Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers

(cytochrome *b*/12S ribosomal DNA/control region/evolutionary genetics/molecular phylogenies)

T. D. KOCHER, W. K. THOMAS, A. MEYER, S. V. EDWARDS, S. PÄÄBO, F. X. VILLABLANCA, AND A. C. WILSON

ABSTRACT **With a standard set of primers directed toward conserved regions, we have used the polymerase chain reaction to amplify homologous segments of mtDNA from more than 100 animal species, including mammals, birds, amphibians, fishes, and some invertebrates. Amplification and direct sequencing were possible using unpurified mtDNA from nanogram samples of fresh specimens and microgram amounts of tissues preserved for months in alcohol or decades in the dry state. The bird and fish sequences evolve with the same strong bias toward transitions that holds for mammals. However, because the light strand of birds is deficient in thymine, thymine to cytosine transitions are less common than in other taxa. Amino acid replacement in a segment of the cytochrome *b* gene is faster in mammals and birds than in fishes and the pattern of replacements fits the structural hypothesis for cytochrome *b*. The unexpectedly wide taxonomic utility of these primers offers opportunities for phylogenetic and population research.**

During the past decade, geneticists and taxonomists have used restriction endonucleases rather than sequencing to examine variation within and between species in specific segments of DNA (1–7). Although the indirect assessment of sequence variation obtained with the restriction endonuclease method is known to have many drawbacks,[§] sequence data have been difficult to obtain. The construction and screening of clone libraries has been too tedious and have demanded too much expertise for routine use by those geneticists and taxonomists who must analyze many individuals.

Dependence on restriction analysis has limited our understanding of the dynamics of DNA sequence evolution. The presence or absence of a restriction site reveals little about the kinds of nucleotide substitutions that have occurred. Thus, although restriction analysis of mtDNA from closely related mammals first showed that these genomes have a higher rate of evolutionary substitution than does nuclear DNA, the demonstration that this acceleration results mainly from an increase in the number of transitions relative to transversions came only from conventional cloning and sequencing (1, 3). Because most studies of animal mtDNA have used restriction analysis, it has been difficult to determine whether a high rate of evolution and a transition bias are characteristic of all animal mtDNAs (8–10). There has been a need for simple methods of sequencing mtDNA to examine the pattern of evolutionary substitution in other animal groups.

A fast alternative to conventional cloning has emerged in the form of the polymerase chain reaction (11). By using this method, unique sequences can be cloned *in vitro* in a matter of hours. Furthermore, the procedure is easily automated, so that hundreds of samples can be amplified each day. The enzymatic amplification of a specific DNA segment is made possible by the highly specific binding of oligonucleotide primers to sequences flanking the segment. These primers allow the binding of a DNA polymerase that then copies the segment. Because each newly made copy can serve as a template for further duplication, the number of copies of the target segment grows exponentially (12). Wrischnik *et al.* (13) found an easy way to sequence the product of this reaction directly. The direct sequencing approach has since been improved (14) and applied to the study of mtDNA variation within the human species (13, 15–17).

Because mtDNA differences among animal species are large (3), it was not expected that primers that amplified specific segments of human mtDNA would amplify the corresponding segments of mtDNA from other species. Nonetheless, we now describe three pairs of primers that amplify homologous sequences from a wide array of animals. This innovation has allowed us to gather sequence data[¶] to evaluate the pattern of molecular evolution in a variety of animal species.

MATERIALS AND METHODS

Primers. To design these primers, we compared the published sequences for mammal (18–20), frog (21), and fly (22) mtDNAs and searched for highly conserved regions. The sequences of the seven primers follow, the letters L and H refer to the light and heavy strands, and the number refers to the position of the 3′ base of the primer in the complete human mtDNA sequence (18): cytochrome *b*, L14841 (5′-AAAAAGCTTCCATCCAACATCTCAGCATGATGAAA-3′) and H15149 (5′-AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA-3′); 12S rRNA, L1091 (5′-AAAAAGCTTCAAACTGGGATTAGATACCCCACTAT-3′) and H1478 (5′-TGACTGCAGAGGGTGACGGGCGGTGTGT-3′); control region, L15926 (5′-TCAAAGCTTACACCAGTCTTGTAAACC-3′), L16007 (5′-CCCAAAGCTAAAATTCTAA-3′), and H00651 (5′-TAACTGCAGAAGGCTAGGACCAAACCT-3′).

[§]The preparation and alignment of restriction maps is itself tedious and subject to errors. Restriction sites offer a biased sample of nucleotide sequences. It is hard to compare results from different laboratories because of variation in the gel separation techniques and restriction enzymes used. When applied to mtDNA, the restriction method is reliable only among closely related taxa. Restriction analysis is limited to the small proportion of species in the biosphere that are big enough to provide milligram amounts of tissue or that can be cultured. The traces of short, modified DNA that survive in museum specimens or ancient remains cannot be analyzed with restriction enzymes.

[¶]The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M25681–M25695).

Many such priming regions exist, so that it is possible to amplify almost any segment of a mtDNA genome at will. In choosing oligonucleotide sequences, we took advantage of the evolutionary stability of regions of rRNA, the anticodon loops of tRNAs, and the active sites of enzymes. The 3′ ends of primers were located on the first or second base of codons for amino acids that are evolutionarily conserved (e.g., tryptophan). Even primers with several mismatches to the template can be used for amplification; the polymerase requires absolute matching of the primer to the template only in the last few bases of the 3′ end of the oligonucleotide.

DNA Extraction. DNA was extracted from tissues by digestion in 100 mM Tris·HCl, pH 8.0/10 mM EDTA/100 mM NaCl/0.1% SDS/50 mM dithiothreitol/proteinase K (0.5 μg/ml) for 2–4 hr at 37°C. The DNA was purified by extracting twice with phenol, once with phenol/chloroform [1:1 (vol/vol)], and once with chloroform. The sample was then concentrated by centrifugal dialysis (Centricon-30, Amicon) or ethanol precipitation.

Polymerase Chain Reaction. Amplification was performed in 100 μl of a solution containing 67 mM Tris (pH 8.8), 6.7 mM $MgSO_4$, 16.6 mM $(NH_4)_2SO_4$, 10 mM 2-mercaptoethanol, each dNTP at 1 mM, each primer at 1 μM, genomic DNA (10–1000 ng), and 2–5 units of *Thermus aquaticus* polymerase (Perkin–Elmer/Cetus). Each cycle of the polymerase chain reaction consisted of denaturation for 1 min at 93°C, hybridization for 1 min at 50°C, and extension for 2–5 min at 72°C. This cycle was repeated 25–40 times depending on the initial concentration of template DNA in the sample.

Generation of Single-Stranded DNA and Sequencing. Electrophoresis of 5 μl of the amplified mixture was done in a 2% agarose gel (NuSieve, FMC) in 40 mM Tris acetate (pH 8.0) and the DNA was stained with ethidium bromide. The gel fragment containing the amplified product was excised from the gel and melted in 1 ml of distilled water, and 1 μl of this mixture was used as the template in a second chain reaction to generate single-stranded DNA for sequencing (14). In this second reaction, the concentration of one or the other primer was reduced 100-fold. After 40 cycles of amplification, free nucleotides and salts were removed by 2–4 cycles of centrifugal dialysis. The DNA was sequenced with a commercial kit (Sequenase, United States Biochemical) and the primer that had been limiting in the second chain reaction.

RESULTS

Primers Amplify a Wide Range of Animal mtDNAs. The first pair of primers amplifies a 307-base-pair segment of the

Table 1. Amplification of mtDNA sequences from 110 animal species by using conserved primers

Type of animal	No. of species/ no. of individuals	Tissue source	Region amplified	Collaborating individual(s)
Mammal				
Rodent	12/81	S, F, P	r, b, d	W.K.T., S.P., F.X.V., and J. Patton (unpublished data)
Carnivore	2/8	F	b, d	T.D.K. and G. Shields (unpublished data)
Ungulate	16/16	S, F	b	T.D.K. and D. Irwin (unpublished data)
Primate	4/20	F, P	b, d	T.D.K. (unpublished data)
Sloth	1/1	F	b	S.P. and A. Sidow (unpublished data)
Marsupial	2/10	F, B	r, b	S.P., A. Sidow, and R. Cann (unpublished data)
Bird				
Songbird	19/22	F, P, B	r, b	S.V.E., S. Pruett-Jones, and R. Cann (unpublished data)
Gamebird	7/7	F	b	T.D.K., L. Williams, and J. Kornegay (unpublished data)
Waterfowl	1/4	F	b	T. Quinn (personal communication)
Amphibian				
Salamander	1/1	F	r, b	T.D.K. (unpublished data)
Frog	11/11	P	b	S. Carr (personal communication)
Reptile				
Crocodile	1/1	F	r, b	D. Mindell (personal communication)
Fish				
Shark	5/8	F	r, b	A. Martin (personal communication)
Cichlid	20/25	A, F, P	r, b, d	T.D.K., A.M., and P. Basasibwaki (unpublished data)
Salmonid	3/3	P	b, d	W.K.T. (unpublished data)
Coryphenid	1/3	F	r, b	R. Cann (personal communication)
Insect				
Cicada	3/12	F, P, A	r	S.P. and C. Simon (unpublished data)
Spider				
Tarantula	1/2	F	r	B. Kessing and C. Simon (personal communication)
Total	110/235			

Sources of DNA include dried skins (S) up to 80 years old, tissues preserved in alcohol (A), frozen tissues (F), blood (B), or mtDNA purified in a cesium chloride gradient (P). The segments of the mitochondrial genome amplified are from the noncoding (D-loop) region (d), and the genes encoding 12S rRNA (r) and cytochrome *b* (b). Regions not listed as amplified generally have not been examined. In some cases we suspect genome rearrangements prevent amplification of genes with flanking tRNA primers. The genera amplified are given below. Full species names are given for the 15 sequences (R1–R5, B1–B5, and F1–F5) analyzed in Figs. 2 and 3. Rodents: *Dipodomys panamintinus* (R1 and R2), *Dipodomys heermanni* (R3), *Dipodomys californicus* (R4), and *Thomomys townsendi* (R5). Carnivores: *Ursus* and *Thalarctos*. Ungulates: *Antilocapra*, *Tayassu*, *Giraffa*, *Tragulus*, *Lama*, *Camelus*, *Hippopotamus*, *Axis*, *Odocoileus*, *Bos*, *Ovis*, *Sus*, *Stenella*, *Equus*, *Diceros*, and *Loxodonta*. Primates: *Homo*, *Pan*, *Gorilla*, and *Pongo*. Sloths: *Bradypus*. Marsupials: *Philander* and *Petrogale*. Songbirds: *Pomatostomus ruficeps* (B1), *Pomatostomus superciliosus* (B2), *Pomatostomus temporalis* (B3), *Pomatostomus isidori* (B4), *Corcorax melanorhamphos* (B5), *Rhipidura*, *Coracina*, *Pachycephala*, *Gymnorhina*, *Microeca*, *Epimachus*, *Cicinnurus*, *Parotia*, *Vestiaria*, *Himatione*, *Hemignathus*, and *Carpodacus*. Gamebirds: *Gallus*, *Alectoris*, *Lophortyx*, *Numida*, *Coturnix*, and *Ortalis*. Waterfowl: *Anser*. Salamanders: *Ambystoma*. Frogs: *Xenopus*. Crocodiles: *Alligator*. Sharks: *Galeocerdo*, *Prionace*, *Sphyrna*, *Heterodontus*, and *Carcharhinus*. Cichlids: *Cichlasoma citrinellum* (F1), *Cichlasoma labiatum* (F2), *Cichlasoma centrarchus* (F3), *Cichlasoma nicaraguense* (F4), and *Julidochromis regani* (F5), *Neetroplus*, *Geophagus*, *Pterophyllum*, *Hemichromis*, *Aequidens*, *Macropleurodus*, *Platytaeniodus*, *Haplochromis*, *Pelvatochromis*, *Crenicichla*, *Astatoreochromis*, and *Oreochromis*. Salmonids: *Oncorhynchus* and *Salmo*. Coryphenids: *Coryphena*. Cicadas: *Banza*, *Magicicada*, and *Okanagana*. Tarantula: *Rhaetostica*.

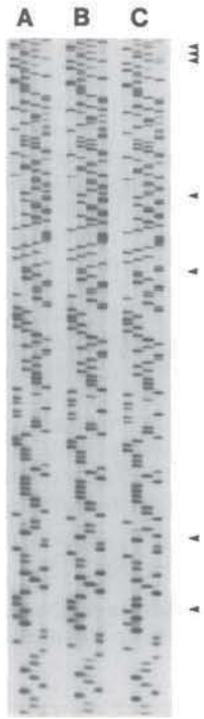

FIG. 1. Sequencing gel for part of the 12S rRNA gene from three kangaroo rats. Lanes: A and B, *Dipodomys agilis*; C, *Dipodomys microps*. Arrowheads mark variable sites.

cytochrome *b* gene not only from humans but also from most other vertebrates tested (Table 1). Likewise, the second pair of primers amplifies a 386-base-pair segment of the small rRNA (12S rRNA) from animals as different as humans, fishes, and insects (Table 1). The third set amplifies the control region of mtDNA (about 1 kilobase) in most mammals and many fishes (Table 1).

It Was Not Necessary to Purify mtDNA Prior to Amplification. Preparations of total cellular DNA are sufficient for amplification using the polymerase chain reaction. Moreover, the amounts of tissue needed to produce a sequence were small, typically a few nanograms of fresh tissue and less than a milligram for old specimens. This is consistent with observations made on single hairs (16) and single sperm (23). Many of the tissues used were from museum specimens stored at room temperature. Dry skins collected in 1911 were the source of some of the rodent results, and fish tissues stored in 70% ethanol for months were also reliable sources of amplifiable mtDNA (Table 1). In each case the amplification products were pure enough to sequence directly, as shown in Fig. 1.

Patterns of mtDNA Sequence Evolution. Fig. 2 presents a subset of the sequences we have obtained to illustrate the utility of the approach. The DNA sequences coding for 80 amino acids of cytochrome *b* are aligned for five rodents, five birds, five fishes, and a human. These sequences are invariant at 111 of the 239 base positions examined. The differences found at the remaining 128 positions are all due to base substitutions of the types expected for a protein-coding gene in an animal mitochondrion (3, 9).

Among very close relatives such as species within a genus (e.g., rodents 1–4, birds 1–4, and fishes 1–4 in Fig. 2) most of the changes are transitions at synonymous sites. By contrast, among more distant relatives such as genera within a family or order (e.g., rodent 5 vs. other rodents, bird 5 vs. other birds, and fish 5 vs. other fishes), transversions are more evident (Fig. 3). Still more distant comparisons, as between orders or classes of vertebrates, reveal that the extent of difference due to transversions reaches a plateau within orders. For example, the number of transversion differences observed between rodent 5 and other rodents (mean = 27 transversions) is nearly the same as between rodent 5 and fish 1 (mean = 38 transversions). Nevertheless, such a plateau or saturation effect is not so evident for amino acid replacements. The number of replacements rises from an

```
                  50                                      60                                      70                                      80
       T   G   L   F   L   A   M   H   Y   S   P   D   A   S   T   A   F   S   S   I   A   H   I   T   R   D   V   N   Y   G   W   I   I   R   Y   L   H   A   N   G
HUMAN ACA GGA CTA TTC CTA GCC ATG CAC TAC TCA CCA GAC GCC TCA ACC GCC TTT TCA TCA ATC GCC CAC ATC ACT CGA GAC GTA AAT TAT GGC TGA ATC ATC CGC TAC CTT CAC GCC AAT GGC
  R1  T.. ..C ... ..T ... ..T ..A ..T ..T A.. ... ... A.. CTC ..G ..A ... ... ... G.T A.. ... ... TG. ... ..T ... ... ... ..A ... C.. ... ... ... A.. ..T ... ..C ..A
  R2  T.. ..C ... ..T ... ..T ..A ..T ..T A.. ... ... A.. CTC ..G ..A ... ... ... G.T A.. ... ... TG. ... ..T ... ... ... ..A ... C.. ... ... ... A.. ..T ... ..C ..A
  R3  T.. ..C ... ... ..G ..T ..A ..T ... A.. ... ... A.. ATC ..G ..A ... ... ... G.T A.. ... ... TGC ... ..T ... ... ... ..A ... C.T ... ..T ... A.. ..T ... ..C ..A
  R4  T.. ..C ... ..T ... ... ..A ... ... A.. ... ..T A.. CTC ..A ..A ..C ... ... G.T ... ..T ... TGC ... ... ... ..C ..C ..T ... C.T ..T ... ... A.A ... ..T ... ..A
  R5  T.T ..C ... ... ... ... ..A ... ..T A.. T.. ..T AG. ..T .A. AG. ... ... ... G.A A.. ..T ... TGC ... ..T ... .G. ... ..T ... C.A ... ..A ..T A.A ..T ..A ... ..A

  B1  ... ..C ... C.. T.. ... GCA ..T ..T A.. G.T ..T A.T ..C CTA ... ..C G.T ..C G.A ... ... ..A TGC ..C A.. ..G C.A .TC ..A ... C.A ... ..A A.. ..G ..T ..T ..C ..A
  B2  ... ..C ... C.. ... ... GCA ... ... A.. G.T ..T A.T ..C CTA ... ..C G.T ..C G.A A.. ... ..A TGC ..C A.T ... C.A .TC ..A ... C.A ... ..A A.. ..A ..T ... ..C ..A
  B3  ... ..C ..G C.. ... ..T GCA ... ... A.. G.T ... A.. ..C CTA ... ... G.C ..C G.A ... ... ..A TGC ..C A.. ... C.A .TC ..A ... C.A ... ..A A.. ..A ..T ..T ..C ..A
  B4  ... ..C ... ... ... ... GCA ..T ... A.. G.C ... A.. ..C CT. ... ..C AAC ..C G.. ... ... .CA TGC ..C A.. ... C.A .TC ..G ... C.A ... ... A.. ..A ..T ... ..C ..A
  B5  ... ..C ... C.A ... ... ... ... ... A.. G.C ... A.. ..C CTA ... ..C A.C ... G.A ... ..T .CA TGC ... A.. ..C C.A .TC ... ... T.A ... ..A A.. ..A ... ..A ..C ..A

  F1  ... ..C ..T ... ..T ..A ..A ... ... A.T T.C ..T AT. G.. ..A ... ... ... ..C G.. ... ... ... TGC ... ..T ... ..C ... ... ... C.A ... ..T A.T A.A ... ... ..C ...
  F2  ... ..C ..T ... ..T ..A ..A ... ... A.T T.C ..T AT. G.. ..A ... ... ... ..C G.. ... ... ... TGC ... ..T ... ..C ... ... ... C.A ... ..T A.T A.A ... ... ..C ...
  F3  ... ..C ..T ... ..T ..A ..A ... ... A.T T.C ..T AT. G.. ..A ... ... ... ..C G.T ... ... ... TGC ... ..T ... ..C ... ... ... C.A ... ..T A.T A.A ..T ... ..C ...
  F4  ... ..C A.. ... ..T ..A ..A ..T ... A.T T.T ... AT. G.. ..A ... ... ... ..C G.T ... ... ... TG. ... ..T ... ... ... ... ... C.. ... ... A.T A.A ... ... ... ...
  F5  ... ..C ..T ... ... ... ..A ... ..T A.C T.C ... AT. G.C ... ... ... ..C ..C G.. ... ... ... TG. ..T ... ..C ..C ..C ..T ... C.. ... ..A A.T A.G ... ... ..C ...

                  90                                      100                                     110                                     120
       A   S   M   F   F   I   C   L   F   L   H   I   G   R   G   L   Y   Y   G   S   F   L   Y   S   E   T   W   N   I   G   I   I   L   L   L   A   T   M   A   T
HUMAN GCC TCA ATA TTC TTT ATC TGC CTC TTC CTA CAC ATC GGG CGA GGC CTA TAT TAC GGA TCA TTT CTC TAC TCA GAA ACC TGA AAC ATC GGC ATT ATC CTC CTG CTT GCA ACT ATA GCA AC
  R1  ..A ... C.T ... ... ... ... ..T .A. ..T ... ..T ..C ... ..A A.C ... ... ..C ... .A. TCT ..T AT. ... ... ... ... ..T ... ... ..T ... ..A T.. CTC ..G ... ..T ..
  R2  ..A ... C.T ... ..C ... ..T ..T .A. ..T ... ... ..C ... ..A A.C ... ... ..C ... .A. TCT ..T AT. ... ... ... ... ..T ... ... ... ... ..A T.. CTC ..G ... ..T ..
  R3  ..A ... C.T ... ..C ... ..T ..T .A. ..T ..T ..T ..C ... ..A A.C ... ... ..C ... .AC TCT ..T AT. ... ... ... ... ..T ... ... ..T ... ..T A.. CTC ..G ... ..T ..
  R4  T.. ... C.T ... ..C ... ... ..T .A. ..T ... ..T ..C ... ..A A.C ..C ..T ..C ... .AC TCT ... A.. ... ..T ... ... ..T ..T ... ..T ... ..T T.. CTT ... ... ... ..
  R5  ..T ... C.. ..T ... ... ... ..A .A. A.C ..T ..T ..A ..T ... A.T ..C ... ..C ..C .AC ..T ..T AAG ... ... ... ..T G.T ... ..C T.G ... ..A T.C TT. T.. ... ..C ..

  B1  ..T ... T.T ... ..C ... ... A.. .A. ... ... ..T ..A ... ..A ..C ..C ... ..C ..T .AC ..A A.. AA. ..G ... ... ... ..T ..A G.C ... ..A ..C ..A A.C CTA ... ... ..
  B2  ... ... T.C ... ... ..T ... A.. .A. ... ... ..T ..A ... ..A ..C ..C ... ..C ..C .AC ..A A.. AA. ..G ... ... ... ... ..A G.C ... ..A ..C ..A A.T CTA ... ... ..
  B3  ..T ... T.C ..T ..C ..T ... A.. .A. ..G ..T ..T ..A ... ..A ..C ..C ... ..C ..C .AC T.A A.. AA. ... ... ... ... ..T ..A G.C ... ..G ..C ..A A.C CTA ... ... ..
  B4  ... ... C.C ... ..C ... ... A.. .AT ... ... ... ..A ... ..A ..C ..C ... ..C ..C .AC ..T A.. AA. ... ... ... ... ..T ..A G.. ... ..A ..C ..G A.C CTA ... ... ..
  B5  ... ..C ... ... ..C ... ... A.. .A. ... ... ... ..C ... ... T.C ..C ... ..C ... .AC ..A A.. AA. ..G ... ... ... ... ..A G.C ... ... ..A ..A A.C CT. ... ... ..

  F1  ..A ..C T.T ... ... ..T ... A.. .A. ..T ... ..T ..C ... ..A ... ..C ... ..C ..C .AC ..T ..T AA. ... ..A ... ..T G.A ..A G.. G.T ... ..C ..C CT. ..C ... AT. ..
  F2  ..A ..C T.T ... ... ..T ... A.. .A. ..T ... ..T ..C ... ..A ... ..C ... ..C ..C .AC ..T ..T AA. ... ..A ... ..T G.A ..A G.. G.T ... ..C ..C CT. ..C ... AT. ..
  F3  ..A ..C T.T ... ... ..T ... A.T .A. ..T ... ..T ..C ... ..A ... ..C ... ..C ..C .AC ..T ..T AA. ... ..A ... ..T G.A ..A G.. G.T ..A ..C ..C CT. ..C ... AT. ..
  F4  ..A ..C T.C ... ..C ... ... A.. .AT ..T ... ..T ..C ... ..A T.. ..C ..T ..C ..C .AC ..T ... AA. ... ..A ... ... G.. ..A G.. ... ... ..C ..C CT. ..C ... AT. ..
  F5  ..A ..C T.C C.. ..C ..T ..T A.T .A. ..C ... ..T ..C ... ..G T.. ..C ..T ..C ..C .A. T.A ..T AA. ... ... ... ..T ..T ..A G.. ... ... ..C ... TT. ... ... AT. ..
```

FIG. 2. DNA sequences of part of the cytochrome *b* gene from five rodents (R1–R5), five birds (B1–B5), and five fishes (F1–F5) aligned with the homologous region in human mtDNA (18). The species code follows the legend of Table 1. Dots indicate sequence identity with human mtDNA. Codons are numbered as in the human sequence.

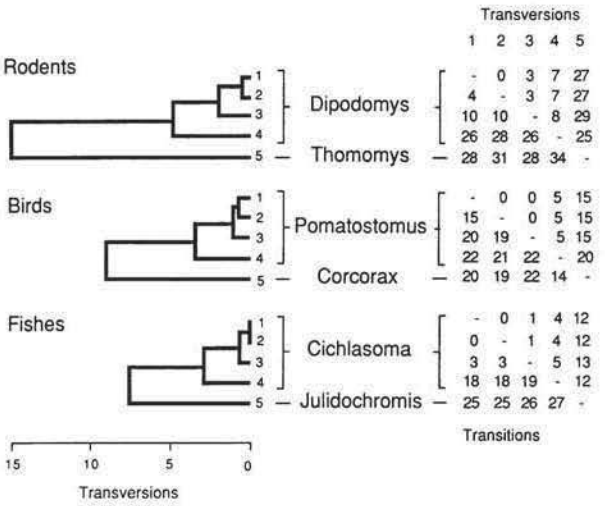

FIG. 3. Tree analysis of cytochrome *b* sequences in Fig. 2. Species labels correspond to Table 1. Numbers of transversion differences among pairs of species appear above the diagonal and numbers of transitions are below. The most parsimonious trees deduced by a character-state analysis of the data are shown. The branch lengths of the trees are drawn proportional to the number of transversions on each lineage, with each transition being considered equivalent to 0.1 transversion (24). These trees are consistent with previous concepts of phylogeny for these groups (25–27).

average of five differences between bird 5 and other birds to an average of 23 differences between birds and fishes. These findings are consistent with the view that transitions occur more often than transversions and that amino acid replacements occur less often than transversions at synonymous positions in cytochrome *b* codons (cf. refs. 3 and 9).

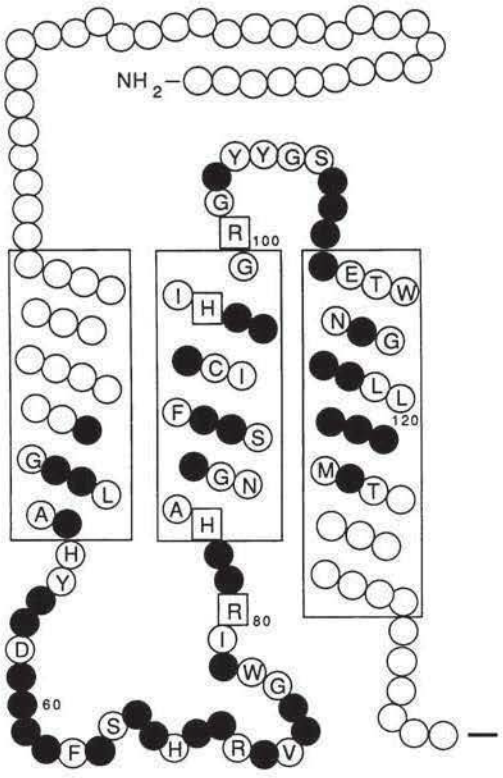

FIG. 4. Model of part of cytochrome *b* showing its position with respect to the inner membrane of the mitochondrion (28). The rectangles enclose α-helical segments spanning the membrane. Solid circles show amino acid residues that vary among the vertebrate sequences in Fig. 3. Circles containing letters were invariant in our study. Squares indicate four invariant residues considered necessary for cytochrome *b* function (28). Open circles refer to residues in regions that were not sequenced for this study.

Pattern of Replacement Fits with the Structural Hypothesis for Cytochrome *b*. Fig. 4 presents a structural hypothesis for cytochrome *b* in the mitochondrial membrane (28). Amino acid replacements are mainly conservative and at positions known to differ between yeast and mammals. No replacements were evident at positions thought to be essential for function (namely, positions 80, 83, 97, and 100). Besides reinforcing the structural model, this result confirms the idea that direct sequencing of enzymatically amplified DNA is at least as reliable as conventional cloning and sequencing (29). If the polymerase chain reaction were generating errors, they would probably be distributed at random with respect to position within codons and to codons within the cytochrome *b* gene.

DISCUSSION

The fact that primers of broad utility could be found for the fast-evolving mtDNA of animals simply by comparing frog and mammal mtDNAs makes it likely that "universal" primers can also be designed for parts of the nuclear, bacterial, chloroplastid, and plant mitochondrial genomes. The only requirements for the construction of such primers are reference sequences from two or three widely diverged creatures. The necessity for cloning is thus bypassed and sequences can then be obtained directly by the polymerase chain reaction for most other members of the clade that contains the reference organisms and frequently for allied clades.

Phylogenetic Trees. These short sequences from a piece of the cytochrome *b* gene contain phylogenetic information extending from the intraspecific level to the intergeneric level. This is evident from comparing trees based on these sequences (Fig. 3) with other evidence concerning the phylogenetic relationships among these animals. In all three cases examined the cytochrome *b* sequences within a genus are more related to one another than to those in other genera. Furthermore, the relationships within genera agree with expectations based on other evidence (see references in Fig. 3). There are also a few base positions at which all members tested within a major group are uniquely alike (e.g., in codons 57, 61, 72, 74, 75, 109, 122, and 123 of birds; Fig. 2). Hence this short sequence is a versatile source of phylogenetic information. By contrast, the restriction endonuclease approach applied to whole mtDNA has a limited phylogenetic range, being useful mainly at or below the genus level (3, 4, 30).

Unusual Base Composition of Bird mtDNA. In Table 2 we present the base composition at the silent positions of cytochrome *b* codons. Rodents, birds, and fishes all exhibit the low incidence of guanines, which has been reported for vertebrate mtDNA (18-20). Striking is the deficiency of thymines at silent positions in the birds studied. A corresponding deficiency exists in the frequency of thymine to cytosine changes during bird evolution. An expected consequence of this compositional bias is that the sequence dif-

Table 2. Average base composition at silent positions in the cytochrome *b* sequences of Fig. 2

Base	Average base composition, % of total		
	Rodents	Birds	Fishes
Guanine	2.3	2.7	0.5
Adenine	28.7	32.0	26.7
Thymine	31.9	11.3	24.3
Cytosine	37.1	54.0	48.5
	$n = 89$	$n = 93$	$n = 90$

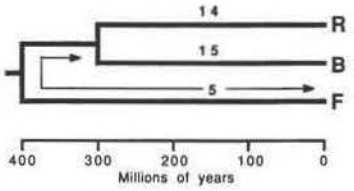

FIG. 5. Numbers of amino acid replacements in a segment of cytochrome *b* in three vertebrate lineages. The rate of replacement is about 5 times lower in the lineage leading to fish (F) than on those leading to birds (B) or rodents (R).

ference measured at such sites can never become very large (cf. ref. 9). This finding points to the need for a more comprehensive study of substitution matrices (31) and compositional bias in avian mtDNA before these sequences are used for studying deep branches in the avian tree.

Slow Amino Acid Change in Fishes. Another notable outcome of this comparative study concerns the tempo of amino acid replacement in mitochondrially encoded proteins. Again, the parsimony principle was used to apportion changes on a tree of known topology relating rodents to birds and fishes. This topology and the times of splitting of the lineages leading to these three groups appear in Fig. 5. During nearly 400 million years of evolution only about 4 changes have occurred on the fish lineages, which contrasts with the 14 to 15 changes on the bird and mammal lineages during the last 300 million years. The results in Fig. 5 suggest a 5-fold higher rate of amino acid substitution on the bird and mammal lineages. This finding fits with other evidence implying low rates of amino acid substitution in the mitochondrial genes of other fishes and of invertebrates (9, 32). Because our results come from only a small segment of one gene, it is clearly desirable to conduct a more comprehensive survey of protein-coding genes. If the unusually high rates of amino acid substitution in birds and mammals are confirmed it will become possible to reconcile the conflicting claims concerning rates of mtDNA evolution in major groups of animals (e.g., refs. 8–10).

Prospects. It is possible to imagine numerous applications of this method. For instance, it will now be possible to follow gene frequency changes through time using both old museum specimens and modern representatives of a population (W.K.T., S.P., F.X.V., and A.C.W., unpublished data). Second, it will be possible to begin to organize knowledge of genetic diversity in natural populations of minute organisms that are not easily grown in the laboratory. A single-cell planktonic organism contains enough mtDNA molecules for successful amplification and sequencing. The ability to compare individuals in this way could have a profound effect on ecological genetics, especially in the marine biosphere. Finally, the ease with which homologous sequences can be gathered will facilitate a synergism between molecular and evolutionary biology, which will lead to insights into genetic structure and function (33) based on the dynamics of molecular change and phylogenetic history.

We thank P. Arctander, P. Basasibwaki, R. Cann, S. Carr, D. Irwin, B. Kessing, J. Kornegay, A. Martin, D. Mindell, S. Pruett-Jones, T. Quinn, G. Shields, A. Sidow, C. Simon, and L. Williams for unpublished results and discussion; B. Malcolm for primer synthesis; J. Patton for access to many museum specimens; S. Ferris for purified primate mtDNAs; and E. Prager, V. Sarich, T. White, and several reviewers for helpful comments on the manuscript. This work received support from the National Science Foundation and the National Institutes of Health.

1. Brown, W. M. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 95–130.
2. Palmer, J. D. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 131–240.
3. Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. & Stoneking, M. (1985) *Biol. J. Linn. Soc.* **26,** 375–400.
4. Avise, J. C. (1986) *Philos. Trans. R. Soc. London Ser. B* **312,** 325–342.
5. Helm-Bychowski, K. M. & Wilson, A. C. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 688–692.
6. Wilson, A. C., Zimmer, E. A., Prager, E. M. & Kocher, T. D. (1989) in *The Hierarchy of Life*, eds. Fernholm, B., Bremer, K. & Jörnvall, H. (Elsevier Science, Amsterdam), pp. 407–419.
7. Aquadro, C. F., Desse, S. F., Bland, M. M., Langley, C. H. & Laurie-Ahlberg, C. C. (1986) *Genetics* **114,** 1165–1190.
8. Vawter, L. & Brown, W. M. (1986) *Science* **234,** 194–196.
9. DeSalle, R., Freedman, T., Prager, E. M. & Wilson, A. C. (1987) *J. Mol. Evol.* **26,** 157–164.
10. Palumbi, S. R. & Wilson, A. C. (1989) *Evolution*, in press.
11. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
12. White, T. J., Arnheim, N. & Erlich, H. A. (1989) *Trends Genet.* **5,** 185–189.
13. Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H. A., Arnheim, N. & Wilson, A. C. (1987) *Nucleic Acids Res.* **15,** 529–542.
14. Gyllensten, U. B. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7652–7656.
15. Vigilant, L., Stoneking, M. & Wilson, A. C. (1988) *Nucleic Acids Res.* **16,** 5945–5955.
16. Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. F. & Erlich, H. A. (1988) *Nature (London)* **332,** 543–546.
17. Pääbo, S., Gifford, J. A. & Wilson, A. C. (1988) *Nucleic Acids Res.* **16,** 9775–9787.
18. 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. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* **290,** 457–465.
19. Anderson, S., de Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F. & Young, I. G. (1982) *J. Mol. Biol.* **156,** 683–717.
20. Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. (1981) *Cell* **26,** 167–180.
21. Roe, B. A., Ma, D.-P., Wilson, R. K. & Wong, J. F.-H. (1985) *J. Biol. Chem.* **260,** 9759–9774.
22. Clary, D. O. & Wolstenholme, D. R. (1985) *J. Mol. Evol.* **22,** 252–271.
23. Li, H., Gyllensten, U. B., Cui, X., Saiki, R. K., Erlich, H. A. & Arnheim, N. (1988) *Nature (London)* **335,** 414–417.
24. Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A. & Wilson, A. C. (1984) *Nature (London)* **312,** 282–284.
25. Patton, J. L., MacArthur, H. & Yang, S. Y. (1976) *J. Mammal.* **57,** 159–163.
26. Ford, J. (1974) *Emu* **74,** 161–168.
27. Regan, C. T. (1905) *Ann. Mag. Nat. Hist. Ser. Seven* **16,** 60–445.
28. Howell, N. & Gilbert, K. (1988) *J. Mol. Biol.* **203,** 607–618.
29. Pääbo, S. & Wilson, A. C. (1988) *Nature (London)* **334,** 387–388.
30. Harrison, R. G. (1989) *Trends Ecol. Evol.* **4,** 6–11.
31. Aquadro, C. F. & Greenberg, B. D. (1983) *Genetics* **103,** 287–312.
32. Thomas, W. K. & Beckenbach, A. T. (1989) *J. Mol. Evol.*, in press.
33. Wilson, A. C. (1988) *Immunogenetics* **28,** 68–69.