Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 8900–8904, September 1994
Evolution

Gene trees and hominoid phylogeny

(hominoid evolution/mitochondrial DNA diversity/cytochrome oxidase subunit II gene)

MARYELLEN RUVOLO*, DEBORAH PAN, SARAH ZEHR, TONY GOLDBERG, TODD R. DISOTELL†, AND MIRANDA VON DORNUM

Department of Anthropology, Harvard University, Cambridge, MA 02138

Communicated by Walter Gilbert, May 26, 1994

ABSTRACT **Here we present a DNA sequence study that incorporates intraspecific variation from all five genera of hominoids (apes and humans). Recently it has been claimed that using single individuals to analyze species' relationships might be misleading if within-species variation is great. Our results indicate that despite high intraspecific variation in mitochondrial cytochrome oxidase subunit II gene sequences of some hominoids, humans and chimpanzees are nonetheless significantly most closely related. We also report the observation that variation within the gorilla species exceeds that between common and pygmy chimpanzee species, a finding with implications for conservation. In contrast, humans are less mitochondrially diverse than lowland gorillas inhabiting western Africa.**

We have investigated intraspecific variability in hominoid mitochondrial DNA using cytochrome oxidase subunit II (COII) gene sequences for two reasons. (*i*) We wanted to compare humans with other hominoids directly at the DNA sequence level to examine further the reportedly low level of genetic variation within *Homo sapiens* (1, 2) and to estimate coalescence times for the human mitochondrial ancestor (3). (*ii*) We wanted to test the assertion (4, 5) that incorporating intraspecific variation could confound phylogenetic discrimination among hominoids. In a study of macaque mitochondrial DNA diversity measured indirectly by restriction site mapping, Melnick *et al.* (5) found intraspecific variation to be sufficiently great in the rhesus monkey that not all of its mitochondrial haplotypes are most closely related as a group distinct from those of other species. Their study indicates that eastern *Macaca mulatta* mitochondrial types are most closely related to *Macaca fuscata* and *Macaca cyclopis* types, whereas western *M. mulatta* types are more distantly related. Although this finding is only weakly supported by cladistic analysis and has not been confirmed with DNA sequence data, nonetheless the possibility of mitochondrial paraphyly needs to be considered because of its occurrence in other organisms [in *Peromyscus* mice (6) and *Anas* ducks (7)]. In such cases, inferences about phylogenetic relationships among species may sometimes be dependent upon the particular individuals chosen to represent them. Previously, using COII gene sequences of single species representatives, we found *Homo* and *Pan* to be the most closely related hominoid genera (8). Will this result still hold if intraspecific variation is taken into account?

Here we present nine mitochondrial COII gene sequences, including three gorilla, three common chimpanzee, and two orangutan subspecies' representatives.‡ These are analyzed along with previously published hominoid COII sequences (3, 8–10). The human sequences represent some of the most mitochondrially different individuals known (3, 11–13) as do the gorilla sequences. In a study of gorilla genetic variation across the species' range, four distinct clades have been identified using mitochondrial hypervariable control region sequences (ref. 14; K. Garner and O. Ryder, personal communication); one representative of each clade was sequenced here for COII: two western lowland gorillas (*Gorilla gorilla gorilla*), one eastern lowland gorilla (*Gorilla gorilla graueri*), and one mountain gorilla (*Gorilla gorilla beringei*). Three common chimpanzee subspecies are also included: central African *Pan troglodytes troglodytes*, eastern African *Pan troglodytes schweinfurthii*, and western African *Pan troglodytes verus*, as are orangutan subspecies from Borneo (*Pongo pygmaeus pygmaeus*) and Sumatra (*Pongo pygmaeus abelii*).

MATERIALS AND METHODS

DNA Samples. These newly reported COII sequences are from *P. troglodytes* (Ptr 4, "Hallie," Yerkes; Ptr 5, "Kirk," Fort Worth Zoo); *Pan paniscus* [Ppa 4, "Kakowet," ISIS no. 160134, SDZ (San Diego Zoo)]; western lowland *G. g. gorilla* (Ggo 3, "Massa," SDZ sample OR287; Ggo 4, "Albert," SDZ sample OR291); eastern lowland *G. g. graueri* (Ggo 5, "Mukisi," Antwerp Zoo, SDZ sample KG083); mountain gorilla *G. g. beringei* (Ggo 6, SDZ sample KG073); Bornean *P. p. pygmaeus* (Ppy2, "Dinah," SDZ sample 5404); Sumatran *P. p. abelii* (Ppy3, "Doris," SDZ sample 4361). From import records, chimpanzee subspecies are *P. t. schweinfurthii* (Ptr 1), *P. t. verus* (Ptr 4), and *P. t. troglodytes* (Ptr 5). Previously reported COII sequences are of orangutan [Ppy1 (10)], siamangs [*Hylobates syndactylus*; Hsy 1 (6); Hsy 2 (8)], humans, chimpanzees, and gorillas (refs. 3 and 8–10; labeled as in ref. 3).

Amplification and Sequencing. DNA sequences were obtained by polymerase chain reaction amplification and direct DNA sequencing using oligonucleotide primers specific for the mitochondrial COII gene, as in Ruvolo *et al.* (8).

RESULTS AND DISCUSSION

The close relationship of humans and chimpanzees, separate from gorillas, is supported by phylogenetic analysis using parsimony (15) (Fig. 1) and distance (18) methods. Notably, inclusion of intraspecific variability provides a high level of support for a human–chimpanzee clade (95% bootstrap value) equivalent to that found in the analysis of single species' representatives (8) (92% bootstrap value). This high level of support exceeds that for the undisputed node uniting humans, chimpanzees, and gorillas (82% bootstrap value). Since both nodes are significantly supported, with probabil-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: COII, cytochrome oxidase subunit II.

*To whom reprint requests should be addressed at: Laboratory of Molecular Evolution, Department of Anthropology, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138.
†Present address: Department of Anthropology, New York University, New York, NY 10003.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U12698–U12706).

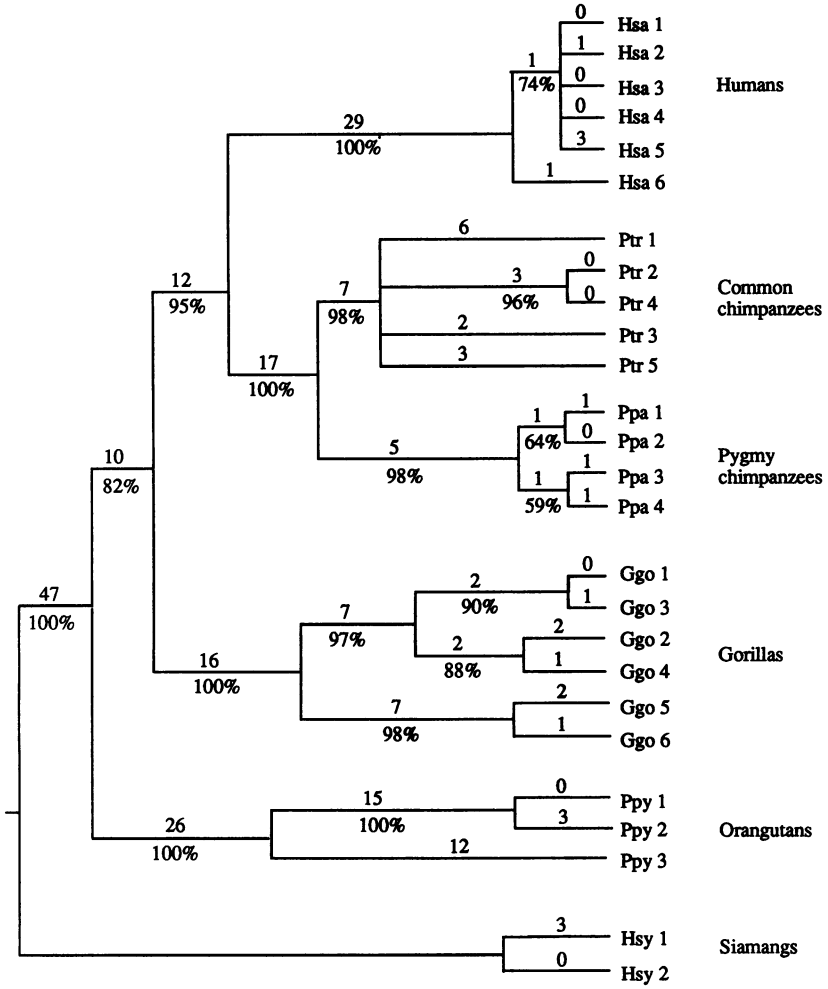

FIG. 1. Maximum parsimony consensus tree of hominoid mitochondrial COII gene sequences. Two most parsimonious trees were generated using the phylogenetic analysis program PAUP (16) (under the branch and bound search option). One tree was identical to the above consensus tree, while the second tree differed slightly in having chimpanzee Ptr 5 as sister to all other *P. troglodytes*. Both trees are of minimum length 325 (consistency index 0.71) and show 12 synapomorphies linking *Homo* and *Pan*. Alternative phylogenies are less parsimonious: shortest trees are 9 steps longer for a *Pan–Gorilla* clade (supported by 3 or 4 synapomorphies); 10 steps longer for a *Homo–Gorilla* clade (2 synapomorphies), using MACCLADE (17) and PAUP (16). Bootstrap percentage values (generated in PAUP, 1000 replications) indicate how often clades appear in repeated data subsamplings: 95% for *Homo–Pan*; 1.2% for *Pan–Gorilla*; 0.3% for *Homo–Gorilla*. Branch lengths indicated are minimum numbers of inferred changes.

ities that the clades are real greater than 0.95 (19), this indicates that the tree is well supported overall. Twelve inferred substitutions link humans and chimpanzees unambiguously in the parsimony analysis (Table 1).

From these COII sequences, the branch separating the *Homo–Pan* ancestor from the common *Gorilla–Homo–Pan* ancestor is relatively long, 42–43% of the averaged *Homo* and *Pan* branch lengths by maximum likelihood corrected distances (20) and by inferred substitutions on the maximum parsimony tree (Fig. 1). This relatively long internode agrees with conclusions based on single species' representatives (8) and with results from a larger mitochondrial dataset including the COII gene of single individuals (10), from DNA hybridization of more than one individual per species (21), and from proteins (22, 23). These mitochondrial and nuclear DNA data are consistent with two successive divergence events leading first to gorilla and then to human and chimpanzee lineages separated by several million years.

Other DNA sequence studies either support the existence of a *Homo–Pan* clade with less separation or cannot resolve relationships among the three species (for reviews see refs. 24 and 25); the latter may be due to slower rates of molecular change. DNA sequence data from one nuclear genomic region support the alternative *Pan–Gorilla* clade (26); however, this conclusion is based on the tandem repeat segment of the involucrin gene, which has undergone extensive deletions/insertions throughout primate evolution, making DNA sequence alignments difficult and reconstruction of molecular evolutionary events ambiguous (27). Furthermore, gorillas are polymorphic for the presence/absence of some involucrin repeats (28) reported to be *Pan–Gorilla* synapomorphies by Djian and Green (26), and these alleles were not included in the phylogenetic analysis. Thus, with this one problematic exception, DNA sequence studies that are able to resolve relationships among the hominoids support a human–chimpanzee clade; COII is among those showing the greatest relative separation from the gorilla lineage.

On the COII gene tree (Fig. 1), hominoid haplotypes cluster by species, unlike the case for macaque mitochondrial haplotypes defined by restriction mapping (5). This difference in mitochondrial structuring between hominoids and macaques may be due to social system. Female macaques are philopatric, and large interpopulational mitochondrial differences exist (29), whereas female hominoids transfer between groups (30, 31), ensuring equivalent degrees of mitochondrial and nuclear gene flow among populations.

Table 1. Hominoid mitochondrial COII gene variation at sites unambiguously linking humans and chimpanzees in a phylogenetic analysis

	60	156	177	204	267	268	342	414	525	549	606	637
Hsa1	T	T	G	G	G	G	C	T	T	T	A	T
Hsa2	.	.	.	.	.	.	.	.	.	.	.	.
Hsa3	.	.	.	.	.	.	.	.	.	.	.	.
Hsa4	.	.	.	.	.	.	.	.	.	.	.	.
Hsa5	.	.	.	.	.	.	.	.	.	.	.	.
Hsa6	.	.	.	.	.	.	.	.	.	.	.	.
Ptr1	.	.	.	A	.	.	.	.	.	.	.	.
Ptr2	.	.	.	.	.	.	.	.	.	.	.	.
Ptr3	.	.	.	.	.	.	.	.	.	.	.	.
Ptr4	.	.	.	.	.	.	.	.	.	.	.	.
Ptr5	.	.	.	.	.	.	.	.	.	.	.	.
Ppa1	.	.	.	.	.	.	.	.	.	.	.	.
Ppa2	.	.	.	.	.	.	.	.	.	.	.	.
Ppa3	.	.	.	.	.	.	.	.	.	.	.	.
Ppa4	.	.	.	.	.	.	.	.	.	.	.	.
Ggo1	A	C	A	A	A	A	T	C	C	A	C	C
Ggo2	A	C	A	.	A	A	T	C	C	A	C	C
Ggo3	A	C	A	A	A	A	T	C	C	A	C	C
Ggo4	A	C	A	.	A	A	T	C	C	A	C	C
Ggo5	A	C	A	A	A	A	T	C	C	A	C	C
Ggo6	A	C	A	A	A	A	T	C	C	A	C	C
Ppy1	G	C	A	A	A	A	.	C	C	A	T	C
Ppy2	G	C	A	A	A	A	.	C	C	A	T	C
Ppy3	G	C	A	A	.	A	T	C	C	A	C	.
Hsy1	A	C	A	A	A	A	T	C	C	A	C	C
Hsy2	A	C	A	A	A	A	T	C	C	A	C	C

On the maximum parsimony tree (Fig. 1), a minimum of 12 inferred nucleotide changes occur along the common ancestral *Homo–Pan* branch at these sites; additional changes (29 maximum) at other COII sites could also have occurred along that branch. Site numbers refer to positions within the COII gene; abbreviations as in Fig. 1.

Mitochondrial sequence diversity suggests a rank ordering of genetic variability levels (Table 2). Differences between species are not expected to change substantially as more individuals representing wider geographic ranges are analyzed; for the two most widely sampled species (humans, gorillas), the difference between any particular human and any particular gorilla COII sequence is a good estimate of the interspecific difference averaged over all individuals. With additional sampling, however, intraspecific estimates for common chimpanzees, pygmy chimpanzees, orangutans, and particularly siamangs may increase.

Because common and pygmy chimpanzees are recognizably distinct species with clear morphological, ecological, and behavioral differences (32), their genetic difference provides a comparative gauge for variation within hominoids. Gorillas were previously found to be the least variable great ape in a mitochondrial restriction study of a limited geographical sample (2). However, our COII sequences from a wider sample show that gorillas are mitochondrially more diverse than the two recognized species of chimpanzees, with the greatest difference between western gorillas (*G. g. gorilla*) on the one hand and eastern gorillas (*G. g. graueri* and *G. g.*

Table 2. Mitochondrial COII gene sequence differences within and between hominoid taxa

Comparison	Average difference*	Maximum difference
Within species		
Within pygmy chimpanzees (*P. paniscus*)	0.5	0.6
Within humans (*H. sapiens*)	0.6	0.9
Within siamangs (*H. syndactylus*)	0.9	0.9
Within common chimpanzees (*P. troglodytes*)	1.0	1.3
Within western lowland gorillas (*G. g. gorilla*)	1.0	1.2
Within eastern gorillas (*G. g. beringei* vs. *G. g. graueri*)	0.6	0.6
Within gorillas (western vs. two eastern subspecies)	3.1	3.5
Within orangutans (*P. p. pygmaeus* vs. *P. p. abelii*)	5.0	5.2
Between species, genera		
Common vs. pygmy chimpanzees	2.7	3.2
Human vs. chimpanzees	9.3	10.4
Gorilla vs. (human, chimpanzees)	11.4	12.9
Orangutan vs. (gorilla, human, chimpanzees)	13.9	15.2

All values are percentages of observed numbers of substitutions, uncorrected for multiple substitutions, of 684 base pairs total.
*Average of pairwise differences calculated through common ancestor.

beringei) on the other. Interestingly, variation in some morphological characters mirrors the mitochondrial DNA results in pattern and degree: cranio-dentally, western lowland gorillas are distinct from eastern lowland and mountain gorillas; common and pygmy chimpanzee species differ to an equivalent or lesser extent (33–35).

From the COII DNA sequence data, orangutan subspecies are the most intraspecifically divergent (Table 2), confirming previous observations based on isozymes (23), two-dimensional protein electrophoresis (23), DNA hybridization (21), and mitochondrial restriction mapping (2). The subspecies differ karyotypically (36) and morphologically in skeletal, cranio-dental, and external features, but not consistently more so than chimpanzee species (34, 35, 37).

Nuclear HOX2 sequences also show relatively high gorilla intraspecific diversity (nine base pairs maximum) but nonetheless link humans and chimpanzees by one substitution (4). Rather than interpreting this as evidence for a human–chimpanzee clade, Ruano *et al.* (4) propose that the polymorphic common African hominoid ancestor speciated trichotomously, and chance fixation of alleles produced the observed HOX2 pattern. However, were this true, it is unlikely that COII sequences would sort on the gene tree according to species. In COII, as with HOX2, the number of nucleotide differences among gorillas is greater than the number of substitutions linking humans and chimpanzees. Yet comparing amounts of within-group to between-group differences does not take the cladistically relevant patterning of such variation into account. Although gorillas vary at 28 COII sites, they are monomorphic at 11 of the 12 unambiguous phylogenetically informative sites linking humans and chimpanzees (Table 1). Intraspecific variation in hominoid COII sequences therefore does not confound phylogenetic reconstruction. Although high levels of intraspecific diversity may exist at some sites within a gene, phylogenetic information can nevertheless be present at other sites, even if those latter sites are fewer in number. This illustrates how a population genetic approach, which focuses on amounts of within-group and between-group variation, views genetic variation in a fundamentally different way from a molecular systematic (in this instance cladistic) one, which links together DNA sequences, regardless of origin, into hierarchically nested groups reflective of shared derived substitutions at particular sites.

CONCLUSIONS

In this phylogenetic study, which includes intraspecific diversity from all hominoid genera measured by DNA sequence comparisons, *Homo* and *Pan* form a clade significantly separated from the *Gorilla* lineage. Inclusion of multiple individuals per species in the phylogenetic analysis does not produce a result different from that based on single individuals representing species (8) in terms of either tree topology or relative branch lengths. The relative degree of separation of the *Homo–Pan* clade measured with the COII gene is among the highest observed for any molecular dataset, similar to that found using sequences from a larger mitochondrial segment (10), DNA hybridization (21), and protein sequences (22, 23).

Mitochondrial diversity within some hominoid species is great, yet this does not obscure phylogenetic relationships. For gorillas and orangutans, within-species variability exceeds that between the different species of chimpanzees. This observation has implications for gorilla and orangutan conservation efforts and zoo management. Members of eastern and western gorilla clades and members of orangutan subspecies should ideally be conserved separately, and interbreeding should be avoided where possible. Establishment of additional gorilla and orangutan species designations (33, 38) may even be warranted.

The limited mitochondrial diversity within humans (1, 2, 11–13) can now be viewed comparatively at the DNA sequence level with that of other hominoid species representing all genera. On average, the most mitochondrially different humans known are less different even than the only two siamangs sequenced to date or than lowland gorillas living in a restricted geographic area of west Africa. This limited genetic diversity becomes equivalent via application of the molecular clock to a recent time for the mitochondrial ancestor of living humans, roughly 200,000–300,000 years ago (3, 11–13). Placing human variability in a comparative primate context helps us recognize that anciently separated mitochondrial lineages, common in other hominoid species, are missing in *H. sapiens*.

We thank Drs. K. Gardner and O. Ryder of the San Diego Zoo; the Virunga Veterinary Clinic, Ruhengeri, Rwanda; the Morris Animal Foundation; Yerkes Primate Center; L. Schobert; D. Pernikoff of the Fort Worth Zoo for DNA, tissue samples, and animal information; and Drs. D. Pilbeam, A. Uchida, A. Yoder, and C.-B. Stewart and an anonymous reviewer for helpful comments and criticism. This work was supported by grants from the Wenner-Gren Foundation for Anthropological Research, the Harvard Human Origins Research Fund, and Elaine and Leon Cohn.

1. Brown, W. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3605–3609.
2. Ferris, S. D., Wilson, A. C. & Brown, W. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2432–2436.
3. Ruvolo, M., Zehr, S., von Dornum, M., Pan, D., Chang, B. & Lin, J. (1993) *Mol. Biol. Evol.* **10**, 1115–1135.
4. Ruano, G., Rogers, J., Ferguson-Smith, A. C. & Kidd, K. K. (1992) *Mol. Biol. Evol.* **9**, 575–586.
5. Melnick, D. J., Hoelzer, G. A., Absher, R. & Ashley, M. V. (1993) *Mol. Biol. Evol.* **10**, 282–295.
6. Avise, J. C., Shapira, J. F., Daniel, S. W., Aquadro, C. F. & Lansman, R. A. (1983) *Mol. Biol. Evol.* **1**, 38–56.
7. Avise, J. C., Ankney, C. D. & Nelson, W. S. (1990) *Evolution* **44**, 1109–1119.
8. Ruvolo, M., Disotell, T. R., Allard, M. W., Brown, W. M. & Honeycutt, R. L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1570–1574.
9. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* **290**, 457–465.
10. Horai, S., Satta, Y., Hayasaka, K., Kondo, R., Inoue, T., Ishida, T., Hayashi, S. & Takahata, N. (1992) *J. Mol. Evol.* **35**, 32–43.
11. Cann, R. L., Stoneking, M. & Wilson, A. C. (1987) *Nature (London)* **325**, 31–36.
12. Vigilant, L., Pennington, R., Harpending, H., Kocher, T. D. & Wilson, A. C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9350–9354.
13. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. (1991) *Science* **253**, 1503–1507.
14. Garner, K. J. & Ryder, O. A. (1992) *Symp. Zool. Soc. London* **64**, 167–181.
15. Fitch, W. (1971) *Syst. Zool.* **19**, 406–416.
16. Swofford, D. L. (1993) PAUP, Phylogenetic Analysis Using Parsimony, Program and Documentation (Center for Biodiversity, Champaign, IL), Version 3.1.1.
17. Maddison, W. P. & Maddison, D. R. (1992) MACCLADE, Analysis of Phylogeny and Character Evolution Program and Documentation (Sinauer, Sunderland, MA), Version 3.0.
18. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
19. Hillis, D. M. & Bull, J. J. (1993) *Syst. Biol.* **42**, 182–192.
20. Felsenstein, J. (1993) PHYLIP, Phylogeny Inference Package, Program and Documentation (University of Washington, Seattle), Version 3.5.
21. Caccone, A. & Powell, J. R. (1989) *Evolution* **43**, 925–942.
22. Goodman, M., Koop, B. F., Tagle, D. A. & Slightom, J. L. (1989) *Genome* **31**, 316–335.

23. Janczewski, D. N., Goldman, D. & O'Brien, S. J. (1990) *J. Hered.* **81,** 375–387.
24. Holmquist, R., Miyamoto, M. & Goodman, M. (1988) *Mol. Biol. Evol.* **5,** 217–236.
25. Bailey, W. J. (1994) *Evol. Anthropol.* **2,** 100–108.
26. Djian, P. & Green, H. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 8447–8451.
27. Ruvolo, M. (1994) *Am. J. Phys. Anthropol.* **94,** 89–113.
28. Teumer, J. & Green, H. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1283–1286.
29. Melnick, D. J. & Hoelzer, G. A. (1992) *Int. J. Primatol.* **13,** 379–393.
30. Pusey, A. E. & Packer, C. (1986) in *Primate Societies*, eds. Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W. & Struhsaker, T. T. (Univ. of Chicago, Chicago), pp. 250–266.
31. Wrangham, R. W. (1980) *Behaviour* **75,** 262–300.
32. Susman, R. L., ed. (1984) *The Pygmy Chimpanzee: Evolutionary Biology and Behavior* (Plenum, New York).
33. Groves, C. P. (1986) in *Comparative Primate Biology: Systematics, Evolution, and Anatomy*, eds. Swindler, D. R. & Erwin, J. (Liss, New York), Vol. 1, pp. 187–217.
34. Shea, B. T. & Coolidge, H. J. (1988) *J. Hum. Evol.* **17,** 671–685.
35. Uchida, A. (1993) *Am. J. Phys. Anthropol.*, Suppl. 16, 198–199 (abstr.).
36. de Boer, L. E. M. & Seuanez, H. (1982) in *The Orang Utan: Its Biology and Conservation*, ed. de Boer, L. E. M. (Junk, The Hague, The Netherlands), pp. 135–170.
37. Groves, C. P., Westwood, C. B. & Shea, B. T. (1992) *J. Hum. Evol.* **22,** 327–340.
38. Andrews, P. (1984) *Primate Eye* **23,** 34–37.