American Journal of Primatology 72:715–724 (2010)

RESEARCH ARTICLE

Phylogeography of the Angolan Black and White Colobus Monkey, *Colobus angolesnsis palliatus*, in Kenya and Tanzania

MONICA M. MCDONALD[1,2*] AND HEALY HAMILTON[2]
[1]*Department of Biology, San Francisco State University, San Francisco, California*
[2]*Center for Applied Biodiversity Informatics, California Academy of Sciences, San Francisco, California*

Little is known about genetic variation in the 6–8 subspecies of *Colobus angolensis*, currently distinguished by pelage differences. We present a comparative genetic analysis of one of these subspecies, *C. a. palliatus*, in Kenya and Tanzania that assesses evolutionary relationships and patterns of mitochondrial genetic diversity in 103 individuals across its geographic range. Fecal samples from approximately 156 individuals were collected in four localities: (1) Diani Forest, Kenya; (2) Shimoni, Kenya; (3) Udzungwa Mountains National Park, Eastern Arc Mountains, Tanzania; and (4) Mount Rungwe, Southern Highlands, Tanzania. These samples represent at least six groups, with 5–15 samples from each. Comparative sequence analysis of a 1,795 base pair mtDNA fragment revealed 19 unique haplotypes in four populations. Phylogenetic analyses suggest that sampled Kenyan haplotypes are paraphyletic, with one Kenyan haplotype basal to all other sampled haplotypes. Analysis of molecular variance (AMOVA) suggests high levels of genetic variation among populations (Φ_{ST} 0.72, $P < 0.001$). Genetic data are concordant with a subspecies level differentiation between *C. a. palliatus* populations in Kenya and those in Central and southern Tanzania, as earlier suggested based on pelage differences. This study highlights the evolutionary distinctiveness of Kenyan populations of *C. a. palliatus* relative to Tanzanian populations. Although *C. a. palliatus* habitat in Tanzania is currently better protected than in Kenya, our results suggest Kenyan and Tanzanian populations should be considered distinct units, and the protection of *C. a. palliatus* habitat in Kenya, as well as habitat connectivity between Kenyan populations, should be prioritized for conservation and management. Am. J. Primatol. 72:715–724, 2010. © 2010 Wiley-Liss, Inc.

Key words: *Colobus angolensis palliatus*; phylogeography; mitochondrial DNA; genetic diversity; conservation

INTRODUCTION

Human activity and habitat decline may drastically alter primate populations and their habitats [Cowlishaw, 1999]. Conservation efforts can help reduce or counteract some of these negative impacts. However, it is important to understand the behavior, ecology, and evolutionary uniqueness of the target primate to support science-based management plans aimed at conserving them and their habitats. For subspecies, such as *Colobus angolensis palliatus*, experiencing habitat fragmentation, genetic studies have become increasingly useful in obtaining this knowledge; however, few behavioral and genetic data exist for most African colobine monkeys [Grubb et al., 2003].

Three genera of African colobines inhabit West Africa, Central Africa, and East Africa: *Colobus, Procolobus,* and *Piliocolobus* [Oates et al., 1994; Oates & Trocco, 1983]. *C. angolensis* is one of four or five species of *Colobus* [Kingdon & Howell, 1993] whose habitat ranges from Angola to Kenya. *C. a. palliatus*, which resides in southern coastal Kenya to southern Tanzania (Fig. 1), is one of six to eight subspecies defined based upon morphology [Dandelot, 1971; Groves, 2001; Napier, 1985; Rahm, 1970].

Rahm [1970] and Napier [1985] recognized eight subspecies of *C. angolensis*, whereas Dandelot [1971] and Groves [2001] recognized only six. Groves and Dandelot suggested that *C. a. ruwenzorii* and *C. a. adolfifriederici* are synonymous because they differ only in pelage density, likely attributable to altitudinal differences in their habitat. They also suggested that *C. a. palliatus* and *C. a. sharpei* should be

Contract grant sponsors: California Academy of Sciences; Institute for Primate Research (IPR).

*Correspondence to: Monica M. McDonald, Department of Anthropology, Washington University, St. Louis, MO 63130-4899.
E-mail: mmcdonald@wustl.edu

Received 1 November 2009; revised 1 March 2010; revision accepted 1 March 2010

DOI 10.1002/ajp.20828
Published online 1 April 2010 in Wiley InterScience (www.interscience.wiley.com).

© 2010 Wiley-Liss, Inc.

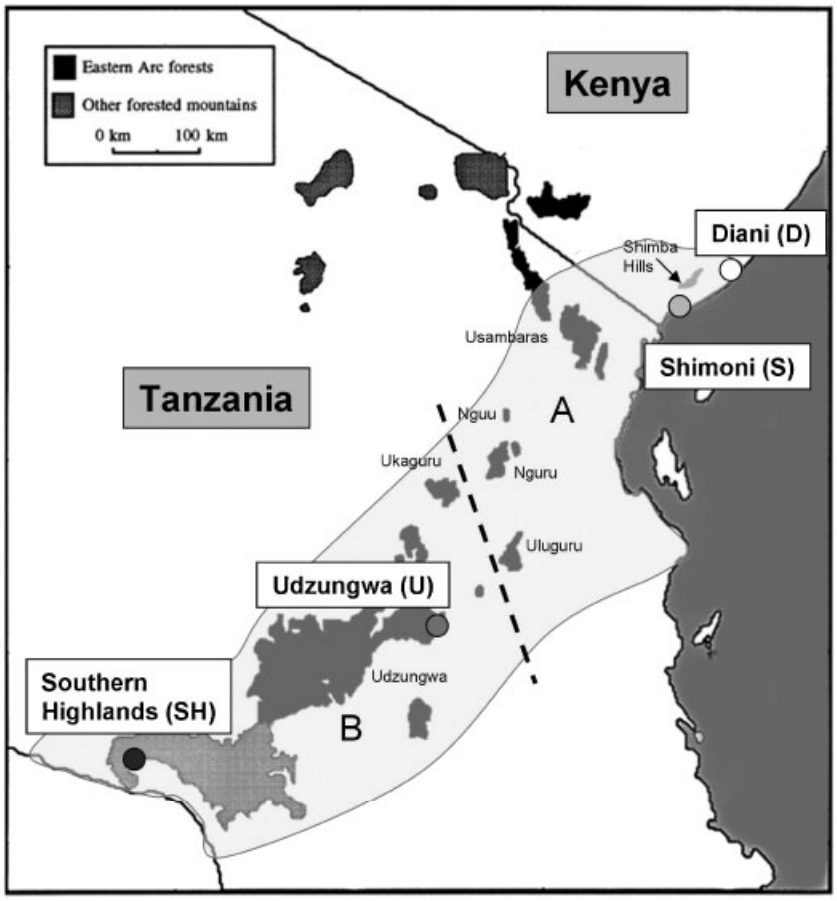

Fig. 1. Map of *Colobus angolensis* habitat in Kenya and Tanzania (lightly shaded region) and the four sampling locations in this study (sample sizes in parentheses): Diani ($N = 22$), Shimoni ($N = 18$), Udzungwa ($N = 30$), Southern Highlands ($N = 33$). The dashed line represents the boundary between *C. a. palliatus* (**A**) and *C. a. sharpei* (**B**) as suggested by Rodgers [1981], with the names of the surrounding Eastern Arc Mountain groups.

combined into a single subspecies, *C. a palliatus*, for similar reasons. The *C. angolensis* populations in this study are referred to as *C. a. palliatus*, following Groves [2001]. A 2003 IUCN/SSC assessment of the diversity of African primates suggested that the systematics of the genus *Colobus* and relationships among its subspecies be further investigated to support science-based conservation planning [Grubb et al., 2003].

There is an urgency to resolve the taxonomic structure of *C. a. palliatus*, as its Kenyan habitat is rapidly deteriorating. *C. a. palliatus* remains only in patches of dry coastal forest in the southernmost tip of the country. Aside from the 192 km^2 Shimba Hills Nature Reserve (Fig. 1), the forests of their Kenyan habitat are unprotected and increasingly fragmented by roads, resorts, and commercial development. Indirect results of this fragmentation include vehicular death and electrocution when monkeys travel among forest fragments [Anderson, 2001]. In 2001, a population density study [Anderson, 2005] estimated that 3,100–4,900 *C. a. palliatus* individuals remain in Kenya. This study focused on animals inhabiting Shimba Hills National Reserve, Diani Forest, and Shimoni whose populations consisted of 1,600–3,300, 330, and 120 individuals, respectively. Since then, the number has decreased, particularly in unprotected areas [Colobus Trust, personal communication, June 2007]. In addition, two *C. a. palliatus* specimens collected in the early 1900s, as well as local interview and archival research carried out by Anderson [2005], confirmed the earlier existence of *C. a. palliatus* in the Kilifi District (50 km north of their present-day habitat). However, they were last sighted there in 1979 and are almost certainly absent from this area today.

C. a. palliatus also inhabits forests throughout Tanzania. It inhabits various forested regions in the Eastern Arc Mountains (Fig. 1), a chain of ''mountain islands'' whose forests have been isolated from one another by woodland savannah, since the Pleistocene [Wasser, 1993] and more recently by anthropogenic impacts. One population of *C. a. palliatus* exists in the Udzungwa Mountains, a central portion of this mountain chain, considered a biodiversity hotspot because of its numerous endemic species [Wasser,

1993]. Another population resides in the Tanzanian Southern Highlands, forested mountains adjacent to the southern end of the Eastern Arc chain. In these two areas, *C. a. palliatus* was earlier assigned to a different subspecies, *C. a. sharpei*, based upon pelage differences [Dandelot, 1971; Groves, 2001; O'Leary, 2003; Rodgers, 1981]. These two forested Tanzanian regions are larger than the colobus habitat in Kenya, are less disturbed, and have varying levels of protection status. However, even in these regions, growing human populations exert increasing pressure on forest resources and hunting may still occur [Rovero et al., 2006].

In 2009, the IUCN altered the status of *C. a. palliatus* from Data Deficient to Least Concern [Kingdon et al., 2008]. Although their listing has changed, few additional studies have been published since the earlier IUCN assessment in 2000, suggesting a continued need for more scientific information [Anderson, 2005]. It is uncertain how genetically and behaviorally different populations once assigned to *C. a. sharpei* are from other populations of *C. a. palliatus*. Furthermore, no genetic data are available on variation within the six putative *C. angolensis* subspecies. Evolutionary relationships among populations of *C. angolensis* across its range remain unclear. Understanding these relationships is essential to support science-based conservation planning to sustain remaining populations of *C. a. palliatus*.

This study is the first effort to analyze phylogeographic structure among putative subspecies *C. a. palliatus* in Kenya and *C. a. sharpei* in Central and southern Tanzania. Our results clarify current taxonomy by assessing the extent of differentiation between populations of *C. a. sharpei* and those known to be *C. a. palliatus*. This study also identifies areas of increased genetic diversity and uniqueness, and examines the relationship between genetic variation and geographic distance. Collectively, our study provides important new insights into the biology of *C. a. palliatus* that are relevant to the development of conservation strategies necessary to sustain its dwindling populations.

METHODS

Sample Collection, DNA Extraction, and Sequencing

Between June and August 2007, we collected 240 fecal samples from approximately 156 individuals using a protocol that adhered to the American Society of Primatologists' Principles for the Ethical Treatment of Non-Human Primates, was approved by the Animal Care and Use Committee (UACUC) of San Francisco State University, and through the permission of appropriate authorities in both Kenya (No. MOST/13/001/37C 94) and Tanzania (No. 2007-226-NA-2007-123). We tracked individuals until they defecated, and immediately collected and stored the samples to ensure freshness. Sampling occurred in four localities spanning the *C. a. palliatus* geographic range: (1) Diani Forest, Kenya; (2) Shimoni, Kenya; (3) Udzungwa Mountains National Park, Eastern Arc Mountains, Tanzania; and (4) Mount Rungwe, Southern Highlands, Tanzania (Fig. 1). Within each locality, the samples represent at least six groups, with 5–15 samples from each. Each sample was stored using two different methods to preserve the quality of mitochondrial DNA: a two-step ethanol–silica method [Nsubuga et al., 2004] and RNA*later* (Ambion, Austin, TX).

We examined the phylogeography of *C. a. palliatus* in 1,795 base pairs (bp) of mitochondrial DNA, spanning the protein-coding cytochrome b and the control region (D-loop). Several studies have examined these gene regions in other taxa with proven utility in intraspecific analyses, and many comparative sequences exist in global genetic databases [Clifford et al., 2004; Jensen-Seaman & Kidd 2001; Telfer et al., 2003].

Using the QIAamp DNA Stool Mini Kit (Qiagen, Valencia, CA), we extracted a 100 mg portion of each sample according to the manufacturer's instructions with minor modifications. We used gel electrophoresis to visualize genomic DNA, and then amplified the DNA by polymerase chain reaction (PCR) using standard amplification methods [Clifford et al., 2004; Jensen-Seaman & Kidd 2001; Marmi et al., 2004; Modolo et al., 2005; Telfer et al., 2003]. Because only one *C. angolensis* cytochrome b sequence and no D-loop sequences exist in GenBank, we designed primers in Primer3 [Rozen & Skaletsky, 2000], using a partial *C. angolensis* mitochondrial genome provided by Dr. Nelson Ting (unpublished data).

For many of the samples, we amplified the entire region of interest using the GluF10 and HV1R primers to reduce the likelihood of amplifying mitochondrial pseudogenes (numts) (Table I) [Triant & DeWoody, 2007]. However, because of some difficulties in cycle sequencing such a large region, we had to reamplify the fragment of interest as two smaller regions, using primers GluF10 or Glu-5F and COB-7R for the first region (896 bp), and primers COB-3F, COB-4F, or COB-8F along with HV1R for the second region (1,231 bp) (Table I). We optimized a general 25 ul PCR protocol based on USB HotStart-IT DNA Polymerase protocol guidelines and ran the samples on a MyCycler thermocycler (Bio-Rad, Hercules, CA). The PCR program began with an initial 2 min denaturation at 95°C, followed by 35 cycles of denaturation (95°C for 30 sec), annealing (55°C for 30 sec), and elongation (68°C for 5 min), and a final 2 min extension at 68°C. PCR products were visualized via gel electrophoresis and amplicons were cleaned before cycle sequencing, using the shrimp alkaline phosphatase–exonuclease I enzyme (ExoSap-IT) protocol. We performed cycle

TABLE I. Primers and Corresponding Sequences from 5′ to 3′, Their Region of Amplification, and Their Location Relative to the Complete Mitochondrial Genome of Sister Taxon *C. guereza* (GenBank Accession Number AY863427)

Primer	Sequence 5′–3′	Region amplified	Location
GluF10	ATGGGTTTTAACCACGACCA	Cytochrome b	14139–14158
Glu-5F	TTTACATGGGTTTTAACCACGAC	Cytochrome b	14134–14156
COB-7R	GCAAACAGAAAATACCACTCTGG	Cytochrome b	15010–15032
COB-4F	CCCACTCTTACACGATTCTTCAC	D-loop	14719–14741
COB-3F	CACCTCCAACTCTGACAAAATTC	D-loop	14835–14857
COB-8F	CCCTTCCACCCCTACTATACAAC	D-loop	14857–14879
HV1R	TTTAAGGGGAACGTGTGAGC	D-loop	16035–16054

sequencing reactions, using BigDye Terminator version 3.1 following instructions for ABI Big Dye version 3.1 (Applied Biosystems, Foster City, CA) and used four primers to cycle sequence the entire mitochondrial fragment of interest. Cycle sequencing products were purified and sequences were read and analyzed on an ABI 3130.

Data Analysis

For most cases, both forward and reverse sequencing was performed; however in a few cases, clean sequences in only one direction were also included. We used Sequencher 4.7 for sequence editing and alignment (Gene Codes Corporation, MI) and codon translation. The absence of stop codons was confirmed in gene-coding regions, to check for nuclear gene copies (pseudogenes) and verify mitochondrial fragments [Collura & Stewart, 1995; Triant & DeWoody, 2007]. We used MacClade 4.08 [Maddison & Maddison, 2003] and PAUP 4.0 [Swofford, 2003] to assess the distribution of characters in this 1,795 bp data set.

To determine diversity indices and phylogeographic structure, we used Arlequin v.3.1 [Excoffier et al., 2005]. We calculated average percent sequence divergence values for each of the four populations using (1) Jukes–Cantor corrected distances for the cytochrome b gene and (2) ''uncorrected p'' distances for the D-loop. We calculated average nucleotide and haplotype diversity values for each subpopulation and gave a range for the overall population. An Analysis of Molecular Variance (AMOVA) was used to determine the extent of population differentiation among the four collection localities: (1) Diani, Kenya; (2) Shimoni, Kenya; (3) Udzungwa Mountains, Tanzania; (4) Southern Highlands, Tanzania. We ran two AMOVA analyses using this grouping scheme, one to assess population differentiation using conventional *F*-statistics, based on allele frequency differences, and the other to assess differentiation using Φ-statistics, which also incorporate haplotype genetic distances [Hartl, 2000]. Finally, we performed a Mantel Test using 1,000 permutations to determine whether a correlation exists between the population Φ_{ST} values and the geographic distances between localities. We used TCS v. 1.21 [Clement et al., 2000] to estimate a minimum spanning network of the mtDNA haplotypes.

For phylogenetic inference, we carried out parsimony and maximum likelihood analyses using PAUP 4.0 [Swofford, 2003] and Bayesian analyses using MrBayes v.3.1.2 [Hall, 2005; Ronquist & Huelsenbeck, 2003]. We used PAUP 4.0 [Swofford, 2003] to run parsimony searches on the 1,795 bp data set of unordered, unweighted characters by using heuristic, step-wise addition of 1,000 random replicates. *C. guereza,* a close relative to *C. angolensis*, was used as the outgroup (accession number AY863427). We used ModelTest 3.7 [Posada & Crandall, 1998] to determine the optimal model of evolution for the maximum likelihood analysis and then performed a heuristic random addition maximum likelihood search in PAUP 4.0 [Swofford, 2003] using a starting neighbor-joining tree. We generated bootstrap support using heuristic, closest step-wise addition of 1,000 random pseudoreplicates for the parsimony search and 1,000 replicates for the maximum likelihood search. For the Bayesian analysis, we divided the data set into three regions: cytochrome b, tRNA, and D-loop, because these three regions can evolve at different rates. We further used MrModelTest 2.3 [Nylander, 2004] to estimate the best fit model of molecular evolution for each of these regions as determined by the Akaike criterion. Finally, we performed a Bayesian search of two parallel runs for 75,000,000 generations of four heated Markov chains per generation, where we sampled trees every 1,000 generations and set burn-in values at 25%, with Tracer v.1.4 used to confirm proper burn-in settings.

RESULTS

The final data set is an alignment of 1,795 bp of mitochondrial DNA for 103 individuals from four localities, two in Kenya (Diani, $N = 22$, Shimoni, $N = 18$) and two in Tanzania (Udzungwa, $N = 30$, Southern Highlands, $N = 33$) (Fig. 1). This data set included the complete protein-coding cytochrome b

gene (1,140 bp) and the first 524 bp of the D-loop (including hypervariable region 1), along with the area separating these two regions that consisted of an mRNA (18 bp) and two tRNAs (tRNA-Thr and tRNA-Pro) (113 bp) (Table II). The final alignment contained 101 variable sites, 79 of which were parsimony informative, and no gapped sites (Table II). In cases where two peaks of similar intensity were found in the electropherogram, the nucleotide was designated with the appropriate ambiguity code; seven ambiguous characters among the *C. a. palliatus* haplotypes were found in the D-loop and were coded as Ns or with the appropriate ambiguity code. This data set revealed 19 unique haplotypes, only two of which were shared between populations. Of these two shared haplotypes, the D2S1 haplotype is shared between the two Kenyan populations: Diani and Shimoni, and the SH1U5 haplotype is shared between the two Tanzanian populations: Udzungwa and Southern Highlands, although only one Udzungwa sample shared this primarily Southern Highlands haplotype. The Kenyan and Tanzanian populations share no haplotypes in this data set.

Population Genetic Analysis

Results from the population genetic analyses show the biggest genetic divide occurring between Kenyan and Tanzanian populations. Average percent sequence divergence between the Kenyan and Tanzanian populations is 1.4% in cytochrome b and 5.0% in D-loop. In comparison, sequence divergence separating the two Tanzanian populations (Udzungwa and Southern Highlands) is only 0.4% in cytochrome b and 2.4% in D-loop, very similar to the genetic distance between the two Kenyan populations (Diani and Shimoni) of 0.6% in cytochrome b and 2.9% in D-loop. We calculated nucleotide and gene diversity for each of the populations as well as the combined data set (Table III). Overall, nucleotide diversity values ranged from 0.0021–0.0113 $\pm$ SD 0.001 to 0.006, whereas overall haplotype diversity across all localities was 0.64–0.85 $\pm$ SD 0.034–0.095. The Udzungwa population in Tanzania showed the highest nucleotide diversity, followed closely by the Shimoni population in Kenya, whereas the Southern Highlands population in Tanzania displayed the highest haplotype diversity. Although greater genetic distance exists between sequences from the Udzungwa and Shimoni populations, many fewer haplotypes exist in these localities. Conversely, the Southern Highlands population displays more haplotype diversity, yet the genetic distance between these haplotypes is small.

For the AMOVA analysis, we divided the sequences into four populations (D, S, U, SH), and combined them into two regional groupings (Kenya = [D, S] and Tanzania = [U, SH]. Results from the AMOVA analysis using conventional *F*-statistics indicate significant structuring of the four populations ($F_{ST} = 0.26$, $P < 0.00001$), as well as significant structuring of populations compared with their regional variation ($F_{SC} = 0.25$, $P < 0.00001$). Most of the genetic variation (73%) was attributed to within population genetic variance, likely a reflection of the several unique haplotypes found in each of these populations. To calculate the AMOVA using Φ-statistics, we applied Tamura and Nei genetic distances with a gamma shape parameter of

TABLE II. Character Status of the Dataset

Gene region	♯ of characters sequenced	Variable characters	Parsimony informative	Gapped characters	Missing or ambiguous
Complete fragment	1795	101	79	0	7
Cyt b	1140	37	30	0	0
D-loop	524	60	45	0	7
mRNA and t-RNAs	131	4	4	0	0

TABLE III. Number of Individuals (*N*), Number of Haplotypes (N_{hap}), Gene Diversity (*h*) with Standard Deviation (SD), Nucleotide Diversity (π) with Standard Deviation (SD), and Number of Transitions (Ti), Transversions (Tv), and Indels

Region	*N*	N_{hap}	$h \pm$SD	$\pi \pm$SD	Ti, Tv, indel
Diani (D)	22	4	0.7229 $\pm$ 0.0638	0.0066 $\pm$ 0.0035	35, 1, 0
Shimoni (S)	18	4	0.6471 $\pm$ 0.0953	0.0105 $\pm$ 0.0055	45, 2, 0
Udzungwa (U)	30	5	0.6782 $\pm$ 0.0572	0.0113 $\pm$ 0.0057	46, 3, 0
Southern Highlands (SH)	33	8	0.8466 $\pm$ 0.0345	0.0021 $\pm$ 0.0012	12, 0, 1
Total	103	21			

*Two shared haplotypes were revealed, one between D and S populations and another between U and SH populations.

0.635; this genetic distance measure was most similar to the HKY85 model used in the phylogenetic analysis. We also measured significant population differentiation compared with the total using Φ-statistics ($\Phi_{ST} = 0.72$, $P < 0.00001$), as well as significant structuring of populations within regional groups ($\Phi_{SC} = 0.49$, $P < 0.00001$). In this analysis, the statistical values were much higher than those found using conventional *F*-statistics. Among population, genetic variation accounted for the majority of genetic diversity (46%) as compared with within population variation (27%) or among population variation within regional groups (28%). These differences between the conventional *F*- and Φ-statistics indicate that most of the sampled genetic variation can be attributed to the mutational differences existing between haplotypes, particularly haplotypes from different geographic regions. These differences also indicate that the genetic variation in the data set is relatively old, making it misleading to attribute equal weighting to all haplotypes, as would be done if only allele frequencies were considered in the analysis. Thus, the Φ-statistical values—rather than the *F*-statistical values—more accurately explain the patterning of genetic variation in this data set.

TCS v.1.21 was used to create a minimum spanning network among unique haplotypes (Fig. 2). A 95% connection limit was used and gaps were treated as missing data. The number of mutations between haplotypes varies markedly from 1 to 40 bp. Geography can explain only some of the large differences between haplotypes. In Kenya, the Diani and Shimoni haplotypes do not cluster into distinct groups, but rather are mixed, with some Kenyan haplotypes differing by very few mutations and others by 30 or more. In Tanzania, the U4 haplotype obtained from individuals in the Udzungwa Mountains is genetically more similar to the Southern Highlands population. Also, although geographically proximate to the other Kenyan individuals, the individuals forming the S2 haplotype are genetically distant and possibly ancestral to all other Kenyan and Tanzanian samples in this study (Figs. 2 and 3).

Finally, the Mantel test indicated a significant correlation ($P < 0.05$) between population Φ_{ST} values and the geographic distances (kilometers) between these four localities, suggesting a simple isolation-by-distance model.

Phylogenetic Analysis

The parsimony analysis yielded 90 equally parsimonious trees, the best of which had a length of 244 steps. A strict consensus of these 90 trees

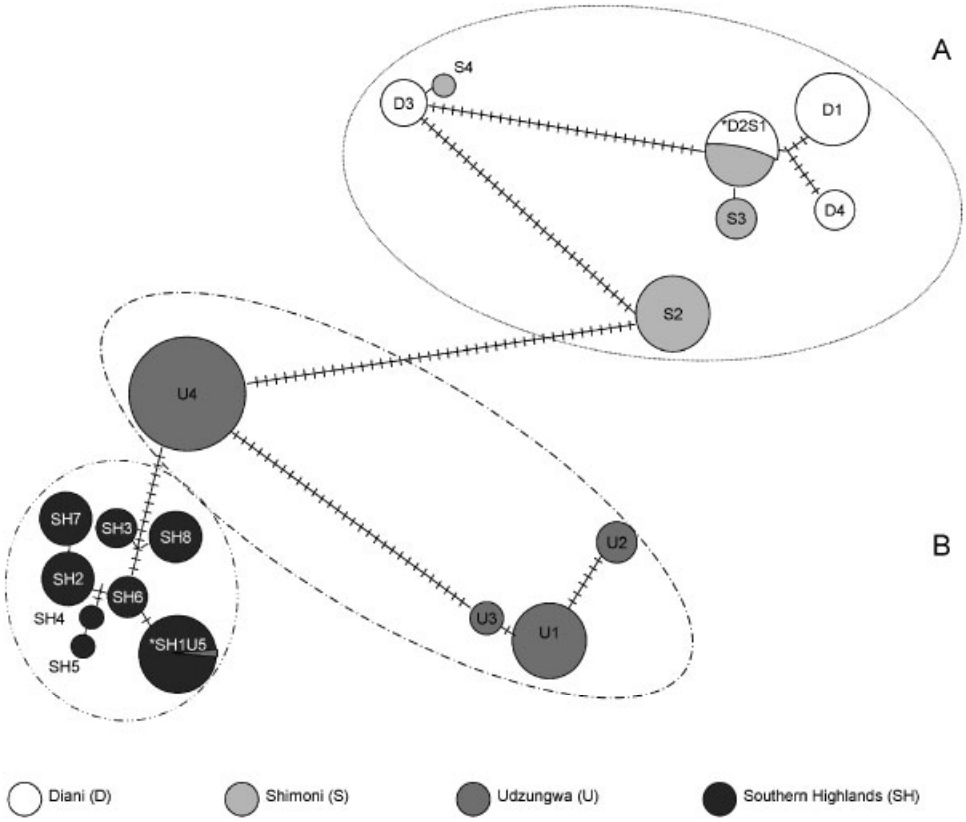

Fig. 2. Minimum-spanning network of the haplotypes created using a 95% connection limit. The size of the circle refers to the relative frequency of the haplotype in the data set. Each tick represents a mutation event. **A** and **B** refer to the Kenyan and Tanzanian haplotypes, respectively. The asterisk denotes the two shared haplotypes, D2S1 and SH1U5.

resulted in six major clades of haplotypes, with most branches receiving strong support. For the maximum likelihood analysis, ModelTest's Akaike criterion [Posada & Crandall, 1998] suggested the HKY85 model with invariant sites and a gamma distribution (HKY85+I+G) as the best fit model of molecular evolution. The maximum likelihood tree topology (−ln 3670.19) was identical to the parsimony tree. We show the likelihood tree with both parsimony and maximum likelihood bootstrap values (Fig. 3).

Both methods retrieved the same six clades. Diani and Shimoni populations from Kenya formed clades I, II, and III, which are paraphyletic. Although both methods suggest generally good support for the individual clades, the relationship among them remains unresolved, with the exception of lineage S2, which seems a highly distinct, possibly basal lineage. The Udzungwa haplotypes form two paraphyletic clades, IV and V, with Clade V clearly basal to all Southern Highlands haplotypes. All Southern Highlands individuals form monophyletic Clade VI. Evolutionary distances are small between Southern Highlands haplotypes compared with all other clades, and all clearly are derived from an Udzungwa ancestor.

For the Bayesian analysis, three different models of molecular evolution best fit the partitioned data set using the AIC criterion in MrModelTest [Nylander, 2004]: HKY85+G model for cytochrome b, HKY85 for the tRNAs, and HKY85+I+G for the D-loop. The Bayesian consensus tree was generally concordant with the likelihood tree. The same six clades were retrieved, but again, low statistical support leaves the relationship among clades poorly unresolved. The clearest evolutionary signals among the clades are the origin of the Southern Highlands haplotypes from one Udzungwa lineage, and the distant, basal position of the S2 lineage (Fig. 3).

DISCUSSION

Population Information, Phylogeography, and Demographic History

A significant amount of genetic variation and population structure exists both within and among these populations of *C. a. palliatus,* and the genetic distance information combined with phylogeographic analyses tell a complex history that can be mapped onto geography.

The extent and distribution of genetic variation in this subspecies suggests a deep evolutionary history. In Kenya, although the Diani and Shimoni populations share only one haplotype, the positioning of the Diani and Shimoni haplotypes on the

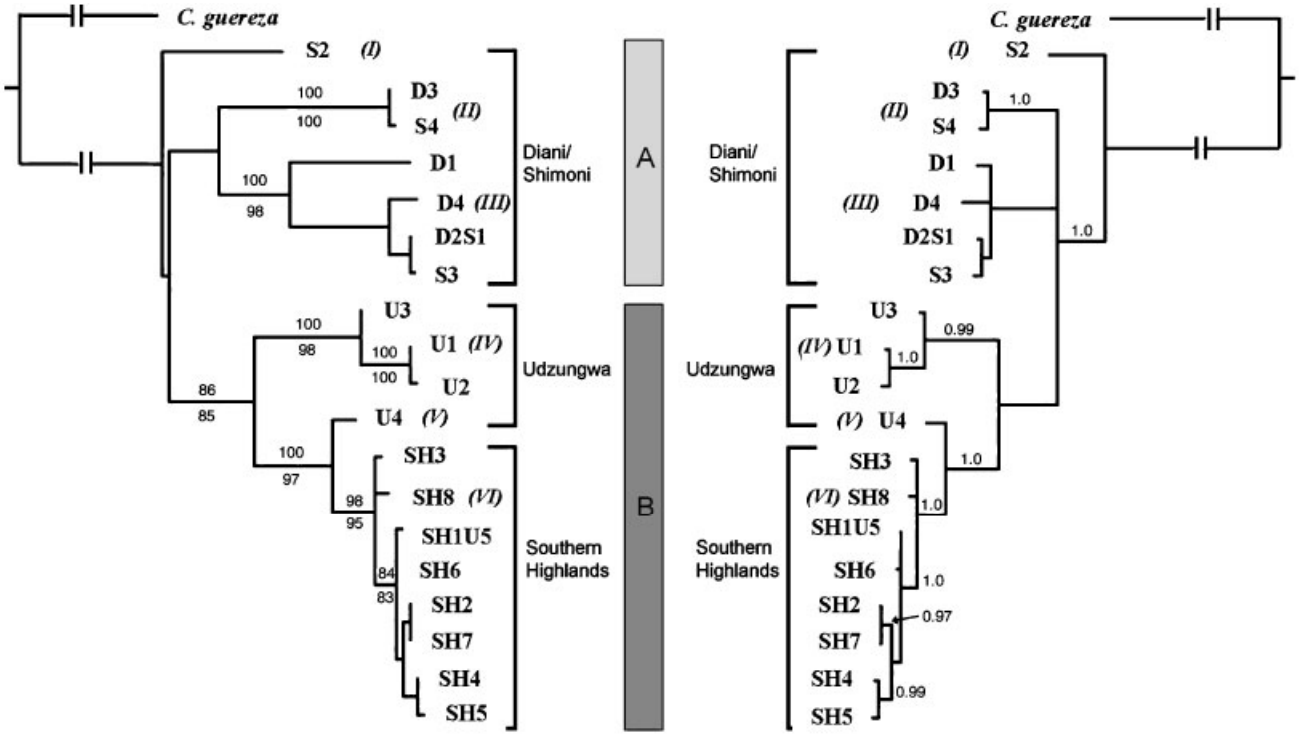

Fig. 3. Maximum-likelihood (ML) tree (left) and Bayesian tree (right) for 19 *C. angolensis* haplotypes with 1 *C. guereza* haplotype as an outgroup. The numbered haplotypes are also coded by region (Diani, D; Shimoni, S; Udzungwa, U; and Southern Highlands, SH). The Roman numerals specify the different clades and **A** and **B** indicate Kenyan and Tanzanian haplotypes, respectively. In the ML tree (left), values above the branches refer to parsimony bootstrap support and values below the branches refer to ML bootstrap support. Only bootstrap support values of 75 or above are shown. In the Bayesian tree (right), posterior probability values greater than or equal to 0.95 are noted near branch nodes.

phylogenetic trees are paraphyletic (Diani: $N_{hap} = 4$, $N_{indiv} = 22$; Shimoni: $N_{hap} = 4$, $N_{indiv} = 18$) and some pairwise haplotype comparisons reveal much more genetic distance than others. The large genetic differences between haplotypes suggest that, at some point in their evolutionary history, these two populations were geographically separated. Conversely, the small genetic differences between other haplotypes in Kenya imply that these two populations have since remixed and have maintained genetic contact until relatively recently. Paraphyly between the Diani and Shimoni populations suggests either continuing gene flow between these populations, or relatively recent isolation with ancestral polymorphisms still shared.

In Tanzania, the Udzungwa clades consist of all Udzungwa individuals ($N_{hap} = 4$, $N_{indiv} = 30$) and the Southern Highlands clade consists of all Southern Highlands individuals ($N_{hap} = 8$, $N_{indiv} = 33$), except for the one Udzungwa individual sharing a Southern Highlands haplotype (SH1U5). The U4 haplotype, although consisting only of Udzungwa individuals, is more closely related to the Southern Highlands clade than to the other Udzungwa haplotypes. This haplotype may be ancestral to the Southern Highlands population, and individuals from this ancestral group diverged into the Udzungwa and Southern Highlands populations observed today. Thus, much of the within population variation from the Udzungwas is old, whereas population variation is much more recent within the Southern Highlands. The SH1U5 shared haplotype suggests a possible recent backwards migration event from the Southern Highlands to the Udzungwas. All methods of phylogenetic reconstruction suggest the S2 haplotype, derived from a subset of the Shimoni population, is basal to all other sampled Kenyan and Tanzanian haplotypes. This subset of Shimoni may be ancestral, having later diverged into the other Kenyan and Tanzanian populations. These findings either suggest a southward dispersal of *C. a. palliatus* from Kenya into Tanzania or several separate vicariant events.

A significant positive correlation exists between Φ_{ST} measures of these four populations and the geographic distances between their localities. This relationship between genetic variation and geographic distance is concordant with the pattern of isolation by distance (IBD), in which individuals disperse slowly across a landscape, and the populations in closer proximity to each other geographically are more genetically similar [Bohonak, 2002]. The phylogenetic analyses also support an IBD pattern in the paraphyly of the two Kenyan populations and the presence of shared haplotypes only between geographically closest regions.

We observed the greatest genetic distance between sampled Kenyan and Tanzanian populations. Cytochrome b sequence divergence was 1.4% and control region distances measured about 5% between the Kenyan and Tanzanian populations of *C. a. palliatus*. This level of genetic distance between Kenyan and Tanzanian *C. a. palliatus* haplotypes is much less than the 6% (cytochrome b) and 10.5% (control region) differences found between *C. angolensis* and *C. guereza* species, and is in the range of cytochrome b evolutionary distances found at the subspecies level in other primate studies. For example, in a phylogeographic examination of the *Trachypithecus* species complex (the silvered langur species group) based on 573 bp of cytochrome b, Roos et al. [2008] recognized a new subspecies of *T. cristatus* from West Malaysia, based in part on $1.4(\pm 0.48)$–$1.9(\pm 0.98)\%$ cytochrome b sequence divergence, which is comparable to the evolutionary difference between Kenyan and Tanzanian clades analyzed here.

Conforming to the general pattern in primate phylogeographical studies in which the hypervariable portion of the control region evolves more rapidly than cytochrome b, control region sequences in Monda et al. [2007] showed greater percent distances between *Nomascus* haplotypes, as did the percent distances of *C. a. palliatus* in our study. However, percent sequence divergence in control region seems a less reliable metric when applied across taxa to determine level of evolutionary distinction. In gibbons, 5% control region divergences corresponded to species level taxonomic distinctions for five distinct species [Monda et al., 2007], whereas in orangutans, distances as large as 6.5% represented only population level variation [Warren et al., 2001]. Clearly, percent sequence divergence is not a single metric that can be applied across distinct primate taxa to designate taxonomic status. Additional information, such as the presence of reciprocal monophyly, the geographic distribution of haplotypes and morphological characters must be integrated with genetic data to form taxonomic hypotheses on a case-by-case basis.

Here, the degree of evolutionary distance between control region and cytochrome b haplotypes in Kenyan and Tanzanian populations, combined with the pelage differences highlighted by Rahm [1970] and Napier [1985], support the earlier assignment of *C. a. palliatus* into two different subspecies in Kenya and Tanzania, *C. a. palliatus* in Kenya and *C. a. sharpei* in Central and southern Tanzania. However, a clear lack of reciprocal monophyly shows the sampled populations may still experience limited gene flow. Comparable genetic studies of nuclear variation and more extensive sampling in the geographic areas between Shimoni, Kenya, and the Udzungwa Mountains, Tanzania, are required to provide further evaluation of the subspecies level division of *C. a. palliatus* in Kenya and *C. a. sharpei* in Central and southern Tanzania.

Comparative phylogeographic patterns across a range of taxa indicate the importance of the Eastern Arc Mountains as a potential barrier to gene flow. In a multispecies comparison, Fjeldså & Bowie [2008] found that seven of eight bird lineages demonstrated a genetic break between the Usambara Mountains in the Northern Eastern Arc Mountains and the closely adjacent Nguu/Nguru Mountains in the Central Eastern Arc (Fig. 1). Avian species in or north of the Usambara Mountains in the Eastern Arc Mountain chain were more genetically similar to those in Southern Kenya, whereas populations residing in Eastern Arc Mountain regions south of the Usambaras were more genetically similar to other central and southern Eastern Arc populations. The genetic relationships of *C. a. palliatus* are consistent with these results [Fjeldså & Bowie, 2008], with the largest break found between the Diani and Shimoni populations north of the Usambaras and the Southern Highlands and Udzungwa populations found south of the Nguu/Nguru. These shared phylogeographic patterns could be further tested with finer scale sampling of *C. a. palliatus*, particularly in the unsampled Northern Eastern Arc Mountains.

Conservation Implications

We found that a large amount of genetic diversity still exists within Kenyan populations of *C. a. palliatus*, and that this population contains ancestral genetic diversity not present in sampled Tanzanian populations. These findings, along with the lack of protection for most *C. a. palliatus* habitat in Kenya, highlight the need to prioritize conservation efforts for the Kenyan populations of *C. a. palliatus*. Within Southern Kenya, there are existing community-based conservation efforts underway, focused on the management and conservation of the colobus and other primates. *The Colobus Trust* currently has public education campaigns and builds ''colobridges'' near their organization to help the monkeys safely cross roads without using electrical lines. Although these colobridges are a significant attempt at connecting nearby habitat fragments, our study highlights the importance not only of establishing a greater amount of protected area in Kenya but also in maintaining and expanding connectivity among anthropogenically isolated *C. a. palliatus* populations in Diani, Shimoni, and the nearby Shimba Hills, in order to ensure the maintenance of the high levels of genetic diversity found in these populations.

ACKNOWLEDGMENTS

This project was funded by the California Academy of Sciences (G. Lindsay Field Research Fund and Center for Comparative Genomics Start-Up Fund); the Genome Analysis Core Facility, Helen Diller Comprehensive Cancer Center, UCSF; and San Francisco State University (IRA Biology Dept. Small Grant 2007 & 2008), and adhered to IACUC protocols, as well as Kenyan and Tanzanian laws. The first author is grateful to the National Museums of Kenya, the Institute for Primate Research (IPR), the Kenya Wildlife Service (KWS), TAWIRI, COSTECH, and TANAPA for permit acquisition and for allowing the collection of fecal material (Kenya—No. MOST/13/001/37C 94, Tanzania—No. 2007-226-NA-2007-123). For logistical planning and sample collection, thanks go to Wakaluzu: Friends of the Colobus Trust in Kenya and the Udzungwa Ecological Monitoring Centre (UEMC) and WCS Southern Highlands Conservation Programme in Tanzania. The first author thanks O. Macharia, P. Gacheru, H. Mwahari, P. Ndungu, F. Rovero, A. Kitale, Saidi, T. Davenport, N. Mpunga, Amnoni, Lusargio, C. Epps, L. Hlusko, R. Lawson, B. Simison, G. Spicer, F. Cipriano, P. Sirya, E. Whitney, J. N. Msindai, S. Cornel, M. Tonione, S. Johnson, J. Phillips-Conroy, individuals at the SFSU Conservation Genetics Laboratory, and the Center for Comparative Genomics. Finally, this project could not have been accomplished without Dr. Nelson Ting's help and access to an unpublished sequence or without the advice and support of Dr. Robert Drewes and Dr. Eric Routman.

REFERENCES

Anderson J. 2001. Status, distribution and conservation of the Angola black-and-white colobus (*Colobus angolensis palliatus*) in coastal Kenya. Diani, Kenya: Wakuluzu, Friends of the Colobus Trust.

Anderson J. 2005. Habitat fragmentation and metapopulation dynamics of the Angolan black-and-white colobus (*Colobus angolensis palliatus*) in Coastal Kenya [Dissertation]. London: UCL.

Bohonak AJ. 2002. IBD (isolation by distance): a program for analyses of isolation by distance. Journal of Heredity 93:153–154.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657–1659.

Clifford SL, Anthony NM, Bawe-Johnson M, Abernethy KA, Tutin CEG, White LJT, Bermejo M, Goldsmith ML, McFarland K, Jeffery KJ, Bruford MW, Wickings EJ. 2004. Mitochondrial DNA phylogeography of western lowland gorillas (*Gorilla gorilla gorilla*). Molecular Ecology 13:1551–1565.

Collura RV, Stewart C-B. 1995. Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. Nature 378:485–489.

Cowlishaw G. 1999. Predicting the pattern of decline in African primate diversity: an extinction debt from historical deforestation. Conservation Biology 13:1183–1193.

Dandelot P. 1971. Order primates, suborder anthropoidea. In: Meester J, Setzer, editors. The mammals of Africa: an identification manual. Washington, DC: Smithsonian Institution. p 1–43.

Excoffier LG, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47–50.

Fjeldså J, Bowie RCK. 2008. New perspectives on the origin and diversification of Africa's forest avifauna. African Journal of Ecology 46:235–247.

Groves C. 2001. Primate taxonomy. Washington, DC: Smithsonian Institution Press.

Grubb P, Butynski TM, Oates JF, Bearder SK, Disotell TR, Groves CP, Struhsaker TT. 2003. Assessment of the diversity of African primates. International Journal of Primatology 24:1301–1357.

Hall BG. 2005. Phylogenetic trees made easy: a how-to manual. Sunderland, MA: Sinauer Associates, Inc.

Hartl DL. 2000. A primer of population genetics. Sunderland, MA: Sinauer Associates, Inc.

Jensen-Seaman MI, Kidd KK. 2001. Mitochondrial DNA variation and biogeography of eastern gorillas. Molecular Ecology 10:2241–2247.

Kingdon J, Howell KM. 1993. Mammals in the forests of Eastern Africa. In: Lovett JC, Wasser SK, editors. Biogeography and ecology of the rainforests of Eastern Africa. New York: Cambridge University Press. p 229–242.

Kingdon J, Struhsaker T, Oates J, Hart JF, Butynski TM, De Jong Y, Groves CP. 2008. *Colobus angolensis* ssp. *palliatus*. In: IUCN 2009. IUCN red list of threatened species. Version 2009.2. www.iucnredlist.org.

Maddison DR, Maddison WP. 2003. MacClade 4, analysis of phylogeny and character evolution. Sunderland, MA: MacClade 4.08 ed.

Marmi J, Bertranpetit J, Terradas J, Takenaka O, Domingo-Roura X. 2004. Radiation and phylogeography in the Japanese macaque, *Macaca fuscata*. Molecular Phylogenetics and Evolution 30:676–685.

Modolo L, Salzburger W, Martin RD. 2005. Phylogeography of Barbary macaques (*Macaca sylvanus)* and the origin of the Gibraltar colony. Proceedings of the National Academy of Sciences of the United States of America 102:7392–7397.

Monda K, Simmons RE, Kressirer P, Su B, Woodruff DS. 2007. Mitochondrial DNA hypervariable region-1 sequence variation and phylogeny of the concolor gibbons, *Nomascus*. American Journal of Primatology 69:1285–1306.

Napier PH. 1985. Catalogue of Primates in the British Museum (Natural History) and elsewhere in the British Isles. Part III: Family Cercopithecidae, Subfamily Colobinae. London: British Museum (Natural History).

Nsubuga AM, Robbins MM, Roeder AD, Morin PA, Boesch C, Vigilant L. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular Ecology 13:2089–2094.

Nylander JAA. 2004. MrModeltest v2.0: program distributed by the author. Evolutionary Biology Centre: Uppsala University.

O'Leary R. 2003. An annotated catalog of African primate genera *Colobus* and *Procolobus* (Cercopithecidae: Colobinae) in the collections of the American Museum of Natural History. American Museum Novitates 3399:1–26.

Oates JF, Trocco TF. 1983. Taxonomy and phylogeny of black and white colobus onkeys: inferences from an analysis of loud call variation. Folia Primatologica 40:83–113.

Oates JF, Davies GA, Delson E. 1994. The diversity of living colobines. In: Davies GA, Oates JF, editors. Colobine monkeys: their ecology, behavior, and evolution. New York, NY: Cambridge University Press. p 45–74.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Rahm UH. 1970. Ecology, zoogeography and systematics of some African forest monkeys. In: Napier JR, Napier PH, editors. Old World monkeys: evolution, systematics and behavior. New York: Academic Press. p 589–626.

Rodgers WA. 1981. The distribution and conservation status of colobus monkeys in Tanzania. Primates 22:33–45.

Ronquist F, Huelsenbeck JP. 2003. MrBayes3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Roos C, Nadler T, Walter L. 2008. Mitochondrial phylogeny, taxonomy and biogeography of the silvered langur species group (*Trachypithecus cristatus*). Molecular Phylogenetics and Evolution 47:629–636.

Rovero F, Struhsaker TT, Marshall AR, Rinne TA, Pedersen UB, Butynski TM, Ehardt CL, Mtui AS. 2006. Abundance of diurnal primates in Mwanihana Forest, Udzungwa Mountains, Tanzania: a multi-observer comparison of line-transect data. International Journal of Primatology 27:675–697.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers In: Krawetz S, Misener S, editors. Bioinformatics methods and protocols: methods in molecular biology. Totowa, NJ: Humana Press. p 365–386.

Swofford D. 2003. PAUP 4.0: phylogenetic analysis using parsimony and other methods. 4th ed. Sunderland, MA: Sinauer Associates.

Telfer PT, Souquiere S, Clifford SL, Abernethy KA, Bruford MW, Disotell TR, Sterner KN, Roques P, Marx PA, Wickings EJ. 2003. Molecular evidence for deep phylogenetic divergence in *Mandrillus sphinx*. Molecular Ecology 12:2019–2025.

Triant DA, DeWoody JA. 2007. The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. Journal of Mammalogy 88:908–920.

Warren KS, Verschoor EJ, Langenhuijzen S, Heriyanto, Swan RA, Vigilant L, Heeney JL. 2001. Speciation and intrasubspecific variation of Bornean orangutans, *Pongo pygmaeus pygmaeus*. Molecular Biology and Evolution 18:472–480.

Wasser SK. 1993. The socioecology of interspecific associations among the monkeys of the Mwanihana rain forest, Tanzania: a biogeographic perspective. In: Lovett JC, Wasser SK, editors. Biogeography and ecology of the rainforests of Eastern Africa. New York: Cambridge University Press. p 267–280.