Multi-locus phylogeny of the tribe Tragelaphini (Mammalia, Bovidae) and species delimitation in bushbuck: Evidence for chromosomal speciation mediated by interspecific hybridization

Alexandre Hassanin[a], Marlys L. Houck[b], Didier Tshikung[c], Blaise Kadjo[d], Heidi Davis[b], Anne Ropiquet[e]

[a] Institut Systématique Evolution Biodiversité (ISYEB), Sorbonne Université, MNHN, CNRS, EPHE; 57 rue Cuvier, CP 51, 75005 Paris, France
[b] San Diego Zoo Institute for Conservation Research; 15600 San Pasqual Valley Road, Escondido, CA 92027, USA
[c] Faculté de médicine vétérinaire; Université de Lubumbashi, 243 BP 1825, The Democratic Republic of the Congo
[d] Université Félix-Houphouët-Boigny, UFR Biosciences; 22 BP 582, Abidjan 22, Cote d'Ivoire
[e] Middlesex University, Department of Natural Sciences, Faculty of Science and Technology, The Burroughs, Hendon, London NW4 4BT, United Kingdom

Abstract
The bushbuck is the most widespread bovid species in Africa. Previous mitochondrial studies have revealed a polyphyletic pattern suggesting the possible existence of two distinct species. To assess this issue, we have sequenced 16 nuclear genes and one mitochondrial fragment (cytochrome b gene + control region) for most species of the tribe Tragelaphini, including seven bushbuck individuals belonging to the two divergent mtDNA haplogroups, Scriptus and Sylvaticus. Our phylogenetic analyses show that the Scriptus lineage is a sister-group of Sylvaticus in the nuclear tree, whereas it is related to Tragelaphus angasii in the mitochondrial tree. This mito-nuclear discordance indicates that the mitochondrial genome of Scriptus was acquired by introgression after one or several past events of hybridization between bushbuck and an extinct species closely related to T. angasii. The division into two bushbuck species is supported by the analyses of nuclear markers and by the karyotype here described for T. scriptus (2n = 57 M/58F), which is strikingly distinct from the one previously found for T. sylvaticus (2n = 33 M/34F). Molecular dating estimates suggest that the two species separated during the Early Pleistocene after an event of interspecific hybridization, which may have mediated massive chromosomal rearrangements in the common ancestor of T. scriptus

Keywords
Spiral-horned antelopes; Species complex; Introgressive hybridization; Chromosomes; Cytogenetics

Introduction
In Sub-Saharan Africa, the bushbuck, Tragelaphus scriptus (Pallas, 1766), is the most widespread species of the family Bovidae (Cetartiodactyla, Ruminantia). It is found in 40 countries, where it occurs in open forests, bush savannahs and woodlands at altitudes between 0 and 4000 m. From a taxonomic point of view, the bushbuck belongs to the tribe Tragelaphini, a monophyletic group of spiral-horned bovids that contains eight other species of the genus Tragelaphus (Kingdon and Hoffmann, 2013; IUCN, 2018): (1) T. angasii (nyala), which lives in Lowveld vegetation in south-eastern Africa (Malawi, Mozambique, South Africa, Swaziland, and Zimbabwe); (2) T. buxtoni (mountain nyala), which is endemic to the highlands of Ethiopia; (3) T. derbianus (giant eland), which is found in the Sudanian savannah woodlands from Senegal to South Sudan; (4) T. eurycerus (bongo), which is distributed in the forest-savannah mosaics of West and Central African lowlands as well as the Kenya highlands; (5) T. imberbis (lesser kudu), which occupies bushlands and thickets in the Horn of Africa; (6) T. oryx (common eland), which occurs throughout the savannah woodlands of eastern and southern Africa; (7) T. spekii (sitatunga), which is found in the lowland forests of the Congo Basin, and near swamp areas within the savannahs of central, eastern and southern Africa, as well as in West Africa where populations are mainly located in forests bordering river deltas; and (8) T. strepsiceros (greater kudu), which lives in thickets and dense woodlands throughout eastern and southern Africa. Morphologically, T. scriptus differs from other tragelaphine species by its lower body weight (females: 24–60 kg; males: 30–80 kg) and its smaller horns (25–57 cm; only found in males) (Kingdon, 1997). There is significant geographic variation in body and horn size, coat coloration, pattern and length of the pelage (Kingdon and Hoffmann, 2013). As a consequence, a high number of subspecies has been described in the literature, e.g., eight in Grubb (2005), 11 in Kingdon and Hoffmann (2013), and 28 in Lydekker (1914). Several of these subspecies were treated as full species in some classifications, e.g., five species in Sclater and Thomas (1894–1900), and eight species in

Groves and Grubb (2011) and Castelló (2016).

Using mitochondrial sequences from the control region and cyto-chrome b gene, Moodley and Bruford (2007) concluded that two divergent haplogroups of bushbuck exist, distributed on either side of the Rift Valley: the Scriptus haplogroup in western-northern sub-Saharan Africa, from Senegal to northern Ethiopia in the North to Angola in the South; and the Sylvaticus haplogroup in eastern-southern sub-Saharan Africa, from southern Ethiopia in the North to South Africa in the South. Analyses of cytochrome b sequences have suggested that these two divergent mitochondrial lineages do not form a monophyletic assemblage (Moodley et al., 2009), but most relationships among species of Tragelaphus were poorly supported due to the use of short sequences (556 nt). The analyses of complete mitochondrial genomes (length of the alignment: 14,902 nt) have recovered the polyphyly of T. scriptus with high support values (Hassanin et al., 2012): the Scriptus haplogroup was allied with T. angasii whereas the Sylvaticus haplogroup was related to the clade uniting T. eurycerus and T. spekii. These mitochondrial results suggest that bushbuck should be classified into two rather than one species.

However, numerous studies on mammals have shown that phylogenies based only on mtDNA data may lead to erroneous taxonomic interpretations. In particular, several studies using both mtDNA and nuDNA markers have shown mtDNA introgression in different species of the family Bovidae, such as the European bison, banteng, and the Pliocene ancestor of wild goat (Ropiquet and Hassanin, 2006; Hassanin and Ropiquet, 2007; Nijman et al., 2008; Hassanin et al., 2013). Therefore, in order to provide definitive taxonomic conclusions, we consider that the mtDNA polyphyly of T. scriptus needs to be confirmed with additional evidence based on nuclear genes.

In this report we sequenced a mtDNA fragment of around 1800 nt, which covers the complete cytochrome b gene and the control region, for several bushbuck collected in Republic of Côte d'Ivoire (CI), and in the province of Haut-Katanga in southern Democratic Republic of the Congo (DRC), where Scriptus and Sylvaticus mtDNA haplogroups may be found in sympatry. Our new mtDNA sequences were compared to those previously published (Moodley and Bruford, 2007; Moodley et al., 2009; Hassanin et al., 2012). In addition, we sequenced 16 nuclear fragments for all species of Tragelaphini except T. buxtoni, as well as for seven bushbuck belonging to the two divergent mtDNA haplogroups, Scriptus and Sylvaticus. The karyotype of six West African bushbuck was investigated through conventional Giemsa staining and GTG-banding techniques. The three main purposes of our study were (i) to establish if there are two species of bushbuck, (ii) to decipher phylogenetic relationships within the tribe Tragelaphini, and (iii) to provide a molecular time scale for this group.

2. Material and methods

2.1. DNA extraction, amplification and sequencing

Total genomic DNA was extracted from muscle or skin samples and cells using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). A mitochondrial DNA (mtDNA) fragment covering the complete cytochrome b gene and the 5′ end of the control region was amplified as detailed in Hassanin et al. (2009; 2012) using three overlapping primer sets: (1) 14,724/CBL402; (2) GluMA/Pro; (3) U844/L482. Sixteen autosomal genes of the nuclear genome were amplified using primers published in previous papers: exon 4 of CSN3 (Ropiquet and Hassanin, 2005b), intron 8 of FGB (Hassanin and Ropiquet, 2007), intron 2 of LALBA (Hassanin and Douzery, 2003), the promoter of LF (Hassanin and Douzery, 1999a), intron 9 of PRKCI (Ropiquet and Hassanin 2005a), intron 13 of SPTBN1 and intron 9 of TG (Ropiquet and Hassanin, 2006), intron 6 of CCAR1, intron 2 of CHPF2, intron 4 of EXOSC9, intron 10 of HDAC2, intron 19 of DIS3, intron 2 of PABPN1, intron 6 of RIOK3, intron 9 of TUFM, and intron 6 of ZFYVE27 (Hassanin et al., 2013).

Amplicons were purified and sequenced by Eurofins Genomics (Ebersberg, Germany) with the primers used for PCR amplification. The sequences were edited and assembled using Sequencher 5.1 (Gene Codes Corporation). Heterozygous sites in the nuclear loci (double peaks) were coded using the IUPAC code. Sequences generated for this study were deposited in the GenBank database (accession numbers MH792162–MH792418; see Appendices A and B for details). Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.08.006.

2.2. Phylogenetic analyses

The 22 new mtDNA sequences (cytochrome b + control region) of bushbuck generated for this study were aligned on AliView 1.22 (Larsson, 2014) with those published in Moodley and Bruford (2007) and Moodley et al. (2009) and those extracted from the complete mitochondrial genomes published in Hassanin et al. (2012). Our final alignment contains 1901 characters for 206 sequences (available upon request to corresponding author) which includes 183 bushbuck and 23 outgroup taxa of the tribes Caprini, Boselaphini, Bovini, and Tragelaphini. A Bayesian tree was constructed as detailed below for other alignments using the resources available from the CIPRES Science Gateway (Miller et al., 2010).

The phylogeny of the tribe Tragelaphini was investigated by analyzing 17 independent genes (the mtDNA fragment and 16 nuclear genes) on a reduced sample of 30 taxa including seven bushbuck. For each gene, DNA sequences were aligned with AliView 1.22 (Larsson, 2014). The best-fitting models of sequence evolution were found under jModelTest 2.1.7 (Darriba et al., 2012). Using the Akaike Information Criterion (AIC), we selected the K80 + G model for LF (342 nt) and TG (786 nt), the HKY+I model for TUFM (860 nt), the HKY+G model for CSN3 (418 nt), PRKCI (525 nt), SPTBN1 (599 nt) and ZFYVE27 (701 nt), the GTR model for EXOSC9 (993 nt), HDAC2 (722 nt), and PABPN1 (896 nt), the GTR+I model for DIS3 (851 nt), FGB (695 nt), LALBA (494 nt), and RIOK3 (722 nt), the GTR+G model for CCAR1 (1031 nt) and CHPF2 (890 nt), and the GTR+I+G model for the mtDNA fragment (1901 nt) and the concatenation of the 16 nuclear markers (nuDNA, 11,525 nt).

Phylogenetic analyses were conducted using Bayesian inference (BI) and maximum likelihood (ML) methods, with gaps treated as missing data. Bayesian inferences were carried out under MrBayes v3.2.1 (Ronquist et al., 2012) on each of the 17 markers and on the nuDNA dataset. The posterior probabilities (PP) were calculated using four independent Markov chains run for 10,000,000 Metropolis-coupled MCMC generations, with tree sampling every 1000 generations and a burn-in of 25%. The nuDNA dataset was also analyzed using the Bayesian approach implemented in BEAST v.2.4.7 (Bouckaert et al., 2014; see the Section 2.3 for details) and for ML reconstruction under PhyML 3.1 (Guindon et al., 2010). ML bootstrap values (BP) were computed after 1000 replicates using the GTR+I+G model and SPR-based tree search algorithm.

The lists of bipartitions obtained from the Bayesian analyses of the 17 independent markers (the mtDNA fragment and 16 nuclear genes) were transformed into a weighted binary matrix using SuperTRI v57 (Ropiquet et al., 2009; Python script available at http://www.normalesup.org/~bli/Programs/programs.html). The matrix includes 2820 binary characters, in which each binary character corresponds to a node, which was weighted according to its probability in one of the 17 lists of bipartitions. In this way, the SuperTRI method takes into account both principal and secondary signals because all phylogenetic hypotheses found during the analyses are represented in the weighted binary matrix used for supertree construction.

Add Figure 1 here

Fig. 1. Mitochondrial tree of 183 mitochondrial haplotypes sequenced for African bushbucks. Most of the mitochondrial sequences of bushbucks were extracted from the study of Moodley and Bruford (2007) and Moodley et al. (2009) (see Appendix A for details on accession numbers). Individuals highlighted in red are those specially analyzed for this study. The asterisk indicates a node supported by PP ≥ 0.95. The map showing the distribution of Scriptus and Sylvaticus groups was interpreted from the IUCN (2018) and the results of Moodley and Bruford (2007). Black circles indicate haplotypes used for mitochondrial dating and red arrows are the individuals used for nuclear phylogeny binary matrix used for supertree construction.

The reliability of the nodes was assessed using three measures proposed in Ropiquet et al. (2009): supertree bootstrap percentages (SBPs) were obtained under PAUP* v4b10 (Swofford, 2003) after 1000 BP replicates of the matrix of 2820 binary characters generated by SuperTRI v57; the mean posterior probabilities (MPP) and reproducibility indices (Rep) were directly calculated on SuperTRI v57. The SBP, MPP and Rep values were reported on the Bayesian tree found with the nuDNA dataset (Fig. 2). SuperTRI analyses were conducted to test for repeated phylogenetic signals in the different independent genes. If a node is recovered with high SBP, MPP, and Rep values (SBP > 50; MPP > 0.5; Rep > 0.5), most markers. If a node is recovered with low Rep values (≤0.12), we can

conclude that the signal is restricted to one or two markers.

Pairwise nucleotide distances were calculated on PAUP* version 4b10 (Swofford, 2003) for both mtDNA and nuDNA datasets.

2.3. Molecular dating

Divergence times were estimated using the concatenated nuclear dataset (30 taxa; 11,525 nt) or a mitochondrial dataset (38 taxa; 1901 nt) and the Bayesian approach implemented in BEAST v.2.4.7 (Bouckaert et al., 2014).

ADD FIGURE 2 HERE

Fig. 2. Multi-locus phylogeny of the tribe Tragelaphini. The Bayesian tree was reconstructed from the concatenated nuclear dataset combining all the 16 autosomal genes (30 taxa and 11,525 nucleotide characters) with a GTR + I + G model. For each internal branch, the two first values correspond to the Bayesian posterior probability (PP, to the left of the slash) and maximum likelihood bootstrap percentage (BP, the right of the slash). The asterisk (*) indicates that the node was supported by maximal values of robustness (PP = 1; BP = 100). The values below are the three measures of reliability obtained from the SuperTRI analyses of the 17 markers (16 nuclear genes and the mitochondrial fragment): from left to right, Supertree Bootstrap percentage (SBP), Mean posterior probability (MPP) and the Reproducibility index (Rep). Thick branches indicate nodes that were considered reliable (for more details, see Results section). The letter “X” indicates that the node was not found in the analysis. For Tragelaphini, the number of exclusive synapomorphies corresponding to either transitions (Ti) and transversions (Tv) are indicated at the nodes, and the position and nature of all diagnostic indels (i: insertion; d: deletion) shared by at least two taxa in the alignments of nuclear genes are highlighted in grey.

For the nuclear dataset, we applied three time constraints modelled using a normal distribution: (1) the most recent common ancestor (MRCA) of Bovidae was set at 20 ± 2 Mya, as the first occurrence of bovids is well documented in the Early Miocene of Africa and Eurasia (Vrba and Schaller, 2000); (2) the MRCA of African Tragelaphini was set at 7 ± 1 Mya, in agreement with paleontological evidence (Kostopoulos and Koufos, 2006; Gentry, 2010) and previous molecular estimations (Hassanin et al., 2012); and (3) the split between American bison and European bison was set at 230 ± 10 ka, as the first entrance of bison into North America was dated between 220 and 240 ka (Scott, 2010). For the mitochondrial dataset, the third calibration point was not used because of the polyphyly of Bos bison, and the MRCA of African Tragelaphini was set at 6.75 ± 0.1 Mya, in agreement with the results obtained from the nuclear dataset.

The analyses were performed with the models of nucleotide substitution selected under jModelTest 2.1 for each marker (see above) and a relaxed-clock model with uncorrelated lognormal distribution for substitution rates. Node ages were estimated using a calibrated Yule speciation prior and 108 generations, with tree sampling every 2000 generations, and a burn-in of 25%. Adequacy of chain mixing and MCMC chain convergence were assessed using the ESS values in Tracer v.1.6. The chronograms were reconstructed with TreeAnnotator and visualized with FigTree v.1.4.1 (software available at http://www. tree. bio.ed.ac.uk/software/).

2.4. Cell cultures and chromosome preparations

Fibroblast cell lines were established from four male and two female T. scriptus skin biopsies and accessioned into San Diego Zoo Global’s Frozen Zoo®. All of the samples were acquired from captive populations in the USA. The exact origin of the founder animals is undocumented, but thought to be from Sierra Leone or Liberia (personal communication K. Benirschke and M. Jones). Metaphase chromosomes of all six individuals were examined by non-differential staining using Giemsa; both females and two males were additionally GTG-banded. Cell culturing, harvesting, and chromosome banding followed the techniques described by Kumamoto et al. (1996). G-banded chromosomes were compared and numbered to the Bos taurus standard karyotype (Popescu et al., 1996). Karyotyping was done using the CytoVision Genus® system by Leica Microsystems.

3. Results

3.1. Mitochondrial haplogroups

The Bayesian tree inferred from the mitochondrial alignment is shown in Fig. 1. The species T. scriptus appears to be polyphyletic, as two divergent haplogroups (p-distances between 12.2% and 16.2%), named Scriptus and Sylvaticus, are related to two different species: Scriptus is a sister-group of T. angasii, whereas Sylvaticus is grouped with T. eurycerus and T. spekii.

Within the Scriptus haplogroup, we identified three subgroups, named Sc1, Sc2 and Sc3, that are separated by p-distances comprised between 5.1% and 9.2% (Appendix C for more details) and are characterized by low levels of inter-individual polymorphism (Sc1: < 2.6%; Sc2: < 3.4%; Sc3: < 3.0%). Within the Sylvaticus haplogroup, we found eight subgroups, named Sy1-Sy8, characterized by low levels of inter-individual polymorphism (Sy1: < 3.4%; Sy2: < 1.3%; Sy3: < 2.5%; Sy4: < 2.8%; Sy5: < 2.4%; Sy6: < 2.0%; Sy7: < 2.4%; Sy8: < 3.0%). The nucleotide p-distances calculated between subgroups Sy4, Sy5, Sy6, Sy7 and Sy8 are comprised between 2.9% and 5.8%. The three other subgroups, Sy1, Sy2 and Sy3, show higher levels of divergence with other subgroups: 6.4–9.2% for Sy1; 5.1–7.1% for Sy2; and 4.9–7.2% for Sy3 (Appendix C for more details).

The 25 bushbuck analyzed for this study are highlighted in red in Fig. 1. Three karyotyped bushbuck from San Diego Zoo belong to three karyotyped bushbuck from San Diego Zoo, four samples from CI, one sample from Cameroon, and seven samples from Haut-Katanga in DRC. The other samples belong to haplogroups Sy1 (five samples from Haut-Katanga), Sy2 (one sample from South Africa), and Sy8 (one sample from Tanzania). Among them, the seven samples indicated with an arrow in Fig. 1 were further sequenced for 16 nuclear markers.

3.2. Multi-locus analyses

The concatenated dataset combining the 16 autosomal nuclear genes contains 11,525 aligned positions. The Bayesian tree reconstructed from this dataset is depicted in Fig. 2. The tribes Bovini, Boselaphini and Tragealphini are found monophyletic with maximal support values (PP = 1; BP = 100), and these three groups are highly supported by the separate analyses of the 17 independent markers (the mt fragment and 16 nuclear genes) (SBP = 100; 0.69 < MPP < 1; 0.76 < Rep < 1; see also the trees of Appendix D). They are also characterized by a large number of exclusive synapomorphies, including both substitutions and deletions (Fig. 2).

Regarding interrelationships between the three tribes of Bovinae, the nuDNA dataset supports the clade uniting Bovini and Tragelaphini (PP = 1; BP = 89; SBP = 61; MPP = 0.37). This node was recovered in six separate Bayesian analyses of the 17 markers (Rep = 0.35), but with PP > 0. 5 for only five markers: CHPF2 (PP = 0.55), CSN3 (PP = 0.88), EXOSC9 (PP = 0.79), RIOK3 (PP = 1), and SPTBN1 (PP = 0.61). The two other hypotheses, i.e., Boselaphini grouped to either Bovini or Tragelaphini, are less supported (BP ≤ 10; SBP ≤ 22; MPP ≤ 0.25). The association between Boselaphini and Bovini is found in five separate analyses of the 17 markers (Rep = 0.29), but with PP > 0.5 for only three markers: CCAR1 (PP = 0.88), FGB (PP = 0.79), and mtDNA (PP = 0.66). The association between Boselaphini and Tragelaphini is found in three separate analyses of the 16 markers (Rep = 0.18): LF (PP = 0.81), PRKCI (PP = 0.61) and ZFYVE27 (PP = 0.78). All other markers provided a weak signal for relationships between the three tribes.

The tribe Tragelaphini is monophyletic with maximum support values (PP = 1; BP = 100) and this group is also supported by all separate analyses of the 17 independent markers (SBP = 100; MPP/ REP = 1; Appendix D). The tribe can be diagnosed by 154 mutations, including 11 deletions and two insertions (Fig. 2).

All the seven bushbuck are grouped together with maximum support values (PP = 1; BP = 100). In addition, this group was found in eight separate analyses of the 17 markers (Rep = 0.47), with PP > 0.5 for six markers: CCAR1 (PP = 0.99), CHPF2 (PP = 0.81), DIS3 (PP = 0.99), PABPN1 (PP = 0.97), SPTBN1 (PP = 0.71) and TUFM (PP = 0.94). Two divergent lineages corresponding to Scriptus and Sylvaticus clusters were also supported by maximum support values (PP = 1; BP = 100). The Scriptus lineage, as represented by the four individuals C7R31, K9x105, CIx1 and CIx4, is found monophyletic with 11 of the 17 markers (Rep = 0.65), with PP > 0.5 for 10 markers: CCAR1 (PP = 1), CHPF2 (PP = 0.90), DIS3 (PP = 0.99), EXOSC9 (PP = 0.98), FGB (PP = 0.99), LALBA (PP = 0.95), LF (PP = 0.97),

mtDNA (PP = 1), PABPN1 (PP = 1), and ZFYVE27 (PP = 1). The Sylvaticus lineage, as represented by the three individuals K9x578, MBL11 and PhC17, is found monophyletic with 12 of the 17 markers (Rep = 0.71), with PP > 0.5 for 11 markers: CCAR1 (PP = 0.96), CHPF2 (PP = 1), FGB (PP = 1), LALBA (PP = 0.96), LF (PP = 1), mtDNA (PP = 1), PABPN1 (PP = 1), RIOK3 (PP = 0.88), TG (PP = 1) and TUFM (PP = 1), and ZFYVE27 (PP = 0.97). Each of these two lineages can be diagnosed by more than 20 specific mutations (Fig. 2).

Interrelationships within both Scriptus and Sylvaticus clusters cannot be resolved. Within Scriptus, the analysis of the nuDNA dataset suggests a basal division separating two groups: (1) CIx1 + C7R31; and (2) CIx4 + K9x105. However, these two groups are supported by PP > 0.5 with only one of the 17 markers, HDAC2 (Appendix D). The analyses of mtDNA data rather favor a grouping of C7R31 with CIx4 (PP = 1; Fig. 1 and Appendix D). Within Sylvaticus, the analysis of the nuDNA dataset suggests a sister-group relationship between K9x576 and MBP11, but here again this node is supported by PP > 0.5 with only one of the 17 markers (RIOK3), and the mtDNA dataset rather supports the grouping of MBP11 with PhC17 (PP = 1; Fig. 1 and Appendix D).

Within the tribe Tragelaphini, the nuclear analyses show strong support (PP = 1; BP/SBP = 99–100; MPP ≥ 0.46; Rep ≥ 0.47) for only two other nodes: the monophyly of the species T. spekii and the sister-group relationship between the two eland species, T. derbianus and T. oryx. Other relationships are generally weakly supported despite the fact that all markers provide several parsimony informative sites for interrelationships within Tragelaphus (between 4 and 29). The lack of robustness is therefore the consequence of discordance between gene trees rather than lack of phylogenetic signal in the molecular markers. For instance, there are six different phylogenetic hypotheses supported by PP > 0.95 for the position of the species T. imberbis (Appendix D): (1) it appears divergent from other species of Tragelaphus with CCAR1 and mtDNA; (2) with DIS3, it is grouped with T. derbianus, T. eurycerus, T. oryx, T. spekii, and T. strepsiceros; (3) with EXOSC9, it is related to T. angasii, T. derbianus, T. eurycerus, T. oryx, T. strepsiceros; (4) with FGB, it is associated to all other species of Tragelaphus, except T. derbianus and T. oryx; (5) with HDAC2, it is enclosed with T. derbianus, T. eurycerus, T. oryx, T. scriptus, and T. spekii; (6) with LF, it is the sister-species of T. angasii.

3.3. Time scale of Tragelaphini

BEAST chronograms inferred using either the concatenation of the 16 nuclear genes or the mitochondrial fragment are depicted in Fig. 3A and B, respectively. Nuclear results suggest that the tribe Tragelaphini diversified rapidly into three major lineages corresponding to T. imberbis, T. angasii, and a third lineage including all other species of Tragelaphus at the end of the Late Miocene epoch, between 6.75 and 5.90 Mya. With the nuDNA dataset, the MRCA of bushbuck is dated at 2.12 Mya, an age which is similar to the divergence between the two eland species, T. derbianus and T. oryx (1.94 Mya). The MRCA of the Scriptus group is estimated at 0.39 Mya and that of the Sylvaticus group at 0.78 Mya.

To better understand mito-nuclear discordance, we also estimated divergence times on a mitochondrial alignment including 38 taxa and 1901 nt using 6.75 ± 0.1 Mya for the MRCA of Tragelaphini, in agreement with the age estimated from the nuDNA dataset. The nuclear and mitochondrial dates are comparable (i.e., overlapping 95% HPD intervals) for most nodes of the nuclear tree. By contrast, the dates estimated for the two bushbuck lineages are not overlapping, as the mitochondrial ages are much older than the nuclear ages: 2.31 versus 0.39 Mya for the MRCA of Scriptus, and 2.61 versus 0.78 Mya for the MRCA of Sylvaticus.

3.4. Cytogenetic analyses of West African bushbuck

The diploid chromosome number of the six West African bushbuck was 57/58 (M/F) (Fig. 4A, non-differentially stained karyotype of a male). Chromosomes of the two females in our study (KB3521 and KB5261) were all acrocentric. The karyotypes of the four males (KB3478, KB3518, KB6342, and KB14341) also consisted of all acrocentric elements with the exception of one large nearly metacentric chromosome derived from fusion of an autosome (BTA 13) to the Y (Fig. 4B, GTG-banded male). The number of chromosomal arms (Fundamental Number, FN) was constant across all six studied individuals,

ADD FIGURE 3 HERE

Fig. 3. Bayesian divergence times (in million years ago, Mya) estimated using the nuclear concatenation of 16 genes (A) or the mitochondrial fragment (B).

Divergence times were estimated with BEAST 2.4.7 (see main text for details). Taxa other than Tragelaphini were removed from the figures. Bold values at the nodes are mean ages. Grey bars and values between brackets represent the 95% Highest Posterior Density (HPD) interval

ADD FIGURE 4 HERE

Fig. 4. Non-differentially stained karyotype (A) and GTG-banded chromosomes (B) of a male of West African bushbuck (Scriptus group), 2n = 57. Chromosomes are aligned to Bos taurus (BTA) standard. The karyotypes show a Y-autosomal translocation involving BTA 13 to the Y (Fig. 4 B, GTG-banded male). The number of chromosomal arms (Fundamental Number, FN) was constant across all six studied individuals, FN = 58.

The cytochrome b gene was sequenced for all the six karyotyped individuals. As anticipated considering the putative geographic origin of the captive population (Sierra Leone or Liberia), we found that all individuals fall into the Scriptus haplogroup (Fig. 1): KB3478, KB5261 and KB6342 individuals belong to the Sc1 subgroup, where they are closely related to bushbuck from Sierra Leone; KB3518, KB3521 and KB14341 individuals share the same haplotype, which occupies a basal position into the Sc3 subgroup.

4. Discussion

4.1. Evidence for a strong mito-nuclear discordance for the position of the Scriptus group

The phylogeny of the tribe Tragelaphini has been previously investigated using mitochondrial or/and nuclear DNA sequences (Hassanin and Douzery, 1999a,b; Matthee and Robinson, 1999; Willows-Munro et al., 2005; Moodley et al., 2009; Hassanin et al., 2012), as well as chromosomal rearrangements (Rubes et al., 2008). Our mitochondrial tree (Fig. 1) strongly supports the monophyly of Tragelaphini, the early divergence of T. imberbis, the polyphyly of T. scriptus, the sister-group relationship between T. angasii and the Scriptus group, the existence of a large clade containing all other taxa of Tragelaphini, the monophyly of eland species (T. derbianus + T. oryx), the sister-group relationship between T. eurycerus and T. spekii, and their association with the Sylvaticus group. All these results are in full agreement with previous mitochondrial studies based on the complete mitochondrial genome (Hassanin et al., 2012), the concatenation of three mitochondrial genes (Cytb, 12S and 16S rRNAs; Willows-Munro et al., 2005), or the Cytb gene (Hassanin and Douzery, 1999a,b; Matthee & Robinson, 1999; Moodley et al., 2009). However, our nuclear tree reconstructed from the concatenation of 16 genes (11,511 nt) shows a different position for the Scriptus lineage. Indeed, it appears as a sister-group of the Sylvaticus lineage (Fig. 2), instead of being related to T. angasii as found in the mitochondrial tree (Fig. 1). The node uniting Scriptus and Sylvaticus lineages is supported by the analyses of six independent nuclear genes. Other nuclear genes do not provide a robust phylogenetic signal (Appendix D). Excepting this strong mito-nuclear discordance, other nodes are congruent with our mitochondrial tree and the previous nuclear tree of Willows-Munro et al. (2005) based on four loci (MGF, PRKCI, SPTBN1, and THY). However, basal relationships within Tragelaphus are poorly supported by our nuclear dataset, which means that the relationships between T. angasii, T imberbis and other tragelaphines are uncertain, as well as the placement of T. strepsiceros (Fig. 2).

4.2. Mitochondrial introgression in the common ancestor of the Scriptus group

In the family Bovidae, comparisons between mtDNA and nuDNA markers have revealed several cases of mitochondrial introgression in which the mtDNA genome from one species has been transferred to another species: (i) from the ancestor of the Himalayan tahr (Hemitragus) to the common ancestor of Capra species (Ropiquet and Hassanin, 2006); (ii) from a species related to the aurochs (Bos primigenius) to the ancestor of European bison, Bos bison bonasus (Verkaar et al., 2004; Hassanin and Ropiquet, 2004; Hassanin et al., 2013), and (iii) from the ancestor of kouprey (Bos sauveli) to the ancestor of Cambodian banteng, Bos javanicus birmanicus (Hassanin and

Ropiquet, 2007). Four characteristics of these bovid species pairs may have favored interspecies hybridization: (i) they are closely related species; (ii) they shared the same habitat, sometimes appearing as mixed groups (e.g. banteng and kouprey in Cambodia); (iii) they are gregarious species with (iv) a polygynous behavior system, in which dominant males mate with multiple females. Finally, we have to assume that their ancestral karyotypes were compatible in order to allow normal meiotic process during gamete production, at least in hybrid females. The mechanism of mitochondrial introgression probably took place in three evolutionary steps: (i) the female(s) of one species crossed with male(s) of another species (interspecies hybridization); (ii) F1 hybrid females were able to produce offspring with the introgressed species (backcross) whereas F1 hybrid males were probably sterile (Haldane's rule) or at least with reduced fitness, preventing introgression of paternal genes; (iii) the new mtDNA genome has been fixed in the introgressed species because of selection and/or genetic drift in small populations due to a severe population bottleneck.

In the case of bushbuck, there are two divergent mitochondrial haplogroups. The Sylvaticus haplogroup is found closely related to T. eurycerus and T. spekii, as in the nuclear tree, whereas the Scriptus haplogroup occupies an anomalous position, as a sister-group of T. angasii. Accordingly, we can assume that the ancestral mitochondrial genome of both Scriptus and Sylvaticus groups had a greater similarity to that currently found in the Sylvaticus lineage, while the Scriptus haplogroup was inherited from another species, now extinct, which diverged from T. angasii around 4.83 Mya (Fig. 3B). The phylogeographic pattern of Scriptus haplotypes (Fig. 1) indicates the existence of three divergent subgroups, Sc1, Sc2 and Sc3. Haplotypes Sc1 were detected in bushbuck of the Guinean forests (Guinea Bissau and Sierra Leone), whereas haplotypes Sc2 were found in those of open forests from Cameroon in the West to Ethiopia in the East through CAR, Chad, northern DRC, Sudan and Uganda. In contrast, the distribution of Sc3 haplotypes can be divided into two geographic regions: (i) West Africa, in which most of the Sc3 haplotypic diversity was detected (mainly in CI and Ghana, but also in Togo and Nigeria), and (ii) Central Africa, where Sc3 haplotypes were found in Cameroon, R. Congo, and southern DRC. Such a phylogeographic pattern indicates that bushbuck possessing Sc3 haplotypes dispersed more recently in southern regions (R. Congo and southern DRC), which means that the Scriptus group originated in the Sudano-Guinean zone of the Northern Hemisphere. Since the common ancestor of the three mitochondrial haplogroups Sc1, Sc2 and Sc3 was dated at 2.31 Mya (Fig. 3B), we can conclude that interspecific hybridization between bushbuck and an extinct species occurred somewhere in the Sudano-Guinean zone during the Early Pleistocene.

4.3. How many species of bushbuck?

Bushbuck populations display an extraordinary amount of variation in body size, coat color and patterns of white and dark brown markings on the head, throat, legs, belly and flanks (Kingdon and Hoffmann, 2013). Many species were described when the interior of Africa was explored by European geographers during the 19th century and at the beginning of the 20th century (e.g., Sclater and Thomas, 1894–1900; Neumann, 1902; Matschie, 1912). They were synonymized with T. scriptus in Lydekker (1914) and subsequent classifications (e.g., Ansell, 1972; Grubb, 2005). More recently, however, several classifications have divided bushbuck populations into eight species (e.g., Groves and Grubb, 2011; Castelló, 2016). It is evident that these authors have been influenced by the high levels of nucleotide diversity found by Moodley and Bruford (2007) in the mitochondrial sequences of bushbuck sampled all over the African continent (11.7% for the control region and 7.4% for the Cytb gene). However, Moodley et al. (2009) have suggested that the two divergent mitochondrial haplogroups Scriptus and Sylvaticus may represent two distinct species. They also cautioned that "such species delineation is not recommendable based on mtDNA data only" and concluded that their mitochondrial results need to be corroborated by nuclear data.

In order to assess this issue, we have sequenced 16 nuclear genes in most species of Tragelaphini, including seven bushbuck individuals representing divergent mitochondrial haplotypes of the two hap logroups Scriptus and Sylvaticus (Fig. 2). Although discordant with the mitochondrial tree for the position of the Scriptus group, our nuclear tree support the existence of two sister-species of bushbuck. Particularly relevant is the fact that the two species T. scriptus and T. sylvaticus are monophyletic in the separate analyses of 10 and 11 markers, respectively (Appendix D; other markers did not provide a robust signal at this level of the tree). Both molecular dating analyses, based on either mitochondrial or nuclear datasets, indicate that the two species T. scriptus and T. sylvaticus

diverged during the Early Pleistocene (Fig. 3). The karyotype here described for T. scriptus (2n = 57 M/58F chromosomes; Fig. 4) is strikingly distinct from the one previously published for T. sylvaticus from South Africa (2n = 33 M/34F chromosomes; Wallace, 1977: Kruger National Park; Rubes et al., 2008: National Zoological Gardens of Pretoria; Western Cape Province). Kingdon (1997) pointed out that "Western forest forms are 'harnessed', with both vertical and horizontal white body stripes and numerous spots on haunches. Eastern and southern 'sylvan' populations are sometimes plain and often sparsely marked with a few light spots and streaks on flanks or haunches." In other words, this means that the two species T. scriptus and T. sylvaticus can be distinguished phenotypically, genetically and cytogenetically.

4.4. Chromosomal speciation mediated by interspecific hybridization?

Our Bayesian estimation of divergence times indicates that the diversification of mitochondrial haplotypes began around 2.61 Mya for T. sylvaticus and around 2.31 Mya for T. scriptus, which is in agreement with the divergence times previously estimated by Moodley and Bruford (2007) using another method and different calibration points (2.7 and 2.5 Mya, respectively for Scriptus and Sylvaticus groups). With the nuclear dataset, we estimated that the separation of T. scriptus from T. sylvaticus took place around 2.12 Mya, i.e., approximately at the same time of the diversification of mitochondrial haplotypes of T. scriptus. Such concomitance suggests that the event of interspecific hybridization, which resulted in the mitochondrial introgression from an extinct species (closely related to T. angasii) into the common ancestor of T. scriptus, was at the origin of the speciation event between T. scriptus and T. sylvaticus. The two species have 2n = 57/58 and 2n = 33/34 chromosomes, respectively, and their karyotypes are divergent from those found in closely related species, such as T. eurycerus and T. spekii, as well as T. derbianus, T. oryx, and T. strepsiceros (see Rubes et al., 2008 for more details). Given the important cytogenetic differences between the two bushbuck species, we suggest that inter-specific hybridization may have mediated massive chromosomal rearrangements involving apparently extensive centric fusion/fission, which in turn have promoted speciation. Karyotypes of both bushbuck species have the same number of chromosomal arms (Fundamental Number; FN = 58) which suggests centric fission as the primary form of chromosomal rearrangement. We plan to further study this issue using a cytogenomic approach combining full genome sequencing and molecular cytogenetic (FISH) to ascertain more precisely the chromosomal homology between the two bushbuck species.

Acknowledgments

We are very grateful to those who provided or helped to collect samples analyzed in our study: Martin Bam, Philippe Blot, Philippe Chardonnet, André Délicat, Bernard Dutrillaux, Céline Canler, Jean-Pierre Hugot, Sylvie Laidebeure, Alexis Lécu, Flobert Njiokou, Célestin Pongombo Shongo, Claire Rejaud, Jacques Rigoulet, Roland Simon, Bettine Van Vurren, and Vitaly Volobouev. We thank Suellen Charter and Julie Fronczek for cell culture and karyotyping, Kurt Benirschke and the late Marvin Jones for research on the origin of the six animals studied at SDZ. This work was supported by the MNHN, CNRS, 'PPF Biodiversité actuelle et fossile', and 'Consortium national de recherche en génomique'. It is part of agreement No. 2005/67 between Genoscope and MNHN on the projects 'SpeedID' and 'Bibliothèque du Vivant'. We acknowledge Huw Jones and Cynthia Steiner for helpful proof reading.

References

Ansell, W.F.H., 1972. Order Artiodactyla. pp. 1-84 in The mammals of Africa: an identification manual. In: Meester, J., Setzer, H.W. (Eds.), Smithsonian Institution Press, Washington, D.C.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST2: a software platform for Bayesian evolutionary analysis. PLOS Comput. Biol. 10, e1003537.

Castelló, J.R., 2016. Bovids of the World: Antelopes, Gazelles, Cattle, Goats, Sheep, and Relatives. Princeton University Press, Princeton, New Jersey (USA), pp. 664.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9, 772.

Gentry, A.W., 2010. Bovidae. In: Werdelin, L., Sanders, W.J. (Eds.), Cenozoic mammals of Africa. University of California Press, Berkeley, California, pp. 741–796.

Groves, C., Grubb, P., 2011. Ungulate Taxonomy. Johns Hopkins University Press, Baltimore, Maryland, pp. 317.

Grubb, P., 2005. Order Artiodactyla. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal Species of the World: A Taxonomic and Geographic Reference, third ed., 2 vols. Washington, DC: Smithsonian Institution Press, 1. pp. 637–722.

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. Systematic Biol. 59 (3), 307–321.

Hassanin, A., Douzery, E.J., 1999a. Evolutionary affinities of the enigmatic saola (Pseudoryx nghetinhensis) in the context of the molecular phylogeny of Bovidae. P. Roy. Soc. B-Biol. Sci. 266, 893–900.

Hassanin, A., Douzery, E.J., 1999b. The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome b gene. Mol. Phylogenet. Evol. 13, 227–243.

Hassanin, A., Douzery, E.J., 2003. Molecular and morphological phylogenies of Ruminantia and the alternative position of the Moschidae. Syst. Biol. 52, 206–228.

Hassanin, A., Ropiquet, A., 2004. Molecular phylogeny of the tribe Bovini (Bovidae,Bovinae) and the taxonomic status of the Kouprey, Bos sauveli Urbain 1937. Mol. Phylogenet. Evol. 33, 896–907.

Hassanin, A., Ropiquet, A., 2007. Resolving a zoological mystery: the kouprey is a real species. P. Roy. Soc. B-Biol. Sci. 274, 2849–2855.

Hassanin, A., Ropiquet, A., Couloux, A., Cruaud, C., 2009. Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe Caprini (Bovidae, Antilopinae). J. Mol. Evol. 68 (4), 293–310.

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, 32–50.

Hassanin, A., An, J., Ropiquet, A., Nguyen, T.T., Couloux, A., 2013. Combining multiple autosomal introns for studying shallow phylogeny and taxonomy of Laurasiatherian mammals: application to the tribe Bovini (Cetartiodactyla, Bovidae). Mol. Phylogenetics Evol. 66, 766–775.

IUCN, 2018. The IUCN Red List of Threatened Species. Downloaded on 29 June 2018. <www.iucnredlist.org>.

Kingdon, J., 1997. The Kingdon Field Guide to African Mammals. Academic Press, San Diego, California, pp. 464.

Kingdon, J.S., Hoffmann, M., 2013. The Mammals of Africa. VI. Pigs, Hippopota M. muses, Chevrotain, Giraffes, Deer, and Bovids. Bloomsbury Publishing, London, UK.

Kostopoulos, D.S., Koufos, G.D., 2006. Pheraios chryssomallos, gen. et sp. nov. (Mammalia, Bovidae, Tragelaphini), from the Late Miocene of Thessaly (Greece): Implications for Tragelaphin biogeography. J. Vertebrate Paleontol. 26 (2), 436–445.

Kumamoto, A.T., Charter, S.J., Houck, M.L., Frahm, M., 1996. Chromosomes of Damaliscus (Artiodactyla, Bovidae): simple and complex centric fusion arrangements. Chrom. Res. 4, 614–621.

Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large data sets.

Bioinformatics 30 (22), 3276–3278.

Lydekker, R., 1914. Catalogue of the ungulate mammals in the British Museum. Natural History) vol. 3. Subfamily xvi. Tragelaphinae. London: British Museum. pp. 150–229.

Matschie, V.P., 1912. Die von Herrn Major P. H. G. Powell-Cotton gesammelten Rassen der Gattung Tragelaphus. Sitzungs-berichte der Gesellschaft Naturforschender Freunde zu Berlin 10, 544–567.

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.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA. pp. 1–8.

Moodley, Y., Bruford, M.W., 2007. Molecular biogeography: towards an integrated framework for conserving pan-African biodiversity. PLoS One 2 (5), 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. Mammalian Biol. 74 (5), 418–422.

Neumann, O., 1902. Sitzung vom 15. April 1902. sprach über neue nordost-und osta-frikanische Säugethiere. Sitzungs-berichte der Gesellschaft Naturforschender Freunde zu Berlin 4, 93–102.

Nijman, I.J., van Boxtel, D.C.J., van Cann, L.M., Marnoch, Y., Cuppen, E., Lenstra, J.A., 2008. Phylogeny of Y chromosomes from bovine species. Cladistics 24, 723–726.

Popescu, C.P., Long, S., Riggs, P., Womack, J., Schmutz, S., Fries, R., Gallagher, D.S.,1996. Standardization of cattle karyotype nomenclature: report of the committee for the standardization of the cattle karyotype. Cytogenet. Cell Genet. 74, 259–261.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61,539–542.

Ropiquet, A., Hassanin, A., 2005b. Molecular phylogeny of Caprines (Bovidae, Antilopinae): the question of their origin and diversification during the Miocene. J. Zool. Syst. Evol. Res. 43, 49–60.

Ropiquet, A., Hassanin, A., 2005a. Molecular evidence for the polyphyly of the genus Hemitragus (Mammalia, Bovidae). Mol. Phylogenet. Evol. 36, 154–168.

Ropiquet, A., Hassanin, A., 2006. Hybrid origin of the Pliocene ancestor of wild goats. Mol. Phylogenet. Evol. 41, 395–404.

Ropiquet, A., Li, B., Hassanin, A., 2009. SuperTRI: a new approach based on branch support analyses of multiple independent data sets for assessing reliability of phylogenetic inferences. C. R. Biol. 332, 832–847.

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. Chrom. Res. 16 (7), 935–947.

Sclater P.L., Thomas O., 1894–1900. The book of Antelopes. vol. IV. Porter, London,
242 pp.

Scott, E., 2010. Extinctions, scenarios, and assumptions: Changes in latest Pleistocene large herbivore abundance and distribution in western North America. Quatern. Int. 217, 225–239.

Swofford, D.L., 2003 PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version

4.0b. Sinauer Associates, Sunderland, MA.

Verkaar, E.L., Nijman, I.J., Beeke, M., Hanekamp, E., Lenstra, J.A., 2004. Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? Mol. Biol. Evol., 21 (2004), pp. 1165-1170

Vrba, E., Schaller, G.B., 2000. Antelopes, deer, and relatives. fossil record, behavioral ecology, systematics and conservation, Yale University Press, New Haven and London, 341 p.

Wallace, C., 1977. Chromosome analysis in the Kruger National Park: the chromosomes of the bushbuck (Tragelaphus scriptus). Cytogenet. Cell Genet. 18 (1), 50–56.

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. Phylogenet. Evol. 35 (3), 624–636