Rapid evolution of the human gene for cytochrome *c* oxidase subunit IV

(respiratory chain/primate/mitochondrial protein/nuclear gene/coevolution)

MARGARET I. LOMAX*[†], DAVID HEWETT-EMMETT[‡], TONY L. YANG*, AND LAWRENCE I. GROSSMAN[§]

*Department of Anatomy and Cell Biology, The University of Michigan Medical School, Ann Arbor, MI 48109; [†]Center for Demographic and Population Genetics, University of Texas Health Sciences Center at Houston, Houston, TX 77225; and [§]Department of Molecular Biology and Genetics, Wayne State University School of Medicine, Detroit, MI 48201

Communicated by James V. Neel, March 5, 1992

ABSTRACT We have compared the DNA sequences of nine mammalian genes for cytochrome *c* oxidase subunit IV (*COX4* genes)—four expressed genes (human, bovine, rat, and mouse) and five pseudogenes (human, chimpanzee, orangutan, squirrel monkey, and bovine)—and constructed the sequence of the ancestral mammalian *COX4* gene. By analyzing these sequences to determine the pattern and rate of nucleotide substitution in each branch of the evolutionary tree, we deduced that the human gene has evolved rapidly since the origin of the primate pseudogene approximately 41 million years ago, and we discuss the suggestion that this results from coevolution of nuclear and mitochondrial genes for cytochrome *c* oxidase.

Cytochrome *c* oxidase (COX; EC 1.9.3.1), the terminal enzyme complex of the mitochondrial electron transport chain, is an example of an unusual class of multisubunit enzyme complex found in both mitochondria and chloroplasts of eukaryotic cells. The novel feature of these complexes is their mixed genetic origins: in each complex, at least one of the polypeptide subunits is encoded in the organelle genome, with the remaining subunits encoded in the nucleus. Thus, two distinct genetic systems, each with its unique features and evolutionary constraints, must interact to produce these essential holoenzymes.

The existence of enzyme complexes encoded in two genetic compartments (1) affords opportunities to investigate the interaction of evolutionary pressures. For example, the mitochondrial genome in primates appears to accumulate mutations at a rate higher than that of single-copy nuclear DNA (2), although the rapid evolution of mitochondrial DNA is known to consist primarily of silent changes. On the other hand, for the nuclear genome, recent studies have shown that the rates of sequence evolution have changed in different phylogenetic groups (3) and have decreased in the primate lineage leading to humans (4–6). These changes in rate affect both substitutions leading to amino acid replacements and silent, presumably neutral, changes, although the mechanism(s) leading to this decreased rate in higher primates is unknown.

To understand whether the high rate of nucleotide substitution observed in primate mitochondrial DNA exerts selective pressure on the nuclear genes for mitochondrial proteins, we have isolated the nuclear genes for COX subunit IV (*COX4* genes) to address questions of rates of evolution of nuclear genes for mitochondrial complexes. An accelerated substitution rate in a human gene would be seen against a relatively low mutational background and would therefore provide a more sensitive test for coevolution. In this study we compared the protein sequences of COX subunit IV from four extant species—cow, man, mouse, and rat—and the

nucleotide sequences of nine *COX4* genes—human, bovine, mouse, and rat cDNAs and human, chimpanzee, orangutan, squirrel monkey, and bovine pseudogenes—to deduce the amino acid and DNA sequences of the ancestral COX IV protein and *COX4* gene at the time of the mammalian radiation and thus determine the rates of nucleotide substitution in each mammalian lineage.

MATERIALS AND METHODS

Protein (7) and most DNA sequence data analyzed herein have been published (8–14). Orangutan (*Pongo pygmaeus*) and squirrel monkey (*Saimiri sciureus*) pseudogenes were isolated by PCR. Two primers based on conserved regions of the *COX4* gene were synthesized. The 5' primer includes the initiation methionine codon (5'-GGCAGAATGGCTAC-CAG-3'); the 3' primer includes the stop codon (5'-CAGCATCTCTCACTTCTTCCACTC-3'). PCR was performed under standard reaction conditions for 30 cycles in a Perkin-Elmer/Cetus thermal cycler with the following cycling conditions: 94°C, 45 sec; 55°C, 1 min; 72°C, 1.5 min. The resulting fragments were cloned in the *Sma* I site of pGEM3zf(+) (Promega) and sequenced (15) on both strands.[¶]

The ancestral protein and DNA sequences were deduced by the method of maximum parsimony (16) from sequence alignments of extant sequences. Evolutionary comparisons were performed according to the methods of Nei and Gojobori (17) and Saitou and Nei (18), using computer programs rewritten for the IBM PC.

RESULTS

Evolution of COX IV Proteins. COX subunit IV is the largest nucleus-encoded subunit of the COX complex. The mature subunit is composed of 147 amino acids, but the initial translation product is a precursor polypeptide 169 amino acids long. The additional 22 residues make up an N-terminal presequence that directs the protein to the mitochondrial outer membrane and facilitates import into the inner membrane. The amino acid sequence of the mature bovine heart COX IV protein has been determined (7) and the sequence of the N-terminal extension has been deduced from the bovine cDNA (8). Similarly, the complete amino acid sequences of the human (10), rat (12), and mouse (13) precursors have been deduced from full-length cDNAs.

The high degree of amino acid conservation in this protein (80–86% identity) allowed us to deduce the ancestral mam-

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.

Abbreviations: COX, cytochrome *c* oxidase; S and N, synonymous and nonsynonymous substitutions; Myr, million years.

[†]To whom reprint requests should be addressed.

[¶]The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M91033 and M91032).

malian protein sequence at the time of the mammalian radiation by the method of maximum parsimony (Fig. 1). It was possible to determine this sequence with certainty at 165 of 169 positions, since at each of these positions two of the three lineages (primate, bovine, or rodent) contained a common residue. At the remaining four positions there were different amino acids in each protein, a circumstance that did not permit unambiguous determination of the ancestral sequence by this method. For three of the four codons, however, we could deduce the likeliest ancestral codon sequence by determining the minimum number of substitutions necessary to convert it into each of the three present codons. The data presented in Fig. 1 indicate that the bovine lineage has sustained the fewest ($n = 8$) substitutions, whereas the human, rat, and mouse proteins have accumulated considerably more substitutions ($n = 20, 15$, and 19 , respectively).

Evolutionary Analysis of *COX4* Genes. The high number of amino acid replacements in the human COX IV protein was surprising in view of the overall slowdown of the evolution of primate nuclear genes (4–6). To gain further insights, we turned to analysis of DNA sequences. By comparing the four extant expressed genes (cDNAs) and several pseudogenes we were able to deduce the nucleotide sequence of four ancestral *COX4* genes (Fig. 2). ("Ancestor" refers to the sequence from which the primate, rodent, and artiodactyl lineages emerge.) We assume for our analysis that the three lineages emerge from a single node (Fig. 3, node 1), although a branched lineage has recently been suggested (20). Primate and bovine ancestors refer to genes present at the time that the respective pseudogenes arose. The rodent ancestor is the sequence present at the time of the mouse/rat divergence.

The deduced sequence shown in Fig. 2 is of the ancestral mammalian *COX4* gene; only the nucleotide substitutions that occurred in each lineage leading to the present gene are presented. Most of the third-position substitutions in the human gene are also present in the primate pseudogenes, a fact indicating that these silent substitutions occurred before the pseudogene arose. Surprisingly, however, most of the substitutions that are unique to the human gene are nonsynonymous. Thus, most of the amino acid substitutions in the human gene occurred *after* the primate pseudogene arose. Although we address the age of the pseudogene below, the preferential occurrence of nonsynonymous substitutions in the human gene is independent of the age of the pseudogene, since the human gene and the pseudogene have been evolving for an equal time since their divergence from a common ancestor.

Statistical Analysis of Nucleotide Substitutions. We determined the evolutionary relationships among seven of these sequences by the neighbor-joining method of Saitou and Nei (18), which groups sequences on the basis of corrected Jukes–Cantor evolutionary distances. These relationships are shown in Fig. 3. We also determined the fraction of nonsynonymous (N) and synonymous (S) substitutions that occurred in each of the *COX4* genes by using the statistical method of Nei and Gojobori (17). The fraction of N or S substitutions per N or S site and the ratio of these values are given on each branch of the evolutionary tree. The ratio N/S, a comparison of replacement substitution with silent substitution, reflects the relative effect of these two rates within any one lineage and is thus independent of time. Striking differences in N/S ratios are observed when these values are computed separately for the intervals (i) between the ancestral mammalian gene (Fig. 3, node 1) and the ancestral primate gene (Fig. 3, node 2), and (ii) between the ancestral primate gene and the present human gene. For the first period, before the origin of the primate ancestral gene, an 8-fold higher fraction of silent vs. replacement substitutions is seen, yielding an N/S ratio of 0.13. During the second period, however, the ratio of N to S (N/S) is 0.59, which compares with the theoretical pseudogene ratio of 1.0. In the primate lineage, therefore, there appear to be two distinct periods of evolution of the *COX4* gene: the period before the pseudogene arose, when there were few replacement substitutions, and the period after the pseudogene arose, when there was a larger number of nonsynonymous substitutions.

The analysis carried out by the Nei and Gojobori method and shown in Fig. 3 was repeated by utilizing two other methods: the Li method assuming unweighted changes, and the Li method using weighted values for amino acid replacements (21) (data not shown). The two methods gave identical values that were within an average of 15% of the Nei–Gojobori values. The values in Fig. 3 were limited by the incomplete bovine cDNA data (8) and were therefore generated with a data set of the 127 N-terminal codons. We constructed the missing bovine cDNA sequence on the basis of the protein data and the assumption that, at each ambiguous position, the conserved nucleotide is used. If a data set of the full 170 codons is now used, the N/S ratio for the higher primate lineage is even higher (0.87). This reflects the fact that of the 6 substitutions in the C-terminal 43 codons, 5 are replacements and only 1 is synonymous. Therefore, neither the method used to analyze the data nor the length of the data set alters the conclusion that the *COX4* gene has accumulated

FIG. 1. Ancestral COX IV protein sequence deduced from the extant human, bovine, mouse, and rat protein sequences. The sequence of the 169 amino acid residues of the ancestral COX subunit IV precursor polypeptide has been deduced from comparison of the proteins and cDNAs from these four extant species. The sequence is numbered from the initiator methionine through the mature protein, which begins at residue 23. The deduced sequence of the ancestral protein is presented on the top line. Only the amino acid substitutions in each extant protein are shown below it. At positions 54, 75, 106, and 136, the ancestral sequence was deduced from the DNA sequence (see text and Fig. 2). It was not possible to deduce the sequence of codon 135; therefore, this ambiguous codon is represented by * and the amino acids shown are not tabulated as changes. The processing site (filled arrow), transmembrane region (box), and clustered human substitutions (horizontal bars) are shown. Open arrows show position of introns in the rat gene (19).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Ancestor (1)	ATG	TTG	GCT	ACC	AGG	GTA	TTT	AGC	CTA	ATT	GGC	AGA	CAG	GCA	ATT	TCC	ACC	TCG	GTG	TGT	GTT	CGG	GCA	CAT	GGA	AGT	GTT	GTG	AAG	AGT	GAA	GAC	TAT	GCT	CTC	CCA	TAT	GTG	GAT	
Primate Ancestor (2)	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Human	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Human Pseudogene	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Chimp Pseudogene	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Orangutan Pseudogene	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Sq Monkey Pseudogene	...	...	...	...	A	...	...	...	...	...	...	...	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Bovine Ancestor (4)	...	...	A	...	A	...	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Bovine	...	...	A	...	A	...	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Bovine Pseudogene	...	...	A	...	A	...	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Rodent Ancestor (5)	...	...	G	...	A	...	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Rat	...	...	G	...	A	...	...	...	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Mouse	...	...	T	...	C	G	C	G	...	...	...	G	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Ancestor (1)	CGG	CCT	GAC	TAC	CCC	TTG	CCC	GAT	GTG	GCC	CAT	GTC	AAG	AAC	CTG	TCT	GCC	AGC	CAG	AAG	GCC	CTG	AAG	GAG	AAG	GAG	AAG	GCC	TCC	TCC	AGC	AGC	CTC	TCC	ATG	GAT	GAG	AAA	GTT	GAG
Primate Ancestor (2)	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Human	...	...	...	C	...	...	G	G	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Human Pseudogene	...	...	...	T	...	G	A	...	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Chimp Pseudogene	...	...	...	T	...	G	A	...	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Orangutan Pseudogene	...	...	...	T	...	G	A	...	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Sq Monkey Pseudogene	...	...	A	...	...	...	...	...	...	...	...	G	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Bovine Ancestor (4)	...	...	...	T	...	...	...	...	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Bovine	...	...	...	T	...	...	...	...	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Bovine Pseudogene	...	...	...	T	...	...	...	...	...	...	...	...	...	C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...							

FIG. 2. Deduced amino acid and nucleotide sequences of the ancestral *COX4* gene and comparison with the nucleotide sequences of seven mammalian *COX4* genes. The complete nucleotide sequence of the deduced ancestral *COX4* gene and its predicted amino acid sequence are shown on the top lines. Codons are numbered with the initiation methionine codon as codon 1. Only nucleotide changes are indicated below this sequence; identical nucleotides are indicated by dots. The ambiguous codon (*) and nucleotides (•) are also denoted. The sequence of the ancestral primate *COX4* gene was deduced from the human cDNA and the human and chimpanzee pseudogenes; likewise, the ancestral bovine gene was deduced from the bovine cDNA and the bovine pseudogene. Since the bovine cDNA sequence is truncated at amino acid residue 106 (8), we used only sequence data up to codon 106 of the mature protein (codon 128 of Fig. 2) for this analysis. Deletions in the pseudogenes and the stop codons are indicated by hyphens (-). Numbers in parentheses after ancestors represent nodes in Fig. 3.

a large number of replacement substitutions in higher primates.

Time of Origin of the Primate Pseudogenes. To estimate the time of origin of the pseudogene we analyzed genomic DNA from several primates by PCR (Fig. 4). The primers used for this analysis should amplify a 525-bp fragment of the pseudogene, but will not amplify the expressed gene (19). A 525-bp product was seen in human and orangutan, as expected, and also in both Old World (macaque) and New World (squirrel) monkeys, but not in a prosimian (galago). Thus, *COX4* pseudogenes were present as early as 35–45 million years (Myr) ago (22, 23).

Comparison of the sequences we obtained by PCR with the human and chimpanzee pseudogenes (Fig. 2) indicates that

both orangutan and squirrel monkey contain orthologous pseudogenes. The orangutan pseudogene contains the two deletions and one insertion (codons 149, 153, and 157) seen in both the human and chimpanzee pseudogenes. Although the squirrel monkey gene lacks these characteristic changes, it has 14 nucleotide changes in common with the other three pseudogenes.

We also have estimated the time of origin of the primate pseudogene by analyzing the rate of synonymous substitution in the human and chimpanzee genes, assuming a constant rate of nucleotide substitution in pseudogenes. If one applies the value of 2×10^{-9} substitution per site per year obtained for primate pseudogenes (6, 23), the time of origin is estimated at 41 Myr ago, in agreement with the physical evidence from PCR.

FIG. 3. Evolutionary relationships of *COX4* genes. The extant *COX4* sequences were analyzed by the neighbor-joining method of Saitou and Nei (18). The lengths of the branches represent percent of total sites substituted, based on corrected Jukes-Cantor evolutionary distances. The ratio of nonsynonymous to synonymous substitutions (N/S ratio) on each lineage is presented above the line, followed by the actual N and S values in parentheses.

DISCUSSION

We have used primate pseudogenes to demarcate two periods in the evolution of the *COX4* gene. These two periods exhibit striking differences in patterns of nucleotide substitutions, as reflected in the N/S ratios. During the approximately 40 Myr since the origin of the pseudogene, the human *COX4* gene has sustained a high proportion of nucleotide substitutions during a period of overall decrease of mutation rate. These nucleotide substitutions result in a high proportion of amino acid replacements, as evidenced by the fact that the N/S ratio for the expressed gene is higher than that of the primate pseudogene.

The high relative rate of nonsynonymous substitution in the human gene could result by two mechanisms. One is that it is a reporter for an increased overall mutation rate. The second is that it is specific to subunit IV and is directed by evolutionary selection. Our data do not support an accelerated substitution rate in higher primates. The alternative view, the directed substitution model, might be implemented

FIG. 4. PCR amplification of *COX4* pseudogenes. Genomic DNA samples (100 ng) from human (H), orangutan (O), pigtail macaque (*Macaca nemestrina*, M), squirrel monkey (S), and bushbaby (*Galago crassicaudatus*, G) were primed and amplified with a primer pair for the human cDNA designed to give a 525-base-pair (bp) product. Aliquots of the PCR products were analyzed on a 2.0% agarose gel with the 100-bp ladder (BRL) as molecular length standards (m). -, No-DNA control.

by the *COX4* gene evolving in concert with a mitochondrially encoded subunit or subunits.

A case that has been proposed as a paradigm of coevolution between a nuclear and a mitochondrial gene is the interaction between cytochrome *c* and subunit II of COX (24). Both proteins appear to have undergone a high rate of substitution in the primate lineage compared with other mammalian lineages (24–26), with the biochemical consequence that primate cytochrome *c* shows a higher reaction rate with primate COX than with nonprimate COX (27). Furthermore, the extensive analysis of cytochrome *c* has allowed two periods of evolution to be observed: a short period of rapid evolution early in the primate lineage and a period of much slower evolution in the higher primates (25). These data contrast with our observation of the period of rapid evolution of *COX4* in the higher primates. Additional data will be required to deduce whether the rapid evolution we observed is the result of a short burst or is part of a continuing selection.

It is difficult to evaluate such a coevolution model without more detailed knowledge of the protein contacts subunit IV makes in the holoenzyme. However, if we picture the mechanism for coevolution as formally analogous to extragenic suppression, then directed substitutions might be expected to be arrayed so as to affect a particular function, as in a cluster that constitutes a protein–protein contact. The amino acid changes in human subunit IV were examined for the presence of clusters. When compared with residues conserved in cow and rat, amino acid changes are found in two clusters, one at the N terminus of the mature protein and containing 7 changes in 24 residues, the other near the end of the hydrophobic transmembrane region and containing 6 changes in 15 residues (Fig. 1). Conversely, the number of amino acid replacements in the bovine and rat COX IV subunits in these regions is minimal. Thus, the coevolution model is consistent with these qualitative considerations. The relation of these clusters to the exon–intron junctions in rat subunit IV (21) was also examined. No correlation between the clusters of changes and the junctions is apparent.

Coevolution may be a mechanism by which multicomponent systems can acquire new or altered functions. A limitation in the analysis of coevolution of COX is that its functions and interactions are only partially understood. Thus, it is difficult, for instance, to evaluate the properties of the complex that are being altered by recent evolutionary changes in lower vs. higher primates. It is nevertheless interesting that at least one other subunit has recently undergone a significant change in humans: subunit VIII, which has at least two tissue isoforms in other animals examined (28–30), appears to contain only one form in humans (31, 32). Taken together, these observations suggest that the rapid evolution of COX II and cytochrome *c* prior to the divergence of humans from Old World monkeys and the more recent rapid evolution of subunit IV that we present here are steps by which different parts of the complex are altering or modulating parts of their function.

The age we calculate for the primate pseudogenes from the number of substitutions depends upon the assumption of substitution rate; it ranges from 17 Myr when we assume a rate based on the average mammalian rate (21) to 45 Myr when we utilize a rate based on the proposed slowdown in the rate of primate evolution (6). The evidence for a rate slowdown in primates has recently been reevaluated and criticized (33). We note, however, that our physical evidence (Fig. 4) indicates the appearance of the pseudogene 35–45 Myr ago and supports the lower rate.

We thank Morris Goodman, Wen-Hsiung Li, and R. S. Seelan for comments on the manuscript. We also thank Pat Venta, Linda Samuelson, and Morris Goodman for orangutan, squirrel monkey, and galago DNAs, respectively. This work was supported by a

University of Michigan Biomedical Research Support Grant, National Institutes of Health Grant GM37086, National Science Foundation Grants DMB-90-05580 (M.I.L.) and DMB-89-05285 (L.I.G.), and the Wayne State University Center for Molecular Biology (RR05384) (L.I.G.).

1. Chomyn, A. & Attardi, G. (1987) *Curr. Top. Bioenerg.* **15**, 295–329.
2. Brown, W. M., Prager, E. M. & Wilson, A. C. (1982) *J. Mol. Evol.* **18**, 225–239.
3. Wu, C. I. & Li, W.-H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1741–1745.
4. Britten, R. J. (1986) *Science* **231**, 1393–1398.
5. Goodman, M. (1986) *Bioessays* **3**, 9–14.
6. Li, W.-H. & Tanimura, M. (1987) *Nature (London)* **326**, 93–95.
7. Sacher, R., Steffens, G. J. & Buse, G. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 1385–1392.
8. Lomax, M. I., Bachman, N. J., Nasoff, M. S., Caruthers, M. H. & Grossman, L. I. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6295–6299.
9. Bachman, N. J., Lomax, M. I. & Grossman, L. I. (1987) *Gene* **55**, 219–229.
10. Zeviani, M., Nakagawa, M., Herbert, J., Lomax, M. I., Grossman, L. I., Sherbany, S., Miranda, A. F., DiMauro, S. & Schon, E. A. (1987) *Gene* **55**, 205–217.
11. Lomax, M. I., Welch, M. D., Darras, B. T., Francke, U. & Grossman, L. I. (1990) *Gene* **86**, 209–216.
12. Gopalan, G., Droste, M. & Kadenbach, B. (1989) *Nucleic Acids Res.* **17**, 4376.
13. Grossman, L. I. & Akamatsu, M. (1990) *Nucleic Acids Res.* **18**, 6454.
14. Ewart, G., Lightowlers, R., Zhang, Y.-Z., Balan, V. J., Kennaway, N. & Capaldi, R. A. (1990) *Biochim. Biophys. Acta* **1018**, 223–224.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Fitch, W. M. (1971) *Syst. Zool.* **20**, 406–416.
17. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3**, 418–426.
18. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
19. Yamada, M., Amuro, N., Goto, Y. & Okazaki, T. (1990) *J. Biol. Chem.* **265**, 7687–7692.
20. Li, W.-H., Gouy, M., Sharp, P. M., O'hUigin, C. & Yang, Y.-W. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6703–6707.
21. Li, W.-H., Luo, C. C. & Wu, C. I. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 1–94.
22. Goodman, M. (1985) in *Evolutionary Perspective and the New Genetics*, eds. Gershowitz, H., Tashian, R. & Rucknagel, D. (Liss, New York), pp. 121–132.
23. Koop, B. F., Tagle, D. A., Goodman, M. & Slightom, J. L. (1989) *Mol. Biol. Evol.* **6**, 580–612.
24. Cann, R. L., Brown, W. M. & Wilson, A. C. (1984) *Genetics* **106**, 479–499.
25. Evans, M. J. & Scarpulla, R. C. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9625–9629.
26. Ramharack, R. & Deeley, R. (1987) *J. Biol. Chem.* **262**, 14014–14021.
27. Osheroff, N., Speck, S. H., Margoliash, E., Veerman, E. C. I., Wilms, J., König, B. W. & Muijsers, A. O. (1983) *J. Biol. Chem.* **258**, 5731–5738.
28. Yanamura, W., Zhang, Y.-Z., Takamiya, S. & Capaldi, R. A. (1988) *Biochemistry* **27**, 4909–4914.
29. Kuhn-Nentwig, L. & Kadenbach, B. (1984) *FEBS Lett.* **172**, 189–192.
30. Lightowlers, R., Ewart, G., Aggeler, R., Zhang, Y.-Z., Calavetta, L. & Capaldi, R. A. (1990) *J. Biol. Chem.* **265**, 2677–2681.
31. Van Kuilenberg, A. B. P., Muijsers, A. O., Demol, H., Dekker, H. L. & Van Beeumen, J. J. (1988) *FEBS Lett.* **240**, 127–132.
32. Rizzuto, R., Nakase, H., Darras, B., Francke, U., Fabrizi, G. M., Megele, T., Walsh, F., Kadenbach, B. & Schon, E. (1989) *J. Biol. Chem.* **264**, 10595–10600.
33. Easteal, S. (1991) *Mol. Biol. Evol.* **8**, 115–127.