Am. J. Hum. Genet. 59:501–509, 1996

How Rapidly Does the Human Mitochondrial Genome Evolve?

Neil Howell,[1] Iwona Kubacka,[1] and David A. Mackey[2]

[1]Department of Radiation Therapy, Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch, Galveston; and [2]Departments of Ophthalmology and Paediatrics, The University of Melbourne, Melbourne, and Menzies Centre for Population Health Research, The University of Tasmania, Launceston

Summary

The results of an empirical nucleotide-sequencing approach indicate that the evolution of the human mitochondrial noncoding D-loop is both more rapid and more complex than is revealed by standard phylogenetic approaches. The nucleotide sequence of the D-loop region of the mitochondrial genome was determined for 45 members of a large matrilineal Leber hereditary optic neuropathy pedigree. Two germ-line mutations have arisen in members of one branch of the family, thereby leading to triplasmic descendants with three mitochondrial genotypes. Segregation toward the homoplasmic state can occur within a single generation in some of these descendants, a result that suggests rapid fixation of mitochondrial mutations as a result of developmental bottlenecking. However, slow segregation was observed in other offspring, and therefore no single or simple pattern of segregation can be generalized from the available data. Evidence for rare mtDNA recombination within the D-loop was obtained for one family member. In addition to these germ-line mutations, a somatic mutation was found in the D-loop of one family member. When this genealogical approach was applied to the nucleotide sequences of mitochondrial coding regions, the results again indicated a very rapid rate of evolution.

Introduction

Sequence analysis of the human mitochondrial genome (mtDNA) has been used by evolutionary and systematic biologists to estimate the geographical origin, early population dynamics, and the time since the most recent common ancestor of modern humans (Cann et al. 1987; Vigilant et al. 1991; Templeton 1993; Penny et al. 1995). A major limitation to further progress in this area is the sketchy understanding of the complex processes through which mtDNA mutations arise, segregate, and are fixed at the levels of the organelle, the cell, the individual, and the population.

Phylogenetic analysis (also termed the "relative branch length" method; see Ruvolo 1996) is the most commonly used approach for estimating the rate of human mtDNA evolution. The mtDNA nucleotide sequences of a sample of individuals are used to construct a phylogenetic tree by using one of the many available methods (maximum parsimony, neighbor joining, or maximum likelihood). The number of mutations that have occurred along each branch of the tree is then estimated. Finally, some independent assumption about the age of the tree (e.g., the time of the human-chimpanzee divergence) is applied to derive a rate of divergence. When this phylogenetic approach has been applied to the noncoding control region (D-loop) of human mtDNA or to one or both of the two hypervariable regions therein (designated "I" and "II"), divergence rates in the range of 7%–22%/million years (Myr) have been obtained (Pesole et al. 1992; Stoneking et al. 1992; Tamura and Nei 1993; Horai et al. 1995). In contrast, divergence rates for the coding region are much lower, ~2%–4%/Myr (Kocher and Wilson 1991; Torroni et al. 1994*b*). One limitation to this approach is that, particularly for the D-loop, there are hypervariable sites that confound accurate estimates of tree topology and branch lengths (Wakeley 1993; Penny et al. 1995; Yang 1995).

This range of mtDNA divergence rates introduces a troubling degree of uncertainty into the question of the time since the origin of modern humans, with most estimates in the range of 0.2–0.6 Myr (e.g., see Penny et al. 1995; Wills 1995; Ruvolo 1996). The issue becomes even murkier when alternative approaches have been used. Thus, Lundstrom et al. (1992) used a finite-sites coalescent model to derive an independent assessment of the rate of mtDNA evolution. Coalescent models assume that the extant alleles of a gene (e.g., the mtDNA is treated as a single "gene") have descended from a single allele; the coalescent time is related to the frequency of nucleotide heterozygosity and population size (females in the case of mtDNA). When their model was applied to the D-loop hypervariable region I sequence data of an Amerindian tribal sample (Ward et al. 1991), they

Received April 10, 1996; accepted for publication June 17, 1996.
Address for correspondence and reprints: Dr. Neil Howell, Biology Division 0656, Department of Radiation Therapy, The University of Texas Medical Branch, Galveston, TX 77555-0656. E-mail: nhowell@mspo3.med.utmb.edu
© 1996 by The American Society of Human Genetics. All rights reserved.
0002-9297/96/5903-0004$02.00

calculated a divergence rate of 110%/Myr, which is much higher than that obtained with phylogenetic approaches. Although their method did not require an external assumption about the age of the mtDNA lineage, Lundstrom et al. did have to estimate the size of the female breeding population.

These discrepancies underscore the importance of developing novel experimental strategies to determine the rate of mtDNA evolution that are based on minimal assumptions. We describe here the use of one such empirical approach for the estimation of the rate of human mtDNA evolution in both the D-loop and coding regions and for the analysis of the origin and fixation of mutations within the mitochondrial genome.

Experimental Procedures

Tissue Samples

Total DNA was isolated, by standard procedures, from the white-blood-cell (WBC)/platelet fraction ("buffy coat") of venous blood, from the platelet-enriched fraction of venous blood, or from fibroblast cultures that had been established by skin biopsy of one family member (individual X-18). PCR amplification was carried out directly with extracts of dried blood spots, for some family members.

Nucleotide Sequence Analysis

The D-loop of the circular human mitochondrial genome spans nucleotides 16024–16569 and 1–648 (numbered according to Anderson et al. 1981); the two halves are contiguous. Two approaches were used to obtain the nucleotide sequence of the D-loop. In the first, which is a minor modification of that described elsewhere (Howell et al. 1995), the D-loop was amplified as two separate fragments that span, respectively, nucleotides 15909–16569 and 1–657. These two D-loop fragments were cloned separately into M13 vectors, and the nucleotide sequence was determined by using a standard dideoxy chain–termination method. Because the two subfragments insert randomly into M13, the nucleotide sequence of ~350 bp was determined from each "end," thereby allowing determination of the sequence of the entire fragment. To confirm the sequence in the middle of these two D-loop regions, internal sequencing primers were used. For each orientation of insertion, the sequences of an average of five or six independent M13 subclones were determined.

It was difficult to PCR-amplify such large fragments of the mitochondrial genome (~660 bp each) from dried blood spots. Therefore, a second strategy was used that involved PCR amplification of the D-loop as a series of four overlapping fragments of ~350 bp each that subsequently were used for M13 cloning. In addition to facilitating PCR amplification, this approach allows determination of the sequence for both strands of the entire D-loop region. For each of these four fragments, 10–12 independent clones were sequenced. Control experiments in which both approaches were used with the same DNA sample yielded identical nucleotide sequences (data not shown).

The nucleotide-sequencing analysis of the coding region has been described elsewhere (Howell et al. 1995). For the USA1 Leber hereditary optic neuropathy (LHON) family (see Results), we have concentrated on the 300-bp mtDNA region that brackets the primary LHON mutation at nucleotide 14484.

Results

Experimental Rationale

The divergence rate that was derived by Lundstrom et al. (1992) predicts that 14 of every 1 million children will differ from their mother at a particular base pair within hypervariable region I. Because Lundstrom et al. sequenced a total span of 360 bp, it may be further predicted that ~1/200 children will differ from their mother at some site within this region. Kocher and Wilson (1991) observed that 63% of all mutations in their sample of D-loop sequences occurred within a span of 400 bp that includes hypervariable region I. Therefore, we estimate that ~1/125 children should differ from their mother at some site within the entire 1.2-kb D-loop; in other words, one mtDNA D-loop mutation should arise every 125 generations if the divergence rate derived by Lundstrom et al. (1992) is accurate. In contrast, the predicted ratio drops to ~1 child/400–1,200 for the slower rates obtained with phylogenetic approaches (Pesole et al. 1992; Stoneking et al. 1992; Tamura and Nei 1993; Horai et al. 1995). One could empirically test these estimates by D-loop sequence analysis for a very large number of mother-child pairs (see Discussion), but we have used a complementary approach that involves the sequencing analysis of the members of a matrilineal pedigree. By analyzing a *single* lineage of mitochondrial genomes, we "know" the mother's mtDNA sequence for all family members, and we need only ascertain those descendants who have accumulated new mutations within the D-loop. This approach is feasible, however, only because of the availability of large, well-characterized LHON matrilineal pedigrees.

The TAS2 LHON Matrilineal Lineage

The TAS2 Australian LHON family spans 12 generations and comprises >700 maternally related individuals whose genealogy can be unambiguously traced back to a woman born in 1714 in London (Mackey and Buttery 1992). A partial pedigree of the TAS2 LHON family is shown in figure 1. This family is homoplasmic for the primary LHON mutation at nucleotide 14484 (Mackey

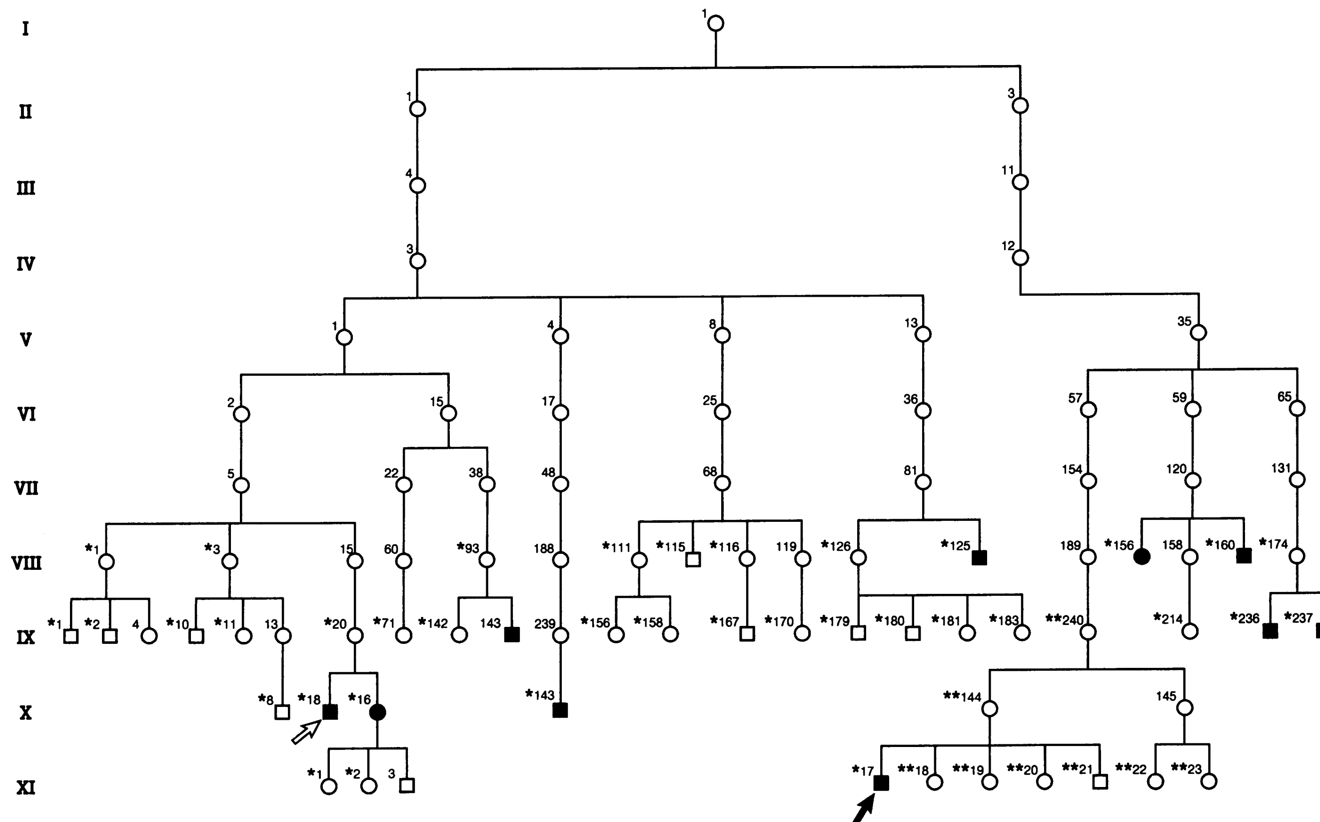

Figure 1 Partial matrilineal pedigree of the TAS2 LHON family. Each family member within a generation is assigned an identification number, and these are shown to the left of the symbols. The blackened symbols denote those individuals affected with the optic neuropathy. The single asterisks indicate individuals whose D-loop sequences were determined in the initial sequencing analysis. The double asterisks indicate the maternal relatives of family member X1-17 (denoted by the blackened arrow) who were subsequently analyzed (table 1). The unblackened arrow indicates family member X-18, who was found to have a somatic D-loop mutation.

and Howell 1992), a point to be discussed later. The TAS2 D-loop sequence has been reported recently (Howell et al. 1995), and it was observed that, relative to the Cambridge reference sequence (Anderson et al. 1981), there were nucleotide substitutions at positions 73, 146, 185, 188, 222, 228, 263, 295, 482, 489, 16069, 16126, 16261, and 16519; this constellation of polymorphisms is defined here as the "TAS2 D-loop haplotype."

The TAS2 mitochondrial genome is a member of a phylogenetic cluster in which the mtDNAs carry the putative secondary LHON mutations at nucleotides 4216 and 13708 (Howell et al. 1995); this cluster is also the European haplotype J that was analyzed by Wallace and coworkers (Torroni et al. 1994*a*). We have sequenced the D-loop region from >40 control and LHON pedigrees that are members of this phylogenetic cluster. The closest "relative" to the TAS2 haplotype (namely, other mitochondrial haplotypes that belong to this phylogenetic cluster) differs by three polymorphisms in the D-loop (data not shown). The fact that the TAS2 haplotype is thus far unique to this LHON matrilineal pedigree gives us unalloyed confidence in the genealogy and in our ability to rule out contamination or mix-up in the DNA samples. In addition to the Australian TAS2 family, we have found three small 14484 English LHON families who also carry the TAS2 D-loop haplotype (data not shown). We are attempting further genealogical studies to link these families to the main TAS2 pedigree. Alternatively, they may represent the descendants of earlier (i.e., pre-1714) branches.

New D-Loop Mutations in the TAS2 LHON Family

In the first analyses, the entire 1.2-kb D-loop sequence was determined for 36 members of the TAS2 pedigree (fig. 1) and for four members of the three English 14484 LHON pedigrees who have the TAS2 haplotype. For these 40 individuals, we analyzed DNA that had been isolated from venous blood samples or from dried blood spots. Furthermore, we sampled widely spaced branches within the TAS2 genealogical tree space, and we were thus able to compare individuals whose most recent common maternal ancestor was the female born in 1714 (individual I-1; fig. 1). No new D-loop mutations were

Table 1

Germ-Line Origin and Transmission of D-Loop Mutations in a Branch of the TAS2 LHON Family

Family Member (No. of Independent Clones)	Frequency of Mitochondrial Genotypes[a] (%)		
	+/+	152/+	+/195
IX-240 (72)	40	33	26
X-144 (102)[b]	14	53	33
XI-17:			
Dried blood spot (58)	16	74	10
Platelet-rich fraction of whole blood (34)	18	65	18
XI-18 (30)	0	7	93
XI-19 (35)	0	57	43
X-20 (32)	13	41	47
X-21 (32)	100	0	0
X-22 (35)	34	6	60
X-23 (35)	94	3	3

[a] +/+ denotes the TAS2 haplotype; and 152/+ and +/195 denote the haplotypes that carry new D-loop mutations.

[b] An additional clone had a recombinant genotype and carried *both* the 152 and 195 substitutions.

detected in the mitochondrial genomes of 39 of these 40 individuals. However, in one TAS2 family member (individual XI-17, indicated by the blackened arrow in fig. 1), *two* new D-loop mutations were detected. Among 58 clones that spanned the appropriate region, 9 (16%) carried the TAS2 haplotype. However, 43 clones from this individual had the TAS2 haplotype plus an additional TA→CG transition at nucleotide position 152, and 6 clones (10%) had the TAS2 haplotype plus a TA→CG transition at nucleotide 195 (table 1). That is, this individual is *triplasmic* and carries three different mtDNA genotypes. It must be emphasized that the region of the D-loop that was analyzed carries a unique constellation of eight substitutions relative to the Cambridge reference: contamination with non-TAS2 sequences cannot explain these findings, because the 58 analyzed clones all contained the complete panoply of TAS2 substitutions.

In addition to the blood samples that were obtained from the TAS2 family members, we also established a fibroblast culture from one of them (individual X-18, indicated by the unblackened arrow in fig. 1). Analysis of mtDNA D-loop sequences from these fibroblasts revealed the presence of a *third* new mutation. Of the 24 clones that were sequenced (pooled results from two different cultures), all carried the TAS2 haplotype but 13 also carried a TA:GC transversion at nucleotide 204. Again, contamination or sample mix-up can be eliminated as an explanation. It appears that the origin of this D-loop mutation is somatic and that it did not arise within the germ line, because the transversion was not detected among the 19 clones of WBC/platelet DNA that were analyzed for this individual. Furthermore, this mutation was not detected in the WBC/platelet DNA of his mother, his sister, or his two nieces (fig. 1).

Follow-Up Studies of the XI-17 TAS2 Branch

It was important to ascertain whether the new D-loop mutations at nucleotides 152 and 195 in family member XI-17 had arisen within the germ line and, if so, to "track" their transmission. To that end, DNA was obtained from his mother, his maternal grandmother, his three sisters, his brother, and two maternal first cousins (fig. 1). Furthermore, an additional, independent DNA sample was obtained from family member XI-17; the same results were obtained as in the first sequencing analyses (table 1).

The complete D-loop was sequenced for these additional eight TAS2 family members; and the relevant results are summarized in table 1. Several interesting results emerged. In the first place, it is clear that *both* mutations arose within the germ line, because their transmission through three generations can be established. Second, the proportions of the three genotypes within branch members provide further evidence for the rapid drift, toward the homoplasmic state, that can occur during transmission from mother to offspring. Thus, family member IX-240 carries approximately equal proportions of the three genotypes, but among her descendants there is preferential accumulation of one or two of the genotypes. This trend is particularly marked for family members X1-18, X-21, and X-23, who are essentially homoplasmic. Drift is apparently random, and there is no observable selection that favors one genotype over another. Finally, in addition to the results shown in table 1, one clone for family member X-144 (representing ~1% of the total number of clones that were analyzed) had the *recombinant* genotype: the mtDNA D-loop carried both the sequence changes, that at nucleotide 152 and that at nucleotide 195.

The presence of the new D-loop mutations in family member IX-240 and her descendants but their absence in the six descendants of family members VI-59 and VI-65 strongly suggests that both mutations arose no earlier than the germ line of family member VI-57 (fig. 1), who was born in 1861. More recent analyses have provided further information on this point. We have analyzed the D-loop from an affected male who is a maternal second cousin of IX-240. A total of 39 clones were sequenced for the appropriate region of the D-loop, *none* of which carried the polymorphisms at nucleotides 152 and 195 (data not shown). Therefore, it appears that both of these sequence changes arose no earlier than the germ line of family member VII-154, who was born in 1885.

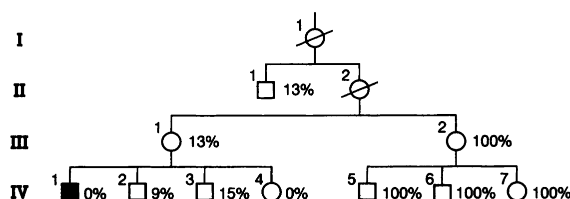

Figure 2 Pedigree of the 14484 USA1 LHON family. A total of 280 independent clones from 10 family members were sequenced for the region of the mitochondrial genome that spans the site of the 14484 primary LHON mutation. All of the clones carried the LHON mutation, and, hence, it is homoplasmic. However, a silent polymorphism at nucleotide 14470 was heteroplasmic, and the numbers indicate the percentage of mitochondrial genomes that carry this new mutation.

These mutations, therefore, have arisen in close proximity, both spatially and temporally.

Analysis of Coding-Region Sequences

During the past several years, we have accumulated nucleotide sequence data for the mitochondrial coding region from many unrelated LHON pedigrees. In most instances, only a single member of each pedigree has been analyzed. In other instances, however, multiple members of the same LHON family have been used for sequencing analysis of regions of the mitochondrial ND (complex I) genes. Some of these families have been described previously: the QLD1 LHON family, who carry the 14484 primary LHON mutation and the pathogenic 4160 mutation (Howell et al. 1991*b;* Howell 1994); the NWC1 3460 LHON family (Howell et al. 1991a); and the TAS1 11778 LHON family (Mackey and Buttery 1992; N.H., unpublished data). In addition to these families, we have also begun the analysis of a small American LHON family who carry the 14484 primary mutation (fig. 2) and who have been designated "USA1." In these families, the primary LHON mutations are homoplasmic.

The cumulative sequencing data for the ND genes from these four LHON pedigrees are summarized in table 2. For each region of the mitochondrial genome, at least two family members were analyzed; and the results for the four families have been pooled. It should be stressed that different numbers of family members have been investigated and that not all four LHON families have been analyzed for each ND-gene region. The results in table 2 indicate that three new coding-region mutations have arisen in these four LHON families. The mutations at nucleotides 4136 and 14560 have been described elsewhere (in Howell et al. 1991*b* and Howell et al. 1992, respectively). The third mutation, a silent polymorphism at nucleotide 14470 (the GLY36 residue of the ND6 subunit is unchanged), is described here for the first time (fig. 2).

How well do these results agree with previous estimates of the rate of mtDNA evolution that were derived by the phylogenetic approach? If the rate of coding-region divergence is 2%/Myr, then 1 child in every 4 million will differ from their mother at a particular base pair within the entire coding region (under the assumption that the mean time in a generation is 26 years). The results in table 2 encompass a cumulative span of 6,773 bp of the coding region. Therefore, it is predicted that ~1/590 children will differ from their mother at some site within this span—or, in other terms, that the rate of mutation in this span of the mitochondrial coding region is 1 mutation/590 generations. However, our sequencing analyses reveal three new mutations among the four LHON pedigrees, even though only 4–23 family members were assayed for each region of the mitochondrial genome. A more quantitative comparison can be derived as follows. A divergence rate of 2%/Myr extrapolates to a mutation frequency of $\sim 2.6 \times 10^{-7}$/bp/generation. In contrast, the results in table 2 yield a mutation frequency of $\sim 5.2 \times 10^{-5}$, a 200-fold-higher rate.

The segregation of these three new mutations differs in each case, thus reinforcing our view (e.g., see Ghosh et al. 1996) that simple models for the segregation and transmission of mtDNA molecules are untenable. The 4136 polymorphism was rapidly fixed in the homoplasmic state, probably within a single generation (Howell et al. 1991*b*). In contrast, the 14560 polymorphism has remained heteroplasmic through three generations

Table 2

Genealogical Analysis of the Mitochondrial Coding Region

Region Analyzed	No. of Family Members[a]	No. of New Mutations
3286–3564	17	0
3540–3815	7	0
3781–4064	5	0
4027–4294	21	1[b]
4441–5533	4	0
9983–10454	9	0
10435–11664	10	0
11630–11925	16	0
11922–12213	10	0
12116–13599	4	0
13576–14138	5	0
14086–14411	4	0
14381–14699	23	2[c]

[a] Members of the QLD1 (Howell et al. 1991*b*), NWC1 (Howell et al. 1991*a*), TAS1 (Mackey and Buttery 1992), and USA1 LHON families were analyzed (see Results).

[b] The Y277C mutation at nucleotide 4136 was found in one small branch of the QLD1 LHON family (Howell et al. 1991*b*).

[c] The silent polymorphism (G36G) at nucleotide 14560 was detected in the NWC1 LHON family (Howell et al. 1992), whereas that at nucleotide 14470 occurs in the USA1 LHON family (fig. 2).

(Howell et al. 1992). The 14470 polymorphism (fig. 2) shows both rapid and slow segregation: it was rapidly fixed in the homoplasmic state within one branch of the family, whereas it has not been fixed in the other branch, probably being lost completely in two members of generation IV.

Discussion

The impetus for this study was the desire to estimate, by using a novel and fully empirical genealogical approach, the rate of divergence within the human mitochondrial genome. If we add together all 45 members of the TAS2 family and include the 4 individuals with the TAS2 D-loop haplotype, and if we consider that the two germ-line mutations were independent events, then we obtain an mtDNA D-loop divergence rate of ~260%/Myr (if we assume that the mean time of a generation is 26 years, as did Lundstrom et al. [1992]) and a rate of ~1 mutation/25 generations. To our knowledge, these rate estimates for the human mitochondrial D-loop are the highest ever reported. Even if one makes the reasonable assumption that all immediate maternal relatives who were not analyzed underwent no new mutations (see, especially, examples such as individuals IX-13, X-145, and XI-3; fig. 1), the derived rate is still very rapid, ~1 mutation/40 generations. In a similar manner, genealogical analysis indicates that the divergence rate in the coding region is ~200-fold higher than that derived from the phylogenetic approach.

We state here our preliminary conclusions, pending further experimental and theoretical analysis, regarding the rapid rate of evolution that was derived here, together with a note of caution. The marked disparity between the genealogical and phylogenetic approaches is striking, and several possible explanations have been considered. For example, it is conceivable that our results represent a statistical "fluke" and that lower divergence rates would have been obtained if the genealogical approach had been extended to a larger data set (however, see below). The complete D-loop was analyzed for 45 family members from a total of >700 (most of whom are still living), a sample size of ~6%. If no additional new mutations have arisen in this family, then the rate of divergence would agree with that obtained elsewhere with the phylogenetic approach. In a similar fashion, the analysis of the coding region represents the cumulative results for *four* different LHON pedigrees, although (again) only a relatively small number of family members have been analyzed. At the same time, this concern must be tempered with the pragmatic realization that the TAS2 D-loop analysis reported here represents the determination and analysis of ~1 million bp of DNA sequence.

Second, it is possible that the rate of evolution may be much higher in some mitochondrial lineages than in others; certain constellations of sequences within the D-loop may increase the rate of additional mutations within this region; that is, the divergence rate obtained with a phylogenetic approach would be an "average" value. In this regard, an important recent study by Nachman et al. (1996) is relevant for our results (also see Ballard and Kreitman 1994; Rand et al. 1994). These investigators have provided evidence that evolution of the human mitochondrial coding region is operating under conditions of nonneutrality and that a substantial proportion of amino acid–replacement substitutions are not fixed at the *population* level. They further note that this nonneutrality will lead to an overestimate of the age of the human ancestor. We had suggested elsewhere that the new mutation at nucleotide 4136 in the QLD1 LHON family was an intragenic suppressor and that it arose under the influence of positive selection (Howell et al. 1991*b*). It also must be borne in mind that the members of the four LHON families carry pathogenic mutations that impair mitochondrial respiratory-chain function, and it is certainly plausible that abnormal mitochondrial metabolism may accelerate the rate of mtDNA mutation. Nachman et al. (1996) also concluded that pathogenic mitochondrial mutations may increase the ratio of replacement polymorphisms to silent polymorphisms, although not all of their "disease"-group individuals (Nachman et al. 1996, table 8) carried mitochondrial mutations whose pathogenicity has been unambiguously established. It will be important, therefore, to apply our approach to other matrilineal pedigrees, including those who carry no known pathogenic mitochondrial mutations.

Until these various issues are resolved, the simplest explanation for our results is that phylogenetic approaches have substantially underestimated the rate of mtDNA divergence and that the more rapid rates derived from the coalescent (Lundstrom et al. 1992) and the present genealogical approaches are more accurate. It is important to note that our estimated rate has received direct support from a recent and independent study that involved a large number of direct sequence comparisons between close maternal relatives from many unrelated maternal lineages: the empirically determined D-loop mutation rate was ~1 mutation/40 generations (Ivanov et al. 1996, p. 419; T. J. Parsons, personal communication).

Our sequencing approach was specifically designed to detect both homoplasmic and heteroplasmic mutations, even if the latter occurred in only 10%–20% of the mtDNA molecules. Because this sensitivity has not been feasible for the large-scale population screenings that were used for the phylogenetic approach, heteroplasmic mutations would have been overlooked in those studies, thus leading to an underestimation of the divergence

rate. In the present studies, six new mutations were detected, five of which were heteroplasmic in at least some of the LHON family members. There is also independent evidence that heteroplasmic D-loop mutations are more frequent than previously had been recognized (e.g., see Comas et al. 1995). We have sequenced the D-loop region from ~100 genealogically unrelated individuals: 2 of these had heteroplasmic mutations (N.H., unpublished data). In addition, the last Russian czar carried a new, heteroplasmic D-loop mutation (Gill et al. 1994; Ivanov et al. 1996). Thus, our empirically derived mutation rate may be higher simply because of a greater mutation-detection sensitivity. However, there is another possible interpretation that could explain the discrepancy between the results of the empirical approach and those of the phylogenetic approaches. The former will detect newly arisen mitochondrial mutations, and many of these, even though they have been transmitted through a few generations, may not be fixed at the *population* level, possibly being lost through the effects of selection (Nachman et al. 1996); that is, the phylogenetic approach may more accurately be said to yield the *fixation* rate of mitochondrial mutations, rather than the rate of mtDNA divergence or mutation.

If a rapid divergence rate of mtDNA is confirmed, previous studies that have used D-loop sequences for geographical and chronological studies of human evolution will have to be reevaluated. As noted earlier, phylogenetic or relative-branch-length approaches indicate that the time since the most recent common ancestor is in the range of 0.2–0.6 Myr. The present results, as well as those of Lundstrom et al. (1992), suggest either a much more recent origin of modern humans or that evolution of the mitochondrial genome is not a reliable "evolutionary clock" (on this latter point, see Donnelly and Tavare 1995). Penny et al. (1995) have pointed out that their phylogenetic analyses of D-loop sequences may be compatible with "cultural" or Pleistocene-explosion models (e.g., see Noble and Davidson 1991) for a very recent major expansion of the human population, perhaps within the past 40,000 years. Analysis of mtDNA mismatch distributions also support a Pleistocene explosion (Sherry et al. 1994; Rogers 1995). Furthermore, analyses of *nuclear* genetic loci are also compatible with a recent origin for modern humans (Whitfield et al. 1995; Armour et al. 1996; Tishkoff et al. 1996; but, for a conflicting result, see Hammer 1995). Overall, therefore, a variety of approaches indicate a recent origin for modern humans, a result that is incompatible with a "slow" rate of mtDNA evolution.

The present studies provide additional evidence that mitochondrial mutations are fixed rapidly in individuals because of "bottlenecking" during oogenesis (Hauswirth and Laipis 1982; Howell et al. 1992). According to this model, only a small subset of the mitochondrial genomes in the female germ-line pool are transmitted to an offspring. As a consequence of bottlenecking, therefore, only a fraction of new mitochondrial mutations will be transmitted to subsequent generations, but the rate of fixation for that transmitted fraction will be much more rapid than it would be if bottlenecking were absent. Thus, the rate of mitochondrial *mutation* is likely to be substantially higher than the rate of divergence, particularly if a significant proportion of mutations that are fixed in individuals are not fixed within the population (see above). In our opinion, the evolutionary implications of bottlenecking have not received sufficient attention.

Our results suggest that recombination of mtDNA molecules may occur in humans, although the detection of a single clone with a recombinant genotype also raises the specter of an artifactual origin such as might occur with "jumping" PCR (Pääbo et al. 1990). Further analysis will be necessary for clarification of the facility with which mtDNA molecules recombine in situ.

Finally, these results provide independent support and additional illustrative detail for the contention that mutations within the D-loop do not occur independently, in time or in space, during evolution (Kocher and Wilson 1991; Wakeley 1993; Cummings et al. 1995; Yang 1995). In the first place, it is striking that, within the TAS2 LHON family, the two germ-line mutations are separated by only 40 bp and that they may have arisen within the same generation (also, the *somatic* mutation in family member X-18 occurs within 10 bp of the mutation at nucleotide 195). Second, the mitochondrial genomes of the TAS2 and VIC2 LHON families descended from a common ancestor, and they differ at nine sites within the D-loop (Howell et al. 1995, fig. 7). In the VIC2 lineage, since the last common ancestor, four sequence changes have occurred in the D-loop, three within a span of 80 bp in hypervariable region I. In the TAS2 mitochondrial genome, five changes have occurred, four within an 80-bp span of the hypervariable region II. This nonrandom pattern of substitution may indicate that there is selective pressure to maintain certain secondary structural features that influence the initiation of replication and/or of transcription, possibly by affecting the binding of regulatory proteins (for a recently described example of protein-mtDNA interactions, see Gadaleta et al. 1996). Discussions of D-loop evolution usually ignore its crucial role in these processes.

The genealogical approach indicates that mutation of the human mitochondrial genome may be even more rapid than previously believed on the basis of phylogenetic or relative-branch-length approaches. In addition to further analyzing the process of mtDNA evolution, we should increase our attention to the broader question of how (or whether) organisms can tolerate, in the sense

of evolution, a genetic system with such a high mutational burden.

Acknowledgments

We gratefully acknowledge the cooperation of the members of the TAS2 LHON family, without whom this study would not have been possible. Dr. Christopher Chan (University of Melbourne) prepared several of the DNA samples, and Mr. Steve Halvorson performed the sequencing analysis of the coding region of the USA1 mitochondrial genome; their efforts constitute a significant contribution to this work. A number of colleagues offered helpful suggestions and comments on this work: we acknowledge, in particular, the discussions with Dr. Thomas J. Parsons (Armed Forces Institute of Pathology) on unpublished studies of mtDNA evolution. This research was supported by grants from the John Sealy Memorial Endowment Fund and the National Eye Institute (grant RO1 EY10758) that were awarded to N.H.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457–465

Armour JAL, Anttinen T, May CA, Vega EE, Sajantila A, Kidd JR, Kidd KK, et al (1996) Minisatellite diversity supports a recent African origin for modern humans. Nat Genet 13: 154–160

Ballard JWO, Kreitman M (1994) Unraveling selection in the mitochondrial genome of *Drosophila*. Genetics 138:757–772

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36

Comas D, Pääbo S, Bertranpetit J (1995) Heteroplasmy in the control region of human mitochondrial DNA. Genome Res 5:89–90

Cummings MP, Otto SP, Wakeley J (1995) Sampling properties of DNA sequence data in phylogenetic analysis. Mol Biol Evol 12:814–822

Donnelly P, Tavare S (1995) Coalescents and genealogical structure under neutrality. Annu Rev Genet 29:401–421

Gadaleta G, D'Elia D, Capaccio L, Saccone C, Pepe G (1996) Isolation of a 25kDa protein binding to a curved DNA upstream the origin of the L strand replication in the rat mitochondrial genome. J Biol Chem 271:13537–13541

Ghosh SS, Fahy E, Bodis-Wollner I, Sherman J, Howell N (1996) Longitudinal study of a heteroplasmic 3460 Leber hereditary optic neuropathy family by multiplexed primer-extension analysis and nucleotide sequencing. Am J Hum Genet 58:325–334

Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, et al (1994) Identification of the remains of the Romanov family by DNA analysis. Nat Genet 6:130–135

Hammer MF (1995) A recent common ancestry for human Y chromosome. Nature 378:376–378

Hauswirth WW, Laipis PJ (1982) Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. Proc Natl Acad Sci USA 79:4686–4690

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial mtDNAs. Proc Natl Acad Sci USA 92:532–536

Howell N (1994) Primary LHON mutations: trying to separate "fruyt" from "chaf." Clin Neurosci 2:130–137

Howell N, Bindoff LA, McCullough DA, Kubacka I, Poulton J, Mackey D, Taylor L, et al (1991*a*) Leber hereditary optic neuropathy: identification of the same mitochondrial ND1 mutation in six pedigrees. Am J Hum Genet 49:939–950

Howell N, Halvorson S, Kubacka I, McCullough DA, Bindoff LA, Turnbull DM (1992) Mitochondrial gene segregation in mammals: is the bottleneck always narrow? Hum Genet 90:117–120

Howell N, Kubacka I, Halvorson S, Howell B, McCullough DA, Mackey D (1995) Phylogenetic analysis of the mitochondrial genomes from Leber hereditary optic neuropathy pedigrees. Genetics 140:285–302

Howell N, Kubacka I, Xu M, McCullough DA (1991*b*) Leber hereditary optic neuropathy: involvement of the mitochondrial ND1 gene and evidence for an intragenic suppressor mutation. Am J Hum Genet 48:935–942

Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ (1996) Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. Nat Genet 12:417–420

Kocher TD, Wilson AC (1991) Sequence evolution of mitochondrial DNA in human and chimpanzees: control region and a protein-coding region. In: Osawa S, Honjo T (eds) Evolution of life: fossils, molecules, and culture. Springer, Tokyo, pp 391–413

Lundstrom R, Tavare S, Ward RH (1992) Estimating substitution rates from molecular data using the coalescent. Proc Natl Acad Sci USA 89:5961–5965

Mackey DA, Buttery RG (1992) Leber hereditary optic neuropathy in Australia. Aust NZ J Ophthalmol 20:177–184

Mackey D, Howell N (1992) A variant of Leber hereditary optic neuropathy characterized by recovery of vision and by an unusual mitochondrial genetic etiology. Am J Hum Genet 51:1218–1228

Nachman MW, Brown WM, Stoneking M, Aquadro CF (1996) Nonneutral mitochondrial variation in humans and chimpanzees. Genetics 142:953–963

Noble W, Davidson I (1991) The evolutionary emergence of modern human behavior: language and its archaeology. Man (ns) 26:223–254

Pääbo S, Irwin DM, Wilson AC (1990) DNA damage promotes jumping between templates during enzymatic amplification. J Biol Chem 265:4718–4721

Penny D, Steel M, Waddell PJ, Hendy MD (1995) Improved analyses of human mtDNA sequences support a recent African origin for *Homo sapiens*. Mol Biol Evol 12:863–882

Pesole G, Sbisa E, Preparata G, Saccone C (1992) The evolution of the mitochondrial D-loop and the origin of modern man. Mol Biol Evol 9:587–598

Rand DM, Dorfsman M, Kann LM (1994) Neutral and non-

neutral evolution of *Drosophila* mitochondrial DNA. Genetics 138:741–756

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. Evolution 49:608–615

Ruvolo M (1996) A new approach to studying modern human origins: hypothesis testing with coalescence time distributions. Mol Phylogenet Evol 5:202–219

Sherry ST, Rogers AR, Harpending H, Soodyall H, Jenkins T, Stoneking M (1994) Mismatch distributions of mtDNA reveal recent human population expansions. Hum Biol 66: 761–775

Stoneking M, Sherry ST, Redd AJ, Vigilant L (1992) New approaches to dating suggest a recent age for the human mtDNA ancestor. Philos Trans R Soc Lond [B] 337:167–175

Tamura K, Nei M (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 (1993) The "Eve" hypothesis: a genetic critique and reanalysis. Am Anthropol 95:51–72

Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, Bonne-Tamir B, et al (1996) Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science 271:1380–1387

Torroni A, Lott MT, Cabell MF, Chen, Y-S, Lavergne L, Wallace DC (1994*a*) mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. Am J Hum Genet 55:760–776

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994*b*) Mitochondrial DNA "clock" for the Amerinds with its implications for timing their entry into North America. Proc Natl Acad Sci USA 91:1158–1162

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Wakeley J (1993) Substitution rate variation among sites in hypervariable region I of human mitochondrial DNA. J Mol Evol 37:613–623

Ward RH, Frazier BL, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 88:8720–8724

Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. Nature 378:379–380

Wills C (1995) When did Eve live? an evolutionary detective story. Evolution 49:593–607

Yang Z (1995) A space-time process model for the evolution of DNA sequences. Genetics 139:993–1005