Mitochondrial COII Sequences and Modern Human Origins[1]

Maryellen Ruvolo, Sarah Zehr,* Miranda von Dornum,* Deborah Pan,* Belinda Chang,† and Jenny Lin**

*Department of Anthropology, Harvard University; and †Program in Neuroscience, Harvard Medical School

The aim of this study is to measure human mitochondrial sequence variability in the relatively slowly evolving mitochondrial gene cytochrome oxidase subunit II (COII) and to estimate when the the human common ancestral mitochondrial type existed. New COII gene sequences were determined for five humans (*Homo sapiens*), including some of the most mitochondrially divergent humans known; for two pygmy chimpanzees (*Pan paniscus*); and for a common chimpanzee (*P. troglodytes*). COII sequences were analyzed with those from another relatively slowly evolving mitochondrial region (ND4-5). From class 1 (third codon position) sequence data, a relative divergence date for the human mitochondrial ancestor is estimated as 1/27th of the human-chimpanzee divergence time. If it is assumed that humans and chimpanzees diverged 6 Mya, this places a human mitochondrial ancestor at 222,000 years, significantly different from 1 Myr (the presumed time of an *H. erectus* emergence from Africa). The mean coalescent time estimated from all 1,580 sites of combined mitochondrial data, when a 6-Mya human-chimpanzee divergence is assumed, is 298,000 years, with 95% confidence interval of 129,000–536,000 years. Neither estimate is compatible with a 1-Myr-old human mitochondrial ancestor. The mitochondrial DNA sequence data from COII and ND4-5 regions therefore do not support this multiregional hypothesis for the emergence of modern humans.

Introduction

The "mitochondrial Eve" hypothesis (Cann et al. 1987) is a statement about both tree topology and time: the common ancestor of all existing human mitochondrial DNA (mtDNA) types originated in Africa 140,000–290,000 years ago. In some ways, the statement about time is the more controversial. If the original claim had posited the same tree topology (in which the basic division on the tree of all human mtDNA sequences is into an African clade and a clade of all other humans including some Africans) but a more ancient origin (say, 1 Myr), it might not have been controversial, since the data could have been interpreted to reflect the initial migration of *Homo erectus* out of Africa, and therefore consistent with the multiregional hypothesis (Wolpoff 1989).

Claims about time are based on interpretations of amounts of DNA sequence differences. The first studies of human mitochondrial diversity relied on indirect measures of DNA sequence difference by using restriction-enzyme site analysis (Brown

1. Key words: human mtDNA variation, cytochrome oxidase subunit II (COII) gene, multiregional hypothesis, hominoid evolution, molecular evolution.

Address for correspondence and reprints: Maryellen Ruvolo, Department of Anthropology, Peabody Museum, Harvard University, 11 Divinity Avenue, Cambridge, Massachusetts 02138.

Mol. Biol. Evol. 10(6):1115–1135. 1993.

0737-4038/93/1006-0001$02.00

1980; Cann et al. 1987). This method has the advantage that it samples the entire mitochondrial genome. In contrast, Vigilant et al. (1989, 1991) directly obtained and compared DNA sequences of a portion of the mitochondrial genome. However, the available mtDNA sequence data, which are direct reflections of genomic diversity and which potentially offer greater resolution than does restriction mapping, do not unambiguously support one topology (Hedges et al. 1992; Maddison et al. 1992; Templeton 1992). Rather, there are (at least) three equally parsimonious classes of trees compatible with the data (Maddison et al. 1992).

This ambiguity is caused by the high rate of molecular evolutionary change demonstrated by the mitochondrial region examined (the control region) and by the presence of few phylogenetically informative characters relative to the number of individuals. On the basis of observed mtDNA sequence differences between pairs of individuals, the hypervariable control subregions evolve ~10 times faster than does the mitochondrial protein-coding gene for cytochrome oxidase subunit II (COII) (K. Garner and O. Ryder, unpublished data; M. Ruvolo, unpublished data). Thus, more slowly evolving protein-coding regions show fewer differences, compared with the control region among humans, and therefore offer potentially fewer phylogenetically informative sites. However, while the slower rate of the protein-coding genes means that relatively few differences are observed between humans and chimpanzees, the chance for the region to become "saturated" with multiple substitutions is reduced, making it more likely that phylogenetic information is preserved. Correction for multiple substitutions is of course still necessary, but, generally, small values of observed sequence difference get corrected very little, if at all, by all correction methods. For greater amounts of observed sequence difference, however, not only is the degree of correction greater, but correction methods vary more in their estimates of actual genetic difference. Therefore, the less quickly evolving portions of the mitochondrial genome showing little difference among humans should potentially provide us with more accurate comparative estimates of sequence divergence among mitochondrial haplotypes than does the control region. Control-region sequences are useful fine-grained indicators of differences among humans (di Rienzo and Wilson 1991; Ward et al. 1991), but, for more distant phylogenetic comparisons, the more slowly evolving regions are preferable.

Here we report results for a slowly evolving mitochondrial protein-coding gene, COII. We have included both the South African !Kung individual found to be most different from other humans on some mitochondrial trees (Cann et al. 1987; Vigilant et al. 1989; Maddison et al. 1992) and some African pygmies from central Africa who were found to be the most divergent individuals on other trees (Vigilant et al. 1991; Maddison et al. 1992). Throughout we compare the COII results with those from another slowly evolving mitochondrial region (the 896-bp segment including partial genes for NADH dehydrogenase subunits 4 and 5, or the ND4-5 region; Kocher and Wilson 1991); this region has been surveyed in some of the same individuals but not in any central African pygmies.

Material and Methods

The new COII sequences reported here are from five humans (*Homo sapiens*); two pygmy chimpanzees, also known as bonobos (*Pan paniscus*); and one common chimpanzee (*P. troglodytes*). For these new sequences, genomic DNA was prepared from hairbulbs (Vigilant et al. 1989) for the Asian Hsa 2 sample and from cultured cells (Maniatis et al. 1989, p. 6.53) for the South African !Kung individual Hsa 6 (cell

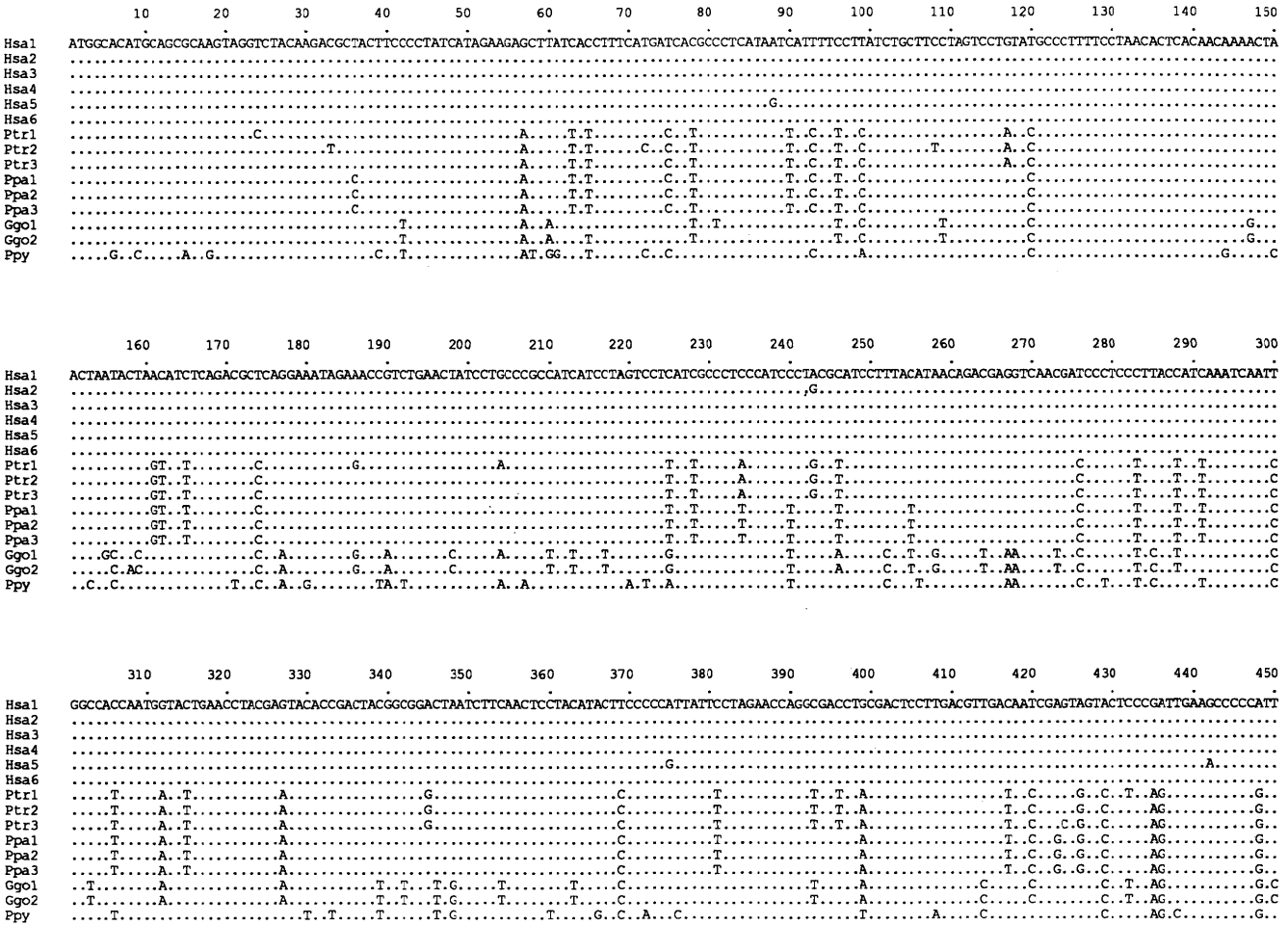
line GM 3043; Human Genetic Mutant Cell Repository, Camden, N.J.). Other genomic DNAs were provided by Dr. L. L. Cavalli-Sforza (from cell lines for humans Hsa 3–5), Dr. R. Honeycutt (from placental tissue for common chimpanzee Ptr 1), and Dr. O. Ryder of the San Diego Zoo (pygmy chimpanzees Ppa 1 and Ppa 3). Total genomic DNA was amplified by the polymerase chain reaction using oligonucleotide primers specific for the COII gene, to create double-stranded and then single-stranded DNA; single-stranded DNA was directly sequenced as described elsewhere (Ruvolo et al. 1991; Disotell et al. 1992). Both DNA strands were sequenced in every case.

Results and Discussion

Mitochondrial COII Gene Sequence Variation

The COII sequences generated are presented in figure 1, together with those previously published hominoid (human and ape) sequences (Anderson et al. 1981; Ruvolo et al. 1991; Horai et al. 1992) used in the analysis. [One COII sequence (Ptr 3), which we previously reported as that of a pygmy chimpanzee (*P. paniscus*) (Ruvolo et al. 1991) is most likely that of a common chimpanzee (*P. troglodytes*). This DNA sequence was generated by R. L. Honeycutt, from a DNA fragment containing the COII gene cloned by W. Brown, and the exact individual from which the DNA was obtained is unknown. When we discovered that the sequence clusters phylogenetically with those of common chimpanzees and not with pygmy chimpanzees (using sequences reported here as well as other unpublished *Pan* sequences), we reexamined available original laboratory notes in which clone "PC-2" was described, in one notebook, as being from a "chimpanzee" and, in other notes relating to DNA sequencing, as being from "common chimpanzee." The clone designation "PC-2" may have been interpreted as an abbreviation for "pygmy chimpanzee" rather than as an abbreviation for the more probable alternative, i.e., "plasmid clone."] Table 1 summarizes the individuals analyzed. Among humans, there are seven variable positions in COII sequences: six transitions (pyrimidine-pyrmidine or purine-purine substitutions) at positions 88, 243, 375, 442, 567, and 666 and one transversion (purine-pyrimidine substitution) at position 528 in !Kung individual Hsa 6 (fig. 2, *top*). Two substitutions occur at first codon positions (88 and 442), causing amino acid replacements in individual Hsa 5; the other five substitutions occur at third codon positions. The mean pairwise difference between humans is 0.34% (2.3 bp), similar to that for the ND4-5 region (0.2%) but less than that for the more quickly evolving control region (1.8%) (Kocher and Wilson 1991).

Humans and chimpanzees differ by an observed average of 9.4% in COII sequence (64 bp of 684 bp, with 61 transitions and 3 transversions), similar to the 9% average for the ND4-5 region (78 bp of 896 bp, with 73 transitions and 5 transversions). The control-region difference is 12% (Kocher and Wilson 1991). Among humans, the control region is 5–9 times more variable than COII and ND4-5 regions but is only 1.3 times more variable between humans and chimpanzees, a good indication that many multiple substitutions have occurred in the control region since the species diverged (Kocher and Wilson 1991). COII and ND4-5 regions are also similar in average transition:transversion (i:v) ratios calculated interspecifically between humans and chimpanzees (20:1 and 15:1, respectively). Hypervariable control-region sequences have a lower interspecific ratio (3:1), again presumably because of multiple transitional substitutions (Kocher and Wilson 1991). In numbers of substitutional differences, the COII and ND4-5 data are similar, and both are different from the control-region data.


10 20 30 40 50 60 70 80 90 100 110 120 130 140 150
Hsa1 ATGGCACATGCAGCGCAAGTAGGTCTACAAGACGCTACTTCCCCTATCATAGAAGAGCTTATCACCTTTCATGATCACGCCCTCATAATCATTTTCCTTATCTGCTTCCTAGTCCTGTATGCCCTTTTCCTAACACTCACAACAAAACTA
Hsa2
Hsa3
Hsa4
Hsa5
Hsa6
Ptr1
Ptr2
Ptr3
Ppa1
Ppa2
Ppa3
Ggo1
Ggo2
Ppy
160 170 180 190 200 210 220 230 240 250 260 270 280 290 300
Hsa1 ACTAATACTAACATCTCAGACGCTCAGGAAATAGAAACCGTCTGAACTATCCTGCCCGCCATCATCCTAGTCCTCATCGCCCTCCCATCCCTACGCATCCTTTACATAACAGACGAGGTCAACGATCCCTCCCTTACCATCAAATCAATT
Hsa2
Hsa3
Hsa4
Hsa5
Hsa6
Ptr1
Ptr2
Ptr3
Ppa1
Ppa2
Ppa3
Ggo1
Ggo2
Ppy
310 320 330 340 350 360 370 380 390 400 410 420 430 440 450
Hsa1 GGCCACCAATGGTACTGAACCTACGAGTACACCGACTACGGCGGACTAATCTTCAACTCCTACATACTTCCCCCATTATTCCTAGAACCAGGCGACCTGCGACTCCTTGACGTTGACAATCGAGTAGTACTCCCGATTGAAGCCCCCATT
Hsa2
Hsa3
Hsa4
Hsa5
Hsa6
Ptr1
Ptr2
Ptr3
Ppa1
Ppa2
Ppa3
Ggo1
Ggo2
Ppy

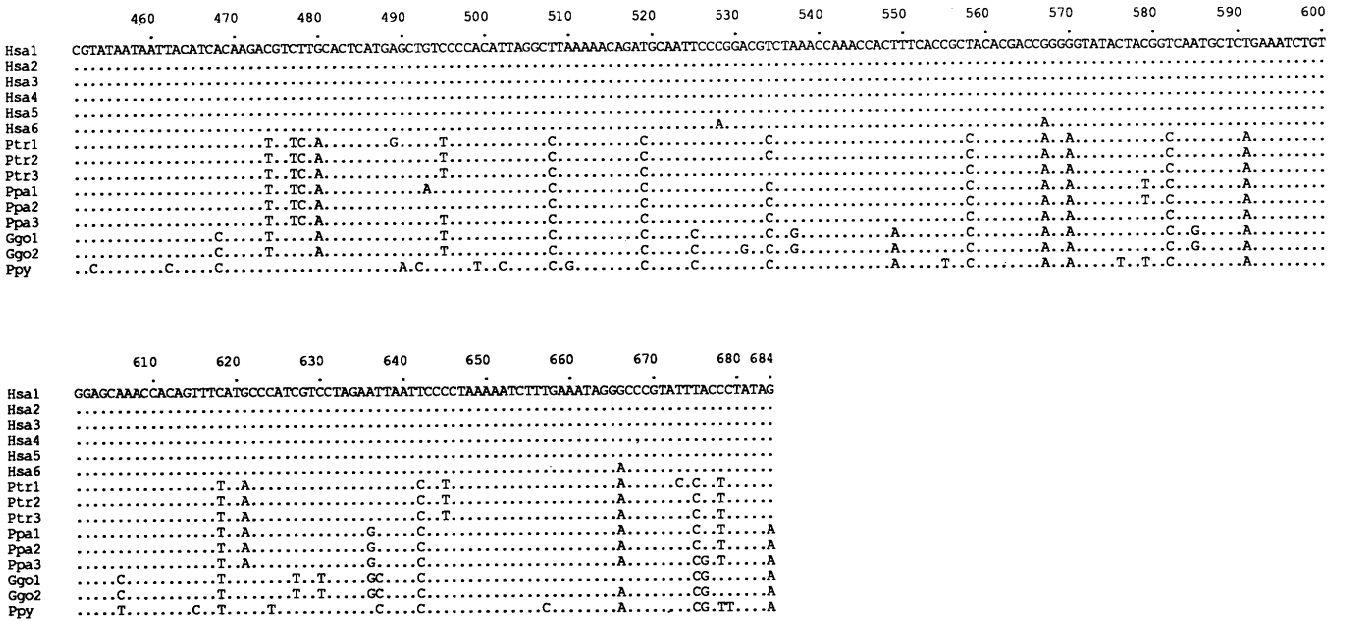


FIG. 1.—DNA sequences of the mitochondrial COII gene from 15 hominoids. Position 1 corresponds to site 7586 in the Anderson et al. (1981) human reference sequence with the COII gene ending at site 8269. Abbreviations for species names and samples are as in table 1.

Table 1
Individuals Studied, and Their Use in Previous Studies

Identification No. in Present Study	Geographic Origin or Species	Previous Study[a] (identification no.[b])	Other Identifier
Humans:			
Hsa 1	Presumably northern European	1, 2 (110), 4 (H1), 6 (118), 7 (ns)	Human or Cambridge reference sequence
Hsa 2	Asian (Taiwan)		
Hsa 3	Zaire (Mbuti pygmy)	6 (5), 7 (ns)	Cell line P45G
Hsa 4	Central African Republic (Biaka pygmy)	6 (2), 7 (ns)	Cell line P116
Hsa 5	Central African Republic (Biaka pygmy)	6 (37), 7 (ns)	Cell line P31
Hsa 6	South African !Kung	2 (1), 4 (H2), 6 (13), 7 (15)	Cell line GM 3043; Wilson lab SA1
Chimpanzees:			
Ptr 1	*Pan troglodytes*		Sally's infant
Ptr 2	*P. troglodytes*	3	
Ptr 3	*P. troglodytes*	5	Clone PC2 (see text)
Ppa 1	*P. paniscus*		Vernon (San Diego Zoo); ISIS 180343
Ppa 2	*P. paniscus*	3	
Ppa 3	*P. paniscus*		Marilyn (San Diego Zoo); ISIS 587376
Gorillas:			
Ggo 1 ...	*Gorilla gorilla*	5	
Ggo 2 ...	*G. gorilla*	3	
Orangutan:			
Ppy	*Pongo pygmaeus*	3	

[a] 1 = Anderson et al. (1981); 2 = Cann et al. (1987); 3 = Horai et al. (1992); 4 = Kocher and Wilson (1991); 5 = Ruvolo et al. (1991); 6 = Vigilant et al. (1991) (for identification nos., see Vigilant 1990); and 7 = Vigilant et al. (1989).
[b] ns = identification no. (order of appearance on tree) was not specified.

Phylogenetic Results

The most parsimonious tree (fig. 3) shows conspecific sequences clustering together despite intraspecific variability, as well as a human-chimpanzee clade as found elsewhere with single individual species representatives (Ruvolo et al. 1991). The same tree topology is found with phenetic methods, by using neighbor-joining (Saitou and Nei 1987) and Fitch and Margoliash (1967) methods. The single most parsimonious tree has length 205, with consistency index 0.849, which means that variable sites change once on average. Human sequences exclusive of the South African !Kung sequence are linked by a minimum of two unambiguous synapomorphies (at positions 666 and 567) at an 85% bootstrap level. This contrasts with a consistency index of 0.34 for the 119 phylogenetically informative sites (each varying three times on average) on one tree of human hypervariable control-region sequences (Vigilant et al. 1991).

The within-human genetic difference can be calculated in a tree-based fashion, as follows: The Hsa 6 (!Kung) COII sequence is cladistically most different from the others, so the average difference between it and other human sequences is taken to represent the maximum difference among humans. This average difference through the root of the human clade is 4 bp (0.58%), consisting of three transitions and one

Hsa 1		0	0	0	0	1	3	3	4	2	2	2	8	9	15
Hsa 2	1		0	0	0	1	3	3	4	2	2	2	8	9	15
Hsa 3	0	1		0	0	1	3	3	4	2	2	2	8	9	15
Hsa 4	0	1	0		0	1	3	3	4	2	2	2	8	9	15
Hsa 5	3	4	3	3		1	3	3	4	2	2	2	8	9	15
Hsa 6	2	3	2	2	5		4	4	5	3	3	3	9	10	16
Ptr 1	65	64	65	65	68	63		0	1	1	1	1	7	8	14
Ptr 2	62	61	62	62	65	60	9		1	1	1	1	7	8	14
Ptr 3	58	57	58	58	61	56	7	4		2	2	2	8	9	15
Ppa 1	61	62	61	61	64	59	21	18	16		0	0	6	7	13
Ppa 2	60	61	60	60	63	58	20	17	15	1		0	6	7	13
Ppa 3	61	62	61	61	64	59	19	16	14	4	3		6	7	13
Ggo 1	75	76	75	75	78	75	68	71	69	66	65	62		1	9
Ggo 2	75	76	75	75	78	73	66	67	65	62	61	58	6		10
Ppy	81	82	81	81	84	79	87	84	84	79	78	79	81	79	

Hsa 1		1, 0	0, 0	0, 0	1, 0	2, 1	57, 3	55, 3	51, 3	53, 2	53, 2	53, 2	63, 8	63, 8	64, 13
Hsa 2	1		1, 0	1, 0	2, 0	3, 1	56, 3	54, 3	50, 3	54, 2	54, 2	54, 2	64, 8	64, 8	65, 13
Hsa 3	0	1		0, 0	1, 0	2, 1	57, 3	55, 3	51, 3	53, 2	53, 2	53, 2	63, 8	63, 8	64, 13
Hsa 4	0	1	0		1, 0	2, 1	57, 3	55, 3	51, 3	53, 2	53, 2	53, 2	63, 8	63, 8	64, 13
Hsa 5	3	4	3	3		3, 1	58, 3	56, 3	52, 3	54, 2	54, 2	54, 2	64, 8	64, 8	65, 13
Hsa 6	3	4	3	3	6		55, 4	53, 4	49, 4	51, 3	51, 3	51, 3	63, 9	61, 9	62, 14
Ptr 1	68	67	68	68	71	67		8, 0	6, 0	19, 1	19, 1	17, 1	56, 7	56, 7	72, 12
Ptr 2	65	64	65	65	68	64	9		4, 0	17, 1	17, 1	15, 1	60, 7	58, 7	70, 12
Ptr 3	62	61	62	62	65	61	8	5		15, 1	15, 1	13, 1	58, 7	56, 7	70, 12
Ppa 1	63	64	63	63	66	62	22	19	18		0, 0	2, 0	54, 6	52, 6	64, 11
Ppa 2	62	63	62	62	65	61	21	18	17	1		2, 0	54, 6	52, 6	64, 11
Ppa 3	63	64	63	63	66	62	20	17	16	4	3		52, 6	50, 6	66, 11
Ggo 1	83	84	83	83	86	84	75	78	77	72	71	68		4, 0	70, 7
Ggo 2	84	85	84	84	87	83	74	75	74	69	68	65	7		70, 7
Ppy	96	97	96	96	99	95	101	98	99	92	91	92	90	89	

FIG. 2.—Sequence differences of the COII gene of 15 hominoids. *Top,* Observed (pairwise) number of nucleotide transitional differences for the 684-bp COII gene sequences (below the diagonal) and transversional differences (above the diagonal). *Bottom,* Total observed nucleotide differences for the COII gene sequences (below the diagonal) and class 1 (third codon positions) transitional and transversional differences of 228 bp total in the COII gene sequences (above the diagonal). Abbreviations are as in table 1.

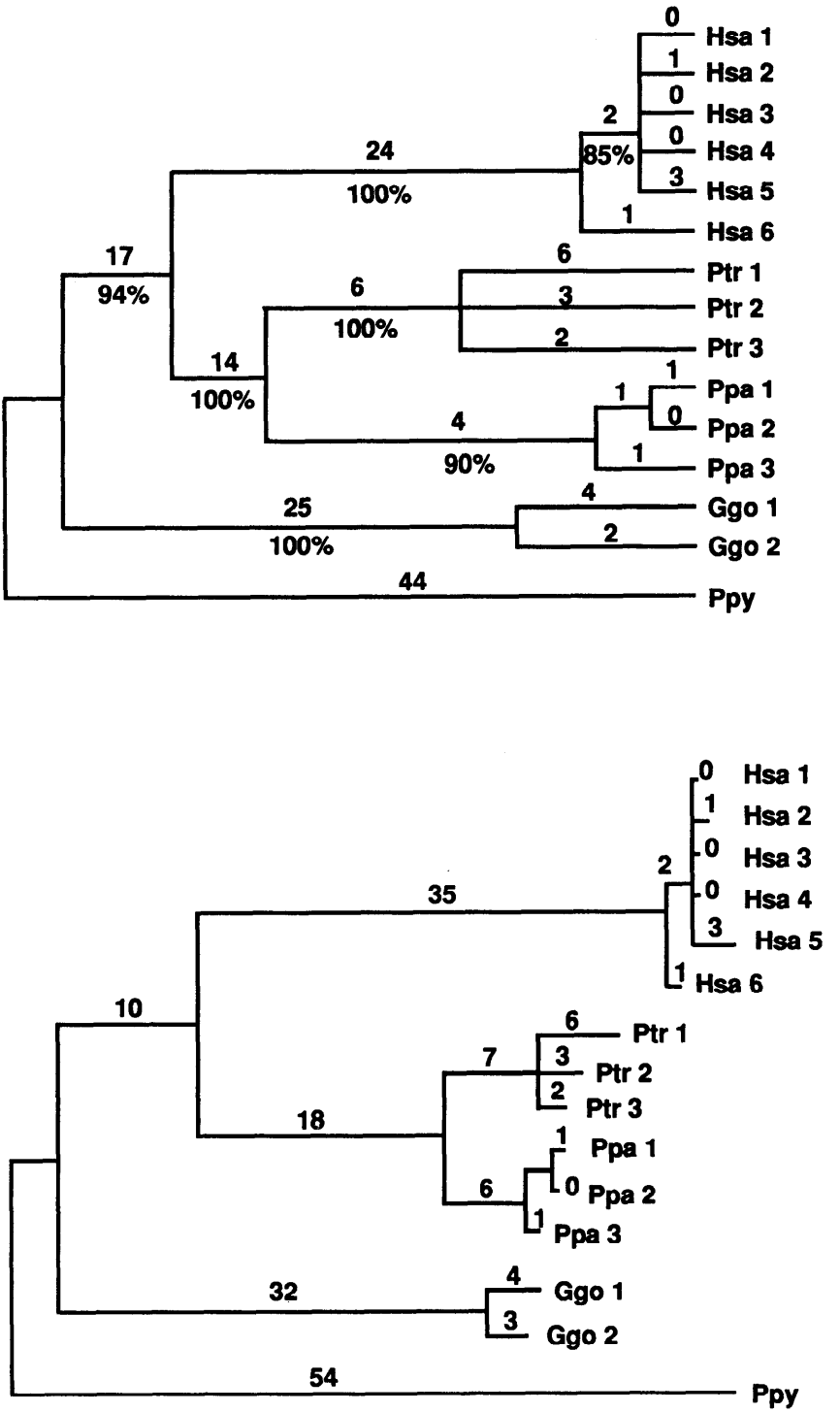


FIG. 3.—Sequence relationships derived from analyses of aligned mitochondrial COII gene sequences (684 bp). *Top,* Maximum parsimony tree (Fitch 1971) constructed using the branch-and-bound search option within PAUP version 3.1.1 (Swofford 1993), with orangutan (Ppy) as outgroup. Tree (length 205, consistency index 0.849) shows minimum possible branch lengths; these unambiguous changes are only a portion of the observed total changes. Bootstrap values indicated were derived from 1,000 replications in PAUP 3.1.1 (Swofford 1993). Species abbreviations are as in table 1. *Bottom,* Distance tree constructed

transversion. This is greater than the average pairwise estimate (0.34%) that includes all pairwise human comparisons, some showing no COII sequence differences. The tree-based value for within-human differences in the ND4-5 region is 0.33% (three transitions of 896 bp total).

Time Scale Based on COII Sequences

Estimating divergence times from molecular data requires, first, the measurement of genetic difference and demonstration of rate constancy; second, estimation of inferred amounts of actual genetic change; and, third, choice of a calibration point and divergence time. Here we avoid error associated with the third step (necessarily reliant on paleontological interpretation) by using relative rather than absolute divergence times. The relative time of the human mitochondrial ancestor is the amount of estimated genetic difference among humans, expressed as a proportion of the estimated genetic difference between human and chimpanzee species. Such "calibration-free" relative divergence times are constant for any given molecular data set. Data sets may differ in their absolute divergence time estimates because of the assumption of different calibration times (compare Ruvolo et al. 1991 with Horai et al. 1992), but using relative date estimates demonstrates their agreement.

Rate Constancy of COII Sequence Data

If DNA evolves at an approximately constant rate, then the number of substitutions that accumulate between two taxa is approximately proportional to their time since divergence. The COII gene has been shown to evolve at a constant rate within higher primates (Ruvolo et al. 1991; Disotell et al. 1992), and the data presented here concur. In the relative-rate test (Sarich and Wilson 1967), distances between an outgroup taxon (in this case, *Pongo*) and different ingroup taxa are compared; equality indicates rate constancy. The average number of observed COII substitutions is 96.5 for *Pongo-Homo,* 99.3 for *Pongo-Pan troglodytes,* 91.6 for *Pongo-Pan paniscus,* and 89.5 for *Pongo-Gorilla,* and all lie within ±5.5% of 94.2, the average for the ingroup species. Therefore, these data exhibit reasonable rate constancy and can be used for divergence time estimates.

Correction Methods for Multiple Substitutions

The observed sequence difference between two taxa is less than or equal to the actual number of substitutions that have occurred since their divergence. This is because, the more ancient the divergence time, the greater the chance of multiple nucleotide substitutions occurring at any given nucleotide position. Actual rather than observed numbers of substitutions are proportional to divergence times (if it is assumed that substitutions occur regularly over time); therefore, a correction method is needed to estimate divergence times.

When no correction is applied, the ratio of *observed* sequence differences provides an upper limit for the relative ancestral human mitochondrial divergence time. This is because (*a*) for the human COII sequences the estimated number of substitutions is equal to or only slightly greater than the observed number but (*b*) between human and chimpanzee the observed difference will be a greater underestimate of the actual substitutional differences. Thus observed differences provide an overestimate; for the COII sequence data, this upper-bound relative date is 0.58%/9.4%, or 1/16.

Several correction methods exist, each reflecting a model of molecular evolutionary change. The methods applied here all assume that transitions and transversions

occur with unequal frequencies and therefore require estimation of the transition: transversion (i:v) ratio. We use a range of i:v ratios consistent with the mtDNA data (see discussion below). The correction methods used here are as follows.

1. *Brown et al.'s (1982) method.*—This is a modification of Jukes and Cantor's (1969) one-parameter model. Transitions and transversions are treated as independent mutational classes; each class is corrected separately, and a weighted average of the corrected values gives the estimated number of actual substitutions. Although this method has been applied frequently to mtDNA data in the literature, it yields, particularly for distantly related taxa, corrected values that are very different from those produced by other correction methods, which generally agree in their estimates. Fitch (1986) has criticized Brown et al.'s method because it treats transitions and transversions as independent processes and is less descriptive of the empirical data than are Kimura's (1980) two-parameter model and Fitch's (1986) nomographic method. Because Brown et al.'s method has been used to estimate the time of the human mitochondrial ancestor (Kocher and Wilson 1991), we apply it here for comparison.

2. *The "transversion method" (Higuchi et al. 1984).*—This assumes that, since, in mtDNA, transitions are much more frequent than transversions, multiple transitional substitutions at any site are more likely than are transversional substitutions. As two taxa diverge, transversional differences should accumulate approximately linearly with time, while observed transitional differences asymptotically level off (Brown et al. 1982). This has been empirically confirmed for mtDNA sequence data (Miyamoto and Boyle 1989; Irwin et al. 1991). Given an estimate of the i:v ratio, the actual number of transitions is roughly the number of transversions times the i:v ratio. Therefore the total number of substitutions can be estimated as

$$\begin{aligned} \text{total substitutions} &= \text{transversions} + \text{transversions} \times (\text{i:v}) \\ &= \text{transversions} \times (1 + \text{i:v}) . \end{aligned} \qquad (1)$$

Like Brown et al.'s (1982) method, the transversion method has the drawback that transitions and transversions are treated as independent classes of substitutional events. We include it because this method has also been used to estimate the time of the human mitochondrial ancestor (Vigilant et al. 1991), with confidence intervals (Nei 1992).

3. *Kimura's (1980) two-parameter method.*—This allows transitions and transversions to have different substitutional frequencies, but these two mutational types are not independent.

4. *Maximum-likelihood correction method (Felsenstein 1990; Kishino and Hasegawa 1989).*—This maximizes the joint probability under the model of pairs of sequences; it is equivalent to constructing two species phylogenies by the maximum-likelihood method and taking the total branch lengths as distances (J. Felsenstein, personal communication). It has a more general underlying model than does Kimura's (1980) two-parameter method, because it allows unequal base frequencies, and, for this reason, we consider it to give the best estimates.

From the analysis of COII, ND4-5, and hypervariable control-region sequences (table 2), we draw the following conclusions. First, relative date estimates vary with correction methods used and with estimated i:v ratios. Second, correction methods differ in how sensitive they are to changes in i:v ratios, with the transversion method being the most sensitive, Brown et al.'s (1982) method intermediately so, and the remaining two methods least sensitive. Third, relative date estimates from the three

Table 2
Relative Divergence Dates for Human Mitochondrial Ancestor

mtDNA Region and Correction Method	Relative Divergence Date[a] When Transition:Transversion Ratio Is		
	15:1	30:1	60:1
COII gene (684 bp; present study):			
Maximum likelihood[b]	1/18	1/19	1/19
Kimura (1980) two-parameter[c]	1/18	1/18	1/19
Transversion method[d]	...[e]	1/23	1/48
Brown et al. (1982)	...[e]	1/19	1/22
ND4-5 region (896 bp; Kocher and Wilson 1991):			
Maximum likelihood	1/29	1/29	1/29
Kimura two-parameter	1/28	1/28	1/29
Transversion method	1/27	1/52	1/103
Brown et al. (1982)	1/30[f]	1/45	1/59
Hypervariable control subregions (Vigilant et al. 1991):[g]			
Transversion method	1/24[f]	1/47	1/93
Brown et al. (1982)	1/13	1/23	1/33

[a] Ratios of within-human to between-human-and-chimpanzee nucleotide substitutions were estimated by different correction methods.

[b] As implemented by Felsenstein (1990) in PHYLIP 3.3.

[c] As implemented by Felsenstein (1990) in PHYLIP 3.3.

[d] Higuchi et al. (1984).

[e] Estimated value is less than that observed because i:v ratio is >15:1.

[f] Previously published value.

[g] Value used for within human sequence divergence in all cases is 2.87% (Vigilant et al. 1991). Hasegawa and Horai (1991) have also analyzed human control region sequences, using a variant of the maximum-likelihood method; on the basis of three different subregions, their relative divergence dates are 1/14, 1/17, and 1/12.5, with transition:transversion ratios 17:1, 27:1, and 14:1, respectively.

mitochondrial regions do not agree when correction method and i:v ratio are held constant, although estimates from ND4-5 and hypervariable control regions tend to be more similar than those from the COII gene. This raises two questions: are the differences between the estimates real in the sense that the mtDNA regions are evolving differently, and which relative date estimate is best?

Substitutional Constraints and Transition:Transversion Ratios

Because of differing constraints on nucleotide substitutions, such as those having to do with codon position and functional properties of encoded proteins, DNA sites evolve at different rates (Li et al. 1985). As R. C. Lewontin (personal communication), has noted, "the sum total of all DNA sequencing studies to date shows that, except for pseudogenes, there is probably no class of DNA not under substitutional constraints." Some of these constraints are understood (e.g., synonymous vs. nonsynonymous changes), while others are not (Gillespie 1991).

From the mitochondrial genetic code, most class 1 (third codon position) substitutions are silent, while most class 2 (first and second codon positions) substitutions lead to amino acid replacements. In mitochondrial protein-coding genes, the observed substitution frequency is far greater for class 1 than for class 2 sites (Brown 1985). This pattern is evident for both COII (figs. 2, *bottom,* and 4A) and ND4-5 protein-coding regions (fig. 4B): class 1 sites accumulate substitutions more quickly than do

class 2 sites. (Control region sequences are noncoding and cannot be analyzed in this way.)

Between COII and ND4-5 protein-coding regions, class 2 sites are accumulating substitutions differently, suggesting that the two regions are under different substitutional/selective constraints (see fig. 4A and B). However, class 1 sites across the two regions are accumulating transitions and transversions similarly (fig. 4C). This fits with the expectation that class 1 substitutions more closely approximate the underlying mutational process than do class 2 substitutions. Since class 1 sites are relatively unconstrained in their substitutions and are also accumulating substitutions similarly over different mitochondrial coding regions, they are likely to provide better relative date estimates than are all sites from an mtDNA region. Viewed as contiguous stretches of DNA, the COII gene and ND4-5 region are evolving differently, but subsets of the two mtDNA regions are evolving in the same way, and we will use these for relative date estimation. Note that this approach is applicable only in cases where sequence differences are small, so that class 1 sites are not saturated with multiple substitutions.

For a best estimate of the i:v ratio, closely related species should be used, since more distantly related species may show lowered i:v ratios because of multiple substitutions at some sites (Simon 1991). However, even closely related species may have multiple substitutions, and within-species comparisons are then preferable. Ideally, the i:v ratio should be calculated as phylogenetic distance approaches zero; here we examine the slope of the transition-transversion curve, while Fitch's (1986) nomographic method uses the intercept on the i:v-ratio axis. For the COII gene, we now have intraspecific sequence data for several hominoids (fig. 5). In class 1 sites, an i:v ratio of 15:1 is a clear underestimate, while estimates of 30:1 and even 60:1 are consistent with the data. The nomographic method (Fitch 1986) estimates that the i:v ratio for these data is >20:1 (W. M. Fitch, personal communication).

To summarize, we would argue that the "best" relative divergence estimate is one based on class I substitutions only, using the maximum-likelihood correction method for multiple substitutions and an i:v ratio in the range of 30:1–60:1. When calculated this way, relative divergence estimates from COII and ND4-5 protein-coding regions now agree (table 3). From these slowly evolving mitochondrial coding regions, the best relative date estimate for the human mitochondrial ancestor as a proportion of the human-chimpanzee divergence time is 1/27.

Paleontological Calibration of Relative Molecular Dates

Testing whether a relative divergence date is consistent with other types of anthropological evidence requires its conversion to an absolute date. This depends on choice of calibration time for some paleontological or prehistorical event—in this case, the human-chimpanzee divergence. If we take the latest possible human-chimpanzee divergence to be 6 Mya (Hill and Ward 1988), the slowly evolving mitochondrial coding regions estimate the human mitochondrial ancestor at 1/27th this time, or 222,000 years. If the species' divergence were as early as 10 Mya (de Bonis et al. 1990), the age indicated by the molecular data would be 370,000 years.

The hypothesis that the human mitochondrial ancestor lived ≥1 Mya (Wolpoff 1989) can be tested. For combined COII and ND4-5 data, there are 460 class I sites; observed differences are 6.4 bp within humans and 108.5 bp between humans and chimpanzees; corrected maximum-likelihood values are 6.5 bp and 174.0 bp (with a 30:1 i:v ratio) and 6.5 bp and 177.6 bp (with a 60:1 i:v ratio), respectively. If the human-chimpanzee divergence is assumed to have occurred 6 Mya, the expected

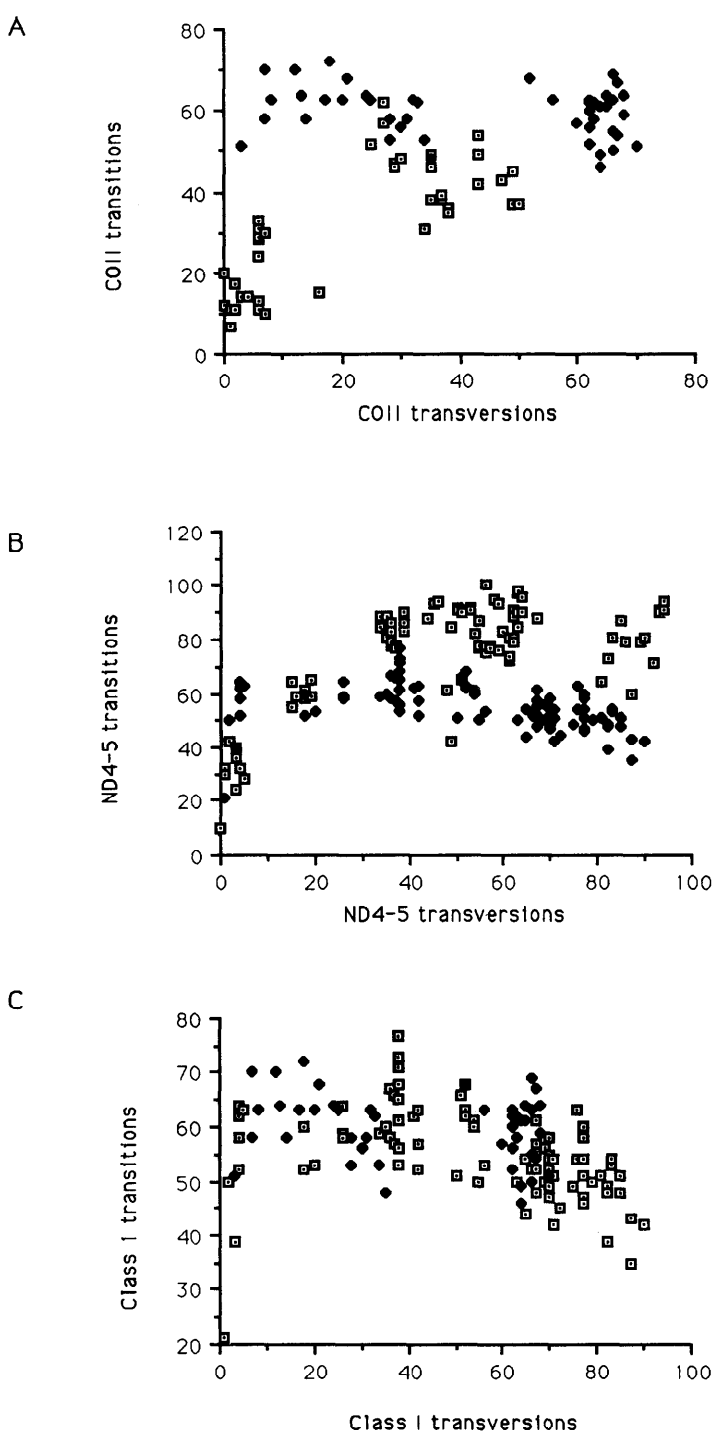


FIG. 4.—Class 1 (♦) and class 2 (⊡) transitional and transversional differences in two slowly evolving mtDNA regions. Class 1 are third codon positions, and class 2 the first and second codon positions. A, COII gene sequences, 684 nucleotides long, are from seven primate species [orangutan (Horai et al. 1992), human, chimpanzee, gorilla, siamang, macaque, and green monkey], mouse, cow, and African clawed toad (Ruvolo et al. 1991, and references therein). B, Protein-coding portions of the mitochondrial ND4-5 region (tRNA sequences are not included). Sequences, 696 nucleotides long, are from 12 primate species, mouse, and cow (Hasegawa et al. 1990, and references therein). C, Class 1 transitions vs. transversions for COII (♦) and

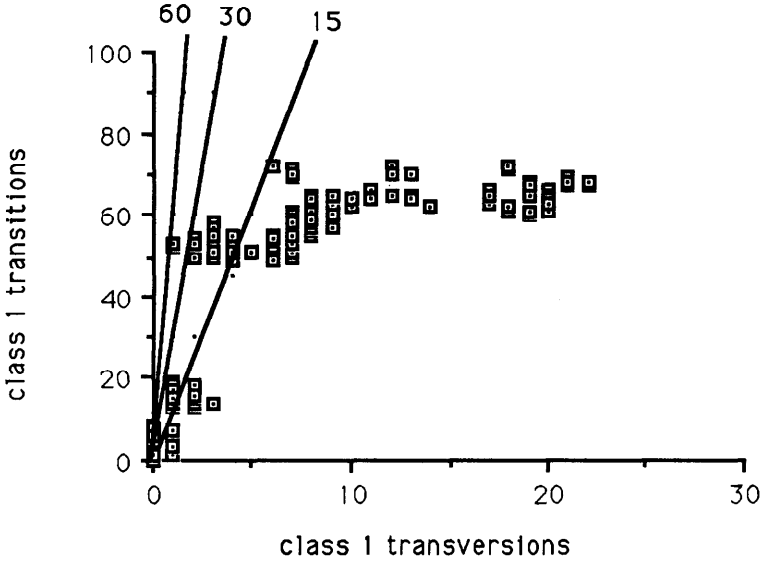


FIG. 5.—Class 1 transitional changes vs. class 1 transversional changes in 24 hominoid COII sequences (Ruvolo et al. 1991; present study; M. Ruvolo, unpublished data), with lines corresponding to i:v ratios of 15:1, 30:1, and 60:1 indicated. These values include intraspecific as well as interspecific pairwise sequence comparisons.

number of combined COII and ND4-5 class I substitutions through the root of a 1-Myr-old human clade would be 1/6 the human-chimpanzee difference, 29.0 bp (with a 30:1 i:v ratio) or 29.6 bp (with a 60:1 i:v ratio). Both expected values are significantly different from the observed corrected 6.5 bp ($\chi^2=17.4$, $P<0.005$; $\chi^2=18.4$, $P<0.005$). For a 10-Mya human-chimpanzee divergence (de Bonis et al. 1990), the expected number of differences for a 1-Myr-old common mitochondrial haplotype is also significantly greater than that observed: 17.4 bp for i:v = 30:1 ($\chi^2=6.8$, $P<0.01$) and 17.8 bp for i:v = 60:1 ($\chi^2=7.2$, $P<0.01$).

As is clear from this analysis, the degree of belief in a human mitochondrial ancestor at 1 Mya is dependent on our choice of a human-chimpanzee divergence time. However, even with a human-chimpanzee divergence as early as 10 Mya, these mitochondrial data are not consistent with a 1-Myr-old common ancestral human mitochondrial haplotype. If an even earlier date for the presumed age of the ancestral human haplotype is tested, such as 1.4 Mya for an *H. erectus* exodus from Africa (Bar-Yosef 1987), the hypothesis is even more strongly rejected.

Coalescence Time Estimates

Templeton (1993) has recently observed that the stochastic nature of the evolutionary process has been ignored in time estimates for the human mitochondrial ancestor. Following his analysis and applying the neutral coalescent model of Tajima (1983), we can estimate mean time to coalescence for human mitochondrial haplotypes that differ most. The method requires specification of a mutation rate, which is a form of calibration. Here we use rates estimated from all 1,580 sites of combined COII and ND4-5 regions. Between humans and chimpanzees, there are 163 inferred substitutions (by maximum-likelihood correction, 30:1 i:v ratio), or 10.3%. For human-chimpanzee divergence times of 4 Mya, 6 Mya, and 10 Mya, nucleotide substitution rates are 1.3×10^{-8}, 0.85×10^{-8}, and 0.5×10^{-8}/site/year/lineage, respectively.

Table 3
Relative Divergence Dates for Human Mitochondrial Ancestor, from Class I Sites (Third Codon Positions) of Protein-coding Genes

mtDNA Region and Correction Method	Relative Divergence Date[a] When Transition: Transversion Ratio Is 15:1	30:1	60:1
COII class I sites only (228 bp;[b] present study):			
Maximum likelihood[c]	1/25	1/27	1/27
Kimura two-parameter[d]	1/21	1/22	1/22
ND4-5 class I sites only (232 bp;[e] Kocher and Wilson 1991):			
Maximum likelihood	1/25	1/27	1/28
Kimura two-parameter	1/22	1/23	1/23

[a] Ratios of within-human to between-human-and-chimpanzee nucleotide substitutions were estimated by different correction methods.

[b] For 228 class I COII sites, there are 3.4 bp of class I substitutions among humans along the tree and 56.2 bp between human and chimpanzee. The corrected maximum-likelihood human value (3.4 bp) equals the observed value, while the human-chimpanzee difference is corrected to 90.9 bp (when a 30:1 i:v ratio is used) or 91.7 bp (when a 60:1 i:v ratio is used).

[c] As implemented by Felsenstein (1990) in PHYLIP 3.3

[d] As implemented by Felsenstein (1990) in PHYLIP 3.3.

[e] For 232 ND4-5 class I sites, the observed 3.0 bp through the root of the human clade and 52.3 bp observed average between humans and chimpanzees get corrected to 3.1 bp for humans, in both cases, and to 83.1 bp and 85.9 bp between species, for i:v ratios 30:1 and 60:1 respectively.

To calculate coalescence times, we need to estimate the expected nucleotide heterozygosity of the combined COII and ND4-5 mitochondrial regions. This is done by calculating the heterozygosities of the two mtDNA regions separately and then taking a weighted average of the two (using both number of individuals and region size in weighting) as the best estimate (R. C. Lewontin, personal communication). Following Templeton (1993), we use Ewens's (1983) formulation of expected nucleotide heterozygosity θ:

$$\theta = k^*/[1+1/2+1/3+\cdots+1/(n-1)], \quad (2)$$

where k^* is the number of sites at which two or more different nucleotides occur and n is the number of "genes" sampled (in this case, the number of individuals). For the first (COII) data set, seven sites vary among six humans, so that $\theta_1 = 3.07$. For the second (ND4-5) data set, six sites vary among seven humans, so that $\theta_2 = 2.45$. The first data set contains 684 bp of DNA from each of six individuals, for a total of 4,104 bp; the second data set contains 896 bp of DNA from each of seven individuals, for a total of 6,272 bp. The weighting factors are then 0.4 (4,104/10,376) and 0.6 (6,272/10,376), respectively, yielding an expected nucleotide heterozygosity of 2.70 for the combined data.

Time to coalescence, T, can be estimated from Templeton's (1993) formulation of Tajima's (1983) equation (20), as

$$T = \theta(1+k)/[2\mu(1+n\theta)], \quad (3)$$

where k is the pairwise divergence among haplotypes (in number of nucleotide dif-

ferences), n is the number of sampled nucleotides, μ is the mutation rate (in substitutions per site per year), and θ is the expected nucleotide heterozygosity. The variance in coalescence time, from Templeton's formulation of Tajima's [1983, eq. (21)], is

$$\sigma^2 = \theta^2(1+k)/[4\mu^2(1+n\theta)^2]\,. \tag{4}$$

For COII and ND4-5 regions, $k = 7$ bp separates the two most different human haplotypes through the root of the human clade, and $n = 1{,}580$ sites are compared. This yields a mean coalescence time of

$$T = 2.53 \times 10^{-3}/\mu \tag{5}$$

with standard deviation

$$\sigma = 8.9 \times 10^{-4}/\mu\,. \tag{6}$$

For $\mu = 1.3 \times 10^{-8}$/year, the mean coalescence time is 195,000 years with standard deviation 68,000 years; for $\mu = 0.85 \times 10^{-8}$/year, the mean coalescence time is 298,000 years with standard deviation 105,000 years; for $\mu = 0.5 \times 10^{-8}$/year, the mean coalescence time is 506,000 years with standard deviation 178,000 years.

As Templeton (1993) observes, 95% confidence limits can be estimated about these coalescence times by using Kimura's (1970) finding that overall distribution of T is approximately gamma distributed. For combined COII and ND4-5 human mtDNA sequences, the estimated mean coalescence time of 195,000 years (corresponding to a 4-Mya human-chimpanzee divergence) has 95% confidence limits of 85,000–349,000 years; the estimated mean of 298,000 years (for a 6-Mya human-chimpanzee divergence) has 95% confidence limits of 129,000–536,000 years; and the estimated mean of 506,000 years (for a 10-Mya human-chimpanzee divergence) has 95% confidence limits of 220,000–910,000 years.

These broad time ranges imposed by the stochastic nature of the evolutionary process notably do not include the timepoint of 1 Mya, although they come close if we assume a human-chimpanzee divergence at 10 Mya. For a 4–6-Mya human-chimpanzee divergence, a multiregional hypothesis that envisions modern *H. sapiens* as emerging from anciently divergent *H. erectus* populations spread throughout the Old World seems unlikely.

Interpreting Other Studies

We emphasize that a molecular data set has to go through several layers of interpretation before even relative divergence times can be estimated and that choice of correction method and i:v ratio can contribute significantly to differences in relative divergence-time estimates. Therefore, it is not instructive to compare estimated dates from existing mtDNA studies (Cann et al. 1987; Vigilant et al. 1989, 1991; Hasegawa and Horai 1991; Kocher and Wilson 1991; Nei 1992; Pesole et al. 1992; Stoneking et al. 1992; Tamura and Nei 1993; Templeton 1993), without consideration of correction methods (equivalent to models of evolutionary change) and estimated parameters (also involving assumptions about how DNA evolves).

For example, there is apparent similarity between published dates estimated from the hypervariable control-region (Vigilant et al. 1991) and ND4-5 sequences (Kocher and Wilson 1991); however, these were made by using different correction methods.

If the same correction method (Brown et al. 1982; as in Kocher and Wilson 1991) and the same 15:1 i:v ratio (found for both data sets) are applied, the relative date estimates differ by more than a factor of two (1/13 vs. 1/30, respectively) (table 2). If, instead, the transversion method (following Vigilant et al. 1991) is applied to both, there is better agreement between data sets (1/24 vs. 1/27). However the transversion method is sensitive to differences in i:v ratio, and this is problematic for date estimation even from a single data set. Analyzing the hypervariable control-region data in different ways, Vigilant (1990, thesis on pp. 72–73) finds that 15:1 and 30:1 i:v ratios are both consistent with the data and not statistically different, but these ratios produce relative dates varying by a factor of two (1/24 and 1/47, respectively; table 2). For human control regions, i:v ratios >15:1 have been estimated by others: 24:1 (Aquadro and Greenberg 1983) and 27:1 (Hasegawa and Horai 1991). Agreement among time estimates does not necessarily signify convergence on an acceptable answer, unless the best possible model of molecular evolutionary change with well-estimated parameters produces those estimates.

For time estimates, confidence limits are also not comparable unless they summarize error in the same variables. From control-region data, Nei (1992) calculates 95% confidence limits of 110,000–504,000 years ago for the time of the common human mitochondrial ancestor. Although it is tempting to compare this with the range derived for COII and ND4-5 data, the two are not equivalent: Nei's estimate is based on the transversion method (which is not as good a model of molecular evolutionary change as are other methods), assumes a 15:1 i:v ratio (although a higher ratio is also compatible with the data and changes the mean considerably), and provides error bars associated only with nucleotide substitution rate (not with stochastic aspects of evolutionary change).

Coalescence times and confidence intervals estimated here for COII and ND4-5 and by Templeton (1993) for control-region data *are* comparable, since both use the same model of evolutionary stochasticity. However, Templeton (1993) does not reject a 1-Myr-old human mitochondrial ancestor, based on estimates using an average mitochondrial mutation rate in the range of $1–2 \times 10^{-8}$/site/year/lineage. While this range is appropriate for more slowly evolving mtDNA regions (e.g., COII and ND4-5), it is an underestimate for hypervariable control subregions, which, as we have shown, evolve roughly 10 times faster. Use of a higher mutation rate will decrease both coalescence time and confidence-interval estimates from control-region data.

By judicious choice of correction method and i:v ratio, we can probably get any two molecular data sets to agree on some predetermined time estimate, but this is counterproductive. What is needed is the development of more generalized (hence better) correction models (e.g., see Hasegawa and Horai 1991; Tamura and Nei 1993), further characterization of molecular evolutionary parameters (e.g., i:v ratios), consideration of all types of error associated with date estimates (Templeton 1993), and consistent application of good methods to different molecular data sets.

Could a Recent Date for the Common Human Mitochondrial Type Be Artifactual?

The mtDNA haplotype date could be later than the actual human ancestral population if mtDNA diversity has been lost during hominid evolution (Wolpoff 1989). Probability of loss is higher in small, nonexpanding populations (Avise et al. 1984), demographic conditions that are thought to be characteristic throughout most of human evolution. However, these demographic conditions are no less characteristic of the

other hominoids, some of which show long branches (ancient mitochondrial lineages) on the COII gene tree. For example, two common chimpanzees surveyed here differ by 8 bp at COII class I sites, more than twice the observed average in humans. In light of the fact that, unlike the humans sampled, the chimpanzee individuals were chosen randomly and probably do not adequately represent total species' genetic variation, this difference is even more impressive, arguing against a solely demographic explanation for reduced human genetic variability.

Because only a small proportion of all living humans have been surveyed, we may have missed sampling someone who is mitochondrially very different from those humans already characterized. While possible, this is unlikely for two reasons. First, examination of the apportionment of genetic diversity within the human species shows that a high proportion (86%) of intraspecific nuclearly encoded variability is contained within populations (Lewontin 1972; Latter 1980), so that humans populations are not highly differentiated. Second, female hominoids generally transfer between groups (Goodall 1986, p. 86; Pusey and Packer 1986; Rodseth et al. 1991; Kano 1992, p. 70), thus insuring mtDNA flow throughout the species. These observations suggest that human populations with mtDNA types highly different from those already discovered are not likely to be found.

Predicted Estimates of Human mtDNA Differences

These mitochondrial sequence data from slowly evolving regions can help us estimate how different the 15,000 bp of non-control-region mtDNA are among the most divergent humans known. Among humans, there is a maximum sequence difference of (*a*) 6 bp in 684 bp of the COII gene and (*b*) 4 bp in 896 bp of the ND4-5 region. If the two slowly evolving mitochondrial regions are assumed to be representative of the 15,000 bp of non-control-region mtDNA, the sequence difference between most different human mitochondrial types is estimated to be ~95 bp, with 19 phylogenetically informative sites in non-control-region mtDNA. These are likely overestimates, since mitochondrially encoded tRNAs and rRNAs evolve more slowly than do mitochondrial protein-coding genes (Cann et al. 1984, 1987). It remains to be seen whether sequencing entire mitochondrial genomes will give sufficient differences among living humans for adequate phylogenetic resolution; it would, however, increase the accuracy of relative date estimates and reduce their confidence limits.

Sequence Availability

The eight newly reported COII gene sequences presented herein have been deposited in GenBank.

Acknowledgments

We would like to thank K. Morrison for culturing cells, J. Tate for preparing DNA, O. Ryder of the San Diego Zoo, R. Honeycutt, and L. L. Cavalli-Sforza for providing DNA samples, Marc Glickman for statistical consulting, and D. Pilbeam, R. C. Lewontin, Walter Fitch, T. F. Smith, C. A. Aquadro, T. Disotell, an anonymous reviewer, and members of my lab for invaluable discussion. This work was supported by the Wenner-Gren Foundation for Anthropological Research, the Harvard Human Origins Research Fund, and Elaine and Leon Cohn.

LITERATURE CITED

ANDERSON, S., A. T. BANKIER, B. G. BARRELL, M. H. L. DE BRUIJN, A. R. COULSON, J. DROUIN, I. C. EPERON, D. P. NIERLICH, B. A. ROE, F. SANGER, P. H. SCHREIER, A. J. H. SMITH, R. STADEN, and I. G. YOUNG. 1981. Sequence and organization of the human mitochondrial genome. Nature **290**:457–465.

AQUADRO, C. F., and B. D. GREENBERG. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. Genetics **103**:287–312.

AVISE, J. C., J. E. NEIGEL, and J. ARNOLD. 1984. Demographic influences of mitochondrial DNA lineage survivorship in animal populations. J. Mol. Evol. **20**:99–105.

BAR-YOSEF, O. 1987. Pleistocene connections between Africa and southwest Asia: an archaeological perspective. Afr. Arch. Rev. **5**:29–38.

BROWN, W. M. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc. Natl. Acad. Sci. USA 77:3605–3609.

———. 1985. The mitochondrial genome of animals. Pp. 95–130 *in* R. J. MACINTYRE, ed. Molecular evolutionary genetics. Plenum, New York.

BROWN, W. M., E. M. PRAGER, A. WANG, and A. C. WILSON. 1982. Mitochondrial DNA sequences of Primates: tempo and mode of evolution. J. Mol. Evol. **18**:225–239.

CANN, R. L., W. M. BROWN, and A. C. WILSON. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics **106**:479–499.

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

DE BONIS, L., G. BOUVRAIN, D. GERAADS, and G. KOUFOS. 1990. New hominid skull material from the later Miocene of Macedonia in northern Greece. Nature **345**:712–714.

DI RIENZO, A., and A. C. WILSON. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc. Natl. Acad. Sci. USA **88**:1597–1601.

DISOTELL, T. R., R. L. HONEYCUTT, and M. RUVOLO. 1992. Mitochondrial DNA phylogeny of the Old-World monkey tribe Papionini. Mol. Biol. Evol. **9**:1–13.

EWENS, W. J. 1983. The role of models in the analysis of molecular genetic data, with particular reference to restriction fragment data. Pp. 45–73 *in* B. S. WEIR, ed. Statistical analysis of DNA sequence data. Marcel Dekker, New York and Basel.

FELSENSTEIN, J. 1990. PHYLIP: phylogenetic inference package, version 3.3. Department of Genetics, University of Washington, Seattle.

FITCH, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. **19**:406–416.

———. 1986. A hidden bias in the estimate of total nucleotide substitutions from pairwise differences. Pp. 315–328 *in* S. KARLIN and E. NEVO, eds. Evolutionary processes and theory. Academic Press, New York.

FITCH, W. M., and E. MARGOLIASH. 1967. A method for estimating the number of invariant amino acid coding positions in a gene using cytochrome *c* as a model case. Biochem. Genet. **1**:65–71.

GILLESPIE, J. H. 1991. The causes of molecular evolution. Oxford University Press, New York.

GOODALL, J. 1986. The chimpanzees of Gombe: patterns of behavior. Harvard University Press, Cambridge, Mass.

HASEGAWA, M., and S. HORAI. 1991. Time of the deepest root for polymorphism in human mitochondrial DNA. J. Mol. Evol. **32**:37–42.

HASEGAWA, M., H. KISHINO, K. HAYASAKA, and S. HORAI. 1990. Mitochondrial DNA evolution in Primates: transition rate has been extremely low in the lemur. J. Mol. Evol. **31**:113–121.

HEDGES, S. B., S. KUMAR, K. TAMURA, and M. STONEKING. 1992. Human origins and analysis of mitochondrial DNA sequences. Science **255**:737–739.

HIGUCHI, R., B. BOWMAN, M. FREIBERGER, O. A. RYDER, and A. C. WILSON. 1984. DNA sequences from the quagga, an extinct member of the horse family. Nature **312**:282–284.

HILL, A., and S. WARD. 1988. Origin of the Hominidae: the record of African large hominoid evolution between 14 my and 4 my. Yearbook Phys. Anthropol. **31**:49–83.

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.

IRWIN, D. M., T. D. KOCHER, and A. C. WILSON. 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. **32**:128–144.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. MUNRO, ed. Mammalian protein metabolism, Academic Press, New York.

KANO, T. 1992. The last ape: pygmy chimpanzee behavior and ecology. Stanford University Press, Stanford.

KIMURA, M. 1970. The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population. Genet. Res. **15**:131–133.

———. 1980. A simple method of estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

KISHINO, H., and M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. **29**:170–179.

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.

LATTER, B. D. H. 1980. Genetic differences within and between populations of the major human subgroups. Am. Nat. **116**:220–237.

LEWONTIN, R. C. 1972. The apportionment of human diversity. Evol. Biol. **6**:381–398.

LI, W.-H., C.-C. LUO, and C.-I. WU. 1985. Evolution of DNA sequences. Pp. 1–94 *in* R. J. MACINTYRE, ed. Molecular evolutionary genetics. Plenum, New York.

MADDISON, D. R., M. RUVOLO, and D. L. SWOFFORD. 1992. Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. Syst. Biol. **41**: 111–124.

MANIATIS, T., E. F. FRITSCH, and J. SAMBROOK. 1989. Molecular cloning: a laboratory manual, 2d ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MIYAMOTO, M. M., and S. M. BOYLE. 1989. The potential importance of mitochondrial DNA sequence data to eutherian mammal phylogeny. Pp. 437–450 *in* B. FERNHOLM, K. BREMER, and H. JORNVALL, eds. The hierarchy of life. Elsevier, Amsterdam.

NEI, M. 1992. Age of the common ancestor of human mitochondrial DNA. Mol. Biol. Evol. **9**:1176–1178.

PESOLE, G., E. SBISÁ, G. PREPARATA, and C. SACCONE. 1992. The evolution of the mitochondrial D-loop region and the origin of modern man. Mol. Biol. Evol. **9**:587–598.

PUSEY, A. E., and C. PACKER. 1986. Dispersal and philopatry. Pp. 250–266 *in* B. B. SMUTS, D. L. CHENEY, R. M. SEYFARTH, R. W. WRANGHAM, and T. T. STRUHSAKER, eds. Primate societies. University of Chicago Press, Chicago.

RODSETH, L., R. W. WRANGHAM, A. M. HARRIGAN, and B. B. SMUTS. 1991. The human community as a primate society. Curr. Anthropol. **32**:221–254.

RUVOLO, M., T. R. DISOTELL, M. W. ALLARD, W. M. BROWN, and R. L. HONEYCUTT. 1991. Resolution of the African hominoid trichotomy by use of a mitochondrial gene sequence. Proc. Natl. Acad. Sci. USA **88**:1570–1574.

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

SARICH, V. M., and A. C. WILSON. 1967. Immunological time scale for hominid evolution. Science **158**:1200–1203.

SIMON, C. 1991. Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA.

Pp. 33–71 *in* G. M. HEWITT, A. W. B. JOHNSTON, and J. P. W. YOUNG, eds. Molecular techniques in taxonomy. NATO ASI ser. H57:33–71. Springer, Berlin and Heidelberg.

STONEKING, M., S. T. SHERRY, A. J. REDD, and L. VIGILANT. 1992. New approaches to dating suggest a recent age for the human mtDNA ancestor. Philos. Trans. R. Soc. Lond. [Biol.] **337**:167–175.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.

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.

TEMPLETON, A. R. 1992. Human origins and analysis of mitochondrial DNA sequences. Science **255**:737.

———. 1993. The "Eve" hypothesis: a genetic critique and reanalysis. Am. Anthropol. **95**:51–72.

VIGILANT, L. 1990. Control region sequences from African populations and the evolution of human mitochondrial-DNA. Ph.D. diss., University of California, Berkeley.

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.

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

WOLPOFF, M. 1989. Multiregional evolution: the fossil alternative to Eden. Pp. 62–108 *in* P. MELLARS and C. STRINGER, eds. The human revolution: behavioral and biological perspectives on the origins of modern humans. Princeton University Press, Princeton, N.J.

WALTER M. FITCH, reviewing editor

Received February 22, 1993; revision received August 2, 1993

Accepted August 2, 1993