Phylogenetic Performance of Mitochondrial Protein-Coding Genes in Resolving Relationships Among Vertebrates

Rafael Zardoya and Axel Meyer

Department of Ecology and Evolution, and Program in Genetics, State University of New York at Stony Brook

A large number of studies in evolutionary biology utilize phylogenetic information obtained from mitochondrial DNA. Researchers place trust in this molecule and expect it generally to be a reliable marker for addressing questions ranging from population genetics to phylogenies among distantly related lineages. Yet, regardless of the phylogenetic method and weighting treatment, individual mitochondrial genes might potentially produce misleading evolutionary inferences and hence might not constitute an adequate representation neither of the entire mitochondrial genome nor of the evolutionary history of the organisms from which they are derived. We investigated the performance of all mitochondrial protein-coding genes to recover two expected phylogenies of tetrapods and mammals. According to these tests, mitochondrial protein-coding genes can be roughly classified into three groups of good (ND4, ND5, ND2, cytb, and COI), medium (COB, COIII, ND1, and ND6), and poor (ATPase 6, ND3, ATPase 8, and ND4L) phylogenetic performers in recovering these expected trees among phylogenetically distant relatives. How general our findings are is unclear. Simple length differences and rate differences between these genes cannot account for their different phylogenetic performance. The phylogenetic performance of these mitochondrial genes might depend on various factors that play a role in determining the probability of discovering the correct phylogeny such as the density of lineage creation events in time, the phylogenetic "depth" of the question, lineage-specific rate heterogeneity, and the completeness of taxa representation.

Introduction

Traditionally, the evolutionary application of mitochondrial DNA (mtDNA) had been limited to population-level questions (reviewed in Avise 1994). However, since the advent of the polymerase chain reaction (PCR; Saiki et al. 1988) and direct sequencing techniques (Kocher et al. 1989) the use of mtDNA as a phylogenetic marker has been extended to much greater levels of phylogenetic inclusiveness (Meyer and Wilson 1990). The relative technical ease with which DNA sequences of mitochondrial genes can be determined through this methodology and special features of mitochondrial DNA (i.e., lack of introns, maternal inheritance, absence of recombination events, and haploidy) have made it the most common type of sequence information used to estimate phylogenies among both closely and distantly related taxa (reviewed in Meyer 1993; Avise 1994). Deservedly, mitochondrial DNA has a reputation of being a reliable tracer of evolutionary history (Meyer 1993, 1994*a*, 1994b; Hillis, Huelsenbeck, and Cunningham 1994; Hillis, Huelsenbeck, and Swofford 1994; Edwards et al. 1995; Moore 1995).

In conducting molecular phylogenetic studies with emphasis more on investigating evolutionary relationships among species and less on the evolution of the sequenced genes itself, a trade-off is exercised at the moment when a particular gene is chosen over an alternative one. In many recent phylogenetic studies, DNA fragments of 300-600 base pairs from a single gene or from a small number of genes are routinely amplified and sequenced from different species in order to infer their evolutionary relationships (Kocher et al. 1989). Consideration in these studies aims to maximize the amount of phylogenetic information per unit effort of sequencing, and tests have been conducted to predict how many nucleotides need to be determined without sacrificing phylogenetic information (e.g., Tajima 1991; Rheztsky and Nei 1994; Cummings, Otto, and Wakeley 1995; Martin et al. 1995). Most studies assume that the information derived from partial mitochondrial data sets is representative of the entire history of both the molecule and the organism from which it is derived. Recently, however, doubts were raised as to whether all mitochondrial genes are equally appropriate for all evolutionary questions and whether individual mitochondrial genes are representative of the entire mitochondrial genome (e.g., Cao et al. 1994; Graybeal 1994; Meyer 1994b; Cummings, Otto, and Wakeley 1995).

We recently determined the complete mitochondrial DNA (mtDNA) sequence of the African lungfish (Zardoya and Meyer 1996). Two data sets, one comprising all mitochondrial tRNA genes and the other combining all mitochondrial protein-coding genes were subjected to phylogenetic analyses using maximum-likelihood (ML), neighbor-joining (NJ), and maximum-parsimony (MP). The resulting identical trees confirmed, with strong statistical confidence, the crucial phylogenetic proximity of lungfish to tetrapods (Zardoya and Meyer 1996) (fig. la). Surprisingly, however, this well-established topology for tetrapod relationships is not always recovered by the same phylogenetic methods with individual mitochondrial protein-coding gene sequences. This might suggest that, regardless of phylogenetic method (Hillis, Huelsenbeck, and Cunningham 1994; Hillis, Huelsenbeck, and Swofford 1994; Edwards et al. 1995) and exclusion treatment (Milinkovitch, Orti, and Meyer 1995), individual mitochondrial genes might potentially produce misleading evolutionary inferences and might not constitute an adequate representation of the

Key words: mtDNA, vertebrate phylogeny, maximum likelihood, maximum parsimony, protein-coding genes.

Address for correspondence and reprints: Rafael Zardoya, Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, New York 11794-5245. E-mail: rzardoya@life.bio.sunysb.edu.

Mol. *Biol. Evol.* 13(7):933–942. 1996
© 1996 by the Society for Molecular Biology and Evolution. ISSN: 0737.4038

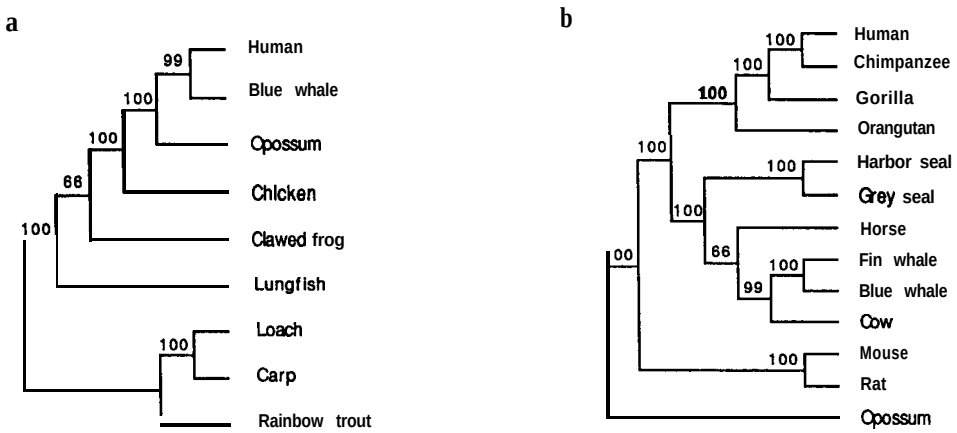

FIG. I.-Majority-rule bootstrap (Felsenstein 1985) consensus tree of (a) tetrapods and ***(b)*** mammals. All **positions of all mitochondrial** protein-coding genes combined were subjected to MP analyses. Numbers above branches indicate bootstrap values for 100 replicates.

entire mitochondrial genome. This may be the case at least for an evolutionary question that involves divergences that date back to the Devonian (fig. la). For population-level questions it probably is less important which protein-coding gene is sequenced since most of the variation is likely to be derived from variation at third codon positions and there these genes will have quite comparable rates of evolution.

To address why particular mitochondrial protein-coding genes can fail to recover the expected phylogenetic tree and, consequently, to evaluate and to delimit the phylogenetic utility of each mitochondrial protein gene, we have analyzed their phylogenetic performance in two different time frames: (1) that among major tetrapod lineages (divergences of >300 MYA) and (2) that among major mammalian groups (divergences of 5-100 MYA).

Materials and Methods

Data Sets

The two data sets analyzed in this study comprise the following 19 complete vertebrate mitochondrial DNA genomes (fig. 1): trout, ***Oncorhynchus mykiss (L2977*** 1; Zardoya, Garrido-Pertierra, and Bautista 1995); carp, ***Cyprinus carpio (X61010;*** Chang, Huang, and Lo 1994); loach, ***Crossostoma lacustre*** (M91245; Tzeng et al. 1992); African lungfish, ***Protopterus dolloi (L42813;*** Zardoya and Meyer 1996); clawed frog, *Xenopus laevis* (M10217; Roe et al. 1985); chicken, *Gallus gallus* ***(X52392;*** Desjardins and Morais 1990); opossum, *Didelphis* ***virginiana (229573;*** Janke et al. 1994); rat, ***Rattus norvergicus (X14848;*** Gadaleta et al. 1989); mouse, *Mus* ***musculus*** (J01420; Bibb et al. 1981); grey seal, ***Hulichoerus grypus (X72004;*** Arnason et al. 1993); harbor seal, ***Phoca vitulina (X63726;*** Arnason and Johnsson 1992); horse, ***Equus caballus (X79547; Xu*** and Arnason 1994); cow, ***Bos taurus (VO0654;*** Anderson et al. 1982); blue whale, ***Balaenoptera musculus (X72204;*** Arnason and Gullberg 1993); fin whale, *Balaenoptera* ***physalus (X61 145;*** Arnason, Gullberg, and Widegren 199 1); orangutan, ***Pongo pygmaeus (D38 115;*** Horai et al. 1995); gorilla, ***Gorilla gorilla (D38114;*** Horai et al. 1995); chimpanzee, ***Pun troglodytes (D38113;*** Horai et al. 1995); human, ***Homo sapiens (501415;*** Anderson et al. 1981).

Phylogenetic Analyses

DNA sequences were aligned using PILEUP (Devereux, Haeberli, and Smithies 1984) and CLUSTAL W (Thompson, Higgins, and Gibson 1994), followed by refinement by eye based on the corresponding deduced amino acid sequences. Gaps resulting from the align-

Table 1
Statistical Confidence of Maximum-Likelihood and Maximum-Parsimony Trees

	ATPASE 6		ATPASE 8		co1		COII		COIII	
	A	SE	A	SE	A	SE	A	SE	A	SE
Kishino-Hasegawa test										
All positions vertebrates	-21	10.8	-2.1	2.9	-8.8	16.4	-14	9.1	-8.6	18.4
mammals	-17	14.1	-8.6	4.6	-7.3	18.4	-14	13.4	-0.3	8.7
No 3rd position. vertebrates	-16	8.5	-3.7	3.6	-5.4	13.4	-11	10.2	-3.3	4.9
mammals	-8.8	15.8	-6.3	6	-34	22.2	-4.3	7.6	-23	19
Templeton test										
All positions . vertebrates	19	8.06	5	5.21	10	11.4	13	9.11	10	8.37
mammals	12	8.13	8	5.11	11	9.11	3	6.4	7	5
No 3rd position. vertebrates	10	4.47	5	4.14	3	5	10	5.1	1	2.23
mammals	5	4.6	7	3.88	14	6.32	0	3.47	9	6.1

ment were treated as missing data. Ambiguous alignments, mainly at 5' and 3' ends of the protein-coding genes, were excluded from the phylogenetic analyses. Overlapping positions (two open reading frames) in several genes (ATPase8/ATPase6, ND4L/ND4, ND5/ND6) were duplicated in the analyses.

All 13 protein-coding genes combined and each one separately were subjected to the MP method (PAUP version 3.1.1., Swofford 1993) using heuristic searches (TBR branch swapping; MULPARS option in effect) with simple stepwise addition of taxa to find the most parsimonious tree. Transitions and transversions were given equal weight (for exceptions see below). When two or more parsimonious trees were produced, a 50% majority-rule consensus tree was constructed. ML (default model of DNAML in PHYLIP, where transversions are given double the weight of transitions and empirical base frequencies are used) and NJ (Saitou and Nei 1987) (based on Kimura distance matrices) analyses of the sequences were performed with PHYLIP (version 3.5) (Felsenstein 1989). Analyses with all phylogenetic methods were also performed, entirely excluding third codon positions in each gene. A third set of MP analyses were performed in which only third-codon-position transitions were excluded. Robustness of the phylogenetic results was tested by bootstrap analyses (Felsenstein 1985) (as implemented in PAUP and PHYLIP with 100 replications each).

Statistical Methods

The statistical confidence of the resulting best trees of each ML analysis with respect to alternative hypotheses was evaluated by calculating the standard deviation of the difference in log-likelihoods between the resulting best tree and the expected alternative hypotheses using the formula of Kishino and Hasegawa (1989) as implemented in the MOLPHY program (Adachi and Hasegawa 1992). If standard deviations where less than 1.96 times the difference in log-likelihoods between two competing phylogenetic hypotheses then the two phylogenies were judged to differ significantly. Similarly, for MP analyses statistical confidence was assessed by calculating the standard deviation of the difference in number of steps between the resulting most parsimonious tree and the expected alternative tree using the method of Templeton (1983) as implemented in PHYLIP (table 1). In this case, competing trees were declared significantly different when standard deviations were found to be less than 1.96 the difference in number of steps (Felsenstein 1989).

Results and Discussion

Combined Protein Genes Analyses

The combined data set of all mitochondrial protein genes yielded, with all three commonly used methods of phylogenetic inference (ML, NJ, and MP), identical and strongly bootstrap-supported trees that recovered the expected (according to morphological, paleontological, and other molecular evidence) tetrapod relationships (fig. 1*A*). Identical trees were recovered by all three phylogenetic methods when only first and second positions were analyzed or by MP when only transversions in third positions were considered.

However, for the corresponding mammalian data set only MP arrived at the expected (according to Janke et al. 1994; Horai et al. 1995; Krettek, Gullberg, and Arnason 1995) topology (fig. 1*B*). With both NJ and ML methods the horse was identified as sister group of seals and not, as expected, of artiodactyls+cetaceans. The same phylogenetic relationships (fig. 1*B*) were obtained when third codon positions were excluded from the analyses (with all phylogenetic methods) or transitions in third codon positions were not considered (this treatment was only conducted for MP). The NJ method supported a (horse+seals) grouping with 89% bootstrap support (fig. 1*B*). It would appear that artiodactyl+cetacean, perisodactyl, and carnivore relationships are still somewhat unclear (e.g., Cao et al. 1994; Krettek, Gullberg, and Arnason 1995) and that differences between the different phylogenetic methods used are indicative of the uncertainty in these relationships and possibly due to the fast, hard-to-resolve origin of these lineages. For the investigation of the performance of individual mitochondrial protein-coding genes with the NJ method the (seal+horse) grouping was used as the null hypothesis against which we tested their performance.

Performance of Individual Genes

The same two sets of phylogenetic analyses as above were performed on both data sets with all 13 mi-

Table 1
Extended

Cytb		ND1		ND2		ND3		ND4		ND4L		ND5		ND6	
A	SE	A	SE	A	SE	A	SE	A	SE	A	SE	A	SE	A	SE
-6.2	14.8	-9.9	8.9	Best		-8.7	5	Best		-15	6.6	-15	16.7	-1.6	4
-2.6	10.5	-13	12.6	-5.1	6.9	-5.4	10.9	Best		-6.2	6.7	-13	15.5	-17	10.8
-7.4	13.1	-7.8	8.7	Best		-16	8.6	Best		-9.4	5.5	Best		-4.3	9
-1.9	6.9	-24	14.9	-9.3	6.3	-4.4	6.9	-0.4	6.3	-9.2	5.7	-9.8	11.8	-12	7.8
11	10.1	6	6.63	6	10.4	18	7.88	Best		8	6	9	12.8	13	10.3
9	8.1	6	10	Best		3	5.75	7	9.54	2	4.25	3	10.7	3	8.78
4	6	4	3.16	3	7.42	13	4.8	3	6.24	9	4.36	Best		2	4.24
0	5.66	14	6	2	4.69	4	3.75	1	5	5	2.65	1	6.4	9	5.57

Table 2
Incorrect High Bootstrap Value Support of Unexpected Grouping Provided by Mitochondrial Genes

Gene	Method	Positions	Odd Grouping	Bootstrap
ATPase 6	MP	All	Chicken+lungfish	73%
ATPase 6	NJ	All	Chicken+lungfish	89%
ATPase 6	NJ	All	Horse+seals	73%
ATPase 6	NJ	No 3rd	Horse+seals	88%
ATPase 8	NJ	All	Marsupial + whale	82%
ATPase 8	NJ	No 3rd	Marsupial + whale	73%
ND3.	MP	No 3rd	Lungfish+mammals	94%
ND3.	NJ	All	Lungfish+mammals	89%
ND3.	NJ	No 3rd	Lungfish+mammals	96%
COII.....	NJ	No 3rd	Marsupial+whale	72%
COII.....	MP	No 3rd	Marsupial + whale	91%
ND4L	NJ	No 3rd	Frog+mammals	75%
ND4L	NJ	All	Frog+mammals	94%
ND1..	MP	All	Frog+lungfish	73%
ND1..	NJ	All	Horse + seals	86%
ND4.	NJ	All	Horse+seals	80%
ND5.	NJ	All	Lungfish+mammals	82%
ND5.	MP	No Ts in 3rd	Lungfish+mammals	74%
ND5.	NJ	All	Horse+ seals	85%
ND5.	NJ	No 3rd	Horse+seals	73%

tochondrial protein-coding genes individually. Of these, only ND4, ND2, and ND5 (and not even with all phylogenetic methods) were able to recover the expected topologies for tetrapods and mammals (table 1). In some cases, the unexpected topologies from some genes (mainly ND3, ND4L, and ATPase 6) were incorrectly supported by relatively high bootstrap values (table 2) that would typically be interpreted as strong support by the data for a particular hypothesis (Hillis and Bull 1993). Several of the individual genes support the (horse+seals) grouping more strongly than the "expected" (horse(artiodactyl+cetacean)) grouping (table 2). Some of these "spurious" results may not be all that surprising since it has been demonstrated that wrong groupings can be recovered if the time interval separating relevant branchings is short (Cao et al. 1994). In some cases, particular clades (i.e., [lungfish+mammals], [marsupial+whales], see table 2) were repeatedly grouped, suggesting specific biases in those lineages. High bootstrap values for these incorrect groupings were observed more often in NJ than in MP tests (table 2). These results may indicate that using resampling methods such as bootstrapping with a purely algorithmic method such as NJ rather than one with an underlying optimality criterion is inappropriate. This is because any incorrect bias that might systematically mislead the algorithm will, with high bootstrap values for incorrect groupings, falsely indicate robustness of the nodes and statistical confidence of the resulting tree (but see Russo, Takezaki, and Nei 1996).

To further evaluate the phylogenetic utility of single mitochondrial protein genes and to investigate potential causes for their phylogenetic failure, the statistical support of the best tree yielded by each gene against the expected tree was estimated. This was done by calculating the standard errors of the differences in log-likelihoods (Kishino and Hasegawa 1989) and in the number of steps (Templeton 1983) between the most likely or most parsimonious tree and the expected tree, respectively (table 1). Both tests yielded largely congruent results, indicating that all single mitochondrial genes with the exception of ATPase6, ND3, and ND4L equally strongly supported the expected and the recovered tetrapod tree (table 1). Similarly, according to both tests, all genes except COI and ND1 (when third positions were excluded from the analyses) did not statistically differ in their support for the recovered and the expected ("best") mammal tree (table 1). Hence, most mitochondrial protein-coding genes had the potential to recover the expected tree since the recovered trees were not significantly different from the expected tree. Nevertheless, most individual genes consistently failed to arrive at the exact expected branching order of both the tetrapod and mammal relationships (fig. 1 and table 1).

In order to test whether the superior performance of some genes was simply due to their lengths and hence a larger number of phylogenetically informative sites (Stewart 1993; Cummings, Otto, and Wakeley 1995), the differences in log-likelihoods (Kishino and Hasegawa 1989) and numbers of steps (Templeton 1983) were standardized for the length of each gene by dividing those values by the number of nucleotides in each gene (fig. 2). The results indicated that the performance of single genes was not simply related to their size but some genes seem to be consistently more reliable tracers of evolutionary history than others (fig. 2). Mitochondrial protein-coding genes might be roughly classified into three groups of relatively good (ND4, ND5, ND2, cytb, and COI), medium (COII, COIII, ND 1, and ND6), and poor (ATPase 6, ND3, ATPase 8, and ND4L) phylogenetic performers in recovering the expected trees from this study (fig. 2). This generalization seems to hold more strongly for the tetrapod than for the mammalian test data set (table 1 and fig. 2). The chances of recovering the expected tree using the first and second categories of genes for the posed phylogenetic problems are higher when third codon positions are completely excluded from the analyses (table 1 and fig. 2). However, it should be noted that cytochrome oxidase genes, due to their highly conserved amino acid sequence, have a poor performance in recovering the mammalian phylogeny if third positions are excluded (Cao et al. 1994). Somewhat unexpectedly, the performance of genes of the third category is improved when all positions are included in the analyses (table 1 and fig. 2).

The ability of mitochondrial genes to infer phylogenetic relationships can be challenged when their rates of evolution differ among lineages, particularly with MP and less so with ML (Adachi and Hasegawa 1995). To address to what extent phylogenetic performance of different genes is related to their particular rates of evolution, we calculated sequence divergences (from Kimura two-parameter distance matrices) for each individual gene among the taxa studied. Divergence times for the taxa analyzed were inferred using information from the fossil record summarized in Benton (1990), and estimates calculated from mitochondrial data by Janke et al. (1994) (fig. 3).

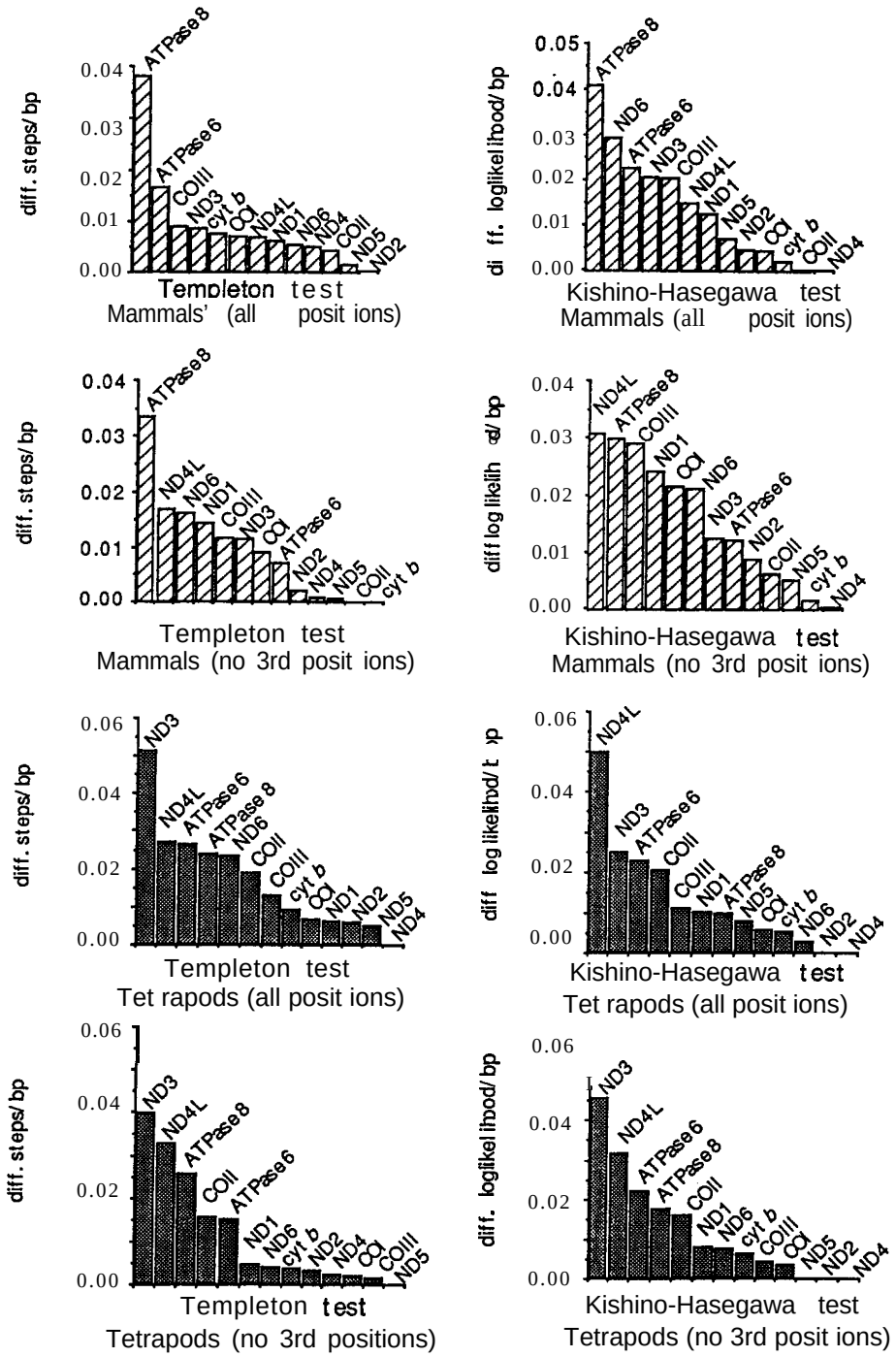

FIG. 2.-Phylogenetic performance of mitochondrial genes. Differences in log-likelihood (Kishino and Hasegawa 1989) and number of steps (Templeton 1983) between the most likely or most parsimonious tree and the recovered tree, respectively, were standardized for the length of each gene. The standardized values estimated for the tetrapod and mammalian data sets (both when all positions and when only first and second positions were included in the analyses) were sorted providing a ranking for the performance and utility of the mitochondrial genes assayed.

For all mitochondrial genes, both considering all positions or excluding third positions, sequence divergence increases linearly until a plateau is reached around a 100 MYA (data not shown) due to the effect of multiple hits. Comparison of the shapes of the curves yielded by each mitochondrial protein-coding gene can provide heuristic information about its phylogenetic utility (Meyer 1993, 1994*a*, 1994b; Graybeal 1994) and is, among other factors, related to functional constraints acting on the gene product (Degli Esposti et al. 1993). In order to facilitate comparisons of rates among different genes, their curves were linearized by calculating and plotting the logarithms of the distance and the divergence time (twice the time back to their inferred

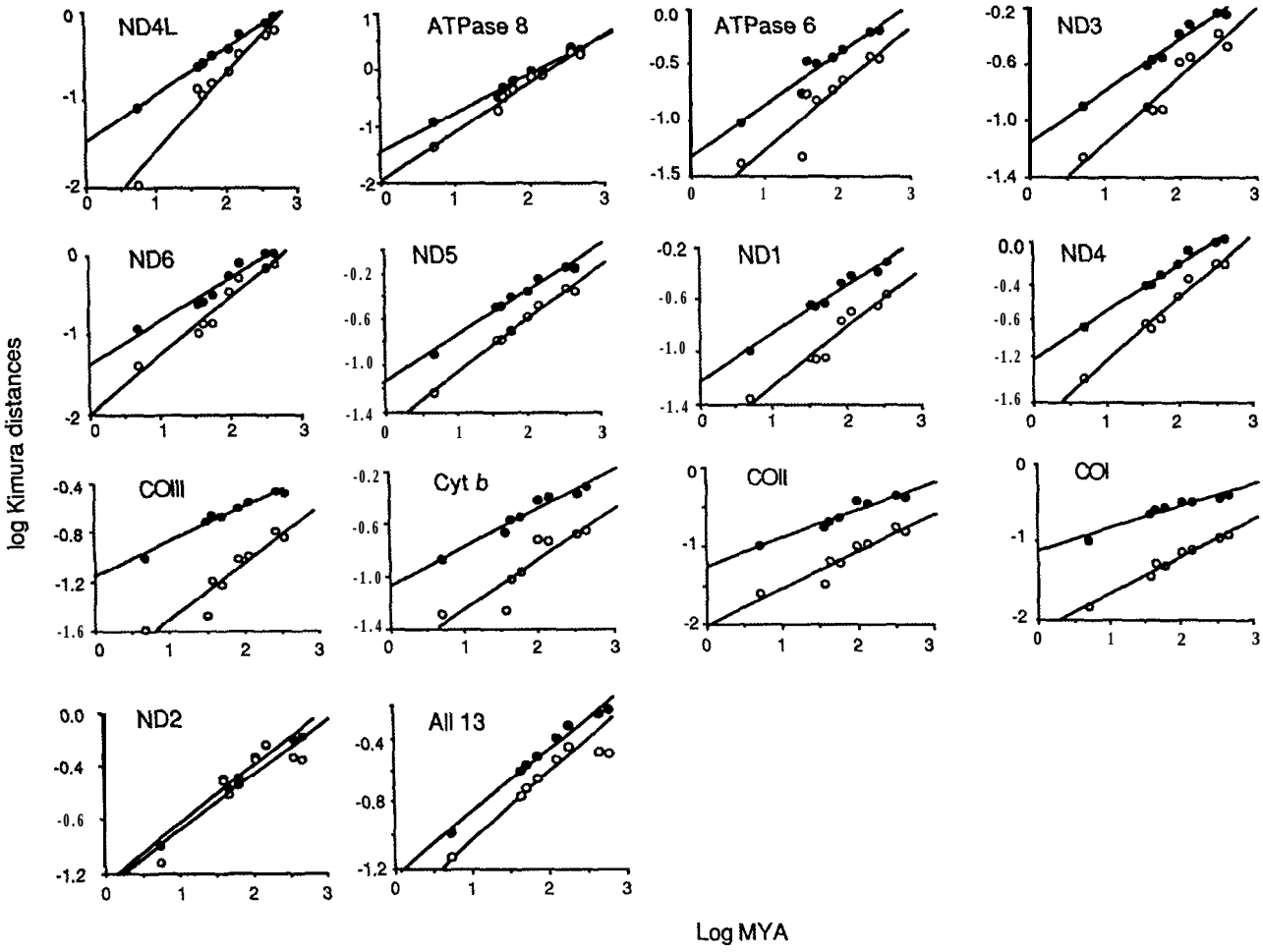

FIG. 3.—Least-squares regression analyses for the logarithm transforms of the distance and the divergence times between eight pairs of all species (human-chimpanzee: 4.9 MYA; mouse-rat: 35 MYA; whales-cow: 41 MYA; carnivores-artiodactyls: 55 MYA; primates-ungulates: 93 MYA; opossum-placental mammals: 130 MYA; chicken-mammals: 300 MYA; lungfish-tetrapods: 398 MYA; see text for details and sources of divergence estimates). Kimura two-parameter distances were estimated including all (solid dots) or only first and second positions (open circles). Divergence times between two taxa are double the time since their divergence time from their common ancestor.

Table 3
Least-Squares Regression Parameters for the Log-Log Transformation of the Distance and Divergence Time Data

	All Positions		No 3rd Positions	
Gene	Slope	Intercepts	Slope	Intercepts
ATPase 8	**0.66**	**-1.53**	**0.91**	-2.41
ATPase 6	0.45	−1.32	0.55	-1.81
COI	0.28	-1.16	0.47	− 1.99
COII	0.36	-1.25	0.46	−1.55
COIII	0.29	-1.15	0.44	-1.71
Cytb	0.29	-1.11	0.55	-1.82
ND1	0.36	-1.23	0.83	-1.99
ND2	0.45	-1.25	0.73	-2.01
ND3	0.37	-1.16	0.45	-2.11
ND4	0.41	-1.25	0.48	-1.64
ND4L	0.55	-1.48	0.37	−1.63
ND5	0.39	-1.17	0.43	− 1.28
ND6	**0.55**	**-1.41**	**0.71**	**-2.38**
All	**0.38**	**-1.22**	**0.33**	**-1.29**

common ancestor) for each pairwise comparison (fig. 3). By this transformation, data points can be easily fitted by a generalized least-squares regression model and slopes for each gene can be calculated (fig. 3). However, because of obvious concerns about the nonindependence of the data points (obtained from related taxa), parameter estimates and tests of significance derived from these regressions need to interpreted with caution (e.g., Graybeal 1994).

Rates of evolution vary considerably between these genes (table 3). In the faster-evolving genes (e.g., ND4L and ATPase 8) the divergence slopes seem to cross earlier, indicating that first and second codon positions, which result in amino acid differences, evolve rather quickly compared to the more conservative genes (e.g., cytochrome oxidase subunits and cytochrome *b)* which show overall much slower rates of evolution at the amino acid level and, hence, in first and second codon positions (table 3 and fig. 3). However, simple rates of evolution would not appear to be a good predictor of the phylogenetic performance of individual mitochondrial protein-coding genes although, generally speaking,

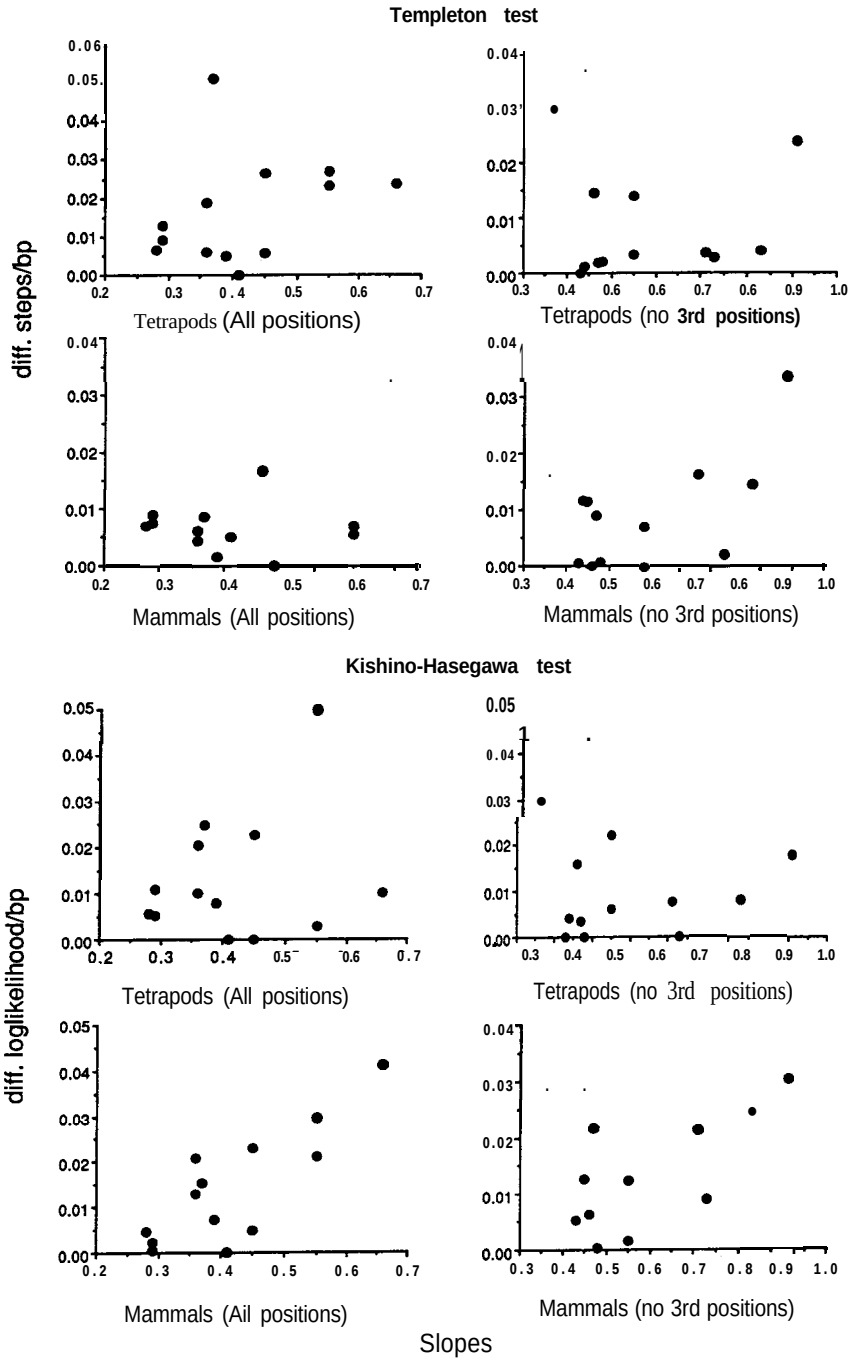

FIG. 4.—Rate of evolution as predictor of phylogenetic performance. The rate of evolution of each mitochondrial gene (represented by its associated log-log slope, see fig. 3) was plotted against its phylogenetic performance in recovering the expected tree (represented by difference in log-likelihood/bp in the Kishino-Hasegawa test and by difference in steps/bp in the Templeton test; see fig. 2).

the genes of intermediate rates might be of better phylogenetic utility for the questions at hand (figs. 2, 3, and 4). There is no strong correlation between the performance of a particular gene and its slope or intercept (figs. 2, 3, 4, table 3). The difference in performance of genes in recovering the expected vertebrate trees seems to be not strongly dependent on their rate of evolution and concomitant saturation processes, but other factors might play a role as well in predicting the phylogenetic utility of these genes for these questions (e.g., Naylor, Collins, and Brown 1995).

The number of phylogenetically informative sites in individual genes is a predictor of the closeness of the recovered to the expected topology (fig. 5). The larger the number of phylogenetically informative sites the closer the recovered phylogeny will be to the expected topology. Third codon positions for both mitochondrial and nuclear genes (Cummings, Otto, and Wakeley 1995;

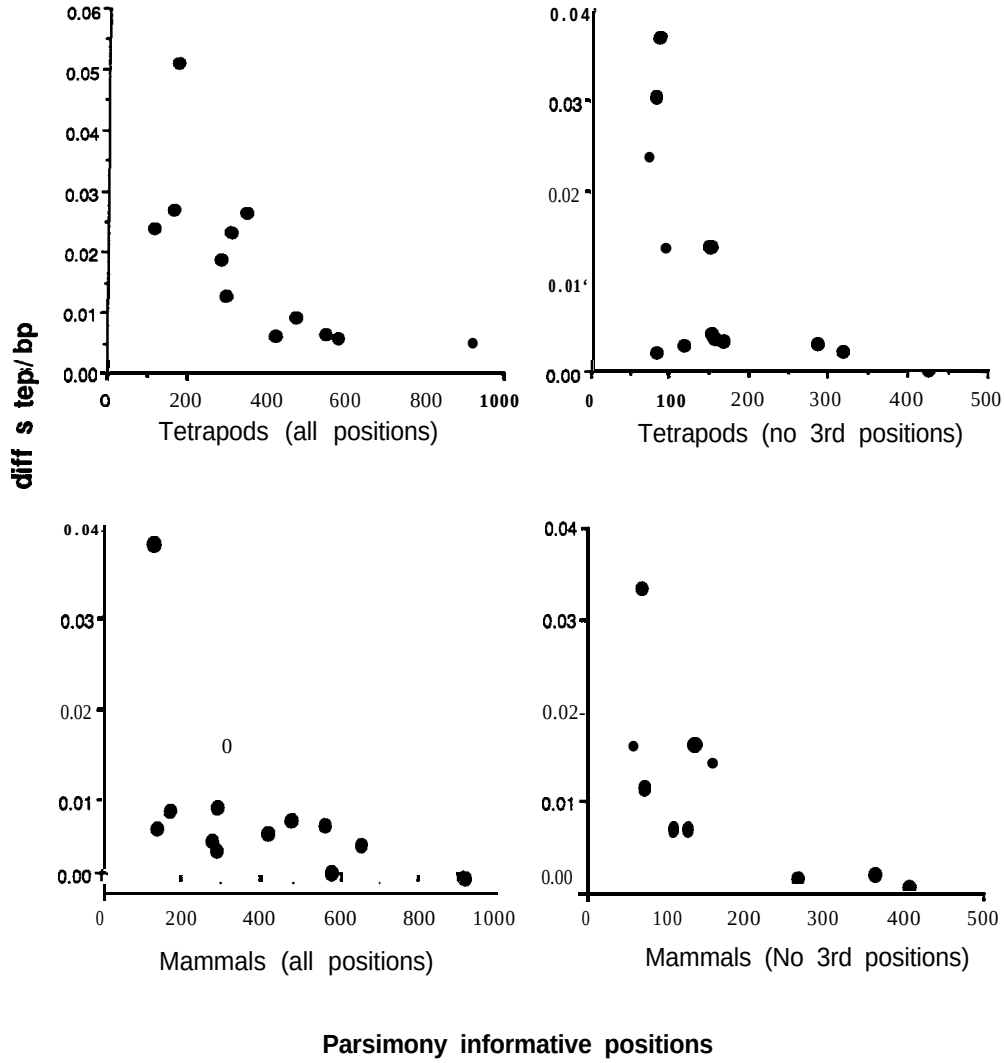

FIG. 5.—Parsimony informative sites as predictors of phylogenetic performance. The difference in number of steps between the most parsimonious tree and the expected tree (Templeton test) per base pair of each mitochondrial gene (see fig. 2) was plotted against the number of parsimony-informative sites of each gene.

Orti and Meyer 1996), despite obvious potential saturation problems (Cao et al. 1994), often contain phylogenetic information, even among distantly related species (e.g., table 1). It is therefore probably unjustified to exclude third codon positions altogether particularly in ML approaches (Cummings, Otto, and Wakeley 1995).

Recently, Cummings, Otto, and Wakeley (1995) conducted a performance analysis of mitochondrial genes. Similar to our findings, they also concluded that single genes do not always recover the expected vertebrate phylogeny, but that they might come close to the "correct" topology. In their analysis they suggested that only a large number (on the order of 2 kb) of nucleotides is able to recover the expected tree. They suggested that individual genes are potentially relatively poor samples of the entire mitochondrial genome and that nucleotides from random samples of the entire mitochondrial genome perform better than continuous nucleotide sequences. Cummings, Otto, and Wakeley (1995) pointed out that DNA sequences inevitably run the danger that neighboring nucleotides do not evolve independently and that, hence, a random sample of nucleotides might provide a better phylogenetic signal than individual genes. These potential drawbacks in terms of nonindependence of individual genes might, however, be outweighed by technical and analytical considerations that would favor a sequencing approach of continuous portions of mitochondrial genes or even entire genes. One additional advantage of sequencing entire genes is that one can thereby study the evolution of the gene itself as well. Importantly, DNA sequences permit the establishment of parameters that shaped the evolution of that gene (e.g., base composition, codon bias, functional constraints on the gene product that feed back to the DNA level, etc.) which might be incorporated into ML models. These detailed ML models will allow for more accurate phylogeny reconstruction. Specific recommendations of which gene to sequence are hence constructive; however, the generality of our recommendations for other phylogenetic questions and taxa is unclear (e.g., if

lamprey is used as outgroup in the analyses, none of the methods, even with all mitochondrial genes combined, is able to recover the expected tree; see Zardoya and Meyer 1996).

After this manuscript was submitted, Russo, Takezaki, and Nei (1996) reported a study related to ours. In their work, the relative efficiencies of different mitochondrial genes and different tree-building methods (MP, NJ, ML, and minimum evolution) in recovering a known vertebrate phylogeny were evaluated. Their results are basically congruent to ours, emphasizing the importance of using large data sets in constructing phylogenetic trees among lineages that are quite distantly related, and stressing that protein rather than nucleotide sequences recover the expected topology more reliably. Similar to our study, they reported the good performance of ND4, ND5, and cytb genes and the poor phylogenetic utility of the ND4L gene. However, in contrast to our findings, they report that NJ seems to perform as well as computationally more intensive methods such as MP or ML. In contrast, our analyses seem to indicate that NJ might incorrectly recover phylogenies when problematic (in terms of heterogeneity of rates of substitution, compositional bias, or taxonomic position) taxa such as the lungfish or the horse are included in phylogenetic analyses of vertebrate relationships.

The analyses reported here raise doubts about the current practice of most studies in molecular systematics that routinely sample relatively short mtDNA sequences to infer phylogeny among lineages of medium (80-300 MYA) or old age (>300 MYA). It might be not without problems to assume that such DNA sequences are representative of the corresponding species' or gene tree's history. However, in most instances the trees recovered by individual mitochondrial protein-coding genes are not statistically distinguishable from the expected trees. Many different considerations factor into the performance of particular genes that continue to make it difficult to predict their utility for particular evolutionary questions. In agreement with previous work (Cao et al. 1994; Cummings, Otto, and Wakeley 1995; Russo, Takezaki, and Nei 1996), our analyses indicate the need for sampling larger sequence data sets from more than one mitochondrial gene (also including nuclear genes). It is likely that the use of combinations of genes with better performance behavior, whether analyzed separately or in a combined approach under currently available methods of phylogenetic inference, will lead to increased accuracy in phylogenetic reconstruction.

Acknowledgments

Thanks to Ehab Abouheif, John Huelsenbeck, Manuel Ruedi, and two anonymous reviewers for discussion and for providing helpful suggestions on the manuscript. R.Z. was sponsored by a postdoctoral grant of the Ministerio de Educacion y Ciencia of Spain. This work received partial financial support from grants from the National Science Foundation (BSR-9107838 and BSR-9119867) and from a collaboration grant with the Max-Planck-Institut für Biology in Tübingen from the Max-Planck Society, Germany. This publication was prepared during A.M.'s tenure as Miller Visiting Research Professor and a Guggenheim Fellow at the University of California at Berkeley. The support of the John Simon Guggenheim Foundation and the Miller Institute and the hospitality of the Museum of Vertebrate Zoology and the Departments of Integrative Biology and Molecular and Cell Biology are gratefully acknowledged.

LITERATURE CITED

ADACHI, J., and M. HASEGAWA. 1992. MOLPHY: programs for molecular phylogenetics I-PROTML: maximum likelihood inference of protein phylogeny. Computer science monographs, 27, Institute of Statistical Mathematics, Tokyo.

———. 199.5. Phylogeny of whales: dependence of the inference on species sampling. Mol. Biol. Evol. **12**:177–179.

ANDERSON, S., A. T. BANKIER, B. G. BARRELL et al. (14 co-authors). 198 1. Sequence and organization of the human mitochondrial genome. Nature **290**:457–464.

ANDERSON, S., M. H. DE BRUIJN, A. R. COULSON, I. C. EPERON, F. SANGER, and I. G. YOUNG. 1982. Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian genome. J. Mol. Biol. **156**:683–717.

ARNASON, U., and A. GULLBERG. 1993. Comparison between the complete mitochondrial sequences of the blue and the fin whale, two species that can hybridize in nature. J. Mol. Evol. **37**:3 12-322.

ARNASON, U., A. GULLBERG, E. JOHNSSON, and C. LEDJE. 1993. The nucleotide sequence of the mitochondrial DNA molecule of the Grey seal, *Halichoerus grypus*, and a comparison with mitochondrial sequences of other true seals. J. Mol. Evol. **37**:323–330.

ARNASON, U., A. GULLBERG, and B. WIDEGREN. 1991. The complete nucleotide sequence of the mitochondrial DNA of the Fin whale, *Balaenopteru physalus*. J. Mol. Evol. 33: 5.56-568.

ARNASON, U., and E. JOHNSSON. 1992. The complete mitochondrial DNA sequence of the Harbor seal, *Phoca vitulina*. J. Mol. Evol. **43**:493–505.

AVISE, J. C. 1994. Molecular markers, natural history, and evolution. Chapman and Hall, New York.

BENTON, M. J. 1990. Phylogeny of the major tetrapod groups: Morphological data and divergence dates. J. Mol. Evol. 30: 409–424.

BIBB, M. J., R. A. VAN ETTEN, C. T. WRIGHT, M. W. WALBERG, and D. A. CLAYTON. 1981. Sequence and gene organization of mouse mitochondrial DNA. Cell **26**:167–180.

CAO, Y., J. ADACHI, A. JANKE, S. PÄÄBO, and M. HASEGAWA. 1994. Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. J. Mol. Evol. **39**:519–527.

CHANG, Y. S., F. L. HUANG, and T. B. Lo. 1994. The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio)* mitochondrial genome. J. Mol. Evol. **38**:138–155.

CUMMINGS, M. P., S. P. OTTO, and J. WAKELEY. 1995. Sampling properties of DNA sequence data in phylogenetic analysis. Mol. Biol. Evol. **12**:814–822.

DEGLI ESPOSTI, M., S. DE VRIES, M. CRIMI, A. GHELLI, T. PATARNELLO, and A. MEYER. 1993. Mitochondrial cytochrome *b:* evolution and structure of the protein. Biochim. Biophys. Acta. (Bioenergetics) **1143**:243–271.

DESJARDINS, I?, and R. MORAIS. 1990. Sequence and gene organization of the chicken mitochondrial genome. J. Mol. Biol. **212**:599–634.

DEVEREUX, J., I? HAEBERLI, and 0. SMITHIES. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.

EDWARDS, A. W. E, M. NEI, N. TAKEZAKI, and T. SITNIKOVA. 1995. Assessing molecular phylogenies. Science **267**:253.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

——— 1989. PHYLIP-phylogeny inference package (version'3.4.). Cladistics **5**:164–166.

GADALETA G., G. PEPE, G. DE CANDIA, C. QUAGLIARIELLO , E. SBISA, and C. SACCONE. 1989. The complete nucleotide sequence of the ***Rattus** norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J. Mol. Evol. **28**:497–516.

GRAYBEAL, A. 1994. Evaluating the phylogenetic utility of genes: a search for genes informative about deep divergences among vertebrates. Syst. Biol. 43: 174-193.

HILLIS, D. M., and J. J. BULL. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42: 182-192.

HILLIS, D. M., J. I? HUELSENBECK, and C. W. CUNNINGHAM. 1994. Application and accuracy of molecular phylogenies. Science **264**:671–677.

HILLIS, D. M., J. P. HUELSENBECK, and D. L. SWOFFORD. 1994. Hobgoblin of phylogenetics? Nature **369**:363–364.

HORAI, S., K. HAYASAKA, R. KONDO, K. TSUGANE, and N. TAKAHATA. 1995. Recent african origin of modem humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc. Natl. Acad. Sci. USA **92**:532–536.

JANKE, A., G. FELDMAIER-FUCHS, K. THOMAS, A. VON HAESELER, and S. PÄÄBO. 1994. The marsupial mitochondrial genome and the evolution of placental mammals. Genetics **137**:243–256.

KISHINO, H., and M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29: 170-179.

KOCHER, T. D., W. K. THOMAS, A. MEYER, S. V. EDWARDS, S. PÄÄBO, F. X. VILLABLANCA, and A. C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals. Proc. Natl. Acad. Sci. USA **86**:6196–6200.

KRETTEK, A., A. GULLBERG, and U. ARNASON. 1995. Sequence analysis of the complete mitochondrial DNA molecule of the hedgehog, ***Erinaceus europaeus,*** and the phylogenetic position of the Lipotyphla. J. Mol. Evol. **41**:952–957.

LYNCH, M., and P. E. JARRELL. 1993. A method for calibrating molecular clocks and its application to animal mitochondrial DNA. Genetics 135: 1197–1208.

MARTIN, M. J., F. GONZALEZ-CANDELAS, F. SOBRINO, and J. DOPAZO. 1995. A method for determining the position and size of optimal sequence regions for phylogenetic analysis. J. Mol. Evol. 41: 1128-l 138.

MEYER, A. 1993. Evolution of mitochondrial DNA in fishes. Pp. l-38 in I? W. HOCHACHKA and T. P. MOMMSEN, eds. Molecular biology frontiers, biochemistry and molecular biology of fishes. Vol. 2. Elsevier Science Publishers, Amsterdam.

———. 1994a. DNA technology and phylogeny of fish: molecular phylogenetic studies of fish. Pp. 219-249 ***in*** A. R. BEAUMONT, ed. Genetics and evolution of aquatic organisms. Chapman and Hall, London.

———. 1994b. Shortcomings of the cytochrome b gene as a molecular marker. Trends Ecol. Evol. **9**:278–280.

MEYER, A. and A. C. WILSON. 1990. Origin of tetrapods inferred from their mitochondrial DNA affiliation to lungfish. J. Mol. Evol. **31**:359–364.

MILINKOVITCH, M. C., G. ORTI, and A. MEYER. 1995. Novel phylogeny of whales revisited but not revised. Mol. Biol. Evol. **12**:518–520.

MOORE, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution **49**:7 18-726.

NAYLOR, G. J., T. M. COLLINS, and W. M. BROWN. 1995. Hydrophobicity and phylogeny. Nature **373**:555–556.

ORTI, G., and A. MEYER. 1996. Molecular evolution of ependymin and the phylogenetic resolution of early divergences among euteleostean fishes. Mol. Biol. Evol. **13**:556–573.

ROE, B. A., M. DIN-POW, R. K. WILSON, and J. F. **WONG.** 1985. The complete nucleotide sequence of the ***Xenopus** laevis* mitochondrial genome. J. Biol. Chem. **260**:9759–9774.

Russo, C. A. M., N. TAKEZAKI, and M. NEI. 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol. Biol. Evol. **13**:525–536.

RZHETSKY, A, and M. NEI. 1994. Unbiased estimates of the number of nucleotide substitutions when substitution rate varies among different sites. J. Mol. Evol. **38**:295–299.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. SCHARF, R. HIGUCHI, G. T. **HORN, K. B.** MULLIS, **and H. A. ERLICH.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–49 1.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

STEWART, C. B. 1993. The powers and pitfalls of parsimony. Nature **361**:603–607.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Illinois Natural History Survey, Champaign, Ill.

TAJIMA, F. 1991. Determination of window size for analyzing DNA sequences. J. Mol. Evol. **33**:470–473.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of human and the apes. Evolution 37: 221-244.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

TZENG, C. S., C. F. HUI, S. C. SHEN, and P. C. HUANG. 1992. The complete nucleotide sequence of the ***Crossostoma** lacustre* mitochondrial genome: conservation and variations among vertebrates. Nucleic Acids Res. **20**:4853–4858.

XU, X., and A. ARNASON. 1994. The complete mitochondrial DNA sequence of the horse, ***Equus** caballus*: extensive heteroplasmy of the control region. Gene **148**:357–362.

ZARDOYA, R., A. GARRIDO-PERTIERRA, and J. M. BAUTISTA. 1995. The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus* ***mykiss.*** J. Mol. Evol. **41**:942–951.

ZARDOYA, R., and A. MEYER. 1996. The complete nucleotide sequence of the mitochondrial genome of the lungfish (***Protopterus** dolloi*), supports its phylogenetic position as a close relative of land vertebrates. Genetics 142: 1249– 1263.

SHOZO YOKOYAMA, reviewing editor

Accepted April 19, 1996