ORIGINAL ARTICLE



Assessing the identity of rare historical museum specimens of the extinct blue antelope (*Hippotragus leucophaeus*) using an ancient DNA approach

Lucy Plaxton¹ · Elisabeth Hempel^{2,3} · William A. Marsh¹ · Roberto Portela Miguez¹ · Isabelle Waurick³ · Andrew C. Kitchener^{4,5} · Michael Hofreiter² · Adrian M. Lister¹ · Frank E. Zachos^{6,7,8} · Selina Brace¹

Received: 2 May 2023 / Accepted: 4 August 2023 / Published online: 24 August 2023 © The Author(s) 2023

Abstract

The blue antelope or bluebuck (*Hippotragus leucophaeus*) is an extinct species of antelope that lived in South Africa until ca. 1799–1800. Disappearing only 34 years after it was described, it was the first large African mammal species to have become extinct in recent times. Therefore, current scientific knowledge of the blue antelope is limited to museum specimens. However, these are extremely rare and further complicated by uncertainties surrounding their taxonomic status due to morphological similarity with the closely related extant roan antelope (*Hippotragus equinus*) and sable antelope (*Hippotragus niger*). A previous study found that only four (out of ten) genetically analysed historical museum specimens could be confirmed as *Hippotragus leucophaeus*. Here, we analyse a further five potential blue antelope specimens, two of which had been morphologically identified as blue antelope (one from the Natural History Museum (NHM) in London and one from National Museums Scotland in Edinburgh), plus an additional three specimens identified here as candidate blue antelope specimens (from the NHM). We find that only one of the five specimens can be genetically validated as *Hippotragus leucophaeus*, whereas the other four specimens were identified as *H. equinus*. Furthermore, using hybridisation capture, we generated two complete mitochondrial genomes for blue antelope specimens that were formerly known only from partial mitochondrial genomes. The diversity analyses including all newly generated genomes further confirm that the mitochondrial nucleotide diversity across this species was low.

Keywords $aDNA \cdot Bluebuck \cdot Extinction \cdot Misidentification \cdot Mitochondrial genome \cdot Museomics \cdot Museum collections \cdot South Africa \cdot Taxonomic identification$

Handling editor: Pamela Burger.

Lucy Plaxton and Elisabeth Hempel contributed equally to this manuscript.

Elisabeth Hempel Hempel.Elisabeth@posteo.org

- ¹ Natural History Museum, London SW7 5BD, UK
- ² Evolutionary Adaptive Genomics, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany
- ³ Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstraße 43, 10115 Berlin, Germany
- ⁴ Department of Natural Sciences, National Museums Scotland, Chambers Street, Edinburgh EH1 1JF, UK

Introduction

The South African blue antelope or bluebuck, *Hippotra*gus leucophaeus (Pallas, 1766), was the first large mammal species to have become extinct in recent times in Africa

- ⁵ School of Geosciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK
- ⁶ Natural History Museum Vienna, Burgring 7, 1010 Vienna, Austria
- ⁷ Department of Evolutionary Biology, University of Vienna, Djerassiplatz 1, 1030 Vienna, Austria
- ⁸ Department of Genetics, University of the Free State, PO Box 339, Bloemfontein, South Africa

(FitzSimons 1920; Harper 1945; Broom 1949; van Bruggen 1959; Robinson et al. 1996). It is generally accepted that the species became extinct around 1799-1800 (Lichtenstein 1811, 1814), only approximately 34 years after it was taxonomically described. Current scientific knowledge is based on a very limited number of historical museum specimens, many of which are of uncertain taxonomic status due to morphological similarity with the closely related extant roan antelope (Hippotragus equinus) and sable antelope (Hippotragus niger) (Hempel et al. 2021). Reassignment of several presumed blue antelope specimens to roan and sable antelope has resulted in the blue antelope being one of the rarest mammal species found in museum collections (Hempel et al. 2021), with currently only around ten potential specimens from historical times, i.e., the eighteenth century (however, there is more (sub)fossil material).

Very few written accounts remain from the seventeenth and eighteenth centuries that describe the blue antelope (Rookmaaker 1992), and the small number of illustrations and paintings from this period are argued to have been created from memory after sightings of the animal, thus possibly featuring inaccuracies and embellishments (Husson and Holthuis 1975). Scientific studies of the blue antelope over the past century are scarce, and recent molecular studies highlight that those based on morphological identification of historical museum specimens may be erroneous (Hempel et al. 2021). Indeed, Hempel et al. (2021) found that only four (out of ten) genetically analysed historical museum specimens could be confirmed as Hippotragus leucophaeus (specimens from Leiden (skull fragments), Vienna, Stockholm and Uppsala). Therefore, there are only four genetically validated specimens and the mounted skin in Paris, that, based on its morphology and the fact that it is a skin, is very likely also a blue antelope, and the mounted skin in Leiden. It highlights a severe lack of reference material available for scientific study and unquestionably limits our ability to understand this species' morphology, ecology, evolution and the causes of its extinction.

The blue antelope was described in 1766 by Peter Simon Pallas as *Antilope leucophaea*, but it is thought to have been first mentioned by Johann Schreyer (1681) and then recorded more elaborately, albeit inaccurately, by Peter Kolb during his travels around South Africa (Kolb 1719). The blue antelope's distinguishing features were a blue-grey tint to its coat, hence its vernacular name, the lack of dark face markings, in contrast to the other two *Hippotragus* species (Lydekker 1914), a white patch in front of the eye (Pallas 1767) and smaller size (Sclater and Thomas 1899). Unlike the roan and sable antelope, which occupy much larger ranges further north (IUCN SSC Antelope Specialist Group 2017a, b), the blue antelope was historically endemic to the Cape region of South Africa, inhabiting an area restricted to ca. 4300 km² even before the arrival of European colonists in the mid-seventeenth century (Kerley et al. 2009). Kerley et al.'s (2009) modelling predicted that its range and census population size, estimated at only ~ 370 individuals in the seventeenth century, were effectively that of an island population trapped in a cycle of decreasing population size due to a combination of biotic and abiotic factors, a concept known as the extinction vortex (Gilpin and Soulé 1986). The low genetic diversity suggested by Kerley et al. (2009) has since been further supported by a study using mitochondrial data (Hempel et al. 2021) and a study using nuclear data (Hempel et al. 2022), both of which also concluded that the historical genetic diversity of the blue antelope is likely to have been low.

The location of late Quaternary fossil remains and rock paintings attributed to the blue antelope suggests that the species previously occupied a significantly larger range across South Africa (Loubser et al. 1990; Plug 1997; Faith 2014; Avery 2019), including the Cape Floristic Region (a land area approximately 90,000 km²) along the southern coast (Skead 1980; Kerley et al. 2003). It has been suggested that the blue antelope may have undertaken seasonal migrations along the coast, to benefit from increased rainfall and seasonal abundance of grasses (Faith and Thompson 2013). If so, rising sea levels in the early Holocene could have disrupted these migratory pathways and contributed to the species' decline (Faith and Behrensmeyer 2013; Faith and Thompson 2013). Habitat competition with and diseases from domestic livestock (Klein 1974) and habitat fragmentation and disruption due to the rise of agriculture (Faith and Thompson 2013) also potentially contributed to its demise. However, it is likely that the most significant contributor to its final extinction was human activity during the seventeenth and eighteenth centuries, particularly overhunting (Harper 1945).

The fossil remains and phylogenetic position within the Hippotragini indicate that the blue antelope was a grazer rather than a browser, favouring predominantly grassland habitats (Stynder 2009; Faith and Thompson 2013). While historically some authors had thought that the blue antelope was actually a subspecies or immature form of the roan antelope (Hippotragus equinus) (e.g., Smith 1849), its status as a distinct species is clear, with mitochondrial and nuclear data continuing to support the placement of the blue antelope as a species distinct from the roan antelope (Robinson et al. 1996; Themudo and Campos 2018; Hempel et al. 2022). This was also already suggested on morphological grounds by Klein (1974). A recent phylogeographical study of roan antelope highlighted significant population structure within the species with four mitochondrial lineages identified: North-west, North-east, East-central and Southern lineages (Gonçalves et al. 2021). Most recent molecular analyses suggest that the sable antelope is sister to the blue antelope, with

the roan antelope positioned outside of this clade (Themudo and Campos 2018; Hempel et al. 2021, 2022).

Historical museum specimens of the blue antelope are extremely rare and confirmed examples have recently been further reduced in number through DNA sequencing, with some specimens now identified as either roan or sable antelope (Hempel et al. 2021). It is sometimes difficult to morphologically distinguish between specimens of the Hippotragine antelopes (Groves and Westwood 1995), especially as, apart from mounted skins, most blue antelope specimens are only pairs of horns or skull fragments (Rookmaaker 1992). No complete skull has yet been DNA-validated as belonging to a blue antelope, and to date the only DNA-validated specimens from historical museum collections comprise two mounted skins in Stockholm and Vienna, skull fragments in Leiden, and a pair of horns in Uppsala (Themudo and Campos 2018; Hempel et al. 2021). Pallas described the blue antelope as Antilope leucophaea in 1766 based on several skins (Mohr 1967; Husson and Holthuis 1969). Whilst the Leiden skull fragment and the Leiden mounted skin (both now assigned as lectotype, Husson and Holthuis 1969) might be the same individual, it is possible that the mounted skin is an additional specimen (although as yet unvalidated through DNA). The blue antelope is a prime example of how misidentifications can cause problems in museum-collectionbased research (Hempel et al. 2021), and highlights that morphology in some cases needs to be complemented by molecular methods of identification.

In an attempt to increase the scarce information available for this species, here we analyse a further five potential blue antelope specimens. Two are morphologically identified as blue antelope and an additional three *Hippotragus* specimens were identified here as candidate blue antelope specimens. In addition, we generated two complete mitochondrial genomes for previously DNA-validated blue antelope specimens, for which only partial mitochondrial genomes were available.

Material and methods

Specimens

Two specimens morphologically identified as blue antelope, a frontlet with horns at the Natural History Museum in London, specimen NHMUK GERM 636e (Fig. 1), and a frontlet with horns at the National Museums of Scotland in Edinburgh, specimen NMS.Z.2004.105 (Supplementary Fig. S4), were the focus of this study to determine if genetic analysis could validate them as blue antelope specimens. The Natural History Museum specimen (NHMUK GERM 636e) has been variously referred to in previous publications as NHMUK GERM 636c, NHMUK GERM



Fig. 1 Specimen NHMUK GERM 636e (A and B) and its specimen labels (C and D), from the Natural History Museum London. Photo credits: Lucy Plaxton

636e and NHMUK GMCM 105e (Supplementary Information for further details). Here, we refer to the specimen only as NHMUK GERM 636e.

Additionally, the entire collection of Hippotragus specimens stored in the Mammal Collection at the Natural History Museum was examined for potential blue antelope specimens based on their historical information. Selection of specimens for sampling was based on the location being within the Cape Colony, South Africa (the historical range of the species); the date of collection, if given, to be no later than the early-1800s to coincide with the species' suspected extinction; and/or if the specimen had been previously labelled as blue antelope. Using these criteria three potential blue antelope specimens were chosen: NHMUK ZD 1997.126/GERM 636a (based on collection date and location), NHMUK ZD 1888.8.20.3 (recorded as both H. leucophaeus and H. equinus on the handwritten specimen labels) and NHMUK ZD 1842.4.11.9 (based on location). These three potential blue antelope specimens alongside the two morphologically identified specimens from London and Edinburgh give a total of five specimens for our analyses (Table 1, Supplementary Table S1 and Supplementary Figures S1-4). Detailed information for each specimen can be found in the Supplementary Information.

Furthermore, we selected two single-stranded libraries, from which previous sequencing efforts had only retrieved partial mitochondrial genomes, to carry out an additional hybridisation capture approach with the aim to recover complete mitochondrial genomes. The two libraries selected were generated from the blue antelope specimens NMW ST 715 from the Natural History Museum in Vienna and RMNH.MAM.20681.a from Naturalis Biodiversity Center in Leiden from Hempel et al. (2021).

Laboratory methods

Sampling and laboratory work for all of the newly extracted DNA samples (Natural History Museum London and National Museums Scotland) were carried out in the dedicated ancient DNA laboratory at the Natural History Museum in London. Two samples, one bone and one horn, were taken from frontlet and horns (NHMUK GERM 636e) and similarly from the skull and horns (NHMUK ZD 1888.8.20.3), and one sample from each of the three remaining specimens (Table 1). Bone/horn powder (23–53 mg per sample) was removed from each specimen using a Dremel drill at low speeds, with the skin being sampled as a ca. 1 cm strip using a scalpel. DNA was extracted from bone/horn following the Dabney et al. (2013) protocol with modifications; the Zymo-Spin V columns were replaced with Roche High Pure Viral Nucleic Acid Large Volume spin columns, and two final elution steps of 50 µl TET buffer (total 100 µl) were performed. For the skin sample the initial overnight extraction buffer (Dabney et al. 2013) was replaced with

Table 1 List of *Hippotragus* specimens analysed in this study

Specimen ID	Species according to museum records	Source location	Specimen prepara- tion	Sample material analysed	Species as geneti- cally identified by this study	Accession number
NHMUK ZD 1997.126/ GERM 636a	H. equinus	Little klibbolik- bonni spring, source of Kru- man River**, S. Africa	Right frontlet with right horn	Horn	H. equinus	OQ587914
NHMUK ZD 1888.8.20.3	H. leucophaeus and H. equinus	Upper banks of the River Gambia, the Gambia	Cranium with horns	Bone and horn	H. equinus	OQ587915
NHMUK ZD 1842.4.11.9	H. equinus	Cape Colony, S. Africa	Skin	Skin	H. equinus	OQ590032
NHMUK GERM 636e	H. leucophaeus	S. Africa	Frontlet with horns	Bone and horn	H. leucophaeus	OQ587911
NMS.Z.2004.105	H. leucophaeus	S. Africa	Frontlet with horns	Bone	H. equinus	OQ590031
NMW ST 715*	H. leucophaeus	-	Mounted skin	Skin	H. leucophaeus	OQ587912
RMNH.MAM.20681.a*	H. leucophaeus	-	Premaxilla and mandible frag- ments	Bone	H. leucophaeus	OQ587913

NHMUK: Natural History Museum, London; NMS: National Museums Scotland

*Two sequences re-analysed with hybridisation capture in this study

**The names are probably misspelled on the label, the correct spelling being "klibbolikhonni (fontein)" and "Kuruman" (this roan antelope maybe the very specimen mentioned by Skead (2011) and Sclater and Thomas (1899))

360 μ l Qiagen ATL buffer and 40 μ l Proteinase K, and the sample was incubated for 24 h in a thermo mixer at 56 °C. Post incubation the protocol for skin was identical to the protocol for bone. Dual indexed libraries for all extracts were constructed based on Meyer and Kircher (2010) and Kircher et al. (2012). Libraries were sequenced on an Illumina Next-Seq 500 producing 76 bp paired-end reads.

Hybridisation capture of the two single-stranded libraries from Hempel et al. (2021) for NMW ST 715 and RMNH. MAM.20681.a was carried out in the NGS laboratory at the Museum für Naturkunde Berlin, Germany. Libraries were re-amplified to obtain a sufficient quantity for subsequent hybridisation capture. Arbor Biosciences myBaits Custom 1-20K 16 Rxn kit (300116.v5-ARB) with 80 nucleotide baits every 20 nucleotides (standard Arbor Biosciences bait design) was used to enrich for targeted sequencing. Synthetic biotinylated RNA baits were designed complementary to sequence targets for blue antelope (MW222233), roan antelope (Hippotragus equinus, NC_020712), sable antelope (Hippotragus niger, NC_020713), takin (Budorcas taxicolor, NC_013069), ibex (Capra ibex, NC_020623), European mouflon (Ovis aries musimon, HM236184), muskox (Ovibos moschatus, NC_020631), and Tibetan antelope (Pantholops hodgsonii, NC_007441) (Xu et al. 2005; Hassanin et al. 2009, 2012; Meadows et al. 2011; Hempel et al. 2021). Assuming a high amount of non-target templates and short target sequences, we followed the High Sensitivity Protocol myBaits - Hybridisation Capture for Targeted NGS - Manual v5.00 September 2020. An average of 1 µg per library was used for capturing. Hybridisation was performed at 62 °C for 24 h for both rounds of enrichment. After each round of enrichment two PCR reactions were generated per enrichment reaction. Libraries were sequenced on an Illumina NextSeq 500 using custom primers producing 75 bp single-end reads at the University Potsdam, Germany (Paijmans et al. 2017).

Bioinformatic methods

For the BWA alignment, fastq files from samples of the same specimen were combined. Adapters (1 bp overlap) and reads shorter than 30 bp were removed with Cutadapt v2.8 (Martin 2011). Subsequently, reads were merged using FLASH v1.2.11 (Magoč and Salzberg 2011) with a maximum overlap of 76 bp for the paired-end reads. All unmerged reads were discarded. The resulting reads were mapped, using BWA aln v0.7.17 (Li and Durbin 2009) with default settings, to the three reference mitochondrial genomes sourced from GenBank: The blue antelope *Hippotragus leucophaeus* (MW222233) from Hempel et al. (2021), and the roan and sable antelope, *Hippotragus equinus* (JN632647) and *Hippotragus niger* (JN632648), respectively, both from Hassanin et al. (2012). Quality filtering was performed using SAMtools view v1.10 (Li et al. 2009) by removing reads with a quality < 30 and discarding unmapped reads. Reads were then sorted using SAMtools sort. Duplicate removal was performed using MarkDupsByStartEnd (https://github.com/dariober/Javacafe/tree/master/MarkDupsByStartEnd). For four samples (NHMUK ZD 1888.8.20.3, NHMUK GERM 636e, NMW ST 715 and RMNH.MAM.20681.a), single-base edits to reads were adjusted by eye, one per sample. Consensus sequences were generated using a 75% majority-rule threshold for base calling, a minimum coverage of $3 \times$ and the "trim to reference" option in Geneious R10 v10.2.3 (Kearse et al. 2012, https://www.geneious.com). A second mapping step was performed to improve the sequence coverage at the ends of the consensus sequences. All reads were mapped and filtered as described above using references with 152 bp or 75 bp, depending on sequencing cycle length, moved from the beginning to the end of the reference to account for the circularity of the mitochondrial genome. A consensus sequence was called and the 152 bp/75 bp shifted back from the end to the beginning of the sequence. In a final step both consensus sequences were aligned using the MAFFT algorithm v7.450 (Katoh et al. 2002; Katoh and Standley 2013) with default parameters as implemented in Geneious, and a consensus sequence was called from both sequences with a 50%majority rule threshold for base calling (option "50%-Strict: Bases matching at least 50% of the sequences").

Alignment statistics and coverage (number of bases aligned/genome length) were obtained using SAMtools depth (Li et al. 2009) and Qualimap (García-Alcalde et al. 2012), respectively. The alignments were visualised and inspected manually using Tablet (Milne et al. 2013) and Geneious, and alignment mapping quality was checked using mapDamage v2.2 (Jónsson et al. 2013), with all samples showing damage patterns indicative of degraded DNA.

To further corroborate species identification and avoid reference bias, a de novo approach to genome assembly was also used. Here, Illumina adapter sequences were trimmed, and paired-end reads merged using AdapterRemoval v2 (Schubert et al. 2016), with a minimum length of read set to 25 base pairs, and a mismatch rate of 3. We then applied mitochondrial baiting and iterative mapping software, MITObim v1.9.1 (Hahn et al. 2013), to de novo assemble mitochondrial genomes from the trimmed reads. To further eliminate potential reference bias, an outgroup species related to the target species, the gemsbok *Oryx gazella* (JN632678, Hassanin et al. 2012), was used as a starting 'seed'. MITObim was run using the -quick option and ten iterations (sufficient for completion) for each sample, with all other settings set to default.

Genome sequences from the BWA mapping approach plus the newly re-sequenced NMW ST 715 and RMNH. MAM.20681.a were aligned using MAFFT with default settings in Geneious, and Bayesian phylogenetic inference was carried out using MrBayes (Huelsenbeck and Ronquist 2001). Before aligning the sequences, control regions were removed from each sequence to ensure proper alignability. The three Hippotragus mitochondrial reference genomes (used for the comparative alignments), the other DNA-validated blue antelope MW222234 (as confirmed by Hempel et al. (2021)), the two re-analysed blue antelope genomes, four H. equinus mitochondrial genomes (representing the four mitochondrial lineages of roan antelope determined by Gonçalves et al. (2021), MZ560850, MZ560862, MZ560873, MZ560907) and the mitochondrial genome of Oryx gazella were included in the MAFFT alignment for phylogenetic analysis. Modeltest-ng (Darriba et al. 2020) was used to calculate the most appropriate substitution model for all alignments, with HKY + I + G4(nst=2, gamma=4) selected for all datasets. MrBayes MCMC was run for 1,000,000 generations, sampling every 1000 generations, discarding the first 25% as burn-in and ensuring convergence after the run. The Oryx gazella genome was used as an outgroup in all cases. The resulting phylogenetic tree was visualised and edited using FigTree v1.4.4 (https://github. com/rambaut/figtree). The process (alignment and phylogenetic analysis) was repeated, but with the sequences from the BWA mapping approach replaced with the sequences from the MITObim mapping approach (all other settings remained as above), to generate a second (MITObim) phylogeny.

To estimate the intraspecific diversity of the blue antelope, including our new sequence, the novel sequence of NHMUK GERM 636e, plus the newly re-sequenced NMW ST 715 and RMNH.MAM.20681.a, were aligned alongside the two existing complete blue antelope mitochondrial genomes from GenBank—MW222233 and MW222234 (Hempel et al. 2021)—using the MAFFT algorithm with default parameters as implemented in Geneious. All ambiguities and gaps were removed, resulting in an alignment length of 16,491 bp. Based on this alignment a median-joining network (Bandelt et al. 1999) was built using default parameters in POPART v1.7 (Leigh and Bryant 2015) and nucleotide diversity and number of segregating sites were calculated. The haplotype network was edited in Inkscape v1.2 (https://inkscape.org).

Results

Mitochondrial DNA genomes were successfully generated for all five specimens extracted for this study. All specimens showed a clear alignment preference in the numbers of quality-filtered, post-duplicate removal reads. Specimen NHMUK GERM 636e aligned most closely to the blue antelope reference genome, while the remaining samples, including Edinburgh specimen NMS.Z.2004.105, aligned more closely to the roan antelope reference genome (Table 2). Furthermore, the hybridisation capture approach yielded two complete mitochondrial genomes for both captured samples NMW ST 715 and RMNH.MAM.20681.a. NHMUK GERM 636e and NMW ST 715 have identical sequences. RMNH.MAM.20681.a is potentially also the same haplotype, but uncertainty at a single base has generated a Y in the control region where the other two samples have a C. There are minor differences in single positions between the re-analysed samples in this study and in Hempel et al. (2021) that are due to the higher sequence coverage in this study (Supplementary Tables S2 and S3).

The non-alignment-based, de novo assembly approach, using MITObim, successfully generated mitochondrial data for all five specimens sampled in this study. Complete mitochondrial genomes were generated for NHMUK ZD 1997.126/GERM 636a and NHMUK GERM 636e, and partial genomes for the remaining three samples (Supplementary Table S4). The complete mitochondrial genome for blue antelope specimen NHMUK GERM 636e was identical to the genome retrieved from the BWA alignment analysis, except for 18 ambiguities within a stretch of 57 bp in the control region. None of the ambiguities contrasted with the bases in the BWA-generated sequence.

The generated Bayesian inference phylogenies show the blue and sable antelope as sister taxa to the exclusion of the roan antelope, with species node posterior probability values all equal to 1. Sample NHMUK GERM 636e was placed with high probability within the blue antelope clade. The remaining samples were placed within the roan antelope clade: specimen NHMUK ZD 1842.4.11.9 groups with the Gonçalves et al.'s (2021) Southern lineage representative (MZ560907), specimen NMS.Z.2004.105 groups with the East-central lineage representative (MZ560873), and NHMUK ZD 1997.126/636a and NHMUK ZD 1888.8.20.3 fall within the roan antelope clade, but outside the Gonçalves et al. (2021) lineage representatives (Fig. 2). The MITObim phylogeny exhibits an identical and similarly well supported topology (Supplementary Fig. S5).

The median-joining network of the five complete blue antelope mitochondrial genomes showed seven segregating sites and a nucleotide diversity (pi) of 0.00025 (Fig. 3). Three of these sites were located in *ND2*, *ND4* and *ND5*, respectively, and four in the control region. The two sites in *ND2* and *ND4* were non-synonymous (Supplementary Table S5).

Discussion

This study has succeeded in generating a mitochondrial genome for specimen NHMUK GERM 636e and validated it genetically as an extinct blue antelope (*Hippotragus*

Table 2 Statistics fo	or all samples aligned usin	using BWA aln to the three reference Hippotragus genomes. The reference	e genomes blue, roan and sable antelope refer to the blue antelope Hippo
gus leucophaeus (N	1W222233) from Hempel	mpel et al. (2021), and the roan and sable antelope, <i>Hippotragus equinus</i> (J	JN632647) and <i>Hippotragus niger</i> (JN632648) from Hassanin et al. (20
The X-fold genome	coverage column denotes	otes the highest/most appropriate reference genome in bold	
	ų L		

Specimen	Reference genome	Raw read pairs (R1/ R2)	Read pairs mapped	% Read pairs mapped	Mapped read pairs after removing dupli- cates	% Mapped read pairs after duplicate removal	X-fold genome cover- age
NHMUK ZD	Blue antelope	2,538,585	6292	0.248	6142	0.242	23.366
1997.126/GERM	Roan antelope	2,538,585	32,026	1.262	31,238	1.231	132.382
050a	Sable antelope	2,538,585	5837	0.23	5700	0.225	21.556
NHMUK ZD	Blue antelope	5,467,375	5275	0.096	5180	0.095	17.531
1888.8.20.3	Roan antelope	5,467,375	22,208	0.406	21,743	0.398	79.682
	Sable antelope	5,467,375	5,032	0.092	4929	0.09	16.580
NHMUK ZD	Blue antelope	7,040,937	163	0.001	163	0.001	0.477
1842.4.11.9	Roan antelope	7,040,937	746	0.003	744	0.003	2.344
	Sable antelope	7,040,937	171	0.001	170	0.001	0.504
NHMUK GERM	Blue antelope	6,949,839	50,432	0.181	48,562	0.175	180.004
636e	Roan antelope	6,949,839	10,562	0.038	10,200	0,037	34.903
	Sable antelope	6,949,839	15,212	0.055	14,651	0.053	50.865
NMS.Z.2004.105	Blue antelope	3,813,191	245	0.002	244	0.002	0.797
	Roan antelope	3,813,191	954	0.006	944	0.006	3.325
	Sable antelope	3,813,191	212	0.001	212	0.001	0.670

Fig. 2 Bayesian inference mitochondrial DNA phylogeny, using the BWA alignment excluding the control region (alignment length 15,437 bp). *Oryx gazella* was used as an outgroup (JN632678, not shown). Samples sequenced for this study are shaded in grey



leucophaeus). This extends the number of DNA-validated historical museum specimens of this poorly documented species to five, and the overall number of specimens likely to six (this includes the Paris mounted skin) or seven (if the Leiden specimens are two individuals). NHMUK GERM 636e has the same, or almost the same, haplotype as the two improved blue antelope mitochondrial genomes for the Vienna and Leiden specimens (Fig. 3). It also provides an important addition to the very limited reference material available for this species. Other taxa that became extinct in the eighteenth and nineteenth centuries, like the quagga (Equus quagga quagga) or Steller's sea cow (Hydrodamalis gigas), are represented in museums by much larger numbers with ~ 34 individuals for the quagga and 27 skeletons, 62 additional skulls and more than 550 bones for Steller's sea cow (Rau 1974, 1978; Mattioli and Domning 2006).

The new mitochondrial genomes further support that mitochondrial nucleotide diversity within the blue antelope is low (nucleotide diversity of 0.00025). The low genetic diversity of the species has recently been confirmed by a nuclear DNA study of historical and prehistoric blue antelope specimens that further suggests that low genetic diversity was prevalent since at least the beginning of the Holocene (Hempel et al. 2022).

The purported (based on morphology) blue antelope specimen at the National Museums of Scotland (NMS.Z.2004.105) has been genetically identified here as a roan antelope. This misidentification occurred despite a direct comparison of the horns with the mounted specimen in Leiden, and demonstrates the difficulties with morphological differentiation between the three Hippotragus species when only horns are available. Similarly, the NHM specimen NHMUK ZD 1888.8.20.3, whose skull with horns (sampled in this study) is registered as H. leucophaeus (an unsampled skin with the same registration number [NHMUK ZD 1888.8.20.3] is labelled as Hippotragus equinus gambianus), has been genetically identified as a roan antelope (Hippotragus equinus). The two additional potential blue antelopes (based on source location) from the NHM collection have also been genetically identified as roan antelopes, increasing London - NHMUK GERM.636.e Vienna - NMW ST 715 Leiden - RMNH.MAM.20681.a



Fig. 3 Median-joining network of the complete mitochondrial genomes of the five currently genetically validated blue antelope specimens including the control region (alignment length 16,491 bp).

Dashes represent mutational steps. Size of circles corresponds to the frequency of the haplotype. Accession numbers are given in parentheses

the number of roan mitochondrial genomes sequenced in this study to four.

The Hippotragus phylogeny produced in this study places the blue and sable antelope as sister taxa, with the roan antelope as a sister to their clade (Fig. 2), which supports recent findings by Themudo and Campos (2018) and Hempel et al. (2021, 2022). The inclusion of the four mitochondrial genomes representing the four roan lineages, as identified by Gonçalves et al. (2021), allowed us to further explore the taxonomy of the roan antelope specimens newly identified by this study. We find that whilst specimen NHMUK ZD 1842.4.11.9 groups with the Gonçalves et al.'s (2021) Southern lineage representative, in accordance with its registered source location as Cape Colony S. Africa, the other three roan antelopes identified in our study did not group with the expected Gonçalves et al.'s (2021) lineages based on their registered source location. Specimen NMS.Z.2004.105 has a registered source location of South Africa and yet groups with the East-central lineage representative, and NHMUK ZD 1997.126/636a (source location South Africa) and NHMUK ZD 1888.8.20.3 (source location upper Gambia River, the Gambia) fall outside the Gonçalves et al. (2021) lineage representatives. Whilst our results are somewhat incongruent with the broad pattern identified in Gonçalves et al. (2021), exceptions to the pattern are not unprecedented as contact zones, multi-lineage locations and extensive Eastcentral haplotype introgression into the Southern lineage range were also reported by Gonçalves et al. (2021). We further note that our incongruent specimens are slightly older-NHMUK ZD 1997.126/636a was collected pre-1817 and NMS.Z.2004.105 in the early nineteenth century (no date is recorded for NHMUK ZD 1888.8.20.3), whereas the oldest Gonçalves et al. (2021) specimen is from 1842. In addition, two of these specimens-NHMUK ZD 1888.8.20.3 and NHMUK ZD 1997.126/636a-form a separate clade from the other lineages. There is a possibility that these could represent the roan lineage that was present at the southern Cape coast from the end of the Pleistocene to the Holocene (Faith 2012) and from which the type of the species originates according to Grubb (1999). However, assuming the labels are correct, one of the specimens is from West Africa (the Gambia) and the other one from further inland in South Africa, which renders the hypothesis that this is the extinct South African coastal lineage highly speculative. A larger phylogenetic study incorporating additional roan samples may help to resolve these issues.

Conclusions

This study highlights the usefulness of natural history collections and ancient DNA sequencing to confirm the taxonomic identity of historical specimens, particularly in cases of morphological similarity and limited specimen availability. It has further identified a novel blue antelope mitochondrial genome, extending the list of DNA-validated historical museum specimens to five and the overall list of blue antelope historical museum specimens to at least six. It also added complete mitochondrial genomes for two of these specimens, for which previously only partial genomes were available. We hope that future studies will verify additional blue antelope specimens to gain further insights into the species' population history and possible historical population structure.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42991-023-00373-4.

Acknowledgements The authors thank Tyler J. Faith and an anonymous reviewer for their helpful and constructive comments on an earlier version of the manuscript. We like to thank the curators Pepijn Kamminga (Naturalis Biodiversity Center), Daniela Kalthoff (Swedish Museum of Natural History), and Hans Mejlon (Museum of Evolution) for their collaboration in earlier projects from which results went into this study.

Author contributions All authors contributed to the study conception and design. Material preparation: RPM and ACK. Data generation and analysis: LP, EH, WM, IW and SB. First draft of the manuscript: LP and SB. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organised by Projekt DEAL.

Data availability The BioProject number in GenBank for this project is PRJNA910172. All untrimmed raw data was uploaded to the Sequence Read Archive under SRR22577643-9 (double-stranded libraries) and SRR23707135-SRR23707136 (single-stranded libraries). The mitochondrial genomes were uploaded to GenBank under OQ587911-OQ587915 and OQ590031-OQ590032. Alignments for the two Bayesian phylogenies and the median-joining network are available as supplementary material.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Avery DM (2019) A fossil history of southern African land mammals. Cambridge University Press, Cambridge
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036
- Broom R (1949) The extinct blue buck of South Africa. Nature 164:1097–1098. https://doi.org/10.1038/1641097b0
- Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, García N, Pääbo S, Arsuaga J-L, Meyer M (2013) Complete mitochondrial genome sequence of a middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc Natl Acad Sci USA 110:15758–15763. https://doi.org/10.1073/pnas. 1314445110
- Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol 37:291–294. https://doi.org/10.1093/molbev/msz189
- Faith JT (2012) Conservation implications of fossil roan antelope (*Hippotragus equinus*) in southern Africa's Cape Floristic Region. In: Louys J (ed) Paleontology in ecology and conservation. Springer, Berlin, pp 239–251
- Faith JT (2014) Late Pleistocene and Holocene mammal extinctions on continental Africa. Earth Sci Rev 128:105–121. https://doi.org/10. 1016/j.earscirev.2013.10.009
- Faith JT, Behrensmeyer AK (2013) Climate change and faunal turnover: testing the mechanics of the turnover-pulse hypothesis with South African fossil data. Paleobiology 39:609–627. https://doi. org/10.1666/12043
- Faith JT, Thompson JC (2013) Fossil evidence for seasonal calving and migration of extinct blue antelope (*Hippotragus leucophaeus*) in southern Africa. J Biogeogr 40:2108–2118. https://doi.org/10. 1111/jbi.12154
- FitzSimons FW (1920) The bluebuck or blaauwbok (*Hippotragus leu-cophaeus*). In: The natural history of South Africa. Longmans, Green and Co., London, pp 97–105
- García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A (2012) Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 28:2678–2679. https://doi.org/10.1093/bioinformatics/ bts503
- Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of extinction. In: Soulé ME (ed) Conservation biology: the science

of scarcity and diversity. Sinauer Associates, Sunderland, pp 19–34

- Gonçalves M, Siegismund HR, Jansen van Vuuren B, Koepfli K-P, Ferrand N, Godinho R (2021) *De novo* whole-genome assembly and resequencing resources for the roan (*Hippotragus equinus*), an iconic African antelope. G3 11:jkab002. https://doi.org/10.1093/ g3journal/jkab002
- Groves CP, Westwood CR (1995) Skulls of the blaauwbok *Hippotragus leucophaeus*. Z Säugetierkd 60:314–318
- Grubb P (1999) Types and type localities of ungulates named from southern Africa. Koedoe 42:13–45
- Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads — a baiting and iterative mapping approach. Nucleic Acids Res 41:e129. https://doi.org/10.1093/nar/gkt371
- Harper F (1945) Blaauwbok; Blue-buck. In: Extinct and vanishing mammals of the old world. The Lord Baltimore Press, Baltimore, pp 698–700
- 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 68:293–310. https://doi.org/10.1007/ s00239-009-9208-7
- Hassanin A, Delsuc F, Ropiquet A, Hammer C, Jansen van Vuuren B, Matthee C, Ruiz-Garcia M, Catzeflis F, Areskoug V, Nguyen TT, 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. https://doi.org/10.1016/j.crvi.2011.11.002
- Hempel E, Bibi F, Faith JT, Brink JS, Kalthoff DC, Kamminga P, Paijmans JLA, Westbury MV, Hofreiter M, Zachos FE (2021) Identifying the true number of specimens of the extinct blue antelope (*Hippotragus leucophaeus*). Sci Rep 11:2100. https://doi.org/10. 1038/s41598-020-80142-2
- Hempel E, Bibi F, Faith JT, Koepfli K-P, Klittich AM, Duchêne DA, Brink JS, Kalthoff DC, Dalén L, Hofreiter M, Westbury MV (2022) Blue turns to gray: paleogenomic insights into the evolutionary history and extinction of the blue antelope (*Hippotragus leucophaeus*). Mol Biol Evol 39:msac241. https://doi.org/10.1093/ molbev/msac241
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755. https://doi.org/10. 1093/bioinformatics/17.8.754
- Husson AM, Holthuis LB (1969) On the type of *Antilope leucophaea* Pallas, 1766, preserved in the collection of the Rijksmuseum van Natuurlijke Historie, Leiden. Zool Meded 44:147–157
- Husson AM, Holthuis LB (1975) The earliest figures of the blaauwbok, *Hippotragus leucophaeus* (Pallas, 1766) and of the greater kudu, *Tragelaphus strepsiceros* (Pallas, 1766). Zool Meded 49:57–65
- IUCN SSC Antelope Specialist Group (2017a) *Hippotragus equinus*. The IUCN red list of threatened species e.T10167A50188287. https://doi.org/10.2305/IUCN.UK.2017-2.RLTS.T10167A501 88287.en
- IUCN SSC Antelope Specialist Group (2017b) *Hippotragus niger*. The IUCN red list of threatened species e.T10170A50188654. https:// doi.org/10.2305/IUCN.UK.2017-2.RLTS.T10170A50188654.en
- Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29:1682–1684. https:// doi.org/10.1093/bioinformatics/btt193
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/ mst010

- Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066. https://doi.org/ 10.1093/nar/gkf436
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https:// doi.org/10.1093/bioinformatics/bts199
- Kerley GIH, Pressey RL, Cowling RM, Boshoff AF, Sims-Castley R (2003) Options for the conservation of large and medium-sized mammals in the Cape Floristic Region hotspot, South Africa. Biol Conserv 112:169–190. https://doi.org/10.1016/S0006-3207(02) 00426-3
- Kerley GIH, Sims-Castley R, Boshoff AF, Cowling RM (2009) Extinction of the blue antelope *Hippotragus leucophaeus*: modeling predicts non-viable global population size as the primary driver. Biodivers Conserv 18:3235–3242. https://doi.org/10.1007/ s10531-009-9639-x
- Kircher M, Sawyer S, Meyer M (2012) Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res 40:e3–e3. https://doi.org/10.1093/nar/gkr771
- Klein RG (1974) On the taxonomic status, distribution and ecology of the blue antelope, *Hippotragus leucophaeus* (Pallas, 1766). Ann S Afr Mus 65:99–143
- Kolb P (1719) Caput Bonae Spei hodiernum. Peter Conrad Monath, Nürnberg
- Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. Methods Ecol Evol 6:1110–1116. https://doi.org/10.1111/2041-210X.12410
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/ bioinformatics/btp352
- Lichtenstein H (1811) Reisen im südlichen Africa in den Jahren 1803, 1804, 1805 und 1806. C. Salfeld, Berlin
- Lichtenstein H (1814) Die Gattung Antilope. Ges Naturf Freunde Berlin Mag Neuesten Entdeck Gesammten Naturk 6(147–160):163–182
- Loubser J, Brink J, Laurens G (1990) Paintings of the extinct blue antelope, *Hippotragus leucophaeus*, in the eastern Orange Free State. S Afr Archaeol Bull 45:106–111. https://doi.org/10.2307/3887969
- Lydekker R (1914) Catalogue of the ungulate mammals in the British Museum (Natural History). Trustees of the British Museum, London
- Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957– 2963. https://doi.org/10.1093/bioinformatics/btr507
- Martin M (2011) Cutadapt removes adapter sequences from highthroughput sequencing reads. Embnet J 17:10–12. https://doi. org/10.14806/ej.17.1.200
- Mattioli S, Domning DP (2006) An annotated list of extant skeletal material of Steller's sea cow (*Hydrodamalis gigas*) (Sirenia: Dugongidae) from the Commander Islands. Aquat Mamm 32:273–288. https://doi.org/10.1578/AM.32.3.2006.273
- Meadows JRS, Hiendleder S, Kijas JW (2011) Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. Heredity 106:700–706. https://doi.org/10.1038/hdy. 2010.122
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring

Harb Protoc 2010:pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

- Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D (2013) Using tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14:193–202. https://doi.org/10.1093/bib/bbs012
- Mohr E (1967) Der Blaubock *Hippotragus leucophaeus* (Pallas, 1766) — Eine Dokumentation. Paul Parey, Hamburg
- Paijmans JLA, Baleka S, Henneberger K, Taron UH, Trinks A, Westbury MV, Barlow A (2017) Sequencing single-stranded libraries on the Illumina NextSeq 500 platform. http://arxiv.org/abs/1711. 11004. Accessed 21 July 2020
- Pallas PS (1767) Spicilegia zoologica quibus novae imprimis et obscurae animalium species. Gottlieb August Lange, Berlin
- Plug I (1997) Late Pleistocene and Holocene hunter–gatherers in the eastern highlands of South Africa and Lesotho: a faunal interpretation. J Archaeol Sci 24:715–727. https://doi.org/10.1006/jasc. 1996.0153
- Rau RE (1974) Revised list of the preserved material of the extinct cape colony quagga, *Equus quagga quagga* (Gmelin). Ann S Afr Mus 65:41–87
- 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. Ann S Afr Mus 77:27–45
- Robinson TJ, Bastos AD, Halanych KM, Herzig B (1996) Mitochondrial DNA sequence relationships of the extinct blue antelope *Hippotragus leucophaeus*. Naturwissenschaften 83:178–182. https://doi.org/10.1007/BF01143060
- Rookmaaker LC (1992) Additions and revisions to the list of specimens of the extinct blue antelope (*Hippotragus leucophaeus*). Ann S Afr Mus 102:131–141
- Schreyer J (1681) Neue Ost-Indianische Reiß-Beschreibung. Johann Christian Wohlfahrt, Leipzig
- Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Not 9:88. https://doi.org/10.1186/s13104-016-1900-2
- Sclater PL, Thomas O (1899) The Blue-buck. In: The book of antelopes. R. H. Porter, London, pp 5–12
- Skead CJ (1980) Blue antelope (extinct blaauwbok), *Hippotragus leucophaeus* (Pallas, 1766). In: Historical mammal incidence in the Cape Province, 2nd edn. Department of Nature and Environmental Conservation of the Provincial Administration of the Cape of Good Hope, Cape Town, pp 526–537
- Skead CJ (2011) Roan (Bastergemsbok). In: Boshoff AF, Kerley GIH, Lloyd PH (eds) Historical incidence of the larger land mammals in the broader western and northern Cape. Centre for African Conservation Ecology, Nelson Mandela Metropolitan University, Port Elizabeth, pp 352–356
- Smith A (1849) Illustrations of the zoology of South Africa. Smith, Elder and Co, London
- Stynder DD (2009) The diets of ungulates from the hominid fossilbearing site of Elandsfontein, Western Cape, South Africa. Quat Res 71:62–70. https://doi.org/10.1016/j.yqres.2008.06.003
- Themudo GE, Campos PF (2018) Phylogenetic position of the extinct blue antelope, *Hippotragus leucophaeus* (Pallas, 1766) (Bovidae: Hippotraginae), based on complete mitochondrial genomes. Zool J Linn Soc 182:225–235. https://doi.org/10.1093/zoolinnean/zlx034
- van Bruggen AC (1959) Illustrated notes on some extinct South African ungulates. S Afr J Sci 55:197–200
- Xu S-Q, Yang Y-Z, Zhou J, Jin G-E, Chen Y-T, Wang J, Yang H-M, Wang J, Yu J, Zheng X-G, Ge R-L (2005) A mitochondrial genome sequence of the Tibetan antelope (*Pantholops hodgsonii*). Genom Proteom Bioinform 3:5–17. https://doi.org/10.1016/ S1672-0229(05)03003-2

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.