Molecular Phylogenetics and Evolution 113 (2017) 150–160

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Reconstructing the molecular phylogeny of giant sengis (Macroscelidea; Macroscelididae; *Rhynchocyon*)

Elizabeth J. Carlen [a,b,1], Galen B. Rathbun [a], Link E. Olson [c], Christopher A. Sabuni [d], William T. Stanley [e], John P. Dumbacher [a,b,*]

[a] Institute for Biodiversity Science and Sustainability, California Academy of Sciences, San Francisco, CA 94118, United States
[b] Department of Biology, San Francisco State University, San Francisco, CA 94132, United States
[c] University of Alaska Museum, Fairbanks, AK 99775, United States
[d] Pest Management Centre, African Centre of Excellence, Sokoine University of Agriculture, Morogoro, Tanzania
[e] Science & Education, Field Museum of Natural History, Chicago, IL 60605, United States

ARTICLE INFO

Article history:
Received 14 February 2017
Revised 5 May 2017
Accepted 12 May 2017
Available online 24 May 2017

Keywords:
Rhynchocyon
Giant sengis
Elephant-shrews
Africa
Macroscelididae
Phylogenetics
Taxonomy

ABSTRACT

Giant sengis (Macroscelidea; Macroscelididae; *Rhynchocyon*), also known as giant elephant-shrews, are small-bodied mammals that range from central through eastern Africa. Previous research on giant sengi systematics has relied primarily on pelage color and geographic distribution. Because some species have complex phenotypic variation and large geographic ranges, we used molecular markers to evaluate the phylogeny and taxonomy of the genus, which currently includes four species: *R. chrysopygus*, *R. cirnei* (six subspecies), *R. petersi* (two subspecies), and *R. udzungwensis*. We extracted DNA from fresh and historical museum samples from all taxa except one *R. cirnei* subspecies, and we generated and analyzed approximately 4700 aligned nucleotides (2685 bases of mitochondrial DNA and 2019 bases of nuclear DNA) to reconstruct a molecular phylogeny. We genetically evaluate *Rhynchocyon* spp. sequences previously published on GenBank, propose that the captive *R. petersi* population in North American zoos is likely *R. p. adersi*, and suggest that hybridization among taxa is not widespread in *Rhynchocyon*. The DNA sample we have from the distinctive but undescribed giant sengi from the Boni forest of northern coastal Kenya is unexpectedly nearly identical to *R. chrysopygus*, which will require further study. Our analyses support the current morphology-based taxonomy, with each recognized species forming a monophyletic clade, but we propose elevating *R. c. stuhlmanni* to a full species.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

The 19 extant species of sengis (elephant-shrews; Rathbun and Kingdon, 2006) in the mammalian Order Macroscelidea are restricted to the African continent and form two well-defined subfamilies, the soft-furred sengis (Macroscelidinae), with 15 extant species in four genera (*Elephantulus*, *Macroscelides*, *Petrodromus,* and *Petrosaltator*), and the giant sengis (Rhynchocyoninae), with four extant species in one genus (*Rhynchocyon*).

Despite their long evolutionary history (Novacek, 1984) and broad African distribution (Corbet and Hanks, 1968) in highly diverse habitats across much of Africa (Rathbun, 2009), sengis have proven to be taxonomically challenging, having relatively few discretely varying morphological traits with which to resolve their phylogeny and taxonomy (Corbet and Hanks, 1968). With the application of molecular genetics in the last several decades, some insights into extant sengi phylogeny and taxonomy have been gained. This work has shown that Macroscelididae are morphologically specialized, yet across a diversity of habitats, they maintain a stable life history and morphology that has masked some of their evolutionary and ecological diversity (Rathbun, 2009).

Abbreviations: 12s16s, 12S rRNA, valine tRNA, and 16S rRNA; AMNH, American Museum of Natural History; BLAST, basic local alignment search tool; BMNH, Natural History Museum, London; Bpp, Bayesian posterior probability; CAS, California Academy of Sciences; CASMAM, California Academy of Sciences Mammalogy; D-loop, hypervariable 5′ end of the control region; ENAM, Enamelin; FMNH, Field Museum of Natural History; IRBP, inter-photoreceptor retinoid-binding protein; MCMC, Markov Chain Monte Carlo; MCZ, Museum of Comparative Zoology; mlb, maximum likelihood bootstrap; MTSN, Museo Tridentino di Scienze Naturali; NCBI, National Center for Biotechnology Information; ND2, NADH dehydrogenase 2; PCR, polymerase chain reaction; RAxML, Random Axelerated Maximum Likelihood; RMCA, Royal Museum of Central Africa; SNP, single nucleotide polymorphism; UAM, University of Alaska Museum; vWF, Von Willebrand factor.

* Corresponding author.
E-mail address: jdumbacher@calacademy.org (J.P. Dumbacher).
[1] Present address: Department of Biology, Fordham University, Bronx, NY 10458, United States.

http://dx.doi.org/10.1016/j.ympev.2017.05.012
1055-7903/© 2017 Elsevier Inc. All rights reserved.

Giant sengis, as their name indicates, are the largest members of the order, with body masses ranging from 300 g to 700 g. They are diurnal, swift quadrupedal forest-floor dwellers with proportionally long legs, a long sparsely-haired tail, and a long snout that can twist and probe in leaf litter in search of invertebrate prey. The golden-rumped sengi (*R. chrysopygus*) is the only giant sengi whose behavioral ecology has been studied in sufficient detail to reveal that its life history is unusual for a small mammal (Rathbun, 2009). Individuals form monogamous pairs on territories, shelter singly in leaf nests on the forest floor, and produce one relatively precocial offspring at a time (FitzGibbon, 1997; Rathbun, 1979).

In the 65 years between 1847 and 1912, ten species and four subspecies of *Rhynchocyon* were described. Corbet and Hanks (1968), using mostly distinctive pelage color patterns (Fig. 1) and allopatric distributions (Fig. 2), conducted a thorough taxonomic revision of the order, resulting in only three recognized giant sengi species. The golden-rumped sengi (*R. chrysopygus*) is monotypic and occurs in coastal Kenya. The black-and-rufous sengi (*R. petersi),* has two subspecies: *R. p. adersi* from islands off Tanzania and *R. p. petersi* from mainland Tanzania and Kenya (Fig. 2). The checkered sengi (*R. cirnei*) has six subspecies: *R. c. cirnei* from Mozambique and southern Malawi, *R. c. shirensis* from the Shire Valley of southern Malawi, *R. c. reichardi* from Tanzania, Malawi, and Zambia highlands, *R. c. hendersoni* from highlands of northern Malawi, *R. c. macrurus* from southeastern Tanzania lowlands, and *R. c. stuhlmanni* from the Congo Basin and western Uganda (Fig. 2). *Rhynchocyon c. shirensis* was a new taxon (Corbet and Hanks, 1968), whereas the other subspecies had previously been described as full species. Corbet and Hanks (1968) also noted that *R. c. stuhlmanni*

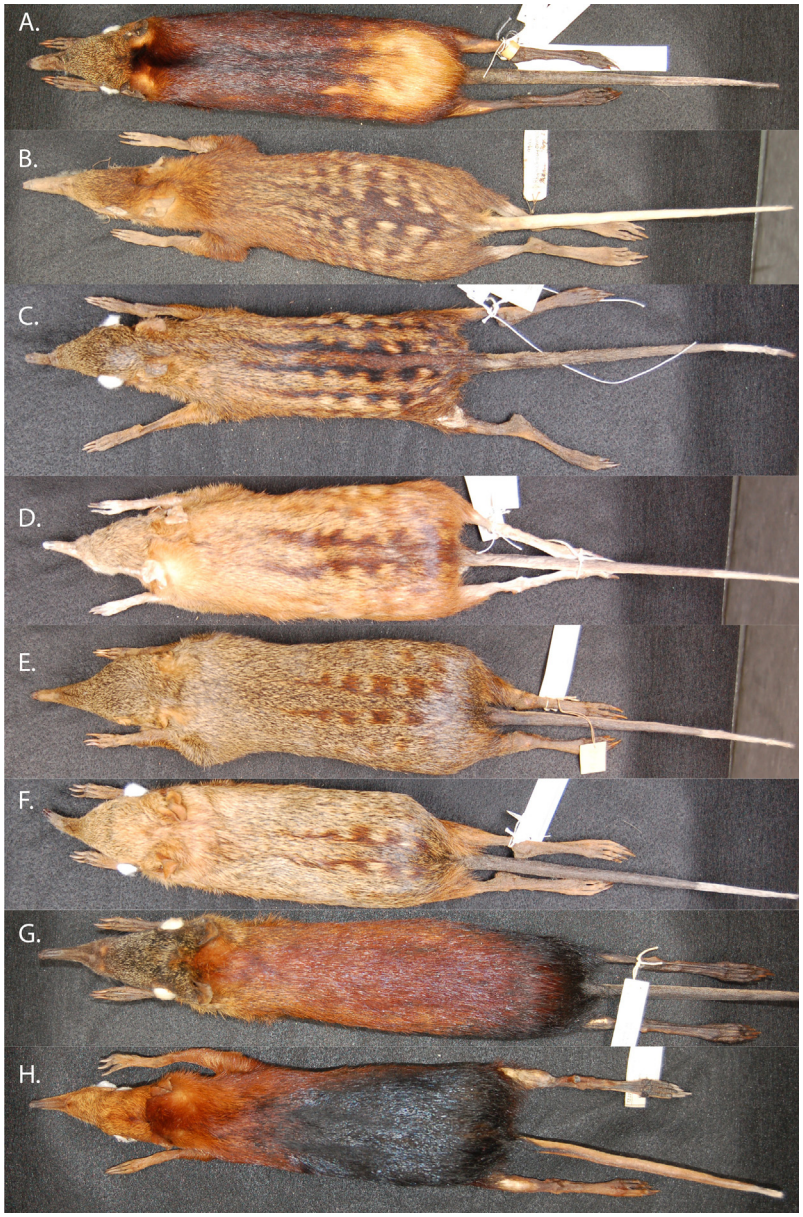

Fig. 1. Study skins showing the color patterns of *Rhynchocyon* taxa (see text for museum abbreviations associated with following catalog numbers). From top to bottom: (A) *R. chrysopygus* CAS MAM 24526; (B) *R. cirnei stuhlmanni* AMNH 49462 (light form of cline), (C) *R. cirnei reichardi* CAS MAM 28535; (D) *R. cirnei macrurus* AMNH 179301 (light form of cline), (E) *R. cirnei shirensis* AMNH 161777; (F) *R. cirnei cirnei* CAS MAM 29358; (G) *R. udzungwensis* CAS MAM 28043; (H) *R. petersi petersi* CAS MAM 30667. The Boni *Rhynchocyon* is not represented, but is superficially similar to *R. udzungwensis*, as are the clinal dark forms of *R. c. macrurus* from southeastern coastal Tanzania and *R. cirnei stuhlmanni* from western Uganda (see Corbet and Hanks, 1968).

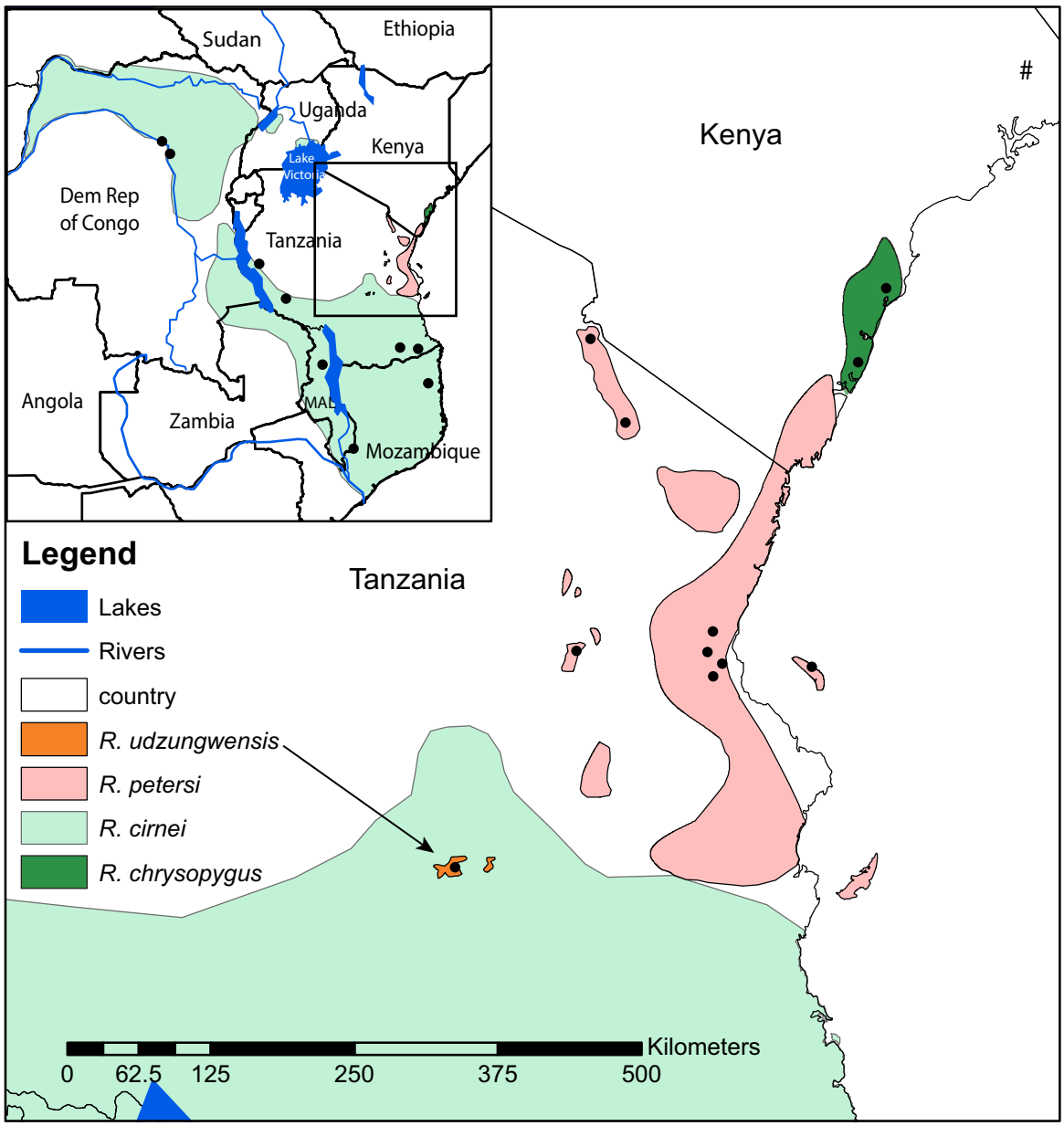

Fig. 2. Distribution of the four species of *Rhynchocyon*, with the polygon for *R. cirnei* in the Democratic Republic of Congo (Congo Basin) and western Uganda representing *R. cirnei stuhlmanni*, which we propose as a fifth species (see text for general distributions of other subspecies). The # is the location of the undescribed form of *Rhynchocyon* in northern coastal Kenya. Distribution polygons are courtesy of IUCN (www.iucnredlist.org), which are based on point data compilation by GBR (www.sengis.org/distribution/). Collection localities (see Table 1) for each taxon are denoted by a ●.

could arguably be elevated to full species based on its short nasal bones, all-white tail, and allopatric distribution in the Congo Basin, though they left this unresolved.

In 2008, Rovero et al. described a fourth species, *R. udzungwensis*, which occurs in two evergreen forests in the Udzungwa Mountains, Tanzania (Figs. 1C and 2). Andanje et al. (2010) reported a potentially new giant sengi from the Boni and Dodori national reserves on the northern coast of Kenya (Fig. 2) that most closely resembles *R. udzungwensis* in coloration, but the phylogeny and taxonomy remains to be studied.

Despite the seminal work by Corbet and Hanks (1968) and recent discoveries, *Rhynchocyon* taxonomy remains problematic. For example, within *R. cirnei*, where nearly all subspecies have distinct checkering patterns on the back (Fig. 1B-F), determining taxonomic status and relationships have been difficult. Similarly, relationships and placement of some *Rhynchocyon* taxa has been difficult, especially those taxa whose checkering is masked with dark pelage (e.g., *R. petersi*, *R. udzungwensis*, Fig. 1G,H, and the dark coastal form of *R. cirnei macrurus*).

A complete and accurate phylogeny for *Rhynchocyon* is needed for several reasons. First, *Rhynchocyon* has important conservation value. Because Rhynchocyoninae taxa are few and distantly related to other mammals (divergence approximately 42.7 ± 4.8 MYA, Douady et al., 2003), they contribute significant ecological and phylogenetic diversity to their communities (Davies and Buckley, 2011; Faith, 1992). Many of the taxa have narrow ranges (e.g. *R. udzungwensis, R. chrysopygus, R. c. shirensis*), and some that appear to have larger ranges are actually restricted to small fragmented patches of montane and coastal forest (e.g. *R. petersi*). Thus, understanding how many distinct *Rhynchocyon* taxa exist

and their relationships to each other within the genus will help determine how to manage populations when they become smaller or as some populations are inevitably extirpated. Second, there is a significant captive population of *R. petersi* living in zoos. These zoos maintain a coordinated breeding program with detailed studbooks, but there is no detailed locality information for the founders collected from eastern Africa. For these to have maximum conservation value for breeding and potential reintroduction, it is imperative to know their source subspecies and population. Finally, there are several sequences of *Rhynchocyon* in GenBank, and we have reason to question the accuracy of these DNA sequence data and metadata. In particular, sequences originally published as *Rhynchocyon* sp. (Douady et al., 2003), have been determined by Smit et al. (2011) to be *R. chrysopygus*. Based upon the likely collecting locality, we doubted this identification, and hoped to test it both genetically and with information from the voucher specimen. Based upon published ribosomal sequence data, we also doubted the accuracy of other sequences, and we hope to provide more accurate data for research. Although many sequences on GenBank have errors, these *Rhynchocyon* sequences are particularly problematic because multiple studies of evolutionary processes have used these to represent the family or subfamily (e.g. Puttick and Thomas, 2015; Smith et al., 2013, 2016), and errors could affect both inferred relationships and evolutionary timing and divergences.

Although several studies on the molecular phylogenetics of soft-furred sengis have been published (Douady et al., 2003; Dumbacher et al., 2012, 2014, 2016; Smit et al., 2007, 2008, 2011), comparatively few have been done on the giant sengis (Lawson et al., 2013; Sabuni et al., 2016; Smit, 2008). Douady et al. (2003) included a single *Rhynchocyon* specimen in their analysis of the role of the Sahara in the diversification of Macroscelidea, but neither a voucher nor the species were identified, although the collection locality was identified as southeastern Tanzania (Douady, 2001). Smit et al. (2011), in their study of the phylogenetic relationships of Macroscelididae, sequenced approximately 2000 bases of the mtDNA gene fragments 12S rRNA, valine tRNA, and 16S rRNA (12s16s) from one each of a *R. chrysopygus*, *R. c. reichardi*, and *R. p. petersi* from the Natural History Museum in London. Based on their phylogenetic analysis, Smit et al. (2011) proposed that *R. petersi* and *R. cirnei* were sister species, and *R. chrysopygus* was sister to them, and further identified the Douady et al. (2003) *Rhynchocyon* sequence as *R. chrysopygus*. However, based on the collection locality of the Douady (2001) tissue in southeastern Tanzania, Smit et al.'s (2011) identification seems unlikely. Most recently, Lawson et al. (2013), examined the interspecific relationship of *R. udzungwensis* and *R. c. reichardi* from four forest sites, including the contact zone between the two taxa, in Tanzania (Fig. 1). They analyzed three mitochondrial loci (ND2, D-loop, 12s) and two nuclear loci (ENAM, vWF) and found the individual nuclear gene trees strongly supported the monophyly of *R. udzungwensis*. However, due to the mixing of mitochondrial clades among species in their phylogeny, Lawson et al. (2013) concluded that ancient (but not current) hybridization occurred between the two taxa because of the reciprocal monophyly of the nuclear alleles and because they did not find morphologically intermediate hybrids. However, it is unclear if historical introgression is widespread among *Rhynchocyon* taxa.

The objective of our research was to generate and analyze DNA sequence data for all named *Rhynchocyon* taxa to reconstruct phylogenetic relationships within the genus. The phylogeny will allow us to determine the authenticity of GenBank *Rhynchocyon* sequences, determine the subspecies of the captive *R. petersi* population, look for evidence of widespread hybridization among *Rhynchocyon* taxa, and assess the currently accepted taxonomy of extant *Rhynchocyon*.

2. Materials and methods

2.1. Taxon sampling

We were able to obtain samples from all currently recognized *Rhynchocyon* taxa (Corbet and Hanks, 1968) except for *R. c. hendersoni* (Table 1). Fresh tissue preserved in alcohol was available with voucher specimens in the mammalogy collections at the California Academy of Sciences (CAS) and the Field Museum of Natural History (FMNH). Unvouchered fresh tissue was also collected for this project (Table 1, and Supplemental Material Table 1). We sampled dried tissue from museum study skins when fresh tissue was not available. Table 1 is a complete list of specimens sampled, including locality coordinates, and Fig. 2 shows the dispersion of collection localities. Additionally, we used GenBank sequences from Douady et al. (2003) and Smit et al. (2011).

2.2. DNA sequencing

We chose to study three independently segregating loci based on previous work done with the family Macroscelididae (Douady, 2001; Douady et al., 2003; Dumbacher et al., 2014; Lawson et al., 2013; Smit et al., 2011; Springer et al., 1997). We sequenced 2685 bases from a mitochondrial region that includes genes for 12s ribosomal RNA, tRNA-valine, and 16s ribosomal RNA (hereafter 12s16s), 976 bases of the nuclear locus inter-photoreceptor retinoid-binding protein exon 1 (IRBP), and 1043 bases of the nuclear locus von Willebrand factor exon 28 (vWF).

We extracted DNA from approximately 25 mg of tissue stored in ethanol and frozen at −80 °C using a DNeasy Blood and Tissue extraction kit (Qiagen, Valencia, CA, USA). Polymerase chain reaction (PCR) was performed on DNA extracts using multiple primer pairs (see Supplementary Material, Table 2). For DNA extractions from fresh tissue, we performed PCR in 25 μl reactions with Invitrogen Taq (Life Technologies, South San Francisco, California, USA) and typical PCR reagents and protocols optimized for each sample type (see Supplementary Material, Detailed Lab Methods).

For historical museum specimens, we sampled approximately 25 mg of dried tissue from the hind foot or the ventral incision of the dried specimen. DNA was extracted in a dedicated ancient DNA laboratory at the California Academy of Sciences (CAS) or the University of Alaska Museum (UAM). For a subset of historical samples, DNA extraction, PCR, and sequencing were performed in both laboratories, providing an independent replication for those individuals. Detailed protocols for historical DNA extraction and PCR can be found in Supplementary Materials, Detailed Lab Methods.

PCR amplicons were Sanger sequenced using BigDye Terminator version 3.1 cycle sequencing chemistry (Life Technologies, South San Francisco, California, USA). Sequences were visualized on an ABI 3130 Genetic Analyzer (Life Technologies, South San Francisco, California, USA) located at CAS's Center for Comparative Genomics or sequenced at the High Throughput Genomics Center, Seattle WA (http://www.htseq.org/).

2.3. Alignment and analysis

Because of a higher likelihood of contamination, all amplicons from historical DNA were checked for contamination using the blastn, megablast, and discontiguous megablast algorithms for the nucleotide Basic Local Alignment Search Tool (BLAST) on the National Center for Biotechnology Information (NCBI) website. Sequences were assembled, edited, and a consensus sequence for each individual was created in Geneious v7.1.4 (Kearse et al., 2012). For heterozygous sequences at nuclear loci, both alleles

Table 1
Data for specimens used for DNA sequencing. Museum numbers are given for vouchered specimens, and field numbers are given for unvouchered specimens and denoted with an asterisk (*). Museum codes are in the abbreviations section. §Denotes sequence was downloaded from GenBank. †Denotes sequence is from historical DNA. Collection locality includes the latitude and longitude in decimal degrees.

Taxon	Voucher/field number	GenBank Accession Number			Collection locality
		12s16s	IRBP	vWF	
E. edwardii	Unknown	AY310885§	AY310899§	AY310892§	South Africa
M. micus	CASMAM27997	KF895104§	KF742665§	KF742645§	Khorixas District, Kunene Region, Namibia; −20.7266, 14.1283
P. tetradactylus	Unknown	AY310883§	AY310897§	AY310890§	Chingulungulu, Tanzania
R. chrysopygus	CASMAM24525	KT348460†	KT348366† & KT358508†	KT358505† & KT358506†	Gedi National Monument, Kilifi District, Kenya; −3.3097, 40.0182
R. chrysopygus	CASMAM24526	KT348461†	None	None	Gedi National Monument, Kilifi District, Kenya; −3.3097, 40.0182
R. chrysopygus	FMNH153106	KT348462†	None	None	Mombasa, Kilifi District, Kenya; −4.05, 39.6667
R. c. cirnei	CASMAM29344	KT348463	KT348372	KT348411	Mareja Reserve, Mozambique; −12.8436, 40.1617
R. c. cirnei	CASMAM29345	KT348466	KT348405	KT348412	Mareja Reserve, Mozambique; −12.8483, 40.1649
R. c. cirnei	CASMAM29351	KT348470	KT348406 & KT348407	KT348413	Mareja Reserve, Mozambique; −12.8440, 40.1609
R. c. cirnei	CASMAM29352	KT348468	KT348375	KT348414	Mareja Reserve, Mozambique; −12.8420, 40.1637
R. c. cirnei	CASMAM29353	KT348469	KT348376	KT348415	Mareja Reserve, Mozambique; −12.8452, 40.1615
R. c. cirnei	CASMAM29355	KT348464	KT348377	KT348423	Mareja Reserve, Mozambique; −12.8420, 40.1637
R. c. cirnei	CASMAM29357	KT348465	KT348378	KT348416	Mareja Reserve, Mozambique; −12.8429, 40.1623
R. c. cirnei	CASMAM29358	KT348467	KT348379	KT348417	Mareja Reserve, Mozambique; −12.8450, 40.1614
R. c. macrurus	RMCA 96.037-M-5388 or RMCA 96.037-M-5390	AY310880§	AY310894§	AY310887§	Chingulungulu region, Tanzania; −10.44, 38.33
R. c. macrurus	FMNH88204	KT348471†	None	None	Mihuru, Newala District, Mtwara Region, Tanzania; −10.6667, 39.5
R. c. reichardi	FMNH171474	KT348474	KT348400	KT348452	Mbizi Mts, Mbizi Forest Reserve, vicinity of Mazumba Hill, Sumbawanga District, Rukwa Region, Tanzania
R. c. reichardi	FMNH171617	KT348475	KT348380	KT348448 & KT348451	Mbizi Mts, vicinity of Mazumba, Sumbawanga District, Rukwa Region, Tanzania
R. c. reichardi	FMNH177823	KT348476	None	KT348449	Mahale Mts, Mahale National Park, 0.5 km NW Nkungwe Hill summit, Kigoma District, Kigoma Region, Tanzania; −6.1043, 29.7790
R. c. reichardi	FMNH178010	KT348477	KT348381	KT348450	Mahale Mts, Mahale National Park, 0.5 km NW Nkungwe Hill summit, Kigoma District, Kigoma Region, Tanzania; −6.1043, 29.7790
R. c. reichardi (labeled *R. c. hendersoni*)	MCZ43732	KT348473†	KT348404†	KT348447†	Vipya Plateau, Malawi
R. c. shirensis	AMNH161777	KT348472†	None	None	Mlanje Plateau, Malawi
R. c. stuhlmanni	M300*	KT348478	KT348409	KT348453	Democratic Republic of the Congo; 0.0131, 25.5565
R. c. stuhlmanni	MK001*	None	KT348409	KT348454	Democratic Republic of the Congo; 0.2946, 25.2917
R. petersi spp.	CASMAM28767	KT348479	KT348382	KT348424	Houston Zoo, Houston, Texas, United States of America
R. petersi spp.	CASMAM29516	KT348480	KT348383	KT348425	Houston Zoo, Houston, Texas, United States of America
R. p. adersi	MCZ22829	KT348481†	None	None	Nyanga Id., Zanzibar, Tanzania
R. p. petersi	FMNH151213	KT348482	KT348384	KT348418	South Pare Mts, Chome Forest Reserve, 5.5 km S Bombo, near Kanza Village, Kilimanjaro Region, Tanzania; −4.32, 38
R. p. petersi	FMNH151214	KT348483	KT348401	KT348419	South Pare Mts, Chome Forest Reserve, 7 km S Bombo, Kilimanjaro Region, Tanzania; −4.33, 38
R. p. petersi	FMNH161311	KT348485	KT348385	KT348420	Nguru Mts, Manyangu Forest Reserve, near Disango, Morogoro District, Morogoro Region, Tanzania; −6.04, 37.5467
R. p. petersi	FMNH161312	KT348486	KT348386	KT348427	Nguru Mts, Manyangu Forest Reserve, near Disango, Morogoro District, Morogoro Region, Tanzania; −6.04, 37.5467
R. p. petersi	FMNH192684	KT348484	KT348402	KT348422	North Pare Mts, Minja Forest Reserve, Mwanga District, Kilimanjaro Region, Tanzania; −3.5815, 37.6773
R. p. petersi	FNMH161395	KT348501	KT348373 & KT358507	KT348421	Nguru Mts, Manyangu Forest Reserve, near Disango, Morogoro District, Morogoro Region, Tanzania; −6.04, 37.5467
R. p. petersi	RP15*	None	KT348387	KT348428	Zaraninge Forest, Tanzania; −6.1367, 38.6055
R. p. petersi	TA1812*	None	KT348374	None	Zaraninge Forest, Tanzania; −6.1055, 38.6158
R. p. petersi	TA1818*	KT348494	KT348408	KT348436	Askari Forest, Tanzania; −5.9955, 38.7607
R. p. petersi	TA1833*	KT348487	KT348388	KT348429	Zaraninge Forest, Tanzania; −6.1056, 38.6167
R. p. petersi	TA1835*	KT348495	KT348389	KT348430	Zaraninge Forest, Tanzania; −6.1126, 38.6211
R. p. petersi	TZ22766*	KT348488	KT348390	KT348431	Gendagenda Forest, Tanzania; −5.5759, 38.6423
R. p. petersi	TZ22767*	KT348499	KT348391	KT348437 & KT348442	Gendagenda Forest, Tanzania; −5.5639, 38.6502
R. p. petersi	TZ22769*	KT348498	KT348392	KT348426 & KT348443	Gendagenda Forest, Tanzania; −5.5871, 38.6395
R. p. petersi	TZ22770*	KT348500	KT348393	KT348432	Gendagenda Forest, Tanzania; −5.5871, 38.6404
R. p. petersi	TZ22774*	KT348491	KT348394	KT348433	Kwamsisi Forest, Tanzania; −5.8909, 38.5928
R. p. petersi	TZ22775*	KT348489	KT348403	KT348434	Kwamsisi Forest, Tanzania; −5.8921, 38.5938
R. p. petersi	TZ22776*	KT348492	KT348395	KT348438 & KT348444	Kwamsisi Forest, Tanzania; −5.8921, 38.5939
R. p. petersi	TZ22778*	KT348493	KT348396	KT348439	Kwamsisi Forest, Tanzania; −5.8938, 38.5949
R. p. petersi	TZ22779*	KT348496	KT348397	KT348440 & KT348445	Kwamsisi Forest, Tanzania; −5.8937, 38.5944
R. p. petersi	TZ22783*	KT348490	KT348398	KT348435	Gendagenda Forest, Tanzania; −5.601, 38.6468
R. p. petersi	TZ22811*	KT348497	KT348399	KT348441 & KT348446	Kwamsisi Forest, Tanzania; −5.8723, 38.5726
R. udzungwensis	CASMAM28043	KT348503	KT348368	KT348455	Udzungwa Mountains, Ndundulu Forest, Tanzania; −7.8045, 36.5059
R. udzungwensis	CASMAM28318	KT348504	KT348369	KT348456	Udzungwa Mountains, Ndundulu Forest, Tanzania; −7.7944, 36.4919

Table 1 (*continued*)

Taxon	Voucher/field number	GenBank Accession Number			Collection locality
		12s16s	IRBP	vWF	
R. udzungwensis	FMNH194127	KT348506	KT348370	KT348457	Udzungwa Mountains, Ndundulu Forest, Tanzania; −7.8045, 36.5059
R. udzungwensis	MTSN6000	KT348505	KT348371	KT348458	Udzungwa Mountains, Ndundulu Forest, Tanzania; −7.8036, 36.5059
R. udzungwensis	BMNH2007.7	KT000011	KT000020	KF202173	Udzungwa Mountains, Ndundulu Forest, Tanzania; −7.8045, 36.5059

were manually phased and given unique names (e.g. allele 1, allele 2). Because so few SNPs were present in nuclear loci and all heterozygotes were restricted to a single change, phasing was done manually. Geneious created alignments using the MAFFT v7.017 alignment plugin (Katoh et al., 2002) for all of our assembled sequences, *Rhynchocyon* sequences downloaded from GenBank, and outgroup sequences. Alignments were checked by eye and exported for analysis. Duplicate haplotypes or allele sequences from multiple individuals were identified and eliminated using FaBox DNAcollapser v1.41 (Villesen, 2007). Aligned matrices for each locus are available as Nexus files and published alongside this manuscript as supplementary materials.

Each of the three independently segregating loci (12s16s, IRBP, vWF) were analyzed independently and as concatenated datasets. Because we were trying to assess the relationship of close relatives and the potential for gene flow, we analyzed each locus separately to specifically look for evidence of introgression and conflicting signal, which may be ignored in analyses of concatenated matrices.

Phylogenetic analyses were performed using both maximum likelihood and Bayesian approaches. First, Nexus files were imported into PAUP* v4.0b10 (Swofford, 2003) and sequences were partitioned into transfer RNAs and ribosomal RNAs for the mitochondrial region, and into codon positions for the nuclear loci. We used MrModelTest v2.3 (Nylander, 2004) and the Akaike Infor-

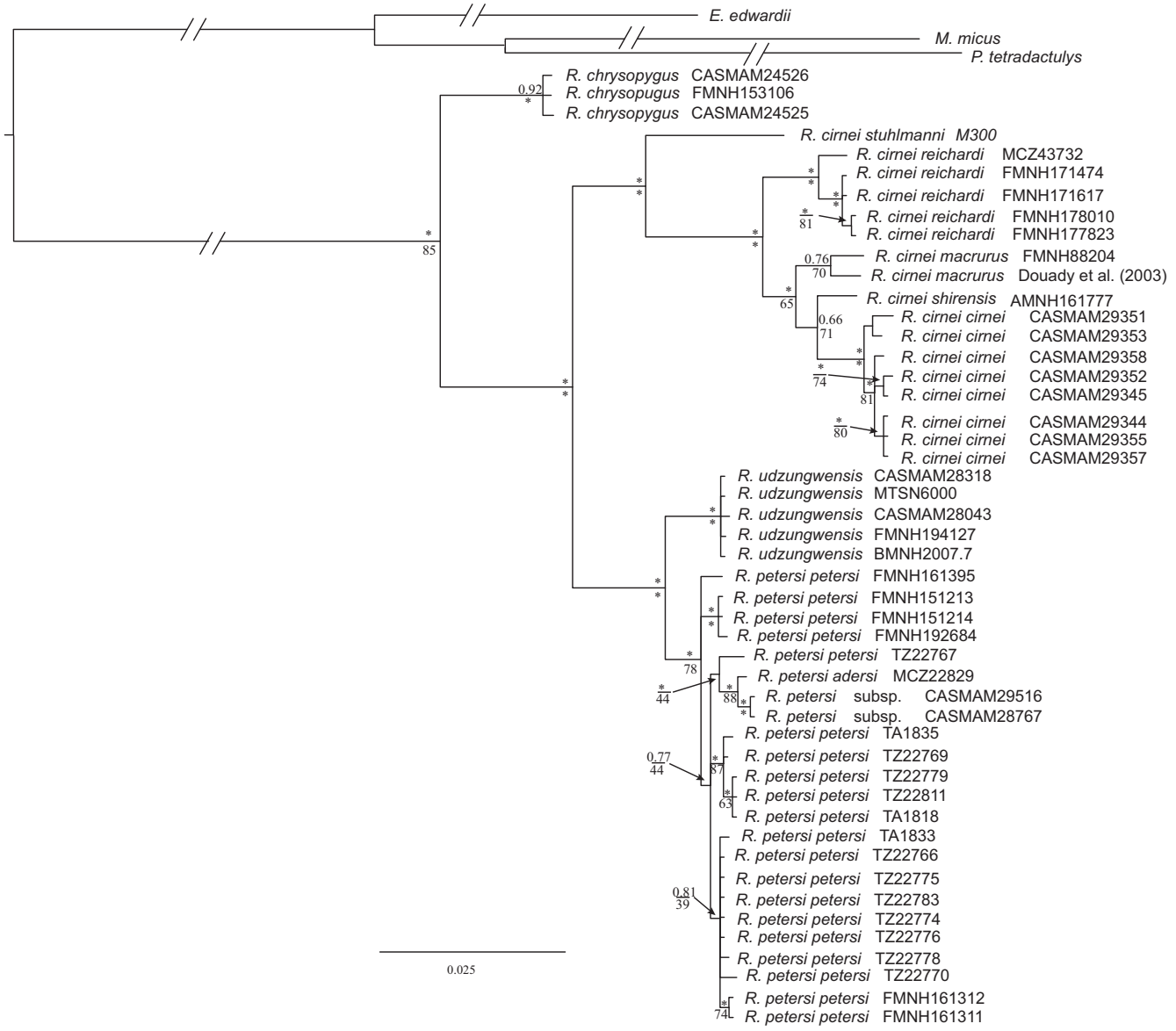

Fig. 3. MrBayes phylogram for *Rhynchocyon* 12s16s mitochondrial region. Bayesian posterior probabilities equal to or above 0.95 were considered significant and are represented by an asterisk (*) above the branch. Maximum likelihood bootstraps equal to or above 90 were considered significant and are represented by an asterisk below the branch.

Table 2
Distance matrix for 12s16s mitochondrial sequences showing uncorrected percent differences (p-distance). The mean genetic distance and standard deviation is shown in the upper right triangle, and range is shown in the lower left triangle.

	E. edwardii	*M. micus*	*P. tetradactylus*	*R. chrysopygus*	*R. cirnei sthulmanni*	*R. cirnei reichardi*	*R. cirnei macrurus*	*R. cirnei shriensis*	*R. cirnei cirnei*	*R. udzungwensis*	*R. petersi adersi*	*R. petersi* sp.	*R. petersi*
E. edwardii		18.10	18.50	19.23 ± 0.37	21.10	21.24 ± 1.03	20.45 ± 1.77	24.50	21.80 ± 0.00	21.50 ± 0.14	19.30	21.40 ± 0.00	21.41 ± 0.21
M. micus	18.10		17.70	20.77 ± 0.06	21.50	21.64 ± 0.25	21.30 ± 0.57	25.60	21.90 ± 0.05	22.18 ± 0.27	20.80	21.70 ± 0.00	21.75 ± 0.18
P. tetradactylus	18.50	17.70		22.53 ± 0.12	22.10	22.78 ± 0.16	22.60 ± 0.14	27.50	22.94 ± 0.05	23.08 ± 0.11	22.80	22.60 ± 0.00	22.68 ± 0.19
R. chrysopygus	19.50–18.80	20.80–20.70	22.60–22.40		3.13 ± 0.15	3.01 ± 0.08	2.85 ± 0.10	8.35 ± 0.07	2.78 ± 0.07	3.00 ± 0.09	2.75 ± 0.07	2.95 ± 0.06	2.72 ± 0.10
R. cirnei sthulmanni	21.10	21.50	22.10	3.00–3.10		2.82 ± 0.18	2.50 ± 0.42	7.80	3.11 ± 0.04	3.30 ± 0.12	3	3.4 ± 0.00	3.24 ± 0.07
R. cirnei reichardi	21.70–19.40	21.80–21.20	22.90–22.50	3.00–2.90	2.90–2.50		1.51 ± 0.38	6.76 ± 0.13	1.58 ± 0.07	3.50 ± 0.43	2.54 ± 0.09	3.46 ± 0.30	3.37 ± 0.36
R. cirnei macrurus	21.70–19.20	21.70–20.90	22.70–22.50	2.90–2.70	2.80–2.20	2.00–1.20		6.05 ± 0.07	1.04 ± 0.47	3.19 ± 0.78	2.20 ± 0.14	3.10 ± 0.81	2.88 ± 0.72
R. cirnei shriensis	24.50	25.60	27.50	8.30–8.40	7.8	7.00–6.70	6.10–6.00		6.09 ± 0/08	8.6 ± 1.57	7.8	7.90 ± 0.00	7.74 ± 0.07
R. cirnei cirnei	21.8	22.00–21.80	23.00–22.90	2.90–2.70	3.20–3.10	1.70–1.40	1.50–0.50	6.20–6.00		3.69 ± 0.15	2.29 ± 0.08	3.53 ± 0.07	3.37 ± 0.08
R. udzungwensis	21.70–21.40	22.60–22.00	23.20–23.00	3.10–2.90	3.50–3.20	4.00–2.60	4.20–2.40	11.40–7.80	4.00–3.60		1.04 ± 0.05	1.24 ± 0.05	0.98 ± 0.08
R. petersi adersi	19.30	20.80	22.80	2.80–2.70	3.00	2.70–2.50	2.30–2.10	7.80	2.40–2.20	1.10–1.00		0.10 ± 0.00	0.22 ± 0.11
R. petersi sp.	21.4	21.7	22.6	3.00–2.90	3.4	3.60–2.90	3.80–2.40	7.9	3.70–3.50	1.30–1.20	0.10		0.42 ± 0.07
R. petersi petersi	21.60–21.00	21.90–21.40	22.80–22.30	3.10–2.50	3.30–3.10	3.60–2.50	3.70–2.10	7.90–7.70	3.60–3.30	1.10–0.80	0.50–0.10	0.50–0.30	

mation Criterion (Akaike, 1974) to assess the rate-specific model of evolution for each partition (Supplementary Materials, Table 3). We performed Bayesian analysis using MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003). Bayesian analysis was run for 10 million Markov Chain Monte Carlo (MCMC) generations, sampling trees and parameters every 1000th generation. The first 25% of the generations sampled were discarded as burnin. We performed maximum-likelihood analysis using Random Axelerated Maximum Likelihood (RAxML) v7.2.6 using the GTR + Γ model (Stamatakis, 2006). The "autoMRE" command, which is a bootstrap convergence test, was used to determine when a sufficient number of bootstrap replicates had been reached (Pattengale et al., 2010).

To test the robustness of the results and the impact of missing data, we removed any individuals with over 50% missing data from the matrix, and repeated the analysis on the reduced matrix. All analyses were performed on the Phylocluster computer at CAS. Support for each node was estimated using Bayesian posterior probabilities in MrBayes (Ronquist and Huelsenbeck, 2003) and bootstrap analysis in RAxML (Stamatakis, 2006). We created a genetic distance matrix for the 12s16s region in Geneious v7.1.4 (Kearse et al., 2012) by subtracting the percent identity provided in the multiple alignment from 100 and averaging across individuals for the same taxon pairs. Nuclear loci were visualized as unrooted TCS allele networks (Clement et al., 2000) using PopART v1 (Leigh et al., 2013).

To assess the support for species delimitation, we used the multispecies coalescent model in the Bayesian reverse-jump MCMC program BPP, version 3.3 (Yang, 2015). We assessed the phylogenetic support for the four recognized *Rhynchocyon* species (*R. chrysopygus, R. cirnei, R. petersi*, and *R. udzungwensis*), but we additionally wanted to assess support for *R. c. stuhlmanni* as a distinct species. To do this, we utilized the A11 model to perform joint species delimitation and species tree estimation, as unguided delimitation is preferred over delimitation using a fixed guide tree and can sometimes show strong species support even when multi-locus phylogenies are uncertain (Yang, 2015). Control files containing the important parameter settings and Bayesian priors are provided in Supplementary Material file: BPP Analyses.

3. Results

Final aligned sequence length for 12s16s equaled 2685 nucleotide base pairs representing 48 specimens across 10 giant sengi and 3 outgroup taxa. Final aligned sequence length for IRBP equaled 976 base pairs representing 45 specimens across 8 giant sengi taxa. Final aligned sequence length for vWF equaled 1043 base pairs representing 45 specimens across 8 giant sengi taxa. Specimens with missing data were mostly from degraded museum skin samples from which our genes were difficult to amplify or sequence (see Table 1). Bayesian and maximum likelihood analyses of the mitochondrial locus 12s16s (Fig. 3) recovered similar trees with consistent support for nodes. We considered node support as significant if the Bayesian posterior probability (Bpp) was equal to or greater than 0.95 and the maximum likelihood bootstrap (mlb) support was equal to or greater than 90. In our 12s16s tree (Fig. 3), there was good phylogenetic resolution at the species level and even at the subspecies level for almost every taxon.

The average pairwise genetic distance matrix for the 12s16s region shows percent differences between 0.1% and 8.8% ($\bar{x}$ = 2.4 ± 1.8) among *Rhynchocyon* taxa (Table 2). The greatest differences are between *R. c. shirensis* and other taxa. However, we are skeptical of these values because *R. c. shirensis* has approximately 1400 fewer bases than other taxa due to poorly preserved DNA, and missing data may contribute to unusual estimates. Excluding

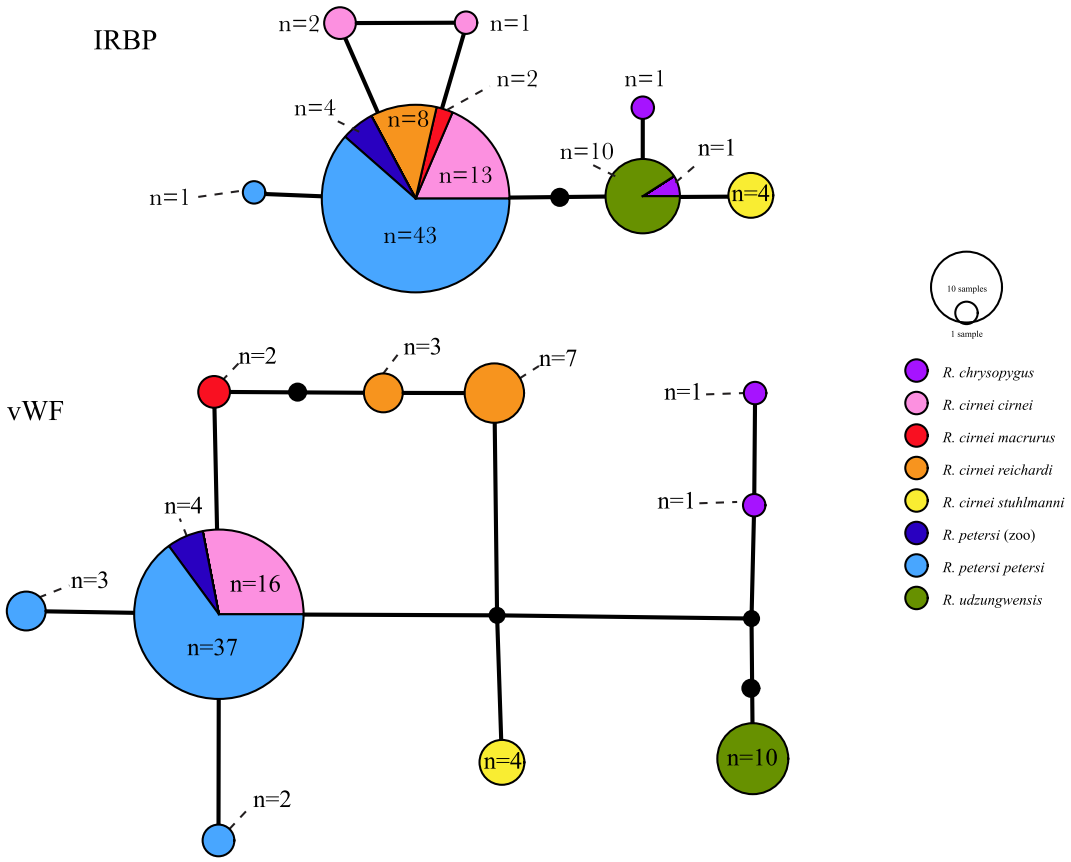

Fig. 4. TCS allelic networks for *Rhynchocyon* nuclear loci IRBP and vWF. The size of the circle is proportional to the number of alleles sampled of that haplotype. Black circles represent inferred alleles that were not recovered in this analysis and 'n' represents the number of alleles recovered for each haplotype.

R. c. shirensis, the remaining *Rhynchocyon* taxa have average pairwise genetic distances ranging from 0.2% to 3.7% ($\bar{x}$ = 2.0 ± 1.4).

Nuclear loci IRBP and vWF were chosen based on previous work with Macroscelididae (Douady et al., 2003; Dumbacher et al., 2014; Lawson et al., 2013; Smit et al., 2011; Springer et al., 1997). However, these loci exhibited low variation among the samples analyzed and we recovered only seven single nucleotide polymorphisms (SNPs) across the 976 IRBP bases and eleven SNPs across the 1043 vWF bases. Thus, we chose to visualize nuclear loci IRBP and vWF as individual allelic networks (Fig. 4).

Within the IRBP allele network, *R. petersi* alleles cluster together with all but one allele from *R. cirnei* subspecies (n = 74 sampled alleles). Most individuals of these taxa are represented by a single shared allele (n = 70 sampled alleles) and only three other alleles that are one nucleotide substitution different (n = 4). *R. c. stuhlmanni*, however, is separated from the remaining *R. cirnei* subspecies by three nucleotide changes, and is the most distant IRBP allele from other *R. cirnei* found in the genus. Additionally, *R. chrysopygus* and *R. udzungwensis* share an allele which differs by one nucleotide change from the *R. c. stuhlmanni* allele and the unique *R. chrysopygus* allele.

The vWF locus contains more genetic variation and taxonomic structure. In the vWF network (Fig. 4), only one allele is present in multiple taxa. This allele is the most common allele overall (n = 57 alleles), and it is shared by *R. c. cirnei*, *R. p. petersi*, and the two captive *R. petersi* from the Houston Zoo (CAS MAM 28767 and CAS MAM 29516). Three additional alleles from *R. c. macrurus* and *R. petersi* diverged from this most common allele by only one nucleotide change. The four remaining taxa, (*R. c. reichardi, R. c. stuhlmanni*, *R. chrysopygus*, and *R. udzungwensis*), each appear relatively distinct; each taxon has unique alleles that are at least two nucleotide changes to any allele belonging to another taxon. Allele networks for both nuclear loci clearly show the distinctness of *R. c. stuhlmanni*, *R. chrysopygus*, and *R. udzungwensis*.

The results of the mitochondrial and nuclear analyses were mostly compatible. Discordance among the analyses came from the close affinity of *R. cirnei* and *R. petersi* in the nuclear genome and shared alleles in both IRBP and vWF (Fig. 4), and yet these two taxa were distant clades in the mitochondrial analysis of 12s16s (Fig. 3). In the mitochondrial analysis, *R. p. petersi* was more closely related to *R. udzungwensis* than any other taxon. All three loci supported the phylogenetic distinctness of *R. c. stuhlmanni*.

Our analysis showed that the Boni tissue sample was 100% identical to *R. chrysopygus* for the mitochondrial locus and vWF. For IRBP, the Boni individual was heterozygous with one allele matching a *R. chrysopygus* allele, while the other allele was new to our analysis and one nucleotide change different from the allele shared by *R. udzungwensis* and *R. chrysopygus*. Thus, the tissue we sequenced and analyzed from the Boni population is indistinguishable from *R. chrysopygus*. We have not included these data in the figures.

BPP analyses using the entire dataset suffered from mixing problems such that initial conditions affected the outcome. Specifically, searches starting at or near the one species model (0000) tended to get stuck in that model, whereas models that initially included multiple species progressed toward strong support for the five species model (1111). Although the manual suggested some potential causes (e.g. too many loci, inappropriate priors), we explored these options without improving the mixing. We additionally explored the role of missing data, and by eliminating individuals with any locus completely missing, we found that mixing and rjMCMC behavior improved. This dataset included 44

individuals (*R. chrysopygus*, n = 1; *R. cirnei*, n = 15, *R. c. stuhlmanni*, n = 1; *R. petersi*, n = 23; and *R. udzungwensis*, n = 4), and was run five times to ensure consistent outcomes. The results consistently supported delimiting all five species (including elevating *R. c. stuhmanni* to full species) with posterior probabilities >0.988 per run (n = 5 runs).

4. Discussion

With nearly complete taxon sampling (we lack *R. c. hendersoni*), our analysis confirms that earlier taxonomists (Corbet and Hanks, 1968; Rovero et al., 2008) accurately inferred taxonomic groups within *Rhynchocyon* using pelage color patterns and geographic range. Below we discuss our findings in relation to previous work on *Rhynchocyon* and include a revised taxonomy.

4.1. Clarifying the taxonomic status of ambiguous GenBank sequences

Douady et al. (2003) posted three sequences for *Rhynchocyon* sp. on GenBank: 12s16s (GenBank accession number AY310880), IRBP (AY310894), and vWF (AY310887). All three sequences cluster with *R. cirnei* in our analyses, in contrast to Smit et al. (2011) that suggested these are *R. chrysopygus*. Thus, we wanted to check for other data confirming the taxonomic identity of this specimen. GenBank entries share a single extraction number (CJD-2003). In his dissertation, Douady (2001) lists two tissues as the source of genetic data for his *Rhynchocyon* sp., (tissue numbers T-1853, T-1854), from the personal collection of François Catzeflis at the Université Montpellier, France, and provides the collection locality for both tissues as Chingulungulu (Tanzania). Douady (2001) does not report which tissue was sequenced and posted on GenBank, but for our purposes it makes little difference because both voucher specimens (collected by Herwig Leirs and Walter Verheyen, Royal Museum of Central Africa, Tervuren, Belgium, catalog numbers 96.037-M-5388 and 96.037- M-5390) came from the same locality (10°44′S, 38°33′E), and a visual inspection of images of the two specimens indicates they were the same taxon (F. Catezflis and H. Leirs, pers. comm.). Based on pelage color and pattern, distribution, and our sequence data, we conclude that the GenBank sequence is from *R. c. macrurus* and not *R. chrysopygus* as proposed by Smit et al. (2011).

4.2. Origins of captive populations

It is not known where the founders of the captive population of *R. petersi* at zoos in the United States (Baker et al., 2005) came from (unpublished correspondence, K. Lengel, P. Riger, and S. Eller). Based on pelage coloration, the sengis are obviously *R. petersi*, but the two subspecies have different geographic distributions, with *R. p. petersi* from mainland Tanzania and southeastern Kenya, and *R. p. adersi* from the islands of Mafia and Zanzibar off the coast of Tanzania. For captive breeding purposes, and if reintroductions should be contemplated in the future, it would be important to know the provenance of the captive population to maintain the genetic integrity of both wild and captive populations. We analyzed the DNA of two captive individuals from the Houston Zoo (Table 1), and these clustered with *R. p. adersi* in the 12s16s mitochondrial phylogram (Fig. 3). We were unable to confirm the clustering at the nuclear loci because we were unable to sequence nuclear DNA from any confirmed *R. p. adersi*. However, our analysis suggests that the zoo specimens were originally taken from *R. p. adersi* exported from Zanzibar or Mafia islands. Because we had only a single *R. p. adersi* sample from the wild population, and because it genetically clustered well within the available variation of the *R. p. petersi* clade, we regard these results as preliminary.

4.3. Hybridization

Lawson et al. (2013) presented data consistent with ancient introgression between *R. c. reichardi* and *R. udzungwensis*, where the distribution of the two taxa meet in the Udzungwa Mountains of Tanzania, calling into question the genetic boundaries of these two taxa. We found no evidence of introgression between any *Rhynchocyon* species, but none of our samples were from adjoining populations like those of Lawson et al. (2013). The differences between our two studies are likely explained by the differences in geographical sampling and perhaps the depth of sampling. Lawson et al. (2013) sampled extensively across a narrow range, targeting the contact zone of *R. c. reichardi* and *R. udzungwensis*. We sampled shallowly across a broad range, mostly away from contact zones. Therefore, if introgression occurs at contact zones, we were less likely to detect it. Indeed, our data suggest that widespread gene flow and panmixia does not occur in *Rhynchocyon*.

Furthermore, our 12s16s phylogeny shows that *R. c. reichardi* and *R. udzungwensis* are not sister taxa, with *R. udzungwensis* being more closely related to *R. petersi* than the *R. cirnei* clade (Fig. 3). Observations of hybridization between non-sister species has been documented in other groups (Dasmahapatra et al., 2007; Good et al., 2003; Larsen et al., 2010; McKitrick and Zink, 1988). Hybrids sometime occur when non-sister species have overlapping ranges or historical contact zones, such as the contact zone between *R. udzungwensis* and *R. c. reichardi*. Thus, studies looking for evidence of *Rhynchocyon* introgression should sample in areas where historical contact between species may have occurred, although these may be exceedingly difficult to find given the disappearance of forest habitats in Africa.

4.4. Current taxonomic status of Rhynchocyon taxa

In their revision of Macroscelididae, Corbet and Hanks (1968) described a new subspecies of *R. cirnei*, *R. c. shirensis,* from the Shire Valley of Malawi, based on a distinct pelage pattern. However, Coals and Rathbun (2013) examined additional museum specimens and observed that the pelage of the Malawi subspecies appeared to be within the variation seen in *R. c. cirnei* specimens from Mozambique. We have only one sample of *R. c. shirensis* in our mitochondrial analysis and no samples in our nuclear analyses; nonetheless, in the 12s16s phylogeny, *R. c. shirensis* falls within the *R. cirnei* clade (Fig. 3) and does not cause any taxa to be paraphyletic, therefore, we recommend continuing to treat *R. c. shirensis* as a subspecies of *R. cirnei* pending additional sampling and analyses.

In Andanje et al. (2010) suggested a potentially new species of *Rhynchocyon* from the Dodori and Boni national reserves on the northern coast of Kenya. A voucher specimen was collected and placed at the National Museums of Kenya (NMK169427), and tissue from this voucher was sent to CAS by the Kenya Wildlife Service. The sequences that we obtained from the tissue were identical to *R. chrysopygus* at 12s16s, vWF, and one of two alleles at IRBP. These data suggest that the specimen we sequenced is genetically very similar to, or perhaps a form of, *R. chrysopygus*. This is surprising given the very different pelage color and patterns between these two allopatric forms (Andanje et al., 2010). Moreover, our work suggests that dorsal pelage pattern and coloration are useful taxonomic characters for other *Rhynchocyon* taxa. Because our results are based upon a single tissue specimen, we are reluctant to draw any conclusions regarding this specimen and the sequences without examining the voucher. More data should be collected and analyzed before any conclusions can be made about the taxonomic status of this morphologically unique giant sengi.

Along with the six *R. cirnei* subspecies, Corbet and Hanks (1968) suggested a potential seventh subspecies based on a single speci-

men collected in northeastern Mozambique, but noted that this specimen might be an intermediate between *R. c. cirnei* and *R. c. macrurus* based on tail coloration. To investigate this potential subspecies, Coals and Rathbun (2013) collected eight *Rhynchocyon* specimens from northeastern Mozambique and compared the pelage of their specimens with several *R. c. cirnei* individuals, including two topotypes. They concluded that the variation in pelage color and pattern within the distribution of *R. c. cirnei* does not justify designation of any new subspecies, and additionally questioned the validity of *R. c. shirensis* pending genetic analyses. Although our genetic analysis of *R. c. shirensis* suggests it is distinct from *R. c. cirnei*, all of our *R. c. cirnei* samples are from northern Mozambique, and we do not have tissue from *R. c. cirnei* topotypes from southern Mozambique. Because we had only one *R. c. shirensis* specimen in our analysis, and no *R. c. cirnei* from nearby southern Mozambique where the type specimens originated, we are unable to assess Corbet and Hanks's (1968) question as to whether the northern Mozambique form of *R. c. cirnei* might be genetically distinct from the southern form from southern Mozambique and Malawi.

Our molecular data suggest that *R. c. stuhlmanni* could be returned to full species, as provisionally proposed by Corbet and Hanks (1968) based upon short nasals and disjunct range. The 12s16s phylogeny (Fig. 3) shows strong support for *R. c. stuhlmanni* as a distinct lineage that is sister to all other *R. cirnei* subspecies. The mean distance matrix for 12s16s (Table 2) shows *R. c. stuhlmanni* as at least 2% divergent from other *R. cirnei*, while the remaining *R. cirnei* subspecies show among subspecies divergences between 1% and 1.6%. Moreover, the nuclear allele networks (Fig. 3) show additional support for the uniqueness of *R. c. stuhlmanni* and support elevating it to full species. In both the IRBP and vWF allele networks, *R. c. stuhlmanni* has a unique allele that is not shared by any other taxa. Furthermore, the *R. c. stuhlmanni* allele in the IRBP network is three steps away from the other *R. cirnei* subspecies, and closer to an allele shared by *R. chrysopygus* and *R. udzungwensis*. Results of BPP analyses additionally corroborate delimiting *R. c. stuhlmanni* as a distinct taxon. Thus *R. c. stuhlmanni* is morphologically, geographically, and genetically distinct from other *R. cirnei*. However, the cline described by Corbet and Hanks (1968) needs genetic examination, as does the cline they describe in southeastern Tanzania for *R. c. macrurus*.

The phylogenetic data and taxonomic revision that we present here will facilitate a future detailed treatment of *Rhynchocyon* phylogeography. For example, it has been proposed that large rivers and their flood plains, as well as lowland ground-water forests, are important limiting factors in the historical and current distribution of *Rhynchocyon* (Corbet and Hanks, 1968; Andanje et al., 2010; Coals and Rathbun, 2013; Rathbun, 2009). It is also possible that the Rift Valley lakes and highlands were prehistorically important vicariant factors, although the current distribution of *Rhynchocyon* taxa suggests that neither the lakes nor elevation completely account for current distributions (Fig. 2, www.sengis.org/distribution). An analysis of the phylogeography of *Rhynchocyon* will need to include a careful assessment of the diversification of other faunal groups in Africa and the likely environmental factors involved, such as climate shifts, tectonics, forest fragmentation, river meanderings, and hydrological basin shifts (Kingdon, 1989; Stanley et al., 2005; deMenocal, 2004; Lawson, 2010; Dimitrov et al., 2012; Taylor et al., 2009; Fjeldsa and Bowie, 2008).

5. Conclusions

Based on our genetic analysis we recommend the following taxonomic treatment for giant sengis:

Family: Macroscelididae Bonaparte, 1838
  Subfamily: Rhynchocyoninae
  Genus: *Rhynchocyon* Peters, 1847
    Rhynchocyon cirnei Peters, 1847
      Rhynchocyon cirnei cirnei Peters, 1847
      Rhynchocyon cirnei shirensis Corbet and Hanks, 1968
      Rhynchocyon cirnei reichardi Reichenow, 1886
      Rhynchocyon cirnei hendersoni Thomas, 1902
      Rhynchocyon cirnei macrurus Günther, 1881
    Rhynchocyon stuhlmanni Matschie, 1893
    Rhynchocyon petersi Bocage, 1880
      Rhynchocyon petersi petersi Bocage, 1880
      Rhynchocyon petersi adersi Dollman, 1912
    Rhynchocyon chrysopygus Günther, 1881
    Rhynchocyon udzungwensis Rathbun & Rovero, 2008

Several important taxonomic issues remain to be resolved that will require further research using molecular genetics in conjunction with morphology and distribution data. These issues include 1. whether the *Rhynchocyon* from the northern coastal area of Kenya represents a new species, 2. whether *R. c. hendersoni* is a valid subspecies rather than only a relatively minor geographic (high elevation) variant of *R. c. reichardi*, 3. whether *R. c. reichardi* should be returned to full species status, 4. whether *R. c. shirensis* represents a minor variant within *R. cirnei* and thus should not be a subspecies, and 5. the genetic nature of the geographic variation in pelage pattern in southeastern Tanzania (*R. c. macrurus*), the Congo Basin (*R. stuhlmanni*), and Mozambique and southern Malawi (*R. c. cirnei* and *shirensis*). In any case, the prediction that continued revisions of the taxonomy of Corbet and Hanks (1968) would result in greater recognized sengi diversity (Rathbun, 2009) is being born out with the recent revision of the genus *Macroscelides* to include three species (Dumbacher et al., 2014), and the creation of a new genus for the North African sengi, *Petrosaltator rozeti* (Dumbacher et al., 2016), and in this paper with the resurrection of *R. stuhlmanni*.

Acknowledgements

Our research was financially supported by a National Science Foundation grant to LEO (DEB-1120904). Additionally, we thank The California Academy of Sciences, the Biology Department and Graduate Student Council in Biology at San Francisco State University, the Society for the Study of Evolution, and the Society for Integrative and Comparative Biology for financial support. We thank B. R. Agwanda of the National Museums of Kenya who captured and prepared the Boni *Rhynchocyon,* S. Andanje of the Kenya Wildlife Service who personally imported *Rhynchocyon* tissue into the United States of America from the wildlife service's collection, and S. Musila of the National Museums of Kenya who encouraged us to include the Boni *Rhynchocyon* in our analyses. Thanks to K. Consolate, who provided us with tissues of *R. c. stuhlmanni*. We are grateful for specimens and photographs provided by N. Duncan at the American Museum of Natural History, and specimens provided by J. Chupasko at the Harvard Museum of Comparative Zoology. We thank K. Hildebrandt, O. Carmi, and A. Sellas who assisted with laboratory work. Thanks to H. A. Smit-Robinson who shared advice and details of her previously published work on *Rhynchocyon*. F. Catzeflis, L. Herwig, and W. Wendelen helped determine the provenance of the Douady et al. (2003) specimen, including providing photographs of specimens in the collection of the Royal Museum of Central Africa. Thanks to K. Lengel and S. Eller, at the Philadelphia Zoo, and P. Riger, at the Houston Zoo who provided information about *Rhynchocyon* zoo specimens. Lastly, we dedicate this paper to our coauthor, friend, and colleague W. T. Stanley, who

died suddenly and unexpectedly in November 2015 while conducting research in Ethiopia.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.05.012.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723. http://dx.doi.org/10.1109/TAC.1974.1100705.

Andanje, S., Agwanda, B.R., Ngaruiya, G.W., Amin, R., Rathbun, G.B., 2010. Sengi (elephant-shrew) observations from northern coastal Kenya. J. East African Nat. Hist. 99, 1–8. http://dx.doi.org/10.2982/028.099.0101.

Baker, A.J., McCafferty, K., Hellmuth, H., 2005. Black-and-rufous sengi (*Rhynchocyon petersi*) at the Philadelphia Zoo. Afrotherian Conservation - Newsletter of the IUCN-SCC Afrotheria Specialist Group 3, 6–7.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659. http://dx.doi.org/10.1046/j.1365-294X.2000.01020.x.

Coals, P.G.R., Rathbun, G.B., 2013. The taxonomic status of giant sengis (genus *Rhynchocyon*) in Mozambique. J. East African Nat. Hist. 101, 241–250. http://dx.doi.org/10.2982/028.101.0203.

Corbet, G.B., Hanks, J., 1968. A revision of the elephant-shrews, family Macroscelididae. Bull. Br. Museum (Natural Hist.) Zool. 16, 45–111.

Dasmahapatra, K.K., Silva-Vásquez, A., Chung, J., Mallet, J., 2007. Genetic analysis of a wild-caught hybrid between non-sister *Heliconius* butterfly species. Biol. Lett. 3, 660–663. http://dx.doi.org/10.1098/rsbl.2007.0401.

Davies, T.J., Buckley, L.B., 2011. Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. Philos. Trans. R. Soc. B 366, 2414–2425. http://dx.doi.org/10.1098/rstb.2011.0058.

deMenocal, P.B., 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. Earth Planet. Sci. Lett. 220, 3–24.

Dimitrov, D., Nogués-Bravo, D., Scharff, N., 2012. Why do tropical mountains support exceptionally high biodiversity? The eastern arc mountains and the drivers of saintpaulia diversity. PLoS ONE 7 (11), e48908. http://dx.doi.org/10.1371/journal.pone.0048908.

Douady, C., 2001. Molecular Phylogenetics of the Insectivora PhD diss. The Queen's University of Belfast.

Douady, C.J., Catzeflis, F., Raman, J., Springer, M.S., Stanhope, M.J., 2003. The Sahara as a vicariant agent, and the role of Miocene climatic events, in the diversification of the mammalian order Macroscelidea (elephant shrews). Proc. Natl. Acad. Sci. U.S.A. 100, 8325–8330. http://dx.doi.org/10.1073/pnas.0832467100.

Dumbacher, J.P., Rathbun, G.B., Smit, H.A., Eiseb, S.J., 2012. Phylogeny and taxonomy of the round-eared sengis or elephant-shrews, genus *Macroscelides* (Mammalia, Afrotheria, Macroscelidea). PLoS ONE 7, e32410. http://dx.doi.org/10.1371/journal.pone.0032410.

Dumbacher, J.P., Rathbun, G.B., Osborne, T.O., Griffin, M., Eiseb, S.J., 2014. A new species of round-eared sengi (genus *Macroscelides*) from Namibia. J. Mammal. 95, 443–454. http://dx.doi.org/10.1644/13-MAMM-A-159.

Dumbacher, J.P., Carlen, E.J., Rathbun, G.B., 2016. *Petrosaltator* gen. nov., a new genus replacement for the North African sengi *Elephantulus rozeti* (Macroscelidea; Macroscelididae). Zootaxa 4136, 567–579.

Faith, D.P., 1992. Conservation evaluation and phylogenetic diversity. Biol. Conserv. 61, 1–10. http://dx.doi.org/10.1016/0006-3207(92)91201-3.

Fjeldsa, J., Bowie, R.C.K., 2008. New perspectives on the origin and diversification of Africa's forest avifauna. Afr. J. Ecol. 46, 235–247.

FitzGibbon, C.D., 1997. The adaptive significance of monogamy in the golden-rumped elephant-shrew. J. Zool. 242, 167–177. http://dx.doi.org/10.1111/j.1469-7998.1997.tb02937.x.

Good, J.M., Demboski, J.R., Nagorsen, D.W., Sullivan, J., 2003. Phylogeography and introgressive hybridization: chipmunks (genus *Tamias*) in the northern Rocky Mountains. Evolution 57, 1900–1916. http://dx.doi.org/10.1111/j.0014-3820.2003.tb00597.x.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucl. Acids Res. 30, 3059–3066. http://dx.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. http://dx.doi.org/10.1093/bioinformatics/bts199.

Kingdon, J., 1989. Island Africa – The Evolution of Africa's Rare Animals and Plants. Princeton University Press, New Jersey. 287 pp.

Larsen, P.A., Marchán-Rivadeneira, M.R., Baker, R.J., 2010. Natural hybridization generates mammalian lineage with species characteristics. Proc. Natl. Acad. Sci. U.S.A. 107, 11447–11452. http://dx.doi.org/10.1073/pnas.1000133107.

Lawson, L.P., 2010. The discordance of diversification: evolution in the tropical-montane frogs of the Eastern Arc Mountains of Tanzania. Mol. Ecol. 19, 4046–4060.

Lawson, L.P., Vernesi, C., Ricci, S., Rovero, F., 2013. Evolutionary history of the grey-faced sengi, *Rhynchocyon udzungwensis*, from Tanzania: a molecular and species distribution modeling approach. PLoS ONE 8, e72506. http://dx.doi.org/10.1371/journal.pone.0072506.

Leigh, J., Bryant, D., Steel, M., 2013. PopART (Population Analysis with Reticulate Trees). <http://www.popart.otago.ac.nz/downloads.shtml> (Accessed 22 February 2015).

McKitrick, M.C., Zink, R.M., 1988. Species concepts in ornithology. Condor 90, 1–14.

Novacek, M., 1984. Evolutionary stasis in the elephant-shrew, *Rhynchocyon*. In: Eldredge, N., Stanley, S.M. (Eds.), Living Fossils. Springer Verlag, New York, New York, pp. 4–22.

Nylander, J.A.A., 2004. MrModeltest v2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University.

Pattengale, N., Alipour, M., Bininda-Emonds, O., Moret, B., Stamatakis, A., 2010. How many bootstrap replicates are necessary? J. Comput. Biol. 17, 337–354. http://dx.doi.org/10.1089/cmb.2009.017.

Puttick, M.N., Thomas, G.H., 2015. Fossils and living taxa agree on patterns of body mass evolution: a case study with Afrotheria. Proc. R. Soc. B-Biol. Sci. 282, 20152023. http://dx.doi.org/10.1098/rspb.2015.2023.

Rathbun, G.B., 1979. The social structure and ecology of elephant-shrews. Z. Tierpsychol. (Supplement 20), 1–77.

Rathbun, G.B., 2009. Why is there discordant diversity in sengi (Mammalia: Afrotheria: Macroscelidea) taxonomy and ecology? Afr. J. Ecol. 47, 1–13.

Rathbun, G.B., Kingdon, J., 2006. The etymology of 'sengi.' Afrotherian Conservation Newsletter IUCN/SSC Afrotheria Specialist Gr. 4, 14–15.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. http://dx.doi.org/10.1093/bioinformatics/btg180.

Rovero, F., Rathbun, G.B., Perkin, A., Jones, T., Ribble, D.O., Leonard, C., Mwakisoma, R.R., Doggart, N., 2008. A new species of giant sengi or elephant-shrew (genus *Rhynchocyon*) highlights the exceptional biodiversity of the Udzungwa Mountains of Tanzania. J. Zool. 274, 126–133. http://dx.doi.org/10.1111/j.1469-7998.2007.00363.x.

Sabuni, C.A., Van Houtte, N., Gryseels, S., Maganga, S.L.S., Makundi, R.H., Leirs, H., Goüy de Bellocq, J., 2016. Genetic structure and diversity of the black and rufous sengi in Tanzanian coastal forests. J. Zool. 1–9. http://dx.doi.org/10.1111/jzo.12384.

Smit, H., 2008. Phylogeography of Three Southern African Endemic Elephant-Shrews and a Supermatrix Approach to the Macroscelidea PhD diss. Stellenbosch University.

Smit, H.A., Robinson, T.J., van Vuuren, B.J., 2007. Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). Mol. Ecol. 16, 2680–2692. http://dx.doi.org/10.1111/j.1365-294X.2007.03334.x.

Smit, H.A., Robinson, T.J., Watson, J., van Vuuren, B.J., 2008. A new species of elephant-shrew (Afrotheria: Macroscelidea: *Elephantulus*) from South Africa. J. Mammal. 89, 1257–1268.

Smit, H.A., van Vuuren, J., O'Brien, P.C.M., Ferguson-Smith, M., Yang, F., Robinson, T. J., 2011. Phylogenetic relationships of elephant-shrews (Afrotheria, Macroscelididae). J. Zool. 284, 133–143. http://dx.doi.org/10.1111/j.1469-7998.2011.00790.x.

Smith, H.F., Parker, W., Kotzé, S.H., Laurin, M., 2013. Multiple independent appearances of the cecal appendix in mammalian evolution and an investigation of related ecological and anatomical factors. C.R. Palevol 12, 339–354. http://dx.doi.org/10.1016/j.crpv.2012.12.001.

Smith, H.F., Parker, W., Kotzé, S.H., Laurin, M., 2016. Morphological evolution of the mammalian cecum and cecal appendix. C.R. Palevol 16, 39–57. http://dx.doi.org/10.1016/j.crpv.2016.06.001.

Springer, M.S., Cleven, G.C., Madsen, O., de Jong, W.W., Waddell, V.G., Amrine, H.M., Stanhope, M.J., 1997. Endemic African mammals shake the phylogenetic tree. Nature 388, 61–64. http://dx.doi.org/10.1038/40386.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690. http://dx.doi.org/10.1093/bioinformatics/btl446.

Stanley, W.T., Rogers, M.A., Hutterer, R., 2005. A new species of *Congosorex* from the Eastern Arc Mountains, Tanzania, with significant biogeographical implications. J. Zool., London 265, 269–280.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taylor, P.J., Maree, S., Van Sandwyk, J., Kerbis Peterhans, J.C., Stanley, W.T., Verheyen, E., Kaliba, P., Verheyen, W., Kaleme, P., Bennett, N.C., 2009. Speciation mirrors geomorphology and palaeoclimatic history in African laminate-toothed rats (Muridae: Otomyini) of the *Otomys denti* and *Otomys lacustris* species-complexes in the 'Montane Circle' of East Africa. Biol. J. Lin. Soc. 96, 913–941.

Villesen, P., 2007. FaBox: an online toolbox for FASTA sequences. Mol. Ecol. Notes 7, 965–968. http://dx.doi.org/10.1111/j.1471-8286.2007.01821.

Yang, Z., 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 61, 854–865.