Molecular Phylogenetics and Evolution 198 (2024) 108131

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Complex patterns of gene flow and convergence in the evolutionary history of the spiral-horned antelopes (Tragelaphini)

Andrinajoro R. Rakotoarivelo [a,b], Thabelo Rambuda [a,c], Ulrike H. Taron [d], Gabrielle Stalder [e], Paul O'Donoghue [f], Jan Robovský [g], Stefanie Hartmann [d], Michael Hofreiter [d], Yoshan Moodley [a,*]

[a] Department of Biological Sciences, University of Venda, Private Bag X5050, Thohoyandou 0950, Republic of South Africa
[b] Department of Zoology and Entomology, University of the Free State: QwaQwa Campus, Private Bag X13, Phuthaditjhaba 9866, Republic of South Africa
[c] Department of Genetics, University of Pretoria, Private Bag X20, Hatfield 0028, Republic of South Africa
[d] Evolutionary Adaptive Genomics, Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany
[e] Department of Interdisciplinary Life Sciences, Research Institute of Wildlife Ecology, University of Veterinary Medicine, Savoyenstraße 1, A-1160 Wien, Austria
[f] Specialist Wildlife Services, St Asaph, United Kingdom
[g] Department of Zoology, Faculty of Science, University of South Bohemia, Branišovská 1760, 37005 České Budějovice, Czech Republic

ARTICLE INFO

Keywords:
Divergence
Gene flow
Convergent evolution
Tragelaphus

ABSTRACT

The Tragelaphini, also known as spiral-horned antelope, is a phenotypically diverse mammalian tribe comprising a single genus, *Tragelaphus*. The evolutionary history of this tribe has attracted the attention of taxonomists and molecular geneticists for decades because its diversity is characterised by conflicts between morphological and molecular data as well as between mitochondrial, nuclear and chromosomal DNA. These inconsistencies point to a complex history of ecological diversification, coupled by either phenotypic convergence or introgression. Therefore, to unravel the phylogenetic relationships among spiral-horned antelopes, and to further investigate the role of divergence and gene flow in trait evolution, we sequenced genomes for all nine accepted species of the genus *Tragelaphus*, including a genome each for the highly divergent bushbuck lineages (*T. s. scriptus* and *T. s. sylvaticus*). We successfully reconstructed the *Tragelaphus* species tree, providing genome-level support for the early Pliocene divergence and monophyly of the nyala (*T. angasii*) and lesser kudu (*T. imberbis*), the monophyly of the two eland species (*T. oryx* and *T. derbianus*) and, importantly, the monophyly of kéwel (*T. s. scriptus*) and imbabala (*T. s. sylvaticus*) bushbuck. We found strong evidence for gene flow in at least four of eight nodes on the species tree. Among the six phenotypic traits assessed here, only habitat type mapped onto the species tree without homoplasy, showing that trait evolution was the result of complex patterns of divergence, introgression and convergent evolution.

1. Introduction

The idea that speciation is a purely bifurcating process in which one species splits into two distinct entities that become reproductively isolated (Mayr, 1963; Coyne and Orr, 2004) is no longer accepted as the norm. Molecular markers in evolutionary studies have provided evidence of hybridization and introgression both between sister and non-sister species (Dowling and Secor, 1997; Arnold, 2006). Advances in DNA sequencing methods and genotyping have already had an impact on our understanding of the complexity of species boundaries (Salazar et al., 2010; Cruickshank and Hahn, 2014; Harrison and Larson, 2014; Gante et al., 2016; Masello et al., 2019). There is also an increasing awareness that interspecific gene flow could play an important role in adaptive radiations, with numerous evolutionary implications (Abbott et al., 2013). Although hybridization can result in species fusion, it has been proposed that interspecific gene flow can also fuel rapid diversification by providing a plentiful source of potentially adaptive variation (Seehausen, 2004; Mallet, 2007; Berner and Salzburger, 2015; Figueiró et al., 2017, Rakotoarivelo et al., 2019a).

Introgression, however, is difficult to detect using single locus markers because differentiating it from allele sharing due to common ancestry, also known as incomplete lineage sorting (ILS), remains a

* Corresponding author.
E-mail address: yoshan.moodley@univen.ac.za (Y. Moodley).

https://doi.org/10.1016/j.ympev.2024.108131
Received 26 December 2023; Received in revised form 19 May 2024; Accepted 15 June 2024
Available online 22 June 2024

major challenge (Holder et al., 2001; Abbott et al., 2013). This problem is compounded in taxa where divergences among lineages were rapid, making them particularly prone to ILS (McPeek, 2008; Wellborn and Langerhans, 2015; Ford et al., 2016). Phenotypic divergence usually develops much faster than reproductive isolation during rapid radiations, allowing for adaptive introgression and hybrid speciation among phenotypically distinct lineages (Masello et al., 2019). Recent developments in genome-wide typing and whole genome sequencing have provided unparalleled resolution for identifying polymorphisms associated with niche adaptation or characterizing patterns of interspecific gene flow (Dasmahapatra et al., 2012, Racimo et al., 2015). Distinguishing introgression from ILS is only possible when large numbers of loci have been sequenced. This is because ILS and gene flow influence the genome differently (Zheng and Janke, 2018), and so a sample of phylogenetic trees reconstructed from across the genome, as well as their branch lengths, could hold clues to their relative influence on evolutionary history (Edelman et al., 2019). Methods have now been developed to identify both the presence (Durand et al., 2011) and direction (Pease and Hahn, 2015) of gene flow, which have been successful in identifying introgression, adaptive introgression, and hybrid speciation, while also distinguishing these reticulate events from ILS (Martin et al., 2013; Liu et al., 2015; Malinsky et al., 2015; Norris et al., 2015; Hibbins and Hahn, 2022).

One of the most successful mammalian adaptive radiations was that of the Bovidae, a family of even-toed ungulates that were able to adapt to changing habitats, radiating into many tribes and becoming one of the most diverse and abundant of mammalian herbivore families (Du Toit and Cumming, 1999) that now comprise some 142 (IUCN SSC Antelope Specialist Group, 2016) or 154 (American Society of Mammalogy Mammal Diversity Database, Burgin et al., 2022) extant species. Among the Bovidae, the Tragelaphini, a monophyletic African tribe that appears in the fossil record during the late Miocene (Gentry, 2010, Bibi, 2013), is characterised by high morphological diversity. Also known as the spiral-horned antelopes or tragelaphines, this tribe is ecologically diverse, with its nine extant species having colonized, except for deserts, nearly every habitat type in sub-Saharan Africa (Fig. 1), from dry scrub and woodlands (*Tragelaphus imberbis*, *T. angasii, T. scriptus*) to open savanna woodland (*T. oryx*, *T. derbianus*, *T. strepsiceros, T. scriptus*) and moist lowland Congolian and montane forests (*T. buxtoni*, *T. scriptus*, *T. eurycerus* and *T. spekii*).

The evolutionary relationships among the extant species of *Tragelaphus* have always been controversial. Morphologically, the tribe was originally classified into three genera, but allozyme and mitochondrial DNA (mtDNA) data reduced these to the single genus *Tragelaphus* (Essop et al., 1997; Gatesy et al., 1997; Georgiadis et al., 1990; Matthee and Robinson, 1999). Further taxonomic conflict based on morphology arose when the first mtDNA sequencing studies revealed polyphyletic relationships between phenotypically similar spiral-horned kudu-like (*T. strepsiceros*, *T. imberbis, T. oryx* and *T. derbianus*) and lyrate-horned nyala-like (*T. angasii*, *T. buxtoni*. *T. spekii, T. scriptus*) taxa, and between species in which females possess horns (*T. oryx*, *T. derbianus*, *T. eurycerus*) (Hassanin and Douzery, 1999; Matthee and Robinson, 1999; Willows-Munro et al., 2005). More slowly evolving nuclear intron sequences largely corroborated the alternative evolutionary history based on mtDNA, thus favouring convergence over ILS, as the basis for the morpho-molecular conflict (Willows-Munro et al., 2005, Rakotoarivelo et al., 2019b). In general, the molecular evidence supported an early divergence of the arid-adapted lesser kudu (*T. imberbis*) and nyala (*T. angasii*), although, depending on DNA marker, these alternate between being recovered either as sister groups or as independently diverging basal taxa (Willows-Munro et al., 2005; Matthee and Robinson, 1999). The remaining seven species were divided into a dry-savanna clade consisting of common and giant eland (*T. oryx* and *T. derbianus*, respectively) and greater kudu (*T. strepsiceros*), and a closed-forest clade comprising mountain nyala (*T. buxtoni*), bushbuck (*T. scriptus*), bongo (*T. eurycerus*) and sitatunga (*T. spekii*) (Hassanin et al., 2018; Rakotoarivelo et al., 2019b). However, the greater kudu is also sometimes placed outside this clade, again depending on marker and analysis. Lastly, a ruminant-scale phylogeny was reconstructed from whole genome sequences for seven of the nine *Tragelaphus* species and corroborated the groupings above, placing the greater kudu as sister to the common eland (Chen et al., 2019). Nevertheless, overall and despite some mito-nuclear conflict, molecular analyses suggested an evolutionary history driven by ecology, with dry- or arid-adapted species being basal, and the ability to inhabit wetter forests being derived. However, the genome-scale study mentioned above did not include the basal *T. angasii* nor *T. derbianus*, nor did it provide analyses on interspecific gene flow.

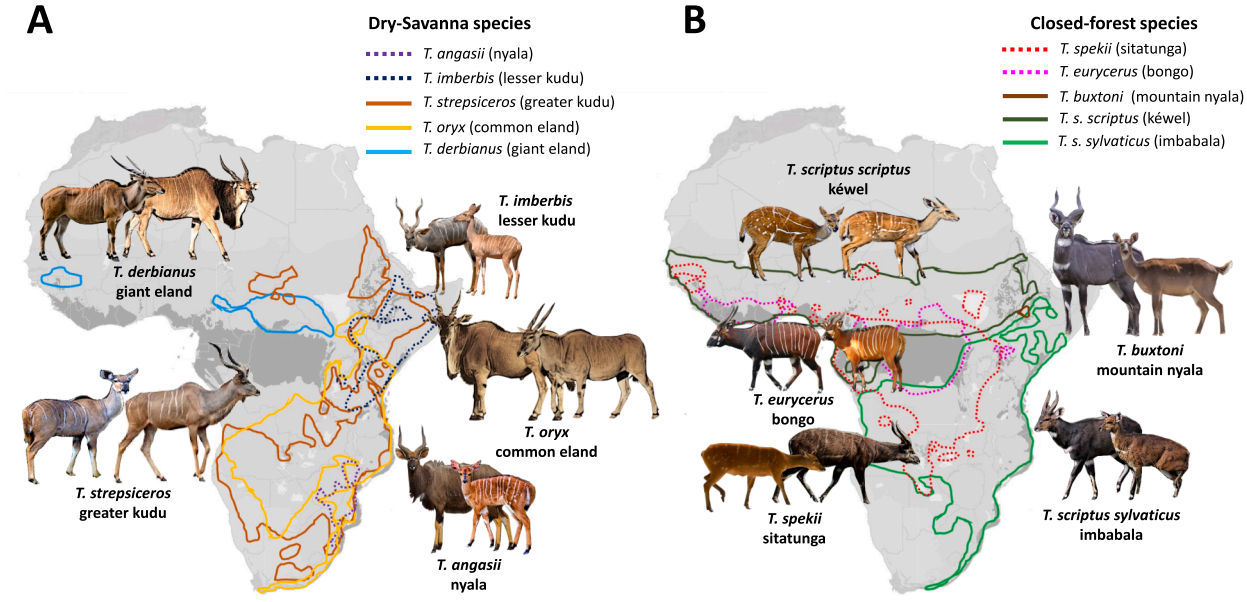

Fig. 1. Distribution and phenotypic variation among *Tragelaphus* species across sub-Saharan Africa. For ease of representation, arid-adapted and dry-savanna species are shown in panel A, and closed-forest species in panel B. Both sexes are portrayed for all species to highlight sexual dimorphism. Both kéwel (*T. scriptus scriptus*) and imbabala (*T. s. sylvaticus*) forms of the bushbuck are shown.

Moreover, not all mito-nuclear discrepancies among the Tragelaphini could be reconciled by hypotheses of ILS or convergence. The type species for this genus, the bushbuck (*T. scriptus*), is the smallest and most adaptable of the spiral-horned antelopes, occurring in several ecoregions from the coast to 4,000 m above sea level, and overlapping in distribution with every other *Tragelaphus* species. Despite being the most widespread of all African bovid species, bushbuck specimens in the studies mentioned above were sampled only in South and/or East Africa. In contrast, a pan-African mtDNA analysis of bushbuck revealed a major split between the harnessed bushbuck or kéwel (*T. s. scriptus*) of West and central Africa and the sylvan bushbuck or imbabala (*T. s. sylvaticus*) from East and southern Africa (Moodley and Bruford, 2007), with the *T. s. scriptus* lineage being most closely related to *T. angasii*, rather than to *T. s. sylvaticus* (Moodley et al., 2009). This mtDNA polyphyly of the bushbuck contrasts dramatically with the monophyly of *T. s. scriptus* and *T. s. sylvaticus* at more slowly evolving nuclear introns (Rakotoarivelo et al., 2019a, Hassanin et al., 2018), providing the "smoking gun" for an ancient hybridisation event between proto-scriptus bushbuck and a now-extinct *Tragelaphus* species, closely related to the nyala. Intriguingly, this Pliocene hybridisation event also resulted in a diploid (2n) chromosome number of 57/58 (males/females) for *T. s. scriptus* (Hassanin et al., 2018), which is similar to that of *T. angasii* (55/56) and different to its sister taxon *T. s. sylvaticus* (33/34). The difference in karyotype between sexes among most tragelaphine species is the result of an ancestral Y-autosomal translocation, and widely varying chromosome number among species (2n: 30 – 57/58) is thought to play a role in reproductive isolation (Rubes et al., 2008). However, the difference in chromosome number between *T. s. scriptus* and *T. s. sylvaticus* does not appear to have hindered gene flow along Africa's rift valley, where the two kinds of bushbuck come into secondary contact, leading to populations of intermediate phenotype (Rakotoarivelo et al., 2019a, Moodley and Bruford, 2007). The detection of ancient interspecific gene flow among the spiral-horned antelopes (Rakotoarivelo et al., 2019a, Hassanin et al., 2018) could provide an alternative explanation for the high prevalence of seemingly convergent phenotypes within this tribe.

In this study we provide the first whole genome data that includes representatives of all species of the genus *Tragelaphus,* plus representatives of both *T. s. scriptus* and *T. s. sylvaticus* forms of the bushbuck. This data set allowed us to test the hypotheses that *T. imberbis* and *T. angasii* are the two basal-most species, that the bushbuck is monophyletic with respect to its nuclear genome, and that gene flow was prevalent during the evolutionary history of the spiral-horned antelopes.

2. Methods

2.1. Taxon sampling

A total of ten samples, one each from the *T. s. scriptus* and *T. s. sylvaticus* bushbucks, as well as one from each of the remaining eight *Tragelaphus* species, were included in this study. Zoo samples (Table 1) were not collected specifically for this study but were taken by respective authorities during general post-mortem examinations or necessary medical interventions and biobanked. They were subsequently provided for use in this study. In all such cases procedures were in line with the national legislation and the institutional ethics and animal welfare guidelines of the different institutions. All other samples were obtained

Table 1
Geographic distribution, sample origin, short read mapping and genome coverage among sampled spiral-horned antelopes.

Species name	Common name (Taxonomic Reference)	Geographic distribution (also see Fig. 1)	Sample origin (sample type)	Voucher	Raw Reads	Autosomal		Mitochondrial		Accession
						Mapped Reads	Coverage (x)	Mapped Reads	Coverage (x)	
Tragelaphus angasii	Nyala (Angas, 1849)	South-east Africa	Zoo, Linzer Tiergarten, Austria (liver)	Z/1152/11	189 M	154 M	9	125 K	1046	SAMN41052117
T. imberbis	Lesser kudu (Blyth, 1869)	Northeast Africa	Zoo, Hannover Zoo, Germany (blood)	XV 2267, 4.4.11, 1892	260 M	209 M	13	36 K	289	SAMN41052118
T. strepsiceros	Greater kudu (Pallas, 1766)	Central, east, and southern Africa	Zoo, La Palmyre Zoo, France (blood)	Koudou #6207	177 M	140 M	9	17 K	143	SAMN41052119
T. oryx	Common eland (Pallas, 1766)	East and southern sub-Saharan Africa	Zoo, Schönbrunn Zoo, Austria (liver)	CB/88/14	159 M	137 M	7	291 K	2266	SAMN41052120
T. derbianus	Giant eland (Gray, 1847)	West and northern sub-Saharan Africa	Wild, Central African Republic, Chinko Project (skin)	R14_0152, 1958/2014-1070399	196 M	177 M	9	246 K	1933	SAMN41052121
T. spekii	Sitatunga (Speke, 1863)	West, central, and east Africa	Zoo, Prague Zoo, Czech Republic (skin)	365 II, F365 9.10.05	206 M	174 M	10	142 K	1140	SAMN41052122
T. buxtoni	Mountain nyala (Lydekker, 1910)	Ethiopian highlands	Wild, Ethiopia, Powell Cotton Museum (skin)	Abyssinia II 15.	234 M	181 M	12	21 K	176	SAMN41052123
T. eurycerus	Bongo (Ogilby, 1837)	Central Africa	Zoo, Port Lympne Zoo, United Kingdom (skin)	B42 II, M42 19.01.06	165 M	140 M	8	201 K	1667	SAMN41052124
T. scriptus scriptus	Kéwel (Pallas, 1766) (Harnessed bushbuck)	Northern sub-Saharan Africa	Wild, Ghana, University of Copenhagen (skin)	4842	195 M	160 M	8	217 K	1694	SAMN41052125
T. s. sylvaticus	Imbabala (Sparrman, 1780) (Sylvan bushbuck)	East and southern Africa	Wild, Eastern Cape, South Africa, Taxidermy Africa (skin)	EC11	152 M	122 M	8	32 K	279	SAMN41052126

M, millions of reads; K, thousands of reads.

from existing collections at museums or taxidermists (Table 1). Although seven other genome samples were potentially available from the data set of Chen et al. (2019), genome quality was low for *T. oryx*, *T. buxtoni* and *T. euryceros* and the genome of *T. spekii* (GCA_006411015) was contaminated. The addition of the remaining three available genomes (*T. imberbis*, *T. strepsiceros* and *T. scriptus*) to this study would have resulted in an imbalanced data set, potentially leading to biases in the estimation of evolutionary rates, without improving our ability to estimate population allele frequencies. Lastly, it was not clear from the available metadata whether the available *T. scriptus* genome was a kéwel or an imbabala.

2.2. Sequencing, raw data processing and mapping

DNA was extracted from the ten spiral-horned antelope samples (Table 1) according to Rohland *et al.* (2010) in a restricted-access, ancient DNA laboratory. Prior to library preparation, DNA extracts were diluted (concentration: 15 – 27 ng/μl) using a Qubit 2.0 Fluorometer (ThermoFisher) with a broad range dsDNA kit. DNA was then sheared using a Covaris S200 sonicator (Covaris Inc.) in 130 μl microTUBE cuvettes following the manufacturer's instructions. Average fragment length after shearing was 577 bp when measured on an Agilent TapeStation 2200 System (Agilent Technologies) using the DNA D1000 ScreenTape assay. Double stranded, double-indexed (dsDNA) libraries were prepared according to Henneberger et al. (2019). Successful library preparation was verified using TapeStation (high sensitivity D1000 ScreenTape) and Qubit (high sensitivity dsDNA) assays. All individual dsDNA libraries were re-amplified to improve library quantity and filtered for fragment size using a Pippin Prep (Sage science) and 2 % agarose gel cassettes (Marker B, no overflow detection) following the manufacturer's instructions. Successful size selection was verified using a TapeStation DNA D1000 ScreenTape assay. The size selected libraries had an average length of 541 bp. Paired-end sequencing was done on an Illumina NextSeq 500 system using 75 or 150 bp libraries.

We used Cutadapt v.1.8.1 (Martin, 2011) to trim Illumina adapter sequences from the ends of reads and remove reads shorter than 30 bp. Unlike other antelope tribes, the Tragelaphini are sister to the Bovini, and are thus closely related to the domestic cow. We therefore used a publicly available *Bos taurus* genome (GCA_000003055.5, also known as UMD3.1) as reference and outgroup in all analyses. Trimmed reads were mapped to the cow reference genome (available https://bovinegenome.org/?q=node/61) using BWA-MEM v0.7.15 with default parameters (Li and Durbin, 2009). Prior to mapping, mitochondrial (MT NC_006853.1) and X-chromosome (X AC_000187.1) scaffolds were removed from the *B. taurus* assembly. Correction of read pairing information and flags (sort -n | fixmate -m), sorting of mapped reads by coordinates (sort) and the removal of read duplicates (markdup -r -S) were performed with SAMtools 1.14 (Li et al., 2009). Variant calling was performed using BCFtools 1.14 in SAMtools (bcftools mpileup -Ou | bcftools call -mv -Oz). Prior to further analyses, we filtered for alignment quality using BCFtools in the following ways: 1) All single nucleotide polymorphisms (SNPs) located within 5 bp of an INDEL variant as well as INDELs themselves were removed (–IndelGap 5); 2) we excluded all sites at which no alternative alleles were called for any of the samples (AC==0); 3) we excluded all sites at which only alternative alleles were called (AC==AN); 4) only biallelic SNPs were retained (–max-alleles 2); 5) minor allele frequency was set to $\geq$0.05 (-i 'MAF $>$ 0.05′) and only SNP positions with less than 50 % missing data were included ('F_MISSING $<$ 0.5′). This resulted in an autosomal data set with 137,074,193 SNPs. The mitochondrial genome was assembled by mapping the trimmed short reads to the mitochondrial genome scaffold (MT NC_006853.1) using BWA-MEM and default settings. We the obtained consensus genome sequences for each individual in ANGSD (Korneliussen et al., 2014) using the maximum effective base depth for each SNP (-doFasta 3) and the following filtering parameters (-minMapQ 25, -minQ 25, -uniqueOnly 1, -remove_bads 1).

2.3. Genetic structure

2.3.1. Mitochondrial genome phylogeny

First, we reconstructed a mitochondrial genome phylogeny for the genus *Tragelaphus*. We used maximum likelihood (ML) in IQ-TREE v.2.0 (Nguyen et al., 2014; Minh et al., 2020a) with the best fit substitution model determined automatically using the new ModelFinder (-m MFP) option (Kalyaanamoorthy et al., 2017). Nodal support for the ML analysis was assessed with 1000 ultra-fast bootstrap (UFBS) replicates (Minh et al., 2013; Hoang et al., 2018) and 1000 SH-like approximate likelihood ratio test (SH-aLRT) replicates (Guindon et al., 2010). This phylogeny was dated using the same fossil calibration points as described below for the species tree.

2.3.2. Principal component analysis of whole genome SNP data

To estimate the overall relatedness among the ten *Tragelaphus* genomes, we performed a model-free principal components analysis (PCA) on the genome alignment in PLINK 1.9 (Chang et al., 2015). Eigenvalues (–pca) were generated, and principal components (PCs) were deduced from the percentage each eigenvalue contributes to the total variance in the data. The resulting eigenvec file contained the coordinates for each sample in component space, and these were plotted using Cubemaker (https://tools.altiusinstitute.org/cubemaker/).

2.3.3. Species tree inference

We used different approaches for estimating the *Tragelaphus* species tree. First, a phylogeny was inferred using a total evidence approach, using ML on a concatenated alignment of the entire SNP data. This approach averages the phylogenetic signal across millions of nucleotide variants but estimates the probability of alternate evolutionary histories for each node through the computation of site concordance factors (Minh et al., 2020b). First, however, we pruned the alignment of linked sites using a 50 kb sliding window, with step size of 10 kb, removing all sites with a squared correlation greater than 0.1 (−-indep-pairwise 50 10 0.1) in PLINK. We then carried out the ML analysis in IQ-TREE v.2.0, with substitution model determined using ModelFinder and nodal support evaluated though 100 bootstrap (BS) replicates, 1000 ultra-fast bootstrap (UFBS) replicates and 1000 SH-like approximate likelihood ratio test (SH-aLRT) replicates. BS values, UFBS and SH-aLRT values were considered strong when higher than 70 %, 95 % and 80 %, respectively (Minh et al., 2013; Hoang et al., 2018; Guindon et al., 2010). We then computed site concordance factors to determine the proportion of aligned sites that supported each node.

We also used the software ASTRAL-III v5.6.3 (Mirarab et al., 2014; Zhang et al., 2018) to infer a phylogeny that was statistically consistent under the Multiple Species Coalescent (MSC). Therefore, in this summary approach a phylogeny may be computed from genome-scale data even in the presence of ILS. We first obtained consensus genome sequences for each individual in ANGSD using the maximum effective base depth for each SNP (-doFasta 3) and the following filtering parameters (-minMapQ 25, -minQ 25, -uniqueOnly 1, -remove_bads 1). We then extracted non-overlapping windows of 100 kbp. Windows for which at least one sample had 50 % or more missing data were discarded. A phylogenetic tree was reconstructed for each 100 kbp window using RaxML v.8.2.12 (Alexandros, 2014), applying a GTRCAT substitution model. We ran ASTRAL-III on the set of 100 kbp window trees to infer a species tree, using "-t 2" to output quartet support values to visualise window tree conflict. Support for this species tree topology was assessed with local posterior probabilities (PP, Sayyari and Mirarab, 2016).

Lastly, we carried out a sliding window analysis to determine the most common tree topologies present in our data set. The idea behind this analysis was that past introgression events are expected to affect discrete parts of the genome, whereas drift affects the entire genome. Therefore, the true evolutionary topology that results from lineage sorting by drift should be among the most-commonly observed window topologies. For this analysis, we used the 100 kbp windows described

above, but also extracted window lengths of 500 kbp and 1 Mbp from the genome alignment. For each genomic window, a ML phylogeny was computed using RaxML as described above. To obtain counts for the observed topologies, all phylogenies for a given window size were imported without branch lengths into R, and tree distances were computed using the multiDist function of the Treespace library (Jombart et al., 2017). Parsing of the exported distance matrix and evaluating the frequency at which certain clades were observed was carried out using custom Perl scripts that made use of BioPerl (Stajich et al., 2002) and the ETE3 software (Huerta-Cepas et al., 2016).

2.4. Molecular dating

As with all molecular dating analyses, reliability of fossil calibration dates and their use can affect the accuracy of divergence time inferences (Donoghue and Benton, 2007; Parham et al., 2012). Furthermore, phylogenetic molecular dating assumes that all alleles in the analysis are derived through genetic drift following a bifurcating process. We therefore discarded those stretches of the genome that potentially contain signatures of alternative, reticulate, evolutionary histories by extracting and concatenating only the 1 Mbp windows that support the single most common tree topology constructed using all three approaches described above. We performed a least-squares dating analysis on this "introgression free" data set using IQ-TREE v.2.0. Constraints on Tragelaphini clade ages were enforced by fossil calibrations for both the mitochondrial genome tree and the autosomal species phylogeny. Since this tribe is well represented in the Plio-Pleistocene fossil record (Gentry, 2010), we used three unambiguous and informative fossil calibration points, originally proposed by Bibi (2013) in his mitochondrial genome analysis of the Bovidae. Calibration dates were parameterized as most recent common ancestor (MRCA) priors, with the minimum and maximum fossil dates adjusted to the 2.5 % and 97.5 % quantiles of a normal distribution, respectively. Stem Bovini (*Tragelaphus* plus cow) was represented by *Selenoportax vexillarius* (10.2–16 Ma), crown Tragelaphini was constrained using a fixed uniform minimum prior of 5.72 Ma, and the *T. lockwoodi* fossil (3.4–4.5 Ma) was used to constrain the stem *T. strepsiceros* lineage (Bibi, 2013, and references therein).

2.5. Genetic introgression among Tragelaphus

Alternate window tree topologies could be the result of ILS or introgression. Since genetic drift is always acting on the genome, regardless of the occurrence of introgression, ILS results in an exponential distribution of branch lengths among alternate window trees, which can be distinguished from the branch length distribution resulting from introgression, which contains the ILS distribution plus that of the introgression event (Edelman et al., 2019). In the presence of only ILS, window trees for any rooted triplet that are concordant with the species tree should be at least as frequent as either of the two alternate topologies (Hibbins and Hahn, 2022). We employed the software quantifying introgression via branch lengths (QuIBL, Edelman et al., 2019) to test for introgression against a null hypothesis of ILS using our 100 kb window trees. Due to the large number of window trees, we first randomly thinned the data set to a manageable size using "awk 'rand() $>$ 0.9'" before running QuIBL.

Then, using the most common tree topology inferred from ML, MSC and window tree methods as our underlying history of divergence and drift, we attempted several approaches to identify potential gene flow between non-sister lineages. The sliding window analysis described above was useful not only for determining the most common tree topology, but also in identifying the most common alternate tree topologies that could have resulted from reticulate (non-bifurcation) evolutionary events. We used ASTRAL to infer the frequency across genomic windows of the main and alternative quartet topologies for those quartets with the highest proportions of the alternative topologies. In addition to analyses of alternative window tree topologies and quartets, patterns of introgression were more formally assessed using Dsuite (Malinsky et al., 2021). D-statistics were calculated from whole genome SNP data for all trios (command 'Dtrios') of *Tragelaphus* species that were compatible with the previously reconstructed species tree, with the outgroup being fixed in position four. A heatmap showing the D-statistic and its p-value was plotted for all pairs of individuals using the plot_d.rb script (https://github.com/mmatschiner/tutorials/blob/master/analysis_of_introgression_with_snp_data/src/plot_d.rb). The Dtrios command also calculated f_4-ratios (f_G) for all ingroup trios. Given that the many resultant f_4-ratios were highly correlated, we used 'f-branch' type calculations (command 'Fbranch', Malinsky et al., 2018) to assign gene flow evidence to specific, possibly internal, branches on a phylogeny, using the –tree option that allows the inclusion of a known species tree, rooted with the cow as outgroup taxon. The output of Dsuite Fbranch was plotted with the dtools.py script (https://github.com/millanek/Dsuite/tree/master/utils/dtools.py).

2.6. Trait evolution

We attempted to map the more commonly known ecological and phenotypic traits among the spiral-horned antelopes onto the species tree using a parsimony approach. These included two ecological traits: habitat type (dry-savanna or closed-forest − based on distribution maps of species and habitat-vegetation types in Africa); and diet specialization (browser, mixed-feeders, or grazers − following Gagnon & Chew (2000), but with a fusion of their categories "browser-grazer intermediates" and "generalists"), and four phenotypic traits: life strategy (generalist or specialist − based on whether or not the taxon was restricted to certain habitats); general phenotype (nyala-like or kudu-like − based mainly on lyrate and spiral horns, respectively, but also the pale colouration of the kudu-like phenotype and the more compact body conformation of the nyala-like phenotype); the presence or absence of horns in both sexes; and pronounced sexual dimorphism in coat colour. These traits were scored for earliest tragelaphines using Gentry (2010) and Haile-Selassie et al. (2009), for extant species using Gagnon & Chew (2000), Groves & Leslie (2011) and Kingdon (2013). The only multi-state trait (diet specialization) was considered an ordered (additive) character. Trait evolution was reconstructed in NONA (v.2.0) with the WINCLADA interface (v.1.00.0884, Nixon, 1999) using unweighted maximum-parsimony, with preference for neither ACCTRAN nor DELTRAN optimization when alternative reconstructions were of equal cost.

3. Results

3.1. Autosomal and mitochondrial data sets

A total of 1933 million (M) raw sequencing reads were generated for this study, with individual species ranging from 152 M to 260 M reads (Table 1). Of these 1595, or between 128 M and 214 M reads (77–90 %) were successfully mapped to the cow reference genome, reflecting its ~13 Ma divergence with the *Tragelaphus* ingroup (Bibi, 2013). The average autosomal coverage attained ranged from 7-fold (*T. oryx*) to 13-fold (*T. imberbis*). One the other hand, between 17 thousand (K) and 291 K reads were mapped onto the mitochondrial reference genome reflecting 143-fold to 2266-fold coverage (Table 1).

3.2. Mitochondrial phylogeny

The reconstructed the first mitochondrial genome phylogeny to include all *Tragelaphus* species (Fig. 2A). As in previous analyses (Hassanin et al., 2012, Bibi, 2013), we found *T. imberbis* to be the basal-most species, followed by *T. angasii* and *T. s. scriptus*, and with clear differentiation between the derived dry savanna and closed forest clades. While *T. strepsiceros*, was included in the dry savanna clade together with *T. oryx* and *T. derbianus*, the bootstrap and SH-aLRT values for this node were markedly lower than other nodes at 79.4 and 76,

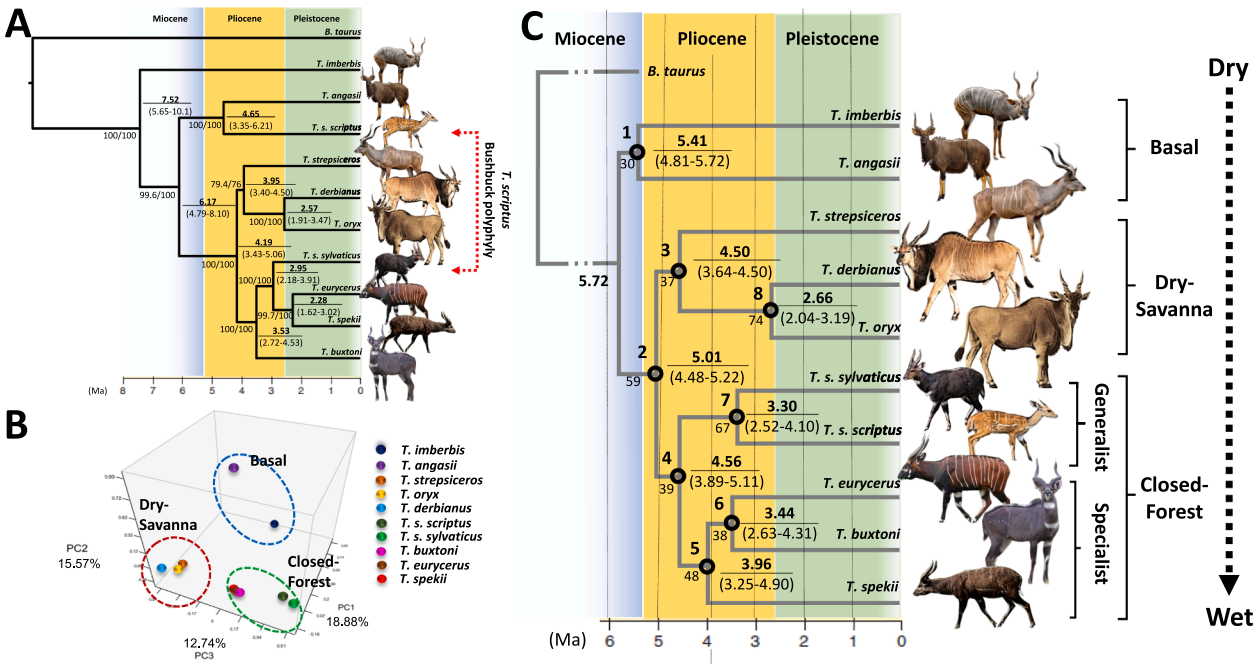

Fig. 2. The structure of genomic variation among the spiral-horned antelopes (*Tragelaphus* spp.) A. Mitochondrial genome maximum likelihood phylogeny reconstructed with IQTREE. SH-aLRT/bootstrap values are given to the left of each node. Estimated node ages [in million years (Ma)] are given in bold to the right of each node, with 95 % highest posterior density (HPD) intervals given below the node age. B. Principal component analysis on the Tragelaphus whole genome data. C. Species tree reconstruction based on whole genome SNP data and whole genome alignments, respectively, inferred with IQtree and ASTRAL. Black circles represent numbered nodes (1–8) with ≥99 % bootstrap, ultrafast bootstrap, SH-aLRT (IQTree) and posterior probability (Astral) support. Site concordance values are given below each node. Estimated node ages [in million years (Ma)] are given in bold to the right of each node, inferred by the LSD2 method, averaged across concatenated 1 Mbp blocks that conform to the species tree topology. 95 % HPD intervals are given below the node age. The phenotypic variation among males of each species is depicted.

respectively. Within the closed forest group, *T. buxtoni*, was placed ancestrally, followed by *T. s. sylvaticus* and with the specialised *T. eurycerus* and *T. spekii* sharing a derived sister relationship. Importantly, the well-known mtDNA polyphyly of the bushbuck was reconstructed, with *T. s. scriptus* sister to *T. angasii* and *T. s. sylvaticus* most closely related to the *T. eurycerus*-*T. spekii* sister grouping.

3.3. Principal component analysis

The principal component analysis aimed to determine the overall structure of the whole genome data (Fig. 2B) and showed that the first three principal components accounted for 18.88 %, 15.57 % and 12.74 % of the overall variability, representing together almost half (47.19 %) of the total variation among the species' genomes. This analysis revealed high levels of genomic distinctiveness among the spiral-horned antelopes, revealing three major clusters. The first component separated *T. imberbis* while the second component separated *T. angasii* from all other species (basal cluster). Principal component three then separated three groups, the greater kudu-eland dry-savanna group (*T. strepsiceros, T. derbianus* and *T. oryx*) and the closed-forest group (*T. spekii*, *T. buxtoni*, *T. eurycerus and T. scriptus*), with the two bushbuck genomes, *T. s. scriptus* and *T. s. sylvaticus*, clearly differentiated from others in the group (Fig. 2B).

3.4. Species tree reconstruction

We used several approaches to infer the phylogenetic relationships among our ten *Tragelaphus* genomes. The ML analysis of the genome-wide SNP data resulted in a strongly supported species tree with bootstrap, ultrafast bootstrapping and approximate likelihood-ratio values ≥99 % (Fig. 2C). Unlike the mitochondrial genome phylogeny (Fig. 2A), the autosomal species tree revealed that *T. angasii* and *T. imberbis* were sister lineages and formed the basal group in the phylogeny, with all other genomes forming two further clades. The first of these clades contained the dry-savanna species, with *T. strepsiceros* most closely related to the eland sister grouping (*T. oryx* and *T. derbianus*). The second clade comprised the specialised closed-forest group of *T. buxtoni*, *T. eurycerus*, *T. spekii*, and, also in contrast to mtDNA, a nested clade containing both *T. s. scriptus* and *T. s. sylvaticus* bushbuck genomes. Species tree reconstruction using the MSC returned the same highly supported tree topology as described above (Fig. 2C). However, despite robust statistical support for all nodes in the ML phylogeny, site concordance values were much lower for all nodes, ranging from 30 % to 75 % (Fig. 2C). The tree topology obtained with both approaches was also the most commonly reconstructed tree in our sliding window analysis (Fig. 3A), albeit with highest frequencies of only 3.57 %, 15 % and 23.17 % of trees, inferred from 100 kbp, 500 kbp and 1 Mbp window sizes, respectively. Since smaller window sizes could potentially identify more ancient reticulation events, we also performed this analysis with a window size of 10 kbp for Chromosome 1, which resulted in the same most common topology but with a frequency of only 0.2 %, suggesting an increase in the number of observed alternate topologies with decreasing window size (data not shown). Thus, although the frequency of the species tree topology decreased with decreasing window size, it was always the most frequently retrieved topology (Fig. 3B). Since all three approaches suggested the same phylogenetic relationships among species, we refer to this topology as the species tree of the spiral-horned antelopes from here on.

3.5. Molecular dating

Molecular dating of both the mitochondrial genome and species trees showed that spiral-horned antelopes coalesced to a late Miocene common ancestor, although the date for the base of the mtDNA tree (Fig. 2A) was older at 7.52 million years ago (Ma), compared to 5.72 Ma) for the species tree. However, this difference was not significant as the 95 %

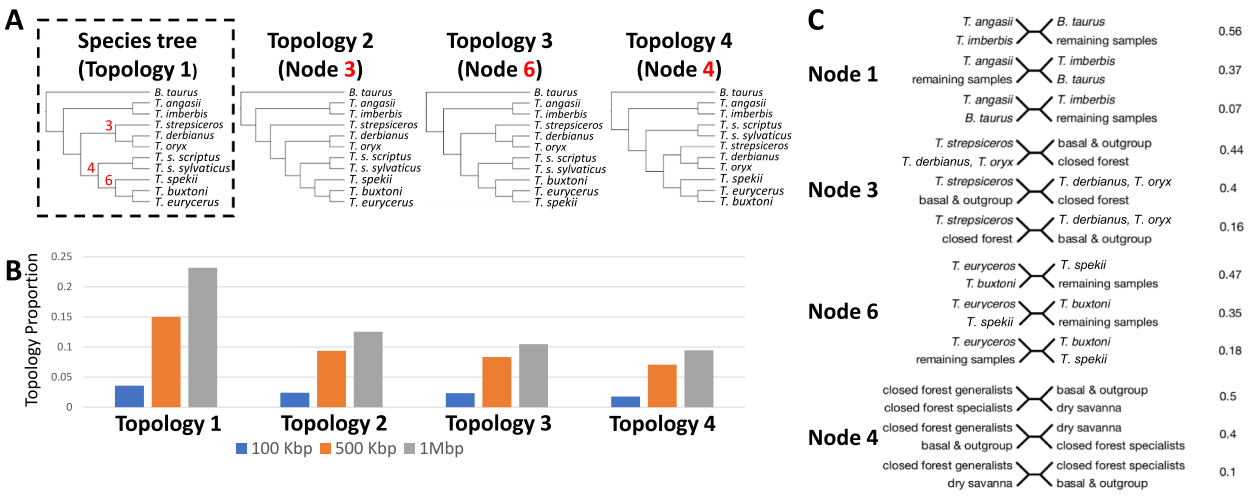

Fig. 3. Characterization of phylogenetic discordance among the spiral-horned antelopes (*Tragelaphus* spp.). A. Tree topologies of the four most common sliding window trees, with the nodes affected in alternative topologies given in red within the species tree (dashed black box and Fig. 2C). B. Frequency of each topology for window sizes of 100 kbp, 500 kbp and 1 Mbp sliding windows. Topologies are ordered as in panel A. C. Frequencies of four sets of main and alternative quartet topologies, inferred from whole genome alignments in ASTRAL. Node 1, Frequency at which the two basal species *T. imberbis* and *T. angasii* are sister taxa; Node 3, Frequency of the occurrence of the dry-savanna clade; Node 4, Frequency of occurrence of the closed-forest clade; Node 6, Frequency at which the bongo (*T. eurycerus*) is sister to the mountain nyala (*T. buxtoni*) relative to the sitatunga (*T. spekii*). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

highest posterior density (HPD) for the older mitochondrial date included the age of the species tree and was likely inflated by an additional branching event at the base of the mitochondrial tree, since the basal *T. imberbis* was not sister to *T. angasii* (Fig. 2A). The HPDs of all other branching events overlapped between mitochondrial and species trees. A series of initial diversification events then occurred in the late Miocene/early Pliocene, with *T. angasii* and *T. imberbis* diverging from a common ancestor, followed by the divergence of the closed-forest clade from dry-savanna ancestors. During the early to middle Pliocene, both eland species (*T. derbianus*, *T. oryx*) diverged from *T. strepsiceros* and the closed-forest clade split into *T. scriptus* and *T. buxtoni-T. eurycerus-T. spekii* clades. Later divergences occurred mostly in the middle Pliocene although the youngest species divergence, between the two eland species, dates to the beginning of the Pleistocene. Within the closed-forest group, the generalist bushbuck lineage (*T. scriptus*) was the first to diverge from the remaining specialist species, with *T. s. scriptus* and *T. s. sylvaticus* lineages diverging subsequently some 3.30 Ma. *T. spekii* was next to diverge from the remaining closed-forest species, leaving the sister taxa *T. eurycerus* and *T. buxtoni* with a divergence of 3.44 Ma.

3.6. Gene flow

All 120 triplet combinations among 100 kb window trees were tested against a null hypothesis of ILS only, in turn using each member of the triplet as the outgroup taxon, resulting in 360 likelihood tests. All but 23 alternate topology triplets had branch length distributions that significantly favoured the introgression model over that of ILS (BIC between models >10, Table S1). The 23 non-significant triplet topologies were rarely detected among sliding windows (count number between 4 and 37). However 21 of these alternate topology triplets still had a greater proportion of loci in favour of introgression compared to ILS, suggesting that their low count number was primarily responsible for their lack of significance. We therefore conclude that the alternate topologies in our data set overwhelmingly favoured a history of introgression over ILS.

Signatures of gene flow among *Tragelaphus* species were inferred through a variety of approaches, all assuming the topology of the inferred species tree (Fig. 2C). We first investigated patterns of genealogical discordance among genomic windows. The most commonly retrieved topology (Topology 1, Fig. 3A) was the same as that inferred by ML and Bayesian multispecies coalescent. Alternative topologies were used to infer reticulate events. Topologies 2, 3 and 4 were the same across the three window sizes (Fig. 3A). Topology 2 differed from that of the species tree in that the greater kudu (*T. strepsiceros*) was no longer sister with the two eland (*T. oryx* and *T. derbianus*, Node 3 in the species tree) but was basal to all species except the basal clade (*T. imberbis* and *T. angasii*) (Fig. 3A). Topology 3 affected Node 6 in the species tree, where the bongo (*T. eurycerus*) was most closely related with the sitatunga (*T. spekii*) rather than with the mountain nyala (*T. buxtoni*) as in the species tree. Finally, Topology 4 broke up the closed-forest clade, which would be otherwise united under Node 4, placing the monophyletic bushbuck (*T. scriptus*) genomes outside this clade, and intermediate between the most basal species and all other species (dry-savanna/closed-forest).

We also used ASTRAL to perform a quartet analysis of genealogical discordance, which ranks quartet topologies according to their frequency of occurrence among sliding window trees of 1 Mbp in size (Fig. 3C). For the investigated four tree topologies (the species tree plus alternative topologies affecting nodes 3, 4, and 6), the quartet representing the species tree topology was the most frequent (>0.44), although the analysis also revealed quartets with appreciable alternative evolutionary histories (<0.40, Fig. 3C). These alternative quartets highlighted similar patterns of discordance as the sliding window trees above, suggesting higher than expected relatedness between: greater kudu (*T. strepsiceros*) and either the nyala (*T. angasii*) and/or the lesser kudu (*T. imberbis* (Node 3)); the two bushbuck (*T. s. scriptus* and *T. s. sylvaticus*) and either basal species *T. angasii* and/or *T. imberbis* (Node 4), and finally the bongo (*T. eurycerus*) and sitatunga (*T. spekii* (Node 6)). However, the ASTRAL analysis revealed additional introgression involving *T. angasii* and other non-sister species (Node 1).

To further investigate reticulate events among *Tragelaphus* taxa, we employed more formal tests of gene flow using the Dsuite package that implements the four-taxon Patterson's D and the f4-ratio for all combinations of trios of species from SNP data. The highest D-statistics were observed almost exclusively among pairwise comparisons with the nyala (*T. angasii*) and all other species except *T. imberbis* and *T. spekii* (Fig. 4A). Also noteworthy were high D values between the dry-savanna clade and the bongo (*T. eurycerus*) and to a lesser extent the mountain nyala (*T. buxtoni*). Unlike D-statistics, the f-branch method uses f4-ratios to

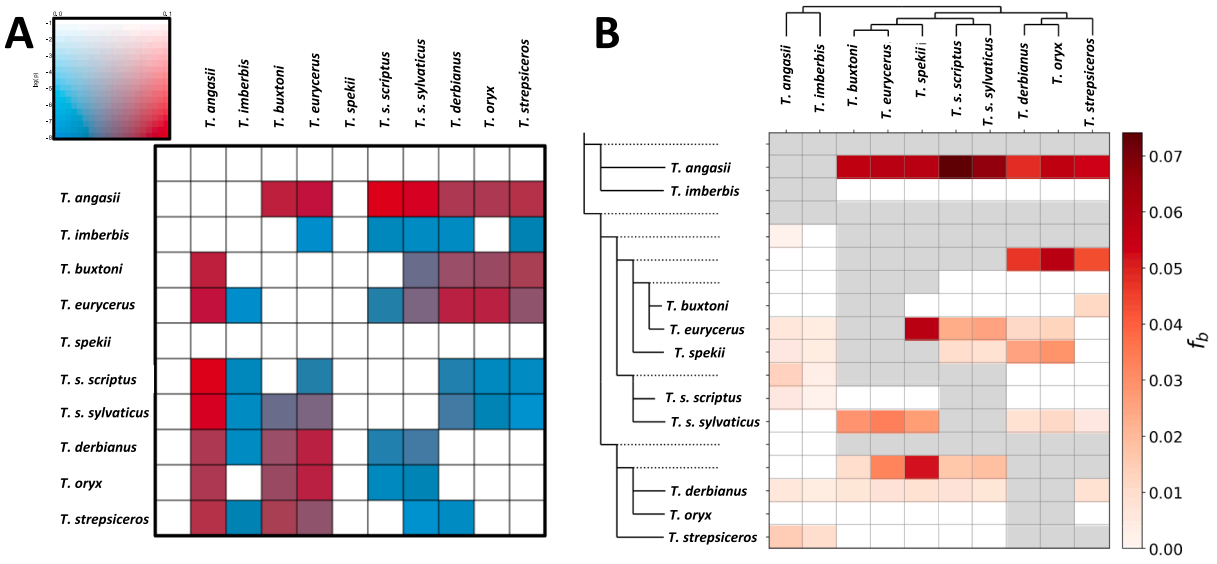

Fig. 4. Post-divergence gene flow among members of the genus *Tragelaphus*. A. Gene flow deduced from raw D-statistics with *p* values inferred from Dsuite Dtrio tests. Red colour indicates higher D-statistics, and more saturated colours indicate greater significance. B. Gene flow conditioned on the species tree topology. This heat map shows the results of the f-branch test on the *Tragelaphus* SNP dataset with the species tree used to simulate the data shown along the x and y axes. The matrix shows the inferred f-branch statistics, denoting excess allele sharing between the branch of the 'laddered' tree on the y axis (relative to its sister branch) and the species identified on the x axis. Red color indicates higher f-branch statistics f_b, and more saturated colours indicate greater value. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

assign gene flow to specific and possibly internal branches on the species tree. This analysis also revealed gene flow between the nyala lineage and other non-sister species, especially from the two bushbuck genomes. Other inferred reticulate events were between the ancestor of the closed forest specialists and the dry-savanna clade, but reciprocally, only between the ancestor of the eland species (*T. oryx* and *T. derbianus*) and *T. spekii* and *T. euryceros*. The f-branch test also showed high gene flow between *T. euryceros* and *T. spekii* but this was not reflected in D-statistics.

3.7. Trait evolution

Of the two ecological and four phenotypic traits that were reconstructed, only habitat type (that is, dry-savanna or closed-forest) was fully consistent with the species tree, requiring a single evolutionary change at the base of the closed-forest clade (Fig. 5A). All other trait reconstructions were indicative of homoplasy. Life strategy, general phenotype and horns in both sexes all required two changes in character state (Fig. 5B–D), whereas sexual dimorphism in coat colour (Fig. 5E) and diet specialisation (Fig. 5F) required three and five character state changes, respectively.

4. Discussion

The evolutionary relationships among the spiral-horned antelopes of Africa have been subject to much debate, mainly owing to the occurrence of several phenotypically similar species pairs that are not sister lineages in the phylogenetic tree (Pallas, 1766; Ogilby, 1837; Gray, 1847; Angas, 1848; Speke, 1863; Blyth, 1869; Sclater and Thomas, 1900; Lydekker, 1910; Hassanin and Douzery, 1999; Matthee and Robinson, 1999; Willows-Munro et al., 2005; Moodley et al., 2009; Hassanin et al., 2018; Rakotoarivelo et al., 2019b). Genetic evidence from traditional molecular markers (mtDNA and nuclear introns) suggested an evolutionary history of divergence, but with a potentially important role for gene flow in the speciation process (Moodley et al., 2009; Hassanin et al., 2012; Hassanin et al., 2018; Rakotoarivelo et al., 2019b). In this study, we generated whole genome sequences for each of the nine extant *Tragelaphus* species and analysed these data to better understand their complex evolutionary history.

4.1. A species tree for Tragelaphus

We inferred the most probable species tree from whole genome data using a concatenation approach (ML) and a summary approach, both of which resulted in the same phylogeny (Fig. 2C). Despite strong statistical support using a range of tests, site concordance factors for the ML tree were lower for all nodes, indicating that the proportion of SNPs supporting each node was much lower. Site concordance values are almost always lower than bootstrap values (Minh et al., 2020b), reflecting alternate evolutionary histories potentially caused by ILS and/or gene flow. Using a sliding window approach, we reconstructed phylogenetic trees of genomic regions to determine the most commonly observed topologies in the genome, with the expectation that the previously inferred species tree would be observed among them. Although it is possible that significant amounts of introgression as well as ILS in a rapid radiation could lead to differences between the most commonly observed tree and the true species tree (Fontaine et al., 2015), that was not the case for this data set. Instead, the most commonly occurring window topology for all window sizes was identical to the ML and summary species trees. We also tested the robustness of this species phylogeny using window tree counts of species quartets, and in all cases, the most commonly observed quartet topologies were in line with those of the inferred species trees. From these results we conclude that we have reconstructed a most likely species tree for *Tragelaphus*.

Our results provided genome-level support for the early Pliocene divergence and monophyly of the nyala (*T. angasii*) and lesser kudu (*T. imberbis*), the monophyly of the two eland species (*T. oryx* and *T. derbianus*) and, importantly, the monophyly of kéwel (*T. s. scriptus*) and imbabala (*T. s. sylvaticus*) bushbuck. These relationships were previously suggested by mtDNA and/or nuclear introns (Willows-Munro et al., 2005, Moodley et al., 2009, Hassanin et al., 2012, Bibi, 2013, Hassanin et al., 2018, Rakotoarivelo et al., 2019b), but not by Chen et al. (2019), a study that did not include nyala, giant eland or kéwel genomes. We were also able to leverage our genome scale data to shed

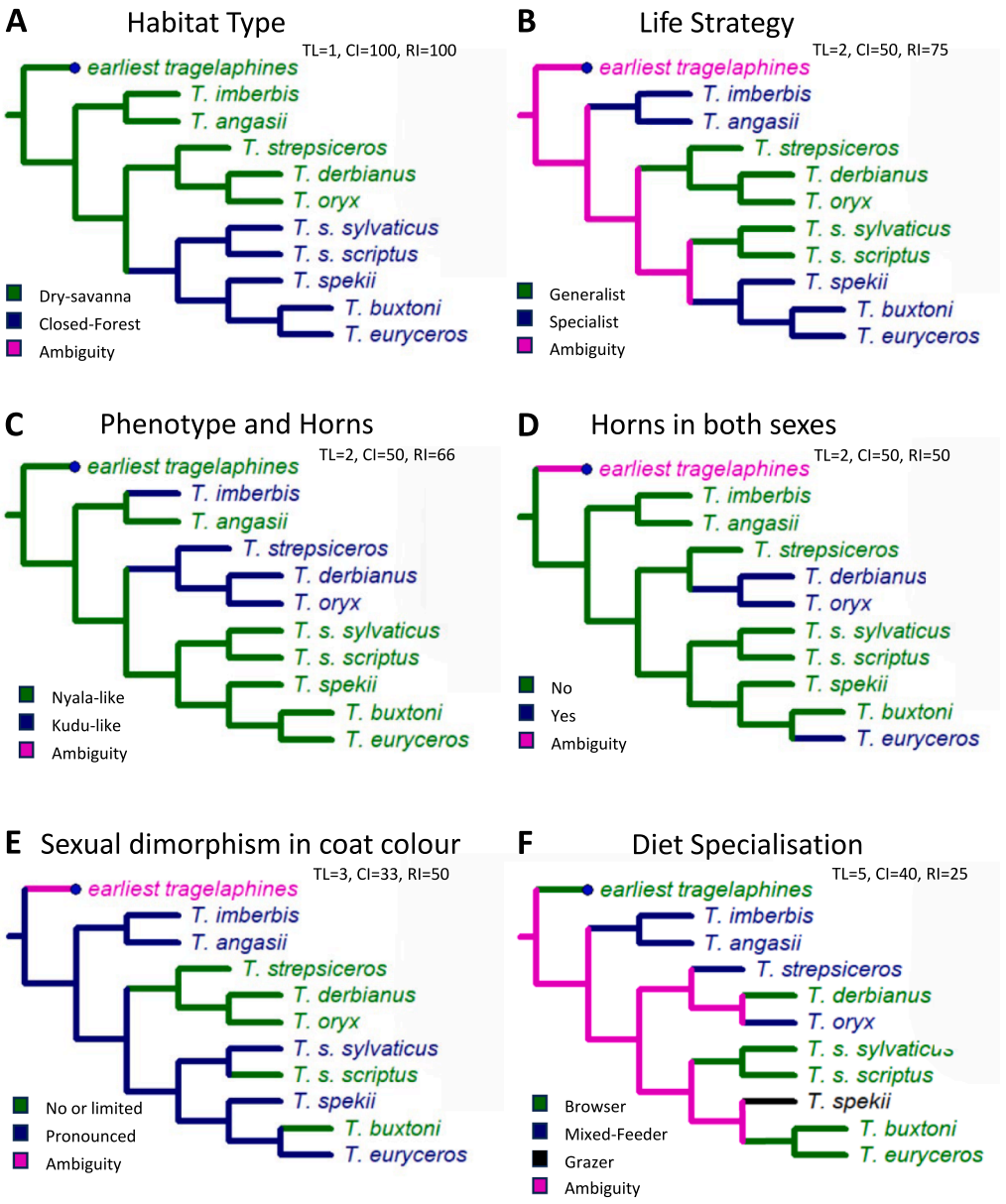

Fig. 5. Reconstructing trait evolution in *Tragelaphus*. Phenotypic traits were mapped onto the species tree using unweighted maximum-parsimony. Characters were scored according to Gagnon & Chew (2000), Gentry (2010), Haile-Selassie et al. (2009) and Groves & Leslie (2011). TL, tree length or the number of character state changes; CI, consistency index; RI, retention index. The ancestral state was determined, where possible, from the earliest known tragelaphine (*T. moroitu*).

further light on uncertainties in phylogenetic positions that were previously flagged by studies using mtDNA (Fig. 2A). Our species tree (Fig. 2C) also showed clear separation between dry-savanna and closed-forest clades, the sister relationship between the greater kudu (*T. strepsiceros*) and the eland (*T. oryx* and *T. derbianus*), and the sister relationship between the bongo (*T. eurycerus*) and the mountain nyala (*T. buxtoni*), rather than to the swamp-dwelling sitatunga (*T. spekii*). These relationships were also suggested by Chen et al. (2019) despite their reduced data set.

4.2. Evolutionary history

Dating the newly reconstructed genome-based species tree allowed us to trace the evolutionary history of the spiral-horned antelopes. Although the Miocene is generally known for its warm climate, by the end of this epoch temperatures had cooled, and dry savanna grasslands increased significantly (Herbert et al., 2016). Environments during the late Miocene must also have been affected by the last of the major episodes of rifting in East Africa, which peaked at around 5.7 Ma (Michon et al., 2022). It is within this period of environmental change that *Tragelaphus* first appears in the fossil record, in the form of *T. moroitu*, which was similar to, but slightly smaller than the present-day *T. angasii* (Haile-Selassie et al., 2009; Gentry, 2010). Soon thereafter, the genus diverged into the common ancestor of the two basal-most species (*T. imberbis* and *T. angasii*) and the common ancestor to all other species (Fig. 2C). Some authors maintain that this early divergence warrants the resurrection of the genera *Ammelaphus* and *Nyala* (Groves and Grubb, 2011). Since the life strategy of *T. moroitu* is not discernible from presently available fossil material, we cannot know whether the generalist strategy of the dry-savanna clade and that of the bushbuck was ancestral or derived (Fig. 5B). It is known, however, that *T. moroitu* was a browser, associated with riparian and drier open woodland (Gentry, 2010), and this suggests that while all spiral-horned antelopes are dependent on wooded areas or thickets for cover, this dependency did not limit their ability to evolve into generalist and highly specialised species (Fig. 5B), nor did an ancestral browsing diet

limit their evolution into mixed feeding (*T. imberbis*, *T. angasii*, *T. strepsiceros*, *T. oryx*) and grazing (*T. spekii*) species (Fig. 5E). Hence, from the species tree and from Fig. 5B and E, it is likely that the specialist, mixed feeding strategy of both basally diverging species, restricted to xeric north-east Africa (*T. imberbis*) and dry woodlands of south-east Africa (*T. angasii*), had already evolved by the time of their common ancestor.

Unlike life and diet strategies that evolved more than once, our results suggest that the closed-forest phenotype was derived only once, and from dry-savanna ancestors (Fig. 5A), around the early-mid Pliocene. Sustained cooling from the middle Pliocene onwards (Herbert et al., 2016) may have isolated forest species (Grubb, 1982) as forested habitats decreased, while the large body sizes of *T. strepsiceros*, *T. oryx* and *T. derbianus* and their generalist life strategy may have been advantageous as grassland and mosaic biomes expanded. The closed-forest clade contains several highly divergent phenotypes that differ widely in size, coloration and habitat specialisation. Within this group, the bushbuck (*T. scriptus*), a browsing forest and mosaic habitat generalist, diverged first in the early-mid Pliocene. The bushbuck also retains the ancestral ability to inhabit drier habitats, can be found in most sub-Saharan ecozones, and remains the most widely distributed bovid species in Africa today.

The common ancestor to the remaining closed-forest species appeared in the mid-Pliocene and rapidly gave rise to the three most uniquely specialised *Tragelaphus* species (Fig. 5B), suggesting the presence of unexploited habitat niches within the forested regions of tropical Africa. Unlike mtDNA and nuclear intron analyses that placed the mountain nyala (*T. buxtoni*) as basal within the closed-forest group (Rakotoarivelo et al., 2019a), genome-scale data placed this species in a derived position, sister to the bongo (Chen et al., 2019 and Fig. 2C), a similarly-sized closed-forest specialist. While *T. buxtoni* is restricted to the montane forests and sub-alpine moorlands of the Ethiopian highlands, *T. eurycerus* occurs on either side of the African rift in the lowland forests of Central and West Africa and on the forested slopes of the more prominent mountains in East Africa, such as Mt Elgon, the Mau Escarpment, the Aberdares and Mt Kenya. The relatively recent common ancestry between *T. eurycerus* and *T. buxtoni* (2.63–4.31 Ma, Fig. 2C) postdates the last of the major rifting events (5.7 Ma) and suggests that forested corridors must have been present between the closed forests of the Congo basin and those of the East African volcanoes and escarpments, allowing for the dispersal of *T. euryceros* into its present day range. Thus, among the closed forest specialists, *T. spekii* was the first to diverge and become specialised (Fig. 2C). Its grossly elongated hooves, elongated legs and oily, water-repellent coat would have been highly advantageous in the permanent swamps and wetlands of central Africa. The results of our analyses suggest that the closed-forest group represents a major transition in the evolution of *Tragelaphus* from its dry ancestral habitats to wetter, more forested ecosystems, in which the generalist *T. scriptus* represents a living relict of this transition between dry and wet habitats.

4.3. Complex patterns of introgression

Although we were able to reconstruct the same species phylogeny with different approaches, low concordance factors showed that as much as 61–70 % of sites supported alternative evolutionary histories for node 1 (the sister relationship between *T. angasii* and *T. imberbis*), node 3 (the placement of *T. strepsiceros* together with *T. oryx* and *T. derbianus*), node 4 (the placement of the closed-forest generalist *T. scriptus* with closed-forest specialists) and node 6 (the sister relationship between *T. eurycerus* and *T. buxtoni*; Fig. 2). Not surprisingly, our sliding window analyses using the full and quartet data sets revealed substantial alternative tree topologies for the very same four nodes described above (Fig. 3). These were also the nodes at which a limited number of traditional loci provided poor resolution/statistical support (that is, Nodes 1, 3 and 6, Willows-Munro et al., 2005, Rakotoarivelo et al., 2019b) or contrasting tree topologies (Node 4, Moodley et al., 2009, Hassanin et al., 2018, Rakotoarivelo et al., 2019b). Since we were able to rule out ILS as a significant factor contributing to these alternate topologies (Table S1), gene flow between non-sister lineages provides an explanation for why a handful of traditional molecular markers were not able to resolve these relationships.

More formal tests of gene flow based on excess allele sharing allowed a greater resolution of these reticulate evolutionary events. The nyala (*T. angasii*) lineage exhibited the most widespread signatures of ancestral gene flow, explaining the low site concordance at node 1 and why 37 % of quartet trees place *T. angasii* closer to the rest of the genus than to its sister taxon *T. imberbis* (Fig. 2C and 3C). It also suggests that *T. angasii*, or a closely related and now-extinct species, ranged widely enough during the early stages of the *Tragelaphus* radiation to have overlapped in distribution with the other speciating lineages. Fossils provide ample evidence for this, as typically nyala-like lyrate and keeled horn cores are abundant in the record, occurring from the late Miocene to the early Pliocene in Ethiopia (Adu-Asu), Kenya (Lukeino) and South Africa (Langebaanweg) (Gentry, 2010). This suggests that there existed nyala-like lineages across Africa that could potentially have been in genetic contact with other speciating *Tragelaphus* lineages. Alternatively, the patterns of widespread introgression involving *T. angasii* could have stemmed from a single gene flow event, possibly into *T. scriptus* (see below). This possibility might arise because f-branch tests carried out on simulated data sets show that introgression may not only be inferred for the taxa between which gene flow actually occurred, but also between closely related taxa where gene flow did not occur, possibly as a result of ancestral population structure (Malinsky et al., 2021). Although unlikely, it is also possible that we have incorrectly inferred the sister relationship between *T. angasii* and *T. imberbis*, and that the nyala is in fact sister to the rest of the species in the genus (alternate topologies for node 1, Fig. 3C). This would explain the high degree of allele sharing between *T. angasii* and all other species, apart from *T. imberbis*.

Nevertheless, non-sister allele sharing was highest between *T. angasii* and the two bushbuck (*T. scriptus*) genomes (Fig. 4A), especially with the kéwel (*T. s. scriptus*, Fig. 4B). This provides the first nuclear genome evidence that introgression was likely responsible for the mtDNA polyphyly of the bushbuck (Fig. 2A). However, we calculated that the mtDNA divergence between sister taxa *T. angasii* and *T. s. scriptus* was ancient (3.35 – 6.21 Ma) and only just overlapping with the nuclear genome divergence between *T. s. scriptus* and *T. s. sylvaticus* (2.52 - 4.10 Ma, Fig. 2C). It is more likely, therefore, that the inferred introgression occurred into the ancestral bushbuck population, prior to its divergence into *T. s. scriptus* and *T. s. sylvaticus* lineages. This gene flow event explains not only the observed nyala-like mtDNA of the kéwel (Fig. 2A), which was lost by the imbabala (*T. s. sylvaticus*) through random genetic drift, but it would also link the pronounced sexual dimorphism in coat colour of the nyala with that of the imbabala (Fig. 5E), which was lost to the kéwel. This scenario would also explain why 40 % of alternative quartet topologies place both bushbuck taxa (*T. s. scriptus* and *T. s. sylvaticus*) as sister to the basal clade and not to the closed forest group (node 4, Fig. 3C). Alternatively, the nyala-like mtDNA might never have been obtained by *T. s. sylvaticus* if the ancestral bushbuck population was already structured phylogeographically at the time of introgression. A larger sample of bushbuck genomes from across that species' range might help to shed more light on these questions.

The f-branch test also suggested gene flow between the ancestor of closed forest specialists with all dry-savanna species, however, the reciprocal part of the plot only identified gene flow between the ancestor of the eland species and *T. spekii* and *T. euryceros* (Fig. 4B). On the other hand, D-statistics identified a similar pattern of allele sharing, but between dry-savanna species with *T. euryceros* and *T. buxtoni*. We suggest, as we do for the widespread introgression events involving *T. angasii* above, that the differences between these similar patterns may have been caused by a single ancient gene flow event and that ancestral

population structure has resulted in the observed differences in allele sharing patterns. Whether this event occurred between the ancestor of the eland species or the ancestor of all dry-forest species is difficult to determine from our data set. The fact that 40 % of quartet topologies place *T. strepsiceros* outside the dry-savanna clade (node 3, Fig. 3C) suggests that gene flow may have occurred only between the eland clade and the closed forest clade. Lastly, gene flow was also inferred between *T. euryceros* and *T. spekii* by the f-branch test and through alternative quartet topologies (node 6, Fig. 3C) but not by D-statistics. This gene flow event may explain why these species are sister taxa for mtDNA. We hope that these complex patterns of introgression may be unravelled in the future through explicit modelling using a more comprehensive dataset, including several genomes per species, sampled from across the range of each species.

4.4. Introgression and the onset of reproductive isolation

The f-branch test inferred ancient gene flow, all of which occurred during the Pliocene and mainly between ancestral lineages. While several extant *Tragelaphus* species are known to hybridise in captivity such as *T. strepsiceros* and *T. oryx* (Boulineau, 1933, Van Gelder, 1977), *T. scriptus* and *T. spekii* (Gray, 1972), *T. euryceros* and *T. spekii* (Koulischer et al., 1973) and *T. strepsiceros* and *T. angasii* (Dalton et al., 2014), interspecific hybrids have not been documented in the wild, despite widespread overlap in the distribution of extant species. We assume, therefore, that these taxa are presently reproductively isolated from each other. Given the range of different chromosome numbers among *Tragelaphus* species, especially among the two basal species *T. imberbis* (2n = 38/38) and *T. angasii* (2n = 55/56), the dry-savanna species (2n = 31/32) and the closed-forest species (2n = 30/30 or 33/34), it is possible at least in part, that present-day reproductive isolation might be linked to variance in karyotype. This would provide one explanation for our inference of Pliocene, and not more recent, gene flow events (Fig. 2C). The only exception to this is the bushbuck, whose kéwel and imbabala forms are in ongoing secondary contact along the African rift system, where populations have been described with intermediate phenotypes (Rakotoarivelo et al., 2019a). Although Hassanin et al. (2018) suggested distinct karyotypes for the two kinds of bushbuck, they also reported the same fundamental number of chromosomes, and this, plus the fact that kéwel and imbabala are sister taxa, allows for the gene flow that produces intermediate phenotypes. Therefore, since the development of karyotype variation and reproductive isolation is partially a product of time since divergence, or last secondary contact, it follows that speciating *Tragelaphus* lineages were probably not yet reproductively isolated during the Pliocene. Hence, the absence of inferred gene flow events in the Pleistocene suggests the onset of reproductive isolation during this period for most species lineages.

4.5. Trait evolution

The nyala-like phenotype is likely ancestral since the ancestral tragelaphine, *T. moroitu*, was cranially similar to *T. angasii*, and with similarly lyrate horns. However, at some point, the lesser kudu (*T. imberbis*) evolved the strikingly different kudu-like phenotype while the nyala (*T. angasii*) retained the presumably ancestral nyala-like form. The next major divergence saw the divergence of a dry-savanna clade that also evolved a kudu-like phenotype and a closed forest clade that probably retained the ancestral nyala-like phenotype (Fig. 5A). The high number of reticulate events involving the *T. angasii* suggests that the nyala-like phenotype could have been conferred to an ancestor of the closed-forest group through gene flow, but lost in the dry-savanna clade despite gene flow with *T. angasii*.

The relative absence of detectable gene flow involving the kudu-like species *T. imberbis* and *T. strepsiceros*, which not only share their tightly spiralled horns, but are highly similar in cranial morphology (Alden et al., 1995; Kingdon, 1982), body form and pelt colouration (Walker et al., 1964) is most surprising given that this phenotype is probably derived and not ancestral. The phylogenetic placement of *T. strepsiceros* has always been erratic (Willows-Munro et al., 2005, Moodley et al., 2009, Hassanin et al., 2012, Rakotoarivelo et al., 2019a and Fig. 2A) and 40 % of the quartets analysed in the present study placed this taxon outside the dry-savanna clade (Fig. 3C). However, if gene flow cannot account for the presence of the kudu-like phenotype in non-sister *T. strepsiceros* and *T. imberbis*, we are left with two possibilities. The first is that the kudu-like phenotype is ancestral to *Tragelaphus* and was retained in the present day lesser and greater kudu. This view goes against the majority of fossil evidence that suggests the earliest tragelaphines were small and nyala-like (Gentry, 2010, Bibi, 2011). Therefore, the second possibility is that the kudu-like phenotype of *T. imberbis* and *T. strepsiceros* is derived and evolved independently in both lineages (Fig. 5A). *T. kyaloae* from the early Pliocene could potentially represent the earliest emergence of a kudu-like phenotype (Gentry, 2010), but later fossils like *Tragelaphus* sp. nov. (Gentry, 1981) and *T. lockwoodi* (Reed and Bibi, 2013), which are more representative of *T. imberbis* and *T. strepsiceros*, respectively, only appear in the record during the middle Pliocene, supporting a derived status for the kudu-like phenotype.

The presence of horns in both sexes of non-sister bongo (*T. eurycerus*) and eland (*T. oryx*, and *T. derbianus*) is another seemingly convergent trait among the spiral-horned antelopes (Fig. 5D). This trait may have resulted from gene flow since 40 % of quartet topologies support the alternate grouping of dry-savanna species and closed-forest specialists as sister clades (node 4, Fig. 3C). Furthermore, the f-branch test and D-statistics inferred gene flow between the ancestor of the eland clade with *T. euryceros* and *T. buxtoni* (Fig. 4), which could have transmitted both the horned-female and monomorphic coat colour phenotypes to the ancestor of the bongo and mountain nyala, with the latter species losing the horned female trait but retaining the monomorphic coat colour trait through genetic drift. Pronounced sexual dimorphism in coat colour is also a trait dispersed among non-sister lineages, however its prevalence at the ancestral position in all subbranches suggests that it may have been the ancestral condition of *Tragelaphus* (Fig. 5E). It is also possible that this trait was conferred through gene flow between *T. angasii* and the ancestors of *T. scriptus*, as suggested above. We caution that many or all of these phenotypic traits are likely polygenic, and although gene flow could still have facilitated their evolution, their acquisition and retention may have been more complex than the scenarios outlined here.

In summary, trait evolution appears highly complex among Tragelaphini. While we show that some characteristics resulted from evolutionary divergence (habitat type, sexual dimorphism in coat colour of dry-savanna species), others could potentially have resulted from gene flow, followed by genetic drift (horns in both sexes, sexual dimorphism/monomorphism in imbabala/mountain nyala coat colour). On the other hand, independent or convergent evolution appears most prevalent in this group. The specialist life strategy has evolved twice independently in the basal sister clade and in the closed forest specialists, with further and extreme specialisation in the latter group to wetlands (*T. spekii*), lowland forests (*T. euryceros*) and mountain forests (*T. buxtoni*). Independent evolution is also the most a likely explanation for the presence of the kudu-like phenotype among non-sister taxa. Diet specialisation was the most variable of all phenotypic traits assessed here, with the ability to browse and graze (mixed feeding) evolving independently up to three times, and grazing evolving once in the highly specialised *T. spekii*. Although more comprehensive sampling, including several individuals of each species, would be required to fully unravel these complex patterns of evolution within this group, the current analysis provides several interesting hypotheses of divergence, gene flow and convergent evolution as potential drivers of phenotypic diversity within the spiral-horned antelopes.

Data and Resource Availability

All short-read data generated for this study were deposited in the

NCBI's short read archive (SRA), with accession numbers SAMN41052117 – SAMN41052126, under BioProject: PRJNA1103378. See also Table 1.

CRediT authorship contribution statement

Andrinajoro R. Rakotoarivelo: Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis. **Thabelo Rambuda:** Methodology, Formal analysis. **Ulrike H. Taron:** Writing – review & editing, Methodology, Formal analysis. **Gabrielle Stalder:** Resources. **Paul O'Donoghue:** Resources. **Jan Robovský:** Writing – review & editing, Visualization, Resources, Methodology, Formal analysis. **Stefanie Hartmann:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Michael Hofreiter:** Writing – review & editing, Supervision, Resources, Investigation, Funding acquisition, Conceptualization. **Yoshan Moodley:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation, 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.

Acknowledgements

We wish to thank the curators and veterinarians at Prague Zoo, namely Barbora Dobiášová, Jaroslav Šimek and Roman Vodička, Czech Republic; Port Lympne Zoo, United Kingdom; Linzer Tiergarten, Austria; Schönbrunn Zoo, Austria; Zoo Hannover, Germany and Zoo La Palmyre, France for their kind help concerning the provided samples and associated information. We further acknowledge the Chinko Project for supplying a wild giant eland sample for this study. H. Siegismund is thanked for providing a kéwel sample from the University of Copenhagen collection. We thank the Centre for High Performance Computing (CHPC) of the Republic of South Africa for access to their servers. TR was supported by a postgraduate bursary and ARR was supported by a postdoctoral fellowship, both from the National Research Foundation of the Republic of South Africa. This project was supported by ERC-Grant 310763 "GeneFlow" to MH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2024.108131.

References

Abbott, R., Albach, D., Ansell, S., et al., 2013. Hybridization and speciation. J. Evol. Biol. 26 (2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x.
Angas, G.F., 1849. Description of *Tragelaphus angasii* Gray, with some account of its habits. Proc. Zool. Soc. London 1848, 89–90.
Arnold, M.L., 2006. Evolution Through Genetic Exchange. Oxford University Press.
Berner, D., Salzburger, W., 2015. The genomics of organismal diversification illuminated by adaptive radiations. Trends Genet. 31 (9), 491–499. https://doi.org/10.1016/j.tig.2015.07.002.
Bibi, F., 2011. *Tragelaphus nakuae*: evolutionary change, biochronology, and turnover in the African Plio-Pleistocene. Zool. J. Linn. Soc. 162 (3), 699–711. https://doi.org/10.1111/j.1096-3642.2010.00691.x.
Bibi, F., 2013. A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics. BMC Evol. Biol. 13, 166. https://doi.org/10.1186/1471-2148-13-166.
Blyth, E., 1869. Notice of two overlooked species of antelope. Proc. Zool. Soc. London 55, 51–55.
Boulineau, P., 1933. Hybridations d'antilopides. La Terre Et La Vie 3, 690–691.
Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4, 7. https://doi.org/10.1186/s13742-015-0047-8.
Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associates Inc, Sunderland.
Cruickshank, T.E., Hahn, M.W., 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. 23 (13), 3133–3157. https://doi.org/10.1111/mec.12796.
Dalton, D.L., Tordiffe, A., Luther, I., Duran, A., Van Wyk, A.M., Brettschneider, H., Oosthuizen, A., Modiba, C., Kotzé, A., 2014. Interspecific hybridisation between greater kudu and nyala. Genetica 142 (3), 265–271. https://doi.org/10.1007/s10709-014-9772-7.
Dasmahapatra, K.K., Walters, J.R., Briscoe, A.D., Davey, J.W., Whibley, A., Nadeau, N.J., Heliconius Genome Consortium, 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487 (7405), 94. https://doi.org/10.1038/nature11041.
Donoghue, P.C., Benton, M.J., 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol. Evol. 22 (8), 424–431. https://doi.org/10.1016/j.tree.2007.05.005.
Dowling, T.E., Secor, C.L., 1997. The role of hybridization and introgression in the diversification of animals. Annu. Rev. Ecol. Syst. 28, 593–619. https://doi.org/10.1146/annurev.ecolsys.28.1.593.
Du Toit, J.T., Cumming, D.H.M., 1999. Functional significance of ungulate diversity in African savannas and the ecological implications of the spread of pastoralism. Biodivers. Conserv. 8, 1643–1661. https://doi.org/10.1023/A:1008959721342.
Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28 (8), 2239–2252. https://doi.org/10.1093/molbev/msr048.
Edelman, N.B., Frandsen, P.B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R.B., García-Accinelli, G., Van Belleghem, S.M., Patterson, N., Neafsey, D.E., Challis, R., 2019. Genomic architecture and introgression shape a butterfly radiation. Science 366 (6465), 594–599.
Essop, M.F., Harley, E.H., Baumgarten, I., 1997. A molecular phylogeny of some Bovidae based on restriction-site mapping of mitochondrial DNA. J. Mammal. 78 (2), 377–386. https://doi.org/10.2307/1382891.
Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., ... & Eizirik, E., 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci. Adv. 3(7), e1700299. Flagstad, A., Syvertsen, P.O., Stenseth, N.C, Jakobsen, K.S., 2001. Environmental change and rates of evolution: the phylogeographic pattern within the hartebeest complex as related to climatic variation. Proc. R. Soc. Lond. B, 268(1468), 667–677. doi: 10.1098/rspb.2000.1416.
Fontaine, M.C., Pease, J.B., Steele, A., Waterhouse, R.M., Neafsey, D.E., Sharakhov, I.V., Jiang, X., Hall, A.B., Catteruccia, F., Kakani, E., Mitchell, S.N., Wu, Y.C., Smith, H.A., Love, R.R., Lawniczak, M.K., Slotman, M.A., Emrich, S.J., Hahn, M.W., Besansky, N.J., 2015. Mosquito genomics. Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science 347 (6217), 1258524. https://doi.org/10.1126/science.1258524.
Ford, A.G., Rüber, L., Newton, J., Dasmahapatra, K.K., Balarin, J.D., Bruun, K., Day, J.J., 2016. Niche divergence facilitated by fine-scale ecological partitioning in a recent cichlid fish adaptive radiation. Evol. Int. J. Organ. Evol. 70 (12), 2718–2735. https://doi.org/10.1111/evo.13072.
Gagnon, M., Chew, A.E., 2000. Dietary preferences in extant African Bovidae. J. Mammal. 81 (2), 490–511.
Gante, H.F., Matschiner, M., Malmstrøm, M., Jakobsen, K.S., Jentoft, S., Salzburger, W., 2016. Genomics of speciation and introgression in Princess cichlid fishes from Lake Tanganyika. Mol. Ecol. 25 (4), 6143–6161. https://doi.org/10.1111/mec.13767.
Gatesy, J., Amato, G., Vrba, E.S., Schaller, G., DeSalle, R., 1997. A cladistic analysis of mitochondrial ribosomal DNA from the Bovidae. Mol. Phylogenet. Evol. 7 (3), 303–319. https://doi.org/10.1006/mpev.1997.0402.
Gentry, A.W., 2010. Bovidae. In: Werdelin, L., Sanders, W.J. (Eds.), Cenozoic Mammals of Africa. University of California Press, Berkeley, California, pp. 747–804.
Georgiadis, N.J., Kat, P.W., Oketch, H., Patton, J., 1990. Allozyme divergence within the Bovidae. Evolution 44 (8), 2135–2149. https://doi.org/10.1111/j.1558-5646.1990.tb04317.x.
Gray, A.P., 1847. Description of a new species of Antelope from West Africa. Ann. Mag. Nat. Hist. [ser. 1] 20 (133), 286.
Gray, A.P., 1972. Mammalian Hybrids. Commonwealth Agricultural Bureau, Edinburgh.
Groves, C.P., Grubb, P., 2011. Ungulate Taxonomy. The John Hopkins University Press, Baltimore.
Groves, C.P., Leslie, D.M., 2011. Family Bovidae (hollow-horned ruminants). In: Wilson, D.E., Mittermeier, R.A. (Eds.), Handbook of the Mammals of the World, Vol. 2. Lynx Edictions, Spain, pp. 447–779.
Grubb, L., 1982. Refuges and dispersal in the speciation of African forest mammals. In: Prance, G.T. (Ed.), Biological Diversification in the Tropics. Columbia University Press, New York, pp. 537–553.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59 (3), 307–321. https://doi.org/10.1093/sysbio/syq010.
Haile-Selassie, Y., Vrba, E.S., Bibi, F., 2009. Bovidae. In: Haile-Selassie, Y., Woldegabriel, G. (Eds.), *Ardipithecus Kadabba:* Late Miocene Evidence from the Middle Awash. Ethiopia. University of California Press, Berkeley, pp. 277–330.
Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species boundaries. J. Hered. 105 (Suppl. 1), 795–809. https://doi.org/10.1093/jhered/esu033.
Hassanin, A., Douzery, E.J., 1999. The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome *b* gene. Mol. Phylogenet. Evol. 13 (2), 227–243. https://doi.org/10.1006/mpev.1999.0619.
Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., van Vuuren, B.J., Matthee, C., Ruiz-Garcia, M., Catzeflis, F., Areskoug, V., Nguyen, T.T., Couloux, A., 2012. Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed

by a comprehensive analysis of mitochondrial genomes. C. R. Biol. 335 (1), 32–50. https://doi.org/10.1016/j.crvi.2011.11.002.
Hassanin, A.M.L., Houck, D., Tshikung, B., Kadjo, H Davis, Ropiquet, A., 2018. Multi-locus phylogeny of the tribe Tragelaphini (Mammalia, Bovidae) and species delimitation in bushbuck: evidence for chromosomal speciation mediated by interspecific hybridization. Mol. Phylogenet. Evol. 129, 96–105. https://doi.org/10.1016/j.ympev.2018.08.006.
Herbert, T.D., Lawrence, K.T., Tzanova, A., Peterson, L.C., Caballero-Gill, R., Kelly, C.S., 2016. Late Miocene global cooling and the rise of modern ecosystems. Nat. Geosci. 9 (11), 843–847. https://doi.org/10.1038/ngeo2813.
Hibbins, M.S., Hahn, M.W., 2022. Phylogenomic approaches to detecting and characterizing introgression. Genetics 220, iyab173.
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 (2), 518–522. https://doi.org/10.1093/molbev/msx281.
Holder, M.T., Anderson, J.A., Holloway, A.K., 2001. Difficulties in detecting hybridization. Syst. Biol. 50 (6), 978–982. https://doi.org/10.1080/106351501753462911.
Huerta-Cepas, J., Serra, F., Bork, P., 2016. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. Mol. Biol. Evol. 33 (6), 1635–1638. https://doi.org/10.1093/molbev/msw046.
IUCN SSC Antelope Specialist Group., 2016. *Tragelaphus* (errata version published in 2017). In: IUCN 2016. Version 2021-1. The IUCN Red List of Threatened Species 2016: e.T22047A115164600. www.iucnredlist.org. Accessed 03 June 2021.
Jombart, T., Kendall, M., Almagro-Garcia, J., Colijn, C., 2017. treespace: Statistical exploration of landscapes of phylogenetic trees. Mol. Ecol. Resour. 17 (6), 1385–1392. https://doi.org/10.1111/1755-0998.12676.
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 (6), 587–589. https://doi.org/10.1038/nmeth.4285.
Kingdon, J., 2013 Mammals of Africa: Volume VI: Hippopotamuses, Pigs, Deer, Giraffe and Bovids. Bloomsbury Publishing, London, United Kingdom, 680 pp.
Korneliussen, T.S., Albrechtsen, A., Nielsen, R., 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinf. 15 (1), 356. https://doi.org/10.1186/s12859-014-0356-4.
Koulischer, L., Tijskens, J., Mortelmans, J., 1973. Chromosome studies of a fertile mammalian hybrid: the offspring of the cross bongo x sitatunga (Bovoidea). Chromosoma 41 (3), 265–270. https://doi.org/10.1007/BF00344021.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Li, H., Handsaker, B., Wysoker, A., et al., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25 (16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
Liu, K.J., Steinberg, E., Yozzo, A., et al., 2015. Interspecific introgressive origin of genomic diversity in the house mouse. Proc. Natl. Acad. Sci. 112 (1), 196–201.
Lydekker, R., 1910. The spotted kudu. Nature 84, 396–397. https://doi.org/10.1038/084396d0.
Malinsky, M., Challis, R.J., Tyers, A.M., Schiffels, S., Terai, Y., Ngatunga, B.P., Miska, E. A., Durbin, R., Genner, M.J., Turner, G.F., 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science 350 (6267), 1493–1498.
Malinsky, M., Svardal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F., Durbin, R., 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nat. Ecol. Evol. 2 (12), 1940–1955.
Malinsky, M., Matschiner, M., Svardal, H., 2021. Dsuite–- fast D-statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. 21 (2), 584–595. https://doi.org/10.1111/1755-0998.13265.
Mallet, J., 2007. Hybrid speciation. Nature 446 (7133), 279–283. https://doi.org/10.1038/nature05706.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet J 17 (1), 10–12. https://doi.org/10.14806/ej.17.1.200.
Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Slazar, C., Walters, J.R., Simpson, F., et al., 2013. *Heliconius* and sympatric speciation. Genome Res. 23 (11), 1817–1828. https://doi.org/10.1101/gr.159426.113.
Masello, J.F., Quillfeldt, P., Sandoval-Castellanos, E., Alderman, R., Calderón, L., Cherel, Y., Cole, T.L., Cuthbert, R.J., Marin, M., Massaro, M., Navarro, J., Phillips, R. A., Ryan, P.G., Shepherd, L.D., Suazo, C.G., Weimerskirch, H., Moodley, Y., 2019. Additive traits lead to feeding advantage and reproductive isolation, promoting homoploid hybrid speciation. Mol. Biol. Evol. 36 (8), 1671–1685. https://doi.org/10.1093/molbev/msz090.
Matthee, C.A., Robinson, T.J., 1999. Cytochrome *b* phylogeny of the family Bovidae: resolution within the Alcelaphini, Antilopini, Neotragini, and Tragelaphini. Mol. Phylogenet. Evol. 12 (1), 31–46. https://doi.org/10.1006/mpev.1998.0573.
Mayr, E., 1963. Animal Species and Evolution. Harvard University Press, Cambridge.
McPeek, M.A., 2008. The ecological dynamics of clade diversification and community assembly. Am. Nat. 172 (6), 270–284. https://doi.org/10.1086/593137.
Michon, L., Famin, V., Quidelleur, X., 2022. Evolution of the East African Rift System from trap-scale to plate-scale rifting. Earth Sci. Rev. 231, 104089.
Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R., 2020a. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37 (5), 1530–1534.
Minh, B.Q., Hahn, M.W., Lanfear, R., 2020b. New methods to calculate concordance factors for phylogenomic datasets. Mol. Biol. Evol. 37 (9), 2727–2733.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30 (17), 541–548. https://doi.org/10.1093/bioinformatics/btu462.
Moodley, Y., Bruford, M.W., 2007. Molecular biogeography: towards an integrated framework for conserving pan-african biodiversity. PLoS One 2, e454.
Moodley, Y., Bruford, M.W., Bleidorn, C., Wronski, T., Apio, A., Plath, M., 2009. Analysis of mitochondrial DNA data reveals non-monophyly in the bushbuck (*Tragelaphus scriptus*) complex. Mamm. Biol. 74 (5), 418–422. https://doi.org/10.1016/j.mambio.2008.05.003.
Nixon, K.C., 1990. Winclada (Beta) v.0.9.9 (Published by the Author).
Norris, L.C., Main, B.J., Lee, Y., et al., 2015. Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. Proc. Natl. Acad. Sci. 112 (3), 201418892. https://doi.org/10.1073/pnas.1418892112.
Ogilby, W., 1837. A view pointing out the characters to which the most importance should be attached in establishing generic distinctions among the Ruminantia. Proc. Zool. Soc. London 4, 131–139.
Pallas, P.S., 1766. Miscellania Zoologica. Apud Petrum van Cleef, Hagae Comitum.
Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A., Inoue, J.G., Irmis, R.B., Joyce, W.G., Ksepka, D.T., Patané, J.S.L., Smith, N.D., Tarver, J.E., van Tuinen, M., Yang, Z., Angielczyk, K.D., Greenwood, J.M., Hipsley, C.A., Jacobs, L., Makovicky, P.J., Müller, J., Smith, K.T., Theodor, J.M., Warnock, R.C.M., Benton, M.J., 2012. Best practices for justifying fossil calibrations. Syst. Biol. 61 (2), 346–359. https://doi.org/10.1093/sysbio/syr107.
Racimo, F., Sankararaman, S., Nielsen, R., Huerta-Sánchez, E., 2015. Evidence for archaic adaptive introgression in humans. Nat. Rev. Genet. 16 (6), 359–371.
Rakotoarivelo, A.R., O'Donoghue, P., Bruford, M.W., Moodley, Y., 2019a. Rapid ecological specialization despite constant population sizes. PeerJ, 7, e6476. doi: 10.7717/peerj.6476.
Rakotoarivelo, A.R., O'Donoghue, P., Bruford, M.W., Moodley, Y., 2019b. An ancient hybridization event reconciles mito-nuclear discordance among spiral-horned antelopes. J. Mammal. 100 (4), 1144–1155. https://doi.org/10.1093/jmammal/gyz089.
Rubes, J., Kubickova, S., Pagacova, E., Cernohorska, H., Di Berardino, D., Antoninova, M., Vahala, J., Robinson, T.J., 2008. Phylogenomic study of spiral-horned antelope by cross-species chromosome painting. Chromosome Res. 16, 935–947.
Salazar, C., Baxter, S.W., Pardo-Diaz, C., et al., 2010. Genetic evidence for hybrid trait speciation in *Heliconius* butterflies. PLoS Genet. 6 (4), e1000930.
Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33 (7), 1654–1668. https://doi.org/10.1093/molbev/msw079.
Sclater, P.L., Thomas, O., 1900. The Book of Antelopes, Vol. 4. RH Porter.
Seehausen, O., 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19 (4), 198–207. https://doi.org/10.1016/j.tree.2004.01.003.
Speke, J.H., 1863. Journal of the Discovery of the Source of the Nile. Blackwood, London, United Kingdom.
Stajich, J.E., et al., 2002. The Bioperl Toolkit: perl modules for the life sciences. Genome Res. 12 (10), 1611–1618. https://doi.org/10.1101/gr.361602#.
Van Gelder, R.G., 1977. An eland x kudu hybrid, and the contents of the genus *Tragelaphus*. Lammergeyer 23, 1–6.
Wellborn, G.A., Langerhans, R.B., 2015. Ecological opportunity and the adaptive diversification of lineages. Ecol. Evol. 5 (1), 176–195. https://doi.org/10.1002/ece3.1347.
Willows-Munro, S., Robinson, T.J., Matthee, C.A., 2005. Utility of nuclear DNA intron markers at lower taxonomic levels: phylogenetic resolution among nine Tragelaphus spp. Mol. Biol. Evol. 35 (3), 624–636. https://doi.org/10.1016/j.ympev.2005.01.018.
Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. 19 (Suppl. 6), 153. https://doi.org/10.1186/s12859-018-2129-y.
Zheng, Y., Janke, A., 2018. Gene flow analysis method, the D-statistic, is robust in a wide parameter space. BMC Bioinf. 19 (1), 1–19. https://doi.org/10.1186/s12859-017-2002-4.