2-2004

Phylogenetic signal and the utility of 12S and 16S mtDNA in frog phylogeny

S. Hertwig

Rafael O. de Sá
University of Richmond, rdesa@richmond.edu

A. Haas

Recommended Citation

Hertwig, S., De Sá, R.O. and Haas, A. (2004), Phylogenetic signal and the utility of 12S and 16S mtDNA in frog phylogeny. Journal of Zoological Systematics and Evolutionary Research, 42: 2-18. https://doi.org/10.1111/j.1439-0469.2004.00225.x

J. Zool. Syst. Evol. Research **42** (2004) 2–18
© 2004 Blackwell Verlag, Berlin
ISSN 0947–5745

Received on 31 July 2002

[1]*Institut für Spezielle Zoologie und Evolutionsbiologie mit Phyletischem Museum, Friedrich-Schiller-Universität Jena, Jena, Germany;* [2]*Department of Biology, University of Richmond, Richmond, VA, USA*

Phylogenetic signal and the utility of 12S and 16S mtDNA in frog phylogeny

S. Hertwig[1], R. O. de Sá[2] and A. Haas[1]

Abstract

Genes selected for a phylogenetic study need to contain conserved information that reflects the phylogenetic history at the specific taxonomic level of interest. Mitochondrial ribosomal genes have been used for a wide range of phylogenetic questions in general and in anuran systematics in particular. We checked the plausibility of phylogenetic reconstructions in anurans that were built from commonly used 12S and 16S rRNA gene sequences. For up to 27 species arranged in taxon sets of graded inclusiveness, we inferred phylogenetic hypotheses based on different *a priori* decisions, i.e. choice of alignment method and alignment parameters, including/excluding variable sites, choice of reconstruction algorithm and models of evolution. Alignment methods and parameters, as well as taxon sampling all had notable effects on the results leading to a large number of conflicting topologies. Very few nodes were supported in all of the analyses. Data sets in which fast evolving and ambiguously aligned sites had been excluded performed worse than the complete data sets. There was moderate support for the monophyly of the Discoglossidae, Pelobatoidea, Pelobatidae and Pipidae. The clade Neobatrachia was robustly supported and the intrageneric relationships within *Bombina* and *Discoglossus* were well resolved indicating the usefulness of the genes for relatively recent phylogenetic events. Although 12S and 16S rRNA genes seem to carry some phylogenetic signal of deep (Mesozoic) splitting events the signal was not strong enough to resolve consistently the inter-relationships of major clades within the Anura under varied methods and parameter settings.

Key words: Lissamphibia – Anura – mitochondrial genes – alignment – rRNA

Introduction

Most contemporary studies in frog systematics have readily assimilated molecular techniques or rely exclusively on them. The application of molecular techniques has given stimulating impulses to frog and amphibian phylogenetics (e.g. De Sá and Hillis 1990; Hedges and Maxson 1993; Hillis et al. 1993; Hay et al. 1995; Ruvinsky and Maxson 1996; Graybeal 1997; Feller and Hedges 1998; Richards and Moore 1998; Garcia-Paris and Jockusch 1999; Clough and Summers 2000; Emerson et al. 2000; Vences et al. 2000; Zardoya and Meyer 2001). In particular, fragments of the mitochondrial 12S and 16S genes have been used ubiquitously and continue to be used in frogs, as well as other vertebrate and invertebrate groups (e.g. Mattern and McLennan 2000; Buckley et al. 2002; Leaché and Reeder 2002). Both genes have been applied at various hierarchical levels of frog phylogeny ranging from intrageneric relationships (e.g. Dawood et al. 2002) to the relationships of the major clades within the Anura (e.g. Hay et al. 1995). On the geological time scale these studies address splitting events covering recent Cenozoic times and deep Mesozoic events, respectively (Sanchiz 1998). Controversial views on the resolving power of mitochondrial rRNA sequences were presented early on (Mindell and Honeycutt 1990; Dixon and Hillis 1993). Yet, the usefulness of the genes to address questions specifically in frog evolution at different hierarchical levels has not been demonstrated.

The alignment of homologous positions within orthologous genes is pivotal in phylogenetic studies of nucleotide sequences. Particularly, the alignment of highly divergent non-protein coding sequences such as rRNAs causes problems with regard to the determination of reliable positional homologies (Simon et al. 1994; Lutzoni et al. 2000). The rRNA sequences can vary considerably in length because of numerous insertions and deletions in fast evolving parts of the genes. Maximum parsimony (MP) and maximum likelihood (ML) methods both depend on correctly homologized positions as represented in the data matrix. The rRNA genes are characterized by regions of highly conserved secondary structure motifs as well as stretches with high rates of sequence evolution (Mindell and Honeycutt 1990; Dixon and Hillis 1993; Simon et al. 1994). In the fast evolving regions indels can cause difficulties in sequence alignments. Different approaches have been proposed to improve the alignment of ambiguously aligned regions: secondary structure-based alignments (Orti et al. 1996; Titus and Frost 1996; Wiens and Reeder 1997), and parsimony-based, or optimization alignments (Wheeler 1996, 1999; Lutzoni et al. 2000; Wheeler 2001). Ambiguously aligned regions have either been excluded altogether from phylogenetic analysis (Gatesy et al. 1993; Leaché and Reeder 2002), coded as missing data, or differentially weighted to reduce the detrimental effect of uncertain positional homologies (e.g. Wheeler et al. 1995; Zardoya and Meyer 2001). These different approaches may lead to inconsistent results in subsequent phylogenetic reconstructions (Vogler and DeSalle 1994; Giribet and Wheeler 1999; Lutzoni et al. 2000).

The phylogenetic signal in 12S and 16S genes with respect to anurans is investigated in this study. The evolution of major clades of frogs presumably took place from 200 to 140 Mio years ago according to the fossil record (Sanchiz 1998). In the light of their long evolutionary history anurans are a good model to test the resolving power of ribosomal genes. Despite many previous studies, large parts of the presumed phylogeny of the Anura are unresolved (Ford and Cannatella 1993) or controversial (Hay et al. 1995). Ascaphids, discoglossids, pipids and pelobatoids are generally considered groups that stem from early splitting events in frog evolution (Sanchiz 1998). The status of the Discoglossidae and Pelobatoidea, however, is uncertain; several alternative phylogenetic hypotheses have been proposed (Ford and Cannatella 1993; Hay et al. 1995; Haas 1997, 2003; Maglia 1998), some based on the genes in question (Hay et al. 1995).

We ask whether these genes can be properly applied over a wide spectrum of evolutionary time to answer questions in frog systematics. As there is no objective criterion to choose between alternative approaches, the effects of various *a priori* decisions of phylogenetic analysis by us on its actual outcome are explored; e.g. the choice of taxa, the alignment method and alignment parameter settings, inclusion/exclusion and weighting, as well as the choice of reconstruction algorithm and models of evolution.

Materials and Methods

Choice of taxa

The set of species examined and the use of species versus higher taxa as terminals in phylogenetic analyses can have tremendous effects on phylogenetic inference (Lecointre et al. 1993; Yeates 1995; Bininda–Emonds et al. 1998; Graybeal 1998; Hillis 1998; Grandcolas and D'Haese 2001). Although single species are commonly used as representatives for species-rich taxa in phylogenetic studies (Hay et al. 1995; Zardoya and Meyer 2000), broader species samples should, in general, lead to more robust hypotheses (Graybeal 1998; Hillis 1998), e.g. amending the problem of long branch attraction (Felsenstein 1978; Swofford et al. 1996). In order to assess the genes' phylogenetic signal for relatively recent splitting events, particularly representatives of the Discoglossidae are used for this study. In Europe, *Alytes*, *Bombina* and *Discoglossus* underwent speciation likely during the Tertiary (Maxson and Szymura 1979, 1984; Sanchiz 1998). Taxa relevant to the issue of the basal branching pattern within the Anuran were combined with them (e.g. archeobatrachians *sensu* Reig 1958). Choice of taxa was inspired by the concept of hierarchical sampling (Graybeal 1993). Accession (GenBank) numbers of sequences examined are summarized in Appendix 1.

In order to explore the effects of taxon sampling on the alignment procedures and the phylogenetic reconstruction three taxon sets were built (see Appendix 2) with graded inclusiveness. Taxon group 1 comprises Dipnoi, *Sphenodon* and Lissamphibia; group 2 includes species representing the Lissamphibia with a caudate, anurans, and a caecilian; whereas group 3 was restricted to discoglossids as ingroup and selected other anurans as outgroup.

Molecular characters and techniques

Muscle tissue of freshly alcohol preserved specimens was excised. Standard proteinase K/PCI. DNA extraction protocols were applied (Maniatis et al. 1982; Hillis et al. 1996). An approximately 400-bp 12S rRNA segment and an approximately 500-bp 16S rRNA segment were amplified using polymerase chain reaction (PCR; Palumbi 1996). Primers were selected according to Goebel et al. (1999): 16S L2a 5′TCGAACTTAGAGATAGCTGGTT3′; 16S H17 5′GCGAATGTT TTTGGTAAACA3′; 12S A-L 5′AAACTGGGATTAGATACCCCA CTAT 3′; 12S B-H 5′GAGGGTGACGGGCGGTGTGT3′. These primers correspond to position 2490–2910 (12SrRNA) and 3458–3963 (16SrRNA) of the *Xenopus laevis* mitochondrial genome (Roe et al. 1985; Gen-Bank no.: M10217). The following PCR temperature cycles were found most efficient in a Robocycler Gradient 96 (Stratagene, La Jolla, California). 12S primers: one cycle (3 min/94°C, 1 min/47°C, 1 min/72°C), followed by 35 cycles (1 min/94°C, 1 min/47°C, 1 min/72°C). 16S primers: 1 cycle (3 min/94°C, 45 s/55°C, 1 min/72°C); followed by 5 cycles (1 min/94°C, 45 s/55°C, 1 min/72°C); followed by 30 cycles (1 min/94°C, 45 s/53°C, 1 min/72°C).

We extracted PCR products from agarose gel electrophoresis using two methods. First, excision of gel pieces containing the DNA with subsequent standard PCI/chloroform extraction technique and precipitation (Maniatis et al. 1982; Hillis et al. 1996). Secondly, trapping of PCR products in a PEG (15% polyethylene glycol) filled gel well during electrophoresis (Hillis et al. 1996).

Purified templates were sequenced in both directions with the Thermo Sequenase Cycle Sequencing Kit (Amersham Pharmacia Biotech, Amersham, UK) and 5′ IRD800 labelled primers (manufacturer's manual LiCor, Lincoln, Nebraska). Sequences were read with a Li-Cor 4000 sequencer (LiCor, Lincoln, Nebraska). Forward and reverse raw sequences were matched with BIOEDIT 5.0.6 (Hall 1999) and GENDOC 2.6 (Nicholas and Nicholas 1997) software. Equivocal positions of the consensus sequences were inspected visually and corrected manually using the original chromatogram files. See Appendix 1 for GenBank accession numbers.

To assess the presumed secondary structure of the sequenced gene stretches we computed the folded structure under given thermodynamic parameters using the Mfold energy-minimization method (folding temperature 25°C; Zuker and Stiegler 1981; Jaeger et al. 1990; Mathews et al. 1999). For this approach we used the complete GenBank sequences for both genes in *Ichthyophis bannanicus*, *Rana catesbeiana* and *Xenopus laevis* as reference taxa (Appendix 1).

Alignment

All 12S and 16S rRNA sequence data were combined in a single data matrix and analysed simultaneously. It was assumed that both 12S and 16S fragments, evolved along the same underlying topology (Buckley et al. 2002), because of their common evolutionary fate as parts of the ribosome, as well as the mitochondrial genome. Both 12S and 16S have similar patterns of high among-site rate variation (Simon et al. 1994; Orti et al. 1996).

Two different methods were used for data analyses. First, a one-step procedure (Sankoff and Rousseau 1975; Wheeler 1996) implemented in the software POY 2.0 (Gladstein and Wheeler 1996). It seeks in a combined analysis (via optimization steps of the nucleotide data) for the tree topology, which is based on the most parsimonious alignment under given parameters (Wheeler 2001). Secondly, a two-step procedure of initial alignment with CLUSTAL X and subsequent phylogenetic reconstruction. In order to identify and delimit ambiguous regions, the sequences were aligned four times with CLUSTAL X 1.8.1. (Thompson et al. 1994; Higgins et al. 1996; Thompson et al. 1997) applying four different sets of multiple alignment parameters: Gap Opening Penalty (GOP) 15/Gap Extension Penalty (GEP) 6.6; GOP 10/GEP 5; GOP 20/GEP 5; and GOP 5/GEP 4. Ambiguously aligned positions were identified by eye and excluded manually using BIOEDIT. The remaining alignment (one in each taxon group) is dubbed CLUSTAL X 'CE' herein, 'E' for excluded.

In a second approach, all sites were retained and three alignments with various gap cost were generated using CLUSTAL X (multiple alignment parameters: 'CA': GOP 15/GEP 5; 'CB': GOP 5/GEP 4; 'CC': GOP 20/GEP 5). Separate analyses were run for gap coding as either fifth character state or missing data.

Finally, in addition to the separate analyses of alignments CLUSTAL X CA, CB, and CC, these alignments were concatenated in a single large matrix. In essence, this procedure weighs sites differentially during tree search depending on their variability of positional homology (Wheeler et al. 1995; Lutzoni et al. 2000).

Outputs of the alignment programs were imported and prepared for phylogenetic analyses in MACCLADE 4.0 (Maddison and Maddison 2000).

The CLUSTAL X CA alignment of taxon group 1 was used to compute pairwise distance measures (distance, number of substitutions) with MEGA 2.01 software (Kumar et al. 2001). '*Complete deletion*' of gaps option was in effect and the TN (Tamura and Nei 1993) model of sequence evolution was used for calculating distances (parameter determined with MODELTEST 3.06, Posada and Crandall 1998).

Tree reconstruction

Maximum parsimony

The parsimony analyses were performed with PAUP 4.0b8 software (Windows Version; Swofford 1998). Initially equally weighted parsimony was applied. Gaps were coded alternatively as fifth character state, assuming that insertions and deletions also represent informative evolutionary changes (Simmons and Ochoterena 2000; Simmons et al. 2001) or as missing data. The shortest trees were sought by heuristic search method (10 000 random addition replicates, TBR branch swapping, *MulTree* in effect). Bootstrap and jackknife (50% deletion) analyses were performed with 2000 replicates (heuristic search, TBR,

10 random additions respectively) to infer branch robustness (Hedges 1992).

Phylogenetic information of transversions was explored separately by re-coding all characters as pyrimidine or purine bases and gaps as missing data (transversion parsimony, Swofford et al. 1996). Additionally, we used a step matrix for the MP analyses with weights for each transformation step according to their frequency distribution within the data matrix (determined with MODELTEST). Substitution costs applied: transversions 4, purine transition 2, pyrimidine transition 1 and indels 2.

Maximum likelihood

MODELTEST was used to determine parameter settings and models of sequence evolution for the different alignments. The ML analyses were performed with TREE-PUZZLE 5.0 (Strimmer and von Haeseler 1996). It uses the heuristic quartet-puzzling algorithm to compute likelihood trees. The TN and the HKY85 (Hasegawa et al. 1985) models were selected in separate analyses of each aligned data set, because the GTR model (general time reversible; Lanave et al. 1984; Rodriguez et al. 1990; Yang 1994) is not implemented in TREE-PUZZLE. Parameters of the models of sequence evolution and rate heterogeneity were estimated by TREE-PUZZLE based on a neighbour-joining tree and the exact likelihood function. The number of replicates of the quartet-puzzling algorithm was set to 10 000. We selected a model of among-site rate heterogeneity of substitutions that consisted of one invariable rate and eight variable rates with gamma distribution.

Bayesian approach

The Bayesian methodology as implemented in MRBAYES 2.01 (Huelsenbeck 2000) estimates the posterior probabilities of the best set of trees for a given model of sequence evolution (Rannala and Yang 1996; Yang and Rannala 1997; Huelsenbeck and Ronquist 2001). The GTR model with eight classes of substitution rates and gamma distribution (GTR + I + gamma) was used; model parameters were estimated by MrBAYES. Analyses were initiated with random starting trees. Additional settings following Huelsenbeck et al. (MrBAYES documentation): program estimated base frequencies; 500 000 generations with four independent Markov Chains were started; every 100th generated topology was saved; the first 500 generated topologies were excluded from the final 50% majority rule consensus tree. In the output files it was controlled whether the Markov chains had become stationary for their log-likelihood scores with 1000 samples and excluded further topologies if necessary. In a further search the parameters were set identical to those of TREE-PUZZLE runs to compare the results of both methods. For that approach, the CLUSTAL X alignment without variable aligned positions and the HKY85 model were used.

One-step analysis with POY

In POY's optimization alignment approach (Wheeler 1996), alignment of sequences and tree reconstruction are performed simultaneously. Commands followed Gladstein and Wheeler (1996): (i) *Diagnose* (prints branch length and apomorphy list derived from a search), (ii) *Impliedalignment* (generates a topology specific multiple alignment based on the synapomorphy scheme), (iii) *Random* 100 (causes 100 random addition sequence searches (build through swapping) to be performed), (iv) *Multibuild* 10 (makes 10 random addition sequence builds on slave nodes, the best ones are submitted to branch swapping), (v) *Maxtrees* 5 (set maximum trees held in buffers to five), (vi) *Slop* 2 (check all tree length which are within 'n' 10th of a per cent of the current minimum value), (vii) *Checkslop* 5 (checks all tree lengths that are within 'n' 10th of a per cent of the current minimum length using an additional tbr branch swapping round), (viii) *Tbr* (tbr branch swapping), (ix) *Nospr* (suppresses spr branch swapping), (x) *Randomizeoutgroup* (randomize the outgroup in (iii) *Random* and (iv) *Multibuild*), and (xi) *Quick* (branch swapping only on minimal-length trees, analogous to *-steepest descent* in PAUP). Gap costs were set to 1, 2, 4 and 8, respectively. Furthermore, the jackboot routines of POY (Gladstein and Wheeler 1996) were used: *-Jackboot -Random* 200 *-Quick -Randomizeoutgroup -Maxtrees* 10 *-Tbr -Nospr*; 50% majority rule consensus trees were computed with CONSENSE in PHYLIP 3.6 (Felsenstein 1989).

Results of analyses were plotted with TREEVIEW 1.61 (Page 1996).

Results

Secondary structure

Comparisons of the complete 12SrRNA sequences of *Ichthyophis bannanicus, Rana catesbeiana* and *Xenopus laevis* (GenBank) showed that differences in the primary structure entail significant differences in the putative secondary structure as inferred from computer folding models. Few positions form highly conserved homologous motifs of the secondary structure in the three taxa. At various sites the computer models reconstructed unpaired bases within conserved regions that otherwise form predominantly helical stems. Some of the paired stem and the unpaired loop regions were shifted in position related to changes in the primary structure. Comparison of these three taxa alone, thus, did not support the notion that secondary structure of rRNA is largely fixed in taxa with highly divergent primary sequences. Furthermore, changes of folding temperature (20, 30°C) in separate runs resulted in different models of secondary structure.

Alignments

The various alignment procedures resulted in data matrices that differed with regard to positional homology hypotheses, the number of variable characters, and the number of parsimony informative characters. The comparison is summarized in Tables 1 and 2. Varying the alignment parameters (gap cost) in each of the alignment procedures gave different and unique alignments. Each of the gap cost regimes applied in CLUSTAL X, e.g. yielded alternative primary homology hypotheses that resulted in different phylogenetic reconstructions. The same effect was evident in the comparison of the implied alignments of POY searches under different parameters. Taxon sampling also had a significant effect of on the outcome of alignments as detailed in Tables 1 and 2.

Congruent nodes in multiple analyses

We sought for nodes with universal support (Tables 3–5) in the n-dimensional space of solutions (Wheeler 1995; Phillips et al. 2000) from all analyses. In all cases, genera represented by more than one species were recovered as monophyletic entities with high support values (*Alytes*, *Bombina*, *Discoglossus*, *Limnodynastes*, *Pelodytes*, *Rana*). The node *Limnodynastes* + *Rana* (representatives of Neobatrachia) was well supported in all analyses. Within discoglossid frogs, methods that resolved intrageneric relationships consistently supported the clades [*Discoglossus montalentii* + (*D. galganoi* + *D. pic-*

Table 1. Alignments with CLUSTAL X. CA, CB and CC represent alignments under different gap cost schemes (see text for further information)

	CA			CB			CC			CE		
G	tp	vp	ip	tp	vp	ip	tp	vp	ip	tp	vp	ip
1	928	689	573	996	748	589	941	666	551	485	270	199
2	915	643	516	953	672	524	913	644	522	542	293	205
3	878	415	308	882	406	302	876	424	312	689	249	166

CE, the alignment without the ambiguously aligned positions. Characterization of the differences: tp, total number of positions of a given alignment; vp, number of variable positions; ip, number of parsimony informative positions (tested in PAUP).

Table 2. Implied alignments of POY (IAP), tested with PAUP. OA GC 1, 2, 4, 8: optimization alignment with gap cost 1, 2, 4, 8

	Group 1			Group 2			Group 3		
Alignment	tp	vp	ip	tp	vp	ip	tp	vp	ip
OA GC 1	1215	981	642	1090	810	568	941	447	304
OA GC 2	1064	843	587	980	710	524	906	431	307
OA GC 2	1065	845	589						
OA GC 4	1009	845	616	955	722	535	904	434	313
OA GC 4							903	431	311
OA GC 8	983	833	655	949	736	592	901	468	337

Characterization of the differences: tp, total number of positions of a given alignment (homology lines); vp, number of variable positions; ip, number of parsimony informative positions (tested in PAUP). Note, that this table contains only implied alignments, which differed in number of homology lines from alignments of equal parsimonious solutions. With gap cost two and group 1 as well as with gap cost four and group 3 more than one alignment based on the same most parsimonious topology was found.

tus + *D. sardus*)] and *Bombina bombina* + *B. variegata*. The signal for *Alytes* + *Bombina* in species groups 1 and 2 was weak. The monophyly of the Pelobatoidea, Pipidae and Discoglossidae was suggested in results from multiple parameter settings and analysis methods. Similarly, within the Pelobatoidea the family Pelobatidae (*s. str.*, i.e. *Pelobates* + *Leptobrachium*) was supported.

In all analyses of taxa group 2 the monophyly of the Anura was robustly supported (Table 4). Yet, analyses of both groups 1 and 2 resulted in numerous conflicting hypotheses of relationships concerning the major anuran clades (deep splits). Nodes connecting major clades were mostly weakly supported in robustness tests or by likelihood values. For example, the relationships of *Ascaphus* appear completely undetermined by these data sets. Its position in the respective topology was highly sensitive to choice of parameters for alignment as well as reconstruction method (ML, MP).

In the two-step procedures, fewer conflicting topologies were derived for species group 3 (discoglossids, *Ascaphus*, *Pelodytes*) than for the two more inclusive taxon samples, probably due to the exclusion of highly divergent sequences in some of their species. In almost all group 3 analyses (except for POY, gap cost 4 and 8), the Discoglossidae was monophyletic and the genus *Discoglossus* was the sister-group of *Alytes* + *Bombina* (Fig. 2, Table 5).

The robustness of these phylogenetic hypotheses (group 3) was assessed by changing the composition of the outgroup. In two sets of analyses *Pelodytes caucasicus* was replaced by either *Xenopus* or *Pelobates* and combined with *Ascaphus* as outgroup. Three CLUSTAL X alignments were performed (parameters as in CA, CB, CC) for each of the outgroup changes and subjected to MP PAUP analyses. In comparison with the original group 3 analyses, some nodes were sensitive to the composition of the outgroup. With *Ascaphus* and *Xenopus* as outgroup, the bootstrap values supporting the Discoglossidae (Table 7) were lower; furthermore, the relationships between *Alytes*, *Bombina* and *Discoglossus* were unresolved. Discoglossidae was monophyletic in only one of the three cases (CA, CB, CC) and the node *Alytes* + *Bombina* received substantially lower support than in original group 3 analyses (Table 7), whereas with *Ascaphus* and *Pelobates* as outgroup.

Reconstruction with POY

Each of the 12 searches with POY found most parsimonious solutions with different topologies and different implied alignments. The jackboot test of POY supported only few nodes robustly in taxon groups 1 and 2, and some of the group 3 analyses (Tables 3–5).

For taxon group 1, gap cost 1 and 2 gave topologies that differed only in the alternative nodes Archaeobatrachia and Neobatrachia + Pipidae (Table 3). Most of the nodes of these topologies were congruent with other results. In contrast, gap cost 4 and 8 resulted in highly implausible nodes, e.g. *Lepidosiren* + *Ambystoma* + *Ichthyophis* as sister group of the Neobatrachia or *Alytes* + [Neobatrachia + (*Lepidosiren* + *Ambystoma* + *Ichthyophis*)].

For taxon group 2, the different gap cost settings each resulted in one most parsimonious solution. Among them, the topology based on gap cost 4 had more nodes congruent with two-step analyses (Table 4) than the other gap cost settings. Among other results, gap cost 1 suggested the clade *Discoglossus* + (*Ascaphus* + Pelobatoidea), whereas with gap cost set to 2, *Ascaphus* was resolved within discoglossids [*Ascaphus* + (*Alytes* + *Discoglossus*)], and, finally for gap cost 8 the clade *Ascaphus* + Pipidae was the sister-group of *Bombina* + Pelobatoidea.

Finally, in taxon group 3 analyses, each of gap cost 1 and 2 gave only one most parsimonious tree, both of identical topology (Table 5). Yet, six different minimum length trees were found in searches with gap cost set to 4, the resulting strict consensus was highly unresolved. Only two shortest topologies were found with gap cost setting 8. In their consensus, the Discoglossidae were paraphyletic [*Pelodytes* + (*Alytes* + *Discoglossus*)] at the exclusion of *Bombina*.

Maximum parsimony PAUP

PAUP. Alignment parameters in CLUSTAL X had a clear effect on the resulting topologies when ambiguously aligned positions and gaps were included in the data matrix. The resulting topologies of these alignments contained conflicting and weakly supported hypotheses of anuran relationships with respect to deep splits in particular. Yet, certain subclades (Fig. 3, Tables 4 and 5) were highly supported in consensus trees. Nodes resolved under MP phylogenetic reconstructions were mostly robust, no matter if gaps were coded as additional fifth character state or as missing data.

The concatenated alignments CABC resulted in increased resolution and higher node support in comparison with analyses dealing with the three alignments separately. In MP analyses of group 1, the combined alignment resulted in 90 minimal-length trees. Their 50% majority rule consensus tree was highly resolved, with early all of the nodes having high

Table 3. Selected clades and their support in the analyses of taxa group 1

Analysis		Lissamphibia	Anura	Archeobatrachia	Discoglossidae	*Alytes* + *Bombina*	*Alytes* + *Discoglossus*	Pipidae	Pelobatoidea	Pelobatidae	Pelobatoide + Pelodytidae	Pelobatoide + scaphiopodidae	Pelodytidae + Scaphiopodidae	Neobatrachia	Neobatrachia + Pipidae
MP															
CA		62	99	69	98	74	95	89	92					100	
CB		94	100		78	82		98	96	98			72	100	
CC		69	99	76	73	61		97	97	97		74		100	
CABC		97	100	91	92	88		100	99	100		68		100	
CA	MD	75	99		64	74			84	93				98	
CB	MD	92	92		72	76		87	87	98			64	99	
CC	MD	93	99	63	70	82		93	97	97	65			99	
CE					70	81								96	
CE	TV														95
CE	SM				57	72		99							
POY															
POY	GC 1	+	+		+	+		+	+	+		+		+	+
POY	GC 2	+	+	+	+	+		+	+	+		+		+ 58	
POY	GC 4						+	+	+	+		+		+ 71	
POY	GC 8								+	+		+		+ 58	
MrBAYES															
MB CA	GTR	50	100		91	100		98	100	100		97		100	90
MB CB	GTR	100	100		87	98		100	100	100			99	100	
MB CC	GTR	100	100	62		99		100	100	100		95		100	
MB CABC GTR		100	100	65	100	100		100	100	100		99		100	
MB CE	GTR				96	98		53						100	72
MB CE	HKY85				92	97								100	
TREE-PUZZLE															
TP CA	TN	89	85			61		94		74			65	66	
TP CB	TN	73	55			71		98	77	96			85	91	
TP CC	TN	91	83			66		95		73			65	78	
TP CABC	TN	95	81			70		96	56	79			79	81	
TP CE	HKY85				53	73		94						63	
TP CE	TN				55	76		93						62	

Numbers represent bootstrap values in parsimony (PAUP, POY), likelihood values of TREE-PUZZLE and posterior probabilities of MRBAYES topologies. Abbreviations: CA, CB, CC, CLUSTAL X alignments under different gap cost; CABC, fusion of CA, CB, CC into a single matrix; CE, CLUSTAL X alignment (ambiguously aligned positions excluded); GC, gap cost; MB, Bayesian analysis; MD, gaps coded as missing data; SM, differential weighting with step matrix; TP, maximum likelihood analysis with TREE-PUZZLE; TV, transversion parsimony; GTR, HKY85, TN are models of sequence evolution. Nodes found in the strict consensus topologies of POY searches are marked by "+".

Table 4. Selected clades and their support derived from the analyses of species group 2

Analysis		Anura	Archeobatrachia	Discoglossidae	*Alytes* + *Bombina*	*Alytes* + *Discoglossus*	Pipidae	Pelobatoidea	Pelobatidae	Pelobatidae + Pelodytidae	Pelobatidae + Scaphiopodidae	Pelodytidae + Scaphiopodidae	Neobatrachia	Neobatrachia + Pipidae
MP														
CA		99		97		63	91	93	94		65		100	
CB		100		99		62	93	71	82			75	100	50
CC		100		88		74	92	97	96		79		100	
RCA		100	70	64	75		94	90	95		75		100	
RCB		100	50	79	81		98	95	97			81	100	
RCC		100	61	70	61		96	96	95		70		100	
CABC		100		100		78	100	97	99		51		100	
CA MD		100		68		55	81	95	90		58		100	
CB MD		100		89		51	81	51	79			70	99	
CC MD		100		77		65	87	95	95		69			
CE		87		67	73				53				93	
CE TV		86											96	
CE SM		92			60				57				98	
POY														
POY GC 1		+	+		+		+	+	+		+		+	
		58											54	
POY GC 2		+				+	+	+	+	+			+	+
													74	
POY GC 4		+	+	+	+		+	+	+		+		+	
													74	
POY GC 8		+	+				+	+	+		+		+	
													74	
MrBAYES														
MB CA	GTR	100		97	75		100	100	100		86		100	
MB CB	GTR	100		100	87		100	100				87	100	
MB CC	GTR	100				96	100	100	100		77		100	68
MB CABC	GTR	100		100	97		100	100	100		95		100	100
MB CE	GTR	57		73	77			64	93	84			100	
MB CE	HKY85	53		63	67			54	67	71			100	
ML														
TREE-PUZZLE														
TP CA	TN	80		59			88	73	90			73	76	
TP CB	TN	90		94		64	97		82			58	75	
TP CC	TN	86		62		70	90	60	81			69	76	
TP CABC	TN	93		86	70		97		74			74	71	
TP CE	HKY85	66			61		88		77				54	
TP CE	TN	68			61		88		78					

Numbers represent bootstrap values in parsimony (PAUP, POY), and neighbour joining (PAUP) analyses, as well as likelihood values of TREE-PUZZLE and posterior probabilities of MRBAYES topologies. Reconstructed nodes in the most parsimonious trees are represented by + CA, CB, CC, CLUSTAL X alignments under different gap cost; CABC: fusion of CA, CB, CC into a single matrix; CE, CLUSTAL X alignment (ambiguously aligned positions excluded); GC, gap cost; MB, Bayesian analysis; MD, gaps coded as missing data; RCA, RCB, RCC, post-alignment reduced samples; SM, differential weighting with step matrix; TP, maximum likelihood analysis with TREE-PUZZLE; TV, transversion parsimony; GTR, HKY85, TN models of sequence evolution.

Table 5. Selected clades and their support from analyses of species group 3

Analysis		Discoglossidae	*Alytes* + *Bombina*	*Alytes*	*Bombina*	*Discoglossus*	*D. galganoi* + *D. pictus* + *D. sardus*	*D. sardus* + *D.pictus*	*D. galganoi* + *D.sardus*	*B. orientalis* + [*B.bombina* + *B.variegata*]	*B.bombina* + *B.variegata*
MP											
CA	BS	95	94	100	100	100	99	61		65	88
	JK	94	94	100	100	100	100	63		65	87
CB	BS	99	99	100	100	100	100	70		79	76
	JK	99	99	100	100	100	100	71		78	75
CC	BS	98	97	100	100	100	99	52		67	88
	JK	99	98	100	100	100	99	53		64	88
CABC	BS	100	97	100	100	100	100	67		71	96
JK		100	97	100	100	100	100	67		72	96
CE	BS	82	88	100	100	100	99	50			61
	JK	80	88	100	100	99	99				61
CE TV	BS		61	100	100	97	55				81
	JK		63	100	100	97					62
CE SM	BS	59	74	100	100	100	98		50		68
	JK	60	73	100	100	99	97		52		67
POY GC 1		+	+	+	+	+	+	+		+	+
	JB	63		100	100	93	78	51			
POY GC 2		+	+	+	+	+	+	+		+	+
	JB			100	100	96	83			60	61
POY GC 4				+	+	+				+	
	JB			100	100	91	78	54		65	68
POY GC 8			+	+	+		+		+	+	
	JB			100	100	71	69			68	58
MrBAYES											
MB CA	GTR	87	99	100	100	100	100		96		90
MB CB	GTR	100	100	100	100	100	100		99		96
MB CC	GTR	85	100	100	100	100	98		95	52	92
MB CABC	GTR	100	100	100	100	100	100		100	65	100
MB CE	GTR	95	68	100	100	100	100		88		85
MB CE	HKY 85	94	67	100	100	100	100		67		85
TREE-PUZZLE											
TP CA	GTR	94	98	100	100	100	100	70		64	79
TP CB	GTR	100	100	100	100	100	100			50	78
TP CC	GTR	93	100	100	100	100	100	67		52	83
TP CABC	GTR	100	100	100	100	100	100	75		57	95
TP CE	HKY85	95	95	100	100	100	100	52			76
TP CE	TN	92	95	100	100	100	100	52			76

Numbers represent bootstrap (BS), jackknife (JK), and jackboot (JB) values in parsimony (PAUP, POY) analyses, as well as likelihood values of TREE-PUZZLE and posterior probabilities of MRBAYES topologies CA, CB, CC, CLUSTAL X alignments under different gap costs; fusion of CA, CB, CC into a single matrix; CE, CLUSTAL X alignment (ambiguously aligned positions excluded); GC, gap cost; MB, Bayesian analysis; MD, gaps coded as missing data; SM, differential weighting with step matrix; TP, maximum likelihood analysis with TREE-PUZZLE; TV, transversion parsimony; GTR, HKY 85, TN, models of sequence evolution. Nodes found in the strict consensus topologies of POY searches are marked by '+'.

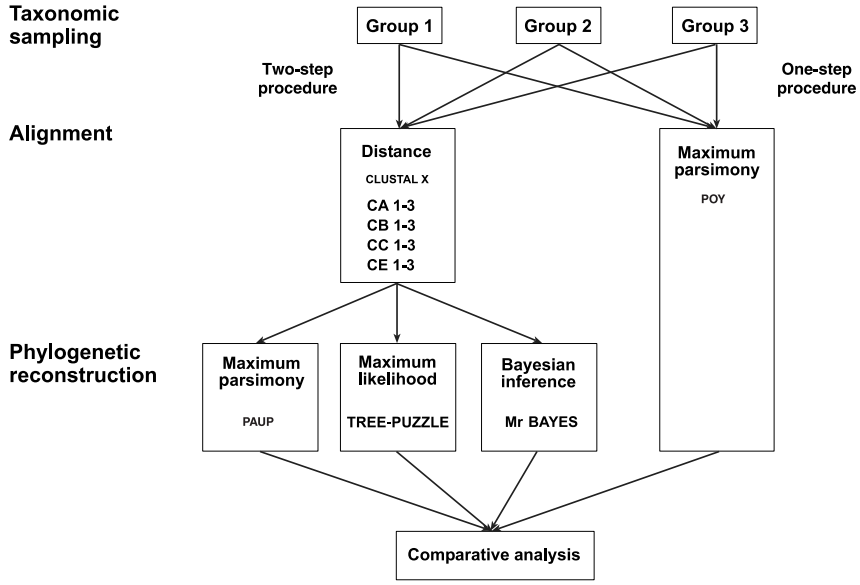

Fig. 1. Study design and methods applied

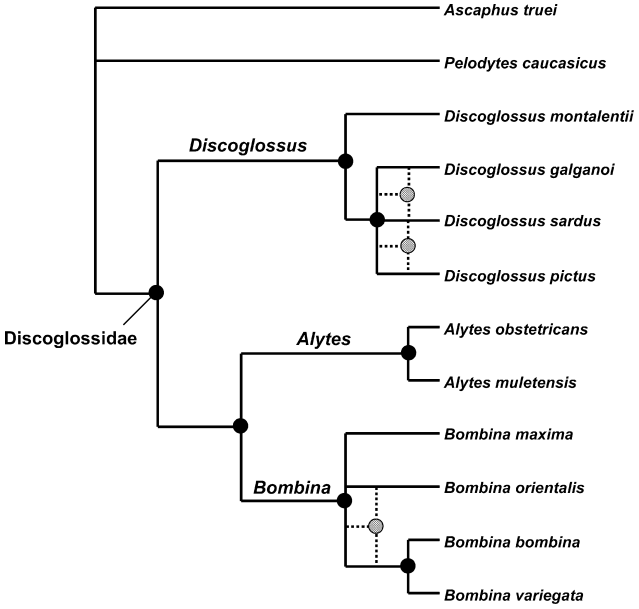

Fig. 2. For species group 3, all analyses reconstructed largely identical topologies, except for alternative subclades indicated by grey nodes and dashed lines

Table 6. Proportion of unresolved quartets in the ML analyses with TREE-PUZZLE (see text). Tree searches under HKY85 model were performed only with CLUSTAL X alignment CE (ambiguous positions excluded)

Data set	HKY85	TN
Group 1		
CA		12.8
CB		10.7
CC		10.8
CABC		2.7
CE	14.0	14.2
Group 2		
CA		15.3
CB		9.3
CC		12.6
CABC		3.3
CE	12.5	13.5
Group 3		
CA		6.5
CB		5.1
CC		5.5
CABC		1.8
CE	9.3	8.7

bootstrap support (Table 3). The search with group 2 yielded 109 equally parsimonious trees with less-resolved consensus and bootstrap trees. Yet, those consensus nodes found had high support values (Table 4).

The exclusion of ambiguously aligned positions not only reduced the number of informative characters (Table 1), but also led to low plausibility of the reconstructed topologies in all analyses of groups 1 and 2. In all iterations of CLUSTAL X alignments the excision of ambiguously aligned positions increased the number of conflicting trees and reduced the resolution in majority rule bootstrap topologies. MP analysis of species group 1 found seven minimum-length trees (ambiguous excluded). Their strict consensus topology contained arrangements in stark conflict with well-established phylogenetic hypotheses, e.g. reconstructing a paraphyletic Anura, a clade consisting of *Sphenodon* + *Leptobrachium* positioned within the Anura, and a sister-group relationship of *Ascaphus* + Neobatrachia. In analyses of species group 2 one most parsimonious tree was found (monophyletic Archaeobatrachia, *Ascaphus* + Pelobatoidea). The support for deep divergence events in these alignments was weak.

Overall, the lower rate of transversions should lead to better resolution of deep divergence events because of low saturation effects. Yet, when using transversions alone, we obtained consensus trees with reduced resolution in bootstrapping (no intrageneric resolution) and highly unlikely topologies (*Sphenodon* + *Leptobrachium* and *Ascaphus* + *Scaphiopus* within the Pelobatoidea). Differential weighting of transformations (stepmatrix) within the species group 1 data set resulted in slightly increased numbers of resolved nodes with $>50\%$ bootstrap and jackknife support. Yet, the most parsimonious topologies were similar to equally weighted parsimony and in conflict with

Table 7. Selected clades and their support in the analyses of species group 3 (*Alytes, Bombina, Discoglossus*) and alternative outgroup composition. Numbers are bootstrap support values in parsimony (PAUP) analyses

Analysis	*Discoglossidae*	*Alytes* + *Bombina*	*Alytes* + *Discoglossus*	*Alytes*	*Bombina*	*Discoglossus*	D.*galganoi* + *D.pictus* + *D.sardus*	D.*sardus* + *D.pictus*	*B.orientalis* + (*B.bombina* + *B.variegata*)	*B.bombina* + *B.variegata*
Ascaphus, Pelodytes										
CA	95	94		100	100	100	99	61	65	88
CB	99	99		100	100	100	100	70	79	76
CC	98	97		100	100	100	99	52	67	88
CABC	100	97		100	100	100	100	67	71	96
Ascaphus, Xenopus										
CA	68			100	100	100	100	59	56	82
CB	75			100	100	100	99	51	58	73
CC	64			100	100	100	100	63	56	82
CABC	89		50	100	100	100	100	61	61	96
Pelobates, Ascaphus										
CA	54			100	100	100	100	57	66	79
CB	71	69		100	100	100	98	51	55	74
CC		53		100	100	100	100	58	65	77
CABC	63	70		100	100	100	100	56	74	94

CA, CB, CC, CLUSTAL X alignments under different gap costs; CABC, fusion of CA, CB, CC into a single concatenated matrix.

well-established hypotheses of amphibian relationships. In species group 2, differential weighting caused lower resolution in the test procedures as compared with equal weights.

Influence of taxon sampling on alignment and phylogenetic reconstruction

We tested the influence of taxon sampling by comparing prealignment taxon exclusion to postalignment taxon exclusion. The first case simply equals group 2 and group 3 analyses. The second case was prepared by aligning group 1 data and subsequently reducing the species to match the species composition of groups 2 and 3. We used CLUSTAL X alignments (CA, CB, CC) and analysed the data with MP in PAUP under described parameters (heuristic search, bootstrapping). The results of group 2 derived from the two approaches were clearly different (Table 4). Some nodes were highly sensitive to taxon sampling, e.g. Archaeobatrachia, and the topology within the Discoglossidae. In group 3, the comparison between the topologies of the postalignment reduction and initial group 3 alignments yielded no differences in the consensus topology, yet, node support (bootstrap) for *Alytes* + *Bombina* dropped from 94–99 to 62–74, respectively.

Maximum likelihood

We used the CLUSTAL X alignments CA, CB, CC, CABC and CE, respectively, as data matrices for TREE-PUZZLE and MRBAYES tree searches.

TREE-PUZZLE. In TREE-PUZZLE branch support values >70% were considered robust. Only nodes with support >50% were considered and shown in the resulting tree topologies. TREE-PUZZLE also computes the percentage of unresolved quartets. This percentage is an indicator for the suitability of the data for the explored phylogenetic problem (Strimmer and von Haeseler, TREE-PUZZLE Manual).

The relatively high percentages of unresolved quartets in taxa groups 1 and 2 (Table 6) seem to indicate that the sequence data and parameters applied were not appropriate to resolve the phylogenetic problem. The ratio of unresolved quartets was lower and the obtained trees were more resolved with alignments derived from the smallest taxa set, group 3. The lowest proportions of unresolved quartets were determined in the concatenated alignments CABC from all groups.

The topologies reconstructed by TREE-PUZZLE from CLUSTAL X alignments under HKY85 and TN models were identical. Only the branch support values of internal nodes differed slightly. Furthermore, the quartet puzzling trees of these alignments of species groups 1 and 2 were highly polytomous (Tables 3 and 4). Using the CLUSTAL X alignments CE some of the few reconstructed nodes based on the species group 1 were considered highly unlikely (paraphyletic Amphibia, *Sphenodon* + *Leptobrachium*, located within the anura, are monophyletic).

MRBAYES

The topology derived from CLUSTAL X CE alignments (ambiguous sites excluded) of group 1 was implausible (Anura paraphyletic and a clade *Sphenodon* + *Leptobrachium* within the Pelobatoidea with high support). The consensus topology derived from species group 2 was not resolved with regard to anuran higher clades (Fig. 4). When CLUSTAL X alignments (CA, CB, CC, CABC, ambiguous sites included) were analysed with Bayesian inference method, the majority rule consensus

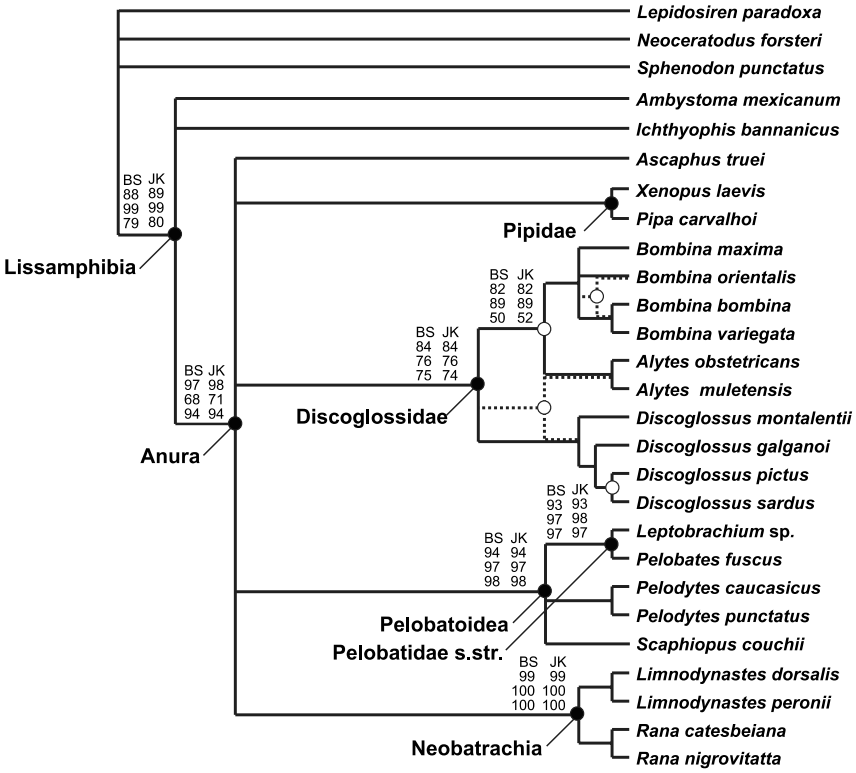

Fig. 3. Summary of the MP analyses based on CLUSTAL X alignments (CA, CB, CC) of species group 1 under different gap costs. Majority rule tree from bootstrap (BS) and jack-knife (JK) runs. The strict consensus of all most-parsimonious trees and the consensus tree from jackknife analyses were identical in topology. White nodes and dashed lines indicate those alternative arrangements in which analyses of species group 2 deviated from analyses of group 1

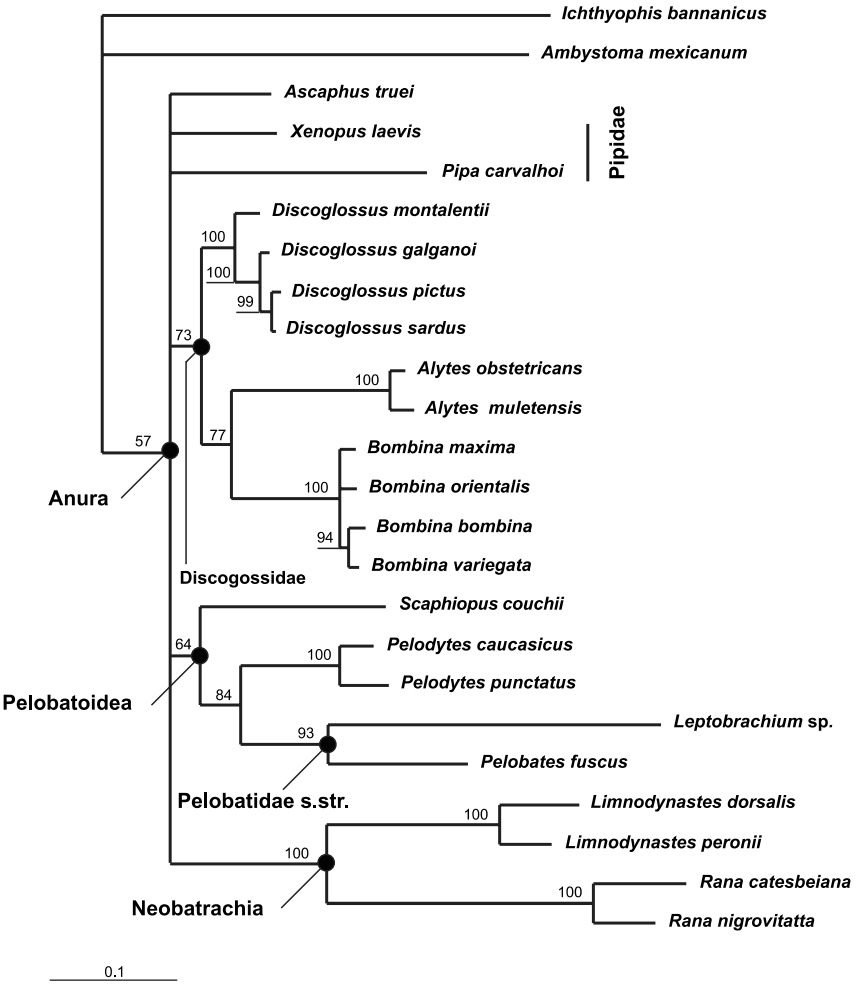

Fig. 4. Majority rule consensus tree of 4500 generated topologies of the Bayesian analysis (MRBAYES) of CLUSTAL X alignment (ambiguously aligned positions excluded); species group 2

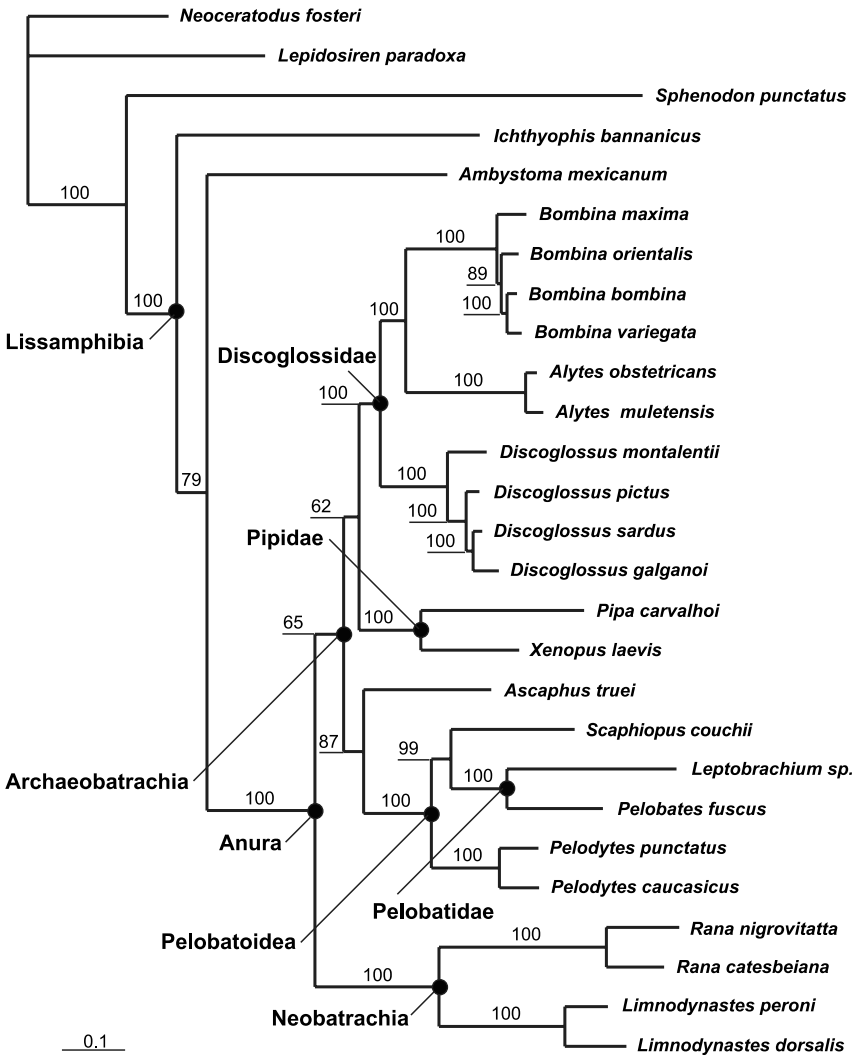

Fig. 5. Majority rule consensus tree of 4500 generated topologies of the Bayesian analysis (MRBAYES) of alignment CABC (CLUSTAL X, concatenated alignments CA, CB, CC); species group 1

trees of groups 1 and 2, respectively, showed considerably more resolved nodes and higher support values (Fig. 5, Tables 3 and 4).

Quartet puzzling resolved a lower number of nodes as compared with Bayesian analyses. The former seemed more conservative in reconstructing nodes from weak phylogenetic signal. Yet, the resulting topologies from both methods were similar for most of their robust nodes. In both TREE-PUZZLE and MRBAYES, the choice of a sequence evolution model (GTR or HKY 85 in MRBAYES, TN or HKY 85 in TREE-PUZZLE) had only minor effects on node support (Table 3–5), without changing the topologies.

Discussion

Secondary structure

RNA secondary structure has been used in different steps of phylogenentic studies (Wheeler and Honeycutt 1988; Simon et al. 1994; Kjer 1995; Orti and Meyer 1997). The phylogenentic approach of inferring secondary structure has been applied repeatedly (de Sá and Hillis 1990; Dixon and Hillis 1993; Alves-Gomes et al. 1995; Kjer 1995; Titus and Frost 1996; Orti and Meyer 1997; Kjer et al. 2001). In this procedure the inference of potential secondary structure motifs depends on the recognition of similarity in the primary structure (Armbruster 2001; Shull et al. 2001). If secondary structure was largely fixed across even distantly related taxa, it could potentially serve as a template for alignments (Kjer 1995; Hancock and Vogler 2000). Yet, there is a high degree of uncertainty about the errors caused by comparing distantly related species with very different primary structures. Mutational processes such as slippage could be agnostic with respect to secondary structure. Also, substitutions within helical stems may not necessarily require a compensatory substitution in the complementary base, if compensated by a shift from a stem to a loop motif (Hancock and Vogler 2000). In contrast to the phylogenetic approach, computer models for RNA folding according to minimized free energy are methodologically independent from alignment. Yet, the real secondary structure may still deviate from the minimal free energy model because of other unaccounted constraints (Zuker and Stiegler 1981; Severini et al. 1996). Different energy optimization algorithms may compute different secondary structures from the same data and rely on unrealistic fixed temperatures (Armbruster 2001). As a result of considerably variation in the primary structures of the three model species (*Xenopus laevis, Rana catesbeiana, Ichthyophis bannanicus*) and temperature dependent variation of models, our computations yielded no reliable secondary structures that were usable for the improvement of alignments or differential weighting schemes.

Alignment of ribosomal genes

Highly variable rRNA regions cause sequence length differences (Indels) among taxa and require gap insertions in alignments. The scope of taxon sampling directly influences the alignment process. Our results, among others (e.g. Wägele and Staniek 1995), indicate the significant influence alignments of ribosomal sequence data have on the results of subsequent phylogenetic reconstructions. Numerous nodes in our topologies were highly sensitive to choice of alignment parameters and methods. In two-step procedures (Fig. 1) and for a given alignment, different reconstruction methods yielded similar topologies that also had comparable bootstrap and likelihood support. The robustness of the phylogenetic signal depended first and foremost on the primary homology hypotheses of nucleotide positions. The reduced taxon sample in group 3 allowed alignment with low ambiguity and yielded topologies with few conflicts, whereas substantial conflicts in topologies prevailed in the larger taxon sets stemming from more ambiguities in their alignments. Furthermore, different reconstruction methods (MP, ML) found the same unlikely nodes, e.g. *Sphenodon* + *Leptobrachium*, *Ascaphus* + *Pelobatoidea*, when the same ambiguous alignment was put in.

The exclusion of gap sites and the coding of gaps as missing data have been discussed as solutions to the dilemma of ambiguously aligned positions (Lutzoni et al. 2000; Cognato and Vogler 2001). Coding gaps as missing data does not solve the more fundamental problem of ambiguous positions, because positional homology remains uncertain (Lutzoni et al. 2000). Also, gaps are a class of potentially informative characters states (Giribet and Wheeler 1999; Lutzoni et al. 2000; Simmons and Ochoterena 2000; Simmons et al. 2001). Coding gaps as either missing data or as a fifth character state had little general effect on inferred topologies in our study. Although, some indels, when coded as fifth character state, were identified as apomorphic character states for certain clades (e.g. *Bombina*, Neobatrachia).

The *de facto* down-weighting of variable aligned positions in the concatenated alignments led to higher resolution in consensus trees (see also Wheeler et al. 1995). This effect could be explained by, first, the higher number of parsimony informative characters that improve the resolution in the phylogenetic hypotheses, and secondly, the implicit weighting in concatenated alignments could reflect more appropriately the relations in substitution rates between fast- and slow-evolving sites.

In our data, resolution decreased with the exclusion of ambiguous positions (see also Cerchio and Tucker 1998; Giribet and Wheeler 1999; Lutzoni et al. 2000; Shull et al. 2001). The shortened sequences did not contain enough informative sites to maintain resolution. Different methods of recognition, delimiting and exclusion of ambiguous regions can introduce subjectivity and lead to conflicting topologies (Lutzoni et al. 2000). Yet, it is not clear to what extent the signal in highly variable regions is perturbed as a result of saturated substitutions and/or false hypothesis of primary homology.

The one-step procedure of POY provides an alternative to distance-based alignment procedures and subsequent tree search (Shull et al. 2001). Although, the method of combined analysis via optimization alignment deviates from the general principle of creating primary hypotheses of homology and subsequent independent tests of these hypotheses in phylogenetic reconstruction (De Pinna 1991; Simmons and Ochoterena 2000). Like all other current methods, POY requires the *a priori* specification of alignment parameters (gap cost). Setting gap cost is subjective; models for the evolution of insertions and deletions are not available (Kluge 1999; Hancock and Vogler 2000; Simmons and Ochoterena 2000). Parsimony-based alignment programs do not test ranges of gap cost in the search for the globally most parsimonious alignment (Shull et al. 2001).

In both one-step and two-step procedures, the variation of parameters and methods for alignment and tree reconstruction lead to numerous conflicting phylogenetic hypotheses. We found no general rule of thumb for setting 'correct' gap cost. In POY, for instance, low gap cost led to more congruent nodes in taxon groups 1 and 3, but more conflicting nodes in taxon group 2.

Suitability of the sequence data

In ribosomal genes, the rate of substitution is site-specific (loops versus stems, domains of tertiary structure of rRNA molecules; Simon et al. 1994). Ribosomal genes should contain information from old splitting events in their conserved regions while fast evolving parts should be useful to resolve more recent events, e.g. intraspecific or intrageneric (Simon et al. 1994). In our study, generally better resolved topologies were obtained when fast evolving sites were included. The degree of noise (saturated sites, wrong homologies) and its mode of distribution (randomly or non-randomly; see Naylor and Brown 1998) is unknown *a priori*; yet, the phylogenetic signal in noisy data sets may be detectable. (Wenzel and Sidall 1999; Broughton et al. 2000; Simmons et al. 2001; Simmons et al. 2002). Otherwise hidden support (Cognato and Vogler 2001) can emerge from the combined analysis of nucleotide stretches with unequal rate of evolution. 'Noise' in the sense of homoplasious characters can contribute to resolution of phylogenetic hypotheses, if an adequate number of terminal taxa were included (Simmons et al. 2002), but resolving deep divergence events requires true signal unmasked by multiple substitutions (Wägele et al. 1999; Wägele and Misof 2001).

In the 12S and 16S sequences examined, genetic distance and substitution plots indicate a high degree of saturation, no matter if the highly variable positions are included or not (Figs 6, 7). Even in the most conserved sequence regions, the phylogenetic information was disturbed by multiple substitutions, as was indicated by the non-linear slope of transitions and transversions (Fig. 7). We assume that the ambiguities of primary homology assessment of large parts of the sequences and noise in the data (in relation to the taxon sample) accounted for the numerous conflicting results and the high sensitivity of the phylogenetic hypotheses to different parameter settings. Particularly, ambiguities in primary homology assessment contribute to the extensive n-dimensional space of possible solutions. The preference of particular sets of analysis parameters remains subjective. Therefore, the use of such highly divergent ribosomal sequences must be considered carefully with respect to the phylogenetic problem in question.

Although, the monophyly of the Lissamphibia and the Anura were supported in numerous analyses, our results suggest that for deep splitting events *within* the Anura the ambiguities in alignments and phylogenetic reconstruction limit the suitability of the gene fragments examined in frog phylogeny (contrary to Hedges and Maxson 1993; Hay et al. 1995; Feller and Hedges 1998). The search for topologies that resolve the basal branching pattern in anurans was sensitive to

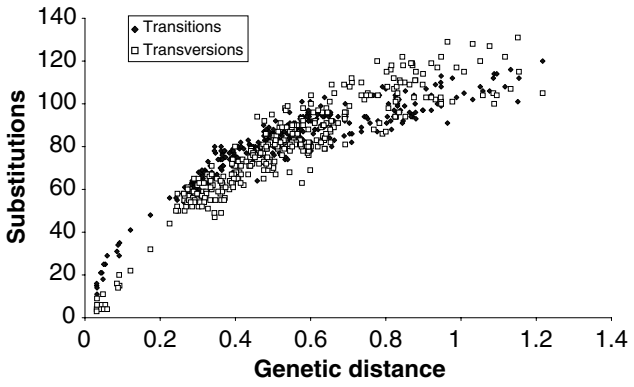

Fig. 6. Plot of pairwise genetic distances [corrected using TN model (Tamura and Nei 1993)] versus number of transitions and transversions based on CLUSTAL X CA alignment (GOP 15/GEP 5)

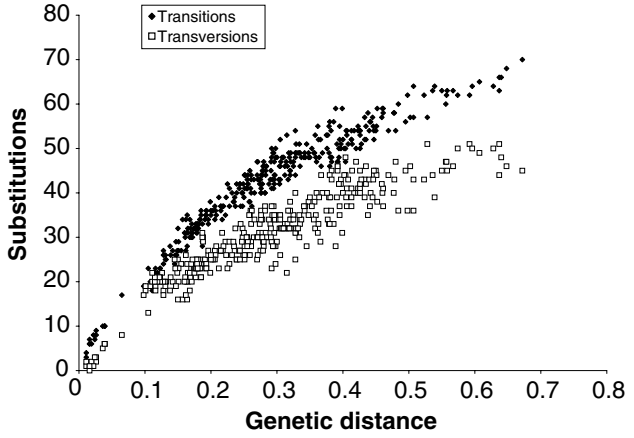

Fig. 7. Plot of pairwise genetic distances [corrected using TN model (Tamura and Nei 1993)] versus number of transitions and transversions based on CLUSTAL X alignment CE (ambiguously aligned positions excluded)

rather small changes in the conjectured homology of nucleotide positions. The limited resolving power of the data could be a consequence of high rates of evolution in both genes or different rates of substitutions between taxa (Simon et al. 1994), possibly responsible for the robust support of *Limnodynastes* + *Rana*. Moreover, it could also be an effect of long branch attraction between taxa with a long history of isolated evolution, such as *Ascaphus* (Swofford et al. 1996; Huelsenbeck 1997; Wiens and Hollingsworth 2000). Such effect is indicated by the varying position of *Ascaphus* in the space of topologies, or the clade *Sphenodon* + *Leptobrachium* in some topologies. A further cause could be a fast radiation of early anurans relative to the substitution rate of our sequences.

We found a robust hypothesis of relationships for species group 3. In the more inclusive groups 1 and 2, the phylogenetic signal seems to be weak for Mesozoic anuran cladogenesis (see also Simon et al. 1994).

Phylogenetic conclusions

The monophyly of the Anura is well supported by numerous apomorphic characters of adults and larvae (Duellman and Trueb 1986; Ford and Cannatella 1993; Haas 1997, 2001, 2003). Previous studies of molecular data corroborated this hypothesis (Hedges and Maxson 1993; Hillis et al. 1993; Hay et al. 1995; Feller and Hedges 1997). The Anura was clearly supported in nearly all of our species group 2 analyses but not in group 1 analyses, which encompassed highly divergent outgroup taxa. We could not reconstruct congruent and robust phylogenetic relationships between the major clades within the Anura. The major clades represented by our taxa likely have split long time ago and long times of separate evolution may have obscured signal.

Previous analyses of molecular data, although in some cases with low support, argued for the monophyly of the Archaeobatrachia consisting of *Ascaphus, Leiopelma, Discoglossidae, Pelobatidae* (*s. lato*), *Pipidae, Pelodytidae* and *Rhinophrynidae* (Hedges and Maxson 1993; Hay et al. 1995; Feller and Hedges 1997). Ford and Cannatella (1993); Hillis et al. (1993) and Haas (1997, 2003) identified archeobatrachians as a paraphyletic group. In our study, there was weak support in only some analyses for the Archaeobatrachia (see Tables 4 and 5). We consider 12S and 16S sequences not suitable to answer this question.

The Pipidae is a well-supported clade within the Anura (Sokol 1977; de Sá and Hillis 1990; Cannatella and de Sá 1993; Ford and Cannatella 1993; Hay et al. 1995; Feller and Hedges 1998; Haas 2003). In our analyses, the monophyly of the Pipidae was robustly supported in many analyses under a broad range of methods and parameter settings.

Numerous studies treated discoglossids as natural group (e.g. Duellman 1975; Laurent 1979; Duellman and Trueb 1986; Sanchiz 1998). *Alytes* and *Barbourula* have not been included in previous molecular studies. Nearly all possible arrangements of discoglossid genera, including paraphyly, had their advocates in previous studies (Lanza et al. 1975; Feller and Hedges, 1998; Maxson and Szymura 1979, 1984; Ford and Cannatella 1993; Hay et al. 1995). In our study, *Alytes* + *Bombina* + *Discoglossus* (monophyletic Discoglossidae) was robustly supported in many of the analyses. Our results obtained from species groups 1 and 2 were ambiguous with regard to intradiscoglossid relationships, whereas analyses of species group 3 gave robust support for a clade *Alytes* + *Bombina* within a monophyletic and robustly supported Discoglossidae. Yet, changing the outgroup composition led to conflicting hypotheses and weakened support for the monophyly of Discoglossidae and *Alytes* + *Bombina* (Table 7). The gene sequences used could not solve the case convincingly.

The Pelobatoidea traditionally includes the Pelodytidae and the Pelobatidae *s. lato* (Duellman 1975; Duellman and Trueb 1986; Ford and Cannatella 1993; Lathrop 1997; Maglia 1998; Sanchiz 1998). The Pelobatidae *s. lato* includes the Pelobatinae (*Pelobates, Spea, Scaphiopus*), Megophryinae and the extinct Eopelobatinae (Duellman and Trueb 1986; Maglia 1998; Sanchiz 1998). Overall, there was support for the monophyly of the Pelobatoidea in a number of analyses. The Pelobatidae *s. str.* containing the Eurasian genera *Pelobates* and *Leptobrachium*, however, was a well-supported clade. The North American *Scaphiopus* and the European *Pelobates* never formed a monophyletic group in our analyses, in contrast to traditional groupings. The results did not resolve the relationships within the Pelobatoidea unambiguously with regard to scaphiopodids, pelodytids and the Eurasian Pelobatidae.

The monophyly of the Neobatrachia, comprising the majority of extant frogs, is widely accepted and supported by

morphological and molecular evidence (Duellman and Trueb 1986; Ford and Cannatella 1993; Hedges and Maxson 1993; Hillis et al. 1993; Hay et al. 1995; Ruvinsky and Maxson 1996; Feller and Hedges 1998; Sanchiz 1998). Despite only two neobatrachians genera were included in the present samples, the clade *Limnodynastes* + *Rana* was robust under a wide range of conditions.

Conclusions

Owing to the complexity of alignments of highly divergent RNA sequences, the current lack of models for the evolution of indels, and the various approaches for phylogenetic reconstruction it is necessary to explore the n-dimensional space of analysis parameters and phylogenetic hypothesis. The space of parameter dependent topologies should be searched for universally supported nodes. Such procedure will lead to rather conservative hypotheses (see also Wheeler 1995; Whiting et al. 1997; Phillips et al. 2000). The search for robust phylogenetic hypotheses makes us considerably more cautious than previous workers to infer the early phylogeny of frogs from 12S to 16S ribosomal genes. Our analyses gave a heterogenous and rather complex picture of noise versus signal. Noise in the data and particularly uncertainties of primary homology necessarily produced numerous conflicting results. Only very few nodes were supported universally under a wide range of *a priori* decisions and analysis paths.

Acknowledgements

The authors are grateful to J. Reiss, K.-D. Kuehnel, S. Voitel, H.-J. Herrmann, N. Sousa, F. Glaw (ZSM), W. Böhme (ZFMK), as well as zoological gardens Stuttgart and Chemnitz for generously provided specimens used in this study. The authors are grateful to M. Goedel for his practical help with laboratory techniques. The authors thank M. Svensson, S. Erlacher, R. Beutel and two anonymous reviewers for their helpful comments on the manuscript. Part of this study was supported by Deutsche Forschungsgesellschaft grant Ha 2323/6-1 to A. Haas and NSF Grant DEB 9815787 to R.O. de Sá.

Zusammenfassung

Phylogenetisches Signal und Eignung von 12S und 16S mtDNA für die Phylogenie der Froschlurche

Zur Anwendung in einer phylogenetische Analyse müssen die ausgewählten Gene konservierte und detektierbare Information zum untersuchten phylogenetischen Niveau enthalten. Ribosomale Gene des Mitochondriums wurden für ein breites Spektrum phylogenetischer Fragestellungen bei verschiedenen Gruppen und insbesondere bei Froschlurchen eingesetzt. Wir untersuchten die Frage, ob Rekonstruktionen der Anuren-Phylogenie, basierend auf 12S und 16S rRNA Gensequenzen, plausibel sind. An einer Auswahl von 27 Arten, arrangiert in Taxa-Gruppen abgestufter Hierarchie, rekonstruierten wir phylogenetische Hypothesen unter verschiedenen, *a priori* festgelegten Bedingungen. Dazu gehörten die Auswahl verschiedener Alinierungsmethoden und—parameter, der Umgang mit variabel alinierten Positionen, die Auswahl der Algorithmen zur Baumkonstruktion sowie die Auswahl alternativer Modelle der Sequenzevolution. Die Methoden und Parameter der Alinierung und der Rekonstruktion, sowie die Auswahl der Taxa, hatten bedeutenden Einfluss auf die Resultate. Daraus resultierte eine große Anzahl alternativer Topologien, in denen nur sehr wenige Knoten in allen Analysen Unterstützung fanden. Ausschluss variabel alinierter Positionen ergaben Topologien mit niedrigem Grad der Auflösung. Die Sequenzen enthielten ein gewisses Signal für die Monophylie von Discoglossidae, Pelobatoidea, Pelobatidae und Pipidae. Der Knoten Neobatrachia wurde deutlich unterstützt. Die robuste Auflösung intragenerischer Phylogenien von *Bombina* und *Discoglossus* weisen auf eine besondere Eignung der Gene für die Untersuchung junger Aufspaltungsereignisse hin. Obwohl 12S und 16S rRNA-Gene eine heterogene Unterstützung für wenige frühe (mesozoische) phylogenetische Ereignisse zeigten, war das Signal nicht geeignet, um die Beziehungen der Taxa höherer Ordnung der Anura unter variierten Parametern und Analysemethoden konsistent aufzulösen.

References

Alves-Gomes, J. A.; Orti, G.; Haygood, M.; Heiligenberg, W.; Meyer, A., 1995: Phylogenetic analysis of the South American electric fishes and the evolution of their electronic system: a synthesis based on morphology, electrophysiology and mitochondrial sequence data. Mol. Biol. Evol. **12,** 298–318.

Armbruster, G. F. J., 2001: Temperature based variation of rRNA secondary stucture models: a case study in the insect *Drosophila simulans*, the land snail *Isabellaria adriani*, and the crustacean *Daphnia pulex*. Can. J. Zool. **79,** 334–345.

Bininda–Emonds, O. R. P.; Bryant, H. N.; Russell, A. P., 1998: Supraspecific taxa as terminals in cladistic analysis: implicit assumptions of monophyly and a comparison of methods. Biol. J. Linn. Soc. **64,** 101–133.

Broughton, R. E.; Stanley, S. E.; Durrett, R. T., 2000: Quantification of homoplasy for nucleotide transitions and transversions and a re-examination of assumptions in weighted phylogenetic analysis. Syst. Biol. **49,** 617–627.

Buckley, T. R.; Arensburger, P.; Simon, C.; Chambers, G. K., 2002: Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. Syst. Biol. **5,** 4–18.

Cannatella, D. C.; de Sá, R. O., 1993: *Xenopus laevis* as a model organism. Syst. Zool. **42,** 476–507.

Cerchio, S.; Tucker, P., 1998: Influence of alignment on the mtDNA phylogeny of Cetacea: questionable support for a Mysticeti/Physeteroidea clade. Syst. Biol. **47,** 336–344.

Clough, M.; Summers, K., 2000: Phylogenetic systematic and biogeography of the poison frogs: evidence from mitochondrial DNA sequences. Biol. J. Linnean Society **70,** 515–540.

Cognato, A. I.; Vogler, A. P., 2001: Exploring data interaction and nucleotide alignment in a multiple gene analysis of Ips (Coleoptera: Scolytinae). Syst. Biol. **50,** 758–780.

Dawood, A.; Channing, A.; Bogart, J. P., 2002: A molecular phylogeny of the frog genus Tomopterna in Southern Africa: examining species boundaries with mitochondrial 12S rRNA sequence data. Mol. Phyl. Evol. **22,** 407–413.

De Pinna, M. C. C., 1991: Concepts and tests of homology in the cladistic paradigm. Cladistics **7,** 367–394.

De Sá, R. O.; Hillis, D. M., 1990: Phylogenetic relationships of the Pipid Frogs *Xenopus* and *Silurana*: an integrating of ribosomal DNA and morphology. Mol. Biol. Evol. **7,** 365–376.

Dixon, M. T.; Hillis, D. M., 1993: Ribosomal secondary structure: compensatory mutations and implications for phylogenetic analysis. Mol. Biol. Evol. **10,** 256–267.

Duellman, W. E., 1975: On the classification of frogs. Occ. Pap. Mus. Nat. Hist. Univ. Kansas **42,** 1–15.

Duellman, W. E.; Trueb L.: 1986; Biology of amphibians. London: John Hopkins Press.

Emerson, S. B.; Inger, R. F.; Iskandar, D., 2000: Molecular systematics of the Fanged Frogs of Southeast Asia. Mol. Phyl. Evol. **16,** 131–142.

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

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

Felsenstein, J., 1989: PHYLIP – phylogeny inference package. Cladistics **5,** 164–166.

Ford, L. S.; Cannatella, D. C., 1993: The major clades of frogs. Herpetol. Monogr. **6,** 94–117.

Garcia-Paris, M.; Jockusch, E. L., 1999: A mitochondrial DNA perspective on the evolution of Iberian *Discoglossus* (Amphibia: Anura). J. Zool. Lond. 209–218.

Gatesy, J.; DeSalle, R.; Wheeler, W., 1993: Alignment–ambiguous nucleotide sites and the exclusion of systematic data. Mol. Phyl. Evol. **2,** 152–157.

Giribet, G.; Wheeler, W. C. 1999: On gaps. Mol. Phyl. Evol. **13,** 132–143.

Gladstein, D.; Wheeler, W. C., 1996: POY. Program and documentation. American Museum of Natural History.

Goebel, A. M.; Donnelly, J. M.; Atz, M. E., 1999: PCR primers and amplification methods for 12S DNA, the control region, cytochrome oxidase 1 and cytochrome b in Bufonids and other frogs, and an overview of PCR primers which have amplified DNA in amphiniands successfully. Mol. Phyl. Evol. **11,** 163–199.

Grandcolas, P.; D'Haese C., 2001: The phylogeny of cockroach families: is the current molecular hypothesis robust? Cladistics **17,** 48–55.

Graybeal, A., 1993: The phylogenetic utility of cytochrome b: lessons from bufonid frogs. Mol. Phyl. Evol. **2,** 256–269.

Graybeal, A., 1997: Phylogenetic relationships of bufonid frogs anf tests of alternative macroevolutionary hypotheses characterizing their radiation. Zool. J. Linn. Soc. **119,** 9–17.

Graybeal, A., 1998: Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. **47,** 9–17.

Haas, A., 1997: The larval hyobranchial apparatus of discoglossoid frogs: its structure and bearing on the systematics of the Anura (Amphibia: Anura). J. Zool. Syst. Evol. Res. **53,** 179–197.

Haas, A., 2001: Mandibular arch musculature of anuran tadpoles, with comments on homologies of amphibian jaw muscles. J. Morph. **247,** 1–33.

Haas, A., 2003: Phylogeny of frogs as inferred from primarily larval characters (Amphibia: Anura). Cladistics **19,** 23–89.

Hall, T. A., 1999: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser **41,** 95–98.

Hancock, J. M.; Vogler, A. P., 2000: How slippage-derived sequences are incorporated into rRNA variable-region secondary structure: implications for phylogeny reconstruction. Mol. Phyl. Evol. **14,** 366–374.

Hasegawa, M.; Kishino, H.; Yano, T., 1985: Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Biol. **22,** 160–174.

Hay, J. M.; Ruvinsky, I.; Hedges, S. B.; Maxson, L. R., 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., 1992: The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies. Mol. Biol. Evol. **9,** 366–369.

Hedges, S. B.; Maxson, L. R., 1993: A molecular perspective on Lissamphibian phylogeny. Herpetol. Monogr. **6,** 27–41.

Higgins, D. G.; Thompson, J. D.; Gibson, T. J., 1996: Using CLUSTAL for multiple sequence alignments. Meth. Enzymol. **266,** 383–402.

Hillis, D. M., 1998: Taxonomic Sampling, phylogenetic accuracy and investigator bias. Syst. Biol. **47,** 3–17.

Hillis, D. M.; Ammerman, L. K.; Dixon, M. T.; Sá, R. O. D., 1993: Ribosomal DNA and the phylogeny of frogs. Herpetol. Monogr. **7,** 118–131.

Hillis, D. M.; Mable, B. K.; Larson, A.; Davis, S. K.; Zimmer, E. A., 1996; Sequencing and cloning. In: Hillis, D. M.; Moritz C.; Mable B. K. (eds), Molecular Systematics Sunderland, Massachusetts: Sinauer.

Huelsenbeck, J. P., 1997: Is the Felsenstein zone a fly trap? Syst. Biol. **46,** 69–74.

Huelsenbeck, J. P., 2000: MrBAYES: Bayesian inference of phylogeny Distributed by the author. New York: Department of Biology, Univ. of Rochester.

Huelsenbeck, J. P.; Ronquist, F. R. 2001: MrBayes: Bayesian inference of phylogeny. Biometrics **17,** 754–755.

Jaeger, J. A.; H.Turner, D.; Zuker, M., 1990: Predicting optimal and suboptimal secondary structure for RNA. Meth. Enzymol 183.

Kjer, K. M., 1995: Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. Mol. Phyl. Evol. **4,** 314–330.

Kjer, K. M.; Blahnik, R. J.; Holzenthal, R. W., 2001: Phylogeny of Trichoptera (Caddisflies): characterization of signal and noise within multiple data sets. Syst. Biol. **50,** 781–816.

Kluge, A. G., 1999: The science of phylogenetic systematics: explanation, prediction and test. Cladistics **15,** 429–436.

Kumar, S.; Tamura, K.; Jakobsen, I. B.; Nei, M., 2001: MEGA2: Molecular Evolutionary Genetics Analysis software, Arizona State University, Tempe, Arizona, USA. Bioinformatics **17,** 1244–1245.

Lanave, C.; Preparata, G.; Saccone, C.; Serio, G., 1984: A new method calculating evolutionary substtution rates. J. Mol. Evol. **20,** 86–93.

Lanza, B.; Cei, J. M.; Crespo, E., 1975: Immunological evidence for the specific status of Discoglossus pictus Otth, 1837 and D sardus Tschudi, 1837, with notes on the families Discoglossidae Günther, 1858 and Bombinidae fitzinger, 1826 (Amphibia: Salentia). Monitore Zool. Ital. (N. S.) **9,** 153–162.

Lathrop, A., 1997: Taxonomic review of the megophryid frogs (Anura: Pelobatoidea). Asiatic Herpetol. Res. **7,** 68–79.

Laurent, R. F., 1979: Esquisse d'une phylogenèse des anoures. Bull. Soc. Zool. France **104,** 397–422.

Leaché, A. D.; Reeder, T. W., 2002: Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of parsimony, likelihood, and Bayesian approaches. Syst. Biol. **51,** 44–68.

Lecointre, G.; Philippe, H.; Le, H. L. V.; Guayader, H. l., 1993: Species sampling has a major impact on phylogenetic inference. Mol. Phyl. Evol. **2,** 205–224.

Lutzoni, F.; Wagner, P.; Reeb, V.; Zoller, S., 2000: Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst. Biol. **49,** 628–651.

Maddison, D. R.; Maddison, W. P., 2000: MacClade 4.0. Sunderland, Massechusetts: Sinauer Associates.

Maglia, A. M., 1998: Phylogenetic relationships of extant Pelobatoid frogs (Anura: Pelobatoidea): evidence from adult morphology. Nat. Hist. Mus. Univ. Kansas Sci. Papers **10,** 1–19.

Maniatis, T.; Fristch, E. F.; Sambrook, J., 1982; Molecular cloning: a laboratorial manual. Cold Spring Harbor, New York: Cold Spring Harbor Publications.

Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D., 1999: Expanded sequence dependence of thermodynamic parameters improves prediction of RNA seqondary structure. J. Mol. Biol. **288,** 911–940.

Mattern, M. Y.; McLennan, D. A., 2000: Phylogeny and speciation of felids. Cladistics **16,** 232–253.

Maxson, L. E. R.; Szymura, J. M., 1979: Quantitative immunological studies of the albumins of several species of fire bellied toads, genus Bombina. Comp. Biochem. Physiol. **63 B,** 517–519.

Maxson, L. E. R.; Szymura, J. M., 1984: Relationships among discoglossid frogs: an albumin perspective. Amphibia—Reptilia **5,** 245–252.

Mindell, D. P.; Honeycutt, R. L., 1990: Ribosomal RNA in vertebrates: evolution and phylogenetic implications. Annu. Rev. Ecol. Syst. **21,** 541–566.

Naylor, G. J.; Brown, W. M., 1998: Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. Syst. Biol. **47,** 61–76.

Nicholas, K. B.; Nicholas, H. B., 1997: GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author. http://www.psc.edu/biomed/genedoc

Orti, G.; Meyer, A., 1997: The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. Syst. Biol. **46,** 75–100.

Orti, G.; Petry, P.; Porto, J. I. R.; Jegu, M.; Meyer, A., 1996: Patterns of nucleotide change in mitochondrial ribosomal genes and the phylogeny of piranhas. J. Mol. Evol. **42,** 169–182.

Page, R. D. M., 1996: TREEVIEW: an application to display phylogenetic trees on personal computers. Comp. Applic. Bios. **12,** 357–358.

Palumbi, S. R., 1996; The polymerase chain reaction. In: Hillis, D. M.; Moritz, C.; Mable B. K. (eds), Molecular Systematics, Sunderland, Massachusetts: Sinauer.

Phillips, A.; Janies, D.; Wheeler, W., 2000: Multiple sequence alignment in phylogenetic analysis. Mol. Phyl. Evol. **16,** 317–330.

Posada, D.; Crandall, K. A., 1998: Modeltest: testing the model of DNA substitution. Bioinformatics **14,** 817–818.

Rannala, B.; Yang, Z., 1996: Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. J. Mol. Evol. **43,** 304–311.

Reig, O. A., 1958: Propositiones para una nueva macrosistematica de los anuros. Nota preliminar Physis **21,** 109–118.

Richards, C. M.; Moore, W., 1998: A molecular phylogeny of the old world tree frog family Racophoridae. J. Herpetol. **8,** 41–46.

Rodriguez, F.; Oliver, J. L.; Marin, A.; Medina, J. R., 1990: The general stochastic model of nucleotide substitution. J. Theor. Biol. **142,** 485–501.

Roe, B. A.; Ma, D. P.; Wilson, R. K.; Wong J. F., 1985: The complete nucleotide sequence of the Xenopus laevis mitochondrial genome. J. Biol. Chem. **260,** 9759–9774.

Ruvinsky, I.; Maxson, L. A., 1996: Phylogenetic relationships among bufonid frogs (Anura: Neobatrachia) inferred from mitochondrial DNA sequences. Mol. Phyl. Evol. **5,** 533–547.

Sanchiz, B., 1998; Encyclopedia of Paleoherpetology: Salentia, München: Verlag, Dr. F. Pfeil, München.

Sankoff, D. D.; Rousseau, P., 1975: Locating the vertices of a Steiner tree in arbitrary space. Math. Prog. **9,** 240–246.

Severini, C.; Silvestrini, F.; Mancini, P.; Rosa, G. L.; Marinucci, M., 1996: Sequence and secondary structure of the rRNA second internal transcribed spacer in the sibling species *Culex pipiens* and *C. quinquefasciatus* (Diptera: Culicidae). Insect Mol. Biol. **5,** 181–186.

Shull, V. L.; Vogler, A. P.; Baker, M. D.; Maddison, D. R.; Hammond, P. M., 2001: Sequence alignment of 18S ribosomal RNA and the basal relationships of adephagan beetles: evidence for monophyly of aquatic families and the placement of Trachypachidae. Syst. Biol.50**,** 945–969.

Simmons, M. P.; Ochoterena, H., 2000: Gaps in sequence-based phylogenetic analysis. Syst. Biol. **49,** 369–381.

Simmons, M. P.; Ochoterena, H.; Carr, T. C., 2001: Incorporation, relative homoplasy and effect of gap characters in sequence-based phylogenetic analyses. Syst. Biol. **50,** 454–462.

Simmons, M. P.; Ochoterena, H.; Freudenstein, J. V., 2002: Amino acids versus nucleotide characters: challenging preconceived notions. Mol. Phyl. Evol. **24,** 78–90.

Simon, C.; Friati, F.; Beckenbach, A.; Crespi, B.; Liu, H.; Flook, P., 1994: Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. **87,** 651–686.

Sokol, O. M., 1977: A subordinal classification of frogs (Amphibia: Anura). J. Zool. Lond. **182,** 505–508.

Strimmer, K.; von Haeseler, A. V., 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). Sunderland, Massachusetts: Sinauer Ass.

Swofford, D. L.; Olsen, G. J.; Waddell, P. J.; Hillis, D. M., 1996; Phylogeny inference. In: Hillis, D. M.; Moritz, C.; Mable, B. K. (eds), Molecular Systematics, Sunderland, Massachusetts: Sinauer.

Tamura, K.; Nei, M., 1993: Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10,** 512–526.

Thompson, J. D.; Higgins, D. G.; Gibson, T. J., 1994: CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucl. Acids Res. **22,** 4673–4680.

Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G., 1997: The CLUSTALX windows interface: flexible strategies for multiple sequence alignments aided by quality analysis tools. Nucl. Acids Res. **24,** 4876–4882.

Titus, T. A.; Frost, D. R., 1996: Molecular homology assessment and phylogeny in the lizard family Opluridae (Squamata: Iguania). Mol. Phyl. Evol. **6,** 49–62.

Vences, M.; Kosuch, J.; Lötters, S.; Widmer, A.; Jungfer, K. H.; Köhler, J.; Veith, M., 2000: Phylogeny and classification of poison frogs (Amphibia: Dendrobatidae), based on mitochondrial 16S and 12S ribosomal gene sequences. Mol. Phyl. Evol. **15,** 34–40.

Vogler, A. P.; DeSalle, R., 1994: Evolution and phylogenetic information content of the ITS-2 region of the tiger beetle *Cicindela dorsalis*. Mol. Biol. Evol. **11,** 393–405.

Wägele, W.; Staniek, G., 1995: Arthopod phylogeny inferred from partial 12S rRNA revisited: monophyly of Tracheata depends on sequence alignment. Zool. Syst. Evol. Research **33,** 75–80.

Wägele, J. W.; Erikson, T.; Lockhart, P.; Misof, W. B., 1999: The Ecdysozoa: artifact or monophylum?. Journal of Zoological Systematics and Evolutionary Research **37,** 211–223.

Wägele, J. W.; Misof, B., 2001: On quality of evidence in phylogeny reconstruction: A reply to Zrzavy's defence of the 'Ecdysozoa' hypothesis. Journal of Zoological Systematics and Evolutionary Research **39,** 165–176.

Wenzel, J. W.; Sidall, M. E., 1999: Noise. Cladistics **15,** 51–64.

Wheeler, W., 1995: Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. Syst. Biol. **44,** 321–331.

Wheeler, W. C., 1996: Optimization alignment: the end of multiple sequence alignment in phylogenetics. Cladistics **12,** 1–9.

Wheeler, W. C., 1999: Fixed character states and the optimization of molecular sequence data. Cladistics **15,** 379–385.

Wheeler, W. C., 2001: Homology and the optimization of DNA sequence data. Cladistics **17,** S3–S11.

Wheeler, W. C.; Honeycutt, R. L., 1988: Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implication. Mol. Biol. Evol. **5,** 90–96.

Wheeler, W. C.; Gatesy, J.; DeSalle, R., 1995: Elision: a method for accommodating multiple molecular sequence alignments with alignment-ambiguous sites. Mol. Phyl. Evol. **4,** 1–9.

Whiting, M. F.; Carpenter, J. C.; Wheeler, Q. D.; Wheeler, W. C., 1997: The strepsiptera problem: phylogeny of the insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Syst. Biol. **46,** 1–67.

Wiens, J. J.; Hollingsthworth, B. D., 2000: War of the iguana: conflicting molecular and morphological phylogenies and long-branch attraction in iguanid lizards. Syst. Biol. **49,** 143–159.

Wiens, J. J.; Reeder, T. W., 1997: Phylogeny of the spiny lizards (Sceloporus) based on molecular and morphological evidence. Herpetol. Monogr. **11,** 1–101.

Yang, Z., 1994: Estimating the pattern of nucleotide substitution. J. Mol. Biol. **39,** 306–314.

Yang, Z.; Rannala, B., 1997: Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. Mol. Biol. Evol. **14,** 717–724.

Yeates, D. K., 1995: Groundplans and exemplars: paths to the tree of life. Cladistics **11,** 345–357.

Zardoya, R.; Meyer, A., 2000: Mitochondrial evidence on the phylogenetic position of Caecilians (Amphibia: Gymnophiona). Genetics **155,** 765–757.

Zardoya, R.; Meyer, A., 2001: On the origin of and phylogenetic relationships among living amphibians. PNAS **98,** 7380–7383.

Zuker, M.; Stiegler, P., 1981: Optimal computer folding of large RNA sequences using the rmodynamics and auxiliary information. Nucl. Acids Res. **9,** 133–148.

Appendix 1

GenBank accession numbers off all sequences examined. AJ numbers refer to our own sequencing

Species	12S	16S
Alytes muletensis	AJ440758	AJ440797
Alytes obstetricans	AJ440759	AJ440798
Ambystoma mexicanum	Y10947	Y10947
Ascaphus truei	X86225, AJ440760	X86293, AJ440799
Bombina maxima	AJ440761	AJ440800
Bombina bombina	AJ440762	AJ440801
Bombina orientalis	AJ440763	AJ440802
Bombina variegata	AJ440764	AJ440803
Discoglossus galganoi	AJ440765	AJ440804
Discoglossus montalentii	AJ440766	AJ440805
Discoglossus pictus	X86235, AJ440767	AJ440806
Discoglossus sardus	AJ440768	AJ440807
Ichthyophis bannanicus	Y10949	Y10949
Lepidosiren paradoxa	Z48715	Z48715
Leptobrachium spec.	AJ440769	AJ440808
Limnodynastes dorsalis	AF261250	AF261268
Limnodynastes peronii	AJ440770	AJ440809
Neoceratodus forsteri	AF302933	AF302933
Pelodytes caucasicus	AJ440771	AJ440810
Pelodytes punctatus	X86236, AJ440772	AJ440811
Pelobates fuscus	AJ440773	AJ440812
Pipa carvalhoi	AJ440774	AJ440813
Rana catesbeiana	X12841	X12841
Rana nigrovittata	AJ440775	AJ440814
Scaphiopus couchii	AJ440776	AJ440815
Sphenodon punctatus	L28076	L28076
Xenopus laevis	M10217	M10217, AJ440816

Appendix 2

Composition of the Taxon Groups (see text for further explanation)

Group 1 (Lissamphibia, Amniota, Dipnoi)
Alytes muletensis, *Alytes obstetricans*, *Ambystoma mexicanum*, *Ascaphus truei*, *Bombina bombina*, *Bombina maxima*, *Bombina orientalis*, *Bombina variegata*, *Discoglossus galganoi*, *Discoglossus montalentii*, *Discoglossus pictus*, *Discoglossus sardus*, *Ichthyophis bannanicus*, *Lepidosiren paradoxa*, *Leptobrachium sp.*, *Limnodynastes dorsalis*, *Limnodynastes peronii*, *Neoceratodus forsteri*, *Pelobates fuscus*, *Pelodytes punctatus*, *Pelodytes caucasicus*, *Pipa carvalhoi*, *Rana catesbeiana*, *Rana nigrovittata*, *Scaphiopus couchi*, *Sphenodon punctatus*, *Xenopus laevis*.

Group 2 (Lissamphibia)
Alytes muletensis, *Alytes obstetricans*, *Ambystoma mexicanum*, *Ascaphus truei*, *Bombina bombina*, *Bombina maxima*, *Bombina orientalis*, *Bombina variegata*, *Discoglossus galganoi*, *Discoglossus montalentii*, *Discoglossus pictus*, *Discoglossus sardus*, *Ichthyophis bannanicus*, *Leptobrachium sp.*, *Limnodynastes dorsalis*, *Limnodynastes peronii*, *Pelobates fuscus*, *Pelodytes punctatus*, *Pelodytes caucasicus*, *Pipa carvalhoi*, *Rana catesbeiana*, *Rana nigrovittata*, *Scaphiopus couchi*, *Xenopus laevis*.

*Group 3 (*Discoglossidae + *Ascaphus, Pelodytes)*
Alytes muletensis, *Alytes obstetricans*, *Ascaphus truei*, *Bombina bombina*, *Bombina maxima*, *Bombina orientalis*, *Bombina variegata*, *Discoglossus galganoi*, *Discoglossus montalentii*, *Discoglossus pictus*, *Discoglossus sardus*, *Pelodytes caucasicus*.

Authors' addresses: Alexander Haas, (for correspondence), Stefan Hertwig, Institut für Spezielle Zoologie und Evolutionsbiologie mit Phyletischem Museum, Friedrich-Schiller-Universität Jena, Erbertstr. 1, 07743 Jena, Germany. E-mail: B5HAAL@rz.uni-jena.de; Rafael O. de Sá, University of Richmond, Department of Biology, Richmond, VA 23173, USA.