First publ. in: Genetics 155 (2000), pp. 765–775

Mitochondrial Evidence on the Phylogenetic Position of Caecilians (Amphibia: Gymnophiona)

Rafael Zardoya* and Axel Meyer†

**Museo Nacional de Ciencias Naturales, 28006 Madrid, Spain and †Department of Biology, University of Konstanz, D-78457 Konstanz, Germany*

Manuscript received November 26, 1999
Accepted for publication February 17, 2000

ABSTRACT

The complete nucleotide sequence (17,005 bp) of the mitochondrial genome of the caecilian *Typhlonectes natans* (Gymnophiona, Amphibia) was determined. This molecule is characterized by two distinctive genomic features: there are seven large 109-bp tandem repeats in the control region, and the sequence for the putative origin of replication of the L strand can potentially fold into two alternative secondary structures (one including part of the $tRNA^{Cys}$). The new sequence data were used to assess the phylogenetic position of caecilians and to gain insights into the origin of living amphibians (frogs, salamanders, and caecilians). Phylogenetic analyses of two data sets—one combining protein-coding genes and the other combining tRNA genes—strongly supported a caecilian + frog clade and, hence, monophyly of modern amphibians. These two data sets could not further resolve relationships among the coelacanth, lungfishes, and tetrapods, but strongly supported diapsid affinities of turtles. Phylogenetic relationships among a larger set of species of frogs, salamanders, and caecilians were estimated with a mitochondrial rRNA data set. Maximum parsimony analysis of this latter data set also recovered monophyly of living amphibians and favored a frog + salamander (Batrachia) relationship. However, bootstrap support was only moderate at these nodes. This is likely due to an extensive among-site rate heterogeneity in the rRNA data set and the narrow window of time in which the three main groups of living amphibians were originated.

LIVING amphibians (Lissamphibia) include three orders: Anura (frogs), Caudata (salamanders), and Gymnophiona (caecilians). Of these, the limbless caecilians, due to their secretive fossorial lifestyle, are the least known group (Carroll 1988). Much of their biology, including aspects of their ecology and behavior, is still poorly understood (Wilkinson and Nussbaum 1997). Caecilians have a distinctive morphology with an elongated limbless body (Wake 1997), a tropical distribution, and are probably of Gondwanan origin (the earliest appearance of this group in the fossil record dates back to the Lower Jurassic; Jenkins and Walsh 1993). Most of the 175 or so caecilian species currently recognized have burrowing habits, and few live in aquatic environments (Taylor 1968; Nussbaum and Wilkinson 1989; Duellman and Trueb 1994). Previous work on caecilian systematics has involved both morphological (Nussbaum and Wilkinson 1989; Wake 1993; Duellman and Trueb 1994) and molecular (Hedges *et al.* 1993; Feller and Hedges 1998) data. Nonetheless, the evolutionary and phylogenetic relationships of the six caecilian families (Ichthyophiidae, Caeciliaidae, Typhlonectidae, Scolecomorphidae, Rhinatrematidae, and Uraeotyphlinae) remain tentative (Nussbaum and Wilkinson 1989; Hedges *et al.* 1993; Wake 1993, 1997).

There still is a lack of consensus regarding living amphibian phylogenetic relationships (reviewed in Trueb and Cloutier 1991b; Carroll *et al.* 1999). The most widely accepted hypothesis, based on morphological data, supports the monophyletic origin in the Late Paleozoic (300 mya) of the three living orders in the class Amphibia and a sister-group relationship between Caudata and Anura (the Batrachia hypothesis) to the exclusion of the Gymnophiona (*e.g.*, Szarski 1962; Parsons and Williams 1963; Milner 1988; Trueb and Cloutier 1991a; Laurin and Reisz 1997). However, both the monophyly of living amphibians and the close phylogenetic relationships of frogs with salamanders are still debated. For instance, Jarvik (1942, 1980) has long maintained, on the basis of morphological analysis of the anterior head region, that tetrapods are polyphyletic in origin. According to this author, salamanders are derived from porolepiform fishes (an extinct group of lobe-finned fishes), whereas frogs are postulated to be descendant from osteolepiform fishes (another extinct group of lobe-finned fishes). Jarvik (1980) was unsure of the origin and phylogenetic relationships of caecilians. Reig (1964) also suggested that Lissamphibia are not a natural group and hypothesized an independent origin of caecilians from the other groups of living am-

Corresponding author: Rafael Zardoya, Museo Nacional de Ciencias Naturales, José Gutierrez Abascal, 2, 28006 Madrid, Spain. E-mail: mcnr154@pinar2.csic.es

phibians. According to this author, caecilian have affinities with microsauria (an extinct group of Lepospondyl amphibians). In a related hypothesis, Carroll and coworkers (Carroll and Currie 1975; Carroll and Holmes 1980; Carroll 1988) advocate a close phylogenetic relationship of both caecilians and salamanders to different lineages of microsaurs. According to this hypothesis, frogs would have evolved from temnospondyl amphibians (Carroll 1988). The monophyly of Lissamphibia and a close phylogenetic relationship between salamanders and caecilians is supported by several morphological (Bolt 1991) and molecular (Larson and Wilson 1989; Hedges and Maxson 1993; Feller and Hedges 1998) phylogenetic studies. A sister-group relationship between frogs and caecilians has apparently never been proposed.

Mitochondrial DNA (mtDNA) has been the molecular marker of choice in numerous phylogenetic analyses of vertebrate relationships and, hence, it was also expected to be appropriate for resolving the question on the origin of the Lissamphibia. Recently, several researchers have demonstrated that individual mitochondrial genes may show a poor performance in recovering the phylogenetic relationships among divergent vertebrate lineages that last shared a common ancestor in the Devonian (Cummings *et al.* 1995; Russo *et al.* 1996; Zardoya and Meyer 1996b). The phylogenetic analysis of complete mitochondrial genome sequences is anticipated to provide more reliable estimations of evolutionary relationships than those of individual genes. But even entire mitochondrial genome sequences are not guaranteed to discover correct topologies among ancient lineages that had a rapid origin. For a detailed discussion of the pitfalls of phylogenetic analyses based on complete mitochondrial genomes, such as among-site rate variation, rate differences among lineages, compositional biases, and saturation effects, see, *e.g.*, Cao *et al.* (1998), Zardoya *et al.* (1998), and Zardoya and Meyer (2000).

Only the complete mtDNA sequence of a single amphibian, the clawed frog *Xenopus laevis* (Anura), is currently available (Roe *et al.* 1985). Unfortunately, this sequence is suspected to contain a sizeable number of sequencing errors, which may reduce its phylogenetic reliability (Feller and Hedges 1998). Here, we present the gene order and complete nucleotide sequence of the mitochondrial genome of a caecilian with the aims of resolving the controversial phylogenetic position of Gymnophiona and of further examining the performance of complete mitochondrial genomes in reconstructing the phylogeny of relatively divergent vertebrate taxa with a rapid adaptive radiation origin (Cao *et al.* 1998; Zardoya *et al.* 1998). We are also presently in the process of sequencing the mitochondrial genome of a salamander, with the objective of establishing the origin and evolutionary relationships of all living orders of amphibians based on complete mitochondrial genome sequences.

MATERIALS AND METHODS

Isolation, PCR, cloning, and sequencing procedures: DNA was purified from the liver and kidneys of a single caecilian (*Typhlonectes natans*) specimen, as previously described (Zardoya *et al.* 1995). The isolated mtDNA was cleaved with the *Hin*dIII restriction enzyme. Six restriction fragments of 2.3, 1.8, 1.7, 1.1, 0.6, and 0.16 kb, which covered a region from 12S to COII genes, were cloned into pUC18. The rest of the molecule (9 kb) was amplified by PCR using the mtDNA extraction as DNA template source. PCR amplification with the primers caecilian COII F (5′-GGT GCC AAC CAC AGC TTT ATG CC-3′) and caecilian 12S R (5′-GTG TAG GGC TGG CCA TAA TTG AT-3′) was achieved by using the Expand long template PCR system of Boehringer Mannheim (Mannheim, Germany; PCR buffer 1) and the following program: 1 cycle of 2 min at 94°; 10 cycles of 10 sec at 94°; 30 sec at 58°, and 6 min at 68°; 20 cycles of 10 sec at 94°, 30 sec at 58°, and 6 min at 68° with an extension step that increases 20 sec for each cycle; and finally, 1 cycle of 5 min at 68°. This 9-kb fragment was used as DNA template source for the subsequent standard PCR amplification of shorter fragments to facilitate the sequencing strategy (see, *e.g.*, Zardoya and Meyer 1997 for standard PCR conditions). PCR fragments were cloned into the pGEM-T vector (Promega, Madison, WI).

Sequencing of the recombinant plasmids was performed with an automated DNA sequencer (Applied Biosystems, Foster City, CA; 373A *Stretch*) using the Taq Dye Deoxy Terminator cycle-sequencing kit (Applied Biosystems). Sequences were obtained using both M13 universal sequencing primers and several specific oligonucleotide primers. The sequences obtained from each clone averaged 450 bp in length, and each sequence overlapped the next contig by ~150 bp. In no case were differences in sequence observed between the overlapping regions.

Phylogenetic analyses: The complete nucleotide sequence of the caecilian mitochondrial genome was aligned to the homologous sequences of other tetrapods using CLUSTAL W (Thompson *et al.* 1994) and refined by eye. Gaps resulting from the alignment were treated as missing data. Ambiguous alignments, mainly in 5′ and 3′ ends of protein-coding genes, in the dihydrouridine (DHU) and TψC arms of the tRNAs, and in several highly variable regions of the rRNA genes, were excluded from the phylogenetic analyses (aligned sequences and exclusion sets are available at http://www.mncn.csic.es/investigacion/bbe/zardoy/primera.htm).

Three data sets were analyzed separately: (1) all protein-coding genes combined (except ND6 because it is encoded by the L strand and therefore has a very different base composition) at the amino acid level; (2) all 22 tRNA gene sequences combined; and (3) 12S and 16S rRNA genes combined. Each of these data sets was subjected to all three commonly used methods of phylogenetic inference [*i.e.*, maximum parsimony (MP), neighbor-joining (NJ), and maximum likelihood (ML)]. MP analyses were performed (PAUP* version 4.0b2a; Swofford 1998) using heuristic searches (TBR branch swapping; MULPARS option in effect) with 10 random stepwise additions of taxa. The tRNA and the rRNA data sets were analyzed using empirical transition:transversion (Ti:Tv) ratios (3:1 and 2:1, respectively). NJ (Saitou and Nei 1987) analyses of the nucleotide and amino acid alignments were based on HKY85 (Hasegawa *et al.* 1985) and mean character distance matrices (PAUP* version 4.0b2a; Swofford 1998), respectively. ML analyses were performed with PAUP* version 4.0b2a (Swofford 1998) and MOLPHY version 2.3 (Adachi and Hasegawa 1996b). In the DNA ML analyses (HKY 85 model; Hasegawa *et al.* 1985), transition:transversion ratios were optimized to maximize the likelihood, and empirical base frequen-

cies were used. In the protein ML analyses, a NJ tree was inferred as starting tree for a local rearrangement search for the ML tree with the mtREV model (Adachi and Hasegawa 1996a).

Robustness of the phylogenetic results was tested by bootstrap analyses (Felsenstein 1985; as implemented in PAUP* version 4.0b2a; Swofford 1998) with 100 pseudoreplications each, the quartet puzzling method (with 1000 puzzling steps; Strimmer and von Haeseler 1996), and the RELL (resampling of the estimated log-likelihood) local bootstrapping method (Kishino *et al.* 1990) with 10,000 pseudoreplications.

The complete mtDNA sequence of the caecilian *T. natans* has been deposited at the EMBL/GenBank data libraries under accession no. AF154051. A figure summarizing the main features of the caecilian genome is also available at http://www.mncn.csic.es/investigacion/bbe/zardoy/primera.htm.

RESULTS AND DISCUSSION

Genome organization: The complete nucleotide sequence of the L strand of the caecilian mt genome was determined. The mitochondrial molecule is 17,005 bp long and encodes for 13 protein-coding, 2 rRNA, and 22 tRNA genes. The organization of the mitochondrial genes and noncoding regions is identical to that of Eutherian mammals, frog, fishes, and sharks (see Table 1). The overall base composition of the L strand of the caecilian mtDNA is skewed, as in most metazoans, against guanine (A, 30%; T, 25%; C, 29%; and G, 16%). There are a total of 28 noncoding intergenic spacer nucleotides with a moderately strong A + C bias (68%). These regions are likely not subjected to strong selection, and this bias is generally interpreted as evidence of an asymmetrical directional mutation pressure (Jermiin *et al.* 1995).

Noncoding sequences: The control region in the caecilian mitochondrial genome is 1630 bp long (Table 1). This unusually large size is mainly due to the existence of seven 109-bp tandem repeats in the right domain of the control region, close to the 3′ end. Of these, six are perfect repeats, whereas the last differs by 12%. Two of the three conserved sequence blocks (CSB-2 and -3) that are involved in the initiation of the mtDNA synthesis (Walberg and Clayton 1981) can be identified in the right domain, upstream of the tandem repeats. A putative CSB-1 could be tentatively identified at position 16,027 but, as in frog (Roe *et al.* 1985), this motif is reduced to only five nucleotides (GACAT). The general absence of CSB-1 in lower vertebrates contrasts with its supposedly essential role in mammalian mitochondrial genomes (Sbisa *et al.* 1997). Furthermore, one termination-associated sequence (TAS-1; Doda *et al.* 1981), as well as the associated complementary motifs 5′-TACAT-3′ and 5′-ATGTA-3′ (Saccone *et al.* 1991) were found close to the 5′ end of the control region. Finally, an interrupted poly-C stretch with moderately high similarity to the CSB-2 is located in position 16,043.

As in most vertebrates, the putative origin of light strand replication (O_L) of the caecilian mitochondrial genome is located in the WANCY region (Seutin *et al.* 1994). This region is 34 bp long and has the potential to fold into a stem-loop secondary structure (Figure 1A). Furthermore, the caecilian L-strand synthesis is likely initiated in a stretch of thymines in the O_L loop (Wong and Clayton 1985). This condition is typical of tetrapods, whereas in fish the O_L loop contains a polypyrimidine tract (Zardoya *et al.* 1995; Zardoya and Meyer 1996a). Interestingly, more than half of the O_L is complementary to part of the sequence of the adjacent $tRNA^{Cys}$ and can potentially fold into an alternative stem-loop secondary structure (Figure 1B). A similar case has been described for the African lungfish (Zardoya and Meyer 1996a). However, both of these unusual L-strand origins have limited sequence similarity to each other.

Coding regions: The caecilian 12S and 16S rRNA genes are 934 and 1571 nucleotides long, respectively (Table 1). Most of the rRNA gene sequence here described showed minor differences (96% similarity) to that previously reported for the *T. natans* mitochondrial rRNA genes, which was obtained via separated PCR reactions (Feller and Hedges 1998). However, a 129-bp fragment (positions 895 to 1023 of our sequence), which comprised the end of the 12S rRNA and the beginning of the $tRNA^{Val}$, was found to be 34% different in both sequences. A search in GenBank revealed that this portion of the caecilian sequence reported by Feller and Hedges (1998) had 96% similarity to the homologous sequence of *X. laevis*.

As in other vertebrates, the caecilian mt genome contains a total of 22 tRNA genes interspersed between ribosomal RNA and protein-coding regions. These tRNA genes, which range in size from 66 to 74 nucleotides (Table 1), were recognized by their capability to fold into a canonical cloverleaf secondary structure (with the exception of $tRNA^{Ser(AGY)}$) and by the presence of specific anticodons (Figure 2). Size variability with respect to homologous vertebrate tRNAs was mainly detected in the DHU and TψC loops whereas the anticodon and acceptor stems were found to be more conserved in general.

All caecilian mtDNA protein-coding genes begin with an ATG start codon except for COI, which initiates with GTG, and ND1, which uses ATT as an initiation codon (see Table 1). The use of GTG as start codon for the COI gene is found in all chordates except mammals, which use ATG [a striking exception to this rule is frog (Roe *et al.* 1985) in which the COI gene starts with ATG]. Most caecilian open reading frames (ORFs) have incomplete stop codons, either a T (ND1, COII, COIII, ND3, ND4, and cyt b) or TA (ATPase 6), that presumably become functional by polyadenylation of the respective mRNAs. Four protein-coding genes use complete TAA stop codons (ATPase 8, ND4L, ND5, and ND6), and two end with TAG (ND2 and COI). This stop codon usage is similar to that reported for the

TABLE 1

Features of the caecilian mitochondrial genome

Feature	From	To	Size (bp)	Codon	
				Start	Stop
tRNA-Phe	1	66	66		
12S rRNA	67	1000	934		
tRNA-Val	1001	1065	65		
16S rRNA	1066	2636	1571		
tRNA-Leu (UUR)	2637	2709	73		
NADH 1	2710	3670	961	ATT	T- -
tRNA-Ile	3671	3741	71		
tRNA-Gln	3812	3743	70 (L)		
tRNA-Met	3812	3879	68		
NADH 2	3880	4911	1032	ATG	TAG
tRNA-Trp	4916	4983	68		
tRNA-Ala	5053	4983	69 (L)		
tRNA-Asn	5127	5055	73 (L)		
Ori L	5161	5128	34 (L)		
tRNA-Cys	5228	5162	67 (L)		
tRNA-Tyr	5298	5229	70 (L)		
COI	5300	6853	1554	GTG	TAG
tRNA-Ser (UCN)	6912	6842	71 (L)		
tRNA-Asp	6913	6981	69		
CO II	6984	7674	691	ATG	T- -
tRNA-Lys	7675	7748	74		
ATPase 8	7750	7914	165	ATG	TAA
ATPase 6	7908	8590	683	ATG	TA-
CO III	8591	9374	784	ATG	T- -
tRNA-Gly	9375	9442	68		
NADH 3	9444	9786	343	ATG	T- -
tRNA-Arg	9787	9853	67		
NADH 4L	9858	10154	297	ATG	TAA
NADH 4	10148	11522	1375	ATG	T- -
tRNA-His	11523	11590	68		
tRNA-Ser (AGY)	11591	11658	68		
tRNA-Leu (CUN)	11659	11729	71		
NADH 5	11730	13517	1788	ATG	TAA
NADH 6	14016	13501	516 (L)	ATG	TAA
tRNA-Glu	14085	14017	69 (L)		
Cyt *b*	14088	15228	1141	ATG	T- -
tRNA-Thr	15229	15296	68		
tRNA-Pro	15375	15308	68 (L)		
Control region	15376	17005	1630		

African lungfish (Zardoya and Meyer 1996a). Interestingly, AGR are not used as stop codons in the caecilian, whereas in both the African lungfish and the frog the ND5 gene ends with AGA.

There are two cases of reading-frame overlap on the same strand. ATPases 8 and 6 share 7 nucleotides, as is the case in the carp (Chang *et al.* 1994) and the lancelet (Spruyt *et al.* 1998; Boore *et al.* 1999), but not in other fishes, frog, reptiles, and birds (10 nucleotides) or mammals (~40 nucleotides). ND4L and ND4 overlap by 7 nucleotides, as in all chordates. Additionally, ND5 and ND6, which are encoded by different strands, share 17 nucleotides. This overlap is similar to that found in the sea lamprey (Lee and Kocher 1995) and considerably larger than that found in most vertebrates (~4 nucleotides). Finally, COI and $tRNA^{Ser(UCN)}$ genes share 12 nucleotides. A similar overlap has been described for the turtle mitochondrial genome (Zardoya and Meyer 1998).

Phylogenetic position of the caecilian: The deduced amino acid sequences of all the mitochondrial ORFs (with the exception of ND6, which is encoded by the L strand) were combined and aligned with 18 representative vertebrate homologous sequences: cod, *Gadus morhua* (X99772, Johansen and Bakke 1996); salmon, *Salmo salar* (U12143, C. D. Hurst, S. E. Bartlett, I. J. Bruce and W. S. Davidson, unpublished results); rainbow trout, *Oncorhynchus mykiss* (L29771, Zardoya *et al.* 1995); carp, *Cyprinus carpio* (X61010, Chang *et al.* 1994); goldfish, *Carassius auratus* (AB006953, Mura-

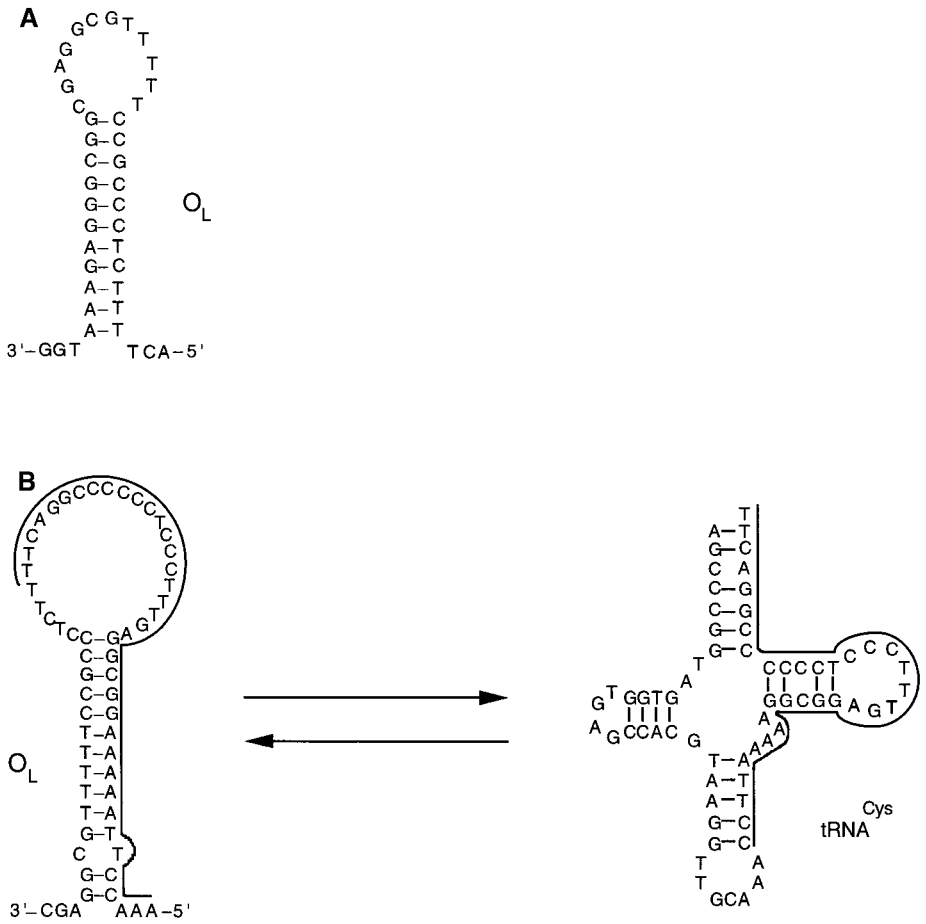

Figure 1.—Proposed alternative stem-loop secondary structures for the L-strand origin of replication. One of these configurations (B) partially shares nucleotides with the $tRNA^{Cys}$ (indicated by a line on its proposed cloverleaf secondary structure).

kami *et al.* 1998); loach, *Crossostoma lacustre* (M91245, Tzeng *et al.* 1992); African lungfish, *Protopterus dolloi* (L42813, Zardoya and Meyer 1996a); coelacanth, *Latimeria chalumnae* (U82228, Zardoya and Meyer 1997); clawed frog, *X. laevis* (M10217, Roe *et al.* 1985); skink, *Eumeces egregius* (AB016606, Kumazawa and Nishida 1999); painted turtle, *Chrysemys picta* (AF069423, Mindell *et al.* 1999); alligator, *Alligator mississippiensis* (Y13113, Janke and Arnason 1997); chicken, *Gallus gallus* (X52392, Desjardins and Morais 1990); ostrich, *Struthio camelus* (Y12025, Härlid *et al.* 1997); platypus, *Ornithorhynchus anatinus* (X83427, Janke *et al.* 1996); 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). A final data set of 3669 positions was gathered, of which 1135 were excluded due to ambiguity. Of the remaining sites, 50% were constant, and 852 were parsimony informative.

The phylogenetic analyses of the combined protein-coding gene data set using teleosts as outgroup (see Cao *et al.* 1998; Zardoya *et al.* 1998; Zardoya and Meyer 2000 for a discussion of outgroup selection for vertebrate phylogenetics using mitochondrial data) produced a single most parsimonious tree with 4370 steps and a consistency index (C.I.) of 0.65. In this tree, the caecilian grouped with the frog to form a monophyletic Amphibia (*e.g.*, Szarski 1962; Parsons and Williams 1963; Trueb and Cloutier 1991a; Laurin and Reisz 1997; Feller and Hedges 1998). This clade was supported by a 66% bootstrap value (Figure 3A). NJ (with mean character distances) and ML (with the mtREV model; −Ln likelihood = 30563.89) analyses also placed the caecilian as sister group of the frog with 72% (Figure 3B) and 95% (Figure 3C) bootstrap support, respectively.

The nucleotide sequences of the 22 tRNAs encoded by the caecilian mt genome were combined and aligned to the homologous sequences of the same 18 vertebrate taxa. Of the tRNA final data set of 1624 positions, 535 were excluded because of ambiguity. Of the remaining 1089 sites, 30% were invariant, and 575 were parsimony informative. When a 3:1 Ti:Tv weighting scheme was used and teleosts were defined as outgroup taxa, MP recovered one most parsimonious tree (3600 steps, C.I. = 0.50) in which a caecilian + frog clade (supported by a 82% bootstrap value) is identified as the sister group of the amniota (Figure 4A). NJ (HKY85 distances) and ML (HKY85 model) analyses of the tRNA data set also supported this clade (Figure 4, B and C) with high bootstrap values (86 and 88%, respectively).

Other vertebrate phylogenetic relationships: Besides the phylogenetic relationships within living amphibians, the recovered topologies reveal the existence of at least

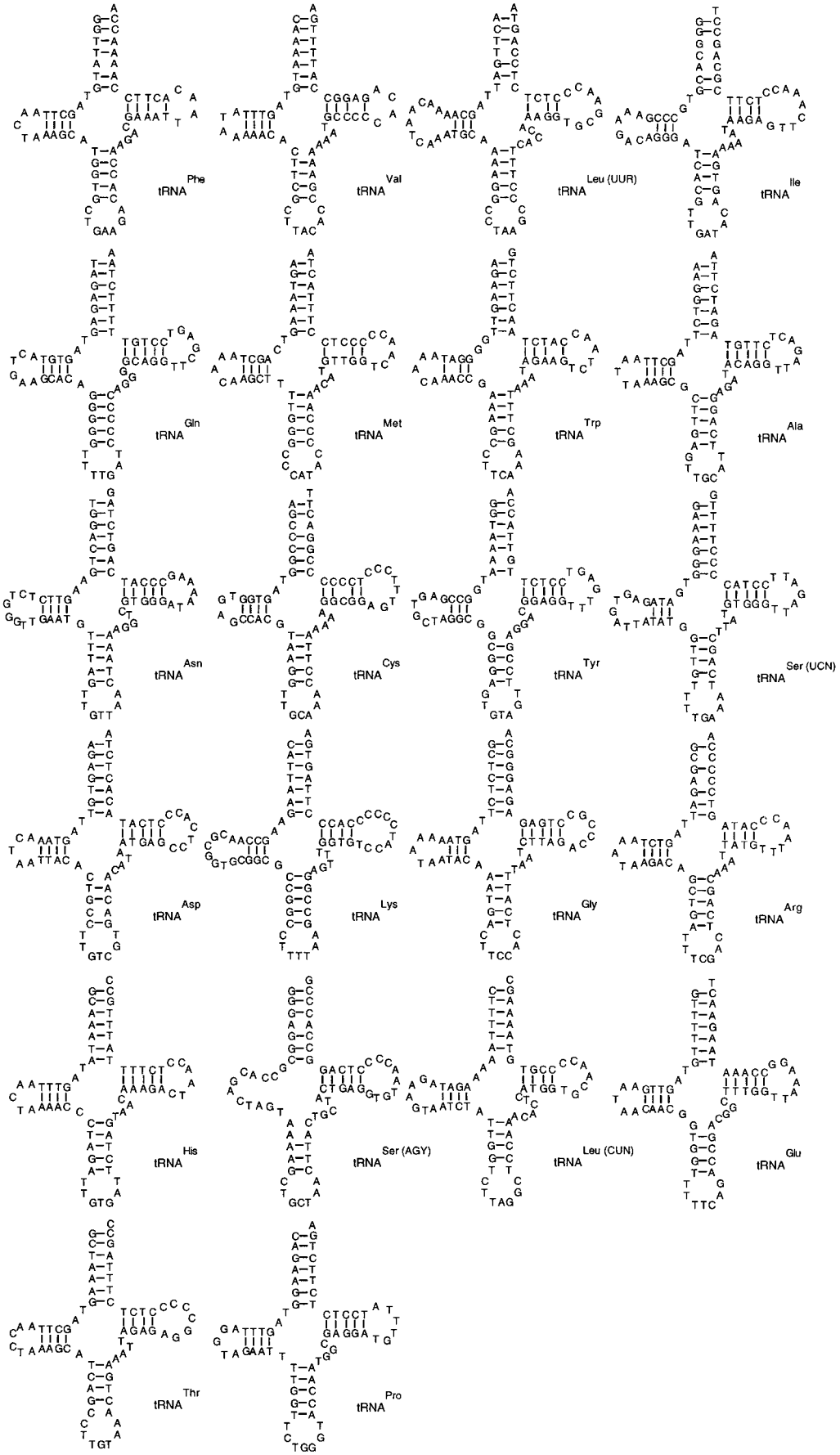

Figure 2.—Proposed cloverleaf secondary structures of the 22 tRNAs deduced from the complete sequence of the mitochondrial genome of a caecilian.

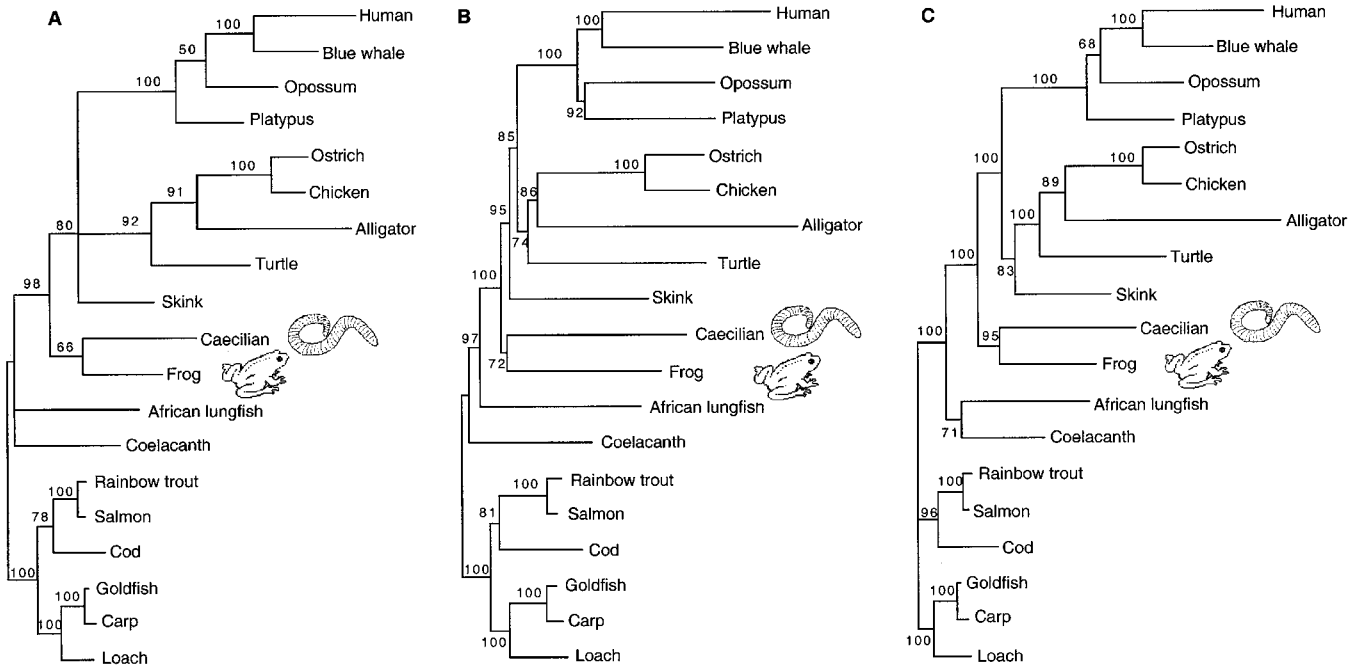

Figure 3.—Phylogenetic position of the caecilian based on a protein data set that combines all mitochondrial protein-coding genes (except ND6). The data set was analyzed with MP (A), NJ (B), and ML (C) methods of phylogenetic inference. Teleosts were used as outgroup taxa. MP and NJ topologies represent 50% majority rule bootstrap consensus trees based on 100 pseudoreplications. The ML tree shows RELL bootstraps based on 10,000 pseudoreplications. In all cases, a caecilian + frog clade is supported, *i.e.*, monophyly of living amphibians.

three controversial nodes in the vertebrate tree. The phylogenetic relationships between sarcopterygian fishes and tetrapods, the phylogenetic relationships among main groups of reptiles, and the phylogenetic position of monotremes with respect to marsupials and placentals cannot be resolved confidently (see Figures 3 and 4). Different methods of phylogenetic inference and different molecular data sets recover alternative hypotheses that can explain the phylogenetic position of these taxa (Figures 3 and 4). The lack of resolution at these nodes likely reflects rapid radiation events in the origin of the corresponding lineages (Carroll 1988).

In an apparent contradiction, previous analyses based on mitochondrial protein data had firmly supported a lungfish + tetrapod clade whereas mitochondrial tRNA evidence had strongly recovered a lungfish + coelacanth clade as sister group of tetrapods (Zardoya *et al.* 1998). In the present study, the addition of the caecilian amino acid and tRNA sequences to the phylogenetic analyses seems to affect the recovery of these groupings by both data sets. The protein data set is unable to resolve confidently the phylogenetic relationships of lungfish, coelacanth, and tetrapods (MP rendered the relationships among these taxa into a polytomy; NJ favored a lungfish + tetrapod clade; ML supported a lungfish + coelacanth grouping; see Figure 3) and the tRNA data set no longer supports with strong statistical confidence any of the alternative hypotheses (Figure 4). Hence, the phylogenetic analyses of both data sets are not anymore in contradiction but reflect a lack of phylogenetically informative positions in the two mitochondrial data sets due to the rapid origin of these vertebrate lineages back in the Devonian (Carroll 1997).

Similarly, the phylogenetic relationships among reptiles varied depending on the phylogenetic analysis performed (Figures 4 and 5). Previous phylogenetic analyses based on mitochondrial data supported a turtle + Archosauria (crocodiles + birds) clade to the exclusion of Lepidosauria (tuatara, snakes, and lizards; Zardoya and Meyer 1998; Kumazawa and Nishida 1999). However, nuclear evidence favored a turtle + crocodile clade (Hedges and Poling 1999). In the new analyses here presented, the protein data set places the turtle as sister group to the Archosauria whereas the tRNA data set supports a turtle + Lepidosauria grouping (Figures 4 and 5). A turtle + crocodile clade was never recovered based on mitochondrial data. Overall, our results strongly support diapsid affinities of turtles (Rieppel and deBraga 1996; Zardoya and Meyer 1998; Kumazawa and Nishida 1999). Although the skink (as representative of the Lepidosauria) is placed generally within the reptiles and basal to Archosauria (Figures 3 and 4), in the MP and NJ analyses of the protein data set, a putative basal position of Lepidosauria to the rest of the Amniota is suggested (Figure 3, A and B). Again, this uncertainty may reflect the rapid origin of the major reptile lineages.

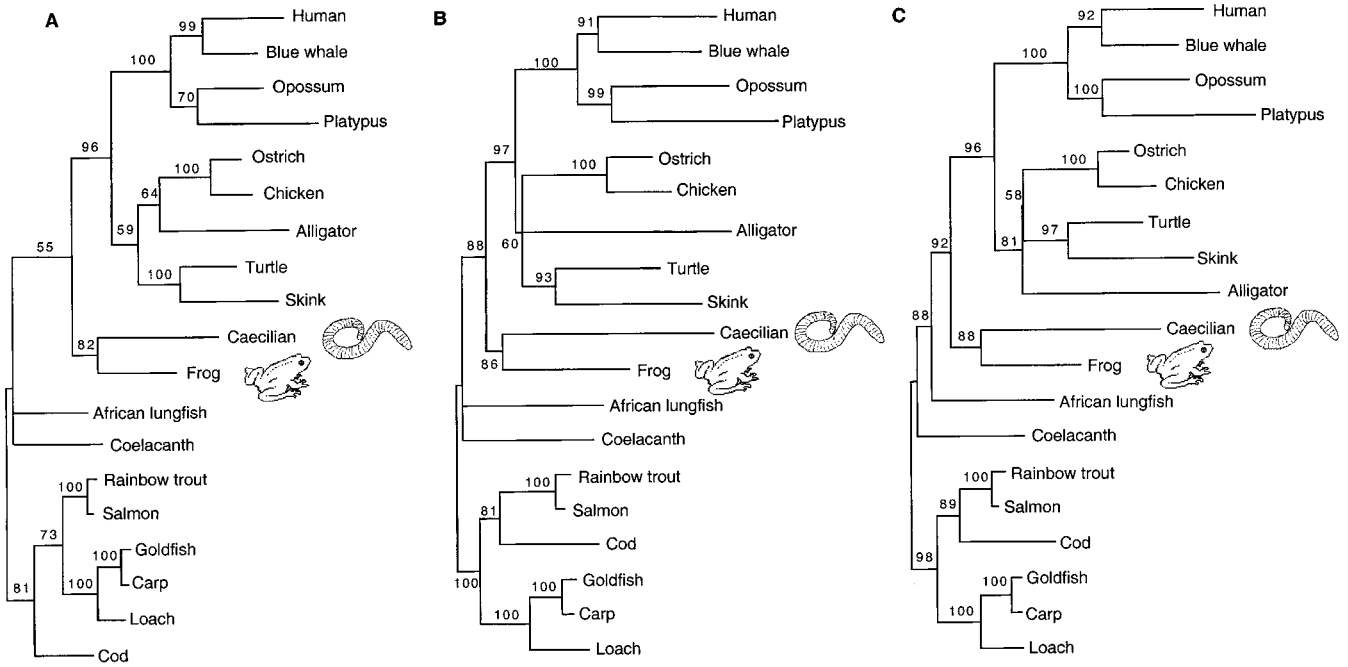

Figure 4.—Phylogenetic position of the caecilian based on a tRNA data set that combines all 22 mitochondrial tRNA genes. The data set was subjected to MP (A), NJ (B), and ML (C) analyses. Teleosts were used as outgroup taxa. MP and NJ topologies represent 50% majority rule bootstrap consensus trees based on 100 pseudoreplications. Support shown for the ML tree is based on 1000 puzzling steps. This data set strongly favors a caecilian + frog grouping, *i.e.*, monophyly of living amphibians.

Finally, the phylogenetic relationships near the root of the mammalian clade also appear to be difficult to resolve (Figures 4 and 5; Waddell *et al.* 1999). The tRNA data set strongly supports the Marsupionta (monotremes + marsupials) hypothesis (Janke *et al.* 1996). However, the protein data set favors moderately the traditional Theria (marsupials + placentals) hypothesis (with the exception of the NJ analysis). This latter result was robust to changes in the sampling of the species representing the main groups of vertebrates (not shown). The resolution of this controversy likely requires the collection and analysis of nuclear data.

Lissamphibia relationships based on mitochondrial rRNA data: To test the phylogenetic relationships among modern amphibians, the complete nucleotide sequences of the caecilian 12S and 16S rRNA genes were aligned to the homologous sequences of another two caecilians (*Ichthyopis bannanicus* and *Epicrionops* sp.; Feller and Hedges 1998), three frogs (*X. laevis*, *Eleutherodactylus cuneatus*, and *Rana pipiens*; Feller and Hedges 1998), and three salamanders (*Siren intermedia*, *Ambystoma mexicanum*, and *Plethodon yonahlossee*; Feller and Hedges 1998). The rRNA data set consisted of 28 taxa (including representatives of the major tetrapod lineages, lobe-finned fishes, and teleosts as outgroup) and 2872 positions. Of these 986 were excluded due to ambiguity and 994 were parsimony informative.

The MP analysis using a 2:1 transversion:transition weighting and teleosts as outgroup recovered a single most parsimonious tree (8216 steps, C.I. = 0.37) that supports monophyly of living amphibians (with a 70% bootstrap value; Figure 5A). The monophyly of the caecilian, salamander, and frog lineages is supported clearly (100% bootstrap values), and a frog + salamander clade is suggested (54% bootstrap support; Figure 5A). This result is in agreement with the Batrachia hypothesis, which is based on morphological data (*e.g.*, Szarski 1962; Parsons and Williams 1963; Trueb and Cloutier 1991a; Laurin and Reisz 1997) and is in disagreement with the results of a related phylogenetic analysis also based on rRNA data that supported a salamander + caecilian relationship (Feller and Hedges 1998). Two important methodological differences between both studies may be responsible for this disagreement: first, the choice of outgroup (teleosts in this work *vs.* amniota in Feller and Hedges 1998) and, second, the representation of amniote lineages (several species per lineage in this work *vs.* a single representative per lineage in Feller and Hedges 1998). NJ and ML analyses of the rRNA data set recovered tree topologies in which the monophyly of living amphibians and the relationships between frogs, salamanders, and caecilians are unresolved (Figure 5B). This may not be surprising since the nodes connecting these lineages are very short and lead to rather long branches. This combination of features is particularly difficult to resolve because it would require a hypothetical molecular marker having different rates of substitution during its history—a fast one during the rapid origin of lineages and a slow one thereafter. Additionally, the mitochondrial rRNA se-

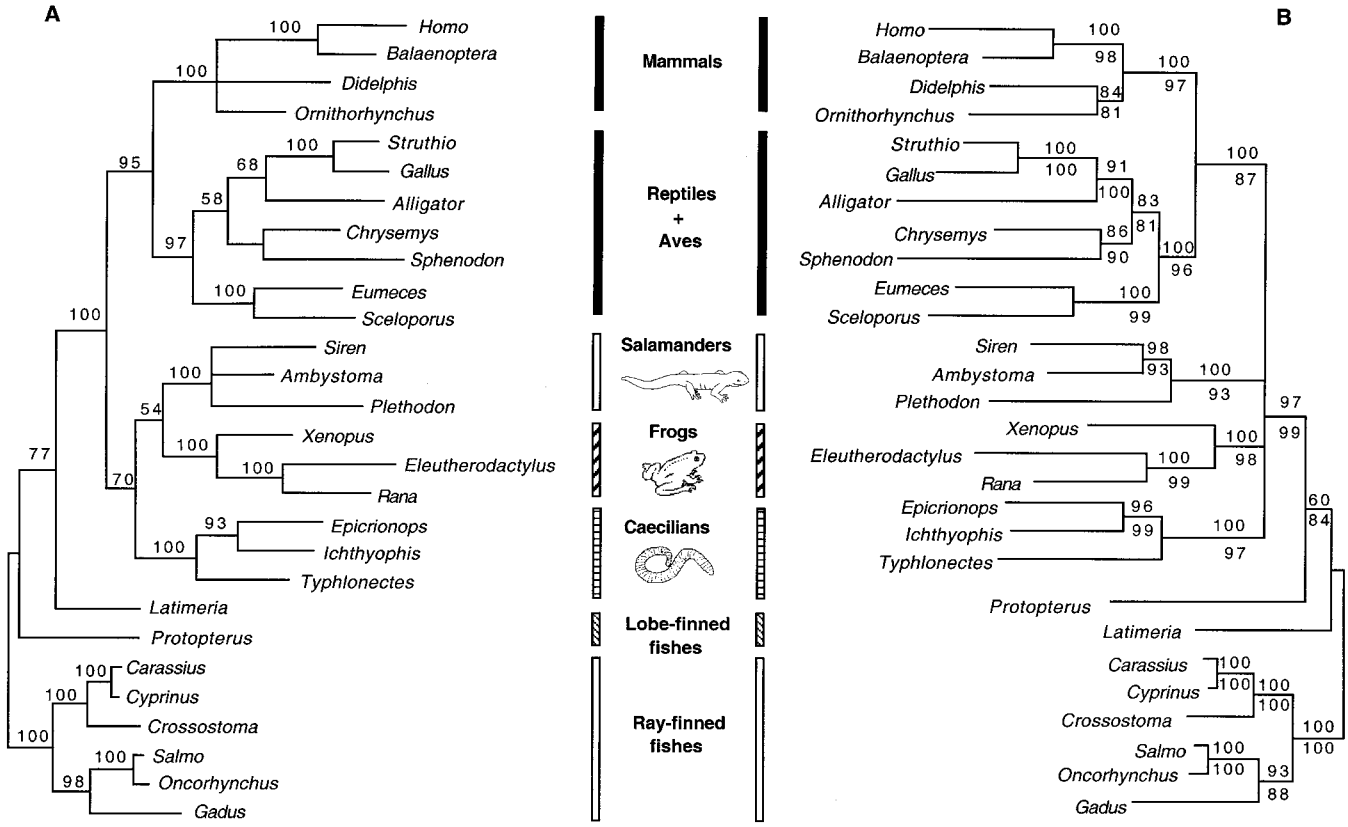

Figure 5.—Phylogenetic relationships among living amphibian orders based on mitochondrial rRNA genes. A data set combining the two rRNA mitochondrial genes (12S and 16S) was analyzed with (A) MP and (B) NJ and ML phylogenetic methods (top and bottom numbers along branches, respectively). Numbers shown represent bootstrap values from 100 pseudoreplicates (MP and NJ) or 1000 puzzling steps (ML). Nodes with a bootstrap value below 50% were collapsed into polytomies. Teleosts were used as outgroup taxa. Monophyly of frogs, salamanders, and caecilians is supported strongly by this data set. A frog + salamander sister-group relationship (Batrachia) is supported by the MP analysis. Overall, interrelationships among the three modern amphibian lineages remain tentative.

quence data set shows a high degree of among-site rate variation between different regions of the 12S and 16S rRNA molecules. As a consequence, only few sites are phylogenetically informative at any particular level of divergence (Olsen and Woese 1993). Interestingly, MP, NJ, and ML analyses of the rRNA data set support a basal position of Lepidosauria within reptiles and recover a tuatara + turtle clade as sister group to Archosauria.

In conclusion, the monophyly of living amphibians is largely supported by mitochondrial evidence. However, the separation of the different lineages of living orders of amphibians (*i.e.*, Anura, Caudata, and Gymnophiona) apparently took place within a narrow window of time (the oldest fossils of the three groups are all from the Jurassic, 213–144 mya; see, *e.g.*, Wake 1997). Hence, the number of phylogenetically informative positions that can resolve these relationships is rather limited, which results in only moderate support for the Batrachia hypothesis with the rRNA data set. It is expected that the future completion of the nucleotide sequence of a salamander mitochondrial genome will be crucial in determining the phylogenetic relationships among all living amphibian orders and in clarifying other controversial phylogenetic relationships of related vertebrates, *e.g.*, lobe-finned fishes and reptiles.

We thank Edward Malaga-Trillo for help with the sequencing of a portion of the mitochondrial genome. Dan Graur and three anonymous reviewers gave insightful comments on the manuscript. R.Z. was sponsored by a postdoctoral contract of the Ministerio de Educacion y Cultura of Spain. This work received partial financial support from grants from the Lion Foundation, the Deutsche Forschungsgemeinschaft, University of Konstanz, and Fond der Chemischen Industrie to A.M.

LITERATURE CITED

Adachi, J., and M. Hasegawa, 1996a Model of amino acid substitution in proteins encoded by mitochondrial DNA. J. Mol. Evol. **42:** 459–468.

Adachi, J., and M. Hasegawa, 1996b *MOLPHY Version 2.3: Programs for Molecular Phylogenetics Based on Maximum Likelihood.* The Institute of Statistical Mathematics, Tokyo.

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.

Bolt, J. R., 1991 Lissamphibian origins, pp. 194–222 in *Origins of the Major Groups of Tetrapods: Controversies and Consensus,* edited by H. P. Schultze and L. Trueb. Cornell University Press, Ithaca, NY.

Boore, J. L., L. L. Daehler and W. M. Brown, 1999 Complete

sequence, gene arrangement, and genetic code of mitochondrial DNA of the cephalochordate *Branchiostoma floridae* (Amphioxus). Mol. Biol. Evol. **16:** 410–418.

Cao, Y., P. J. Waddell, N. Okada and M. Hasegawa, 1998 The complete mitochondrial DNA sequence of the shark *Mustelus manazo*: evaluating rooting contradictions to living bony vertebrates. Mol. Biol. Evol. **15:** 1637–1646.

Carroll, R. L., 1988 *Vertebrate Paleontology and Evolution*. Freeman, New York.

Carroll, R. L., 1997 *Patterns and Processes of Vertebrate Evolution*. Cambridge University Press, Cambridge, UK.

Carroll, R. L., and P. J. Currie, 1975 Microsaurs as possible apodan ancestors. J. Linn. Soc. Lond. Zool. **57:** 229–247.

Carroll, R. L., and R. Holmes, 1980 The skull and jaw musculature as guides to the ancestry of salamanders. J. Linn. Soc. Lond. Zool. **68:** 1–40.

Carroll, R. L., A. Kuntz and K. Albright, 1999 Vertebral development and amphibian evolution. Evol. Dev. **1:** 36–48.

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.

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

Doda, J. N., C. T. Wright and D. A. Clayton, 1981 Elongation of displacement loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. USA **78:** 6116–6120.

Duellman, W. E., and L. Trueb, 1994 *Biology of Amphibians*. Johns Hopkins University Press, Baltimore.

Feller, A. E., and S. B. Hedges, 1998 Molecular evidence for the early history of living amphibians. Mol. Phylogenet. Evol. **9:** 509–516.

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

Härlid, A., A. Janke and U. Arnason, 1997 The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. Mol. Biol. Evol. **14:** 754–761.

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., and L. R. Maxson, 1993 A molecular perspective on lissamphibian phylogeny. Herpetol. Monogr. **7:** 27–42.

Hedges, S. B., and L. L. Poling, 1999 A molecular phylogeny of reptiles. Science **283:** 998–1001.

Hedges, S. B., R. A. Nussbaum and L. R. Maxson, 1993 Caecilian phylogeny and biogeography inferred from mitochondrial DNA sequences of the 12SrRNA and 16SrRNA genes (Amphibia: Gymnophiona). Herpetol. Monogr. **7:** 64–76.

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., and U. Arnason, 1997 The complete mitochondrial genome of *Alligator mississipiensis* and the separation between recent Archosauria (birds and crocodiles). Mol. Biol. Evol. **14:** 1266–1272.

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.

Janke, A., N. J. Gemmell, G. Feldmaier-Fuchs, A. von Haeseler and S. Paabo, 1996 The mitochondrial genome of a monotreme—the platypus (*Ornithorhynchus anatinus*). J. Mol. Evol. **42:** 153–159.

Jarvik, E., 1942 On the structure of the snout of crossopterygians and lower gnathostomes in general. Zool. Bidr. Upps. **21:** 235–675.

Jarvik, E., 1980 *Basic Structure and Evolution of Vertebrates*. Academic Press, London.

Jenkins, F. A., and D. M. Walsh, 1993 An early Jurassic caecilian with limbs. Nature **365:** 246–249.

Jermiin, L., D. Graur and R. H. Crozier, 1995 Evidence from analyses of intergenic regions for strand-specific directional mutation pressure in metazoan mtDNA. Mol. Biol. Evol. **12:** 558–563.

Johansen, S., and I. Bakke, 1996 The complete mitochondrial DNA sequence of Atlantic cod, *Gadus morhua*: relevance to taxonomic studies among cod-fishes. Mol. Mar. Biol. **5:** 203–214.

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, 1999 Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. Mol. Biol. Evol. **16:** 784–792.

Larson, A., and A. C. Wilson, 1989 Patterns of ribosomal RNA evolution in salamanders. Mol. Biol. Evol. **6:** 131–154.

Laurin, M., and R. Reisz, 1997 A new perspective on tetrapod phylogeny, pp. 9–59 in *Amniote Origins*, edited by S. S. Suemida and K. L. Martin. Academic Press, New York.

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.

Milner, A. R., 1988 The relationships and origin of living amphibians, pp. 59–102 in *The Phylogeny and Classification of the Tetrapods*, edited by M. J. Benton. Clarendon Press, Oxford.

Mindell, D. P., M. D. Sorenson, D. E. Dimcheff, M. Hasegawa, J. C. Ast *et al.*, 1999 Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. Syst. Biol. **48:** 138–152.

Murakami, M., Y. Yamashita and H. Fujitani, 1998 The complete sequence of mitochondrial genome from a gynogenetic triploid 'ginbuna' (*Carassius auratus langsdorfi*). Zool. Sci. **15:** 335–337.

Nussbaum, R. A., and M. Wilkinson, 1989 On the classification and phylogeny of caecilians (Amphibia: Gymnophiona), a critical review. Herpetol. Monogr. **3:** 1–42.

Olsen, G. J., and C. R. Woese, 1993 Ribosomal RNA: a key to phylogeny. FASEB J. **7:** 113–123.

Parsons, T. S., and E. E. Williams, 1963 The relationships of the modern Amphibia: a re-examination. Q. Rev. Biol. **38:** 26–53.

Reig, O., 1964 El problema del origen monofilético o polifilético de los anfibios, con consideraciones sobre las relaciones entre Anuros, Urodelos y Apodos. Ameghiniana **3:** 191–211.

Rieppel, O., and M. deBraga, 1996 Turtles as diapsid reptiles. Nature **384:** 453–455.

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.

Saccone, C., G. Pesole and E. Sbisa, 1991 The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. J. Mol. Evol. **33:** 83–91.

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

Sbisa, E., F. Tanzariello, A. Reyes, G. Pesole and C. Saccone, 1997 Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene **205:** 125–140.

Seutin, G., B. Franz Lang, D. P. Mindell and R. Morais, 1994 Evolution of the WANCY region in amniote mitochondrial DNA. Mol. Biol. Evol. **11:** 329–340.

Spruyt, N., C. Delarbre, G. Gachelin and V. Laudet, 1998 Complete sequence of the amphioxus (*Branchiostoma lanceolatum*) mitochondrial genome: relations to vertebrates. Nucleic Acids Res. **26:** 3279–3285.

Strimmer, K., and A. von Haeseler, 1996 Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13:** 964–969.

Swofford, D. L., 1998 *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0*. Sinauer Associates, Sunderland, MA.

Szarski, H., 1962 The origin of the Amphibia. Q. Rev. Biol. **37:** 189–241.

Taylor, E. H., 1968 *The Caecilians of the World: A Taxonomic Analysis*. University of Kansas Press, Lawrence, KS.

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.

Trueb, L., and R. Cloutier, 1991a A phylogenetic investigation of the inter- and intrarelationships of the Lissamphibia (Amphibia: Temnospondyli), pp. 223–313 in *Origins of the Major Groups of Tetrapods: Controversies and Consensus*, edited by H. P. Schultze and L. Trueb. Cornell University Press, Ithaca, NY.

Trueb, L., and R. Cloutier, 1991b Toward an understanding of the amphibians: two centuries of systematic history, pp. 175–193 in *Origins of the Major Groups of Tetrapods: Controversies and Consensus*, edited by H. P. Schultze and L. Trueb. Cornell University Press, Ithaca, NY.

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.

Waddell, P. J., Y. Cao, J. Hauf and M. Hasegawa, 1999 Using novel phylogenetic methods to evaluate mammalian mtDNA, including amino acid-invariant sites-LogDet plus site stripping, to detect internal conflicts in the data, with special reference to the positions of the hedgehog, armadillo, and elephant. Syst. Biol. **48:** 31–53.

Wake, M. H., 1993 Non-traditional characters in the assessment of caecilian phylogenetic relationships. Herpetol. Monogr. **7:** 42–55.

Wake, M. H., 1997 Amphibian locomotion in evolutionary time. Zoology **100:** 141–151.

Walberg, M. W., and D. A. Clayton, 1981 Sequence and properties of the human KB cell and mouse L cell D-Loop regions of mitochondrial DNA. Nucleic Acids Res. **9:** 5411–5421.

Wilkinson, M., and R. A. Nussbaum, 1997 Comparative morphology and evolution of the lungless caecilian Atretochoana eiselti (Taylor) (Amphibia: Gymnophiona: Typhlonectidae). Biol. J. Linn. Soc. **62:** 39–109.

Wong, T. W., and D. A. Clayton, 1985 In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. Cell **42:** 951–958.

Zardoya, R., and A. Meyer, 1996a 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.

Zardoya, R., and A. Meyer, 1996b Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. Mol. Biol. Evol. **13:** 933–942.

Zardoya, R., and A. Meyer, 1997 The complete DNA sequence of the mitochondrial genome of a "living fossil," the coelacanth (*Latimeria chalumnae*). Genetics **146:** 995–1010.

Zardoya, R., and A. Meyer, 1998 Complete mitochondrial genome suggests diapsid affinities of turtles. Proc. Natl. Acad. Sci. USA **95:** 14226–14231.

Zardoya, R., and A. Meyer, 2000 Vertebrate phylogeny: limits of inference of mitochondrial genome and nuclear rDNA sequence data due to an adverse phylognetic signal/homoplasy ratio. In *Major Events in Early Vertebrate Evolution: Paleontology, Phylogeny, and Development*, edited by Per Ahlberg. Taylor & Francis, London.

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., Y. Cao, M. Hasegawa and A. Meyer, 1998 Searching for the closest living relative(s) of tetrapods through evolutionary analyses of mitochondrial and nuclear data. Mol. Biol. Evol. **15:** 506–517.

Communicating editor: N. Takahata