

Nucleotide Diversity in Gorillas

Ning Yu,^{*,1} Michael I. Jensen-Seaman,^{†,1} Leona Chemnick,[‡] Oliver Ryder[‡] and Wen-Hsiung Li^{*,2}

^{*}Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, [†]Human and Molecular Genetics Center, Medical College of Wisconsin, Milwaukee, Wisconsin 53226 and [‡]Center for Reproduction of Endangered Species, Zoological Society of San Diego, San Diego, California 92101

Manuscript received June 26, 2003

Accepted for publication November 19, 2003

ABSTRACT

Comparison of the levels of nucleotide diversity in humans and apes may provide valuable information for inferring the demographic history of these species, the effect of social structure on genetic diversity, patterns of past migration, and signatures of past selection events. Previous DNA sequence data from both the mitochondrial and the nuclear genomes suggested a much higher level of nucleotide diversity in the African apes than in humans. Noting that the nuclear DNA data from the apes were very limited, we previously conducted a DNA polymorphism study in humans and another in chimpanzees and bonobos, using 50 DNA segments randomly chosen from the noncoding, nonrepetitive parts of the human genome. The data revealed that the nucleotide diversity (π) in bonobos (0.077%) is actually lower than that in humans (0.087%) and that π in chimpanzees (0.134%) is only 50% higher than that in humans. In the present study we sequenced the same 50 segments in 15 western lowland gorillas and estimated π to be 0.158%. This is the highest value among the African apes but is only about two times higher than that in humans. Interestingly, available mtDNA sequence data also suggest a twofold higher nucleotide diversity in gorillas than in humans, but suggest a threefold higher nucleotide diversity in chimpanzees than in humans. The higher mtDNA diversity in chimpanzees might be due to the unique pattern in the evolution of chimpanzee mtDNA. From the nuclear DNA π values, we estimated that the long-term effective population sizes of humans, bonobos, chimpanzees, and gorillas are, respectively, 10,400, 12,300, 21,300, and 25,200.

THE amount and pattern of genetic diversity in a species can provide valuable information for deducing the evolutionary history of the species, including past changes in population size, effects of social structure on genetic diversity, patterns of past migration, and signatures of past selection events. For these reasons, numerous studies of genetic diversity have been conducted on humans (*e.g.*, CANN *et al.* 1987; TISHKOFF *et al.* 1996; HARPENDING *et al.* 1998). Recently, there has also been considerable interest in the level and pattern of nucleotide diversity in the African apes (see below). These studies have revealed that although humans currently are geographically widespread and number in the billions, they show reduced genetic variation compared to the geographically more restricted African apes (RUVOLO *et al.* 1994; DEINARD and KIDD 1998, 1999, 2000; GAGNEUX *et al.* 1999; KAESSMANN *et al.* 1999, 2001; JENSEN-SEAMAN *et al.* 2001) and have a long-term effective population size of only $\sim 10,000$. It seems that humans are unusual compared to African apes in this respect, which suggests that the last common ancestor of Homo, Pan, and Gorilla was probably much more similar to the extant apes than to modern humans.

FERRIS *et al.*'s (1981) study of ape mitochondrial DNA (mtDNA) by restriction enzyme mapping was the first to suggest a much higher level of mtDNA variation in great apes than in humans. In that study, the mtDNA genome of chimpanzees was found to be three times more variable than that of humans; even gorillas, which had the least amount of mtDNA variation among the apes, exhibited twice as much variation as humans. This view was supported by sequence data from COII and 16S rRNA (RUVOLO *et al.* 1994; NODA *et al.* 2001). Further, in the first hypervariable segment (~ 300 bp) of the D-loop, gorillas also carried twice as much nucleotide diversity as humans, and chimpanzees had three times that of humans (MORIN *et al.* 1994; GARNER and RYDER 1996; DEINARD and KIDD 2000; JENSEN-SEAMAN and KIDD 2001). In a recent study of five autosomal loci among African apes, common chimpanzees carried the greatest amount of nucleotide diversity, with bonobos and gorillas possessing somewhat less variation (JENSEN-SEAMAN *et al.* 2001). Data from a 10-kb X-linked noncoding region also revealed a three- to fourfold higher nucleotide diversity in both gorillas and chimpanzees than in humans (KAESSMANN *et al.* 1999). Although data from gorillas are lacking, substantially greater diversity was found in the Y chromosome of chimpanzees and bonobos than in that of humans (STONE *et al.* 2002). Therefore, DNA sequence data from both the mitochondrial and nuclear genomes strongly suggested that this greater

¹These authors contributed equally to this work.

²Corresponding author: Department of Ecology and Evolution, University of Chicago, 1101 E. 57th St., Chicago, IL 60637.
E-mail: whli@uchicago.edu

amount of diversity is a general feature of the African apes (JENSEN-SEAMAN *et al.* 2001).

Noting that the nuclear DNA polymorphism data in apes were from only a few loci, we decided to do a further investigation (YU *et al.* 2003). We sequenced 50 DNA segments in nine bonobos and 17 chimpanzees from East, Central, and West Africa. These 50 segments were the same as in YU *et al.* (2002), who studied 30 humans from various localities around the world; the 50 segments were randomly chosen from the noncoding, nonrepetitive parts of the human genome (CHEN and LI 2001). Unexpectedly, the new data revealed a considerably smaller difference between the levels of nucleotide diversity in chimpanzees and humans (YU *et al.* 2003), with the former possessing only ~ 1.5 -fold greater diversity than the latter. Thus, with a clearer view obtained with a much larger number of loci, we found that autosomal DNA and mtDNA actually gave different pictures of the levels of nucleotide diversity in humans and chimpanzees. This result raised the question, Is this also true for other apes? To gain a better understanding of the amount of intraspecific genetic variation in the gorilla genome, it is necessary to obtain data from a similarly large number of loci.

Gorillas are found discontinuously in the tropical forests of equatorial Africa. The taxonomy of gorillas has long been debated (GROVES 2003). They have traditionally been considered a single species (COOLIDGE 1929); however, more researchers are considering them as two separate species, *Gorilla gorilla* for western gorillas and *G. beringei* for eastern gorillas (GROVES 2001). Wild gorilla populations are diminishing due to destruction of the tropical forests, habitat loss, and poaching. Mountain and eastern lowland gorillas are classified by IUCN—The World Conservation Union as endangered, while western lowland gorillas are considered threatened (LEE *et al.* 1988). In the present study we have obtained DNA samples from 15 captive western lowland gorillas (*G. gorilla gorilla*) to estimate nucleotide diversity at the same 50 loci as used previously for humans, chimpanzees, and bonobos (YU *et al.* 2002, 2003).

MATERIALS AND METHODS

Sample sources: DNA from 15 western lowland gorillas was used in this study. Seven individuals (named Massa, Samson, Dolly, Tuffi, Porta, Freddy, and OR 802) were from the San Diego Zoo. Blood from three individuals (Holoko, Choomba, and Mumbah) was a generous gift from George Amato of the Wildlife Conservation Society. DNA from two individuals (Abe and Oko) was a generous gift from Amos Deinard and Kenneth Kidd of Yale University. Blood obtained during routine veterinary examinations from one individual (Moka) was kindly donated by the National Zoo in Washington, DC, and that from two more individuals (Josephine and Jimmie) was kindly donated by the Miami Metro Zoo in Miami. Except for Or802 (no name), who has unrelated parents and whose grandparents were all wild born, all individuals were originally wild born and all samples were independent.

PCR amplification and sequencing of DNA segments: The 50 noncoding, nonrepetitive genomic segments (each ~ 1 kb) were originally selected randomly from the human genome (CHEN and LI 2001; YU *et al.* 2002). All were chosen to avoid coding regions or close linkage to any coding regions. In each segment and its nearby regions there was no registered gene in GenBank and no potential coding region was detected by either GenScan or GRAIL-EXP.

Touch-down PCR (DON *et al.* 1991) was used and the reactions were carried out following the condition described in ZHAO *et al.* (2000). The PCR products were purified by Wizard PCR Preps DNA purification resin kit (Promega, Madison, WI). Sequencing reactions were performed according to the protocol of the ABI Prism BigDye terminator sequencing kits (Perkin-Elmer, Norwalk, CT) modified by one-quarter reaction. The extension products were purified with Sephadex G-50 (DNA grade, Pharmacia), and run on an ABI 377XL DNA sequencer using 4.25% gels (Sooner Scientific). About 500 bp of each segment was sequenced in both directions.

ABI DNA Sequence Analysis 3.0 was used for lane tracking and base calling. The data were then proofread manually and heterozygous sites were detected as double peaks. The forward and reverse sequences were assembled automatically in each individual using SeqMan (DNASTar, Madison, WI). The assembled files were carefully checked by eye. Fluorescent traces for each variant site were rechecked again in all individuals. All singletons, which are variants that appear only once in the entire sample, were verified by PCR reamplification and resequencing of the PCR products in both directions. No attempt was made to determine gametic phase (haplotypes) of individuals with multiple polymorphic sites per locus. Rather, the segments within an individual were concatenated in a random manner into two continuous sequences using DAMBE (XIA and XIE 2001). The new sequences have been deposited in GenBank (accession nos. AY447025–AY447950).

Data analysis: The sequences were aligned by SeqMan. Nucleotide diversity values and the average percentage distances between species were calculated using DNASP version 3.14 (ROZAS and ROZAS 1999).

RESULTS AND DISCUSSION

Distribution of single nucleotide polymorphisms: Because one of the 50 segments could not be amplified in four individuals, this segment was not included in this study. We sequenced the remaining 49 segments in 15 western lowland gorillas. The total number of nucleotide sites sequenced, after exclusion of deletions and insertions, is $\sim 23,056$ bp. A total of 138 single nucleotide polymorphisms (SNPs) were found in the 15 gorilla samples (30 sequences); 29 of them (21%) were observed only once (*i.e.*, singletons) and 21 (15%) only twice (doubletons). Interestingly, in gorillas more than half (64%) of the variants were intermediate or high-frequency variants. This excess of intermediate frequency variants is also seen in the values of Tajima's *D* statistic (TAJIMA 1989), where for the concatenated sequences only gorillas have a positive value of *D*, while humans, chimpanzees, and bonobos have negative values, implying that ancient gorilla populations may have been subdivided.

Adequacy of the samples: Since our sample size is relatively small, we need to consider the problem of

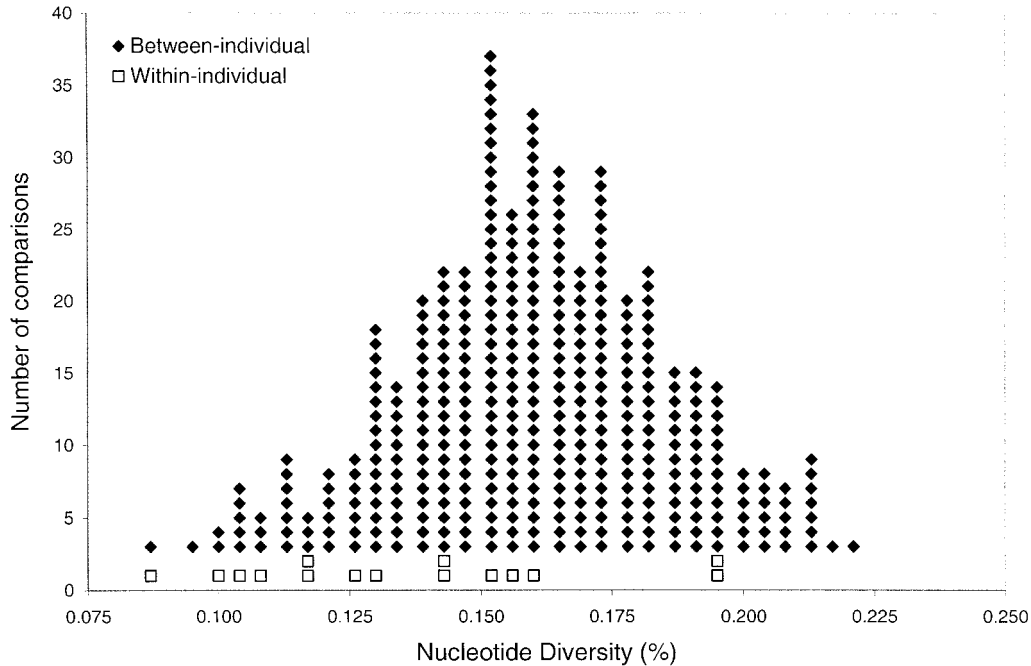


FIGURE 1.—Distributions of the within-individual and between-individual nucleotide diversity in gorillas.

sampling bias. For this purpose, we consider the effect of sampling on nucleotide diversity (π) because π is the quantity of our primary interest in this study; π is defined as the number of nucleotide differences per site between two randomly chosen sequences in a population. As noted in Yu *et al.* (2002), an estimation bias may be detected by comparing within-individual π values (π_w) with between-individual π values (π_b). Ideally, each sequence in a sample should be taken randomly from the population, but we have included the two sequences within each of the individuals sampled. It is possible that the two sequences in an individual are not completely independent if the individual is “inbred” to some extent, in the sense that both sequences within an individual likely came from the same subpopulation, rather than from true random mating throughout the larger population. Therefore, the within-individual π values (π_w) should tend to be smaller than the between-individual π values (π_b) and their inclusion should tend to give an underestimate of π . However, if the average π_b and π_w values are similar, then the sampling scheme would seem largely adequate and the inclusion of π_w values in the estimation of π should produce no substantial bias.

Figure 1 shows that the distribution of π_b values is like a normal distribution, except that one point ($\pi_b = 0.078\%$) is substantially lower than the others. This observation suggests that there was no strong sampling bias. Moreover, excluding the “exceptional” point affects little the average π value. The distribution of the 15 π_w values, which range from 0.078 to 0.195%, is somewhat narrower than that of the π_b values, which range from 0.078 to 0.221%, and the average π_w (0.136%) is lower than the average π_b (0.159%; $P < 0.01$, one-tailed *t*-test).

This comparison suggests that some of the individuals may have been inbred to some extent. However, excluding the 15 π_w values from comparison increases the average value only from 0.158 to 0.159%. We therefore take 0.158% as our estimate of the nucleotide diversity in western lowland gorillas.

The present study included individuals from only one of the gorilla subspecies, the western lowland gorilla; furthermore, since little is known of the geographic origin of these individuals they may not represent the full range of variation in this subspecies. As several studies have shown the amount of genetic distance between eastern and western gorillas to be as much as or greater than that between chimpanzees and bonobos from mtDNA loci such as COII, D-loop, and NADH5 (RUVOLO *et al.* 1994; GARNER and RYDER 1996; JENSEN-SEAMAN *et al.* 2001), we may be missing substantial variation in the genus *Gorilla*. On the other hand, data from eight independent nuclear loci suggested that the difference between the nuclear genomes of eastern and western gorillas was actually rather small compared to that between chimpanzees and bonobos (JENSEN-SEAMAN 2001, 2003) and that the inclusion of eastern gorilla samples at a few nuclear loci made almost no difference in the estimate of π for gorillas as a whole (JENSEN-SEAMAN 2000). A study of a 10-kb noncoding region in Xq13.3 confirmed the much smaller divergence between the nuclear genomes of eastern and western gorillas when compared to the Pan species (KAESSMANN *et al.* 2001). Nonetheless, without more data it is impossible to speculate on the potential effect of the inclusion of eastern gorilla individuals.

Nucleotide diversity: For the 49 DNA segments we studied, the range of π is from 0 (6 segments) to 0.49%

(Table 1). Such large fluctuations were also observed in humans, chimpanzees, and bonobos (Yu *et al.* 2002, 2003). These observations are not surprising because the nucleotide diversity in a short DNA region is subject to strong stochastic effects. In addition, variation in π may also arise from variation in mutation rate among genomic regions. Low π could also result from a recent selective sweep, but since these 49 segments were drawn from 16 different chromosomes, with most chosen to be millions of nucleotides from the next nearest segment, selection is not likely having any strong impact on the diversity values. Gorillas have the highest average π value (0.158%), which is close to twice that of humans (0.087%, Table 2). In contrast, the π value of bonobos (0.077%) is somewhat lower than that of humans, and that of chimpanzees (0.134%) is only 50% higher than that of humans.

Some reports have suggested that at autosomal loci gorillas have up to three times greater sequence diversity than humans (DEINARD and KIDD 1999; JENSEN-SEAMAN *et al.* 2001). Sequences from a 10-kb X-linked noncoding region revealed nucleotide diversity five times higher in gorillas than in humans (KAESSMANN *et al.* 2001). However, several mtDNA studies, each using different gorilla samples, showed a pattern similar to that observed in this study. Results using mtDNA sequence data from the COII gene, 16S rRNA, and the first hypervariable segment of the D-loop all revealed approximately twice as much diversity in western lowland gorillas as in humans (RUVOLO *et al.* 1994; JENSEN-SEAMAN and KIDD 2001; NODA *et al.* 2001).

Among previous nuclear loci studied, the highest nucleotide diversity was found in chimpanzees at ADH1, APOB, DRD2, and DRD4, while gorillas carried the highest variation at HOXB6 and Xq13.3. The nucleotide diversity in gorillas from the 49 segments in the present study is the highest, followed by chimpanzees. However, at 20 of the 49 segments (41%), chimpanzees had greater nucleotide diversity than gorillas, demonstrating the importance of examining a large number of loci to obtain a reliable conclusion. Bonobos carry the lowest nucleotide diversity, lower than that in humans. Therefore, having a much greater amount of nucleotide diversity than humans is not a general feature of the African apes.

Effective population sizes: To estimate effective population size (N_e), we calculated the average mutation rate, which is 1.0469×10^{-9} /site/year and determined the mutation rate per nucleotide site per generation (u) by using the sequence divergence (d) between species (Table 3) and assuming that the divergence time between human and gorilla and between human and the chimpanzee-bonobo lineage is 8 and 6 million years, respectively (BRUNET *et al.* 2002; VIGNAUD *et al.* 2002). Of course, using different divergence dates will yield slightly different estimates of u and N_e . Since we are interested in the long-term effective population size, we

use TAJIMA's (1983) estimator $\pi = 4N_e u$ and we assume that the generation time is 15 years for gorillas. The N_e for humans is estimated to be 10,400 (Table 2), which is similar to the commonly used value (10,000) in the literature (NEI and GRAUR 1984; TAKAHATA *et al.* 1995; ZHAO *et al.* 2000), while that for bonobos (12,300) is only slightly larger, that for chimpanzees (21,300) is about twice as large, and the N_e for western lowland gorillas (25,200) is ~ 2.5 times larger.

Differences in N_e between species could be due to several factors, including differences in present census size, past changes in population size, mating system, and population substructure. The relatively small population size in humans, especially considering their large census size, has most often been attributed to a large expansion, possibly following a bottleneck, from a much smaller population at some time in the recent past (HARPENDING *et al.* 1998). That gorillas and chimpanzees have N_e at least twice as large as humans would suggest that these apes have not experienced similar dramatic population bottlenecks. The larger effective population size of gorillas relative to chimpanzees is intriguing and not likely due to a larger census size, given that at least at present gorillas have a more restricted geographic range than chimpanzees and in most habitats live at similar population densities as chimpanzees (KURODA *et al.* 1996; YAMAGIWA *et al.* 1996). Also, it is unlikely that the larger gorilla N_e is due to differences in mating system between chimpanzees and gorillas. In fact, given the single-male polygyny of gorillas (WATTS 1996) and its associated high variance in male reproductive success, one may predict the opposite—that chimpanzees would have a larger N_e since their promiscuous mating (DIXSON 1998) would lead to more males contributing to the next generation and a higher male effective population size.

Therefore, it is possible that the larger gorilla N_e may be due to their greater population subdivision. The excess of intermediate frequency variants in our gorilla sample supports this notion. Also, ecological evidence suggests that gorilla populations may be more subdivided than chimpanzee populations inhabiting the same area. Chimpanzees are able to live in a wider range of habitats including open woodland and savanna (KORTLAND 1983) and therefore may be capable of maintaining long distance gene flow between forests. In contrast, gorilla populations are restricted to forests and therefore may be unable to share migrants with other populations across open habitats. Indeed, genetic studies have revealed that chimpanzees share mtDNA haplotypes over 900 km (MORIN *et al.* 1994; GOLDBERG and RUVOLO 1997). The same has not been found for gorillas (JENSEN-SEAMAN and KIDD 2001; CLIFFORD *et al.* 2003), although far fewer wild populations of gorillas have been sampled. The increased N_e of gorillas may therefore be due to increased population subdivision relative to chimpanzees, with the caveat that although it has been shown that popu-

TABLE 1

Nucleotide diversity in each of the 49 DNA segments studied in gorillas, chimpanzees, bonobos, and humans

Segments (bp)	Human	Gorilla	Bonobo	Chimpanzee
NT2041 (440)	0	0.201	0.025	0.072
NA1364 (417)	0	0.401	0.125	0.206
NT0953 (472)	0	0.069	0	0.161
NT2012 (470)	0	0.330	0.063	0.192
NT2064 (521)	0	0	0.021	0.075
NT1584 (522)	0.006	0.320	0.040	0.295
NT2563 (522)	0.013	0.108	0.061	0.123
NT0946 (424)	0.015	0.440	0	0.217
NT2191 (391)	0.017	0.493	0	0.213
NT2659 (452)	0.015	0.171	0	0.124
NT1419 (448)	0.022	0.192	0.025	0.122
NT1506 (427)	0.023	0.227	0.110	0
NT24894 (423)	0.024	0.192	0	0.069
NT2472 (422)	0.023	0.148	0.070	0.083
NT2984 (468)	0.035	0.325	0.223	0.463
NT2609 (432)	0.038	0.071	0.363	0.300
NT812 (491)	0.014	0.161	0	0.190
NT2906 (483)	0.04	0.087	0.090	0.093
NT2986 (492)	0.043	0.014	0.096	0.077
NT2265 (451)	0.036	0.015	0.065	0.130
NT866 (447)	0.050	0.349	0	0.122
NT2019 (441)	0.049	0.436	0	0.326
NT2266 (471)	0.057	0.239	0.193	0.189
NT2018 (517)	0.066	0.090	0	0.032
NT1482 (460)	0.073	0.110	0.254	0.063
NT2963 (541)	0.068	0.245	0.098	0.150
NT2020 (456)	0.087	0	0.024	0.219
NT2568 (426)	0.057	0.244	0.026	0.028
NT813 (526)	0.093	0.070	0.080	0.022
NT1469 (467)	0.099	0.055	0	0.131
NT2294 (458)	0.124	0	0.093	0.049
NT10604 (489)	0.113	0.115	0	0.012
NT787 (442)	0.127	0.077	0.136	0.067
NT2560 (477)	0.122	0	0.190	0.012
NT2987 (446)	0.121	0	0.047	0.000
NT0151 (421)	0.122	0.139	0.053	0.424
NT2085 (506)	0.139	0.082	0.044	0.248
NT2352 (493)	0.138	0.213	0	0.068
NT1251 (493)	0.152	0.130	0.097	0.055
NT2021 (477)	0.042	0.128	0.222	0.132
NT2558 (424)	0.187	0.170	0	0.140
NT1386 (371)	0.159	0.280	0	0.000
NT1412 (553)	0.180	0.043	0.038	0.260
NA2920 (488)	0.209	0.298	0	0.127
NT784 (495)	0.222	0.088	0.165	0.101
NT864 (492)	0.213	0.249	0	0.117
NT2988 (588)	0.228	0.099	0.178	0.039
NT1636 (503)	0.252	0.040	0.127	0.134
NT2924 (468)	0.299	0	0.204	0.105
49 segments	0.087	0.158	0.076	0.134

lation subdivision can lead to an increase in N_e (WRIGHT 1943), it can also lead to a decreased N_e depending on actual levels of migration and variance in reproductive success between subpopulations (WHITLOCK and BARTON 1997; LAPORTE and CHARLESWORTH 2002).

It is especially interesting to compare the levels of diversity and estimates of N_e at the subspecies level between our sample of western lowland gorillas (*G. g. gorilla*) and our previous data from the sympatric Central African chimpanzee (*Pan t. troglodytes*; YU *et al.* 2003).

TABLE 2
Average nucleotide diversity in gorillas, chimpanzees, bonobos, and humans and effective population sizes estimated from π

	<i>n</i>	<i>s</i>	π (%)	N_e (π) ^a	Reference
49 segments (noncoding)					
<i>G. g. gorilla</i>	30	138	0.158	25,200	This study
<i>P. troglodytes</i>	34	183	0.134	21,300	This study
<i>P. paniscus</i>	13	57	0.077	12,300	This study
<i>H. sapiens</i>	60	132	0.087	10,400	This study
<i>HOXB6</i> (intergenic)					
<i>G. g. gorilla</i>	30	7	0.195	50,800	DEINARD and KIDD (1999)
<i>P. troglodytes</i>	82	8	0.176	45,900	DEINARD and KIDD (1999)
<i>P. paniscus</i>	36	6	0.175	45,600	DEINARD and KIDD (1999)
<i>H. sapiens</i>	210	4	0.060	11,700	DEINARD and KIDD (1999)
<i>ADHI</i> (intronic)					
<i>G. g. gorilla</i>	28		0.083	13,500	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>P. troglodytes</i>	22		0.101	15,900	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>P. paniscus</i>	10		0	0	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>H. sapiens</i>	20		0	0	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>APOB</i> (intronic)					
<i>G. g. gorilla</i>	30		0.133	14,700	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>P. troglodytes</i>	90		0.232	25,500	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>P. paniscus</i>	36		0.015	1,700	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>H. sapiens</i>					
<i>DRD2</i> (intronic)					
<i>G. g. gorilla</i>	27		0	0	DEINARD and KIDD (1998)
<i>P. troglodytes</i>	71		0.247	38,600	DEINARD and KIDD (1998)
<i>P. paniscus</i>	33		0.113	17,700	DEINARD and KIDD (1998)
<i>H. sapiens</i>	190		0	0	DEINARD and KIDD (1998)
<i>DRD4</i> (intergenic)					
<i>G. g. gorilla</i>	57		0.140	10,700	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>P. troglodytes</i>	32		0.332	33,300	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>P. paniscus</i>	12		0.212	21,300	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>H. sapiens</i>	110		0.097	5,900	JENSEN-SEAMAN <i>et al.</i> (2001)
<i>Xq13.3</i> (noncoding)					
<i>G. g. gorilla</i>	10	39	0.168	51,700	KAESSMANN <i>et al.</i> (1999)
<i>P. troglodytes</i>	30	85	0.131	39,000	KAESSMANN <i>et al.</i> (1999)
<i>P. paniscus</i>	5	13	0.022	7,000	KAESSMANN <i>et al.</i> (1999)
<i>H. sapiens</i>	73	33	0.033	7,000	KAESSMANN <i>et al.</i> (2001)
mtDNA D-loop					
<i>G. g. gorilla</i>	20		5.19		GARNER and RYDER (1996); JENSEN-SEAMAN and KIDD (2001)
<i>P. troglodytes</i>	41	108	7.75		DEINARD and KIDD (2000)
<i>P. paniscus</i>	17	53	3.87		DEINARD and KIDD (2000)
<i>H. sapiens</i>	54	79	2.63		INGMAN <i>et al.</i> (2000)
mtDNA <i>16S</i>					
<i>G. g. gorilla</i>	12		0.24		NODA <i>et al.</i> (2001)
<i>P. troglodytes</i>	37		0.27		NODA <i>et al.</i> (2001)
<i>P. paniscus</i>	14		0.33		NODA <i>et al.</i> (2001)
<i>H. sapiens</i>	53		0.16		NODA <i>et al.</i> (2001)
mtDNA <i>COII</i>					
<i>G. g. gorilla</i>	6		1.0		RUVOLO <i>et al.</i> (1994)
<i>P. troglodytes</i>	5		1.0		RUVOLO <i>et al.</i> (1994)
<i>P. paniscus</i>	4		0.5		RUVOLO <i>et al.</i> (1994)
<i>H. sapiens</i>	6		0.6		RUVOLO <i>et al.</i> (1994)

s, segregating sites.

^aThe average mutation rate (*u*) is 1.0469×10^{-9} /site/year used for estimating N_e for the 49 segments. Generation lengths are assumed to be 15 and 20 years for apes and humans, respectively.

TABLE 3

Average sequence divergence (%) between taxa estimated from the 49 DNA segments studied

Species	Chimpanzee	Bonobo	Gorilla
Bonobo	0.377		
Gorilla	1.656	1.74	
Human	1.229	1.31	1.64

The geographic range of the western gorilla is contained entirely within that of this chimpanzee subspecies (GROVES 1971; KORTLAND 1983). *P. t. troglodytes* has the highest levels of nucleotide diversity among the three common chimpanzee subspecies as estimated from mtDNA data (MORIN *et al.* 1994; DEINARD and KIDD 2000), X chromosomal data (KAESSMANN *et al.* 1999), Y chromosomal data (STONE *et al.* 2002), and autosomal data (YU *et al.* 2003). Similarly, it is apparent that *G. g. gorilla* has substantially more nucleotide diversity than other gorilla subspecies on the basis of mtDNA data (GARNER and RYDER 1996; JENSEN-SEAMAN and KIDD 2001). Thus, these sympatric African ape subspecies are each the most diverse representatives of their respective genera, suggesting that perhaps ecological conditions in this part of equatorial West Africa have been more conducive to maintaining long-term effective population sizes. Several forest refuges within the ranges of these subspecies have been proposed, which may have buffered them against population reductions during climatic fluctuations of the African Quaternary (LIVINGSTONE 1982; MALEY 1996). Where they differ, however, is that western lowland gorillas have somewhat greater π and N_e than Central African chimpanzees. More strikingly, these gorillas have an excess of intermediate frequency mutations, while this chimpanzee subspecies has an excess of singletons (YU *et al.* 2003), suggesting a greater level of population subdivision in gorillas within the same geographic area as chimpanzees (AVISE 2000). This may have been especially true during periods of forest reduction and fragmentation associated with global cooling and drying over the last several hundred thousand years because chimpanzees are capable of living in dry savanna or open woodland environments by maintaining communities with very large home ranges, while gorillas are not found in such open environments (YAMAGIWA 1999). Of course, other alternative explanations could be invoked; for example, the eastern gorilla subspecies may have originated as migrants from West Africa and therefore the reduced variation in the former may be a result of a population bottleneck associated with colonization.

Our estimates of the N_e of African apes are at most only 2.5 times larger than that of modern humans. This estimate is close to the average of five nuclear loci of

JENSEN-SEAMAN *et al.* (2001) but is considerably lower than that of CHEN and LI (2001), who estimated the N_e of the ancestral human-chimpanzee population to be between five and nine times that of modern humans. On the other hand, a reanalysis of Chen and Li's data set using the maximum-likelihood method suggested an effective population size of $\sim 12,000$ for the ancestor of humans and chimpanzees (YANG 2002). This estimate is only slightly larger than the estimate (10,400) of N_e for modern humans and much lower than that (21,300) for extant chimpanzees. The N_e for the common ancestor of all three species was estimated to be $\sim 38,000$ (YANG 2002), which is higher than the present estimates of N_e for any living African ape. It is tempting to speculate that this last common ancestor of the African apes and humans, with its large effective population size, may have shared some ecological characteristics with gorillas, the most diverse living African ape. Furthermore, one may speculate that some of the social or ecological changes resulting in a lower N_e in extant humans and chimpanzees may have already begun to occur in their common ancestor following its divergence from gorilla. Of course, we recognize that the long-term N_e of any species may be influenced by unrecoverable idiosyncrasies in the species' unique history. Further research is needed to reconcile the different estimates proposed for the effective population sizes of our ancestors and to test hypotheses seeking to explain the differences among species.

Mitochondrial vs. nuclear DNA: As one can see from the above studies, most of the data disclosed twice as much nucleotide diversity in gorillas as in humans, assessed using both mitochondrial and nuclear DNA. Thus, unlike chimpanzees, there is a similar ratio of nucleotide diversity between humans and gorillas in both nuclear and mtDNA data (Table 2). WISE *et al.* (1997) pointed out a disparity in using mtDNA vs. nuclear DNA between humans and chimpanzees and in further comparisons to other nonhuman primates they suggested that humans, not chimpanzees, were unusual in possessing such low levels of mtDNA diversity relative to that of the nuclear genome. In the 49 autosomal segments studied, the difference in nucleotide diversity between humans and chimpanzees is considerably smaller for nuclear DNA than for mtDNA data (YU *et al.* 2003), while a similar level was observed in gorillas in this study. Therefore, this disparity may arise from the unusually high estimates of chimpanzee mtDNA diversity. There are several possible explanations. First, not all loci are expected to give the same results because of stochastic effects. The fourfold smaller effective population size in mtDNA compared to nuclear DNA will increase the stochastic aspects of drift. Second, the cause of this disparity could be a reduction in the effective population size (N_e) in the human lineage since the human-chimp divergence; a reduction in N_e causes a

larger decrease in nucleotide diversity for mtDNA than for nuclear DNA. This is also true for bonobos, which may also have experienced a population reduction during the recent past. Third, different studies have used different samples, which can strongly affect the results of diversity studies. Finally, it may be possible that balancing selection or local directional selection has acted to increase variation in the chimpanzee mitochondrial genome.

Conclusion: Gorillas possess the greatest amount of autosomal nucleotide diversity and the largest effective population size among all of the living species in the African ape-human clade, with about twice as much diversity as modern humans. Gorillas also show the greatest evidence of population subdivision. A reduction in effective population size may have occurred in the common ancestor of humans and chimpanzees following divergence from the gorillas. Finally, we hope that this understanding of the amount and pattern of genetic variation in the African apes, along with future studies that sample wild populations, can help in establishing conservation priorities for these endangered species.

We thank George Amato, Amos Deinard, and Kenneth Kidd for their generous donation of blood samples or DNA. We thank the National Zoo and Miami Metro Zoo for their donation of gorilla blood samples. This study was supported by National Institutes of Health grants GM55759 and GM30998.

LITERATURE CITED

- AVISE, J. C., 2000 *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.
- BRUNET, M., F. GUY, D. PILBEAM, H. T. MACKAYE, A. LIKIUS *et al.*, 2002 A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* **418**: 145–151.
- CANN, R., M. STONEKING and A. C. WILSON, 1987 Mitochondrial DNA and human evolution. *Nature* **325**: 31–36.
- CHEN, F.-C., and W.-H. LI, 2001 Genomic divergences between human and other hominoids and the effective population size of the common ancestor of human and chimpanzees. *Am. J. Hum. Genet.* **68**: 444–456.
- CLIFFORD, S. L., K. A. ABERNETHY, L. J. T. WHITE, C. E. G. TUTIN, M. W. BRUFORD *et al.*, 2003 Genetic studies of western gorillas, pp. 269–292 in *Gorilla Biology: A Multidisciplinary Perspective*, edited by A. B. TAYLOR and M. L. GOLDSMITH. Cambridge University Press, Cambridge, UK.
- COOLIDGE, H. J., Jr., 1929 A revision of the genus *Gorilla*. *Memoirs Museum Comp. Zool. Harvard Coll.* **50**: 292–381.
- DEINARD, A. S., and K. K. KIDD, 1998 Evolution of a D2 dopamine receptor intron within the great apes and humans. *DNA Seq.* **8**: 289–301.
- DEINARD, A. S., and K. K. KIDD, 1999 Evolution of a HOXB6 intergenic region within the great apes and humans. *J. Hum. Evol.* **36**: 687–703.
- DEINARD, A. S., and K. K. KIDD, 2000 Identifying conservation units within captive chimpanzee populations. *Am. J. Phys. Anthropol.* **111**: 25–44.
- DIXSON, A. F., 1998 *Primate Sexuality*. Oxford University Press, Oxford.
- DON, R. H., P. T. COX, B. J. WAINWRIGHT, K. BAKER and J. S. MATTICK, 1991 “Touchdown” PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* **19**: 4008.
- FERRIS, S. D., W. M. BROWN, W. S. DAVIDSON and A. C. WILSON, 1981 Extensive polymorphism in the mitochondrial DNA of apes. *Proc. Natl. Acad. Sci. USA* **78**: 6319–6323.
- GAGNEUX, P., C. WILLS, U. GERLOFF, D. TAUTZ, P. A. MORIN *et al.*, 1999 Mitochondrial sequences show diverse evolutionary histories of African hominoids. *Proc. Natl. Acad. Sci. USA* **96**: 5077–5082.
- GARNER, K. J., and O. A. RYDER, 1996 Mitochondrial DNA diversity in gorillas. *Mol. Phylogenet. Evol.* **6**: 39–48.
- GOLDBERG, T. L., and M. RUVOLO, 1997 The geographic apportionment of mitochondrial genetic diversity in East African chimpanzees, *Pan troglodytes schweinfurthii*. *Mol. Biol. Evol.* **14**: 976–984.
- GROVES, C. P., 1971 Distribution and place of origin of the gorilla. *Man* **6**: 44–51.
- GROVES, C. P., 2001 *Primate Taxonomy*. Smithsonian Press, Washington, DC.
- GROVES, C. P., 2003 A history of gorilla taxonomy, pp. 15–34 in *Gorilla Biology: A Multidisciplinary Perspective*, edited by A. B. TAYLOR and M. L. GOLDSMITH. Cambridge University Press, Cambridge, UK.
- HARPENDING, H. C., M. A. BATZER, M. GURVEN, L. B. JORDE, A. R. ROGERS *et al.*, 1998 Genetic traces of ancient demography. *Proc. Natl. Acad. Sci. USA* **95**: 1961–1967.
- INGMAN, M., H. KAESSMANN, S. PÄÄBO and U. GYLLENSTEN, 2000 Mitochondrial genome variation and the origin of modern humans. *Nature* **408**: 708–713.
- JENSEN-SEAMAN, M. I., 2000 Evolutionary genetics of gorillas. Ph.D. Thesis, Yale University, New Haven, CT.
- JENSEN-SEAMAN, M. I., and K. K. KIDD, 2001 Mitochondrial DNA variation and biogeography of eastern gorillas. *Mol. Ecol.* **10**: 2241–2247.
- JENSEN-SEAMAN, M. I., A. S. DEINARD and K. K. KIDD, 2001 Modern African ape populations as genetic and demographic models of the last common ancestor of humans, chimpanzees, and gorillas. *J. Hered.* **92**: 475–480.
- JENSEN-SEAMAN, M. I., A. S. DEINARD and K. K. KIDD, 2003 Mitochondrial and nuclear DNA estimates of divergence between western and eastern gorillas, pp. 247–268 in *Gorilla Biology: A Multidisciplinary Perspective*, edited by A. B. TAYLOR and M. L. GOLDSMITH. Cambridge University Press, Cambridge, UK.
- KAESSMANN, H., V. WIEBE and S. PÄÄBO, 1999 Extensive nuclear DNA sequence diversity among chimpanzees. *Science* **286**: 1159–1162.
- KAESSMANN, H., V. WIEBE, G. WEISS and S. PÄÄBO, 2001 Great ape DNA sequences reveal a reduced diversity and an expansion in humans. *Nat. Genet.* **27**: 155–156.
- KORTLAND, A., 1983 Marginal habitats of chimpanzees. *J. Hum. Evol.* **12**: 231–278.
- KURODA, S., T. NISHIHARA, S. SUZUKI and R. A. OKO, 1996 Sympatric chimpanzees and gorillas in the Ndoki Forest, Congo, pp. 71–81 in *Great Ape Societies*, edited by W. C. MCGREW, L. F. MARCHANT and T. NISHIDA. Cambridge University Press, Cambridge, UK.
- LAPORTE, V., and B. CHARLESWORTH, 2002 Effective population size and population subdivision in demographically structured populations. *Genetics* **162**: 501–519.
- LEE, P. C., J. THORNBACK and E. L. BENNETT, 1988 *Threatened Primates of Africa: The IUCN Red Data Book*. IUCN—The World Conservation Union, Gland, Switzerland/Cambridge, UK.
- LIVINGSTONE, D. A., 1982 Quaternary geography of Africa and the refuge theory, pp. 523–535 in *Biological Diversification in the Tropics*, edited by G. T. PRANCE. Columbia University Press, New York.
- MALEY, J., 1996 The African rain forest: main characteristics of changes in vegetation and climate change from the Upper Cretaceous to the Quaternary. *Proc. R. Soc. Edinb. [Biol]* **104**: 31–37.
- MORIN, P. A., J. J. MOORE, R. CHAKRABORTY, L. JIN, J. GOODALL *et al.*, 1994 Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* **265**: 1193–1201.
- NEI, M., and D. GRAUR, 1984 Extent of protein polymorphism and the neutral mutation theory. *Evol. Biol.* **17**: 73–118.
- NODA, R., C. G. KIM, O. TAKENAKA, R. E. FERRELL, T. TANOUÉ *et al.*, 2001 Mitochondrial 16S rRNA sequence diversity of hominoids. *J. Hered.* **92**: 490–496.
- ROZAS, J., and R. ROZAS, 1999 DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**: 174–175.
- RUVOLO, M., D. PAN, S. ZEHR, T. GOLDBERG, T. DISOTELL *et al.*, 1994

- Gene trees and hominoid phylogeny. *Proc. Natl. Acad. Sci. USA* **91**: 8900–8904.
- STONE, A. C., R. C. GRIFFITHS, S. L. ZEGURA and M. F. HAMMER, 2002 High levels of Y-chromosome nucleotide diversity in the genus *Pan*. *Proc. Natl. Acad. Sci. USA* **99**: 43–48.
- TAJIMA, F., 1983 Evolution relationship of DNA sequences in finite populations. *Genetics* **105**: 437–460.
- TAKAHATA, N., Y. SATTA and J. KLEIN, 1995 Divergence time and population size in the lineage leading to modern humans. *Theor. Popul. Biol.* **48**: 198–221.
- TISHKOFF, S. A., E. DIETZSCH, W. SPEED, A. J. PAKSTIS, J. R. KIDD *et al.*, 1996 Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* **271**: 1380–1387.
- VIGNAUD, P., P. DURINGER, H. T. MACKAYE, A. LIKIUS, C. BLONDEL *et al.*, 2002 Geology and palaeontology of the upper Miocene Toros-Menalla hominid locality, Chad. *Nature* **418**: 152–155.
- WATTS, D. P., 1996 Comparative socio-ecology of gorillas, pp. 16–28 in *Great Ape Societies*, edited by W. C. MCGREW, L. F. MARCHANT and T. NISHIDA. Cambridge University Press, Cambridge, UK.
- WHITLOCK, M. C., and N. H. BARTON, 1997 The effective size of a subdivided population. *Genetics* **146**: 427–441.
- WISE, C. A., M. SRAML, D. C. RUBINSZTEIN and S. EASTEAL, 1997 Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees. *Mol. Biol. Evol.* **14**: 707–716.
- WRIGHT, S., 1943 Isolation by distance. *Genetics* **28**: 114–138.
- XIA, X., and Z. XIE, 2001 DAMBE: data analysis in molecular biology and evolution. *J. Hered.* **92**: 371–373.
- YAMAGIWA, J., 1999 Sociological factors influencing population structure of gorillas and chimpanzees. *Primates* **40**: 87–104.
- YAMAGIWA, J., T. MARUHASHI, T. YUMOTO and N. MWANZA, 1996 Dietary and ranging overlap in sympatric gorillas and chimpanzees in Kahuzi-Biega National Park, Zaire, pp. 82–98 in *Great Ape Societies*, edited by W. C. MCGREW, L. F. MARCHANT and T. NISHIDA. Cambridge University Press, Cambridge, UK.
- YANG, Z., 2002 Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics* **162**: 1811–1823.
- YU, N., F.-C. CHEN, S. OTA, L. JORDE, P. PAMILO *et al.*, 2002 Larger genetic differences within Africans than between Africans and Eurasians. *Genetics* **161**: 269–274.
- YU, N., M. I. JENSEN-SEAMAN, L. CHEMNICK, J. R. KIDD, A. S. DEINARD *et al.*, 2003 Low nucleotide diversity in chimpanzees and bonobos. *Genetics* **164**: 1511–1518.
- ZHAO, Z., L. JIN, Y.-X. FU, M. RAMSAY, T. JENKINS *et al.*, 2000 World-wide DNA sequence variation in a 10 kilobase noncoding region on chromosome 22. *Proc. Natl. Acad. Sci. USA* **97**: 11354–11358.

Communicating editor: Y. X. Fu

