Syst. Biol. 0(0):1–17, 2020

DOI:10.1093/sysbio/syaa034

Phylogenomics Reveals Ancient Gene Tree Discordance in the Amphibian Tree of Life

PAUL M. HIME[1,2,*,†], ALAN R. LEMMON[3,*,†], EMILY C. MORIARTY LEMMON[4], ELIZABETH PRENDINI[5], JEREMY M. BROWN[6], ROBERT C. THOMSON[7], JUSTIN D. KRATOVIL[2,8], BRICE P. NOONAN[9], R. ALEXANDER PYRON[10], PEDRO L. V. PELOSO[5,11], MICHELLE L. KORTYNA[4], J. SCOTT KEOGH[12], STEPHEN C. DONNELLAN[13,14], RACHEL LOCKRIDGE MUELLER[15], CHRISTOPHER J. RAXWORTHY[5], KRUSHNAMEGH KUNTE[16], SANTIAGO R. RON[17], SANDEEP DAS[18], NIKHIL GAITONDE[16], DAVID M. GREEN[19], JIM LABISKO[20,21], JING CHE[22,23] AND DAVID W. WEISROCK[2,*]

[1]*Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA;* [2]*Department of Biology, University of Kentucky, Lexington, KY 40506, USA;* [3]*Department of Scientific Computing, Florida State University, Tallahassee, FL 32306, USA;* [4]*Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA;* [5]*Division of Vertebrate Zoology: Herpetology, American Museum of Natural History, New York, NY 10024, USA;* [6]*Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA;* [7]*School of Life Sciences, University of Hawai'i, Honolulu, HI 96822, USA;* [8]*Department of Entomology, University of Kentucky, Lexington, KY 40546, USA;* [9]*Department of Biology, University of Mississippi, Oxford, MS 38677, USA;* [10]*Department of Biological Sciences, The George Washington University, Washington, DC 20052, USA;* [11]*Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, 66075-750, Brazil;* [12]*Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, 2601, Australia;* [13]*South Australian Museum, North Terrace, Adelaide 5000, Australia;* [14]*School of Biological Sciences, University of Adelaide, Adelaide 5005, Australia;* [15]*Department of Biology, Colorado State University, Fort Collins, CO 80523, USA;* [16]*National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560065, India;* [17]*Museo de Zoología, Escuela de Biología, Pontificia Universidad Católica del Ecuador, Quito, Ecuador;* [18]*Forest Ecology and Biodiversity Conservation Division, Kerala Forest Research Institute, Peechi, Kerala 680653, India;* [19]*Redpath Museum, McGill University, Montreal, Quebec H3A 0C4, Canada;* [20]*The Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, The University of Kent, Canterbury, Kent, CT2 7NR, UK;* [21]*Island Biodiversity and Conservation Centre, University of Seychelles, PO Box 1348, Anse Royale, Mahé, Seychelles; and* [22]*State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Kunming 650223, China; and* [23]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China*
**Correspondence to be sent to: Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA;*
E-mail: paul.hime@ku.edu.
Paul M. Hime and Alan R. Lemmon contributed equally to this article.

Received 26 April 2019; reviews returned 14 April 2020; accepted 16 April 2020
Associate Editor: Adam Leaché

Abstract.—Molecular phylogenies have yielded strong support for many parts of the amphibian Tree of Life, but poor support for the resolution of deeper nodes, including relationships among families and orders. To clarify these relationships, we provide a phylogenomic perspective on amphibian relationships by developing a taxon-specific Anchored Hybrid Enrichment protocol targeting hundreds of conserved exons which are effective across the class. After obtaining data from 220 loci for 286 species (representing 94% of the families and 44% of the genera), we estimate a phylogeny for extant amphibians and identify gene tree–species tree conflict across the deepest branches of the amphibian phylogeny. We perform locus-by-locus genealogical interrogation of alternative topological hypotheses for amphibian monophyly, focusing on interordinal relationships. We find that phylogenetic signal deep in the amphibian phylogeny varies greatly across loci in a manner that is consistent with incomplete lineage sorting in the ancestral lineage of extant amphibians. Our results overwhelmingly support amphibian monophyly and a sister relationship between frogs and salamanders, consistent with the Batrachia hypothesis. Species tree analyses converge on a small set of topological hypotheses for the relationships among extant amphibian families. These results clarify several contentious portions of the amphibian Tree of Life, which in conjunction with a set of vetted fossil calibrations, support a surprisingly younger timescale for crown and ordinal amphibian diversification than previously reported. More broadly, our study provides insight into the sources, magnitudes, and heterogeneity of support across loci in phylogenomic data sets.[AIC; Amphibia; Batrachia; Phylogeny; gene tree–species tree discordance; genomics; information theory.]

Understanding the extent of gene tree-species tree discordance in empirical studies can have important ramifications beyond the resolution of the species tree (Page and Charleston 1997; Degnan and Rosenberg 2006). Gene tree discordance has numerous causes (Degnan and Rosenberg 2009; Edwards 2009), and when a product of incomplete lineage sorting (ILS) or introgressive hybridization, can affect downstream interpretations about organismal evolution by obscuring divergence time estimates (Burbrink and Pyron 2011), or inducing apparent (but spurious) substitution rate variation (Mendes and Hahn 2016). A focus on the species tree and not on its discordant gene trees can lead to inaccurate interpretations of character evolution (Hahn and Nakhleh 2016), including artifactual patterns of diminishing molecular convergence over time (Mendes et al. 2016).

Gene tree discordance has been studied most intensively at relatively shallow phylogenetic scales (Maddison 1997; Pollard et al. 2006; Carstens and Knowles 2007; Avise and Robinson 2008; White et al. 2009). However, simulation results suggest that given particular combinations of effective population size and speciation rates, patterns of ILS may persist over tens or hundreds of millions of years (Oliver 2013). In line with these expectations, some empirical studies have purportedly found a signature of ILS at deep phylogenetic scales on the order of tens of millions of years (Pollard et al. 2006; Ebersberger et al. 2007; Hobolth et al. 2011; Salichos and Rokas 2013;

MANUSCRIPT CATEGORY: Systematic Biology

Suh et al. 2015; Sun et al. 2015). Although considerable attention has been paid to the overall reconstruction of very deep phylogenetic relationships using total evidence approaches (e.g., Rokas and Carroll 2006; Dunn et al. 2008), few studies have intensively examined the potential for gene tree–species tree discordance at deep timescales on a gene-by-gene basis. At these deep phylogenetic scales, distinguishing true discordance from gene tree estimation error can be challenging (Gatesy and Springer 2014).

Loci that are genealogically concordant with the species tree can lose phylogenetic information over time through multiple substitutions at sites, potentially leading to the spurious estimation of a discordant gene tree. Signals from actual discordance may similarly erode over time. Consequently, genes with high information content may be required to detect and quantify gene tree discordance at deep scales (Townsend 2007; Salichos and Rokas 2013). Scrutinizing signal from individual gene alignments and their contributions to the overall estimate of the species tree is also important (e.g., Brown and Thomson 2016). The topology of the gene tree alone may hide signal that is distributed across sites within a gene, signal that can be extracted to evaluate alternative gene-tree topologies and their support for a species-tree hypothesis (Arcila et al. 2017). Individual genes may be subject to systematic error due to a wide range of phenomena (Rannala and Yang 2008), and evidence for deep discordance may also provide insight into demographic histories.

Early amphibian evolution represents a unique empirical challenge for analyses of deep genealogical discordance (San Mauro et al. 2005; Siu-Ting et al. 2019). Here, we examine these issues using phylogenomic data assembled from extant amphibians (Lissamphibia) to estimate phylogenetic history across this clade and focus on the contentious relationships among the three deeply diverged amphibian orders. Amphibians are a diverse and imperiled class of vertebrates, with more than 8100 species described to date (AmphibiaWeb 2020; Frost 2020). Three amphibian orders are recognized: Anura (frogs and toads, approximately 7200 species), Caudata (salamanders and newts, approximately 750 species), and Gymnophiona (caecilians, approximately 200 species). Despite extensive investigation, phylogenetic affinities of some amphibian groups remain unresolved, including disagreement about deep interordinal relationships among the three extant amphibian orders, as well as relationships among more shallowly diverged lineages (Larson and Wilson 1989; Feller and Hedges 1998; Frost et al. 2006; Roelants et al. 2007; Fong et al. 2012; Chen et al. 2015; Jetz and Pyron 2018; Siu-Ting et al. 2019). While much research has centered on the use of morphology and paleontological approaches to resolve deep amphibian relationships (e.g., Carroll 2007; Sigurdsen and Green 2011; Pardo et al. 2017; Matsumoto and Evans 2018; Schoch 2019), here we place our focus on molecular phylogenetic approaches.

Although the monophyly of each amphibian order is not disputed (Frost et al. 2006; Roelants et al. 2007; Anderson et al. 2008; Pyron and Wiens 2011), relationships among them have remained murky. Three possible topologies exist (assuming a monophyletic Amphibia): the Batrachia hypothesis (e.g., Haeckel 1866; Duellman and Trueb 1994; Zardoya and Meyer 2001; Frost et al. 2006; Anderson et al. 2008; Siu-Ting et al. 2019), which places frogs and salamanders as each other's closest relatives, the Procera hypothesis (Feller and Hedges 1998; Siu-Ting et al. 2019), which suggests a sister relationship between salamanders and caecilians, and the Acauda hypothesis (named here), which proposes that frogs and caecilians form a clade.

Although different regions of the genome may support different topologies, a single species tree exists for interordinal amphibian relationships; thus, the question of the relationships between frogs, salamanders, and caecilians becomes a model selection problem. The Batrachia and Procera hypotheses have received the most support in previous studies (Frost et al. 2006; Roelants et al. 2007; Pyron and Wiens 2011; Siu-Ting et al. 2019), although recent genome-scale analyses with modest numbers of taxa (Fong et al. 2012; Chen et al. 2015) have found some loci that appear to support each of the alternative hypotheses. These studies (Fong et al. 2012; Chen et al. 2015) have even revealed some loci that appear to support the nonmonophyly of amphibians with respect to amniotes, further widening the hypothesis space for the deep evolution of Amphibia (Supplementary Fig. S1 available on Dryad at https://doi.org/10.5061/dryad.9kd51c5dc).

The technical hurdles to generating genome-scale data sets for phylogenetic inference in amphibians are still nontrivial, due primarily to their deep evolutionary divergences and inordinately large genomes (e.g., salamander genomes are ~15–120 Gb; Gregory 2020; Weisrock et al. 2018). Targeted sequence-capture approaches have grown in popularity in recent years (e.g., Bi et al. 2012; Faircloth et al. 2012; Lemmon et al. 2012) and have the potential to streamline the sequencing of large numbers of loci in parallel. Yet, the constraints imposed by sequence divergence across broad taxonomic scales limit the applicability of sequence capture approaches when capture probes are designed from a single or a few model taxa. Additionally, in clades with extreme variation in genome size, sequence capture efficiency can be highly variable because target loci are effectively diluted in large genomes.

Recent studies using large molecular data sets in amphibians have demonstrated the power of genomic data for resolving relationships, divergence times, and reconstructing shared patterns of diversification across amphibians (Roelants et al. 2007; Peloso et al. 2016; Feng et al. 2017; Irisarri et al. 2017; Streicher et al. 2018; Jetz and Pyron 2018; Hutter et al. 2019; Siu-Ting et al. 2019). However, to date no study has combined dense sampling from the nuclear genome with comprehensive taxon sampling at the subfamily level across all three orders of

extant amphibians. To address these needs across extant amphibians, we designed a taxon-specific variation of the Anchored Hybrid Enrichment (AHE) protocol (Lemmon et al. 2012). By including probes designed from multiple model species spanning the three orders, we provide a novel sequence capture system for effectively targeting and sequencing hundreds of nuclear protein-coding loci across all extant amphibian species. We employed this capture system to collect AHE data and reconstruct the phylogeny of extant amphibians, assess the degree of deep-time genealogical discordance across their genomes, and evaluate the alternative hypotheses described above.

Materials and Methods

Amphibian-Specific Sequence Capture Probe Design

We designed a probe set that targets amphibian-specific orthologs for 366 of the 512 AHE loci developed by Lemmon et al. (2012). This probe set is designed from a diverse array of representatives of each of the three amphibian orders. Following Barrow et al. (2018), Heinicke et al. (2018), and Yuan et al. (2019), we mined the publicly available genome sequence for the model frog *Xenopus tropicalis* (Hellsten et al. 2010) and published transcriptomes for the salamanders *Ambystoma mexicanum* (Wu et al. 2013) and *Notophthalmus viridescens* (Abdullayev et al. 2013). To increase taxon representation in our probe design, we also developed and mined genomic resources *de novo* for six additional frogs (*Ascaphus montanus*, *Gastrophryne carolinensis*, *Mixophyes schevilli*, *Pseudacris feriarum*, *Pseudacris nigrita*, and *Rana sphenocephala*), one salamander (*Desmognathus fuscus*), and one caecilian (*Ichthyophis multicolor*), as well as transcriptomic resources for two additional salamanders (*Cryptobranchus alleganiensis* and *Ensatina eschscholtzii*). For each of these 13 amphibian taxa, we identified putative orthologs for the 366 amphibian-specific AHE loci from Lemmon et al. (2012) and designed RNA capture probes specific to these loci. Although not all of the 366 target loci were identified in all 13 model taxa, each locus was represented by, on average, 11.1 model taxa. Because it targets a subset of protein-coding exons, this probe set represents $<1\%$ of the *Xenopus* genome. After retaining a set of 220 putatively single-copy orthologous loci (see below in *Nuclear Locus Assembly and Alignment*), we identified human orthologs for these loci (see Supplementary Table S1 available on Dryad).

We designed a set of 120 base pair (bp) RNA probes tiled across each of these loci for each of the 4061 locus-by-model-taxon combinations. Each locus consists of an evolutionarily conserved core exonic region flanked by more variable regions on either side. Our probes for each model taxon covered these core regions and also extended into the flanks in order to increase the lengths of captured loci across diverse taxa. Across all 13 model taxa and 366 target loci, the region covered by our probes was 1,090 bp per locus on average. In practice, longer assemblies are generated from this type of data because the use of paired-end sequencing allows for the extension of sequenced regions beyond the core conserved regions covered by the probes. This set of 57,750 unique 120-mer probes was synthesized by Agilent Technologies and sequences of these probes are included in this study's Dryad package (https://doi.org/10.5061/dryad.9kd51c5dc).

Taxon Sampling for Amphibian Phylogenetics

We assembled tissues and/or genomic DNA, nearly all from museum vouchers, for a set of 286 amphibian species broadly covering family- and subfamily-level diversity and performed targeted sequence capture using our amphibian-specific probe set. Including eight of the 13 model taxa used in the probe kit design, our ingroup taxa consisted of 15 species and nine families of caecilians (out of 214 species and 10 families), 41 species and 10 families of salamanders (out of 740 species and 10 families), and 230 species and 52 families of frogs (out of 7,193 species and 54 families) (detailed in Supplementary Table S2 available on Dryad; species and family counts follow AmphibiaWeb.org, last accessed April 13, 2020). Within each of the three amphibian orders, we attempted to sample representatives from each recognized family, and from multiple subfamilies in the case of particularly diverse families (taxonomy used here follows AmphibiaWeb 2020). We sampled taxa in rough proportion to the species richness of their respective families, but we were also constrained by the availability of tissues and the quality of genomic DNA available for sequencing. To guide taxon choice, we consulted previously published phylogenies and, where possible, attempted to include similar taxonomic coverage of the different amphibian families.

We specifically attempted to sample deeply divergent lineages, as well as taxa that would break up long branches deep in the amphibian phylogeny or potentially provide resolution of recalcitrant nodes. Capturing large numbers of loci from the deeply divergent salamander *Siren intermedia* was problematic due to the combination of an extremely large genome, and deep divergence from any of the model probe taxa. Because of the importance of *Siren* for resolving salamander phylogeny, we sequenced a multi-tissue transcriptome for this species and mined orthologs from this assembly to include with our alignments. For the early-branching caecilian *Rhinatrema bivittatum*, we mined a published transcriptome sequence (Irisarri et al. 2017) for orthologs to our target loci.

Outgroup choice can have important downstream implications for phylogenetic inference (e.g., Wilberg 2015; Grant 2019), so to counter the potential for long branch attraction and to better estimate model parameters and divergence times for deep branches, we included multiple outgroup taxa. We included gene sequences mined from available genomic resources on GenBank from four amniote outgroups (*Anolis carolinensis*, taxon ID: 28377; *Chrysemys picta*, taxon ID:

8478; *Gallus gallus*, taxon ID:9031; and *Homo sapiens*, taxon ID 9606) and the coelacanth (*Latimeria chalumnae*, taxon ID: 7897). To identify putatively orthologous sequences to our target loci for these five outgroup taxa, we performed blastn searches with the NCBI BLAST algorithm (Johnson et al. 2008) against our target loci.

Genomic Library Preparation and High-Throughput Sequencing

Genomic DNA was extracted using DNeasy silica column kits (Qiagen). We performed 2% agarose gel electrophoresis for each sample to confirm the presence of high molecular weight genomic DNA. Library preparation and enrichment were performed at Florida State University's Center for Anchored Phylogenomics (www.anchoredphylogeny.com). Each sample was fragmented with a Covaris sonicator to a mean fragment size of 300–500 bp. DNA samples with signs of degradation were subjected to less fragmentation during library preparation. Individual samples were uniquely barcoded by ligation of 8 bp single index oligonucleotides and then pooled together in batches of 12–16 for multiplexed target capture. Capture reactions were carried out following Lemmon et al. (2012). The enriched, captured products were amplified by low-cycle PCR with high fidelity polymerase. Resulting genomic libraries were bead-cleaned and pooled for paired-end 150 bp sequencing on the Illumina platform (12–24 caecilians and salamanders per Illumina lane, and 24–60 frogs per lane) and sequenced across 14 lanes of an Illumina HiSeq2500 sequencer at the Florida State University Translational Lab.

Nuclear Locus Assembly and Alignment

Loci were assembled using a quasi-*de novo* strategy employing spaced kmer matching and extension assembly from merged paired-end Illumina reads (as in Rokyta et al. 2012; Prum et al. 2015; Hamilton et al. 2016). Orthology was established using a neighbor-joining algorithm based on pairwise distance matrices that accommodate high variation at third codon positions (see Hamilton et al. 2016). Ortholog filtering resulted in a set of 220 putatively single-copy nuclear loci for which all five outgroup taxa were also sampled. Of this set of 220 loci, 195 contained representatives of each of the three amphibian orders (Supplementary Tables S3 and S4 available on Dryad). We used the approach of Pyron et al. (2016) to recover both allele copies for diploid individuals (and multiple allele copies for ploidy levels greater than two). In the case of phase ambiguities, nucleotides were randomly resolved to one of their potential states. Although we generated phased data for every individual, we retained only one randomly chosen haplotype for each individual at each locus in order to greatly reduce the computational burden of downstream analyses. Multiple sequence alignment was performed in a nested procedure. We first performed four separate alignments for frogs, salamanders, caecilians, and the amniote and *Latimeria* outgroups in MAFFT v7.221 (Katoh and Standley 2013) with the L-INS-i parameter settings. These subalignments were then combined using the MAFFT—merge function.

Preliminary examination of gene trees for these alignments revealed that some taxa had clearly erroneous placements in a small number of gene trees (e.g., a salamander placed within frogs, or a caecilian nested within amniotes). These taxa were typically characterized by very long branch lengths. Further scrutiny revealed that in nearly every case, large numbers of ambiguous or missing sites apparently drove this pattern. This effect of missing data was, however, not unexpected (Lemmon et al. 2009). To clean up these alignments, we implemented a taxon-filtering procedure for each locus that culled any ingroup taxon with greater than 50% missing and/or ambiguous sites across an alignment, or with a terminal gene tree branch length greater than five times the average branch length for that tree. This filtering procedure removed less than 1% of the taxon-by-locus combinations and greatly improved the consistency of estimated gene trees.

Culled alignments were examined by eye to correct obvious misalignment issues (e.g., large gaps preceded by a single leading nucleotide), and to establish reading frames across protein-coding portions of each locus. Between zero and two base pairs were trimmed from the upstream ends of each alignment so that the first nucleotide represented the first codon position and the last nucleotide represented the third codon position of each locus. Establishment of reading frame and alignment corrections were performed in Geneious R8 (Kearse et al. 2012) with hydrophobicity display enabled for translated amino acid sequences. In many cases, manual alignment correction substantially increased polarity conservation, especially around gaps where the alignment may be particularly uncertain.

MtDNA Assembly and Sample Vetting

To verify the identities of samples, we assembled complete and partial mitochondrial genomes for nearly all taxa from off-target bycatch reads. These mitochondrial data served as integrated "barcodes" with which we could verify the integrity of our taxon identifiers and flag potential cases of misidentified taxa. Raw read data were reassembled *de novo* with trinityrnaseq v2.0.3 (Grabherr et al. 2011) and the resulting assemblies were mined for mtDNA regions using blastn searches against known mitogenomes. This procedure identified six instances of apparent pairwise transposition of sample labels, corroborated by aberrant placements in preliminary gene trees (e.g., placement of two taxa appeared transposed). This mislabeling likely occurred after DNA extraction but prior to library preparation. When mtDNA confirmed sample switching, taxon identifiers for these pairs of taxa were amended accordingly.

Gene Tree Estimation

Based on strict orthology assessment, we retained 220 loci for gene tree estimation. Individual locus

MANUSCRIPT CATEGORY: Systematic Biology

alignments had on average 5.88% missing sites (range 0.36–19.75% missing sites). Because these AHE loci are exclusively protein-coding, we used a codon-based partitioning strategy for identifying and modeling substitution variation. For each locus, we used the corrected Akaike Information Criterion (AIC_c) calculated in PartitionFinder2 (Lanfear et al. 2016) to simultaneously select among possible partitioning schemes and models of molecular evolution using the greedy search algorithm (Lanfear et al. 2012). We chose partitioning schemes for each locus using AIC_c for use during gene tree estimation (detailed in Supplementary Table S3 available on Dryad). Maximum likelihood (ML) estimates of each individual gene tree were obtained in RAxML v8.2.11 (Stamatakis 2014) under several different topological constraints. For each RAxML analysis, 500 rapid bootstraps were conducted followed by a series of 20 slow ML optimization steps before the full ML analysis with branch length optimization. For each locus, we performed 17 separate ML tree searches. The first ML search was for a completely topologically unconstrained analysis. The second ML search was performed using a topological constraint that enforced amphibian monophyly but did not enforce any further topological constraints. These monophyletic-Amphibia constraint trees were used in downstream analyses of the support across loci for either the Batrachia, Procera, or Acauda hypotheses. Additionally, we inferred ML trees for each of the 15 backbone constraints representing the 15 possible tree topologies relating frogs, salamanders, caecilians, amniotes, and the outgroup *Latimeria* (Supplementary Fig. S1 available on Dryad). These 15 constrained trees were used in downstream analyses of the support for all competing interordinal models (see below). For 15 backbone constraints, intraordinal amphibian monophyly and the monophyly of amniotes were both enforced, but no constraints were imposed on relationships within the amphibian orders or within amniotes.

Species Tree Estimation

We employed several methods to reconstruct the topology of the amphibian species tree. As a first pass, we concatenated all 220 loci together into a single alignment with 291,282 nucleotide characters for 291 taxa (286 amphibians, four amniotes, and the *Latimeria* outgroup). Overall, this concatenated alignment was 86.4% complete and had 199,328 variable sites and 176,483 parsimony-informative sites. We also explored optimal partitioning schemes for the concatenated alignment in two different ways. One approach combined first, second, and third codon positions across all loci into their own respective subsets and identified an optimal partitioning scheme based on a greedy search in PartitionFinder2 where aggregate first, second, and third codon positions each comprised their own partitions. Next, we used the rcluster algorithm (Lanfear et al. 2014) in PartitionFinder2 to heuristically search model space for 660 maximum potential partitions from 220 loci each partitioned by codon position resulting in 76 separate partitions. The best partitioning schemes for both approaches were used for concatenated ML phylogeny estimation in RAxML. RAxML analyses employed 500 rapid bootstrap replicates, followed by 20 slow ML optimization steps before the full ML analysis with branch length optimization. We enforced a single topological constraint on these concatenated ML analyses by setting *Latimeria* as the outgroup to all other taxa.

To account for different coalescent histories among loci, we performed species tree estimation in ASTRAL2 v5.6.1 (Mirarab and Warnow 2015) using the best ML tree for each of the 220 loci from the unconstrained gene tree analysis. Using ASTRAL, we produced trees with local posterior estimates of branch support (Sayyari and Mirarab 2016), as well as estimates of the effective number of genes supporting each bipartition and the relative support for each of the alternative quartet resolutions. Although ASTRAL is technically not a coalescent method, it is statistically consistent with the coalescent model for increasing numbers of loci and sites (Mirarab and Warnow 2015).

Divergence Time Estimation

We estimated divergence times in the MCMCTree program from PAML v4.9f (Yang 2007) using a set of 19 fossil calibrations, which were also applied by Feng et al. (2017) (see Supplementary Table S5 available on Dryad and Supplementary Fig. S2 available on Dryad). These fossil calibration points cover many of the deep branches within tetrapods and within amphibians. We began by estimating the substitution rate for each of the 220 loci partitioned by codon position in the PAML program baseml under a GTR+Γ model of nucleotide substitution with five discretized rate categories. For each locus, the gene tree topology was fixed to that from the unconstrained RAxML gene tree analyses and the root age for the divergence between *Latimeria* and tetrapods was set to 450 million years ago (Ma) (Benton et al. 2015). Based on all loci and codon positions, we estimated a mean substitution rate of 0.899 substitutions per billion years and used this estimate to parameterize a diffuse gamma Dirichlet prior on locus rates (rgene_gamma) as Γ(1 1.11).

We used the concatenated alignment of 220 loci, partitioned into aggregate first, second, and third codon positions as input for MCMCTree and fixed the tree topology to the ASTRAL species tree estimate. We then applied 19 fossil calibration points to constrain nodes such that 95% of the prior probability mass fell between the lower and upper (soft) bounds of each fossil calibration, with 2.5% prior density extending above and below these bounds. Next, we obtained maximum likelihood estimates (MLEs) of branch lengths by approximate likelihood and calculated the gradient and Hessian matrices at MLEs of branch lengths

with the usedata=3 option. Using inferred ML branch lengths, we estimated divergence times in MCMCTree with the usedata=2 option, uncorrelated rates across loci, and a GTR+Γ model of nucleotide substitution. MCMCTree analyses used two independent MCMC chains with a 1,000,000 generation burnin, subsequently sampling every 100 generations until 10,000 samples were collected. The outputs of these analyses were combined and summarized together using the print=-1 setting in MCMCTree to calculate the posterior mean divergence times and 95% highest posterior density (HPD) credible intervals on divergence times. Effective sample size (ESS) values of parameters were calculated for the combined results in Tracer v1.7.1 (Rambaut et al. 2018). We also performed fossil cross-validation analyses (Near and Sanderson 2004) of our divergence time estimates to quantify the sensitivity of our results to individual fossil calibration points by systematically omitting one of each of the 19 fossil calibrations in each run. The time-calibrated phylogeny was visualized with the MCMCtreeR package (Puttick 2019).

Support across Loci for Interordinal Amphibian Relationships

One hundred ninety-five loci included representatives of all three amphibian orders and were thus suitable for investigating support for competing interordinal models (Supplementary Fig. S1 available on Dryad and Supplementary Table S4 available on Dryad). We initially quantified support across these 195 loci for the three competing interordinal models that support a monophyletic Amphibia (Batrachia, Procera, and Acauda). We counted how many of these constrained ML gene trees recovered topologies consistent with Batrachia, Procera, or Acauda. For each of these 195 genes, we also calculated the proportion of gene tree bootstrap replicates that supported Batrachia, Procera, or Acauda and plotted these values across loci. Additionally, we used a model-selection framework to calculate the relative support for each of these three competing interordinal topologies. For each locus, we compared the log likelihoods of constrained gene trees and calculated the Akaike Information Criterion (AIC; Akaike 1974) for each topological model relative to the best model [the model with the lowest $-\ln(L)$]. AIC was calculated as $2k - 2\ln(L)$, where k is the number of free parameters in the model of sequence evolution and the tree (e.g., branch lengths) and L is the maximum likelihood under the constraints for a given hypothesis. There is an additional unknown number of free parameters associated with topological freedom in our hypotheses (i.e., we constrain monophyly of orders, but not relationships within orders) that we do not include in k, because the appropriate number is unclear. However, this value should be the same across hypotheses, because each has the same amount of topological freedom, therefore the associated penalty term would cancel when calculating differences in AIC scores. These differences between the best model and the other i competing models for each locus were calculated as $|\mathrm{AIC}_{\min} - \mathrm{AIC}_i|$. Because k is identical across hypotheses, calculating these differences in AIC is equivalent to examining the log of the maximum likelihood ratio between the best model and each alternative. We also repeated these analyses for the full 15-model set depicted in Supplementary Figure S1 available on Dryad.

To further quantify support for competing interordinal models for amphibian relationships, we conducted approximately unbiased (AU) tests of topology (Shimodaira 2002) across the 15 backbone constraints. To examine support across loci for the interordinal amphibian relationships, we applied the method of gene genealogy interrogation (GGI; Arcila et al. 2017). For each locus, we inferred an ML topology and then plotted the cumulative number of loci supporting each alternative topological hypothesis, rank ordered by decreasing statistical significance of the AU P-value. We performed AU tests in PAUP* (v4a.151) (Swofford 2003).

RESULTS

Nuclear Locus Assembly and Alignment

After filtering for potential paralogs and missing data in alignments, we retained a set of 220 loci, of which 195 had representatives of all three amphibian orders present. This resulted in total of 291,282 aligned nucleotide positions across all loci. On average, individual frogs, caecilians, and salamanders had 214, 165, and 146 loci, and 97%, 75%, and 66% of sites in the alignments present, respectively.

Gene Tree Estimation

Codon-based partitioning analyses and substitution model selection for individual loci identified 217 loci for which a single partition containing all codon positions and a GTR+Γ model of sequence evolution was applied (Supplementary Table S3 available on Dryad). The remaining three loci (L123, L262, and L285) were best modeled with two partitions, one combining first and second codon positions and the second partition comprising third codon positions and a GTR+Γ model. RAxML bootstrap support (BS) for individual gene trees was relatively high, with an average BS value of 73.8 across all branches in all gene trees (median value of 90). Individual unconstrained gene trees varied according to which of the 15 possible interordinal models of amphibian and amniote relationships were present in the ML tree. Among unconstrained gene trees, 143 loci identified a monophyletic Amphibia, 67 loci recovered Batrachia, 38 loci supported Procera, and 38 loci supported Acauda. All gene trees are included in this study's Dryad accession.

MANUSCRIPT CATEGORY: Systematic Biology

Topology and Timescale for the Amphibian Tree of Life

Concatenated RAxML and ASTRAL analyses produced similar estimates of the species tree (normalized Robinson–Foulds distance = 0.027), and we present the ASTRAL species tree here (Fig. 1; concatenated RAxML tree in Supplementary Fig. S3 available on Dryad). In both estimates of the species trees, we recovered a monophyletic Amphibia, monophyly of each amphibian order, and a frogs-sister-to-salamanders relationship (the Batrachia hypothesis), all with BS = 100. MCMCTree results indicated that the Amniota-Lissamphibia split occurred 347–352 Ma, during the early Carboniferous, that the ancestor of caecilians and an ancient batrachian diverged 287–303 Ma, in the late Carboniferous or early Permian, and that frogs split from salamanders by 269–275 Ma, during the middle Permian (Fig. 2).

The estimated phylogenetic relationships within each of the three amphibian orders were largely consistent with those from previous studies, with most interfamilial branches receiving maximal bootstrap support in both the RAxML and ASTRAL analyses (Supplementary Figs. S3 and S4 available on Dryad, respectively). Within caecilians, we recovered the deepest divergence event as being between the Rhinatrematidae and all other caecilians examined (Figs. 1 and 2, Supplementary Figs. S3 and S4 available on Dryad), with this divergence occurring 102–130 Ma in the early Cretaceous. All remaining family-level relationships within caecilians were resolved with high branch support (Supplementary Figs. S3 and S4 available on Dryad).

Our results unequivocally supported an initial divergence event in crown salamanders between Cryptobranchioidea (Cryptobranchidae and Hynobiidae) and a clade comprising Salamandroidea and Sirenidae (Fig. 1, Supplementary Figs. S3 and S4 available on Dryad), with a divergence occurring 148–171 Ma in the early to middle Jurassic (Fig. 2). After divergence with Sirenidae, the first split within Salamandroidea involved a clade comprising Ambystomatidae and Dicamptodontidae and Salamandridae, and a clade comprising successive divergence events splitting off Proteidae, Rhyacotritonidae, Amphiumidae, and Plethodontidae.

Within frogs, we recovered a nonmonophyletic Archaeobatrachia (Figs. 1 and 3, Supplementary Figs. S3 and S4 available on Dryad). At the root of the frog clade, the Leiopelmatoidea (Leiopelmatidae and Ascaphidae, Green and Cannatella 1993) diverged from all other frogs 205–223 Ma in the late Triassic (Fig. 3). Additional archaeobatrachian superfamily-level groups were also resolved, including Bombinatoroidea (Bombinatoridae and Alytidae) and the Pipoidea (Rhinophrynidae and Pipidae), and Pelobatoidea, the clade containing (Scaphiopodidae, Pelodytidae, Pelobatidae, and Megophryidae). Our results suggest that *Leiopelma* and *Ascaphus* diverged 187–209 Ma in the early Jurassic, and that Neobatrachia split from Pelobatoidea (Scaphiopodidae, Pelobatidae, and Pelodytidae) 169–185 Ma in the middle Jurassic.

Neobatrachia was recovered as monophyletic in all analyses (Fig. 1), and all methods supported *Heleophryne* as sister to all other neobatrachians (Figs. 1 and 3, Supplementary Fig. S3 and S4 available on Dryad). The remaining neobatrachian lineages formed two clades: (Hyloidea, (Myobatrachidae, *Calyptocephalella*)) and (Ranoidea, (*Sooglossus*, *Nasikabatrachus*)). Within Ranoidea, we recovered a clade comprising (Microhylidae, (Natatanura, Afrobatrachia)). Within Neobatrachia, we found that Heleophrynidae split from all other neobatrachians 140–152 Ma in the late Jurassic or early Cretaceous (Fig. 3). Our results also suggest that the major neobatrachian crown lineages Hyloidea, Ranoidea, Afrobatrachia, and Natatanura arose 67–77 Ma, 99–111 Ma, 81–92 Ma, and 61–69 Ma, respectively (Figs. 3 and 4). We found support for a sister relationship between *Nasikabatrachus* and *Sooglossus* and estimated the divergence time between these two deeply divergent anuran families at 62–93 Ma (Fig. 4). This particularly large range of dates for this node may be associated with an elevated percentage of missing loci for *Nasikabatrachus*. Within Ranoidea, we recovered Microhylidae sister to a clade containing Afrobatrachia and Natatanura (Fig. 4). Posterior mean divergence times, upper and lower bounds of the 95% HPD confidence intervals, and ESS values for all nodes are provided in Supplementary Table S6 available on Dryad. Our divergence time estimation results were largely robust to the sensitivity analyses wherein we reran the MCMCTree analyses excluding a single fossil calibration in each run. Results of the exclusion analyses are provided in Supplementary Table S7 available on Dryad.

Placement of Nasikabatrachidae and Sooglossidae

Our final ASTRAL and concatenated RAxML analyses both supported a sister relationship between the two deeply divergent anuran families Nasikabatrachidae and Sooglossidae (Supplementary Fig. S5 available on Dryad). Initial ASTRAL and RAxML analyses based on gene alignments that were not filtered for missing sites (see Materials and Methods section) disagreed with respect to the placement of *Nasikabatrachus*, with ASTRAL analyses placing it as sister to all Microhylidae and RAxML supporting a placement sister to *Sooglossus*. However, upon further scrutiny of individual gene alignments, we noted a pattern of short, gap-laden loci supporting *Nasikabatrachus* as sister to the microhylid genus *Oreophryne*, whereas loci with fewer missing sites supported the conventionally accepted placement of *Nasikabatrachus* sister to *Sooglossus* (Supplementary Fig. S5 available on Dryad). Final analyses that filtered out individual sequences with > 50% of sites missing for any locus resulted in ASTRAL species tree placement of *Nasikabatrachus* and *Sooglossus* as sister clades.

MANUSCRIPT CATEGORY: Systematic Biology

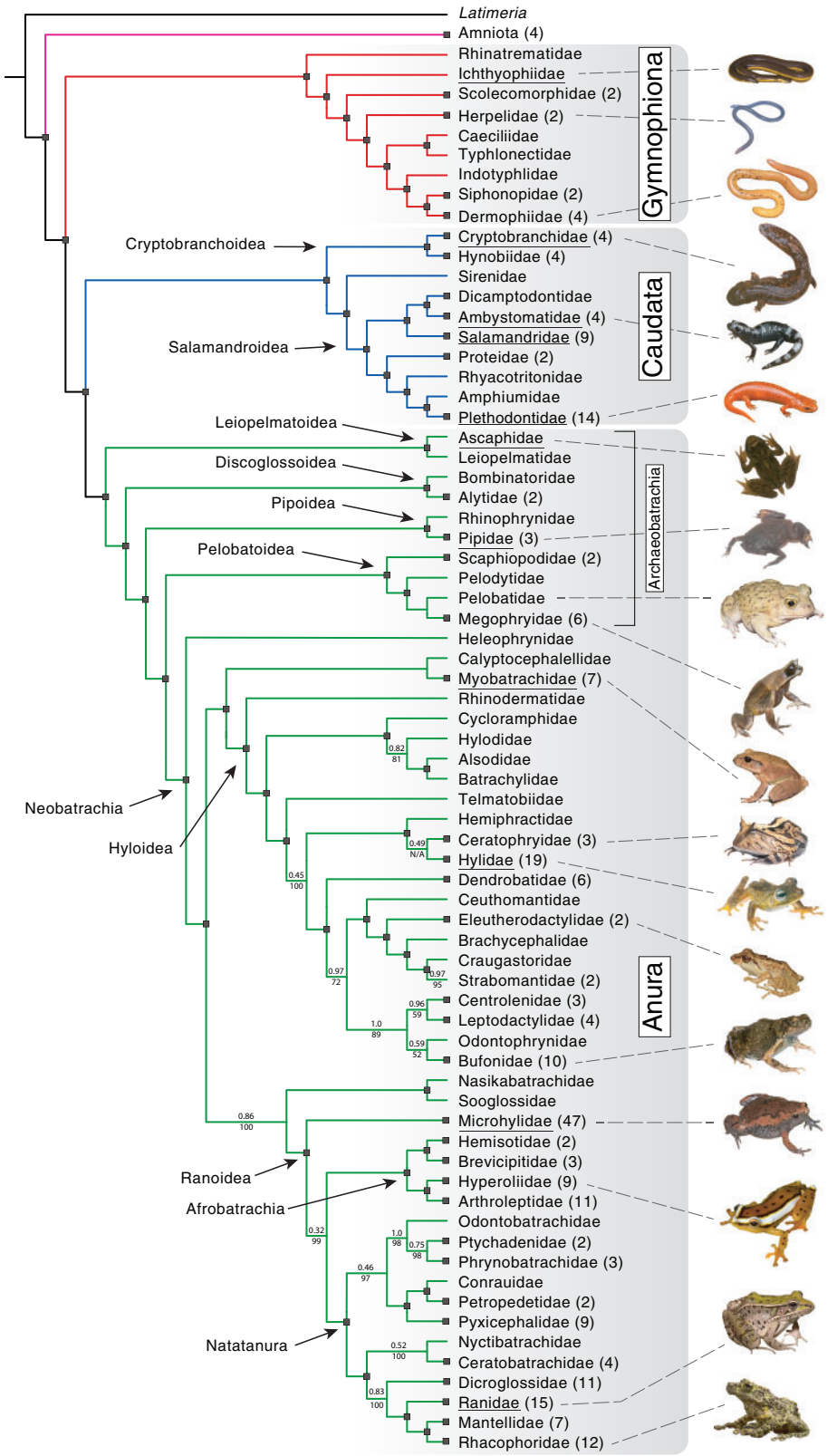

FIGURE 1. Topology of family-level amphibian relationships inferred from ASTRAL and RAxML analyses. Gray boxes denote branches with ASTRAL local posterior values of 1 and RAxML bootstrap percentages of 100. For branches without unanimous support, upper values indicate ASTRAL local posterior probabilities and lower values indicate RAxML bootstrap percentages. N/A indicates that a branch was recovered by ASTRAL but not by RAxML. Terminal branches without support values are represented by a single taxon from a given family. Values in parentheses after a family name indicate the number of species sampled if greater than one. Families from which species were selected for probe design are underlined. Dashed gray lines connect exemplar photos to their respective families. [Color figure can be viewed in the online pdf version]

MANUSCRIPT CATEGORY: Systematic Biology

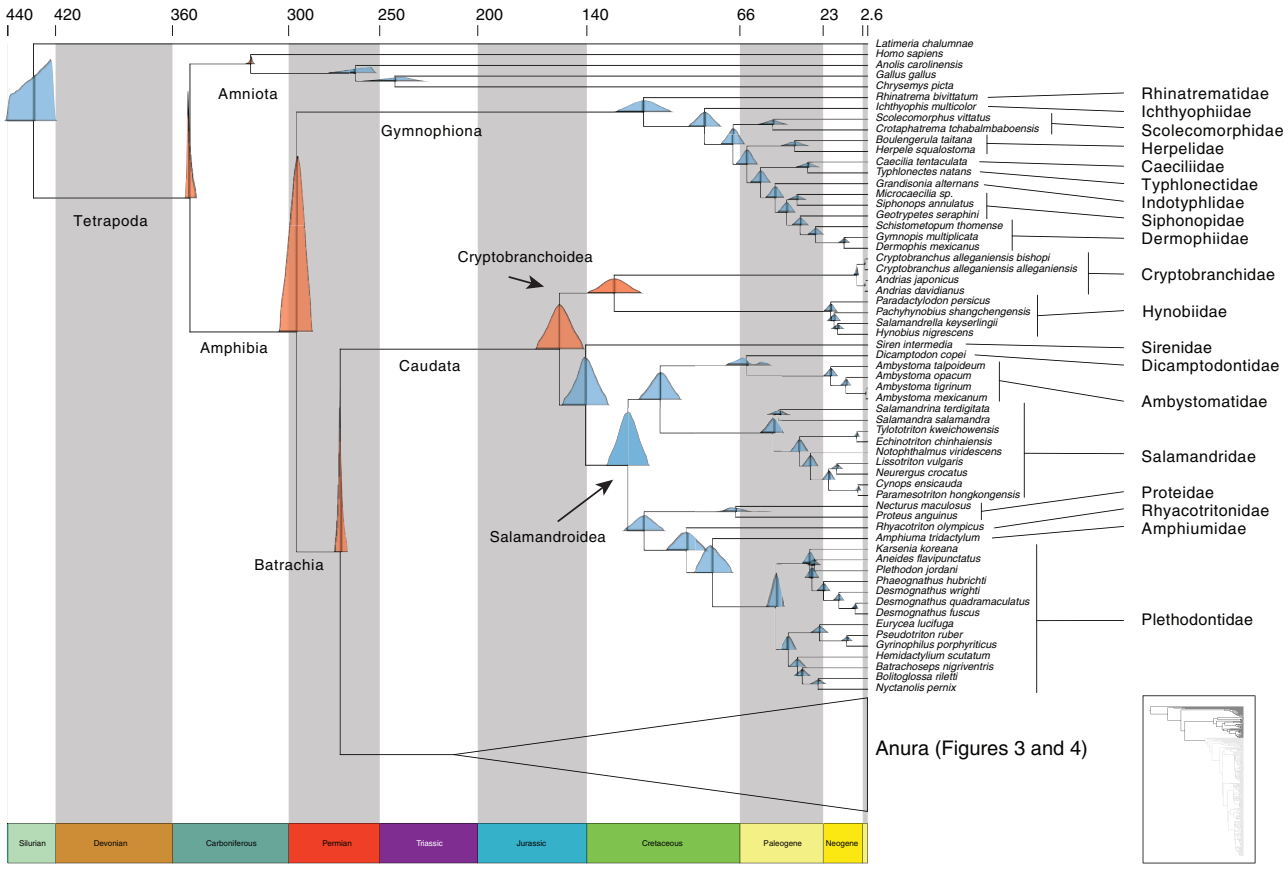

FIGURE 2. Time-calibrated phylogeny for Amphibia. Divergence times among amniotes, caecilians, and salamanders. Divergence times across Amphibia were inferred in MCMCTree with 19 fossil calibration points. Posterior distributions are plotted for each node, representing the 95% highest posterior density confidence interval of divergence times. Fossil calibrated nodes are denoted with orange posterior distributions and noncalibrated nodes are shown in blue. The tree topology is based on the results of the ASTRAL analysis. The heights of the posterior distributions are scaled to their corresponding branches. [Color figure can be viewed in the online pdf version]

Gene Tree Discordance and Genomic Support for Deep Amphibian Relationships

When considering only the Batrachia, Procera, and Acauda interordinal models, each of the three alternative interordinal hypotheses for a monophyletic Amphibia was supported by different subsets of the genome. Quantifying bootstrap support for interordinal models involving a monophyletic Amphibia revealed substantial variation across loci (Fig. 5a). A greater number of gene trees exhibited high bootstrap support for the Batrachia hypothesis relative to the Acauda or Procera hypotheses. However, each of these three interordinal hypotheses had at least three gene trees with BS $>$ 90 at the interordinal nodes. Of 195 loci with at least one member of each amphibian order included, 89 loci supported Batrachia, 51 supported Procera, and 55 supported Acauda, respectively. Similarly, a rank-ordered plot of BS by interordinal hypothesis revealed multiple loci with low to moderate bootstrap support for each hypothesis (Fig. 5a). In fact, only 28 loci had BS $>$ 90 for any of the deep interordinal nodes, suggesting that many loci might lack sufficient information content for resolving these deepest nodes in the amphibian phylogeny (Siu-Ting et al. 2019).

Measures of ΔAIC-based model support across the full set of 15 models that also account for the possibility of a nonmonophyletic Amphibia found very strong support for models involving a monophyletic Amphibia, with ΔAIC values as high as 1,400 against some nonmonophyletic models. Of the 195 interordinal-informative loci, 144 loci supported monophyly of Amphibia (68 loci supported Batrachia, 38 loci supported Procera, and 38 loci supported Acauda). The remaining 51 loci supported one of the 12 alternative models of a nonmonophyletic Amphibia (Supplementary Fig. S6 available on Dryad). The raw log likelihoods of the 15 constrained models across 195 loci used as input for AIC-based measures of support for the 3- and 15-hypothesis analyses are included in Supplementary Table S8 available on Dryad.

Results from GGI analyses (Fig. 6) demonstrated that the majority of loci support a monophyletic Amphibia, although support for either the Acauda, Procera, or Batrachia hypotheses varied across loci. Approximately unbiased test *P*-values represent how strongly competing topological models can be rejected

MANUSCRIPT CATEGORY: Systematic Biology

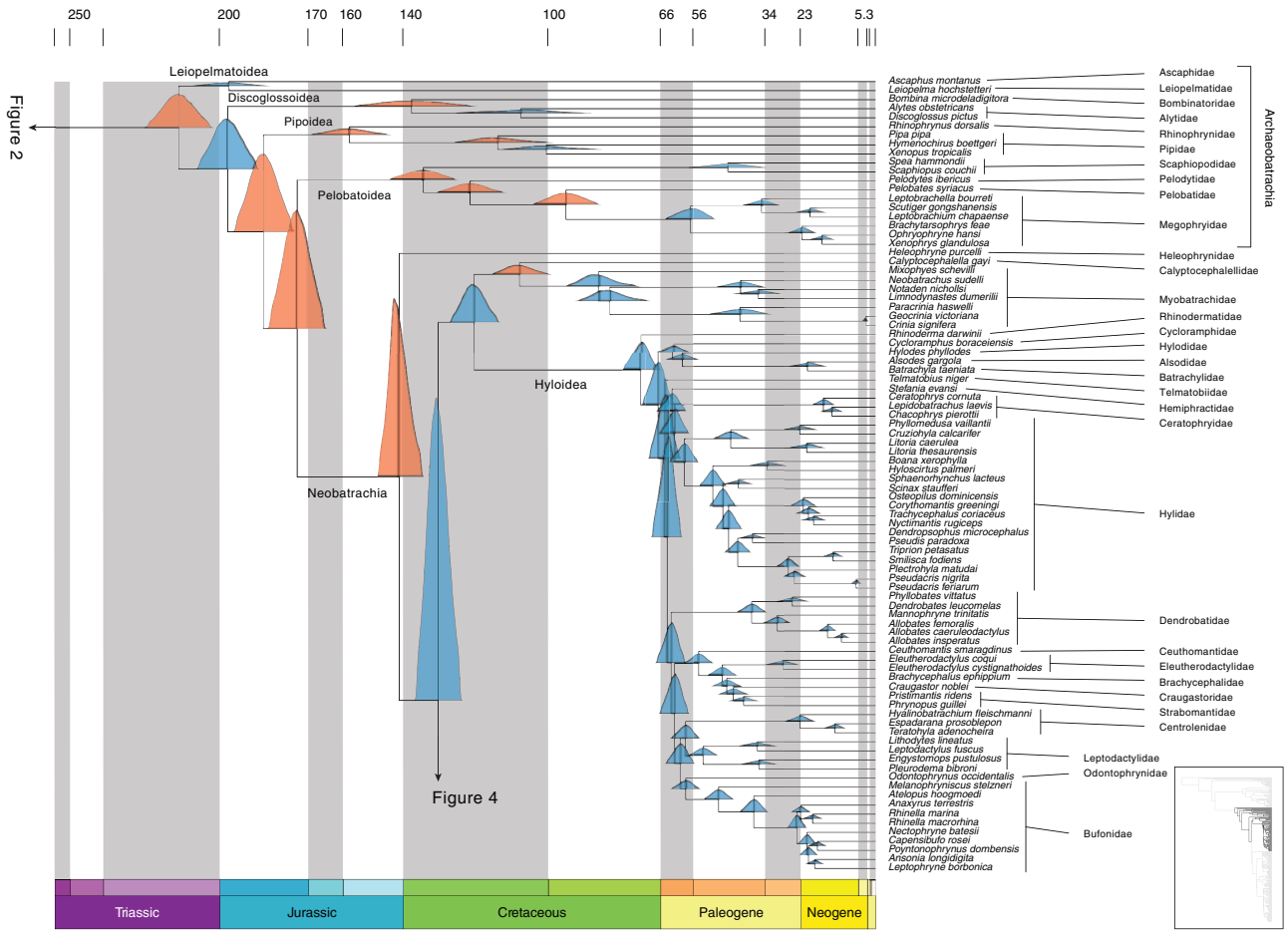

FIGURE 3. Time-calibrated phylogeny for Anurans I. Divergence times for a portion of Anura, containing the Archeobatrachia and some lineages and clades of Neobatrachia. Posterior distributions are plotted for each node, representing the 95% highest posterior density confidence interval of divergence times. Fossil calibrated nodes are denoted with orange posterior distributions and noncalibrated nodes are shown in blue. The tree topology is based on the results of the ASTRAL analysis. The heights of the posterior distributions are scaled to their corresponding branches. [Color figure can be viewed in the online pdf version]

relative to the preferred model, with a canonical value of 0.05 reflecting strong support. Very few loci had strong AU support ([1 - AU *P*-value] > 0.95), but most gene trees supported the Batrachia hypothesis (Fig. 6). Across loci supporting each of the 15 possible interordinal topologies rank-ordered by decreasing AU test values, support dropped precipitously for Procera and Acauda and nononophyletic Amphibia models. Notably, only a small number of loci supporting any of the 15 possible models received support above our threshold of $P \leq 0.05$. Figure 6a shows GGI support for the three monophyletic Amphibia models and the 12 aggregated nonmonophyletic Amphibia models, whereas Figure 6b depicts GGI support across the 12 alternative models individually.

DISCUSSION

Using data generated with a newly developed, amphibian-specific Anchored Hybrid Enrichment probe set, we produced the most comprehensive phylogenomic inference of the amphibian Tree of Life to date, both in terms of taxa (286 species, ~95% of amphibian families) and genetic loci (220 nuclear loci). Although there were few surprising topological relationships uncovered by this work (see below), we provide the most robust assessment to date of amphibian relationships and their divergence times, placing all major amphibian lineages together in a comprehensive phylogenomic framework. We recognize that future study is warranted in several clades (e.g., Plethodontidae, Microhylidae) to better resolve more shallow-scale relationships.

Overall, our species tree results firmly corroborate the monophyly of Amphibia and strongly supported the Batrachia hypothesis, grouping frogs and salamanders as a clade. Perhaps the most surprising aspect of our results was the identification of significant variation across the genome in the strength of support for competing hypotheses about the deepest divergences in the evolutionary history of amphibians. Some loci even exhibited strong support for a nonmonophyletic Amphibia. In the context of a monophyletic Amphibia,

MANUSCRIPT CATEGORY: Systematic Biology

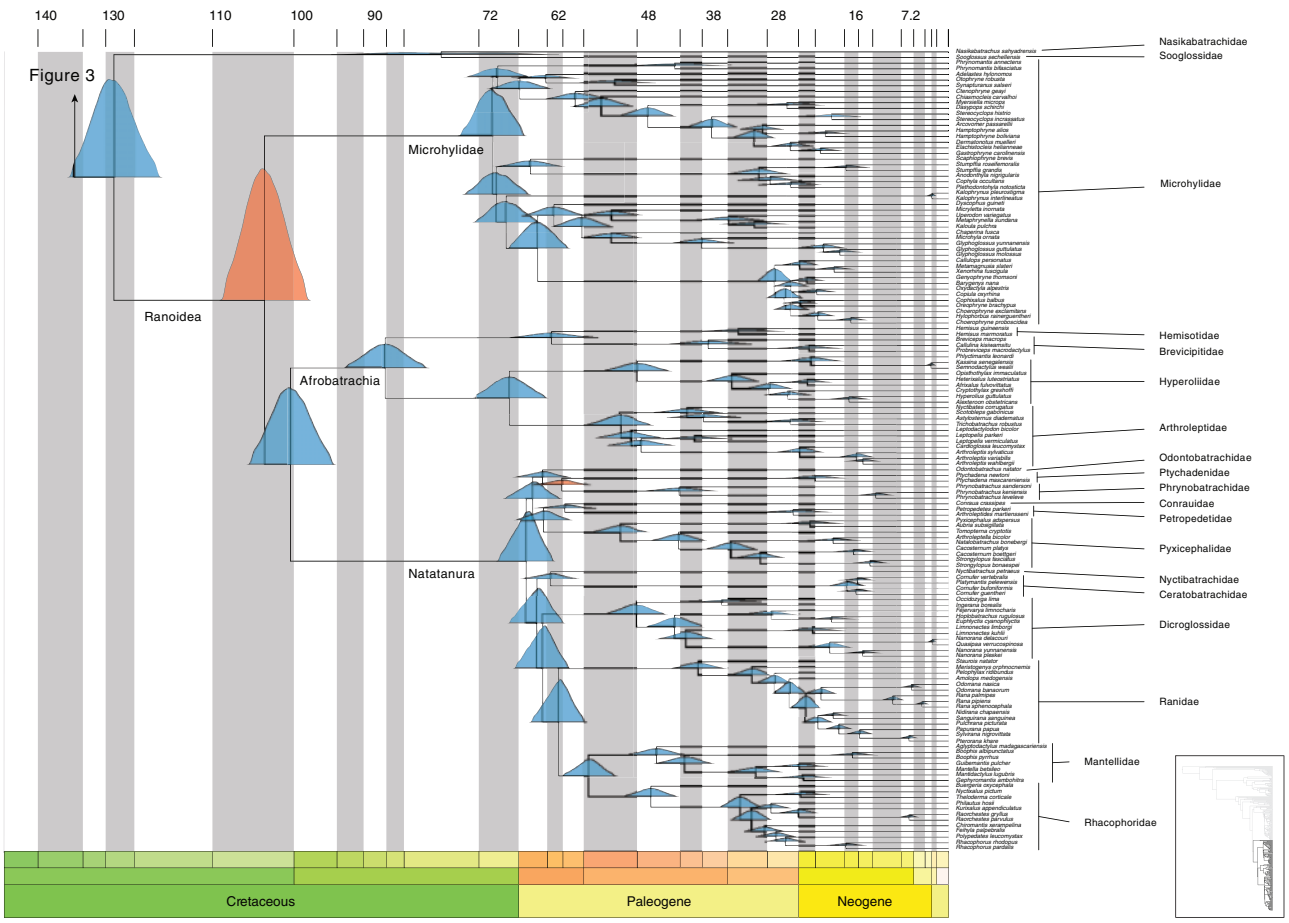

FIGURE 4. Time-calibrated phylogeny for Anurans II. Divergence times for the remaining portion of Anura, containing the remaining clades from Neobatrachia. Posterior distributions are plotted for each node, representing the 95% highest posterior density confidence interval of divergence times. Fossil calibrated nodes are denoted with orange posterior distributions and noncalibrated nodes are shown in blue. The tree topology is based on the results of the ASTRAL analysis. The heights of the posterior distributions are scaled to their corresponding branches. [Color figure can be viewed in the online pdf version]

each of the three possible interordinal hypotheses received strong support from at least some loci (see below). Yet, at the species-tree level, our genome-wide data strongly supported the monophyly of Amphibia and a sister relationship between frogs and salamanders (the Batrachia hypothesis).

Collectively, our results imply either the persistence of patterns of evolutionary mechanisms that contribute to genealogical discord (e.g., ILS or gene flow) over extremely deep timescales, and/or systematic error in phylogenetic estimates from some loci, leading to strongly supported, but inaccurate, estimates of gene trees. The observed heterogeneity in gene tree topologies in deep evolutionary history is obscured by species tree analyses, potentially leading to oversimplification of genomic evolutionary history. Our findings underscore the importance of conducting an in-depth examination of genome-wide phylogenetic signal as opposed to utilizing summary approaches such as gene concatenation or species tree analyses alone.

Our results further support recent analyses presented by Burbrink et al. (2020) for squamates. In that study, genomic interrogation using machine learning showed strong overall congruence for a single topological resolution of lizard and snake relationships that is obscured for some early nodes by apparent ILS, as well as poor or degraded phylogenetic signal in some loci. These congruent results across taxa and timescales for similarly constructed genomic data sets suggests that these processes of signal decay and genealogical discordance driven by rapid radiations and short internodes may be a ubiquitous feature of early branches in the Tree of Life (e.g., Rokas and Carroll 2006; Dunn et al. 2008). However, approaches such as ours and those of Burbrink et al. (2020) nonetheless show the tractability of resolving strong support for these relationships given sufficient data.

Evaluating Interordinal Amphibian Models and Deep ILS

Deep-time gene tree–species tree discordance was evident from multiple analyses. Bootstrap results provided some indication for variation across loci in

MANUSCRIPT CATEGORY: Systematic Biology

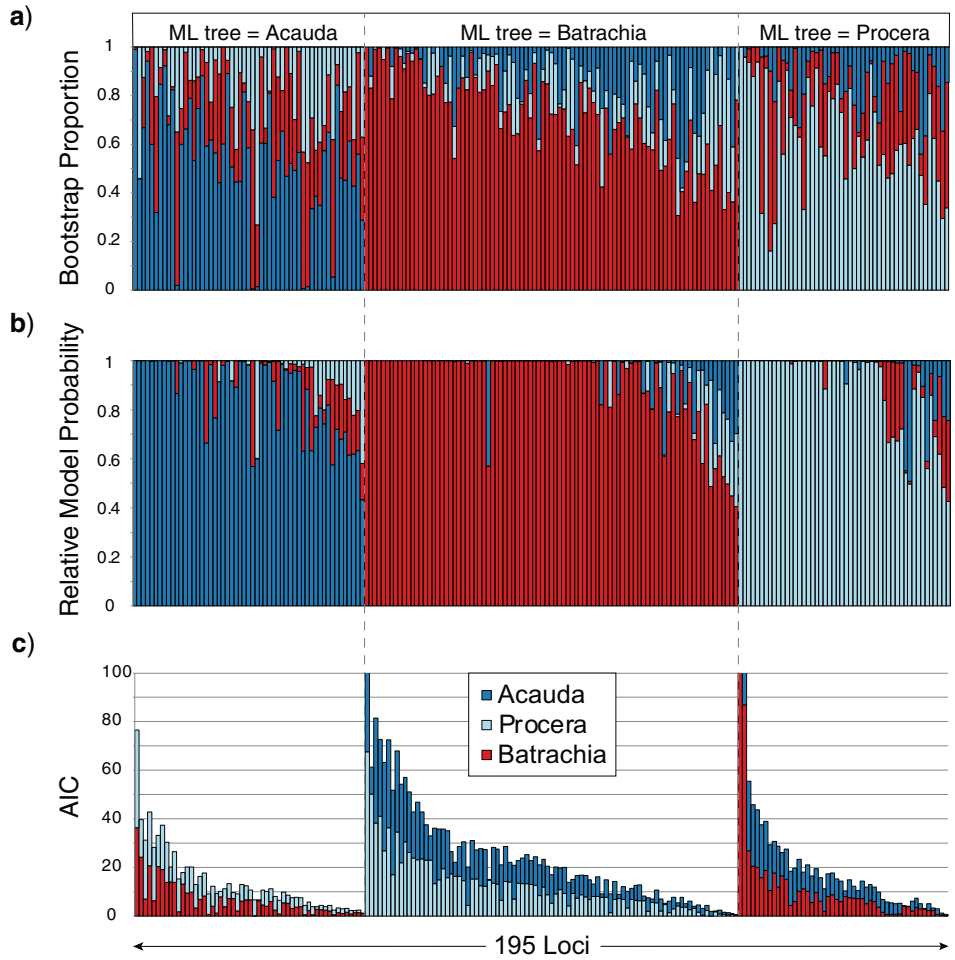

FIGURE 5. Evaluation of the magnitude and direction of support for competing models of a monophyletic Amphibia. a) Proportions of RAxML gene tree bootstrap replicates supporting the three possible interordinal topologies, assuming a monophyletic Amphibia. b) Relative model probabilities among three possible interordinal models. c) Differences in Akaike Information Criterion (ΔAIC) between best model and alternative models quantifying the magnitude of support against the two alternative models for interordinal amphibian relationships (bars are stacked for clarity and values are not cumulative). In all three panels, 195 loci with representatives of each amphibian order are plotted along the horizontal axis. Alternative models for Acauda, Batrachia, and Procera are colored in dark blue, red, and light blue, respectively. [Color figure can be viewed in the online pdf version]

support of different interordinal models involving a monophyletic Amphibia. Many loci were equivocal in their bootstrap support for the reconstruction of any of the three interordinal models involving a monophyletic Amphibia. Yet, each of the three interordinal hypotheses was reconstructed from at least a small number of loci with high bootstrap support. For example, three loci supported an Acauda gene tree with BS $>$ 90 and eight loci supported a Procera gene tree with BS $>$ 90, both in contrast to the robust species tree reconstruction of a Batrachia relationship (with 17 loci with concordant gene trees at BS $>$ 90). The bootstrap represents a bounded measure of branch support, with the ability to tell us something about the direction, but not the magnitude, of support for a particular hypothesis.

The use of GGI/AU tests and information-theoretic relative model probabilities provided two explicit statistical approaches for comparing competing interordinal models, although these provided somewhat contrasting results. GGI/AU tests were quite conservative, with small numbers of loci statistically supporting all of the three monophyletic amphibian models. For example, a total of just four loci supported a Batrachia gene tree topology with an AU test value $\geq$0.95, and no loci produced gene trees with significant AU Test support for Acauda or Procera topologies. While the GGI analyses indicated that Batrachia is favored for gene tree relationships among orders (roughly half of all loci have "best" gene trees consistent with Batrachia), most loci lacked definitive signal this far back in the evolutionary history of amphibians.

Our GGI results for the deep (250–300 Ma) divergences in Amphibia mirror those found in other anciently diverging clades (such as early animal relationships), in that very few individual gene trees are statistically significant for any one model (Arcila et al. 2017). In

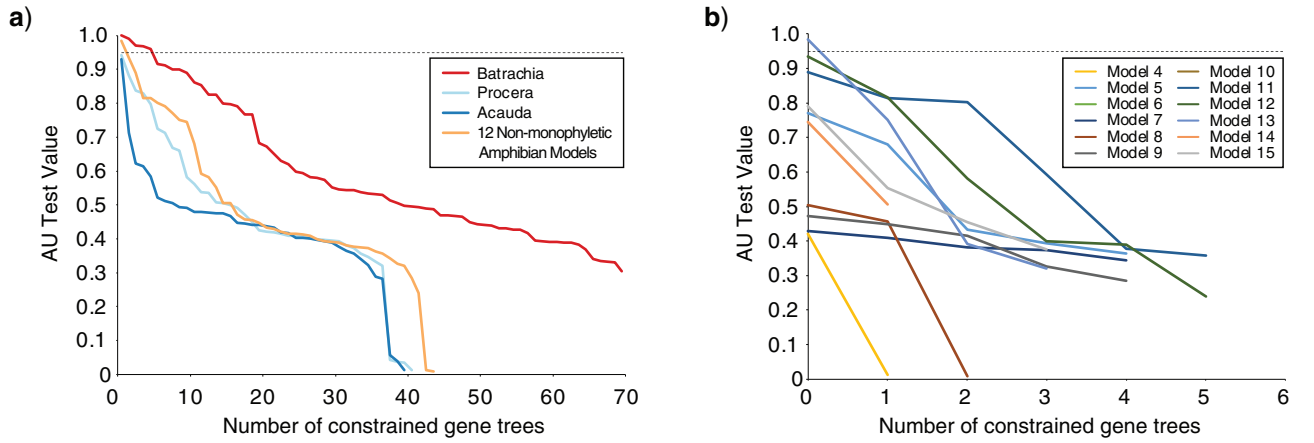

FIGURE 6. Gene genealogy interrogation with the approximately unbiased (AU) test across 195 loci. Results from AU tests performed for each of 195 loci with representatives of each amphibian order. Loci are binned according to which of the 15 possible topologies are supported by unconstrained RAxML gene trees, and then ranked according to the AU test *P*-value supporting their preferred hypothesis. a) Most unconstrained gene trees support a monophyletic Amphibia, but very few are significantly better than alternative trees at a threshold of 0.95. Alternative models for Acauda, Batrachia, and Procera are colored in dark blue, red, and light blue, respectively, and loci supporting any of the 12 models for a nonmonophyletic Amphibia are colored in orange. b) Loci supporting a nonmonophyletic Amphibia are plotted by which of the alternative models are supported. Model numbers correspond to Supplementary Figure S1 available on Dryad. [Color figure can be viewed in the online pdf version]

contrast to the conservative GGI/AU tests, information-theoretic relative model probabilities provided stronger support across loci for all three monophyletic amphibian models. For example, nearly all of the loci supporting each of the three monophyletic amphibian models received relative model probabilities greater than 0.5 and many loci supported one particular model with a relative model probability of 1.0. However, both the GGI/AU tests and the relative model probabilities, like the bootstrap support values, are bounded metrics of support ranging from 0 to 1 and, thus, may potentially obscure some patterns of support across loci.

In contrast, all three of our competing interordinal models had best gene trees that were supported with $\Delta AIC > 40$ ($\Delta AIC > 2$ is typically viewed as strong support for one model over another; Burnham and Anderson 2002). We note, however, that in cases where the degree of topological constraint varies among hypotheses, AIC may be challenging, because the appropriate number of free parameters is unknown. Nonetheless, many studies may benefit from focusing on a set of topological hypotheses with the same degree of topological freedom, as we do here.

When considering all results, we find strong support for amphibian monophyly and the Batrachia hypothesis at the species tree level, and substantial support for discordant gene trees across this deep evolutionary split. Our interpretation of this genealogical discord is ultimately based on the extremely strong measures of ΔAIC support for particular models. We find no correlation between the best supported interordinal models and numbers of taxa from each order, percent missing data across loci, GC content, or the variability of loci (Supplementary Fig. S7 available on Dryad). Several potential and nonmutually exclusive explanations may account for this observation. Our detection of strongly supported discordant gene trees may reflect one of the potential evolutionary phenomena known to yield such patterns (Maddison 1997), including ILS, horizontal transfer, natural selection, saturation, or a gene duplication and loss resulting in errors in orthology assignment (Siu-Ting et al. 2019).

We are unable to definitively discriminate among these possible mechanisms of genealogical discordance. However, the preservation of a historical demographic signal of ILS seems most likely because the number of loci supporting Batrachia is roughly equal to the numbers of loci supporting Procera and Acauda, and because the number of loci supporting these two minor topologies (Procera and Acauda) are also roughly equal, a specific prediction of ILS (Degnan and Rosenberg 2009). Nearly 12 myr separate the 95% credible intervals for the divergences between the lower bound of the divergence between Gymnophiona and the ancestor to Batrachia and the upper bound of the divergence between Caudata and Anura (Fig. 2), and it might be surprising to consider a pattern of ILS persisting over such a long period of time. However, at least one simulation study has demonstrated the potential for ILS to persist over tens to hundreds of millions of years (Oliver 2013). At least some empirical evidence has indicated that the preservation of deep ILS may be present in other major vertebrate clades (e.g. birds, Poe and Chubb 2004; Edwards et al. 2005). Although we observe a pattern of gene tree discordance that is consistent with expectations of ILS. Regardless of the specific mechanism(s) underlying deep gene tree discordance, a sole focus on the estimation and analysis of the species tree topology (rather than also scrutinizing individual gene tree topologies) may lead to a failure to

consider more nuanced aspects of evolutionary history (Hahn and Nakhleh 2016).

Our detection of genealogical discordance across the amphibian genome also suggests that if one were to sample only a few loci at random, the genealogical variation across loci could lead to either weak support for deep nodes or spurious evidence for an alternative interordinal model. At least one other empirical amphibian study has reported discordance across loci, including a finding similar to ours, namely that some loci may support a nonmonophyletic Amphibia (Fong et al. 2012), and stochastic sampling effects combined with deep ILS may together account for the variation across previous studies in terms of which interordinal amphibian topologies have been supported.

Effects of Missing Data in Phylogenomic Data Sets

Our study also emphasized the importance of scrutinizing the effects of large amounts of missing data from individual gene alignments on species tree topologies. We found the effects of missing data to be a particularly important issue in the resolution of relationships among an important set of frog clades, which is expected when branches are very strongly supported (Lemmon et al. 2009). Our initial analyses led to a disagreement between RAxML and ASTRAL trees with respect to the placement of *Nasikabatrachus* (Supplementary Fig. S5a available on Dryad). Although the concatenated ML tree strongly supported the canonical placement (Biju and Bossuyt 2003) of *Nasikabatrachus* as sister to *Sooglossus* (BS= 100), the initial ASTRAL tree supported a surprising and previously unreported sister relationship between *Nasikabatrachus* and the family Microhylidae, with relatively strong support (ASTRAL local posterior = 0.71). Upon closer examination of the placement of *Nasikabatrachus* in the 194 gene trees containing both *Nasikabatrachus* and *Sooglossus*, we found that loci with greater proportions of missing sites (n = 98 loci) consistently favored *Nasikabatrachus* as sister to the microhylid *Oreophryne*, whereas loci with fewer missing data tended to favor traditional placements (Biju and Bossuyt 2003; Feng et al. 2017) of *Nasikabatrachus* sister to *Sooglossus* (n = 94 loci) (Supplementary Fig. S5b available on Dryad). Subsequent filtering of individual gene alignments to exclude taxa with > 50% of sites missing brought the ASTRAL and RAxML topologies into agreement (Supplementary Fig. S5c available on Dryad). Thus, ASTRAL was likely misled by the fact that gene trees supporting the novel placement of *Nasikabatrachus* slightly outnumbered gene trees supporting the previously identified placement as sister to *Sooglossus*. Other studies have also reported conflicting placements for taxa with large amounts of missing data when using either supermatrix (e.g., Lemmon et al. 2009) or species tree methods (Hosner et al. 2016; Moyle et al. 2016). An early expectation for genome-scale phylogenetics was that massive genetic data sets, regardless of levels of missing data, would provide unwavering reconstructions for even the most recalcitrant branches of the Tree of Life through the sheer size of the data. However, whereas adding more genetic loci to species tree analyses can increase the accuracy of the resulting inferences, it is becoming clear that phylogenomic data, on their own, are no panacea for difficult phylogenetic problems (Jeffroy et al. 2006; Brown and Thomson 2016; Shen et al. 2017; Walker et al. 2018).

Genomic Perspectives on the Amphibian Tree of Life

Beyond the firm establishment of the Batrachia relationship at the base of the amphibian Tree of Life, our study also cemented a number of important relationships between and within family-level lineages. A notable exception is that our results support a sister clade relationship between Afrobatrachia and Natatanura (local posterior support = 0.32), with the two as sister of Microhylidae (Fig. 4). This result contrasts with previous hypotheses of amphibian phylogenies based on relatively small amounts of DNA sequence data (e.g., less than 20 loci, Frost et al. 2006; Pyron and Wiens 2011) but agrees with more recent genomic studies (Feng et al. 2017; Siu-Ting et al. 2019; Yuan et al. 2019). Some amphibian clades received relatively high bootstrap support from concatenated RAxML analyses but are nonetheless not well supported by ASTRAL analyses. Within some families (e.g., Plethodontidae and Microhylidae), phylogenetic relationships are less well supported than relationships between families, suggesting some recalcitrant portions of the amphibian phylogeny still remain to be addressed.

Our divergence time estimates for Lissamphibia and Batrachia are somewhat younger than in Roelants et al. (2007), but largely in line with Feng et al. (2017). However, perhaps the most surprising aspects of our study are the relatively younger divergence times among families within each of the three amphibian orders. The Late Cretaceous-Paleogene boundary has been suggested as a critical epoch for amphibian evolution, reflected in the origin and early diversification of several hyperdiverse clades (i.e., Hyloidea, Microhylidae, Afrobatrachia, Natatanura, Salamandridae, and Plethodontidae) (Roelants et al. 2007; Pyron 2014; Feng et al. 2017). Our divergence time estimates, based on the most comprehensive amphibian dataset assembled to date, push the date of origin and initial diversification of frogs, caecilians, and salamanders, later by several million years, and may thus reshape the way we think about the evolution and biogeography of amphibians. This is especially the case concerning hypotheses of continental vicariance and dispersal (e.g., Pyron 2014) and the emphasis currently placed on the Cretaceous–Paleogene boundary (Feng et al. 2017). Understanding the main drivers of amphibian diversification in the Paleogene and Neogene, which contributed so significantly to the

radiation that we see today, will be a major focus of future systematics research on amphibians.

Supplementary Material

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.9kd51c5dc. The sequence data generated in this study are also available from NCBI under BioProject accession PRJNA627509.

Funding

This work was supported by grants from a graduate student research award from the Society of Systematic Biologists and the University of Kentucky G.F. Ribble Endowment (to P.M.H.), by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/BEX 2806/09-6 to P.L.V.P.), and by the National Science Foundation (DEB-0949532 and DEB-1355000 to D.W.W., DEB-1120516 to E.M.L., IIP-1313554 to A.R.L. and E.M.L, DEB-1355071 to J.M.B., DEB-1441719 to R.A.P., DEB-1311442 to P.L.V.P., DEB-1354506 to R.C.T., DEB-1021247 to E.P. and C.J.R., DEB-1021299 to K.M. Kjer, and DEB-1257610, DEB-0641023, DEB-0423286, and DEB-9984496 to C.J.R.), and the Australian Research Council (DP120104146 to J.S.K. and S.C.D.). S.R.R. thanks SENESCYT (Arca de Noé Initiative; SRR and O. Torres-Carvajal principal investigators) for funding for tissue collection. J.L. was supported by the Systematics Association and the Linnean Society Systematics Research Fund. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program (DGE-3048109801 to P.M.H.) and by the National Science Foundation-supported National Center for Supercomputing Applications Blue Waters Graduate Research Fellowship Program (under Grant No. 0725070, subaward 15836, to P.M.H.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Acknowledgements

We thank the following institutions and individuals for providing access to critical tissues samples: American Museum of Natural History (Darrel Frost, David Kizirian, Julie Feinstein); California Academy of Sciences (David Blackburn, Jens Vindum); Florida Museum of Natural History, University of Florida (Pamela Soltis); Biodiversity Institute and Natural History Museum, University of Kansas (Rafe Brown, Linda Trueb, Andrew Campbell); Louisiana State Museum of Natural History, Louisiana State University (Robb Brumfield, Donna Dittmann); Museum of Comparative Zoology, Harvard University (Jim Hanken, José Rosado, Breda Zimkus); Museum of Vertebrate Zoology, University of California, Berkeley (Jim McGuire, Carol Spencer, Ted Papenfuss, Marvalee Wake, Sima Bouzid); Museum Victoria (Jane Melville, Joanna Sumner); National Museum of Natural History (Kevin De Queiroz, Addison Wynn); South African National Biodiversity Institute (Zoe Davids); Saint Louis Zoological Park (Jeffrey Ettling, Mark Wanner, Randall Junge); Universidade de Brasília (Guarino Colli); University of Michigan Museum of Natural History (Ronald A. Nussbaum, Gregory Schneider); Yale Peabody Museum (Gregory Watkins-Colwell); as well as J.J. Apodaca, Alan Channing, Becky Chong, Tyler Frye, S. Blair Hedges, Elizabeth Jockusch, Jarrett Johnson, Christopher McNamara, Eric O'Neill, Todd Pierson, Steve Richards, and Kelly Zamudio. Indian work was conducted with collection permits from state forest departments in Kerala and Maharashtra, with government funding to NCBS and KFRI. JL thanks the Seychelles Islands Foundation, the Seychelles National Parks Authority, the Seychelles Bureau of Standards, the Seychelles Environment Department, Richard A. Griffiths, Jim J. Groombridge, Lindsay Chong-Seng, and Simon T. Maddock. We thank the University of Kentucky Information Technology Department and Center for Computational Sciences for computing time on the Lipscomb High Performance Computing Cluster, and the University of Kansas Center for Research Computing for computing time on the KU CRC Community Cluster. We thank Sean Holland at the FSU Center for Anchored Phylogenomics for assistance with data collection and analysis. We thank Todd Pierson, Andrew Snyder, Brian Gratwicke, Luke Mahler, Michael Durham, Danté Fenolio, Jake Hutton, John Measey, and Alfredo Colón for photograph usage permission. We also thank three anonymous reviewers, and the associate editor and editor-in-chief for comments that improved the manuscript. All work was conducted in accordance with applicable institutional guidelines for animal welfare under IACUC protocols SLZ-2009-04 and SLZ-2010-07 to P.M.H., and UKY-2012-0952 to D.W.W.

References

Abdullayev I., Kirkham M., Björklund, Å.K., Simon, A., Sandberg, R. 2013. A reference transcriptome and inferred proteome for the salamander *Notophthalmus viridescens*. Exp. Cell Res. 319:1187–1197.

Akaike H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19:716–723.

AmphibiaWeb. 2020. Available from: https://amphibiaweb.org, University of California, Berkeley, CA, USA. Accessed April 13, 2020.

Anderson J.S. 2008. Focal review: the origin(s) of modern amphibians. Evol. Biol. 35:231–247.

Anderson J.S., Reisz R.R., Scott D., Fröbisch N.B., Sumida S.S. 2008. A stem batrachian from the Early Permian of Texas and the origin of frogs and salamanders. Nature 45:515–518.

Arcila D., Ortí G., Vari R., Armbruster J.W., Stiassny M.L., Ko K.D., Sabaj M.H., Lundberg J., Revell L.J., Betancur R.R. 2017. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. Nat. Ecol. Evol. 1:0020.

Avise J.C., Robinson T.J. 2008. Hemiplasy: a new term in the lexicon of phylogenetics. Syst. Biol. 57:503–507.

Barrow L.N., Lemmon A.R., Lemmon E.M. 2018. Targeted sampling and target capture: assessing phylogeographic concordance with genome-wide data. Syst. Biol. 67:979–996.

Benton M.J., Donoghue P.C., Asher R.J., Friedman M., Near T.J., Vinther J. 2015. Constraints on the timescale of animal evolutionary history. Palaeontol. Electron. 18:1–106.

Bi K., Vanderpool D., Singhal S., Linderoth T., Moritz C., Good J.M. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13:403.

Biju S.D., Bossuyt F. 2003. New frog family from India reveals an ancient biogeographical link with the Seychelles. Nature 425:711–714.

Brown, J.M., Thomson R.C. 2016. Bayes factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. Syst. Biol. 66:517–530.

Burbrink F.T., Pyron R.A. 2011. The impact of gene-tree/species-tree discordance on diversification-rate estimation. Evolution 65:1851–1861.

Burbrink F.T., Grazziotin F.G., Pyron R.A., Cundall D., Donnellan S., Irish F., Keogh J.S., Kraus F., Murphy R.W., Noonan B.P., Raxworthy C.J., Ruane S., Lemmon A.R., Lemmon E.M., Zaher H. 2020. Interrogating genomic-scale data for Squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. Syst. Biol. doi:10.1093/sysbio/syz062.

Burnham K.P., Anderson D.R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. New York: Springer.

Carstens B.C., Knowles L.L. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. Syst. Biol. 56:400–411.

Carroll, R.L, 2007. The Palaeozoic ancestry of salamanders, frogs and caecilians. Zool. J. Linn. Soc. 150:(suppl_1):1–140.

Chen M.Y, Liang D., Zhang P. 2015. Selecting question-specific genes to reduce incongruence in phylogenomics: a case study of jawed vertebrate backbone phylogeny. Syst. Biol. 64:1104–1120.

Degnan J.H., Rosenberg N.A. 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2:e68.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Duellman, W.E., Trueb L. 1994. Biology of amphibians. Baltimore, USA: Johns Hopkins University Press.

Dunn C.W., Hejnol A., Matus D.Q., Pang K., Browne W.E., Smith S.A., Seaver E., Rouse G.W., Obst M., Edgecombe G.D., Sørensen M.V. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452:745–750.

Ebersberger I., Galgoczy P., Taudien S., Taenzer S., Platzer M., Von Haeseler A. 2007. Mapping human genetic ancestry. Mol. Biol. Evol. 24:2266–2276.

Edwards S.V., Jennings W.B., Shedlock A.M. 2005. Phylogenetics of modern birds in the era of genomics. Proc. R. Soc. Lond., B. Biol. Sci. 272:979–992.

Edwards, S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution 63:1–19.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61:717–726.

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

Feng Y.J., Blackburn D.C., Liang D., Hillis D.M., Wake D.B., Cannatella D.C., Zhang P. 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous-Paleogene boundary. Proc. Natl. Acad. Sci. USA 114:E5864–E5870.

Fong J.J., Brown J.M., Fujita M.K., Boussau B. 2012. A phylogenomic approach to vertebrate phylogeny supports a turtle-archosaur affinity and a possible paraphyletic Lissamphibia. PLoS One 7:e48990.

Frost, D.R., Grant T., Faivovich J., Bain R., Haas A., Haddad C.F.B., de Sá R., Channing A., Wilkinson M., Donnellan S.C., Raxworthy C., Campbell J.A., Blotto B.L., Moler P., Drewes R.C., Nussbaum R.A., Lynch J.D., Green D.M., Wheeler W. 2006. The amphibian tree of life. Bull. Am. Mus. Nat. Hist. 297:1–370.

Frost, D.R. 2020. Amphibian species of the world: an online reference. Version 6.1. Electronic Database accessible. Available from: https://amphibiansoftheworld.amnh.org/index.php. American Museum of Natural History, New York, USA. Accessed April 13, 2020.

Gatesy J., Springer M.S. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80:231–266.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644–652.

Grant T. 2019. Outgroup sampling in phylogenetics: severity of test and successive outgroup expansion. J. Zool. Syst. Evol. Res. 57:748–763.

Green, D.M., Cannatella, D.C. 1993. Phylogenetic significance of the amphicoelous frogs, Ascaphidae and Leiopelmatidae. Ecol. Ethol. Evol. 5:233–245.

Gregory, T.R. 2020. Animal genome size database. Available from: https://www.genomesize.com. Accessed March 20, 2020.

Haeckel, E. 1866. Generelle Morphologie der Organismen: Allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie. Berlin: G. Reimer.

Hahn M.W., Nakhleh L. 2016. Irrational exuberance for resolved species trees. Evolution 70:7–17.

Hamilton C.A., Lemmon A.R., Lemmon E.M., Bond J.E. 2016. Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider tree of life. BMC Evol. Biol. 16:212.

Hellsten U., Harland R.M., Gilchrist M.J., Hendrix D., Jurka J., Kapitonov V., Ovcharenko I., Putnam N.H., Shu S., Taher L., Blitz I.L. 2010. The genome of the western clawed frog *Xenopus tropicalis*. Science 328:633–636.

Heinicke M.P., Lemmon A.R., Lemmon E.M., McGrath K., Hedges S.B. 2018. Phylogenomic support for evolutionary relationships of New World direct-developing frogs (Anura: Terraranae). Mol. Phylogenet. Evol. 118:145–155.

Hobolth A., Dutheil J.Y., Hawks J., Schierup M.H., Mailund T. 2011. Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. Genome Res. 21:349–356.

Hosner P.A., Faircloth B.C., Glenn T.C., Braun E.L., Kimball R.T. 2016. Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Mol. Biol. Evol. 33:1110–1125.

Hutter C.R., Cobb K.A., Portik D.M., Travers S.L., Wood P.L., Brown R.M. 2019. FrogCap: a modular sequence capture probe set for phylogenomics and population genetics for all frogs, assessed across multiple phylogenetic scales. bioRxiv 825307. Available from: https://www.biorxiv.org/content/early/2019/10/31/825307.

Irisarri I., Baurain D., Brinkmann H., Delsuc F., Sire J., Kupfer A., Petersen J., Jarek M., Meyer A., Vences M., Philippe H. 2017. Phylotranscriptomic consolidation of the jawed vertebrate timetree. Nat. Ecol. Evol. 1:1370–1378.

Jeffroy O., Brinkmann H., Delsuc F., Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22:225–231.

Jetz W., Pyron R.A. 2018. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nat. Ecol. Evol. 2:850.

Johnson M., Zaretskaya I., Raytselis Y., Merezhuk Y., McGinnis S., Madden, T. L. 2008. NCBI BLAST: a better web interface. Nucleic Acid. Res. 36:W5–W9.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Lanfear R., Calcott B., Ho S.Y., Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29:1695–1701.

Lanfear R., Calcott B., Kainer D., Mayer C., Stamatakis A. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14:82.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott, B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34:772–773.

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

Lemmon A.R., Emme S.A., Lemmon E.M. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61:727–744.

Lemmon A.R., Brown J.M., Stanger-Hall C., Lemmon E.M. 2009. The effect of ambiguous data on phylogenetic estimates obtained by obtained by maximum-likelihood and Bayesian inference. Syst. Biol. 58:130–145.

Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

McCartney-Melstad E., Mount G.G., Shaffer H.B. 2016. Exon capture optimization in amphibians with large genomes. Mol. Ecol. Res. 16:1084–1094.

Mendes F.K., Hahn M.W. 2016. Gene tree discordance causes apparent substitution rate variation. Syst. Biol. 65:711–721.

Mendes F.K., Hahn Y., Hahn M.W. 2016. Gene tree discordance can generate patterns of diminishing convergence over time. Mol. Biol. Evol.33:3299–3307.

Mirarab S., Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31:i44–i52.

Matsumoto, R. and Evans, S.E. 2018. The first record of albanerpetontid amphibians (Amphibia: Albanerpetontidae) from East Asia. PLoS One 13(1):e0189767.

Moyle R.G., Oliveros C.H., Andersen M.J., Hosner P.A., Benz B.W., Manthey J.D., Travers S.L., Brown R.M., Faircloth B.C. 2016. Tectonic collision and uplift of Wallacea triggered the global songbird radiation. Nat. Commun. 7:12709.

Near T. J., Sanderson M. J. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. Philos. Trans. R. Soc. Series B: Biological Sciences. 359: 1477–1483.

Oliver J.C. 2013. Microevolutionary processes generate phylogenomic discordance at ancient divergences. Evolution 67:1823–1830.

O'Neill E.M., Schwartz R., Bullock C.T., Williams J.S., Shaffer H.B., Aguilar-Miguel X., Parra-Olea G., Weisrock D.W. 2013. Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. Mol. Ecol. 22:111–129.

Page R.D., Charleston M.A. 1997. From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. Mol. Phylogenet. Evol. 7:231–240.

Pardo J.D., Small B.J., Huttenlocker A.K. 2017. Stem caecilian from the Triassic of Colorado sheds light on the origins of Lissamphibia. Proc. Natl. Acad. Sci. USA 114:E5389–E5395.

Peloso P.L., Frost D.R., Richards S.J., Rodrigues M.T., Donnellan S., Matsui M., Raxworthy C.J., Biju S.D., Lemmon E.M., Lemmon A.R., Wheeler W.C. 2016. The impact of anchored phylogenomics and taxon sampling on phylogenetic inference in narrow-mouthed frogs (Anura, Microhylidae). Cladistics 32:113–140.

Poe S., Chubb A.L. 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution 58:404–415.

Pollard D.A., Iyer V.N., Moses A.M., Eisen M.B. 2006. Widespread discordance of gene trees with species tree in *Drosophila*: evidence for incomplete lineage sorting. PLoS Genet. 2:e173.

Portik D.M., Smith L.L., Bi K. 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). Mol. Ecol. Res. 16:1069–1083.

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569–573.

Puttick M.N. 2019. MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. Bioinformatics 35:5321–5322.

Pyron R.A., Wiens J.J. 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. Mol. Phylogenet. Evol. 61:543–583.

Pyron R.A. 2014. Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. Syst. Biol. 65:779–797.

Pyron R.A., Hendry C.R., Hsieh F., Lemmon A.R., Lemmon E.M. 2016. Integrating phylogenomic and morphological data to assess candidate species-delimitation models in brown and red-bellied snakes (*Storeria*). Zool. J. Linn. Soc. 177:937–949.

Rambaut A., Drummond A.J., Xie, D., Baele, G., Suchard, M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67:901–904

Rannala B., Yang Z. 2008. Phylogenetic inference using whole genomes. Annu. Rev. Genomics Hum. Genet. 9:217–231.

Roelants K., Gower D.J., Wilkinson M., Loader S.P., Biju S.D., Guillaume K., Moriau L., Bossuyt F. 2007. Global patterns of diversification in the history of modern amphibians. Proc. Natl. Acad. Sci. USA 104:887–892.

Rokas, A., Carroll, S.B. 2006. Bushes in the tree of life. PLoS Biol. 4:e352.

Rokyta D.R., Lemmon A.R., Margres M.J., Aronow K. 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). BMC Genomics 13:312.

Salichos L., Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497:327–331.

San Mauro D., Vences M., Alcobendas M., Zardoya R., Meyer A. 2005. Initial diversification of living amphibians predated the breakup of Pangaea. Am. Nat. 165:590–599.

Sayyari E., Mirarab S. 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33:1654–1668.

Schoch, R.R. 2019. The putative lissamphibian stem-group: phylogeny and evolution of the dissorophoid temnospondyls. J. Paleo. 93(1):137–156.

Shen X.X., Hittinger C.T., Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. 1:0126.

Sigurdsen, T., Green, D.M. 2011. The origin of modern amphibians: a re-evaluation. Zool. J. Linn. Soc. 162:457–469.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492–508.

Siu-Ting K., Torres-Sánchez M., San Mauro D., Wilcockson D., Wilkinson M., Pisani D., O'Connell M.J., Creevey C.J. 2019. Inadvertent paralog inclusion drives artifactual topologies and timetree estimates in Phylogenomics. Mol. Biol. Evol. 36:1344–1356.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Streicher J.W., Miller E.C., Guerrero P.C., Correa C., Ortiz J.C., Crawford A.J., Pie M.R., Wiens J.J. 2018. Evaluating methods for phylogenomic analyses, and a new phylogeny for a major frog clade (Hyloidea) based on 2214 loci. Mol. Phylogenet. Evol. 119:128–143.

Suh A., Smeds L., Ellegren H. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. PLoS Biol. 13:e1002224.

Sun M., Soltis D.E., Soltis P.S., Zhu X., Burleigh J.G., Chen Z. 2015. Deep phylogenetic incongruence in the angiospersm clade Rosidae. Mol. Phylogenet. Evol. 83:156–166.

Swofford, D. L. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, MA: Sinauer Associates.

Townsend J.P. 2007. Profiling phylogenetic informativeness. Syst. Biol. 56:222–231.

Walker J.F., Brown J.W., Smith S.A. 2018. Analyzing contentious relationships and outlier genes in phylogenomics. Syst. Biol. 67:916–924.

Weisrock D.W., Hime P.M., Nunziata S.O., Jones K.S., Murphy M.O., Hotaling S., Kratovil J.K. 2018. Surmounting the large-genome "problem" for genomic data generation in salamanders. In: Hohenlohe P., Rajora O., editors. Wildlife genomics. New York, USA: Springer.

White M.A., Ané C., Dewey C.N., Larget B.R., Payseur B.A. 2009. Fine-scale phylogenetic discordance across the house mouse genome. PLoS Genet. 5:e1000729.

Wilberg E.W. 2015. What's in an outgroup? The impact of outgroup choice on the phylogenetic position of *Thalattosuchia* (Crocodylomorpha) and the origin of Crocodyliformes. Syst. Biol. 64:621–637.

Wu C.H., Tsai M.H., Ho C.C., Chen C.Y., Lee H. S. 2013. *De novo* transcriptome sequencing of axolotl blastema for identification of differentially expressed genes during limb regeneration. BMC Genomics 14:434.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yuan Z., Zhang B., Raxworthy C.J., Weisrock D.W., Hime P.M., Jin J., Lemmon A.R., Lemmon E.M., Holland S.D., Kortyna M.L., Zhou W., Peng M., Che J., Prendini E.S. 2019. Natatanuran frogs used the Indian Plate to step-stone disperse and radiate across the Indian Ocean. Nat. Sci. Rev. 6:10–14.

Zardoya R. and Meyer A., 2001. On the origin of and phylogenetic relationships among living amphibians. Proc. Natl. Acad. Sci. USA 98:7380–7383.