Biodiversity and Conservation: Electronic Supplementary Material

Exploring the bushmeat market in Brussels, Belgium: a clandestine luxury business

Sophie Gombeer*, Casimir Nebesse, Prescott Musaba, Steve Ngoy, Marc Peeters, Ann Vanderheyden, Kenny Meganck, Nathalie Smitz, Frank Geers, Sarah Van Den Heuvel, Thierry Backeljau, Marc De Meyer and Erik Verheyen

* Corresponding author: Sophie Gombeer
Royal Belgian Institute of Natural Sciences
sophie.gombeer@naturalsciences.be

ESM_1: Additional information on the DNA-based identifications

Materials and methods

PCR reactions were carried out in a 20 µl final volume, containing 1.8 µl of template DNA (regardless of initial concentration), 0.2 mM dNTPs (each), 0.2 µM primers (each), 1x PCR buffer, 1.5 mM MgCl2 and 0.8 units of Platinum™ Taq DNA polymerase (Invitrogen, Thermo Fisher Scientific). PCR cycling conditions were 95 °C for 5 min followed by 35 cycles of 95 °C for 30 s, 40 °C (COI) or 51 °C (cytb) for 30 s, and 72 °C for 45 s (COI) or 1 min (cytb) with a final extension time of 72 °C for 10 min. Amplification success was checked on a 1 % agarose gel using UV-light and MidoriGreen Direct (Nippon Genetics Europe). PCR products were purified using the ExoSAP-IT™ (Thermo Scientific) method and send to Macrogen Europe (Netherlands) for sequencing. Raw sequences were trimmed and assembled using Geneious Prime® (Biomatters Ltd., Auckland, New Zealand). The consensus sequences were aligned with ClustalW (as implemented in Geneious Prime®) to control for gaps, translate the sequences into amino acids, map the primer pairs and trim the consensus sequences.

We used sequences present in both BOLD and GenBank for all members of the families to which the collected bushmeat samples belonged to construct the Neighbour-Joining trees. All available COI and cytb data (including sequences extracted from full mitochondrial genomes) were imported into Geneious Prime (Biomatters Ltd., Auckland, New Zealand). These datasets (one per family) were cleaned by removing sequences with incomplete species identifications, sequences that were defined as unverified in the sequence description, as well as very short sequences (< 270 and 300 nt for COI and cytb, respectively). Because the Thryonomyidae and Hystricidae datasets were very

small, they were analysed together. We deleted identical sequences per species to allow for the detection of eventual shared sequences among species within each family. In some cases the resulting datasets contained a large number of sequences for some species (e.g. *Bos taurus* and *Ovis aries*). In those instances we only kept the longest sequences, preferably from different projects (based on GenBank accession numbers), in our data set.

The respective sample sequences were added to each dataset, together with three to five outgroup sequences from species belonging to another family. Each dataset was aligned using ClustalW (as implemented in Geneious Prime) and subsequently trimmed to the full length of the generated fragments. In instances where a large number of sequences only covered part of the full fragment length, which was especially true for cytb, additional datasets were generated in which the alignment (including the sample sequence) was trimmed to a shorter length based on the length of the downloaded sequences (Table ESM_1a). Sequences of less than one third of the alignment length (after trimming), and sequences that contained a large number of ambiguity codes or caused gaps in the alignment, were deleted. An exception was made if it concerned sequences for a species in the genus of the bushmeat sample (as identified by the search engines) and no other sequences were available for that species. Sequences that did not align with any of the other sequences in the dataset were deleted, even if that removed the species from the dataset. The final dataset was realigned using ClustalW, after which the alignment was used to construct a Neighbour-Joining tree with Geneious Prime (genetic distance model: Tamurei-Nei). Bootstrapping (500 replicates) was used to assess tree node support. Trees were rooted with appropriate outgroup sequences.

Our final identification of the bushmeat samples was based on the combined outcome of both search engines and the most inclusive taxon assignment inferred from the NJ-trees. This was done taking the reliability and the comprehensiveness of the reference databases into account by examining the origin of sequences (e.g. published in a peer-reviewed paper vs. direct submission to GenBank), comparing the number of species and genera in the final NJ-trees with the number of presently accepted taxa as mentioned on the Integrated Taxonomic Information System website (ITIS; www.itis.gov/), and by examining if the closest relatives were present (based on phylogenetic studies in literature). The species assignments were considered correct when the searches using BLAST and BOLD resulted in a match with a single species with a sequence match > 97 % (Johnston et al. 2011) and the sample sequence belonged to a single-species cluster with a bootstrap support > 95 (Eaton et al. 2010). If either condition was not fulfilled, the sample sequence was assigned to multiple species or to the genus level.

Results & Discussion

We encountered several issues that may lead to erroneous species assignments or hamper accurate identifications. The first encountered issue concerns the taxonomic gaps in the reference databases, as also reported by other studies as the major cause for not reaching species-level identifications (Olayemi et al. 2011; D'Amato et al. 2013; Gaubert et al. 2015). Taxonomic gaps are common and problematic for poorly documented faunas like rare species and species-rich genera (Morf et al. 2013). *Cephalophus* and *Cercopithecus*, the two most speciose genera in our study, are well represented in the sequence datasets due to genus-wide studies (van Vliet et al. 2008; Ntie et al. 2010; Johnston et al. 2011; Johnston and Anthony 2012; Guschanski et al. 2013). An example in which the under-representation in reference sequence databases hinders species level identification is the case of our *Potamochurus* samples. In addition, DNA reference databases may contain erroneous species assignments. After examining the source of the sequences, it seems that reference sequences for *Atherurus africanus* and *Tragelaphus eurycerus*, were misidentified while they match some of our samples (see also Online Resources ESM_3). A final issue results from taxonomic uncertainties due to e.g. incomplete lineage sorting of recently diverged taxa, recent or ongoing hybridisation, or disagreements between gene trees and species trees due to marker resolution (Ntie et al. 2010; Eaton et al. 2010; D'Amato et al. 2013; Gaubert et al. 2015). This problem prevents species-level identification for several bovids (see also Online Resources ESM_3) and is especially relevant for the *Cephalophus* samples. Identifying duikers using DNA markers is not straightforward due to their recent diversification (Jansen van Vuuren and Robinson 2001; Johnston and Anthony 2012) and high intra-specific variability (Ntie et al. 2010).

Remarks accompanying Table ESM_1b: numbers refer to superscript numbers in Table ESM_1b, NJ-trees can be found in Online Resource 3.

(1) Two sequences (*Potamochoerus porcus*, AY534299; *P. larvatus*, AY534300) from the same study (Niebert and Tönjes 2005) do not cluster with other *Potamochoerus* sequences (Figure S3 NJ-trees 33-35). These sequences instead cluster with *Sus* sequences and are considered potential misidentifications.

(2) The *Bos* cluster in which the sample sequence cluster does not contain all *Bos* sequences (Figure S3 NJ-trees 1 & 2). There is no separation of the different species. The cluster also contains one *Bubalus bubalis* sequence (KT375483) which is probably a misidentification.

(3) The *Bos* cluster containing the sample sequence contains all *Bos taurus, B. indicus* and *B. primigenius* sequences, as well as sequences of several other *Bos* species (*B. frontalis, B. gaurus, B. javanicus*), but does not contain all *Bos* sequences (Figure S3 NJ-trees 14-17).

(4) Selected to be the most probable species considering the distribution of the other *Bos* species in the cluster.
(5) A *Sus scrofa* (KJ192746) sequence with 99.24 % ID was the next best match, yet this is considered a potential misidentification as no other *Sus* sequences closely match the sample sequence.
(6) One *Atherurus africanus* sequence (KJ192745) does not cluster with the other *A. africanus* sequences (Figure S3 NJ-tree 13) and probably involves a misidentification since the same pattern is observed for cytb (Figure S3 NJ-tree 36), while all other *A. atherurus* sequences from the same study (Gaubert et al. 2015) cluster.
(7) This sequence of *Atherurus macrourus* is the only full length cytb sequence for the *Atherurus* genus available on GenBank. All other *Atherurus* cytb sequences deposited in GenBank seem to be 402 bp or shorter. This is why the query coverage for the 'BLAST Best hit' is 35 % only.
(8) One *Atherurus africanus* sequence (KJ193303) does not cluster with the other *A. africanus* sequences (Figure S3 NJ-trees 36-37) and probably involves a misidentification since the same pattern is observed for COI (Figure S3 NJ-tree 13), while all other *A. atherurus* sequences from the same study (Gaubert et al. 2015) cluster.
(9a) The BOLD search output states: "a species level match could not be made, the queried specimen is likely to be one of the following: *Cephalophus callipygus, Cephalophus ogilbyi, Cephalophus weynsi*."
(9b) The BOLD search output states: "a species level match could not be made, the queried specimen is likely to be one of the following: *Cephalophus callipygus, Cephalophus ogilbyi*."
(10) The sequence belongs to a supported cluster containing all sequences from *Cephalophus callipygus, C. ogilbyi* and *C. weynsi* (Figure S3 NJ-trees 1-2, 5-6).
(11) The sequence belongs to a cluster containing all sequences from *Cephalophus callipygus, C. ogilbyi* and *C. weynsi* (Figure S3 NJ-trees 14-16, 19-21). The bootstrap support for this cluster ranges from 77 to 91.6 depending on the used dataset; e.g. all Bovidae *vs.* only Cephalophinae sequences, and long *vs.* trimmed sequences.
(12) Considered a potential misidentified sequence (see remark 6) because other *Atherurus africanus* sequences are available on GenBank, but no other appear in the search result.
(13) One *Thryonomys swinderianus* sequence (KJ192955) does not cluster with the other *T. swinderianus* sequences (Figure S3 NJ-tree 13) and probably involves a misidentification since the same pattern is observed for cytb (Figure S3 NJ-trees 36-37), while all other *T. swinderianus* sequences from the same study (Gaubert et al. 2015) cluster.
(14) One *Thryonomys swinderianus* sequence (KJ193490) does not cluster with the other *T. swinderianus* sequences (Figure S3 NJ-trees 36-37) and probably involves a misidentification since

the same pattern is observed for COI (Figure S3 NJ-tree 13), while all other *T. swinderianus* sequences from the same study (Gaubert et al. 2015) cluster.

(15) A *Tragelaphus eurycerus* (JF728784) sequence with 99.54 % ID is the next best match, yet this is considered a possible misidentification (see remark 16).

(16) One *Tragelaphus spekii* sequence (JF728788) does not cluster with all others (Figure S3 NJ-trees 14, 16-17, 23-25) and may involve a misidentification (e.g. sample swap with a *T. eurycerus* (JF728784) specimen from the same study (Schikora 2012)).

(17) The BOLD search output states: "a species level match could not be made. The nearest match is with *Cercopithecus neglectus*".

(18) The sample sequence always clusters with (a few) *Philantomba monticola* sequences (Figure S3 NJ-trees 14-17, 19-22). The bootstrap support for this cluster ranges from 76.6 to 100 and depends on the used database used; e.g. all Bovidae *vs.* only Cephalophinae sequences, and long *vs.* trimmed sequences.

(19) The sample sequence clusters with all available *Cercopithecus ascanius* sequences (Figure S3 NJ-trees 9-11), yet in two of the three analyses, this cluster also includes of a number of *C. cephus* sequences (GQ144551, GQ144553, GQ144554, GQ144556; Eaton et al., 2010). When using only *Cercopithecus* sequences to build the NJ-tree (Figure S3 NJ-tree 11), *Cercopithecus ascanius* is monophyletic and supported with a high bootstrap value (95.2).

Identification details

Potamochoerus sp. (BXL001, BXL005 & BXL007)

An underrepresentation of sequences from the whole geographic range prevents a species-level identification for the *Potamochoerus* samples. In the NJ-trees the sample sequences (BXL001, BXL005 & BXL007) cluster with *Potamochoerus* sequences, forming a sister clade to *P. larvatus* for COI and to *P. porcus* for cytb. Considering this difference between the two markers, the low number of available sequences for *P. larvatus*, and the low sequence similarities from the search engines, the sample sequences could only be reliably identified to the genus level. Indeed, for *P. larvatus*, having a larger and partially overlapping range distribution with *P. porcus*, only one sequence was available for each of the markers, originating from Zambia (COI) or Zimbabwe (cytb). This incomplete geographic sampling of *P. larvatus* might have resulted in an underrepresentation of intra-specific genetic variation. This might in turn explain the observed low % sequence similarities and why the sample sequences cluster differently depending on the fragment used. This example highlights the importance of knowledge on intra-species variability and a well-represented presence of closely related species sequences.

Bos taurus / indicus (BXL002 & BXL006)

The samples are identified as *Bos taurus* using both search engines, yet with a next closest species with (near) maximum sequence identity. In the NJ-trees the two sample sequences cluster with several other *Bos* species, including the next best match identified by BLAST. The inability to distinguish among the *Bos* species is due to the poor taxonomic resolution, especially for COI. Tobe et al. (2010) found low inter-species genetic distances among the domesticated species *B. taurus* and *B. indicus* and suggested it might be due to misidentification of some reference samples or to potential hybrid individuals. For our study, it is sufficient that we can rule out the possibility that these two samples originate from wild animals. Considering the different species in the cluster and their geographic distribution the unknown samples most probably belong to the domesticated *B. taurus* or *B. indicus*.

Atherurus africanus (BXL003)

The identification is clear, especially when considering the outcome of the NJ-trees in which the sample sequences cluster with *Atherurus africanus* with maximal support. This interpretation is based, however, on the assumption that some sequences in the NJ-trees are misidentified (Table S1b). The downloaded sequences of *A. africanus* that do not cluster with all other *A. africanus* sequences concern a case where species identifications were based on secondary morphological features of smoked bushmeat samples (Gaubert et al. 2015).

Cephalophus callipygus / ogilbyi / weynsi (BXL004 & BXL008)

Two samples are identified by BLAST as *Cephalophus callipygus* with *C. ogilbyi* as a next best match, while the BOLD engine indicates it is probably one of the following species: *C. callipygus, C. ogilbyi* or *C. weynsi* (Table S1b). When considering the Bovidae and Cephalophinae NJ-trees, the sample sequences cluster with the three *Cephalophus* species identified by BOLD with limited (cytb) to high (COI) support; neither COI, nor cytb trees allow to identify the *Cephalophus* samples to the species level. This agrees with earlier studies (Ntie et al. 2010; Johnston et al. 2011) that demonstrated that *C. callipygus, C. ogilbyi* and *C. weynsi* cannot be differentiated. Other studies have shown that also other mitochondrial markers such as 12S rRNA (van Vliet et al. 2008), the control region (Ntie et al. 2010) or the whole mitochondrial genome (Hassanin et al. 2012) fail to distinguish among these three African duiker species. In addition, Johnston and Anthony (2012) observed that the inclusion of multiple nuclear markers does not solve this issue, however, *C. weynsi* was missing from their analyses. Hassanin et al. (2012) therefore questioned the species status of these three taxa and suggested a re-analysis of the alpha taxonomy and geographic ranges of the red duiker species.

Thryonomys swinderianus (BXL009 & BXL012)

The identification is straightforward, especially when considering the outcome of the NJ-trees in which the sample sequences clustered with *Thryonomys swinderianus* with maximal support. This interpretation assumes that some sequences in the NJ-trees are misidentified (Table S1b). The identifications of the downloaded sequences of *T. swinderianus* that do not cluster with the other *T. swinderianus* sequences turn out to be based on secondary morphological features of smoked bushmeat samples (Gaubert et al. 2015).

The absence of data for the congener *T. greogorianus* may be a problem if this non-represented congener shares identical sequences with *T. swinderianus*. The geographic distribution range of *T. gregorianus* virtually overlaps with the range of *T. swinderianus*. Therefore this identification should be regarded cautiously, notwithstanding the high bootstrap supports of the recovered clusters as well as the near maximum % pairwise identity (for COI).

Tragelaphus spekii (BXL010)

For the sample sequence identified as *Tragelaphus spekii*, there is a clear drop in % pairwise identity between best hit and next closest species, and clustering of *T. spekii* sequences in the NJ-trees with (near) maximum support. This interpretation is based, however, on the assumption that the identity of one *Tragelaphus spekii* and one *T. eurycerus* sequences may have been swapped before submission to GenBank (Table S1b).

Cercopithecus neglectus (BXL011 & VRT1)

The different identification methods applied agree for the two *Cercopithecus neglectus* samples with (near) maximal support for the cluster in all NJ-trees and a gap in the % pairwise identity between best hit and next closest species for both COI and cytb.

Philantomba monticola (VRT2)

The unknown sample assigned to *Philantomba monticola* can only be identified based on cytb due to the failure to sequence COI. The *P. monticola* cytb sequences do not always form a monophyletic group, but the sample sequence always clusters with *P. monticola* sequences, whereas the support of the cluster depends on the used dataset.

Cercopithecus ascanius (RTBF1)

Four COI sequences of *Cercopithecus cephus* from the same study (Eaton et al. 2010) cluster with *C. ascanius* instead of with the other *C. cephus* sequences. However, when considering the genus-level

COI NJ-tree, as well as all cytb NJ-trees, *C. ascanius* forms a monophyletic group with high support which includes the sample sequence.

References

D'Amato ME, Alechine E, Cloete KW, et al (2013) Where is the game? Wild meat products authentication in South Africa: a case study. Investig Genet 4:. https://doi.org/10.1186/2041-2223-4-6

Eaton MJ, Meyers GL, Kolokotronis S-O, et al (2010) Barcoding bushmeat: molecular identification of Central African and South American harvested vertebrates. Conserv Genet 11:1389–1404

Gaubert P, Njiokou F, Olayemi A, et al (2015) Bushmeat genetics: setting up a reference framework for the DNA typing of African forest bushmeat. Mol Ecol Resour 15:633–651

Guschanski K, Krause J, Sawyer S, et al (2013) Next-Generation Museomics Disentangles One of the Largest Primate Radiations. Syst Biol 62:539–554

Hassanin A, Delsuc F, Ropiquet A, et al (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

Jansen van Vuuren B, Robinson TJ (2001) Retrieval of Four Adaptive Lineages in Duiker Antelope: Evidence from Mitochondrial DNA Sequences and Fluorescence in Situ Hybridization. Mol Phylogenet Evol 20:409–425

Johnston AR, Anthony NM (2012) A multi-locus species phylogeny of African forest duikers in the subfamily Cephalophinae: evidence for a recent radiation in the Pleistocene. BMC Evol Biol 12:120. https://doi.org/10.1186/1471-2148-12-120

Johnston AR, Morikawa MK, Ntie S, Anthony NM (2011) Evaluating DNA barcoding criteria using African duiker antelope (Cephalophinae) as a test case. Conserv Genet 12:1173–1182

Morf NV, Wood KL, Köppel R, et al (2013) A multiplex PCR method to identify bushmeat species in wildlife forensics. Forensic Sci Int Genet Suppl Ser 4:. https://doi.org/10.1016/J.FSIGSS.2013.10.104

Niebert M, Tönjes RR (2005) Evolutionary Spread and Recombination of Porcine Endogenous Retroviruses in the Suiformes. J Virol 79:649–654. https://doi.org/10.1128/jvi.79.1.649-654.2005

Ntie S, Johnston AR, Mickala P, et al (2010) A molecular diagnostic for identifying central African forest artiodactyls from faecal pellets. Anim Conserv 13:80–93

Olayemi A, Oyeyiola A, Antunes A, et al (2011) Contribution of DNA-typing to bushmeat surveys: assessment of a roadside market in south-western Nigeria. Wildl Res 38:696–716

Schikora TF (2012) Climate-linked temporal and spatial patterns in the evolution of African Bovidae. Goethe Universitat, Frankfurt am Main, Germany

Tobe SS, Kitchener AC, Linacre AMT (2010) Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome b and Cytochrome Oxidase Subunit I Mitochondrial Genes. PLoS

One 5:. https://doi.org/10.1371/journal.pone.0014156

van Vliet N, Zundel S, Miquel C, et al (2008) Distinguishing dung from blue, red and yellow-backed duikers through noninvasive genetic techniques. Afr J Ecol 46:411–417

Table ESM_1a: Summary of the datasets used to construct the NJ-trees for each of the families and each of the markers specifying the number of sequences, the number of represented genera and species, the length of the alignment and the samples identified using this dataset. The numbers under the (sub)family and genus names in the first two columns indicate the number of known genera and species in these taxa according to the Integrated Taxonomic Information System (ITIS; itis.gov).

			COI					cytb					
Family	Genus/Subfamily	Dataset [a]	NJ-tree [b]	# seq	# gen	# pp.	align (bp)	NJ-tree [b]	# seq	# gen	# spp.	align (bp)	Samples
Bovidae		all	1	982	52	134	658	14	788	60	160	1140	BXL002, BXL004, BXL006, BXL008, BXL010, VRT2
(50-144)		long	2	825	51	130	658	15	621	55	151	1140	
		trimmed A						16	729	55	151	708	
		trimmed B						17	642	57	151	439	
	Bos	all	3	195	1	8	658	18	72	1	9	1140	BXL002, BXL006
	(5)	long	4	176	1	8	658						
	Cephalophinae	all	5	126	3	18	658	19	117	3	20	1140	BXL004, BXL008, VRT2
	(3-18)	long	6	116	3	18	658	20	54	3	19	1140	
		trimmed A						21	97	3	19	514	
		trimmed B						22	74	3	20	402	
	Tragelaphus	all	7	39	1	8	658	23	91	1	8	1140	BXL010
	(7)	long	8	31	1	8	658	24	44	1	8	1140	
		trimmed						25	90	1	8	556	
Cercopithecidae		all	9	541	23	98	658	26	511	22	115	1140	BXL011, VRT1, RTBF1
(23-154)		long	10	511	23	98	658	27	345	22	101	1140	
		trimmed A						28	437	22	108	597	
		trimmed B						29	429	22	110	402	
	Cercopithecus	all	11	93	1	19	658	30	75	1	18	1140	BXL011, VRT1, RTBF1
	(20)	long						31	57	1	17	1140	
		trimmed						32	72	1	18	402	
Suidae		all	12	149	4	8	659 [c]	33	123	5	13	1140	BXL001, BXL005, BXL007
(5-19)		long						34	83	5	13	1140	
		trimmed						35	121	5	13	439	
Thryonomyidae (1-2)		all	13	35	4	5	658	36	56	4	7	1140	BXL003, BXL009, BXL012
& Hystricidae (3-11)		trimmed						37	56	4	7	402	

[a] all: all sequences that passed through the selection criteria (see methods section) are included in the NJ-tree; long: only (nearly) full length sequences are included in the NJ-tree; trimmed: the alignment of the 'all' dataset was trimmed because multiple sequences only covered a fragment of the marker; A and B: the dataset was trimmed in two different ways.

[b] The number refers to the number in the NJ-tree file name in Online Resource 3.

[c] One sequence which was causing a single gap in the alignment was withheld for NJ-tree construction in order to keep the species in the analysis.

Table ESM_1b: Details on the outcome of the search engines BLAST and BOLD as well as on the position of the sample sequences in the NJ-trees for both COI and cytb.

Sample	Sold as	Scientific name	English name	Marker	Seq length (bp)	BLAST* Best hit	Percent Identity (%)	Query Coverage (%)	Accession nr	Next closest species	Percent Identity (%)
BXL001	Ngulu ya zamba	*Potamochoerus porcus*	Red river hog	COI	658	*Potamochoerus larvatus*	96.64	76	JQ690390	*Potamochoerus porcus*	95.24
				cytb	1128	*Potamochoerus larvatus*	94.85	99	GQ338966	*Potamochoerus porcus*	93.70
BXL002	Pakasa	*Syncerus caffer*	African buffalo	COI	658	*Bos taurus*	100	100	MN200938	*Bos indicus*	100
				cytb	1132	*Bos taurus*	100	100	MN200921	*Bos indicus*	99.91
BXL003	Simbiliki	*Thryonomys* sp.	Cane rat	COI	658	*Atherurus africanus*	99.70	100	KJ192734	*Hystrix cristata* [5]	84.49
				cytb	1107	*Atherurus africanus*	99.75	35	KJ193297	*Atherurus macrourus* [7]	83.82
BXL004	Antilope	*Cephalophus* sp.	Duiker	COI	658	*Cephalophus callipygus*	100	100	JN632614	*Cephalophus ogilbyi*	99.85
				cytb	1101	*Cephalophus callipygus*	99.82	100	JN632614	*Cephalophus ogilbyi*	96.46
BXL005	Ngulu ya zamba	*Potamochoerus porcus*	Red river hog	COI	658	*Potamochoerus larvatus*	96.64	76	JQ690390	*Potamochoerus porcus*	95.24
				cytb	1065	*Potamochoerus larvatus*	94.84	100	GQ338966	*Potamochoerus porcus*	94.08
BXL006	Pakasa	*Syncerus caffer*	African buffalo	COI	658	*Bos taurus*	100	100	MN200938	*Bos indicus*	100
				cytb	1112	*Bos taurus*	100	77	KT260196	*Bos indicus*	99.91
BXL007	Antilope	*Cephalophus* sp.	Duiker	COI	658	*Potamochoerus larvatus*	96.64	76	JQ690390	*Potamochoerus porcus*	95.24
				cytb	1105	*Potamochoerus larvatus*	95.02	99	GQ338966	*Potamochoerus porcus*	93.94
BXL008	Antilope	*Cephalophus* sp.	Duiker	COI	658	*Cephalophus callipygus*	100	100	HQ644088	*Cephalophus ogilbyi*	98.78
				cytb	1114	*Cephalophus callipygus*	98.92	100	JN632612	*Cephalophus ogilbyi*	97.76
BXL009	Simbiliki	*Thryonomys* sp.	Cane rat	COI	658	*Thryonomys swinderianus*	100	100	KJ192912	*Atherurus africanus* [12]	99.70
				cytb	1110	*Thryonomys swinderianus*	90.65	100	AJ301644	*Sorex fumeus*	80.26
BXL010	Antilope	*Cephalophus* sp.	Duiker	COI	658	*Tragelaphus spekii*	99.39	100	KJ192918	*Tragelaphus eurycerus*	95.28
				cytb	1090	*Tragelaphus spekii*	99.72	100	AJ222680	*Tragelaphus eurycerus* [15]	94.68
BXL011	Makako	*Cercocebus* sp.	Mangabey	COI	658	*Cercopithecus neglectus*	97.43	82	JQ256929	*Cercopithecus roloway*	89.55
				cytb	1110	*Cercopithecus neglectus*	96.94	100	JQ256930	*Cercopithecus nicitans*	91.57
BXL012	Simbiliki	*Thryonomys* sp.	Cane rat	COI	658	*Thryonomys swinderianus*	100	100	KJ192912	*Atherurus africanus* [12]	99.70
				cytb	1116	*Thryonomys swinderianus*	90.69	100	AJ301644	*Sorex fumeus*	80.32
VRT1	Makako	*Cercocebus* sp.	Mangabey	COI	658	*Cercopithecus neglectus*	97.43	82	JQ256929	*Cercopithecus solatus*	90.82
				cytb	1107	*Cercopithecus neglectus*	96.84	100	JQ256930	*Cercopithecus nicitans*	91.57
VRT2	Antilope	‡	Antelope	COI	-	-	-	-	-	-	-
				cytb	1086	*Philantomba monticola*	99.82	100	JN632686	*Philantomba maxwellii*	93.74
RTBF1	Antilope	‡	Antelope	COI	658	*Cercopithecus ascanius*	99.78	81	JQ256938	*Cercopithecus cephus*	97.05
				cytb	1111	*Cercopithecus ascanius*	98.78	96	JQ256938	*Cercopithecus erythrotis*	94.96

* Details of the BLAST searches: performed October 2019, blastn, nucleotide collection (nr/nt) database, optimisation for highly similar seq (megablast), ranking of sequences by percent identity.

^ Details of the BOLD searches: animal identification [COI], current database (October 2019), species level barcode records (3,573,026 Sequences/211,694 Species/92,285 Interim Species), hits sorted by similarity (%).

° For a detailed discussion on the final identification, please refer to ESM_1.

‡ Samples were not collected by co-authors PM, CN and SN, so no additional information to interpret the local name was available. Therefore we choose to include all antelopes instead of restricting the translation to

† long = bootstrap value for the cluster using the dataset with long sequences; max = maximal support considering all different datasets.

Query Coverage (%)	Accession nr	BOLD^ Best ID	ID%	Sample ID	NJ-trees Species	Monophyletic Cluster	Bootstrap Support†	FINAL ID°
98	KJ192898	No match	-	-	*Potamochoerus larvatus*	Y	99.4	*Potamochoerus* sp.
99	DQ315602				*Potamochoerus porcus*	Y[1]	99.4	
100	KY650680	*Bos taurus*	100%	KU947028	*Bos* sp.	N[2]	100 (long)	*Bos taurus / indicus* [4]
100	JN817330				*Bos* sp.	N[3]	100	
96	KX241536	*Atherurus africanus*	99.69%	KJ192737	*Atherurus africanus*	Y[6]	100	*Atherurus africanus*
98	FJ931121				*Atherurus africanus*	Y[8]	99.4	
100	HQ644109	*Cephalophus callipygus* [9a]	100%	GQ144489	*Cephalophus* species	N[10]	100 (long)	*Cephalophus callipygus / ogilbyi / weynsi*
100	AF153897				*Cephalophus* species	N[11]	max 91.6	
98	KJ192898	No match	-	-	*Potamochoerus larvatus*	Y	99.4	*Potamochoerus species*
100	DQ315602				*Potamochoerus porcus*	Y[1]	99.4	
100	KY650680	*Bos taurus*	100%	KU947028	*Bos* sp.	N[2]	100 (long)	*Bos taurus / indicus* [4]
100	JN817330				*Bos* sp.	N[3]	100	
98	KJ192898	No match	-	-	*Potamochoerus larvatus*	Y	99.4	*Potamochoerus* sp.
100	EU189380				*Potamochoerus porcus*	Y[1]	99.4	
100	KJ192789	*Cephalophus callipygus* [9b]	100%	HQ644088	*Cephalophus* species	N[10]	100 (long)	*Cephalophus callipygus / ogilbyi / weynsi*
100	JN632620				*Cephalophus* species	N[11]	max 91.6	
100	KJ192745	*Thryonomys swinderianus*	100%	KJ192912	*Thryonomys swinderianus*	Y[13]	100	*Thryonomys swinderianus*
94	KF302841				*Thryonomys swinderianus*	Y[14]	100	
99	KY650692	*Tragelaphus spekii*	99.39%	KJ192918	*Tragelaphus spekii*	Y	100	*Tragelaphus spekii*
100	JN632703				*Tragelaphus spekii*	Y[16]	100	
98	JQ256923	*Cercopithecus neglectus* [17]	97.35%	AY972784	*Cercopithecus neglectus*	Y	100	*Cercopithecus neglectus*
76	JQ256952				*Cercopithecus neglectus*	Y	100	
100	KJ192745	*Thryonomys swinderianus*	100%	KJ192912	*Thryonomys swinderianus*	Y[13]	100	*Thryonomys swinderianus*
94	KF302841				*Thryonomys swinderianus*	Y[14]	100	
29	JQ256920	*Cercopithecus neglectus* [17]	97.35%	AY972784	*Cercopithecus neglectus*	Y	100	*Cercopithecus neglectus*
77	JQ256952				*Cercopithecus neglectus*	Y	100	
-	-	-	-	-	-	-	-	*Philantomba monticola*
100	JF728780				*Philantomba monticola*	Y[18]	max 100	
98	GQ144556	*Cercopithecus ascanius*	99.44%	Private sequence	*Cercopithecus ascanius*	N[19]	max 95.2	*Cercopithecus ascanius*
100	JQ256936				*Cercopithecus ascanius*	Y	100	

.

duikers only.