Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 3246–3250, May 1982
Genetics

Novel features of animal mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes

(rate of evolution/silent substitution/mutation rate/fixation/sequence divergence estimation)

GREGORY G. BROWN* AND MELVIN V. SIMPSON†

*Department of Biology, McGill University, Montreal, P. Q. H3A 1B1, Canada; and †Department of Biochemistry, State University of New York, Stony Brook, New York 11794

Communicated by Philip Siekevitz, February 22, 1982

ABSTRACT **The sequence of the region of the mitochondrial genome that encodes cytochrome oxidase subunit II (COII) has been determined for each of two closely related rat species, *Rattus norvegicus* and *R. rattus*. Comparison of the two sequences shows that 94.4% of the nucleotide substitutions are silent. The occurrence of this high proportion of silent substitutions leads us to propose that the rapid evolution of mtDNA relative to nuclear DNA is due only to silent changes and that amino acid-altering substitutions accumulate in nuclear and mtDNA at comparable rates. Other novel features of the nucleotide substitution pattern in the rat COII gene are a high transition/transversion ratio (8.0:1) and a strong bias toward C ↔ T transitions in the light strand. Comparison of the *R. norvegicus* COII sequence with the bovine and human sequences shows that there may be selective constraints on some silent positions within the gene and that its rate of evolution may be different in different mammalian lineages.**

The degree to which a sequence is conserved through the course of evolution is frequently used as a means of assessing its functional significance (see, e.g., ref. 1). It is generally thought that the greater the functional importance a sequence has, the less likely it is that changes in that sequence will be tolerated by the organism. In the case of proteins, numerous examples can be cited to support this contention (2). Animal mtDNA, however, appears to be an exception to this rule. In spite of the fact that it is essential for cellular respiration (3), it exhibits a large degree of sequence variability both within (4–6) and between (7, 8) species.

By comparing restriction maps of 11 different rat mtDNAs, we have previously ruled out the possibility that this variation is concentrated in functionally unimportant evolutionary "hot spots" (4, 5). The variant restriction sites were found to be evenly distributed on the molecule in the region outside the rRNA coding sequences. Since ≈80% of this region codes for proteins, we concluded that the phenomenon of rapid evolution pertained to the protein coding genes. It remained unclear, however, why the protein coding sequences should exhibit this high degree of variation.

We have suggested one possibility for this—i.e., that an exceptionally high percentage of the base substitutions may occur at silent positions within codons (4, 5). If this were the case, then the fast rate of nucleotide sequence evolution would not be such an anomaly. The gene products themselves, upon which selection acts, would be relatively more conserved. Sequence analyses (5, 9–11) of two laboratory rat mtDNA variants (12, 13) provided preliminary support for this hypothesis; all three substitutions observed within a 169-base-pair stretch of DNA were silent. Three substitutions, of course, are too few from which to draw statistically valid conclusions. In addition, the gene product corresponding to this 169-base-pair sequence is unknown, and we considered it vital to study the sequence of a gene with a known protein product.

We report here the sequences of the cytochrome oxidase subunit II (COII) genes of two closely related rat species, *Rattus norvegicus* and *R. rattus*. Comparison of the sequences shows that the proportion of nucleotide substitutions that are silent is extraordinarily high, higher in fact than has been observed in any protein coding sequences, mitochondrial or nonmitochondrial, to date (see *Discussion*). Comparison of the rat sequences with other COII sequences reveals several other interesting features of mtDNA evolution, among them that the rate of evolution of the molecule appears to differ in different mammalian lineages.

MATERIALS AND METHODS

Source of mtDNAs. Type NA mtDNA, one of the two mtDNA types found in the laboratory rat population, was obtained from livers of Sprague–Dawley laboratory rats (*R. norvegicus*). Type RE mtDNA was obtained from the liver and kidneys of a single wild *R. rattus* animal trapped in Fort Lauderdale, FL (4). The *R. norvegicus* mtDNA, although obtained from laboratory rats, differs in sequence from the mtDNAs of wild *R. norvegicus* from either Japan or the United States by approximately the same amount as these wild rat mtDNAs differ from each other (4, 5). It can therefore be considered to be an appropriate representative of the mtDNA of this species. mtDNAs were purified as described (4).

Cloning. The *Hin*dIII fragments produced from each mtDNA were inserted into the *Hin*dIII site of plasmid pBR322 by the ligation procedure described previously (10). Prior to ligation the *Hin*dIII-digested pBR322 DNA was treated with alkaline phosphatase to prevent recirculation. Ligation mixtures were used to transform competent *Escherichia coli* cells, and ampicillin-resistant transformants were screened for the presence of mitochondrial *Hin*dIII fragments by the method of Birnboim and Doly (14). The clones designated NA330 and RE217 carried the conserved 2.1-kilobase fragments from types NA and RE mtDNAs, respectively.

Restriction Mapping. The locations of restriction sites in the inserted fragments of pNA330 and pRE217 were determined by multiple digestion procedures. Whenever possible, digests of the two plasmids were subjected to electrophoresis in adjacent lanes of the same gel so that an accurate assessment of whether or not a particular restriction fragment was conserved between the two DNAs could be made.

DNA Sequence Analyses. Restriction fragments falling in the left-most region of the inserts (see Fig. 1), were labeled either

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: COII, cytochrome oxidase subunit II.

at their 5′ ends with [γ-^{32}P]ATP and polynucleotide kinase (15) or at their 3′ ends with [α-^{32}P]dATP and the Klenow fragment of DNA polymerase I (16). Labeled ends of DNA fragments were separated and analyzed according to specified protocols (15).

Estimation of Sequence Divergence. Sequence divergence was estimated from comparison of restriction maps and fragments as described (4). For cases in which the divergence values obtained in nucleotide sequence comparisons were corrected for multiple hits and reversions, the procedure of Perler *et al.* (17) was used; K values (18), the corrected number of nucleotide substitutions per base, were calculated for category 1 and 3 silent substitutions and category 2 and 3 replacement substitutions (see ref. 17 for definitions of substitution categories). In some cases, the K values for the category 1 and 2 substitutions were further corrected for the unequal frequencies of transitions and transversions by using the equations

$$K' = \frac{r + 2K}{3r} \quad \text{(category 1)}$$

and

$$K' = \frac{r + 2K}{3} \quad \text{(category 2)},$$

where r = rate of occurrence of transitions/rate of occurrence of transversions and K' is the corrected value of K. The equation for the category 1 sites was derived by combining Salser's equation 3 (19), which makes allowance for unequal rates of occurrence for transitions and transversions, with the corresponding equation 1 of Perler *et al.* (17), which does not. A correction equation, which took into account the effects of unequal rates of transitions and transversions, was derived for the category 2 replacement substitutions. This was combined with equation 2 of Perler *et al.* (17) to give the above equation for category 2 sites. The total silent (K_s, K_s') and replacement (K_r, K_r') substitution values were calculated from the weighted averages of the values obtained for each category (17).

The nodal sequence used in the calculations for Fig. 3 was derived from the *R. norvegicus*, human, and bovine sequences by the procedure outlined by Fitch (20).

RESULTS

Restriction Maps and Sequence Analysis Strategy. Two considerations prompted the choice of the mtDNAs of *R. norvegicus* and *R. rattus* as the subject of this study. The rate of mtDNA evolution begins to decrease with increasing evolutionary distance after ≈15% of the bases have undergone substitution (8). To identify the type of base substitution characteristic of the rapid early phase, it seemed essential, therefore, to choose mtDNAs that differed in sequence by ≈15% or less. On the other hand, if the mtDNAs chosen were too closely related, a relatively large amount of sequence would have to be determined to obtain a fair statistical picture of the base substitution pattern. The estimated sequence divergence of 16% between the mtDNAs of these two rat species (4, 5) indicated that they were highly suitable for this purpose.

From its location in human mtDNA (21, 22), it appeared that the COII gene would fall at least partly within a *Hin*dIII fragment that spans from 51.5 to 63 units on the maps of the two rat mtDNAs. The *Hin*dIII fragments of a *R. norvegicus* mtDNA (type NA) and *R. rattus* mtDNA (type RE) (4) were cloned in pBR322. Recombinant plasmids carrying the COII gene containing fragments of *R. norvegicus* and *R. rattus* mtDNAs were identified and designated pNA330 and pRE217, respectively. The detailed restriction maps constructed for the inserted fragments in both plasmids are shown in Fig. 1, together with the *Hin*dIII maps of types NA and RE mtDNAs. From the fraction of conserved restriction sites in the detailed maps, we calculate that the amount of sequence divergence between the cloned fragments is 15%, in close agreement with the previous estimate of 16% for the two intact mtDNAs (4, 5).

The sequences of the two left-most *Hin*fI fragments in the inserts of both pNA330 and pRE217 were determined by the Maxam and Gilbert method (15). Areas analyzed in individual runs, together with the direction of the run, are shown in Fig. 1. About 70% of the *R. norvegicus* sequence and 40% of the *R. rattus* sequence were obtained on both DNA strands. Sequences were not determined through one restriction site in *R. norvegicus* and two restriction sites in *R. rattus* because a continuous sequence was obtained in one of the species at these sites that aligned with the sequence of the other without discontinuity. Furthermore, both rat COII sequences could be continuously aligned with the sequences of other mammalian COII genes (21–23).

Organization of the Rat COII Region. The sequence of 702 of the 829 nucleotides determined for each species is given in Fig. 2. The location of the COII genes was determined by alignment with the human sequence (21, 22). The genes are 684 nucleotides long and fall completely within the analyzed fragments; their 5′ ends (sense strand) are 128 nucleotides from the more internal *Hin*fI sites (Fig. 1) and their 3′ ends are 11 bases from the *Hin*dIII sites that form the boundaries of the inserts. Both rat COII sequences exhibit 72% homology with the human sequence and 80% homology with the bovine sequence (23).

The locations of tRNA genes occurring in the fragments analyzed were also determined by alignment with the human and bovine sequences. The organization of the COII gene with respect to these genes is identical in the two rat species and identical to the arrangement found in bovine mtDNA (23). In the two rat species and in the cow, one nucleotide separates the 3′ end of the $tRNA^{Asp}$ gene from the 5′ end of the COII gene and three nucleotides separate the 3′ end of the COII gene from the 5′ end of the $tRNA^{Lys}$ gene. A different arrangement occurs in human mtDNA, where the $tRNA^{Lys}$ gene is "butt-jointed" to the COII gene and 25 nucleotides separate the COII and $tRNA^{Lys}$ genes (21, 22).

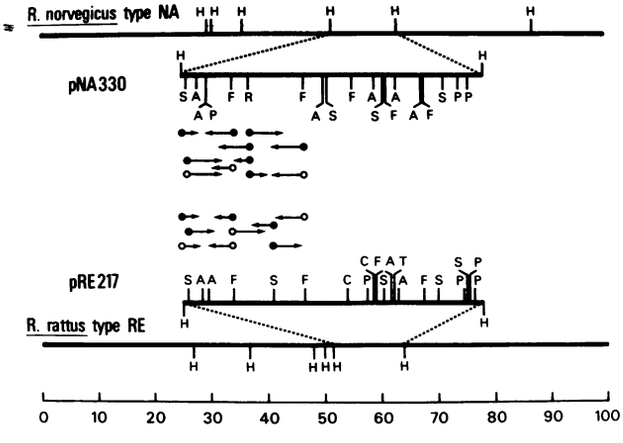

FIG. 1. *Hin*dIII maps of *R. norvegicus* type NA and *R. rattus* type RE mtDNAs and detailed maps of the inserted *Hin*dIII fragments of pNA330 and pRE217. The origin of heavy-strand replication is at zero, and its replication proceeds from left to right. Arrows indicate the direction and length of individual sequence analyses. ●, Analyzed fragment was labeled at its 5′ end; ○, analyzed fragment was labeled at its 3′ end. Cleavage sites: *A*, *Hae* III; *C*, *Hin*cII; *F*, *Hin*fI; *H*, *Hin*dIII; *P*, *Hpa* II; *R*, *Eco*RI; *S*, *Sau*3A; *T*, *Sst* I.

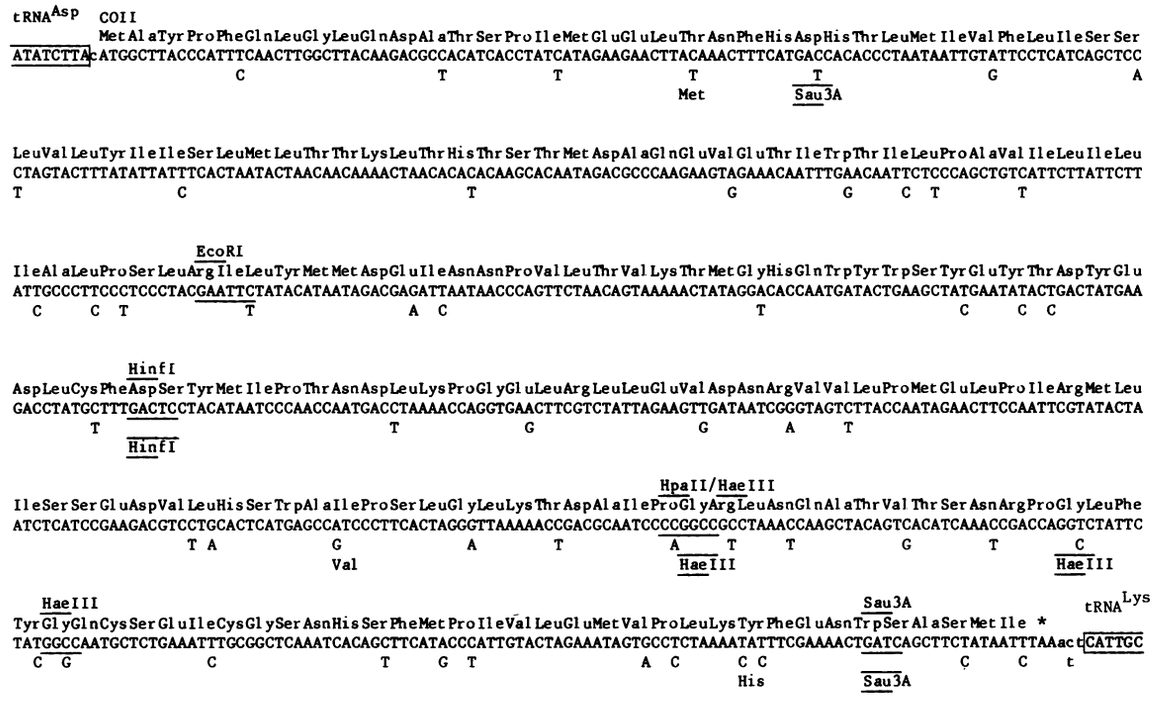

FIG. 2. Nucleotide sequence of the rat COII gene region. The *R. norvegicus* sequence is shown with the changes occurring in the *R. rattus* sequence given below it. Sequences presented are of the light strand, oriented 5′ → 3′, left to right. Brackets indicate the flanking tRNA genes; *, termination codon. Restriction sites indicated in the maps of Fig. 1 are underlined and identified. Nucleotides designated by lower case letters are not part of the structural genes. The amino acid sequence was determined by using the mitochondrial coding rules given in Anderson *et al.* (22).

Comparison of the Rat COII Sequences. In contrast to the 15% sequence divergence estimated from restriction data, only 8% of the bases in the two rat COII genes are different. The locations of the nucleotide substitutions are indicated in Fig. 2, and the substitutions are classified according to type of the base change produced in Table 1. The changes are distributed evenly throughout the gene and 51 out of a total of 54, or 94%, are silent. Forty-seven of these silent substitutions occur in the third codon position, while four occur in the first codon position (CTPu ↔ TTPu). Eighty-nine percent of the substitutions are transitions, and 77% of these transitions are C ↔ T replacements in the sense, or light strand.

DISCUSSION

Rapidly Accumulating Substitutions in mtDNA Are Silent. The pattern of nucleotide substitution observed in the rat COII sequences is unusual in a number of respects. Perhaps the most striking of these is the extent to which the substitutions are silent. Ninety-four percent of the substitutions do not result in amino acid changes. This figure is higher than that observed in any of the genes considered in a recent comparison by Jukes (24) and only in the sea urchin histone genes are comparable values obtained. Since the nucleotide sequence divergences are approximately the same for the various mitochondrial genes (4, 5, 22) and COII has the least highly conserved amino acid sequence of the known mitochondrially encoded proteins in mammals (22), we can reasonably assume that similar or even greater levels of silent substitution characterize other mitochondrial genes.

Table 1. Nucleotide differences in the mitochondrial light strand between the COII genes of *R. norvegicus* and *R. rattus*

	Silent	Replacement
Transitions		
C ↔ T	35	2
G ↔ A	10	1
Total	45	3
Transversions		
G ↔ C	2	0
G ↔ T	1	0
A ↔ C	2	0
A ↔ T	1	0
Total	6	0
Total substitutions	51	3

Thirty to fifty percent of the evolutionarily fixed nucleotide substitutions in most nuclear genes lead to amino acid replacement (24). In comparison, only 6% of the substitutions in the COII gene cause such replacement. Thus, even if the COII gene were to evolve 6–9 times faster than an average nuclear gene, the extent of the change resulting in the gene products would be similar. As the rate of mtDNA evolution has been estimated to be 5–10 times that of single copy nuclear genes (8), it is probable that no great discrepancy exists between the rates of evolution of typical nuclear and mitochondrially encoded proteins.

Restriction Estimates of Sequence Divergence. An assumption made in estimating sequence divergence from restriction data is that all nucleotides have an equal chance of undergoing substitution (25). The limitation of nucleotide substitutions to the third position in codons may affect the accuracy with which sequence differences between mtDNAs can be approximated from restriction analyses. For example, one can envision an extreme situation of two DNA sequences that differ at every third base; no restriction sites could be held in common by these DNAs and hence the restriction site estimate of sequence divergence would be 100% instead of the actual figure of 33%. A comparable, although less extreme situation may account, at least in part, for the discrepancy between the restriction-esti-

mated sequence divergence for the rat COII genes (15%) and the actual divergence (8%). It is possible that other sequence divergence estimates for mtDNA based on restriction data may be similarly in error.

The Mutation Rate for mtDNA Appears to Be High. Without prior knowledge of the base substitution pattern, the rapid evolution of mtDNA could be ascribed to either a high mutation rate or a high level of fixation of mutations (8). Two possibilities by which increased fixation could give rise to the high rate of evolution seem consistent with the high proportion of silent mitochondrial substitutions: (*i*) selection for silent changes that is higher in the mitochondrion and (*ii*) selection against silent changes that is stronger in the nucleus. In general, however, we would expect most silent changes to be neutral, or near neutral (26), and therefore not highly susceptible to selection in either the mitochondrion or the nucleus. We consider it far more likely, therefore, that it is a high mutation rate that is primarily responsible for the fast rate of mtDNA evolution and that the high proportion of silent substitutions arises from very strong selection against amino acid-altering mutations in the mitochondrion.

Several possible sources for a high mitochondrial mutation rate have been discussed (5, 8). Recently, Kunkel and Loeb (27) have reported that the mtDNA polymerase, γ, incorporates mispaired bases $\approx$5 times more frequently than polymerase α, which is responsible for nuclear DNA replication. This provides us with another possible explanation for the high mitochondrial mutation rate—namely, that mtDNA replication itself is inherently inaccurate.

Transitions vs. Transversions. When globin or other nuclear encoded mRNAs from different species are compared, the number of substitutions that are transitions is usually approximately the same as the number that are transversions (20). In contrast, in the rat COII genes, the number of transitions exceeds the number of transversions by a factor of 8.0/1. This high proportion of transitions is in part due to the extent of silent substitution. However, if we consider only those third-position substitutions in the gene, in those codon families in which transitions and transversions stand an equal chance of being silent, the transitions still outnumber the transversions, in this case by a factor of 3.2/1. Since the opportunity for transversions to occur at these positions is twice as great as for transitions, the rate of transition substitutions must be 6.4 times that of transversions. An additional bias is seen in the high proportion (77%) of the transitions that are observed as C $\leftrightarrow$ T changes in the light strand.

This nonrandomness does not seem to be the result of selection for preferentially used codons. If this were the case, many of the changes would involve adenosine residues, since the most frequently used codons in mitochondria have adenosine or cytidine in the third position (22). The predominance of C $\leftrightarrow$ T changes in only the light strand cannot be accounted for by the Py/Pu ratio in the third codon position (which is 1.1:1) and therefore seems particularly difficult to explain, since it indicates that some types of mutations occur more frequently in one DNA strand than in the other. In the case of animal mtDNA, however, this may be possible, because the heavy strand remains unpaired during much of the replication cycle while the light strand does not (28). Thus, changes that occur preferentially on single-stranded DNA, such as cytosine deamination, will occur more frequently on the heavy than on the light strand. Interestingly, Fitch (20) has noticed a similar bias among the silent substitutions occurring in mammalian globin genes.

Closely vs. Distantly Related COII Sequences. The types of base changes observed when the *R. norvegicus* COII gene sequence is compared with the *R. rattus*, bovine, and human COII sequences are summarized in Table 2. The substitution pattern observed in the rat vs. rat comparison differs significantly from that observed when more distantly related species are compared. Twenty-two percent of the base differences between the rat and bovine sequences and 44% of the base differences between the rat and human sequences result in amino acid replacements. In contrast, only 6% of the substitutions in the two rat sequences do so. In addition, the number of transitions observed in comparisons of more distantly related species is approximately equal to the number of transversions.

The differences between the comparisons of closely and more distantly related species can be understood if we adopt the view that the rat/rat substitution pattern is the pattern that best represents the process of mtDNA evolution. In a relatively short evolutionary period after the divergence of two mtDNAs, the rapid accumulation of silent substitutions will lead to "saturation" of the silently mutatable sites. Subsequent silent substitutions will not be observed when the DNA sequences are compared, because they will simply be additional "hits" at already substituted sites. Replacement substitutions, however, will gradually continue to accumulate in mtDNA after the silent sites are saturated. The proportion of silent substitutions will therefore appear lower when more distantly related species are compared. Thus, to obtain an accurate picture of certain aspects of mtDNA evolution one must choose closely related species for comparison, in order that the short-term events are not obscured by longer-term events.

One might suppose that the need to use closely related species could be circumvented if the base-substitution values obtained in the comparisons of the more distantly related species were corrected for multiple hit and reversion events. Equations by which such corrections can be made have been proposed (17–19), and we have applied these to the different sequence

Table 2. Comparison of mammalian COII gene sequences

	Nucleotide substitutions, no.				Nucleotide sequence homology, %	Replacement substitutions, %	Substitutions per nucleotide†					
	Silent*	Replacement*	Transition	Transversion			K_s	K'_s	K_r	K'_r	K_s/K_r	K'_s/K'_r
R. norvegicus/human	107.5	84.5	96	96	71.8	44.0	1.21	0.81	0.17	0.24	7.1	3.3
R. norvegicus/bovine	104.5	29.5	64	70	80.3	22.0	1.33	0.82	0.07	0.10	19.0	8.2
R. norvegicus/ *R. rattus*	51	3	48	6	92.1	5.6	0.51	0.28	0.005	0.005	102.0	56.0

* Nonintegral values arise because some substitutions are scored as silent according to the codon of one species but as replacement according to the codon of the other.

† Average number of base substitutions per nucleotide occurring at silent (K_s, K'_s) or replacement (K_r, K'_r) positions within codons, after correction for multiple substitution and reversion events. For the K-value calculations, transitions and transversions were assumed to occur at the same rate, whereas for the K' calculations, transitions were assumed to be 6.4 times as frequent as transversions, their apparent relative frequencies according to the category 3 silent substitutions in the rat/rat comparison.

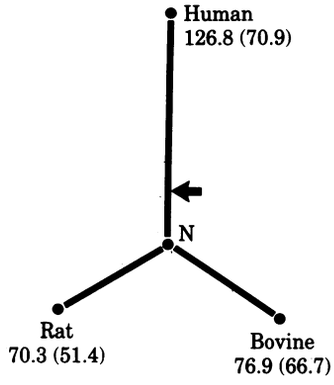

FIG. 3. Unrooted tree depicting relative distances of rat, human, and bovine COII sequences from the nodal sequence (designated N). The length of each branch is proportional to the number of nucleotide differences, as given below the species name. Values in parentheses represent number of silent differences. The arrow indicates the point most equidistant from all ends.

comparisons. The corrected values for silent (K_s) and replacement (K_r) changes, expressed in terms of substitutions per nucleotide, are given in Table 2. It can be seen that, among the corrected values, considerable discrepancy still exists between comparisons of closely and more distantly related species regarding the relative frequency of the two types of substitutions (K_s/K_r values). Correction of the K values for the high frequency of transitions relative to transversions (K' values) does not markedly affect this discrepancy.

It is interesting that even corrected values for silent substitutions display some evolutionary saturation (17). This phenomenon may be due to unknown selective constraints on a fraction of the total potentially silent sites (17). Such constraints seem to occur in globin genes (19, 20) and the large discrepancy between the K_s/K_r ratios for closely and distantly related species suggests they occur in the COII genes as well.

Rates of mtDNA Evolution May Vary. An unrooted evolutionary tree for the mammalian COII gene is shown in Fig. 3. On it are indicated the number of silent and replacement substitutions occurring between each existent sequence and the nodal sequence—i.e., the ancestral sequence at the node N (20). The number of substitutions between the nodal and human sequences is greater than between the nodal and either of the other two. Unless we maintain that the most recent common ancestor for the three present-day genes is represented by the point at the arrow, which seems improbable on phylogenetic grounds (29), we must conclude that the gene has evolved at a different rate, at least in the primate lineage.

The approximate constancy of evolutionary rates (2) has been cited as evidence that most of the substitutions that accumulate in the course of evolution are selectively neutral (30). The apparent variability of rates in the case of the COII gene might therefore be taken as evidence to the contrary: that selection for favorable substitutions is the driving force for fixation in this case. The observation that the greater distance between the human and nodal sequences can be accounted for almost entirely by replacement substitutions lends support to the view that the primate rate has been accelerated by increased positive (or decreased negative) selection for amino acid replacements. Clearly, however, more information on the processes involved in the generation and fixation of mitochondrial mutations is necessary to substantiate this view. It is worth noting that, in contrast to that of the COII gene, the rates of evolution of nuclear genes appear to be slower in the primate lineage (31).

Note Added in Proof. When the rat COII sequences are compared with those of the COII gene of the recently analyzed mouse mtDNA (32), it is found that 97% of the changes in the case of *R. norvegicus* and 98% of the changes in the case of *R. rattus* are silent. A slightly different sequence for the *R. norvegicus* COII gene has recently been reported by Grosskopf and Feldman (33).

We thank Drs. S. P. Gibbs, Y. Nishioka, and D. P. S. Verma for critical comments on the manuscript; D. P. S. Verma and K. Marcu for use of equipment; and Frank Castora, Normand Brisson, and Mike Kuehn for helpful discussions. This work was supported by an American Cancer Society Professorship Grant to M.V.S. and grants from the National Institutes of Health (GM 22333-15); the National Science Foundation (PCM-8004738); the Natural Sciences and Engineering Research Council, Canada (A7624); and the Graduate Faculty of McGill University. Part of the work was performed while G.G.B. was a National Institutes of Health Postdoctoral Fellow (GM 06972).

1. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.
2. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639.
3. Tzagaloff, A., Macino, G. & Sebald, W. (1979) *Annu. Rev. Biochem.* **48**, 419–442.
4. Brown, G. G. & Simpson, M. V. (1981) *Genetics* **97**, 125–143.
5. Brown, G. G., Castora, F. J., Frantz, S. C. & Simpson, M. V. (1981) *Ann. N.Y. Acad. Sci.* **361**, 135–153.
6. Ferris, S. D., Brown, W. M., Davidson, W. S. & Wilson, A. C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6319–6323.
7. Dawid, I. B. (1972) *Dev. Biol.* **29**, 139–151.
8. Brown, W. M., George, M., Jr., & Wilson, A. C. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1967–1971.
9. Castora, F. J., Arnheim, N. A. & Simpson, M. V. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6415–6419.
10. de Vos, W. M., Bakker, H., Saccone, C. & Kroon, A. M. (1980) *Biochim. Biophys. Acta* **607**, 1–9.
11. Goddard, J. M., Masters, J. N., Jones, S. S., Ashworth, W. D., Jr., & Wolstenholme, D. R. (1981) *Chromosoma* **82**, 595–609.
12. Francisco, J. F. & Simpson, M. V. (1977) *FEBS Lett.* **79**, 291–294.
13. Francisco, J. F., Brown, G. G. & Simpson, M. V. (1979) *Plasmid* **2**, 426–436.
14. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **62**, 1511–1523.
15. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
16. Drouin, J. (1980) *J. Mol. Biol.* **140**, 15–34.
17. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. & Dodgson, J. (1980) *Cell* **20**, 555–566.
18. Kimura, M. & Ohta, T. (1972) *J. Mol. Evol.* **2**, 87–90.
19. Salser, W. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **42**, 985–1002.
20. Fitch, W. M. (1980) *J. Mol. Evol.* **16**, 151–209.
21. Barrell, B. G., Bankier, A. T. & Drouin, J. (1979) *Nature (London)* **282**, 189–194.
22. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* **290**, 457–465.
23. Young, I. G. & Anderson, S. (1980) *Gene* **12**, 257–265.
24. Jukes, T. H. (1980) *Science* **210**, 973–977.
25. Upholt, W. B. (1977) *Nucleic Acids Res.* **4**, 1257–1265.
26. Kimura, M. (1968) *Genet. Res. Cambridge* **11**, 247–269.
27. Kunkel, T. A. & Loeb, L. A. (1981) *Science* **213**, 765–767.
28. Kasamatsu, H. & Vinograd, J. (1974) *Annu. Rev. Biochem.* **43**, 695–719.
29. Van Valen, L. (1971) *Evolution* **25**, 420–428.
30. Kimura, M. (1968) *Nature (London)* **217**, 624–626.
31. Langley, C. H. & Fitch, W. M. (1974) *J. Mol. Evol.* **3**, 161–177.
32. Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. (1981) *Cell* **26**, 167–180.
33. Grosskopf, R. & Feldman, H. (1981) *Curr. Genet.* **4**, 151–158.