Molecular Phylogenetics and Evolution 204 (2025) 108272

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

A time-calibrated salamander phylogeny including 765 species and 503 genes

Alexander A. Stewart *, John J. Wiens

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721-0088, USA

ARTICLE INFO

Keywords:
Divergence dating
Missing data
Phylogeny
Phylogenomics
Salamander
Supermatrix

ABSTRACT

Recent time-calibrated amphibian phylogenies agree on the family-level relationships among extant salamanders but had disparate sampling regimes and inferred very different divergence times. For example, a recent phylogenomic study based on 220 nuclear loci had limited taxon sampling (41 species) and estimated relatively young divergence dates, whereas a more extensive supermatrix study based on 15 genes and 481 species estimated dates that were 22–45 million years older for major clades. Here, we combined phylogenomic and supermatrix approaches to estimate the largest salamander phylogeny to date based on molecular markers. Our matrix contained 765 salamander species and 503 genes (with 92.3% missing data overall). We included 284 more species than the previous largest salamander phylogeny (59% increase) and sampled approximately 93% of all currently described salamander species. Our dating analyses incorporated more than twice as many fossil calibration points within salamanders as previous studies. Maximum-likelihood estimates of tree topology yielded family-level relationships that were consistent with earlier studies. Nearly all species were placed in the expected genera, despite extensive missing data in many species. Bootstrap support was generally high across the tree but was poor in some clades where sampling of genes was limited (e.g., among some bolitoglossine salamanders). The dating analyses yielded age estimates for major clades that were generally intermediate between those from the previous phylogenomic and supermatrix analyses. We also provide a set of 200 time-calibrated trees for use in comparative analyses.

1. Introduction

Salamanders are one of the three major groups of living amphibians. With 816 species currently described, they have far fewer species than anurans (7,682 species) but far more than caecilians (222 species; AmphibiaWeb, 2024; 4 Feb 2024). They include 10 families and 68 genera. They are distributed mostly in the temperate Northern Hemisphere (North America, Europe, and Asia), but with a major radiation (tropical bolitoglossines) that extends from tropical Mexico into northern South America.

Salamanders are pivotal research subjects for many topics in ecology and evolution, including the evolution of paedomorphosis and direct development (Bonett et al., 2014; Liedtke et al., 2022), genome-size evolution (Sessions, 2008; Sun et al., 2012; Liedtke et al., 2018), morphological evolution (Mueller et al., 2004; Bonett and Blair, 2017; Bonett et al., 2018), transitions from ZW to XY sex determination (Hime et al., 2019), the origins of species richness patterns (Wiens et al., 2007; Kozak and Wiens, 2012), community assembly (Kozak et al., 2005, 2009), the evolution of lunglessness (Lewis et al., 2022), and hybridization (Melander and Mueller, 2020; Pyron et al., 2022b; Pierson et al., 2024). For many of these topics, having accurate estimates of salamander phylogeny is crucial.

In some ways, the current resolution of salamander phylogeny is relatively good. For example, relationships among salamander families have been relatively stable for decades (at least based on molecular data). A recent phylogenomic analysis (Hime et al., 2021) included all 10 families (34 genera and 41 species) based on data from 220 nuclear loci from anchored hybrid enrichment (AHE). Furthermore, a recent supermatrix study (Jetz and Pyron, 2018) sampled 481 salamander species based on sequence data (with 178 more added randomly based on taxonomy), including more than 50% of all known species. Both studies yielded congruent estimates of relationships among salamander families.

At the same time, there are important issues in salamander

Abbreviations: AHE, anchored hybrid enrichment.

* Corresponding author.
E-mail address: alexanderstewart@arizona.edu (A.A. Stewart).

https://doi.org/10.1016/j.ympev.2024.108272
Received 29 March 2024; Received in revised form 23 November 2024; Accepted 9 December 2024
Available online 15 December 2024

phylogeny that remain unresolved. First, these two recent studies (Jetz and Pyron, 2018; Hime et al., 2021) estimated very different divergence dates for these major salamander clades (Table 1). For these clades (Table 1), the estimates from Jetz and Pyron (2018) were ~22–45 million years older than the estimates from Hime et al. (2021). These differences are especially striking given that crown-group salamanders seem to be less than 200 million years old (Table 1). Further, both studies included only a limited number of fossil calibration points. The number of calibration points may be far more important in determining the ages estimated than other factors, like the number of genes sampled or the amount of missing data (Zheng and Wiens, 2015). Second, 63 % of salamander species and 41 % of salamander genera belong to one family (Plethodontidae; AmphibiaWeb, 2024). Therefore, the extent to which salamander phylogeny can be considered resolved depends (in part) on the agreement among studies about relationships in this clade. Unfortunately, Hime et al. (2021) sampled only 12 of the 28 genera in this family, and only two genera of tropical bolitoglossines (a clade containing the majority of plethodontid genera and species; 14/28 genera and 327/516 species; AmphibiaWeb, 2024). Further, some relationships in Plethodontidae were clearly discordant between Jetz and Pyron (2018) and Hime et al. (2021), despite the limited taxon sampling in the latter study. For example, Hime et al. (2021) placed Hemidactylium as the sister group to the tribe Bolitoglossini, whereas Jetz and Pyron (2018) placed *Hemidactylium* with the tribe Spelerpini (including *Eurycea*, *Gyrinophilus*, and *Pseudotriton*). Similarly, Hime et al. (2021) placed Plethodon as the sister group to the clade of *Phaeognathus* + *Desmognathus*, whereas Jetz and Pyron (2018) placed *Aneides* in this position instead. It is unclear whether these differences are explained by the limited sampling of genes in Jetz and Pyron (2018), limited sampling of taxa in Hime et al. (2021), or some other factor.

Here, we provide an improved estimate of salamander phylogeny. First, we combine the phylogenomic and supermatrix approaches to maximize the number of genes sampled and the number of species sampled (e.g., Cho et al., 2011; Portik et al., 2023a,b; Talavera et al., 2022; Wiens et al., 2005; Zheng and Wiens, 2016). Specifically, we start with the phylogenomic dataset of Hime et al. (2021), which spans all salamander families. We also include data from several other phylogenomic studies within families (Ambystomatidae: Everson et al., 2021; Plethodontidae: Pyron et al., 2020, 2022b; Salamandridae: Rancilhac et al., 2021). We then query GenBank using gene names from the markers in the Hime et al. (2021) dataset, gene names from the markers in Shen et al. (2013, 2016), and for seven mitochondrial and nine nuclear markers commonly used in previous studies (e.g., Wiens et al., 2007; Kozak et al., 2009; Vieites et al., 2011; Pyron and Wiens, 2011; Pyron, 2014; Rovito et al., 2015; Jetz and Pyron, 2018). We then combine these datasets into a single matrix for maximum-likelihood analysis, using the data-matrix assembly program SuperCRUNCH (Portik and Wiens, 2020). We also time calibrate this tree (using penalized likelihood; Sanderson, 2002; Smith and O'Meara, 2012), with a more extensive set of fossil calibration points than used in previous studies of salamander phylogeny. Finally, we generate a set of 200 time-calibrated trees for use in comparative analyses.

Table 1
Comparing estimated ages of major salamander clades among selected studies, including analyses with all 13 fossil calibration points within salamanders and alternative analyses with only 10.

Clade	This study (13 points)	Hime et al. (2021)	Jetz & Pyron (2018)	This study (10 points)
Salamander crown	175.0	159.2	197.0	174.1
Cryptobranchoidea	112.2	130.9	152.8	106.3
Sirenidae + Salamandroidea	167.2	145.6	182.8	168.5
Salamandroidea	154.1	123.8	165.0	157.6
Salamandridae + Dicamptodontidae + Ambystomatidae	140.3	107.2	149.2	142.7
Dicamptodontidae + Ambystomatidae	90.2	62.8	96.5	91.1
Salamandridae	77.5	49.1	74.9	77.9
Proteidae + Rhyacotritonidae + Amphiumidae + Plethodontidae	148.2	115.4	153.7	152.1
Rhyacotritonidae + Amphiumidae + Plethodontidae	128.6	93.8	138.5	138.8
Amphiumidae + Plethodontidae	119.8	80.2	123.2	129.6
Plethodontidae	92.0	47.7	88.6	104.4

2. Materials and methods

2.1. Outgroup selection

We selected amphibian outgroups from among the species sampled by Hime et al. (2021). Specifically, we selected a single representative species from each of the 11 most basal frog families (those closest to the root in terms of number of nodes). We also included all 14 caecilian species with valid binomial names used in that study (representing 9 of the 10 caecilian families). In addition to the amphibian species from Hime et al. (2021), we selected additional outgroup species that had whole genomes available, including species from Gymnophiona (*Geotrypetes seraphini*, *Microcaecilia unicolor*, and *Rhinatrema bivittatum*), Anura (*Xenopus tropicalis*), Squamata (*Anolis carolinensis*), Testudines (*Chrysemys picta*), Aves (*Gallus gallus*), Mammalia (*Homo sapiens*), and Coelacanthiformes (*Latimeria chalumnae*). Among the outgroup species that had whole genomes available, three overlapped with the species sampled by Hime et al. (2021), and both AHE and genomic data were considered for those species (*G. seraphini*, *R. bivittatum*, and *X. tropicalis*). In total, we included 31 outgroup species, including 11 frogs, 15 caecilians, and five more distant outgroup species.

2.2. Sequence assembly

We identified 16 genetic markers that were commonly utilized in other studies of caudate molecular phylogenetics (e.g., Wiens et al., 2007; Kozak et al., 2009; Vieites et al., 2011; Pyron, 2014; Rovito et al., 2015). These included seven mitochondrial markers (12S, 16S, *CO1*, *CO2*, *CYTB*, *ND2*, and *ND4*) and nine nuclear genes (*BDNF*, *CXCR4*, *H3A*, *NCX1*, *POMC*, *RHO*, *SIA*, *TYR*, and *RAG1*). In addition to this core set of genes, we targeted two other sets of nuclear genes that have been sequenced at a broad phylogenetic scale in caudates. The first set consisted of 56 PCR-based nuclear protein-coding loci sequenced by Shen et al. (2013, 2016). The second set consisted of 220 AHE loci sequenced by Hime et al. (2021). Between these two datasets, 25 genes had identical or synonymous names. Sequence alignment via BLASTn (Camacho et al., 2009) revealed that 16 of these 25 redundant genes had complete or partial sequence overlap, indicating they could effectively be treated as the same marker. A summary of all markers is given in Supplementary Table S1, and a summary of merged markers is given in Supplementary Table S2.

The 263 non-redundant gene names we obtained were used to query GenBank, and an additional search term was added to capture whole mitogenomes. Searches were performed for each locus individually at each taxonomic level (i.e., "Caudata" for salamanders and specific names for each outgroup without a sequenced genome). Searches were performed on 12 June 2024, and search results from all markers were concatenated into a single file. Because all markers sequenced by Shen et al. (2013, 2016) were available on GenBank, they were captured at this point and included in the GenBank dataset. The resulting sequences were processed to remove loci that did not match search terms, to remove sequences without a binomial name (i.e., sequences with open nomenclature qualifiers), and to reformat the file for SuperCRUNCH

using custom scripts written in Python (available at https://github.com/aast242/salamander-timetree). For sequences identified to the subspecies level, if a sequence for the nominate subspecies was present, only nominate sequences were retained for the GenBank dataset. If the nominate subspecies did not have an available sequence, all sequences for the species were retained. In all cases, subspecies were truncated to the species level after processing. In cases of obvious nomenclatural error on GenBank, species names were changed to match AmphibiaWeb taxonomy (e.g., *Triton torosa* changed to *Taricha torosa*). While we mainly utilized the taxonomy presented by AmphibiaWeb (2024), we considered *Hydromantes* and *Speleomantes* as separate genera, following Amphibian Species of the World (Frost, 2024). We also included three species (*Hynobius amabensis*, *Hynobius miyazakiensis*, and *Tylototriton sini*) not present in AmphibiaWeb (2024) but present in Amphibian Species of the World (Frost, 2024). The AHE sequences from Hime et al. (2021) were directly downloaded from GenBank via BioProject accession PRJNA627509 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA627509) and similarly reformatted for SuperCRUNCH.

We also searched for additional species that were not found in our initial GenBank searches. For all salamander species that were present on AmphibiaWeb (2024) but did not have any associated sequences on GenBank we searched the literature and museum databases to identify sequences for that species. For example, all sequences on GenBank that should have been labelled as *Aneides klamathensis* were actually labelled as *A. flavipunctatus* (the species that *A. klamathensis* was split from). Some salamander species also had sequences present on GenBank, but they were initially labelled with an open-taxonomy modifier that was not updated after the specimen was assigned to the species level (e.g., GenBank accession AY728235 is listed as *Bolitoglossa n. sp*. RLM-2004, but is listed as *B. sombra* in the associated museum database).

Primary articles describing species that lacked correctly labeled GenBank sequences were located via AmphibiaWeb (2024), Amphibian Species of the World (Frost, 2024), or manual searches. Manual searches were performed by searching Google Scholar using the species name and "description" as query terms, and all literature searches were performed from 7 June 2024 to 12 June 2024. We then used these primary articles to identify the GenBank numbers of specimens assigned to these species. Relevant sequences for these species were then renamed in our dataset following the nomenclature in the article. In cases where specimen vouchers for sequence data were provided, the voucher number was searched for on Arctos (Cicero et al., 2024) and the nomenclature for the relevant specimens was followed. In the case of the *Aneides flavipunctatus* complex, many specimens were still labeled as *A. flavipunctatus* on Arctos (Cicero et al., 2024), so taxonomy was assigned based on the clear geographic boundaries between the four species (Reilly and Wake, 2019). Any individuals collected at contact zones between species were excluded. Manual taxonomy modifications added 47 salamander species to the tree, and all sequences renamed via this method are presented in Supplementary Table S3.

Four additional datasets were sourced from recent salamander phylogenomics studies. The first dataset consisted of 14 nuclear loci sequenced in various *Ambystoma* and *Dicamptodon* species by Williams et al. (2013). Because these loci were not all available on GenBank, sequences were directly downloaded from the Dryad repository associated with that study (https://doi.org/10.5061/dryad.2gq14). The second dataset was derived from two AHE studies on the genera *Desmognathus* and *Phaeognathus* by Pyron et al. (2020, 2022b). Sequences were downloaded from the Dryad repositories associated with each study (https://doi.org/10.5061/dryad.34tmpg4g1 and https://doi.org/10.5061/dryad.f4qrfj6x8, respectively). All candidate mitonuclear species of *Desmognathus* were renamed following recent refinement of the taxonomy within the genus (Means et al., 2017; Pyron and Beamer, 2022a, 2022b, 2022c, 2023a, 2023b; Pyron et al., 2022a, 2023). Candidate species that did not have a binomial name (*D. orestes* A/C and *D. orestes* B) were treated as distinct species in subsequent analyses. The third dataset was procured from a recent study on tiger salamander (*Ambystoma*) biogeography (Everson et al., 2021). This study sequenced a previously developed panel of 95 nuclear loci (O'Neill et al., 2013) in a broad selection of species within the *Ambystoma tigrinum* complex. Sequences were directly downloaded from the GitHub repository associated with that project (https://github.com/kelly-sovacool/tiger_salamander_project). The final dataset was from a recent transcriptomics study in salamandrids (Rancilhac et al., 2021). Data were acquired via direct correspondence with the authors. BLASTn (Camacho et al., 2009) searches were utilized to match orthologous transcripts to existing markers in the matrix. All directly downloaded datasets were reformatted for SuperCRUNCH using custom Python scripts (available at https://github.com/aast242/salamander-timetree).

At this point, all the datasets were combined into a single fasta file. These included the data from GenBank (including sequences from Shen et al., 2013, 2016), Williams et al. (2013), Hime et al. (2021), Everson et al. (2021), Pyron et al. (2020, 2022b), and Rancilhac et al. (2021). Sequence overlaps between markers from different datasets were examined using BLASTn (Camacho et al., 2009). Overlapping markers were merged into a single marker. A detailed breakdown of the data source for each marker is presented in Supplementary Table S1. A summary of merged markers is presented in Supplementary Table S2. After markers from different datasets were merged, the single file containing all sequences was split into locus-specific fasta files that could enter the SuperCRUNCH pipeline at the "Similarity Filtering" step.

2.3. SuperCRUNCH analysis

We utilized SuperCRUNCH (Portik and Wiens, 2020) to process our sequences and prepare them for concatenated phylogenetic analyses. All data files associated with the construction of the salamander supermatrix are available at https://github.com/aast242/salamander-timetree. A more detailed description of the SuperCRUNCH pipeline can be found in the recent analysis of frog phylogeny by Portik et al. (2023a), but we provide a brief overview here.

Sequence-similarity filtering was performed to exclude sequences not sharing significant identity with other sequences assigned to the same marker. Mitochondrial genes were filtered using *Reference_Blast_Extract.py* with default settings and reference files for each gene. These reference files were manually constructed from 15 complete mitogenomes spanning all salamander families and from five complete mitogenomes from outgroup species (see Supplementary Table S4). After filtering, mitochondrial genes were checked for human contamination using *Contamination_Filter.py* with default settings. Nuclear genes were filtered using *Cluster_Blast_Extract.py* with default settings.

After similarity filtering, the longest sequence for each marker in each species was chosen as the representative sequence using *Filter_Seqs_and_Species.py*. We used a minimum length requirement of 150 bp (−m 150). Loci represented by at least 15 species were retained in the analysis pipeline. These last two criteria are somewhat arbitrary, but the goal was to exclude markers that were extremely short within a species or that were extremely incomplete among species. All representative sequences were oriented in the same direction using *Adjust_Direction.py* and open reading frames were identified in coding sequences using *Coding_Translation_Tests.py* with the appropriate translation table (nuclear vs. mitochondrial). Coding sequences were aligned using MACSE v2 (Ranwez et al., 2018) through *Align.py* with thorough search settings and the dual-alignment option enabled. Markers represented solely by sequences that failed translation tests and sequences of 12S and 16S rRNAs were aligned using MAFFT (Katoh et al., 2002) through *Align.py* with thorough search settings. Any markers identified as problematic via gene-tree validation (see Section 2.5 below) that were not replaced by an acceptable sequence after two rounds of validation were excluded at this point. Sequence alignments were trimmed using trimAl (Capella-Gutiérrez et al., 2009) through *Trim_Alignments_Trimal.py* to remove columns with gaps in more than 90% of the sequences across species. This was followed by the removal of columns with 100% missing data.

Prior to sequence concatenation, we renamed sequences to contain only the species name. We also removed identical sequences by checking if each marker alignment could be properly read by RAxML (Stamatakis, 2014) and using the resulting "reduced" file. Markers with at least 10 salamander sequences were retained and *Concatenation.py* was used to generate the final concatenated alignment file.

2.4. Sequence selection from outgroup genomes

After a single representative sequence was chosen in each species for each marker by SuperCRUNCH, sequences were combined into a single fasta file and used to query outgroup genomic coding sequences using BLASTn (Camacho et al., 2009) with a 1e-10 E-value cutoff. The outputs from BLASTn searches were used to extract the best matching sequence for each marker from each of the outgroup genomes using custom scripts written in Python (available at https://github.com/aast242/salamander-timetree). These extracted outgroup sequences were added to the SuperCRUNCH starting sequences, and all SuperCRUNCH analyses were rerun with the added sequences.

2.5. Sequence validation

We also checked the estimated gene tree from each marker to identify sequences that potentially represented contamination, incorrect identifications, or other problems. We estimated a maximum-likelihood gene tree for each marker with 100 rapid bootstraps in RAxML (Stamatakis, 2014). These trees were visualized using ETE 3 (Huerta-Cepas et al., 2016). We checked these trees for species that were associated with extremely long branches or extremely short branches (i.e. identical to another species), and those that were placed in the wrong genus or family. Whenever possible, problematic sequences were replaced with a different sequence for the same marker by removing the sequence from the initial dataset and allowing SuperCRUNCH to choose another representative. Genomic outgroup sequences with discordant gene names (e.g., marker is *FAT4* but outgroup annotation is *FAT1*) were also removed from the initial dataset. This process was repeated twice in an attempt to find an appropriate sequence for each species from each marker. A third, final round of validation was performed in which any remaining anomalous sequences were removed from the single-sequence SuperCRUNCH dataset and not replaced. GenBank accessions and unique identifiers for the final sequences included in concatenated analyses are presented in Supplementary Table S5.

We also attempted to limit the impact of using sequences from taxa that have undergone extensive taxonomic revision. Specifically, we removed all sequences from recently revised *Desmognathus* species that predated the taxonomic revision (e.g., *D. fuscus* previously represented three lineages that have been subsequently elevated to species). We also removed all sequences from *Paramesotriton labiatus* from before the delineation of *Paramesotriton* and *Pachytriton* by Nishikawa et al. (2011). Neither of these removals excluded all sequences for a given species, and representative sequences were still chosen by SuperCRUNCH for all loci affected by these removals.

2.6. Partitioning

We used PartitionFinder 2 (Lanfear et al., 2017) to identify an optimal partitioning strategy for the concatenated set of all 512 markers. Following Portik et al. (2023a), we performed two partitioning analyses. Our first analysis used a single partition per codon position per protein-coding gene, with a single partition for each rRNA gene. The second used one partition per marker. These analyses enabled selection of optimal partitioning strategies with 1179 and 398 partitions, respectively. Preliminary maximum-likelihood (ML) analyses of the partitioned datasets using RAxML (Stamatakis, 2014) proved to be extremely slow, with neither strategy producing any trees after running for 48 h on a server with 64 threads and 128 GB of RAM. We therefore used a strategy with a more limited set of partitions, following Portik et al. (2023a). This strategy included eight partitions: one partition for each rRNA (12S and 16S), one partition for each codon position across all the mitochondrial protein-coding loci, and one partition for each codon position across all the nuclear protein-coding loci.

2.7. Maximum-likelihood tree estimation

Concatenated phylogenetic analyses were performed in RAxML (Stamatakis, 2014). We used the GTR + CAT substitution model (general-time reversible with the CAT approximation of the gamma distribution of among-site rate heterogeneity) with branch-length calculations. We performed maximum-likelihood (ML) estimation on 25 distinct starting topologies to find the best scoring ML tree. Estimates were conducted in three separate analyses (two runs with ten trees each, and one run with five trees), each with unique starting seeds and a final branch-length optimization step for the best scoring tree in each run. We also performed rapid bootstrapping analyses using the GTR + CAT substitution model with branch-length calculations to generate 200 trees (with branch lengths) for use in comparative analyses. The 200 bootstrap trees also allowed us to generate confidence intervals for the dating analyses and to gauge branch support for the best-scoring ML topology.

2.8. Divergence-dating analyses

We used penalized likelihood (Sanderson 2002) implemented in treePL (Smith and O'Meara, 2012) to estimate divergence-times for the best scoring ML tree and the 200 bootstrap trees. We used 22 calibration points based on fossil evidence to constrain the ages of nodes within the trees. The collection of fossil calibrations from Portik et al. (2023a) was used as a starting point, but we surveyed other studies to refine our fossil calibrations (Gao and Shubin, 2012; Feng et al., 2017; Jetz and Pyron, 2018; Hime et al., 2021). A detailed description of the fossil calibrations is provided in Supplementary File S1. We included a total of 22 fossil calibration points, including 13 within salamanders.

To determine the optimal smoothing parameter for the best-scoring ML tree, we performed a thorough random subsample and replicate cross-validation analysis in treePL. Smoothing parameter values ranged from 1e-15 to 1e+10 in tenfold increments (1e-15, 1e-14, 1e-13, etc.). To ensure that the results of these analyses were stable, the cross-validation analysis was repeated ten times, and the smoothing parameter with the lowest chi-square value across runs was kept as the optimal value for the ML tree.

The optimal smoothing parameter was then used to estimate divergence-times for the best scoring ML tree and all 200 bootstrap replicates. Following Portik et al. (2023a), we did not repeat the cross-validation analysis on each bootstrap replicate: since the bootstrap replicates were derived from the original data, we assumed that the best-fit smoothing parameter for the original data should have the best fit. Furthermore, the cross-validation analyses were very computationally intensive. The 200 dated bootstrap trees were summarized onto the ML tree topology using TreeAnnotator (Suchard et al., 2018) to generate 95% confidence intervals for node ages. An additional set of dating analyses was performed as described above but with the fossil constraints associated with Cryptobranchidae, Proteidae, and Sirenidae removed (see justification in Supplementary File S1).

For each major group of salamanders, we illustrated the optimal maximum-likelihood tree and estimated divergence dates, along with the bootstrap values for each node. We present the confidence intervals for each node in a series of supplementary figures matching those in the main text. We found that figures containing both bootstrap values and confidence intervals were too difficult to read, given the many nodes in most figures.

2.9. Potential biases in branch lengths and divergence times

Many species in the tree were represented only by mitochondrial data, and many species had extensive missing data. Therefore, we tested how missing data and mitochondrial data might have influenced the estimates of divergence times. Previous analyses suggest that there should be no impact of missing data on estimated divergence times (Zheng and Wiens, 2015), but species with only mitochondrial data might have longer estimated branch lengths (given faster expected rates of change in mitochondrial markers; Mulcahy et al., 2012). We examined the relationships between each species' maximum-likelihood terminal (tip) branch lengths, their proportion of missing data, and their proportion of mitochondrial markers present (among all markers present in that species). More relevant to potential biases, we also examined these relationships using the terminal branch lengths from the time-calibrated tree. All statistical tests were performed using SciPy version 1.10.0 (Virtanen et al., 2020). First, we tested for normality in the distribution of raw and time-calibrated branch lengths using the Shapiro-Wilk test, both for untransformed and log-transformed data. Non-parametric tests were used in further analyses, since both branch-length measures failed this test ($p < 0.001$ for both measures). All tests included only the ingroup (salamander) species, for a total of 765 species.

We performed Mann-Whitney U tests to determine whether the mean branch lengths of species represented solely by mitochondrial markers ($n = 263$) were significantly longer than the branch lengths of those species represented by one or more nuclear markers in addition to the mitochondrial markers ($n = 502$). Mann-Whitney U tests were performed using both raw and time-calibrated branch lengths. Kendall's tau was calculated to assess correlations between branch lengths (maximum likelihood and time-calibrated), missing data, and the proportion of mitochondrial markers in each species.

2.10. Nuclear-only phylogeny

We also tested if the overall tree topology might have been influenced by mitochondrial markers. Therefore, we constructed a small matrix without mitochondrial markers (nuclear only). Given that many species lacked substantial numbers of nuclear markers, we included only one species per genus and included the species with the largest number of nuclear markers in each genus. For the three genera inferred to be paraphyletic (*Cynops*, *Tylototriton*, and *Pseudoeurycea*) we included one species from each clade of species from these genera (two species sampled per genus). Sequence files from the main SuperCRUNCH analysis were filtered to contain only the species described above; then, markers with at least five salamander species were retained for subsequent analyses. Those markers passing this filtering step were trimmed an additional time and concatenated as previously described to generate the final concatenated alignment. We used one partition for each codon position, performed ML estimation on 25 distinct starting topologies using RAxML, and generated 100 rapid bootstrap trees to gauge branch support. We then compared the overall topology of this tree to the main tree including all markers and taxa.

3. Results

3.1. Properties of the dataset

The final dataset entering the SuperCRUNCH pipeline consisted of 575 markers (568 nuclear markers and 7 mitochondrial markers), including 14 nuclear markers from Williams et al. (2013), 56 nuclear markers from Shen et al. (2013, 2016), 220 nuclear markers from Hime et al. (2021), 95 nuclear markers from Everson et al. (2021), 381 nuclear and 2 mitochondrial markers from Pyron et al. (2020), 233 nuclear markers from Pyron et al. (2022b), and 81 nuclear markers from Rancilhac et al. (2021). After filtering and processing these markers, 503 were included in the final matrix, including 7 mitochondrial markers and 496 nuclear markers (Supplementary Table S1). A summary of the types of markers representing the species in each family is supplied in Supplementary Table S6. We note that Rancilhac et al. (2021) included 5455 nuclear markers but only for 40 species, so most markers that were unique to that study across salamanders were excluded given our filtering criterion.

The final GAMMA-based score of the best tree resulting from maximum-likelihood estimation in RAxML was −8,964,262. The search for this best-fit tree took approximately two weeks running on a server with 32 threads and 128 GB of RAM. The estimation of 200 rapid bootstrap trees took approximately nine days running on the same system. The cross-validation analyses in treePL showed that a smoothing parameter of 1e-10 had the lowest chi-square value out of all 10 runs. This value was the lowest in three runs, and was within two points of the lowest chi-square value in six of the remaining seven runs. Smoothing parameters and the associated chi-square values for each of the 10 runs are available as Supplementary Table S7. The maximum-likelihood tree is available as Supplementary File S2, the time-calibrated tree as Supplementary File S3, the set of 200 bootstrapped, time-calibrated trees is available as Supplementary File S4, and the time-calibrated tree with 95 % confidence intervals is available as Supplementary File S5.

3.2. Higher-level relationships and divergence dates

Relationships among salamander families (Fig. 1) were strongly supported (bootstrap, bs = 99–100%), and closely matched those from previous studies. The Cryptobranchoidea (Cryptobranchidae, Hynobiidae) was strongly supported as monophyletic. Within the sister group to Cryptobranchoidea, Sirenidae was the sister group to the clade containing all other families (Salamandroidea). Within Salamandroidea, there was a strongly supported clade uniting Salamandridae, Ambystomatidae, and Dicamptodontidae, with the latter two families as sister taxa. Proteidae was the sister group to the remaining three families, which consisted of Rhyacotritonidae and the sister families Amphiumidae + Plethodontidae. Overall, these family-level relationships were identical to those in Jetz and Pyron (2018) and Hime et al. (2021). Earlier studies have also found identical relationships, such as Shen et al. (2013).

There was considerably more discordance among previous studies in the divergence dates for these clades. As summarized in Table 1, the dates for these major clades estimated by Hime et al. (2021) were much younger than those estimated by Jetz and Pyron (2018), by roughly 22–45 million years (mean = 37 million years). The dates estimated here are summarized in Fig. 1, with confidence intervals on dates for comparable clades in Supplementary Fig. S1. These dates were generally intermediate between these two sets of estimates (8 out of 11 clades; Table 1). The only exceptions were Cryptobranchoidea (considerably younger than both of these previous estimates, by 18.7–40.6 million years), Salamandridae (older by 2.6 million years than Jetz and Pyron, 2018), and Plethodontidae (older by 3.4 million years than Jetz and Pyron, 2018). On average, the ages estimated here for these 11 clades (Table 1) were 26.3 Myr older than those of Hime et al. (2021) and 10.6 Myr younger than those estimated by Jetz and Pyron (2018). Our estimated dates are also broadly similar to those of Shen et al. (2016).

We also performed alternative dating analyses on the best maximum-likelihood tree with fossil calibration points removed for Cryptobranchidae, Proteidae, and Sirenidae. In this alternative dated tree, the crown-group age of Cryptobranchidae was 13.3 Mya (versus 56.9 Mya in the main analysis), Proteidae was 106.8 Mya (versus 104.1 Mya), and Sirenidae was 38.5 Mya (versus 46.2 Mya). Estimates for 11 major clades (Table 1) were older than the main analyses for 9 clades, but generally within 10 Mya. Trees showing the relationships between the salamanders in the main and alternative dating analyses are presented as *Supplementary Figs. S2 and S3*, respectively. For ease of visualization and comparison, only relationships among genera are shown. The full tree

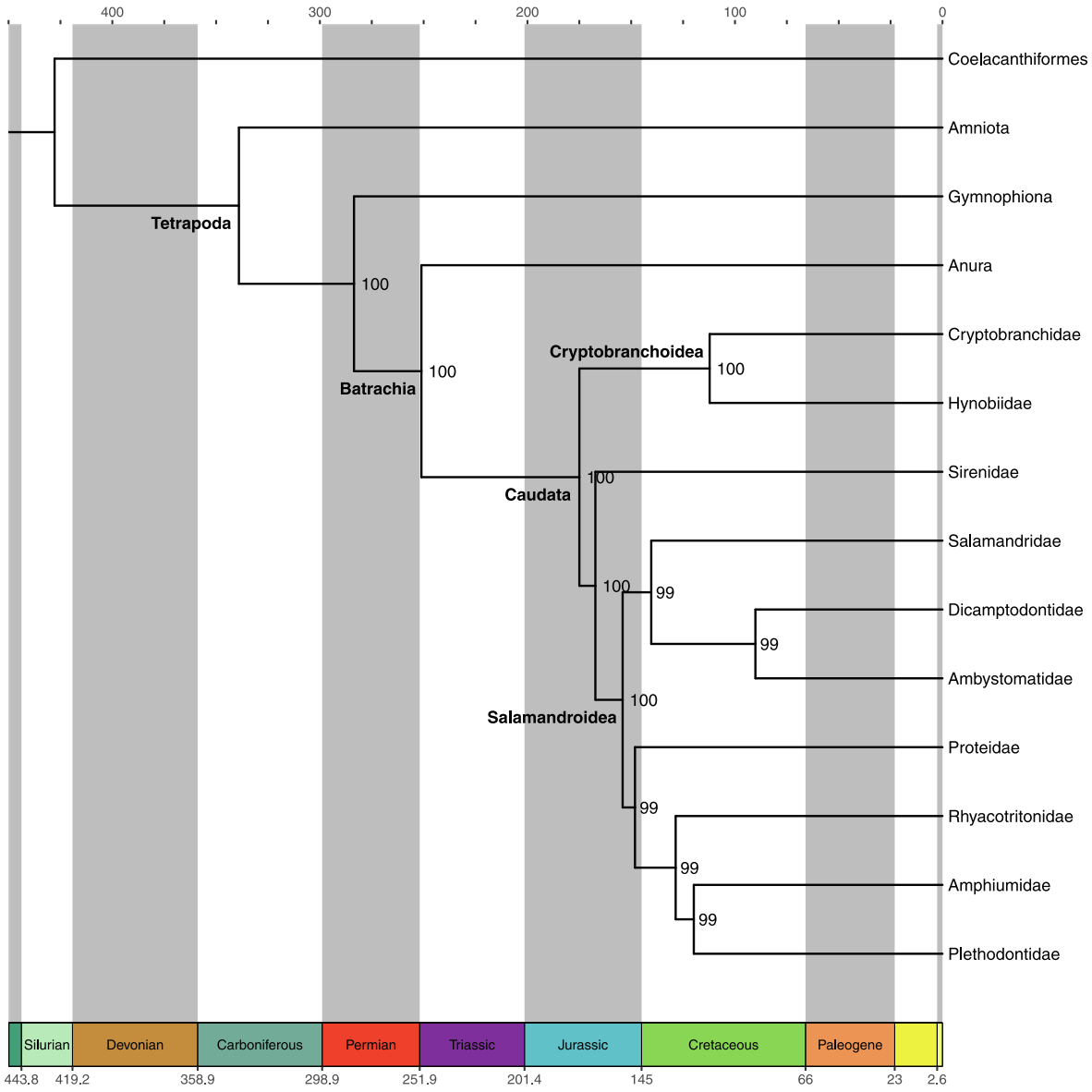

Fig. 1. Higher-level relationships and divergence times among salamander families and outgroups inferred here. Numbers at nodes indicate bootstrap values. Relationships within families are shown in Figs. 2–9. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S1. The full tree is available as Supplementary File S3.

generated from the alternative dating analysis is available as *Supplementary File S6*.

3.3. Relationships within families

In the sections below, we give highlights of the major relationships within some of the larger families (>10 species). We also briefly compare our results to those of Jetz and Pyron (2018) and Hime et al. (2021), and a few other studies. However, for the sake of brevity, we do not verbally describe all species-level relationships, nor do we compare our phylogenetic results to every study published on relationships within each family.

Hynobiidae. Within Hynobiidae (Fig. 2; confidence intervals on estimated dates in Supplementary Fig. S4), all relationships among genera were strongly supported (bs = 100%). *Onychodactylus* (subfamily Onychodactylinae) was placed as the sister taxon to all other genera (subfamily Hynobiinae). The clade of *Paradactylodon* + *Ranodon* was strongly supported as the sister group to the remaining genera, and *Pachyhynobius* was the sister taxon to the remaining genera (*Batrachuperus*, *Hynobius*, *Liua*, *Pseudohynobius*, and *Salamandrella*). The large genus *Hynobius* was the sister taxon to the clade containing *Salamandrella* (*Batrachuperus* (*Liua* + *Pseudohynobius*)).

Jetz and Pyron (2018) found broadly similar relationships in this family. However, they placed *Salamandrella* with *Pachyhynobius* rather than with *Batrachuperus*, *Liua,* and *Pseudohynobius*, as we do here. Zheng et al. (2011) found relationships similar to ours (using nuclear data; their Fig. 1), but placed *Salamandrella* as the sister taxon to the clade of *Pachyhynobius, Hynobius*, and (*Batrachuperus* (*Liua* + *Pseudohynobius*)). Weisrock et al. (2013) also found similar relationships using mitochondrial data alone, but placed *Pachyhynobius* with *Salamandrella* as the sister group to the clade of (*Batrachuperus* (*Liua* + *Pseudohynobius*). The relationships that we found were consistent with those from Hime et al. (2021), who sampled only four genera.

Salamandridae. Many relationships among genera within Salamandridae were strongly supported (Fig. 3; confidence intervals on dates in Supplementary Fig. S5). However, there was weak support (bs = 36%) for a clade placing the subfamilies Salamandrinae and Pleurodelinae as sister taxa, with the genus *Salamandrina* (subfamily

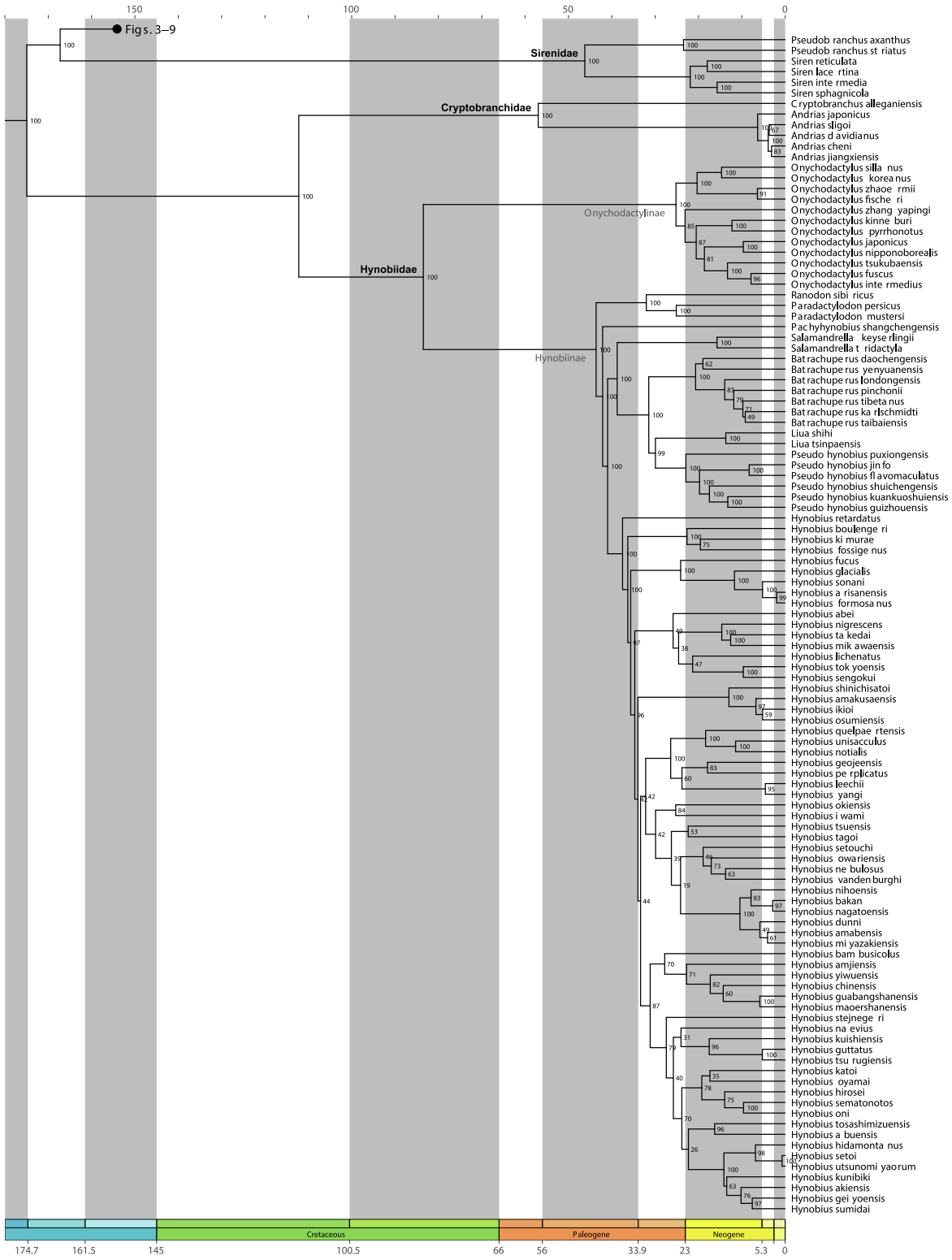

Fig. 2. Relationships and divergence times within Cryptobranchoidea and its sister group. Numbers at nodes indicate bootstrap values, bolded annotations in black indicate families, and annotations in grey indicate subfamilies. Relationships within Salamandroidea (the sister taxon to Sirenidae) are shown in Figs. 3–9. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S4.

Salamandrininae) as the sister group to these two subfamilies. Subfamily Salamandrinae is a strongly supported clade consisting of *Chioglossa* + *Mertensiella* and *Salamandra* + *Lyciasalamandra*. The remaining genera make up the subfamily Pleurodelinae. Within Pleurodelinae the sister group to all other genera was a clade containing *Pleurodeles* + (*Echinotriton* + *Tylototriton*). We found *Echinotriton* nested within *Tylototriton*, but with only moderate support (there is strong support for the reciprocal monophyly of these genera in an extensive molecular study by Dufresnes and Hernandez, 2023). The remaining genera were in a large clade of newts from North America, Europe, and Asia. The clade of

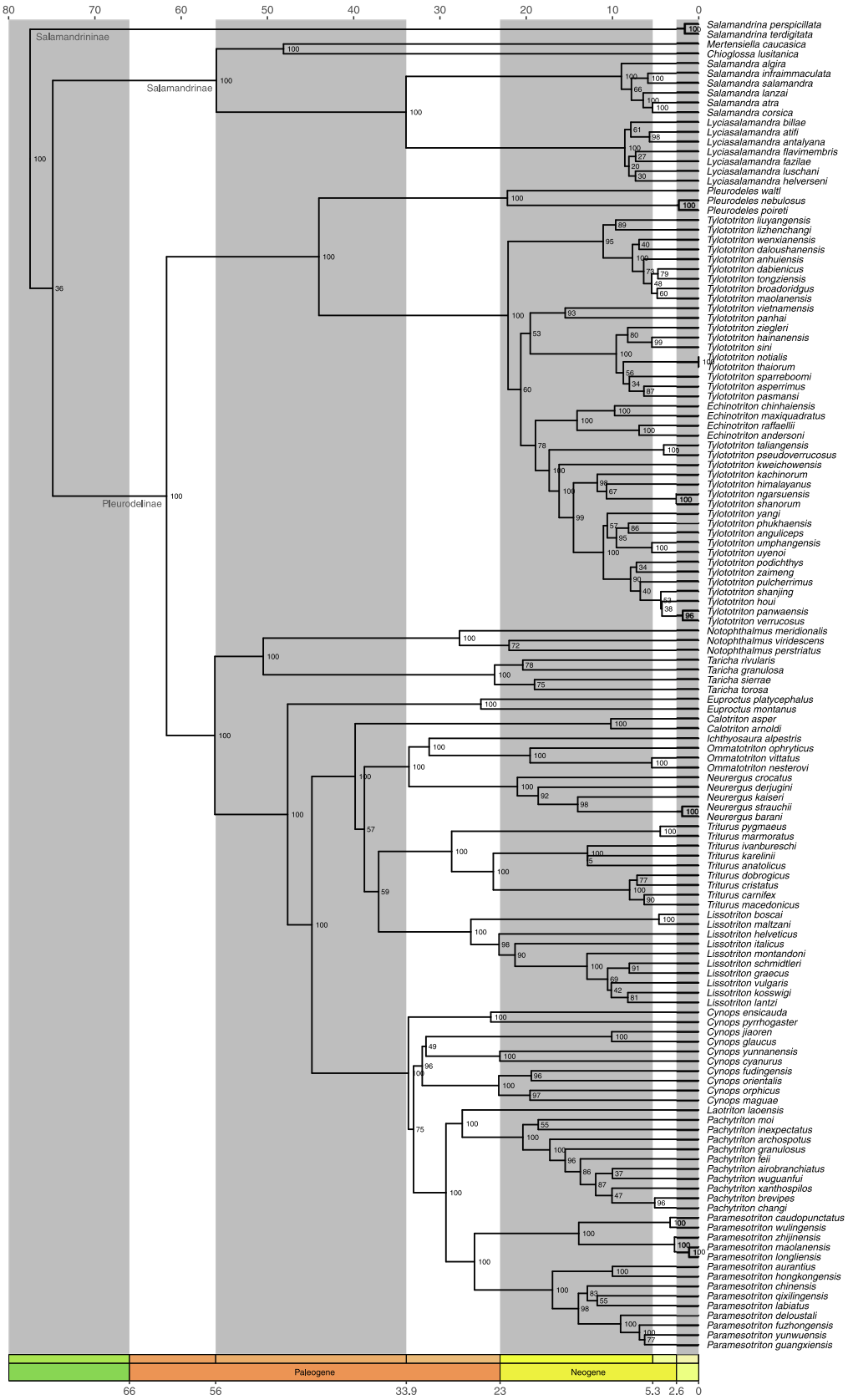

Fig. 3. Relationships and divergence times within Salamandridae. Numbers at nodes indicate bootstrap values and annotations in grey indicate subfamilies. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S5.

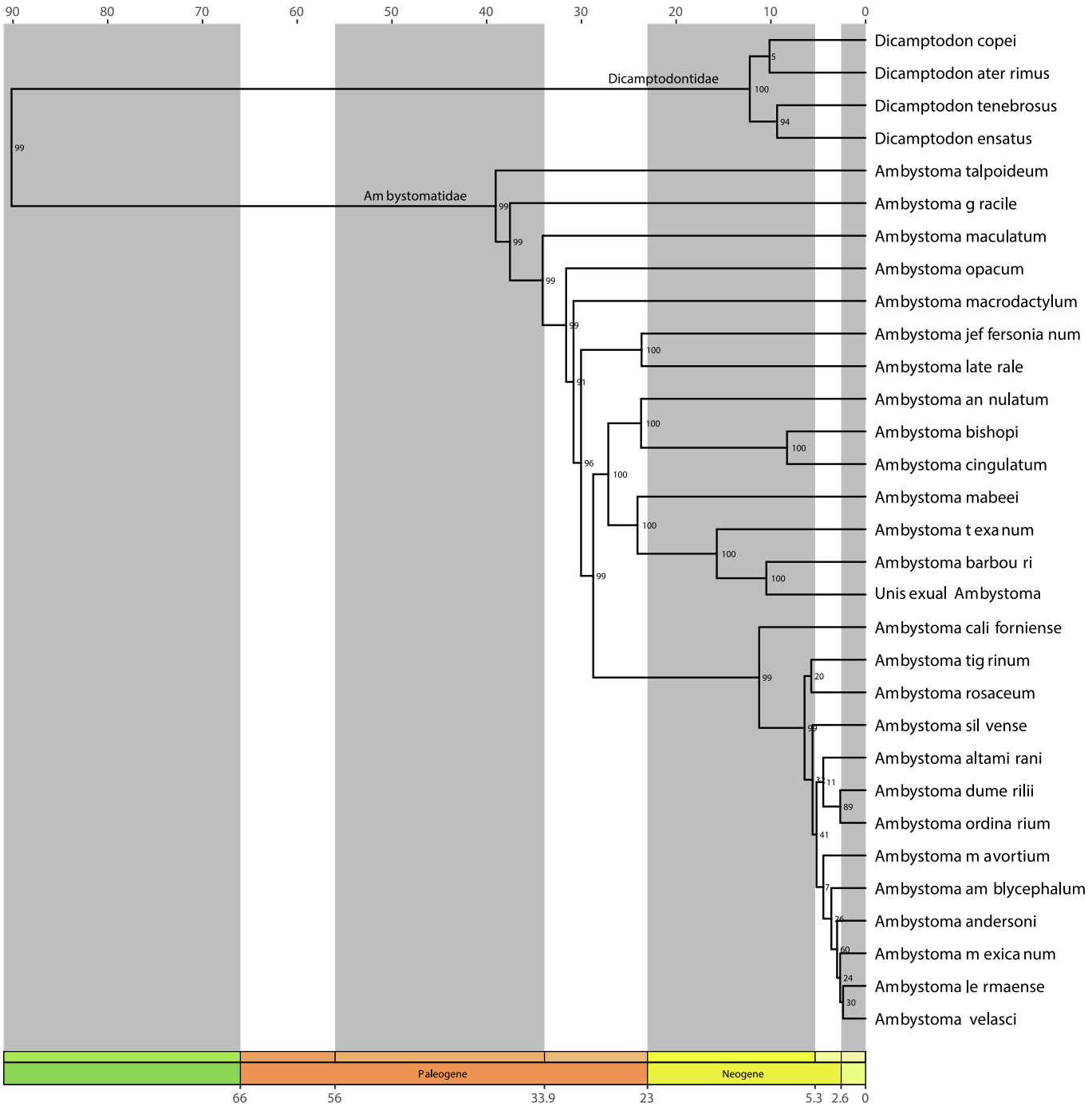

Fig. 4. Relationships and divergence times within Dicamptodontidae and Ambystomatidae. Numbers at nodes indicate bootstrap values and bolded annotations in black indicate families. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S6.

North American newts (*Notophthalmus* + *Taricha*) was the sister group to the remaining genera in this clade, and the European genus *Euproctus* was the sister taxon to the remaining genera. Those genera included a clade of Asian newts, consisting of *Cynops* (*Paramesotriton* (*Laotriton* + *Pachytriton*)). The genus *Cynops* was paraphyletic with respect to the other genera in this clade, with *C. ensicauda* and *C. pyrrhogaster* as the sister group to a moderately supported clade (bs = 75%) containing the other members of *Cynops* and the other genera. These Asian genera were the sister group to a strongly supported clade (bs = 100%) of mostly European and Western Asian newts, with *Calotriton* as the sister taxon to a pair of sister clades, one containing *Neuregus* (*Ichthyosaura* + *Ommatotriton*)) and the other containing *Lissotriton* + *Triturus*. However, not all relationships were strongly supported within this clade of newts. The sister taxon of *Calotriton* was only weakly supported (bs = 61%), as was the clade uniting *Lissotriton* and *Triturus*. The relationships among *Neuregus*, *Ichthyosaura*, and *Ommatotriton* were all strongly supported (bs = 100%).

These generic-level relationships were very similar to those in the tree of Jetz and Pyron (2018). However, in their tree, the sister group to all other salamandrids was Salamandrinae, the clade containing ((*Chioglossa* + *Mertensiella*) + (*Salamandra* + *Lyciasalamandra*)), whereas Salamandrininae (*Salamandrina*) was the sister taxon to Pleurodelinae. These two clades have their positions reversed in our tree. Hime et al. (2021) sampled only nine salamandrid genera and placed

Salamandrininae as the sister taxon to Salamandrinae. Otherwise, the relationships in their tree were identical to ours. Similar to Hime et al. (2021), Rancilhac et al. (2021) placed Salamandrininae as the sister group to Salamandrinae, with the clade *Salamandrina* (*Chioglossa* + (*Salamandra* + *Lyciasalamandra*)) as the sister group to all other salamandrids. Note that Rancilhac et al. (2021) included many additional nuclear loci that were not included here (5455 in total, 81 included here), but we excluded these loci here because data were available for only a limited number of species.

Relationships among genera of European newts were very different between our tree and that of Jetz and Pyron (2018). They found the relationships: (*Ichthyosaura* (*Lissotriton* ((*Ommatotriton* + *Neuregus*) + (*Calotriton* + *Triturus*)). We found: (*Calotriton* ((*Lissotriton* + *Triturus*) + (*Neuregus* (*Ichthyosaura* + *Ommatotriton*)). Importantly, the relationships that we found among these European newt genera were identical to those from phylogenomic data from the study by Rancilhac et al. (2021). This is almost certainly because we incorporated phylogenomic data from that study. We also note that the non-monophyly of *Cynops* found here was also found in that study. The conflicts between our results and those of Jetz and Pyron (2018) seem to arise from conflict between the mitochondrial genes and nuclear genes, since Rancilhac et al. (2021) found similar relationships to those of Jetz and Pyron (2018) when analyzing mitochondrial genes alone, including monophyly of *Cynops*, and the clades *Ommatotriton* + *Neuregus* and *Calotriton* + *Triturus*. Further, Wiens et al. (2011) analyzed only mitochondrial data and found relationships similar to those found by Jetz and Pyron (2018). Rancilhac et al. (2021) suggested that the conflicts between trees from nuclear data and mitochondrial data in European newts were explained by ancient introgression of mitochondrial genes among lineages.

Ambystomatidae. Ambystomatidae presently consists of one genus, *Ambystoma* (AmphibiaWeb, 2024). Our tree strongly resolved relationships among the species found primarily in the U.S. and Canada at the base of the tree, but yielded only weak support for many relationships in the young clade of recently diverged Mexican species (Fig. 4; confidence intervals on dates in Supplementary Fig. S6). The analysis strongly placed *Ambystoma talpoideum* as the sister taxon to the remaining *Ambystoma* species, followed successively by *A. gracile*, *A. maculatum*, *A. opacum*, *A. macrodactylum,* and *A. laterale* + *A. jeffersonianum*. All these relationships were strongly supported (bs $\geq$ 90%). The sister group of the latter species pair was divided into two strongly supported clades, one consisting of northern North American species (including *A. annulatum, A. bishopi, A. cingulatum, A. mabeei, A. texanum,* and *A. barbouri*), and one consisting mostly of Mexican species. Within the northern North American clade, we placed the unisexual *Ambystoma* as the sister group to *A. barbouri* and estimated a divergence time of 10.5 Myr ago between the two. This placement is consistent with previous work, but our divergence estimate is twice as old as previous estimates (Bi and Bogart, 2010). Within the clade of mostly Mexican species, the basal splits separate the more northern species, *A. californiense* and the clade of *A. tigrinum* + *A. rosaceum*, from the remaining species. The remaining species consist mostly of recently diverged species from southern Mexico (along with *A. mavortium*), with weakly supported relationships among them.

The relationships among *Ambystoma* species in the tree of Jetz and Pyron (2018) were mostly unresolved, and were quite different from those supported here. For example, *A. gracile* and *A. mavortium* were sister taxa, and some species from the southern Mexican clade were grouped with northern North American species (e.g., *A. maculatum* + *A. rivulare*, *A. macrodactylum* + *A. silvense*). Based on our strongly supported results, these relationships are most likely incorrect.

Our phylogeny among North American taxa (Fig. 4) broadly resembled that of Williams et al. (2013), from which much of our nuclear data for this family were taken. However, a notable difference is that those authors placed *A. talpoideum* and *A. gracile* as sister taxa at the base of the genus (using species-tree methods), whereas we placed *A. talpoideum* and *A. gracile* as successive sister groups to all other *Ambystoma*. Furthermore, we supported a large North American clade (including *A. annulatum, A. bishopi, A. cingulatum, A. mabeei, A. texanum,* and *A. barbouri*) as the sister group to the clade including *A. tigrinum* and the Mexican species, whereas those authors placed the latter clade inside the former (i.e. Mexican clade inside the northern North American clade). We note that there was also disagreement about relationships among these species in the different species-tree analyses of Williams et al. (2013).

Our phylogeny within *Ambystoma* (Fig. 4) also broadly resembled that of Everson et al. (2021), from which we also obtained nuclear data. However, they included only some of the northern North American species, and there was considerable discordance between their results and ours involving relationships of the recently diverged Mexican species. These relationships were generally only weakly supported in our tree, and in theirs. It should be noted that our tree followed the taxonomic recommendations put forward by Everson et al. (2021), which synonymized many of the Mexican *Ambystoma* species and demoted some to subspecies (which are not represented in our tree).

Plethodontidae. Within Plethodontidae (Fig. 5), the subfamily Plethodontinae was the strongly supported sister group (bs = 100%) to all the remaining genera (currently classified as Hemidactyliinae by AmphibiaWeb, 2024). Plethodontinae was divided into two subclades (Fig. 5; confidence intervals on dates in Supplementary Fig. S7), each with only moderate bootstrap support in our tree. One consisted of *Karsenia* and *Plethodon* (bs = 77%). Within *Plethodon*, we found strong support for traditional relationships among the species groups, with the western *Plethodon* clade (*P. neomexicanus* and relatives) as the sister taxon to the rest of the genus, the *cinereus* group as the sister taxon to other eastern *Plethodon*, and the *wehrlei-welleri* group and *glutinosus* group as sister taxa. *Plethodon jordani* was strongly supported as the sister taxon to the rest of the *glutinosus* group (with *P. metcalfi* strongly supported as the sister taxon to the remaining species), but many other relationships in this group were weakly supported.

The other clade within Plethodontinae (Fig. 5; Supplementary Fig. S7) was more strongly supported (bs = 90%). This clade consisted of one strongly supported subgroup containing *Ensatina* + (*Speleomantes* + *Hydromantes*), and another strongly supported subgroup containing *Aneides* (*Phaeognathus* + *Desmognathus*). Most relationships within these genera were also very strongly supported.

The relationships within Plethodontinae showed several notable differences with those from Jetz and Pyron (2018). For example, those authors placed *Ensatina* with the clade of *Aneides* (*Phaeognathus* + *Desmognathus*), whereas we placed *Ensatina* with *Speleomantes* and *Hydromantes*. They placed *Karsenia* with *Speleomantes* + *Hydromantes,* whereas *Karsenia* was placed with *Plethodon* in our tree. Furthermore, in their tree, the clade of *Karsenia*, *Speleomantes,* and *Hydromantes* is the sister group to all other Plethodontinae.

The tree of Hime et al. (2021) contained only five genera of Plethodontinae. Those authors placed *Karsenia* as the sister group to other plethodontines, rather than with *Plethodon* as we do here. They also placed *Aneides* with *Plethodon*, rather than with *Desmognathus* + *Phaeognathus* as we do here.

Many previous studies have addressed these relationships, but with more limited sampling of taxa and genes. For example, Vieites et al. (2011) found different placements for *Aneides, Ensatina*, and *Karsenia* depending on the data (nuclear vs. mitochondrial) and analysis. Kozak et al. (2009) placed *Plethodon* as the sister taxon to all other Plethodontinae, with strong support, and *Karsenia* with *Hydromantes* + *Speleomantes* with weak support, and *Ensatina* with *Aneides* and the *Phaeognathus* + *Desmognathus* clade with weak support.

Within Hemidactyliinae (Fig. 6; confidence intervals on dates in Supplementary Fig. S8), we found strong support for placing the clade of *Eurycea*, *Pseudotriton* and relatives (tribe Spelerpini of some authors) as the sister group to all other members of the subfamily. Within this clade, all relationships among genera were strongly supported (bs = 100%), including *Stereochilus* (*Gyrinophilus* + *Pseudotriton*) and *Urspelerpes* +

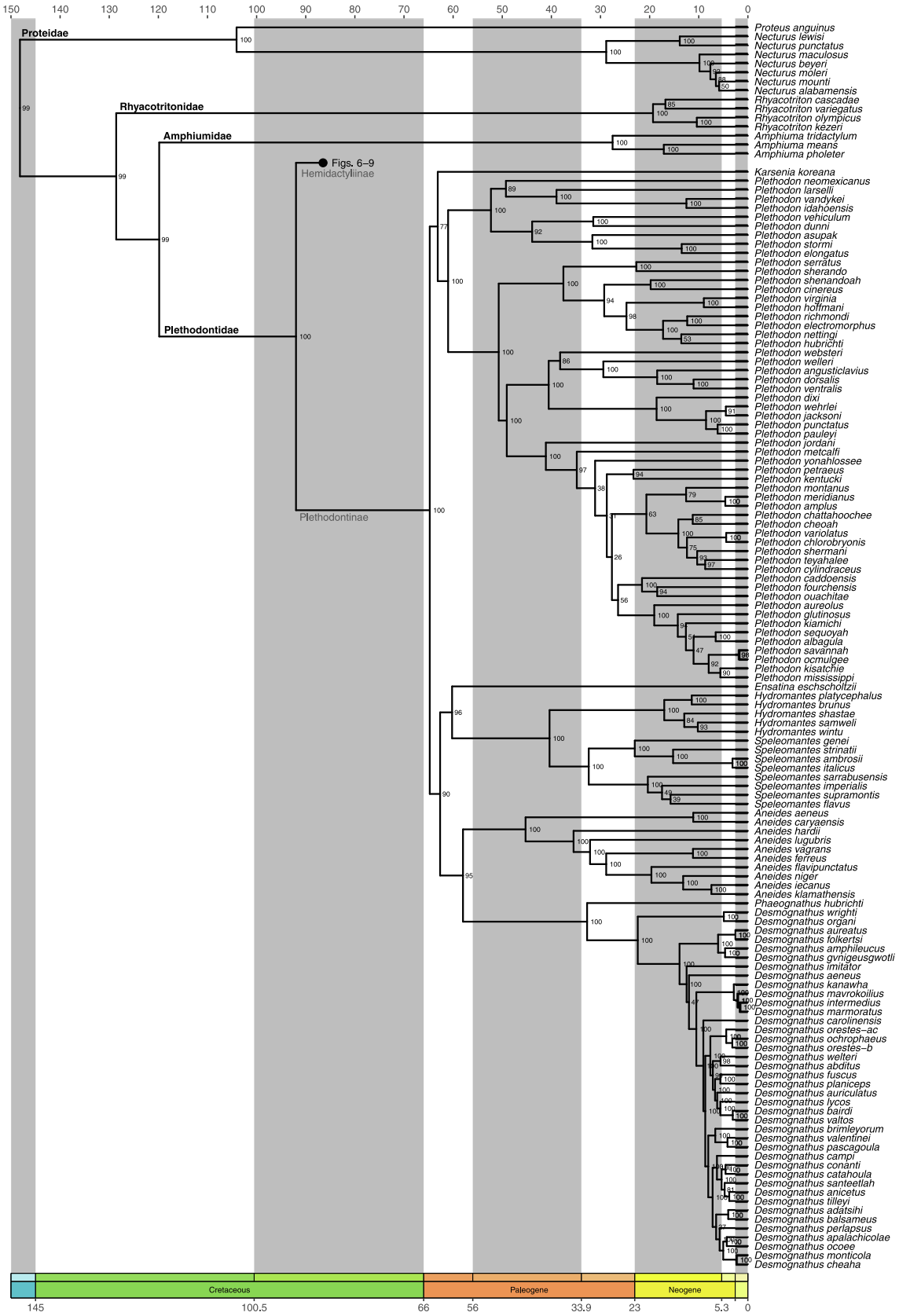

Fig. 5. Relationships and divergence times within Proteidae, Rhyacotritonidae, Amphiumidae, and the subfamily Plethodontinae (Plethodontidae). Numbers at nodes indicate bootstrap values, bolded annotations in black indicate families, and annotations in gray indicate subfamilies. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S7.

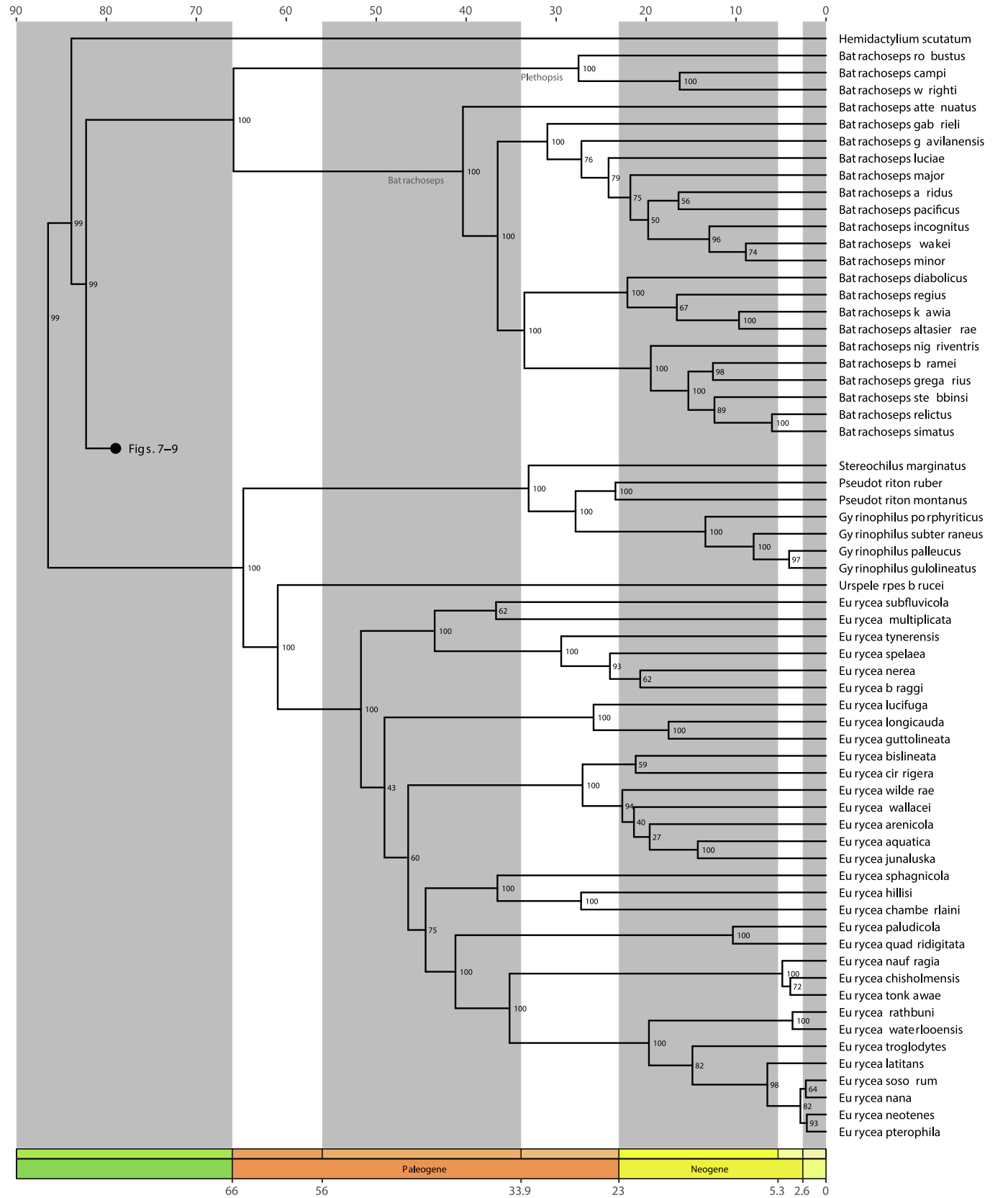

Fig. 6. Relationships and divergence times among the oldest clades within the subfamily Hemidactyliinae of the family Plethodontidae, especially the tribe Spelerpini and the genera *Hemidactylium* and *Batrachoseps*. Numbers at nodes indicate bootstrap values and annotations in gray indicate subgenera. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S8.

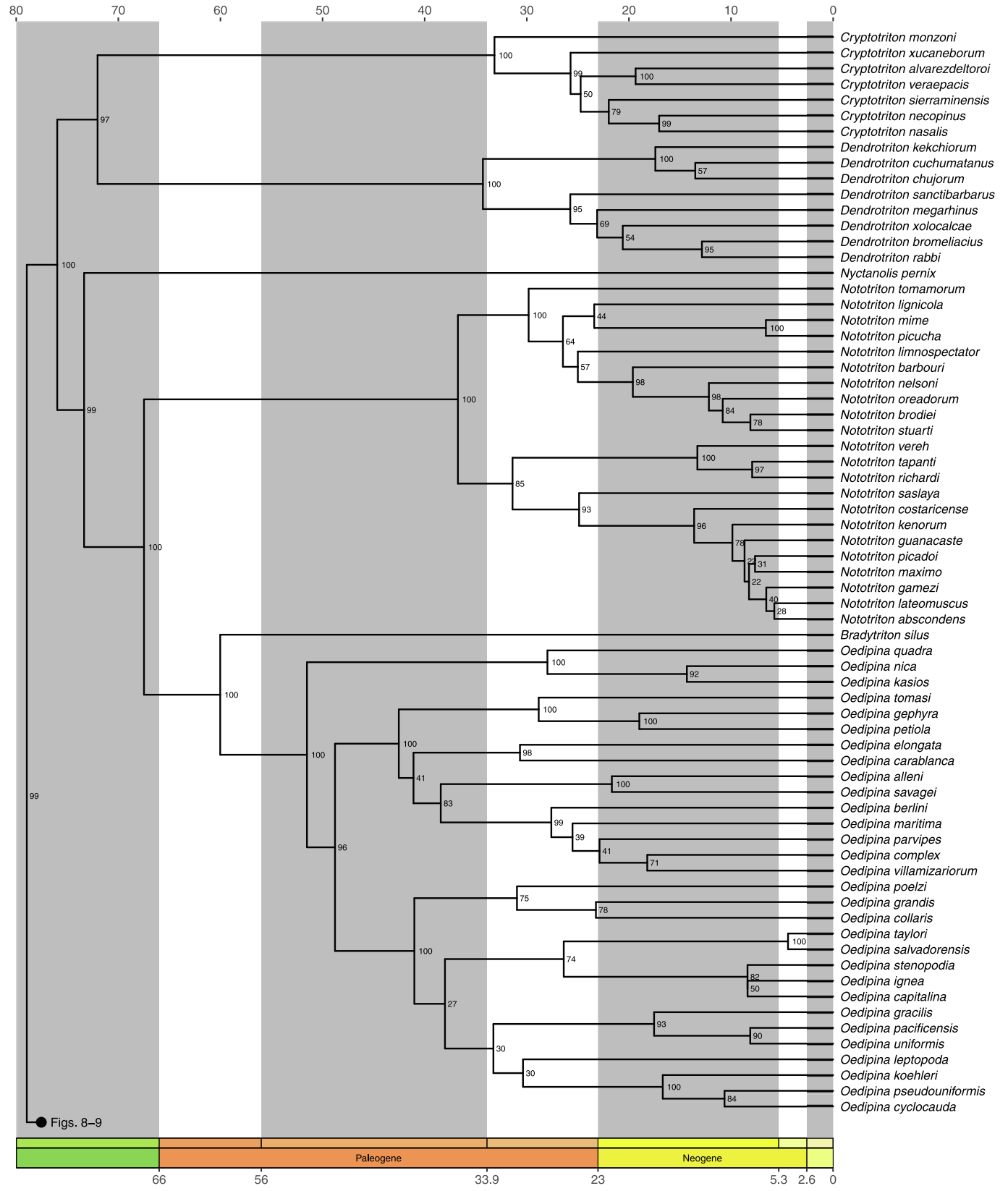

Fig. 7. Relationships and divergence times among the deepest clades of tropical bolitoglossine salamanders (Hemidactyliinae: Plethodontidae). Numbers at nodes indicate bootstrap values. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S9.

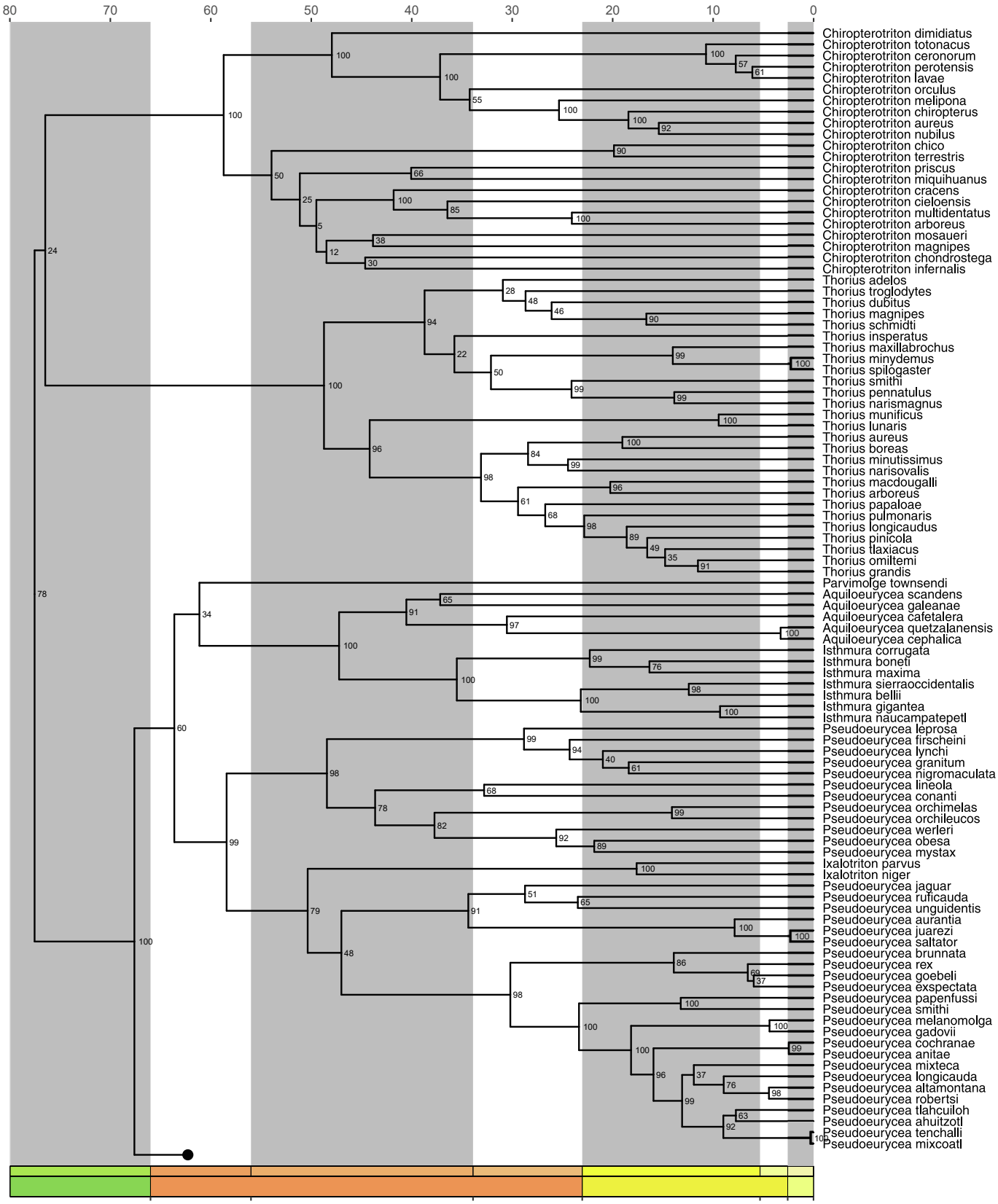

Fig. 8. Relationships and divergence times among additional tropical bolitoglossine salamanders (Hemidactyliinae: Plethodontidae). Numbers at nodes indicate bootstrap values. Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S10.

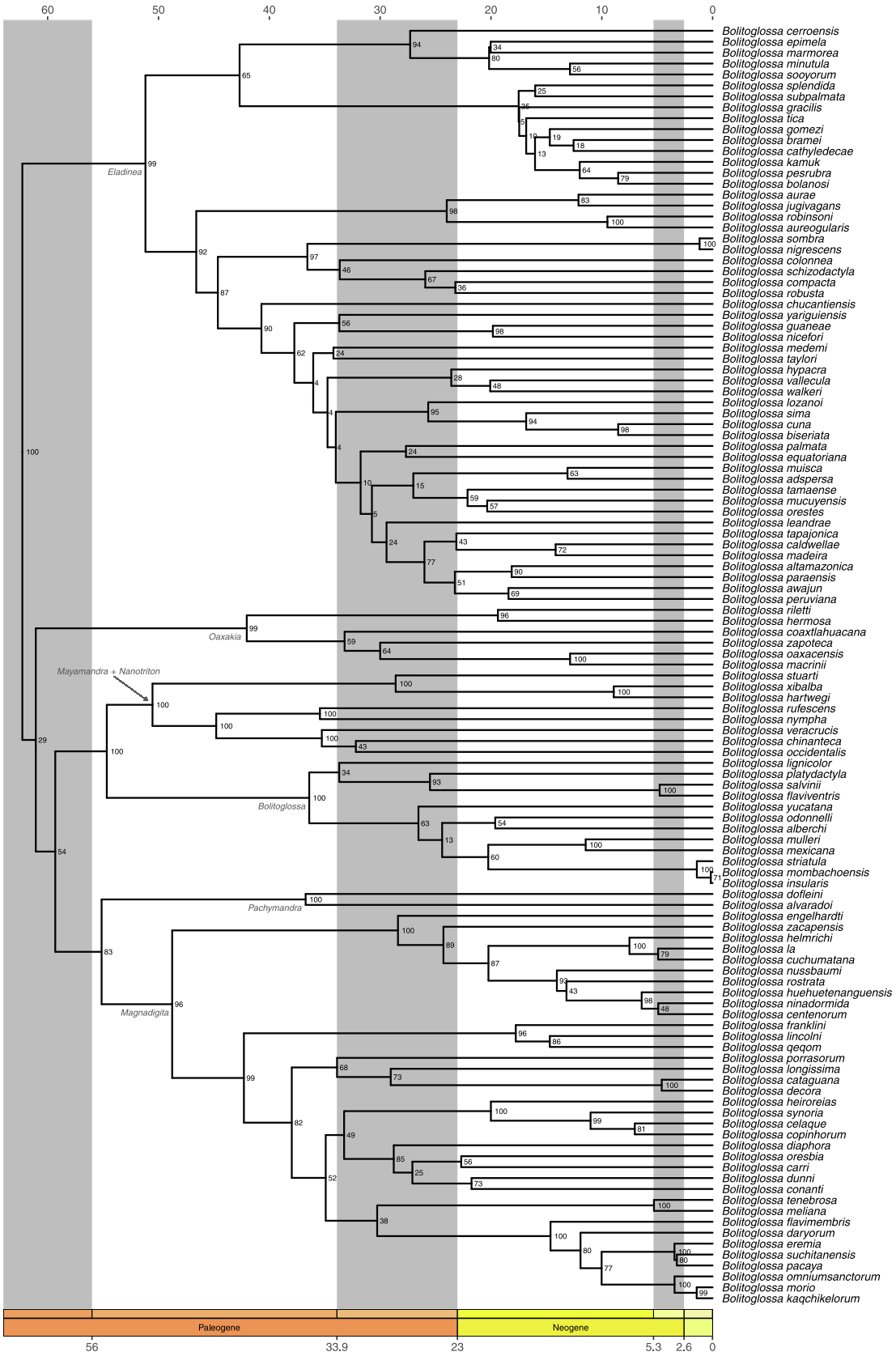

Fig. 9. Relationships and divergence times in the tropical bolitoglossine genus Bolitoglossa (Hemidactyliinae: Plethodontidae). Numbers at nodes indicate bootstrap values and annotations in grey indicate subgenera. All subgenera are monophyletic except for Mayamandra and Nanotriton (combined into a single clade in our tree). Confidence intervals on estimated ages for these nodes are shown in Supplementary Fig. S11.

Eurycea. The monotypic genus *Hemidactylium* was strongly supported (bs = 99%) as the sister group to a well-supported clade (bs = 99%) containing *Batrachoseps* and the tropical bolitoglossines (Fig. 6). Within *Batrachoseps* (Fig. 6), we found relationships that were generally strongly supported and very similar to those of Jockusch et al. (2015), with the genus divided into two subgenera (*Plethopsis* and *Batrachoseps*), and the subgenus *Batrachoseps* divided into three species groups: *pacificus* + (*nigriventris* + *diabolicus*).

We briefly compare these higher-level hemidactyliine relationships to those from other studies. In the tree of Jetz and Pyron (2018), the relationships of *Hemidactylium* and the Spelerpini were reversed relative to our tree, with *Hemidactylium* as the sister group to all other Hemidactyliinae (as in Kozak et al., 2009). The relationships for these taxa found here were consistent with those of Hime et al. (2021) and earlier studies by Vieites et al. (2011; nuclear data trees) and Shen et al. (2016). Overall, our strongly supported placement for the controversial *Hemidactylium* was consistent with previous studies of multiple nuclear markers, and conflicted with those studies that sampled a higher proportion of mitochondrial genes.

Among the tropical bolitoglossines (Fig. 7; confidence intervals on dates in Supplementary Fig. S9), the sister group to all other species was a strongly supported clade containing the genera *Dendrotriton* + *Cryptotriton* and the genera *Nyctanolis* (*Nototriton* (*Bradytriton* + *Oedipina*)). All of these relationships were strongly supported. The sister group to that clade was a moderately supported one (bs = 78%) containing the remaining genera (Fig. 8; confidence intervals on dates in Supplementary Fig. S10). Within the latter clade was a smaller, weakly supported clade (bs = 24%) containing the genera *Chiropterotriton* + *Thorius* (Fig. 8).

The clade of *Chiropterotriton* + *Thorius* was the sister taxon to a large, strongly supported clade containing the very large genus *Bolitoglossa* and the genera *Pseudoeurycea*, *Ixalotriton*, *Parvimolge*, *Isthmura*, and *Aquiloeurycea* (Fig. 8; Supplementary Fig. S10). The latter five genera formed a weakly supported clade (bs = 60%). Within this clade was one strongly supported subgroup (bs = 99%) containing the genera *Pseudoeurycea* and *Ixalotriton,* with *Ixalotriton* nested inside of *Pseudoeurycea*. The sister taxon to that subgroup was a weakly supported clade (bs = 34%) containing the genera *Parvimolge*, *Isthmura*, and *Aquiloeurycea.* The genera *Isthmura* and *Aquiloeurycea* were strongly supported as sister taxa in our tree.

Within *Bolitoglossa* (117 species sampled) we found a mixture of strongly supported and weakly supported relationships at all levels within the genus (Fig. 9; confidence intervals on dates in Supplementary Fig. S11). The higher-level relationships within the genus were discordant with those estimated by Parra-Olea et al. (2004), who proposed the subgeneric taxonomy that we follow here. Parra-Olea et al. (2004), in their maximum-likelihood tree (their Fig. 4), found weak support (bs < 50%) for the relationships: *Oaxakia* (*Pachymandra* (*Magnadigita* (*Eladinea* (*Bolitoglossa* (*Mayamandra, Nanotriton*))))).

Here we placed the large subgenus *Eladinea* (53 out of 70 species sampled) as the sister taxon to all other *Bolitoglossa*. Monophyly of *Eladinea* was strongly supported, but the sister group to *Eladinea* was a weakly supported clade (bs = 29%). Within this latter clade, the subgenus *Oaxakia* (6 of 6 species sampled) was strongly supported as monophyletic (bs = 99%) and was placed as the sister group to the remaining species, which formed a weakly supported clade (bs = 54%).

Within this latter clade, there was a strongly supported clade (bs = 100%) that united the subgenera *Bolitoglossa*, *Mayamandra*, and *Nanotriton*. The latter two subgenera formed a strongly supported clade (bs = 100%). However, *Mayamandra* (3 out of 4 species sampled) was not supported as monophyletic because *B. veracrucis* was placed within the subgenus *Nanotriton*. Because of this, the subgenus *Nanotriton* (4 out of 4 species sampled) was not monophyletic either. The subgenus *Bolitoglossa* was strongly supported as monophyletic (12 of 13 species sampled; all but *B. jacksoni*). Finally, the subgenera *Magnadigita* and *Pachymandra* were moderately supported (bs = 83%) as sister taxa. *Magnadigita* (36 out of 38 species sampled) was strongly supported as monophyletic, as was *Pachymandra* (2 out of 3 species sampled).

In summary, we found the relationships: *Eladinea* (*Oaxakia* ((*Bolitoglossa* (*Mayamandra*, *Nanotriton*)) + (*Magnadigita*, *Pachymandra*))). By contrast, Parra-Olea et al. (2004) found the relationships: *Oaxakia* (*Pachymandra* (*Magnadigita* (*Eladinea* (*Bolitoglossa* (*Mayamandra, Nanotriton*))))). The only relationships shared are *Bolitoglossa* (*Mayamandra, Nanotriton*). However, many conflicting relationships among subgenera were weakly supported in both trees.

We briefly compare these relationships among bolitoglossine genera to those estimated by Jetz and Pyron (2018) and other authors. Many of the relationships that we found among tropical bolitoglossine genera conflicted somewhat with those found by Jetz and Pyron (2018). For example, we found a strongly supported clade of six genera (Fig. 7) as the sister group to other bolitoglossines: (*Dendrotriton* + *Cryptotriton*) (*Nyctanolis* (*Nototriton* (*Bradytriton* + *Oedipina*))), whereas Jetz and Pyron (2018) found the sister group to other tropical bolitoglossines to be a clade containing (*Thorius* (*Cryptotriton*, *Chiropterotriton*)), ((*Dendrotriton* + *Nyctanolis*), (*Nototriton* (*Bradytriton* + *Oedipina*))). Thus, these clades are similar, but in their tree this clade contains the genera *Thorius* and *Chiropterotriton*.

The other major clade found by Jetz and Pyron (2018) in tropical bolitoglossines contained the genera *Bolitoglossa*, *Pseudoeurycea*, *Ixalotriton*, and *Parvimolge*. This was largely in agreement with our results (although we recognized *Isthmura* and *Aquiloeurycea* within *Pseudoeurycea*, following Rovito et al., 2015). Our results also agree with theirs in placing the clade of *Pseudoeurycea (sensu lato), Ixalotriton*, and *Parvimolge* as the sister taxon of *Bolitoglossa*, but they disagree about relationships within that clade. Specifically, those authors placed *Ixalotriton* and *Parvimolge* as sister taxa, whereas we did not. Jetz and Pyron (2018) had many relationships unresolved within the genera *Bolitoglossa*, *Oedipina*, and *Pseudoeurycea*. Our trees within these genera are fully resolved, but with variable levels of branch support.

We also compare our results to a study that focused on tropical bolitoglossine salamanders (Rovito et al., 2015), using both nuclear and mitochondrial data (but with somewhat limited taxon sampling, 57 species). Their trees were generally more similar to ours than to the tree of Jetz and Pyron (2018) for these genera. Rovito et al. (2015) supported a clade that was the sister group to other tropical bolitoglossines that included *Dendrotriton, Cryptotriton, Nyctanolis, Nototriton, Bradytriton,* and *Oedipina*. This is the same clade of six genera found in our study, although their estimated relationships within this clade were not identical to ours. For the remaining genera, they found the relationships: (*Thorius* (*Chiropterotriton* (all other genera))), whereas we found: ((*Thorius* + *Chiropterotriton*) (all other genera)), but in our tree the *Thorius* + *Chiropterotriton* clade was only weakly supported. All three studies agree on a large clade containing the genera *Bolitoglossa*, *Pseudoeurycea* (*sensu lato*), *Ixalotriton*, and *Parvimolge*. However, there is disagreement about relationships within this clade. Notably, the different analyses of Rovito et al. (2015) did not agree on these relationships either (i.e. their species-tree analyses with and without mitochondrial data, and their concatenated analyses). We note that our analysis placed *Ixalotriton* within *Pseudoeurycea* (sensu stricto), but this was not found in the analyses of Rovito et al. (2015) or Jetz and Pyron (2018).

3.4. Potential biases in branch lengths and divergence times

Species with only mitochondrial data had significantly longer maximum-likelihood branch lengths (mean = 0.021, n = 263) than species with at least one nuclear marker (mean = 0.018, n = 502; p < 0.001, Mann-Whitney U test). However, when time-calibrated branch lengths were examined instead, there was no significant difference in the distribution of branch lengths (p = 0.91, Mann-Whitney U test), and species with only mitochondrial markers actually had shorter mean branch lengths (mean = 14.63, n = 263) than species with at least one

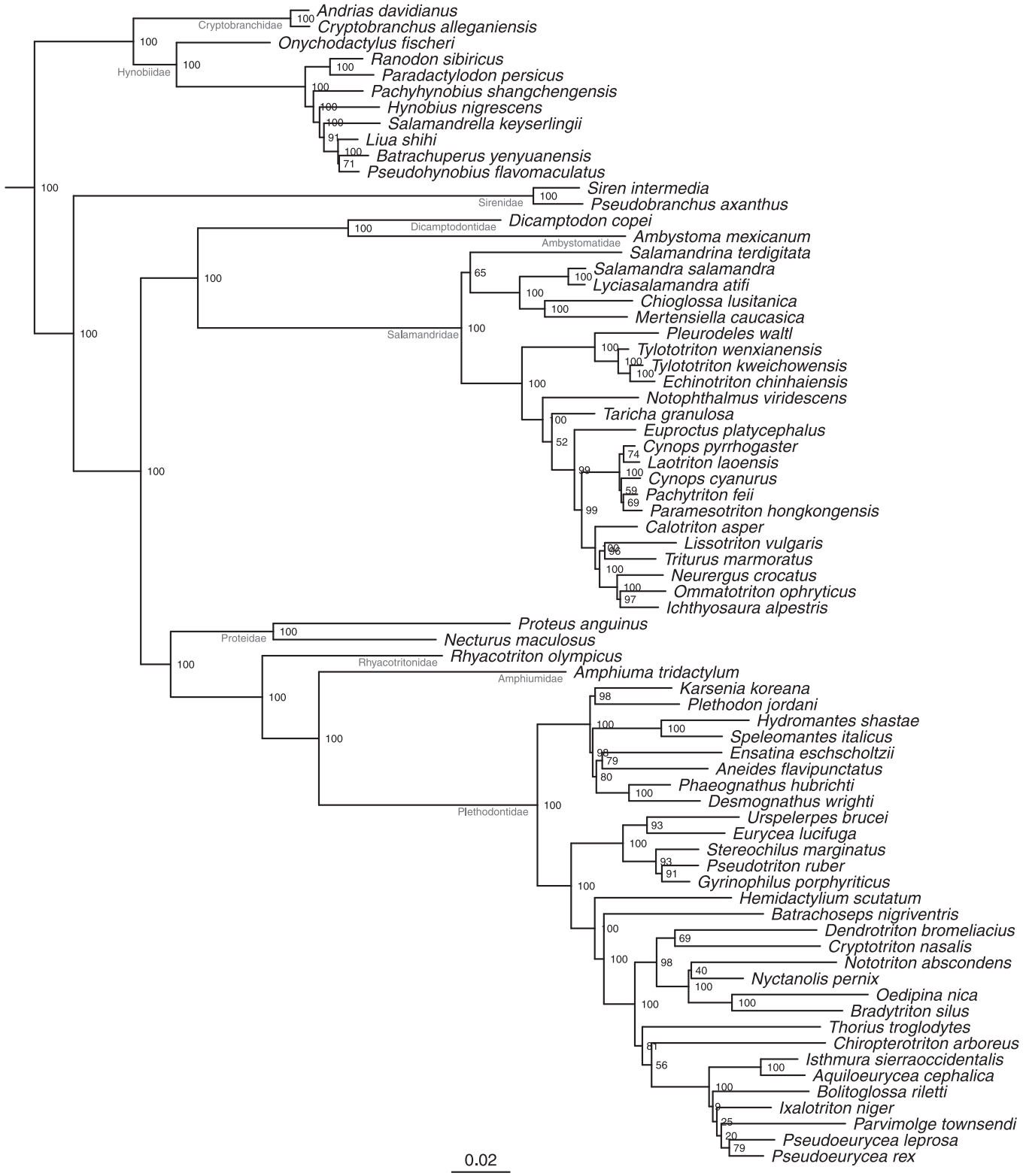

Fig. 10. Relationships among salamander genera inferred from nuclear markers only. Numbers at nodes indicate bootstrap values and annotations in grey indicate families. The scale represents substitutions per site.

nuclear sequence (mean = 15.56; $n = 502$).

There were significant positive correlations between species maximum-likelihood branch lengths and both their proportion of missing data ($n = 765$, Kendall's $\tau = 0.20$, $p < 0.001$) and their proportion of mitochondrial markers ($n = 765$, Kendall's $\tau = 0.19$, $p < 0.001$). However, when time-calibrated branch lengths were examined instead, these correlations were extremely weak. There was a barely significant positive correlation between branch lengths and missing data ($n = 765$, Kendall's $\tau = 0.05$, $p = 0.048$), and a non-significant, positive correlation between branch lengths and the proportion of mitochondrial markers ($n = 765$, Kendall's $\tau = 0.04$, $p = 0.150$). Therefore, we infer that the inclusion of taxa with considerable missing data and

predominantly mitochondrial markers had very limited impact on the estimated divergence times. All statistical tests are available on the GitHub repository associated with the project (https://github.com/aast242/salamander-timetree), and all data utilized are available in Supplementary Table S8.

3.5. Nuclear-only phylogeny

We generated a tree based only on nuclear markers to assess the impact of the mitochondrial data on the generic-level relationships. The family-level relationships were identical between the main and nuclear-only trees. Therefore, we focus on relationships among genera within the largest families (i.e., those with multiple genera). This tree is presented in Fig. 10 and is available as Supplementary File S7.

Hynobiidae: Within Hynobiidae, relationships were generally identical between the main tree (including mitochondrial data) and the nuclear-only tree (Fig. 10). However, within the clade of *Batrachuperus*, *Liuia*, and *Pseudohynobius*, the main tree strongly supported the clade *Liuia* and *Pseudohynobius* (bs = 99%), whereas the nuclear tree more weakly supported (bs = 71%) the clade *Batrachuperus* + *Pseudohynobius*.

Salamandridae: Within Salamandridae, the main tree shows weak support for a clade uniting all other genera exclusive of *Salamandrina* (bs = 36%), whereas the nuclear-only tree (Fig. 10) showed weak support (bs = 65%) for placing *Salamandrina* with the Salamandrinae. The two trees agreed on the relationships within Salamandrinae: (*Salamandra* + *Lyciasalamandra*) + (*Mertensiella* + *Chioglossa*).

Within the salamandrid subfamily Pleurodelinae, both trees agreed that the clade of *Pleurodeles*, *Echinotriton*, and *Tylototriton*, is the sister taxon to the remaining genera, that *Pleurodeles* is the sister taxon to *Echinotriton* and *Tylototriton*, and that *Tylototriton* is paraphyletic with respect to *Echinotriton*. The main tree shows strong support for placing *Notophthalmus* with *Taricha* as sister taxa (bs = 100%), together forming the sister taxon to the remaining genera (bs = 100%), whereas the nuclear-only tree shows weak support (bs = 52%) for placing *Taricha* as the sister taxon to the other newts. Both trees agree that *Euproctus* is the sister taxon to the remaining newts, and that these remaining newts are divided into two strongly supported clades: an Asian clade (including *Cynops* and relatives) and a primarily European clade (including *Triturus* and relatives). Within the Asian clade, the main tree has *Cynops* as paraphyletic with respect to a strongly supported clade (bs = 100%) consisting of *Paramesotriton*, *Laotriton*, and *Pachytriton*, with *Laotriton* and *Pachytriton* weakly supported as sister taxa (bs = 75%). In the nuclear-only tree, *Cynops* is also paraphyletic, but *Laotriton* is weakly placed with *Cynops pyrrhogaster* (bs = 74%) and *Cynops cyanurus* is weakly placed (bs = 59%) with the weakly supported clade of *Pachytriton* + *Paramesotriton* (bs = 69%). In the primarily European clade of newts, both the main tree and nuclear-only tree agree that *Calotriton* is the sister to the remaining genera, which consist of a clade of *Lissotriton* + *Triturus* and a clade of *Neurergus* (*Ichthyosaura* + *Ommatotriton*). All these relationships are strongly supported in the nuclear-only tree, but the sister group to *Calotriton* and, within it, the *Lissotriton* + *Triturus* clade are each only weakly supported in the main tree (bs = 57–59%).

Plethodontidae: Within Plethodontidae, the broad-scale relationships are similar between the main and nuclear-only trees (Fig. 10), but there are some differences. Both show strong support for monophyly of the subfamilies Plethodontinae and Hemidactyliinae. Within Plethodontinae, both trees agree that *Karsenia* and *Plethodon* are sister taxa and form the sister group to the other plethodontine genera, although the nuclear-only tree shows stronger support for these relationships. Among the other plethodontine genera, the main tree shows strong support for placing *Ensatina* with *Hydromantes* + *Speleomantes* (bs = 96%), and *Aneides* with *Desmognathus* + *Phaeognathus* (bs = 100%). By contrast, the nuclear-only tree shows relatively weak support (bs = 79%) for placing *Aneides* with *Ensatina*. Both trees show strong support for the clades *Hydromantes* + *Speleomantes* and *Desmognathus* + *Phaeognathus*.

Within Hemidactyliinae, both the main tree and nuclear-only trees agree that the Spelerpini (*Eurycea*, *Pseudotriton*, and relatives) is the sister taxon to the other genera, that *Hemidactylium* is the sister taxon to the hemidactyliine genera outside Spelerpini, and that *Batrachoseps* is the sister taxon to the tropical bolitoglossines. These relationships are strongly supported in both trees (bs = 99–100%). Furthermore, relationships are fully congruent within Spelerpini, with the relationships: (*Eurycea* + *Urspelerpes*) + (*Stereochilus* (*Gyrinophilus* + *Pseudotriton*)). These relationships are strongly supported by both datasets (bs = 91–100% nuclear only; bs = 100% main tree).

Among tropical bolitoglossines, there is considerably more disagreement, but many of these disagreements are only weakly supported. First, both trees agree that there is a strongly supported clade consisting of the genera *Cryptotriton, Dendrotriton, Nyctanolis, Nototriton, Bradytriton*, and *Oedipina*. They also agree that there is a clade consisting of *Cryptotriton* and *Dendrotriton*, and that this clade is the sister group to the other genera. They also agree that *Bradytriton* and *Oedipina* are sister taxa (with strong support). They disagree about the placement of *Nyctanolis*, which is strongly supported as the sister taxon to the other four genera in the main tree, and weakly supported as the sister taxon to *Nototriton* by the nuclear-only data (bs = 40%).

The two trees also agree that the remaining genera form a clade. In the main tree, *Chiropterotriton* and *Thorius* are weakly supported as sister taxa (bs = 24%), whereas in the nuclear-only tree, *Chiropterotriton* is weakly placed as the sister taxon to the remaining bolitoglossine genera (including *Bolitoglossa* and *Pseudoeurycea*).

These remaining bolitoglossine genera form a strongly supported clade in both the main tree and the nuclear-only tree (bs = 100%). Within this clade, most relationships are only weakly supported in both trees. Both trees agree that *Aquiloeurycea* and *Isthmura* are strongly supported as sister taxa (bs = 100%). In the main tree, *Parvimolge* is weakly placed with these two genera, but weakly placed with *Pseudoeurycea* in the nuclear-only tree. *Ixalotriton* is weakly placed inside *Pseudoeurycea* in the main tree, and weakly supported as the sister taxon to *Pseudoeurycea* + *Parvimolge* in the nuclear-only tree.

In summary, the main tree and the nuclear-only tree are broadly congruent, and agree for 80% of the 70 comparable nodes. There are some disagreements about relationships among genera within families (in Hynobiidae, Salamandridae, and Plethodontidae). However, these conflicts were not strongly supported in the nuclear-only tree. For all 14 of the conflicting nodes, the nuclear-only tree showed only weak support for these conflicting relationships (bs = 9–80%), whereas the main tree showed either strong support (bs > 94%, 8 out of 14 nodes) or weak support (bs = 24–79%, 6 nodes). Such a pattern is inconsistent with the idea that the nuclear-only tree reflects the true relationships whereas the main tree does not.

4. Discussion

In this study, we combined phylogenomic and supermatrix approaches to provide a new estimate of salamander relationships. We sampled 765 species here, including 284 species (with sequence data) not sampled in the largest previous supermatrix study (59% increase). We also included more than twice the number of fossil calibration points within salamanders (13 vs. 6). Our results showed strong support for most relationships both among and within families, but with some remaining areas of uncertainty among and within genera. The divergence dates inferred here (Table 1) were largely intermediate between the relatively ancient dates estimated by Jetz and Pyron (2018) and the much younger ones from Hime et al. (2021).

We included a large number of taxa and markers, but the data matrix was dominated by missing data cells (92%). There were extensive missing data because the phylogenomic dataset included many genes, but relatively few species had data for all these genes. Instead, many species were included in the combined matrix based on a limited number of mitochondrial markers (or mitochondrial and nuclear markers). Thus, most species lacked data for hundreds of nuclear markers, leading

to >90% missing data for most species. The potential impact of missing data on model-based phylogenetics has been a subject of considerable debate (Wiens, 2003; Lemmon et al., 2009; Sanderson et al., 2010; Wiens and Morrill, 2011; Roure et al., 2013; Jiang et al., 2014; Hosner et al., 2016; Xi et al., 2016; Talavera et al., 2022). Yet, we found few obvious negative impacts of missing data here. For example, almost all genera were inferred to be monophyletic. The three exceptions were found to be non-monophyletic in previous analyses or in analyses of single genes. For example, we found that the genus *Pseudoeurycea* had the genus *Ixalotriton* placed inside it. However, previous studies have also suggested that *Pseudoeurycea* is not monophyletic (Rovito et al., 2015), including studies placing *Ixalotriton* within *Pseudoeurycea* (Wiens et al., 2007). Therefore, this result does not appear to be an artifact of missing data. Similarly, we found the genus *Cynops* to be non-monophyletic, but this was supported in previous phylogenomic analyses (Rancilhac et al., 2021). Finally, we also found *Echinotriton* to be nested inside of *Tylototriton*. This pattern appeared in many individual gene trees from the dataset of Rancilhac et al. (2021). However, that study (and others) strongly suggest that these two genera are more likely to be monophyletic (e.g., Dufresnes and Hernandez, 2023).

Furthermore, we found strong bootstrap support for most nodes throughout the tree. Thus, most of the taxa with extensive missing data seemed to be placed in the expected clades with strong support. These results are consistent with other empirical studies (Wiens et al., 2005; Cho et al., 2011; Zheng and Wiens, 2016; Talavera et al., 2022; Portik et al., 2023b) showing that extensive missing data are not necessarily problematic for phylogenetic inference. They are also consistent with many previous simulation studies (Wiens, 2003; Gouveia-Oliveira et al., 2007; Wiens and Morrill, 2011; Talavera et al., 2022), which suggest that highly incomplete taxa can be placed accurately in phylogenies, as long as they have adequate non-missing data. The tradition of sampling the same fast-evolving mitochondrial markers across many salamander groups (e.g., cytochrome *b*) may have been particularly helpful in resolving relationships at the species level. Thus, given that at least one gene was sampled that allowed the correct placement of a given species, the missing data for hundreds of other genes did not seem to have a strong negative impact on the analyses.

An obvious question arising from our study is: why do the divergence dates differ so much among recent studies of salamander phylogeny? A second, related question is: given this uncertainty, which estimate should be considered more reliable? Subsampling analyses suggest that the number of fossil calibration points is a crucial variable for estimating divergence dates, more so than the number of markers or the amount of missing data (Zheng and Wiens, 2015). Given this, it is important to note that Hime et al. (2021) included only one fossil calibration point within salamanders. This might have caused their divergence dates within salamanders to be insufficiently constrained and underestimated. Jetz and Pyron (2018) included six fossil calibration points within salamanders and estimated much older dates for comparable clades (Table 1). Here, we used a total of 13 fossil calibration points within salamanders. Our divergence-date estimates for major clades here were generally intermediate between those of Jetz and Pyron (2018) and Hime et al. (2021). Specifically, they were intermediate for 73% (8/11) of the clades in Table 1, younger than both for one clade (Cryptobranchoidea), and older than both for two others (Plethodontidae, Salamandridae). Overall, our larger sampling of fossil calibration points should yield more accurate divergence-date estimates. However, other factors could also be relevant. For example, all three studies used somewhat different data and methods to estimate both the phylogeny and the divergence dates. Along these lines, the dominance of fast-evolving mitochondrial markers in the supermatrix datasets might lead to older estimates, whereas the exclusive use of slow-evolving nuclear markers in the phylogenomic datasets might lead to younger estimates. However, our comparison of estimated branch lengths among species showed that those species with a higher proportion of mitochondrial markers did not have significantly older divergence times.

We appreciate that some readers may be concerned that many species ($n = 263$) are included based only on mitochondrial data, and that mitochondrial data might have had a negative impact on the estimated phylogeny and divergence times. First, although there are some documented cases where mitochondrial data can give seemingly misleading results in salamanders (e.g., Rancilhac et al., 2021), we do not know of studies suggesting that mitochondrial data are generally misleading for phylogenetic inference. Indeed, our results showed that species represented only by mitochondrial data were consistently placed in the genera expected based on prior taxonomy (with the caveat that mitochondrial data may have been used in the prior genus-level taxonomy in some cases). Therefore, including some species based on mitochondrial DNA alone seems strongly preferable to the current alternatives, which are: (1) adding these species randomly within genera based on taxonomy alone (e.g., Jetz and Pyron 2018), and (2) excluding them entirely until some point in the future when they can be included based on multiple nuclear loci. Second, we performed analyses to address the possibility that mitochondrial data may have negatively influenced the overall concatenated analyses. Specifically, we generated a tree among genera based on nuclear data alone (Fig. 10) and compared this to the main concatenated tree. We found that this nuclear-only tree was broadly congruent with that based on the combined nuclear and mitochondrial data (80 % of nodes shared) and that all conflicting nodes were only weakly supported in the nuclear-only tree. This latter result suggests that many of these conflicts did not arise solely because the nuclear-only tree was right and the concatenated tree was wrong, but instead suggests that many conflicts arose because the nuclear data alone were insufficient to resolve the phylogeny with strong branch support (even when only considering relationships among genera). For the majority of these conflicting nodes (8/14; 57%), the resolution in the combined-data tree was strongly supported. Of course, this is no guarantee that all of these conflicting nodes are resolved correctly by the combined data. However, there is little justification for dismissing our results on the grounds that they are generally distorted by misleading mitochondrial data. A meta-analysis in vertebrates suggested that conflicts between trees from mitochondrial and nuclear data are common (and are associated with shorter branches) but strongly supported conflicts are uncommon and can be resolved in favor of mitochondrial data or nuclear data with similar frequency (Fisher-Reid and Wiens, 2011). Third, we specifically tested whether species with only mitochondrial data had longer estimated branch lengths and older divergence times. We found that species with only mitochondrial data tended to have longer maximum-likelihood branch lengths (as expected given their faster rates), but that this difference disappeared entirely when comparing the estimated divergence times.

We note several limitations of our study and areas for future research on salamander phylogeny. First, phylogenomic data (i.e. hundreds of nuclear loci) are lacking for most species. Indeed, 263 species were included based on mitochondrial data alone. Adding large numbers of nuclear loci should help to resolve many parts of the phylogeny that remain uncertain, especially if the same loci are consistently sampled across species. These uncertain parts include relationships among some bolitoglossine genera (e.g., *Chiropterotriton*, *Parvimolge*, *Thorius*) and at least some relationships within many genera (e.g., *Ambystoma*, *Bolitoglossa*, *Chiropterotriton*, *Eurycea*, *Hynobius*, *Plethodon*). Unfortunately, we also note that phylogenomic data have a somewhat mixed record for fully resolving all within-genus relationships. For example, they appear to have been very successful in *Desmognathus* (Pyron et al., 2020, 2022b) and less so for the very recent splits in Mexican *Ambystoma* (Everson et al., 2021). Addition of more nuclear genes for all species could further resolve concerns about potential nuclear-mitochondrial discordance (e.g., Rancilhac et al., 2021), and the potential impacts of mitochondrial data on divergence-date estimation (e.g., Mulcahy et al., 2012). A second major area for future research is to include data on the remaining 7% of described salamander species. Over 80% of the remaining unsampled species are bolitoglossine salamanders. Third, we also note

that the tree was estimated using concatenated analyses, rather than species-tree methods. However, the application of species-tree methods may not be practical until most species are included based on a consistent set of nuclear loci. Overall, we acknowledge that our estimate of salamander phylogeny is not perfect, but we consider this tree a strong improvement over trees based on relatively few nuclear markers and in which many species are added randomly based on taxonomy rather than based on sequence data (Jetz and Pyron, 2018).

In summary, we provide a new estimate of salamander phylogeny that combines phylogenomic and supermatrix approaches, and includes 59% more species than the previous largest study. This tree is generally strongly supported and consistent with earlier molecular studies, with most genera recovered as monophyletic. The new divergence dates estimated are based on many more fossil calibration points than previous studies and help resolve the striking differences between two earlier estimates.

CRediT authorship contribution statement

Alexander A. Stewart: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **John J. Wiens:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Science Foundation Graduate Research Fellowship Grant DGE-2137419. We thank two anonymous reviewers and A. Larson for many helpful comments on the manuscript, and R. Bonett for input on fossil calibrations. We also thank L. Rancilhac for providing salamandrid transcriptome assemblies. Most importantly, we thank the many authors that generated the original sequence data that this study is based on.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2024.108272.

References

AmphibiaWeb. 2024. <https://amphibiaweb.org> University of California, Berkeley, CA, USA. Accessed 4 February 2024.
Bi, K., Bogart, J.P., 2010. Time and time again: unisexual salamanders (genus *Ambystoma*) are the oldest unisexual vertebrates. BMC Evolutionary Biology 10, 238.
Bonett, R.M., Blair, A.L., 2017. Evidence for complex life cycle constraints on salamander body form diversification. In: Proceedings of the National Academy of Sciences, U.S. A. 114,, pp. 9936–9941.
Bonett, R.M., Phillips, J.G., Ledbetter, N.M., Martin, S.D., Lehman, L., 2018. Rapid phenotypic evolution following shifts in life cycle complexity. Proceedings of the Royal Society of London B. Biological Sciences 285, 20172304.
Bonett, R., Steffen, M., Lambert, S., Wiens, J.J., Chippindale, P.T., 2014. Evolution of paedomorphosis in plethodontid salamanders: ecological correlates and re-evolution of metamorphosis. Evolution 68, 466–482.
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421.
Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
Cho, S., Zwick, A., Regier, J.C., Mitter, C., Cummings, M.P., Yao, J., Du, Z., Zhao, H., Kawahara, A.Y., Weller, S., Davis, D.R., Baixeras, J., Brown, J.W., Parr, C., 2011. Can deliberately incomplete gene sample augmentation improve a phylogeny estimate for the advanced moths and butterflies (Hexapoda: Lepidoptera)? Systematic Biology 60, 782–796.
Cicero, C., Koo, M.S., Braker, E., Abbot, J., Bloom, D., Campbell, M., Cook, J.A., Demboski, J.R., Doll, A.C., Frederick, L.M., Linn, A.J., Mayfield-Meyer, T.J., McDonald, D.L., Nachman, M.W., Olson, L.E., Roberts, D., Sikes, D.S., Witt, C.C., Wommack, E.A., 2024. Arctos: Community-driven innovations for managing natural and cultural history collections. PLoS ONE 19, E0296478.
Dufresnes, C., Hernandez, A., 2023. Towards completing the crocodile newts' puzzle with all-inclusive phylogeographic resources. Zoological Journal of the Linnean Society 197, 620–640.
Everson, K.M., Gray, L.N., Jones, A.G., Lawrence, N.M., Foley, M.E., Sovacool, K.L., Kratovil, J.D., Hotaling, S., Hime, P.M., Storfer, A., Parra-Olea, G., Percino-Daniel, R., Aguilar-Miguel, X., O'Neill, E.M., Zambrano, L., Shaffer, H.B., Weisrock, D.W., 2021. Geography is more important than life history in the recent diversification of the tiger salamander complex. Proceedings of the National Academy of Sciences, U.S.A. 118, E2014719119.
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. Proceedings of the National Academy of Sciences, U.S.A. 114, E5864–E5870.
Fisher-Reid, M.C., Wiens, J.J., 2011. What are the consequences of combining nuclear and mitochondrial data for phylogenetic analysis? Lessons from *Plethodon* salamanders and 13 other vertebrate clades. BMC Evolutionary Biology 11, 300.
Frost, D. R., 2024. Amphibian Species of the World: an Online Reference. Version 6.2 (Accessed 4 February 2024). Electronic Database accessible at https://amphibiansoftheworld.amnh.org/index.php. American Museum of Natural History, New York, USA.
Gao, K.-Q., Shubin, N.H., 2012. Late Jurassic salamandroid from western Liaoning, China. Proceedings of the National Academy of Sciences, U.S.A. 109, 5767–5772.
Gouveia-Oliveira, R., Sackett, P.W., Pedersen, A.G., 2007. MaxAlign: maximizing usable data in an alignment. BMC Bioinformatics 8, 312.
Hime, P.M., Briggler, J.T., Reece, J.S., Weisrock, D.W., 2019. Genomic data reveal conserved female heterogamety in giant salamanders with gigantic nuclear genomes. G3 Genes|genomes|genetics 9, 3467–3476.
Hime, P.M., Lemmon, A.R., Lemmon, E.C.M., Prendini, E., Brown, J.M., Thomson, R.C., Kratovil, J.D., Noonan, B.P., Pyron, R.A., Peloso, P.L.V., Kortyna, M.L., Keogh, J.S., Donnellan, S.C., Mueller, R.L., Raxworthy, C.J., Kunte, K., Ron, S.R., Das, S., Gaitonde, N., Green, D.M., Labisko, J., Che, J., Weisrock, D.W., 2021. Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. Systematic Biology 70, 49–66.
Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T., 2016. Phylogenomic inference of landfowl (Aves: Galliformes): effects of missing data in concatenated and coalescent inference and evidence for a bias in gene tree reconciliation approaches. Molecular Biology and Evolution 33, 1110–1125.
Huerta-Cepas, J., Serra, F., Bork, P., 2016. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. Molecular Biology and Evolution 33, 1635–1638.
Jetz, W., Pyron, R.A., 2018. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nature Ecology and Evolution 2, 850–858.
Jiang, W., Chen, S.-Y., Wang, H., Li, D.-Z., Wiens, J.J., 2014. Should genes with missing data be excluded from phylogenetic analyses? Molecular Phylogenetics and Evolution 80, 308–318.
Jockusch, E.L., Martínez-Solano, I., Timpe, E.K., 2015. The effects of inference method, population sampling, and gene sampling on species tree inferences: an empirical study in slender salamanders (Plethodontidae: *Batrachoseps*). Systematic Biology 64, 66–83.
Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30, 3059–3066.
Kozak, K.H., Wiens, J.J., 2012. Phylogeny, ecology, and the origins of climate-richness relationships. Ecology 93, S167–S181.
Kozak, K.H., Larson, A., Bonett, R.M., Harmon, L.J., 2005. Phylogenetic analysis of ecomorphological divergence, community structure, and diversification rates in dusky salamanders (Plethodontidae: *Desmognathus*). Evolution 59, 2000–2016.
Kozak, K.H., Mendyk, R.W., Wiens, J.J., 2009. Can parallel diversification occur in sympatry? Repeated patterns of body-size evolution in co-existing clades of North American salamanders. Evolution 63, 1769–1784.
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution 34, 772–773.
Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Systematic Biology 58, 130–145.
Lewis, Z.R., Kerney, R., Hanken, J., 2022. Developmental basis of evolutionary lung loss in plethodontid salamanders. Science Advances 8, eabo6108.
Liedtke, H.C., Gower, D.J., Wilkinson, M., Gomez-Mestre, I., 2018. Macroevolutionary shift in the size of amphibian genomes and the role of life history and climate. Nature Ecology and Evolution 2, 1792–1799.
Liedtke, H.C., Wiens, J.J., Gomez-Mestre, I., 2022. The evolution of reproductive modes and life cycles in amphibians. Nature Communications 13, 7039.
Means, D.B., Lamb, J., Bernardo, J., 2017. A new species of dusky salamander (Amphibia: Plethodontidae: *Desmognathus*) from the Eastern Gulf Coastal Plain of the United States and a redescription of *D. auriculatus*. Zootaxa. 4263, 467–506.

Melander, S.L., Mueller, R.L., 2020. Comprehensive analysis of salamander hybridization suggests a consistent relationship between genetic distance and reproductive isolation across tetrapods. Copeia 108, 987–1003.
Mueller, R.L., Macey, J.R., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proceedings of the National Academy of Sciences, U.S.A. 101, 13820–13825.
Mulcahy, D.G., Noonan, B.P., Moss, T., Townsend, T.M., Reeder, T.W., Sites Jr., J.W., Wiens, J.J., 2012. Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. Molecular Phylogenetics and Evolution 65, 974–991.
Nishikawa, K., Jiang, J.-P., Matsui, M., Mo, Y.-M., 2011. Unmasking *Pachytriton labiatus* (Amphibia: Urodela: Salamandridae), with description of a new species of *Pachytriton* from Guangxi, China. Zoological Science 28, 453–461.
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. Molecular Ecology 22, 111–129.
Parra-Olea, G., García-París, M., Wake, D.B., 2004. Molecular diversification of salamanders of the tropical American genus *Bolitoglossa* (Caudata: Plethodontidae) and its evolutionary and biogeographical implications. Biological Journal of the Linnean Society 81, 325–346.
Pierson, T.W., Kozak, K.H., Glenn, T.C., Fitzpatrick, B.M., 2024. River drainage reorganization and reticulate evolution in the two-lined salamander (*Eurycea bislineata*) species complex. Systematic Biology 73, 26–35.
Portik, D.M., Streicher, J.W., Blackburn, D.C., Moen, D.S., Hutter, C.R., Wiens, J.J., 2023b. Redefining possible: combining phylogenomic and supersparse data in frogs. Molecular Biology and Evolution 40, msad109.
Portik, D.M., Streicher, J.W., Wiens, J.J., 2023a. Frog phylogeny: A time-calibrated, species-level tree based on hundreds of loci and 5,242 species. Molecular Phylogenetics and Evolution 188, 107907.
Portik, D.M., Wiens, J.J., 2020. SuperCRUNCH: A bioinformatics toolkit for creating and manipulating supermatrices and other large phylogenetic datasets. Methods in Ecology and Evolution 11, 763–772.
Pyron, R.A., 2014. Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. Systematic Biology 63, 779–797.
Pyron, R.A., Beamer, D.A., 2022a. Nomenclatural solutions for diagnosing 'cryptic' species using molecular and morphological data facilitate a taxonomic revision of the black-bellied salamanders (Urodela, *Desmognathus 'quadramaculatus'*) from the southern Appalachian Mountains. Biotaxa 27, 1–43.
Pyron, R.A., Beamer, D.A., 2022b. Allocation of *Salamandra auriculata* Holbrook, 1838, with a new species of swamp-dwelling dusky salamander (Plethodontidae: *Desmognathus*) from the Atlantic Coastal Plain. Zootaxa 5188, 587–5s95.
Pyron, R.A., Beamer, D.A., 2022c. Systematics of the Ocoee salamander (Plethodontidae: *Desmognathus ocoee*), with description of two new species from the southern Blue Ridge Mountains. Zootaxa 5190, 207–240.
Pyron, R.A., Beamer, D.A., 2023a. A systematic revision of the shovel-nosed salamander (Plethodontidae: *Desmognathus marmoratus*), with re-description of the related *D. aureatus* and *D. intermedius*. Zootaxa 5270, 262–280.
Pyron, R.A., Beamer, D.A., 2023b. Systematic revision of the spotted and northern dusky Salamanders (Plethodontidae: *Desmognathus conanti* and *D. fuscus*), with six new species from the eastern United States. Zootaxa 5311, 451–504.
Pyron, R.A., O'Connell, K.A., Lemmon, E.M., Lemmon, A.R., Beamer, D.A., 2020. Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*). Molecular Phylogenetics and Evolution 146, 106751.
Pyron, R.A., O'Connell, K.A., Lamb, J.Y., Beamer, D.A., 2022a. A new, narrowly endemic species of swamp-dwelling dusky salamander (Plethodontidae: *Desmognathus*) from the Gulf Coastal Plain of Mississippi and Alabama. Zootaxa 5133, 53–82.
Pyron, R.A., O'Connell, K.A., Lemmon, E.M., Lemmon, A.R., Beamer, D.A., 2022b. Candidate-species delimitation in *Desmognathus* salamanders reveals gene flow across lineage boundaries, confounding phylogenetic estimation and clarifying hybrid zones. Ecology and Evolution 12, e8574.
Pyron, R.A., O'Connell, K.A., Duncan, S.C., Burbrink, F.T., Beamer, D.A., 2023. Speciation hypotheses from phylogeographic delimitation yield an integrative taxonomy for seal salamanders (*Desmognathus monticola*). Systematic Biology 72, 179–197.
Rancilhac, L., Irisarri, I., Angelini, C., Arntzen, J.W., Babik, W., Bossuyt, F., Künzel, S., Tim Lüddecke, T., Pasmans, F., Sanchez, E., Weisrock, D., Veith, M., Wielstra, B., Steinfartz, S., Hofreiter, M., Philippe, H., Vences, M., 2021. Phylotranscriptomic evidence for pervasive ancient hybridization among Old World salamanders. Molecular Phylogenetics and Evolution 155, 106967.
Ranwez, V., Douzery, E.J.P., Cambon, C., Chantret, N., Delsuc, F., 2018. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Molecular Biology and Evolution 35, 2582–2584.
Reilly, S.B., Wake, D.B., 2019. Taxonomic revision of black salamanders of the *Aneides flavipunctatus* complex (Caudata. Plethodontidae). PeerJ e7370.
Roure, B., Baurain, D., Philippe, H., 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Molecular Biology and Evolution 30, 197–214.
Rovito, S.M., Parra-Olea, G., Recuero, E., Wake, D.B., 2015. Diversification and biogeographical history of Neotropical plethodontid salamanders. Zoological Journal of the Linnean Society 175, 167–188.
Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Molecular Biology and Evolution 19, 101–109.
Sanderson, M.J., McMahon, M.M., Steel, M., 2010. Phylogenomics with incomplete taxon coverage: the limits to inference. BMC Evolutionary Biology 10, 155.
Sessions, S.K., 2008. Evolutionary cytogenetics in salamanders. Chromosome Research 16, 183–201.
Shen, X.X., Liang, D., Feng, Y.J., Chen, M.Y., Zhang, P., 2013. A versatile and highly efficient toolkit including 102 nuclear markers for vertebrate phylogenomics, tested by resolving the higher level relationships of the Caudata. Molecular Biology and Evolution 30, 2235–2248.
Shen, X.-X., Liang, D., Chen, M.-Y., Mao, R.-L., Wake, D.B., Zhang, P., 2016. Enlarged multilocus data set provides surprisingly younger time of origin for the Plethodontidae, the largest family of salamanders. Systematic Biology 65, 66–81.
Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28, 2689–2690.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus. Evolution 4, vey016.
Sun, C., López Arriaza, J.R., Mueller, R.L., 2012. Slow DNA loss in the gigantic genomes of salamanders. Genome Biology and Evolution 4, 1340–1348.
Talavera, G., Lukhtanov, V., Pierce, N.E., Vila, R., 2022. DNA barcodes combined with multilocus data of representative taxa can generate reliable higher-level phylogenies. Systematic Biology 71, 382–395.
Vieites, D.R., Román, S.N., Wake, M.H., Wake, D.B., 2011. A multigenic perspective on phylogenetic relationships in the largest family of salamanders, the Plethodontidae. Molecular Phylogenetics and Evolution 59, 623–635.
Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribero, A. H., Pedregosa, F., van Mulbregt, P., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17, 261–272. SciPy 1.0 Contributors.
Weisrock, D.W., Macey, J.R., Matsui, M., Mulcahy, D.G., Papenfuss, T.J., 2013. Molecular phylogenetic reconstruction of the endemic Asian salamander family Hynobiidae (Amphibia, Caudata). Zootaxa 3626, 77–93.
Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Systematic Biology 52, 528–538.
Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Systematic Biology 60, 719–731.
Wiens, J.J., Fetzner, J.W., Parkinson, C.L., Reeder, T.W., 2005. Hylid frog phylogeny and sampling strategies for speciose clades. Systematic Biology 54, 719–748.
Wiens, J.J., Parra-Olea, G., García-París, M., Wake, D.B., 2007. Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. Proceedings of the Royal Society B: Biological Sciences 274, 919–928.
Wiens, J.J., Sparreboom, M., Arntzen, J.W., 2011. Crest evolution in newts: implications for reconstruction methods, sexual selection, phenotypic plasticity and the origin of novelties. Journal of Evolutionary Biology 24, 2073–2086.
Williams, J.S., Niedzwiecki, J.H., Weisrock, D.W., 2013. Species tree reconstruction of a poorly resolved clade of salamanders (Ambystomatidae) using multiple nuclear loci. Molecular Phylogenetics and Evolution 68, 671–682.
Xi, Z., Liu, L., Davis, C.C., 2016. The impact of missing data on species tree estimation. Molecular Biology and Evolution 33, 838–860.
Zheng, Y., Peng, R., Kuro-o, M., Zeng, X., 2011. Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: a case study of salamanders (Order Caudata). Molecular Biology and Evolution 28, 2521–2535.
Zheng, Y., Wiens, J.J., 2015. Do missing data influence the accuracy of divergence-time estimation with BEAST? Molecular Phylogenetics and Evolution 85, 41–49.
Zheng, Y., Wiens, J.J., 2016. Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. Molecular Phylogenetics and Evolution 94, 537–547.