

## Mitogenomic Perspectives on the Origin and Phylogeny of Living Amphibians

PENG ZHANG, HUI ZHOU, YUE-QIN CHEN, YI-FEI LIU, AND LIANG-HU QU

*Key Laboratory of Gene Engineering of the Ministry of Education, Biotechnology Research Center, Zhongshan University, Guangzhou 510275, P. R. China; E-mail: lsbrc04@zsu.edu.cn (L.-H. Q.)*

**Abstract.**—Establishing the relationships among modern amphibians (lissamphibians) and their ancient relatives is necessary for our understanding of early tetrapod evolution. However, the phylogeny is still intractable because of the highly specialized anatomy and poor fossil record of lissamphibians. Paleobiologists are still not sure whether lissamphibians are monophyletic or polyphyletic, and which ancient group (temnospondyls or lepospondyls) is most closely related to them. In an attempt to address these problems, eight mitochondrial genomes of living amphibians were determined and compared with previously published amphibian sequences. A comprehensive molecular phylogenetic analysis of nucleotide sequences yields a highly resolved tree congruent with the traditional hypotheses (Batrachia). By using a molecular clock-independent approach for inferring dating information from molecular phylogenies, we present here the first molecular timescale for lissamphibian evolution, which suggests that lissamphibians first emerged about 330 million years ago. By observing the fit between molecular and fossil times, we suggest that the temnospondyl-origin hypothesis for lissamphibians is more credible than other hypotheses. Moreover, under this timescale, the potential geographic origins of the main living amphibian groups are discussed: (i) advanced frogs (neobatrachians) may possess an Africa-India origin; (ii) salamanders may have originated in east Asia; (iii) the tropic forest of the Triassic Pangaea may be the place of origin for the ancient caecilians. An accurate phylogeny with divergence times can be also helpful to direct the search for “missing” fossils, and can benefit comparative studies of amphibian evolution. [Amphibian; mitochondrial genome; molecular dating; phylogeny; timescale.]

The Amphibia first appeared in the late Devonian and then became the dominant land vertebrates in the following Carboniferous. Although most ancient amphibian lineages became extinct before the Jurassic, some amphibian lineages survived and are represented today by the three distinctly different groups of living amphibians (Lissamphibia), the Anura (frogs), Caudata (salamanders), and Gymnophiona (caecilians). However, a large gap, both in time and in morphology, separates the modern amphibians from the varied Paleozoic amphibians. Thus the questions remain, “When did the modern amphibians originate and who were their closest Paleozoic ancestors?” The highly specialized anatomy of amphibians makes it difficult to find unambiguous clues to their ancestry. Moreover, early lissamphibian fossils are very rare. It is also difficult to establish a convincing evolutionary pattern from ancient amphibians to modern ones based only on the relatively sparse fossil record. Paleontologists are still debating whether the lissamphibians are monophyletic or polyphyletic (Milner, 1988; Trueb and Cloutier, 1991; Carroll and Holmes, 1980; Bolt, 1991). Even if the monophyly of the Lissamphibia is accepted, it is still controversial whether the extinct temnospondyls or the lepospondyls are the sister-group of Lissamphibia (Milner, 1988; Trueb and Cloutier, 1991; Laurin and Reisz, 1997).

Dating information extracted from molecular data is an alternative method to improve our hypotheses of branching order when fossil records are insufficient. However, a clocklike substitution rate is often required under the traditional dating method, but this assumption is usually unrealistic. Fortunately, newly developed molecular dating techniques (Thorne and Kishino, 2002; Sanderson, 2002) enable us to infer reliable dating information among lineages with different evolutionary rates.

To infer an accurate time frame for modern amphibians, a convincing phylogeny of lissamphibians (that is

to say, finding the relationships among the three living orders) is first required. However, there is no generally accepted consensus regarding the phylogenetic relationships among salamanders, caecilians, and frogs. Much morphological evidence (Milner, 1988; Trueb and Cloutier, 1991) favors the close relationship between frogs and salamanders, whereas most molecular phylogenetic analyses based on rRNA genes (Hedges et al., 1990; Larson, 1991; Hay et al., 1995; Feller and Hedges, 1998) suggest a caecilian-salamander sister-group. Even studies using complete mitochondrial genomes have failed to reach an agreement (Zardoya and Meyer, 2001; Zhang et al., 2003b). In an attempt to resolve this problem, we sequenced eight additional mitochondrial genomes of living amphibians. By combining these sequences with previously published amphibian mitochondrial genomes, we have been able to bring more data to bear on amphibian phylogeny than have been used in previous studies.

### MATERIALS AND METHODS

#### *Sequence Data Preparation*

A total of eight mitochondrial genomes of living amphibians were sequenced. The amphibian species were carefully selected so that every major amphibian group contained at least two species (in an effort to reduce long-branch attraction artefacts). We also tried to include more primitive lineages (whenever possible) to make subsequent molecular dating more accurate. With all samples, mitochondrial DNA was amplified in two fragments (A and B) longer than 5 kb (to avoid amplifying nuclear copies) by using the Long-and-Accuracy PCR method (LA-PCR). Methods for DNA extraction, amplification, and sequencing are given elsewhere (see Zhang et al., 2003a, for details). The primers used in the LA-PCR amplification are given in Table 1. To perform

TABLE 1. List of the primers used in the LA-PCR and their usage.

Primer ID	Sequence 5' → 3'	Species	Fragment A	Fragment B
LX16S1	GGTTTACGACCTCGATGTTGGATCA	<i>Ichthyophis bannanicus</i>	LX16S1-ICHCOIIR	ICHCOII-LX16S1R
LX16S1R	GACCTGGATTACTCCGGTCTGAACTC	<i>Paramesotriton hongkongensis</i>	LX16S1-PARCOIIR	PARCOII-LX16S1R
LX9844	AGCGACAGCCTTTTAAAGCTGTAGA	<i>Bombina fortinuptialis</i>	LX16S1-LX11932H	LX9844-LX16S1R
LX11932H	AACCATAATTWGYTGAGCCGAAAT	<i>Polypedates megacephalus</i>	LX16S1-LX11932H	LX9844-LX16S1R
KALCOII	CCCATCTAATGACCTTACACCA	<i>Kaloula pulchra</i>	LX16S1-KALCOIIR	KALCOII-LX16S1R
KALCOIIR	TGTGGTTAGCTCCGCAGATT	<i>Microhyla heymonsi</i>	LX16S1-LX11932H	LX9844-LX16S1R
ICHCOII	AAGTGGACAATCGGATGGTGGT	<i>Bufo melanostictus</i>	LX16S1-LX11932H	LX9844-LX16S1R
ICHCOIIR	TTAAGCGCCCTGGAATAGCATCTG	<i>Hyla chinensis</i>	LX16S1-LX11932H	LX9844-LX16S1R
PARCOII	GCGACTAAACCAAACAACCTTC			
PARCOIIR	TACACCTATAGATGGGACAGTTCA			

a comprehensive phylogenetic analysis, another 12 representative vertebrate complete mitochondrial genomes (6 amphibians, 1 bird, 1 crocodile, 2 mammals, 1 lobe-finned fish, and 2 ray-finned fish) were retrieved from the Genbank. The details for all sequences used in this study are given in Table 2. The ray-finned fishes (carp, *Cyprinus carpio*; dogfish, *Scyliorhinus canicula*) were used as outgroup species.

TABLE 2. List of species used in this study, along with GenBank accession numbers and references.

Species	Accession no.	Reference
<b>Gymnophiona</b>		
<i>Typhlonectes natans</i>	NC_002471	Zardoya and Meyer (2000)
<i>Ichthyophis bannanicus</i>	AY458594	This study
<b>Caudata Cryptobranchoidea</b>		
<i>Andrias davidianus</i>	AJ492192	Zhang et al. (2003a)
<i>Ranodon sibiricus</i>	AJ419960	Zhang et al. (2003b)
<b>Caudata Salamandroidea</b>		
<i>Paramesotriton hongkongensis</i>	AY458597	This study
<i>Mertensiella luschni</i>	NC_002756	Zardoya and Meyer (2001)
<b>Anura "Archaeobatrachia"<sup>†</sup></b>		
<i>Bombina fortinuptialis</i> <sup>*</sup>	AY458591	This study
<i>Xenopus laevis</i>	NC_001573	Roe et al. (1985)
<b>Anura Neobatrachia Ranoidea</b>		
<i>Rana nigromaculata</i>	NC_002805	Sumida et al. (2001)
<i>Polypedates megacephalus</i>	AY458598	This study
<b>Anura Neobatrachia Microhyloidea</b>		
<i>Kaloula pulchra</i>	AY458595	This study
<i>Microhyla heymonsi</i>	AY458596	This study
<b>Anura Neobatrachia Bufonoidea</b>		
<i>Bufo melanostictus</i>	AY458592	This study
<i>Hyla chinensis</i>	AY458593	This study
<b>Amniota Diapsida</b>		
<i>Gallus gallus</i>	NC_001323	Desjardins and Morais (1990)
<i>Alligator mississippiensis</i>	NC_001922	Janke and Arnason (1997)
<b>Amniota Mammalia</b>		
<i>Bos Taurus</i>	NC_001567	Anderson et al. (1982)
<i>Homo sapiens</i>	NC_001807	Andrews et al. (1999)
<b>Dipnoi</b>		
<i>Protopterus dolloi</i>	NC_001708	Zardoya and Meyer (1996a)
<b>Actinopterygii</b>		
<i>Scyliorhinus canicula</i>	NC_001950	Delarbre et al. (1998)
<i>Cyprinus carpio</i>	NC_001606	Chang et al. (1994)

\*Complete mtDNA sequences except for a portion of the control region.

<sup>†</sup>Archaeobatrachia is not a natural group.

### Phylogenetic Analysis

Multiple alignments were prepared for both rRNAs and for 12 protein-coding genes (ND6 was excluded) by using Clustal X (Thompson et al., 1997) at default settings. All third codon positions of protein-coding genes were excluded from our analyses. To avoid bias in refining alignments, we used Gblocks (Castresana, 2000) to extract regions of defined sequence conservation from the alignments. We used stringent parameter settings: minimum number of sequences for a conserved position 17; minimum number of sequences for a flanking position 21; maximum number of contiguous nonconserved positions 8; minimum length of a block 10). Finally, all alignments (12 proteins and 2 rRNAs) were combined. One species (*Polypedates megacephalus*) lacked ATP8 and ND5 genes in its mitochondrial genome. Thus, gaps were added to the corresponding alignments and treated as missing data in the following analyses. Amino acid sequences were not used in phylogenetic analysis, as the resulting trees were not well resolved.

DNA molecular phylogenies were derived by using PAUP\*, version 4.0b8 (Swofford, 2001), with maximum parsimony (MP), maximum likelihood (ML), and neighbor-joining (NJ) analyses. MP analyses were performed by using heuristic searches (TBR branch swapping; MULPARS option in effect) with 100 random-addition sequences. All sites were given equal weight in the parsimony analysis. Support for internal branches in the parsimony analysis was assessed by using 1000 bootstrap pseudoreplicates, with 10 random-addition sequences performed in each replicate. In the ML analyses, the nucleotide substitution model selection was carried out by using ModelTest version 3.06 (Posada and Crandall, 1998). A GTR+I+ $\Gamma$  model was selected by using Akaike Information Criterion (AIC = 137833.6). Heuristic ML analyses were conducted with TBR branch swapping (10 random addition sequences), and the substitution model parameter values were set according to ModelTest's results. NJ analyses were based on ML distance matrices taking account of the heterogeneity of rates among sites with a discrete gamma distribution. Nonparametric bootstrapping analyses for NJ were based on 1000 pseudoreplicate data sets. The Bayesian analysis (BA) was performed with MrBayes (Huelsenbeck and Ronquist, 2001) by using a setting corresponding to GTR+I+ $\Gamma$  model. Model parameter

values were treated as unknown and were estimated in each analysis. Random starting trees were used, and analyses were run for 1 million generations, sampling the Markov chains every 100 generations. Bayesian posterior probabilities were then calculated from the sample points after the MCMC algorithm started to converge. To ensure that our analyses were not trapped in local optima, three independent Markov chain Monte Carlo (MCMC) runs were performed. Topologies and posterior clade probabilities from different runs were compared for congruence (Huelsenbeck and Imennov, 2002).

Tests of alternative phylogenetic hypotheses among living amphibians were accomplished by using the CONSEL program (Shimodaira and Hasegawa, 2001). The first step was to reconstruct alternative tree topologies. PAUP\* heuristic searches under a GTR+I+ $\Gamma$  model and incorporating a topological constraint were conducted in order to identify the highest likelihood topology that satisfied a given hypothesis (e.g., the affinity between salamanders and caecilians). Second, PAUP\* was used to produce a log file for the sitewise log-likelihoods of alternative trees given the concatenated data set with a GTR+I+ $\Gamma$  model. The generated log file was submitted to the CONSEL program to calculate the p-value for each alternative topology by using the approximately unbiased (AU) test of Shimodaira (2002) and the Shimodaira-Hasegawa (SH) test (Shimodaira and Hasegawa, 1999). It has been shown that rate heterogeneity among genes is not approximated well enough by a single discrete  $\Gamma$ -distribution model of concatenated sequences, and that the summation from the separate analyses of individual genes is preferable to a single analysis of concatenated sequences (Cao et al., 1999). Therefore, log files for sitewise log-likelihood for individual gene alignments were generated with a different shape parameter for the  $\Gamma$ -distribution and with different base frequencies estimated for each gene. These log files were summarized to a single one and then subjected to the CONSEL analysis as mentioned above.

#### Molecular Dating

The Bayesian molecular dating was carried out by using the MultiDivtime software package described by Thorne and Kishino (2002). The same DNA alignments used in the phylogenetic analyses were used in the dating analyses. The lack of ATP8 and ND5 genes in *Polypedates megacephalus* did not matter because the program allowed some species to be absent in the alignments. The tree inferred from the nucleotide sequence phylogenetic analysis was used as the reference topology. The teleost sequence (*Cyprinus carpio*) served as the outgroup allowing the tree relating the remaining 19 ingroup sequences to be rooted. The Bayesian molecular dating process was performed with a prior of 400 Mya for the ingroup root (the split between Dipnoi and Rhipidistia + Tetrapoda; Benton, 1990) and a standard deviation of 10 Mya (i.e.,  $rttm = 4$ ,  $rtmsd = 0.1$ ). The prior mean and standard deviation for the gamma distribution describing the rate at the root node ( $rtrate$  and  $rtratesd$ ) were both set to 0.06.

The prior mean and standard deviation for the gamma distribution of the parameter controlling rate variation over time (i.e.,  $brownmean$  and  $brownsd$ ) were both set to 0.5. As the constraint points, we used  $310 \pm 10$  Mya for the Synapsida + Sauropsida separation (Kumar and Hedges, 1998; Benton, 1997), dates reliably estimated from fossil evidence (i.e., a minimum age of 300 Mya and a maximum age of 320 Mya for this node). To allow the Markov chain to reach stationarity, the Markov chain Monte Carlo algorithm completed 200,000 initial cycles before the state of the Markov chain was sampled. Thereafter, the Markov chain was sampled every 100 cycles until a total of 10,000 samples were collected. To test whether or not the Markov chain was converging, three single runs were performed. Similar results from the three runs were observed.

#### RESULTS

Our new mitochondrial genomes have been deposited in the GenBank under the accession numbers AY458591 to AY458595. Most amphibian sequences have the standard gene content (2 rRNAs, 22 tRNAs, and 13 protein-coding genes) of higher vertebrates. However, there are some traits in the new sequences that deserve to be mentioned. (1) The long noncoding spacer between tRNA-Thr and tRNA-Pro genes in other salamanders was also present in our new salamander sequence. (2) The sequence for Banna caecilian (*Ichthyophis bannanicus*) (15,983 bp) is of the smallest size among known amphibian sequences because its control region is very short (616 bp). (3) For all Neobatrachia frog sequences, the tRNA-Leu (CUN), tRNA-Thr, and tRNA-Pro genes form a concatenation unit and locate upstream to tRNA-Phe gene. This situation of tRNA gene rearrangement is only found in published Neobatrachia frog sequences, which would favor a monophyletic origin for the "new frogs" (Neobatrachia). (4) The tree frog, *Polypedates megacephalus*, possesses a novel mitochondrial gene order in amphibians. Unlike other Neobatrachia frogs, the tRNA-Leu<sup>CUN</sup> and tRNA-Thr genes exchange their positions in *Polypedates megacephalus* and form a Thr-Leu<sup>CUN</sup>-Pro-Phe tRNA gene tetrad. Moreover, the ATP8 and ND5 genes are absent in *P. megacephalus* mitochondrial genome; a noncoding sequence of 853 nt long has replaced the original position of ATP8 gene.

The final concatenated alignment contained 7659 nucleotide sites for the 21 taxa listed in Table 2. Of these sites, 3953 were constant, 857 were variable, and 2849 were informative for parsimony. Heuristic MP analysis yielded a single most-parsimonious tree (Fig. 1) with a length 14,259 steps (retention index = 0.4001). ML, NJ, and Bayesian analyses of the same data set produced trees with exactly the same topology as that found in the MP analysis (Fig. 1). In the Bayesian analyses, the three independent MCMC runs resulted in concordant joint posterior probability distributions for the topology and the estimated parameters of the model of sequence evolution. This result suggested that the chains were run for a sufficient number of generations.

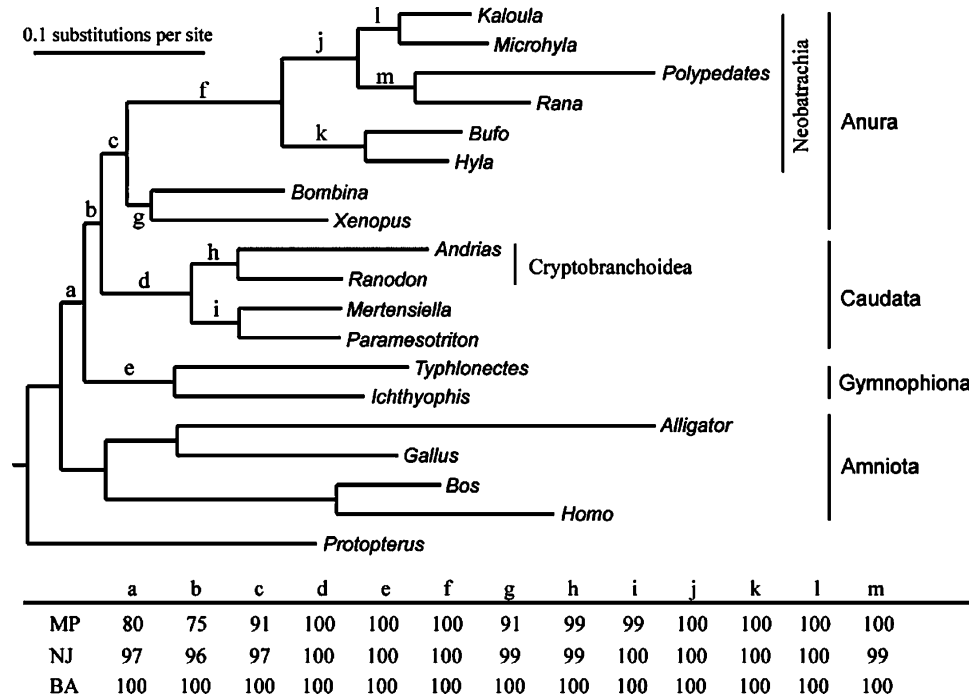


FIGURE 1. Phylogenetic relationships of frogs, salamanders, and caecilians inferred from analyses of mitochondrial genome sequences. Branches with letters have branch support values given below the tree for maximum parsimony (MP), neighbor-joining (NJ), and the likelihood-based Bayesian method (BA). Scale bar represents 0.1 substitutions per site. Outgroup species are not shown. Branch lengths were estimated by using maximum-likelihood inference.

In the recovered tree, all acknowledged natural groups (amphibians and amniotes) are well recognized. All methods highly suggested the monophyly of living amphibians (node “a”) and that of frogs (node “c”), salamanders (node “d”), caecilians (node “e”), and Neobatrachia (node “f”). In agreement with most morphological analyses, our tree suggested a close relationship between frogs and salamanders (Batrachia hypothesis, Milner, 1988; Trueb and Cloutier, 1991) with high support (node “b”). Although the *Xenopus-Bombina* clade (Archaeobatrachia) is strongly supported in our tree (node “g”), the Archaeobatrachia is actually not regarded as a natural group (Ford and Cannatella, 1993; Roelants and Bossuyt, 2005). When adding another partial “Archaeobatrachia” frog mitochondrial genome sequence (*Brachytarsophrys carinensis*, our unpublished results; only the 12S and cytb genes are incomplete) to the data set, the obtained tree does show a paraphyletic origin for the Archaeobatrachia (results not shown). In the topological analyses on the concatenated nucleotide sequences, the frog-caecilian tree and the paraphyly tree were always rejected ( $P < 0.05$ ). No statistically significant differences were found between the frog-salamander tree and the salamander-caecilian tree (Table 3). When using the separate analysis strategy (see Materials and Methods), we still cannot reject the salamander-caecilian hypothesis, but the  $P$  values of the salamander-caecilian tree dropped considerably ( $\sim 0.1$ ) (Table 3).

In our Bayesian molecular dating process, we calculated both the prior distribution (results not shown) and the posterior distribution (results shown in Fig. 2)

for all nodes. Similar results were observed in both analyses (for example, the posterior distribution for the lissamphibian origin is 337 [321–353] Mya and the prior distribution for this node is 339 [323–354] Mya). When allowing that all genes have same tendency to change rate (i.e., commonbrown = 1), similar results were also observed. The mean and 95% credibility intervals of molecular divergence times for the Amphibia-Amniota/reptiliomorphs separation and Bird-Crocodile separation were 354 (341–367) Mya and 259 (236–282) Mya, respectively, which are in close agreement with fossil-based estimates (365 Mya and 254 Mya, respectively). Moreover, recent findings on earliest known cryptobranchoids from the late Jurassic indicate that the Hynobiidae + Cryptobranchidae separation began 160 Mya (Gao and Shubin, 2001), which is very close to our estimate (158 [135–181] Mya). Consistency between the fossil- and molecular-based dating has

TABLE 3. Statistical confidence ( $P$ -values) for alternative relationships of the three orders of living amphibians AU: approximately unbiased test; SH: Shimodaira-Hasegawa test. For each test, the left value refers to the concatenated analysis, the right value refers to the sum of the separate analyses.

Topologies	$P$ values	
	AU	SH
((salamanders, frogs), caecilians)	0.818/0.919	0.895/0.955
((salamanders, caecilians), frogs)	0.223/0.120	0.264/0.134
((frogs, caecilians), salamanders)	0.029* / 0.025*	0.045* / 0.034*
(frogs, caecilians, salamanders)	<0.01* / <0.01*	0.011* / <0.01*

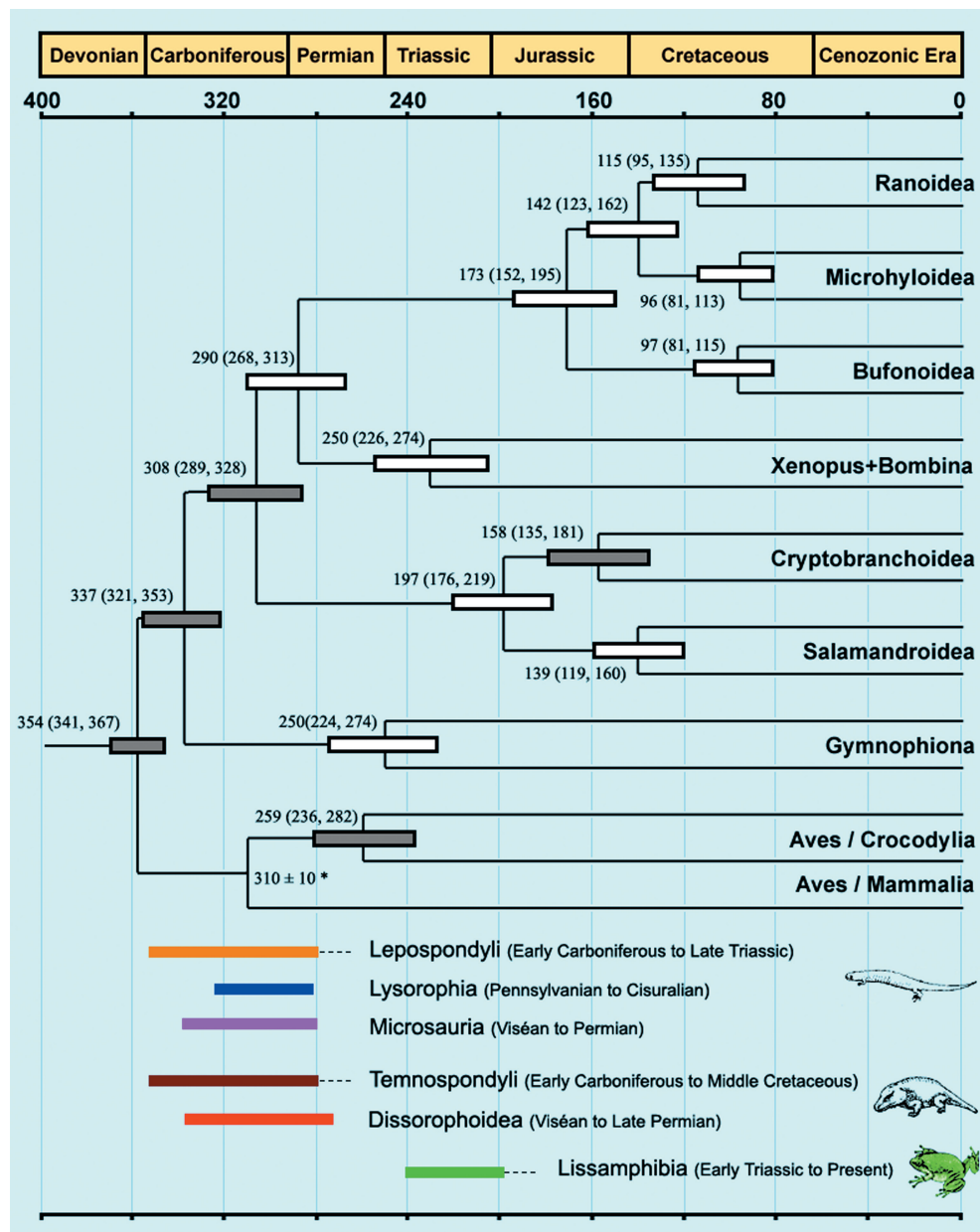


FIGURE 2. Molecular tree topology combined with dating of the phylogenetic nodes. Branch lengths are proportional to divergence times. The two dates used to constraint the tree were the Dipnoi/Rhipidistia + Tetrapoda separation ( $400 \pm 10$  Mya) (not shown) and the Synapsida + Sauropsida separation ( $310 \pm 10$  Mya) (indicated by an asterisk). Numbers above the nodes are the mean estimated divergence time (in Mya). Numbers in parentheses represent 95% credibility intervals (represented by horizontal bars). Blank horizontal bars indicate clades for which we lacked basal lineages, hence inferred ages for these clades are likely to be underestimates. The rough stratigraphic duration of related and possible ancestral fossil stocks are given below the tree.

dramatically raised confidence in our date estimates. The details of our molecular dating are given in Figure 2.

## DISCUSSION

### *Amphibian Phylogeny: Batrachia versus Procerata*

Historically, perhaps the most controversial question in living amphibian phylogeny has been the relationships among the three living orders: Anura (frogs), Caudata (salamanders), and Gymnophiona (caecilians).

Much morphological evidences suggest that salamanders are the closest living relatives of frogs (and form the clade Batrachia) to the exclusion of caecilians (Milner, 1988; Benton, 1990; Trueb and Cloutier, 1991; Laurin and Reisz, 1997). However, some authors have interpreted other morphological data from both living and fossil amphibians as supporting a close phylogenetic relationship between salamanders and caecilians (Carroll and Holmes, 1980; Bolt, 1991; Carroll et al., 1999).

Earlier studies of nuclear ribosomal genes (Hedges et al., 1990; Larson, 1991) and mitochondrial ribosomal genes (Hay et al., 1995; Feller and Hedges, 1998) concluded that, contrary to the commonly accepted Batrachia hypothesis, salamanders and caecilians are sister taxa (Procerata hypothesis). Zardoya and Meyer (2001) analyzed a mitochondrial genomic data set and concluded that the Batrachia hypothesis was also supported by molecular evidence. On the other hand, Zhang et al. (2003b) reanalyzed the mitochondrial genomic data set but with another frog (*Rana nigromaculata*) as the representative for anurans and suggested a close relationship between salamanders and caecilians again.

The incongruence of molecular phylogenies of living amphibians by using different data sets has long caused confusion and argument among molecular systematists. Recently, several researchers have demonstrated that short mtDNA fragments may perform poorly in inferring phylogeny among old lineages (i.e., >300 Mya) (Cao et al., 1994; Cummings et al., 1995; Zardoya and Meyer, 1996b). Thus, longer sequences may provide new insights. However, analyses of two different mitochondrial genomic alignments still resulted in contrary conclusions (as mentioned above). The possible source of this incongruence may be a long-branch-attraction (LBA) artefact (Felsenstein, 1978). This is likely to arise when internodes are short (relative to long terminal branches), and limited taxon sampling does not sufficiently truncate these long terminal branches (Phillipe, 2000). When LBA is occurring, the observed order of branching in molecular phylogeny probably partially reflects relative rates of evolution; i.e., the faster a lineage evolved for a given gene, the earlier it branches in the tree. In both published mitochondrial genomic trees (Zardoya and Meyer, 2001; Zhang et al., 2003b) only a single sequence for frogs or caecilians was included and the first amphibian lineage to branch off had the longest branch length within the lissamphibian clade, suggesting to us that perhaps LBA has occurred. We believe that sequencing more mitochondrial genomes in those amphibian lineages with highly variable evolutionary rates (especially frogs and caecilians) will help to truncate the long terminal branches among those lineages.

#### *Lissamphibian Origin: Monophyly versus Polyphyly*

The three living orders of amphibians differ significantly in their body plans; thus, the origin of Lissamphibia is another controversial topic in the tetrapod evolutionary history. Our molecular tree strongly supports a close relationship between frogs and salamanders (Batrachia hypothesis). Therefore, the polyphyletic origin theory that salamanders and caecilians evolved from lepospondyl amphibians (Carroll and Holmes, 1980; Bolt, 1991; Carroll et al., 1999) would be indirectly rejected by this result. Based on this hypothesis, the salamanders should be the sister-group to caecilians, not to frogs.

With a constraint of the frog-salamander clade (Batrachia), there are three main paleontological theories about the origin of the lissamphibians: (1) lissamphibians are

derived from the temnospondyl dissorophids (Milner, 1988; Trueb and Cloutier, 1991) (Fig. 3a); (2) lissamphibians are derived from the lepospondyl lysorophids (Laurin and Reisz, 1997) (Fig. 3b); (3) caecilians are most closely related to lepospondyl microsaurids, but salamanders and frogs are derived from different families of temnospondyl dissorophids (Carroll, 2001) (Fig. 3c). The temnospondyls, originating in the Mississippian period (Early Carboniferous, about 355 Mya), are an important order in the history of the amphibians. The intercentra of their vertebrae are large wedge-shaped elements, and the pleurocentra are comparatively small blocks that fitted in between the intercentra. These amphibians were most characteristically developed in the subsequent Pennsylvanian and Permian periods and became extinct in the middle Cretaceous. The dissorophids is a large superfamily of temnospondyl amphibians described from the Early Carboniferous (Viséan, about 340 Mya) to the Late Permian basin of North America. Those who favor a temnospondyl origin for extant amphibians generally identify the dissorophids as the ancestral stock. The lepospondyls are a highly varied amphibian group that appeared as early as the Mississippian period, and before the close of Paleozoic times became extinct. In the lepospondyls, the vertebrae are not preformed in cartilage, but rather are formed directly as spool-like, bony cylinders around the notochord. However, the lepospondyls are possibly a paraphyletic group united mostly by relatively small size and lack of labyrinthodont dentition. The microsaurids are a diverse group of small lepospondyl amphibians from the Carboniferous (Viséan, about 340 Mya) to the Permian periods. The lysorophids were a highly derived relative to other Paleozoic amphibians, most closely related to Microsauria, living from the Pennsylvanian (about 325 Mya) to the Cisuralian (about 270 Mya) periods. The rough stratigraphic duration for these related fossil stocks are illustrated in Figure 2.

Generally, our molecular date for the divergence of the three orders of living amphibians (337 [321–353] Mya) (Fig. 2) could be compatible with all the hypotheses above. But with this molecular time range, we can evaluate the credibility of each hypothesis. To simplify the mathematic deduction, we can simply regard the molecular time range as a mathematic model and treat the fossil time ranges (hypotheses) as data sets. The fitness ( $P$  value) between the model and the data set is the overlap integral. According to the Center Limit Theorem, the distribution of the mean of a variable will tend to follow the normal distribution when samples are adequate ( $n > 30$ ) no matter what distribution the variable follows. Thus, as an assumed model, the mean of our molecular time for an individual node tends to follow the normal distribution in the 95% credible interval because the number of samples in the Bayesian analysis is far beyond what is necessary ( $n = 10,000 \gg 30$ ). Although the age distributions of some nodes are a little skewed, the distribution for the lissamphibian divergence estimation (337 [321–353] Mya), our node of interest, is not skewed. Therefore, the molecular estimate for the lissamphibian divergence can be

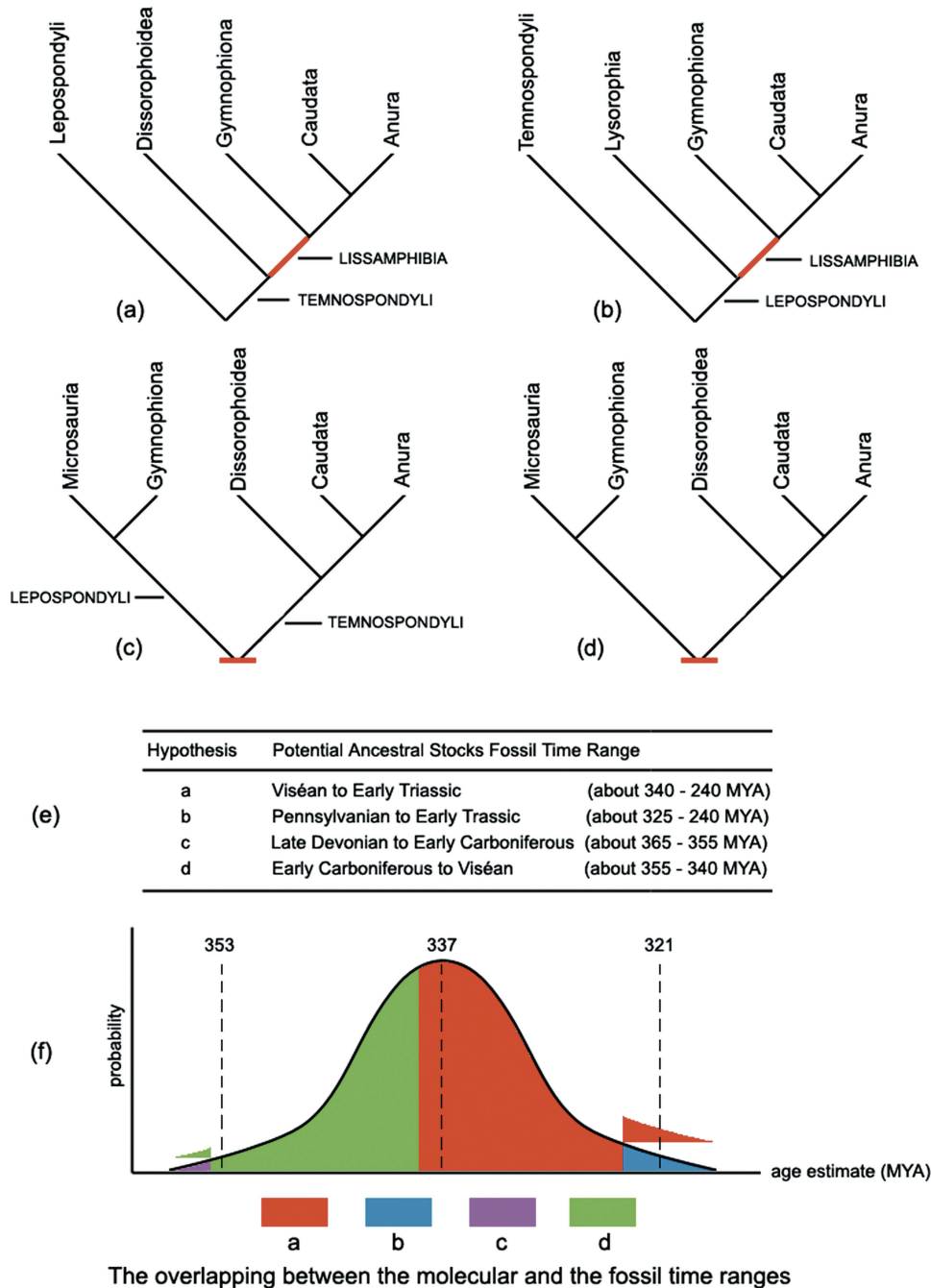


FIGURE 3. Evaluation of four alternative hypotheses of phylogenetic relationships among recent and fossil amphibians based on molecular time estimation. (a) Lissamphibia is monophyletic with Temnospondyli as sister group. (b) Lissamphibia is monophyletic with Lepospondyli as sister group. (c) Lissamphibia is not monophyletic. Gymnophiona is related to Microsauria (Lepospondyli), whereas Anura and Caudata are related to Temnospondyli. (d) Similar to (c), but the divergence of the lissamphibians is thought to be only a reflection of the microsaur-dissorophid separation. Ancestral stems for lissamphibians are indicated by red lines on the trees, the fossil time range inference is given in the table below (e). The credibility of each hypothesis is a function of how much the corresponding fossil time range (e) overlaps with the molecular estimate of the age of the lissamphibian divergence (337 [321–353] Mya) (f). Overlapping beyond the 95% credible interval can be considered as a rejection.

described using a normal distribution as in Figure 3f. The greater the overlap between the fossil time and the molecular time, the higher probability a hypothesis receives. The time ranges for the potential ancestral stocks

for each hypothesis are shown in the table of Figure 3e. It is obvious that the dissorophid hypothesis (red color) covers more area of the molecular time range than the lysorophid hypothesis (blue color), which suggests that

the temnospondyl theory (Fig. 3a) is more probable than the lepospondyl theory (Fig. 3b). As to the third hypothesis (Fig. 3c), the initial divergence of the lissamphibians is a reflection of the temnospondyls + lepospondyls separation, which took place beyond the Early Carboniferous. The fossil time range of this hypothesis is beyond the 95% credible interval of the molecular time range (pink color); thus, we can reject this hypothesis. However, because the lepospondyls are possibly paraphyletic and the phylogenetic relationships among Paleozoic amphibians are still tentative, the divergence of the lissamphibians under this hypothesis may be only a reflection of the microsaur-dissorophids separation (instead of the temnospondyls-lepospondyls separation) that might take place at a much later time (possibly in the Early Carboniferous, before the first appearance of the microsaur or the dissorophids) (Fig. 3d). Hence the independent origin hypothesis has moderate credibility (green color). In conclusion, the dissorophid hypothesis is the most probable theory under our molecular time evaluation.

Indeed, only paleontological data can provide direct evidence to support which extinct amphibian group is actually most closely related to living amphibians. One problem in the comparative studies among amphibians is that morphological characters are easily affected by the complications of adaptive convergence and it is therefore difficult to distinguish between primitive and convergent characters. For example, the array of features related to hearing (otic notch and slender stapes, etc.) is thought to be powerful evidence supporting a temnospondyl origin for lissamphibians. However, the lepospondyl theory considers such features as homoplastic (but not homologous) characters. (Coates and Milner, 2000; Laurin et al., 2000). With an accurate estimate of phylogeny and divergence times, comparative biologists are now able to focus their attention on certain stocks within a fixed period. This will help paleontologists to modify their character database and improve the reliability of subsequent phylogenetic analyses.

#### *Interpretation and Hypothesis of Biogeographic Patterns*

Although our analyses cannot include the most primitive amphibian lineages for some clades (i.e., dates for those divergence events may be underestimated), the new timescale can still provide new perspectives on the origin of the main living amphibian groups. Our molecular time for the origin of Salientia (290 [268–313] Mya) is much older than the previous molecular estimate ( $197 \pm 43.2$  Mya) (Kumar and Hedges, 1998) and the fossil-based time (about 240 Mya) (Benton, 1990). This dating is probably an underestimate because primitive frog lineages such as *Ascaphus* and *Leiopelma* are not included. A possible interpretation of this gap in fossil records is that molecular times are overestimated. However, this is unlikely, as earlier (Late-Devonian) and later (Mid-Jurassic) fossil and molecular dates show a close agreement. Recently, Roelants and Bossuyt have also used the Bayesian approach to address the origin of frogs; they have obtained a different result ( $225 \pm 28$  Mya) (Roelants and

Bossuyt, in press). The reasons for such incongruence may be due to: (i) different data sizes (7659 base pairs in our study, 3963 in Roelants and Bossuyt); (ii) the heterogeneity of data (ours: mitogenomes; theirs: nuclear and mitochondrial gene fragments). To get more accurate results in future research, larger data sets from different sources (both mtDNA and nuclear DNA) should be used. The earliest fossils of the Salientia from both Madagascar (*Triadobatrachus*) (Rage and Rocek, 1989) and Poland (*Czatkobatrachus*) (Evans and Borsuk-Bialynicka, 1998) indicate that at least in the early Triassic the stem stock of Salientia had possessed a worldwide distribution. Our results put this time even further to the early Permian. The Permian-Triassic gap in our record of Salientia evolution may be filled by future fossil finds of this period.

Fossil records of “advanced” frogs (Neobatrachia) are only known from the Cenozoic. The apparent relatively late origin of this group is at odds with their generally wide distribution across the main continental masses. Such distributions of animals that cross saltwater with difficulty would appear to be incompatible with a time of origin when the separation of the continents was already well advanced. Our timescale indicate that the Neobatrachia originated in at least the mid-Jurassic (perhaps the early Jurassic). Obviously, a Jurassic origin would enable the ancestors of Neobatrachia to disperse to other continental masses via intercontinental corridors. Although no fossil records of advanced frogs are known from the Mesozoic period, the modern distribution of primitive lineages of the Neobatrachia can show some clues on their early evolution. The basal lineages of advanced frogs (Sooglossidae and Nasikabatrachidae) are now distributed in the India-Seychelles; other primitive lineages are distributed in Africa (Heleophrynidae) and Australia (Myobatrachidae) (according to the phylogram of Biju and Bossuyt, 2003). Based on a world map of the early Jurassic (Scotes, 2002), an Africa-India origin may be the most parsimonious dispersal pattern for advanced frog evolution (Fig. 4a).

The new timescale indicates that modern salamanders originated in at least the early Jurassic (perhaps the late Triassic), a time when Pangaea just began to break. This time is much later than the supposed Permian origin (Milner, 1983), but it is in congruence with a strong Laurasian pattern of distribution of modern salamanders. The earliest known fossil record of salamanders is from north China (Gao and Shubin, 2001) but North America possesses most salamander families, including the most primitive lineage of modern salamanders—Sirenidae. Where is the right place of the radiation for the early salamander evolution? The continent map of the early Jurassic may show some clues (Fig. 4). A North American origin for salamanders will enable salamanders to disperse into Gondwana easily before the mid-Jurassic. However, there is no fossil evidence of this period from Gondwana, making this hypothesis somewhat suspect. On the contrary, a Far East origin will inhibit the Gondwanan dispersal from Laurasia (because of long distance) until the breakup of Pangaea had been well established (Fig. 4b). The high diversity of modern



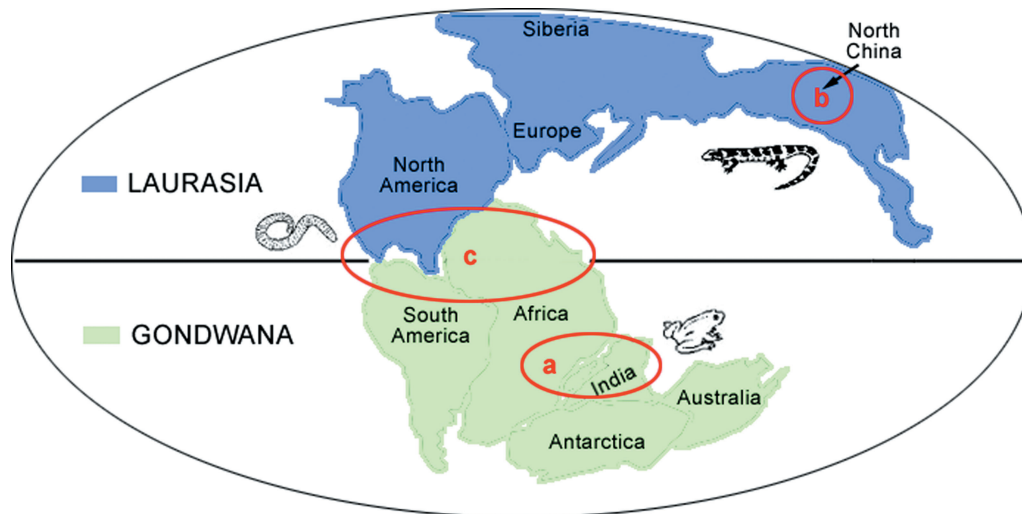


FIGURE 4. The rough continental map from the Triassic to the early Jurassic with potential regions of origin for main living amphibian groups indicated by circles. We tentatively propose (a) an Africa-India origin for advanced frogs, and (b) a Far East origin for modern salamanders. The tropic-forest of Triassic Pangaea (southern North America, northern South America and Northern Africa in the modern world) may be the place of origin of living caecilians (c).

salamander families in North America may be a consequence of continental vicariance when North America was separated from Eurasia in the early Cretaceous.

The living caecilians are almost endemic to Gondwanan areas. The only exception—Ichthyophiidae—which is found on Gondwanan (India) and Laurasian (Southeast Asia) landmasses, is thought to be the result of continental drift on the Indian subcontinent (Duellman and Trueb, 1986). Therefore, a Gondwanan origin after the breakup of Pangaea for modern caecilians would be justified. However, the discovery of a caecilian fossil (*Eocaecilia micropodia*) (Jenkins and Walsh, 1993) from the Lower Jurassic of Arizona (Laurasia) began to shed doubts on the authenticity of such a hypothesis. Regarding this contradiction, Feller and Hedges (1998) argued that the common ancestors of living caecilians were presumably limbless and the limbed caecilian fossil might have been just close to the divergence of salamanders and caecilians. In agreement with the fossil evidence, our molecular dating further shows that the common ancestors of living caecilians arose at least in the Lower Triassic. Such an early origin of caecilians would make their continental dispersal possible. The ancient caecilians might have been similar to their living descendants, living in a tropic-forest burrowing lifestyle. If so, the tropic forest of the Triassic Pangaea could have been the original place for the ancient caecilian occurrence (Fig. 4c).

#### ACKNOWLEDGMENTS

We are very grateful to Xiao-Hong Chen for technical assistance and Hong-Liang Pang for support in sequencing some mitochondrial genomes. Dr. Shao-Qin Tang helped to collect most of amphibian samples. Constructive comments by Rod Page, Karl Kjer, and Jeff Thorne greatly improved this manuscript. We also thank Professor Mohsen Ghadessy for revising the English text. This work is supported by the National Natural Science Foundation of China (No. 30370754) and the

Program for Changjiang Scholars and Innovative Research Team in University from the Ministry of Education of China.

#### REFERENCES

- Anderson, S., M. H. DE Bruijn, A. R. Coulson, I. C. Eperon, F. Sanger, and I. G. Young. 1982. Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. *J. Mol. Biol.* 156:683-717.
- Andrews, R. M., I. Kubacka, P. F. Chinnery, R. N. Lightowers, D. M. Turnbull, and N. Howell. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23:147.
- Benton, M. J. 1990. Phylogeny of the major tetrapod groups: Morphological data and divergence dates. *J. Mol. Evol.* 30:409-424.
- Benton, M. J. 1997. *Vertebrate palaeontology*, 2nd edition. Chapman & Hall, London, UK.
- Biju, S. D., and F. Bossuyt. 2003. New frog family from India reveals an ancient biogeographical link with the Seychelles. *Nature* 425:711-714.
- Bolt, J. R. 1991. Lissamphibian origins. Pages 194-222 in *Origins of the major groups of Tetrapods: Controversies and consensus* (H. P. Schultze and L. Trueb, eds.). Cornell University Press, Ithaca, New York.
- Cao, Y., J. Adachi, A. Janke, S. Pääbo, and M. Hasegawa. 1994. Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: Instability of a tree based on a single gene. *J. Mol. Evol.* 39:519-527.
- Cao, Y., K. Kim, J. Ha, and M. Hasegawa. 1999. Model dependence of the phylogenetic inference: Relationship among carnivores, perissodactyls and cetartiodactyls as inferred from mitochondrial genome sequences. *Genes Genet. Syst.* 74:211-217.
- Carroll, R. L. 2001. The origin and early radiation of terrestrial vertebrates. *J. Paleontol.* 75:1202-1213.
- 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.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540-552.
- 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.

- Coates, M. I., and A. R. Milner. 2000. Early tetrapod evolution. *Trends Ecol. Evol.* 15:327–328.
- 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.
- Delarbre, C., N. Spruyt, C. Delmarre, C. Gallut, V. Barriel, P. Janvier, V. Laudet, and G. Gachelin. 1998. The complete nucleotide sequence of the mitochondrial DNA of the dogfish, *Scyliorhinus canicula*. *Genetics* 150:331–344.
- Desjardins, P., and R. Morais. 1990. Sequence and gene organization of the chicken mitochondrial genome, a novel gene order in higher vertebrates. *J. Mol. Biol.* 212:599–634.
- Duellman, W. E., and L. Trueb. 1986. *Biology of amphibians*. McGraw-Hill, New York.
- Evans, S. E., and M. Borsuk-Bialynicka. 1998. A stem-group frog from the Early Triassic of Poland. *Acta Palaeontol. Polon.* 43:573–580.
- 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. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401–410.
- Ford, L. S., and D. C. Cannatella. 1993. The major clades of frogs. *Herpetol. Monogr.* 7:94–117.
- Gao, K. Q., and N. Shubin. 2001. Late Jurassic salamanders from northern China. *Nature* 410:574–576.
- Hay, J. M., I. Ruvinsky, S. B. Hedges, and L. R. Maxson. 1995. Phylogenetic relationships of amphibian families inferred from DNA sequences of mitochondrial 12S and 16S ribosomal RNA genes. *Mol. Biol. Evol.* 12:928–937.
- Hedges, S. B., K. D. Moberg, and L. R. Maxson. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol. Biol. Evol.* 7:607–633.
- Huelsenbeck, J. P., and N. S. Imennov. 2002. Geographic origin of human mitochondrial DNA: Accommodating phylogenetic uncertainty and model comparison. *Syst. Biol.* 51:155–165.
- Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
- Janke, A., and U. Arnason. 1997. The complete mitochondrial genome of Alligator mississippiensis and the separation between recent archosauria (birds and crocodiles). *Mol. Biol. Evol.* 14:1266–1272.
- Jenkins, F. A., and D. M. Walsh. 1993. An early Jurassic caecilian with limbs. *Nature* 365:247–250.
- 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.
- Kumar, S., and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:197–200.
- Larson, A. 1991. A molecular perspective on the evolutionary relationships of the salamander families. *Evol. Biol.* 25:211–277.
- Laurin, M., M. Girondot, and A. de Ricqlès. 2000. Early tetrapod evolution. *Trends Ecol. Evol.* 15:118–123.
- Laurin, M., and R. Reisz. 1997. A new perspective on tetrapod phylogeny. Pages 9–59 in *Amniote origins*, (S. S. Suemida and K. L. Martin, eds.). Academic Press, New York.
- Milner, A. R. 1983. *Evolution, time and space: The emergence of the biosphere*. Academic Press, New York.
- Milner, A. R. 1988. The relationships and origin of living amphibians. Pages 59–102 in *The phylogeny and classification of the Tetrapods* (M. J. Benton, ed.). Clarendon Press, Oxford.
- Philippe, H. 2000. Opinion: Long branch attraction and protist phylogeny. *Protist.* 151:307–316.
- Posada, D., and K. A. Crandall. 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Rage, J. C., and Z. Roček. 1989. Redescription of *Triadobatrachus massinoti* (Piveteau, 1936) an anuran amphibian from the early Triassic. *Palaeontographi. Abt.* 206:1–16.
- 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.
- Roelants, K., and F. Bossuyt. 2005. Archaeobatrachian paraphyly and Pangaeian diversification of crown-group frogs. *Syst. Biol.* 54:111–126.
- Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Mol. Biol. Evol.* 19:101–109.
- Scotese, C. R. 2002. <http://www.scotese.com> (PALEOMAP website).
- Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16:1114–1116.
- Shimodaira, H., and M. Hasegawa. 2001. CONSEL: For assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246–1247.
- Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* 51:492–508.
- Sumida, M., Y. Kanamori, H. Kaneda, Y. Kato, M. Nishioka, and H. Yonekawa. 2001. Complete nucleotide sequence and gene organization of the mitochondrial genome of the Japanese pond frog *Rana nigromaculata*. *Genes Genet. Syst.* 76:311–325.
- Swofford, D. L. 2001. PAUP\*: Phylogenetic analysis using parsimony (\*and other methods), version 4.0b8. Sinauer Associates, Sunderland, Massachusetts.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids. Res.* 24:4876–4882.
- Thorne, J. L., and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51:689–702.
- Trueb, L., and R. Cloutier. 1991. A phylogenetic investigation of the inter- and intrarelations of the Lissamphibia (Amphibia: Temnospondyli). Pages 223–313 in *Origins of the major groups of Tetrapods: Controversies and consensus* (H. P. Schultze and L. Trueb, eds.). Cornell University Press, Ithaca, New York.
- 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. 2000. Mitochondrial evidence on the phylogenetic position of caecilians (Amphibia: Gymnophiona). *Genetics* 155:765–775.
- Zardoya, R., and A. Meyer. 2001. On the origin of and phylogenetic relationships among living amphibians. *Proc. Natl. Acad. Sci. USA* 98:7380–7383.
- Zhang, P., Y. Q. Chen, Y. F. Liu, H. Zhou, and L. H. Qu. 2003a. The complete mitochondrial genome of the Chinese giant salamander, *Andrias davidianus* (Amphibia: Caudata). *Gene* 311:93–98.
- Zhang, P., Y. Q. Chen, H. Zhou, X. L. Wang, and L. H. Qu. 2003b. The complete mitochondrial genome of a relic salamander, *Ranodon sibiricus* (Amphibia: Caudata) and implications for amphibian phylogeny. *Mol. Phylogenet. Evol.* 28:620–626.

First submitted 24 March 2004; reviews returned 29 September 2004;

final acceptance 14 December 2004

Associate Editor: Karl Kjer