Low Nucleotide Diversity in Chimpanzees and Bonobos

Ning Yu,* Michael I. Jensen-Seaman,*,[1] Leona Chemnick,† Judith R. Kidd,‡ Amos S. Deinard,§ Oliver Ryder,† Kenneth K. Kidd‡ and Wen-Hsiung Li*,[2]

*Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, †Center for Reproduction of Endangered Species, Zoological Society of San Diego, San Diego, California 92101, ‡Department of Human Genetics, Yale University School of Medicine, New Haven, Connecticut 06520-8005 and §School of Veterinary Medicine and Department of Anthropology, University of California, Davis, California 95616

ABSTRACT

Comparison of the levels of nucleotide diversity in humans and apes may provide much insight into the mechanisms of maintenance of DNA polymorphism and the demographic history of these organisms. In the past, abundant mitochondrial DNA (mtDNA) polymorphism data indicated that nucleotide diversity (π) is more than threefold higher in chimpanzees than in humans. Furthermore, it has recently been claimed, on the basis of limited data, that this is also true for nuclear DNA. In this study we sequenced 50 noncoding, nonrepetitive DNA segments randomly chosen from the nuclear genome in 9 bonobos and 17 chimpanzees. Surprisingly, the π value for bonobos is only 0.078%, even somewhat lower than that (0.088%) for humans for the same 50 segments. The π values are 0.092, 0.130, and 0.082% for East, Central, and West African chimpanzees, respectively, and 0.132% for all chimpanzees. These values are similar to or at most only 1.5 times higher than that for humans. The much larger difference in mtDNA diversity than in nuclear DNA diversity between humans and chimpanzees is puzzling. We speculate that it is due mainly to a reduction in effective population size (N_e) in the human lineage after the human-chimpanzee divergence, because a reduction in N_e has a stronger effect on mtDNA diversity than on nuclear DNA diversity.

SINCE the discovery of extensive mitochondrial DNA (mtDNA) polymorphism in apes by restriction enzyme mapping (Ferris *et al.* 1981), it has been known that the nucleotide diversity (π) in mtDNA is at least threefold higher in chimpanzees than in humans. This view has been confirmed by recent sequence data from the control region (Wise *et al.* 1997) and from synonymous sites in the *ND*2 gene (Stone *et al.* 2002). On the other hand, since the late 1970s it has been known that the level of heterozygosity at protein coding loci is higher in humans than in chimpanzees (King and Wilson 1975; Lucotte 1983). Supporting this view that humans may have as much or greater diversity in the nuclear genome was the observation that humans had higher levels of heterozygosity at microsatellite loci than chimpanzees (Wise *et al.* 1997), although such a difference seen in similar studies was potentially attributable to ascertainment bias (Ellegren *et al.* 1995; Crouau-

Roy *et al.* 1996). Therefore, it was commonly thought that mtDNA and nuclear DNA gave different pictures of polymorphism in humans and chimpanzees. However, recent DNA polymorphism data from a 10-kb X-linked noncoding region, two intergenic (*HOXB6, DRD4,* ~1 kb each), two intronic (*ADH1,* ~600 bp; *DRD2,* ~300 bp) regions, and 5.8-kb silent sites in genes at six nuclear loci revealed a three- to fourfold higher nucleotide diversity in chimpanzees than in humans (Deinard and Kidd 1999, 2000; Kaessmann *et al.* 1999; Jensen-Seaman *et al.* 2001; Satta 2001), leading to the view that, like mtDNA, nuclear DNA sequence diversity is also much higher in chimpanzees than in humans.

However, as the data are limited, this issue deserves further investigation. In a recent study Yu *et al.* (2002) sequenced 50 DNA segments randomly chosen from the noncoding, nonrepetitive parts of the human genome in 30 humans from various localities around the world. In the present study we have sequenced the same 50 segments in 9 bonobos and 17 chimpanzees from East, Central, and West Africa. Unexpectedly, the new data reveal a small difference between the levels of nucleotide diversity in chimpanzees and humans. Therefore, nuclear DNA and mtDNA actually give different pictures of the levels of nucleotide diversity in humans and chimpanzees. How this difference arose is a puzzling question and we will attempt to find an answer.

MATERIALS AND METHODS

Sample sources: DNA from nine bonobos (*Pan paniscus*) and 17 common chimpanzees (six *P. troglodytes verus*, five *P. t. troglodytes*, two *P. t. schweinfurthii*, and four individuals of unknown subspecies) was used in this study. The five *P. t. troglodytes* individuals (named Cheetah, Dodo, Bakoumba, Julie, and Noemie) were from J. Wickings, CIRMF, Gabon. Three of the six *P. t. verus* individuals (Rinus, Anita, and Hannibal) were from A. Prince, Vialab, Liberia, one (Herman) was from the Lowery Zoo, and two (Bert and Tate) were from the New Iberia Research Center; four individuals (Carl, Kasey, Harv, and Tank) were of unknown geographic origin. The two *P. t. schweinfurthii* individuals (Harriet and Kobi) were from J. Fritz, Arizona Primate Foundation. Although no geographical information was available for the individuals housed at the New Iberia Research Center, they were regarded as *P. t. verus* on the basis of their mitochondrial D-loop sequences (MORIN *et al.* 1994), and the nuclear sequences generated from these samples did not contradict this classification (DEINARD and KIDD 2000). Of the nine *P. paniscus* individuals, two (Bosondjo and Matata) were from the Atlanta Zoo/Yerkes Regional Primate Center; two (Lody and Maringa) from the Milwaukee Zoo, and five (Kakowet, Lokalema, Charlie, Vernon, and Linda) from the San Diego Zoo. However, all nine individuals were originally caught in Zaire and all of them were chosen to be 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.

Touchdown PCR (DON *et al.* 1991) was used and the reactions were carried out following the conditions 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, Piscataway, NJ) 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.

Data analysis: The sequences were aligned by SeqMan. Nucleotide diversity values were calculated using DNASP version 3.14 (ROZAS and ROZAS 1999) and the average percentage distances between species were calculated using DAMBE (XIA and XIE 2001).

RESULTS AND DISCUSSION

Distribution of SNPs: We sequenced 50 noncoding segments in nine bonobos and 17 chimpanzees from East, Central, and West Africa. The total number of nucleotide sites sequenced, after exclusion of deletions and insertions (mostly single nucleotide indels), is ~23,500.

A total of 186 single nucleotide polymorphisms (SNPs) were found in the 17 chimpanzee samples (34 sequences); 51 of them were observed only once (*i.e.*, singletons) and 15 only twice (doubletons). The number of variant sites found was 54 in the 12 West African chimpanzee (*P. t. verus*) sequences, 101 in the 10 Central African chimpanzee (*P. t. troglodytes*) sequences, and 39 in the 4 East African chimpanzee (*P. t. schweinfurthii*) sequences. Thus, many more variants were found in the Central African subspecies than in the West and East African subspecies, indicating a much higher DNA diversity in the Central African subspecies. The numbers of singletons were 14, 58, and 27 in the West, Central, and East African chimpanzee sequences, respectively. Thus, in the Central and East African subspecies, more than half of the variants were singletons, whereas in the West African subspecies less than one-third of the variants were singletons with an equal number of doubletons. A total of 63 SNPs were found in the 18 bonobo sequences; there were 21 singletons and 11 doubletons. Clearly, bonobos are less polymorphic than each of the three chimpanzee subspecies.

Adequacy of the sample sizes: As our sample sizes are relatively small, we need to consider the problem of 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. The two sequences in an individual are not completely independent if the individual is "inbred" to some extent. Anyway, 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. To simplify the analysis, we concatenate the segments in an individual in a random manner into two continuous sequences and these sequences are then used to compute the π_b and π_w values.

For the Central African chimpanzee sequences, the distribution of π_b values, which ranges from 0.098 to 0.174%, is only somewhat wider than that of the five π_w values, which ranges from 0.102 to 0.153%. Since the average π_b (0.131%) is <10% higher than the average π_w (0.121%), the sampling bias should not be strong.
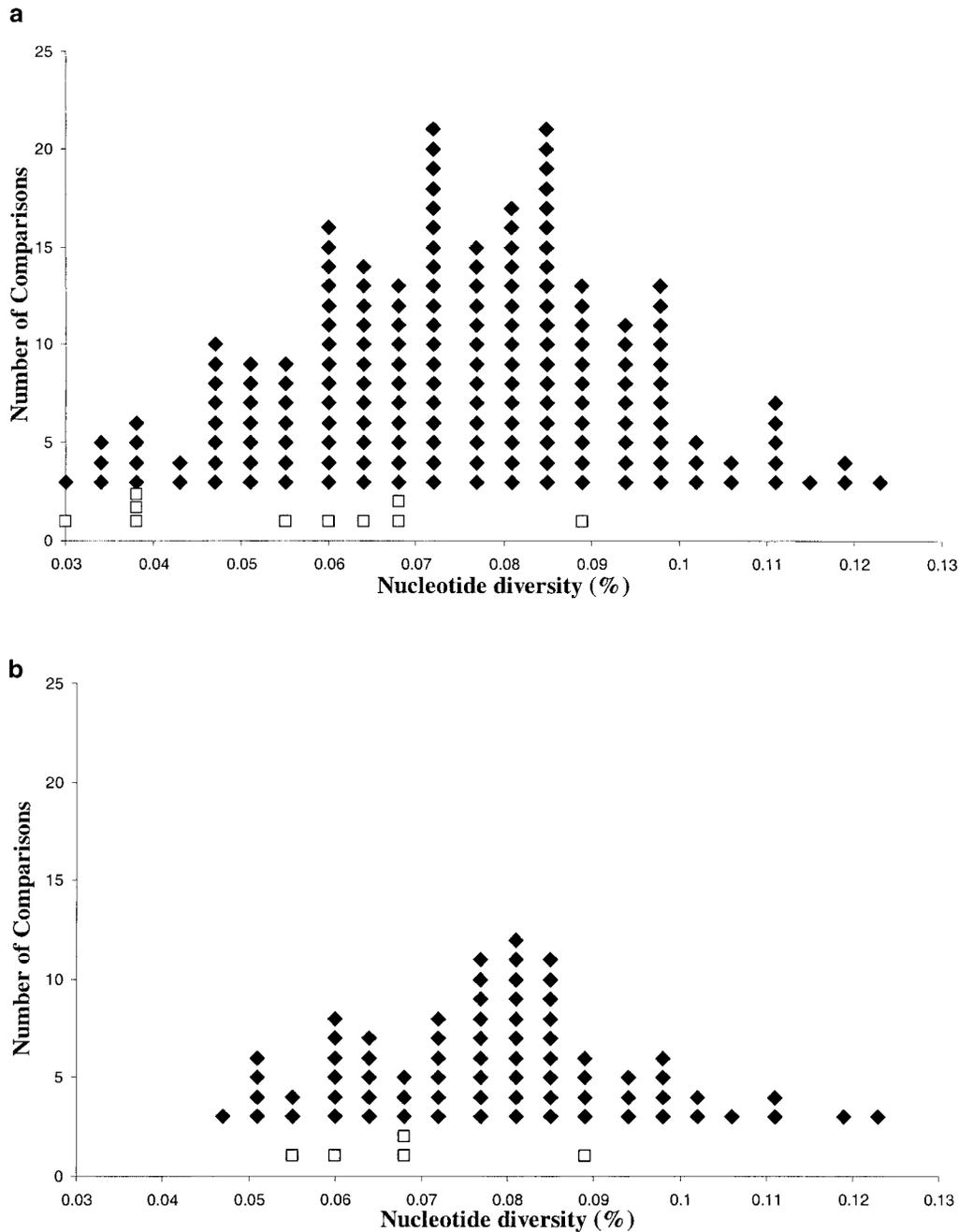
FIGURE 1.—Distributions of the within-individual (π_w; □) and between-individual (π_b; ◆) nucleotide diversity values in bonobos. (a) All of the 18 sequences are included. (b) One individual (Bosondjo) and one sequence (randomly chosen) from each of three individuals (Kakowet, Lody, and Linda) are excluded.

A similar comment applies to the West African chimpanzee sample. There is one very low π_w value (0.051%) and the average π_w value (0.072%) is ∼10% lower than the average π_b value (0.082%). The average π_w without the outlier becomes 0.077%, which is not far from the average π_b (0.082%). This comparison suggests that the estimated π value (0.082%) in this subspecies is probably somewhat biased downward. For the East African chimpanzee sample, the average π_w value (0.088%) is close to the average π_b value (0.092%). This may suggest no substantial bias. However, because only two individuals were sampled, the estimate may not be reliable and should be taken with caution.

For the bonobo sample, the distributions of π_w and π_b values are given in Figure 1a. Several π_w and π_b values are <0.04%, whereas most of the others are considerably >0.04%. This observation suggests that some of the individuals are fairly closely related to each other or inbred, although they were originally chosen to be independent. Indeed, the four π_b values between Bosondjo and Maringa (studbook numbers 64 and 60, respectively) were only 0.034, 0.034, 0.038, and 0.047%. We therefore excluded Bosondjo from comparison. Moreover, the π_w value is only 0.030% for Kakowet (studbook no. 34) and 0.038% for Lody and Linda (studbook nos. 68 and 23). We therefore excluded one sequence (randomly chosen) from each of these three individuals. After the exclusion of these sequences, only 13 se-

quences remain and the new distributions of π_w and π_b values are given in Figure 1b. The exclusion of the above sequences increases the average π value from 0.075 to 0.078%. The latter value will be used as our estimate.

Nucleotide diversity: For the 50 DNA segments we obtained, the range of π is from 0 (14 segments) to 0.39% in the West African chimpanzee sample, from 0 (7 segments) to 0.46% in the Central African chimpanzee sample, from 0 (23 segments) to 0.45% in the East African chimpanzee sample, and from 0 (3 segments) to 0.46% in the entire chimpanzee sample (Table 1). The range of π is from 0 to 0.36% in the bonobo sample and from 0 to 0.30% in the human sample (Table 1). Such large fluctuations 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.

Table 2 shows that Central African chimpanzees have the highest average π value (0.130%), followed by East African chimpanzees (0.092%), and then West African chimpanzees (0.082%). As mentioned above, the π value estimated from the East African chimpanzee sample may not be reliable because of a small sample size and that from the West African chimpanzee sample might be biased downward. We note further that in previous studies the East African chimpanzee had a greater π at *APOB* and *PABX* but a smaller π at the *HOXB6* intergenic region and at the mtDNA *ND2* and the mtDNA control region than did the West African chimpanzee (Table 2; WISE *et al.* 1997; DEINARD and KIDD 2000; STONE *et al.* 2002). Therefore, further data are needed to see whether the π value for the East African chimpanzee is actually higher than that for the West African chimpanzee.

Surprisingly, the average π value in bonobos (0.078%) is somewhat lower than that in humans (0.088%; Table 2). The observation that bonobos have lower nucleotide diversity than humans is in agreement with the Xq13.3 data, but contrary to the *HOXB6*, *DRD4*, *DRD2*, and *NRY* regions. Furthermore, the average π values in the East and West African chimpanzee subspecies (0.092 and 0.082%) are similar to that in humans. The Central African chimpanzee is the only subspecies that has a π value (0.130%) higher than that in humans and the difference is only 50%. We note further that even the highest π_b value for the concatenated sequences in the Central African chimpanzee sample is only 0.174% (see above), which is only two times the average π value in humans. When all chimpanzee sequences are considered together, the average π value (0.132%) is again only 50% higher than that in humans. Actually, if we consider African humans only, the π value for the 50 DNA segments becomes 0.115% (YU *et al.* 2002), which is only 13% lower than that for chimpanzees.

Previous reports have suggested that chimpanzees have two to four times greater amounts of π than hu-

mans at several autosomal nuclear loci, including the noncoding intergenic regions near *HOXB6* and *DRD4* (DEINARD and KIDD 1999; JENSEN-SEAMAN *et al.* 2001), the intronic regions of *DRD2* and *AHD1* (DEINARD and KIDD 1998; JENSEN-SEAMAN *et al.* 2001), and synonymous sites at six protein coding loci (SATTA 2001). Similar results were also observed at the X chromosomal locus Xq13.3 (KAESSMANN *et al.* 1999) and the Y chromosomal *NRY* locus (STONE *et al.* 2002). However, since our data set has a much wider genomic representation, it should be more reliable. The stochastic effects of examining only a small number of loci can be seen, for example, in that in one-third (17 of 50) of the segments that we examined chimpanzees had more than three times higher π than humans (Table 1), which is approximately the value that was found by others when looking at a single locus (Table 2; DEINARD and KIDD 1999; KAESSMANN *et al.* 1999). Thus, the difference in nucleotide diversity between humans and chimpanzees is considerably smaller for nuclear DNA than for mtDNA data (Table 2). The disparity in using mtDNA *vs.* nuclear DNA in comparisons between humans and chimpanzees was originally pointed out by WISE *et al.* (1997), who in further comparisons to other nonhuman primates suggested that humans, not chimpanzees, were unusual in possessing such low levels of mtDNA diversity relative to that of the nuclear genome.

Another measure of genetic variability is the number of segregating alleles in the sample. However, because the sample sizes are different for different populations, we consider $\theta = 4N_e u$, where N_e is the effective population size and u is the mutation rate per site per generation. The θ values estimated from the numbers of segregating sites by Watterson's estimator (WATTERSON 1975) are given in Table 2. We note that, compared to the difference in π (0.088 *vs.* 0.078%) between humans and bonobos, the difference in θ (0.123 *vs.* 0.082%) is even larger (Table 2), probably reflecting an increase in low-frequency alleles due to a recent population expansion in humans. The highest θ value for the three chimpanzee subspecies (0.152%) is that for Central African chimpanzees, but it is only 24% higher than that for humans. When all chimpanzee sequences are considered together, θ becomes 0.194%, which is only 50% higher than that in humans.

Effective population sizes: To estimate effective population size (N_e) we estimate 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 the human and chimpanzee-bonobo lineages is 6 MY (BRUNET *et al.* 2002; VIGNAUD *et al.* 2002). Since we are interested in the long-term effective population size, we use Tajima's estimator, $\pi = 4N_e u$ (TAJIMA 1983), and we assume that the generation time is 15 years for chimpanzees and bonobos and 20 years for humans. The N_e for humans is estimated to be 10,500 (Table 2), which is similar to

TABLE 1

Nucleotide diversity in each of the 50 DNA segments studied in chimpanzees, bonobos, and humans

Segments (bp)	Nucleotide diversity (%)					
	Human[a]	Bonobo	Chimp	$P.t.v.$[b]	$P.t.t.$[b]	$P.t.s.$[b]
NT2041 (440)	0	0.025	0.072	0.048	0.081	0.114
NA1364 (417)	0	0.125	0.206	0	0.405	0.36
NT0953 (472)	0	0	0.161	0.111	0.198	0
NT2012 (471)	0	0.062	0.192	0	0.245	0.106
NT2064 (521)	0	0.021	0.075	0.07	0.077	0.096
NT1584 (521)	0.006	0.04	0.295	0.122	0.294	0
NT2563 (523)	0.013	0.061	0.123	0.14	0.102	0.096
NT0946 (424)	0.015	0	0.217	0.25	0.047	0.118
NT2191 (392)	0.017	0	0.213	0	0.312	0.425
NT2659 (452)	0.015	0	0.124	0.081	0.167	0.111
NT1419 (448)	0.022	0.025	0.122	0.025	0.193	0
NT1506 (427)	0.023	0.11	0	0	0	0
NT24894 (423)	0.024	0	0.069	0.049	0.142	0
NT2472 (422)	0.023	0.07	0.083	0.076	0.095	0
NT2984 (470)	0.035	0.223	0.461	0.375	0.459	0
NT2609 (432)	0.038	0.363	0.3	0.354	0.129	0.116
NT812 (490)	0.014	0	0.19	0.173	0.15	0.238
NT2906 (483)	0.04	0.09	0.093	0.126	0.034	0
NT2986 (492)	0.043	0.096	0.077	0.06	0.108	0
NT2265 (451)	0.036	0.065	0.13	0.065	0.197	0.148
NT866 (447)	0.05	0	0.122	0.082	0.194	0
NT2019 (440)	0.049	0	0.322	0.231	0.217	0.455
NT2266 (471)	0.057	0.193	0.189	0.14	0.075	0.106
NT2018 (517)	0.066	0	0.032	0	0.069	0.097
NT1482 (460)	0.073	0.254	0.063	0	0.174	0
NT2963 (541)	0.068	0.098	0.15	0.137	0.086	0.123
NT2020 (456)	0.087	0.024	0.219	0.115	0.102	0.11
NT2568 (426)	0.057	0.026	0.028	0	0.047	0
NT813 (526)	0.093	0.08	0.022	0	0.068	0
NT1469 (467)	0.099	0	0.131	0.21	0	0
NT2294 (458)	0.124	0.093	0.049	0.024	0.078	0.109
NT10604 (489)	0.113	0	0.012	0	0.041	0
NT787 (422)	0.125	0.136	0.067	0.096	0	0.118
NT2560 (477)	0.122	0.19	0.012	0.023	0	0
NT2987 (448)	0.112	0.047	0	0	0	0
NT0151 (421)	0.122	0.053	0.424	0.394	0.132	0.119
NT2085 (506)	0.139	0.044	0.248	0.063	0.246	0
NT2352 (493)	0.138	0	0.068	0	0.113	0.135
NT1251 (493)	0.152	0.097	0.055	0.074	0.041	0
NT2021 (477)	0.042	0.222	0.132	0	0.382	0.105
NT2558 (427)	0.186	0	0.139	0.069	0.125	0.117
NT1386 (371)	0.159	0	0	0	0	0
NT1412 (552)	0.18	0.038	0.26	0.02	0.233	0.181
NA2920 (470)	0.209	0	0.127	0.06	0.146	0.273
NT784 (495)	0.222	0.165	0.101	0.085	0.094	0
NT2557 (431)	0.212	0	0.041	0.052	0	0.116
NT864 (492)	0.213	0	0.117	0.173	0.041	0
NT2988 (584)	0.229	0.179	0.04	0.019	0.061	0
NT1636 (505)	0.251	0.127	0.133	0.116	0.11	0.297
NT2924 (468)	0.299	0.204	0.105	0	0.185	0.214
Average	0.0882	0.073	0.132	0.086	0.130	0.092

[a] he human data are from YU *et al.* (2002).
[b] *P.t.v.*, West African chimpanzee; *P.t.t.*, Central African chimpanzee; *P.t.s.*, East African chimpanzee.

TABLE 2

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

	n	s	π (%)	θ (%)	N_e $(\pi)^a$	References
50 segments (noncoding)						
P.t.s.	4	46	0.092	0.110	14,600	1
P.t.t.	10	101	0.130	0.152	20,100	1
P.t.v.	12	54	0.082	0.076	13,000	1
P. troglodytes	34	186	0.132	0.194	20,900	1
P. paniscus	13	60	0.078	0.082	12,400	1
H. sapiens	60	118	0.088	0.123	10,500	1
HoxB6 (intergenic)						
P.t.s.	6	0	0	0		2
P.t.t.	18	5	0.119	0.143	31,000	2
P.t.v.	58	5	0.105	0.106	27,400	2
P. troglodytes	82	8	0.176	0.158	45,900	2
P. paniscus	36	6	0.175	0.142	45,600	2
H. sapiens	210	4	0.060	0.066	11,700	2
Xq13.3 (noncoding)						
P.t.t.	12	64	0.175	0.209	52,000	3
P.t.v.	17	24	0.050	0.07	15,000	3
P. troglodytes	30	85	0.131	0.211	39,000	3
P. paniscus	5	13	0.022	0.024	7,000	3
H. sapiens	69	33	0.033	0.068	7,000	4
mtDNA ND2 (synonymous)						
P.t.s.	6	5	0.718	0.837	7,000	5, 6
P.t.t.	17	13	1.200	1.489	12,000	5, 6
P.t.v.	22	25	2.400	2.697	23,000	5, 6
P. troglodytes	45	52	4.320	4.613	41,000	5, 6
P. paniscus	5	13	2.879	2.365	28,000	5, 7
H. sapiens	73	33	0.609	2.565	4,000	6, 8
mtDNA control region (partial 342 bp)						
P.t.s.	3	13	2.56	2.53		9
P.t.t.	9	45	5.59	4.84		9
P.t.v.	29	74	5.3	5.51		9
P. troglodytes	41	108	7.75	7.38		9
P. paniscus	17	53	3.87	4.58		9
H. sapiens (340 bp)	54	79	2.63	5.10		8

References: 1, present study; 2, DEINARD and KIDD (1999); 3, KAESSMANN *et al.* (1999); 4, KAESSMANN *et al.* (2001); 5, STONE *et al.* (2002); 6, WISE *et al.* (1997); 7, HORAI *et al.* (1992); 8, INGMAN *et al.* (2000); 9, DEINARD and KIDD (2000). Abbreviations of chimpanzee subspecies follow those of Table 1. *s*, segregating site.

^{a} Generation lengths are assumed to be 15 and 20 years for apes and humans, respectively.

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 (12,400) is only slightly larger and that for the Central African chimpanzees (20,900), or for the entire species of chimpanzees, is about twice as large.

Causes for different patterns of nuclear DNA and mtDNA diversity: As noted above, for nuclear DNA the nucleotide diversity in humans is only 50% lower than that in chimpanzees, whereas previous studies have found the mtDNA nucleotide diversity in humans to be at most only one-third of that found in chimpanzees

(WISE *et al.* 1997; STONE *et al.* 2002). What are the causes for this sharp contrast? A highly plausible cause is a reduction in the effective population size in the human lineage since the human-chimp divergence. There is strong evidence for this putative reduction (RUVOLO 1997; HARPENDING *et al.* 1998; CHEN and LI 2001). As noted by FAY and WU (1999), a reduction in N_e causes a larger decrease in nucleotide diversity for mtDNA than for nuclear DNA. This follows from the theory that the effect of a bottleneck on π is proportional to T/N_1, where T is the time since the bottleneck and N_1 is the new effective population size, and from the fact that the

TABLE 3

Average sequence divergence (%) between taxa estimated
from the 50 DNA segments studied

Species	P.t.s.	P.t.t.	P.t.v.	Chimp	Bonobo
P.t.t.	0.133				
P.t.v.	0.103	0.125			
Bonobo	0.353	0.370	0.376	0.373	
Human	1.194	1.214	1.233	1.221	1.30

Abbreviations of chimpanzee subspecies follow those of
Table 1.

effective population size for mtDNA is usually smaller than that for nuclear DNA (TAKAHATA 1993). An additional factor is that there has been an increase in generation time in the human lineage, which leads to a higher mutation rate per generation and also to a higher expected π value in the case of nuclear DNA but to little increase in the case of mtDNA because of maternal inheritance and limited germ cell divisions per generation in the female germline.

Another possibility is that population subdivision might have been stronger in humans than in chimpanzees or the female migration rate might have been lower in humans than in chimpanzees, so that the effective population size for mtDNA is considerably smaller in humans than in chimpanzees (BIRKY *et al.* 1989). Evidence to support this argument may come from the observation that chimpanzee females uniformly disperse from their natal troop, in contrast to the typical mammalian (and primate) pattern of female philopatry with male dispersal (GREENWOOD 1980). The potential for high rates of female migration in chimpanzees is seen in the long-distance (>900 km) sharing of mtDNA haplotypes (MORIN *et al.* 1994; GOLDBERG and RUVOLO 1997).

Divergences between subspecies and species: The average sequence divergence (*d*; average nucleotide differences per site) between chimpanzee subspecies is the smallest between the West and East African chimpanzees (Table 3), although in terms of geographic distance it should be the largest. The smaller divergence occurs because the levels of nucleotide diversity in the East and West African chimpanzees are smaller than that in the Central African chimpanzees. For the same reason, the *d* value is the largest between the Central and East African subspecies. The between-subspecies *d* values are only slightly higher than the within-subspecies π values, indicating a small separation between subspecies.

The sequence divergence between the bonobo and the chimpanzee (0.373%) is about three times higher than the nucleotide diversity within the chimpanzee species (0.132%), indicating a much longer separation time between the two species than between the chimpanzee subspecies. The sequence divergence between the human and the chimpanzee is 1.22% (Table 3), which is similar to the values ($\sim$1.2%) obtained on the basis of

>2 million bp (CHEN *et al.* 2001; EBERSBERGER *et al.* 2002; FUJIYAMA *et al.* 2002). The sequence divergence between the human and the bonobo (1.30%) is somewhat higher than that between the human and the chimpanzee, probably because the bonobo has a smaller effective population size than the chimpanzee and so has been subject to a stronger effect of random drift.

These data also provide the largest and most comprehensive estimate of divergence time between chimpanzees and bonobos, estimated here to be 1.8 MYA, assuming a 6-MYA Homo-Pan split (BRUNET *et al.* 2002; VIGNAUD *et al.* 2002). This estimate is the same as that by STONE *et al.* (2002), based on their data from the nonrecombing region of Y, and is intermediate between published data using mtDNA (2.5 MYA; GAGNEUX *et al.* 1999) and X chromosomal DNA (0.9 MYA; KAESSMANN *et al.* 1999). Also, the date of the formation of the Congo River, which currently prevents contact between these species, has been estimated from geological and limnological evidence to have formed $\sim$1.5 MYA (BEADLE 1981). Perhaps the formation of the Congo River initiated the separation between chimpanzees and bonobos. This was also a time of potentially large climatic changes in the African climate and changes in vegetation patterns, proposed by some to have catalyzed the origins of several hominid species and human innovations (VRBA 1995).

LITERATURE CITED

BEADLE, L. C., 1981 *The Inland Waters of Tropical Africa.* Longman, London.

BIRKY, C. W., JR., P. FUERST and T. MARUYAMA, 1989 Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics **121:** 613–627.

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.

CHEN, F. C., and W.-H. LI, 2001 Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am. J. Hum. Genet. **68:** 444–456.

CHEN, F. C., E. J. VALLENDER, H. WANG, C. S. TZENG and W.-H. LI, 2001 Genomic divergence between human and chimpanzee estimated from large-scale alignments of genomic sequences. J. Hered. **92:** 481–489.

CROUAU-ROY, B., S. SERVICE, M. SLATKIN and N. FREIMER, 1996 A fine-scale comparison of the human and chimpanzee genomes: linkage, linkage disequilibrium and sequence analysis. Hum. Mol. Genet. **5:** 1131–1137.

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. Physiol. Anthropol. **111:** 25–44.

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.

Ebersberger, I., D. Metzler, C. Schwarz and S. Paabo, 2002 Genomewide comparison of DNA sequences between humans and chimpanzees. Am. J. Hum. Genet. **70:** 1490–1497.

Ellegren, H., C. R. Primmer and B. C. Sheldon, 1995 Microsatellite 'evolution': Directionality or bias? Nat. Genet. **11:** 360–362.

Fay, J. C., and C.-I Wu, 1999 A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. Mol. Biol. Evol. **16:** 1003–1005.

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.

Fujiyama, A., H. Watanabe, A. Toyoda, T. D. Taylor, T. Itoh *et al.*, 2002 Construction and analysis of a human-chimpanzee comparative clone map. Science **295:** 131–134.

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.

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.

Greenwood, P. J., 1980 Mating systems, philopatry and dispersal in birds and mammals. Anim. Behav. **28:** 1140–1162.

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.

Horai, S., Y. Satta, K. Hayasaka, R. Kondo, T. Inoue *et al.*, 1992 Man's place in Hominoidea revealed by mitochondrial DNA genealogy. J. Mol. Evol. **35:** 32–43.

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.

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. Paabo, 2001 Great ape DNA sequences reveal a reduced diversity and an expansion in humans. Nat. Genet. **27** (2): 155–156.

Ingman, M., H. Kaessmann, S. Pääbo and U. Gyllensten, 2000 Mitochondrial genome variation and the origin of modern humans. Nature **408:** 708–713.

King, M. C., and A. C. Wilson, 1975 Evolution at two levels in humans and chimpanzees. Science **188:** 107–116.

Lucotte, G., 1983 Chimpanzees show less variation than man: electrophoretic evidence. J. Hum. Evol. **12:** 419–424.

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.

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., 1997 Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. Mol. Biol. Evol. **14:** 248–265.

Satta, Y., 2001 Comparison of DNA and protein polymorphisms between humans and chimpanzees. Genes Genet. Syst. **76:** 159–168.

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., 1993 Allelic genealogy and human evolution. Mol. Biol. Evol. **10:** 2–22.

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.

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.

Vrba, E. S., 1995 On the connection between paleoclimate and evolution, pp. 24–48 in *Paleoclimate and Evolution, With Emphasis on Human Origins*, edited by E. S. Vrba, G. H. Denton, T. C. Partridge and L. H. Burckle. Yale University Press, New Haven, CT.

Watterson, G. A., 1975 On the number of segregation sites. Theor. Popul. Biol. **7:** 256–276.

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.

Xia, X., and Z. Xie, 2001 DAMBE: Data analysis in molecular biology and evolution. J. Hered. **92:** 371–373.

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.

Zhao, Z., L. Jin, Y.-X. Fu, M. Ramsay, T. Jenkins *et al.*, 2000 Worldwide DNA sequence variation in a 10 kilobase noncoding region on chromosome 22. Proc. Natl. Acad. Sci. USA **97:** 11354–11358.

Communicating editor: N. Takahata