Comparative Nuclear and Mitochondrial Genome Diversity in Humans and Chimpanzees

Cheryl A. Wise, Michaela Sraml,*[1] David C. Rubinsztein,† and Simon Easteal**

*The John Curtin School of Medical Research, The Australian National University; and †Department of Medical Genetics, Addenbrooke's Hospital, Cambridge, U.K.

Restriction mapping and sequencing have shown that humans have substantially lower levels of mitochondrial genome diversity (d) than chimpanzees. In contrast, humans have substantially higher levels of heterozygosity (H) at protein-coding loci, suggesting a higher level of diversity in the nuclear genome. To investigate the discrepancy further, we sequenced a segment of the mitochondrial genome control region (CR) from 49 chimpanzees. The majority of these were from the *Pan troglodytes verus* subspecies, which was underrepresented in previous studies. We also estimated the average heterozygosity at 60 short tandem repeat (STR) loci in both species. For a total sample of 115 chimpanzees, $d = 0.075 \pm 0.037$, compared to 0.020 ± 0.011 for a sample of 1,554 humans. The heterozygosity of human STR loci is significantly higher than that of chimpanzees. Thus, the higher level of nuclear genome diversity relative to mitochondrial genome diversity in humans is not restricted to protein-coding loci. It seems that humans, not chimpanzees, have an unusual d/H ratio, since the ratio in chimpanzees is similar to that in other catarrhines. This discrepancy in the relative levels of nuclear and mitochondrial genome diversity in the two species cannot be explained by differences in mutation rate. However, it may result from a combination of factors such as a difference in the extent of sex ratio disparity, the greater effect of population subdivision on mitochondrial than on nuclear genome diversity, a difference in the relative levels of male and female migration among subpopulations, diversifying selection acting to increase variation in the nuclear genome, and/or directional selection acting to reduce variation in the mitochondrial genome.

Introduction

Comparisons of genetic diversity between closely related species can contribute importantly to a number of issues in population genetics and evolution. Such comparisons have played a role in understanding aspects of genome evolution, including the evolutionary dynamics of transposable elements (Dowsett and Young 1982) and the evolutionary patterns and dynamics of short tandem repeat (STR) loci (Rubinsztein et al. 1995). They have also been important sources of evidence for the action of natural selection. Balancing selection has been indicated by polymorphism maintained between species (Mayer et al. 1992), and discrepancies in levels of diversity within and between species have been interpreted as indicating various forms of natural selection (McDonald and Kreitman 1991; Ballard and Kreitman 1994; Nachman et al. 1996).

In this context, the increasing wealth of information about different kinds of genetic diversity in humans (*Homo sapiens*) provides a valuable basis for comparison with related species, particularly with chimpanzees (*Pan troglodytes* and *P. paniscus*). Understandably, the investigation of genetic diversity in chimpanzees has not been as comprehensive as in humans. Most intensively investigated have been electrophoretically detectable protein polymorphisms (King and Wilson 1975; Bruce and Ayala 1979) and mitochondrial genome variation (Ferris et al. 1981; Morin et al. 1994).

The results have suggested a discrepancy between the relative levels of mitochondrial and nuclear genome diversity in humans and chimpanzees (Wilson et al. 1985). Chimpanzees appear to have more mitochondrial genome diversity, whereas humans have more nuclear genome diversity.

The ratio of gene diversities in mitochondrial and nuclear genomes, d/H, is approximately $N_f\mu/2N_e\nu$, where N_e is the effective population size, N_f is the number of breeding females, μ is the mutation rate for the mitochondrial genome, and ν is the mutation rate of genes in the nuclear genome (Birky, Maruyama, and Fuerst 1983). This ratio assumes selective neutrality and depends on sex ratio, population structure, and mutation rates. Thus, if there is a discrepancy in the ratio between two species, it is of considerable interest, as it implies a difference in some aspect of their evolutionary dynamics.

Here we further examine levels of mitochondrial and nuclear genome diversity in humans and chimpanzees. We extend the number of described *P. troglodytes* mitochondrial control region (CR) sequences to include a better representation of the subspecies *P. t. verus*, and we determine levels of variation at STR loci. By investigating this type of nuclear genome variation, which is very different from the protein variation previously studied, we can determine whether the results indicated by the protein variation can be generalized to the nuclear genome.

Materials and Methods

Samples

The 49 chimpanzee blood samples used for mitochondrial CR sequencing and the samples used for eight of the STR loci (table 1) were obtained from animals held under long-term observation in one of several pri-

[1] Present address: Woden Valley Hospital, Canberra, Australia.

Key words: human, chimpanzee–*Pan troglodytes*, DNA polymorphism, mtDNA, STR polymorphism.

Address for correspondence and reprints: Cheryl Wise, Human Genetics Group, John Curtin School of Medical Research, The Australian National University, Canberra, ACT 0200, Australia. E-mail: cheryl.wise@anu.edu.au.

Mol. Biol. Evol. 14(7):707–716. 1997

Table 1
Tandem Repeat Loci Analyzed in this Study

Locus	Chromosomal Location	Primer (5′–3′)	Reference for Human
COL2A1	12q14.3	CCAGGTTAAGGTTGACAGCT GTCATGAACTAGCTCTGGTG	a
D7S460	7	AATACCCCAAGGGGTGGTAA CATTGATGAACAGTTCAAGCA	b
D8S342	8q1	CAGCCTGGGCAATAGAAAGAGAC CAGTGCTCCCTTCCTTGAAGTTTC	c
MYCN.....	2p24	GGAGGCTGAAAGCACAGTTG TGGGCAACAAGAGCAAAACT	d
RENA4	1q32	AGAGTACCTTCCCTCCTCTACTCA CTCTATGGAGCTGGTAGAACCTGA	e
TH04	11p15.5	CAGCTGCCCTAGTCAGCAC GCTTCCGAGTGCAGGTCACA	f
BG01	11p15.5	ATAGACTGGAGTAAAGGAA CTTCTACTCTGTGAATGGA	g
EG01	11p15.5	CAAAGTAGTGGGAAGCTGT ATGGATAGGATAAGTCCCC	g

NOTE.—References are as follows: a, Wu, Seino, and Bell (1990); b, Hudson et al. (1992); c, Lu et al. (1993); d, Fougerousse et al. (1992); e, Edwards et al. (1992); f, Polymeropoulos et al. (1991); and g, This study—previously uncharacterized short tandem repeat loci. BG01 is a dinucleotide repeat $(TG)_n$, which starts at position 1479 of the human and chimpanzee β-globin genomic sequence (Savatier et al. 1985). EG01 is a tetranucleotide repeat $(GTAT)_n$, which starts at position 2939 of the human and chimpanzee ψη-globin genomic sequence (Miyamoto, Slightom, and Goodman 1987).

mate colonies at the Laboratory of Slow, Latent and Temperate Virus Infection of the National Institute of Health (NIH). They were supplied by Drs. D. C. Gajdusek and C. J. Gibbs Jr. All individuals were drawn from a larger sample of 102 individuals, the majority of which were wild-caught (Board, Gibbs, and Gajdusek 1981). The exact sources of all these animals are not known. However, almost all appear to belong to the west African subspecies *P. t. verus* (see *Results and Discussion*). Chimpanzee samples used for analysis of an additional 27 STR loci were from animals wild-caught in Sierra Leone and thus also from the *P. t. verus* subspecies (Slierendregt et al. 1993; Rubinsztein et al. 1995). These animals appear, from their MHC haplotype heterogeneity, to be relatively outbred. The human blood samples were from a variety of sources.

Mitochondrial Control Region Sequencing

Polymerase chain reaction (PCR) amplification of the mitochondrial CR was performed in two stages. A 1,309-bp product encompassing the entire CR was generated using primers L15926 (5′-TACACCAGTCTTGTAAACC-3′) and H629 (5′-TGTTTATGGGGTGATGTGA-3′). This product was used as the template for subsequent amplification using the nested primers L15997 (Ward et al. 1991) and H16401 (Vigilant et al. 1989). In each PCR reaction, one primer was biotinylated and the other was incorporated with an M13 sequencing primer. For each individual, both H (heavy) and L (light) strands were sequenced on an ABI Model 373A Automated DNA Sequencer.

The 49 chimpanzee mitochondrial CR sequences reported here have been deposited in the DDBJ/EMBL/GenBank International Nucleotide Sequence Database under accession numbers U84293–U84341.

STR Genotyping

The procedures used to analyze 27 STR loci in chimpanzees were as described previously (Rubinsztein et al. 1995). Eight additional loci were analyzed (table 1), including five tetranucleotide repeats and one 31–34-bp repeat originally characterized in humans, and one tetra- and one dinucleotide repeat that had not previously been characterized in either species.

Data Analysis

Consensus sequences for each individual were obtained by aligning forward- and reverse-complement sequences in the SeqEd™ 675 DNA Sequence Editor (ABI) program. Chimpanzee sequences from published sources were also used in the comparative analyses. Three sequences were obtained from Kocher and Wilson (1991), and 63 were obtained from the DDBJ/EMBL/GenBank International Nucleotide Sequence Database (published in Morin et al. 1994) through the Australian National Genomic Information Service (ANGIS). Sequence alignment was performed manually using Genetic Data Environment (GDE) 2.2 (Smith et al. 1994). The alignment can be obtained from: http://jcsmr.anu.edu.au/dmm/humgen.html.

The numbers of variable sites in the CR sequences and the numbers of differences between sequence pairs were counted. Between-sequence distances were computed using Kimura's (1980) two-parameter method, with a transition/transversion ratio of nine estimated directly from the sequence data. Distances were also estimated using Tamura and Nei's (1993) method, assuming a Γ distribution with $\alpha = 0.5$ (Wakeley 1993), as implemented in MEGA 1.0 (Kumar, Tamura, and Nei 1993). Alignment gaps and sites with missing information were ignored. The mean of the pairwise distances (d) was estimated, and the standard deviation of this

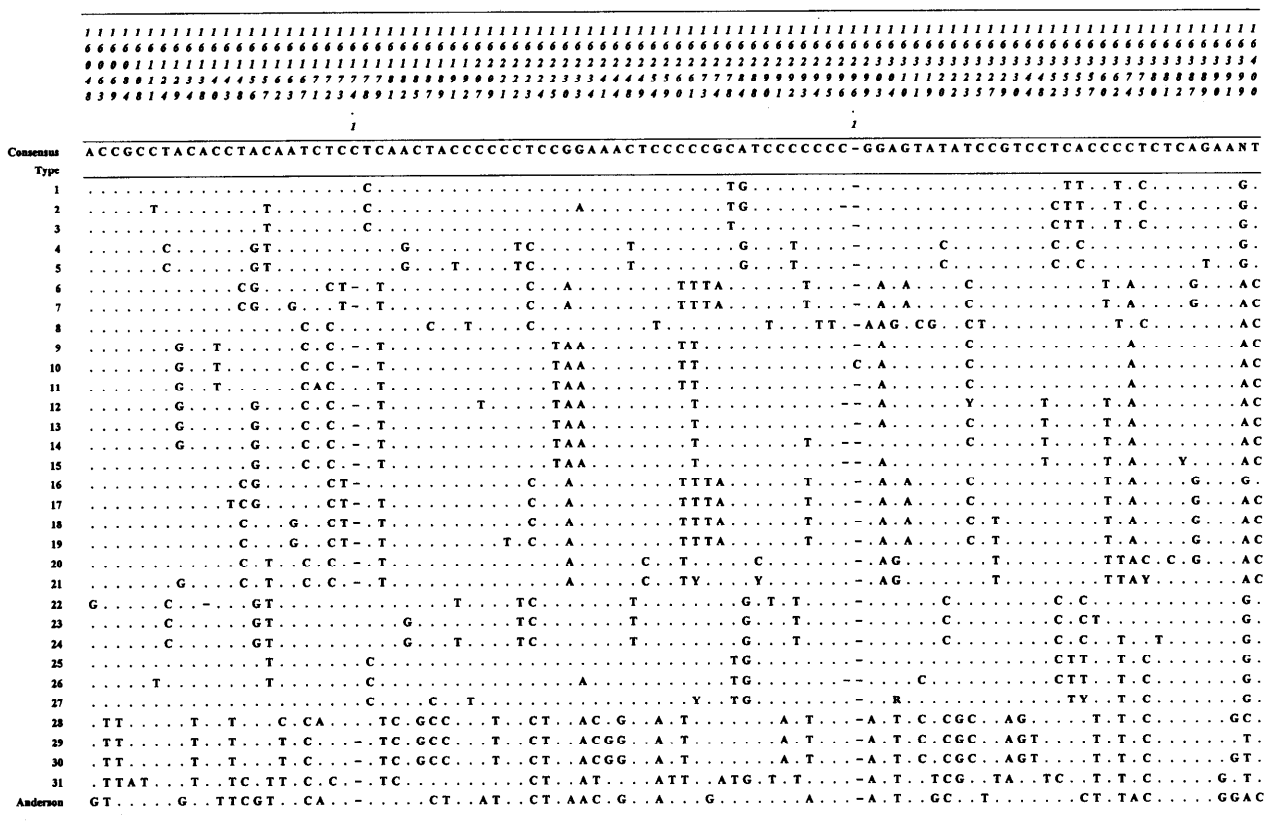

FIG. 1.—Nucleotide sequence differences in a 377-bp segment of the control region for 52 chimpanzee individuals. A total of 94 variable sites were observed, shown as differences from a chimpanzee consensus sequence. Nomenclature is in accordance with Anderson et al. (1981), and numbers followed by a decimal point indicate additional nucleotides not found in this sequence. Dots indicate sequence identity, dashes denote indels, and the IUB single-letter codes for nucleotide bases have been used for ambiguous bases. The number of individuals for each sequence type (*n*) is shown at the right of each sequence. Each sequence type corresponds to the following individuals in figure 2. Type 1—A-26, A-69, A-179, A-242, A-283, and A-292; Type 2—A-50, A-52, and A-199; Type 3—A-268 and A-290; Type 4–A-94, A-128, and A-130; Type 5–A-208 and A-285; Type 6—A-90 and A-139; Type 7—A-137 and A-286; Type 8—A-172, A-192, A-197, and A-281; Type 9—A-42, A-89, A-118, A-136, A-230, and A-291; Type 10—A-131; Type 11—C2; Type 12—A-101; Type 13—A-129; Type 14—A-235; Type 15—A-282; Type 16—A-33; Type 17—A-97; Type 18—A-56; Type 19—A-288; Type 20—A-62; Type 21—A-182; Type 22—C3; Type 23—A-102; Type 24—A-231; Type 25—A-108; Type 26—A-239; Type 27—A-284; Type 28—C1; Type 29—A-60; Type 30—A-175; Type 31—A-176.

estimate was calculated using equation (30) of Tajima (1983). The frequency distributions of the distances were displayed as histograms (mismatch distributions), as proposed by Rogers and Harpending (1992).

The phylogeny of the *P. troglodytes* CR sequences was estimated using the neighbor-joining (NJ) method (Saitou and Nei 1987) as implemented in PHYLIP 3.5c (Felsenstein 1993), with a *P. paniscus* sequence (Foran, Hixson, and Brown 1988) as the outgroup. The programs SEQBOOT, DNADIST, NEIGHBOR, and CONSENSE were used with 1,000 replicates, based on both Kimura two-parameter and Tamura-Nei Γ distances, with a randomized input order.

The average heterozygosity (*H*) across all STR loci was estimated, and the standard error of *H* was obtained, as described by Nei and Roychoudhury (1974*a*).

Results and Discussion

Mitochondrial Genome Sequence Diversity

We determined the nucleotide sequence of a 377-bp segment of the noncoding CR (positions 16024 to 16400 in the reference sequence of Anderson et al. [1981]) for 49 chimpanzee individuals. We compared these sequences with three sequences reported elsewhere (Kocher and Wilson 1991). The sites that vary among these sequences are shown in figure 1. There were 31 distinctive sequence types defined by 94 variable sites, including four indels (fig. 1). Nine sequence types were shared among two to six individuals, while the remaining 22 sequence types were observed only once. At two sites (16078 and 16166), all chimpanzee individuals lacked the A residue present in the human reference sequence (Anderson et al. 1981). Most differences between the sequences (89.7%) result from transition-type mutations. Moreover, transitions between pyrimidines (72.6%) occurred much more frequently than those between purines, probably reflecting the low G content in the L-strand of the mitochondrial genome. In this respect, the variation is similar to that of other published chimpanzee sequences (Morin et al. 1994) and to that of humans (Vigilant et al. 1989; Horai and Hayasaka 1990; Horai et al. 1993).

A 345-bp segment of the noncoding CR has previously been sequenced for 63 chimpanzee individuals (Morin et al. 1994), the majority of which were of known geographic origin. With the addition of the pres-

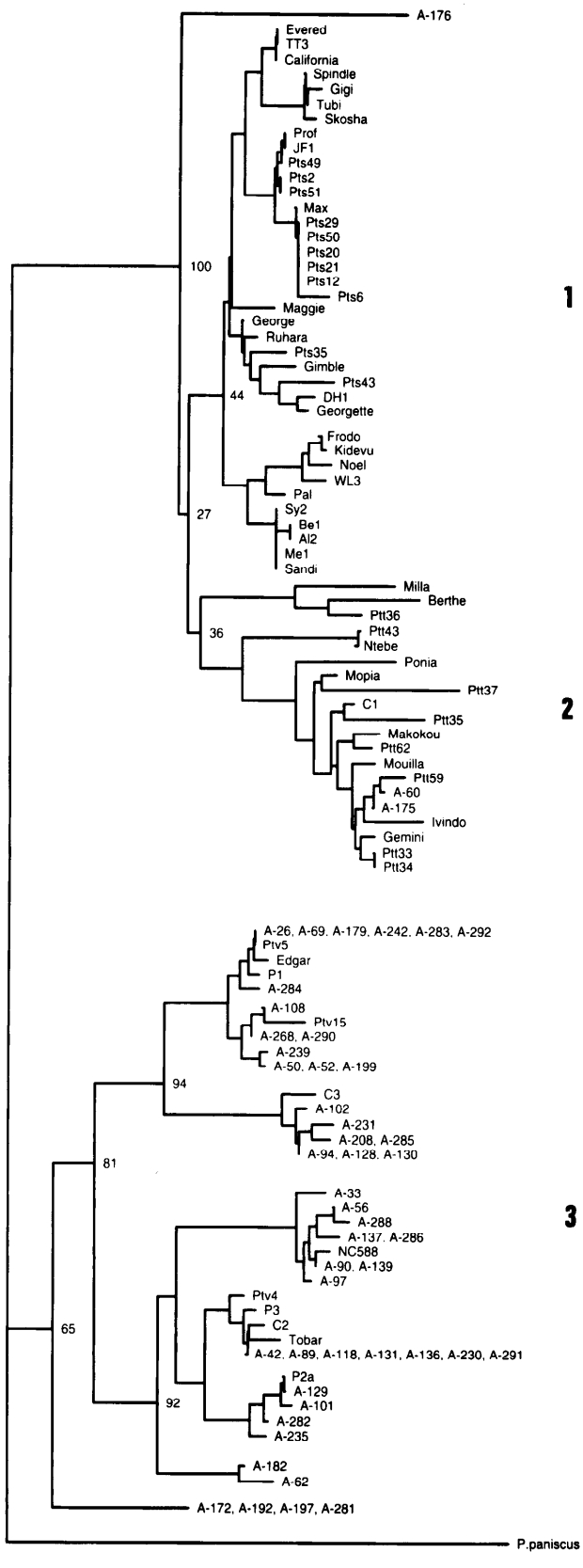

FIG. 2.—Phylogenetic tree relating the sequences of this study and other published chimpanzee sequences using *P. paniscus* as the outgroup. Samples from this study have the prefix "A"; samples taken from Kocher and Wilson (1991) are C1 to C3; and samples taken from Morin et al. (1994) are named according to that study. The phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei 1987) based on Kimura (1980) two-parameter distances between control region sequences. The reliability of each interior branch was tested by 1,000 bootstrap replications, and reliabilities are shown as percentage values next to the major branches. Based on previous characterization of control region sequences (Morin et al. 1994), the three distinct clades (1 to 3) correspond respectively to the three subspecies *P. t. schweinfurthii, P. t. troglodytes,* and *P. t. verus.* The majority of the samples in the present study appear to be from the west African subspecies *P. t. verus,* as they fall within clade 3.

ent 52 sequences, the CR sequences from 115 chimpanzees were aligned and compared, and the number of nucleotide differences per site between each pair of sequences was estimated. The mean nucleotide distance (d) between all pairs of the 115 sequences was estimated as 0.075, 0.083, and 0.103 using a simple proportional distance, Kimura's (1980) two-parameter distance, and Tamura and Nei's (1993) Γ distance methods, respectively. The respective maximum distances are 0.139, 0.163, and 0.227.

Morin et al. (1994) found that chimpanzee CR sequences cluster into three distinct subspecies clades. The bootstrapped NJ tree presented in figure 2 implies that the majority of the sequences in the present study are from the west African subspecies *P. troglodytes verus,* as they cluster with the sequences from this subspecies reported by Morin et al. (1994). Two sequences appear to be derived from the central African subspecies *P. t. troglodytes* (A-60 and A-175 in fig. 2), and five sequences cannot be ascribed to any of the three subspecies (A-172, A-192, A-197, A-281, and A-176 in fig. 2). However, analysis of the data using a transversion network approach (H.-J. Bandelt, personal communication) indicates that A-172, A-192, A-197, and A-281 are from the west African subspecies *P. t. verus,* and A-176 is from the east African subspecies *P. t. schweinfurthii.*

The pairwise nucleotide differences between sequences in each of the three subspecies (fig. 3) are significantly different from the Poisson distributions predicted for the observed mean values using a chi-square test ($P < 0.001$). The distributions for *P. t. verus* and *P. t. troglodytes* are clearly multimodal (fig. 3*A* and *B*), a pattern consistent with stable populations (Slatkin and Hudson 1991; Rogers and Harpending 1992). The *P. t. schweinfurthii* distribution is approximately unimodal (fig. 3*C*), which Rogers and Jorde (1995) have interpreted as indicating that this subspecies has experienced a recent population expansion. The uncorrected mean (0.022 ± 0.012) and maximum (0.043) pairwise differences in *P. t. schweinfurthii* are lower than in *P. t. verus* and *P. t. troglodytes* (means: 0.051 ± 0.025 and 0.040 ± 0.021, respectively; maxima: 0.093 and 0.091, respectively).

The mean and maximum pairwise differences in a diverse sample of 1,554 humans are 0.020 ± 0.011 and 0.067, respectively. In a sample of 389 Africans, who contain the most divergent of the human mtDNA lineages (Cann, Stoneking, and Wilson 1987; Vigilant et al. 1989, 1991; Horai et al. 1993), they are 0.025 ± 0.013 and 0.062, respectively (Watson 1996, pp. 90–91).

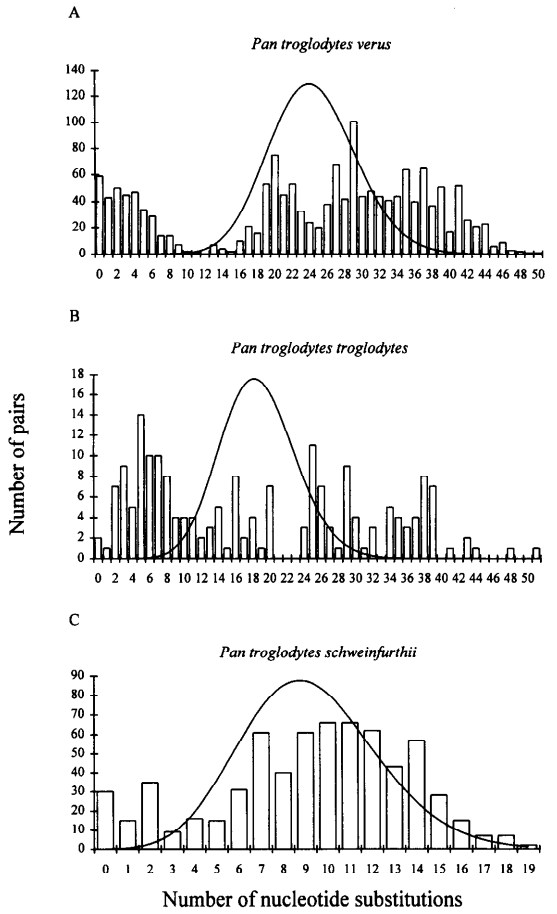

FIG. 3.—Frequency distributions of nucleotide distances (Tamura and Nei 1993) between all pairs of individuals in three chimpanzee subspecies. *A*, Fifty-seven west African chimpanzees (*P. t. verus*). *B*, Twenty central African chimpanzees (*P. t. troglodytes*). *C*, Thirty-seven east African chimpanzees (*P. t. schweinfurthii*). The expected Poisson distributions having the same means as the observed distributions are indicated by solid lines.

Thus, the nucleotide distances within chimpanzee subspecies are the same as or greater than the distances within the entire human species, and the overall distances in chimpanzees are substantially greater than in humans. This is consistent with the greater levels of variation observed in earlier restriction mapping studies of the mitochondrial genomes of various ape species (Ferris et al. 1981). Rogers and Jorde (1995) have also suggested that chimpanzee mitochondrial diversity exceeds that of humans by nearly an order of magnitude.

Nuclear Genome Diversity

We compared levels of heterozygosity in humans and chimpanzees at 35 STR loci analyzed in the present study and 33 loci for which variation in the two species had been described elsewhere (table 2). The heterozygosity of human STR loci is significantly higher than that of chimpanzees using the nonparametric Wilcoxon's signed-ranks test ($z = -2.99$, $P < 0.005$; Seigel 1956, pp. 75–83).

All but two of the loci analyzed were already known to be polymorphic in humans. The possibility of ascertainment bias resulting in the relative overestimation of heterozygosity in humans must therefore be considered. Although bias in heterozygosity estimates cannot be completely ruled out, there are two reasons why it is unlikely. First, the mean and the range of heterozygosities for the analyzed loci are similar in humans to those found in a sample of randomly selected loci (Hudson et al. 1992). Thus, they do not appear to be biased in favor of loci with high heterozygosity. Second, ascertainment bias will only result in a pronounced estimate bias when heterozygosity in the source species is less than approximately 0.4; the bias is negligible when heterozygosity exceeds approximately 0.5 (Rogers and Jorde 1996). Thus, the tandem repeat loci we analyzed that are polymorphic in both humans and chimpanzees should be essentially free of the effect of ascertainment bias because of their high heterozygosities in humans. Bias may still occur, however, if loci are selected that are highly polymorphic in humans but essentially monomorphic in chimpanzees because of base substitutions that interfere with the generation of repeat-length polymorphism (Crouau-Roy et al. 1996). To avoid such bias, we excluded from analysis the eight loci that were highly polymorphic in humans but monomorphic, or nearly monomorphic, in chimpanzees.

The higher level of diversity at STR loci in humans is consistent with data from other kinds of loci. Chimpanzees also have fewer polymorphic red cell enzyme and serum protein loci, and only 0.1–0.2 of the average heterozygosity (H) found at these loci in humans (King and Wilson 1975; Bruce and Ayala 1979). There is a possibility of ascertainment bias in these comparisons since loci may have been selected for analysis because they are polymorphic in humans. However, Harris and Hopkinson (1972) obtained an estimate of $H = 0.072$ in Europeans for 26 enzyme-coding loci chosen without any prior knowledge of polymorphism. This was very similar to the estimate of 0.065 obtained for 45 other enzyme-coding loci from data compiled from the literature, indicating that, at least for enzyme-coding loci, the heterozygosity estimates in humans are not biased. In later studies (Nei and Roychoudhury 1974*b*, 1982), H was estimated as 0.10–0.14, the difference being attributed to the inclusion of non-enzyme-coding loci in the latter studies, which are generally more polymorphic than enzyme-coding loci. The H values obtained for chimpanzees include data on non-enzyme-coding loci; thus, the human values that are appropriate for comparison with chimpanzees are 0.10–0.14, reported by Nei and Roychoudhury (1974*b*, 1982). However, the more conservative estimate of 0.067 (Harris and Hopkinson 1972) is also substantially greater than the range of estimates (0.01–0.02) for chimpanzees (King and Wilson 1975; Bruce and Ayala 1979).

In addition, of five minisatellite loci compared between humans and chimpanzees (Ely et al. 1992), three were polymorphic in both species, with average heterozygosities of 0.72 and 0.61 in humans and chimpanzees, respectively. There are also indications of lower levels of variation in chimpanzees at other minisatellite loci (Wolff et al. 1991) and at Major Histocompatibility

Table 2
Numbers of Alleles and Heterozygosity at 68 Short Tandem Repeat Loci in Humans and Chimpanzees

	Human			Chimpanzee		
Locus	*n*	No. of Alleles	Hetero-zygosity	*n*	No. of Alleles	Hetero-zygosity
D13S71[a]	138	5	0.747	186	3	0.187
D13S118[a]	154	8	0.727	168	7	0.738
D13S121[a]	156	8	0.772	156	14	0.883
D13S122[a]	154	12	0.838	178	5	0.628
D13S124[a]	156	6	0.673	178	8	0.824
D13S193[a]	152	10	0.745	182	13	0.723
D13S197[a]	154	22	0.880[b]	108	2	0.018[b]
FES[c]	48	6	0.766	90	4	0.506
MBP1[c]	72	6	0.801	90	7	0.526
MBP2[c]	72	7	0.811	90	7	0.657
SE33[c]	78	21	0.942	90	3	0.364
TH01[c]	1,280	8	0.790	90	3	0.293
VWF[c]	200	7	0.734	90	8	0.667
Mfd 3[d]	218	10	0.748	29	10	0.885
Mfd 32[d]	220	11	0.708	32	6	0.825
Mfd 38[d]	220	14	0.844	32	16	0.928
Mfd 59[d]	225	13	0.879	32	14	0.896
Mfd 75[d]	219	15	0.883	32	7	0.719
Mfd 104[d]	182	15	0.837	32	15	0.928
Mfd 139[d]	213	20	0.797	32	6	0.677
Mfd 142[d]	204	10	0.758	32	4	0.708
D4S174[e]	100	11	0.826	56	17	0.974
D4S190[e]	100	8	0.804[b]	55	6	0.266[b]
D4S230[e]	100	12	0.824	56	12	0.886
D4S391[e]	100	13	0.927	56	13	0.975
D4S404[e]	100	5	0.701[b]	56	3	0.107[b]
D4S405[e]	100	9	0.866	56	8	0.797
D4S418[e]	100	9	0.824	56	7	0.886
D4S419[e]	100	6	0.727	54	9	0.798
D4S425[e]	100	7	0.715	56	7	0.926
D4S551[e]	100	12	0.712	56	11	0.891
D4S616[e]	100	10	0.828	56	12	0.752
D4S885[e]	100	8	0.748[b]	56	2	0.132[b]
D2S119[f]	14	8	1.000	16	7	1.000
D2S123[f]	16	7	0.933	16	3	0.133
D2S391[f]	16	4	0.533	16	1	0.000
D3S1029[f]	16	6	0.933	16	5	0.267
D3S1038[f]	16	6	0.800	16	3	0.267
D3S1076[f]	16	5	0.667	16	6	0.667
D3S1298[f]	16	9	0.800	16	5	0.133
D3S1561[f]	16	8	0.933	16	4	0.267
D9S66[f]	16	9	1.000	8	3	0.286
D9S104[f]	16	6	0.400	16	3	0.800
D9S122[f]	16	6	0.933[b]	16	1	0.000[b]
D9S150[f]	16	6	0.800[b]	16	1	0.000[b]
D9S298[f]	16	4	0.533	16	3	0.933
DXS3[f]	8	4	1.000	16	4	0.533
DXS207[f]	8	4	1.000	16	7	0.800
DXS228[f]	8	6	1.000[b]	16	2	0.000[b]
DXS453[f]	8	5	0.857	16	6	0.667
DXS1110[f]	8	5	0.571	16	7	0.800
DYS II[f]	8	5	0.571	16	3	0.800
CYBB[f]	8	7	0.857	16	4	1.000
DBH[f]	14	3	0.769	16	3	0.400
H1[f]	8	3	0.571	16	6	0.667
L128[f]	6	5	1.000	16	5	0.533
MAOB[f]	8	7	1.000	16	3	0.400
PFC[f]	8	9	1.000	14	5	1.000
STR44[f]	8	8	1.000	16	5	0.800
X75b[f]	16	6	0.933	16	3	0.400
COL2A1[f]	32	5	0.836	100	5	0.505
D7S460[f]	48	?	0.970[b]	78	1	0.000[b]
D8S342[f]	66	10	0.619	96	5	0.566
MYCN[f]	118	3	0.424	102	3	0.613
RENA4[f]	1,250	6	0.399	104	2	0.308

Table 2
Continued

Locus	Human			Chimpanzee		
	n	No. of Alleles	Heterozygosity	*n*	No. of Alleles	Heterozygosity
TH04[f]	70	5	0.791	102	2	0.357
BG01[g]	98	7	0.717	104	5	0.444
EG01[g]	102	2	0.040	102	1	0.000
Mean		8.033	0.776		6.383	0.630
Standard error		0.498	0.023		0.501	0.035

NOTE.—*n*, number of chromosomes examined. If two related individuals shared one allele at a locus, one copy was excluded, causing allele frequencies (and thereby the expected heterozygosity) to be based on an odd number of chromosomes (J. C. Garza, personal communication). The heterozygosity of each locus is weighted according to sample size (Nei and Roychoudhury 1974*a*).

[a] Data from Deka et al. (1994).

[b] Data excluded from analysis because of negligible variation in chimpanzees despite high heterozygosity in humans.

[c] Data from Pascall et al. (1994) and references therein.

[d] Data from Garza, Slatkin, and Freimer (1995).

[e] Data from Crouau-Roy et al. (1996).

[f] Only chimpanzees analyzed in the present study. Data sources for humans are as in table 1 and Rubinsztein et al. (1995).

[g] Both humans and chimpanzees analyzed in the present study.

Complex (*Mhc*) loci (Kenter et al. 1992; Ayala and Escalante 1996).

Discrepancies of Mitochondrial/Nuclear Genome Diversity Ratios

The consistency of STR and other kinds of loci in showing a higher level of nuclear genome diversity in humans than in chimpanzees implies that the nuclear genome as a whole is more variable in humans. In contrast, the mitochondrial genome is less variable in humans than in chimpanzees. However, our finding of a discrepancy between the levels of mitochondrial and nuclear genome diversity in humans and chimpanzees is based on samples that were collected rather differently in the two species. The possibility that the result is a sampling artifact needs to be considered.

The microsatellite diversity was estimated largely from samples of European humans and west African chimpanzees, although in both species individuals from other geographical areas were analyzed for some loci. European humans do not appear to have higher levels of microsatellite diversity than humans from other regions (Jorde et al. 1995; Richards et al. 1996). It is possible that west African chimpanzees have lower levels of nuclear genome diversity than chimpanzees from other regions (although they do not have lower levels of mitochondrial genome diversity).

We have discussed above how the mitochondrial diversity is lower in humans regardless of whether comparisons are made between samples collected from one geographical region in both species (e.g., west Africa in chimpanzees and Africa in humans), or across the geographical range of both species. Thus, this difference does not appear to be a sampling artifact.

The discrepancy could be a consequence of one or more factors that differ between humans and chimpanzees. To investigate this further, it would be useful to know if either one of the species has an unusual *d/H* ratio. We compared the levels of nuclear and mitochondrial genome diversity in humans and chimpanzees with those in six other catarrhine taxa (table 3). Nuclear genome diversity was estimated as H for protein-coding loci. Mitochondrial genome diversity was estimated as either d_1, the mean pairwise distance between CR sequences, or, in the cases where CR sequence data were not available, d_2, the mean pairwise distances of the entire mitochondrial genome based on restriction maps. The human d_1/H and d_2/H ratios are both substantially less than those of all other species. In contrast, the chimpanzee ratios are similar to those of the other nonhuman species. The discrepancy thus appears to be a result of an unusually low *d/H* ratio in humans rather than an unusually high ratio in chimpanzees.

Our estimate of the difference in *d/H* ratio between humans and chimpanzees is conservative. European humans have unusually low levels of mitochondrial genome diversity ($d = 0.013 \pm 0.007$; Watson 1996, p. 90), so by restricting the nuclear genome comparison to a largely European human sample but extending the mitochondrial genome comparison to include individuals from other regions, we may have underestimated the extent of the difference in *d/H* ratios between the species. If the comparison were restricted largely to European humans ($d_1/H = 0.19$) and west African chimpanzees ($d_1/H = 2.43$), the discrepancy in the ratios between these representative groups of the two species would be 12.8-fold, compared to 11.9-fold (table 3).

It should be noted that some previous studies (Hammer 1995; Nei 1995) have concluded that in humans d and H give consistent estimates of effective population size, implying that the *d/H* ratio in humans is not unusual. These conclusions, however, are based on mutation rate estimates that may not be accurate.

A low human *d/H* ratio could result from an unusual mutation rate, either an unusually high nuclear genome mutation rate or an unusually low mitochondrial genome mutation rate. Mutation rate differences, how-

Table 3
Mitochondrial and Nuclear Genome Diversity in Catarrhines

Species	d_1	d_2	H	d_1/H	d_2/H
Homo sapiens	0.020[a] ± 0.011	0.0032[d]	0.067[h] ± 0.02	0.30	0.05
Pan troglodytes	0.075[b] ± 0.037	0.0133[e]	0.021[i] ± 0.01	3.57	0.63
Gorilla gorilla	0.099[c] ± 0.050	—	0.049[j] ± 0.03	2.02	—
Pan paniscus	—	0.0100[e]	0.022[j] ± 0.02	—	0.45
Pongo pygmaeus abelii	—	0.0210[e]	0.048[j] ± 0.03	—	0.44
Pongo pygmaeus pygmaeus	—	0.0050[e]	0.025[j] ± 0.03	—	0.20
Macaca fuscata	—	0.0132[f]	0.013[k] ± 0.001	—	1.02
Macaca fascicularis	—	0.0410[g]	0.096[l]	—	0.43

NOTE.—d_1, mean pairwise distances based on nucleotide sequences from an approximately 300- to 400-bp segment of the mitochondrial control region. Distances were computed using an uncorrected proportional distance, and their standard deviations were calculated using equation (30) of Tajima (1983). The standard deviation estimates include both sampling and stochastic variance. d_2, mean pairwise distances based on restriction maps of the entire mitochondrial genome. H, average heterozygosity for protein-coding loci weighted according to sample size (Nei and Roychoudhury 1974*a*). H is estimated in all species except humans using data on both enzyme-coding and non-enzyme-coding loci. The human estimate is based only on the less variable enzyme-coding loci since this estimate appears not to be biased by inclusion of a disproportionate number of polymorphic loci (Harris and Hopkinson 1972). The estimate is conservative, and inclusion of data for non-enzyme-coding loci gives human H estimates in the range 0.10 to 0.14 (Nei and Roychoudhury 1974*b*, 1982).

[a] 1,554 worldwide humans (Watson 1996, p. 91).

[b] Data combined from the present study and Morin et al. (1994).

[c] Garner and Ryder (1996).

[d] Cann, Stoneking, and Wilson (1987).

[e] Ferris et al. (1981).

[f] Hayasaka et al. (1986).

[g] Harihara et al. (1988).

[h] Harris and Hopkinson (1972).

[i] King and Wilson (1975). Bruce and Ayala (1979) obtained a chimpanzee H estimate of 0.01. This value gives a d_1/H ratio of 7.50.

[j] Bruce and Ayala (1979).

[k] Nozawa et al. (1982).

[l] Nei and Graur (1984).

ever, would result in substitution rate differences in the human lineage compared with the chimpanzee lineage, and no such differences are apparent (Sibley and Ahlquist 1987; Horai et al. 1992; Easteal, Collet, and Betty 1995, pp. 49–61). It has been suggested that microsatellite mutation rates are higher in humans than in chimpanzees (Rubinsztein, Leggo, and Amos 1995), possibly related to differences in allele length (Rubinsztein et al. 1995). However, differences in both allele length and mutation rate have been questioned by Ellegren, Primmer, and Sheldon (1995), who suggest that the apparent allele length difference may be due to ascertainment bias. This issue has yet to be resolved.

Evidence for a mutation rate difference is based on discrepancies of the estimated and expected genetic distances between humans and chimpanzees and between two human groups (Europeans and sub-Saharan Africans). No estimate was made of divergence within chimpanzees. The validity of the suggested mutation rate difference depends on many factors, including the extent of divergence within chimpanzees, the relationship between genetic distance and separation time over a period of several million years, the stochastic error in the estimate of genetic distance, the assumed divergence time of humans and chimpanzees, and the assumed divergence time of sub-Saharan Africans and Europeans. Although the possibility of differences in mutation rates cannot be excluded, further investigation is needed before it is clearly demonstrated.

Other factors that could explain the low human d/H ratio includc: (1) A relatively higher ratio of males to females in humans—d depends only on the effective number of females, while H depends on the effective number of both males and females (Birky, Maruyama, and Fuerst 1983). (2) Less population subdivision in humans–population subdivision tends to increase diversity, and the effect on the mitochondrial genome is greater than on the nuclear genome (Birky, Fuerst, and Maruyama 1989). (3) A relatively higher rate of female than of male migration among subpopulations in humans (i.e., relatively less female subdivision and relatively more male subdivision). This will tend to reduce d relative to H (Birky, Fuerst, and Maruyama 1989). (4) Diversifying selection acting to increase variation in the human nuclear genome. (5) Directional selection acting to reduce variation in the human mitochondrial genome.

The significance of the results presented here is that they establish the generality of the discrepancy, and they contribute further to its quantification. This provides the basis for a detailed investigation of the possible ways in which these factors might have interacted to produce the discrepancy. This will have important implications with respect to the dynamics of human evolution.

Acknowledgments

We are grateful to D. C. Gajdusek and C. J. Gibbs Jr. for supplying the chimpanzee blood samples, to L. Croft, J. Leggo, B. Matheson, and P. Munroe for technical assistance, and to D. Betty, C. Groves, A. Thorne, and two anonymous reviewers for comments on a previous version of this manuscript. This work was funded

by Australian Research Council grant # A59332440 to S.E.

LITERATURE CITED

ANDERSON, S., A. T. BANKIER, B. G. BARRELL et al. (14 co-authors). 1981. Sequence and organisation of the human mitochondrial genome. Nature **290**:457–465.

AYALA, F. J., and A. A. ESCALANTE. 1996. The evolution of human populations: a molecular perspective. Mol. Phylogenet. Evol. **5**:188–201.

BALLARD, J. W. O., and M. KREITMAN. 1994. Unraveling selection in the mitochondrial genome of Drosophila. Genetics **138**:757–772.

BIRKY, C. W., P. A. 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.

BIRKY, C. W., T. MARUYAMA, and P. A. FUERST. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics **103**:513–527.

BOARD, P. G., C. J. GIBBS, and D. C. GAJDUSEK. 1981. Polymorphism of erythrocyte glyoxalase II in anthropoid primates. Folia Primatol. **36**:138–143.

BRUCE, E. J., and F. J. AYALA. 1979. Phylogenetic relationships between man and the apes: electrophoretic evidence. Evolution **33**:1040–1056.

CANN, R. L., M. STONEKING, and A. C. WILSON. 1987. Mitochondrial DNA and human evolution. Nature **325**:31–36.

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.

DEKA, R., M. D. SHRIVER, L. M. YU, L. JIN, C. E. ASTON, R. CHAKRABORTY, and R. E. FERRELL. 1994. Conservation of human chromosome 13 polymorphic microsatellite $(CA)_n$ repeats in chimpanzees. Genomics **22**:226–230.

DOWSETT, A. P., and M. W. YOUNG. 1982. Differing levels of dispersed repetitive DNA among closely related species of *Drosophila.* Proc. Natl. Acad. Sci. USA **79**:4570–4574.

EASTEAL, S., C. COLLET, and D. BETTY. 1995. The mammalian molecular clock. Springer, Austin, Tex.

EDWARDS, A., H. A. HAMMOND, L. JIN, C. T. CASKEY, and R. CHAKRABORTY. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics **12**:241–253.

ELLEGREN, H., C. R. PRIMMER, and B. C. SHELDON. 1995. Microsatellite 'evolution': directionality or bias? Nat. Genet. **11**:360–362.

ELY, J., R. DEKA, R. CHAKRABORTY, and R. E. FERRELL. 1992. Comparison of five tandem repeat loci between humans and chimpanzees. Genomics **14**:692–698.

FELSENSTEIN, J. 1993. PHYLIP: phylogeny inference package. Version 3.5c. University of Washington, Seattle.

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.

FORAN, D. R., J. E. HIXSON, and W. M. BROWN. 1988. Comparisons of ape and human sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. Nucleic Acids Res. **16**:5841–5861.

FOUGEROUSSE, F., R. MELONI, C. ROUDAUT, and J. S. BECKMANN. 1992. Tetranucleotide repeat polymorphism at the human N-MYC gene (MYCN). Nucleic Acids Res. **20**:1165.

GARNER K. J., and O. A. RYDER. 1996. Mitochondrial DNA diversity in gorillas. Mol. Phylogenet. Evol. **6**:39–48.

GARZA, J. C., M. SLATKIN, and N. B. FREIMER. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Mol. Biol. Evol. **12**:594-603.

HAMMER, M. F. 1995. A recent common ancestry for human Y chromosomes. Nature **378**:376–378.

HARIHARA, S., N. SAITOU, M. HIRAI, N. AOTO, K. TERAO, F. CHO, S. HONJO, and K. OMOTO. 1988. Differentiation of mitochondrial DNA types in *Macaca fascicularis.* Primates **29**:117–127.

HARRIS, H., and D. A. HOPKINSON. 1972. Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphisms. Ann. Hum. Genet. Lond. **36**:9–20.

HAYASAKA, K., S. HORAI, T. SHOTAKE, K. NOZAWA, and E. MATSUNAGA. 1986. Mitochondrial DNA polymorphism in Japanese monkeys, *Macaca fuscata.* Jpn. J. Genet. **61**:345–359.

HORAI, S., and K. HAYASAKA. 1990. Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. Am. J. Hum. Genet. **46**:828–842.

HORAI, S., R. KONDO, Y. NAKAGAWA-HATTORI, S. HAYASHI, S. SONODA, and K. TAJIMA. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol. Biol. Evol. **10**:23–47.

HORAI, S., Y. SATTA, K. HAYASAKA, R. KONDO, T. INOUE, T. ISHIDA, S. HAYASHI, and N. TAKAHATA. 1992. Man's place in Hominoidea revealed by mitochondrial DNA genealogy. J. Mol. Evol. **35**:32–43.

HUDSON, T. J., M. ENGELSTEIN, M. K. LEE, E. C. HO, M. J. RUBENFIELD, C. P. ADAMS, D. E. HOUSMAN, and N. C. DRACOPOLI. 1992. Isolation and chromosomal assignment of 100 highly informative human simple sequence repeat polymorphisms. Genomics **13**:622–629.

JORDE, L. B., M. J. BAMSHAD, W. S. WATKINS, R. ZENGER, A. E. FRALEY, P. A. KRAKOWIAK, K. D. CARPENTER, H. SOODYALL, T. JENKINS, and A. R. ROGERS. 1995. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am. J. Hum. Genet. **57**:523–538.

KENTER, M., N. OTTING, J. ANHOLTS, M. JONKER, R. SCHIPPER, and R. E. BONTROP. 1992. *Mhc*-DRB diversity of the chimpanzee (*Pan troglodytes*). Immunogenetics **37**:1–11.

KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

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

KOCHER, T. D., and A. C. WILSON. 1991. Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein-coding region. Pp. 391–413 *in* S. OSAWA and T. HONJO, eds. Evolution of life: fossils, molecules and culture. Springer, Tokyo.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Version 1.0. The Pennsylvania State University, University Park.

LU, J., R. RILEY, M. ROBERTSON, L. NELSON, and K. WARD. 1993. Tetranucleotide repeat polymorphisms at the D8S342, D8S323, D8S345, D8S315 and D8S347 loci on 8q. Hum. Mol. Genet. **2**:1743.

MAYER, W. E., C. O'HUIGIN, Z. ZALESKA-RUTCZYNSKA, and J. KLEIN. 1992. Trans-species origin of *Mhc-DRB* polymorphism in the chimpanzee. Immunogenetics **37**:12–23.

McDONALD, J. H., and M. KREITMAN. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila.* Nature **351**:652–654.

MIYAMOTO, M. M., J. L. SLIGHTOM, and M. GOODMAN. 1987. Phylogenetic relations of humans and African apes from DNA sequences in the ψη-globin region. Science **238**:369–373.

MORIN, P. A., J. J. MOORE, R. CHAKRABORTY, L. JIN, J. GOODALL, and D. S. WOODRUFF. 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science **265**:1193–1201.

NACHMAN, M. W., W. M. BROWN, M. STONEKING, and C. F. AQUADRO. 1996. Nonneutral mitochondrial DNA variation in humans and chimpanzees. Genetics **142**:953–963.

NEI, M. 1995. Genetic support for the out-of-Africa theory of human evolution. Proc. Natl. Acad. Sci. USA **92**:6720–6722.

NEI, M., and D. GRAUR. 1984. Extent of protein polymorphism and the neutral mutation theory. Evol. Biol. **17**:73–118.

NEI, M., and A. K. ROYCHOUDHURY. 1974*a.* Sampling variances of heterozygosity and genetic distance. Genetics **76**: 379–390.

———. 1974*b.* Genic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids. Am. J. Hum. Genet. **26**:421–443.

———. 1982. Genetic relationship and evolution of human races. Evol. Biol. **14**:1–59.

NOZAWA, K., T. SHOTAKE, Y. KAWAMOTO, and Y. TANABE. 1982. Population genetics of Japanese monkeys: II. Blood protein polymorphisms and population structure. Primates **23**:252–271.

PASCALL, V. L., G. DESTRO BISOL, M. DOBOSZ, E. D'ALOJA, G. PAONESSA, and U. MEREU. 1994. Chimpanzee DNA profiles on trial. Nature **367**:692–693.

POLYMEROPOULOS, M. H., H. XIAO, D. S. RATH, and C. R. MERRIL. 1991. Tetranucleotide repeat polymorphism at the human tyrosine hydroxylase gene (TH). Nucleic Acids Res. **19**:3753.

RICHARDS, R. I., J. CRAWFORD, K. NARAHARA et al. (15 co-authors). 1996. Dynamic mutation loci: allele distributions in different populations. Ann. Hum. Genet. **60**:391–400.

ROGERS, A. R., and H. HARPENDING. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. **9**:552–569.

ROGERS, A. R., and L. B. JORDE. 1995. Genetic evidence on modern human origins. Hum. Biol. **67**:1–36.

———. 1996. Ascertainment bias in estimates of average heterozygosity. Am. J. Hum. Genet. **58**:1033–1041.

RUBINSZTEIN, D. C., W. AMOS, J. LEGGO, S. GOODBURN, S. JAIN, S.-H. LI, R. L. MARGOLIS, C. A. ROSS, and M. A. FERGUSON-SMITH. 1995. Microsatellite evolution—evidence for directionality and variation in rate between species. Nat. Genet. **10**:337–343.

RUBINSZTEIN, D. C., J. LEGGO, and W. AMOS. 1995. Microsatellites evolve more rapidly in humans than in chimpanzees. Genomics **30**:610–612.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SAVATIER, P., G. TRABUCHET, C. FAURE, Y. CHEBLOUNE, M. GOUY, G. VERDIER, and V. M. NIGON. 1985. Evolution of the primate β-globin gene region. High rate of variation in CpG dinucleotides and in short repeated sequences between man and chimpanzee. J. Mol. Biol. **182**:21–29.

SEIGEL, S. 1956. Nonparametric statistics for the behavioural sciences. McGraw-Hill, Tokyo.

SIBLEY, C. G., and J. E. AHLQUIST. 1987. DNA hybridization evidence of hominoid phylogeny: results of an expanded data set. J. Mol. Evol. **26**:99–121.

SLATKIN, M., and R. R. HUDSON. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics **129**:555–562.

SLIERENDREGT, B. L., M. KENTER, N. OTTING, J. ANHOLTS, M. JONKER, and R. E. BONTROP. 1993. Major histocompatibility complex class II haplotypes in a breeding colony of chimpanzees (*Pan troglodytes*). Tissue Antigens **42**:55–61.

SMITH, S. W., R. OVERBEEK, C. R. WOESE, W. GILBERT, and P. M. GILLEVET. 1994. The genetic data environment and expandable GUI for multiple sequence analysis. Comput. Appl. Biosci. **10**:671–675.

TAJIMA, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics **105**:437–460.

TAMURA, K., and M. NEI. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10**:512–526.

VIGILANT, L., R. PENNINGTON, H. HARPENDING, T. D. KOCHER, and A. C. WILSON. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proc. Natl. Acad. Sci. USA **86**:9350–9354.

VIGILANT, L., M. STONEKING, H. HARPENDING, K. HAWKES, and A. C. WILSON. 1991. African populations and the evolution of human mitochondrial DNA. Science **253**:1503–1507.

WAKELEY, J. 1993. Substitution rate variation among sites in hypervariable region I of human mitochondrial DNA. J. Mol. Evol. **37**:613–623.

WARD, R. H., B. L. FRAZIER, K. DEW-JAGER, and S. PÄÄBO. 1991. Extensive mitochondrial diversity within a single American tribe. Proc. Natl. Acad. Sci. USA **88**:8720–8724.

WATSON, E. E. 1996. Threads from the past: a genetic study of African ethnic groups and human origins. Thesis, Massey University.

WILSON, A. C., R. L. CANN, S. M. CARR et al. (11 co-authors). 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. **26**:375–400.

WOLFF, R., Y. NAKAMURA, S. ODELBERG, R. SHIANG, and R. WHITE. 1991. Generation of variability at VNTR loci in human DNA. Pp. 20–38 *in* T. BURKE, G. DOLF, A. J. JEFFREYS, and R. WOLFF, eds. DNA fingerprinting: approaches and applications. Birkhauser Verlag, Basel.

WU, S., S. SEINO, and G. I. BELL. 1990. Human collagen, type II, alpha 1, (COL2A1) gene: VNTR polymorphism detected by gene amplification. Nucleic Acids Res. **18**:3102.

CRAIG MORITZ, reviewing editor

Accepted March 19, 1997