffe populations to the west and South African giraffe populations to the north, south and east (Fig. 1). MtDNA analyses identified individuals from CKGR as Angolan giraffe (Winter et al., 2018), potentially connecting to the eastern occurrence in the of the Angolan haplotype in the BVC, resulting in an unexpected wedge-like split in the geographical distribution of the South African giraffe.

The enigmatic desert-dwelling Angolan giraffe inhabits the hyper-arid northern Namib, the oldest known desert at 55–80 Myr old (Goudie, 2010). The giraffe there have evolved a lighter, paler coat pattern than other Angolan giraffe populations (Fennessy, 2004). The population is currently estimated at 455 individuals (GCF pers. comm.), historically reduced by poaching to $<$ 100 individuals during the Namibian War of Independence 1966—1990, and likely resulting in a genetic bottleneck (Fennessy, 2004; Reardon, 1986). Initial MtDNA analyses suggested that the desert-dwelling population is genetically distinct from other Angolan giraffe, including those living approximately 300 km east in the Etosha National Park (ENP) (Winter et al. 2018). In fact, Winter et al. (2018) suggested that attempts to augment the desert-dwelling giraffe population with individuals from ENP in 1991 (Fennessy, 2004) were largely unsuccessful, as the ENP haplotype was found in only two of the desert-dwelling giraffe sampled. This study also estimated that these two populations have been distinct (separated) for more than 30,000 years.

As with many small and isolated populations, including the desert-dwelling Angolan giraffe, inbreeding and loss of genetic variability are common conservation concerns and potential threats to their long-term survival. Many small populations of the southern giraffe are isolated in private and public parks and reserves and may suffer from the loss of genetic variability and/or inadvertently be hybridized with other subspecies through transboundary translocations. However, the impact of this is unknown for giraffe.

To investigate the genetic health, taxonomic assignment, and geographic distribution of key southern giraffe populations at the genomic level, and to estimate their genomic diversity, we genome-sequenced an additional 85 southern giraffe individuals with a 10-fold Illumina short-read coverage goal. The datasets were analyzed together with twelve publicly available southern giraffe sequences (Coimbra et al., 2021). To complete (root) some phylogenomic analyses, we added available genomes from the other three giraffe species (*G. camelopardalis*, *G. reticulata*, *G. tippelskirchi*) as well as the okapi (*Okapia johnstoni*). For phylogenomic analysis, we supplemented with Masai giraffe samples from Coimbra et al. (2021), as well as one Masai giraffe and an Okapi sequence from Agaba et al. (2016).

These analyses allow for a detailed assessment of population genomics, such as admixture, genetic structure and heterozygosity of the two southern giraffe subspecies and their key populations in areas across their natural occurrence.

2. Materials and methods

2.1. Sampling

Skin biopsies from 85 southern giraffe were obtained from Southern Africa using remote biopsy darts and preserved in $\geq$ 97 % ethanol (Fig. 1; Table 1). The Giraffe Conservation Foundation (GCF) and partners collected biopsies with country-specific permissions, following ethical guidelines under each country's national law and GCF IACUC protocol. DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany).

The map (Fig. 1) was created in R with ggplot2 v3.4.4 (Wickham, 2011), maps v3.4.2 (Deckmyn, 2018), rnaturalearth v1.0.1 (Massicotte and South, 2023), and sf v1.0–15 (Pebesma et al., 2020) with shapefiles from GCF (Brown et al., 2022).

2.2. Genome sequencing

Whole genomes of the Angolan and South African giraffe individuals were sequenced by Novogene Europe (Cambridge, UK) on the Illumina NovaSeq 6000 platform (2 $\times$ 150 bp, 350 bp insert size) to about 10-fold coverage. For the downstream analyses, we added short-read sequences from 12 southern giraffe and 11 Masai giraffe (Coimbra et al., 2021), as well as an Okapi and one Masai giraffe (Agaba et al., 2016) (**Table S2**).

2.3. Filtering and mapping

To filter the raw FASTQ files, fastp v.0.23.4 (Chen et al., 2018) was run with base correction and low complexity filter enabled. Adapters and polyG tails were detected and removed by default. Reads with base quality $<$ 15 in a sliding window of four bp, reads shorter than 36 bp, and reads with more than five unknown bases were also removed.

The remaining reads were mapped to a chromosome-level Masai

Table 1
Population abbreviations and locations. The table contains the full names and locations of the studied populations, as well as the respective species assignment.

Population ID	Full name	Country	(SUB)Species
HSB	Hoarusib River catchment	Namibia	*Giraffa giraffa angolensis*
HNB	Hoanib River catchment	Namibia	*Giraffa giraffa angolensis*
CKGR	Central Kalahari Game Reserve	Botswana	*Giraffa giraffa giraffa*
ENP	Etosha National Park	Namibia	*Giraffa giraffa angolensis*
BNP	Bwabwata National Park	Namibia	*Giraffa giraffa giraffa*
BVC	Bubye Valley Conservancy	Zimbabwe	*Giraffa giraffa giraffa*
KKR	Kalahari Khamab Reserve	South Africa	*Giraffa giraffa giraffa*
MTNP	Mosi-oa-tunya National Park	Zambia	*Giraffa giraffa giraffa*
SUN	Zambezi Sun Hotel (Avani Victoria Falls Resort)	Zambia	*Giraffa giraffa giraffa*
V	Vumbura Concession	Botswana	*Giraffa giraffa giraffa*
LVNP	Luangwa Valley National Park	Zambia	*Giraffa tippelskirchi thornicrofti*
MA	Masai Mara National Reserve	Kenya	*Giraffa tippelskirchi tippelskirchi*
SGR	Selous Game Reserve	Tanzania	*Giraffa tippelskirchi tippelskirchi*
WOAK (outgroup)	White Oak Holdings	United States	*Okapia johnstoni*

giraffe s. str. genome assembly (Farré et al., 2019)(GCA_013496395) using BWA mem v.0.7.17 (Li, 2013). The BAM files were deduplicated with picard v.2.20.8 (Broad Institute, 2019) and Qualimap v.2.2.2 and multiqc v.1.14 (Ewels et al., 2016; García-Alcalde et al., 2012) were used to create and combine mapping statistics for the BAM files. Indels were realigned using GATK v.3.8.1 (McKenna et al., 2010). Repeats in the reference were identified and masked using the Cetartiodactyla database of RepeatMasker v.4.1.2, and a de-novo RepeatModeler v.2.0.3 library (Flynn et al., 2020; Smit et al., 2020) and a BED file containing the repetitive regions was created using bedtools v.2.28.0 (Quinlan and Hall, 2010). The BAM files were then cleaned and only proper read pairs mapped to non-repetitive regions of the autosomes were retained (Coimbra et al., 2021).

2.4. Genotype calling and downstream analyses

SNP calling was performed with ANGSD v.0.935 (Korneliussen et al., 2014) using the default SAMTools model for genotype likelihood estimation. Extended BAQ calculation (flag −-baq 2) was enabled, with a minimum mapping and base quality of 30 (−minMapQ 30 −minQ 30). Depth statistics for each individual combined was inferred with Sambamba v.1.0.0 (Tarasov et al., 2015). Optimal maximum and minimum depths were calculated using the site depth distribution of all individuals: median ± the median's absolute deviation. Sites with a strand bias p-value, HWE, heterozygous bias $< 1 \times 10^{-6}$ were removed (−doHWE 1 −hwe_pval 1e-6 −sb_pval 1e-6 −hetbias_pval 1e-6). Only biallelic SNPs, those called with a p-value $< 1 \times 10^{-6}$ and with less than 10 % missingness, were retained (−snps_pval 1e-6). Genotype likelihoods output was enabled to output posterior probabilities of all possible genotypes (−doGeno 8). The analysis results were output in a BEAGLE file and a BCF file (−doGLF 2 −doBcf 1). To reduce the size of the BCF file, the genotype likelihood and genotype probability tags were removed from the file with Bcftools v1.15 (Danecek et al., 2021). A second dataset was created by extracting only southern giraffe from the BEAGLE file and the VCF/BCF file. NGSRelate v.2.0 (Korneliussen et al., 2014) was used to analyze kinship among the southern giraffe individuals. NATora v.001105 (Leal et al., 2022) was used for relatedness pruning by setting the cutoff for first-degree relationships (0.1768).

Linkage disequilibrium (LD) pruning was conducted with ngsLD v.1.1.1 (Fox et al., 2019) for both datasets. LD was estimated for SNPs up to 200 kb apart. SNP pairs with r2 > 0.1 in a 150 kbp window were discarded with prune graph (Vieira, 2024).

We utilized four main datasets (both as VCF file and BEAGLE file): 1. complete data set with all individuals and all sites (dataset 1), 2. sub-dataset containing all individuals and only unlinked sites (LD-pruned) sites (dataset 2), 3. sub-dataset containing only southern giraffe and all sites (dataset 3), and 4. sub-dataset containing only unrelated southern giraffe individuals and unlinked (LD-pruned) SNPs (dataset 4).

2.5. Population structure & gene flow

Principal component analysis (PCA) was performed using PCAngsd v.1.10 (Meisner and Albrechtsen, 2018) with the LD-pruned SNP-dataset (dataset 4) of 66 unrelated southern giraffe (dataset 4; unrelated individuals, unlinked sites). The results were plotted in R using modified scripts by Coimbra et al. (2021) using plot3d (Soetaert, 2024). Population admixture was inferred using NGSadmix, running K=1 to K=6 with 100 bootstrap replicates per K. The K=1 and K=6 runs were run only for statistical reasons as CLUMPAK does not recognize the lowest and highest values in the BestK pipeline. The admixture plots were plotted in R using ggplot2 v.3.4 (Wickham, 2011) and manually edited. The CLUMPAK BestK pipeline was used to identify the BestK using Evanno's method. The BestK results and the bootstrap likelihoods of each K were plotted in R with ggplot2 (Wickham, 2011) based on modified versions of the scripts of Coimbra et al. (2021). An additional analysis with NGSadmix omitting the admixed Angolan giraffe individuals was run as described above.

The VCF file (dataset 2; all individuals, unlinked sites) was prepared for SambaR v.1.10 (de Jong et al., 2021) R v4.1.3 (R Core Team, 2022) using plink v.1.9 and vcftools v.0.1.17 (Danecek et al., 2011; Purcell et al., 2007). The input was filtered (indmiss = 0.25; snpmiss = 0.1) with adegenet (Jombart, 2008; Jombart and Ahmed, 2011). Phylogenetic trees were compared with the 'comparetrees' function of SambaR to infer an unrooted tree with the least conflict between pathlengths and the given distance matrix using poppr (Kamvar et al., 2014), phangorn v.2.8.1 (Schliep, 2011), SambaR internal functions, and hclust (R Core Team, 2022). This compares clustering methods such as (Bio)Neighbour Joining and Ordinary Least Squares and distance calculation methods such as Pi distance and Euclidean distances among others (Fig S8). The tree was then rooted and visualized in R with a script from Coimbra et al. (2023) with tidyverse (Wickham et al., 2019), ape v.5.7.1 (Paradis and Schliep, 2019), treeio v.1.2 (Wang et al., 2020), ggtree v.3.1 (Yu et al., 2023) and patchwork v1.2 (Pedersen, 2022).

Dsuite v.0.5 (Malinsky et al., 2021) was run with default settings on the complete dataset (dataset 1; all individuals, all sites) to infer gene flow based on d-statistics and plotted with the included tools. Two separate runs based on two differing topologies were run. The first topology placed the CKGR in the Angolan giraffe, while the second topology placed them in the South African giraffe.

2.6. Mitochondrial analysis

Short-read sequences of 51 Masai giraffe, 42 reticulated giraffe, 53 northern giraffe, and one okapi from GenBank were added (Agaba et al., 2016; Coimbra et al., 2021, 2022, 2023) for de-novo assembly of mitogenomes (**Table S3**).

Mitogenomes were assembled using Mitoflex v.0.2.9 and GetOrganelle v.1.7.7.0 (Jin et al., 2020; Li et al., 2021) and annotated with Mitos2 v.2.1.3 embedded in Galaxy v.23.1 (Arab et al., 2017; Donath et al., 2019; The Galaxy Community, 2022), and MitoZ v3.6 (Meng et al., 2019) (see **Table S3** for details). Nucleotide sequences of the 13 protein-coding genes for each individual were renamed and reordered using parallel, seqtk v1.3-r106, and seqkit v2.0.0 (Li, 2023; Shen et al., 2016; Tange, 2011). The sequences were aligned using MAFFT v.7.475 (Katoh et al., 2002), concatenated with FASconcat v.1.04 (Kück and Meusemann, 2010) and clipped with clipkit v.1.4.1 (Steenwyk et al., 2020). A maximum Likelihood phylogeny was then computed using IQTree v.2.2.2.3 (Minh et al., 2020) with 1000 ultrafast bootstrap replicates (Hoang et al., 2018), BNNI enabled (−-bnni; optimization of ultrafast bootstrap by NNI), and the HKY+F+I substitution model inferred with modelfinder (Kalyaanamoorthy et al., 2017). In addition, a median-joining (mjn) haplotype network was generated in PopArt v.1.7 (Bandelt et al., 1999; Leigh and Bryant, 2015) from the concatenated alignment (without the Okapi outgroup), which was converted into Nexus format using seqmagick (FHCRC, 2024).

For submission of the mitogenomes to NCBI we used MitoAnnotator v.4.00 (Iwasaki et al., 2013; Zhu et al., 2023) to reorient the sequences, reannotated these with MitoZ (Meng et al., 2019) and manually curated the annotations.

2.7. Heterozygosity and ROH (runs of homozygosity)

Heterozygosity was calculated by generating a consensus sequence and estimating the folded site frequency spectrum with 200 bootstrap replicates using realSFS embedded in ANGSD (Korneliussen et al., 2014), using modified versions ('-baq2′ instead of '-baq 1′; removal of '-C 50′) of the scripts of Coimbra et al. (2021).

The southern giraffe only dataset (dataset 3) was used as input for the Runs of Homozygosity (ROH) analysis containing 97 southern giraffe individuals. The BCF file was converted to PLINK format using PLINK and then to the Oxford geno format, needed as input for RZooRoH v.0.3.1 (Druet et al., 2021) (plink −bfile angsd.snps −recode oxford −autosome −out angsd.snps). RoH was inferred separately for each population in RZooRoH v.0.3.1, setting 16 pre-defined classes (15 HBD and 1 non-HBD).

The data was visualized following the scripts by Coimbra et al. (2021) in R for plotting with tidyverse packages (Wickham et al., 2019), viridis v.0.4.2 (Garnier et al., 2021), reshape2 v.1.4.4 (Wickham, 2007), RColorBrewer v.1.1–3 (Neuwirth, 2022), and patchwork v.1.1.3 (Pedersen, 2022).

3. Results

3.1. Sequencing

Short-read sequences were obtained from 85 individuals and supplemented with 15 sequences from databases (**Table S1**, **Table S2**). On average 42.9 GB bp of raw data was produced, with 42.3 bp remaining after filtering The short reads were mapped to the reference genome, the Masai giraffe (Farré et al., 2019), resulting in an average mean mapping depth of 10-fold after deduplication and filtering (**Table S2**).

3.2. Genotype calling, relatedness analysis and linkage pruning

Genotype calling with ANGSD of 110 Individuals resulted in 3,351,815 SNPs across the genome (dataset 1). LD-Pruning of the complete dataset resulted in 179,925 unrelated SNPs across the genome (dataset 2). The removal of related southern giraffe resulted in 66 unrelated southern giraffe individuals remaining (dataset 3; Relatedness analysis: Table S4). Linkage disequilibrium pruning (LD-Pruning) of the 66 unrelated southern giraffe individuals resulted in 144,868 unlinked SNPs (dataset 4).

3.3. Phylogeny

The mitogenomic analyses (**Figure S5, S6**) confirmed the previous matrilineal identification of the Angolan giraffe individuals (Coimbra et al., 2023; Winter et al., 2018). Based on the mtDNA analyses, the BVC and CKGR populations were firmly placed within the Angolan giraffe, both in the tree and the haplotype network. The South African giraffe populations instead clustered with the Masai giraffe. Mitogenomic discordance could be found within the other giraffe species as well. Two Masai giraffe individuals (GF005, GF007) were found within the reticulated giraffe and two reticulated giraffe individuals (RetRot1, GF292) were found within the northern giraffe.

The IQTree phylogeny (Maximum Likelihood; dataset 1, based on 3,197,800 after removal of invariant sites; Figure **S7**) and the genomic distance tree (Ordinary Least Squares; Pi distance; dataset 3,179,911 SNPs after filtering) (Fig. 2) show differing placements of the CKGR population. The IQtree phylogeny places the CKGR with the Angolan giraffe, while the distance tree places them with the South African giraffe. The placement of the BVC with the South African giraffe is congruent among both trees. The populations of the South African giraffe appear to be more isolated, while the Angolan giraffe in Namibia showcases many Hoanib River (HNB) and Hoarusib River (HSB) individuals clustering within the other and in the ENP population.

3.4. Admixture analyses and gene flow

At K=2, the admixture analysis of genomic sequences identified two clusters representing the respective subspecies, Angolan and South African giraffe (Fig. 3**A**). Four ENP individuals showed slight admixture with the South African giraffe. The CKGR giraffe showed more pronounced signals of admixture between the two subspecies with admixture proportions of ∼ 33 % Angolan and ∼ 66 % South African giraffe origin. The BVC giraffe clustered with the South African subspecies with no evidence of admixture. The other South African giraffe populations exhibited no signals of shared ancestry with the Angolan giraffe either. K=2 is the BestK with Evanno's method and the increase in mean likelihood decreases at K=2 (Fig. 3**B-C**), likewise identifying K=2 as the best fit.

Sub-structuring among the Angolan giraffe began at K=3, separating individuals from the northwest Namibian desert-dwelling population (HNB, HSB) and the ENP. Admixture with ENP was particularly frequent within the HSB giraffe population, while only three individuals within the HNB population showed admixture with the ENP. HNB101 does not have a signal of the desert-dwelling giraffe at all and appears to be a pure ENP. The individuals of the CKGR retained the admixture signal from the Angolan giraffe (specifically ENP).

Further sub-structuring within the South African giraffe was evident at K=4. BVC and Zambezi Sun Hotel (SUN), Zambia individuals clustered separately from the remaining South African populations. The CKGR giraffe exhibited admixture signals from the ENP cluster and both South African clusters. The Kalahari Khamab Reserve (KKR), South Africa population showed mixed ancestry among the southern giraffe clusters. Therefore, giraffe from the CKGR appear to be a genetic melting pot with shared ancestry of various South African giraffe populations as well as Angolan giraffe from ENP. At K=5, CKGR became its own cluster. The Mosi-oa-Tunya NP (MTNP) individual showed admixture with the CKGR, Vumbura Concession (V), Botswana and Bwabwata NP (BNP), Namibia populations. The KKR contained signals of all three South African giraffe clusters.

An admixture analysis of non-admixed desert-dwelling giraffe individuals showed a separation of the desert populations and the ENP population (**Figure S9**), while no changes were observed in the South African giraffe clusters. This additional analysis indicates that the desert-dwelling giraffe is, in fact, its own distinct population, likely with adaptations to the extreme environment.

The Dsuite analysis (Fig. 4) shows weak signals of introgression between the Luangwa Valley (LVNP), Zambia lineage, and the southern giraffe. Within the southern giraffe, Dsuite indicates gene flow between the ENP and HSB populations in both topologies but does not detect any between the ENP and the HNB. The CKGR population and the Angolan giraffe indicate strong gene flow patterns between them, as well as with

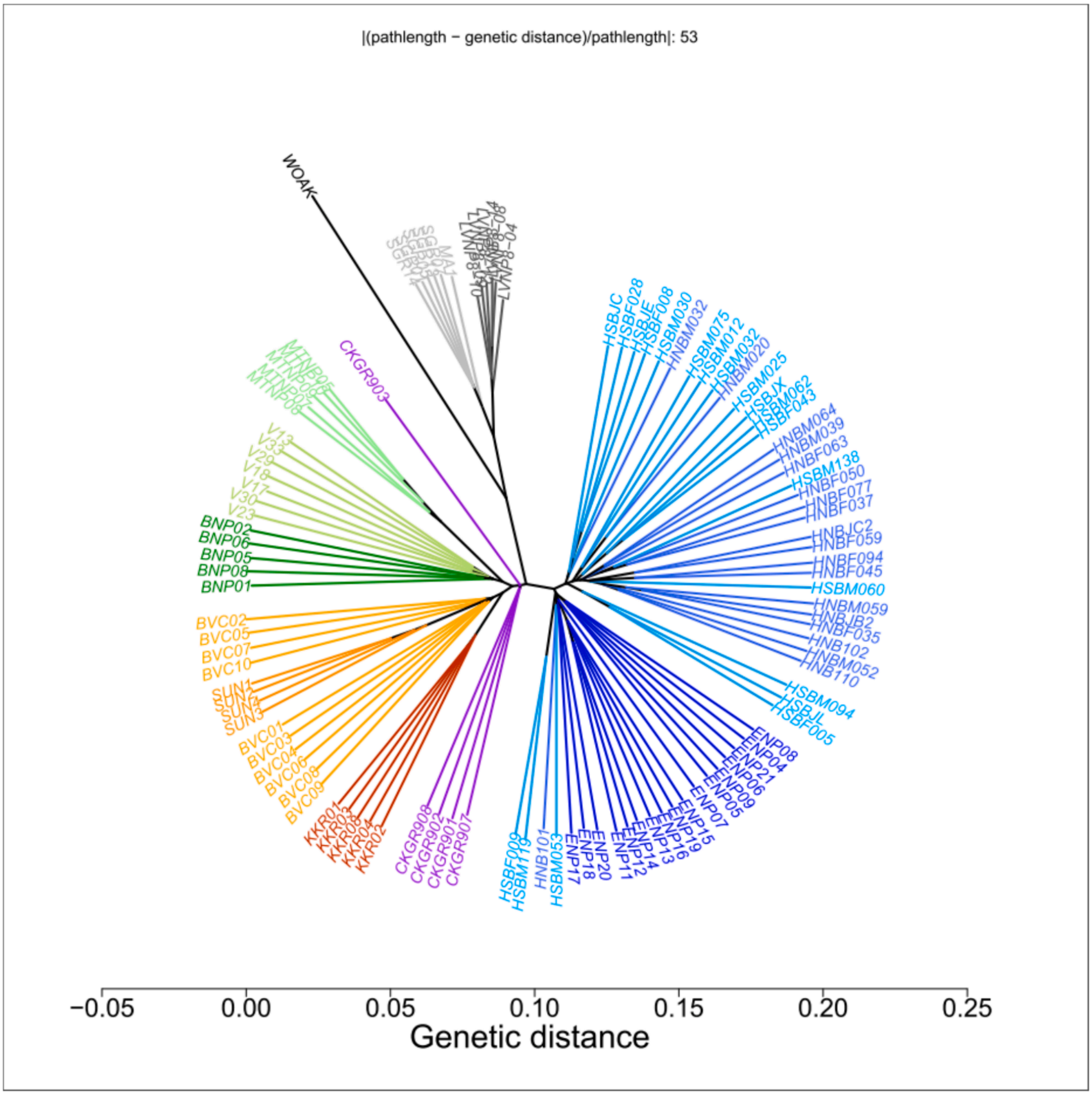

Fig. 2. Dendrogram based on Pi-distance and OLS (ordinary least squares). The dendrogram depicts the genomic differences, Fst, and bottle-necks unlike a classic tree (de Jong et al., 2023). While CKGR individuals branch off from a common point, indicating admixture, while other, isolated populations (MTNP, KKR, SUN) have internal branches, which indicate loss of genetic variability, indicating bottle-neck events. The distance-clustering comparison can be found in S8.

the South African giraffe, depending on the topology (up to and more than 25 %). The CKGR and the MTNP also exhibited minor gene flow (~7%).

3.5. PCA analysis

A principal component (PC) analysis (Fig. 5) based on 144,868 SNPs was congruent with the results of the structure and phylogenetic analyses. The Angolan giraffe (circles) are clearly separated from the South African giraffe, with individuals from the ENP being distinct from the HNB and HSB populations. Individuals from the HNB and HSB form largely separate groups. Interestingly, among the desert-dwelling giraffe, individuals from the HSB show greater spread of variation (light blue circles) along the second PC (PC2), placing them between the ENP and HNB populations. This is consistent with the greater admixture of the ENP genotype into the HNB genotype.

The South African giraffe individuals (diamonds) are tightly grouped within their origin, showing less intrapopulation variation than the Angolan giraffe. Interpopulation variation is larger, with the populations being spread in all three PC axes, especially along PC2. The BNP and V populations cluster together, which supports the findings of the admixture plot, most likely due to their geographical proximity to one another and resulting genetic exchange. The same tight clustering can be found in the BVC and SUN, though the origin of the latter is unknown (J. Fennessy pers. comm.). The MTNP population similarly is of unknown origin (Giraffe Conservation Foundation, 2022), though the PCA and admixture suggest a close relationship with the BNP and V populations.

Along PC1, CKGR is closer to the other South African giraffe, as is the case for BVC. Both BVC and SUN are tightly clustered along all three axes, as are the BNP and V populations. As expected from the mixed ancestry the genotype of the CKGR individuals (purple diamonds) differentiates itself from the South African giraffe at PC1 toward the Angolan giraffe and is placed in the middle of both subspecies.

3.6. Heterozygosity and runs of homozygosity (RoH)

The Angolan and South African giraffe genomes show an approximately equal degree of heterozygosity among all individuals, varying from 0.05 % to 0.08 % (Fig. 6**A**). Even the relatively isolated and small desert-dwelling populations (HNB, HSB) have approximately the same level of heterozygosity as the more connected and denser giraffe populations of the ENP and the South African giraffe.

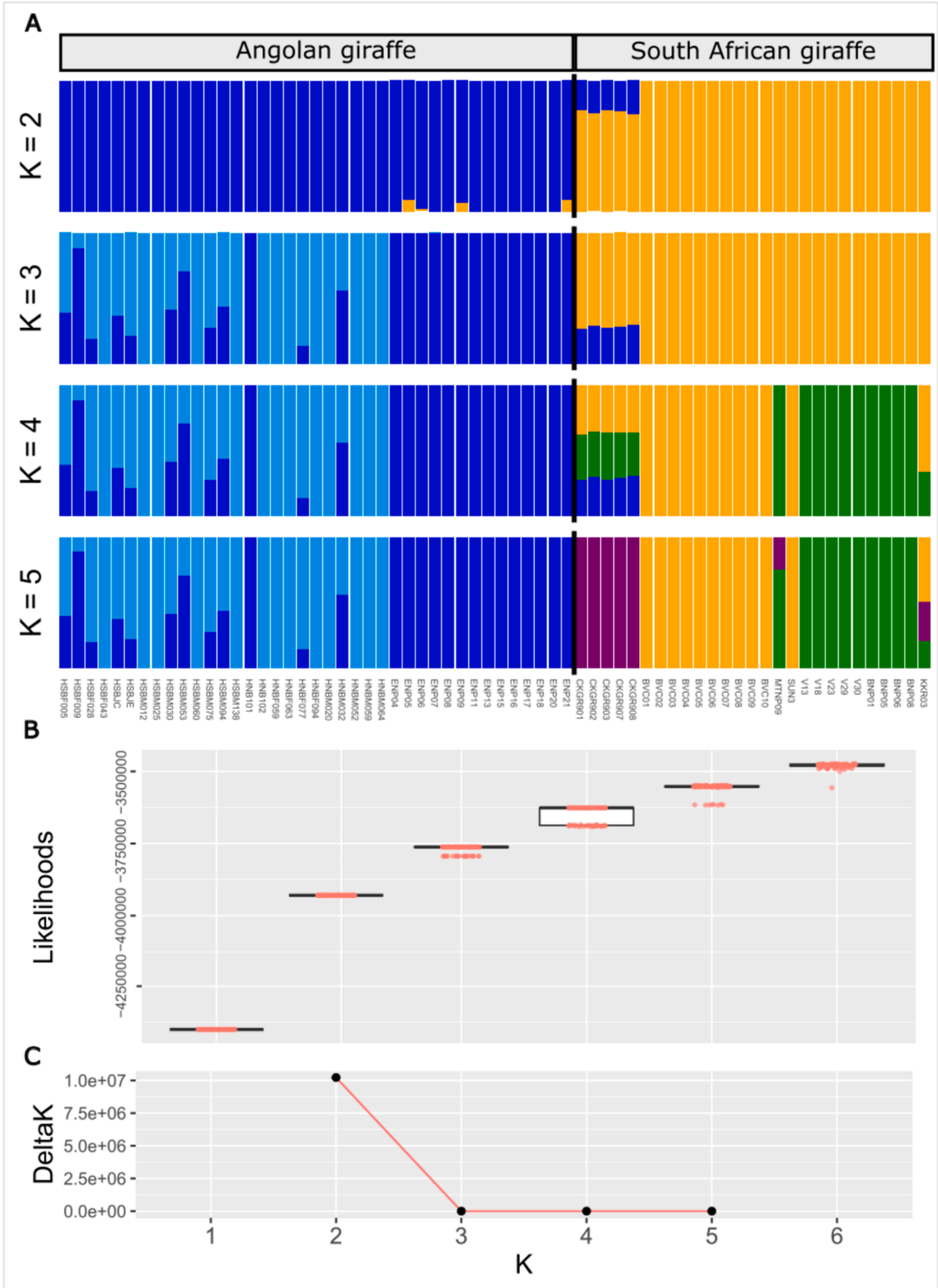

Fig. 3. Structure analysis, likelihoods and DeltaK analysis. A Admixture for 66 unrelated Angolan and South African giraffe for different numbers of assumed populations for K=2 to K=5 (K1 and K6 not shown, only used for statistical reasons). Colors indicate a cluster, abbreviations for the localities see Table 1. At K=2 the two subspecies are well separated, with four ENP individuals showing slight admixture from the South African giraffe. The CKGR giraffe shows roughly 1/3 of admixture from the Angolan giraffe. At K=3 the ENP giraffe becomes a distinct cluster and reveals extensive introgression of ENP in the desert giraffe population. At K=4 the MTNP, V and BNP giraffe form a separate cluster from the BVC and SUN giraffe. CKGR shows admixture and retains the signal of the ENP giraffe, and also shows admixture of the two South African giraffe clades. KKR shows a mixed signal from the South African giraffe clusters. At K=5 CKGR forms a separate cluster, while MTNP and KKR show mixed signals from the South African giraffe clusters. **B** Mean likelihoods boxplot. The difference in median likelihood of 100 runs per K starts plateauing at K=2, while statistical dispersion starts increasing. **C** BestK by Evanno plot, identifying K=2 as the best K value.

Analyses of RoH show that all southern giraffe populations, as well as the desert-dwelling Angolan individuals have a relatively low number of long homozygous stretches in their genome (Fig. 6**B**). Among the southern giraffe, individuals from the MTNP and SUN showed elevated levels of the amount and the length of their RoH. This may be explained by the high relatedness within the populations, especially the MTNP as four (out of five total) individuals were removed due to relatedness. The CKGR population exhibits lower levels of F_{ROH} and no recent RoHs, as long stretches (yellow) are absent compared to other populations.

The number and total length (Fig. 6**C**) of RoHs in individuals from each population roughly reflect the variability in their HBD class composition and inbreeding coefficient. This is particularly evident in the desert-dwelling giraffe (HNB, HSB), and BVC. Compared to the Angolan giraffe populations, the South African giraffe populations appear to possess a larger diversity in the total length (sum) and number of their RoH.

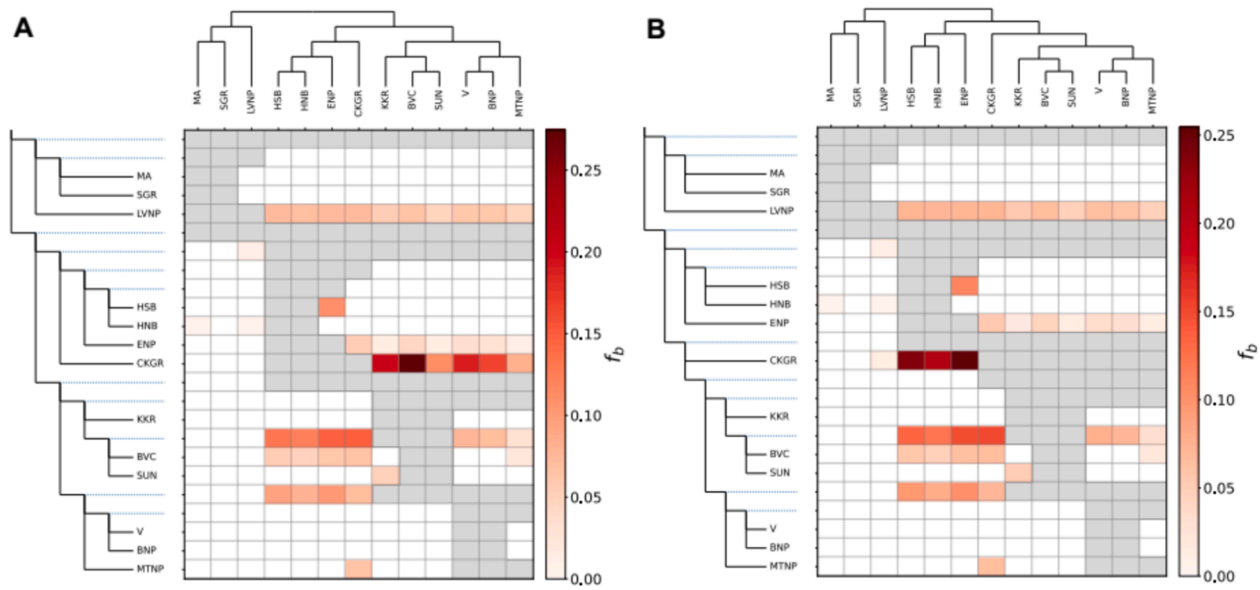

Fig. 4. Dsuite Fbranch test. The red boxes indicate gene flow signal, the darker the stronger the signal. Both variations show similar results with weak signal of introgression from the Masai giraffe and Luangwa giraffe to the southern giraffe. Within the southern giraffe, we see a strong signal of gene flow across the populations, with very strong signals between the Angolan populations and CKGR. Weaker signals can be observed between the Angolan and remaining South African giraffe. Grey boxes signify that for the given topology Fbranch cannot be calculated **A** Based on the IQTree phylogeny with CKGR related to the Angolan populations. **B** Based on pi distance phylogeny with CKGR closely related to the South African populations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

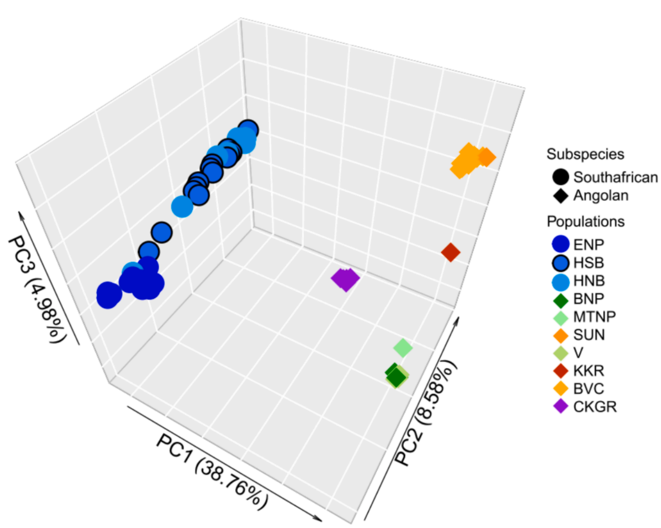

Fig. 5. PCA analyses of 66 unrelated giraffe individuals. The PCA shows a clear separation along PC1 of two major clusters corresponding to the two subspecies, Angolan (circles) and South African giraffe (diamonds). The wider spread desert-dwelling Angolan giraffe individuals (HSB,HNB; blue and circled blue) indicates greater genomic variation. In contrast, the South African giraffe populations are more closely clustered. The BVC population is clearly clustering with the South African giraffe. CKGR individuals take an intermediary position between the subspecies clusters, but is closer to the South African giraffe. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

4.1. Classification of the Bubye Valley Conservancy (BVC) and Central Kalahari Game Reserve (CKGR)

In contrast to mtDNA analyses (Hassanin et al., 2007; Winter et al., 2018), genomic sequence analyses recover the entire genetic heritage of an individual. This study unambiguously identified the BVC giraffe individuals as of South African giraffe subspecies (*G. g. giraffa*) and highlights the limitation of mtDNA analyses for taxonomic purposes when confronted with cases of mitochondrial capture or incomplete lineage sorting.

The genome of the CKGR giraffe individuals, previously identified as Angolan giraffe by mtDNA analyses (Bock et al., 2014; Winter et al., 2018), shows signals of mixed ancestry between South African and Angolan giraffe (specifically ENP population, Fig. 3**A**). The process of mitochondrial capture, leading to conflicts between the phylogenetic signal from the matrilineally inherited mtDNA and the nuclear genome, has been observed in a previous study of giraffe (Petzold and Hassanin, 2020) and other species such as bears, parrots, and frogs (Hailer et al., 2012; Shipham et al., 2015; Zhou et al., 2012).

The BVC individuals have the genomic characteristics of the South African giraffe genotype with only traces of genomic admixture from the Angolan giraffe. Thus, the mitochondrial genotype appears to have entered the population in the past and has been maintained, probably by the philopatry of females with South African giraffe nuclear genotype. Over time, the Angolan nuclear genotype has been lost due to backcrossing with the South African population, while the Angolan giraffe mitochondrial genotype has been maintained.

Similarly, the classification of individuals from the CKGR as Angolan giraffe based on mtDNA analyses (Bock et al., 2014; Winter et al., 2018) and historical and current distribution maps (Brown et al., 2022; O'Connor et al., 2019) can, therefore, no longer be maintained. Interestingly, the genomes of CKGR individuals harbor only about one-third of the Angolan giraffe genotype that is ostensibly of ENP origin, even though it may also originate from unsampled other Angolan giraffe populations. Two-thirds of the genomes of CKGR individuals are clearly of the South African giraffe genotype, with one-third originating from the Angolan giraffe (Fig. 3**A**). This also makes it difficult to place in a bifurcating tree as results can vary between different methods (Fig. 2). The mixed genomic ancestry of the CKGR is a unique feature among southern giraffe that should be considered in future conservation efforts to: (a) preserve this naturally evolved feature that is likely a consequence of the geographic location between the Angolan and the South

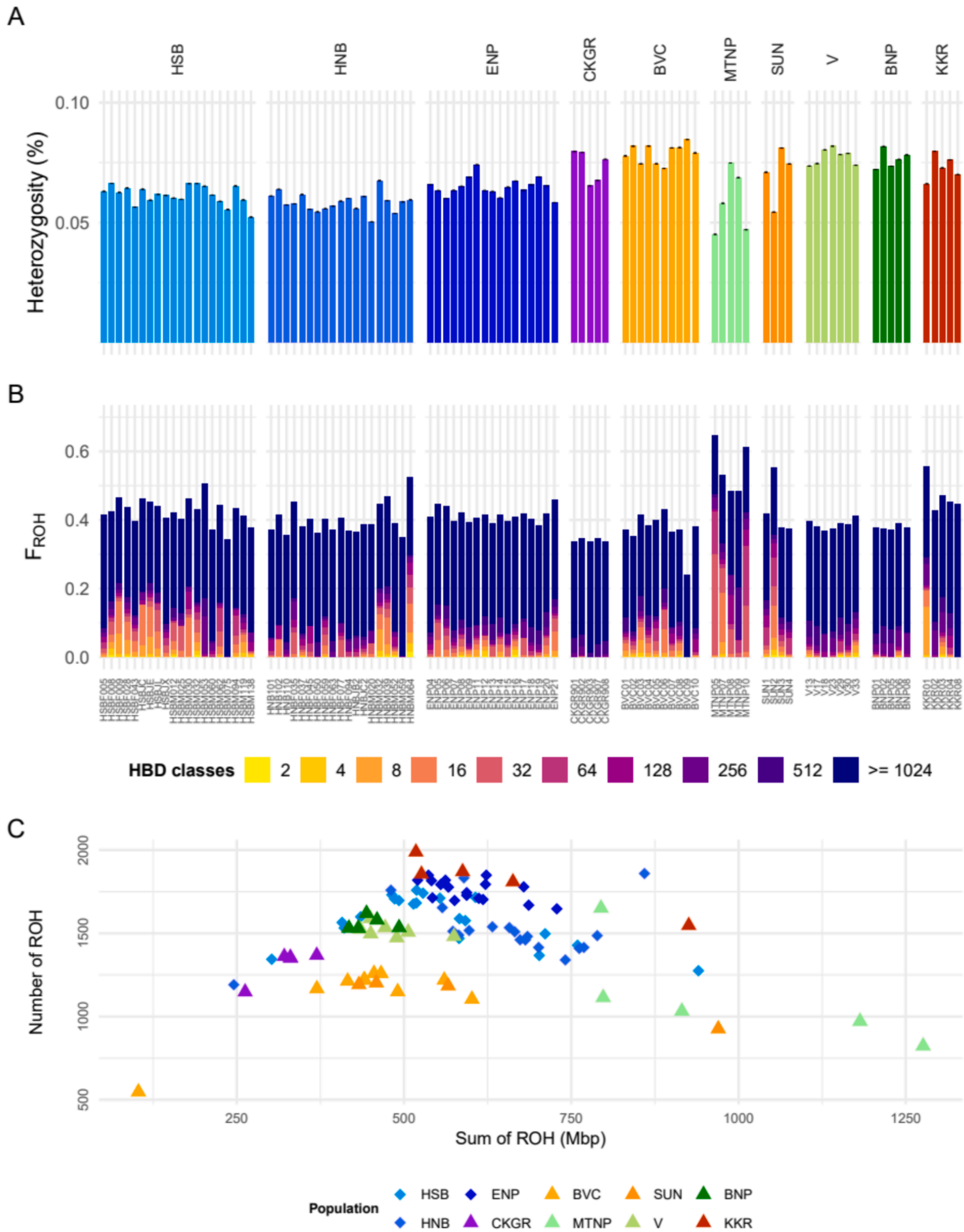

Fig. 6. RoH and heterozygosity analysis. A Heterozygosity in percent per individual, separated by populations. Heterozygosity is similar across all populations and subspecies, with the desert-dwelling HSB and HNB Angolan giraffe populations showing only slightly lower heterozygosity. **B** Runs of homozygosity (realized Inbreeding coefficients) per individual and population. HBD classes are roughly equal to twice the number of ancestors associated with them. Higher classes (≥1024) correspond to shorter RoH stretches, while lower classes correspond to longer stretches. **C** Number and sum of RoH lengths in Mbps per individual and population.

African giraffe, and (b) individuals from this location with this unique genotype should not be the first choice to augment other southern giraffe populations in the future. However, we hesitate to label them hybrids, as this admixture likely was a natural, long-term process and calling them 'hybrids' may hinder future conservation efforts.

While natural migration and large-scale movements of giraffe individuals do occur (Brown and Bolger, 2020; Flanagan et al., 2016), the possibly undocumented, albeit highly unlikely, giraffe translocation(s) into the CKGR may also have resulted in the mixing of different genotypes. The biogeographic patterns presented by Lorenzen et al. (2012) place the central occurrence of the CKGR between the southern and southwestern regions. This coincides with the occurrence of the Angolan and the South African giraffe, while there is no current barrier to their dispersal, although Bock et al. (2014) suggest that this may have been the case in the past. Regardless, the CKGR appears to be a melting pot of the two southern giraffe subspecies.

The maternally inherited Angolan mtDNA genotype preserved in the CKGR and BVC may stem from an ancestral, natural migration of Angolan giraffe eastwards, which may have genetically swamped the South African giraffe due to male dispersal from other regions. The CKGR is located south of major geographic depressions (Caprivi Depression, Makgadikgadi Basin, Mababe Basin) that are characterized by temporal flooding and shallow mega-lake formation during the Late Pleistocene (Riedel et al., 2014). It is conceivable that the disappearance of the mega-lake led to giraffe moving into the Kalahari region. This allowed the Angolan giraffe, which is currently more westerly distributed, to have met and interbred with South African giraffe, leading to the admixture we see today.

The disappearance of the mega-lake allowed giraffe populations of both subspecies to migrate into the Kalahari region. The persistent, homogenous, and high proportion of Angolan giraffe genotype among all five CKGR individuals may be evidence of such a scenario. The geographically wedge-shaped distribution of the Angolan mitochondrial genome coincides with the cryptic rift valley, as suggested in Bock et al. (2014) and the above-mentioned mega lake. The cryptic rift valley extends from Namibia into Botswana and Zimbabwe and splits the Kalahari Desert in Botswana into two separated zones, which act as a barrier in the north and the Limpopo River acting as a barrier in the South. The corridor in between may have facilitated the migration of the Angolan giraffe eastward toward Zimbabwe and the introduction of the Angolan mtDNA genotype.

The persistence of the matrilineal mitochondrial genotype, especially in BVC and CKGR, might be attributed to the strong philopatry in giraffe, where females tend to remain at their birthplace while males migrate over greater distances in search of mates (Bock et al., 2014; Bond et al., 2021; Van der Waal et al., 2014; Winter et al., 2018).

In other South African giraffe, pheno- and genotype populations that are geographically close to the current Angolan giraffe genotype (BNP, V), any accidental admixture with Angolan may have been lost due to backcrossing with *G. g. giraffa*. Alternatively, the areas occupied currently by BNP and V giraffe, are north of the former mega-lake systems, and may have been void of giraffe during flooding and been recolonized almost exclusively by the South African giraffe genotype.

To settle the relationship between the genotype distribution and past geological processes, a wider sampling from across the southern giraffe distribution is required. The additional data will allow modelling of past population structure and migration events among populations in southern Africa.

4.2. Classification of the Angolan giraffe

The Angolan giraffe is divided into two main genotypes, the ENP giraffe and the northern Namib desert-dwelling giraffe (HNB, HSB). It is noteworthy that the genotype compositions of individuals from the two adjacent ephemeral river catchments in northwestern Namibia, ~70 km apart, differ markedly from each other (Fig. 3**A**). For K=3, the HSB individuals show overall more admixture of the genotype found in the ENP (dark blue) and fewer genotypes of the desert-dwelling Angolan giraffe (light blue) found predominantly in the HNB. This is likely a non-intended consequence of a human-facilitated giraffe translocation in 1991 to augment the severely depleted desert-dwelling giraffe of the HSB population with individuals from the ENP (Fennessy, 2004).

The released ENP giraffe individuals subsequently spread their genotype (dark blue) among the then exclusively desert-dwelling giraffe genotype (light blue) as shown in the additional Structure analysis, which numbered as few as one hundred individuals at the time (Fennessy, 2004). The same can be observed in the Dsuite results, which shows a clear signal (10 %) of recent introgression between the HSB and ENP populations, while none can be observed between the ENP and HNB populations. However, this can also be attributed to the software not being able to detect all signals reliably.

The life expectancy of wild giraffe is poorly documented, with Angolan giraffe in northwestern Namibia known to reach > 30 years (J. Fennessy pers. comm.). However, no known individuals translocated from ENP to northwestern Namibia in the 1980 s survive today (Fennessy, 2004).

Interbreeding with ENP individuals has clearly left a major signature of the ENP genotype in the HSB population but not in the HNB population, likely as a result of philopatry. Though we find no evidence of genetic differentiation, we expected a larger impact in the HNB population and speculate this could be due to limited mating between the two populations, especially due to their close geographic proximity and large movement ranges (Fennessy, 2004). While ongoing field studies since the late 1990 s have shown that giraffe do move between the HSB and HNB populations, it is predominantly males that migrate, and only little genetic exchange seems to be taking place (Fennessy, 2009; Hart et al., 2021). In addition, the results explain the presence of the matrilineal ENP mitochondrial genotype in Namibia's desert-dwelling giraffe (Winter et al., 2018), which is derived from introduced ENP females that had mated with desert-dwelling males.

5. Conclusion

To date, less than one percent of giraffe and other wildlife have been genotyped for their entire genome. However, it is evident that traditional genotyping using a single or even a few genetic markers does not truthfully represent the genomic diversity of a population or taxa. In some cases, single locus genotyping has led to misclassification and erroneous assignment of (sub)species and worse, failure to identify distinct taxa. Furthermore, whole-genome sequencing makes it possible to reveal the genetic impact of human-facilitated (giraffe) animal translocations that influence the genotype of local populations. Despite the introgression of the ENP genotype into Namibia's desert-dwelling Angolan giraffe with the aim to augment their numbers, long-term genetic consequences are unknown and unpredictable. The CKGR population warrants further studies on southern giraffe populations to identify more admixed populations within the southern giraffe distribution. Their existence also confirms the need for discussion on how these admixed populations need to be handled in future giraffe conservation management.

We believe our findings also warrant further detailed assessments of giraffe populations across species, subspecies and country-wide populations to better inform future conservation measures.

CRediT authorship contribution statement

David Prochotta: Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Sven Winter:** Writing – review & editing, Visualization, Formal analysis, Conceptualization. **Julian Fennessy:** Writing – review & editing, Resources, Funding acquisition, Conceptualization. **Axel Janke:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Sequence data is available at NCBI through BioProject PRJNA1033849. Genotype datasets are available at Zenodo (https://zenodo.org/records/11280653).

Acknowledgements

We thank an array of partners, government and non-government, across Africa who collaborated with the Giraffe Conservation Foundation (GCF) to permit, collect, and include samples in this analysis. The southern giraffe painting was drawn by J.B. Hlíðberg (www.fauna.is). The Galaxy server that was used is in part funded by Collaborative Research Centre 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and German Federal Ministry of Education and Research (BMBF grants 031 A538A/A538C RBC, 031L0101B/031L0101C de.NBI-epi, 031L0106 de.STAIR (de.NBI)).

Funding

The present study is a product of the Centre for Translational Biodiversity Genomics (LOEWE-TBG) as part of the "LOEWE—Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz" program of Hesse's Ministry of Higher Education, Research, and the Arts as well as the Leibniz Association. Funding and collection of field samples was facilitated by the Giraffe Conservation Foundation and its partners.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2024.108198.

References

Agaba, M., Ishengoma, E., Miller, W.C., McGrath, B.C., Hudson, C.N., Bedoya Reina, O. C., Ratan, A., Burhans, R., Chikhi, R., Medvedev, P., Praul, C.A., Wu-Cavener, L., Wood, B., Robertson, H., Penfold, L., Cavener, D.R., 2016. Giraffe genome sequence reveals clues to its unique morphology and physiology. Nat. Commun. 7, 11519. https://doi.org/10.1038/ncomms11519.

Al Arab, M., Zu Siederdissen, C.H., Tout, K., Sahyoun, A.H., Stadler, P.F., Bernt, M., 2017. Accurate annotation of protein-coding genes in mitochondrial genomes. Mol. Phylogenet. Evol. 106, 209–216.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Bánki, O., Roskov, Y., Döring, M., Ower, G., Hernández Robles, D.R., Plata Corredor, C. A., Stjernegaard Jeppesen, T., Örn, A., Vandepitte, L., Hobern, D., Schalk, P., DeWalt, R.E., Ma, K., Miller, J., Orrell, T., Aalbu, R., Abbott, J., Adlard, R., Aedo, C., et al., 2024. Catalogue of Life (Version 2024–07-18). Catalogue of Life, Amsterdam, Netherlands. https://doi.org/10.48580/dgbqz.

Bock, F., Fennessy, J., Bidon, T., Tutchings, A., Marais, A., Deacon, F., Janke, A., 2014. Mitochondrial sequences reveal a clear separation between Angolan and South African giraffe along a cryptic rift valley. BMC Evol. Biol. 14. https://doi.org/10.1186/s12862-014-0219-7.

Boddaert, P., 1785. Elenchus animalium (Vol. 1) . Hake.

Bond, M.L., Lee, D.E., Ozgul, A., Farine, D.R., König, B., 2021. Leaving by staying: Social dispersal in giraffes. J. Anim. Ecol. 90, 2755–2766. https://doi.org/10.1111/1365-2656.13582.

Broad Institute, 2019. Picard Toolkit. Broad Institute, https://broadinstitute.github.io/picard/.

Brown, M.B., Bolger, D.T., 2020. Male-biased partial migration in a giraffe population. Front. Ecol. Evol. 7.

Brown, M.B., Kulkarni, T., Ferguson, S., Fennessy, S., Muneza, A., Stabach, J.A., Fennessy, J., 2022. Conservation Status of Giraffe: Evaluating Contemporary Distribution and Abundance with Evolving Taxonomic Perspectives. In: DellaSala, D. A., Goldstein, M.I. (Eds.), Imperiled: the Encyclopedia of Conservation. Elsevier, Oxford, pp. 471–487. https://doi.org/10.1016/B978-0-12-821139-7.00139-2.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34. https://doi.org/10.1093/bioinformatics/bty560 i884–i890.

Coimbra, R.T.F., Winter, S., Kumar, V., Koepfli, K.-P., Gooley, R.M., Dobrynin, P., Fennessy, J., Janke, A., 2021. Whole-genome analysis of giraffe supports four distinct species. Curr. Biol. 31, 2929–2938.e5. https://doi.org/10.1016/j.cub.2021.04.033.

Coimbra, R.T.F., Winter, S., Mitchell, B., Fennessy, J., Janke, A., 2022. Conservation genomics of two threatened subspecies of Northern Giraffe: the west African and the kordofan giraffe. Genes 13, 221. https://doi.org/10.3390/genes13020221.

Coimbra, R., Winter, S., Muneza, A., Fennessy, S., Otiende, M., Mijele, D., Masiaine, S., Stacy-Dawes, J., Fennessy, J., Janke, A., 2023. Genomic analysis reveals limited hybridization among three giraffe species in Kenya. BMC Bio. https://doi.org/10.1101/2023.09.12.557229.

Dagg, A.I., 1971. Giraffa camelopardalis. Mamm. Species 1–8. https://doi.org/10.2307/3503830.

Dagg, A.I., Foster, J.B., 1982. The giraffe: its biology, behavior, and ecology. R.E. Krieger Pub, Co, Huntington, N.Y.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. 1000 genomes project analysis group the variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. https://doi.org/10.1093/gigascience/giab008.

de Jong, M.J., de Jong, J.F., Hoelzel, A.R., Janke, A., 2021. SambaR: an R package for fast, easy and reproducible population-genetic analyses of biallelic SNP data sets. Mol. Ecol. Resour. 21, 1369–1379. https://doi.org/10.1111/1755-0998.13339.

de Jong, M.J., Niamir, A., Wolf, M., Kitchener, A.C., Lecomte, N., Seryodkin, I.V., Fain, S. R., Hagen, S.B., Saarma, U., Janke, A., 2023. Range-wide whole-genome resequencing of the brown bear reveals drivers of intraspecies divergence. Commun Biol 6, 1–16. https://doi.org/10.1038/s42003-023-04514-w.

Deckmyn, O.S. code by R.A.B. and A.R.W.R. version by R.B.E. by T.P.M. and A., 2018. maps: Draw Geographical Maps.

Donath, A., Jühling, F., Al-Arab, M., Bernhart, S.H., Reinhardt, F., Stadler, P.F., Middendorf, M., Bernt, M., 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Res. 47, 10543–10552. https://doi.org/10.1093/nar/gkz833.

Druet, T., Kadri, N.K., Gautier, A.B. and M., 2021. RZooRoH: Partitioning of Individual Autozygosity into Multiple Homozygous-by-Descent Classes.

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

Farré, M., Li, Q., Darolti, I., Zhou, Y., Damas, J., Proskuryakova, A.A., Kulemzina, A.I., Chemnick, L.G., Kim, J., Ryder, O.A., Ma, J., Graphodatsky, A.S., Zhang, G., Larkin, D.M., Lewin, H.A., 2019. An integrated chromosome-scale genome assembly of the Masai giraffe (Giraffa camelopardalis tippelskirchi). GigaScience 8, giz090. https://doi.org/10.1093/gigascience/giz090.

Fennessy, J., 2009. Home range and seasonal movements of *Giraffa camelopardalis angolensis* in the northern Namib Desert. Afr. J. Ecol. 47, 318–327. https://doi.org/10.1111/j.1365-2028.2008.00963.x.

Fennessy, J., Bidon, T., Reuss, F., Kumar, V., Elkan, P., Nilsson, M.A., Vamberger, M., Fritz, U., Janke, A., 2016. Multi-locus analyses reveal four giraffe species instead of one. Curr. Biol. 26, 2543–2549. https://doi.org/10.1016/j.cub.2016.07.036.

Fennessy, J.T., 2004. Ecology of desert-dwelling giraffe Giraffa camelopardalis angolensis in northwestern Namibia (Thesis).

FHCRC, 2024. https://github.com/fhcrc/seqmagick.

Flanagan, S.E., Brown, M.B., Fennessy, J., Bolger, D.T., 2016. Use of home range behaviour to assess establishment in translocated giraffes. Afr. J. Ecol. 54, 365–374. https://doi.org/10.1111/aje.12299.

Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., Smit, A.F., 2020. RepeatModeler2 for automated genomic discovery of transposable element families. PNAS 117, 9451–9457. https://doi.org/10.1073/pnas.1921046117.

Giraffe Conservation Foundation, 2022. Country Profile: Republic of Zambia– Giraffe Conservation Status Report.

Fox, E.A., Wright, A.E., Fumagalli, M., Vieira, F.G., 2019. ngsLD: evaluating linkage disequilibrium using genotype likelihoods. Bioinformatics 35, 3855–3856. https://doi.org/10.1093/bioinformatics/btz200.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T.F., Conesa, A., 2012. Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 28, 2678–2679. https://doi.org/10.1093/bioinformatics/bts503.

Garnier, S., Ross, N., Rudis, B., Sciaini, M., Camargo, A.P., Scherer, C., 2021. viridis: Colorblind-Friendly Color Maps for R.

Goudie, A., 2010. Namib Sand Sea: Large Dunes in an Ancient Desert. In: Migon, P. (Ed.), Geomorphological Landscapes of the World. Springer, Netherlands, Dordrecht, pp. 163–169. https://doi.org/10.1007/978-90-481-3055-9_17.

Hailer, F., Kutschera, V.E., Hallström, B.M., Klassert, D., Fain, S.R., Leonard, J.A., Arnason, U., Janke, A., 2012. Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science 336, 344–347. https://doi.org/10.1126/science.1216424.

Hart, E.E., Fennessy, J., Wells, E., Ciuti, S., 2021. Seasonal shifts in sociosexual behaviour and reproductive phenology in giraffe. Behav. Ecol. Sociobiol. 75, 15. https://doi.org/10.1007/s00265-020-02954-6.

Hassanin, A., Ropiquet, A., Gourmand, A.-L., Chardonnet, B., Rigoulet, J., 2007. Mitochondrial DNA variability in *Giraffa camelopardalis*: consequences for taxonomy, phylogeography and conservation of giraffes in West and central Africa. C. r. Biol. 330, 265–274. https://doi.org/10.1016/j.crvi.2007.02.008.

Hausen, N., 2017. Computer-aided photographic pelage pattern analysis of Giraffa camelopardalis (Artiodactyla:Girrafidae). University of London [Unpublished Master's thesis].

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.
ITIS, 2024. Integrated Taxonomic Information System (ITIS) [WWW Document]. URL http://www.itis.gov (accessed 3.14.24).
Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T.P., Sado, T., Mabuchi, K., Takeshima, H., Miya, M., Nishida, M., 2013. MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Mol. Biol. Evol. 30, 2531–2540. https://doi.org/10.1093/molbev/mst141.
Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C.W., Yi, T.-S., Li, D.-Z., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21, 241. https://doi.org/10.1186/s13059-020-02154-5.
Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129.
Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btr521.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.
Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., 2014. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2, e281.
Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066. https://doi.org/10.1093/nar/gkf436.
Korneliussen, T.S., Albrechtsen, A., Nielsen, R., 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinf. 15, 356. https://doi.org/10.1186/s12859-014-0356-4.
Kück, P., Meusemann, K., 2010. FASconCAT: convenient handling of data matrices. Mol. Phylogenet. Evol. 56, 1115–1118. https://doi.org/10.1016/j.ympev.2010.04.024.
Leal, T.P., Furlan, V.C., Gouveia, M.H., Duarte, J.M.S., Fonseca, P.A., Tou, R., de Oliveira Scliar, M., de Araujo, G.S., Costa, L.F., Zolini, C., Peixoto, M.G.C.D., 2022. NAToRA, a relatedness-pruning method to minimize the loss of dataset size in genetic and omics analyses. Comput. Struct. Biotechnol. J. 20, 1821–1828. https://doi.org/10.1016/j.csbj.2022.04.009.
Leigh, J.W., Bryant, D., 2015. popart: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410.
Lesson, R.P., 1842. Nouveau Tableau Du Règne Animal: Mammifères. Arthus Bertrand.
Li, J.-Y., Li, W.-X., Wang, A.-T., Zhang, Y., 2021. MitoFlex: an efficient, high-performance toolkit for animal mitogenome assembly, annotation and visualization. Bioinformatics 37, 3001–3003. https://doi.org/10.1093/bioinformatics/btab111.
Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. doi: 10.48550/arXiv.1303.3997.
Li, H., 2023. https://github.com/lh3/seqtk/.
Lorenzen, E.D., Heller, R., Siegismund, H.R., 2012. Comparative phylogeography of African savannah ungulates1. Mol. Ecol. 21, 3656–3670. https://doi.org/10.1111/j.1365-294X.2012.05650.x.
Lydekker, R., 1903. Local variation in the Giraffe. In: Animal Life and the World of Nature. Hutchinson & Co., London, pp. 78–84.
Lydekker, R., 1904. On the Subspecies of Giraffa camelopardalis. Proceedings of the Zoological Society of London 74, 202–229. doi: 10.1111/j.1469-7998.1904.tb08288.x.
Malinsky, M., Matschiner, M., Svardal, H., 2021. Dsuite - Fast D-statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. 21, 584–595. https://doi.org/10.1111/1755-0998.13265.
Mammal Diversity Database, 2024. Mammal Diversity Database. doi: 10.5281/ZENODO.10595931.
Massicotte, P., South, A., 2023. rnaturalearth: World Map Data from Natural Earth.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110.
Meisner, J., Albrechtsen, A., 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics 210, 719–731. https://doi.org/10.1534/genetics.118.301336.
Meng, G., Li, Y., Yang, C., Liu, S., 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 47, e63–e. https://doi.org/10.1093/nar/gkz173.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.
Neuwirth, E., 2022. RColorBrewer: ColorBrewer Palettes. doi: 10.32614/CRAN.package.RColorBrewer.
Noack, Th., 1908. Die Giraffe Des Sambesi-Gebietes. Zoologischer Anzeiger 33, 354–356.
O'Connor, D., Stacy-Dawes, J., Muneza, A., Fennessy, J., Gobush, K., Chase, M.J., Brown, M.B., Bracis, C., Elkan, P., Zaberirou, A.R.M., Rabeil, T., Rubenstein, D., Becker, M.S., Phillips, S., Stabach, J.A., Leimgruber, P., Glikman, J.A., Ruppert, K., Masiaine, S., Mueller, T., 2019. Updated geographic range maps for giraffe, Giraffa spp., throughout sub-Saharan Africa, and implications of changing distributions for conservation. Mammal Rev. 49, 285–299. https://doi.org/10.1111/mam.12165.
Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528.
Pebesma, E., Bivand, R., Racine, E., Sumner, M., Cook, I., Keitt, T., Lovelace, R., Wickham, H., Ooms, J., Müller, K., Pedersen, T.L., Baston, D., 2020. sf: Simple Features for R.
Pedersen, T.L., 2022. patchwork: The Composer of Plots. doi: 10.32614/CRAN.package.patchwork.
Petzold, A., Hassanin, A., 2020. A comparative approach for species delimitation based on multiple methods of multi-locus DNA sequence analysis: a case study of the genus Giraffa (Mammalia, Cetartiodactyla). PLoS One 15, e0217956.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. https://doi.org/10.1086/519795.
Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033.
R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.
Reardon, M., 1986. The besieged desert- war, drought, poaching in the namib desert. HarperCollins Publishers.
Riedel, F., Henderson, A.C.G., Heußner, K.-U., Kaufmann, G., Kossler, A., Leipe, C., Shemang, E., Taft, L., 2014. Dynamics of a Kalahari long-lived mega-lake system: hydromorphological and limnological changes in the Makgadikgadi Basin (Botswana) during the terminal 50 ka. Hydrobiologia 739, 25–53. https://doi.org/10.1007/s10750-013-1647-x.
Schliep, K.P., 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27, 592–593. https://doi.org/10.1093/bioinformatics/btq706.
Seymour, R., 2001. Patterns of subspecies diversity in the giraffe, Giraffa camelopardalis (L. 1758): comparison of systematic methods and their implications for conservation policy (phd). University of Kent. https://doi.org/10.22024/UniKent/01.02.86053.
Shen, W., Le, S., Li, Y., Hu, F., 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One 11, e0163962.
Shipham, A., Schmidt, D.J., Joseph, L., Hughes, J.M., 2015. Phylogenetic analysis of the Australian rosella parrots (Platycercus) reveals discordance among molecules and plumage. Mol. Phylogenet. Evol. 91, 150–159. https://doi.org/10.1016/j.ympev.2015.05.012.
Smit, A., Hubley, R., Green, P., 2020. RepeatMasker Open-4.1.1.
Soetaert, K., 2024. plot3D: Plotting Multi-Dimensional Data. doi: 10.32614/CRAN.package.plot3D.
Steenwyk, J.L., Iii, T.J.B., Li, Y., Shen, X.-X., Rokas, A., 2020. ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. PLoS Biol. 18, e3001007.
Tange, O., 2011. GNU parallel - the command-line power tool. The USENIX Magazine 36, 42–47.
Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., Prins, P., 2015. Sambamba: fast processing of NGS alignment formats. Bioinformatics 31, 2032–2034. https://doi.org/10.1093/bioinformatics/btv098.
The Galaxy Community, 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Research 50, W345–W351. doi: 10.1093/nar/gkac247.
Van der Waal, K.L., Wang, H., McCowan, B., Fushing, H., Isbell, L.A., 2014. Multilevel social organization and space use in reticulated giraffe (Giraffa camelopardalis). Behav. Ecol. 25, 17–26. https://doi.org/10.1093/beheco/art061.
Vieira, F.G., 2024. prune_graph [WWW Document]. URL https://github.com/fgvieira/prune_graph (accessed 3.20.24).
Wang, L.-G., Lam, T.-T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C.W., Jones, B.R., Bradley, T., Zhu, H., Guan, Y., Jiang, Y., Yu, G., 2020. Treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. Mol. Biol. Evol. 37, 599–603. https://doi.org/10.1093/molbev/msz240.
Wickham, H., 2007. Reshaping data with the reshape package. J. Stat. Softw. 21, 1–20. https://doi.org/10.18637/jss.v021.i12.
Wickham, H., 2011. ggplot2. Wires Comput. Stat. 3, 180–185. https://doi.org/10.1002/wics.147.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. J. Open Source Soft. 4, 1686. https://doi.org/10.21105/joss.01686.
Winter, S., Fennessy, J., Fennessy, S., Janke, A., 2018. Matrilineal population structure and distribution of the Angolan giraffe in the Namib desert and beyond. Ecol. Genetics Genomics 7–8, 1–5. https://doi.org/10.1016/j.egg.2018.03.003.
Yu, G., Lam, T.T.-Y., Xu, S., Li, L., Jones, B., Silverman, J., Iwasaki, W.M., Xia, Y., Huang, R., 2023. ggtree: an R package for visualization of tree and annotation data. doi: 10.18129/B9.bioc.ggtree.
Zhou, W.-W., Wen, Y., Fu, J., Xu, Y.-B., Jin, J.-Q., Ding, L., Min, M.-S., Che, J., Zhang, Y.-P., 2012. Speciation in the Rana chensinensis species complex and its relationship to the uplift of the Qinghai-Tibetan Plateau. Mol. Ecol. 21, 960–973. https://doi.org/10.1111/j.1365-294X.2011.05411.x.
Zhu, T., Sato, Y., Sado, T., Miya, M., Iwasaki, W., 2023. MitoFish, MitoAnnotator, and MiFish pipeline: updates in 10 years. Mol. Biol. Evol. 40, msad035. https://doi.org/10.1093/molbev/msad035.