Rapid ecological specialization despite constant population sizes

Andrinajoro R. Rakotoarivelo[1, 3] , Michael W. Bruford[2] and Yoshan Moodley[1]

[1]Department of Zoology, University of Venda, Private Bag X5050, Thohoyandou 0950, Republic of South Africa.

[2]Cardiff School of Biosciences, Sir Martin Evans Building, Cardiff University, Museum Avenue, Cardiff, CF10 3AX, United Kingdom.

[3]Natiora Ahy Madagasikara, Antananarivo 101, Madagascar.

Corresponding author:
Andrinajoro R. Rakotoarivelo[1]
Email address: andrinajoro@moov.mg

ABSTRACT

Background. The bushbuck, *Tragelaphus scriptus*, is the most widespread and ecologically diverse ungulate species complex within the spiral-horned antelopes, occurring in approximately 73% of the total land area of sub-Saharan Africa. This species was found to consist of two genetically divergent lineages based on the mitochondrial (mt)DNA control region. One lineage inhabited the north-western half of the African continent (*T. scriptus*) while the other lineage (*T. sylvaticus*) was found in the south-eastern half. The complex was also found to comprise an unprecedented example of 23 phylogenetically distinct groups ('ecotypes'), with montane and desert phenotypes potentially resulting from convergent evolution. The current study aim to test hypotheses regarding historical demography and adaptation of bushbuck using a higher-resolution framework, with faster evolving nuclear markers(MGF, PRKCI, SPTBN, and THY) as well as three further mitochondrial markers (cytochrome b, 12S rRNA, and 16S rRNA).

Methods. Genealogies were reconstructed for the nuclear and mitochondrial data sets and for each gene independently to test the non-monphyly of the bushbuck complexe in a multi loci framework. In addition, we reconstruct the phylogeographic history of the bushbuck complex by a Bayesian discrete phylogeographic approach of our nucDNA data set to investigate its geographic diffusion and ancestral sequence location.

Results. We uncovered two evolutionarily divergent lineages and geographically restricted lineages (*Sylvaticus* and *Scriptus*) of bushbuck using phylogenetics. Molecular dating indicates that these lineages last shared a common ancestor ~2.54 million years ago. Summary statistics and analysis of the frequency

distributions of DNA polymorphisms do not have any support for expanding population. Both BSPs and EBSPs indicate that the *Scriptus* and *Sylvaticus* lineages have remained relatively stable during the last 225-450Kya.

Discussion. Both nucDNA and mtDNA support previously findings of two genetically divergent *Sylvaticus* and *Scriptus* lineages, despite them coming into secondary contact in several geographic regions. The three mtDNA loci confirmed 15 of the previously defined ecotypes, including those with convergent phenotypes. However, the nuclear tree showed less phylogenetic resolution at the more derived parts of the genealogy, possibly due to incomplete lineage sorting of the slower evolving nuclear genome. The only exception to this was the montane ecotype *meneliki* of the Ethiopian highlands, which formed a monophyletic group at three of the four nucDNA loci. The independent evolution of this group relative to phenotypically similar montane ecotypes in Africa confirm previously suggestions of convergence within the bushbuck complex.

Keywords: Phylogeography, Ecological adaptation, Phenotypic convergence,

INTRODUCTION

The bushbuck (*Tragelaphus scriptus*) is a well-known, highly diverse species complex of spiral-horned antelopes. This species is unique, being the most widespread and ecologically diverse of any ungulate species, occurring in approximately 73% of the total land area of sub-Saharan Africa. Across this vast and heterogeneous region, bushbuck occur in most habitat types (Moodley & Bruford 2007) from forested to xeric zones and ranging in altitude from sea-level to 4000m.

Phenotypic diversity among bushbuck populations is unprecedented, with at least 42 subspecies describe across its range (Lydekker, 1914; Grubb, 1985). The complex can be subdivided into two divergent morphological groups which inhabit the western and northern (*Scriptus* group) and eastern and southern (*Sylvaticus* group) parts of the species range (Fig 1). *Scriptus* is smaller and less dimorphic, but it possesses a heavily striped white harness-like pattern, whereas most populations of the larger *Sylvaticus* have little to no striping at all. Although known to favour areas of thick cover wherever they occur, bushbuck do not inhabit the dense rainforest of the Congo basin, preferring the mosaic landscapes at its fringe. The two groups are therefore separated in the west and south by the Lower Congo valley and the Congo basin respectively, but in eastern Africa *Scriptus* and *Sylvaticus* come into secondary contact from the northern end of the Albertine rift along the Imatong and Didinga Mountains of South Sudan following the rift into the Ethiopian Highlands (white arrows, Fig. 1). Within this zone of contact, the phenotypic integrity of each form may be maintained through habitat preference; the *Scriptus* form inhabits the lowlands while the large, dark and heavy-coated *Sylvaticus* montane ecotype inhabits the high altitude forests, although evidence of gene flow has been observed (Moodley & Bruford 2007).

Genetically, the bushbuck complex comprises a paraphyletic mitochondrial (mt)DNA clade (Moodley et al., 2009), with divergent *Scriptus* and *Sylvaticus* super-lineages (Moodley & Bruford, 2007). The complex was also found to be highly structured into 23 phylogenetically distinct haplogroups (*Scriptus* 8; *Sylvaticus* 15), each with differing levels of ecological specialization. Among the more specialized haplogroups, the montane ecotypes (*meneliki*, *powelli*, *barkeri* and *delamerei*), and more xeric-adapted ecotypes (*decula*, *dodingae*, *fasciatus1*, *fasciatus2* and *roualeyni*) appear to have evolved more than once through convergent evolution (Moodley & Bruford 2007). Much of the mtDNA variation in the complex is structured according to ecoregion (Olson et al., 2001), suggesting local ecological conditions as a driver for the evolution of specialization. Ecological conditions are in turn driven by a combination of local geology and an oscillating Pleistocene paleoclimate (Vrba 1995; Bobe &

Behrensmeyer, 2004; Fernández & Vrba, 2005). However, where the species evolved and its subsequent routes of colonization and diversification are still a matter of speculation.

Despite the research potential of this system, only mtDNA data have been generated for this species to date. Not only is the mitochondrial genome a single locus, it is also maternally inherited so mtDNA structure may not be representative of nuclear DNA (nucDNA) structure in species with sex biases in dispersal/phylopatry. Genetic drift is also more effective in sorting non-segregating mtDNA lineages as their effective population size is four times smaller than segregating nucDNA. Therefore, whether the nuclear genome is structured similarly, or even whether *Scriptus* and *Sylvaticus* constitute different nuclear lineages, is unknown. This problem reflects the difficulty in obtaining a representative sample for a species that ranges across such a wide distribution, and in which many regions are politically unstable. Furthermore, demographic analyses that may evidence population responses to paleo-environmental conditions and a spatially-informed phylogeographic analysis of origins and colonisation routes have never been carried out.

To test the hypotheses of variation, structure and potential adaptation purported by previous mtDNA work, we sequenced representative bushbuck from across the species range using a higher-resolution multilocus framework of four nuclear introns, complemented by three further mtDNA markers. We further reconstructed both the demographic and phylogeographic histories of the bushbuck complex using this new data set to shed further light on the evolution of this remarkable species.

MATERIALS & METHODS

Taxon sampling

A total of 27 bushbuck individuals (excluding outgroups) were included in this study (Table S1). Samples sourced previously by Moodley & Bruford (2007) were re-extracted and representatives of all 23 ecotypes were selected Fig. 1; Supplementary Table S1). As outgroup, we used both the distantly related *Bos taurus* as well as the most closely related lesser kudu (*Tragelaphus imberbis*) to root trees in several of the phylogenetic analyses.

DNA sequencing

Four nuclear intron DNA markers (MGF - mast cell growth factor, PRKCI - protein-kinase CI, B-spectrin non-erythrocytic 1 - SPTBN, and THY - thyrotropin) were amplified and sequenced in the 27 individuals above using published primers (Matthee et al., 2001). Additionally mtDNA sequences were

amplified and sequenced from three mtDNA cytochrome b (Cyt b), 12S rRNA, and 16S rRNA (for mtDNA PCR and primer details see (Arnason, Gullberg & Widegren, 1993; Simonsen, Siegismund & Arctander, 1998). In order for downstream comparison of summary statistics, the same number of individuals were sequenced for each locus. Sequences from each gene were first aligned using ClustalW (Thompson et al., 1994) as implemented in BioEdit (Hall, 1999), using default settings and thereafter manually to optimize homology. All heterozygous sites in the nucDNA were coded using the appropriate IUB code. Model selection for the best fitting substitution model for each gene was conducted in jModelTest (Posada, 2008; Darrida et al., 2012) under the Bayesian information criterion, which was preferred over the Akaike information criterion, to guard against over parameterization by averaging the likelihood over all included parameters.

Analysis of Genetic Diversity and positive selection

The seven markers were analyzed separately and then concatenated into a single data matrix, an mtDNA and a nucDNA matrix. The number of variable sites, number of parsimony informative sites and nucleotide frequencies were estimated for each data matrix in MEGA 7 (Kumar et al., 2016). Further, for each gene we calculated standard diversity statistics for each locus in DnaSP 5.0 (Librado and Rozas, 2009). These include: number of polymorphic sites (s), number of haplotypes, haplotype diversity (Hd), nucleotide diversity (Pi), and average number of pairwise differences per sequence (k). Summary statistics were also calculated for the total data and for each major clade inferred form phylogenetic analyses.

We used several analyses to test our seven loci for neutrality. The McDonald and Kreitman test (MKT) was used to detect signatures of selection and measure the amount of adaptive evolution within a species at the molecular level. Under this test, a neutrality index (NI) quantifies the direction of departure from neutrality, comparing the ratio of non-synonymous to synonymous variation between species (Dn/Ds) with the ratio of non-synonymous to synonymous variation within species (Pn/Ps). NI was calculated using the Standard and Generalized McDonald-Kreitman Test (MKT; Egea et al., 2008) website. Because silent mutations are neutral, a neutrality index lower than 1 (i.e. NI < 1) indicates an excess of non-silent divergence, which occurs when positive selection is at work in the population. When positive selection is acting on the species, natural selection favors a specific phenotype over other phenotypes, and the favored phenotype begins to go to fixation in the species as the allele frequency for that phenotype increases (Biswas and Akey, 2006). Furthermore, we used the coalescent parameters Tajima's D (Tajima, 1989) and Fu's Fs (Fu 1997) to test for deaprtures from the neutral theory and these were calculated in DnaSP v5.

Phylogenetic analyses

Phylogenetic reconstruction was performed using both maximum likelihood (ML) and Bayesian approaches using the software Garli 2.0 (Zwickl, 2006) and BEAST v2.4.5 (Bouckaert et al., 2014) respectively. The total data matrix was partitioned by gene, with the parameters of nucleotide substitution models (12S - HKY + I + G, 16S – HKY, Cyt b- HKY+I, MGF - TIM1 + I, PRKCI – HKY, SPTBN – HKY, THY - TIM1ef + I) and unlinked across partitions. Each ML analysis was initiated from a random starting tree, with nodal support assessed using 1000 bootstrap replicates. A 50 % majority rule consensus tree was constructed using the CONSENSE program in the PHYLIP package (Felsenstein, 2005). Using BEAST, five independent runs of 1 billion generations each were performed; each run consisted of four Monte Carlo Markov chains (MCMC), with topologies sampled every 100000 generations. The program Tracer 1.6 (Rambaut et al., 2014) was used to determine that the effective sample size (ESS) had reached > 200 for all parameters. In each simulation the first 20% of generations were discarded as burn-in. Genealogies were also reconstructed for the nuclear and mitochondrial data sets and for each gene independently using the same MCMC parameters.

Molecular dating

We dated our nuclear phylogeny using only the nuclear DNA data set, since the mtDNA of bushbuck are paraphyletically related (Moodley et al. 2009), and so mitochondrial branch lengths may be upwardly biased. Multiple fossil calibration points were used to scale nodal height estimation. We calibrated the bushbuck divergence based on the earliest appearance of *T. scriptus* s.l. in the fossil record known from Kenya (Leakey and Harris, 2003), and Ethiopia (Kalb et al, 1982) as early as 3.9 Mya and a minimum age of constraint of 2.58 Mya as suggested by Hassanin and Douzery (1999). An exponential distribution is used with a 2.5% probability quantile set at the age of the fossil with hard bound at the youngest bound and a soft maximum bound, beyond which it is unlikely that the divergence actually occurred. Our last calibration point constrained the evolution of the tribe Tragelaphinii 5.72 Mya (95% probability, 4.7-6.7 Mya; Deino et al., 2002). In the latter case normal distribution was used allowing for the actual node age to be equally younger or older than the fossil record. Phylogenetic relationships and divergence times were estimated using an uncorrelated relaxed lognormal Bayesian molecular clock approach in BEAST v. 2.4.5 software (Bouckaert et al., 2014). A Yule speciation process was applied to the tree inference through the MCMC (Markov chain Monte Carlo) with a random starting tree. All other parameters were the same as in previous analysis.

Inferring historical demography

In addition to Tajima's D and Fu's Fs, which may also be used to infer demography in neutrally evolving loci, demographic changes in both clades were also inferred from the observed mismatch distribution for each of the populations, calculating the raggedness index (R2) according to the population expansion model in DnaSP. This measure quantifies the smoothness of the observed mismatch distribution, with lower raggedness characterizing population that experienced sudden expansion, whereas higher raggedness values suggest stationary or bottlenecked populations (Harpending et al., 1993; Harpending, 1994). Lastly, changes in effective population size were inferred using Bayesian Skyline Plots (BSP: Drummond et al., 2005). These plots utilize the coalescent properties of gene trees to plot population size changes over time, and were inferred for the *Scriptus* and *Sylvaticus* clades using BEAST (Bouckaert et al., 2014). In order to incorporate stochastic differences between gene genealogies in the estimation of population parameters, we constructed multi-locus Extended Bayesian Skyline Plots (EBSP; Heled and Drummond, 2008) for each clade. In addition, EBSP estimates posterior probabilities for the number of population size change events. A mitochondrial divergence rate of 0.056 per million years was used (Arbogast & Slowinski, 1998) as well as appropriate inheritance scalars were used to account for potential difference in effective population size between mtDNA and nucDNA. The lengths of the MCMC chains were set to 1 billion to achieve effective sample sizes (ESS) and proper mixing of Markov chains. To account for biases due to genetic structure (Ho and Shapiro, 2011), we divided the data into *Scriptus* and *Sylvaticus* groups and reconstructed their demographic histories separately.

Bayesian phylogeographic reconstruction

We attempted to reconstructed the phylogeographic history of two major clades of the bushbuck complex using our nucDNA data set within a Bayesian discrete phylogeographic approach of in BEAST 1.8.4 (Lemey et al., 2009; Drummond et al., 2012). We used three geographical states corresponding to the continental regions where both lineage is present: west (W), east (E), and south (S). These phylogeographic analyses were run under a constant-size coalescent model, molecular clock parameterised as described above and with a random starting tree as tree model. Bayesian Stochastic Search Variable Selection (BSSVS) was used to identify those rates (colonization routes) that were frequently invoked to explain the diffusion process (Lemey et al., 2009). The maximum clade credibility (MCC) tree was computed and annotated using the BEAST module TreeAnnotator v1.8.4 (Drumond et al. 2012). We then used SpreaD3 v0.9.6 (Bielejec et al 2016; available: https://github.com/phylogeography/SpreaD3) to project the MCC phylogeny onto a spatial framework and summarized the full posterior distribution of trees to calculate the 95% highest posterior density (HPD) of node locations. The final result were visualized in Google Earth (http://earth.google.com/).

Genetic variation and its relationship to taxonomy and biogeography

To test whether nucDNA supported the hypothesis that ecology has driven genetic diversification in this complex (Moodley & Bruford, 2007), we tested the fit of a comprehensive biogeographic model (Olsen et al. 2001) to the nucDNA data, relative to that of taxonomic and geographic models using a multiple regression on genetic distance matrices (MRM), implemented in DISTLM (Anderson, 2004). MRM involves a multiple regression of a response matrix on any number of explanatory matrices, where each matrix contains distances or similarities. Pair-wise genetic distances of nucDNA data between all 27 samples was used as the response matrix. The MRM method also allows the use of covariables to assess a models conditional effect on that of explanatory matrices. We defined the basic units for the taxonomy model relative to the proposed phenotypic classification of the bushbuck based on the combined classifications of Grubb-Best (Best, 1962; and Grubb ,1985) used in Moodley & Bruford (2007). A matrix of geographic coordinates (latitude and longitude) was included as a covariable to assess the possible the effect of isolation-by-distance (IBD) on the model being tested. In a widely ranging species, IBD may significantly influence genetic structure due to the geographic distance separating the widely distributed sampling locations. MRM method allows the quantification of this effect, conditional on that of biogeography and taxonomy. .

RESULTS

This study generated a total DNA sequence alignment of 4675 bp, of which ingroup taxa accounted for 387 segregating sites. Nuclear introns were less diverse (2596 bp, 33 segregating sites) than mitochondrial genes (2596 bp, 33 segregating sites, see Table 2). All DNA sequences were found to be evolving neutrally (MKT: $\chi 2$ P>0.1).

Structure and divergence

Phylogenetic analyses of mitochondrial (Fig 2A) and nuclear (Fig. 2B) multilocus alignments yielded highly concordant ML topologies. Both marker sets recovered two well supported *Scriptus* and *Sylvaticus* lineages, although the level of phylogenetic resolution was much higher for mtDNA, which recovered all 23 clades originally identified using control region DNA (Moodley & Bruford 2007). By contrast nuclear introns identified the lineage of the Kidepo bushbuck (*T. s. dodingae*) as well as a Nile-Abyssinian (*bor-decula*) bushbuck clade within *Scriptus.* The *Sylvaticus* clade was also less structured, with the montane Menelik`s bushbuck (*T. s. meneliki*) being ancestral and the only resolvable clade. However, montane *T. s. barkeri* and *T. s. delamerei*, both lineages of the xeric –zone Somali bushbuck (*T.*

252 *s. fasciatus*), as well as Luangwa and Angolan bushbuck lineages were characterized by higher nuclear
253 divergence (Fig. 2B).
254
255 Bayesian dating of nuclear DNA loci estimated the coalescence of all ingroup gene tree lineages
256 to the late Pliocene-early Pleistocene 2.5 – 2.62 Mya (95% HPD, Fig. 3). Divergence within each group
257 occurred relatively recently in the Late Pleistocene. *Scriptus* lineages coalesced between 0.10 – 0.48 Mya
258 (95% HPD) and the Nile-Abyssinian bushbuck clade to 0.03 – 0.22 Mya (95% HPD). Divergence within
259 *Sylvaticus* was slightly earlier between 0.33 – 0.95 Mya (95% HPD) and 0.16 – 0.47 Mya (95% HPD) for
260 non-Menelik's bushbuck lineages.
261
262 **Demographic analyses**
263 We found both Fu's Fs and Tajima's D indices to be slightly negative among nuclear and
264 mitochondrial loci, at the group and species levels (Table 3). However, only locus SPTBN1 returned
265 statistically significant indices, allowing a rejection of the neutrality/constant population size null
266 hypothesis at the species level. Furthermore, the frequencies of pair-wise differences within each
267 population were also consistent with a null hypothesis of constant population size, with non-significant
268 raggedness indices (R2) for all mismatch distributions (Table 3). Additionally, single locus Bayesian
269 skyline analyses based on mtDNA indicated that the effective population sizes of both *Scriptus* and
270 *Sylvaticus* have remained relatively stable during the last 225-450Kya (Fig. 4A). Finally, multilocus
271 extended Bayesian skyline analyses of nuclear introns supported mtDNA results showing a relatively
272 stable effective population size for both lineages (Fig. 4B).
273
274 **Bayesian phylogeographic reconstruction**
275 We used a discrete Bayesian phylogeographic approach to reconstruct patterns of spatial dispersal
276 and the ancestral location for the origin of the species complex. Results (not shown) were very similar
277 with or without geographically informed priors. Within *Scriptus*, the analysis separated a well-supported
278 *dodingae-decula* clade in the east, from bushbuck inhabiting regions across the Nile and further west
279 (including the Nile bushbuck (*T. s. bor*, Fig 5A). *Sylvaticus* also comprised significant phylogeographic
280 structuring, with Menelik's bushbuck most ancestral, but forming a geographic clade with other East
281 African lineages, separated from coastal and southern African lineages (Fig 5B). North-East Africa,
282 specifically Ethiopia, was identified as the most likely ancestral location for the origin of the bushbuck
283 radiation (Fig 5C). From this origin, dispersal events were invoked in a westward direction for *Scriptus*
284 and in a southward direction for *Sylvaticus*, both events occurring on either side of the Congo basin.
285

Ecological adaptation

MRM analysis revealed that biogeography explained a significant 95% of the nuclear genetic variation within the species complex (Table 4). Taxonomic designation and geographic distance accounted respectively for 77% and 26% of the variation, with only the latter significant. Under the conditional influence of isolation by distance, both biogeographic and taxonomic models account for 41% and 65% of the genetic variation respectively.

DISCUSSION

Patterns of genetic diversity

Genetic diversity was high across the species complex. Similar display high levels of genetic diversity have been observed in leopard and African buffalo (Spong, Johansson & Björklund, 2000; Smitz et al., 2013). In addition, genetic diversity was higher for organelle mtDNA than nucDNA. This is expected since is generally higher than that of the nucDNA (Nei and Kumar 2000). Effective mtDNA population sizes are also a quarter than of nuclear DNA populations, encouraging lineage sorting through stronger genetic drift. The higher diversity of *Sylvaticus* is consistent with an earlier divergence time relative to *Scriptus* (Fig. 3).

Origins, divergence and secondary contact

Fossil records from the mid-Pliocene (approximately 3.9 Mya) of proto-bushbuck are known from several sites in eastern and southern Africa. *T. scriptus* remains were recovered in Ethiopia (Kalb et al. 1982) and Kenya (Harris et al., 1998; Leake and Harris, 2003). We observed a more recent diversification of *Sylvaticus* and *Scriptus* lineages. Since these fossils predate the estimated divergence within the bushbuck, they suggest a possible ancestral origin from north-east Africa. This is indeed the inference from our Bayesian phylogeographic reconstruction, supporting an origin for the species in Ethiopia. Until the late Pliocene, north-east Africa was densely forested habitat (Partridge, Wood & deMenocal, 1995; Reed, 1997), supporting the idea that ancestral bushbuck were both forest dwelling and used its peculiar harnessed striping pattern as an adaptation for camouflage in closed habitats (Moodley & Bruford 2007). There is some evidence that striping patterns among other bovids are also associated with living in forest habitat (Stoner et al., 2003).

The past 2 - 3 Mya has seen a major paleoclimatic shift that led to the expansion of grassland habitats in Africa, consequently inducing a drastic change in bovid species composition, specifically in north-east Africa (Bobe and Behrensmeyer, 2004; Hernandez Fernandez and Vrba, 2006; Maslin, 2007).

This also coincided with major geomorphological processes along the Gregory and Albertine Rifts (Vrba, 1995; Reed, 1997). The combination of paleoclimatic shifts and tectonic uplift have shaped the phylogeography of terrestrial African vertebrates (Flagstad et al., 2001; Trauth et al., 2007; Lorenzen et al., 2010; Voelker et al., 2010; Faulkes et al., 2011; Barlow et al., 2013; Jacobs et al., 2013). The *Scriptus-Sylvaticus* divergence can also be traced back to this time, and their extant distributions on either side of the Rift Valley (Fig. 1) suggest vicariance of the two lineages, on the basis of the major tectonic uplift events along the East African Rift system. Since divergence, however, *Scriptus* and *Sylvaticus* appear to have remained geographically isolated. The expansion of the rainforest belt in Central Africa could potentially have limited gene flow during wet interglacial cycles. However, increased secondary contact may have been possible during glacial cycles, especially between lowland and montane ecotypes *dodingae/barkeri* and *decula/meneliki* in East Africa and *bor/dianae* and *phaleratus/ornatus* south. Nevertheless, we found no evidence of haplotype/allele sharing between *Scriptus* and *Sylvaticus*, suggesting that gene flow between them was limited. A further analysis with whole genome sequences may yet shed further light on the evolution of resilience in this species.

A stable Pleistocene demographic history

Both bushbuck lineages appear to have been demographically stable through the mid to late Pleistocene (Table 3, Fig. 4), despite most of the diversity within each lineage having evolved during this time. This is a surprising result, as the Pleistocene is known for its dramatic climatic fluctuations. Ungulate population sizes are inherently linked with climate change over evolutionary timescales (Lorenzen et al. 2011), and the distributions of herbivores would presumably have shifted in accordance with vegetation change. Yet, during this time of evolutionary change, bushbuck little evidence of demographic change since the *Scriptus-Sylvaticus* divergence.

Rapid ecological specialization

Demographic stability also appears to be at odds with high levels of variation observed both morphologically and genetically. The extant genetic diversity in both *Sylvaticus* and *Scriptus* was generated in the late Pleistocene, <1 Mya, but with most divergences occurring within the last 0.5 Mya. Much of this diversity is reflected in mitochondrial DNA (Fig 2A), and has been described previously (Moodley & Bruford 2007). Although, fewer divergence events were identified with nuclear intron sequences, a large proportion of the nuclear sequence diversity could be attributed to biogeography, even when conditioned on geography (Table 4). This lends strong support to the hypothesis that local ecology has helped shape the structure of genetic diversity in this species.

By dating our nuclear tree we were also able to estimate a reliable timeframe for the onset of divergence events in the species complex. Within *Sylvaticus*, Menelik's bushbuck (*T. s. meneliki*) was first to diverge into cooler habitats of the Ethiopian massif. Larger size, a darker and thicker coat are typical of several mammalian montane ecotypes (egs. Red squirrel, *Paraxerus palliates;* Saola, *Pseudoryx nghetinhensis*). Bergman's rule predicts an increase in size among colder-adapted species (Bergmann, 1847; Freckleton et al., 2003; Clauss et al., 2013), whereas darker and thicker coats help in thermoregulation (Mills and Hes, 1997; Amy & Kunz, 2012). The early differentiation of montane Menelik's bushbuck, and the more recent evolution of other montane ecotypes (eg. *T. s. barkeri*, *T. s. delamerei*) strengthens evidence for the independent convergence of the montane phenotype among *Sylvaticus* bushbuck.

The Somali bushbuck (*T. s. fasciatus*) is also large in size and is able to survive deep into the xeric interior of the Horn of Africa along the watercourses of the Wabi Shebelle and the Juba River. This ecotype comprises two paraphyletic mtDNA lineages (Fig 2A) and independent nuclear lineages (Fig. 2B), suggesting the bushbuck colonized the Somali arid zone through two migration or range expansion events of different coastal bushbuck populations from the south.

Within *Scriptus*, the Nile-Abyssinian bushbuck (*T. s. bor*-*T. s. decula)*) clade diverged into the more open, drier habitats of the mosaic region on the fringes of the Sahel, whereas other populations that remained in more closed forested regions remained undifferentiated at the nuclear sequence level. This is reflected in phenotype, as most *Scriptus* populations are strikingly patterned with the typical bushbuck "harness", striping is reduced in those *Scriptus* populations in more open habitats such as *T. s. bor*, *T. s. decula* and *T. s. dodingae*. There is also a suggestion of reduced patterning among *Sylvaticus* bushbuck. Although much less strikingly coloured, individuals in some *Sylvaticus* populations such as the Chobe bushbuck (*T. s. ornatus*) and the Ituri bushbuck (*T. s. dianae*) may be more heavily patterned with vertical and horizontal stripes and spots. However, such individuals become rarer in populations to the south where habitats are drier and more open. A similar loss of patterning occurs across the north-south range of the plains zebra, which is also suggested to be in response to open drier environments (Rau 1978, Leonard 2005).

Conclusions

In the present study, we sequenced mitochondrial and nuclear DNA 27 individuals representing the range of distinct ecotypes previously described within the bushbuck complex. Phylogenetic congruence was observed between mitochondrial and nuclear markers, both identifying two genetically divergent lineages (*Scriptus* and *Sylvaticus*) in the late Pliocene, with further diversification into more specialised groupings

387 during the Pleistocene. Although climatic upheaval during the Pleistocene may have promoted one of the
388 most astonishing examples of incipient speciation among mammals in Africa, we do not observe evidence
389 that these changes were effected by decreases in population size (genetic drift). The strong association
390 between genetic diversity and ecological region suggests that the exceptional diversity within the
391 bushbuck complex may have been driven, at least in part, by parapatric speciation.
392

REFERENCES

Anderson MJ. 2004. DISTLM v.5: a FORTRAN computer program to calculate a distance-based multivariate analysis for a linear model. Department of Statistics, University of Auckland, New Zealand.

Amy LN, Kunz TH. 2012. Effects of Solar Radiation on Animal Thermoregulation. In: Babatunde EB, ed. *Solar Radiation*. InTech, ISBN: 978-953-51-0384-4. Available at http://www.intechopen.com/books/solar-radiation/effects-of-solar-radiation-on-animal-thermoregulation

Arnason U, Gullberg A, Widegren B. 1993. Cetacean mitochondrial DNA control region: Sequences of all extant baleen whales and two sperm whale species. *Molecular Biology and Evolution* 10: 960–970.

Avise JC. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, Massachusetts.

Barlow A, Baker K, Hendry CR, Peppin L, Phelps T, Tolley KA, Wüster CE, Wüster W. 2013. Phylogeography of the widespread African puff adder (*Bitisarietans*) reveals multiple Pleistocene refugia in southern Africa. *Molecular Ecology* **22**:1134–1357 DOI 10.1111/mec.12157 PMID: 23286376.

Bergmann C. 1847. Ueber die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien* 1:595–708.

Best GA. 1962. Rowland Ward's records of big game. XIth Edition: Africa. London: Rowland Ward Ltd. pp 198–208.

Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. 2016. SpreaD3: interactive visualisation of spatiotemporal history and trait evolutionary processes. *Molecular Biology and Evolution* **33**(8):2167-9. doi: 10.1093/molbev/msw082.

Biswas S, Akey JM. 2006. Genomic insights into positive selection. *Trends in Genetics* **22**(8):437–446. DOI 10.1016/j.tig.2006.06.005.

Bobe R, Behrensmeyer AK. 2004. The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus Homo. *Palaeogeography, Palaeoclimatology, Palaeoecology* **207**:399–420.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard, MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology* **10**(4): e1003537 DOI 10.1371/journal.pcbi.1003537.

Broennimann O, Ursenbacher S, Meyer A, Golay P, Monney J-C, Schmocker H, Guisan A, Dubey S. 2014. Influence of climate on the presence of colour polymorphism in two montane reptile species. *Biology Letters* **10**: 20140638. DOI 10.1098/rsbl.2014.0638.

Brown DM, Brenneman RA, Koepfli K-P, Pollinger JP, Milá B, Georgiadis NJ, LouisJr EE, Grether GF, Jacobs DK, Wayne RK. 2007. Extensive population genetic structure in the giraffe. BMC Biology, 5, 57, DOI 10.1186/1741-7007-5-57.

Caro T. 2005. The adaptive significance of coloration in mammals. *Bioscience* **55**(2):125-136. DOI 10.1641/0006-3568(2005)055[0125:TASOCI]2.0.CO;2.

Clauss M, Dittmann MT, Müller DWH, Meloro C, Codron D. 2013. Bergmann's rule in mammals: A cross-species interspecific pattern. *Oikos* **122**(10): 1465–1472 DOI 10.1111/j.1600-0706.2013.00463.x.

Cerling TE. 1992. Development of grasslands and savannas in East Africa during the Neogene. *Palaeogeography, Palaeoclimatology, Palaeoecology* 97(3):241-247 DOI 10.1016/0031-0182(92)90211-M.

Clusella-Trullas S, Terblanche JS, Blackburn TM, Chown SL. 2008. Testing the thermal melanism hypothesis: a macrophysiological approach. *Functional Ecology* **22**:232– 238 DOI 10.1111/j.1365-2435.2007. 01377.x

Cowling SA, Cox PM, Jones CD, Maslin MA, Peros M, Spall SA. 2008. Simulated glacial and interglacial vegetation across Africa: implications for species phylogenies and trans-African migration of plants and animals. *Global Change Biology* **14**:827–840.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8): 772 DOI: 10.1038/nmeth.2109

deMenocal PB. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* **220**:3–24.

Deino AL, Tauxe L, Monaghan M, Hill A. 2002. Ar-40/Ar-30 geochronology and paleomagnetic stratigraphy of the Lukeino and lower Chemeron formations at Tabarin and Kapcheberek, Tugen Hills, Kenya. *Journal of Human Evolution* **42**:117-140.

Dorst J, Dandelot P. 1970. *A field guide to the larger mammals of Africa*. London: Collins.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22**:1185–1192.

Drummond AJ, Suchard MA, Xie D, Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7 *Molecular Biology And Evolution* **29**: 1969-1973.

Dupont LM. 2011. Orbital scale vegetation change in Africa. *Quaternary Science Reviews* **30**:3589–3602.

Edwards EJ, Osborne CP, Strömberg CA, Smith SA; C4 Grasses Consortium, Bond WJ, Christin PA, Cousins AB, Duvall MR, Fox DL, Freckleton RP, Ghannoum O, Hartwell J, Huang Y, Janis CM, Keeley JE, Kellogg EA, Knapp AK, Leakey AD, Nelson DM, Saarela JM, Sage RF, Sala OE, Salamin N, Still CJ, Tipple B. 2010. The origins of C4 grasslands: integrating evolutionary and ecosystem science. *Science* **328** 587–591.

Egea R, Casillas S, Barbadilla A. 2008. Standard and Generalized McDonald-Kreitman Test: a website to detect selection by comparing different classes of DNA sites. *Nucleic Acids Research* **36** (Web Server issue): W157-W162. Available at http://mkt.uab.cat/mkt/.

Faulkes CG, Bennett NC, Cotterill, FPD, Stanley W, Mgode GF, Verheyen E. 2011. Phylogeography and cryptic diversity of the solitary-dwelling silvery mole-rat, genus *Heliophobius* (family: Bathyergidae). *Journal of Zoology* **285**: 324–338.

Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package). Distributed by the author. Seattle, WA: Department of Genome Sciences, University of Washington.

Fernández MH, Vrba ES. 2005. A complete estimate of the phylogenetic relationships in Ruminantia: a dated species level supertree of the extant ruminants. *Biological Reviews of the Cambridge Philosophical Society **80***:269–302.

Flagstad A, Syvertsen PO, Stenseth NC, Jakobsen KS. 2001. Environmental change and rates of evolution: The phylogeographic pattern within the hartebeest complex as related to climatic variation. *Proceedings Royal Society, London B* **268**: 667–677. PMID: 11321054.

Freckleton RP, Harvey PH, Pagel M. 2003. Bergmann's rule and body size in mammals. *The American Naturalist* **161** (5): 821–825.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**:915-925.

Grubb P. 1985. Geographical variation in the bushbuck of eastern Africa (*Tragelaphus scriptus*; Bovidae). In: Schuchmann KL, ed. *Proc Intern Symp African Vertebr*. Bonn: Museum A König, 11–26.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**:95–8.

Haltenorth T, Diller H. 1980. *A field guide to the mammals of Africa including Madagascar*. London: Collins. pp 57–58.

Harpending HC.1994. Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**:591-600.

Harpending HC, Sherry ST, Rogers AR, Stoneking M. 1993. The genetic structure of ancient human populations. *Current Anthropology* **34**:483-496.

Harris JM, Brown FH, Leake MG. 1988. Stratigraphy and paleontology of Pliocene and Pleistocene localities west of Lake Turkana, Kenya. *Natural History Museum of Los Angeles County, Contribution in Science* **399**: 1–128.

Heled J, Drummond AJ. 2008. Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* **8**: 289.

Hewitt GM. 2004. The structure of biodiversity—insights from molecular phylogeography. *Frontiers in Zoology* **1**: 4. PMID: 15679920

Ho SYW, Shapiro B. 2011. Skyline Plot Methods for Estimating Demographic History from Nucleotide Sequences. *Molecular Ecology Resources* **11**(3):423-34. DOI 10.1111/j.1755-0998.2011.02988.x.

IUCN SSC Antelope Specialist Group. 2016. *Tragelaphus scriptus* (errata version published in 2017). The IUCN Red List of Threatened Species 2016: e.T22051A115165242. DOI 10.2305/IUCN.UK.2016-3.RLTS.T22051A50196111.en.

Jacobs BF. 2004. Paleobotanical studies from Tropical Africa: Relevance to the evolution of forest, woodland and savannah biomes. *Philosophical Transactions of the Royal Society B* **359**:1573–1583.

Jacobs DS, Babiker H, Bastian A, Kearney T, van Eeden R, Bishop JM. 2013. Phenotypic convergence in genetically distinct lineages of a *Rhinolophus* species complex (Mammalia, Chiroptera). *PLoS ONE* **8**(12): e82614. DOI: 10.1371/journal.pone.0082614. PMID: 24312666.

Kalb JE, OswaldEB, Tebedge S, Mebrate A, Tola E, et al.,1982. Geology and stratigraphy of Neogene deposits, Middle Awash Valley, Ethiopia. *Nature* **298**:98–106.

Kingdon J .1997. *The Kingdon field guide to African mammals*. London: Academic Press. pp 476.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **33** (7): 1870-1874. DOI 10.1093/molbev/msw054.

Leake MG, Harris JM. 2003. *Lothagam: the dawn of humanity in eastern Africa*. New York: Columbia University Press. pp 678.

Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* **5**:e1000520.

Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution* **27**:1877–1885.

Leonard JA, Rohland N, Glaberman S, Fleischer RC, Caccone A, Hofreiter M. 2005. A rapid loss of stripes: the evolutionary history of the extinct quagga. *Biology Letters* **1**:291–295 DOI 10.1098/rsbl.2005.0323

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451-1452.

Lorenzen ED, Simonsen BT, Kat PW, Arctander P, Siegismund HR. 2006. Hybridization between subspecies of waterbuck (*Kobus ellipsiprymnus*) in zones of overlap with limited introgression. *Molecular Ecology* **15**:3787–3799.

Lorenzen ED, De Neergaard R, Arctander P, Siegismund HR. 2007. Phylogeography, hybridization and Pleistocenerefugia of the kob antelope (*Kobus kob*). *Molecular Ecology* **16**:3241–3252.

Lorenzen E D, Masembe C, Arctander P, Siegismund H R. 2010. A long-standing Pleistocene refugium in Southern Africa and a mosaic of refugia in East Africa: insights from mtDNA and the common eland antelope. *Journal of Biogeography* 37:571–581.

Lydekker R. 1914. *Catalogue of the ungulate mammals in the British Museum Natural History Vol. 3*. London: British Museum. pp 317–326.

Matthee CA, Burzlaff JD, Taylor JF, Davis SK. 2001. Mining the mammalian genome for artiodactyl systematics. *Systematic Biology* **50**:1-24.

Matthee CA, Davis SK. 2001. Molecular insights into the evolution of the family Bovidae: a nuclear DNA perspective. *Molecular Biology and Evolution* **18**: 1220-1230.

Mayaux P, Bartholome E, Fritz S, Belward A. 2004. A new land-cover map of Africa for the year 2000. *Journal of Biogeography* **31**:861–877.

Mills MGL, Hes L. 1997. *Complete Book of Southern African Mammals*. South Africa: Struik Winchester . pp 356.

Moodley Y, Bruford MW. 2007. Molecular biogeography: towards an integrated framework for conserving pan-African biodiversity. *PloS One* **5**: e454.

Moodley Y, Bruford MW, Bleidorn C, Wronski T, Apio A, Plath M. 2009. Analysis of mitochondrial DNA data reveals non-monophyly in the bushbuck (*Tragelaphus scriptus*) complex. *Mammalian Biology* **74**:418-422.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, et al. 2001. Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* **51**:933–937.

Partridge TC, Wood B, deMenocal PB. 1995. The influence of global climatic change and regional uplift on large-mammalian evolution in East and Southern Africa. In: Vrba E, Denton G, Partridge TC,

Burckle L, eds. *Paleoclimate and Evolution With Emphasis of Human Origins*. New Haven: Yale Univ Press. pp 330–355.

Plumptre AJ, Wronski T. 2013. *Tragelaphus scriptus*. In: Kingdon JS and Hoffmann M, eds. *The Mammals of Africa*. VI. Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer, and Bovids. London :Bloomsbury Publishing.

Posada D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* **25**: 1253-1256.

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer

Rau RE. 1978. Additions to the revised list of preserved material of the extinct Cape colony quagga and notes on the relationship and distribution of southern plains zebras. *Annals of the South African Museum* **77**:27–45.

Reed KE. 1997. Early hominid evolution and ecological change through the African Plio-Pleistocene. *Journal of Human Evolution* **32**: 289–322.

Simonsen BT, Siegismund HR, Arctander P. 1998. Population structure of African buffalo inferred from mtDNA sequences and microsatellite loci: high variation but low differentiation. *Molecular Ecology* **7**: 225–237

Smit H A, Robinson T J, Van Vuuren BJ. 2007. Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Molecular Ecology* **16**: 2680–2692.

Smitz N, Berthouly C, Cornélis D, Heller R, Van Hooft P, Chardonnet P, Caron A, Prins H, Jansen van Vuuren B Delongh H, Michaux J. 2013. Pan-African genetic structure in the African Buffalo (*Syncerus caffer*): Investigating intraspecific divergence. *PLoS ONE* **8**(2):e56235. DOI:10.1371/journal.pone.0056235.

Spong G, Johansson M, Björklund M. 2000. High genetic variation in leopards indicates large and long-term stable effective population size. *Molecular Ecology* **9**:1773–1782.

574 **Stoner CJ, Caro TM, Graham CM. 2003.** Ecological and behavioral correlates of coloration in
575 artiodactyls: systematic analyses of conventional hypotheses. *Behavioral Ecology* **14**:823–840.

576 **Szabo B, Haynes C, Maxwell TA. 1995.** Ages of Quaternary pluvial episodes determined by uranium-
577 series and radiocarbon dating of lacustrine deposits of Eastern Sahara. *Palaeogeography,*
578 *Palaeoclimatology, Palaeoecology* **113**: 227–242.

579 **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: Molecular Evolutionary
580 Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**:2725-2729.

581 **Trauth MH, Maslin MA, Deino AL, Strecker MR, Bergner AGN, Dünforth M. 2007**. High- and low-
582 latitude forcing of Plio-Pleistocene East African climate and human evolution. *Journal of Human*
583 *Evolution* 53:475–486.

584 **Thompson JD, Higgins DG, Gibson TJ. 1994.** CLUSTAL W: improving the sensitivity of progressive
585 multiple sequence alignment through sequence weighting, position-specific gap penalties and weight
586 matrix choice. *Nucleic Acids Research* **22**:4673-4680.

587 **Voelker G, Outlaw RK, Bowie RC. 2010**. Pliocene forest dynamics as a primary driver of African bird
588 speciation. *Global Ecology and Biogeography* **19**:111–121.

589 **Vrba E. 1995.** The fossil record of African antelopes (Mammalia, Bovidae) in relation to human
590 evolution and paleoclimate, In: Vrba E, Denton G, Burckle L, Partridge T, eds. *Paleoclimate and*
591 *Evolution With Emphasis on Human Origins*. New Haven: Yale University Press, 385–424.

592 **Woldegabriel G, Haile-Selassie Y, Renne PR, Hart WK, Ambrose SH, Asfaw B, Heiken G, White T.**
593 **2001.** Geology and palaeontology of the late Miocene middle Awash valley, Afar rift, Ethiopia. *Nature*
594 **412**:175-178.

595 **Zwickl DJ. 2006.** Genetic algorithm approaches for the phylogenetic analysis of large biological
596 sequence datasets under the maximum likelihood criterion. D. Phil. Thesis , The University of Texas.

Table 1.Species-wide genetic sampling of bushbuck across sub-Saharan Africa.

Locality number	Sample	HaplogroupSample[1]	Taxonomic Subspecies	Common name[2]	Locality	Latitude	Longitude	Country	Voucher	Source
1	*scriptus_SL*	*scriptus*	*scriptus*	Senegal bushbuck	Sierra Leone	7.54	-11.12	Sierra Leone	20.7.10.21	Natural History Museum, London
2	*dodingae*	*dodingae*	*dodingae*	Kidepo bushbuck	Kedef Valley, western Dodinga Hills	4.45	33.31	South Sudan	Uganda 368	Powell Cotton Museum, Birchington
3	*bor*	*bor*	*bor*	Nile bushbuck	Bouroum	10.45	18.8	Chad	Chad 116	Powell Cotton Museum, Birchington
4	*decula2*	*decula*2	*decula*	Abyssinian bushbuck	Din Din	8.45	40.1	Ethiopia	AD2	Travel Ethiopia, Addis Ababa
5	*decula1*	*decula*1	*decula*	Abyssinian bushbuck	Din Din	8.45	40.1	Ethiopia	AD1	Travel Ethiopia, Addis Ababa
6	*Lowervolta*	Lower Volta	*scriptus*	Lower Volta bushbuck	Ejura, Ashanti Region	7.38	-1.37	Ghana	GH4849	Department of Evolutionary Biology, Copenhagen
7	*Niger*	Niger	*scriptus*	Niger bushbuck	Aningo	8.6	8.85	Nigeria	26344	Nationaal Natuurhistorisch Museum,
8	*phaleratus*	*phaleratus*	*phaleratus*	Cabinda bushbuck	Tshimbali	-4.72	13.1	DRC	17820	Royal Museum for Central Africa, Te
9	*UpperVolta*	Upper Volta	*scriptus*	Upper Volta bushbuck	Kasana, Upper West Region	10.88	-1.99	Ghana	GH6335	Department of Evolutionary Biology, Copenhagen
10	*Angola*	Angola	*ornatus*	Angolan bushbuck	Lifune	-8.4	13.45	Angola	B14201	Staatliche Naturhistorische Sammlun
11	*ornatus*	*ornatus*	*ornatus*	Chobe bushbuck	Kazungula	-17.78	25.27	Zimbabwe	Zimbabwe 07	Bromley Game Skin Tannery, Harare
12	scriptus2	scriptus2	*sylvaticus*	South African bushbuck	South Africa	-30.64	29.29	South Africa	ref 16	
13	*sylvaticus*	*sylvaticus*	*sylvaticus*	South African bushbuck	Humansdorp, Eastern Cape	-34.02	24.77	South Africa	Eastern Cape 04	Taxidermy Africa, Humansdorp, Sou
14	*meneliki1*	*meneliki*1	*meneliki*	Menelik's bushbuck	Cure Rey, Arussi Mountains	7.05	39.42	Ethiopia	Abyssinia II 30	Powell Cotton Museum, Birchington
15	*meneliki2*	*meneliki*2	*meneliki*	Menelik's bushbuck	Boare, Arussi Mountains	7.45	39.45	Ethiopia	Abyssinia II 56	Powell Cotton Museum, Birchington
16	*dianae*	*dianae*	*dianae*	Ituri bushbuck	Kasindi	-0.04	29.71	DRC	Congo 329	Powell Cotton Museum, Birchington
17	*dama*	*dama*	*dama*	Kavirondo bushbuck	Irumu	1.45	29.87	DRC	Congo 159	Powell Cotton Museum, Birchington
18	*barkeri*	*barkeri*	*barkeri*	Barker's bushbuck	Lomuleng, Imatong Mountains	3.95	33	South Sudan	Sudan I 27	Powell Cotton Museum, Birchington
19	scriptus1	*delamerei*2	*delamerei*	Lord Delamere's bushbuck	Kenyan	-0.28	37.02	Kenya	ref 10	
20	*haywoodi*	*delamerei*1	*meruensis*	Lord Delamere's bushbuck	Mount Meru	-3.23	36.75	Tanzania	MM0555	Department of Evolutionary Biology, Copenhagen
21	*massaicus*	*massaicus*	*massaicus*	Massai bushbuck	Chiredzi	-21	31.5	Zimbabwe	Zimbabwe 10	Bromley Game Skin Tannery, Harare
22	*fasciatus1*	*fasciatus*1	*fasciatus*	Somali bushbuck	Mona Mofa Camp, Jubaland	0	42.12	Somalia	Jubaland 34	Powell Cotton Museum, Birchington
23	*roualeyni*	*roualeyni*	*roualeyni*	Limpopo bushbuck	Thabazimbi	-24.6	27.4	South Africa	Limpopo 12	Nico van Rooyen Taxidermy, Rossly

24	*fasciatus2*	*fasciatus*2	*fasciatus*	Somali bushbuck	Mona Mofa Camp, Jubaland	0	42.12	Somalia	Jubaland 14	Powell Cotton Museum, Birchington
25	*Luangwa*	Luangwa	*ornatus*	Luangwa bushbuck	Msandile	-13.5	32.75	Zambia	17001	Livingstone Museum, Livingstone, Z
26	*Zambezi1*	Zambezi1	*ornatus*	Zambezi bushbuck	Kanyemba	-15.7	30.32	Zimbabwe	Zimbabwe 17	Taxidermy Enterprises, Bulawayo, Z
27	*Zambezi2*	Zambezi2	*ornatus*	Zambezi bushbuck	Mhangura	-16.9	30.15	Zimbabwe	Zimbabwe 06	Bromley Game Skin Tannery, Harare

1. After Moodley and Bruford (2007)

2. After Halternorth (1963). Where no common name exists the dominant geographic feature of the area was used

DRC Democratic Republic of the Congo

Table 2 Genetic diversity for mtDNA regions (12S rRNA, 16S rRNA, and *Cyt b*), nucDNA regions (MGF, PRKCI, SPTBN, and THY) for all ingroup sequences and the two major *Scriptus* and *Sylvaticus* clades.

	Locus	n	Size (bp)	S	π	h	Hd	k	S/k
Entire species complex	12SrRNA	27	593	63	0.036	21	0.98	21.348	2.951
	16SrRNA	27	347	35	0.038	17	0.954	13.137	2.664
	Cytochrome*b*	27	1140	255	0.072	24	0.991	82	3.11
	MGF	27	671	10	0.003	5	0.635	1.852	5.399
	PRCK1	27	498	2	0.0003	3	0.145	0.148	13.51
	SPTBN1	27	764	12	0.001	7	0.456	0.957	12.539
	THY	27	663	2	0.0008	3	0.501	0.541	3.696
Scriptus clade	12SrRNA	27	593	17	0.012	8	0.972	7.167	2.371
	16SrRNA	27	347	3	0.003	3	0.667	1	3
	Cytochrome*b*	27	1140	90	0.028	8	0.972	32.389	2.778
	MGF	27	671	0	0	1	0	0	2.712
	PRCK1	27	498	2	0.001	3	0.556	0.611	0
	SPTBN1	27	764	0	0	1	0	0	3.273
	THY	27	663	0	0	1	0	0	0
Sylvaticus clade	12SrRNA	27	593	27	0.01	13	0.961	5.81	4.64
	16SrRNA	27	347	23	0.02	14	0.974	6.843	3.361
	Cytochrome*b*	27	1140	158	0.035	16	0.987	40.333	3.917
	MGF	27	671	10	0.002	4	0.399	1.601	6.246
	PRCK1	27	498	0	0	1	0	0	0
	SPTBN1	27	764	13	0.002	7	0.634	1.542	9.155
	THY	27	663	1	0.0003	2	0.209	0.209	4.785

S - number of polymorphic sites; π - nucleotide diversity; h - number of haplotypes; Hd - haplotype diversity; k - average number of nucleotide differences; S/k - expansion coefficient.

Statistically significant results were indicated by asterisks: * $P < 0.05$, **$P < 0.01$.

Table 3 Demography and tests of the neutral modelfor mtDNA regions (12S rRNA, 16S rRNA, and *Cyt b*), nucDNA regions(MGF, PRKCI, SPTBN, and THY), and defined major clades of Bushbuck based on nucDNA sequences.

	Locus	Fu's Fs	Tajima's D	Raggedness (R2)	Mismatch distribution	Tau (τ)
Entire species complex	12SrRNA	-2.04	1.02	0.163	Multimodal	5.154
	16SrRNA	-1.007	1.244	0.185	Multimodal	5.302
	Cytochrome *b*	0.074	0.606	0.153	Multimodal	33.927
	MGF	0.93	-1.15678	0.107	Multimodal	0.607
	PRCK1	-2.223	-1.511	0.131	Unimodal	0.148
	SPTBN1	-3.091*	-2.312**	0.088	Unimodal	0
	THY	0.15	0.091	0.135	Unimodal	
*Scriptus*clade	12SrRNA	-1.788	0.401	0.186	Multimodal	4.105
	16SrRNA	-0.707	-0.359	0.229	Unimodal	1
	Cytochrome *b*	1.138	-0.113	0.17	Multimodal	13.51
	MGF	-	-	-	-	-
	PRCK1	-0.532	-0.583	0.185	Unimodal	0.611
	SPTBN1	-	-	-	-	-
	THY	-	-	-	-	-
*Sylvaticus*clade	12SrRNA	-3.842	-1.036	0.097	Multimodal	3.057
	16SrRNA	-4.371	-0.076	0.146	Multimodal	4.327
	Cytochrome *b*	-0.382	-0.562	0.113	Multimodal	22.63
	MGF	1.007	-1.618	0.106	Multimodal	0
	PRCK1	-	-	-	-	-
	SPTBN1	-2.257	-2.207**	0.1	Unimodal	0.303
	THY	-0.011	-0.529	0.104	Unimodal	0.209

Statistically significant results were indicated by asterisks: * $P < 0.05$, ** $P < 0.01$.

Table 4. Fitting of bushbuck nucDNA genetic distance data against taxonomic, biogeographic, and geographic models.

		Multivariate matrix regression				
Predictors	**Model**	**df**	**pseudo-F**	**Marginal**	**pseudo-F**	**Conditional**
Taxonomy	All subspecies	25	2.049	0.770	1.886	0.657
Biogeography	Olson et al.	25	10.121	0.953**	7.892	0.414
Geography	Coordinates	25	4.130	0.264 *	-	-

permutation P <0.05 *; <0.01 **

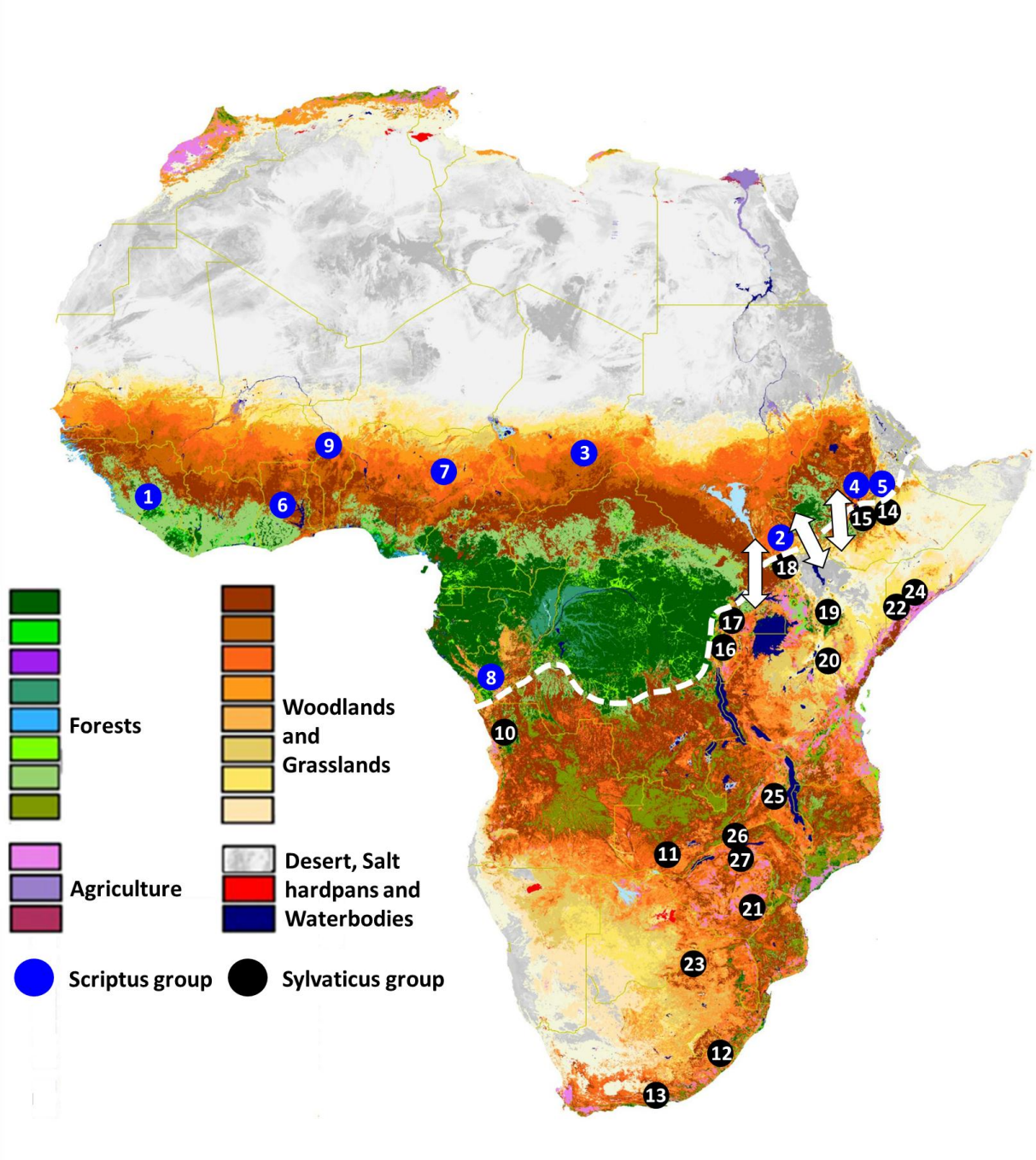

Figure 1. The land cover of Africa reconstructed from remotely sensed data (redrawn from Mayaux et al., 2004). The geographical distribution of sampling localities included in the present study are shown on the map. Taxa are plotted as dots and designated either blue for the *Scriptus* group or black for the *Sylvaticus* group. Samples are numbered according to Table 1. A dashed white line divides the distributions of both groups and white arrows show zones of potential gene flow.

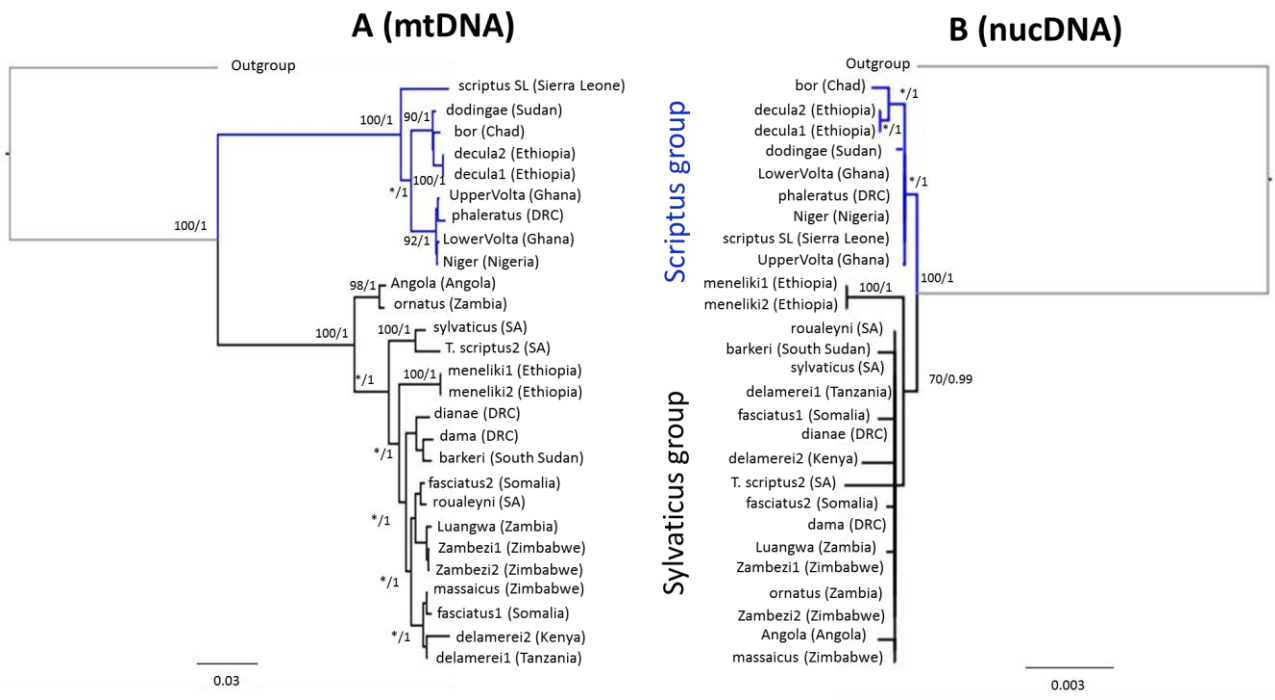

Figure 2.Tree topologies based on maximum likelihood retrieved from (A) the combined mtDNA data and (B) the combined nucDNA data. Values given above the branches represent maximum likelihood bootstrap values and maximum clade probabilities.

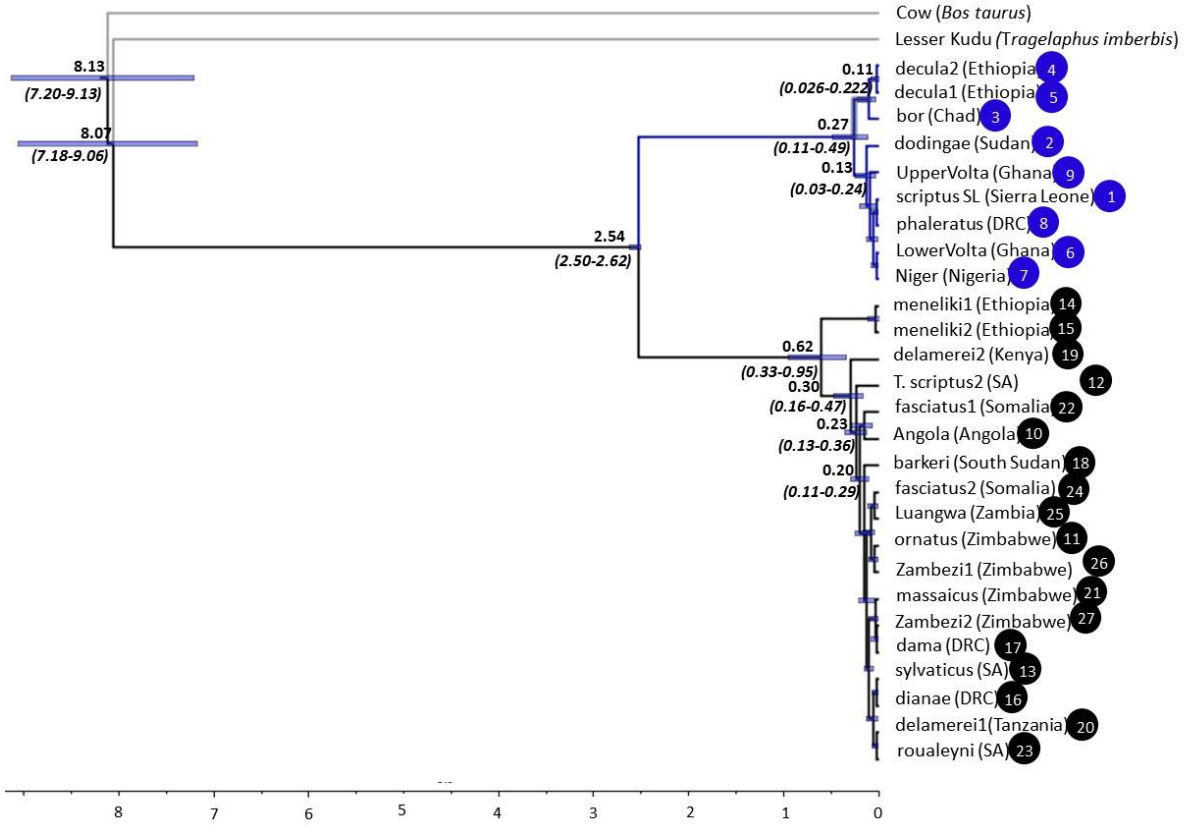

Figure 3. Dated Bayesian phylogeny of four bushbuck nuclear DNA sequences reconstructed in BEAST. Fossil calibration points are indicated by letters A and B. Median divergence time estimates (in MYA) and 95% HPD values are adjacent to their respective nodes. Purple nodal bars correspond to the 95% HPD. Major bushbuck groups are colored as in Fig. 1.

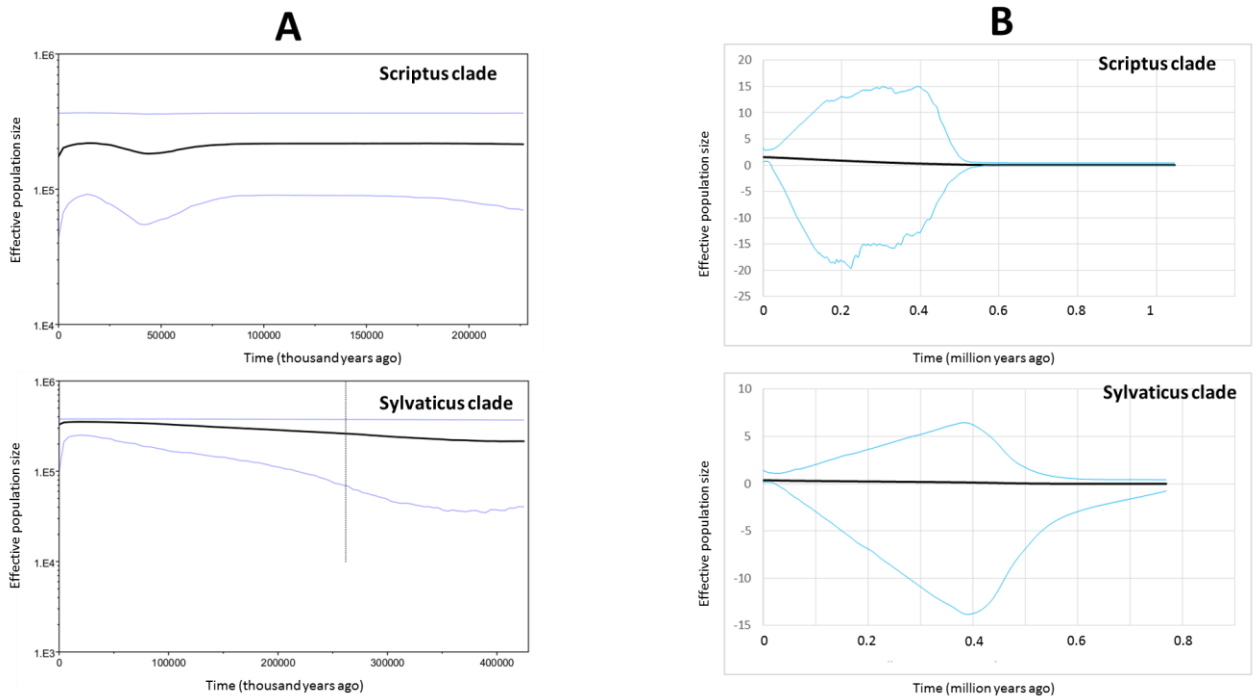

Figure 4. (A) Bayesian Skyline Plots (BSPs). BSPs represent population size changes over time, inferred with mtDNA and an assumed divergence rate of 0.056 per million years. The X-axes are time in thousands of years. Y-axes are mean effective population sizes log-scale. Solid black lines represent median height and areas between blue lines encompass the 95% highest posterior density (HPD). (B) Extended Bayesian Skyline Plots (EBSPs). EBSPs represent population size changes over time in two of the mtDNA clades, inferred by mtDNA and nucDNA. X-axes are time in millions of years,Y-axes are effective population size divided by generation time.

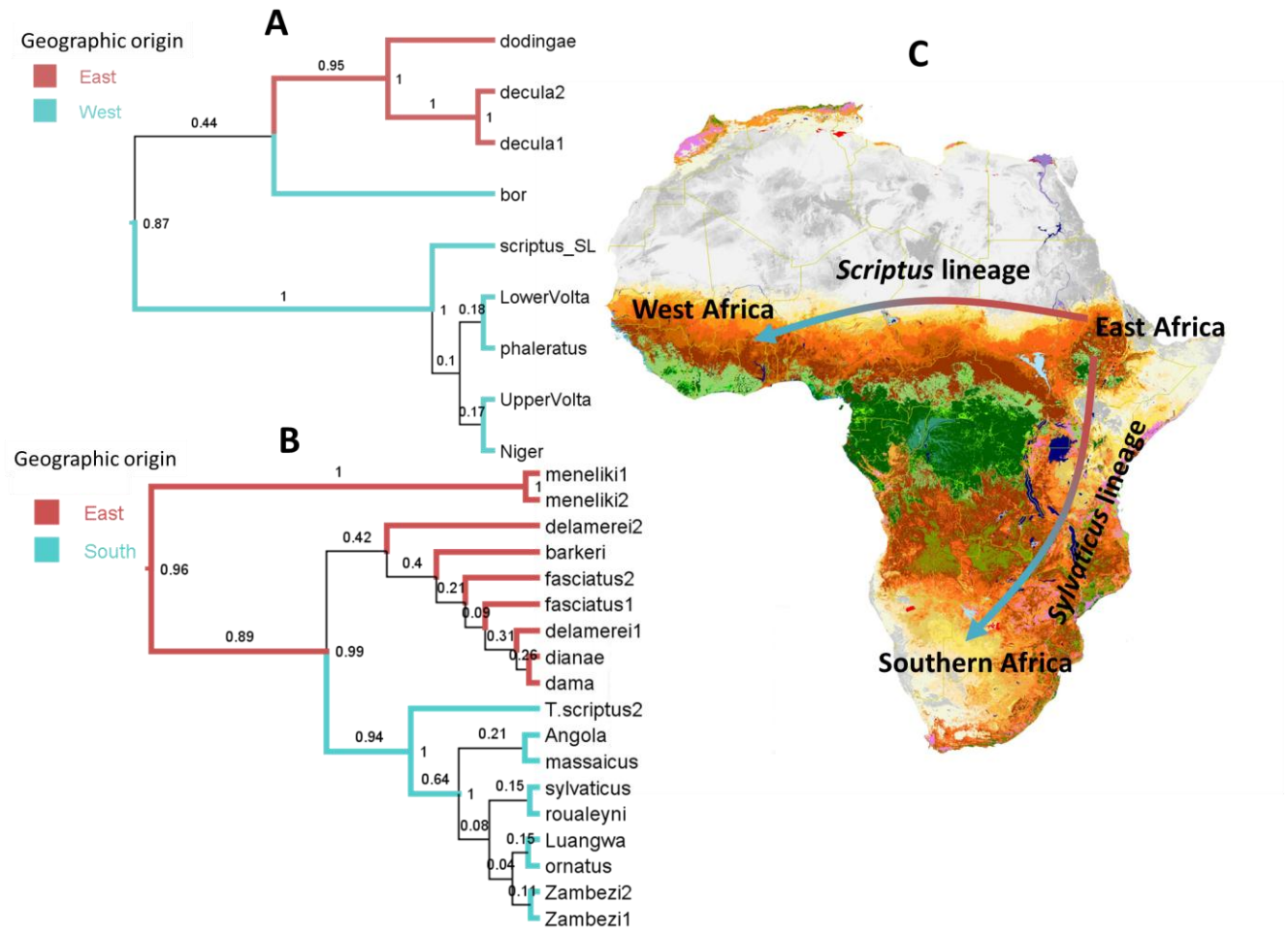

Figure 5. Bayesian ancestral range reconstruction and colonization history of bushbuck based on nucDNA markers. (A) *Scriptus* lineage, (B) *Sylvaticus* lineage. (C) Colonization routes ofbushbuck species complex identified by BSSVS. Lines between geographic regions represent branches in the MCC tree along which the relevant location transition occurs. The map is based on redrawn map from Mayaux et al., 2004. Numbers above branches are Bayesian posterior probabilities (PP). The coloured branch lengths represent the ancestral range with highest marginal probability for each lineage as inferred in BEAST (only branches with PP > 0.5). Numbers at each node represent marginal probabilities foralternative ancestral locations.