

Mitochondrial Phylogeny and Systematics of Baboons (*Papio*)

Timothy K. Newman,^{1*} Clifford J. Jolly,² and Jeffrey Rogers^{1,3}

¹Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas 78227

²Department of Anthropology, New York University, New York New York 10003

³Southwest National Primate Center, San Antonio, Texas 78227

KEY WORDS taxonomy; primate; Cercopithecine; mtDNA; ND4; ND5

ABSTRACT Baboons (*Papio*, s.s.) comprise a series of parapatric allotaxa (subspecies or closely related species) widely distributed in sub-Saharan Africa. Despite extensive studies of their ecology, morphology, and behavior, disagreement about their phylogenetic relationships continues, as expressed in the current coexistence of at least three major, competing taxonomic treatments. To help resolve this situation, we sequenced ~900 bases of mitochondrial DNA of 40 individuals from five of the widely recognized “major” allotaxa. Total sequence diversity (>5%) is high compared to most primate species. Major mitochondrial clades correspond to recognized allotaxa, with the important exception that haplotypes from yellow and olive baboons form a single, monophyletic clade within which the two allotaxa do not comprise mutually exclusive clusters. The major clades fall unambiguously

into the pattern: (chacma (Guinea (hamadryas (yellow + olive))). This phylogeny does not support taxonomies that oppose hamadryas to all other baboons (“desert” vs. “savanna”), but is compatible with the view that all definable allotaxa should be recognized as coordinates, either as “phylogenetic” species or “biological” subspecies. The close relationship and unsegregated distribution of haplotypes from Kenyan and Tanzanian yellow and olive baboons are unexplained, but may reflect introgression across the documented hybrid zone. The overall phylogeny, when combined with paleontological data, suggests a southern African origin for extant *Papio* baboons, with all extant lineages sharing a common mitochondrial ancestor at approximately 1.8 Ma. *Am J Phys Anthropol* 124:17–27, 2004.

© 2004 Wiley-Liss, Inc.

Baboons of the genus *Papio* are distributed throughout most of sub-Saharan Africa, outside of the lowland equatorial forest belt (Fig. 1). Parapatric, morphologically definable geographic populations, henceforth called allotaxa (Grubb, 1999), are numerous (Hill, 1970; Jolly, 1993), exhibit major differences in social organization (Kummer, 1968, 1995; Altmann, 1980; Anderson, 1983), and have adapted to a wide range of habitats (Popp, 1983). Baboons are among the most studied of nonhuman primates, yet understanding and communication are hampered by a lack of taxonomic consensus (Groves, 2001). There is general, though not universal (Jolly, 1993, 2001), agreement that five “major forms” can be recognized within *Papio*, which are most unambiguously labeled by their vernacular names: chacma, Guinea (or red), olive (or anubis), yellow, and hamadryas (or sacred). One question, not addressed by this paper, is whether some allotaxa additional to the “basic five” (e.g., “grey-footed baboons” traditionally included among chacmas, and “kinda baboons,” generally lumped with yellows) merit similar recognition (Jolly, 1993, 2001). This problem aside, currently competing baboon taxonomies imply disagreements of two kinds. The first is whether the recognizable allotaxa should be regarded as full species (*Papio ursinus*, *P. papio*, *P.*

anubis, *P. cynocephalus*, and *P. hamadryas*, respectively) or as subspecies of a single, polytypic species, which by priority must be called *P. hamadryas*. This decision depends upon taxonomic taste, and in particular, upon the choice between “biological” and “phylogenetic” species concepts. The case for and against species rank is discussed by Jolly (1993) and Groves (2001), and is unaffected by the present analysis.

A second disagreement concerns relationships among the recognizable allotaxa within *Papio*. The striking morphological and behavioral uniqueness of the hamadryas baboon has often been recognized by allocating it to a distinct species (e.g., Thorington and Groves, 1970; Richard, 1985; Smuts et al., 1986), superspecies (Hill, 1967), or even genus (see

Grant sponsor: NIH; Grant numbers: HL28972, RR13986.

*Correspondence to: Tim Newman, Laboratory of Neurogenetics, NIAAA/NIH, 12420 Parklawn Drive, Park 5 Building, Room 451, Rockville, MD 20852. E-mail: tknewman@mail.nih.gov

Received 8 October 2002; accepted 28 April 2003.

DOI 10.1002/ajpa.10340

Published online 20 August 2003 in Wiley InterScience (www.interscience.wiley.com).

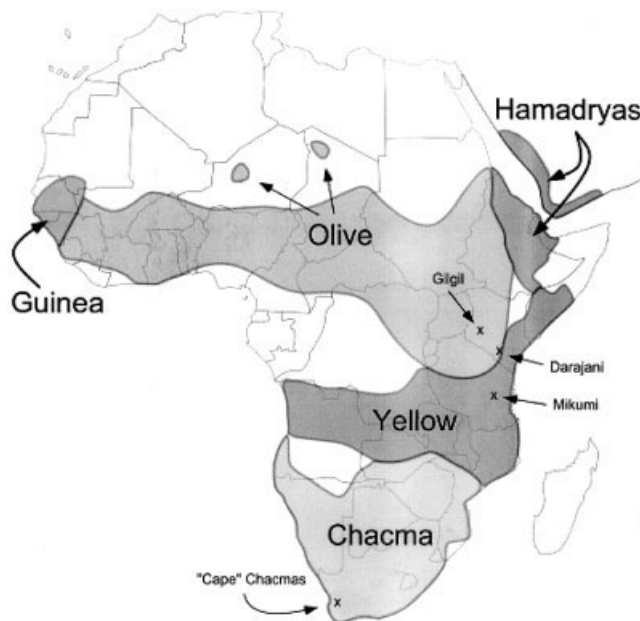


Fig. 1. Map of Africa, showing approximate distribution of five allotaxa used in this study. "X" refers to trapping locations for samples of olive baboons from Gilgil, olive and yellow baboons from Darajani, and yellow baboons from Mikumi National Park. For chacma baboon samples, "X" indicates approximate geographic location for "Cape" chacma baboons. See text for further discussion of animals sampled.

discussion by Roth, 1965), and lumping all other forms in a coordinate taxon. One version of this two-taxon scheme, in which hamadryas (as *Papio hamadryas*) is opposed at species rank to all other forms (lumped as *P. cynocephalus*) (Buettner-Janusch, 1966), has wide currency, and seems especially favored by socioecologists (e.g., Smuts et al., 1986 and references therein). Other relationships have been suggested. Ellerman et al. (1953) considered chacma and olive baboons conspecific, and distinct from both hamadryas and yellow baboons (they did not discuss Guinea baboons). While maintaining their specific distinctness, Kingdon (1997) also suggested a sister-group relationship between olive and chacma baboons. Several authors (e.g., Hill, 1970) pointed out phenotypic and behavioral traits linking hamadryas and Guinea baboons, while one study (Williams-Blangero et al., 1990) found that Guinea baboons were the most divergent of the five allotaxa. Jolly (1965, 1993) suggested that the primary phylogenetic split within *Papio* might separate the respective ancestors of forms with a mainly northern distribution and manes of waved hair (Guinea, olive, hamadryas) and southern forms lacking manes (chacmas, yellows).

While molecular data have been widely used to investigate the population structure of wild baboon populations (see review by Rogers, 2000), including a naturally occurring hybrid zone (Newman, 1997; Brett et al., 1977; Woolley-Barker, 1999), little is known about patterns of genetic differentiation

among baboon populations. One important study by Williams-Blangero et al. (1990) compared allele frequencies from nine blood proteins in all five generally recognized baboon taxa. Although their study was not phylogenetic in nature, they did generate a two-dimensional measure of genetic distance between taxa suggesting that, in general, olive, yellow, and hamadryas baboons were more closely related to one another than to chacma or Guinea baboons. Here, we use nucleotide sequences of mitochondrial DNA (mtDNA) from animals representing each of the five generally recognized allotaxa to assess these alternative phylogenetic hypotheses.

MATERIALS AND METHODS

Subjects

Most of the 40 *Papio* samples used in this study came from individuals housed at the Southwest Foundation for Biomedical Research (SFBR) in San Antonio, Texas (Table 1). Nine captive-born olive and yellow baboon subjects were derived from stocks captured in southern and western Kenya more than 25 years ago (Maples and McKern, 1967). Four wild-caught olive baboons were imported from an area around Gilgil, Kenya, and five wild-caught yellow baboon samples (not SFBR colony animals) were collected in Mikumi National Park, Tanzania (Rogers and Kidd, 1993). Provenience for the six chacma baboons was unavailable, but morphology and pelage were typical for the nominate South African form from the Cape. Guinea baboons were derived from a captive colony once held at the San Antonio Zoo (see Maples and McKern, 1967, p. 21). Although no further information on provenience was available, visual inspection of three of the Guinea baboons alive at the time this study was undertaken suggested typical pelage and morphology for this taxon. For all captive-born animals, colony records were used to ensure that none of the study animals had mixed ancestry or were closely related. As representatives of outgroup taxa, we included four gelada baboon (*Theropithecus gelada*) samples, obtained from the San Antonio Zoo, and a published sequence (Hayasaka et al., 1996) for the rhesus macaque (*Macaca mulatta*).

Laboratory methods

DNA was isolated from whole blood, hair follicles, or (in the case of archived tissues from deceased animals) frozen liver or blood cells, using a commercial DNA extraction kit (Gentra Systems, Inc., Minneapolis, MN) or standard proteinase K and phenol-chloroform extraction (Rogers and Kidd, 1993; Hillis et al., 1996). Two baboon samples, one olive and one hamadryas, were extracted by a method that enriches the ratio of mtDNA to nuclear DNA (Arnason et al., 1991). These were used as a control for detecting spurious PCR products resulting from nuclear pseudogene copies of functional mtDNA genes.

TABLE 1. Specimen localities and haplotype designations

Animal ID	Taxon	Common name	Locality of origin	Animal location	Haplotype ⁵	GenBank accession no.
3X0001	<i>Papio hamadryas ursinus</i>	Chacma	Southern Africa	SFBR ²	U1*	AY212057
3X0007	<i>P. h. ursinus</i>	Chacma	Southern Africa	SFBR	U1	
3X0010	<i>P. h. ursinus</i>	Chacma	Southern Africa	SFBR	U2*	AY212058
3X0020	<i>P. h. ursinus</i>	Chacma	Southern Africa	SFBR	U3*	AY212059
3X0021	<i>P. h. ursinus</i>	Chacma	Southern Africa	SFBR	U2	
3X0023	<i>P. h. ursinus</i>	Chacma	Southern Africa	SFBR	U1	
8058	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P1*	AY212049
7411	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P2*	AY212050
8282	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P3*	AY212051
5X0015	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P4*	AY212052
9915	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P5*	AY212053
9882	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P6*	AY212054
9643	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P7*	AY212055
11391	<i>P. h. papio</i>	Guinea	West Africa	SFBR	P8*	AY212056
2X0120	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia ¹	SFBR	H5	
8813	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H2 ⁴	
8247	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H1*	AY212034
8295	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H2*	AY212035
6919	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H4	
9430	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H3*	AY212036
8480	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H4*	AY212037
9047	<i>P. h. hamadryas</i>	Hamadryas	Africa or Arabia	SFBR	H5*	AY212038
1X1146	<i>P. h. anubis</i>	Olive	Kenya	SFBR	A1*	AY212040
1X4853	<i>P. h. anubis</i>	Olive	Kenya	SFBR	A2*	AY212041
6521	<i>P. h. anubis</i>	Olive	Kenya	SFBR	A6	
9362	<i>P. h. anubis</i>	Olive	Kenya	SFBR	A2	
9347	<i>P. h. anubis</i>	Olive	Kenya	SFBR	A2	
7126	<i>P. h. anubis</i>	Olive	Kenya	SFBR	A3*	AY212042
14702	<i>P. h. anubis</i>	Olive	Gilgil, Kenya	SFBR ³	A6	
14710	<i>P. h. anubis</i>	Olive	Gilgil, Kenya	SFBR ³	A4*	AY212097
14722	<i>P. h. anubis</i>	Olive	Gilgil, Kenya	SFBR ³	A5*	AY212098
14739	<i>P. h. anubis</i>	Olive	Gilgil, Kenya	SFBR ³	A6*	AY212099
1X3337	<i>P. h. cynocephalus</i>	Yellow	Kenya	SFBR	C1*	AY212043
9536	<i>P. h. cynocephalus</i>	Yellow	Kenya	SFBR	C2*	AY212044
1X4315	<i>P. h. cynocephalus</i>	Yellow	Kenya	SFBR	C3*	AY212045
IK05mk	<i>P. h. cynocephalus</i>	Yellow	Mikumi NP, Tanzania	Wild	C4*	AY212046
3117mk	<i>P. h. cynocephalus</i>	Yellow	Mikumi NP, Tanzania	Wild	C5*	AY212047
2008mk	<i>P. h. cynocephalus</i>	Yellow	Mikumi NP, Tanzania	Wild	C6*	AY212048
5002mk	<i>P. h. cynocephalus</i>	Yellow	Mikumi NP, Tanzania	Wild	C6	
5012mk	<i>P. h. cynocephalus</i>	Yellow	Mikumi NP, Tanzania	Wild	C6	
T1	<i>Theropithecus gelada</i>	Gelada	Central Ethiopia	San Antonio Zoo	T1*	AY212060
T2	<i>Theropithecus gelada</i>	Gelada	Central Ethiopia	San Antonio Zoo	T1	
T3	<i>Theropithecus gelada</i>	Gelada	Central Ethiopia	San Antonio Zoo	T1	
T4	<i>Theropithecus gelada</i>	Gelada	Central Ethiopia	San Antonio Zoo	T1	

¹ "Africa" refers to Ethiopia, Eritrea, and Somalia; "Arabia" includes areas in Yemen and Saudi Arabia.

² SFBR, Southwest Foundation for Biomedical Research.

³ Wild-caught, currently housed at SFBR.

⁴ A single hamadryas sequence (H2) contained stop codons and other mutations that differed markedly from the mitochondrially enriched hamadryas control sequence (see Methods), suggesting that it may be a nuclear pseudogene. This sequence was not included in phylogenetic analyses.

⁵ Starred specimens are designated "type" sequences, as submitted to GenBank.

The target region within the mitochondrial genome (the "896" or "Brown" region) was previously shown (Newman, 1997) to include phylogenetically informative, intraspecific variation in baboons. It is bounded by two *Hind*III restriction sites and spans portions of the ND 4 and ND 5 genes, as well as three tRNAs (Brown et al., 1982; Anderson et al., 1981).

A test panel (N = 14) composed of two samples from each of the five analyzed subspecies, one sample from each of the outgroup taxa (gelada and macaque), a human sample, and a DNA-free negative control was assembled for optimizing PCR conditions and designing baboon-specific primers. Initially, an ~1,100 base-pair fragment was PCR-am-

plified from total genomic DNA with primers L11574 and H12569 (see Table 2 for all primer sequences and reaction conditions) in a Perkin-Elmer 9700 Thermocycler, using the following cycling conditions: 94°C denature for 2 min (94°C, 15 sec; 56°C, 15 sec; and 72°C, 30 sec) for 30 cycles, plus a final 72°C extension for 3 min. Amplification reactions contained 0.75 units *Taq* polymerase (Promega Corp., Madison, WI), 1× manufacturer's buffer (Promega Corp.), 2.5 mM final concentration MgCl₂, 5 μM each primer, and 50 ng template DNA in a 25-μl final volume reaction. Amplicons were separated by electrophoresis in 1.2% TBE agarose gels, and visualized by staining with ethidium bromide. Bands were excised from the gel and purified using a gel

TABLE 2. Primer sequences and reaction conditions

Primer name ¹	Primer sequence	Reaction conditions	
		Annealing temperature	Cycle parameters
L11574	5'-CTATCCCTATGAGGGATAATTATAAC-3'	56°C	94-15", 56-15", 72-30"
H12652	5'-AATGTTTGGGCTGAGTTTATATATCA-3'	56°C	94-15", 56-15", 72-30"
L896-275F	5'-TAACATACTAATCACAGCCCTC-3'	56°C	94-15", 56-15", 72-30"
A896LF	5'-CTAGTAATGTAGCCTCCCTC-3'	58°C	95-10", 58-10", 72-25"
A896HR	5'-CACAGTCTAACGTTTTGATTA-3'	58°C	95-10", 58-10", 72-25"
B896LF	5'-CAACTATTCTCCTATCTCTCAACC-3'	58°C	95-10", 58-10", 72-25"
B896HR	5'-TAGACCAGGTAATGAATAGTGC-3'	58°C	95-10", 58-10", 72-25"

¹ L, light strand; H, heavy strand; F, forward; R, reverse; A, fragment A; B, fragment B. For first two primers, numbers following L or H refer to position of 3' end of primer in human mtDNA sequence by Anderson et al. (1982). Primers L11574 and H12652 produce a fragment ~1,150 bp that spans "896" region. Internal primers A896LF/HR and B896LF/HR produce fragments ~550 bp each and overlap one another by ~100 bp. Magnesium chloride final concentration for all reactions was 2.5 mM.

purification kit (Qiagen, Inc., Valencia, CA). A volume equivalent to 20 ng of purified product (typically 3-5 μ l) was used as a template for cycle sequencing, using FS Dye Terminators (Applied Biosystems, Foster City, CA), following the manufacturer's protocols and using L11574 and H12569 as sequencing primers. Sequences were visualized using an ABI 373 automated sequencer (Applied Biosystems). A new primer, L896-275F, was designed to initiate sequencing approximately 275 bases downstream from L11574 in order to confirm terminal sequences produced by L11574 and H12569. To facilitate sequencing of the entire panel of 44 individuals, full-length sequences (~1,100 bp) from the test panel were aligned and used to design two sets of nested, baboon-specific primer pairs. These primers amplify two overlapping fragments of ~550 bases, resulting in 900 total bases of sequence, including the two *Hind*III sites. Subsequent cycle-sequencing reactions were carried out directly from PCR products. Unincorporated primers and nucleotides were removed using a PCR Purification Kit (Qiagen, Inc.). Nucleotide sequences were determined using BigDye Terminators (Applied Biosystems) and an ABI 377 Automated Sequencing System. All sequences were verified by sequencing both the forward (L) and reverse (H) strands. Finally, the full set of sequences was assembled using DNASIS-Mac version 2.0 (Hitachi, Ltd.), and aligned by eye.

Phylogenetic methods

Once aligned, unique sequences were imported into MacClade 3.08 (Maddison and Maddison, 1992) and PAUP* 4.0b10 (Swofford, 1999), and their phylogenetic relationships were determined by maximum parsimony, maximum likelihood, and distance methods. For parsimony analysis, we found all trees of the shortest length by implementing PAUP's "Branch and Bound" (B & B) algorithm. All sites were weighted equally, and gaps were treated as missing data. Additional analyses employing various weighting schemes produced inconsequential differences in tree topology. Support for the branching structure of the parsimony tree was evaluated by 1,000 bootstrap pseudoreplicates, using tree bisec-

tion reconstruction (TBR) branch-swapping with 100 random addition replicates.

Maximum likelihood (ML) analysis was conducted by selecting one tree from several equally parsimonious trees produced using heuristic methods (TBR branch-swapping, simple addition sequence). That tree was then used as the starting point for an iterative search strategy. Under PAUP's maximum likelihood options, the transition/transversion (ti/tv) ratio for the substitution model, gamma distribution (α) for among-site rate variation, and proportion of invariable sites (p^{inv}) were estimated. Base frequencies were determined empirically. Using the resulting estimated parameters, a heuristic ML search was then conducted (equal weights, as-is addition, TBR branch-swapping). This tree was then used to reestimate model parameters, and the process was repeated until the $-\ln L$ score did not change.

For distance-based analyses, matrices of pairwise distances between haplotypes were constructed using distances uncorrected for multiple substitutions, as well as two common models of substitution: HKY85 (Hasegawa et al., 1985) and the two-parameter model of Kimura (1980). Trees based on calculated distances were constructed and evaluated using PAUP's neighbor-joining (NJ) algorithm.

Five constraint trees depicting the five morphologically defined taxa, arranged according to the competing taxonomic hypotheses described above, were constructed in MacClade then imported into PAUP. $-\ln L$ scores for each tree were determined and compared statistically, using the likelihood ratio test of Kishino and Hasegawa (1989). This test sums the difference in log likelihoods for each site between two trees, and assumes that if the two trees differ significantly, then the sum of the likelihood differences will be significantly different from zero (Page and Holmes, 1998). This test method was also applied to parsimony scores (numbers of steps) between pairs of trees. Again, the null hypothesis is that the sum of differences at each site between trees is not significantly different from zero if the trees are not significantly different.

Age calibration

Fossil evidence (Delson, 1993; Jablonski, 1993 and references therein) suggests that *Theropithecus* and *Papio* appeared as distinct evolutionary lineages by 3.5–4.0 million years ago (Ma). By using the observed minimum sequence divergence of 11.4% between *Theropithecus gelada* and *Papio* mtDNA lineages, we estimated an approximate rate of sequence divergence of 2.85–3.26% per million years in this clade. This is in general agreement with the rate of sequence divergence (2–4%) reported for the “896” region in other mammalian taxa (Hayasaka et al., 1996; Kocher et al., 1989; Brown et al., 1979). We then used these values as a local molecular clock to estimate the age of each node in the phylogeny of *Papio* haplotypes. Although it is typically unnecessary to test whether the rate of sequence variation between subspecies or closely related species would violate the hypothesis of a molecular clock, we nevertheless tested for this, using a likelihood ratio test where $2\Delta = \log L_{\text{no clock}} - \log L_{\text{clock}}$ (Page and Holmes, 1998). We constructed two ML trees, one with and one without a molecular clock enforced, but otherwise with identical settings. The resulting value can be tested statistically, since it is distributed as a χ^2 with $(n - 2)$ degrees of freedom, where n = the number of branches in the trees.

RESULTS

Phylogenetics

Twenty-nine unique sequences were detected among the 44 DNA samples examined (Table 1). Every population was polymorphic, no haplotype was found in more than one population, and all phylogenetic analyses revealed tree topologies that were characterized by four major clades.

Of 896 bases, 142 were variable, and 106 were parsimony-informative. Branch and bound analysis returned 24 equally parsimonious trees (329 steps, consistency index = 0.80, retention index = 0.87). Figure 2 shows a strict consensus of the 24 trees, with bootstrap values above branches and the number of substitutions below branches. All haplotypes from chacma (U1, U3, and U4), Guinea (P1–P8), and hamadryas (H3–H6, and H6–H8) baboons cluster into well-supported monophyletic groups. Another strongly supported clade (five synapomorphies) unites all olive and yellow baboon haplotypes. This clade exhibits little internal resolution, except that yellow baboon haplotypes from Mikumi in central Tanzania consistently form a monophyletic subclade defined by three synapomorphies. The basal structure of the tree, which takes the form (chacma (Guinea (hamadryas (yellow/olive))), is remarkably robust, with 78–100% bootstrap support on all major branches. All differences among alternative, equally parsimonious trees occur in the terminal nodes, primarily within the clade that includes sequences from olive and yellow baboons.

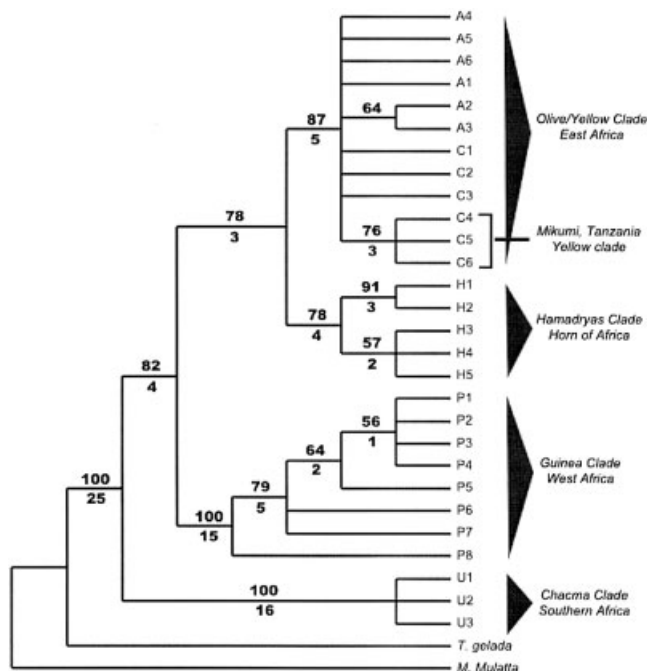


Fig. 2. Phylogenetic relationships among *Papio* haplotypes, based on a strict consensus of 24 trees, all 329 steps long, derived using a “Branch and Bound” algorithm with sites unweighted as implemented in PAUP* 4.0b10. Numbers above branches are bootstrap scores, and numbers below branches are numbers of synapomorphies supporting each major clade. See Table 1 for taxa abbreviations. General geographic provenience is indicated for each major clade.

Maximum likelihood analysis required three iterations for the estimation of model parameters, resulting in a single ML topology ($-\ln L = 2,734.563$, $Ti/Tv = 9.945$, $\alpha = 0.267$, $p^{inv} = 0.0402$), with a branching structure nearly identical to the B&B tree. Distance-based NJ trees were nearly identical in branching order and branch length regardless of substitution model, as well as exhibiting nearly identical major branching structure to both parsimony and ML analyses. Chacma haplotypes once again comprise the sister group to all other *Papio* types, with Guinea baboon haplotypes the next internal lineage to diverge. The clade containing all olive and yellow baboon haplotypes has much shorter branch lengths than the clades of Guinea and chacma haplotypes, which exhibit more than 2–3 times the number of changes along their branches.

Comparisons of sequence diversity among all samples as measured by percent pairwise distances, uncorrected for multiple substitutions, are presented in Table 3. Values along the diagonal are the mean percent pairwise differences among haplotypes within a clade, and the other values are the mean percent pairwise differences between clades. These values, condensed and grouped by allotaxa for clarity, are nearly identical to those resulting from the application of the two-parameter model of Kimura (1980) and the HKY85 substitution model. Intra-

TABLE 3. Mean percent pairwise matrix

	Olive/yellow	Hamadryas	Guinea	Chacma	Gelada
Olive/yellow	0.0049				
Hamadryas	0.0188	0.0085			
Guinea	0.0431	0.0401	0.0066		
Chacma	0.0483	0.0495	0.0597	0.0090	
Gelada	0.1161	0.1140	0.1291	0.1258	

clade variation ranges from 0.49% in the olive/yellow clade to 0.90% in the chacma clade. Average pairwise distances between clades range from 1.88% between the olive/yellow and hamadryas haplotypes to 5.97% between the chacma and Guinea haplotypes. The latter value exceeds within-species sequence variation for the same mitochondrial region reported in several species of *Macaca* (Hayasaka et al., 1996), such as crab-eating macaques (*M. fascicularis*, 3.9%), Japanese macaques (*M. fuscata*, 1.5%), and rhesus macaques (*M. mulatta*, 4.2%). However, this diversity is not unprecedented in primates (e.g., chimpanzees, *Pan troglodytes*; Morin et al., 1994; Gagneux et al., 1999).

In Table 4 are the results of comparing the preferred mtDNA tree with trees that represent the four alternative taxonomic hypotheses described earlier. Both ML and parsimony scores support the phylogeny presented in Figure 2. All other taxonomic schemes score significantly worse ($P < 0.001$). The next best tree (10 steps longer) is one in which hamadryas and Guinea lineages form a clade that is sister to an ((olive/yellow) chacma) clade, followed by the socioecological model in which hamadryas forms the sister taxon to all other *Papio* (12 steps longer).

Age calibration

The difference in log L scores between the ML trees calculated with and without a molecular clock enforced ($2\Delta = 26.43$) was not statistically significant ($\chi^2_{(28, 0.05)} = 41.33$), and thus we assume our sequences to be suitable for use in calibrating a local molecular clock. Using a *Papio-Theropithecus* divergence date of 3.5–4.0 million years yields the estimates for internal divergences that are shown against the nodes in Figure 3. The initial divergence of chacma baboon haplotypes from the common stem yields a date of 1.79 (1.69–2.09) Ma. The lineage leading to Guinea baboons diverged 1.37 (1.23–1.51) Ma, and the clade formed by hamadryas and olive/yellow lineages diverged 619 (577–660) Ka. The mean percent pairwise differences among haplotypes within the olive/yellow clade suggests a coalescent age of 161 (150–172) Ka. (Note added in proofs: Recently refined estimates of *Papio-Theropithecus* mtDNA distance, based on unpublished data, suggest that divergences among lineages within *Papio* may be slightly older (at most 5–7%) than those reported here. The adjusted dates lie within the confidence limits reported above.)

DISCUSSION

Phylogeny

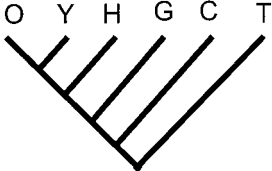
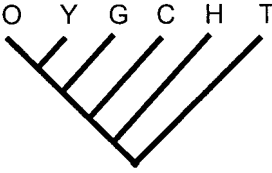
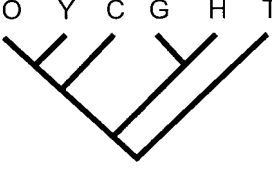

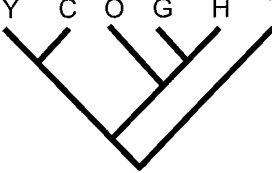
The mtDNA sequence data presented here provide new insights into the evolutionary relationships among the five “major forms” of *Papio* baboons. Our results suggest that chacma baboons are the sister group to all other extant baboon populations. This cladistic arrangement was not proposed in any previous study, but we know of no suite of characters that provides strong evidence against this conclusion (see discussion of comparative baboon morphology in Jolly, 1993). Our findings do not support exclusive relationships either between chacma and olive baboons, or between hamadryas and Guinea baboons. The character states cited by Ellerman et al. (1953) (large size, dark pelage) as linking chacmas and olives are superficial and easily explained as homoplasies. The Guinea-hamadryas resemblances (small size, arched rather than kinked tail, well-developed mane, multiringed hairs, and tendency to form permanent male-female bonds within a multimale society) are more numerous and seem less likely to be homoplastic. They merit attention as possible plesiomorphies retained from an early, though not necessarily basal, ancestor.

The study of protein variation by Williams-Blangero et al. (1990), like the present analysis, found close similarity among olive, yellow, and hamadryas baboons. The major difference between our results and those of Williams-Blangero et al. (1990) is that the protein data suggest that Guinea baboons are the most divergent of all five “major forms.” This inferred evolutionary distance between Guinea baboons and other forms is driven by the presence at high frequency among Guinea baboons of unique (private) alleles at 2 of the 9 loci surveyed. In one of those cases (GPI), all Guinea baboons were fixed for one allele, while all other baboons in the study shared a different allele. Though clearly a meaningful difference, these private alleles in Guinea baboons may be derived, autapomorphic character states for baboons, and do not necessarily link the remaining four “major forms” phylogenetically.

Neither the first- nor the second-ranking tree (Table 4) supports the two-taxon, hamadryas-vs.-the-rest taxonomic scheme (Simpson, 1945; Buettner-Janusch, 1966; Thorington and Groves, 1970; Smuts et al., 1986). Allocation of hamadryas baboons to one species, with placement of the other four allotaxa into a second species (*P. cynocephalus*, sensu lato), establishes the latter as a nonholophyletic, “wastebasket” taxon for all “savannah” baboons. Such usage is discouraged by current taxonomic practice.

The sequence and age of nodes on the mtDNA tree have zoogeographic implications. Molecular (Disotell, 1996) and fossil (Delson, 1993; Jablonski, 1993) evidence suggests that the lineages leading respectively to extant *Papio*, to the gelada baboon (*Theropithecus gelada*), and to the black mangabey

TABLE 4. Constraint tree tests of competing taxonomic hypotheses for Papio

Hypotheses	Trees ¹	K-H likelihood ratio test			K-H and Templeton parsimony test						
		-ln L	$\Delta -\ln L$	<i>P</i>	Length	Δ length	SD difference	<i>t</i>	<i>P</i>	Templeton	
1) Molecular based: simplified cladogram representing four main mtDNA lineages.		2,854.192	(Best)		336	(Best)					
2) Socioecological model (Beuttner-Janusch (1966), Smuts et al. (1986), and others) hamadryas is sister taxon to all other lineages (((O-Y)G)C)H).		2,877.087	22.895	<0.01*	348	12	3.44	3.49	<0.01*	<0.01*	
3) Hill (1967), and others: Guinea/hamadryas clade is sister to olive, yellow and, chacma clade ((O-Y)C)(H-G)).		2,874.020	19.827	<0.01*	346	10	3.73	2.68	<0.01*	<0.01*	
4) Ellerman, et al. (1953) and Kingdon (1971): olive/chacma clade is sister to hamadryas, Guinea, and yellow clade (O-C)/((H-G)Y).		2,891.016	36.824	<0.01*	353	17	4.97	3.42	<0.01*	<0.01*	
5) Jolly (1993): north/south split: yellow/chacma clade is sister to olive, hamadryas, and Guinea clade ((Y-C)(O(G-H))).		2,891.016	36.824	<0.01*	352	16	4.87	3.28	<0.01*	<0.01*	

¹ O, olive; Y, yellow; H, hamadryas; G, Guinea; C, chacma; T, gelada.

* Statistically significantly worse than molecular-based taxonomic model.

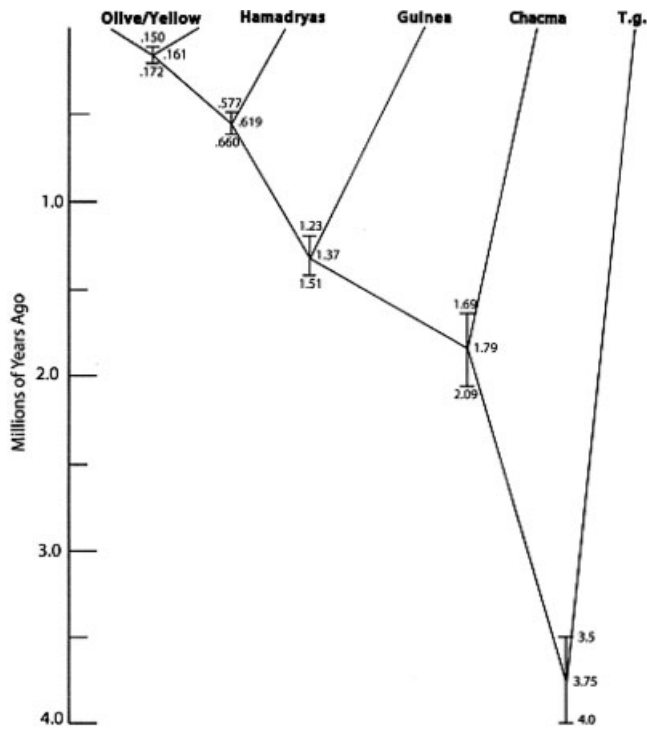


Fig. 3. Phylogram showing estimated divergence dates for four major mitochondrial lineages within *Papio*.

genus (*Lophocebus*) diverged almost simultaneously, and that this divergence had occurred by about 4 Ma. The Plio-Pleistocene fossil record suggests that after this divergence, *Papio* and its probable close relatives, *Parapapio* and *Dinopithecus*, were relatively rare in East Africa, but were common and diverse in southern Africa. Fossils of *Papio* proper first appear in South Africa about 2.5 Ma, and document a basal radiation within the genus that includes species (e.g., *P. izodi*) thought to be sister taxa to all extant members of the genus (Delson, 1984). When combined with the basal divergence of the chacma mitochondrial clade at ~2 Ma, these considerations suggest that the species ancestral to all extant *Papio* inhabited southern Africa. In this case, the primary divergence presumably resulted from the isolation, at the northerly end of this range, of an offshoot population from which all non-chacma populations were ultimately derived. The next divergence, which separates Guinea baboons from the yellow-olive-hamadryas clade, may have occurred as an east/west bifurcation, within this northern expansion.

Phylogeographic complications

While most haplotypes in our sample cluster into clades coincident with traditional, pelage-based taxa, the East African yellow and olive baboons are an exception, forming a single, undifferentiated, and relatively shallow-rooted clade, within which Mikumi yellow baboons form a minor subgroup. We can

suggest three possible explanations for this lack of separation between the two allotaxa:

1. The observed simplicity throughout the rest of the phylogeny is an artifact of sampling bias. Overlapping, nontaxon-specific mitochondrial clades may be characteristic of other, or perhaps all, pairs of baboon allotaxa, but were not found in our study because of limited geographic sampling of mtDNA from other allotaxa.
2. Yellow and olive baboons, as whole taxa, are more closely related to each other than to any other pair of taxa in our analysis, and the lack of differentiation between them reflects polymorphism retained from a relatively recent, common ancestral population.
3. The observed pattern is the result of local gene flow in East Africa between differentiated yellow and olive baboons.

The first of these possibilities cannot be excluded on the basis of our data, and will require testing against much larger samples and more geographically dispersed representatives of all taxa. However, it is very unlikely that the complexity of the yellow-olive clade is the result of using captive animals from the SFBR colony, because 1) several wild-caught individuals of each subspecies were used (4 wild-caught olive and 5 wild yellow baboons), and 2) only captive-born animals whose pedigree records unambiguously indicated subspecies identity were included. It should also be noted that the shallow root of the olive-yellow clade could not be duplicated by artificially combining any other pair of samples. The haplotype lineages in the olive-yellow clade are not simply mixed; they are very closely related. We believe that this first explanation is unlikely.

Similarly, the second possibility cannot be definitively excluded without more information, but seems implausible as well. Yellow and olive baboons are widely distributed, morphologically distinct forms, yet in our sample their most similar haplotypes differ by only a single base-pair substitution. If mitochondrial diversity faithfully reflected a population history of simple divergence, a single, panmictic population from which both yellows and olive are descended would have lived only 150,000–175,000 years ago, less than one-tenth the time that has elapsed since the earliest (chacma vs. the rest) mitochondrial divergence documented in our data. Moreover, if the yellow-olive-hamadryas cluster of populations (as well as their mitochondria) comprises a true clade, this implies homoplasy in pelage features. If the stem form of the clade had a mane (like hamadryas and olive baboons), yellow baboons must have acquired their “southern” pelage features secondarily. Conversely, if the stem form resembled a yellow baboon in pelage, the maned pelage that hamadryas and olive share with Guinea baboons was an independent, parallel development. Either of these scenarios seems unlikely, but they will be crit-

ically tested only when samples from other, geographically distant populations (e.g., West African olive baboons and yellow baboons from south and southwest of Mikumi) are examined. If they do not reveal greater mtDNA diversity, then a recent common ancestry for the yellow-olive-hamadryas cluster will be suggested.

The third explanation, that the observed pattern is a local phenomenon related to comparatively recent hybridization at the border between taxa, seems to us the most plausible on present evidence. Since mtDNA does not recombine and is maternally inherited, it can retain a signal of past episodes of hybridization (especially sex-biased gene flow) that have been phenotypically obscured through backcrossing. This process has been widely documented in mammals, including wild canids (Lehman et al., 1991), deer (Carr et al., 1986), and Scandinavian house mice (Gyllensten and Wilson, 1987). In macaques, Melnick et al. (1993), Evans et al. (1999), and others described patterns of mitochondrial gene flow and introgression that reflect complex phylogeography in species such as *Macaca mulatta*, *M. nemestrina*, *M. tonkeana*, and other closely related species. Indeed, Tosi et al. (2000) argued that one macaque species (*Macaca arctoides*) may have arisen as a result of hybridization between *M. assamensis/thibetana* and *M. fascicularis*, based on topological discrepancies between mtDNA and Y chromosome-based phylogenetic analyses.

Yellow and olive baboons are known to meet and hybridize at many sites along an extensive contact zone that runs southwest to northeast across Tanzania and Kenya (Kingdon, 1971; Maples and McKern, 1967; Samuels and Altmann, 1986). The SFBR yellow baboons were captured within a few tens of miles of the boundary, near Darajani, Kenya (C. Bramblett, personal communication), where hybridization was actively occurring at that time (Maples and McKern, 1967). Moreover, populations of yellow baboons in this region (sometimes distinguished as the "ibean" baboon, *Papio ibeanus* or *P. cynocephalus ibeanus*) exhibit pelage features recalling those of olive baboons, suggesting a long history of sporadic introgression. If this interpretation is correct, all the haplotypes seen in our yellow and olive baboons originated in only one of the two taxa, and are found in the other as a result of secondary introgression, either of individual haplotypes, or of one or more ancestral types that subsequently diversified. Much more work in the field will be needed to clarify the details of the situation, especially to determine 1) whether the observed cluster of yellow-olive haplotypes originated in a yellow or an olive population, 2) how far from the allotaxon border the "invading" haplotype(s) spread, and 3) whether indigenous haplotypes also persist in the recipient population. Clearly, the phylogeography of mtDNA diversity in olive and yellow baboons can only be fully assessed by sampling these allotaxa over a

much broader geographic range, and especially from sites far removed from the East African hybrid zone. Olive baboons from West and Central Africa (Uganda to Sierra Leone) and yellow baboons from Zambia, southern Tanzania, and Mozambique will be especially interesting.

CONCLUSIONS

Considering the genus as a whole, there is no reason to believe that the genetic complexity of the olive-yellow clade in East Africa is unique. All baboon allotaxa appear to be capable of producing viable and fertile offspring when crossed. Active, natural, and genetically complex hybrid zones have been found wherever the critical interfaces have been investigated in the field (Jolly, 1993). Moreover, in historical perspective, the present interglacial ("pluvial") period is merely the most recent interval in a series of climatic cycles extending back at least 1.5 Ma (Bromage and Schrenk, 1999). By alternately shrinking and extending population sizes and ranges, and successively promoting and restricting interdemographic gene flow, these cycles must have drastically altered the composition of baboon gene pools in ways analogous to processes documented in the better-known fauna of glacial Eurasia and North America (Hewitt, 1996). In this regard, it should be noted that several distinct baboon allotaxa have yet to be genetically investigated. These include the gray-footed chacma baboons, which are geographically and morphologically intermediate between "typical" chacmas and yellows, and Kinda baboons, which in some respects are extreme and miniaturized versions of the yellow phenotype. We are confident that more detailed information will not overturn the outline of baboon mitochondrial phylogeny established here, but strongly suspect that many more complications remain to be discovered in its details.

ACKNOWLEDGMENTS

We thank the San Antonio Zoo for permission to collect hair and blood samples from their captive gelada baboons for DNA extraction, and Dr. Michelle Leland of the Southwest Foundation for Biomedical Research for her help in coordinating the collection. Thanks go to Mary Sparks (SFBR) for her immeasurable help in sorting through sometimes very old data records in the process of selecting samples for this analysis. We also thank Eric Delson, Todd Disotell, Karen Chambers, Tim Anderson, and especially Derek Wildman for stimulating comments and thought-provoking criticisms on earlier drafts of this manuscript. Thanks to Kirstin Sterner for access to unpublished gelada sequence. We also thank two anonymous reviewers for their helpful comments. This research was supported in part by NIH grants HL28972 and RR13986 to J.R.

LITERATURE CITED

- Altmann J. 1980. Baboon mothers and infants. Cambridge, MA: Harvard University Press.
- Anderson CM. 1983. Levels of social organization and male-female bonding in the genus *Papio*. *Am J Phys Anthropol* 60:15–22.
- Anderson S, Bankier AT, Barrell BG, et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.
- Arnason U, Gullberg A, Widegren B. 1991. The complete nucleotide sequence of the mitochondrial DNA of the fin whale *Balaenoptera physalus*. *J Mol Evol* 33:556–568.
- Brett FL, Jolly CJ, Socha WW, Wiener AS. 1977. Human-type ABO blood groups in Ethiopian baboons. *Yrbk Phys Anthropol* 20:276–289.
- Bromage TG, Schrenk F, editors. 1999. African biogeography, climate change, and human evolution. New York: Oxford University Press.
- Brown WM, George M Jr, Wilson AC. 1979. Rapid evolution of an animal mitochondrial DNA. *Proc Natl Acad Sci USA* 76:1967–1971.
- Brown WM, Prager EM, Wang A, Wilson AC. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol* 18:225–239.
- Buettner-Janusch J. 1966. A problem in evolutionary systematics: nomenclature and classification of baboons, genus *Papio*. *Folia Primatol (Basel)* 4:288–308.
- Carr SM, Ballinger SW, Deer JN, Blankenship LH, Bickham JW. 1986. Mitochondrial DNA analysis of hybridization between sympatric white-tailed and mule deer in West Texas. *Proc Natl Acad Sci USA* 83:9576–9580.
- Delson E. 1984. Cercopithecoid biochronology of the African Pliocene: correlation among eastern and southern hominid-bearing localities. *Cour Forsch Inst Senckenberg* 69:199–218.
- Delson E. 1993. *Theropithecus* fossils from Africa and India and the taxonomy of the genus. In: Jablonski NG, editor. *Theropithecus: the rise and fall of a primate genus*. Cambridge: Cambridge University Press. p 157–189.
- Disotell TR. 1996. The phylogeny of the Old World monkeys. *Evol Anthropol* 5:18–24.
- Ellerman JR, Morrison-Scott TCS, Hayman RW. 1953. Southern African mammals, 1758–1951: a reclassification. London: British Museum (Natural History).
- Evans BJ, Morales JC, Supriatna J, Melnick DJ. 1999. Origin of the Sulawesi macaques (Cercopithecidae: Macaca) as suggested by mitochondrial DNA phylogeny. *Biol J Linn Soc* 66:539–560.
- Gagneux P, Wills C, Gerloff U, et al. 1999. Mitochondrial sequences show diverse evolutionary histories of African hominoids. *Proc Natl Acad Sci USA* 96:5077–5082.
- Groves CP. 2001. Primate taxonomy. Washington, DC: Smithsonian Institution Press.
- Grubb P. 1999. Evolutionary processes implicit in distribution patterns of modern African mammals. In: Bromage TG, Schrenk F, editors. African biogeography, climate change, and human evolution. New York: Oxford University Press. p 150–164.
- Gyllenstein UB, Wilson AC. 1987. Interspecific mitochondrial DNA transfer and the colonization of Scandinavia by mice. *Genet Res Camb* 49:25–29.
- Hasegawa M, Kishini H, Yano T-A. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22:160–174.
- Hayasaka K, Fujii K, Horai S. 1996. Molecular phylogeny of macaques: implications of nucleotide sequences from an 896-base pair region of mitochondrial DNA. *Mol Biol Evol* 13:1044–1053.
- Hewitt GM. 1996. Some genetic consequences of Ice Ages, and their role in divergence and speciation. *Biol J Linn Soc* 58:247–276.
- Hill WCO. 1967. Taxonomy of the baboon. In: Vagtborg H, editor. The baboon in medical research, volume 2. Austin: University of Texas Press. p 4–11.
- Hill WCO. 1970. Primates: comparative anatomy and taxonomy VIII, Cynopithecinae (*Papio*, *Mandrillus*, *Theropithecus*). New York: Wiley-Interscience.
- Hillis DM, Moritz C, Mable BK. 1996. Molecular systematics, 2nd ed. Sunderland, MA: Sinauer Associates, Inc.
- Jablonski NG. 1993. *Theropithecus: the rise and fall of a primate genus*. Cambridge: Cambridge University Press.
- Jolly CJ. 1965. Origins and specialisations of the long-faced cercopithecoidea. PhD dissertation, London University.
- Jolly CJ. 1993. Species, subspecies and baboon systematics. In: Kimbel W, Martin L, editors. Species, species concepts and primate evolution. New York: Wiley. p 67–107.
- Jolly CJ. 2001. A proper study for mankind: analogies from the Papionin monkeys and their implications for human evolution. *Yrbk Phys Anthropol* 33:177–204.
- Kimura M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120.
- Kingdon J. 1971. East African mammals: an atlas of evolution in Africa. New York: Academic Press.
- Kingdon J. 1997. The Kingdon field guide to African mammals. San Diego: Academic Press.
- Kishino H, Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order of the Hominoidea. *J Mol Evol* 29:170–179.
- Kocher TD, Thomas WK, Meyer A, et al. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Natl Acad Sci USA* 86:6196–6200.
- Kummer H. 1968. Social organization of hamadryas baboons. Basel: Karger.
- Kummer H. 1995. In quest of the sacred baboon. Princeton, NJ: Princeton University Press.
- Lehman NA, Eisenhower K, Hansen LD, Mech RO, Peterson PJ, Gogan P, Wayne RK. 1991. Introgression of coyote mitochondrial DNA in sympatric North American grey wolf populations. *Evolution* 45:104–119.
- Maddison WP, Maddison DR. 1992. MacClade: an analysis of phylogeny and character evolution, version 3.03. Sunderland, MA: Sinauer Associates.
- Maples WR, McKern TW. 1967. A preliminary report on classification of the Kenya baboon. In: Vagtborg H, editor. The baboon in medical research, volume 2. Austin: University of Texas Press. p 13–22.
- Melnick DJ, Hoelzer GA, Absher R, Ashley MV. 1993. MtDNA diversity in rhesus monkeys reveals overestimates of divergence time and parapatry with neighboring species. *Mol Biol Evol* 10:282–295.
- Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS. 1994. Kin selection, social structure, gene flow and the evolution of chimpanzees. *Science* 265:1192–1201.
- Newman TK. 1997. Mitochondrial DNA analysis of intraspecific hybridization in *Papio hamadryas anubis*, *P. h. hamadryas* and their hybrids in Awash National Park, Ethiopia. Ph.D. dissertation, New York University.
- Page DMP, Holmes EC. 1998. Molecular evolution: a phylogenetic approach. Oxford: Blackwell Science, Ltd.
- Popp JL. 1983. Ecological determinism in the life histories of baboons. *Primates* 24:198–210.
- Richard A. 1985. Primates in nature. New York: W.H. Freeman and Co.
- Rogers J. 2000. Molecular genetic variation and population structure in *Papio* baboons. In: Whitehead P, Jolly CJ, editors. Old World monkeys. Cambridge: Cambridge University Press. p 57–76.
- Rogers J, Kidd KK. 1993. Nuclear DNA polymorphisms in a wild population of yellow baboons (*Papio hamadryas cynocephalus*) from Mikumi National Park, Tanzania. *Am J Phys Anthropol* 90:477–486.
- Roth WT. 1965. The taxonomy of the baboon and its position in the order of primates. In: Vagtborg H, editor. The baboon in medical research, volume 1. Austin: University of Texas Press. p 3–16.

- Samuels A, Altmann J. 1986. Immigration of a *Papio anubis* male into a group of *Papio cynocephalus* baboons and evidence for an anubis-cynocephalus hybrid zone in Amboseli, Kenya. *Int J Primatol* 7:131–138.
- Simpson GG. 1945. Principles of classification and a classification of mammals. *Bull Am Mus Nat Hist* 85:1–350.
- Smuts BB, Seyfarth RM, Wrangham RW, Struhsaker TT. 1986. *Primate societies*. Chicago: University Press.
- Swofford D. 1999. PAUP* 4.0b10: phylogenetic analysis using parsimony. Sunderland, MA: Sinauer Associates, Inc.
- Thorington RW, Groves CP. 1970. An annotated classification of the Cercopithecoidea. In: Napier JR, Napier PH, editors. *The Old World monkeys*. New York: Academic Press. p 629–647.
- Tosi AJ, Morales JC, Melnick DJ. 2000. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. *Mol Phylogenet Evol* 17:133–144.
- Williams-Blangero S, VandeBerg JL, Blangero J, Konigsberg L, Dyke B. 1990. Genetic differentiation between baboon subspecies: relevance for biomedical research. *Am J Primatol* 20:67–81.
- Woolley-Barker T. 1999. Social organization and genetic structure in a baboon hybrid zone. Ph.D. dissertation, New York University.