J Mol Evol (1993) 36:327–334

Journal of Molecular Evolution

Structure and Evolution of Opossum, Guinea Pig, and Porcupine Cytochrome *b* Genes

Din-Pow Ma,[1,2] Andrey Zharkikh,[2] Dan Graur,[3] John L. VandeBerg,[4] and Wen-Hsiung Li[2]

[1]Department of Biochemistry and Molecular Biology, Dorman Hall, Mississippi State University, Mississippi State, MS 39762, USA, [2]Center for Demographic and Population Genetics, University of Texas, P.O. Box 20334, Houston, TX 77225, USA, [3]Department of Zoology, George S. Wise Faculty of Sciences, Tel Aviv University, Ramat Aviv 69978, Israel, and [4]Department of Genetics, Southwest Foundation for Biomedical Research, P.O. Box 28147, San Antonio, TX 78284, USA

Summary. We have sequenced the mitochondrial cytochrome *b* gene from the guinea pig, the African porcupine, and a South American opossum. A phylogenetic analysis, which includes 22 eutherian and four other vertebrate cytochrome *b* sequences, indicates that the guinea pig and the porcupine constitute a natural clade (Hystricomorpha) that is not a sister group to the clade of mice and rats (Myomorpha). Therefore, the hypothesis that the Rodentia is paraphyletic receives additional support. The artiodactyls, the perissodactyls, and the cetaceans form a group that is separated from the primates and the rodents. The 26 sequences are used to study the structure/function relationships in cytochrome *b*, whose function is electron transport. Most of the amino acid residues involved in the two reaction centers are well conserved in evolution. The four histidines that are believed to ligate the two hemes are invariant among the 26 sequences, but their nearby residues are not well conserved in evolution. The eight transmembrane domains represent some of the most divergent regions in the cytochrome *b* sequence. The rate of nonsynonymous substitution is considerably faster in the human and elephant lineages than in other eutherian lineages; the faster rate might be due to coevolution between cytochrome *b* and cytochrome *c*.

Key words: Cytochrome *b* — Mitochondrial DNA — Mammalian phylogeny — Functional constraints — Coevolution — *Cavia porcellus* — *Monodelphis domestica* — *Hystrix africaeaustralis*

Offprint requests to: W.-H. Li

The Rodentia is the most speciose mammalian order, consisting of more than half of all the extant eutherian species. (See Luckett and Hartenberger 1985.) In the past, this order used to be divided into three suborders: Sciuromorpha (squirrel-like rodents), Myomorpha (rat-like rodents), and Hystricomorpha (porcupine-like rodents). (See Nowak and Paradiso 1983.) More recently, however, the first two suborders have been clustered together into a single one called Sciurognathi. (See Luckett and Hartenberger 1985.)

Although the monophyly of rodents has long been considered a firmly established fact, recent molecular data suggest that the guinea pig (*Cavia porcellus*) and the other New World hystricomorphs may not be a sister group of the myomorphs (Graur et al. 1991, 1992; Li et al. 1992a). Rather, the myomorphs may be more closely related to the primates and the artiodactyls than to the hystricomorphs. This phylogenetic hypothesis raises some interesting issues, such as the possibility that extensive morphological parallelism or convergence has occurred during the evolution of eutherians (Allard et al. 1991) or the possibility that diagnostic morphological features used in the classification of rodents are primitive rather than derived (Li et al. 1992a). Since Graur et al.'s (1991) suggestion of the paraphyly of the order Rodentia is extremely con-

troversial (Allard et al. 1991; Hasegawa et al. 1992; Li et al. 1992b; Novacek 1992), we have made efforts to increase both the data base and the taxonomic range on which to base our phylogenetic conclusions. For these purposes we sequenced the mitochondrial cytochrome *b* (COB) genes from the guinea pig and a South American opossum (*Monodelphis domestica*); the latter will be used as an outgroup to the eutherians. We chose to sequence the COB gene because it can be readily amplified by the polymerase chain reaction (PCR) and because this gene has been sequenced in many vertebrate species. (See Irwin et al. 1991 and references therein.)

Another question related to the issue of rodent phylogeny concerns the relationships between the New World families of the Hystricomorpha (the Caviomorpha or guinea-pig-like rodents) and the Old World hystricomorphs. These two groups are sometimes treated as separate taxa within the Rodentia. (See Romer and Parsons 1977.) If indeed the guinea pig and its caviomorph relatives do not belong to Rodentia but represent an independent order of eutherians, we must find out whether or not this order includes the Old World hystricomorphs as well. For this purpose, we also sequenced the COB gene from the African porcupine (*Hystrix africaeaustralis*).

The present three sequences, together with 23 published vertebrate COB sequences, provide a good data set for identifying highly conserved regions and for inferring the structure/function relationships in this protein.

Materials and Methods

DNA and Sequence Data Sources. Total genomic DNA was extracted from liver tissues of a South American opossum (*Monodelphis domestica*) and an African porcupine (*Hystrix africaeaustralis*), and from a blood sample of the guinea pig (*Cavia porcellus*). The porcupine tissue was a gift from Dr. Rodney Honeycutt. The DNA sequences for human, mouse, rat, chicken, *Xenopus*, and sturgeon cytochrome *b* were taken from GenBank and the other sequences were from Irwin et al. (1991).

DNA Amplification and Sequencing. Mitochondrial sequences containing the cytochrome *b* gene were amplified by PCR. The primers used for amplification were the flanking tRNA sequences L14724 (5'-CGAAGCTTGATATGAAAAACCATCGTTG-3') and H15915 (5'-AACTGCAGTCATCTCCGGTTTACAAGAC-3') (Irwin et al. 1991). The PCR products were purified by preparative electrophoresis on a 1% agarose gel and eluted by a modified freeze-thaw method (Thurning et al. 1975). The purified COB DNA fragments were then digested with *Hin*dIII and *Pst*I and cloned into M13mp18 and M13mp19 cleaved with *Hin*dIII and *Pst*I or *Pst*I alone. The COB fragments cloned in both orientations were sequenced by the dideoxy chain termination method (Sanger et al. 1977). The single-stranded recombinant phage DNAs were isolated by phenol extraction of the PEG-precipitated phage and sequenced using a universal primer (Anderson et al. 1980). In some cases, a sequential series of overlapping clones was also produced using recombinant single-stranded M13 DNA and complementary 22- or 29-mers as described (Dale et al. 1985). To avoid sequencing errors introduced by PCR amplification, multiple M13 clones were sequenced.

Data Analysis. The number of nucleotide substitutions per nonsynonymous site (K_A) between every two sequences was computed by the method of Li et al. (1985). The number of nucleotide substitutions per synonymous site (K_S) was very large and could not be reliably estimated between most sequence pairs. The K_A values were used to reconstruct a phylogenetic tree by using the neighbor-joining method (Saitou and Nei 1987). Bootstrap estimates of the confidence level of subsets of taxa (Felsenstein 1985) were obtained by using a program modified from that of T.S. Whittam.

Results and Discussion

Cytochrome b Sequences

The entire cytochrome *b* genes from the guinea pig, the African porcupine, and the opossum were amplified by PCR and completely sequenced. The three sequences are shown in Fig. 1; the initiation codon is not presented. The guinea pig and the African porcupine sequences are 381 codons long while the opossum sequence has three additional codons at the 3' end.

Phylogentic Analysis

To the 20 sequences previously used by Irwin et al. (1991), we add two eutherian sequences (the guinea pig and the African porcupine), a metatherian sequence (the opossum), and also the homologous sequences from chicken, *Xenopus*, and sturgeon. The last four sequences are used as outgroups. In comparison, no outgroup was used for phylogenetic reconstruction in Irwin et al.'s study. A phylogenetic tree for the 26 sequences is inferred by applying the neighbor-joining method to the K_A values between sequences; the distance matrix is large and is not presented here.

There are similarities between Irwin et al.'s tree and ours. The branching order for pronghorn, fallow deer, giraffe, black-tailed deer, goat, sheep, cow, and chevrotain is the same, and so is the branching order for the three dolphin sequences. In addition, both trees suggest the following clades: mouse and rat, zebra and rhinoceros, pig and peccary, and human and elephant. Finally, our tree agrees with Irwin et al.'s in that both trees suggest a superordinal clade consisting of artiodactyls, perissodactyls and cetaceans to the exclusion of primates, rodents, and proboscids. However, there

```
Guinea Pig  ACC CAC CTA CGA AAA TCA CAC CCA CTC ATC AAA ATC ATT AAC CAC TCC CTA ATT GAC CTC CCA GCT CCA TCC AGC ATT TCA
Porcupine   ACA AAC ATC CGA AAA TCC CAT CCT CTT CTC AAA ATT ATT AAC CAC TCA TTC ATT GAC CTT CCC ACC CCA TCT AAC ATC TCA
Opossum     ACC AAC CTA CGA AAA AAC TAC CCC TTA ATA AAA ATT ATT AAC CAC TCA TTC ATT GAC CTA CCA GCC CCT TCC AAT ATC TCT

ACG TGA TGA AAC TTC GGC TCC CTC TTA GGC ATC TGT CTA GGC CTG CAA ATT ATT ACA GGA CTC TTC CTA GCA ATA CAC TAT ACT GCA GAC
ACA TGA TGA AAC TTC GGC TCA CTT TTA GGA GCC TGC TTA ATT ATC CAA ATC CTT ACA GGT CTA TTC CTA GCA ATA CAT TAC ACT GCC TAC
GCT TGA TGG AAC TTT GGA TCA CTT TTA GGC ATA TGT TTA ATT ATC CAA ATC CTA ACA GGA CTA TTT CTA GCC ATA CAT TAT ACA TCA GAC

ACT TCC ACG GCA TTC TCG TCT GTC GCC CAC ATT TGC CGA GAC GTA AAC TAT GCC TGA CTG ATC CGA TAT CTA CAT GCC AAC GGA GCA TCC
ACA ACT ACA GCA TTT TCA TCA GTA GCC CAT ATT TGC CGA GAC GTC AAT TAC GGA TGA CTA ATT CGT TAC CTC CAC GCT AAC GGA GCT TCA
ACT CTA ACC GCT TTT TCA TCA GTA GCA CAT ATT TGC CGA GAT GTT AAC TAC GGA TGA CTG ATC CGA AAT CTA CAT GCT AAC GGA GCT TCA

ATA TTC TTT ATT TTC CTA TAT CTA CAC ATC GGA CGA GGT ATT TAC TAC GGA TCA TAC ACA TTC CTA GAG ACA TGA AAT ATT GGA ATT GCT
ATA TTC TTT ATC TGT CTA TAC CTC CAC GTA GGC CGA GGG TTA TAC TAT GGA TCC TAC ATA TTT ACA GAA ACT TGA AAT ATC GGA ATT CTC
ATA TTC TTT ATA TGT TTA TTC CTT CAC GTA GGA CGA GGA ATT TAC TAT GGT TCC TAC CTG TAT AAA GAA ACC TGA AAC ATC GGA GTG ATT

GTT CTG TTC ACA GTT ATG GCT ACC GCA TTC ATA GGG TAC GTA TTG CCA TGG GGT CAA ATA TCC TTT TGA GGT GCT ACC GTT ATT ACT AAT
TTA CTG TTT ACA GTA ATG GCT ACT GCC TTC ATA GGA TAC GTC CTC CCA TGA GGA CAA ATA TCT TTT TGA GGG GCT ACT GTC ATT ACC AAT
CTC ATA CTT ACT GTA ATA GCC ACT GCC TTC GTA GGC TAT GTA CTC CCA TGA GGA CAA ATA TCC TTT TGA GGG GCT ACA GTT ATT ACT AAC

GTT CTA TCA GCT ATC CCC TAC ATC GGG ACA ACC CTT GTA GAG TGA ATC TGA GGG GGG TTC TCA GTA GAC AAA GCC ACC CTA ACA CGA TTC
TTA TTA TCA GCA ATC CCC TAT ATT GGC ACA ACA CTA GTT GAG TGA ATC TGA GGG GGG TTT TCA GTA GAT AAA GCA ACT TTA ACA CGA TTC
GTT TTA TCA GCC ATT CCA TAC ATT GGT AAT ACT CTA GTA GAA TGA ATC TGA GGC GGA TTT TCA GTT GAC AAA GCT ACA TTA ACT CGA TTC

TTT GCC TTC CAC TTT ATT GTT CCA TTC ATC ATC ACC GCC CTA GTG ATA GTC CAC CTC TTA TTC CTC CAC GAG ACA GGA TCA AAC AAC CCA
TTT GCT TTC CAC TTC AGC CTG CCA TTC ATC ATC ACA GCC CTA GTA CTA GTT CAT CTA CTA TTT TTA CAC GAA ACA GGG TCA AAC AAC CCA
TTC GCA TTC CAT TTT ATT TTA CCT TTC ATT ATC CTT GCA TTA GTT ATT GTT CAC CTC CTA TTC CTA CAT GAG ACC GGA TCA AAT AAT CCT

TCA GGA CTA AAC TCA GAC TCC GAC AAA ATC CCA TTC CAC CCT TAT TAC ACA ATC AAA GAC ATT TTA GGA GCC TTA TTT ATA ATA CTA GCT
TCA GGC ATT GAC TCA AAC TCA GAC AAA ATT CCA TTC CAC CCC TAT TAC ACA ATT AAA GAT ATT CTA GGC CTC CTA TTA ATA CTA ACA GCC
ACA GGA ATT AAC CCC AAC TCA GAC AAA ATT CCA TTT CAC CCC TAC TAC ACT ATC AAA GAC GCC CTA GGC CTA ATC CTT ATA CTT CTT ATT

CTT CTA TGC CTA GTA CTC TTT ACA CCC GAC CTA TTA GGA GAC CCA GAT AAC TAC ACA CCC GCC AAC CGG CTG AAT ACG CCA CCA CAC ATT
CTA CTA ATC CTA GTA CTA TTT TCC CCA GAC CTT TTA GGA GAC CCG GAT AAC TAT ACT CCA GCA AAC CCC TTA AAT ACT CCT CCC CAT ATT
TTA ATA TCA CTA GCA ATA TTC TCA CCC GAT ATA CTA GGT AAC CCA GAC AAC TTT ACA CCA GCC AAC CCA TTA AAC ACT CCT CCA CAT ATT

AAA CCA GAG TGG TAT TTC TTA TTT GCC TAC GCA ATC CTC CGC TCT ATC CCT AAC AAA CTA GGA GGG GTC CTA GCC CTA GTT CTC TCT ATT
AAA CCA GAA TGA TAT TTC CTA TTC GCT TAC GCT ATC CTA CGC TCA ATC CCT AAT AAA CTC GGA GGA GTA TTA GCC CTT ATC TTC TCT ATC
AAA CCA GAA TGA TAT TTT CTA TTT GCC TAT GCA ATC TTA CGA TCA ATT CCA AAC AAA TTA GGA GGA GTA TTA GCT CTC TTA GCA TCT CTC

CTA ATC CTA GCC CTA TTC CCC ATA CTC CAC ACA TCA AAA CAA CGT AGC ATA CGA TTC CGC CCC CTC AGC CAA TGC CTT CTA TGA TTA CTA
CTA ATC CTA GCA ATC ATT CCC CTT CTT CAT ACA TCA AAA CAA CGA AGC ATA CTA TTT CGC CCT TTC AGC CAA TGC TTA TTC TGA ATC CTA
TTA ATC CTA CTA ATC ATC CCA CTG CTC CAC ACA TCA AAA CAA CGA AGC CTA ATG TTC CGA CCA ATT TCA CAG ATT ATG TTC TGA TTA CTA

GCA GCC AAT CTC CTC ATC CTC ACA TGA ATC GGA GGA CAA CCC GTT GAG CAT CCC TAC ATC ACC ATT GGC CAG TCG GCC TCC ATC CCC TAC
GCT GCC AAC CTA CTT ATC CTT ACA TGA ATT GGA GGC CAA CCA GTT GAA CAC CCA TAC ATT ACC ATT GGT CAA CTA GCA TCC ATC TCC TAC
GTA GCC AAC CTT TTA ACC CTT ACA TGA ATT GGA GGA CAA CCA GTA GAA CAA CCT TTT ATT ATT ATT GGT CAA CTA GCC TCA ACC CTA TAT

TTC TTC ATT ATC TTA ATC CTT TTC CCC CTG ACG AGC CTA TTA GAA AAC AAA ATA TTA AAA TGA ··· ··· ··· AGA
TTC TCT ATT CTA CTA ATT ATT ATA CCC CTA ACT AGC ATT ATA GAA AAC AAA CTA CTT AAA TGA ··· ··· ··· AGA
TTT TCA CTC ATT ATT ATC TTT ATG CCA TTA GCA GGT ATA TAT GAA GAT CAT TTA CTT GAA CCA AAA TTT CCA TAG
```

Fig. 1. The nucleotide sequences of the cytochrome *b* genes from the guinea pig, the African porcupine, and a South American opossum (*Monodelphis domestica*). The initiation codon ATG is omitted from the alignment and the stop codon is underlined. The GenBank accession numbers for the new sequences are X70673 and X70674.

are also differences between the trees. First, in Irwin et al.'s tree, the three dolphin sequences are clustered inside the artiodactyl sequences, whereas in our study they are outside the artiodactyls, though the distance separating the two groups is very small. In this respect, our tree is more reasonable than Irwin et al.'s tree, for the latter suggests that the camel is closer to dolphins than to other

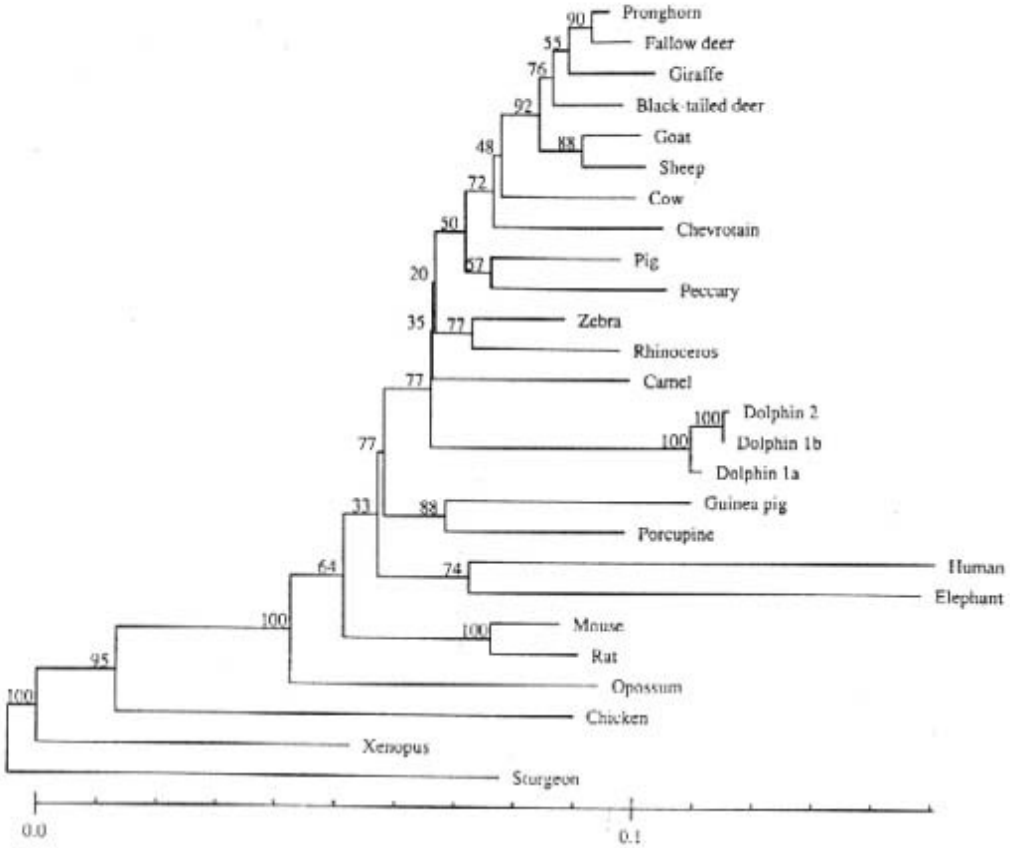

Fig. 2. A phylogenetic tree inferred from the cytochrome *b* sequences by the neighbor-joining method. The branch lengths are proportional to the number of nucleotide substitutions per nonsynonymous site. The number on each branching point indicates the proportion of bootstrap replicates (100 replicates) in which the subset of taxa were clustered as a group.

artiodactyls. Second, in Irwin et al.'s tree, the perissodactyl (zebra-rhinoceros) cluster branches off first, before the pig-peccary cluster, which in turn branches off before the camel. In our tree, the camel branches off first, then the zebra-rhinoceros cluster, and then the pig-peccary cluster. In this respect, Irwin et al.'s tree is more reasonable, but the internal branches separating these groups are very short in both trees. Third, Irwin et al.'s tree is unrooted, and therefore it is not possible to infer the order in which the various eutherian taxa diverged from one another. In comparison, our tree is rooted by four outgroup sequences, and therefore the sequence of divergence can be inferred. From Fig. 2 we see that the myomorph (mouse and rat) cluster is the first eutherian lineage to have branched off, in agreement with Li et al. (1990).

In Fig. 2, goat, sheep, and cow are not placed in the same clade. Despite the fact that this separation is supported by 92% of the bootstrap replicates, it is probably an erroneous arrangement because the three species are commonly thought to belong to the same clade (e.g., Young 1981), and the monophyly of the family Bovidae is supported by mtDNA sequence data from a 2.7-kb stretch that covers the 12S and 16S rRNA genes and three adjacent tRNA genes (Allard et al. 1992).

A major question we wanted to address is whether or not the guinea pig and the porcupine are sister taxa. The tree in Fig. 2 supports this grouping; in 88 out of the 100 bootstrap replicates the two taxa ended up clustered together. However, Fig. 2 also suggests that the guinea pig and the porcupine lineages are quite distantly related. Another major question was whether or not the hystricomorphs (porcupines and guinea pigs) and the myomorphs (mice and rats) are sister groups. In our tree, these groups are not clustered together. The result is in agreement with the hypothesis that the hystricomorphs and the myomorphs do not belong to the same order (Graur et al. 1991; Li et al. 1992a). However, in the present tree the hystricomorph se-

quences branch off after the divergence of the human sequence, and this does not support the view that the hystricomorph lineage has branched off earlier than the divergence between the primate and the myomorph lineages (Graur et al. 1991; Li et al. 1992a). In summary, the COB sequences support neither the traditional tree, in which the myomorphs and the hystricomorphs are monophyletic, nor Graur et al.'s (1991) tree, in which the hystricomorphs are an outgroup to the myomorphs, the primates, and the artiodactyls.

Note, however, that all the above suggestions are based on a single sequence, and so the conclusions should be regarded as tentative. The bootstrap resampling indicates that high confidence can be placed on only a few branching points: the separation of tetrapods from the other vertebrates, the separation of mammals and avians from amphibians, the separation of mammals from birds, the clustering of mouse and rat, and the two clusterings of the three dolphin sequences. All other subsets of mammalian taxa appeared in less than 95% of the bootstrapping replicates, and should, therefore, be treated with caution.

Relationship Between Function and Sequence Conservation

Cytochrome *b* is involved in electron transport. It is thought to contain two quinone reaction sites—one (Q_o) located on the proton output side of the mitochondrial membrane and the second (Q_i) on the proton input side of the membrane (Crofts et al. 1987; Howell and Gilbert 1988; Howell 1989). The Q_o site constitutes the ubiquinol oxidizing portion of the Q cycle, while the Q_i site functions as ubiquinone reductase. Q_o is composed of several segments. The segment that extends approximately from residue 139 to 149 (or 159) is involved in the binding of or interaction with Q_o inhibitors. This region is very well conserved (Fig. 3; see also Irwin et al. 1991). However, the residues at positions 158 and 159 are not well conserved, for several substitutions are observed—e.g., from nonpolar threonine (T) to polar aspartic acid (D). Thus, the segment involved in the Q_o site may not extend to positions 158 and 159. On the other hand, it might extend to position 126 on the N-terminal side because the residues at positions 126, 127, and 128 are invariant among the 26 sequences (Fig. 3). The second putative Q_o region, which extends approximately from residue 269 to residue 289 (or 294), is probably involved in the redox catalysis. This region is as well conserved as the previous region (Fig. 3). In this region the tripeptide proline–glutamic acid–tryptophan (PEW) at residues 270–272 is thought to constitute the center of the catalytic part. The tripeptide was suggested to be invariant in evolution (Howell 1989), but Fig. 3 shows that in the black-tailed deer W has been replaced by C (cysteine). This region also contains the amino acid leucine (L) at residue 293 that might be involved in binding the polyisoprenoid side chain of quinone/quinol at the Q_o site. This residue has been conserved among all the 26 sequences (Fig. 3). Other segments in the Q_o site are less well defined but may contain the segment from residues 69 to 80 (Howell 1989), which is not so well conserved as the previous two regions (Fig. 3).

The other quinone redox site of cytochrome *b*, Q_i, is less well defined but probably contains the segment from residues 21–41. This segment is also well conserved among the 26 sequences, though not so well as the Q_o segments. Positions 223 and 224, which are almost invariably occupied by a tyrosine-tyrosine (YY) dipeptide (Fig. 3), and position 231, which is invariably a glycine (G) in mammals, have also been postulated to be involved in the formation of the Q_i site.

The four histidines (H) at positions 83, 97, 182, and 196 are believed to ligate the two heme groups. These histidines are invariant in all the sequences so far examined (Fig. 3; Howell 1989; Irwin et al. 1991). However, the residues around these histidines, except those around His182, are not well conserved.

Cytochrome *b* has been suggested to contain eight transmembrane domains, denoted for historical reason as I, II, III, V, VI, VII, VIII, and IX (Crofts et al. 1987; Howell and Gilbert 1988; Howell 1989). These transmembrane domains are less well conserved than the sequences forming the Q_o and Q_i sites. In fact, domains VI and IX appear to be the most divergent parts of the protein (Fig. 3). The segment of the first 17 residues at the N-terminal end is also a divergent part of the protein.

Cann et al. (1984) showed that subunit II of cytochrome oxidase (COII), which is encoded by the mitochondrial genome, evolves much faster in the human lineage than in the rodent and artiodactyl lineages. (See also Brown and Simpson 1982.) A similar observation (Carlson et al. 1977; Evans and Scarpulla 1988) has been made concerning the rate of evolution of cytochrome *c*, which is encoded by the nuclear genome. Cytochrome *c* receives electrons indirectly from cytochrome *b* and passes them directly to COII. According to Cann et al., these "coordinated" accelerations in rates represent a case of coevolution between nuclear and mitochondrial components of a biochemical complex. In the lineage leading to the apes and humans, both cytochrome *c* and COII have undergone complementary functional changes away from the ancestral mammalian state. An interesting observation from the present phylogenetic analysis (Fig. 2) is that the

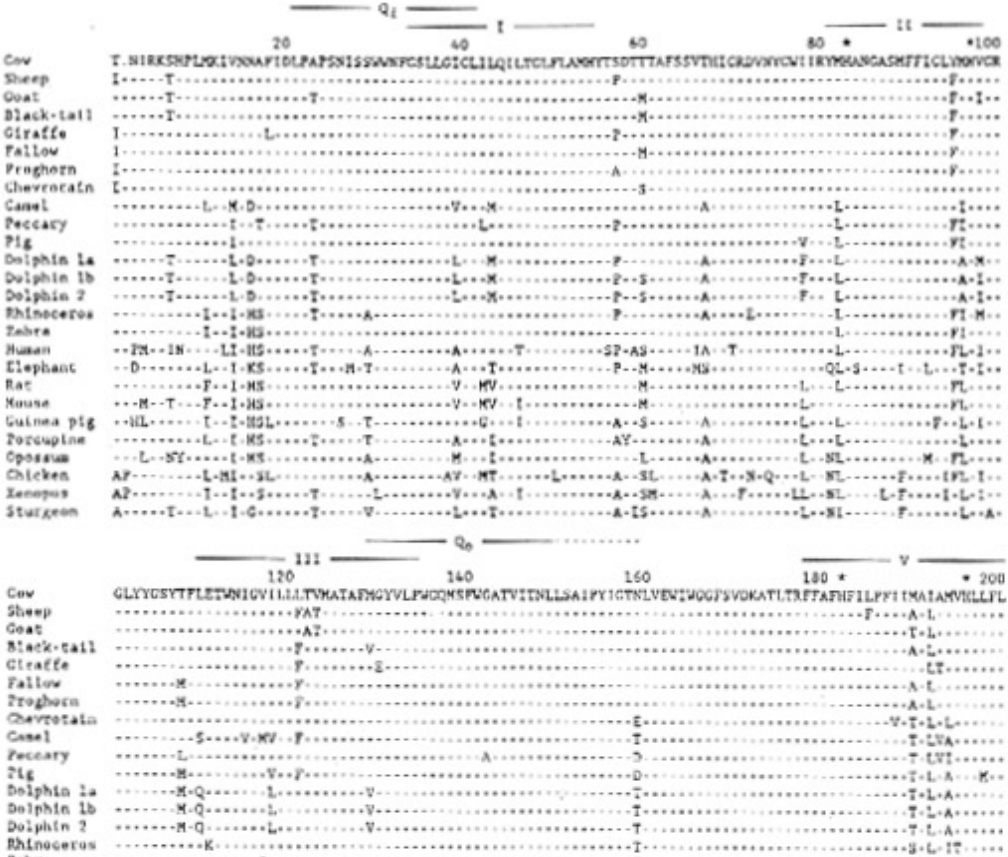

Fig. 3. Alignment of the 26 cytochrome *b* sequences used in the present study. *A dash* signifies that the amino acid at that position is the same as that of the first (cow) sequence and *a dot* indicates a deletion. The two reaction sites are denoted by Q_i and Q_o and the eight transmembrane domains are denoted by I, II, III, V, VI, VII, VIII, and IX. The four histidines at positions 83, 97, 182, and 196 are indicated by *asterisks* and the tripeptide PEW at positions 270–272 is indicated by #.

rate of substitution in the cytochrome *b* genes from human and elephant is much higher than that in other mammals. Given that cytochrome *b* is also involved in mitochondrial electron transport and oxidative phosphorylation, and must interact with proteins encoded by the nuclear genome (Hatefi 1985), it is possible that the acceleration in rates seen in humans and elephants represents another case of coevolution between mitochondrial and nuclear genes. It is therefore interesting to see whether there are unique changes in the reaction sites in the human and elephant cytochrome *b*. Indeed, in the Q_o site, there are two unique changes in human cytochrome *b:* at position 70, c (cysteine) is replaced by T (threonine), and at position 279, A (alanine) is replaced by T. In the Q_i site, there is a unique change in the elephant cytochrome *b:* at position 27, I (isoleucine) is replaced by M (methionine). However, whether any of these changes have a significant effect on the interaction between cy-

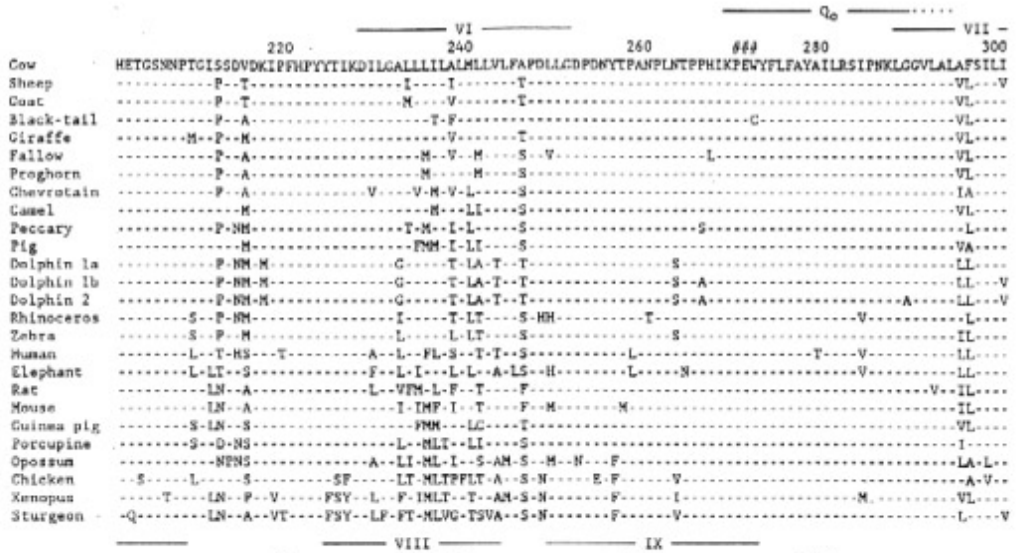

Fig. 3. Continued.

tochrome *b* and cytochrome *c* needs to be tested experimentally. We also note that most of the changes in the human and elephant cytochrome *b* sequences occurred outside the reaction sites, and thus whether many of these changes had arisen from coevolution remains to be determined.

Acknowledgments. We thank Dr. R. Honeycutt for the African porcupine liver tissue, Dr. D.M. Irwin for sending us his cytochrome *b* sequences, and Dr. T.S. Whittam for sending us his bootstrap program for the neighbor-joining method. This study was supported by grants from NIH.

References

Allard MW, Miyamoto MM, Honeycutt RL (1991) Tests for rodent polyphyly. Nature 353:610–611

Allard MW, Miyamoto MM, Jarecki L, Kraus F, Tennant MR (1992) DNA systematics and evolution of the artiodactyl family Bovidae. Proc Natl Acad Sci USA 89:3972–3976

Anderson S, Gait MJ, Mayol L, Young IG (1980) A short primer for sequencing DNA cloned in the single-stranded phage vector M13mp2. Nucleic Acids Res 8:1731–1743

Brown GG, Simpson MV (1982) Novel features of animal mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes. Proc Natl Acad Sci USA 79:3246–3250

Cann RL, Brown WM, Wilson AC (1984) Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106:479–499

Carlon SS, Mross GA, Wilson AC, Mead RT, Wolin LD, Bowers SF, Foley NT, Muijsers AO, Margoliash E (1977) Primary structure of mouse, rat, and guinea pig cytochrome C. Biochemistry 16:1437–1442

Crofts A, Robinson H, Andrews K, Van Doren S, Berry E (1987) Catalytic sites for reduction and oxidation of quinones. In: Papa S, Chance B, Ernster L (eds) Cytochrome systems: molecular biology and bioenergetics. Plenum Press, New York, pp 617–624

Dale RMK, McClure BA, Houchins JP (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid 13: 31–40

Evans M, Scarpulla RC (1988) The human somatic cytochrome c gene: two classes of processed pseudogenes demarcate a period of rapid molecular evolution. Proc Natl Acad Sci USA 85:9625–9629

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Graur D, Hide WA, Li W-H (1991) Is the guinea-pig a rodent? Nature 351:649–652

Graur D, Hide WA, Zharkikh A, Li W-H (1992) The biochemical phylogeny of guinea-pigs and gundis, and the paraphyly of the order rodentia. Comp Biochem Physiol 101B:495–498

Hasegawa M, Cao Y, Aduchi J, Yano T-A (1992) Rodent polyphyly? Nature 355:595

Hatefi Y (1985) The mitochondrial electron transport and oxidative phosphorylation system. Annu Rev Biochem 54:1015–1069

Howell N (1989) Evolutionary conservation of protein regions in the protonmotive cytochrome *b* and their possible roles in redox catalysis. J Mol Evol 29:157–169

Howell N, Gilbert K (1988) Mutational analysis of the mouse mitochondrial cytochrome *b* gene. J Mol Biol 203:607–617

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. J Mol Evol 32:128–144

Li W-H, Luo C-C, Wu C-I (1985) Evolution of DNA sequences. In: MacIntyre RJ (eds) Molecular evolutionary genetics. Plenum Press, New York, pp 1–94

Li W-H, Gouy M, Sharp PM, O'hUigin C, Yang Y-W (1990) Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. Proc Natl Acad Sci USA 87:6703–6707

Li W-H, Hide WA, Graur D (1992b) Origin of rodents and guinea-pigs. Nature 359:277–278.

Li W-H, Hide WA, Zharkikh A, Ma D-P, Graur D (1992a) The molecular taxonomy and evolution of the guinea pig. J Heredity 83:174–181

Luckett WP, Hartenberger (eds) J-L (1985) Evolutionary relationships among rodents. Plenum, New York

Novacek MJ (1992) Mammalian phylogeny: Shaking the tree. Nature 356:121–125

Nowak RM, Paradiso JL (1983) Walker's mammals of the world. Johns Hopkins University Press, Baltimore and London

Romer AS, Parsons TS (1977) The vertebrate body. Saunders, Philadelphia

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibition. Proc Natl Acad Sci USA 74: 5463–5467

Thuring RWJ, Sanders JPM, Borst P (1975) A freeze-thaw method for recovering long DNAs from agarose gels. Analyt Biochem 66:213–220

Young JZ (1981) The life of vertebrates. Clarendom Press, Oxford, England

Received July 22, 1992/Revised October 30, 1992