Searching for the Closest Living Relative(s) of Tetrapods Through Evolutionary Analyses of Mitochondrial and Nuclear Data

Rafael Zardoya,[*1] *Ying Cao,*[†] *Masami Hasegawa,*[†] *and Axel Meyer*[*2]

*Department of Ecology and Evolution and Program in Genetics, State University of New York, Stony Brook; and †The Institute of Statistical Mathematics, Tokyo, Japan

The phylogenetic relationships of the African lungfish (*Protopterus dolloi*) and the coelacanth (*Latimeria chalumnae*) with respect to tetrapods were analyzed using complete mitochondrial genome DNA sequences. A lungfish + coelacanth clade was favored by maximum parsimony (although this result is dependent on which transition : transversion weights are applied), and a lungfish + tetrapod clade was supported by neighbor-joining and maximum-likelihood analyses. These two hypotheses received the strongest statistical and bootstrap support to the exclusion of the third alternative, the coelacanth + tetrapod sister group relationship. All mitochondrial protein coding genes combined favor a lungfish + tetrapod grouping. We can confidently reject the hypothesis that the coelacanth is the closest living relative of tetrapods. When the complete mitochondrial sequence data were combined with nuclear 28S rRNA gene data, a lungfish + coelacanth clade was supported by maximum parsimony and maximum likelihood, but a lungfish + tetrapod clade was favored by neighbor-joining. The seemingly conflicting results based on different data sets and phylogenetic methods were typically not statistically strongly supported based on Kishino-Hasegawa and Templeton tests, although they were often supported by strong bootstrap values. Differences in rate of evolution of the different mitochondrial genes (slowly evolving genes such as the cytochrome oxidase and tRNA genes favored a lungfish + coelacanth clade, whereas genes of relatively faster substitution rate, such as several NADH dehydrogenase genes, supported a lungfish + tetrapod grouping), as well as the rapid radiation of the lineages back in the Devonian, rather than base compositional biases among taxa seem to be directly responsible for the remaining uncertainty in accepting one of the two alternate hypotheses.

Introduction

Lobe-finned fishes (Sarcopterygii) were a highly successful group during the Devonian, about 400 MYA. At that time, up to 10 Dipnoi (lungfishes), 3 Actinistia (coelacanths), 1 Onychodontida, and about 8 Rhipidistia (Porolepiformes, Osteolepiformes, Rhizodontida, Elpistostegalia) families with hundreds of species lived in the oceans and river systems of the Gondwana supercontinent (Cloutier and Ahlberg 1996; Maisey 1996). By the end of the Devonian, circa 360 MYA, the first tetrapods diverged from the Rhipidistia and colonized land (Ahlberg, Clack, and Luksevics 1996; Cloutier and Ahlberg 1996). The diversity of all lobe-finned fishes began to decline after the mass extinction in the Permian (290 to 245 MYA), which only three groups of sarcopterygians (Tetrapoda, Dipnoi, and Actinistia) survived. Today, there are about 23,500 living species of tetrapods but only five species of lungfishes and one extant coelacanth, the last survivor of the Actinistia (Maisey 1996).

After many decades of debate, most paleontologists now agree that the rhipidistian fishes, and within them, the elpistostegids (also known as panderichthyids) are the extinct relatives of tetrapods (Vorobyeva and Schultze 1991; Ahlberg, Clack, and Luksevics 1996; Cloutier and Ahlberg 1996) (fig. 1). However, there is neither such consensus regarding the relationships of Actinistia, Dipnoi, and Rhipidistia nor agreement on the relationships among basal Rhipidistia (fig. 1). Most recent paleontological analyses support a Dipnoi + Rhipidistia sister group relationship (Panchen and Smithson 1987; Ahlberg 1991; Forey, Gardiner, and Patterson 1991; Long 1995). However, there is also paleontological opinion for a sister group relationship between the Actinistia and Rhipidistia (Schultze 1987; Long 1989) and support for a sister group relationship between the Actinistia and Dipnoi (e.g., Chang 1991; Schultze 1994). The discrepancy among these hypotheses largely stems from differences in character definition and character coding between researchers (summarized by Cloutier and Ahlberg 1996). This debate is likely to continue unless new relevant fossils are discovered and agreement among paleontologists about some characters is achieved.

A perspective based on molecular rather than on phenotypic data from lungfishes, the coelacanth, and tetrapods, the only living representatives of sarcopterygians, can aid in this regard. Both mitochondrial (Meyer and Wilson 1990; Meyer and Dolven 1992; Hedges, Hass, and Maxson 1993; Yokobori et al. 1994) and nuclear (Zardoya and Meyer 1996*b*) nucleotide sequences have been collected with the specific goal of resolving the relationships among living sarcopterygians (reviewed in Meyer 1995; Zardoya and Meyer 1997*b*), but the available molecular data have not provided complete resolution of this controversy. The phylogenetic analysis of nuclear 28S rRNA and mitochondrial COI gene sequences favored a lungfish + coelacanth relationship but could not reject the hypothesis of a lungfish + tetrapod clade (Yokobori et al. 1994; Zardoya and Meyer 1996*b*). Mitochondrial rRNA and cytochrome *b* data supported the hypothesis of lungfish as the closest living relatives

[1] Present address: Museo Nacional de Ciencias Naturales, Madrid, Spain.

[2] Present address: Department of Biology, University of Konstanz, Konstanz, Germany.

Key words: mtDNA, 28S rRNA, sarcopterygians, coelacanth, lungfish, tetrapods.

Address for correspondence and reprints: Rafael Zardoya, Museo Nacional de Ciencias Naturales, José Gutierrez Abascal, 2, 28006 Madrid, Spain. E-mail: mcnr154@pinar2.csic.es.

Mol. Biol. Evol. 15(5):506–517. 1998

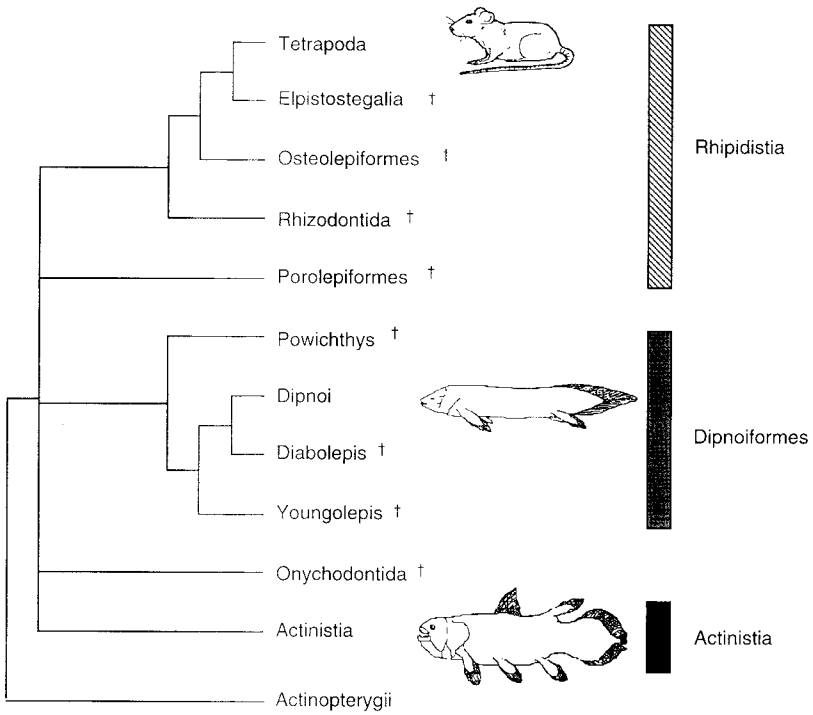

FIG. 1.—Consensus tree summarizing the phylogenetic relationships among major sarcopterygian lineages based on paleontological data (e.g., Cloutier and Ahlberg 1996). Relationships among advanced rhipidistia, porolepiformes, dipnoiformes, actinistia, and onychodontia are still not resolved and are therefore depicted as a polytomy. † = extinct taxa.

of tetrapods (Meyer and Wilson 1990; Meyer and Dolven 1992) without being able to statistically reject the other two competing hypotheses (Zardoya and Meyer 1996*b*).

It is likely that many of the phylogenetically informative sites that could be used to trace the origins of the lineages leading to the three living groups of sarcopterygians in the late Devonian have been obliterated by the accumulation of subsequent changes in these short-sequence data sets. The phylogenetic noise that accumulated during the last 360 MYA or so is complicating the successful recovery of the true relationships among these three taxa. In order to obtain better resolution of the question at hand, larger data sets that maximize the number of phylogenetically informative sites need to be collected and analyzed (Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*). With this aim, we recently completed the sequencing of the mitochondrial genomes of the African lungfish (*Protopterus dolloi*) (Zardoya and Meyer 1996*a*) and the coelacanth (*Latimeria chalumnae*) (Zardoya and Meyer 1997*a*), and here, we provide a summary of detailed phylogenetic analyses of the relationships among the living sarcopterygian lineages based on these two new complete mitochondrial DNA data sets.

Materials and Methods

Mitochondrial Genomes

The data sets analyzed in this study comprise the following 12 complete vertebrate mitochondrial DNA genomes: lamprey, *Petromyzon marinus* (U11880; Lee and Kocher 1995); bichir, *Polypterus ornatipinnis* (U62532; Noack, Zardoya, and Meyer 1996); rainbow trout, *Oncorhynchus mykiss* (L29771; 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*a*); coelacanth, *Latimeria chalumnae* (U82228; Zardoya and Meyer 1997*a*); clawed frog, *Xenopus laevis* (M10217; Roe et al. 1985); chicken, *Gallus gallus* (X52392; Desjardins and Morais 1990); opossum, *Didelphis virginiana* (Z29573; Janke et al. 1994); blue whale, *Balaenoptera musculus* (X72204; Arnason and Gullberg 1993); human, *Homo sapiens* (D38112; Horai et al. 1995).

Phylogenetic Analyses

DNA sequences were aligned using CLUSTAL W (Thompson, Higgins, and Gibson 1994) followed by refinement by eye based on the corresponding deduced amino acid sequences and rRNA and tRNA secondary structures. Gaps resulting from the alignment were treated as missing data. Ambiguous alignments were excluded from the phylogenetic analyses (aligned sequences and exclusion sets are available from the authors on request).

Outgroup selection was performed using RASA based on a statistical criterion that estimates ingroup pleisomorphy (Lyons-Weiler, Hoelzer, and Tausch 1998). Initially, four distinct mitochondrial DNA data sets were analyzed: (1) all 13 protein coding genes combined, excluding third codon positions; (2) 12S and 16S rRNA genes combined; (3) all 22 tRNA gene sequences

combined; and (4) protein-coding (excluding third codon positions), rRNA, and tRNA genes combined. Each of these DNA data sets was subjected to the maximum-parsimony (MP) method (PAUP* version d54; Swofford 1997) using heuristic searches (TBR branch swapping; MULPARS option in effect) with 10 random stepwise additions of taxa to find the most parsimonious trees. Transitions and transversions were always given equal weight (for alternative weighting schemes, see below). Neighbor-joining (NJ) (Saitou and Nei 1987) analyses (based on Kimura two-parameter distance matrices) of the sequences were performed with PHYLIP (version 3.55) (Felsenstein 1989) and PAUP* version d54 (Swofford 1997). To account for the variation of substitution rates among sites (Yang and Kumar 1996), the α shape parameter of the gamma distribution of rate variation was estimated based on the MP tree by the method of Yang and Kumar (1996) with PAUP* version d54 (Swofford 1997). Maximum-likelihood (ML) analyses were performed with PAUP* version d54 (Felsenstein's 1984 model), and MOLPHY version 2.3 (Adachi and Hasegawa 1996*a*) (HKY 85 model; Hasegawa, Kishino, and Yano 1985). In both cases, transition/transversion ratios were optimized to maximize the likelihood, and empirical base frequencies were used.

Additionally, the protein-coding-gene data set was analyzed at the amino acid level with all three phylogenetic methods using PAUP* version d54 (Swofford 1997), PHYLIP version 3.55 (Felsenstein 1989), and MOLPHY version 2.3 (Adachi and Hasegawa 1996*a*). In the protein ML analyses, an NJ tree was inferred as the starting tree for a local rearrangement search for the ML tree with the JTT and mtREV models (which better approximate the evolution of the individual proteins encoded by the mitochondrial DNA; Adachi and Hasegawa 1996*b*). Finally, all mitochondrial data sets were also combined with a nuclear 28S rRNA gene data set (Zardoya and Meyer 1996*b*) and analyzed with MP, NJ, and ML.

Robustness of the phylogenetic results was tested by bootstrap analyses (Felsenstein 1985) (as implemented in PAUP* version d54 and PHYLIP version 3.55 with 100 pseudoreplications each) and the RELL (resampling of the estimated log-likelihood) method (Kishino, Miyata, and Hasegawa 1990) (as implemented in MOLPHY version 2.3 with 10,000 pseudoreplications).

Statistical Methods

Statistical confidence of the resulting best trees of each ML analysis was evaluated by calculating the standard error of the difference in log-likelihood between the resulting best tree and the competing hypotheses using the formula of Kishino and Hasegawa (1989) as implemented in the MOLPHY version 2.3 program (Adachi and Hasegawa 1996*a*) and PAUP* version d54 (Swofford 1997). Similarly, for MP analyses, statistical confidence in the results was assessed with a two-tailed Wilcoxon matched-pairs signed-ranks test (Siegel 1956) by calculating the standard deviation of the difference in number of steps between the resulting most parsimonious tree and the alternative trees using the method of Templeton (1983) as implemented in PHYLIP version 3.55 (Felsenstein 1989). If the difference in log-likelihoods or in number of steps between two competing phylogenetic hypotheses was more than 1.96 times the standard deviation, then the two phylogenies were declared significantly different ($P < 0.05$) (Felsenstein 1989).

Results and Discussion

Performance of Lamprey and Bichir as Outgroup Taxa

According to several lines of morphological evidence, lamprey and bichir represent the most basal vertebrates from which complete mitochondrial genome sequences have been obtained so far. Therefore, trees were initially rooted using these two taxa as outgroups. However, none of the analyses using lamprey or bichir as outgroup were able to recover some of the well-established relationships among vertebrates (Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*a*; but see Noack, Zardoya, and Meyer 1996) (figs. 2 and 3). The phylogenetic relationships within amniota and teleosts were consistently recovered regardless of the phylogenetic method of inference utilized when these two taxa were used as outgroups (figs. 2 and 3). However, the relative phylogenetic positions of the frog, the coelacanth, the lungfish, and the bichir varied depending on the method of phylogenetic inference and outgroup taxon utilized (figs. 2 and 3). As expected, the conflicting phylogenetic positions assigned to these problematic taxa were accompanied by only low bootstrap values, indicating weak statistical support of such nodes (figs. 2 and 3). One likely explanation for these alternative topologies is that many substitutions have accumulated along the branch connecting lamprey and bichir to the ingroup taxa, leading to outgroup–ingroup attractions that particularly affect those ingroup taxa that have longer branches (Felsenstein 1978; Swofford et al. 1996), i.e., the frog, the coelacanth, the lungfish, and the bichir.

Recently, it has been indicated that mitochondrial hydrophobic amino acids contain relatively high levels of homoplasy due to compositional constraints in second codon positions (Naylor, Collins, and Brown 1995; Naylor and Brown 1997). However, the performance of lamprey and bichir as outgroup taxa was not improved when the mitochondrial protein-coding-gene data set was analyzed at the amino acid level excluding potentially homoplasious amino acids (G, A, V, L, S, K, Y, I, F, W, D, E, R, H), regardless of the phylogenetic method of inference utilized.

The unsatisfactory performance of lamprey and bichir as outgroup taxa was confirmed by a relative apparent synapomorphy analysis (Lyons-Weiler, Hoelzer, and Tausch 1998). This analysis determines which of a set of candidate outgroups maximizes character covariation in the ingroup and, therefore, provides the best possible estimate of plesiomorphy for the ingroup, estimated by a test statistic (tRASA) (Lyons-Weiler, Hoelzer, and Tausch 1998). The best outgroup taxon is the one that results in a higher tRASA value after all potential out-

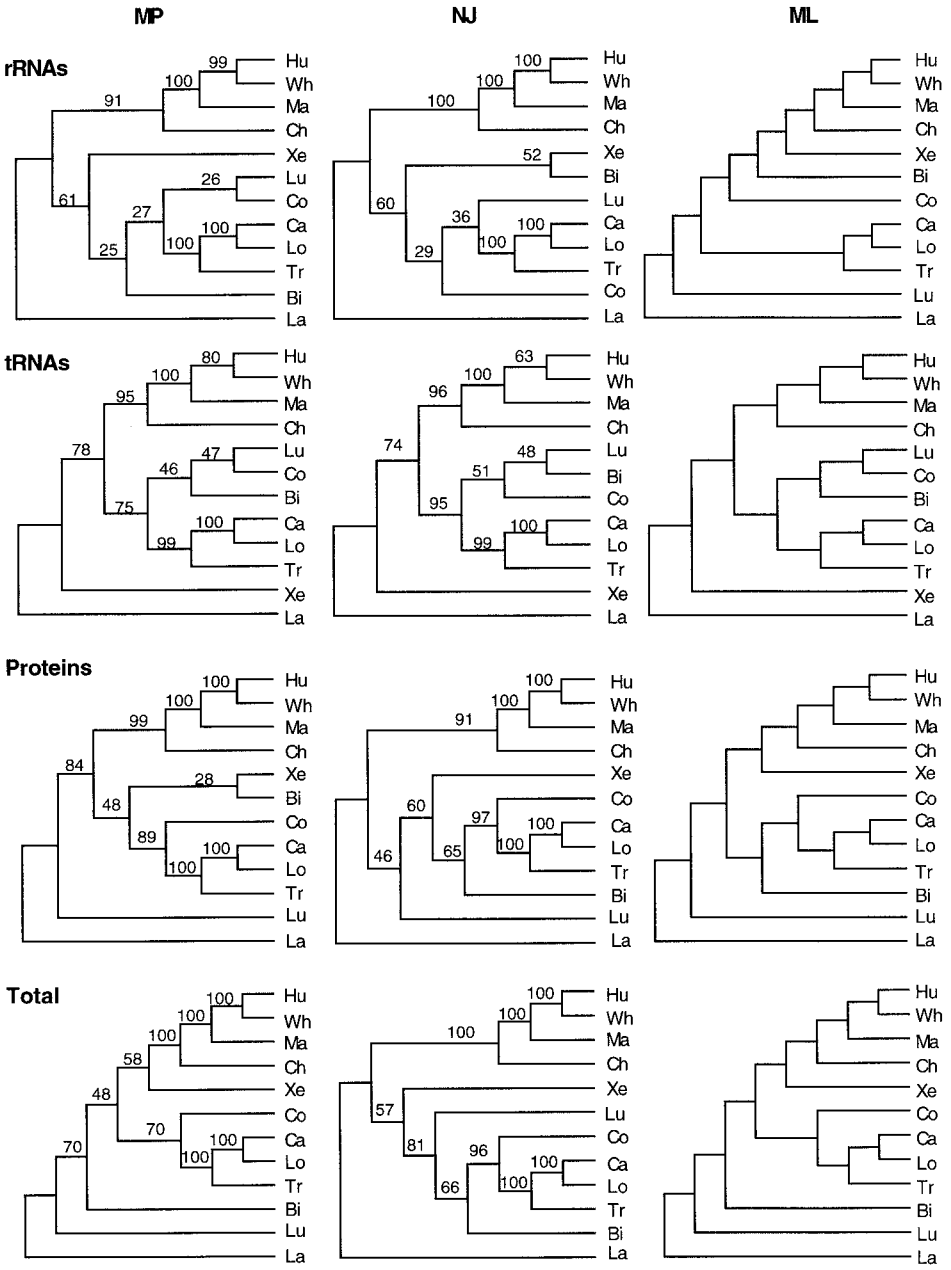

FIG. 2.—Phylogenetic performance of the lamprey (*Petromyzon marinus*) mitochondrial genome (Lee and Kocher 1995) as outgroup. Four mitochondrial DNA data sets were analyzed (all 13 protein-coding genes combined, excluding third codon positions; 12S and 16S rRNA genes combined; all 22 tRNA genes combined; the protein-coding [excluding third codon positions], rRNA, and tRNA data sets combined) with all three commonly used methods of phylogenetic inference (MP, NJ, ML) using the lamprey as outgroup taxon. Unorthodox vertebrate phylogenetic relationships (see Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*a*) are obtained regardless of the phylogenetic method used and the mitochondrial data set analyzed. Hu, human; Wh, whale; Ma, opossum; Ch, chicken; Xe, frog; Lu, lungfish; Co, coelacanth; Ca, carp; Lo, loach; Tr, trout; Bi, bichir; La, lamprey.

groups have been tested. According to the relative apparent synapomorphy analysis, the trout (tRASA = 18.70) represents a better outgroup for the data set combining rRNA, tRNA, and protein-coding gene sequences than bichir (tRASA = 9.81) or lamprey (tRASA = 6.60).

Phylogenetic Analysis of the Entire Mitochondrial DNA Data Set

The vertebrate mitochondrial DNA data set which combines all protein-coding (excluding third codon positions), rRNA, and tRNA gene sequences (16,140 characters) was analyzed with MP, NJ, and ML using teleosts (trout, carp, and loach) as outgroup taxa and excluded bichir and lamprey.

Parsimony Analyses

A single most parsimonious tree supporting a lungfish + coelacanth clade, 8,468 steps long (consistency index [CI] = 0.65) was obtained when transitions in first codon positions of the protein-coding genes were excluded from the analysis (fig. 4*C*). The lungfish + coe-

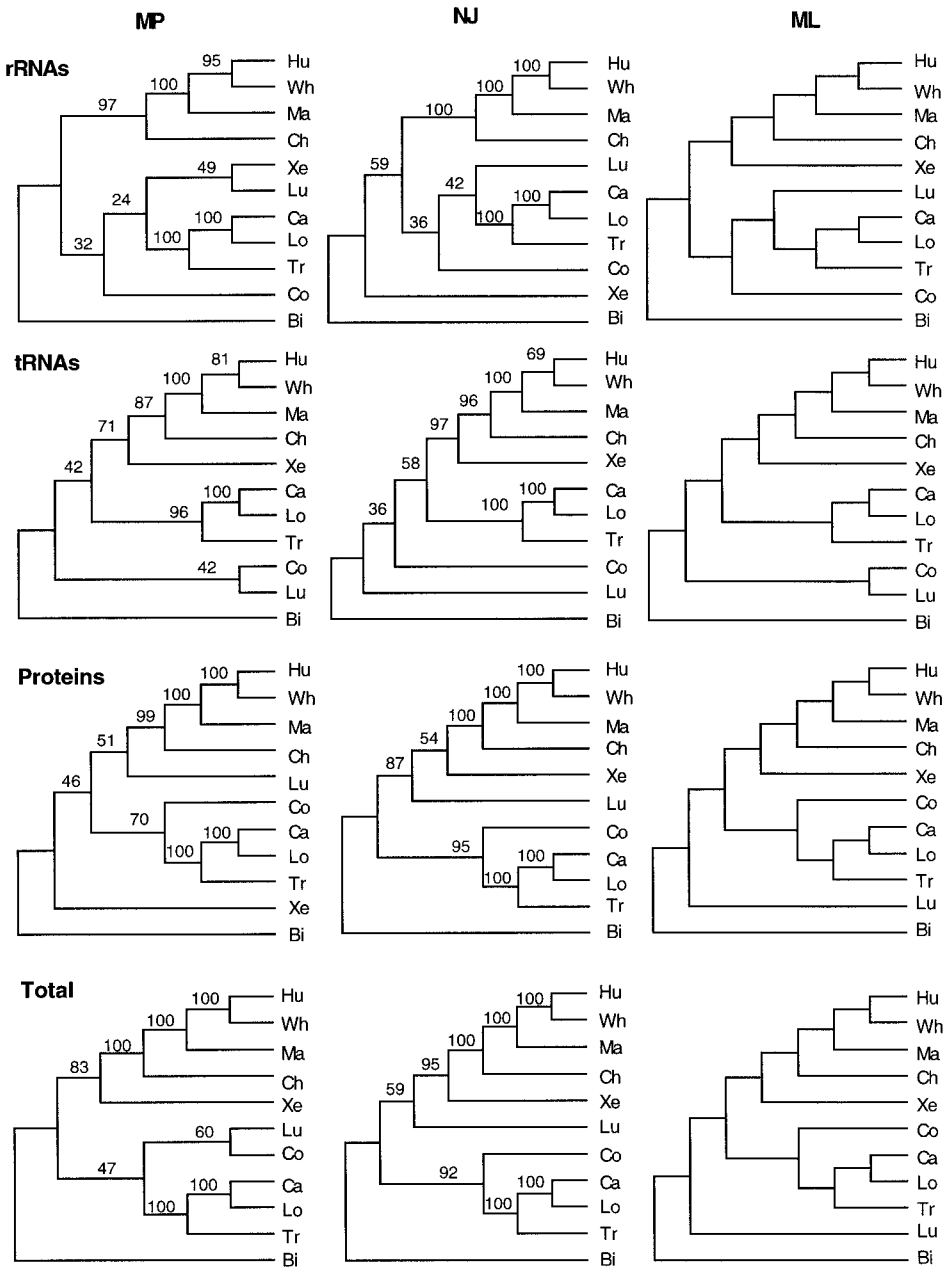

FIG. 3.—Phylogenetic performance of the bichir (*Polypterus ornatipinnis*) mitochondrial genome (Noack, Zardoya, and Meyer 1996) as outgroup. As in the case of lamprey (see fig. 2), well-established vertebrate phylogenetic relationships (see Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*a*) cannot be recovered regardless of the phylogenetic method (MP, NJ, ML) used and the mitochondrial data set (protein-coding, rRNA, tRNA, total) analyzed. Hu, human; Wh, whale; Ma, opossum; Ch, chicken; Xe, frog; Lu, lungfish; Co, coelacanth; Ca, carp; Lo, loach; Tr, trout; Bi, bichir; La, lamprey.

lacanth node was supported by a 56% bootstrap value. However, if transitions in first codon positions of protein-coding genes were included in the analysis or a transition : transversion weight of 1:2 was adopted for the whole data set, a lungfish + tetrapod clade was favored (fig. 4*A*). In these cases, the lungfish + tetrapod clade had bootstrap support of 51% and 68%, respectively. If third codon positions of the protein-coding genes were not excluded from the analysis, one single most parsimonious tree with a lungfish + frog clade was recovered. This result is likely due to the addition of a considerable amount of noise when third codon positions are included in the analysis (Swofford et al. 1996), which results in long-branch attraction (Felsenstein 1978) and particularly affects MP (Russo, Takezaki, and Nei 1996).

Neighbor-Joining and Maximum-Likelihood Analyses

Interestingly, NJ (Kimura two-parameter distances corrected for gamma-distributed rates across sites; $\alpha = 0.53$) and ML (empirical base frequencies: A = 0.26, C = 0.25, G = 0.20, T = 0.29; transitions/transversions [ti/tv] = 1.45; F84 model: −Ln likelihood = 61,722.5; HKY85 model: −Ln likelihood = 61,702.9) also arrived

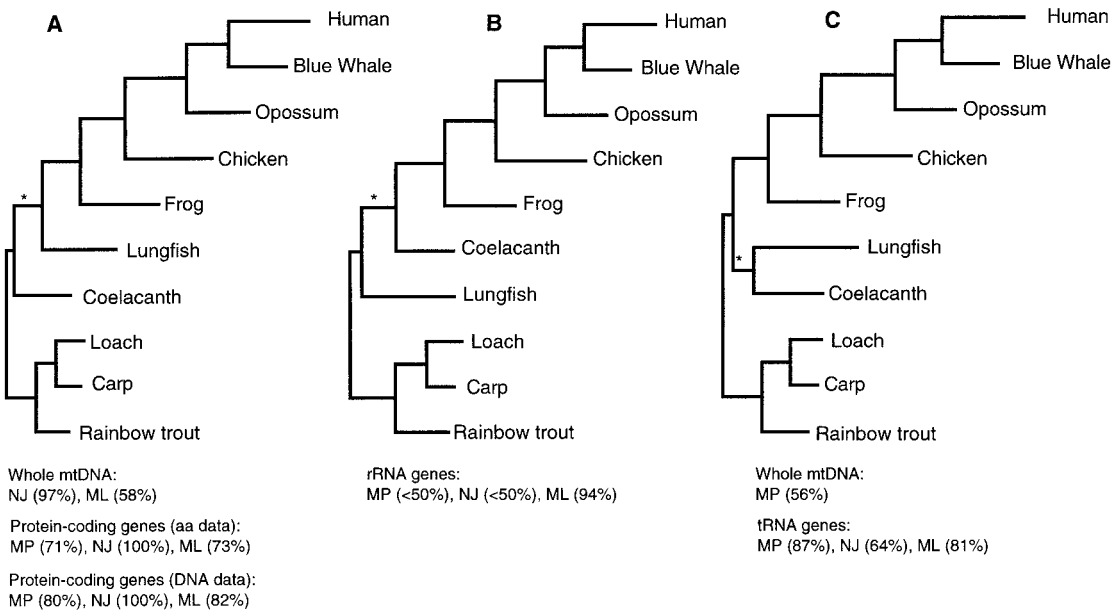

FIG. 4.—Mitochondrial DNA support for the three alternative hypotheses on the phylogenetic relationships of the living sarcopterygian lineages. *A,* Lungfish as the closest living sister group to tetrapods is supported by the whole mitochondrial data set when analyzed with MP (ti/tv = 1; transitions in first codon positions are included, and third codon positions are excluded), NJ, and ML, and by the protein-coding gene data set (both at the amino acid and at the DNA level) when analyzed with MP, NJ, and ML. *B,* Coelacanth as closest living relative of tetrapods is favored by the mitochondrial rRNA gene data set with all three commonly used phylogenetic methods of inference. *C,* A lungfish + coelacanth clade is supported by the whole mitochondrial data set when analyzed with MP (ti/tv = 1; no third codon positions; no transitions in first codon positions), and by the mitochondrial tRNA data set regardless of the phylogenetic method of inference. Bootstrap values supporting the nodes indicated by an asterisk are given for each of the phylogenetic analyses.

at a topology in which the lungfish is placed as the closest living relative of tetrapods (fig. 4*A*). In the NJ analysis, the lungfish + tetrapod clade was supported by a 97% bootstrap value, whereas in the ML analysis, the same node has 58% bootstrap support.

Because of the apparent discrepancy in the resulting topologies between the different methods of phylogenetic inference, we conducted Kishino-Hasegawa (Kishino and Hasegawa 1989) and Templeton (1983) tests (table 1). According to these tests, none of the three competing hypotheses about the relationships of sarcopterygians could be statistically rejected with the phylogenetic information contained in the whole mitochondrial DNA genome data sets.

Recent studies (Kumazawa and Nishida 1996; Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*) have demonstrated that the entire set of mitochondrial genes was able to recover a well-established vertebrate phylogeny with strong bootstrap support regardless of the phylogenetic method of inference utilized. In our case, all interior branches with the exception of those leading to lungfish, coelacanth, and frog had bootstrap values close to 100%, regardless of the method used and in support of previous findings (Kumazawa and Nishida 1996; Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*). However, lungfish and coelacanth nodes received only moderate bootstrap support with MP (56%, 51%, and 58% depending on the weighting scheme) and ML (58%) and high bootstrap support with NJ (97%) (fig. 4). The moderate bootstrap values and the discrepancy between methods suggest that mitochondrial sequences may not able to resolve deep-branch phylogenetic relationships of species that have diverged within a short period of time.

Separate Phylogenetic Analysis of the Protein-Coding, rRNA, and tRNA Mitochondrial DNA Data Sets

To better understand the performance of the mitochondrial data set in discerning between competing hypotheses, the protein-coding, rRNA, and tRNA data sets were analyzed separately. The analyses of the protein-coding-gene DNA sequences (11,736 characters) with all three commonly used methods of phylogenetic inference (MP, NJ, and ML) favored lungfish as the closest living relatives of tetrapods (fig. 4*A*). The lungfish + tetrapod node was supported by a 80% bootstrap value in the MP analysis when transitions in first codon positions were excluded from the analysis (4,210 steps; CI = 0.67). The bootstrap support for the same node in the NJ analysis (gamma distribution shape parameter α = 0.41) was 100%. In the ML analysis (empirical base frequencies: A = 0.24, C = 0.26, G = 0.18, T = 0.32; ti/tv = 1.17; F84 model: −Ln likelihood = 39,126.9; HKY85 model: −Ln likelihood = 39,110.4), the bootstrap value for the lungfish + tetrapod node was 82%. According to the Kishino-Hasegawa and Templeton tests, the protein-coding-gene data set, although overall strongly in support of a lungfish + tetrapod clade, cannot statistically reject a lungfish + coelacanth clade (table 1). However, the coelacanth closest relationship to tetrapods is clearly rejected with this data set (table 1).

Table 1
Phylogenetic Relationships Among Living Sarcopterygians

	Kishino-Hasegawa Test		Templeton Test	
	Ln L	Δ ± SE	Steps	Δ ± SE
Whole mtDNA				
Lungfish + tetrapods	−61,702.91	—	10,812	—
Coelacanth + tetrapods	−61,722.32	−19.5 ± 22	10,832	20 ± 15.43
Lungfish + coelacanth	−61,713.81	−10.9 ± 22.3	10,813	1 ± 16.03
Protein-coding genes (DNA data)				
Lungfish + tetrapods	−38,220.01	—	6,549	—
Coelacanth + tetrapods	−39,265.61	−45.6 ± 20.9	6,584	35 ± 11.6
Lungfish + coelacanth	−38,240.31	−20.3 ± 16.5	6,565	16 ± 12.4
Protein-coding genes (aa data)				
Lungfish + tetrapods	−34,029.41	—	5,920	—
Coelacanth + tetrapods	−34,067.71	−37.7 ± 19.4	5,959	37 ± 10.8
Lungfish + coelacanth	−34,036.41	−7 ± 14.4	5,920	—
rRNA genes				
Lungfish + tetrapods	−13,485.41	−16.5 ± 10	2,520	10 ± 8.13
Coelacanth + tetrapods	−13,468.91	—	2,510	—
Lungfish + coelacanth	−13,486.61	−17.7 ± 9.8	2,524	14 ± 7.88
tRNA genes				
Lungfish + tetrapods	−9,361.01	−13.4 ± 7.8	1,743	19 ± 7.14
Coelacanth + tetrapods	−9,355.91	−8.3 ± 9.1	1,738	14 ± 7.48
Lungfish + coelacanth	−9,347.61	—	1,724	—

Note.—The differences in log-likelihoods (Ln L) of alternative trees from that of the ML tree (Δ) are shown with their SEs (following ±), which were estimated by the formula of Kishino and Hasegawa (1989). The differences in number of steps from that of the MP tree (Δ) are shown with their SEs (±), which were estimated by the Wilcoxon rank test (Templeton 1983). Third codon positions of protein-coding genes were excluded from the analyses at the DNA level. Two phylogenies were declared significantly different when the difference in log-likelihoods or numbers of steps was more than 1.96 times the SE.

The analysis was also performed at the inferred-amino-acid level (3,887 characters). Again, MP arrived at a single most parsimonious tree 5,093 steps long (CI = 0.79) that supported a lungfish + tetrapod grouping (71% bootstrap value) (fig. 4*A*). NJ (mean character distance; 100% bootstrap value for the lungfish + tetrapod node) and ML (JTT, −Ln likelihood = 36,753.63; mtREV, −Ln likelihood = 34,029.4) analyses also arrived at this topology (73% bootstrap value for the lungfish + tetrapod node). However, again and despite the quite strong bootstrap support, the mitochondrial protein amino acid data set could not statistically reject the alternative hypotheses based on the Kishino-Hasegawa test. According to the Templeton test of the alternative hypotheses, a coelacanth + tetrapod clade could be ruled out based on this data set, but not a lungfish + coelacanth clade.

Following Naylor, Collins, and Brown (1995) and Naylor and Brown (1997), to improve the phylogenetic inference, hydrophobic residues (all except M, N, T, Q, P, C) were excluded from the analyses at the amino acid level because of their relatively high levels of homoplasy. One single most parsimonious tree (576 steps; CI = 0.83) that supported a lungfish + tetrapod grouping was recovered. However, this grouping was not strongly supported (<50% bootstrap value), suggesting that not only

Table 2
Statistical Support of Each of the 13 Mitochondrial Protein Genes for the Three Competing Hypotheses on the Relationships Among Living Sarcopterygians Lineages

	ATPase6	ATPase8	COI	COII	COIII	Cytochrome *b*
Tree (aa level)						
Lungfish + tetrapods	−5.4 ± 4.1	−0.1 ± 0.4	−5.8 ± 6.9	−3 ± 2.9	−2.4 ± 2.9	ML (3,414.2)
Coelacanth + tetrapods	−4.8 ± 4.5	−0.1 ± 0.6	−3.3 ± 7.8	−0.1 ± 4.4	−1.5 ± 3.6	−9.8 ± 7.6
Lungfish + coelacanth	ML (1,770.9)	ML (451.6)	ML (2,934.8)	ML (1,907.4)	ML (1,920.4)	−12 ± 6.6
Length	181	34	502	214	258	376
Tree (DNA level)						
Lungfish + tetrapods	−4.3 ± 4.4	−1.1 ± 1.5	−7.6 ± 9.1	−4.5 ± 4.7	−8.8 ± 5.5	ML (3,823.8)
Coelacanth + tetrapods	−4.7 ± 4.2	ML (516.4)	−9.1 ± 8.6	ML (2,105.3)	−7.8 ± 6.0	−11.5 ± 10.2
Lungfish + coelacanth	ML (2,705.4)	−1.1 ± 1.5	ML (3,462.2)	−1.2 ± 5.9	ML (2,186.0)	−18.9 ± 8.0
Length	362	68	1,004	428	516	752

Note.—The −log-likelihoods of the maximum-likelihood (ML) trees are given in parentheses, and the differences in log-likelihoods of alternative trees from that of the ML tree are shown with their SEs (following ±), which were estimated by the formula of Kishino and Hasegawa (1989). Third codon positions of protein-coding genes were excluded from the analyses at the DNA level. Two phylogenies were declared significantly different when the difference in log-likelihoods was more than 1.96 times the SE.

noise but also considerable phylogenetic information was lost when hydrophobic amino acids were omitted in the analysis. Similar results (i.e., lower bootstrap values when hydrophobic residues were excluded) were obtained for the NJ and ML analyses.

Interestingly, a coelacanth + tetrapod clade (fig. 4*B*) was favored when the rRNA gene sequences (2,101 characters) were analyzed with MP (2,510 steps; CI = 0.65), NJ (gamma distribution shape parameter α = 0.73), and ML (empirical base frequencies: A = 0.33, C = 0.24, G = 0.21, T = 0.22; ti/tv = 1.55; F84 model: −Ln likelihood = 13,448.3; HKY model: −Ln likelihood = 13,468.9). However, the coelacanth + tetrapod node had a bootstrap value below 50% in both the MP and the NJ analyses. None of the two alternative hypotheses could be rejected when Kishino-Hasegawa and Templeton tests were conducted (table 1). The lack of resolution of the rRNA data set is likely due to the relatively high mutation rate of the two rRNA genes, which is not appropriate for the question at hand, and the alignment problems associated with several regions of these genes, which considerably reduce the number of phylogenetically informative sites included in the analyses. It is also interesting to note that our analysis of the rRNA data set seemingly differs from a similar one reported by Hedges, Hass, and Maxson (1993). In that analysis, a lungfish + tetrapod clade was clearly supported by the rRNA data set with a 91% bootstrap value in the NJ analysis (the bootstrap value for this node for the MP analysis was not reported). Our alignment and exclusion set was essentially the same as that utilized by Hedges, Hass, and Maxson (1993) (kindly provided by S. B. Hedges). Therefore, the different outcomes are due to the different taxa utilized, as well as the fact that Hedges, Hass, and Maxson's (1993) analyses were performed with a scrambled coelacanth/alligator 16S rRNA gene sequence (see Zardoya and Meyer 1997*a*).

Finally, the tRNA data set (1,604 characters) supported a lungfish + coelacanth clade (fig. 4*C*) when analyzed with MP (1,724 steps; CI = 0.62), NJ (gamma distribution shape parameter α = 1.26), and ML (empirical base frequencies: A = 0.29, C = 0.20, G = 0.23, T = 0.28; ti/tv = 3.05; F84 model: −Ln likelihood = 8,674.5; HKY85 model: −Ln likelihood = 8,677.0). The lungfish + coelacanth node was supported by 87%, 64%, and 81% bootstrap values in the MP, NJ, and ML analyses, respectively. Mitochondrial tRNA genes were found to be particularly appropriate for reconstructing deep-branch phylogenetic relationships which exceed divergence times of 100 Myr (Kumazawa and Nishida 1993, 1996). In such cases, combined tRNA gene sequences showed better performance than protein-coding gene sequences in recovering a well-established vertebrate phylogeny with MP and NJ (Kumazawa and Nishida 1996). However, according to the Kishino-Hasegawa and Templeton tests, neither of the two alternative hypotheses could be rejected with this data set (table 1).

Base compositional differences among species could have been responsible for the ambiguity shown among analyses. However, no significant bias was shown for the three data sets (see table 3 in Zardoya and Meyer 1997*a*). Nevertheless, an NJ analysis using the logDet distance (Lockhart et al. 1994), which is robust to base compositional bias among taxa, was also performed. A lungfish + tetrapod node was supported with a 98% bootstrap value when all data sets were combined, and with 100% and 53% bootstrap values when the protein-coding data set and the rRNA data set were analyzed, respectively. However, a lungfish + coelacanth clade was favored with a 68% bootstrap value when the tRNA data set was analyzed. Therefore, in general, the logDet transformation results agreed with those obtained without considering species-related compositional bias.

Maximum-Likelihood Analyses of Individual Mitochondrial Protein-Coding Genes

The support of individual mitochondrial protein-coding genes for each of the competing hypotheses on the phylogenetic relationships among the coelacanth, lungfishes, and tetrapods (which does not necessarily imply the most likely tree favored by each gene) was estimated with the ML method at both the amino acid and the DNA levels (table 2). According to the results

Table 2
Extended

ND1	ND2	ND3	ND4	ND4L	ND5	ND6
ML (3,047.6)	ML (4,235.1)	ML (1,071.2)	ML (5,012.3)	ML (1,268.5)	−7.7 ± 8.9	−4 ± 3.7
−15 ± 7.7	−0.6 ± 4.7	−4.5 ± 4.6	−7.2 ± 5.7	−1.3 ± 2.0	−14 ± 7.2	−4.1 ± 3.5
−13 ± 8.2	−2.2 ± 3.9	−1.6 ± 6.0	−5.1 ± 6.3	−1.3 ± 2.0	ML (5,440.3)	ML (1,526.6)
311	313	103	433	96	490	118
ML (3,519.7)	ML (4,774.2)	ML (1,306.1)	ML (5,595.9)	ML (1,390.5)	−4.6 ± 8.7	ML (1,582.8)
−14 ± 8.7	−6.5 ± 5.2	−1.8 ± 2.4	−5.4 ± 6.0	−3.5 ± 3.0	−11 ± 7.0	−1.2 ± 2.3
−15 ± 8.4	−4.8 ± 5.8	−0.1 ± 3.9	−6.4 ± 5.6	−3.4 ± 3.2	ML (5,850.6)	−0.3 ± 2.9
622	626	206	866	192	980	236

obtained, at the amino acid level, the ATPase and cytochrome oxidase mitochondrial genes preferentially supported a coelacanth + lungfish clade, whereas the NADH dehydrogenase and the cytochrome *b* mitochondrial genes mainly favored a lungfish + tetrapod relationship (with the exceptions of ND5 and ND6, which support a coelacanth + lungfish clade). At the DNA level, ATP6, COII, COIII, and ND5 favored a lungfish + coelacanth clade; ND1, ND2, ND3, ND4, ND4L, ND6, and cytochrome *b* supported a lungfish + tetrapod clade, whereas ATP8 and COI supported a coelacanth + tetrapod clade (table 2). In all cases, none of the alternative hypotheses could be statistically ruled out (with the exception of cytochrome *b* at the amino acid level, which clearly rejects a lungfish + coelacanth clade). Previous studies (Cao et al. 1994; Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*) showed that ND5, ND4, COI, and cytochrome *b* genes are the most appropriate for reconstructing reliable trees. In contrast, ATPase8 and ND4L were the least accurate. Moreover, nucleotide sequences were found to be less appropriate than amino acid sequences (Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*). In our case, at the amino acid level, ND5 and COI support a lungfish + coelacanth clade, whereas ND4 and cytochrome *b* form lungfish + tetrapod grouping. Again, two different phylogenetic signals are found even in the ''most reliable'' genes (according to Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*), suggesting that mitochondrial sequences are not appropriate for the question at hand.

Combined Nuclear and Mitochondrial Evidence on the Identification of the Closest Living Relative(s) of Tetrapods

All currently available molecular data on the relationships among living sarcopterygians were finally combined: the mitochondrial DNA data presented in this paper and the nuclear (28S rRNA) data previously published (Zardoya and Meyer 1996*b*) (20,926 characters). Analyses were performed with MP and NJ on the five taxa (trout, coelacanth, lungfish, frog, and human) for which both data sets are available. However, for the ML analyses, the additive capability of the estimated log-likelihood for different genes allowed us to evaluate the total evidence of several independent analyses of different genes (Adachi and Hasegawa 1996*a*; Hasegawa, Adachi, and Milinkovitch 1997) and, accordingly, to combine the analysis of the 10-taxon mitochondrial data set of this paper and the 12-taxon 28S rRNA data set (Zardoya and Meyer 1996*b*).

When (1) third codon positions of mitochondrial protein-coding genes were excluded, (2) a transition : transversion weight of 1:2 was applied for the 28S rRNA data set (based on an estimated α/β ratio of 3.85), and (3) rainbow trout was used as outgroup, the MP analysis of the combined nuclear and mitochondrial data set arrived at a single most parsimonious tree (7,075 steps, CI = 0.84) in which a lungfish + coelacanth clade was favored, but with bootstrap support below 50%. Under the same conditions, however, the NJ analysis (Kimura two-parameter; α = 0.62) supported a lungfish + tetrapod clade with a bootstrap value of 90%. Finally, the ML analysis of this data set favored a lungfish + coelacanth relationship when the mitochondrial protein-coding gene subset was analyzed both at the amino acid level (79% bootstrap support) and at the DNA level (61% bootstrap support) (table 3). However, if a Kishino-Hasegawa test is performed, none of the two alternative hypotheses could be statistically ruled out with this data set (table 3).

Conclusion

We presented a detailed phylogenetic analysis of the largest data set (up to 20,926 characters) collected so far to address the question of the evolutionary relationships among the coelacanth, lungfishes, and tetrapods. Among the three competing hypotheses that can explain the relationships among the three living sarcopterygian lineages, the overall evidence mainly supported a lungfish + coelacanth and a lungfish + tetrapod grouping, but not the coelacanth + tetrapod relationship. The lungfish + coelacanth clade was supported by the nuclear 28S rRNA gene data (Zardoya and Meyer 1996*b*) and the mitochondrial ATPase6, ATPase8, COI, COII, COIII, ND5, and tRNA data (Zardoya and Meyer 1997*a*). However, the alternative hypotheses, i.e., lungfish as sister group to tetrapods and coelacanth as sister group to tetrapods, could not be statistically ruled out. In fact, the mitochondrial ND1, ND2, ND3, ND4L, ND4, ND6, and cytochrome *b* genes (and all mitochondrial protein-coding genes combined) strongly supported a lungfish + tetrapod clade, whereas the rRNA data set weakly favored a coelacanth + tetrapod clade. Although it cannot be statistically rejected in all cases, it seems from our analyses that the coelacanth + tetrapod hypothesis is the most unlikely of the three. Both the lungfish + coelacanth and the lungfish + tetrapod hypotheses are also the most preferred from a morphological point of view (Forey 1987; Panchen and Smithson 1987; Ahlberg 1991; Chang 1991; Forey, Gardiner, and Patterson 1991; Schultze 1994; Long 1995; Cloutier and Ahlberg 1996).

The contradictory evidence that is obtained from this large molecular data set and the presence of two different conflicting signals in the same data set reflects the difficulty of finding unequivocal molecular traces of the rapid origin of sarcopterygian lineages within a narrow window of time (15 million years) that dates back to the Devonian (Meyer 1995). Differences in rates of evolution in different genes likely account in part for the conflicting support for alternative topologies (Russo, Takezaki, and Nei 1996; Zardoya and Meyer 1996*c*), since it appears that, in general, the more slowly evolving genes (cytochrome oxidase subunits, tRNAs, and 28S rRNA) support the lungfish + coelacanth hypothesis, whereas relatively faster evolving genes (e.g., NADH subunits) support the lungfish + tetrapod grouping (table 4). As exceptions to this rule, the relatively slowly evolving gene cytochrome *b* favors a lungfish + tetrapod clade, and the relatively fast genes ND6, ATPase8, and ATPase6 support a lungfish + coelacanth grouping (table 4). On the other hand, no significant

Table 3
Nuclear and Mitochondrial Support to the Three Competing Hypotheses on the Relationships of Living Sarcopterygians

Tree	rRNA		tRNA		28S rRNA		Protein-Coding (aa)		Total with Protein-Coding (aa)		Protein-Coding (DNA)		Total with Protein-Coding (DNA)	
	Δ ± SE	BP	Δ ± SE	BP	Δ ± SE	BP	Δ ± SE	BP	Δ ± SE	BP	Δ ± SE	BP	Δ ± SE	BP
(Lungfish, tetrapods)	−16.5 ± 10	37	−13.4 ± 7.8	1	−16 ± 9.9	3	ML (34,029.4)	64	−21.1 ± 20.8	18	ML (38,220.0)	83	−7.9 ± 22.2	35
(Coelacanth, tetrapods)	ML (13,468.9)	94	−8.3 ± 9.1	17	−14.6 ± 10.2	5	−37.7 ± 19.4	0	−35.8 ± 23.7	3	−45.6 ± 20.9	0	−30.5 ± 25	4
(Lungfish, coelacanth)	−17.7 ± 9.8	21	ML (9,347.6)	82	ML (11,505.5)	92	−7 ± 14.4	36	ML (68,376.2)	79	−20.3 ± 16.5	17	ML (72,580.0)	61
Length	2,101		1,417		3,221		3,429		10,168		6,858		13,597	

Note.—The −log-likelihoods of the maximum-likelihood (ML) trees are given in parentheses, and the differences in log-likelihoods of alternative trees from that of the ML tree (Δ) are shown with their SEs (following ±), which were estimated by the formula of Kishino and Hasegawa (1989). Third codon positions of protein-coding genes were excluded from the analyses at the DNA level. Two phylogenies were declared significantly different when the difference in log-likelihoods was more than 1.96 times the SE. BP = Bootstrap probabilities.

Table 4
Favored Hypotheses and Relative Rates of Evolution of the Different Protein-Coding Mitochondrial Genes

Gene	TBL	Maximum *p* distance	Hypothesis
ATPase8	429.96	79	Lungfish + coelacanth
ND6	367.7	67	Lungfish + coelacanth
ND4L	361.8	42	Lungfish + tetrapods
ND2	316.72	58	Lungfish + tetrapods
ND4	225.08	60	Lungfish + tetrapods
ND3	204.82	44	Lungfish + tetrapods
ND5	203.02	44	Lungfish + coelacanth
ATPase6	175.79	50	Lungfish + coelacanth
ND1	170.33	33	Lungfish + tetrapods
COII	147.48	33	Lungfish + coelacanth
Cytochrome *b*	140.73	28	Lungfish + tetrapods
COIII	91.91	25	Lungfish + coelacanth
COI	54.99	14	Lungfish + coelacanth

Note.—Relative rates of evolution of the different genes are shown as total branch lengths (TBLs) of the preferred hypothesis of each gene estimated with maximum likelihood and maximum (uncorrected) *p* distances as calculated by Russo, Takezaki, and Nei (1996).

base compositional differences were found among the taxa studied (Zardoya and Meyer 1997*a*). Therefore, this points to a rate-related bias rather than a base compositional bias in the outcome of this phylogenetic question, although additional factors that affect the accuracy of the phylogenetic reconstruction should also be involved (e.g., structural constraints or concerted evolution of ATPase versus cytochrome oxidase versus NADH subunits might also be related to the existence of two different and conflicting phylogenetic signals in the same data set).

It seems evident from our results that new molecular data need to be found that can confidently resolve the true phylogeny among the three living sarcopterygian lineages. Future studies on this question will need to focus on nuclear protein-coding genes and search for phylogenetically informative insertion/deletion events in both coding and noncoding nuclear regions. From a mitochondrial point of view, the sequencing of new amphibian mitochondrial genomes (unpublished data) is desirable, because the frog (*Xenopus laevis*) mitochondrial genome (the only one yet sequenced from amphibians) (Roe et al. 1985) shows a long branch which potentially can bias phylogenetic inferences regarding this elusive evolutionary issue.

Acknowledgments

We thank David Mindell and an anonymous reviewer for providing useful suggestions on the manuscript. David Swofford kindly granted permission to publish results based on the test version of his PAUP* program. R.Z. was sponsored by a postdoctoral grant from the Ministerio de Educacion y Ciencia of Spain. This work received partial financial support from grants from the National Science Foundation (BSR-9107838, BSR-9119867, DEB-9615178) and from a collaboration grant with the Max-Planck-Institut für Biology in Tübingen from the Max-Planck Society, Germany, to A.M. Moreover, this work was supported by grants from the

Ministry of Education, Science, and Culture of Japan to M.H.

LITERATURE CITED

ADACHI, J., and M. HASEGAWA. 1996*a*. MOLPHY version 2.3: Programs for molecular phylogenetics based on maximum likelihood. Comput. Sci. Monogr. **28**:1–150.

———. 1996*b*. Model of amino acid substitution in proteins encoded by mitochondrial DNA. J. Mol. Evol. **42**:459–468.

AHLBERG, P. E. 1991. A re-examination of sarcopterygian interrelationships, with special reference to the porolepiformes. Zool. J. Linn. Soc. **103**:241–288.

AHLBERG, P. E., J. A. CLACK, and E. LUKSEVICS. 1996. Rapid braincase evolution between *Panderichthys* and the earliest tetrapods. Nature **381**:61–63.

ARNASON, U., and A. GULLBERG. 1993. Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. J. Mol. Evol. **37**:312–322.

CAO, Y., J. ADACHI, A. JANKE, S. PAABO, 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, M. M. 1991. Rhipidistians. Pp. 1–28 *in* H. P. SCHULTZE and L. TRUEB, eds. Origins of the higher groups of tetrapods. Controversy and consensus. Cornell University Press, Ithaca, N.Y.

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.

CLOUTIER, R., and P. E. AHLBERG. 1996. Interrelationships of basal sarcopterygians. Pp. 445–479 *in* M. L. J. STIASSNY, L. R. PARENTI, and G. D. JOHNSON, eds. Interrelationships of fishes. Academic Press, San Diego.

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

FELSENSTEIN, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. **27**:401–410.

———. 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.

FOREY, P. L. 1987. Relationships of lungfishes. J. Morphol. **1**(Suppl.)**:**75–91.

FOREY, P. L., B. G. GARDINER, and C. PATTERSON. 1991. The lungfish, the coelacanth and the cow revisited. Pp. 145–174 *in* H. P. SCHULTZE and L. TRUEB, eds. Origins of the higher groups of tetrapods. Controversy and consensus. Cornell University Press, Ithaca, N.Y.

HASEGAWA, M., J. ADACHI, and M. MILINKOVITCH. 1997. Novel phylogeny of whales supported by total molecular evidence. J. Mol. Evol. **44**(Suppl. 1)**:**117–120.

HASEGAWA, M., H. KISHINO, and T. YANO. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22**:160–174.

HEDGES, S. B., C. A. HASS, and L. R. MAXSON. 1993. Relations of fish and tetrapods. Nature **363**:501–502.

HORAI, S., K. HAYASAKA, R. KONDO, K. TSUGANE, and N. TAKAHATA. 1995. Recent African origin of modern 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. PAABO. 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.

KISHINO, H., T. MIYATA, and M. HASEGAWA. 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. J. Mol. Evol. **30**:151–160.

KUMAZAWA, Y., and M. NISHIDA. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. **37**:380–398.

———. 1996. Phylogenetic utility of mitochondrial tranfer RNA genes for deep divergence in animals. Pp. 29–35 *in* M. NEI and N. TAKAHATA, eds. Current topics on molecular evolution. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, and the Graduate School for Advanced Studies, Hayama, Japan.

LEE, W. J., and T. D. KOCHER. 1995. Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. Genetics **139**:873–887.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

LONG, J. A. 1989. A new rhizodontiform fish from the early Carboniferous of Victoria, Australia with remarks on the phylogenetic position of the group. J. Vertebr. Paleontol. **9**: 1–17.

———. 1995. The rise of fishes: 500 million years of evolution. The John Hopkins University Press, Baltimore and London.

LYONS-WEILER, J., G. A. HOELZER, and R. J. TAUSCH. 1998. Optimal outgroup analysis. Biol. J. Linn. Soc. (in press).

MAISEY, J. G. 1996. Discovering fossil fishes. Holt, New York.

MEYER, A. 1995. Molecular evidence on the origin of tetrapods and the relationships of the coelacanth. Trends Ecol. Evol. **10**:111–116.

MEYER, A., and S. I. DOLVEN. 1992. Molecules, fossils and the origin of tetrapods. J. Mol. Evol. **35**:102–113.

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

NAYLOR, G. J., and W. M. BROWN. 1997. Structural biology and phylogenetic estimation. Nature **388**:527–528.

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

NOACK, K., R. ZARDOYA, and A. MEYER. 1996. The complete mitochondrial DNA sequence of the bichir (*Polypterus ornatipinnis*), a basal ray-finned fish: ancient establishment of the consensus vertebrate gene order. Genetics **144**:1165–1180.

PANCHEN, A. L., and T. R. SMITHSON. 1987. Character diagnosis, fossils and the origin of tetrapods. Biol. Rev. **62**:341–438.

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.

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

SCHULTZE, H. P. 1987. Dipnoans as sarcopterygians. J. Morphol. **1**(Suppl.)**:**39–74.

———. 1994. Comparison of hypotheses on the relationships of sarcopterygians. Syst. Biol. **43**:155–173.

SIEGEL, S. 1956. Nonparametric statistics. McGraw-Hill, New York.

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

———. 1997. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer, Sunderland, Mass.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. HILLIS, C. MORITZ, and B. K. MABLE, eds. Molecular systematics. Sinauer, Sunderland, Mass.

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.

VOROBYEVA, E., and H. P. SCHULTZE. 1991. Description and systematics of panderichthyid fishes with comments on their relationship to tetrapods. Pp. 68–109 *in* H. P. SCHULTZE and L. TRUEB, eds. Origins of the major groups of tetrapods: controversies and consensus. Cornell University Press, Ithaca, N.Y.

YANG, Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. **11**:367–372.

YANG, Z., and S. KUMAR. 1996. Approximate methods for estimating the pattern of nucleotide substitution and the variation of substitution rates among sites. Mol. Biol. Evol. **13**: 650–659.

YOKOBORI, A. I., M. HASEWAGA, T. UEDA, N. OKADA, K. NISHIKAWA, and K. WATANABE. 1994. Relationship among coelacanths, lungfishes, and tetrapods: a phylogenetic analysis based on mitochondrial cytochrome oxidase I gene sequences. J. Mol. Evol. **38**:602–609.

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*a.* 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.

———. 1996*b.* Evolutionary relationships of the coelacanth, lungfishes, and tetrapods based on the 28S ribosomal RNA gene. Proc. Natl. Acad. Sci. USA **93**:5449–5454.

———. 1996*c.* Phylogenetic performance of mitochondrial protein coding genes in resolving relationships among vertebrates. Mol. Biol. Evol. **13**:933–942.

———. 1997*a.* The complete DNA sequence of the mitochondrial genome of a ''living fossil'', the coelacanth (*Latimeria chalumnae*). Genetics **146**:995–1010.

———. 1997*b.* Molecular phylogenetic information on the identity of the closest living relative(s) of land vertebrates. Naturwissenschaften **84**:389–397.

SHOZO YOKOYAMA, reviewing editor

Accepted January 9, 1998