*Proc. Natl. Acad. Sci. USA*
Vol. 93, pp. 5449–5454, May 1996
Evolution

# Evolutionary relationships of the coelacanth, lungfishes, and tetrapods based on the 28S ribosomal RNA gene

**(lobe-finned fishes/molecular phylogeny/colonization of land/preadaptation)**

RAFAEL ZARDOYA* AND AXEL MEYER

Department of Ecology and Evolution and Program in Genetics, State University of New York, Stony Brook, NY 11794-5245

*Communicated by Ernst Mayr, Harvard University, Cambridge, MA, January 25, 1996 (received for review July 31, 1995)*

**ABSTRACT** **The origin of land vertebrates was one of the major transitions in the history of vertebrates. Yet, despite many studies that are based on either morphology or molecules, the phylogenetic relationships among tetrapods and the other two living groups of lobe-finned fishes, the coelacanth and the lungfishes, are still unresolved and debated. Knowledge of the relationships among these lineages, which originated back in the Devonian, has profound implications for the reconstruction of the evolutionary scenario of the conquest of land. We collected the largest molecular data set on this issue so far, about 3,500 base pairs from seven species of the large 28S nuclear ribosomal gene. All phylogenetic analyses (maximum parsimony, neighbor-joining, and maximum likelihood) point toward the hypothesis that lungfishes and coelacanths form a monophyletic group and are equally closely related to land vertebrates. This evolutionary hypothesis complicates the identification of morphological or physiological preadaptations that might have permitted the common ancestor of tetrapods to colonize land. This is because the reconstruction of its ancestral conditions would be hindered by the difficulty to separate uniquely derived characters from shared derived characters in the coelacanth/lungfish and tetrapod lineages. This molecular phylogeny aids in the reconstruction of morphological evolutionary steps by providing a framework; however, only paleontological evidence can determine the sequence of morphological acquisitions that allowed lobe-finned fishes to colonize land.**

The origin of land vertebrates is a question that has fascinated paleontologists and comparative morphologists for several decades. However, this issue is still debated because of the complexity of the series of morphological and physiological modifications that were involved in the transition of life in water to life on land, the difficulty in identifying homologous characters in fragmentary fossils, a general paucity of fossils, and the rapidity with which land was conquered (for reviews, see refs. 1–4). Nonetheless, it is generally accepted that lobe-finned fishes (Sarcopterygii), which also include the tetrapod lineage (5), form a monophyletic group. Ray-finned fishes (Actinopterygii) are only distantly related to tetrapods (6, 7). To better understand the origin of tetrapods, their phylogenetic relationships to other lineages of lobe-finned fishes need to be agreed upon. It is likely that an extinct lineage of rhipidistians (a major group of Sarcopterygian fishes) might have been the direct ancestor of tetrapods rather than any living lineage of lobe-finned fishes (refs. 3 and 8; for review, see ref. 4). Coelacanths (Actinistia) were traditionally classified with rhipidistians in the Crossopterygii (e.g., ref. 8) and, therefore, many believed that they are the closest living relatives of tetrapods (e.g., ref. 9). However, other researchers (e.g., refs. 7 and 10), based on morphological evidence, also proposed that lungfishes are the closest living relatives of tetrapods. It still remains unclear which of the two extant groups of lobe-finned fishes, the lungfish or the coelacanth, is the living sister group of land vertebrates or whether both are equally closely related to tetrapods (Fig. 1).

During the last decade, molecular data have been collected with the explicit goal of discriminating between these three competing hypotheses (Fig. 1). There is evidence from molecular data, particularly mitochondrial DNA sequences (11–13), that links lungfishes rather than the coelacanth as closest living relatives, among fishes, to tetrapods (Fig. 1*A*). Conversely, based on the molecular studies published to date, evidence for the sister group relationship of the coelacanth and tetrapods is weak (Fig. 1*B*) (for review, see ref. 4). But there are morphological (14) and, recently, also molecular (15) data supporting a coelacanth/lungfish grouping; hence those studies suggest that both the coelacanth and lungfishes are equally closely related to tetrapods (Fig. 1*C*) (for review, see ref. 4). However, this molecular study (15) could not statistically rule out one of the other hypotheses (for review, see ref. 4). Although many morphology and molecule-based phylogenetic studies have attempted to settle this question, the results obtained so far do not provide unequivocal evidence as to whether the lungfishes or the coelacanth or both lineages are equally (Fig. 1) related to tetrapods and these alternative hypotheses are still debated based on both kinds of evidence (for review, see ref. 4). It is clear that this will continue to be a difficult phylogenetic question to resolve because the radiation of lungfishes, coelacanth, and early tetrapods occurred within a narrow window of time (of about 20–30 million years) back in the Devonian. Studies that use molecules to investigate these evolutionary events are faced with the difficulty that mutations that mark these lineages have been overlaid by newer mutations during the last 360 million years (for review, see ref. 4).

It is apparent (4, 12, 15) that more sequence information than is available today is needed to establish a molecule-based hypothesis of the phylogeny of tetrapods. In addition, not only more sequences but also more species should ideally be included in studies on this issue (12, 15, 16) and a gene of appropriate mutation rate needs to be identified. Studies based on nuclear ribosomal genes have a long and successful history in establishing phylogenetical relationships among distantly related groups of species (for reviews, see refs. 17 and 18). Fragments of both 18S and 28S ribosomal RNA genes were assayed in the past to address the question on the origin of tetrapods (16, 19, 20). These studies supported the sister group relationship of lobe-finned fishes to tetrapods, but they were typically unable to distinguish which sarcopterygian fishes were more akin to tetrapods because they often did not include either the lungfishes or the coelacanth (for review, see ref. 4).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Abbreviations*: MP, maximum parsimony; ML, maximum likelihood; NJ, neighbor joining.

*Data deposition*: The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U34336–U34342).

*To whom reprint requests should be addressed. e-mail: rzardoya@life.bio.sunysb.edu.

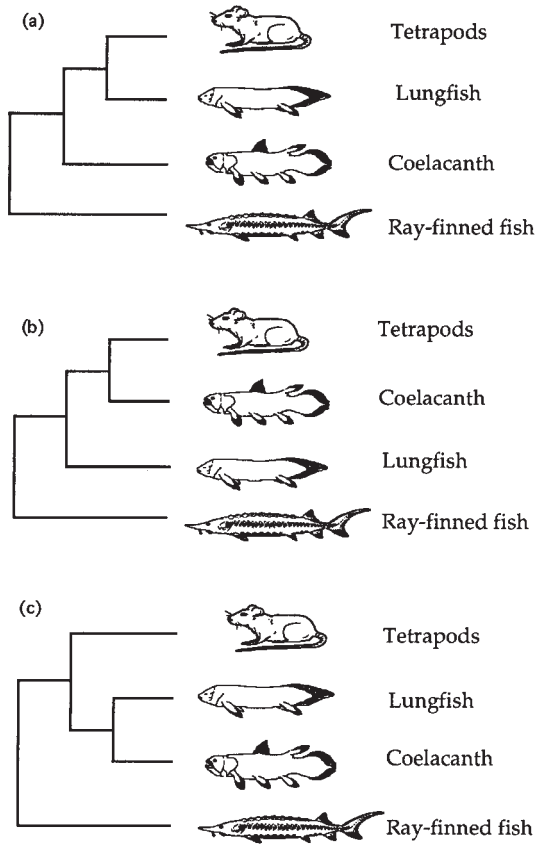


FIG. 1. Alternative hypotheses of sister group relationships between sarcopterygii and tetrapods. (*A*) Lungfish as the sister group of tetrapods. (*B*) Coelacanth as the closest living relative of tetrapods. (*C*) Coelacanth and lungfish equally closely related as sister groups of tetrapods.

Nuclear rRNA genes are found in multiple copies that are organized in tandem arrays separated by nontranscribed spacers along the genome. Each repeat unit of the array consists of three rRNA genes (18S, 5.8S, and 28S) and two internal and one external transcribed spacers (e.g., refs. 17, 18, and 21). Each rRNA gene is organized into several divergent domains, also called expansion segments, interspersed between slowly evolving and, therefore, highly conserved cores (22). Divergent domains are commonly subjected to insertion and deletion events, which coupled with substitution rates that are at least two orders of magnitude higher than that for cores (23), account for the large overall size variation of rRNA genes among eukaryotes. A clear bias toward nucleotide substitutions (especially transitions) rather than insertions or deletions is found in the highly conserved core sequences (23). These dual modes of evolution in the expansion domains and the cores of nuclear rRNA molecules (particularly 28S) make these genes useful for phylogenetic analyses. This is because relationships can be addressed among distantly related groups (by focusing on the slowly evolving core regions) (e.g., refs. 24 and 25) and relatively closely related taxa (26) (by focusing on variation in the more rapidly evolving expansion segments). However, this mode of rRNA gene evolution also means that, at any given evolutionary distance, the phylogenetic signal is constrained to relatively few sites in the molecule because conserved positions are invariant, whereas variable sites, which might also be ambiguously aligned, obscure the phylogenetic signal by back-mutations and hence noise (23).

The vast majority of phylogenetic analyses that are based on rRNA genes have typically used the 18S rRNA gene, which is the more slowly evolving one (compiled in ref. 18). Evolutionary studies using the 18S rRNA gene successfully documented the early branching of eukaryotes but had limited success with more closely related taxa such as vertebrates and the origin of tetrapod question (19). The 28S rRNA gene is a molecule that seemed particularly promising for this issue (for review, see ref. 4) because it had been used to suggest a clear link between the coelacanth and tetrapods (18). Unfortunately, until now no lungfishes had been investigated; this prevented the discrimination between the three competing hypotheses (Fig. 1). Moreover, only five complete 28S rRNA gene sequences have been reported for vertebrates so far (22, 27–30).

We amplified and sequenced by PCR (31) almost the complete 28S rRNA gene in seven fish species: three lungfish species, the coelacanth, the rainbow trout, an eel, and a sturgeon. Thereby, we more than doubled the number of vertebrate species for which the entire 28S rRNA gene has been sequenced. The new sequences were compared to the complete sequences already reported for five tetrapods [mouse (22); clawed frog (27); rat (28); human (29); Kenya smooth clawed frog (30)]. The phylogenetic analyses of these sequences indicate that the 28S rRNA gene can be used to address the question of the relations among ray-finned fishes, lobe-finned fishes, and tetrapods.

## MATERIAL AND METHODS

**DNA Sources and Extraction.** Total cellular DNA was extracted from muscle of rainbow trout (*Oncorhynchus mykiss*), eel (*Anguilla rostrata*), short nose sturgeon (*Acipenser brevirostrum*), the coelacanth (*Latimeria chalumnae*) and the South American (*Lepidosiren paradoxa*), African (*Protopterus aethiopicus*) and Australian (*Neoceratodus forsteri*) lungfish as described (32).

**PCR Amplification, Cloning and DNA Sequencing.** A combination of eight sets of primers (Fig. 2) was designed to amplify contiguous and overlapping fragments (averaging about 450 bp) of almost the entire 28S rRNA gene. These primers were designed based on highly conserved regions of an alignment of published sequences from several eukaryotes (33). Special primers were needed to amplify the D2 divergent domain of eel and coelacanth (D2′F, 5′-GGTGGTAAACTCCATCTAAGGCTA-3′; D2′R, 5′-ATAGTTCACCATCTTTCGGGTCC-3′) and the D8 divergent domains of coelacanth, eel, and sturgeon (D8′R, 5′-AGTGGGAATCTCGTTCATCCA-3′) because the original primer sets did not amplify in these species. Amplifications were done in 50 $\mu$l of Tris (67 mM, pH 8.8) containing 2 mM $MgCl_2$, 0.3 mM of each dNTP, 150 ng of each primer, template DNA (10–1000 ng), and AmpliTaq DNA polymerase (2.5 units, Perkin/Elmer–Cetus). Thirty cycles of PCR (denaturing at 94°C for 60 s, annealing at 50°C for 60 s, and extending at 72°C for 60 s) were performed to generate double-stranded DNA fragments. An aliquot of the PCR product (5 $\mu$l) was cloned in pGEM-T Vector (Promega) following the manufacturer's instructions. Typically, two to four clones were sequenced for each PCR product. Recombinant plasmids were sequenced on an Applied Biosystems model 373A Stretch DNA sequencer using the Taq Dye Deoxy Terminator Cycle Sequencing Kit (Applied Biosystems) and M13 universal (−40) and reverse primers following manufacturer's instructions.

**Sequence Analysis.** A multiple-sequence alignment was performed using CLUSTAL W (34) followed by refinement by eye based on 28S rRNA secondary structure models. The aligned data set can be requested from the authors. The aligned sequences were subjected to maximum-parsimony (MP) method [PAUP 3.1.1 (35)] using branch and bound to find the most parsimonious tree. Maximum-likelihood (ML) and neighbor-joining (NJ) (36) (based on Kimura distance matrixes) analyses of the sequences were performed with PHYLIP 3.5 (37). Robustness of the MP and NJ phylogenetic results was tested by bootstrap analyses (38) (PAUP and PHYLIP, 500 replications). The likelihood of the three alternative hypoth-

| Name | Sequence (5′→3′) | Sense | Divergent domain |
|---|---|---|---|
| 28S D1F | CCCGCTGAATTTAAGCATATAAGTAAGCGG | Forward | D1 |
| 28S D1R | AACGGTTTCACGCCCTCTTGAACT | Reverse | |
| 28S D2F | AAGTTGAAAAGAACTTTGAAGAGA | Forward | D2 |
| 28S D2R | GTTAGACTCCTTGGTCCGTGT | Reverse | |
| 28S D3F | CCCACCCCGACCCGTCTTGAA | Forward | D3, D4 and D5 |
| 28S D3R | TCGGTTCATCCCGCAGCG | Reverse | |
| 28S D6F | GGTAAAGCGAATGATTAGAGGTCTT | Forward | D5 and D6 |
| 28S D6R | GACTGACCCATGTTCAACTGCTGT | Reverse | |
| 28S D7F | AAGTGGAGAAGGGTTCCATGTGA | Forward | D7 |
| 28S D7R | AGAGCCAATCCTTATCCCGAAGTT | Reverse | |
| 28S D8F | AAGGGAAGTCGGCAAGTCAGATCCGT | Forward | D8 |
| 28S D8R | TAATTAGATGACGAGGCATTTGGC | Reverse | |
| 28S D9F | CGGCGGGAGTAACTATGACTCTCTTAAGGT | Forward | D9 and D10 |
| 28S D9R | CCGCCCCAGCCAAACTCCCCA | Reverse | |
| 28S D11F | TGAAATACCACTACTCTTATCGTT | Forward | D10 and D11 |
| 28S D11R | GGATTCTGACTTAGAGGCGTTCAG | Reverse | |

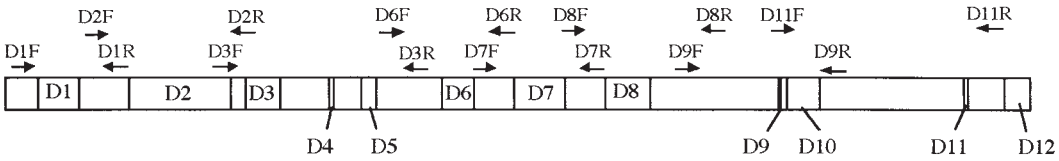


FIG. 2. Sets of primers designed to amplify contiguous and overlapping fragments covering almost (>95%) the complete 28S rRNA gene. Position, orientation, and divergent domains amplified by each set of primers are depicted.

esis on the origin of tetrapods based on this data set was evaluated by the method of Kishino and Hasegawa (39) using the MOLPHY package (40).

## RESULTS

**Amplification of the 28S Gene by PCR.** The designed sets of primers consistently amplified by PCR (31), fragments which partially overlap each other and cover almost the whole (more than 95% of the gene) 28S gene from seven species of fishes. The sequence of these primers is conserved among the studied species; it is expected that they will work as well in other vertebrate taxa. Variation among PCR products for several clones of a given fragment when found was located in variable expansion segments that were subsequently excluded from phylogenetic analysis.

**Size and Base Composition of Fish 28S rRNA Genes.** The almost complete sequence of the 28S rRNA gene, covering 11 divergent domains and their flanking conserved regions, was determined in seven fish species (for a total of about 24,000 bp determined for this study). The average overall length of these 28S gene fragments was of ≈3,500 bp, whereas in amphibians and mammals these fragments are ≈3,700 and 4,400 bp long, respectively (Table 1). This increase in size of almost 1 kb between 28S rRNA genes of fishes and mammals is mainly due to length variation in divergent domains D2 and D8. All divergent domains combined make up about half of fish 28S rRNA genes. In fishes, the overall base composition of 28S

Table 1. Base frequencies of aligned 28S rRNA sites used in the phylogenetic analyses

| Species | A | C | G | T | No. of sites analyzed | Total no. of sites |
|---|---|---|---|---|---|---|
| *Acipenser* | 0.218 | 0.256 | 0.330 | 0.196 | 3172 | 3431 |
| *Anguilla* | 0.212 | 0.267 | 0.340 | 0.180 | 3177 | 3662 |
| *Oncorhynchus* | 0.217 | 0.254 | 0.340 | 0.195 | 3177 | 3573 |
| *Latimeria* | 0.214 | 0.258 | 0.340 | 0.187 | 3169 | 3410 |
| *Neoceratodus* | 0.232 | 0.238 | 0.314 | 0.216 | 3164 | 3400 |
| *Lepidosiren* | 0.228 | 0.240 | 0.326 | 0.205 | 3167 | 3472 |
| *Protopterus* | 0.231 | 0.238 | 0.323 | 0.209 | 3103 | 3351 |
| *Xenopus laevis* | 0.206 | 0.271 | 0.346 | 0.177 | 3160 | 3725 |
| *Xenopus borealis* | 0.208 | 0.271 | 0.346 | 0.175 | 3178 | 3759 |
| *Rattus* | 0.208 | 0.268 | 0.348 | 0.175 | 3180 | 4351 |
| *Mus* | 0.209 | 0.268 | 0.346 | 0.176 | 3180 | 4283 |
| *Homo* | 0.209 | 0.267 | 0.349 | 0.174 | 3168 | 4592 |
| *Means* | 0.216 | 0.258 | 0.337 | 0.189 | 3166 | 3751 |

rRNA genes is characterized by an overrepresentation of guanines (32–34%) whereas the other nucleotides are almost equally distributed (Table 1). In tetrapods, on the other hand, 28S rRNA genes have a slight bias toward cytosines (30–33%) as well as guanines (36%). The conserved core regions of the 28S rRNA molecule are particularly rich in purines; a pattern that is conserved and found in all vertebrates. The divergent domains show a high G+C content (of about 70% in fishes and 80% in tetrapods due to the increase in cytosines) with a marked compositional bias against adenines and thymines. This feature is more pronounced in tetrapods than in fishes.

**Phylogenetic Analyses of 28S rRNA: The Coelacanth and Lungfishes Are Equally Closely Related to Tetrapods.** The new nucleotide sequences of the fish 28S rRNA genes (short nose sturgeon, eel, rainbow trout, coelacanth, and the South American, African, and Australian lungfish), were aligned to those already reported for mouse (22), clawed frog (27), rat (28), human (29), and Kenya smooth clawed frog (30) in order to determine which of the lineages of lobe-finned fishes is the closest living relative of tetrapods. Unambiguous alignments were obtained for most of the 28S rRNA gene, not only in the conserved core but also including large portions of the expansion segments. However, several divergent domains (in particular D2, D6, and D8) showed definite regions of ambiguous alignment due to the presence of large insertions in tetrapods. Although some of these regions could be aligned among closely related species (e.g., within lungfishes, amphibians, or rodents), they were excluded from subsequent phylogenetic analyses (23). Our data set was large and consisted of 4,786 sites of which 1,565 were conservatively excluded due to ambiguous alignment, 2,414 were invariant, and 807 were phylogenetically informative. Differences in base compositional biases between species under consideration can potentially interfere with phylogenetic reconstruction. However, for all nucleotide sites of all species included in these analyses, differences in base composition were minimal (Table 1) and should not have interfered with the recovery of phylogenetic signal.

Similar to previous phylogenetic analyses of the 28S rRNA gene (18), the same weight was given to paired (stems) and unpaired (loops, bulges) positions. To account for the faster rate of transitions over transversions in nuclear rRNA genes (17, 26), a transition/transversion ratio of 2.0 was used in the phylogenetic analyses because this was the ratio observed among relatively closely related species. Analyses of the data set with MP, NJ, and ML methods using sturgeon as outgroup yielded identical and congruent topologies (Fig. 3).

All three phylogenetic methods unequivocally support a coelacanth/lungfish clade, and thereby suggest that both groups, the lungfishes and the coelacanth, are equally closely related to tetrapods (Figs. 1*C* and 3). The monophyly of tetrapods is confirmed and that of living lobe-finned fishes is strongly supported. Among lungfishes, the South American and African species also form a monophyletic group as expected. The robustness of these results was confirmed by the high bootstrap values (38) obtained in NJ and MP trees (Fig. 3). The strength of the phylogenetic signal of 28S rRNA gene sequences was also measured by the $g_1$ test statistic (41). The tree-length distribution was strongly skewed, an indication of significant phylogenetic signal ($g_1 = -1.1$; $10^4$ random trees). The tree-length distribution of the topologically constrained analyses clustering tetrapods ($g_1 = -0.94$) was also highly skewed. In the ML analysis, all branch lengths were found to be significantly greater than zero ($P < 0.01$); a consistency index of 0.77 was obtained for the MP tree (without transition/transversion weighting).

To further assess the robustness of the result obtained from the phylogenetic analyses of the 28S rRNA genes, the standard errors of the difference in log-likelihood of the alternative hypotheses on the origin of tetrapods (i.e., either lungfishes or coelacanth as closest living relative of tetrapods) compared with the favored ML tree (Fig. 1*C*) were calculated by the formula of Kishino and Hasewaga (39). The alternative trees had log-likelihoods significantly lower than that of the ML tree (Fig. 4), again supporting the hypothesis shown in Fig. 1*C*. The favored MP tree (1762 steps) is 9–10 steps shorter (Fig. 4) than the alternative solutions (Fig. 1).

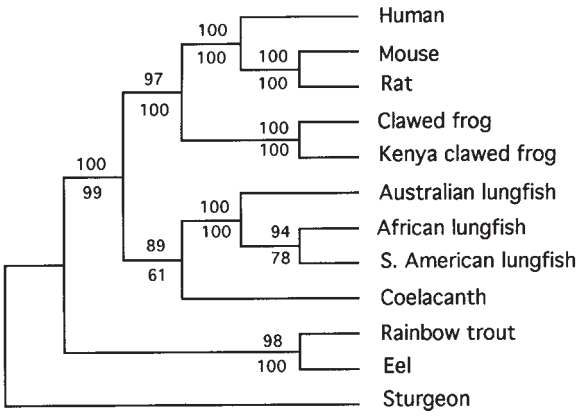


FIG. 3. Majority rule bootstrap consensus tree of tetrapods phylogenetic relationships based on 500 replications (38). The 28S rRNA data set was subjected to MP (bootstrap values above branches) and NJ (bootstrap values below branches) analyses.

## DISCUSSION

The 28S rRNA gene sequences reported in this work more than double the sequence information known on this gene for vertebrates. These new sequences will contribute to a better understanding of the patterns of evolution of this scarcely studied molecule, especially in vertebrates. Furthermore, a list of specific primers that reliably amplify this gene is provided for future studies (Fig. 2).

The 28S rRNA genes of fishes are slightly shorter than those of amphibians and much shorter than those of mammals. This increase in size of the 28S rRNA among vertebrates is precisely localized in the expansion segments of the molecule as was noted (22). The expansion segments of vertebrate 28S rRNAs are characterized by a slight bias to exclude adenine and for high G+C content. This base composition pattern reflects their potential capability to fold into secondary structures with particularly large GC-rich stems (42). However, the conserved core of the molecule has an almost even distribution of the different types of nucleotides (Table 1) that is correlated with the presence not only of stem-loop structures but also long single-stranded A-rich regions (43) that are thought to interact with other rRNA subunits and ribosomal proteins (42). Although divergent domains may differ extensively among distantly related taxa, they are a good source of phylogenetically informative sites (26) and a considerably large fraction of these domains could be unambiguously aligned in our data set and therefore included in the phylogenetic analyses.

The identical topology that resulted from MP, NJ, and ML phylogenetic analyses of the 28S rRNA gene data set is supported by high bootstrap values and favor a coelacanth/lungfish grouping (Figs. 1*C*, 3, and 4). Also, the reliability of the most parsimonious tree seems high as demonstrated by a $g_1$ test statistic (41). The strong resolution of the parsimony analysis and the presence of a clear phylogenetical signal in the 28S rRNA gene sequences are also confirmed by the fact that even by constraining the analysis by grouping those best-resolved internal branches of the most parsimonious tree (i.e., tetrapods) the skewness of the tree-length distribution is maintained. The favored ML tree inferred from this data set was found to be a significantly better estimate (39) than the two alternative topologies (Figs. 1 and 4). According to this

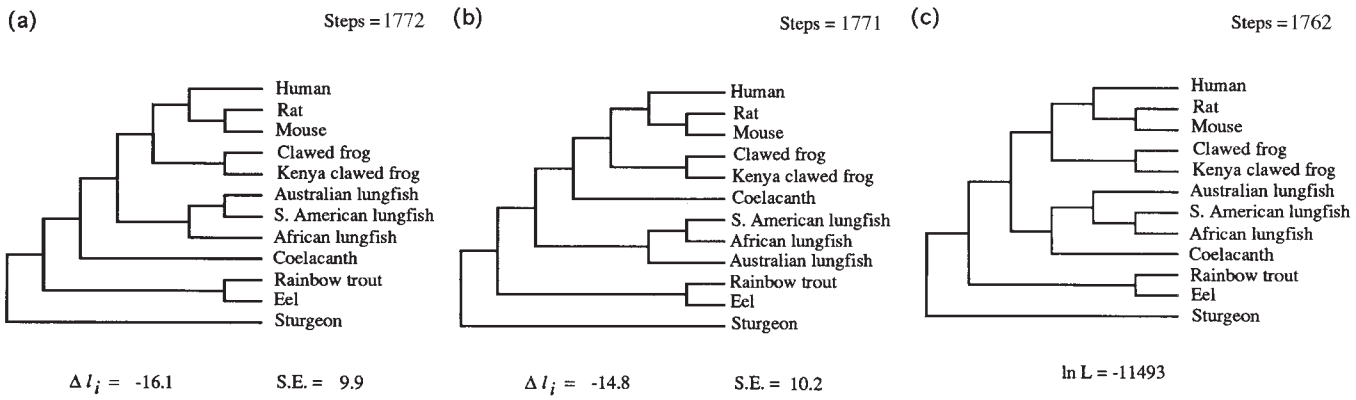


FIG. 4. ML analyses of the three alternative hypotheses (see Fig. 1) explaning the sister group relationships of lobe-finned fish groups and tetrapods. The ML of the three competing hypotheses was calculated based on the 28S rRNA data set by the formula of Kishino and Hasewaga (39). $\Delta l_i$, Difference in log-likelihood between the best tree (Figs. 1*C* and 3) and the alternative topologies. SE is the standard error. The lengths of the corresponding MP trees (Ti:Tv 2:1) for each hypothesis are also shown. Because the $\Delta l_i$ for hypotheses a and b are larger than their corresponding SE, they have a significantly less likely fit of the data to the model than hypothesis c.

topology, both lungfishes and the coelacanth are equally closely related to tetrapods and should both be considered as the closest living relatives of tetrapods.

Our result agrees with the most recently conducted morphological (14, 44) and molecular (15) analyses on this issue. Chang (14) pointed out several morphological traits that support the coelacanth/lungfish grouping. Chang's analyses contradicted previous morphological studies that identified lungfishes as the sistergroup of tetrapods (7) but support other earlier morphological analyses (e.g., ref. 45) of neural systems in these species.

The lungfish+tetrapod sistergroup relationship, originally proposed by Rosen *et al.* (7), had been supported by analyses of cytochrome *b* and rRNA mitochondrial DNA data sets (11–13). We have reanalyzed the largest mitochondrial DNA data set available so far (13) and found that none of the three hypotheses can be ruled out by it according to the Kishino-Hasewaga test (39). (Fig. 1 *A*, $\Delta l_i = -0.4$, SE = 7.4; *B*, $\Delta l_i = -0.7$, SE = 7.3; *C*, favored ML tree). However, because the origin of sarcopterygian lineages dates back to the Devonian this might be a question that is at the limit of resolution of fast-evolving mitochondrial genes such as 12S and 16S rRNA genes (unpublished data).

Conversely, the high statistical confidence of our results suggest that the phylogenetical signal derived from the complete 28S rRNA gene, unlike that of the 18S rRNA gene and mitochondrial rRNA genes, can successfully resolve phylogenetic problems that occurred 400 million years ago. This might be due to a favorable mix of slowly and rapidly evolving regions in the 28S rRNA gene. The ability of this gene to identify the living piscine sister group of tetrapods is an important contribution to the larger issue of clarifying the relationships among all sarcopterygian groups, including extinct ones.

The conflict between various morphology-based phylogenetic studies on this issue points to the problem of identifying homologous structures and the difficulty of how to deal with missing information from fragmentary fossils. Since the coelacanth and the lungfishes represent highly derived lineages, the coelacanth/lungfish grouping implies that it might be more difficult to identify phenotypic traits in all three extant lineages of lobe-finned fishes that might have been present in their common ancestor. The task of identifying preadaptations that facilitated the colonization of land of this common ancestor is rendered more complex if the coelcanth/lungfish grouping is correct. This is because, according to this grouping, shared traits between tetrapods and either lungfishes or the coelacanth could equally parsimoniously be interpreted to have been either lost or gained independently in one of these three lineages. Therefore, the morphology of the common ancestor of the coelacanth/lungfish and the tetrapod lineages is more difficult to reconstruct based on phenotypic traits. Whereas, if either hypothesis shown in Fig. 1 *A* or *B* were correct, shared derived traits of either lungfishes or the coelacanth and tetrapods would be assumed to have been present in their common ancestor and hence be interpreted as preadaptations for the colonization of land. Some extinct lineages of rhipidistian fishes are now known (46) to be more closely to tetrapods than either the lungfishes or the coelacanth and therefore hold the key to the reconstruction of the morphology of the ancestor of land vertebrates. Since only living representatives of lobe-finned fishes are available for molecular phylogenetic work, only paleontological work can establish the phylogeny and the character evolution from extinct species along the lineages leading to the three extant groups of sarcopterygians.

We thank Pete Lockhart (Massey University) for further analyses and Ernst Mayr, Scott Edwards, and David Rand for providing helpful suggestions on the manuscript. The tissues were kindly provided by David Reznick (South American lungfish), Robert Murphy (coelacanth), Blair Hedges (Australian lungfish), and William Bemis (African lungfish). This work received partial support by grants from the National Science Foundation to A.M. (BSR-9107838 and BSR-9119867). R.Z. was sponsored by a postdoctoral grant of the Ministerio de Educacion y Ciencia of Spain.

1. Forey, P. L. (1988) *Nature (London)* **336,** 727–732.
2. Ahlberg, P. E. & Milner, A. R. (1994) *Nature (London)* **368,** 507–514.
3. Schultze, H. P. (1994) *Syst. Biol.* **43,** 155–173.
4. Meyer, A. (1995) *Trends Ecol. Evol.* **10,** 111–116.
5. Ahlberg, P. E. (1991) *Zool. J. Linnean Soc.* **103,** 241–288.
6. Patterson, C. (1980) in *Origin of Tetrapods: Historical Introduction to the Problem*, ed. Panchen, A. L. (Academic, New York), pp. 159–175.
7. Rosen, D. E., Forey, P. L., Gardiner, B. G. & Patterson, C. (1981) *Bull. Am. Natl. Mus. Nat. Hist.* **167,** 159–276.
8. Romer, A. S. (1966) *Vertebrate Paleontology* (Univ. Chicago Press, Chicago).
9. Fritsch, B. (1987) *Nature (London)* **327,** 153–154.
10. Forey, P. L. (1987) *J. Morphol.* **1,** Suppl., 75–91.
11. Meyer, A. & Wilson, A. C. (1990) *J. Mol. Evol.* **31,** 359–364.
12. Meyer, A. & Dolven, S. I. (1992) *J. Mol. Evol.* **35,** 102–113.
13. Hedges, S. B., Hass, C. A. & Maxson, L. R. (1993) *Nature (London)* **363,** 501–502.
14. Chang, M. M. (1991) in *Origins of the Higher Groups of Tetrapods: Controversy and Consensus,* eds. Schultze, H. P. & Trueb, L. (Cornell Univ. Press, Ithaca, NY), pp. 3–28.
15. Yokobori, A. I., Hasewaga, M., Ueda, T., Okada, N., Nishikawa, K. & Watanabe, K. (1994) *J. Mol. Evol.* **38,** 602–609.

16. Le, H. L., Lecointre, G. & Perasso, R. (1993) *Mol. Phyl. Evol.* **2,** 31–51.
17. Mindell, D. P. & Honeycutt, R. L. (1990) *Annu. Rev. Ecol. Syst.* **21,** 541–566.
18. Hillis, D. M. & Dixon, M. T. (1991) *Q. Rev. Biol.* **66,** 411–453.
19. Stock, D. W., Moberg, k. D., Maxson, L. R. & Whitt, G. S. (1991) *Environ. Biol. Fish.* **32,** 99–117.
20. Hillis, D. M., Dixon, M. T. & Ammerman, L. K. (1991) *Environ. Biol. Fish.* **32,** 119–130.
21. Long, E. O. & Dawid, I. B. (1980) *Annu. Rev. Biochem.* **49,** 727–764.
22. Hassouna, N., Michot, B. & Bachellerie, J. P. (1984) *Nucleic Acids Res.* **12,** 3563–3583.
23. Olsen, G. J. & Woese, C. R. (1993) *FASEB J.* **7,** 113–123.
24. Gouy, M. & Li, W. H. (1989) *Mol. Biol. Evol.* **6,** 109–122.
25. Hasewaga, M., Iida, Y., Yano, T., Takaiwa, F. & Iwabuchi, M. (1985) *J. Mol. Evol.* **22,** 32–38.
26. Larson, A. & Wilson, A. C. (1989) *Mol. Biol. Evol.* **6,** 131–154.
27. Ware, V. C., Tague, B. W., Clark, C. G., Gourse, R. L., Brand, R. C. & Gerbi, S. A. (1983) *Nucleic Acids Res.* **11,** 7795–7817.
28. Hadjiolov, A. A., Georgiev, O. I., Nosikov, V. V. & Yavachev, L. P. (1984) *Nucleic Acids Res.* **12,** 3677–3693.
29. Gonzalez, I. L., Gorski, J. L., Campen, T. J., Dorney, D. J., Erickson, J. M., Sylvester, J. E. & Schmickel, R. D. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7666–7670.
30. Ajuh, P. M., Heeney, P. A., Maden, B. E. & Edward, H. (1991) *Proc. R. Soc. London B* **245,** 65–71.
31. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
32. Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X. & Wilson, A. C. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6196–6200.
33. Lenaers, G., Maroteaux, L., Michot, B. & Herzog, M. (1989) *J. Mol. Evol.* **29,** 40–51.
34. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
35. Swofford, D. L. (1993) PAUP, Phylogenetic Analysis Using Parsimony (Illinois Nat. Hist. Surv., Champaign IL).
36. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol* **4,** 406–425.
37. Felsenstein, J. (1989) *Cladistics* **5,** 164–166.
38. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
39. Kishino, H. & Hasewaga, M. (1989) *J. Mol. Evol.* **29,** 170–179.
40. Adachi, J. & Hasewaga, M. (1992) MOLPHY, Programs for Molecular Phylogenetics I-PROTML: Maximum Likelihood Inference of Protein Phylogeny (Inst. Stat. Math., Tokyo).
41. Hillis, D. M. & Huelsenbeck, J. P. (1992) *J. Hered.* **83,** 189–195.
42. Vawter, L. & Brown, W. M. (1993) *Genetics* **134,** 597–608.
43. Gutell, R. R., Weiser, B., Woese, C. R. & Noller, H. F. (1985) *Prog. Nucleic Acid Res. Mol. Biol.* **32,** 155–216.
44. Forey, P. L., Gardiner, B. G. & Patterson, C. (1991) in *Origins of the Higher Groups of Tetrapods: Controversy and Consensus.* eds. Schultze, H. P. & Trueb, L. (Cornell Univ. Press, Ithaca, NY), pp. 145–174.
45. Northcutt, R. G. (1987) *J. Morphol.* **1,** Suppl., 277–297.
46. Ahlberg, P. E. (1995) *Nature (London)* **373,** 420–425.