Nonneutral Mitochondrial DNA Variation in Humans and Chimpanzees

Michael W. Nachman,* Wesley M. Brown,† Mark Stoneking‡ and Charles F. Aquadro*

**Section of Genetics and Development, Cornell University, Ithaca, New York 14853, †Department of Biology, University of Michigan, Ann Arbor, Michigan 48109, and ‡Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania 16802*

Manuscript received May 20, 1995
Accepted for publication November 18, 1995

ABSTRACT

We sequenced the NADH dehydrogenase subunit 3 (*ND3*) gene from a sample of 61 humans, five common chimpanzees, and one gorilla to test whether patterns of mitochondrial DNA (mtDNA) variation are consistent with a neutral model of molecular evolution. Within humans and within chimpanzees, the ratio of replacement to silent nucleotide substitutions was higher than observed in comparisons between species, contrary to neutral expectations. To test the generality of this result, we reanalyzed published human RFLP data from the entire mitochondrial genome. Gains of restriction sites relative to a known human mtDNA sequence were used to infer unambiguous nucleotide substitutions. We also compared the complete mtDNA sequences of three humans. Both the RFLP data and the sequence data reveal a higher ratio of replacement to silent nucleotide substitutions within humans than is seen between species. This pattern is observed at most or all human mitochondrial genes and is inconsistent with a strictly neutral model. These data suggest that many mitochondrial protein polymorphisms are slightly deleterious, consistent with studies of human mitochondrial diseases.

A basic aim of population genetics is to understand the forces that shape and maintain genetic variability in nature. This is important not only for understanding how evolution proceeds, but also for choosing appropriate genetic markers for population studies. The neutral theory serves as a useful null hypothesis for studies of genetic variation because it makes simple, testable predictions.

The neutral theory asserts that most mutations are not in fact neutral, but rather, deleterious (KIMURA 1983). However, those mutations that survive long enough to be sampled are postulated to be equivalent with respect to fitness. Under this hypothesis, the amount of variation within a species is determined by the neutral mutation rate and the effective population size. The amount of divergence between species is determined by the neutral mutation rate and the time since divergence (and to a small extent, the effective population size). One prediction of the neutral theory is that the amount of variation within species will be correlated with the amount of divergence between species for different genes or gene regions. This prediction forms the basis of several statistical tests of neutrality that utilize DNA sequence data (*e.g.*, HUDSON *et al.* 1987; MCDONALD and KREITMAN 1991).

Mitochondrial DNA (mtDNA) is used extensively as a marker in evolutionary studies; implicit in these studies is the assumption that it is a neutral marker. This assumption is important for such things as measuring gene flow (SLATKIN 1985), estimating effective population size (WILSON *et al.* 1985), detecting population subdivision (AVISE *et al.* 1987), and dating events using a molecular clock (CANN *et al.* 1987).

A number of studies have compared patterns of RFLP variation in human mtDNA to the predictions of neutral, equilibrium models (*e.g.*, WHITTAM *et al.* 1986; EXCOFFIER 1990; MERRIWETHER *et al.* 1991) and found fewer intermediate-frequency variants than expected using WATTERSON's test of homozygosity (1978) or TAJIMA's test (1989). While these findings are inconsistent with a neutral, equilibrium model, it is unclear whether the deviations arise from nonequilibrium conditions (*e.g.*, a recent population expansion), nonneutral conditions (*e.g.*, a selective sweep of a mtDNA haplotype), or both.

More recent studies have used DNA sequence data to test the hypothesis that mtDNA variation is neutral. NACHMAN *et al.* (1994) sequenced the NADH dehydrogenase subunit 3 (*ND3*) gene in 56 wild *Mus domesticus*, two *M. musculus* and one *M. spretus* and found that the ratio of replacement to silent nucleotide differences was significantly greater within species than between species. This result is inconsistent with neutral expectations and cannot be explained by nonequilibrium conditions alone. In Drosophila, nonneutral patterns have been documented for NADH dehydrogenase subunit 5 (*ND5*) (RAND *et al.* 1994), cytochrome b (BALLARD and KREITMAN 1994), and ATPase6 (KANEKO *et al.* 1993). These studies detected a variety of complicated patterns that may be due to several different factors. Common to all three studies, however, was the finding of higher ratios of replacement to silent changes within species than between species either for all or part of the genes in question.

Corresponding author: Michael W. Nachman, Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721.

The present study was undertaken to further address the generality of this result. We are interested in two questions: Is the high ratio of replacement to silent polymorphisms that is seen in mice and flies also seen in humans and chimpanzees? Is this nonneutral pattern common to other mitochondrial genes? To address these questions, we have sequenced the *ND3* gene in 61 humans from around the world, five common chimpanzees, and one gorilla. Comparisons of replacement and silent nucleotide differences within and between species reject a strictly neutral model as an explanation for the evolution of this gene in humans and chimpanzees. To see if this result is common to other mitochondrial genes, we reanalyzed previously published RFLP data from the entire mitochondrial genome of 147 humans (CANN *et al.* 1987). Specific nucleotide changes were inferred from gains of restriction sites relative to the Cambridge sequence (ANDERSON *et al.* 1981). We have also compared the entire mtDNA sequence of three humans (ANDERSON *et al.* 1981; OZAWA *et al.* 1991; HORAI *et al.* 1995). Comparison of these polymorphism data with divergence data from published chimpanzee and gorilla sequences (HORAI *et al.* 1995) reveal an excess of amino acid polymorphisms throughout the human mitochondrial genome.

MATERIALS AND METHODS

Samples and DNA preparation: DNA samples for humans (*Homo sapiens*) and the single gorilla (*Gorilla gorilla*) were prepared and used in earlier studies as described by BROWN (1980) and STONEKING *et al.* (1990). The 61 human samples included 10 Negroid Africans, 10 Khoisan Africans, 10 Papua New Guineans, 4 African Americans, 23 Caucasians, and 4 Chinese. The five chimpanzees (*Pan troglodytes*) were wild caught from unknown locations in West Africa; tissue samples were obtained from North American primate research centers and DNA samples were kindly provided by T. MELTON. Three major subspecies of *Pan troglodytes* are currently recognized: *P.t. troglodytes, P.t. schweinfurthii,* and *P.t. verus.* In the wild, these subspecies are geographically isolated yet morphologically indistinguishable. The geographic origin of most captive chimpanzees in the U.S., including those in our sample, is unknown. However, the amount of sequence divergence between chimpanzee haplotypes reported below (see RESULTS) suggests that at least two subspecies are present among the five individuals (MORIN *et al.* 1994).

PCR amplification and DNA sequencing: A 563-bp fragment encompassing the entire *ND3* gene was PCR amplified (SAIKI *et al.* 1986) in all individuals. The following primers were used in all amplifications: L9986 (5′-GTATGTCTCCATCTATTGATGAGGG-3′) and H10500 (5′-TTCCTAGAAGTGAGATGGTAAATGC-3′), where L and H denote the light and heavy strand of the mitochondrial genome, and numbers correspond to the position of the 3′ base of each primer in the published human sequence (ANDERSON *et al.* 1981). DNA was amplified in 40 thermal cycles of 94° for 1 min, 55° for 1 min, and 72° for 3 min in 100 μl reaction volume with Taq polymerase (Perkin-Elmer-Cetus) with conditions as specified by the supplier. PCR products were precipitated with ammonium acetate and resuspended in 30 μl ddH_20 for direct double stranded di-deoxy sequencing (SANGER *et al.* 1977) as previously described (NACHMAN *et al.* 1994). Each PCR reaction provided sufficient product for two sequencing reactions. The *ND3* gene is 345 bases long, and both strands were sequenced. Sequencing primers were: L9986 (25 mer), L10023 (17 mer), H10500 (25 mer), H10353 (17 mer), H10354 (18 mer), and H10195 (19 mer), with primer designations as described above and primer sequences taken from ANDERSON *et al.* (1981). All primers worked on each species except H10353, which did not sequence chimpanzee DNA and was therefore substituted with H10354.

Analysis of *ND3* data: Sequences were aligned by hand and the number of replacement and silent nucleotide substitutions were counted. Nucleotide diversity (π) was calculated for each species (NEI and LI 1979). To test the neutral expectation that the ratio of replacement to silent substitutions should be the same within and between species (MCDONALD and KREITMAN 1991), replacement and silent polymorphisms and fixed differences were calculated at *ND3.* The use of fixed differences in this test has strong theoretical support (SAWYER and HARTL 1992). Within species, polymorphisms were counted as the number of segregating sites. Between species, differences were counted two different ways; namely as fixed differences between humans and chimpanzees and the total number of between-species substitutions occurring along the branches of a tree connecting humans, chimpanzees, and gorilla. Fisher's exact tests were used to test the null hypothesis of equal ratios of replacement to silent substitutions within and between species.

Analysis of RFLP and sequence data from the literature: Two different measures of mtDNA polymorphisms within humans were obtained from analyses of two different datasets. First, we reanalyzed the published RFLP data of CANN *et al.* (1987). Their survey included mtDNA from a worldwide sample of 147 humans, including 20 Africans, 34 Asians, 46 Caucasians, 21 aboriginal Australians, and 26 aboriginal New Guineans. MtDNA was purified, cut with 12 restriction enzymes (*Hpa*I, *Ava*II, *Fnu*DII, *Hha*I, *Hpa*II, *Mbo*I, *Taq*I, *Rsa*I, *Hin*fI, *Hae*III, *Alu*I, and *Dde*I), and subjected to high resolution mapping by comparison with the complete human mtDNA sequence (ANDERSON *et al.* 1981). Restriction sites in CANN *et al.* (1987, Figure 3) that were gained relative to the Cambridge sequence (ANDERSON *et al.* 1981) were used to infer specific nucleotide substitutions. Fifty-six nucleotide substitutions were inferred this way, lying in 12 of the 13 protein coding genes. To investigate departures from neutrality in the frequency distribution of polymorphisms within humans, we used the test of Tajima (1989). This was done for replacement and silent sites in African and non-African populations. For a second estimate of replacement and silent polymorphisms within humans, we compared the complete mtDNA sequences of three humans: a European (ANDERSON *et al.* 1981), a Japanese (OZAWA *et al.* 1991) and an African (HORAI *et al.* 1995). While this sample is very small ($n = 3$) compared with the RFLP sample ($n = 147$), the number of polymorphic sites observed among the three sequences (86 sites, see RESULTS) is actually greater than the number of polymorphic sites observed among the 147 restriction-mapped mtDNAs (56 sites, see RESULTS). Comparisons between human, chimpanzee, and gorilla for the entire mtDNA genome were done as described above for *ND3,* using published sequences (HORAI *et al.* 1995).

RESULTS

***ND3*:** The aligned nucleotide and amino acid *ND3* sequences for human, chimpanzee, and gorilla are shown in Figure 1. For humans and chimpanzees, consensus sequences are shown; for gorilla, the sequence from the single individual in our sample is shown.

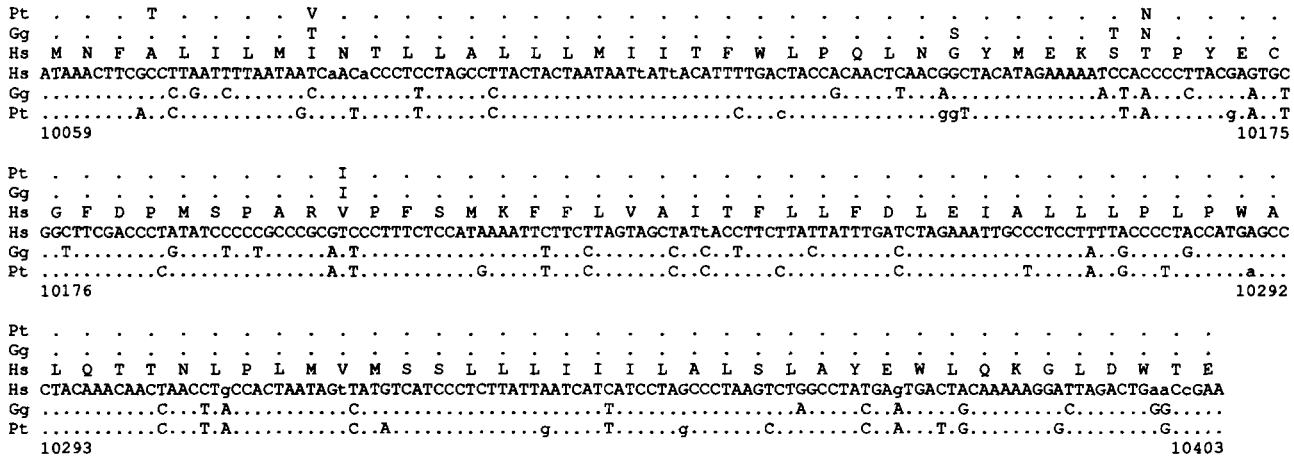

FIGURE 1.—Aligned nucleotide and amino acid sequences of human (Hs), gorilla (Gg), and chimpanzee (Pt). Single letter amino acid codes are used. Dots indicate identity; only differences are shown. The human and chimpanzee sequences are consensus sequences. Polymorphisms within species are indicated by lower case letters, and alternate states are shown in Tables 1 and 2. Coordinates are from the Cambridge sequence (ANDERSON *et al.* 1981).

Within humans (n = 61) we observed 14 haplotypes and 11 polymorphic sites (Table 1). Four of the 11 polymorphisms were at replacement sites, and nucleotide diversity for the entire sample was π = 0.39%. Phylogenetic analysis of these data with branch-and-bound searches using PAUP (SWOFFORD 1993) revealed 27 equally parsimonious unrooted networks, one of which is shown in Figure 2. There are two common haplotypes (1 and 2), distinguished by a single replacement substitution, and most other haplotypes are present at low frequencies. In the network shown, convergent substitutions occur at sites 10321 and 10373 (underlined). Other equally parsimonious trees include multiple substitutions at sites 10321, 10373, or 10398, and differ in the placement of haplotypes 7, 8, 9, 10, or 14. There is some geographic structure to the data, with one-half of the network (haplotypes 1, 3, 6, 10, 12, 14) deriving largely, though not exclusively, from non-African individuals, and the other half of the network (haplotypes 2, 4, 5, 7, 8, 9, 11, 13) deriving largely, though not exclusively, from African individuals. The fixation index (F_{ST}) calculated for African *vs.* non-African populations is F_{ST} = 0.185.

Within chimpanzees (n = 5) we observed three hap-

TABLE 1

***ND3* polymorphisms among 61 humans**

Haplotype	Sites[a] 00112333334 88113127990 69588013780 rrssssrssrs	Sample size (N)[b] Afr	Ind	Chi	Cau	Total
1	AATTTGTGAAC	2	1	0	18	21
2	.....G.	14	1	0	3	18
3	.....A.....	0	2	2	0	4
4	.....GT	0	1	1	1	3
5	..C.....G.	3	0	0	0	3
6	....C.....	0	3	0	0	3
7	.....C..G.	2	0	0	0	2
8	.....A.G.	1	0	0	0	1
9	G.....A.G.	1	0	0	0	1
10	.....C....	1	0	0	0	1
11	.....GGT	0	1	0	0	1
12	...C.....	0	1	0	0	1
13	.T.....GGT	0	0	1	0	1
14	.....A...	0	0	0	1	1

[a] Polymorphic sites are denoted by last three digits from the numbers in ANDERSON *et al.* (1981); *i.e.*, 086 denotes 10086. Replacement (r) and silent (s) sites are indicated.

[b] Sample sizes are given separately for Africans (Afr), Indonesians (Ind), Chinese (Chi), and Caucasians (Cau).

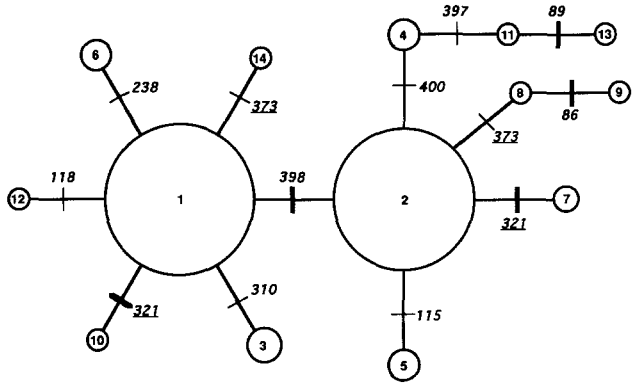

FIGURE 2.—One of 27 most-parsimonious unrooted networks of 14 human ND3 haplotypes. The length of this tree is 13 steps and the consistency index is CI = 0.846. Circles denote haplotypes, numbered 1–14 as in Table 1. The size of each circle is proportional to the frequency of the haplotype in the sample. Substitutions between haplotypes are indicated by a bold hash mark (replacement) or a fine hash mark (silent) and numbered as in Table 1. Parallel substitutions are underlined. Addition of outgroup sequences (chimpanzee or gorilla) does not provide additional resolution to this network, and still results in 27 equally parsiomious trees.

lotypes and seven polymorphic sites (Table 2). Four of the seven polymorphisms were at replacement sites, and nucleotide diversity for the entire sample was π = 0.93%. This high level of nucleotide diversity derives mainly from the presence of two very divergent haplotypes which differ at seven out of 345 sites (2.03%). This is considerably greater than the two most divergent human haplotypes, which differ at four out of 345 sites (1.16%), and is similar to the level of divergence reported for a small section of the mtDNA cytochrome *b* gene between *P.t. verus* and either *P.t. troglodytes* or *P.t. schweinfurthii* (2.8%) (MORIN *et al.* 1994). Divergence between *P.t. troglodytes* and *P.t. schweinfurthii* at cytochrome *b* is <0.5%. It is likely, therefore, that our sample includes *P.t. verus* and at least one of the two closely related subspecies, *P.t. troglodytes* and *P.t. schweinfurthii.* This has important implications for the interpretation of our results and will be discussed below.

The numbers of replacement and silent substitutions within humans, within chimpanzees, and between species are summarized in Table 3. Between species, there are approximately seven times as many silent changes as replacement changes. Within species, there are more nearly equal numbers of replacement and silent substitutions, in contrast to the predictions of the neutral theory. The null hypothesis is rejected when polymorphism data from both species are pooled together and compared with differences between species, as done by MCDONALD and KREITMAN (1991). The null hypothesis is also rejected when chimpanzee polymorphisms alone are compared with interspecific differences, but not when human polymorphisms alone are compared with interspecific differences (Table 3).

All mitochondrial genes: Table 4 shows the nucleotide substitutions inferred among 147 humans from the RFLP data of CANN *et al.* (1987). For each of the 13 protein-coding genes, Table 5 shows the number of replacement and silent polymorphisms within humans (for both the RFLP and the sequence data) and the number of replacement and silent differences among species. The results of the genome-wide MCDONALD-KREITMAN tests are given in Table 6. For the RFLP data, 23 of 56 substitutions (41.1%) were at replacement sites, and for the sequence data, 31 of 86 substitutions (36.0%) were at replacement sites. These proportions, which are similar to each other, are also similar to the proportion of replacement substitutions at the *ND3* gene alone (36.4%). In contrast, between humans and chimps, 179 of 1094 substitutions (16.4%) were at replacement sites, and among human, chimpanzee, and gorilla, 328 of 1881 (17.4%) were at replacement sites. MCDONALD-KREITMAN tests reject the null hypothesis of strict neutrality (Table 6). To see if this pattern is due to the contribution of one or a few unusual genes, we looked at subsets of the data. Patterns of replacement and silent substitution for the cytochrome oxidase genes and for the NADH dehydrogenase genes are shown separately in Table 6. In both cases, the null hypothesis is rejected, indicating that this pattern is widespread in the human mitochondrial genome. However, the ratios of replacement to silent substitution are approximately three times greater for the NADH dehydrogenase genes than for the cytochrome oxidase genes (both within and between species), consistent with a greater degree of selective constraint for the cytochrome oxidase genes. Other genes (*ATP, Cytb*) were not analyzed separately because of the small number of polymorphisms (see Table 5).

The frequencies of most (but not all) polymorphic sites in the RFLP data were low (Table 4). To see if the frequency distribution is consistent with a neutral, equilibrium model, we calculated TAJIMA's D statistic for various subsets of the data (Table 7). An excess of rare sites is observed in the non-African sample, but not in the African sample, consistent with earlier findings of nonequilibrium conditions in non-African human

TABLE 2

***ND3* polymorphisms among five chimpanzees**

Haplotype	Sites[a] 1111233 2447835 8340993 srrrssr	Sample size (*N*)
1	CGGGAGG	3
2	.....AA	1
3	TAAAGAA	1

[a] Polymorphic sites are denoted by last three digits from the numbers in ANDERSON *et al.* (1981) as in Table 1.

TABLE 3

Number of replacement and silent nucleotide differences within and between species at the *ND3* gene

	Within			Between	
	Humans	Chimps	Total	Human and chimp	Human, chimp, and gorilla
Replacement	4	4	8	4	7
Silent	7	3	10	31	45

Fisher's exact tests were used to test the null hypothesis that the ratio of replacement to silent differences is the same within and between species. The null hypothesis was rejected in tests comparing the total within-species variability with differences between humans and chimpanzees ($P < 0.02$) and to differences among human, chimpanzee, and gorilla ($P < 0.02$). The null hypothesis was also rejected when the within-species variability from chimpanzees alone was used ($P < 0.02$ for both tests), but not when the within-species variability from humans alone was used ($0.05 < P < 0.20$ for both tests).

mtDNA samples (WHITTAM *et al.* 1986; EXCOFFIER 1990; ROGERS and HARPENDING 1992). However, the African sample in the data of CANN *et al.* (1987) is much smaller (20) than the non-African sample (127); it is possible that the nonsignificant TAJIMA's D in the African sample is due to the lower power of the test with a smaller sample. The ratio of replacement to silent polymorphisms is nearly identical in African (7:10) and non-African (19:29) samples.

DISCUSSION

Amino acid polymorphisms in humans: The results from the *ND3* gene alone are generally consistent with the genome-wide results from the RFLP data of CANN *et al.* (1987) and the sequence data of HORAI *et al.* (1995). Although the *ND3* data (with human polymorphisms alone) do not show a significant departure from neutrality while the genome-wide data do, this difference is probably due to the fewer numbers of polymorphic sites in the *ND3* data compared to the genome-wide data. The overall ratios of replacement to silent substitutions within humans and between species are quite similar in all three datasets (compare Tables 3 and 6). In all cases, a greater number of amino acid substitutions is seen within humans than expected based on interspecific comparisons. What might account for this pattern?

One formal possibility is that some form of balancing or diversity-enhancing selection is maintaining amino acid variability within humans. Because each individual possesses effectively one mtDNA haplotype, heterosis is impossible. Selection would therefore have to vary either temporally or spatially to maintain variation. We consider this hypothesis unlikely because of the large number of low frequency polymorphisms. In the RFLP data, we observed 23 replacement polymorphisms within humans, while only about six would be expected based on interspecific comparisons. This is an excess of 17. Because <5% of the genome has been sampled, we estimate that there may be ~340 replacement polymorphisms in the entire mitochondrial genome that are not accounted for by interspecific comparisons. This is not meant to be an accurate estimate, but only to illustrate that the actual number of excess replacement polymorphisms is not small. It is difficult to imagine how selection could maintain polymorphisms at so many linked sites. Furthermore, diversity enhancing selection may be expected to maintain polymorphisms at moderate frequencies, yet we see just the opposite in the RFLP data. TAJIMA's D is negative for all subsets of the data, and is significantly negative for the non-African sample, indicating an excess of rare sites (Table 5).

A second explanation for the data is that there has been a recent and dramatic relaxation of selective constraint in the human lineage. Under this scenario, the excess amino acid polymorphisms seen within humans were formerly deleterious and are now neutral. Such a change in selective constraint would have to have occurred recently, because relatively few amino acid changes have become fixed in the human lineage. This hypothesis says that we sampled the population during the interval when sufficient time has elapsed since the relaxation of constraint so that replacement substitutions have risen to detectable frequencies but not yet become fixed. TAKAHATA (1993) has argued that the accumulation of formerly deleterious mutations in human populations since the Pleistocene may explain patterns of polymorphism at nuclear genes as well. Interestingly, a large survey of nuclear genes also reveals a high ratio (11:15) of replacement to silent polymorphisms (LI and SADLER 1991).

A third possibility is that amino acid substitutions are mildly deleterious (*e.g.*, OHTA 1972, 1992). Mutations with selection coefficients (s) in the neighborhood of $s = 1/N_e$, where N_e is the effective population size, will contribute more to heterozygosity than to substitution (*i.e.*, fixation), relative to strictly neutral mutations (KIMURA 1983). These mutations may appear within species as polymorphisms, but they are unlikely to become fixed and thus will rarely appear as differences between species. The behavior of slightly deleterious alleles has served as the basis for several models of molecular evolution (*e.g.*, KIMURA 1979; OHTA 1972, 1973, 1976, 1992; OHTA and TACHIDA 1990; TACHIDA 1991). One general criticism of these models is that they require a very narrow range of s and N_e to work (GILLESPIE 1991, 1994). If such models are to serve as general explana-

TABLE 4

Nucleotide substitutions inferred from RFLP data

Gene	Enzyme site[a]	Enzyme	Frequency (%)	Nucleotide site[b]	Nucleotide substitution	Silent or replacement	Amino acid substitution
ND1	3391	*Hae*III	2.0	3394	TAT-CAT	R	Tyr-His
	3592	*Hpa*I	8.2	3594	GTC-GTT	S	
	3842	*Hae*III	0.7	3843	TGA-TGG	S	
	4092	*Hin*fI	0.7	4095	ACC-ACT	S	
ND2	4643	*Rsa*I	1.4	4646	TAT-TAC	S	
	4732	*Rsa*I	1.4	4732	AAT-AGT	R	Asn-Ser
	4793	*Hae*III	1.4	4793	ATA-ATG	S	
	5351	*Hha*I	2.0	5351	CTA-CTG	S	
COI	5984	*Ava*II	0.7	5984	GGA-GGG	S	
	5985	*Rsa*I	4.8	5987	GTC-GTA	S	
	6166[c]	*Hha*I	0.7	6167	GGT-GGC	S	
	6356	*Dde*I	1.4	6357	CTA-TTA	S	
	6409[c]	*Taq*I	0.7	6411	AAT-GAT	R	Asn-Asp
	6501	*Hpa*II	0.7	6503	CCA-CCG	S	
	6610	*Hin*fI	0.7	6614	TTT-TTC	S	
	6915	*Rsa*I	1.4	6917	GTG-GTA	S	
	7025	*Alu*I	17.7	7028	GCC-GCT	S	
	7241	*Rsa*I	0.7	7241	GCA-GCG	S	
	7347	*Hae*III	0.7	7348	GTC-GGC	R	Val-Gly
COII	7617	*Hha*I	1.4	7617	GAC-GGC	R	Asp-Gly
	7970	*Hin*fI	4.8	7973	CCA-TCA	R	Pro-Ser
	8165	*Hae*III	0.7	8167	GGT-GGC	S	
ATP8	none						
ATP6	8852[c]	*Hha*I	1.4	8853	TGA-TGC	R	Trp-Cys
	9009	*Alu*I	1.4	9009	ACC-ACA	S	
	9070	*Taq*I	2.0	9072	TCA-TCG	S	
	9150	*Mbo*I	2.0	9150	TTA-TTG	S	
COIII	9429	*Rsa*I	1.4	9431	GTC-GTA	S	
	9714	*Hae*III	0.7	9716	GGT-GGC	S	
	9859	*Hin*fI	0.7	9860	TGC-TGA	R	Cys-Trp
ND3	10066	*Hha*I	0.7	10066	TTC-TGC	R	Phe-Cys
	10084	*Taq*I	0.7	10086	AAC-GAC	R	Asn-Asp
	10352	*Alu*I	0.7	10355	GCC-GCT	S	
	10394	*Dde*I	39.5	10398	ACC-GCC	R	Thr-Ala
ND4L	10644	*Rsa*I	1.4	10646	GTG-GTA	S	
	10694	*Alu*I	0.7	10697	GCC-GCT	S	
ND4	10806	*Hin*fI	6.8	10810	CTT-CTC	S	
	10893	*Taq*I	0.7	10895	AAC-GAC	R	Asn-Asp
	11146	*Dde*I	1.4	11149	TTG-TTA	S	
	11161	*Hpa*II	1.4	11164	CGA-CGG	S	
	11806	*Alu*I	2.0	11807	ACT-GCT	R	Thr-Ala
	12026	*Hpa*I	0.7	12026	ATT-GTT	R	Ile-Val
ND5	12345[c]	*Rsa*I	9.5	12346	CAC-TAC	R	His-Tyr
	12810	*Rsa*I	2.0	12810	TGA-TGG	S	
	12925	*Hin*fI	0.7	12928	CCA-TCA	R	Pro-Ser
	13068	*Alu*I	0.7	13071	GCC-GCT	S	
	13096	*Rsa*I	0.7	13099	GCA-CCA	R	Ala-Pro
	13100	*Hpa*II	0.7	13101	GCA-GCC	S	
ND6	14279	*Hae*III	0.7	14280	TCA-CCA	R	Ser-Pro
	14279	*Mbo*I	0.7	14281	GGG-GGA	S	
	14322	*Alu*I	0.7	14324	AAC-AGC	R	Asn-Ser
	14385	*Dde*I	0.7	14389	GGG-GGC	S	
	14509	*Alu*I	3.4	14510	GTT-GCT	R	Val-Ala
	14567	*Hpa*II	0.7	14570	AGC-ACC	R	Ser-Thr
CYTB	14749	*Hae*III	0.7	14750	ACC-GCC	R	Thr-Ala
	15005	*Hin*fI	1.4	15006	GCC-GAC	R	Ala-Asp
	15606	*Alu*I	10.9	15607	AAA-AAG	S	

Data inferred from CANN *et al.* 1987.

[a] Enzyme site is the position of the first nucleotide in the recognition sequence; site coordinates are from the Cambridge sequence (ANDERSON *et al.* 1981)

[b] Nucleotide site indicates the position of the inferred substitution.

[c] Denotes an ambiguous site gain; the other possible positions for the ambiguous replacement sites also result in replacement substitutions and the other possible positions for ambiguous silent sites also result in silent substitutions. Ambiguous sites that contained both silent and replacement substitutions in alternative positions were excluded.

TABLE 5

Number of replacement and silent nucleotide substitutions within and between species for each of the 13 mitochondrial protein coding genes

	Polymorphisms within humans[a]		Differences between species[a]	
	RFLP	Sequence	Human and chimp	Human, chimp and gorilla
ND1				
Replacement	1	2	18	27
Silent	3	5	71	123
ND2				
Replacement	1	3	11	30
Silent	3	4	86	148
COI				
Replacement	2	3	6	12
Silent	9	5	125	214
COII				
Replacement	2	0	5	9
Silent	1	2	59	102
ATP8				
Replacement	0	1	4	11
Silent	0	3	13	20
ATP6				
Replacement	1	3	13	26
Silent	3	3	45	87
COIII				
Replacement	1	1	7	12
Silent	2	4	66	115
ND3[b]				
Replacement	3	1	6	11
Silent	1	2	34	50
ND4L				
Replacement	0	0	1	3
Silent	2	4	21	33
ND4				
Replacement	3	2	23	41
Silent	3	10	107	184
ND5				
Replacement	3	5	51	91
Silent	3	7	143	248
ND6				
Replacement	4	5	7	11
Silent	2	1	43	69
CYTB				
Replacement	2	5	27	44
Silent	1	5	102	160

[a] All data taken from the literature as described in MATERIALS AND METHODS.

[b] The between-species numbers for *ND3* are not the same here and in Table 3 because these are fixed differences (see MATERIALS AND METHODS) and depend on the variation sampled within species.

tions of molecular evolution, then in species with larger N_e, s would have to be correspondingly smaller so that a substantial fraction of new mutations fall within the neighborhood of $s = 1/N_e$. In this regard, it may seem unlikely that a slightly deleterious model could explain the excess amino acid polymorphisms observed in species with such seemingly different levels of N_e as flies, humans, and mice. However, mitochondrial N_e may be roughly similar among these species (*i.e.*, within an order of magnitude) because levels of mtDNA variability are similar. Species-wide estimates of mitochondrial nucleotide diversity based on RFLP data are $\pi = 0.21\%$ for *Drosophila melanogaster* (HALE and SINGH 1987), $\pi = 0.77\%$ for *M. domesticus* (FERRIS *et al.* 1983), and $\pi = 0.32\%$ for *H. sapiens* (CANN *et al.* 1987), although mitochondrial mutation rates in Drosophila may be approximately one order of magnitude lower than in mammals (RAND 1994, and references therein). It is difficult to evaluate the fit of our data to models of slightly deleterious evolution because most published models have focused on rates and patterns of divergence rather than on levels and patterns of polymorphism or the relation-

TABLE 6

McDONALD-KREITMAN tests comparing replacement and silent nucleotide substitutions within and between species

	Polymorphisms within humans[a]		Differences between species[a]	
	RFLP	Sequence	Human and chimp	Human, chimp and gorilla
All genes[b]				
Replacement	23	31	179	328
Silent	33	55	915	1,553
CO genes[c]				
Replacement	5	4	18	33
Silent	12	11	250	431
ND genes[c]				
Replacement	15	18	117	214
Silent	17	33	505	855

[a] All data taken from the literature as described in MATERIALS AND METHODS.

[b] Fisher's exact tests were used to test the null hypothesis that the ratio of replacement to silent differences is the same within and between species. Four different tests were performed using all genes (RFLP *vs.* human and chimp; RFLP *vs.* human, chimp and gorilla; sequence *vs.* human and chimp; sequence *vs.* human, chimp and gorilla). All rejected the null hypothesis ($P \leq 0.00005$ for each).

[c] To see if the overall result is caused by the contribution of one or a few particularly deviant loci, we analyzed cytochrome oxidase (CO) genes and NADH dehydrogenase (ND) genes separately. The null hypothesis was also rejected for the eight tests involving each of these groups of genes ($P \leq 0.02$ for each test). *ATP* and *Cytb* are not shown separately because of the small number of polymorphisms at these loci (Table 5).

ship between polymorphism and divergence. Slightly deleterious models have previously been invoked as potential explanations for patterns of mtDNA evolution in Hawaiian Drosophila (DESALLE and TEMPLETON 1988), *D. melanogaster* and *D. simulans* (BALLARD and KREITMAN 1994), and *M. domesticus* (NACHMAN *et al.* 1994), although in none of these cases has there been an attempt to fit the data to a model in a quantitative manner. Computer simulations of slightly deleterious protein evolution in which patterns of replacement and silent polymorphism and divergence are considered would be valuable for evaluating existing datasets.

While we consider the first hypothesis (balancing selection) implausible, it is more difficult to distinguish between the other two hypotheses (relaxation of constraint and slightly deleterious mutations). It appears unlikely that the observed patterns derive from nonequilibrium conditions per se, such as changes in population size, although nonequilibrium conditions are known to characterize human populations (WHITTAM *et al.* 1986; EXCOFFIER 1990; DI RIENZO and WILSON 1991; ROGERS and HARPENDING 1992). This is because the test proposed by MCDONALD and KREITMAN is a test of neutrality and does not require populations to be at equilibrium (MCDONALD and KREITMAN 1991; SAWYER and HARTL 1992; BROOKFIELD and SHARP 1994). In fact, the African and non-African subsets of the RFLP data show nearly identical ratios of replacement to silent polymorphisms, although a neutral, equilibrium model is not rejected for the African data and is rejected for the non-African data (Table 7).

TABLE 7

TAJIMA'S D (1989) calculated from human RFLP

	TAJIMA'S D	*P* value
African sample		
Replacement	−1.446	>0.10
Silent	−0.526	>0.10
Non-African sample		
Replacement	−1.857	<0.05
Silent	−2.287	<0.01

Data calculated from CANN *et al.* 1987. African sample, $N = 20$; non-African sample, $N = 127$.

The pattern of excess replacement polymorphism shown here for humans is qualitatively similar to that seen in mtDNA of mice (NACHMAN *et al.* 1994) and flies (KANEKO *et al.* 1993; BALLARD and KREITMAN 1994; RAND *et al.* 1994), and raises the question of whether a general mechanism can explain this result in these different species. *M. domesticus, D. melanogaster,* and *D. simulans* are all commensal with humans and it is unclear how this may affect such things as changes in selection pressures over time. It will be useful to study patterns of mtDNA polymorphism and divergence in species that are not associated with humans and whose habitats and ranges are thought to have been stable for several million years. Such species may be less likely to have experienced a recent relaxation of constraint.

Slightly deleterious mutations and human disease: The slightly deleterious model proposed here is consistent with some recent views on human mitochondrial diseases (WALLACE 1992a,b, 1994; LERTRIT *et al.* 1994). A variety of defects in oxidative phosphorylation have

TABLE 8

Ratio of replacement to silent substitutions among healthy and disease individuals

	Polymorphisms within humans		Divergence
	Disease[a]	Healthy[b]	Human-chimp
Replacement (R)	91	31	179
Silent (S)	99	55	915
Ratio (R/S)	0.92	0.56	0.20

[a] Disease sequences are from the literature: $N = 10$ from MARZUKI *et al.* (1991), $N = 9$ from OZAWA *et al.* (1991), $N = 2$ from BROWN *et al.* (1992a), $N = 1$ from BROWN *et al.* (1992b), $N = 1$ from WALLACE *et al.* (1988), $N = 1$ from YONEDA *et al.* (1990), $N = 1$ from KOBAYASHI *et al.* (1991).

[b] Healthy sequences are from the literature as described in the MATERIALS AND METHODS.

$N = 25$ for disease sequences; $N = 3$ for healthy sequences.

been associated with mtDNA mutations. Many of these mutations are thought to be only mildly or moderately deleterious and are associated with adult-onset or degenerative diseases such as Leber's hereditary optic neuropathy (LHON). Because symptoms can appear relatively late in life, the reproductive output of individuals carrying such mutations may not be strongly affected (WALLACE 1994). There are ~20 mtDNA missense mutations that have been associated with human disease (WALLACE *et al.* 1993; LERTRIT *et al.* 1994), although there are conflicting opinions on the relative pathogenicity of a number of these mutations (WALLACE *et al.* 1993). In general, identification of disease missense mutations has been based on their association with disease phenotypes and relative absence in control individuals, although some putative disease mutations appear at low frequencies in control samples (WALLACE *et al.* 1993). We looked for disease missense mutations among the replacement polymorphisms identified from our analysis of the RFLP and sequence data. Two of the 20 appear in these data. The T-C substitution at site 3394 in the *ND1* gene identified in the RFLP data (Table 4) replaces a tyrosine with a histidine and is associated with LHON (BROWN *et al.* 1992a; JOHNS *et al.* 1992). The G-A substitution at site 5460 in the *ND2* gene replaces an alanine with a threonine and differs between the Cambridge sequence (ANDERSON *et al.* 1981) and the other two complete human mtDNA sequences (OZAWA *et al.* 1991; HORAI *et al.* 1995). This mutation was found at high frequency in a sample of patients with Alzheimer's disease (LIN *et al.* 1992), although a subsequent study failed to confirm this association (PETRUZZELLA *et al.* 1992).

To see if the presence of deleterious mutations may inflate the ratio of replacement to silent substitutions, we have compared 25 complete or nearly complete human mtDNA sequences among individuals with disease phenotypes (Table 8). Each of these sequences includes >90% of the protein-coding regions in the human mitochondrial genome, although the missing portions are not the same among the different sequences. The sample includes a heterogeneous mixture of disease phenotypes including LHON, MERRF (myoclonic epilepsy with ragged red fibers), MELAS (mitochondrial encephalomyopathy, lactic acidosis, stroke-like episodes), and cardiomyopathies. The ratio of replacement to silent substitutions (0.92) among the individuals in this sample is greater than the corresponding ratio among healthy individuals (0.56), and this difference is marginally significant (Fisher's exact test, $P = 0.05$). This suggests that a number of mutations of small effect may contribute to disease phenotypes. It also illustrates that the ratio of replacement to silent polymorphisms may be increased by the presence of deleterious mutations.

Amino acid polymorphisms in chimpanzees: The amount of sequence divergence observed between chimpanzee haplotypes 1 and 3 (Table 2) is 2.03%. It is clear that this level of divergence corresponds to that seen between allopatric subspecies from West and Central Africa (MORIN *et al.* 1994). Thus, the ratio of replacement to silent substitutions observed within chimpanzees (Table 3) does not reflect variation from a single panmictic population. The differences between haplotypes 1 and 3 almost certainly accrued in allopatry (MORIN *et al.* 1994). It is therefore probable that different forces have produced high ratios of replacement to silent substitutions within humans and within chimpanzees. In particular, it is difficult to envision how slightly deleterious mutations could produce the pattern seen in chimpanzees.

Even though the chimpanzee sample does not represent a single population, the observed variation is inconsistent with strict neutrality. As pointed out by BALLARD and KREITMAN (1994), one can construct a MCDONALD-KREITMAN type of test for any partitioning of a neutral genealogy. The observed distribution of replacement and silent variation within chimpanzees (but between subspecies) thus requires some explanation.

Two alternative hypotheses may explain these data. First, the higher ratio of replacement to silent substitution within chimpanzees than between species may reflect the fixation of adaptive substitutions during the evolution of chimpanzee subspecies. This hypothesis predicts that within subspecies, the ratio of replacement to silent changes should be smaller than between subspecies. Second, the ratio seen between subspecies may reflect a relaxation of selective constraint that occurred before the splitting of subspecies, but long after the separation of humans and chimpanzees. This hypothesis predicts that within subspecies, the ratio of replacement to silent changes should be the same as between subspecies. Distinguishing between these hypotheses will require further sampling.

Implications for mtDNA as a marker: During the last 10 years, there has been an accumulation of data on

mtDNA variation in flies, mice, and humans suggesting that the evolution of this molecule does not fit the patterns predicted by a strictly neutral model of evolution (*e.g.*, WHITTAM *et al.* 1986; DESALLE and TEMPLETON 1988; EXCOFFIER 1990; MERRIWETHER *et al.* 1991; BALLARD and KREITMAN 1994; NACHMAN *et al.* 1994; RAND *et al.* 1994). In some situations it appears that nonequilibrium conditions may be partly responsible for the patterns (*e.g.*, WHITTAM *et al.* 1986) while in others it appears that some form of natural selection is at least partly responsible (*e.g.*, BALLARD and KREITMAN 1994). Regardless of the cause of the observed deviations, these patterns have the potential to affect the use of mtDNA as a neutral marker in evolutionary and population level studies. While mtDNA will undoubtedly continue to be useful in many situations, it is reasonable to test for neutrality before assuming, for example, the clock-like behavior of a mitochondrial gene in any particular situation. Such tests not only provide information about the forces acting during the evolutionary history of a locus, but they also allow us to improve our estimates of parameters of interest. For example, the age of the most recent human mitochondrial ancestor proposed by CANN *et al.* (1987) was based on comparison of human and chimpanzee mtDNA sequences and the assumed clock-like behavior of mtDNA. The results presented here indicate that a substantial proportion of the replacement substitutions found within the human lineage never become fixed. This will lead to an overestimate of the age of the human mtDNA ancestor. In particular, the interspecific ratio of replacement to silent substitutions in the RFLP data (179:915) can be used to estimate the number of replacement substitutions within humans expected under a strictly neutral model (6.5). The difference between the observed value (23) and the expected value (6.5) represents polymorphisms that are not accounted for by the interspecific molecular clock and accounts for 29.5% of the total variability observed within humans. Thus, the age of the human mtDNA ancestor may be overestimated by ~30%.

We thank H. SOODYALL and T. MELTON who helped prepare some of the DNA. We thank D. J. BEGUN, M. J. FORD, M. T. HAMBLIN, C. MORITZ, K. S. PHILLIPS, and T. G. SCHURR for comments or discussions. This research was supported by a National Institutes of Health postdoctoral fellowship and National Science Foundation grant to M.W.N. and a NIH grant to C.F.A.

LITERATURE CITED

ANDERSON, S., A. T. BANKIER, B. G. BARRELL, M. H. L. DEBRUIJN, A. R. COULSON *et al.*, 1981 Sequence and organization of the human mitochondrial genome. Nature **290:** 457–465.

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB *et al.*, 1987 Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. **18:** 489–522.

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

BROOKFIELD, J. F. Y., and P. M. SHARP, 1994 Neutralism and selection face up to DNA data. Trends Genet. **10:** 109–111.

BROWN, M. D., A. S. VOLJAVEC, M. T. LOTT, I. MACDONALD and D. C. WALLACE, 1992a Leber's hereditary optic neuropathy: a model for mitochondrial neurodegenerative diseases. FASEB J. **6:** 2791–2799.

BROWN, M. D., A. S. VOLJAVEC, M. T. LOTT, A. TORRONI, C.-C. YANG *et al.*, 1992b Mitochondrial DNA complex I and III mutations associated with Leber's Hereditary Optic Neuropathy. Genetics **130:** 163–173.

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

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

DESALLE, R., and A. R. TEMPLETON, 1988 Founder effects and the rate of mitochondrial DNA evolution in Hawaiian *Drosophila*. Evolution **42:** 1076–1084.

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

EXCOFFIER, L., 1990 Evolution of human mitochondrial DNA: evidence for departure from a pure neutral model of populations at equilibrium. J. Mol. Evol. **30:** 125–139.

FERRIS, S. D., R. D. SAGE, E. M. PRAGER, U. RITTE and A. C. WILSON, 1983 Mitochondrial DNA evolution in mice. Genetics **105:** 681–721.

GILLESPIE, J. H., 1991 *The Causes of Molecular Evolution.* Oxford University Press, Oxford.

GILLESPIE, J. H., 1994 Substitution processes in molecular evolution. III. Deleterious alleles. Genetics **138:** 943–952.

HALE, L. R., and R. S. SINGH, 1987 Mitochondrial DNA variation and genetic structure in populations of *Drosophila melanogaster*. Mol. Biol. Evol. **4:** 622–637.

HORAI, S., K. HAYASAKA, R. KONDO, K. TSUGANE and N. TAKAHATA, 1995 Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc. Natl. Acad. Sci. USA **92:** 532–536.

HUDSON, R. R., M. KREITMAN and M. AGUADE, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

JOHNS, D. R., M. J. NEUFELD and R. D. PARK, 1992 An ND-6 mitochondrial DNA mutation associated with Leber Hereditary Optic Neuropathy. Biochem. Biophys. Res. Comm. **187:** 1551–1557.

KANEKO, M., Y. SATTA, E. T. MATSURA and S. CHIGUSA, 1993 Evolution of the mitochondrial ATPase 6 gene in *Drosophila*: unusually high level of polymorphism in *D. melanogaster*. Genet. Res. **61:** 195–204.

KIMURA, M., 1979 Model of effectively neutral mutations in which selective constraint is incorporated. Proc. Natl. Acad. Sci. USA **76:** 3440–3444.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

KOBAYASHI, Y., M. Y. MOMOI, K. TOMINAGA, H. SHIMOIZUMI, K. NIHEI *et al.*, 1991 Respiration-deficient cells are caused by a single point mutation in the mitochondrial tRNA-Leu (UUR) gene in mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes (MELAS). Am. J. Hum. Genet. **49:** 590–599.

LERTRIT, P., R. M. I. KAPSA, M. J. B. JEAN-FRANCOIS, D. THYAGARAJAN, A. S. NOER *et al.*, 1994 Mitochondrial DNA polymorphism in disease: a possible contributor to respiratory dysfunction. Hum. Mol. Genet. **3:** 1973–1981.

LI, W.-H., and L. A. SADLER, 1991 Low nucleotide diversity in man. Genetics **129:** 513–523.

LIN, F.-H., R. LIN, H. M. WISNIEWSKI, Y.-W. HWANG, I. GRUNDKE-IQBAL *et al.*, 1992 Detection of point mutations in codon 331 of mitochondrial NADH dehydrogenase subunit 2 in Alzheimer's brains. Biochem. Biophys. Res. Comm. **182:** 238–246.

MARZUKI, S., A. S. NOER, P. LERTRIT, D. THYAGARAJAN, R. KAPSA *et al.*, 1991 Normal variants of human mitochondrial DNA and translation products: the building of a reference data base. Hum Genet. **88:** 139–145.

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

MERRIWETHER, D. A., A. G. CLARK, S. W. BALLINGER, T. G. SCHURR,

H. SOODYALL *et al.*, 1991 The structure of human mitochondrial DNA variation. J. Mol. Evol. **33:** 543–555.

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

NACHMAN, M. W., S. N. BOYER and C. F. AQUADRO, 1994 Nonneutral evolution at the mitochondrial NADH dehydrogenase subunit 3 gene in mice. Proc. Natl. Acad. Sci. USA **91:** 6364–6368.

NEI, M., and W.-H. LI, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76:** 5269–5273.

OHTA, T., 1972 Population size and rate of evolution. J. Mol. Evol. **1:** 305–314.

OHTA, T., 1973 Slightly deleterious mutant substitutions in evolution. Nature **246:** 96–98.

OHTA, T., 1976 Role of very slightly deleterious mutations in molecular evolution and polymorphism. Theor. Popul. Biol. **10:** 254–275.

OHTA, T., 1992 The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. **23:** 263–286.

OHTA, T., and H. TACHIDA, 1990 Theoretical study of near neutrality. I. Heterozygosity and rate of mutant substitution. Genetics **126:** 219–229.

OZAWA, T., M. TANAKA, S. SUGIYAMA, H. INO, K. OHNO *et al.*, 1991 Patients with idiopathic cardiomyopathy belong to the same mitochondrial DNA gene family of Parkinson's disease and mitochondrial encephalomyopathy. Biochem. Biophys. Res. Comm. **177:** 518–525.

PETRUZZELLA, V., X. CHEN and E. A. SCHON, 1992 Is a point mutation in the mitochondrial ND2 gene associated with Alzheimer's disease? Biochem. Biophys. Res. Comm. **186:** 491–497.

RAND, D. M., 1994 Thermal habit, metabolic rate and the evolution of mitochondrial DNA. Trends Ecol. Evol. **9:** 125–131.

RAND, D. M., M. DORFSMAN and L. M KANN, 1994 Neutral and nonneutral evolution of Drosophila mitochondrial DNA. Genetics **138:** 741–756.

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

SAIKI, R. K., T. L. BUGAWAN, G. T. HORN, K. B. MULLIS and H. A. ERLICH, 1986 Analysis of enzymatically amplified Beta-globin and HLA-DQalpha DNA with allele-specific oligonucleotide probes. Nature **324:** 163–166.

SANGER, F., S. NICKLON and A. R. COULSON, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SAWYER, S. A., and D. L. HARTL, 1992 Population genetics of polymorphism and divergence. Genetics **132:** 1161–1176.

SLATKIN, M., 1985 Gene flow in natural populations. Annu. Rev. Ecol. Syst. **16:** 393–430.

STONEKING, M., L. B. JORDE, K. BHATIA and A. C. WILSON, 1990 Geographic variation in human mitochondrial DNA from Papua New Guinea. Genetics **124:** 717–733.

SWOFFORD, D. L., 1993 PAUP (Phylogenetic analysis using parsimony) Illinois Natural History Survey, Champaign, IL.

TACHIDA, H., 1991 A study on a nearly neutral mutation model in finite populations. Genetics **128:** 183–192.

TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.

TAKAHATA, N., 1993 Relaxed natural selection in human populations during the Pleistocene. Jpn. J. Genet. **68:** 539–547.

WALLACE, D. C., 1992a Mitochondrial genetics: a paradigm for aging and degenerative diseases? Science **256:** 628–632.

WALLACE, D. C., 1992b Diseases of the mitochondrial DNA. Annu. Rev. Biochem. **61:** 1175–1212.

WALLACE, D. C., 1994 Mitochondrial DNA mutations in diseases of energy metabolism. J. Bioenerg. Biomembr. **26:** 241–250.

WALLACE, D. C., G. SINGH, M. T. LOTT, J. A. HODGE, T. G. SCHURR *et al.*, 1988 Mitochondrial DNA mutation associated with Leber's Hereditary Optic Neuropathy. Science **242:** 1427–1430.

WALLACE, D. C., M. T. LOTT, A. TORRONI and M. D. BROWN, 1993 Report of the committee on human mitochondrial DNA. Genome Priority Rep. **1:** 727–757.

WATTERSON, G. A., 1978 The homozygosity test of neutrality. Genetics **88:** 405–417.

WHITTAM, T. S., A. G. CLARK, M. STONEKING, R. L. CANN and A. C. WILSON, 1986 Allelic variation in human mitochondrial genes based on patterns of restriction site polymorphism. Proc. Natl. Acad. Sci. USA **83:** 9611–9615.

WILSON, A. C., R. L. CANN, S. M. CARR, M. GEORGE, U. B. GYLLENSTEN *et al.*, 1985 Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. **26:** 375–400.

YONEDA, M., Y. TANNO, S. HORAI, T. OZAWA, T. MIYATAKE *et al.*, 1990 A common mitochondrial DNA mutation in the t-RNA^{Lys} of patients with myoclonus epilepsy associated with ragged-red fibers. Biochem. Int. **21:** 789–796.

Communicating editor: A. G. CLARK