Am J Hum Genet 35:167–176, 1983

Amino Acid Change Associated with the Major Polymorphic Hinc II Site of Oriental and Caucasian Mitochondrial DNAs

HUGUES BLANC,[1,2] KUANG-HO CHEN,[1,3] MELANIE A. D'AMORE,[1] AND DOUGLAS C. WALLACE[1,4]

SUMMARY

The mitochondrial DNAs (mtDNAs) from 116 Oriental and Caucasian blood samples were analyzed for their Hinc II restriction endonuclease cleavage patterns using Southern analysis and ^{32}P human mtDNA probes. Seven distinct patterns were found, all of which could be interrelated by single nucleotide changes. The predominant pattern (mtHinc II-2) was found in 97% of the Caucasians and 73% of the Orientals. This mtDNA morph had one more Hinc II site than did the second most common morph (mtHinc II-1), which was found only in 20% of the Orientals. Three additional patterns were in a single Oriental sample, a fourth in a single Caucasian sample, and a fifth in one member of each population. The polymorphic site that differentiated mtHinc II-1 and mtHinc II-2 was cloned and sequenced. A single nucleotide change was found that created an Hinc II site and changed the amino acid sequence of the URF5 gene. Comparison of these sequences with those of other primates [15] revealed that the Asian mtHinc II-1 and mtHinc II-4 mtDNAs were identical in this region with those of chimpanzees and orangutans. These results suggest that the Asian mtHinc II-1 mtDNA may have been ancestral to other human mtDNAs.

Received June 8, 1982; revised September 7, 1982.

This work was supported by grants GM-28428 and GM-24285 from the National Institutes of Health, grant PCM 80-21871 from the National Science Foundation, and March of Dimes Basic Research Grant 1-778 awarded to D. C. W.

[1] Department of Genetics, Stanford School of Medicine, Stanford, CA 94305. Request reprints from D. C. W.

[2] Present address: Centre National de La Recherche Scientifique, Centre de Genetique Moleculaire, 91190 Gif-Sur-Yvette, France.

[3] Present address: Department of Mathematics, California State University, Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840.

[4] Present address: Departments of Biochemistry and Pediatrics, Woodruff Memorial Building, Emory Medical School, Atlanta, GA 30322.

INTRODUCTION

Extant human mtDNAs have been proposed to have radiated from a single mtDNA type in the relatively recent past [1]. This conclusion was based on the frequency of variation of restriction endonuclease sites in human mtDNA [1–3] and the conclusion that mammalian mtDNA sequence varies more rapidly than comparable nuclear DNA sequences [4].

Analysis of Hpa I-site variation in blood cell mtDNAs derived from 235 individuals of five ethnic groups (Oriental, Caucasian, and three African) revealed that certain restriction patterns correlate highly with the ethnic origin of the individual. Six Hpa I restriction patterns were observed in this sample, each related to all of the others by a series of single site changes. The simplest Hpa I restriction pattern (mtHpaI-1) resulted from mtDNAs with only two sites and was found in 12.5% of the Orientals. The most common mtDNA differed from mtHpaI-1 by the acquisition of a new restriction site. This mtDNA was found in 80% of the Orientals and 98% of the Caucasians. A third mtDNA (mtHpaI-3) contained yet another site and was found only in Africans [5].

Comparison of the human Hpa I restriction patterns with those of other primates [6, 7] suggested that the Oriental mtHpaI-1 mtDNA was most closely related to other primate mtDNAs and thus possibly ancestral to human mtDNA radiation [5]. The Hpa I site change between mtHpaI-1 and mtHpaI-2 was also associated with the appearance of a new Hinc II restriction site. Since the recognition site of Hpa I (GTTAAC) is closely related to that of Hinc II (GTYRAC, where Y is either C or T and R is either G or A), it seemed likely that the same base change created both new sites.

To obtain more information on this pivotal mutation, we analyzed the Hinc II cleavage patterns of mtDNAs from 116 Oriental and Caucasian blood samples. These results confirmed the importance of this polymorphic site in the differentiation of the Asian and European populations. We then cloned and sequenced the mtDNA region containing this site from one Caucasian who had the site and one Oriental who lacked it. These sequences not only revealed the basis of the mutation but also provided further support for the hypothesis that the mtHpa I-1 and mtHinc II-1 mtDNA was ancestral to the other human mtDNAs.

MATERIALS AND METHODS

MtDNA Preparation and Analysis

MtDNAs were extracted from platelets isolated from peripheral blood [8]. Blood samples were obtained from independent donors of the following origins: 61 Caucasians (U.S.A. and Europe) and 55 Orientals (53 from Taiwan and mainland China and two from Japan).

Restriction endonuclease digests were incubated at 37°C for 4 hrs using the buffers recommended by the manufacturers. Approximately 0.2 μg of mtDNA was used in single digests and 0.5 μg in double digests. Double Hinc II and Bgl I and Hinc II and Hind III digests were done in Hinc II buffer. The double Hinc II and Xho I digests were done sequentially. Restriction fragments were separated on 0.7% horizontal agarose slab gels [9] and transferred to cellulose nitrate [10]. Hybridizations of the filters were performed with ^{32}P nick-translated mtDNA purified from HeLa cells [5].

MtDNA Cloning and Sequencing

Circular mtDNAs were isolated from platelets obtained from 400 ml of blood [8]. Blood from donors was tested to verify the absence of hepatitis surface antigen prior to cloning experiments. Closed circular DNA was purified from total platelet DNA by two successive CsCl-ethidium bromide gradients [9]. The 890 base pair (bp) Hind III fragment containing the polymorphic Hinc II *a* site (see fig. 1) was cloned from a Caucasian mtHinc II-2 mtDNA that contains the *a* site and an Oriental mtHinc II-4 mtDNA that lacks the site. These fragments were ligated into the unique Hind III site of *E. coli* plasmids pBR322 or pBR325 [11]. The recombinant plasmids were transformed into *E. coli* strain HB101 [12], and the clones screened for ampicillin resistance and tetracycline sensitivity. Restriction analysis of clones ST17/7.1 (Caucasian mtHinc II-2) and ST36/12.6 (Oriental mtHinc II-4) verified that the correct 0.89 kilobase (kb) Hind III fragments had been inserted.

The nucleotide sequence in the vicinity of the polymorphic *a* site was determined using the chemical method of Maxam and Gilbert [13]. After digestion by Hind III, recombinant plasmid DNAs ST17/7.1 and ST36/12.6 were 5′ end-labeled with T4 kinase and ^{32}P ATP (Amersham, Arlington Heights, Ill.). The two 0.89-kb mtDNA fragments were purified by electroelution from a 4% acrylamide gel and digested by Hae III. Hae III cleaves the

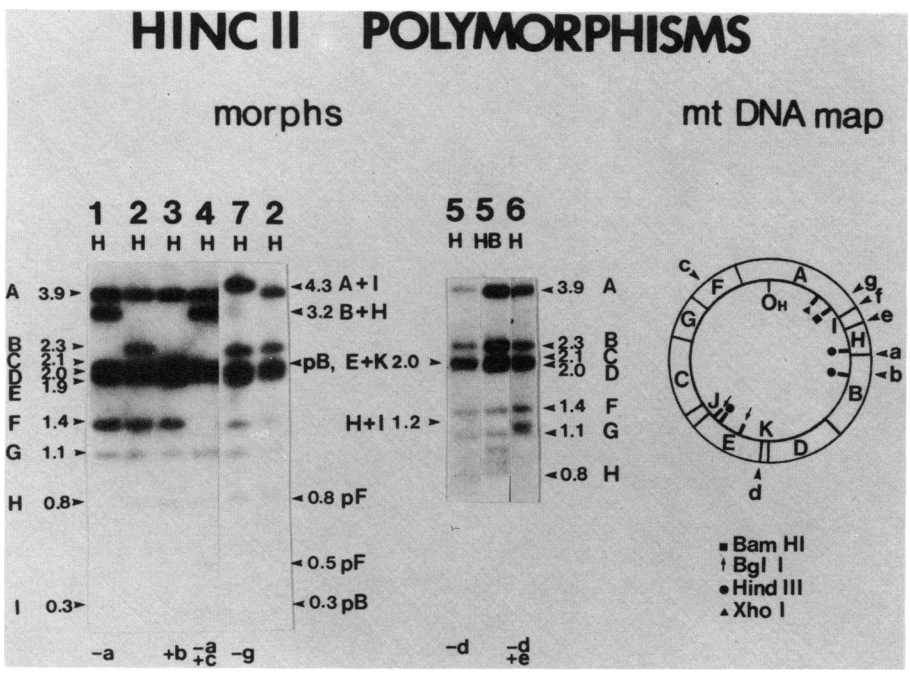

FIG. 1.—Fragment patterns and site changes of MtHinc II morphs. *Circular map* shows the fragment locations from an Hinc II digestion of the morph mtHinc II-2 mtDNA. *Peripheral arrows with letters a–q* indicate location of polymorphic sites. Site *q* is present in the published sequence [14] but was not observed in our study. *Autoradiographs* show mtDNA morphs 1–7 digested with Hinc II (*H*) and morph mtHinc II-5 digested with Hinc II and Bgl I (*HB*). *Left column of letters A–I* and *corresponding molecular weights* are those previously reported for mtHinc II-2 [2]. *Right column of letters A–D, F–H*, and *corresponding molecular weights* are characteristic of mtHinc II-5. *Center column of letters* and *molecular weights* indicate locations of polymorphic fragments in various morphs. *A+I* = mtHinc II-7, *B+H* = mtHinc II-1 and -4, *pB* = mtHinc II-3, *E+K* = mtHinc II-5 and -6, *H+I* = mtHinc II-6, and *pF* = mtHinc II-4. *Small letters at bottom of gel lanes* indicate the loss (−) or gain (+) of a site in that morph and correspond to *small letters on the mtDNA map*.

0.89-kb Hind III fragment once, generating two fragments 620 and 270 bp. The 270-bp fragment containing the Hinc II polymorphic site *a* was purified by electroelution from an acrylamide gel and sequenced.

RESULTS

Hinc II mtDNA Polymorphisms

MtDNAs from 116 individuals were analyzed for Hinc II polymorphic sites. Seven distinct cleavage patterns were observed (see fig. 1). Each pattern was assigned a morph number (mtHinc II-1 to mtHinc II-7). The mtDNA circular map of morph mtHinc II-2 has already been established [2]. Morph mtHinc-2 has 11 fragments labeled A to K and was used as the reference morph (see fig. 1).

MtHinc II-1 differs from mtHinc II-2 by the absence of fragments B and H. These are combined in a new fragment at 3.2-kb (fig. 1). This change has been confirmed because the Hpa I site *c* is also lost [5].

Similarly, mtHinc II-5 differs from mtHinc II-2 in that fragments E and K are missing and a new 2.0-kb fragment is found (fig. 1). The loss of the Hinc II *d* site would account for this change. This was verified by a Hinc II and Bgl I double digest, which yielded the predicted 1.0-kb and 0.9-kb fragments (fig. 1, column 5HB).

Morph mtHinc II-7 differs from mtHinc II-2 by the loss of fragments A and I in the acquisition of a new band at 4.3 kb. A double digest with Hinc II and Xho I gave fragments of 1.6 and 2.6 kb. Thus, the 4.3-kb fragment is the fusion of A plus I due to the loss of Hinc II site *f*.

Morph mtHinc II-3 is characterized by the absence of mtHinc II-2 fragment B (2.3 kb) (fig. 1). This could be explained by the acquisition of a new Hinc II site *b* in the B fragment yielding a new 2.0-kb fragment in the region of the C, D, and E bands and a 0.3-kb fragment in the region of the I fragment. This interpretation was confirmed because the change corresponded to the previously mapped Hpa I site *e* [5].

Morphs mtHinc II-4 and -6 probably differ from mtHinc II-2 by two mutations. MtHinc II-4 has the 3.2-kb fragment of mtHinc II-1 resulting from the absence of the *a* site and fusion of morph mtHinc II-2 fragments B and H. In addition, it has lost the fragment F, probably because of the acquisition of a new Hinc II site, site *c* (fig. 1), generating a fragment at the same position as H and another at about 0.5 kb. This mutation also created the new Hpa I site *q*, permitting its confirmation [5]. Morph mtHinc II-6 can be linked to mtHinc II-2 through mtHinc II-5. Both mtHinc II-5 and -6 possess a fragment corresponding to the fusion of E and K. Moreover, mtHinc II-6 exhibits a new fragment at 1.2 kb that was generated by the fusion of fragments H and I at site *e* (fig. 1). This was confirmed by double digestion with Hinc II and Hind III yielding, among others, the expected 1.0- and 0.2-kb fragments. Each of the seven Hinc II morphs are thus interrelated by single site changes. This is diagrammed in figure 2.

Frequency of the Different Hinc II Morphs

Several Hinc II morphs differed both in frequency and distribution within the two populations (table 1). Morph mtHinc II-2 was the most frequently encountered,

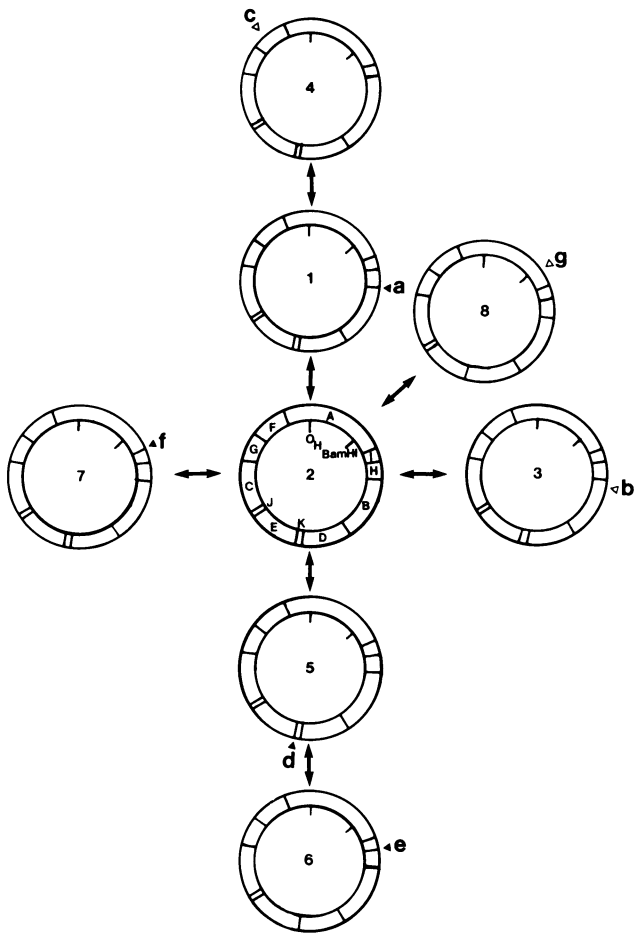

FIG. 2.—Relation between the mtHinc II morphs. *Each circle* represents an mtHinc II morph with *the morph no. in center of circle. Letters A through K* (mtHinc II-2) show locations of the Hinc II fragments. *Closed triangles* on the perimeter of the circles indicate loss of an Hinc II site (*a, d, e, f*). *Open triangles* indicate acquisition of an Hinc II site (*b, c, q*). *Double ended arrows* connect those morphs that could be interconverted by single site changes.

being present in 85% of all individuals sampled. However, the Caucasian and Oriental samples had significantly different frequencies of this morph. Ninety-seven percent of the Caucasian mtDNAs were mtHinc II-2 as compared to 73% of the Oriental mtDNAs. Morph mtHinc II-1 was found in 9.5% of the total population sampled and was the second most common mtDNA pattern. This morph was unique to Orientals, in whom it was present in 20% of the population. This was the most striking difference between the two groups.

The five other morphs were rare. Morphs mtHinc II-3, -4, and -6 were each found in one Oriental sample. Morph mtHinc II-7 was observed in a single Caucasian sample, while morph mtHinc II-5 was found in one Caucasian and one Oriental sample. Of all seven morphs observed, six (1–6) were found among

TABLE 1

FREQUENCY OF MTHINC II MORPHS

ETHNIC GROUP	PORTION OF SAMPLE	MTHINC II MORPH NO. 1	2	3	4	5	6	7	TOTAL
Caucasians	No. individuals.....	0	59	0	0	1	0	1	61
	%	0.00	96.7	0.00	0.00	1.64	0.00	1.64	100
Orientals	No. individuals.....	11	40	1	1	1	1	0	55
	%	20.0	72.7	1.82	1.82	1.82	1.82	0.00	100
Total	No. individuals.....	11	99	1	1	2	1	1	116
	%	9.48	85.4	0.86	0.86	1.72	0.86	0.86	100

NOTE: For each morph, table indicates no. individuals with that mtDNA in each population and percentages of the population they represent. Tests of homogeneity revealed that the Caucasian and Oriental populations are significantly different at the 95% level with a χ^2 of 18.37 at 6 df. Further, they differ significantly in the frequency of the mtHinc II-2 mtDNA with a χ^2 of 3.84 at 1 df.

Orientals but only three (2, 5, and 7) were found among Caucasians. Thus, there is much more variation among Asian mtDNAs than among European mtDNAs. Within each population, all of the morphs observed can be interrelated by single nucleotide changes.

Nature of the Variation at Hinc II Site a

To determine the molecular basis of the major polymorphic Hinc II site *a*, the smallest Hind III fragment was cloned from a representative Caucasian mtDNA with the site (mtHinc II-2) and an Oriental mtDNA without the site (mtHinc II-4). Comparison of the nucleotide sequences of these two mtDNA fragments over more than 100 bp, from nucleotide 12328 to nucleotide 12434 (numbered according to [14]), showed only one difference: a C to T transition at position 12406 on the H strand of the Oriental mtDNA (fig. 3). This C to T transition alters the Hinc II recognition sequence resulting in the loss of site *a* and the fusion of Hinc II fragments B and H.

In Caucasians, the 5′ to 3′ sequence is GTTAAC. However, in the Oriental morph, the sequence is 5′ ATTAAC 3′. This transition occurs in URF5, an amino acid coding region of unknown function [14]. The nucleotide change alters the

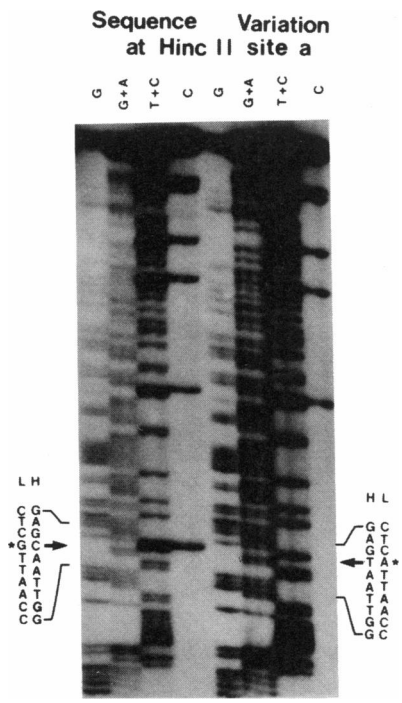

FIG. 3.—Sequence variation at Hinc II site *a*. Sequencing gel of Caucasian morph mtHinc II-2 (*left*) and Oriental morph mtHinc II-4 (*right*) showing nucleotides 12328 to 12434 (numbering according to [14]). Sequence in the region of Hinc II *a* site is interpreted with *letters*: *A* = adenylate, *G* = guanidylate, *C* = cytidylate, and *T* = thymidylate. *Asterisk* shows location of nucleotide change. *H* indicates heavy strand and *L* indicates light strand.

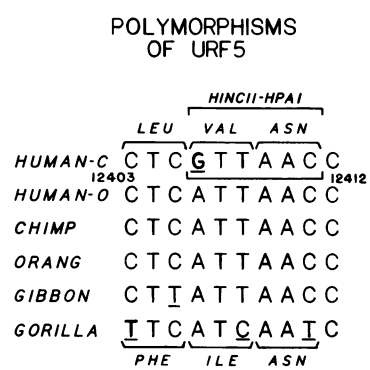

FIG. 4.—Polymorphisms of URF5. Sequence of primate mtDNAs from nucleotides 12403 to 12412 [14] that includes the polymorphic Hinc II *a* site in URF5. *Human-C* is the Caucasian mtHinc II-2 morph (this paper and [14, 15]). *Human-O* is the Oriental mtHinc II-4 morph (this paper). *Remaining primate sequences* are from Brown et al. [15]. Amino acids corresponding to the codons are shown *above and below the sequences*. G at 12406 shows the nucleotide transition. This causes the URF5 amino acid change from isoleucine in the primate and Oriental mtHinc II-4 mtDNAs to a valine in the Caucasian mtHinc II-2 mtDNA.

amino acid coding sequence, a valine codon occurring in the Caucasian mtDNA and an isoleucine codon in the Oriental mtDNA (fig. 4).

DISCUSSION

A high degree of polymorphism has been found in the Hinc II restriction sites of Oriental mtDNAs relative to those of Caucasians. This confirms similar observations made with other enzymes ([3, 5] and B. Cann and A. Wilson, personal communication, 1981).

Of the seven site changes observed, four (*a, b, e,* and *f*) were located in URF5. Site *d* is located in the region coding for cytochrome c oxidase subunit II and site *c* was located in the 16S rRNA gene. The site *q* in the eighth morph is located in URF6. Morph mtHinc II-8 was not found in our population screen but was deducted from the mtDNA sequence reported by Anderson et al. [14]. No genomic rearrangements, deletions, or insertions were detected during the course of this analysis.

The most striking feature of this study was the absence in 20% of the Oriental mtDNAs of the Hinc II site *a*, a feature common to almost all other mtDNAs. To determine the molecular nature of this Oriental variant, we sequenced this polymorphic region and discovered that this change not only altered the Hinc II and Hpa I restriction sites, but also changed the amino acid sequence of URF5. To date, the sequence of all mtDNA restriction-site polymorphisms in protein coding regions examined in man [5, 15] and rat [16] have been found not to alter the amino acid sequence of the protein. Consequently, this unique amino acid substitution raises the distinct possibility that this replacement could be affected by selective pressures.

From the sequence alone, it is impossible to determine whether the presence or absence of site *a* was ancestral. However, during our study of this region in humans, a parallel sequence analysis of this region was being conducted in apes

[17]. Comparison of the sequences from humans and apes between nucleotides 12403 and 12412 revealed that the Oriental mtHinc II-1 and mtHinc II-4 region is identical with that of chimpanzees and organutans (fig. 4). Further, the nucleotide at position 12406 of these Oriental morphs is identical with all other primates and differs only from human mtHinc II-2. Although this data cannot rule out the possibility that the A to G transition discovered in humans is not also found in other primates, it is important to note that a survey of the mtDNAs of 13 great apes did not reveal a single individual with a Hpa I or Hinc II site at this location [7]. Consequently, the most parsimonious conclusion is that this sequence was carried over from the protohominid lineage to the Oriental mtHinc II-1 mtDNA. Subsequently, this mtDNA mutated to mtHinc II-2 to yield the present predominant mtHinc II-2 mtDNA. It is unclear whether the replacement of morph mtHinc II-1 by morph mtHinc II-2 in the human population was the product of genetic drift or selection.

This result, together with the observation that there is greater variation in Hinc II sites in Orientals than in Caucasians, provides strong support for our previous proposal that the Oriental mtDNA mtHpa I-1 was ancestral to human mtDNA radiation [5]. These observations are consistent with the hypothesis that modern human ethnic groups diverged from an Asian origin [18].

ACKNOWLEDGMENTS

We recognize the excellent assistance of Mr. W. L. Hinton, III, and Ms. S. Grant and express our special appreciation to Drs. Wilson, Cann, Brown, Prager, and Wang for making their data available to us prior to publication.

REFERENCES

1. Brown W: Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc Natl Acad Sci USA* 77:3605–3609, 1980
2. Brown WM, Goodman HM: Quantitation of intrapopulation variation by restriction endonuclease analysis of human mitochondrial DNA, in *Extrachromosomal DNA*, edited by Cummings DJ, Borst P, Dawid IB, Weissman SM, Fox CF, New York, Academic Press, 1979, pp 485–499
3. Potter SS, Newbold JE, Hutchison CA III, Edgell MH: Specific cleavage analysis of mammalian mitochondrial DNA. *Proc Natl Acad Sci USA* 72:4496–4500, 1975
4. Brown WM, George M Jr, Wilson AC: Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA* 76:1967–1971, 1979
5. Denaro M, Blanc H, Johnson MJ, et al.: Ethnic variation in HpaI endonuclease cleavage patterns of human mitochonrial DNA. *Proc Natl Acad Sci USA* 78:5768–5772, 1981
6. Ferris SD, Wilson AC, Brown WM: Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. *Proc Natl Acad Sci USA* 78:2432–2436, 1981
7. Ferris SD, Brown WM, Davidson WS, Wilson AC: Extensive polymorphism in the mitochondrial DNA of apes. *Proc Natl Acad Sci USA* 78:6319–6323, 1981
8. Giles RE, Blanc H, Cann HM, Wallace DC: Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA* 77:6715–6719, 1980
9. Giles RE, Stroynowski I, Wallace DC: Characterization of mitochondrial DNA in chloramphenicol-resistant interspecific hybrids and a cybrid. *Somat Cell Genet* 6:543–554, 1980
10. Southern EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503–517, 1975

11. BLANC H, WRIGHT CT, BIBB MJ, WALLACE DC, CLAYTON DA: Mitochondrial DNA of chloramphenicol-resistant mouse cells contains a single nucleotide change in the region encoding the 3′ end of the large ribosomal RNA. *Proc Natl Acad Sci USA* 78:3789–3793, 1981
12. COHEN SN, CHANG ACY, HSU L: Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci USA* 69:2110–2114, 1972
13. MAXAM AM, GILBERT W: Sequencing end-labeled DNA with base-specific chemical cleavage. *Methods Enzymol* 65:499–560, 1980
14. ANDERSON S, BANKIER AT, BARRELL BG, ET AL.: Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465, 1981
15. JOHNSON MJ, WALLACE DC, FERRIS SD, RATTAZZI MC, LAYRISSE Z, CAVALLI-SFORZA LL: Radiation of human mitochondrial DNA types analyzed by restriction endonuclease cleavage patterns. Submitted for publication
16. CASTORA FJ, ARNHEIM N, SIMPSON MV: Mitochondrial DNA polymorphism: evidence that variants detected by restriction enzymes differ in nucleotide sequence rather than methylation. *Proc Natl Acad Sci USA* 77:6415–6419, 1980
17. BROWN WM, PRAEGER EM, WANG A, WILSON AC: Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol* 18:225–239, 1982
18. PIAZZA A, MENOZZI P, CAVALLI-SFORZA LL: Synthetic gene frequency maps of man and selective effects of climate. *Proc Natl Acad Sci USA* 78:2638–2642, 1981