University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

USGS Staff -- Published Research | US Geological Survey

2010

Barcoding bushmeat: molecular identification of Central African and South American harvested vertebrates

Mitchell J. Eaton
University of Colorado, Boulder & U.S. Geological Survey, meaton@usgs.gov

Greta J. Meyers
Barnard College

Sergios-Orestis Kolokotronis
American Museum of Natural History

Matthew S. Leslie
American Museum of Natural History

Andrew P. Martin
University of Colorado, Boulder

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unl.edu/usgsstaffpub

Part of the Geology Commons, Oceanography and Atmospheric Sciences and Meteorology Commons, Other Earth Sciences Commons, and the Other Environmental Sciences Commons

Eaton, Mitchell J.; Meyers, Greta J.; Kolokotronis, Sergios-Orestis; Leslie, Matthew S.; Martin, Andrew P.; and Amato, George, "Barcoding bushmeat: molecular identification of Central African and South American harvested vertebrates" (2010). *USGS Staff -- Published Research*. 1127.
https://digitalcommons.unl.edu/usgsstaffpub/1127

This Article is brought to you for free and open access by the US Geological Survey at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in USGS Staff -- Published Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
Mitchell J. Eaton, Greta J. Meyers, Sergios-Orestis Kolokotronis, Matthew S. Leslie, Andrew P. Martin, and George Amato

This article is available at DigitalCommons@University of Nebraska - Lincoln: https://digitalcommons.unl.edu/usgsstaffpub/1127

Conserv Genet (2010) 11:1389–1404
DOI 10.1007/s10592-009-9967-0

RESEARCH ARTICLE

Barcoding bushmeat: molecular identification of Central African and South American harvested vertebrates

Mitchell J. Eaton · Greta L. Meyers · Sergios-Orestis Kolokotronis · Matthew S. Leslie · Andrew P. Martin · George Amato

Received: 28 February 2009 / Accepted: 23 July 2009 / Published online: 1 September 2009
© Springer Science+Business Media B.V. 2009

U.S. government works are not subject to copyright.

Abstract The creation and use of a globally available database of DNA sequences from a standardized gene region has been proposed as a tool for species identification, assessing genetic diversity and monitoring the legal and illegal trade in wildlife species. Here, we contribute to the Barcode of Life Data System and test whether a short region of the mitochondrial cytochrome *c* oxidase subunit 1 (*COX1*) gene would reliably distinguish among a suite of commonly hunted African and South American mammal and reptile species. We used universal primers to generate reference barcode sequences of 645 bp for 23 species from five vertebrate families (*Crocodilidae*, *Alligatoridae*, *Bovidae*, *Suidae* and *Cercopithecidae*). Primer cocktails yielded high quality barcode sequences for 179 out of 204 samples (87.7%) from all species included in the study. For most taxa, we sequenced multiple individuals to estimate intraspecific sequence variability and document fixed diagnostic characters for species identification. Polymorphism in the *COX1* fragment was generally low (mean = 0.24%), while differences between congeneric species averaged 9.77%. Both fixed character differences and tree-based maximum likelihood distance methods unambiguously identified unknown and misidentified samples with a high degree of certainty. Barcode sequences also differentiated among newly identified lineages of African crocodiles and identified unusually high levels of genetic diversity in one species of African duiker. DNA barcoding offers promise as an effective tool for monitoring poaching and commercial trade in endangered species, especially when investigating semi-processed or morphologically indistinguishable wildlife products. We discuss additional benefits of barcoding to ecology and conservation.

Keywords Barcode of life · Caiman · Crocodiles · Cytochrome *c* oxidase subunit 1 (*COX1*) · Hunting · Molecular forensics · Primates · Wildlife monitoring · Maximum likelihood phylogeny · Ungulates

Electronic supplementary material The online version of this article (doi:10.1007/s10592-009-9967-0) contains supplementary material, which is available to authorized users.

M. J. Eaton · A. P. Martin
Department of Ecology and Evolutionary Biology, N122 Ramaley, University of Colorado, Boulder, CO 80309, USA

G. L. Meyers
Department of Environmental Science, Barnard College, 3009 Broadway, 404 Altschul Hall, New York, NY 10027, USA

S.-O. Kolokotronis · M. S. Leslie · G. Amato
Sackler Institute for Comparative Genomics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA

M. J. Eaton (✉)
Patuxent Wildlife Research Center, US Geological Survey, 12100 Beech Forest Road, Laurel, MD 20708-4017, USA
e-mail: meaton@usgs.gov; eaton.mitch@gmail.com

Introduction

The hunting of tropical wildlife has historically been conducted for subsistence consumption and for local trade. But current trends in wildlife harvest from across the globe suggest that the volume of extraction of wild game, or “bushmeat”, has increased considerably, and many species are in sharp decline due to over exploitation (Albrechtsen et al. 2007; Bennett et al. 2007; Milner-Gulland et al. 2003; Redford 1992). Former locally-based subsistence economies have become global, and bushmeat is now a significant export product traded at regional and international

scales. Of equal importance to the deleterious impacts on wild populations and ecosystems promoted by unregulated harvest are the potential human and agricultural heath impacts stemming from the increased potential for zoonotic disease transfer (Chomel et al. 2007; Milius 2005).

Considering the international trade in bushmeat and fisheries, estimates of its worth are in excess of US$60 billion per year, with wildlife and wildlife products contributing US$5–15 billion (Baker 2008). A significant portion of this trade is illegal (US$5–8 billion, in Baker 2008), involving species that are protected by national laws and international conventions governing the use of wildlife and wildlife products. Given the illicit nature of the trade, it is difficult to accurately assess and monitor the volumes and species involved (Chomel et al. 2007; Milius 2005) and thus fully understand existing and potential impacts on economies, wildlife populations and health.

Existing legislation and treaties governing the trade in wildlife, such as the Convention on the International Trade of Endangered Species (CITES) and the United States Endangered Species Act (ESA), are based on the recognition of distinct population or taxonomic units. At a minimum, enforcement of regulations depends upon an ability to identify suspected illegal products at the species level. Accurate identification is often impaired due to the types of products involved, which are typically processed and difficult to identify using morphological techniques. To improve our ability to detect, monitor and control the trade in wildlife and wildlife products, more accurate and efficient methods of species identification are required. Identification of fish and wildlife species targeted for commercial trade is considered among the most useful applications of molecular ecology (Baker 2008).

The ability to identify wildlife products, whether as processed meat, skins or whole animals, is being formalized by the development of DNA sequence databases using a standardized gene fragment (Ratnasingham and Hebert 2007; Ross et al. 2003). A database of single-gene "barcodes" has been proposed to classify the complete diversity of life (Hebert et al. 2003a; Ratnasingham and Hebert 2007) and proponents argue that such a tool could variously be employed for defining taxonomic units for conservation (Neigel et al. 2007; but see Rubinoff 2006), biological inventory (Janzen et al. 2005) and species discovery (Bickford et al. 2007; Hebert et al. 2004; Kaila and Stahls 2006; Witt et al. 2006). The 5′ Folmer region of the mitochondrial (mtDNA) cytochrome *c* oxidase I (*COX1*) gene has been recommend as a standard for DNA barcoding (Folmer et al. 1994; Hebert et al. 2003a, b; Ivanova et al. 2007). Although there has been considerable criticism of the philosophical and practical underpinnings of DNA barcoding (DeSalle 2006; Fitzhugh 2006; Rubinoff 2006; Rubinoff et al. 2006; Song et al. 2008), its application for species identification has largely been uncontested (e.g. Rubinoff et al. 2006).

The use of sequence data for investigations of endangered species collected in commercial markets is widely reported. While the *COX1* region has not been used as a standard metric, these studies have employed mitochondrial markers to identify samples to the species level. Yan et al. (2005) used the cytochrome *b* (*CYTB*) gene to identify Chinese alligators (*Alligator sinensis*) from fresh and partially cooked meat found in Chinese markets. Baker and colleagues used mtDNA sequences and microsatellites to identify endangered whale species sampled from markets in Korea and Japan (Baker et al. 1996, 2002). Using diagnostic characters in *CYTB*, Birstein et al. (1998) identified three species of caviar-producing sturgeon and discovered that nearly a quarter of commercially available caviar lots sold in New York City were mislabeled. Marko et al. (2004) also used *CYTB* sequences to determine that 77% of fish sold in the US labeled as red snapper were actually other species. Martin (1991) and Moura et al. (2008) used mtDNA sequences, including *COX1*, to identify commercially fished shark species when morphological characters (e.g. fins, heads) were equivocal in discerning among and within genera. The identification of endangered wildlife products in each of these cases would not have been possible without molecular methods because vendors had a vested interest in concealing the identity of the species being sold.

Ours is the first study to examine the utility of universal *COX1* primers as a standard metric to identify multiple species for monitoring the global trade in wildlife, with particular emphasis on species commonly traded in bushmeat markets. Our study includes species from five taxonomic families: bovids (duikers and spiral-horned antelope; genera: *Cephalophus*, *Tragelaphus*), suids (red river hog; genus: *Potamochoerus*), cercopithecoid primates (old world monkeys and mangabeys; genera: *Cercopithecus*, *Lophocebus*), alligators (genera: *Caiman*, *Melanosuchus*, *Paleosuchus*) and crocodiles (genera: *Crocodylus*, *Osteolaemus*, *Mecistops*) (Table 1). We obtained samples either from museum collections or from in situ captures. We use fixed diagnostic characters to compare interspecific and intergeneric levels of variation and bootstrap node support from a maximum likelihood (ML) tree-based approach to assess the monophyly of closely related species. Diagnostic characters and phylogenetic support were also used to assign a small number of unidentified samples to species and to examine the museum collection and field specimens for possible errors in species identification or labeling. The poor quality of some tissue used in this study reflects the suboptimal conditions encountered when working with material sampled from bushmeat markets, processed wildlife products, or reference museum samples

Table 1 Sample list, sample size, number of samples that failed to sequence, tissue type, sample origin, and conservation status for 11 genera included in the *COX1* barcoding study. Unknown samples are listed with their respective known family classification

Order	Family	Species	Museum	Wild	Failed	Storage method	Origin	Status[a]
Crocodilia	Alligatoridae	*Paleosuchus palpebrosus*	4	–	2	Tissue (blood) in buffer	Brazil	Least concern
		Paleosuchus trigonatus	3	–	–	Tissue (blood) in buffer	Brazil	Least concern
		Caiman crocodilus chiapasius	4	–	–	Tissue (blood) in buffer, dried blood on gauze	Costa Rica	Least concern
		Caiman yacare[b]	8	–	3	Dried blood on gauze	Bolivia; Brazil; Peru	Least concern
		Caiman c. crocodilus × *yacare*	1	–	–	Tissue (blood) in buffer	Brazil	Least concern
		Melanosuchus niger	4	–	–	Dried tissue (non-blood)	Peru	Least concern
		Caiman latirostris	8	–	2	Tissue (blood) in buffer, dried blood on gauze	Bolivia; Brazil; Paraguay	Least concern
		Unknown caiman	1	–	–	Tissue (blood) in buffer, dried blood on gauze	Bolivia	
	Crocodylidae	*Crocodylus acutus*	1	–	–	Dried tissue (non-blood)	Oaxaca, Mexico	Vulnerable
		Mecistops cataphractus	–	10	–	Dried tissue (non-blood)	Gabon; Congo	Data deficient
		Crocodylus niloticus	–	8	–	Dried tissue (non-blood)	Gabon	Least concern
		Crocodylus niloticus	–	1	–	Dried tissue (non-blood)	Congo	Least concern[c]
		Osteolaemus tetraspis	1	16	3	Dried tissue (non-blood)	Gabon; Cameroon	Vulnerable
		Osteolaemus osborni	–	11	1	Dried tissue (non-blood), tissue in buffer	Congo	Vulnerable
		Unknown African crocodile[d]	5	–	4	Dried tissue (processed leather)	Unknown	
Artiodactyl	Bovidae	*Tragelaphus eurycerus eurycerus*	–	1	–	Dried blood on filter paper	Ndoki, Congo	Near threatened
		Cephalophus nigrifrons	–	5	–	Tissue in buffer	Ndoki, Congo	Least concern
		Cephalophus leucogaster	–	8	1	Tissue in buffer	Ndoki, Congo	Least concern
		Cephalophus dorsalis[e]	–	10	1	Tissue in buffer	Kabo, Congo	Least concern
		Cephalophus callipygus	–	26	3	Dried blood on filter paper	Ndoki, Congo	Least concern
		Cephalophus monticola	–	26	2	Dried blood on filter paper, tissue in buffer	Ndoki, Congo	Least concern
		Unknown ungulate	–	5	1	Dried blood on filter paper	Congo	
	Suidae	*Potamochoerus porcus*	–	8	1	Tissue in buffer	Ndoki, Congo	Least concern
Primates	Cercopithecidae	*Cercopithecus cephus*	–	7	–	Dried blood on filter paper, tissue in buffer	Ndoki, Congo	Least concern
		Cercopithecus nictitans	–	8	1	Dried blood on filter paper	Ndoki, Congo	Least concern
		Cercopithecus pogonias	–	6	–	Dried blood on filter paper, tissue in buffer	Ndoki, Congo	Least concern
		Lophocebus albigena	–	8	–	Dried blood on filter paper	Ndoki, Congo	Least concern
Total			40	164	25			

[a] As established by the International Union for the Conservation of Nature's (IUCN) Red List of Threatened Species, www.iucnredlist.org

[b] Includes one sample originally identified as *M. niger*

[c] The IUCN has yet to recognize the new crocodile species and their status remains the same as the original taxa

[d] From US Fish & Wildlife Service, National Wildlife Property Repository (http://www.fws.gov/mountain-prairie/law/property_repository.html)

[e] Includes one sample originally identified as *C. callipygus*

and, therefore, provides a robust evaluation of the utility of DNA barcoding for wildlife monitoring and investigations.

Materials and methods

Species identification and sample collection

Specimens used in this study were identified and collected under four different scenarios. The first involved harvested mammals surveyed during a prior study of bushmeat hunting in the Republic of Congo (Table 1; Eaton 2002). Species identifications were made by MJE and trained field assistants using African mammal guides (Estes 1991; Kingdon 1997) in conjunction with the knowledge of local hunters. Voucher photographs were taken of all species and later confirmed using Nowak (1999). All samples, with the exception of five unidentified ungulates, were collected from freshly killed, whole animals in which species identification was unambiguous. The five unidentified samples were collected from partially butchered animals that could only be confirmed as belonging to ungulates based on hair pattern and coloration. Samples of Osborn's dwarf crocodiles (*Osteolaemus osborni*) and a single Nile crocodile (*Crocodylus niloticus*) were collected from intact, hunted animals during a subsequent study in the Republic of Congo (Eaton et al. 2009; Thorbjarnarson and Eaton 2004). The second sample collection scenario consisted of capture and identification of live African crocodiles for systematics and ecological research in the Republics of Congo and Gabon (Eaton 2006; Eaton and Barr 2005; Eaton et al. 2009; Thorbjarnarson and Eaton 2004). Although there have been recent revisions to the taxonomy of all African crocodiles (Eaton et al. 2009; Hekkala 2004; McAliley et al. 2006; Schmitz et al. 2003), identification to genus (and in most cases to species) was unambiguous. The third source of samples was voucher crocodilian specimens obtained from museum collections. These included dried blood and tissue from six species of South American caiman, the American crocodile (*Crocodylus acutus*) and the Central African dwarf crocodile (*Osteolaemus tetraspis*) provided by the Yale Peabody Museum (YPM) and the American Museum of Natural History (AMNH; see Table 2 for details and accession numbers). One sample was identified only as 'caiman' (YPM 15394) and analyzed as an unknown species. The final source of samples used in this study was wildlife products confiscated by the US Fish and Wildlife Service as illegal imports and donated to this project as research material. Aside from originating in Africa, the specimens (five crocodile skin handbags) contained no additional information on location or species. These unidentified samples were included in the study to test the ability of standard barcode primers to amplify highly degraded material. Field-collected tissue samples were stored in 10% buffered EDTA-DMSO and kept at room temperature for up to several months before being stored at −20°C; blood samples were applied directly to Whatman filter paper (Florham Park, NJ), then dried and stored at room temperature for nearly 10 years.

DNA extraction, amplification and sequencing

Samples included in the study had been stored as preserved fresh tissue, dried tissue, blood dried in buffer, blood dried on filter paper, and processed skin products. DNA was extracted from tissue and blood using DNEasy kits (Qiagen) in a pre-PCR laboratory to prevent contamination. Extractions followed the manufacturer's protocol for buffered animal tissues but were modified slightly for extraction of DNA from dried blood or older tissue to ensure maximum yield from low-quality samples. Modifications included incubating the lysed tissue at 65°C for 15 min after adding AL buffer and incubating again at 4°C for 1 h after adding ethanol. To maximize final yield of genomic DNA, 75 μl of AE buffer, preheated to 70°C, was added and left to incubate for 45 min before centrifuging and collecting flow-through. Because the processed crocodile leather products were the most difficult to extract, we assessed their DNA yield using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific).

Sequence data were generated using one of three universal *COX1* primer "cocktails" (Table 3; Ivanova et al. 2007). Primer selection for each taxon was made through polymerase chain reaction (PCR) optimization of cocktails 'COI-1,' 'COI-2' and 'COI-3' and one individual primer (VF1d_t1/VR1d_t1) using representative samples from each genus of artiodactyl, primate and crocodilian included in the study, including positive control samples previously sequenced at other gene regions. Table 3 summarizes the taxon-specific primer combinations. PCR was performed in Mastercycler EP gradient S thermocyclers (Eppendorf) in a 25 μl reaction volume containing 1.0 μl genomic DNA (~25 ng/μl), 1× PCR Buffer (Fisher Scientific), 0.24 μM dNTPs, 15 ng BSA, 1 U Taq polymerase (Fisher Scientific), and 0.4 μM of each forward and reverse primer or primer cocktail (Integrated DNA Technologies). Failed amplifications were repeated under the same conditions with 2 μl of genomic DNA. Generally, the optimized PCR thermal cycling profile was a step-up protocol of 94°C for 3 min, 5 cycles of 94°C for 30 s, 51.1°C for 40 s and 72°C for 1 min, followed by 30 cycles of 94°C for 30 s, 56.9°C for 40 s and 72°C for 1 min, with a final extension at 72°C for 10 min. The COI-1 primer cocktail required a different annealing temperature, (54.2°C) for the first 5 cycles. PCR products were visualized by agarose gel electrophoresis, and the cocktail or primer pair yielding the brightest, thinnest band

Table 2 Accession and locality information for museum and US Fish and Wildlife crocodilian specimens sequenced for *COX1* barcoding

Study ID	Institution accession #[a]	Species	Country	Locality	Collection year	Sex
Caiman1	YPM 15394	Caiman	Bolivia	Tarija Dept	1986	
C. c. chiapasius3	YPM 15709	*Caiman crocodilus chiapasius*	Costa Rica	Puntarenas Province	1993	M
C. c. chiapasius4	YPM 15713	*Caiman crocodilus chiapasius*	Costa Rica	Puntarenas Province	1993	M
C. c. chiapasius16	YPM 15741	*Caiman crocodilus chiapasius*	Costa Rica	Alajuela Province	1993	
C. c. chiapasius17	YPM 15743	*Caiman crocodilus chiapasius*	Costa Rica	Alajuela Province	1993	
C. yacare7	YPM 15402	*Caiman crocodilus yacare*	Bolivia	SantaCruz Dept	1989	
C. yacare8	YPM 15401	*Caiman crocodilus yacare*[b]	Bolivia	El Beni Dept	1986	
C. yacare15	YPM 15683	*Caiman crocodilus yacare*	Brazil	Mato Grosso State	1987	
C. yacare32	YPM 15570	*Caiman crocodilus yacare*	Brazil	Mato Grosso Estate	1987	F
C. yacare33	YPM 15410	*Caiman crocodilus yacare*	Bolivia	Pando Dept	1986	
C. yacare34	YPM 15775	*Caiman crocodilus yacare*[b]	Paraguay	Misiones Dept	1986–1987	
C. yacare35	YPM 15669	*Caiman crocodilus yacare*[b]	Brazil	Mato Grosso State	1986–1987	
C.c. crocodilus × yacare2	YPM 15547	*C.c.crocodilus* × *yacare*	Brazil	Amazonas State	1987	F
C. latirostris5	YPM 15754	*Caiman latirostris*	Paraguay	Presidente Hayes Dept	1986–1987	
C. latirostris6	YPM 15755	*Caiman latirostris*	Paraguay	Neembucu Dept	1986–1987	
C. latirostris18	YPM 15392	*Caiman latirostris*	Bolivia	Tarija Dept	1986	
C. latirostris19	YPM 15554	*Caiman latirostris*[b]	Brazil	Espirito Santo State	1989	
C. latirostris20	YPM 15551	*Caiman latirostris*[b]	Brazil	Espirito Santo State	1989	
C. latirostris21	YPM 15548	*Caiman latirostris*	Brazil	Mato Grosso du Sol	1987	M
C. latirostris22	YPM 15754	*Caiman latirostris*	Paraguay	Presidente Hayes Dept	1986–1987	
C. latirostris23	YPM 15393	*Caiman latirostris*	Bolivia	Tarija Dept	1986	
M. niger9	YPM 15833	*Melanosuchus niger*	Peru	Madre de Dios Dept	1984	
M. niger10	YPM 15834	*Melanosuchus niger*	Peru	Madre de Dios Dept	1984	
M. niger24	YPM 15695	*Melanosuchus niger*	Peru	Madre de Dios Dept	1987	F
M. niger25	YPM 15832	*Melanosuchus niger*[c]	Peru	Madre de Dios Dept	1984	
M. niger26	YPM 15835	*Melanosuchus niger*	Peru	Madre de Dios Dept	1984	
P. palpebrosus11	YPM 15703	*Paleosuchus palpebrosus*	Brazil	Mato Grosso State	1987	
P. palpebrosus12	YPM 15702	*Paleosuchus palpebrosus*[b]	Brazil	Rondonia State	1987	
P. palpebrosus28	YPM 15697	*Paleosuchus palpebrosus*[b]	Brazil	Rondonia State	1987	
P. palpebrosus29	YPM 15704	*Paleosuchus palpebrosus*	Brazil	Mato Grosso State	1988	
P. trigonatus13	YPM 15705	*Paleosuchus trigonatus*	Brazil	Rodinia State	1987	F
P. trigonatus31	YPM 15700	*Paleosuchus trigonatus*	Brazil	Rodinia State	1987	
P. trigonatus14	YPM 15699	*Paleosuchus trigonatus*	Brazil	Rodinia State	1987	
C. acutusM7	AMNH R100634	*Crocodylus acutus*	Mexico	Oaxaca	1967	
O. tetraspisM11	AMNH R75421	*Osteolaemus tetraspis*	Cameroon	Unknown	Unknown	
1USFWS	USFWS PB800	African crocodile1[d]	Africa	Unknown	Unknown	
2USFWS	USFWS C5236	African crocodile2[b]	Africa	Unknown	Unknown	
3USFWS	USFWS C7848	African crocodile3[b]	Africa	Unknown	Unknown	
4USFWS	USFWS C6688	African crocodile4[b]	Africa	Unknown	Unknown	
5USFWS	USFWS C3508	African crocodile5[b]	Africa	Unknown	Unknown	

[a] Yale Peabody Museum (YPM); American Museum of Natural History (AMNH); US Fish & Wildlife Service National Wildlife Property Repository (USFWS)

[b] Unsuccessfully sequenced

[c] Determined to be *C. yacare*

[d] Determined to be *C. niloticus*

Table 3 Primers, primer cocktails and universal M13 tails used to sequence all samples included in this study. Table and primer information modified from Ivanova et al. (2007). Position of M13 tail indicated by *[M13F]* or *[M13R]*

Name	Ratio	Cocktail name/Primer sequence 5′-3′
COI-1 duikers (*Cephalophus* spp.), red river hog (*Potamochoerus porcus*), gray-cheeked mangabey (*Lopohocebus albigena*)		
FF2d	1	TTCTCCACCAACCACAARGAYATYGG
FR1d	1	CACCTCAGGGTGTCCGAARAAYCARAA
COI-2 guenons (*Cercopithecus* spp.), in bold; bongo (*Tragelaphus eurycerus*)		
LepF1_t1	1	*[M13F]*ATTCAACCAATCATAAAGATATTGG
VF1_t1	1	*[M13F]*TCTCAACCAACCACAAAGACATTGG
VF1d_t1	1	***[M13F]*TCTCAACCAACCACAARGAYATYGG**
VF1i_t1	3	*[M13F]*TCTCAACCAACCAIAAIGAIATIGG
LepRI_t1	1	*[M13R]*TAAACTTCTGGATGTCCAAAAAATCA
VR1d_t1	1	***[M13R]*TAGACTTCTGGGTGGCCRAARAAYCA**
VR1_t1	1	*[M13R]*TAGACTTCTGGGTGGCCAAAGAATCA
VR1i_t1	3	*[M13R]*TAGACTTCTGGGTGICCIAAIAAICA
COI-3 (crocodilians)		
VF2_t1	1	*[M13F]*CAACCAACCACAAAGACATTGGCAC
FishF2_t1	1	*[M13F]*CGACTAATCATAAAGATATCGGCAC
FishR2_t1	1	*[M13R]*ACTTCAGGGTGACCGAAGAATCAGAA
FR1d_t1	1	*[M13R]*ACCTCAGGGTGTCCGAARAAYCARAA
M13F (−21)		TGTAAAACGACGGCCAGT
M13R (−27)		CAGGAAACAGCTATGAC

was chosen for each taxon. PCR products were purified with Ampure magnetic beads (Agencourt Bioscience) on a BioMek FX robotic platform (Beckman Coulter Inc.). Primer cocktails COI-2 and COI-3 included a modified forward and reverse universal M13 tag (Ivanova et al. 2007). A single primer pair, M13(−21) and M13(−27), was used for bidirectional sequencing reactions of PCR products generated from these two cocktails using BigDye v1.1 chemistry (Applied Biosystems Inc.) on a 3730xl DNA Analyzer (Applied Biosystems, Inc.). Forward and reverse sequences were assembled and edited in Sequencher 4.6 (Gene Codes Corp.) and verified by eye. Contig sequences were aligned using ClustalW (Thompson et al. 1994) as implemented in MEGA 4.0 (Tamura et al. 2007).

Sequence analysis

Unidentified crocodilian and ungulate samples were sequenced but then set apart from the remaining data set. All sequences obtained from species identified through museum collections or by independent morphological evaluation were grouped into their respective taxa and examined for diagnostic molecular characters that could be used as identifiers in future studies. Because species misidentifications are possible in field studies, as well as in the accession of voucher museum specimens, we examined both field and museum samples for obvious errors in species designation. Samples with possible mistaken identities were compared with all other known species to determine if a match could be made based on diagnostic characters. Because of small sample sizes for several taxa and the associated problems of overestimating diagnostic character sites (Brower 1999; Davis and Nixon 1992), we also confirmed misidentified specimens by means of their placement on phylogenetic trees. We used maximum likelihood (ML) phylogenetic inference as implemented in RAxML 7.0.4 (Stamatakis 2006) using the general time-reversible (GTR) substitution model (Lanave et al. 1984; Rodriguez et al. 1990) with rate heterogeneity parameters modeled by the Gamma (Γ) distribution and four rate categories (Yang 1994). Node support was evaluated with 100 rapid bootstrap replicates (Stamatakis et al. 2008). Outgroups included the American alligator (*Alligator mississippiensis*) for both Crocodylidae and Alligatoridae, the bongo (*Tragelaphus eurycerus eurycerus*) for the ungulate phylogeny, and the rhesus macaque (*Macaca mulatta*) for primates. Resulting trees were inspected for monophyletic groupings and for phylogenetic support values of nodes subtending possibly misidentified samples. Mistakes that could unambiguously be attributed to a species based on matching of diagnostic characters and phylogenetic placement were included in their respective taxonomic group for subsequent analyses.

Sequence variability in *COX1* was evaluated at three hierarchical levels: among conspecifics, among congeneric species and among genera within each of the three orders included in this study. The bongo (*T. eurycerus eurycerus*) and the American crocodile (*Crocodylus acutus*), each represented by only one sample, were excluded from these analyses. Intraspecific comparisons quantified nucleotide diversity for each species based on average nucleotide

substitutions per site (transitions + transversions) using a Tamura-Nei model with pairwise sequence comparisons (Tamura and Nei 1993). To compare congeneric species, we identified fixed nucleotide character differences for every species pair within each represented genus and calculated species divergence using the same Tamura-Nei substitution model. This same approach was used to compare average genus-level nucleotide divergence among orders. Positions that contained a fixed character state among all individuals within a species and an alternate but fixed state at the homologous site in the second species were considered diagnostic at the species level. We used MEGA for genetic diversity and divergence calculations and for visual examination of homologous character state positions among congeneric species. We assessed species monophyly and divergence among sister taxa by observing the resulting ML trees and node support values.

Treating the unknown ungulate, caiman and crocodile samples as a test case for species identification, we compared their sequences against diagnostic characters from our set of known species and included these samples in a reanalysis of ML phylograms. Unidentified samples were assigned to described species based on three criteria. First, an unidentified sample had to be included within a monophyletic group with bootstrap support exceeding 95%. Second, its sequence had to be at least 98% similar to the most common haplotype from a described taxon. Third, its inclusion into a particular monophyletic group had to preserve diagnostic character sites previously identified in that group. We considered these criteria sufficiently conservative to prevent type I errors (incorrectly attributing a sample to the wrong species).

Results

Primer selection and sequencing of degraded tissue and blood

The universal primers of Ivanova et al. (2007) reliably sequenced a 645 bp fragment of *COX1* mtDNA for all mammal and reptile species included in this study. Based on gel visualization, the COI-I cocktail worked optimally for duikers (*Cephalophus* spp.), the red river hog (*Potamochoerus porcus*) and the gray-cheeked mangabey (*Lophocebus albigena*). COI-2 worked best for the bongo, while a single primer pair from this cocktail [V(F,R)1d_t1] was optimal for the guenons (*Cercopithecus* spp.). Cocktail COI-3 worked for all species of crocodilians (Table 3).

Because monitoring of the bushmeat trade using DNA barcoding identification will often require amplification of degraded tissue samples, we evaluated the success of universal *COX1* primers on samples varying widely in age and curation method. The crocodile skin products proved difficult, with widely varying but generally low average DNA concentrations (13.5 ng/μl, SD = 13.9) and low purity (average 260:280 nm = 1.46, SD = 0.34). We were able to extract template DNA from two of the five leather products but only one of these produced a bidirectional *COX1* sequence (DNA yield = 16.35 ng/μl). We were able to obtain high-quality sequence data from a total of 179 of 204 samples tested (87.7%; Table 1). Of 43 samples that originally failed to sequence, seven were successfully sequenced following a second round of PCR amplification and 11 others produced sequences after a re-extraction of genomic DNA. Of the 25 samples that ultimately failed to yield sequences, nine produced visible PCR bands of the approximate molecular weight of the *COX1* fragment, but failed to sequence even after a second round of extraction and amplification. Final clean trace files of all sequenced samples (with the exception of six previously unidentified specimens, Table 1) were contributed to the BOLD database (Ratnasingham and Hebert 2007) at http://www.barcodinglife.org and sequences deposited on NCBI GenBank under accession numbers: GQ144467–GQ144639.

Correction of misidentified specimens

Based on diagnostic sites and branch placement on the ML tree, we discovered two discrepancies in species identification. One caiman sample (*M. niger*25) aligned unambiguously with *Caiman yacare* (Fig. 1a), but was labeled as belonging to *Melanosuchus niger*. It is unclear whether the specimen's label was switched in the lab or the collections, misidentified in the field or incorrectly accessioned; the YPM reptile staff is investigating (G.J. Watkins-Colwell, pers. comm.). One ungulate sample collected in the field and labeled as Peter's duiker (*C. callipygus*YF42) was identical to the most common haplotype of the bay duiker (*Cephalophus dorsalis*; Fig. 1b).

Intraspecific sequence variation

Within-species nucleotide diversity (average rate of nucleotide substitutions per site) in the *COX1* gene ranged from 0.0 to 1.92% (Table 5). The greatest amount of variability was observed in the Nile crocodile (*Crocodylus niloticus*). However, recent studies of this species using several mtDNA and nuDNA markers support a species-level division between northwestern and southeastern Africa (Hekkala 2004; Schmitz et al. 2003), though the geographic extent of each clade is not yet resolved. Eight of our Nile crocodile samples were captured in Gabon and, when we analyzed this geographic group separately, exhibited <0.01% nucleotide diversity (Table 5).

(a)

C. yacare32
92 M. niger25
C. yacare15
99 C. yacare33
C.c.crocodilusXyacare2
C. yacare7
99 C. c. chiapasius4
87 C. c. chiapasius3
96 C. c. chiapasius17
100 C. c. chiapasius16
80
C. latirostris5
C. latirostris6
Caiman1
100 C. latirostris22
C. latirostris18
C. latirostris23
97 C. latirostris21
M. niger24
M. niger9
100 M. niger26
M. niger10
100 P. palpebrosus11
P. palpebrosus29
100
P. trigonatus14
99 P. trigonatus31
P. trigonatus13
A. mississippiensis
0.01

(b)

C. leucogaster2157
C. leucogaster2220
Uniden11
Uniden10
C. callipygus2232
C. callipygus0351
C. callipygus0920
C. callipygus1453
C. callipygus1468
C. callipygus0402
C. callipygus0367
C. callipygus0410
C. callipygus2121
C. callipygus0910
Uniden15
C. callipygus2823
C. callipygus0003
C. callipygus0417
C. callipygus0404
C. callipygus0412
90 C. callipygus2153
C. callipygus0398
C. callipygus0406
91 C. callipygus0353
C. callipygus0399
86 C. callipygus0356
C. callipygus2246
C. callipygus0922
C. monticola0395
C. monticola0946
C. monticola0394
C. monticola0941
C. monticola5801
C. monticola0420
96 C. monticola2295
C. monticola2156a
C. monticola2265
C. monticola0947
C. monticola0421
C. monticola0409
C. monticola0872
C. monticola0396
C. monticola0918
Uniden2
C. monticola0912
C. monticola2110
C. monticola2148
C. monticola0002
C. monticola0363
C. monticola2222
C. monticola0397
C. monticola5802
C. monticola0416
C. dorsalis1010
C. callipygusYF42
C. dorsalis1031
C. dorsalisYF15
100 C. dorsalisYF115
C. dorsalisYF75
C. dorsalis0901
C. dorsalis0887
C. dorsalisYF111
C. nigrifrons2213
C. nigrifrons2101
100 C. nigrifrons2218
C. nigrifrons1004
C. nigrifrons0293
89 C. leucogaster2206
C. leucogaster2191
C. leucogaster0251
C. leucogaster0854
C. leucogaster0234
T. e. eurycerus9
0.01

(c)

O. tetraspis556
O. tetraspis557
O. tetraspis156
O. tetraspis554
O. tetraspis201
O. tetraspis146
O. tetraspis220
O. tetraspisM11
100 O. tetraspis102
O. tetraspis139
O. tetraspis551
O. tetraspis187
O. tetraspis127
O. tetraspis101
100
O. osborni92
O. osborni93
O. osborni120
O. osborni117
O. osborni108
100 O. osborni95
O. osborni46
O. osborni115
O. osborni94
O. osborni109
100
M. cataphractusG13
M. cataphractusG9
M. cataphractusG15
M. cataphractusG14
M. cataphractusG5
M. cataphractusC4
M. cataphractusC1
M. cataphractusG7
96 M. cataphractusC1M
M. cataphractusC2M
C. nicoticusG44
C. nicoticusG37
C. nicoticusG48
C. nicoticusG47
C. nicoticusG46
100 C. nicoticusG36
59 C. nicoticusG17
C. nicoticusG13
C. acutusM7
99 C. niloticus Congo
1USFWS
A. mississippiensis
0.01

◂ **Fig. 1** Maximum likelihood phylograms for **a** South American caimans, **b** African duikers, and **c** African crocodiles. Unknown, blind or misidentified samples are shown in *bold* type. Node support values are based on filtering the best maximum likelihood tree through 100 rapid bootstrap replicate trees. Log-likelihood and alpha shape parameter values, respectively, were **a** −1,949.953015 and 0.199779, **b** −1,888.9429 and 0.02, and **c** −1,965.181261 and 0.248764

We observed the second greatest amount of intraspecific diversity (1.18%) in 26 specimens of Peter's duiker (*Cephalophus callipygus*; Table 5). A subset of eight *C. callipygus* exhibited fixed differences at 13 out of 21 variable sites found within the larger group (Table 5) and formed a paraphyletic assemblage with moderate bootstrap support (<85%, Fig. 1b). When polymorphisms were evaluated separately for the two groups, diversity was 0.41% for the eight individuals and 0.05% in the remaining 18 individuals (Table 5). To reduce the likelihood that we had sequenced a pseudogene in any of the *C. callipygus* samples, we inspected the reading frames and found no premature stop codons. All substitutions were synonymous, thus not resulting in amino acid replacement.

Congeneric species differences

Based on the results of previous phylogenetic (Eaton et al. 2009) and morphological (Brochu 2007) studies, we treated African dwarf crocodiles (genus *Osteolaemus*) from Congo and Gabon as distinct species. We also considered the spectacled caiman (*Caiman crocodilus chiapasius*) and Yacare caiman (*C. yacare*) to be distinct taxa (Busack and Pandya 2001; Vasconcelos et al. 2006) and allowed for the possibility that our Nile crocodile (*Crocodylus niloticus*) samples represent distinct geographic lineages, based on the recent studies noted above.

The number of potentially diagnostic nucleotide characters between congeneric species ranged from 87 sites (13.5% of the sequenced fragment length) between mustached *(Cercopithecus cephus*) and crowned (*C. pogonias*) guenons, to only four characters (0.6% of the *COX1* fragment) between *Caiman yacare* and *Caiman crocodilus chiapasius* (Table 4). Nucleotide divergence values for these species pairs were 0.164 and 0.013, respectively (Table 5). *C. yacare* showed some evidence of paraphyly with *C. c. chiapasius* (bootstrap support = 87%), based on one *C. yacare* sample (C. yacare7), while the *C. c. chiapasius* samples themselves grouped more strongly (96%; Fig. 1a).

The average number of diagnostic sites for the 17 congeneric species pairs was 52.3 (8.1% of fragment length; SD = 18.4 sites), corresponding to an average genetic divergence of 0.098 (SD = 0.036). Position and character states of fixed nucleotide sites are presented in Table 4. Among the three orders compared, average congeneric nucleotide divergence was smallest in the crocodilians (0.065, SD = 0.03; excluding monotypic genera *Melanosuchus* and *Mecistops*) and highest among the three cercopithecoid primate species (0.144, SD = 0.029; excluding monotypic *Lophocebus*). Average pairwise divergence among the five duiker species (0.104, SD = 0.02) was intermediate between primates and crocodilians. Figure 2 demonstrates the range of intraspecific nucleotide diversity found in this study as compared to sequence divergence between congeneric species pairs.

Higher-order comparisons

Comparing sequences between genera revealed average divergence values ranging from 0.084 (*Caiman–Melanosuchus*) to 0.212 (*Cercopithecus–Lophocebus*) within families, to a high of 0.295 between Suidae (*Potamochoerus*) and Bovidae (*Cephalophus*; Table 6). The greatest divergence between Alligatoridae and Crocodylidae was *Paleosuchus* to *Osteolaemus* (0.244; Table 6).

Diagnosis of unknown samples

We determined the species identity of the unknown crocodilian and ungulate samples by matching ostensibly fixed character sites and by placement and likelihood support in the phylogenetic trees. The sequence for the unknown caiman (Caiman1) matched the most common haplotype of *Caiman latirostris* and fell within this monophyletic group with 100% node support (Fig. 1a). The museum sample identified as a hybrid caiman (*C. c. crocodilus* × *yacare*) grouped with *C. c. chiapasius* + C. yacare7, but with low bootstrap support (34%; Fig. 1a, node value not shown). This sample, however, matched with *C. yacare* at all diagnostic sites distinguishing the two species (Table 4). The crocodile skin handbag (1USFWS) matched closely (99.8% identity) with the one Nile crocodile sampled in the Congo. Differentiation of these two samples from Nile crocodiles collected in Gabon received high bootstrap support (99%, Fig. 1c). Comparing these putative groupings revealed 31 potentially fixed character differences and a sequence divergence of 0.053 between Nile crocodile clades (Table 4, 5). One unidentified ungulate sample (Uniden2) matched the most common *Cephalophus monticola* haplotype, while the remaining three (Uniden10, 11, 15) matched the most common *C. callipygus* haplotype. Phylogenetic placement of all four unknown ungulate samples was unambiguous (Fig. 1b). No unknown or misidentified primate samples were detected and therefore the phylogeny is not shown.

Table 4 Nucleotide positions and character states of diagnostic sites in a 645 bp *COX1* gene fragment. Twenty-one species from five tropical vertebrate mammals and reptiles are represented (see Table 1 for details on taxa and sample numbers and localities). Ambiguity codes denote that this position for a given species is not fixed and that the homologous position for another taxon is diagnostic

Artiodactyla	1	5	7	10	11	13	14	16	19	22	25	31	34	40	43	49	50	52	55	56	58	64	67	73	74	76	82	85	91	94	100	103	109	112	115	118	122	124	127	130	133	139	142	151	154	157	160	163	166
C. callipygus1	T	C	A	T	C	C	T	A	T	T	T	T	C	A	A	T	C	A	C	T	A	C	C	A	T	A	A	T	C	G	A	T	A	T	C	C	A	C	A	C	A	A	C	T	C	C	A	A	A
C. callipygus2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
C. monticola	.	.	.	C	.	T	.	.	.	Y	.	.	T	T	.	Y	.	.	.	.	.	.	.	R	C	G	.	.	.	.	.	C	.	C	.	.	G	T	.	.	.	.	.	.	.	T	.	.	.
C. dorsalis	.	.	G	C	.	.	C	.	C	.	.	C	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	C	.	A	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	T	.	.	G
C. nigrifrons	.	T	.	.	.	.	C	G	C	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
C. leucogaster	.	.	R	C	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	.	C	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	R
P. porcus	C	T	G	.	T	G	C	.	.	.	C	A	A	.	T	C	T	G	.	C	G	T	T	.	C	G	G	C	T	A	T	.	.	C	T	T	.	T	T	A	C	C	.	C	T	T	G	R	.

Artiodactyla	169	172	178	181	184	187	190	193	196	197	199	202	205	206	208	214	217	220	223	226	229	232	235	238	241	244	247	250	253	256	265	268	271	274	277	278	280	281	283	284	286	289	290	292	295	298	301	307	313	316	319	322	328	334	340
C. callipygus1	A	C	T	A	C	C	C	C	G	C	A	T	C	C	G	T	T	C	A	T	A	A	T	C	A	G	T	C	A	T	T	C	T	C	T	T	C	C	R	C	A	T	C	G	A	T	T	A	T	G	A	A	C	A	C
C. callipygus2	C	.	C	.	.	.	.	.	.	Y	.	.	.	.	R	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	.
C. monticola	T	Y	.	R	R	.	.	.	A	.	R	.	Y	.	A	.	.	.	.	C	G	.	.	T	.	A	.	T	.	.	.	.	.	.	.	.	T	.	.	T	.	C	.	A	.	C	A	T	C	A	.	.	.	C	.
C. dorsalis	T	T	.	G	T	.	.	.	A	.	.	C	T	.	A	.	.	.	.	.	.	.	.	.	G	A	C	.	.	.	.	.	.	.	.	.	.	.	G	T	G	C	T	A	G	.	C	T	C	A	.	.	.	C	.
C. nigrifrons	T	.	.	.	.	.	.	.	A	.	.	C	T	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	T	.	.	.	C	.	.	.	.	.	T	.	A	.	.	.	C	T
C. leucogaster	C	.	.	.	.	.	.	.	A	.	.	C	T	.	.	.	R	.	.	.	.	.	C	.	.	A	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	C	.	A	.	.	.	T	.	R	G	G	.	T	T
P. porcus	T	.	C	G	T	T	T	T	A	.	T	A	G	T	A	.	A	A	T	.	.	C	C	A	T	A	C	.	G	C	A	T	A	A	C	C	T	.	G	.	.	A	.	A	G	C	A	.	.	.	G	T	A	C	T

Artiodactyla	343	346	347	352	355	358	359	361	364	367	373	376	382	385	386	388	391	394	400	401	403	406	409	412	415	418	421	424	427	428	433	436	439	445	448	451	457	460	463	466	469	481	484	487	490	493	496	500	502	505	508	511	517	518	523	526
C. callipygus1	C	T	C	A	C	C	C	A	C	T	A	C	A	C	C	A	C	C	T	C	T	C	G	A	T	T	T	A	T	T	G	T	T	T	C	C	T	T	T	G	A	A	T	A	C	A	C	T	A	T	R	A	A	T	C	T
C. callipygus2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	A	C	.	.	T	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	R	R	.	.	.	.
C. monticola	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	C	.	A	T	R	.	.	.	C	.	C	.	A	C	.	C	.	T	.	C	.	.	.	.	C	.	T	.	T	C	.	W	.	.	.	.	Y	.
C. dorsalis	.	.	.	.	T	.	T	.	Y	.	.	.	.	.	.	.	T	.	.	.	A	.	A	.	.	.	.	G	.	.	A	.	.	.	T	.	C	.	.	A	.	.	.	.	.	.	.	.	G	.	.	.	.	C	T	.
C. nigrifrons	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	G	T	.	.	T	A	.	A	G	.	C	.	.	.	.	.	.	.	.	T	T	.	.	.	.	G	.	.	G	.	G	T	.	.	.	.	.	G	C	.	.
C. leucogaster	.	.	.	.	M	.	.	G	.	C	.	.	.	.	.	G	T	.	.	.	A	.	T	.	C	.	T	.	.	Y	.	.	.	C	T	T	C	.	Y	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
P. porcus	A	C	T	T	G	T	.	.	.	C	C	T	T	T	T	.	A	T	C	.	A	T	T	.	A	A	A	.	C	.	.	.	C	.	T	.	C	.	C	A	.	G	C	.	.	.	A	C	.	.	C	.	.	C	T	A

Artiodactyla	529	532	536	538	541	544	547	550	553	556	557	562	571	574	577	578	581	586	589	592	595	596	598	601	604	613	619	625	628	631	635	637	643
C. callipygus1	C	A	T	A	T	C	C	T	T	A	C	A	T	A	A	T	C	A	C	A	T	T	A	T	A	C	A	A	C	A	C	T	G
C. callipygus2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	Y	G	.	.	.
C. monticola	Y	.	C	.	C	.	T	.	C	.	T	.	Y	.	.	.	T	.	.	.	C	.	.	C	.	T	G	.	A	G	.	.	A
C. dorsalis	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	C	.	T	.	.	.	.	.	C	G	T	.	.	T	.	.	.	A
C. nigrifrons	.	G	.	.	.	.	.	C	.	.	T	.	.	.	.	C	.	T	.	.	.	C	G	C	.	.	.	.	T	G	.	.	A
C. leucogaster	.	.	.	.	C	.	.	.	C	.	T	.	C	G	.	C	T	T	T	.	.	C	.	C	.	T	.	.	T	.	.	C	A
P. porcus	.	.	C	T	A	G	T	G	A	T	T	G	.	C	G	C	T	.	T	C	.	C	.	C	.	T	.	T	A	.	Y	A	C

Primates	1	4	5	7	10	13	16	19	20	22	34	35	37	43	45	46
C. cephus	A	T	C	G	C	A	A	C	G	T	A	A	C	C	C	T
C. pogonias	.	.	.	A	.	.	.	.	K	G	G	G	T	.	.	A
C. nictitans	R	.	.	A	.	.	R	Y	.	.	.	G	.	.	.	.
L. albigena	C	C	T	A	T	G	G	T	.	.	.	G	T	T	T	R

Primates	49	55	58	59	61	64	70	73	74	76	79	82	85	90	91	92	94	95	97	100	102	106	109	110	111	112	115	118	121	124	127	130	142	143	145	151	154	160	163	169	175	184	190	193	197	199	202	205	206	208	214	217
C. cephus	T	T	T	C	T	T	T	A	T	A	C	A	C	G	T	C	G	T	G	T	G	C	T	G	T	T	C	T	T	C	A	A	C	A	C	T	C	A	C	C	G	R	G	C	T	A	G	C	T	A	T	T
C. pogonias	.	.	.	Y	C	C	Y	G	C	.	T	.	T	.	.	.	.	C	A	C	.	.	C	A	.	C	T	.	C	.	.	.	.	T	G	.	.	.	G	.	.	A	T	A	.	C	.	A	.	Y	G	.
C. nictitans	.	.	.	.	.	.	.	G	C	G	T	.	.	.	.	A	.	.	C	.	T	.	A	.	.	.	.	.	.	.	G	.	T	.	T	.	.	.	T	.	A	G	.	.	C	R	C	.	.	.	.	.
L. albigena	C	C	C	.	.	.	.	.	C	G	A	G	.	A	C	T	.	C	A	C	A	.	C	A	C	C	.	C	C	T	.	G	.	G	.	C	T	.	.	T	A	C	A	T	C	.	A	T	.	.	C	C

Primates	220	229	235	241	247	250	253	262	265	268	271	274	277	280	281	283	286	287	292	294	299	302	304	310	313	322	328	331	334	343	346	347	350	358	359	361	367	373	376	379	382	385	386	388	391	400	403	406	407	415	416	424	428	430
C. cephus	C	A	C	C	T	T	A	G	T	T	A	A	T	C	C	G	G	T	A	C	G	A	A	G	T	A	T	G	A	T	T	T	G	C	C	C	T	A	C	T	G	C	T	G	T	C	C	T	T	A	G	T	T	A
C. pogonias	T	G	T	T	C	C	G	.	C	.	C	C	C	.	.	A	A	C	G	.	.	.	C	A	.	.	C	A	G	G	C	.	.	.	.	.	C	G	.	.	A	T	.	R	.	.	.	.	.	.	.	C	C	.
C. nictitans	Y	.	T	T	.	C	.	.	T	C	T	C	.	T	.	.	A	C	.	.	A	G	R	.	.	T	.	A	.	C	.	.	.	T	.	.	C	.	.	C	A	.	.	A	A	.	.	.	.	.	A	.	C	.
L. albigena	T	.	T	T	C	.	.	A	C	C	C	C	.	.	T	A	A	C	.	M	A	G	.	A	C	.	.	A	.	C	.	C	T	.	T	T	C	.	T	C	A	.	C	A	C	T	T	C	C	C	A	C	C	G

Primates	433	436	442	445	448	454	460	463	466	472	475	478	487	490	493	496	500	502	505	508	514	515	518	523	526	529	532	535	538	541	547	548	556	557	559	562	565	568	571	574	577	589	592	601	604	607	620	621	622	628	631	634	637
C. cephus	G	C	C	T	C	T	T	C	G	T	C	T	A	Y	A	A	T	R	T	A	T	G	C	T	A	A	C	A	A	T	C	C	C	T	G	T	C	C	T	T	G	C	A	C	C	C	A	T	T	G	A	C	C
C. pogonias	.	G	T	C	.	C	C	.	.	C	T	.	.	C	G	C	Y	G	.	C	.	.	.	.	.	.	.	.	.	C	T	T	T	C	A	C	T	T	.	.	A	T	C	T	.	T	.	C	.	A	G	T	T
C. nictitans	A	.	.	C	T	.	.	T	.	C	.	.	.	.	.	C	.	.	C	T	.	.	T	C	.	R	T	.	.	C	.	.	.	.	.	.	.	T	C	C	A	.	C	.	T	T	.	C	C	A	.	.	T
L. albigena	A	T	T	C	T	.	C	.	A	C	T	A	G	C	.	C	C	.	.	C	A	A	T	.	G	.	.	C	G	C	T	.	.	C	A	C	.	.	C	.	.	T	C	.	T	T	G	.	.	A	.	.	T

Crocodylidae	5	7	10	16	25	29	31	36	37	40	43	50	52	58	59	61	67	73	76	82	85	88	91	92	94	100	106	109	118	121	124	127	145	151	154	160	163	172	175	178	193	196	199	202	206	211	214	217	223	226
C. niloticus GA	T	G	T	T	C	G	C	T	A	A	C	A	A	A	T	A	A	G	C	G	A	C	C	T	C	A	C	A	T	T	T	Y	C	T	C	A	T	T	G	C	T	A	A	C	T	A	T	A	C	C
C. niloticus CO	.	.	.	.	.	.	.	.	.	G	.	.	G	.	.	G	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.
M. cataphractus	.	A	C	.	.	.	T	W	G	.	A	C	.	C	C	.	G	A	.	.	.	T	A	.	T	.	.	.	C	.	C	T	.	C	T	.	A	C	A	.	C	.	.	A	C	.	.	.	A	.
O. tetraspis	.	A	C	C	A	R	T	.	.	R	A	C	.	T	Y	.	G	A	A	A	T	.	T	C	T	.	T	G	.	C	.	C	T	.	.	G	A	C	A	.	C	G	.	G	C	.	C	.	.	T
O. osborni	C	A	.	C	.	.	T	.	.	.	A	C	.	T	C	.	.	A	.	A	T	.	G	.	T	G	T	.	.	.	.	A	.	C	.	.	A	C	A	T	C	.	G	A	C	.	.	.	A	T

Crocodylidae	235	238	241	247	253	263	265	266	268	271	274	282	284	287	289	290	292	295	298	304	307	313	316	319	322	325	328	334	337	343	346	355	358	361	367	370	373	385	386	388	391	394	403	406	415	418	421	424	427	428	430	433	439	448	451	454
C. niloticus GA	C	T	C	C	A	T	G	C	A	C	A	C	C	C	T	C	C	T	C	T	T	T	G	T	C	C	A	A	C	G	G	A	C	A	C	C	A	C	C	C	T	T	T	C	G	G	A	C	C	C	T	A	T	C	C	G
C. niloticus CO	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	.	.	.	.	.	G	.	.	.	A	.	.	.	.	.	.	.	.	T	.	C	.	.	A	A	.	.	.	.	C	.	.	T	T	.
M. cataphractus	.	C	A	.	G	C	A	.	A	T	.	.	T	A	T	A	.	.	.	A	T	C	.	.	C	.	T	A	A	.	G	.	A	.	.	.	.	.	.	A	.	C	C	T	A	.	C	A	.	T	A	G	.	T	.	C
O. tetraspis	T	C	A	T	.	.	A	T	.	G	C	T	T	.	.	.	A	.	.	C	.	A	C	C	T	A	.	T	T	A	A	G	T	G	.	T	G	T	T	A	.	C	.	T	.	T	.	A	T	T	A	G	C	.	.	A
O. osborni	T	.	A	T	.	C	A	T	.	A	T	.	T	.	.	.	A	.	.	C	C	C	C	C	T	A	.	.	T	.	A	G	T	.	T	T	G	T	T	A	C	C	.	T	A	T	.	A	T	T	A	.	C	.	.	A

Table 4 continued

	457	458	460	463	472	476	478	481	487	496	499	505	508	511	514	517	518	520	523	526	532	535	539	541	544	545	547	550	553	559	560	562	565	568	571	577	578	580	583	586	589	596	598	604	607	610	613	616	619	622	625	631	634	640	641
	T	A	C	C	A	G	A	G	A	G	C	T	G	A	T	T	C	G	T	A	C	C	C	G	C	T	A	A	C	A	G	T	A	A	T	A	C	A	A	T	T	T	G	C	C	C	T	C	C	A	A	A	C	C	C
	.	.	.	.	C	.	.	A	.	A	.	.	.	R	.	.	.	A	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	T	C	C	C	T	.	T	.	.	.	T	.	.	.	.	.	.
	.	.	T	.	.	T	.	A	G	A	A	.	.	.	.	G	T	A	C	T	A	.	.	.	.	G	C	G	A	.	.	.	.	T	.	G	.	G	T	.	C	.	.	.	.	.	.	.	.	.	.	.	T	.	T
	C	C	T	T	.	T	G	.	.	A	A	C	A	.	C	A	.	A	C	G	A	.	.	.	.	G	C	.	G	.	R	G	.	.	.	.	T	.	Y	.	C	.	A	T	.	T	C	T	T	T	.	G	T	.	.
	C	.	T	T	.	T	.	.	.	A	A	.	.	G	.	A	T	A	C	.	A	T	.	A	A	G	C	T	A	G	.	.	.	.	.	G	T	.	C	.	C	C	A	T	.	T	.	.	T	.	G	.	T	T	.

Alligatoridae	7	13	22	23	25	28	29	31	41	43	49	51	52	58	82	85	91	94	100	106	112	118	124	127	130	136	139	142	143	145	154	166	170	181	187	202	205	206	208	217	229	238	247	263	265	268	277	286	292	298	301	304
C. yacare	C	T	A	A	C	R	G	C	G	A	A	T	T	C	G	C	C	C	A	C	C	C	T	T	C	C	C	C	A	T	T	A	G	C	T	C	C	T	A	A	G	A	C	C	T	A	T	A	T	C	T	C
C. latirostris	.	.	.	.	T	A	.	.	R	.	C	W	.	.	.	.	.	.	G	.	.	.	.	C	.	T	T	T	.	.	C	.	A	Y	.	T	.	.	.	G	.	C	T	T	A	G	C	.	.	.	.	T
C. c. chiapasius	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
M. niger	A	.	G	.	T	A	.	.	.	.	.	.	C	.	A	.	T	A	.	.	.	T	.	C	.	.	T	T	.	C	.	.	A	.	.	T	A	C	G	.	.	.	.	T	A	C	.	.	A	.	.	.
P. palpebrosus	G	C	G	G	.	A	K	.	.	C	.	.	C	T	A	A	.	A	R	Y	.	T	C	C	T	.	.	T	.	.	.	.	.	.	C	R	.	.	.	R	A	G	.	.	A	C	C	G	C	T	C	.
P. trigonatus	A	C	G	G	T	A	T	T	K	C	.	.	C	.	A	A	.	A	.	.	T	T	C	C	.	.	.	.	G	C	C	.	.	.	.	.	.	C	.	.	A	.	.	.	A	C	C	.	C	T	C	T

	310	313	316	319	328	340	343	346	352	359	364	370	376	379	382	385	386	388	391	394	400	403	406	424	427	430	433	439	442	445	451	454	457	460	463	472	481	496	499	502	508	511	514	518	520	526	535	539	541	544	545	547	550	553	556	557	562	565
C. yacare	A	A	A	T	A	C	C	T	C	C	T	C	A	C	A	T	T	A	A	C	A	A	T	T	C	A	A	C	C	C	A	A	C	C	C	C	A	A	A	A	T	A	T	T	A	A	C	T	G	A	G	C	A	G	A	C	C	T
C. latirostris	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	T	C	C	.	G	G	T	T	.	.	.	T	T	.	.	G	.	.	G	.	G	.	.	G	.	T	C	A	.	T	A	.	A	G	.	T	C
C. c. chiapasius	.	.	.	.	.	.	.	.	.	.	Y	.	G	.	.	C	.	.	.	.	.	.	.	.	T	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.
M. niger	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	C	.	.	.	.	.	T	.	C	.	C	.	T	.	T	T	G	T	.	T	A	.	.	.	.	.	.	C	C	.	.	.	C	T	.	T	A	.	A	.	T	.	C
P. palpebrosus	.	G	G	C	.	T	T	C	T	A	C	.	.	A	.	C	C	.	C	.	C	.	C	C	.	R	C	T	.	.	.	.	.	T	.	.	G	.	C	.	.	.	G	.	.	.	A	.	A	C	T	.	C	A	.	.	T	C
P. trigonatus	G	G	G	C	G	T	.	C	T	.	C	.	.	.	.	C	C	T	C	T	C	.	C	C	.	.	C	.	.	.	.	.	.	T	.	.	.	C	C	.	C	.	A	.	.	G	.	C	A	C	T	.	C	A	.	.	T	.

	571	574	583	596	598	601	607	613	616	619	622	625	628	638
C. yacare	T	T	T	T	A	C	T	K	C	T	G	A	A	A
C. latirostris	.	C	.	.	.	.	C	C	.	C	A	.	.	.
C. c. chiapasius	.	.	.	.	.	.	.	.	.	.	.	.	.	.
M. niger	.	C	.	.	.	.	A	C	T	A	A	.	.	G
P. palpebrosus	.	.	C	C	.	T	C	.	.	C	A	.	G	.
P. trigonatus	C	.	C	Y	G	.	C	.	.	A	A	G	.	.

Discussion

Methodological considerations

Universal primer cocktails compiled by Ivanova et al. (2007; Table 3) successfully amplified the targeted barcode region for all eleven genera included in this study. High quality sequence data were generated from older specimens and those stored under suboptimal conditions, including dried tissue and blood maintained at room temperature for more than 20 years. DNA extraction and sequencing of confiscated crocodile leather products proved difficult, however, with only one of five skins identified. The tanning process used to preserve the crocodile leather, although apparently minimal, likely contributed to DNA degradation. Following additional optimization of PCR thermal profiles and reaction volumes, and re-extraction of difficult samples, we were ultimately unable to produce sequences from 12.3% of tissue samples. Tissue stored in buffered solution required the least procedural modifications for amplification and sequencing. Specimens stored as dried tissue or dried blood on gauze stored longer than 15 years were the most unreliable in producing high quality sequences. Problems associated with amplification of degraded DNA are well known (Deagle et al. 2006; Shapiro et al. 2004; Vuissoz et al. 2007) and it is unlikely that recovering full-length *COX1* barcodes ($\sim$650 bp) from processed animal products will be routinely successful. Hajibabaei et al. (2006) recognized this limitation and designed compatible primers to sequence short *COX1* barcode fragments (221 and 134 bp) in Lepidoptera and Hymenoptera. Appreciating the importance of a varied molecular toolkit, we are using the taxa included in the current study to develop and test a suite of mini-barcode primers to be paired with the existing universal $5'$ primer set of Ivanova et al. (2007) to sequence fragments in the range of 150–350 bp. Initial trials on samples with low probabilities of producing full barcode sequences (leather products and museum specimens) are promising and appear to capture sufficient numbers of diagnostic characters for species identification (data not shown). The primer sets and our results will be reported in a subsequent publication.

Intraspecific variation

Intraspecific variation in the barcode region was low and generally fell within reported ranges of within-species mtDNA divergence (rarely exceeding 2%, Avise 2000). In this study, we encountered two species with genetic distances approaching the expected upper range of intraspecific variation. We observed the greatest amount of nucleotide diversity (1.92%) in the Nile crocodile (*C. niloticus*). This crocodilian is now considered to be two genetically distinct lineages based on multiple lines of evidence (Hekkala 2004; Schmitz et al. 2003). Although the previous studies did not include *COX1*, we sequenced the

Table 5 Inter- and intraspecific nucleotide differences in the *COX1* gene for three orders of tropical mammals and reptiles. The lower triangular matrix quantifies the number of fixed, diagnostic nucleotide positions between species pairs within each order; values in the upper matrix represent pairwise nucleotide divergence (Tamura-Nei) across the 645 bp fragment of *COX1*. Diagonal values (in *bold*) are the average number of base substitutions per site (×100) for intraspecific comparisons

	Crocodilia	1	2	3	4	5	6	7	8	9	10	11
1.	*M. niger* (4)	**0.000**	0.132	0.129	0.083	0.082	0.086	0.229	0.222	0.212	0.243	0.232
2.	*P. palpebrosus* (2)	73	**0.310**	0.066	0.120	0.110	0.118	0.229	0.214	0.213	0.230	0.254
3.	*P. trigonatus* (3)	74	37	**0.100**	0.133	0.110	0.119	0.217	0.212	0.208	0.235	0.256
4.	*C. latirostris* (6)	49	66	75	**0.100**	0.078	0.082	0.225	0.208	0.211	0.235	0.247
5.	*C. yacare* (5)	48	62	66	45	**0.230**	0.013	0.212	0.217	0.208	0.235	0.246
6.	*C. c. chiapasius* (4)	50	64	67	46	4	**0.420**	0.225	0.227	0.216	0.244	0.249
7.	*M. cataphractus* (10)	126	125	121	124	122	126	**0.030**	0.136	0.136	0.135	0.152
8.	*C. niloticus*-Gabon (8)[a]	123	118	119	116	123	128	79	**0.070**	0.053	0.173	0.182
9.	*C. niloticus*-Congo (2)[a]	114	113	112	114	113	118	76	31	**0.160**	0.178	0.205
10.	*O. osborni* (10)	131	125	128	128	129	134	77	96	95	**0.000**	0.098
11.	*O. tetraspis* (14)	131	139	111	138	136	139	85	100	106	57	**0.280**

	Artiodactyla	12	13	14	15	16	17	18	19
12.	*C. nigrifrons* (5)	**0.00**	0.103	0.072	0.131	0.082	0.082	0.085	0.327
13.	*C. dorsalis* (9)	61	**0.06**	0.091	0.130	0.097	0.098	0.098	0.289
14.	*C. leucogaster* (7)	41	51	**0.34**	0.124	0.093	0.092	0.099	0.290
15.	*C. monticola* (24)	74	69	66	**0.53**	0.111	0.109	0.119	0.305
16.	*C. callipygus* (23)	44	46	46	54	**1.18**	–	–	0.264
17.	*C. callipygus* grp1 (15)	50	56	52	58	–	**0.05**	0.024	0.318
18.	*C. callipygus* grp2 (8)	49	55	54	59	–	13	**0.41**	0.328
19.	*P. porcus* (7)	140	128	125	120	124	132	134	**0.13**

	Primates	20	21	22	23
20.	*C. cephus* (7)	**0.150**	0.111	0.164	0.240
21.	*C. nictitans* (7)	62	**0.540**	0.157	0.203
22.	*C. pogonias* (6)	87	72	**0.340**	0.194
23.	*L. albigena* (8)	118	99	102	**0.110**

[a] Analyzed as a single species, intraspecific variability of *C. niloticus* was 1.92%

12S rRNA gene for Nile crocodiles from Congo and Gabon (Eaton, unpublished data) to compare with the lineages described by Schmitz et al. (2003) and Hekkala (2004). We found that Gabon crocodiles grouped strongly (ML bootstrap support = 96%) with the assemblage including South Africa, Madagascar, Kenya, Zimbabwe, Sudan and Egypt (GenBank accession numbers: AY195943, AY195945, AY195946, AY195950–AY195955). The Congo specimen and handbag aligned with samples from Senegal, Chad, the Gambia and Mauritania (AY195944, AY195947–AY195949, AY195956, AY195957) with similar bootstrap support (96%). Although it is not possible to compare Nile crocodile *COX1* barcodes directly to clades produced by the 12S rRNA gene, we infer that the differentiation of barcode samples from Congo and Gabon corroborate the earlier studies and represent a geographic division between the former conspecific crocodiles.

Peter's duiker (*Cephalophus callipygus*) was the second species with relatively high levels of intraspecific diversity. A subgroup of eight individuals displayed putative fixed characters at more than half of the variable sites for this species and formed a paraphyletic assemblage (Fig. 1b). Although NuMts (mitochondrial pseudogenes in the nucleus) are not easy to diagnose in cases of noncoding DNA (Kolokotronis et al. 2007), they often contain premature stop codons and frame-shift mutations (e.g. Chung and Steiper 2008; Lemos et al. 1999). Our sequence data contained neither, suggesting NuMts are unlikely to explain the observed sequence variation and phylogenetic pattern. In an extensive DNA barcode survey of birds, Kerr et al. (2007) were able to detect a low prevalence of pseudogenes due to their generally reduced size (100–200 bp) and, presumably, disrupted reading frames. Hybridization with another duiker species is also an unlikely explanation for

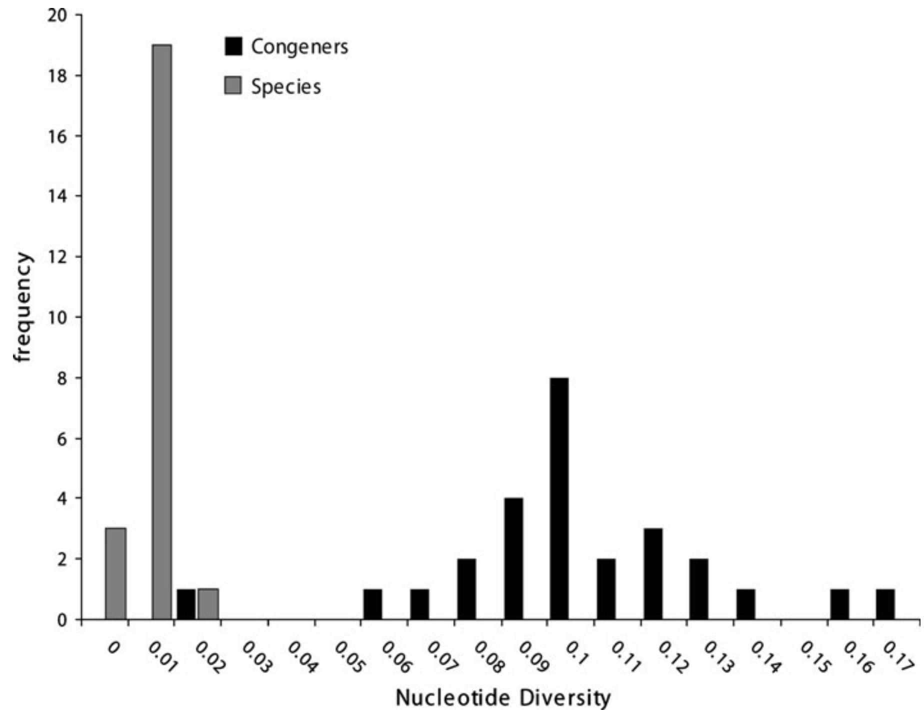

Fig. 2 Frequency histogram of intraspecific *COX1* sequence diversity ("species") compared to nucleotide divergence between pairs of congeneric sister species analyzed in this study. Only sister species *Caiman yacare* and *C. crocodilus chiapasius* had a divergence value overlapping the range of intraspecific nucleotide polymorphisms. The taxonomy of these two species is still being investigated

Table 6 Average nucleotide divergence (Tamura-Nei) between genera within orders for a 645 bp fragment of *COX1* mtDNA

Order (Family)	Genera compared	Divergence
Artiodactyla	*Cephalophus–Potamochoerus*	0.295
Primates	*Cercopithecus–Lophocebus*	0.212
Crocodilia	*Caiman–Crocodylus*	0.214
	Caiman–Mecistops	0.221
	Caiman–Osteolaemus	0.243
	Melanosuchus–Crocodylus	0.217
	Melanosuchus–Mecistops	0.229
	Melanosuchus–Osteolaemus	0.238
	Paleosuchus–Crocodylus	0.212
	Paleosuchus–Mecistops	0.223
	Paleosuchus–Osteolaemus	0.244
(Alligatoridae)	*Caiman–Melanosuchus*	0.084
	Caiman–Paleosuchus	0.118
	Melanosuchus–Paleosuchus	0.131
(Crocodylidae)	*Crocodylus–Osteolaeumus*	0.184
	Crocodylus–Mecistops	0.136
	Osteolaemus–Mecistops	0.143

these findings, as our sample set included all sympatric duikers found in the Republic of Congo with the exception of the yellow-backed duiker (*C. silvicultor*), a much larger species not likely to interbreed with *C. callipygus*. Our observation of cryptic genetic variation in Peter's duiker is corroborated by at least one other molecular study which found two major clusters of *C. callipygus* in forests of central Gabon, although the authors offered no discussion of this observation (van Vliet et al. 2008).

Because inference on divergence was based on a single mitochondrial locus within a sympatric population of *C. callipygus*, we cannot reject the possibility that our observations may conflict with the genome tree or true phylogeny for this species due to introgressive hybridization or incomplete lineage sorting (Funk and Omland 2003). Such concerns are reduced in the case of the Nile crocodile, for which analysis of additional genes (including nuclear loci) corroborate those of *COX1* in refuting an imperfect species taxonomy (Hekkala 2004; Schmitz et al. 2003). However, since our results are reported for allopatric populations of Nile crocodiles, introgression may in fact be an important consideration in future analyses if single-gene phylogenies are constructed from specimens sampled in putative hybrid zones.

Interspecific variation

While a character-based approach has been recommended as an improvement over distance-based thresholds for species identification (Rach et al. 2008; Rubinoff et al. 2006), the interspecific differences we present here are based on both distance metrics and diagnostic nucleotide characters. Mean congeneric nucleotide divergence in our study was 9.8% and the average number of fixed characters was 52.3. Variation between congeneric species pairs substantially exceeded levels of intraspecific polymorphism and fell within the range of genetic divergences

reported for a wide range of vertebrate taxa. Hebert et al. (2003b) found that 93.8% of vertebrate congeners had a sequence divergence between 4 and 32% (mean = 9.6%) for the same region of *COX1* while Johns and Avise (1998) reported average genetic distances for congeneric mammal and reptile species generally exceeded ~3% in the mitochondrial cytochrome *b* gene. One exception observed here was between *Caiman yacare* and *C. crocodilus chiapasius*, whose taxonomy is still under debate, but for which a small number of diagnostic characters ($n = 4$) may reliably identify individuals to species (Table 4).

Identification of unknown and misidentified samples

The primary intent of this work was to evaluate whether *COX1* barcoding would serve as a reliable means to identify wildlife species sampled during bushmeat monitoring or while investigating trade in embargoed wildlife products. Matching homologous diagnostic sites and phylogenetic methods unequivocally assigned all unknown samples to the correct species, including caiman, crocodiles and ungulates. Barcode sequences also identified errors in identification or labeling that occurred during field collection or in museum accessioning.

Conclusions

Our results suggest that with minimal effort and simple refinements to DNA extraction and PCR protocols, accurate barcode sequence data can be obtained from most wildlife products encountered in bushmeat monitoring programs and wildlife investigations. Sequencing shorter barcode fragments should increase the success of working with degraded DNA samples and we have begun work to modify universal primers that will contribute to this effort for a wide range of taxa. Strong phylogenetic support and the high frequency of fixed character states between closely related taxa offer convincing evidence that *COX1* barcoding gene will reliably diagnose many common African and neotropical bushmeat species. Bushmeat monitoring and investigations of wildlife commercialization and trade are likely to benefit from this molecular approach, especially when sampling from semi-processed products (e.g. from urban or import markets), when working with products that are difficult to identify (e.g. bird and reptile eggs, fish species, skins), or when focusing in areas containing morphologically cryptic species. In an effort to reduce inevitable errors while conducting field research and in museum curation, mtDNA barcoding offers a simple, low-cost and accurate method for verifying species identities. Although additional independent lines of evidence are needed to substantiate the levels of divergence observed in Peter's duiker (*C. callipygus*), barcodes may have highlighted a novel evolutionary lineage worthy of further investigation. Finally, generating a database of barcode sequences for tropical wildlife will offer researchers, conservationists and managers an effective tool for more precisely delineating the extent, range and genetic diversity of species of concern.

Acknowledgments This work was supported by the Alfred P. Sloan Foundation and the Richard Lounsbery Foundation. The National Science Foundation and the American Museum of Natural History's Research Experience for Undergraduates Program supported the laboratory work of GLM. The AMNH sponsored Aritra Datta and Arlene Amador to extract and sequence the USFWS crocodile skin products and we thank them for their efforts. Field support was provided to MJE by the Wildlife Conservation Society's Congo and Gabon programs, the National Geographic Society, the Rufford Foundation, Lincoln Park Zoo's Asia & Africa Fund and the MacArthur Program of the University of Minnesota. MJE thanks Paul and Sarah Elkan (WCS-Congo), Debora Pires, and Congo field assistants Yamba Flavien, Bienvenu Kimbembe and Rufin Lekana. New World crocodilian samples were collected by Peter Brazaitis, Carlos Yamashita (IBAMA, Brazil) and George Rebelo (INPA, Brazil) and provided by G. J. Watkins-Colwell of the Peabody Museum. Ellen Bean and three anonymous reviewers greatly improved the clarity and scope of this manuscript.

References

Albrechtsen L, David WMA, Paul JJ et al (2007) Faunal loss from bushmeat hunting: empirical evidence and policy implications in Bioko island. Environ Sci Policy 10:654–667

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

Baker CS (2008) A truer measure of the market: the molecular ecology of fisheries and wildlife trade. Mol Ecol 17:3985–3998

Baker CS, Cipriano F, Palumbi SR (1996) Molecular genetic identification of whale and dolphin products from commercial markets in Korea and Japan. Mol Ecol 5:671–685

Baker CS, Dalebout ML, Lento GM et al (2002) Gray whale products sold in commercial markets along the pacific coast of Japan. Mar Mamm Sci 18:295–300

Bennett EL, Blencowe E, Brandon K et al (2007) Hunting for consensus: reconciling bushmeat harvest, conservation, and development policy in west and central Africa. Conserv Biol 21:884–887

Bickford D, Lohman DJ, Sodhi NS et al (2007) Cryptic species as a window on diversity and conservation. Trends Ecol Evol 22:148–155

Birstein VJ, Doukakis P, Sorkin B et al (1998) Population aggregation analysis of three caviar-producing species of sturgeons and implications for the species identification of black caviar. Conserv Biol 12:766–775

Brochu CA (2007) Morphology, relationships, and biogeographical significance of an extinct horned crocodile (Crocodylia, Crocodylidae) from the Quaternary of Madagascar. Zool J Linn Soc 150:835–863

Brower AVZ (1999) Delimitation of phylogenetic species with DNA sequences: a critique of Davis and Nixon's population aggregation analysis. Syst Biol 48:199–213

Busack SD, Pandya S (2001) Geographic variation in *Caiman crocodilus* and *Caiman yacare* (Crocodylia: Alligatoridae): systematic and legal implications. Herpetologica 57:294–312

Chomel BB, Belotto A, Meslin FX (2007) Wildlife, exotic pets, and emerging zoonoses. Emerg Infect Dis 13:6–11

Chung W, Steiper M (2008) Mitochondrial *COII* introgression into the nuclear genome of *Gorilla gorilla*. Int J Primatol 29:1341–1353

Davis JI, Nixon KC (1992) Populations, genetic variation, and the delimitation of phylogenetic species. Syst Biol 41:421–435

Deagle BE, Eveson JP, Jarman SN (2006) Quantification of damage in DNA recovered from highly degraded samples–a case study on DNA in faeces. Front Zool 3:11

DeSalle R (2006) Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff. Conserv Biol 20:1545–1547

Eaton MJ (2002) Subsistence wildlife hunting in a multi-use forest of the Republic of Congo: monitoring and management for sustainable harvest (Unpublished MS thesis), University of Minnesota, St. Paul, p 90

Eaton MJ (2006) Ecology, conservation and management of the Central African dwarf crocodile (*Osteolaemus tetraspis*), a progress report pp 84–95. Crocodiles: Proceedings of the 18th working meeting of the IUCN-SSC Crocodile Specialist Group, IUCN, Gland, Switzerland

Eaton MJ, Barr B (2005) Regional Africa report: Lac Tele, Rep. of Congo. Crocodile Specialist Group Bull 24:18–20

Eaton MJ, Martin AP, Thorbjarnarson J et al (2009) Species-level diversification of African dwarf crocodiles (Genus *Osteolaemus*): a geographic and phylogenetic perspective. Mol Phylogenet Evol 50:496–506

Estes RD (1991) The behavior guide to African mammals: including hoofed mammals, carnivores, primates. The University of California Press, Berkeley

Fitzhugh K (2006) DNA barcoding: an instance of technology-driven science? Bioscience 56:462–463

Folmer O, Black M, Hoeh W et al (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu Rev Ecol Evol Syst 34:397–423

Hajibabaei M, Smith MA, Janzen DH et al (2006) A minimalist barcode can identify a specimen whose DNA is degraded. Mol Ecol Notes 6:959–964

Hebert PDN, Cywinska A, Ball SL et al (2003a) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci 270:313–321

Hebert PDN, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B Biol Sci 270:S96–S99

Hebert PDN, Penton EH, Burns JM et al (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci USA 101:14812–14817

Hekkala E (2004) Conservation genetics at the species boundary: case studies from African and Caribbean crocodiles (genus *Crocodylus*) (Unpublished Ph.D. thesis), Columbia University, New York, p 142

Ivanova NV, Zemlak TS, Hanner RH et al (2007) Universal primer cocktails for fish DNA barcoding. Mol Ecol Notes 7:544–548

Janzen DH, Hajibabaei M, Burns JM et al (2005) Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Phil Trans R Soc Lond B 360:1835–1845

Johns GC, Avise JC (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. Mol Biol Evol 15:1481–1490

Kaila L, Stahls G (2006) DNA barcodes: Evaluating the potential of *COX1* to diffentiate closely related species of *Elachista* (Lepidoptera: Gelechioidea: Elachistidae) from Australia. Zootaxa 1170:1–26

Kerr KCR, Stoeckle MY, Dove CJ et al (2007) Comprehensive DNA barcode coverage of North American birds. Mol Ecol Notes 7:535–543

Kingdon J (1997) The Kingdon field guide to African mammals. Academic Press, London

Kolokotronis SO, MacPhee RDE, Greenwood AD (2007) Detection of mitochondrial insertions in the nucleus (NuMts) of Pleistocene and modern muskoxen. BMC Evol Biol 7:67

Lanave C, Preparata G, Saccone C et al (1984) A new method for calculating evolutionary substitution rates. J Mol Evol 20: 86–93

Lemos B, Canavez F, Moreira M (1999) Mitochondrial DNA-like sequences in the nuclear genome of the opossum genus *Didelphis* (Marsupialia : Didelphidae). J Hered 90:543–547

Marko PB, Lee SC, Rice AM et al (2004) Mislabeling of a depleted reef fish. Nature 430:309–310

Martin AP (1991) Application of mitochondrial DNA sequence analysis to the problem of species identification of sharks. NOAA NMFS 115:53–59

McAliley LR, Willis RE, Ray DA et al (2006) Are crocodiles really monophyletic? Evidence for subdivisions from sequence and morphological data. Mol Phylogenet Evol 39:16–32

Milius S (2005) Bushmeat on the Menu: untangling the influences of hunger, wealth, and international commerce. Science News 167:138

Milner-Gulland EJ, Bennett EL, The SCB 2002 Annual Meeting Wild Meat Group (2003) Wild meat: the bigger picture. Trends Ecol Evol 18:351–357

Moura T, Silva MC, Figueiredo I et al (2008) Molecular barcoding of north-east Atlantic deep-water sharks: species identification and application to fisheries management and conservation. Mar Freshwater Res 59:214–223

Neigel J, Domingo A, Stake J (2007) DNA barcoding as a tool for coral reef conservation. Coral Reefs 26:487–499

Nowak RM (1999) Walker's mammals of the world. John Hopkins University Press, Baltimore

Rach J, DeSalle R, Sarkar IN et al (2008) Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. Proc R Soc Lond B Biol Sci 275:237–247

Ratnasingham S, Hebert PDN (2007) BOLD: The barcode of life data system. Mol Ecol Notes 7:355–364

Redford KH (1992) The empty forest. Bioscience 42:412–422

Rodriguez F, Oliver JL, Marin A et al (1990) The general stochastic model of nucleotide substitution. J Theor Biol 142:485–501

Ross HA, Lento GM, Dalebout ML et al (2003) DNA surveillance: web-based molecular identification of whales, dolphins, and porpoises. J Hered 94:111–114

Rubinoff D (2006) Utility of mitochondrial DNA barcodes in species conservation. Conserv Biol 20:1026–1033

Rubinoff D, Cameron S, Will K (2006) A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. J Hered 97:581–594

Schmitz A, Mansfeld P, Hekkala E et al (2003) Molecular evidence for species level divergence in African Nile Crocodiles *Crocodylus niloticus* (Laurenti, 1786). CR Palevol 2:703–712

Shapiro B, Drummond A, Rambaut A et al (2004) Rise and fall of the Beringian steppe bison. Science 306:1561–1565

Song H, Buhay JE, Whiting MF et al (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc Natl Acad Sci USA 105:13486–13491

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758–771

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

Tamura K, Dudley J, Nei M et al (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Thorbjarnarson JB, Eaton MJ (2004) Preliminary examination of crocodile bushmeat issues in the Republic of Congo and Gabon, pp 236–247. Crocodiles: Proceedings of the 17th working meeting of the IUCN-SSC Crocodile Specialist Group, IUCN, Gland, Switzerland

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

Vasconcelos WR, Hrbek T, Da Silveira R et al (2006) Population genetic analysis of *Caiman crocodilus* (Linnaeus, 1758) from South America. Genet Mol Biol 29:220–230

Vuissoz A, Worobey M, Odegaard N et al (2007) The survival of PCR-amplifiable DNA in cow leather. J Archeol Sci 34: 823–829

Witt JDS, Threloff DL, Hebert PDN (2006) DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. Mol Ecol 15:3073–3082

Yan P, Wu X-B, Shi Y et al (2005) Identification of Chinese alligators (*Alligator sinensis*) meat by diagnostic PCR of the mitochondrial cytochrome *b* gene. Biol Conserv 121:45–51

Yang ZH (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39:306–314