PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Chan KO, Hutter CR, Wood Jr PL, Grismer LL, Brown RM. 2020 Target-capture phylogenomics provide insights on gene and species tree discordances in Old World treefrogs (Anura: Rhacophoridae). *Proc. R. Soc. B* **287**: 20202102. https://doi.org/10.1098/rspb.2020.2102

Received: 27 August 2020
Accepted: 13 November 2020

Subject Category:
Evolution

Subject Areas:
evolution, genomics, taxonomy and systematics

Keywords:
incomplete lineage sorting, anomaly zone, bootstrap, parsimony informative sites, concordance factor, branch support

Author for correspondence:
Kin Onn Chan
e-mail: cko@nus.edu.sg

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5218608.

THE ROYAL SOCIETY PUBLISHING

Target-capture phylogenomics provide insights on gene and species tree discordances in Old World treefrogs (Anura: Rhacophoridae)

Kin Onn Chan[1], Carl R. Hutter[2], Perry L. Wood Jr[3], L. Lee Grismer[4] and Rafe M. Brown[5]

[1]Lee Kong Chian Natural History Museum, National University of Singapore, 2 Conservatory Drive, Singapore 117377, Republic of Singapore
[2]Museum of Natural Sciences and Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA
[3]Department of Biological Sciences and Museum of Natural History, Auburn University, Auburn, AL 36849, USA
[4]Herpetology Laboratory, Department of Biology, La Sierra University, Riverside, CA 92505, USA
[5]Biodiversity Institute and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA

KOC, 0000-0001-6270-0983; CRH, 0000-0001-6381-6339; PLW, 0000-0003-3767-5274; LLG, 0000-0001-8422-3698; RMB, 0000-0001-5338-0658

Genome-scale data have greatly facilitated the resolution of recalcitrant nodes that Sanger-based datasets have been unable to resolve. However, phylogenomic studies continue to use traditional methods such as bootstrapping to estimate branch support; and high bootstrap values are still interpreted as providing strong support for the correct topology. Furthermore, relatively little attention has been given to assessing discordances between gene and species trees, and the underlying processes that produce phylogenetic conflict. We generated novel genomic datasets to characterize and determine the causes of discordance in Old World treefrogs (Family: Rhacophoridae)—a group that is fraught with conflicting and poorly supported topologies among major clades. Additionally, a suite of data filtering strategies and analytical methods were applied to assess their impact on phylogenetic inference. We showed that incomplete lineage sorting was detected at all nodes that exhibited high levels of discordance. Those nodes were also associated with extremely short internal branches. We also clearly demonstrate that bootstrap values do not reflect uncertainty or confidence for the correct topology and, hence, should not be used as a measure of branch support in phylogenomic datasets. Overall, we showed that phylogenetic discordances in Old World treefrogs resulted from incomplete lineage sorting and that species tree inference can be improved using a multi-faceted, total-evidence approach, which uses the most amount of data and considers results from different analytical methods and datasets.

1. Introduction

One of the main foci of phylogenomics is to elucidate ambiguous relationships that Sanger-derived markers were unable to resolve. For Sanger-based phylogenetic analysis, the gold standard for resolving relationships is to obtain strong branch support, usually measured in the form of high bootstrap proportions or Bayesian posterior probabilities [1,2]. However, phylogenomic studies have demonstrated that these measures do not necessarily reflect confidence for the correct species tree topology. High branch support can be positively misleading and adding orders of magnitude more data can counterintuitively increase spurious support for the wrong tree [3–5]. This is corroborated by in-depth examinations of genome-scale data that have revealed high levels of gene tree discordance [6–8] and well-supported but conflicting evolutionary

© 2020 The Author(s) Published by the Royal Society. All rights reserved.

histories among different data partitions, portions of the genome or classes of markers [9,10].

Species tree inference using genomic data can be biased by systematic errors resulting from insufficient informative data leading to unresolved gene trees [11], different data filtering strategies or datatypes [7,12–15], and the implementation of different analytical methods such as concatenation versus coalescent methods [16–18]. Although various analytical strategies have been implemented to reduce systematic error, a best-practice consensus has yet to be reached, indicating a lack of understanding on how to effectively analyse large genomic datasets, thereby hampering our ability to fully harness the power of genome-scale data.

Discordance between gene trees and the species tree can also result from biological processes such as introgression, gene duplication and loss, horizontal gene transfer and incomplete lineage sorting (ILS) [19–21]. ILS occurs when ancestral polymorphisms do not reach fixation between successive divergence events. This can occur during periods of rapid diversification, particularly when effective population size is large relative to its associated branch length [22]. One of the adverse effects of extreme ILS is that the most common gene trees may not be congruent with the species tree topology. In such instances, the tree-space where these so-called anomalous gene trees predominate is known as the anomaly zone [23]. Consequently, species tree inference in the anomaly zone can be challenging due to insufficient amounts of informative sites at short branches [24], stochastic mutational variance [25] and inappropriate methods that do not account for ILS [16]. Although the conceptual basis of the anomaly zone has been well established, its occurrence and effects on empirical systems have rarely been demonstrated [26,27]. As such, it is important for empirical phylogenomic studies to examine the numerous underlying factors that could potentially bias species tree inference [7,28]. A multifarious approach that thoroughly explores potential biases, considers ameliorative steps and assesses the consistency of results across different treatments of data and comparisons of analytical methods should therefore be the new benchmark for phylogenomic studies.

Old World treefrogs of the family Rhacophoridae consist of 428 species that are widely distributed across Asia and Southeast Asia, with a disjunct occurrence in Africa [29]. Although this charismatic family has been the focus of many phylogenetic studies, the relationships of several major clades have yet to be unambiguously resolved [30–33]. In particular, the placements of the genera *Gracixalus*, *Philautus*, *Feihyla*, *Polypedates* and *Rhacophorus sensu lato* (s.l.) have never been completely resolved [33–36]. Additionally, placement of the genus *Nyctixalus* has, alternatively, been inferred with strong support in two different topological positions [37–39]. All phylogenetic studies on rhacophorids have so far focused on obtaining strong branch support. By contrast, no studies have been performed to assess the robustness of inferred relationships (despite ostensibly strong support) and causes of phylogenetic discordance, which can be used to provide a deeper understanding on the evolutionary processes that affect species tree inference. As such, we employed a multifarious approach using a large and diverse set of phylogenomic markers consisting of exons, introns, UCEs and a set of 30 Sanger-based markers generated using FrogCap [40], to compare the consistency of inferred relationships among rhacophorid genera. We also assessed the effects of various data filtering strategies and analytical methods, in addition to an in-depth examination of gene tree and branch length variation, in hopes of determining underlying causes of phylogenetic discordance.

2. Materials and methods

(a) Taxon sampling and DNA extraction

Datasets consist of five outgroup (*Arthroleptis variabilis*, *Scaphiophryne marmorata*, *Cornufer guentheri*, *Boophis tephraeomystax* and *Mantidactylus melanopleura*) and 45 ingroup samples. The ingroup is represented by 10 of the 18 known genera: *Nyctixalus*, *Theloderma*, *Gracixalus*, *Kurixalus*, *Raorchestes*, *Philautus*, *Chiromantis*, *Feihyla*, *Polypedates* and *Rhacophorus*. The outgroup comprises taxa from the family Mantellidae (sister to Rhacophoridae), followed by taxa from progressively distant families such as Ceratobatrachidae, Arthroleptidae and Microhylidae [41]. Tissue samples for molecular work were obtained from the museum holdings of the University of Kansas (KU), California Academy of Sciences (CAS), La Sierra University Herpetological Collection, Riverside, California (LSUHC), Field Museum of Natural History (FMNH), Australian Museum (AMH) and the Museum of Vertebrate Zoology, Berkeley (MVZ). The list of samples and their associated metadata are presented in electronic supplementary material, table S1.

(b) Bioinformatics

Details on library preparation, sequencing and bioinformatics are provided in the electronic supplementary material. Alignments were generated and processed using custom scripts (https://github.com/chutter/FrogCap-Sequence-Capture) and saved separately into usable datasets for phylogenetic analyses and data type comparisons: (i) introns (exons were trimmed from the original contig and the two remaining intronic regions were concatenated); (ii) exons (each alignment was adjusted to be in an open-reading frame and trimmed to the largest reading frame that accommodated greater than 90% of the sequences); (iii) UCEs; and (iv) 'legacy', commonly used nuclear markers from amphibian phylogenetic studies (full list of markers can be obtained from [40]).

(c) Data filtering and phylogenetic analysis

To assess potential biases arising from different datatypes and data filtering strategies, we analysed all datasets separately. Because markers with missing taxa and low phylogenetic information can potentially increase gene tree estimation error [11], each dataset was filtered at 75% taxon sampling completeness (markers with 25% or more missing data were discarded) and 50% parsimony informative sites (PIS; bottom 50% of markers with the least PIS were discarded). Data filtering was not performed on the legacy dataset due to insufficient numbers of markers. Finally, we performed a total-evidence analysis on a combined dataset consisting of unfiltered exon, intron and UCE markers (all-combined). Sampling completeness and PIS were calculated using the summary function in the program AMAS [42].

Phylogenetic estimation was performed using concatenation and summary coalescent-based methods. For the analysis of concatenated data, we used the program IQ-TREE v. 1.7 [43] to perform two sets of analyses on each dataset. The first was an unpartitioned analysis using the GTR + GAMMA substitution model, while the second was a partitioned analysis (partitioned by marker) with the GTR + GAMMA model applied to each partition. We applied a parameter-rich substitution model for both analyses (as opposed to model testing) due to computational limitations, and because this strategy has been shown to produce similar, if not better results for large phylogenomic datasets [44,45]. Branch support was assessed using ultrafast bootstrapping (1000 replicates) [46]. We also tested for saturation at the first,

Table 1. List of datasets analysed in this study with their associated summary statistics. Datasets with the prefix unf = unfiltered; mis75 = 75% taxon completeness; pis50 = top 50% markers with highest parsimony informative sites.

dataset	number of markers (mean length)	base pairs	variable sites			PIS		
			total	mean	mean prop.	total	mean	mean prop.
unf-exon	12 370 (208)	2 579 130	1 173 581	95	0.43	767 630	62	0.27
mis75-exon	8328 (229)	1 904 991	902 898	108	0.44	609 565	73	0.30
pis50-exon	6185 (277)	1 715 499	863 701	140	0.50	582 889	94	0.33
unf-intron	12 257 (424)	5 191 144	4 443 393	362	0.86	3 387 639	276	0.65
mis75-intron	7558 (452)	3 414 409	2 981 322	394	0.88	2 364 933	313	0.70
pis50-intron	6129 (531)	3 256 482	2 887 746	471	0.89	2 283 964	373	0.71
unf-UCE	650 (744)	483 713	247 439	381	0.50	150 023	231	0.30
mis75-UCE	533 (777)	414 172	217 542	408	0.50	136 055	255	0.32
pis50-UCE	325 (878)	285 325	163 895	504	0.58	106 704	328	0.38
legacy	30 (1150)	34 506	13 046	435	0.35	8405	280	0.22

second and third codon positions using the index for substitution saturation implemented in the program DAMBE7 [47]. To overcome the issue of saturation, we conducted an IQ-TREE analysis on the amino acid alignment of the unfiltered exon dataset.

We performed a summary-based species tree analysis using ASTRAL-III [48]. IQ-TREE was used to estimate gene trees for each individual marker. To assess the effects of model selection on gene tree and species tree inference, we implemented two strategies. First, gene trees were estimated using the best-fit substitution model for each individual marker as determined by the program ModelFinder [49]. Second, we applied a parameter-rich model (GTR + GAMMA) on all gene trees. In both analyses, the best ML tree was retained as input for the ASTRAL analysis. We also implemented the site-based method SVDQuartets [50] that bypasses the use of gene trees altogether. The Quartet Fiduccia–Mattheyses (QFM) assembly algorithm was implemented without local searching. Topology tests were performed in IQ-TREE to assess the fit of inferred topologies against the all-combined alignment (additional details in the electronic supplementary material).

(d) Branch support and congruence

In addition to bootstrap (IQ-TREE) and posterior probabilities (ASTRAL), we used gene (gCF) and site concordance factor (sCF) to investigate topological conflict around each branch of the species tree. For every branch of the species tree, the gCF and sCF represents the percentage of decisive gene trees and alignment sites, respectively, containing that branch [51]. Assuming that the only source of discordance between gene and species trees is ILS, the multi-species coalescent model predicts that the probability of a gene tree quartet matching the species tree topology is higher than the probability of matching the two alternatives. Additionally, the two alternatives will have similar frequencies [52,53]. Using this expectation, we calculated the relative frequency of branch quartets surrounding focal clades. We then used a χ^2-test to determine whether the frequency of gene trees (gCF) and sites (sCF) supporting the two alternate topologies was significantly different [54]. Non-significant p-values ($p > 0.05$) indicate a failure to reject the hypothesis of equal frequencies, implying that discordance among gene trees and/or sites is likely due to ILS. Concordance factors were calculated in IQ-TREE [51] and the χ^2-test was performed in R using scripts available here: http://www.robertlanfear.com/blog/files/concordance_factors.html. Because ILS is often associated with rapid radiations, we tested focal branch lengths (d, in coalescent units) for polytomies ($d = 0$), where all three frequencies are expected to be close to one-third if polytomous [55].

Short internal branches in a species tree are also in danger of being in the anomaly zone and can generate gene tree topologies that do not match the species tree [23]. We followed the approach by [26,56] to estimate the presence of the anomaly zone in four-taxon subclades by comparing consecutive, ancestor and descendant, internal branch lengths. If these sets of internodes were in the anomaly zone, at least one anomalous gene tree would be expected to be present. Relative branch quartet frequency analyses were performed using the program DiscoVista [53]. Polytomy tests were performed and branch lengths (in coalescent units) were obtained with ASTRAL-III.

3. Results

(a) Genomic data

After quality filtering, loci matching and alignment, a total of 13 221 loci were obtained, including 12 370 exons, 12 257 introns, 650 UCEs and 30 legacy loci after trimming (table 1). On average, legacy markers were longest, followed by UCEs, introns and exons. Intron datasets had the most characters and the highest number of variable sites and PIS. Although UCEs had similar average PIS compared to introns, the average proportion of PIS per marker was approximately half that of introns (table 1).

(b) Phylogenetic inference

Saturation tests showed non-significant saturation at all three codon positions (electronic supplementary material, table S2 and figure S1). Partitioned and unpartitioned IQ-TREE analyses yielded congruent topologies. Similarly, ASTRAL species tree topologies derived from gene trees estimated using either model testing or GTR + GAMMA produced congruent topologies. Insignificant changes in branch lengths and branch support were detected in both comparisons (partitioned versus unpartitioned; model testing versus GTR + GAMMA).

Overall, five different topologies (referred to as T1–T5; figure 1) were inferred across our three classes of phylogenetic

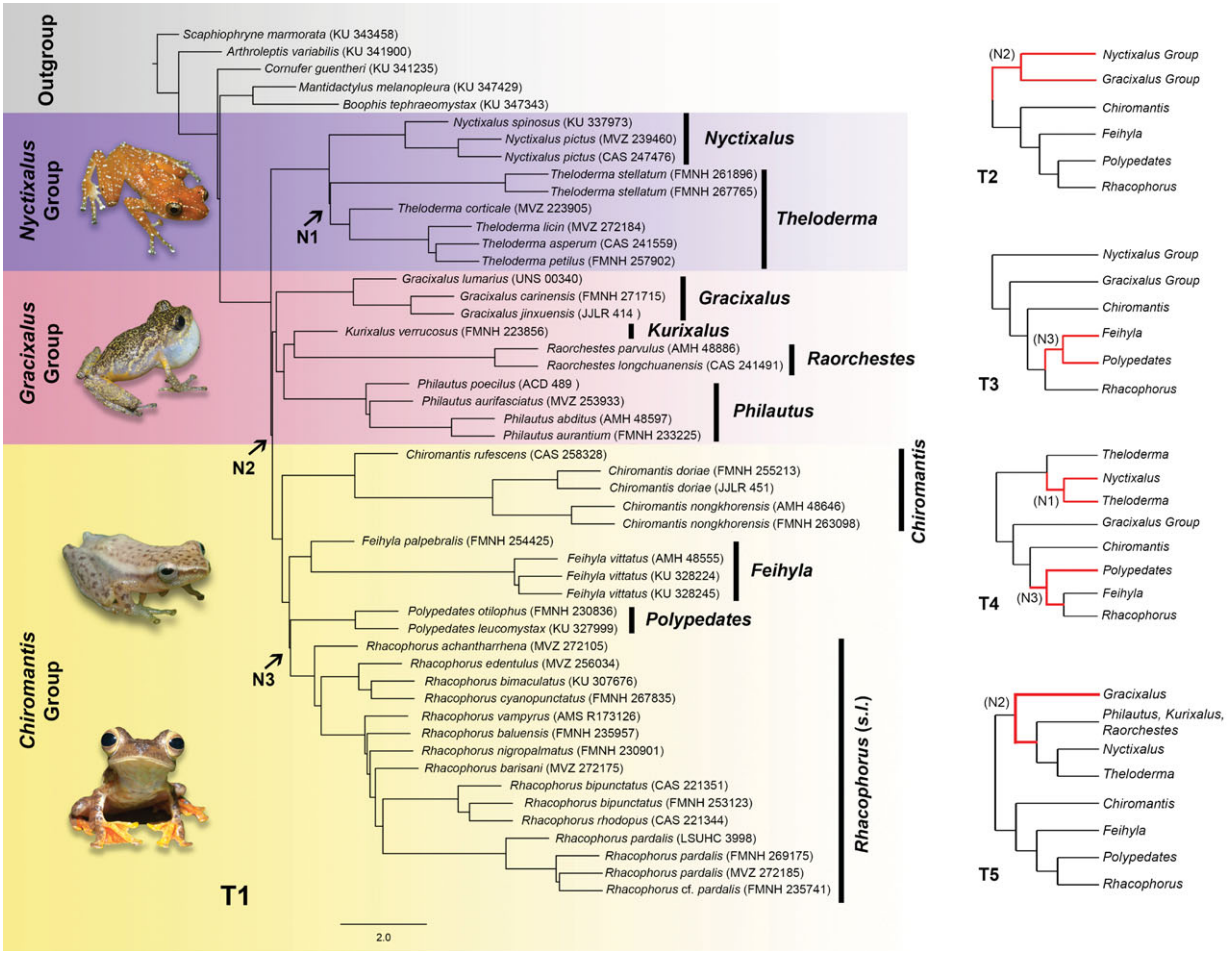

Figure 1. Inferred topologies (T1–T5) from IQ-TREE, ASTRAL-III and SVDQuartets across all datasets (datasets and their corresponding topologies are summarized in table 2). Left: representative ASTRAL species tree from the all-combined dataset depicting the T1 topology; N1–N3 are nodes where discordance occur across the five topologies. Right: cladograms depicting the T2–T5 topologies, where red branches represent discordant clades with respect to T1. Inset photos (from top to bottom) are representative species for each of the major groups: *Nyctixalus pictus*, *Philautus vermiculatus*, *Chiromantis marginis* and *Rhacophorus pardalis*. All photos were taken by the first author. (Online version in colour.)

analyses and 11 datasets. IQ-TREE and ASTRAL produced congruent topologies except for the exon datasets (table 2). The exon datasets produced variable results and inferred either T1 or T3 depending on the type of analytical method and filtering scheme. SVDQuartets consistently inferred T5 across all datasets (figure 1 and table 2). The topology tests showed that T1 was the optimal topology, while topologies T2–T5 were strongly rejected (electronic supplementary material, table S3).

(c) Phylogenetic discordance

Discordance across the five topologies revolved around three nodes (N1–N3; figure 1). A comparison between bootstrap and concordance factors showed that there was no relationship between bootstrap support and congruence (gCF and sCF; figure 2*a*). All focal nodes had high bootstrap values despite high levels of discordance, as indicated by low gCF and sCF values (figure 2*a*) and there was no correlation between branch length and its associated bootstrap value (figure 2*b*). However, there was a strong positive correlation between branch length and concordance factors (figure 2*c*,*d*). Gene concordance did not significantly improve when data was filtered by taxon completeness but showed marginal improvements when filtered by PIS (electronic supplementary material, figure S2).

The relative frequency of gene trees at N1 revealed that most gene trees supported the T1 topology across all but the legacy dataset, which had a significantly larger proportion of gene trees supporting the T4 topology (figure 3). At N2, equal proportions of gene trees supported the T1 and T2 topologies in the all-combined dataset. Most gene trees supported the T1 topology in the exon, UCE and legacy datasets, whereas only 2% more gene trees supported the T2 topology across all intron datasets. At N3, exon datasets inferred either the T1 or T3 topologies depending on the filtering scheme and the analytical method. For exon datasets filtered at 50% PIS and 75% completeness, the T3 topology was inferred when analysed using IQ-TREE, but the T1 topology was inferred when analysed with ASTRAL (table 2). The relative frequency analysis of those datasets revealed that equal proportions of gene trees supported either topology. For the unfiltered exon dataset, most gene trees supported the T1 topology, but the IQ-TREE analysis inferred the T3 topology. Node 3 was further revealed as a polytomy in all exon datasets (polytomies were not detected in the other datasets).

The χ^2-test failed to reject ($p > 0.05$) the hypothesis of equal gene tree frequencies (gEF) for most datasets at N1 but rejected ($p < 0.05$) gEF for most datasets at N2 and N3. The hypothesis of equal site frequencies (sEF) could not be

Table 2. Summary of inferred topologies, branch support at each focal node (N1–N3) and results of the anomaly zone analysis (figure 1 for topological relationships and location of focal nodes). Datasets with the prefix unf = unfiltered; mis75 = 75% taxon completeness; pis50=top 50% markers with highest parsimony informative sites. QS = maximum normalized quartet score; BS = bootstrap support from IQ-TREE; PP = local posterior probability from ASTRAL; AZ = anomaly zone (yes/no).

	topology			N1		N2		N3	
dataset	IQ-TREE	ASTRAL-III (QS)	SVDQ	BS/PP	AZ	BS/PP	AZ	BS/PP	AZ
all-combined	T1	T1 (0.71)	T5	100/1	y	100/1	y	100/1	y
unf-exon	T3	T1 (0.68)	T5	100/1	y	100/1	y	100/1	y
mis75-exon	T3	T1 (0.68)	T5	100/1	y	100/1	y	100/0.75	y
pis50-exon	T3	T3 (0.71)	T5	100/1	y	100/1	y	100/0.23	y
unf-intron	T2	T2 (0.76)	T5	100/1	y	100/1	y	100/1	y
mis75-intron	T2	T2 (0.76)	T5	100/1	y	100/1	y	100/1	y
pis50-intron	T2	T2 (0.77)	T5	100/1	y	100/1	y	100/1	n
unf-UCE	T1	T1 (0.77)	T5	100/1	y	99/0.67	y	60/0.89	n
mis75-UCE	T1	T1 (0.77)	T5	99/1	y	93/0.63	y	72/0.91	n
pis50-UCE	T1	T1 (0.80)	T5	99/1	y	98/0.81	y	76/0.79	n
legacy	T4	T4 (0.81)	NA	99/1	n	59/0.76	n	53/0.43	y

rejected across most datasets at N1 and N3 but was rejected across most datasets at N2 (electronic supplementary material, table S3). Anomaly zone calculations showed that the focal nodes N1–N3 were in the anomaly zone for most datasets except N2 for the legacy dataset and N3 for all UCE datasets and one intron dataset filtered for PIS (table 2).

4. Discussion

(a) Impact of model selection and data filtering on phylogenetic inference

Studies have showed that unpartitioned analyses of concatenated data can be statistically inconsistent in the presence of ILS and produce erroneous topologies [18,45,57]. However, our results from partitioned and unpartitioned analyses inferred similar topologies across all datasets, indicating that the differential effects of these methods are still not fully understood. ASTRAL species tree topologies were also congruent when gene trees were estimated using model testing or GTR + GAMMA, corroborating previous studies that demonstrated the insignificant impacts of model testing on species tree inference when large amounts of data are used [44,45]. These studies, including ours, show that applying a parameter-rich substitution model such as GTR + GAMMA in lieu of model testing is a viable strategy that can significantly reduce the computation time and load of large phylogenomic analyses.

Within each class of data (intron, exon and UCE), the implementation of different filtering strategies did not alter the final topology or significantly affect branch support. Although the final topology was unchanged, filtering by PIS was clearly more effective at improving concordance compared to taxon completeness (as assessed by ASTRAL quartet scores; table 2). Interestingly, at N2, the majority of gene trees across all intron datasets supported a suboptimal topology. This mismatch between the most probable gene tree topology (T2; figure 3) and the optimal species tree (T1) indicates that introns could be more susceptible to the adverse effects of ILS compared to exons or UCEs. Another noteworthy result is that all UCE datasets inferred the optimal topology despite containing substantially less markers compared to introns and exons (weaker branch support is presumably due to the fewer number of markers). Quartet scores for UCE datasets were also higher compared to introns and exons, indicating that UCE markers may be less affected by ILS (table 2). In agreement with [45], we showed that data filtering does not confer appreciable benefits in terms of species tree inference and that analyzing more (unfiltered) data produces the best results. Surprisingly, the SVDQuartets analysis inferred a novel topology across all datasets (T5) which was not supported by the topology test, possibly due to high ILS and large numbers of sites per locus that is known to affect the accuracy of this analysis [58].

(b) Assessing discordance in phylogenomic datasets

Our study also empirically demonstrates that traditional bootstrapping is not an appropriate measure to assess branch support. Bootstrap values were not correlated with topological concordance [51] and routinely produced strong support at highly discordant nodes. This is because resampling procedures, such as non-parametric bootstrapping, essentially measure site-sampling variance as opposed to observed variance in the data. Because site-sampling variance is an inverse function of sample size (amount of data), bootstrap values will concomitantly increase with the amount of analysed data [4,59], this tendency does not necessarily reflect variation in the data themselves. Although it has long been known that the traditional bootstrap cannot be interpreted as the probability of a branch belonging to the true species tree and is biased by the total number of characters in the data [60], it remains one of the most widely used measures of branch support, even in the phylogenomic era. Hence, it is worth reiterating here that traditional bootstrapping should not be used as a measure of branch support in large phylogenomic datasets. Our

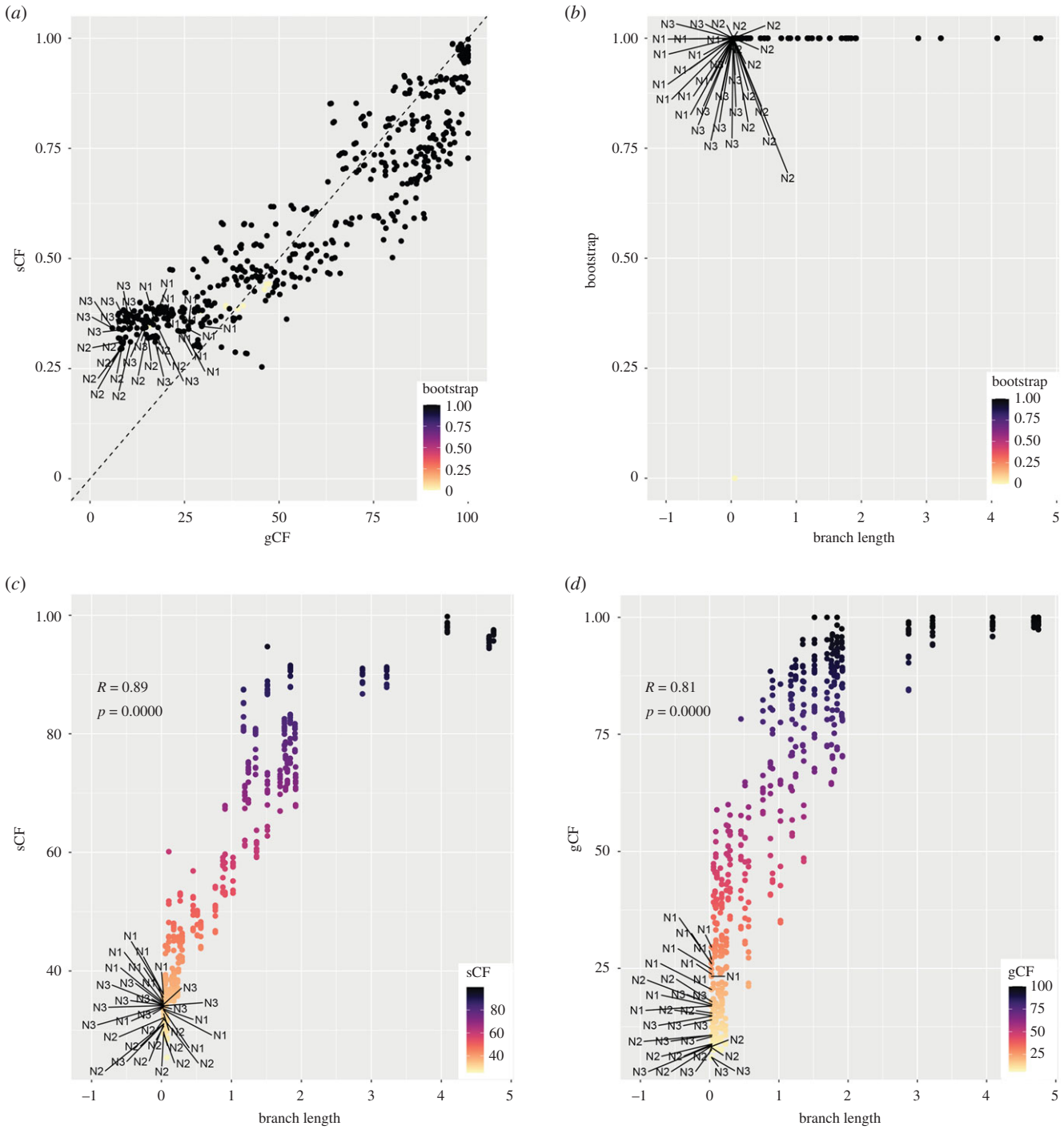

Figure 2. Scatter plots showing the relationships between (*a*) gCF, sCF and bootstrap values; (*b*) branch length (in coalescent units) and bootstrap values; (*c*) branch length and sCF; and (*d*) branch length and gCF. Points corresponding to the focal nodes (N1–N3) are labelled. (Online version in colour.)

study clearly shows that while low bootstrap support does indeed reflect uncertainty, high values can be positively misleading. Other measures such as local posterior probabilities and concordance factors provide a more accurate representation of topological variation and uncertainty in genomic datasets. However, higher concordance values can also be an artefact of small datasets [51] (our legacy datasets produced the highest overall concordance factors; electronic supplementary material, figure S2). Therefore, our results indicate that traditional bootstrapping is more appropriate for small datasets compared to concordance factors, while the converse should be applied for larger, genome-scale data.

Our results also showed that discordance was strongly correlated with branch length, where nodes with the highest discordance also had the shortest branches (figure 2). Similar to results from other phylogenomic studies, we showed that resolving very short internodes resulting from consecutive rapid diversification events remains a challenge, even with the availability of orders of magnitude more data than ever before [5,28,61].

(c) Causes of phylogenomic discordance in Rhacophoridae

All discordant nodes were associated with extremely short internal branches (figure 2), indicating that rapid diversification events were the likely cause of ILS. We further showed that discordant nodes were in the anomaly zone for most of the tested datasets, which explains the high incidence of anomalous gene trees (most probable gene trees that do not match the

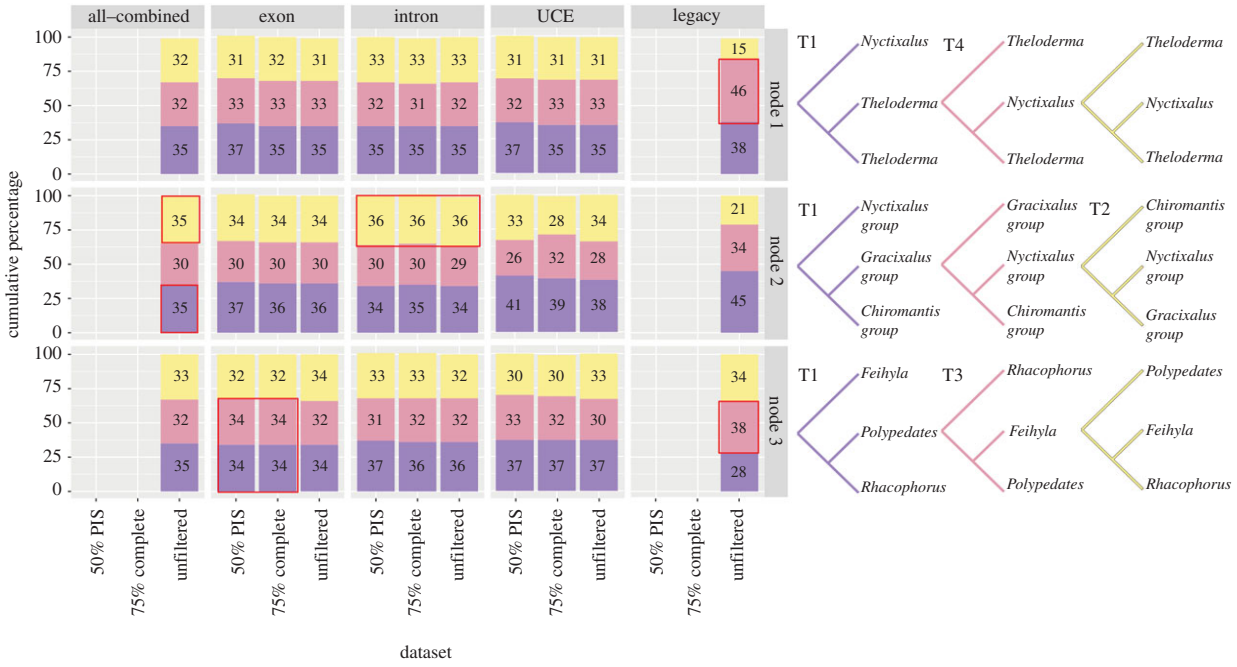

Figure 3. Relative frequency (in percentage) of gene trees supporting the three possible topologies surrounding each of the focal nodes of the ASTRAL species tree. Cladograms representing each inferred topology are colour-coded to match the stacked bars. Red boxes highlight overwhelming support for the suboptimal topology, or equal support for conflicting topologies. (Online version in colour.)

underlying species tree; figure 3). However, although present in most of our datasets, ILS had distinct effects on different nodes, datasets and analysis. At N1, only the legacy dataset inferred a different topology (T4), whereas all genomic datasets consistently inferred the T1 topology, indicating that the optimal topology was able to be inferred by increasing the amount of data. However, despite using ~580 k–3.3 million PIS, exon and intron datasets inferred conflicting topologies with strong support at N2 and N3. At N2, the majority of gene trees across all intron datasets strongly supported the T2 topology, while SVDQuartets inferred the T5 topology, indicating that discordance at N2 could be a result of both ILS and gene tree estimation error. Although the exon datasets inferred the T3 topology at N3 (all other datasets and analyses inferred T1), the gene tree frequency analysis and polytomy test revealed that N3 was a polytomy across all exon datasets, which would explain the incongruence surrounding that node for those datasets. However, we were unable to determine whether the node constituted a hard or soft polytomy. Notwithstanding ILS, different classes of data are also likely undergoing different selection pressures [62,63], which could contribute to the conflicting but strongly supported alternate topologies.

Another possible cause of discordance is hidden paralogy stemming from ancestral genome or gene duplications and subsequent losses in different gene copies resulting in single gene copies that do not reflect its genealogy [64,65]. We assessed hidden paralogy (described in electronic supplementary material) and did not uncover evidence of genome duplications nor large amounts of gene duplications in Rhacophoridae. Instead, most detectable paralogues were novel gene duplications in extant taxa not shared with other taxa (electronic supplementary material, table S5). Because of the small number of paralogues overall, hidden paralogy is unlikely to drive widespread patterns of discordance. However, without complete genomes, we were unable to rule out completely the possibility that hidden paralogy could be present, such as in cases where a duplication occurs and different duplicates are lost resulting in a single copy for all extant species.

Overall, our systematic analyses of different classes of data support the hypothesis that ILS (caused by rapid diversification events) is likely one of the dominant underlying factors responsible for most of the deep-level discordances in Old World treefrogs. Thus, our ability to establish direct links between phylogenetic uncertainty, ILS and the anomaly zone provides valuable insights into how underlying microevolutionary processes can complicate species tree estimation, despite using large amounts and different classes of genomic data.

Data accessibility. Raw sequence reads are available at the GenBank SRA: BioProject PRJNA633673 (outgroup samples); BioProject PRJNA659075 (ingroup samples; raw sequences to be uploaded upon acceptance of manuscript). FrogCap bioinformatic scripts are available at: https://github.com/chutter/FrogCap-Sequence-Capture. Custom scripts used for analyses are available at: https://github.com/chankinonn-/Phylogenomic-scripts. All relevant alignments, gene trees, consensus/species trees and partitioning files that are required to reproduce this study have been uploaded to the Dryad Digital Repository: https://doi.org/10.5061/dryad.8cz8w9gn7 [66].

Authors' contributions. K.O.C. designed the study, carried out laboratory molecular work, performed analysis and wrote the manuscript; C.R.H. designed the probeset and bioinformatics pipeline, provided bioinformatic support and critically revised the manuscript; P.L.W. helped with molecular laboratory work and revised the manuscript; L.L.G. provided tissue samples; R.M.B. critically revised the manuscript and provided funding for molecular work.

Competing interests. We declare we have no competing interests.

Funding. We received no funding for this study.

Acknowledgements. Illumina sequencing was partially funded by NSF support to R.M.B. (DEB 1654388) and our fieldwork was supported by DEB 0743491 and 1557053. We thank La Sierra University, The Field Museum of Natural History (Chicago), Museum of Vertebrate Zoology (Berkeley, California), California Academy of Sciences and Jodi Rowley (Australian Museum) for providing tissues. This paper is contribution number 932 of the Auburn University Museum of Natural History. We thank the Philippine Department of the Environment and Natural Resources, the Solomon Islands Forestry

Department, Madagascar's Ministry of the Environment, Forests and Tourism, and Thailand's Ministry of Natural Resources and Environment for issuing collecting and export permits to R.M.B. and colleagues at KU. We thank the Malagasy authorities for approving research permits (no. 298/13/MEF/SG/DGF/DCB.SAP/SCBSE, no. 303/14/MEEMF/SG/DGF/DAPT/SCBT, no. 329/15/MEEMF/SG/DGF/DAPT/SCBT); specimens were exported under permits (no. 017N-EV01/MG14, no. 055NEA02/MG15, no. 041N-EA01/MG16).

References

1. Hillis DM, Bull JJ. 1993 An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* **42**, 182–192. (doi:10.1093/sysbio/42.2.182)
2. Wilcox TP, Zwickl DJ, Heath TA, Hillis DM. 2002 Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Mol. Phylogenet. Evol.* **25**, 361–371. (doi:10.1016/S1055-7903(02)00244-0)
3. Yang Z, Zhu T. 2018 Bayesian selection of misspecified models is overconfident and may cause spurious posterior probabilities for phylogenetic trees. *Proc. Natl. Acad. Sci.* **115**, 1854–1859. (doi:10.1073/pnas.1712673115)
4. Kumar S, Filipski AJ, Battistuzzi FU, Kosakovsky Pond SL, Tamura K. 2012 Statistics and truth in phylogenomics. *Mol. Biol. Evol.* **29**, 457–472. (doi:10.1093/molbev/msr202)
5. Roycroft EJ, Moussalli A, Rowe KC. 2019 Phylogenomics uncovers confidence and conflict in the rapid radiation of Australo-Papuan rodents. *Syst. Biol.* **69**, 431–444. (doi:10.1093/sysbio/syz044)
6. Dell'Ampio E *et al.* 2014 Decisive data sets in phylogenomics: lessons from studies on the phylogenetic relationships of primarily wingless insects. *Mol. Biol. Evol.* **31**, 239–249. (doi:10.1093/molbev/mst196)
7. Jarvis ED *et al.* 2014 Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320–1331. (doi:10.1126/science.1251385)
8. Jeffroy O, Brinkmann H, Delsuc F, Philippe H. 2006 Phylogenomics: the beginning of incongruence? *Trends Genet.* **22**, 225–231. (doi:10.1016/j.tig.2006.02.003)
9. Platt RN *et al.* 2018 Conflicting evolutionary histories of the mitochondrial and nuclear genomes in New World *Myotis* bats. *Syst. Biol.* **67**, 236–249. (doi:10.1093/sysbio/syx070)
10. Reddy S *et al.* 2017 Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. *Syst. Biol.* **66**, 857–879. (doi:10.1093/sysbio/syx041)
11. Rokas A, Williams BI, King N, Carroll SB. 2003 Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* **425**, 798–804. (doi:10.1038/nature02053)
12. Chan KO, Alexander AM, Grismer LL, Su Y-C, Grismer JL, Quah ESH, Brown RM. 2017 Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian torrent frogs. *Mol. Ecol.* **26**, 5435–5450. (doi:10.1111/mec.14296)
13. Hosner PA, Faircloth BC, Glenn TC, Braun EL, Kimball RT. 2016 Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). *Mol. Biol. Evol.* **33**, 1110–1125. (doi:10.1093/molbev/msv347)
14. Huang H, Knowles LL. 2016 Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Syst. Biol.* **65**, 357–365. (doi:10.1093/sysbio/syu046)
15. Molloy EK, Warnow T. 2017 To include or not to include: the impact of gene filtering on species tree estimation methods. *Syst. Biol.* **67**, 285–303. (doi:10.1093/sysbio/syx077)
16. Mendes FK, Hahn MW. 2018 Why concatenation fails near the anomaly zone. *Syst. Biol.* **67**, 158–169. (doi:10.1093/sysbio/syx063)
17. Mirarab S, Bayzid MS, Warnow T. 2016 Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. *Syst. Biol.* **65**, 366–380. (doi:10.1093/sysbio/syu063)
18. Roch S, Steel M. 2015 Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. *Theor. Popul. Biol.* **100**, 56–62. (doi:10.1016/j.tpb.2014.12.005)
19. Léveillé-Bourret É, Starr JR, Ford BA, Moriarty Lemmon E, Lemmon AR. 2018 Resolving rapid radiations within angiosperm families using anchored phylogenomics. *Syst. Biol.* **67**, 94–112. (doi:10.1093/sysbio/syx050)
20. Meiklejohn KA, Faircloth BC, Glenn TC, Kimball RT, Braun EL. 2016 Analysis of a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some multispecies coalescent methods. *Syst. Biol.* **65**, 612–627. (doi:10.1093/sysbio/syw014)
21. Hahn MW. 2007 Bias in phylogenetic tree reconciliation methods: implications for vertebrate genome evolution. *Genome Biol.* **8**, R141. (doi:10.1186/gb-2007-8-7-r141)
22. Maddison WP. 1997 Gene trees in species trees. *Syst. Biol.* **46**, 523–536. (doi:10.1093/sysbio/46.3.523)
23. Degnan JH, Rosenberg NA. 2009 Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340. (doi:10.1016/j.tree.2009.01.009)
24. Xi Z, Liu L, Davis CC. 2015 Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. *Mol. Phylogenet. Evol.* **92**, 63–71. (doi:10.1016/j.ympev.2015.06.009)
25. Huang H, Knowles LL. 2009 What is the danger of the anomaly zone for empirical phylogenetics? *Syst. Biol.* **58**, 527–536. (doi:10.1093/sysbio/syp047)
26. Linkem CW, Minin VN, Leaché AD. 2016 Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Syst. Biol.* **65**, 465–477. (doi:10.1093/sysbio/syw001)
27. Cloutier A, Sackton TB, Grayson P, Clamp M, Baker AJ, Edwards S V. 2019 Whole-genome analyses resolve the phylogeny of flightless birds (Palaeognathae) in the presence of an empirical anomaly zone. *Syst. Biol.* **68**, 937–955. (doi:10.1093/sysbio/syz019)
28. Burbrink FT *et al.* 2020 Interrogating genomic-scale data for squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. *Syst. Biol.* **69**, 502–520. (doi:10.1093/sysbio/syz062)
29. Frost DR. 2020 Amphibian species of the world: an online reference. Version 6.0 (accessed 21 April 2020). See http//research.amnh.org/herpetology/amphibia/index.html. New York, USA: Am. Museum Nat. Hist..
30. Hertwig ST, Schweizer M, Das I, Haas A. 2013 Diversification in a biodiversity hotspot—the evolution of Southeast Asian rhacophorid tree frogs on Borneo (Amphibia: Anura: Rhacophoridae). *Mol. Phylogenet. Evol.* **68**, 567–581. (doi:10.1016/j.ympev.2013.04.001)
31. Nguyen TT, Matsui M, Eto K. 2015 Mitochondrial phylogeny of an Asian tree frog genus *Theloderma* (Anura: Rhacophoridae). *Mol. Phylogenet. Evol.* **85**, 59–67. (doi:10.1016/j.ympev.2015.02.003)
32. Chan KO, Grismer LL, Brown RM. 2018 Comprehensive multi-locus phylogeny of Old World tree frogs (Anura: Rhacophoridae) reveals taxonomic uncertainties and potential cases of over- and underestimation of species diversity. *Mol. Phylogenet. Evol.* **127**, 1010–1019. (doi:10.1016/j.ympev.2018.07.005)
33. Chen JM *et al.* 2020 An integrative phylogenomic approach illuminates the evolutionary history of Old World tree frogs (Anura: Rhacophoridae). *Mol. Phylogenet. Evol.* **145**, 106724. (doi:10.1016/j.ympev.2019.106724)
34. Meegaskumbura M, Senevirathne G, Biju SD, Garg S, Meegaskumbura S, Pethiyagoda R, Hanken J, Schneider CJ. 2015 Patterns of reproductive-mode evolution in Old World tree frogs (Anura, Rhacophoridae). *Zool. Scr.* **44**, 509–522. (doi:10.1111/zsc.12121)
35. Li JT, Li Y, Klaus S, Rao D-Q, Hillis DM, Zhang Y-P. 2013 Diversification of rhacophorid frogs provides

evidence for accelerated faunal exchange between India and Eurasia during the Oligocene. *Proc. Natl. Acad. Sci. USA* **110**, 3441–3446. (doi:10.1073/pnas.1300881110)

36. Pyron AR, Wiens JJ. 2011 A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Mol. Phylogenet. Evol.* **61**, 543–583. (doi:10.1016/j.ympev.2011.06.012)
37. Poyarkov NA, Orlov NL, Moiseeva A V, Galoyan EA, Nguyen TT, Gogoleva SS. 2015 Sorting out moss frogs: mtDNA data on taxonomic diversity and phylogenetic relationships of the Indochinese species of the genus *Theloderma* (Anura, Rhacophoridae). *Russ. J. Herpetol.* **22**, 241–280.
38. Sivongxay N, Davankham M, Phimmachak S, Phoumixay K, Stuart BL. 2016 A new small-sized *Theloderma* (Anura: Rhacophoridae) from Laos. *Zootaxa* **4147**, 433–442. (doi:10.11646/zootaxa.4147.4.5)
39. Rowley JJL, Le DTT, Hoang HD, Dau VQ, Cao TT. 2011 Two new species of Theloderma (Anura: Rhacophoridae) from Vietnam. *Zootaxa* **3098**, 1–20. (doi:10.11646/zootaxa.3098.1.1)
40. Hutter CR, Cobb KA, Portik DM, Travers SL, Wood PL, Brown RM. 2019 FrogCap: a modular sequence capture probe set for phylogenomics and population genetics for all frogs, assessed across multiple phylogenetic scales. *bioRxiv* **825307**. (doi:10.1101/825307)
41. Feng Y-J, Blackburn DC, Liang D, Hillis DM, Wake DB, Cannatella DC, Zhang P. 2017 Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. *Proc. Natl. Acad. Sci. USA* **114**, E5864–E5870. (doi:10.1073/PNAS.1704632114)
42. Borowiec ML. 2016 AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* **4**, e1660. (doi:10.7717/peerj.1660)
43. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015 IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274. (doi:10.1093/molbev/msu300)
44. Abadi S, Azouri D, Pupko T, Mayrose I. 2019 Model selection may not be a mandatory step for phylogeny reconstruction. *Nat. Commun.* **10**, 934. (doi:10.1038/s41467-019-08822-w)
45. Chan KO, Hutter CR, Wood PL, Grismer LL, Brown RM. 2020 Larger, unfiltered datasets are more effective at resolving phylogenetic conflict: introns, exons, and UCEs resolve ambiguities in golden-backed frogs (Anura: Ranidae; genus *Hylarana*). *Mol. Phylogenet. Evol.* **151**, 106899. (doi:10.1016/j.ympev.2020.106899)
46. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Le SV. 2017 UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522. (doi:10.1093/molbev/msx281)
47. Xia X. 2018 DAMBE7: new and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* **35**, 1550–1552. (doi:10.1093/molbev/msy073)
48. Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018 ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinf.* **19**, 15–30. (doi:10.1186/s12859-018-2129-y)
49. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. (doi:10.1038/nmeth.4285)
50. Chifman J, Kubatko L. 2014 Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**, 3317–3324. (doi:10.1093/bioinformatics/btu530)
51. Minh BQ, Hahn MW, Lanfear R. 2020 New methods to calculate concordance factors for phylogenomic datasets. *Mol. Biol. Evol.* **37**, 2727–2733. (doi:10.1093/molbev/msaa106)
52. Allman ES, Degnan JH, Rhodes JA. 2011 Identifying the rooted species tree from the distribution of unrooted gene trees under the coalescent. *J. Math. Biol.* **62**, 833–862. (doi:10.1007/s00285-010-0355-7)
53. Sayyari E, Whitfield JB, Mirarab S. 2018 DiscoVista: interpretable visualizations of gene tree discordance. *Mol. Phylogenet. Evol.* **122**, 110–115. (doi:10.1016/j.ympev.2018.01.019)
54. Huson DH, Klopper T, Lockhart PJ, Steel MA. 2005 Reconstruction of reticulate networks from gene trees. In *Proceedings of the Ninth International Conference on Research in Computational Molecular Biology*, pp. 233–249. Berlin, Germany: Springer.
55. Sayyari E, Mirarab S. 2018 Testing for polytomies in phylogenetic species trees using quartet frequencies. *Genes* **9**, 132. (doi:10.3390/genes9030132)
56. Rosenberg NA. 2013 Discordance of species trees with their most likely gene trees: a unifying principle. *Mol. Biol. Evol.* **30**, 2709–2713. (doi:10.1093/molbev/mst160)
57. Warnow T. 2015 Concatenation analyses in the presence of incomplete lineage sorting. *PLoS Curr. Tree Life* **7**, 1–10.
58. Chou J, Gupta A, Yaduvanshi S, Davidson R, Nute M, Mirarab S, Warnow T. 2015 A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics* **16**, 1–11. (doi:10.1186/1471-2164-16-S10-S2)
59. Felsenstein J. 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791. (doi:10.1111/j.1558-5646.1985.tb00420.x)
60. Soltis PS, Soltis DE. 2003 Applying the bootstrap in phylogeny reconstruction. *Stat. Sci.* **18**, 256–267. (doi:10.1214/ss/1063994980)
61. Streicher JW, Schulte JA, Wiens JJ. 2016 How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in Iguanian lizards. *Syst. Biol.* **65**, 128–145. (doi:10.1093/sysbio/syv058)
62. Bravo GA *et al.* 2019 Embracing heterogeneity: coalescing the tree of life and the future of phylogenomics. *PeerJ* **2019**, 1–60. (doi:10.7717/peerj.6399)
63. Shen XX, Hittinger CT, Rokas A. 2017 Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat. Ecol. Evol.* **1**, 1–10. (doi:10.1038/s41559-017-0126)
64. Struck TH. 2013 The impact of paralogy on phylogenomic studies: a case study on annelid relationships. *PLoS ONE* **8**, e0062892. (doi:10.1371/journal.pone.0062892)
65. Siu-Ting K, Torres-Sanchez M, Mauro DS, Wilcockson D, Wilkinson M, Pisani D, O'Connell MJ, Creevey CJ. 2019 Inadvertent paralog inclusion drives artifactual topologies and timetree estimates in phylogenomics. *Mol. Biol. Evol.* **36**, 1344–1356. (doi:10.1093/molbev/msz067)
66. Chan KO, Hutter CR, Wood Jr PL, Grismer LL, Brown RM. 2020 Data from: Target-capture phylogenomics provide insights on gene and species tree discordances in Old World treefrogs (Anura: Rhacophoridae). Dryad Digital Repository. (doi:10.5061/dryad.8cz8w9gn7)

Supplementary Material

Probe design, library preparation, and sequencing

Probe design follows [1] and is summarized here. Probes were synthesized as biotinylated RNA oligos in a myBaits kit (Arbor Biosciences™, formerly MYcroarray® Ann Arbor, MI) by matching 25 publicly available transcriptomes to the *Nanorana parkeri* and *Xenopus tropicalis* genomes using the program BLAT 3.50 [2]. Duplicate matches to the genome, ensuring the removal of any potential paralogs. Additional analyses to detect hidden paralogy were performed and described below. Matching sequences were clustered by their genomic coordinates to detect presence/absence across species with the goal of including markers represented across the majority of species. To narrow the locus selection to coding regions, each cluster was matched to available coding region annotations from the *Nanorana parkeri* genome using the program EXONERATE [3]. Loci from all matching species were then aligned using MAFFT 7.313 [4], and markers were selected based on phylogenetic representation, informativeness, and other filters explained in [1]. For each marker, we used the *N. parkeri* genome sequence as the chosen design marker. To create a .fasta file of bait sequences as individual entries, we separated the selected markers into 120 bp-long sequences with 2x tiling (50% overlap among baits) using an R script. These markers have an additional bait at each end extending into the intronic region to increase the coverage and capture success of these areas. Baits were then filtered, retaining those: without sequence repeats; a GC content of 30%–50%; and baits that did not match to their reverse complement or multiple genomic regions. Additionally, 646 UCEs that contain at least 10% informative sites were included [5]. Finally, we used the myBaits-2 kit (40,040 baits) with 120mer sized baits with our chosen filtered bait sequences.

Library preparation was performed by Arbor Biosciences following the myBaits V3.1 manual and briefly follows: (1) genomic DNA was sheared to 300–500 bp; (2) adaptors were ligated to DNA fragments; (3) each library was amplified for 6 cycles using unique combinations of i7 and i5 indexing primers attached to the adapters to later identify individual samples; (4) biotinylated 120mer

RNA library baits were hybridized to the sequences; (5) target sequences were selected by adhering to magnetic streptavidin beads; (6) Enrichment incubation times ranged 18–21 hours; (7) target regions were amplified via PCR for 10 cycles; and (7) samples were pooled and sequenced on an Illumina HiSeq PE-3000 with 150 bp paired-end reads. Sequencing was performed at the Oklahoma Medical Research Foundation DNA Sequencing Facility.

Data Processing Pipeline

A bioinformatics pipeline for filtering adapter contamination, assembling contigs, and exporting alignments is scripted in R and available at (bioinformatics-pipeline_stable-v1; https://github.com/chutter/FrogCap-Sequence-Capture). Raw reads were cleaned of adapter contamination, low complexity sequences, and other sequencing artifacts using the program FASTP (default settings; [6]). Adapter-cleaned reads were then matched to a database of bacterial (human skin, ultra-pure water contamination, and other common bacteria) and other genomes (*C. elegans, Drosophilia*) to ensure that no contamination persisted in our final dataset (see Supplementary Table S3 from [1] for GenBank Accession Numbers of reference genomes). We decontaminated the adapter-cleaned reads with the program BBMAP from BBTools (https://jgi.doe.gov/data-and-tools/bbtools/), by matching cleaned reads to each reference contaminant genome (reads removed if they matched >95 percent similarity). After this step, final reads were saved separated as "cleaned-reads."

Prior to assembly, the "cleaned-reads" were further processed to decrease computational load and increase accuracy. We merged paired-end reads using BBMerge [7] from BBTools. BBMerge also fills in missing gaps between non-overlapping paired-end reads by assembling missing data from the other paired-end reads using the "Tadpole" program. Next, exact duplicates were removed if both read pairs were duplicated, using "dedupe" from BBTools. Additionally, duplicates from the set of merged paired-end contigs were removed if they were exact duplicates or were contained within another merged contig. Merged singletons and paired-end reads were assembled *de novo* using the program SPADES v.3.12 [8], which runs BAYESHAMMER [9] error correction on the reads

internally. Data were assembled using several different k-mer values (21, 33, 55, 77, 99, 127), in which orthologous contigs resulting from the different k-mer assemblies were merged. We used the DIPSPADES [10] function to assemble contigs that were polymorphic by generating a consensus sequence from both haplotypes from orthologous regions such that polymorphic sites were resolved randomly.

Consensus haplotype contigs were then matched against reference marker sequences used to design the probes separately for our three probe sets with BLAST (*dc-megablast*). Contigs were discarded if they failed to match ≥ 30% of the reference marker, and contig matches fewer than 50 bp were removed. Contig matches to a given reference marker were discarded if more than one contig matched to the marker and were overlapping. For non-overlapping matches to the same reference marker, we merged these contigs by joining them together (Ns inserted in matching positions). The final set of matching contigs were labelled with the name of each marker, followed by each sample's unique institutional identifier (i.e., the corresponding museum voucher catalog number, field collector number, etc.), and assembled in a single file to be parsed out separately for multiple sequence alignment in the next step.

Alignment and Trimming

Next, our final set of matching markers was aligned on a marker-by-marker basis using MAFFT local pair alignment (settings: max iterations = 1000; ep = 0.123; op = 3; --adjust-direction). We screened each alignment for samples ≥40% divergent from consensus sequences, which were almost always incorrectly assigned contigs. Alignments were retained if they included four or more taxa, had ≥100 bp length, and mean sample specificities (i.e., the "breadth of coverage" of the sample; see below) ≥50% across the alignment (to prevent non-overlapping segments of the alignment). We then separated alignments into two initial datasets: (1) "Exons-Only," which included only exon contigs with intronic region trimmed from each alignment using the *Nanorana* genome sequence reference exon as a guide; and (2) "All-Markers," which included the entire matching contig to the reference marker. These two sets of alignments were only externally trimmed, using a custom R script, resulting in alignments in which at least 50% of the samples have sequence data at the both alignment ends.

The Exons-Only and All-Markers alignment sets were further trimmed into usable phylogenetic analyses datasets, and data type comparisons: (1) Introns (exons were trimmed from the original contig and the two remaining intronic regions were concatenated); (2) Exons (each alignment was adjusted to be in an open-reading frame and trimmed to the largest reading frame that accommodated >90% of the sequences); (3) UCEs; and (4) "Legacy," commonly used nuclear markers from amphibian phylogenetic studies (full list of markers can be obtained from [1]).

Anomaly zone calculation

The boundary of the anomaly zone *a(x)* in a four-taxon quartet is defined by:

$$a(x) = \log\left[\frac{2}{3} + \left(\frac{3e^{2x} - 2}{18(e^{3x} - e^{2x})}\right)\right]$$

where *x* is the length of a branch in the species tree that has a descendent internal branch [11]. The species tree is in the anomaly zone if the length of the descendent internal branch is less than *a(x)* [11,12]. We used this equation to estimate the presence of the anomaly zone in four-taxon subclades by comparing consecutive, ancestor-and-descendent, internal branch lengths [11,13]. If these sets of internodes were in the anomaly zone, at least one anomalous gene tree would be expected to be present.

Paralogy detection

We assessed each sample and genus for paralogs to detect potential genome duplication events within Rhacophoridae, as these could be a source of hidden paralogy. Hidden paralogy is where different duplicate copies are lost within different lineages, resulting in a gene history that does not reflect the actual history of that gene and could appear to be ILS. Detecting genome duplications within Rhacophoridae and looking at the pattern of shared paralogs among lineages could indicate the presence of hidden paralogs; however, hidden paralogs resulting from a genome duplication that

occurred before the MRCA of Rhacophoridae with different copies were lost in different lineages resulting in all extant species having a single copy could not be detected nor ruled out hidden paralogy.

Potential paralogs were assessed using several approaches: 1) Marker and probe design explicitly excluded duplicate exon and even probe matches in the *Xenopus* and *Nanorana* genomes, excluding any potential paralogs from genome duplications that occurred across the span of most frogs; 2) for each Rhacophoridae sample, we matched the sample contigs to the targets, and assessed duplicate matches of the targets to multiple contigs, which would indicate either lineage specific or Rhacophorida-wide paralogs. We collected data on the number of paralogs per sample and genus as well as counted whether the paralogs were shared among closely related species / clades or dispersed phylogenetically. A dominant dispersed phylogenetic pattern would suggest that most paralogs are lineage-specific, and that there are not any widespread paralogs resulting from duplication events earlier in the history of Rhacophoridae. Additionally, the presence of a substantial number of paralogs would suggest a genome duplication specific to Rhacophoridae.

Our results from these analyses are summarized in Table S5 and suggest that paralogs occur at a rate ~1-6% per sample (124-885 paralogs; 3655 different paralogs), which indicates that there have not been any widespread genome or gene duplication events within Rhacophoridae. We instead show that duplication mostly occurs near the tips of the phylogeny in extant species (2029 paralogs), in a phylogenetically dispersed pattern (i.e. only 1629 paralogs shared among two or more species). Additionally, 96 paralogs are shared by the majority of members of a genus. These results do not suggest a pattern of hidden paralogy via ancestral duplication events earlier in the history of Rhacophoridae as paralogs tend to not be shared by larger clades. However, we cannot rule out that different duplicate copies were all lost in different lineages, which would be difficult to detect with current methods and available data.

Topology tests

Likelihood-based tree topology tests were conducted to assess the fit of alternative hypotheses of evolutionary relationships to the All-combined dataset. These include the approximately unbiased (AU) test [14] and tests implementing the RELL approximation [15] including bootstrap proportion (BP), Kishino-Hasegawa test [16], Shimodaira-Hasegawa test [17], and expected likelihood weights [18]. Model parameters were estimated based on the maximum likelihood tree from the All-combined dataset and the number of RELL replicates was set at 1000. We tested all unique consensus topologies (T1–T5) and results are presented in Table S3.

Supplementary figures and tables

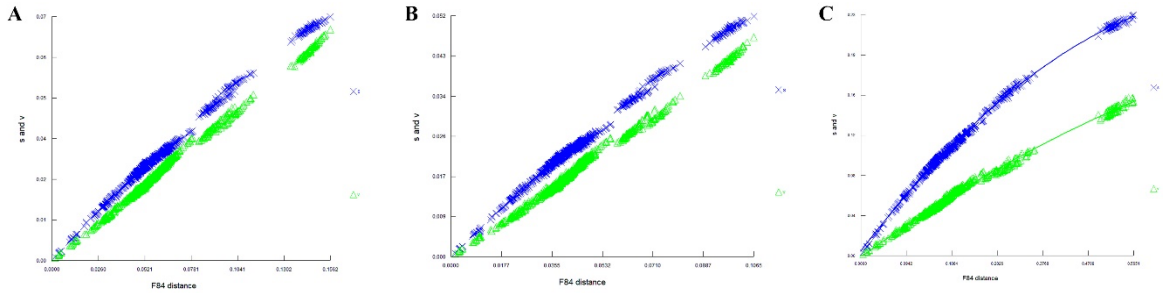

Fig. S1 Saturation plots showing transitions (blue) and transversions (green) vs. phylogenetic distance (F84 model) for the first (A), second (B), and third (C) codon positions.

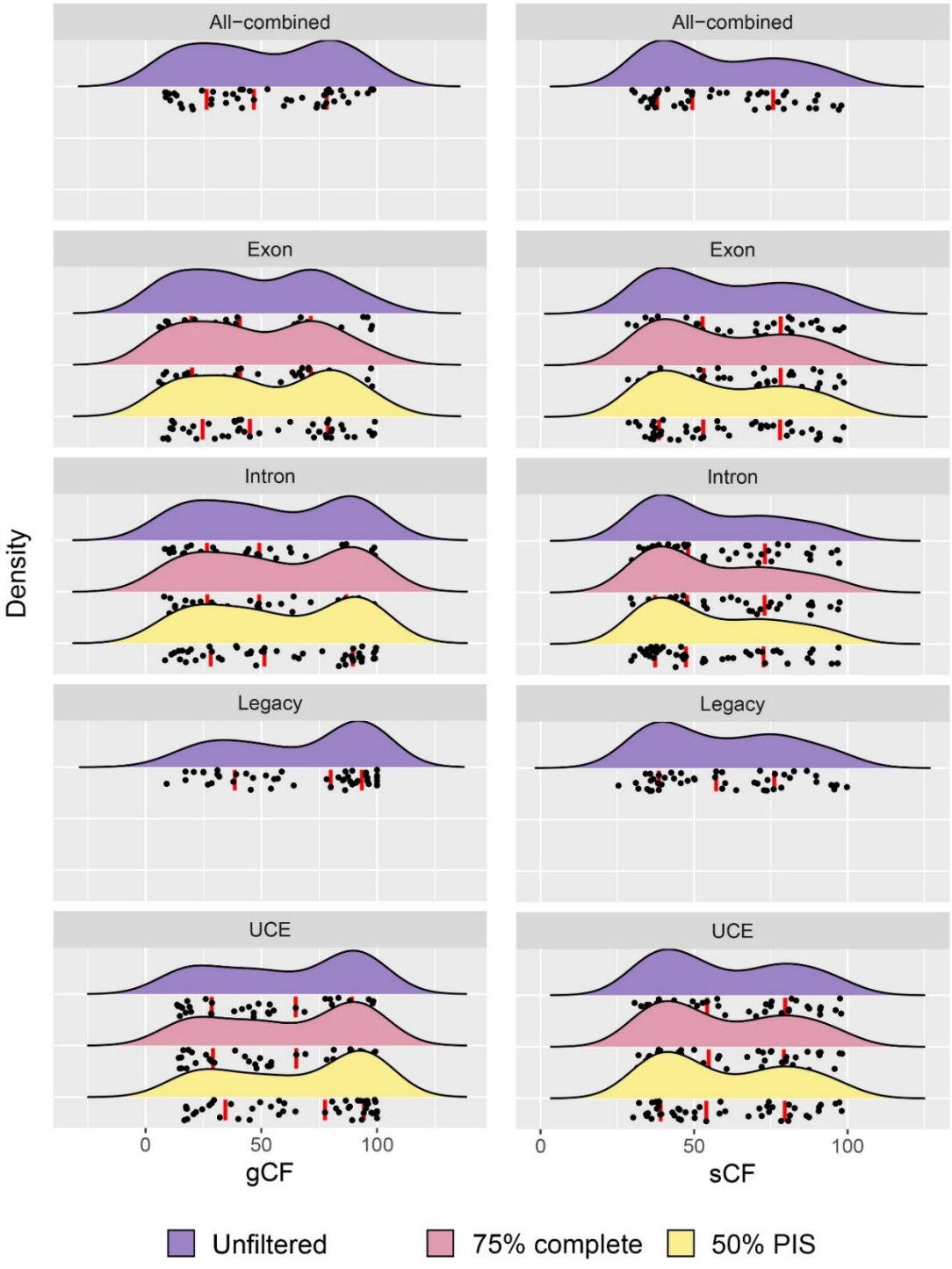

Fig. S2. Density plots of gCF and sCF values across different data types and filtering strategies. Red vertical lines represent the median, first, and third quartile. Colors represent the different filtering strategies applied to each dataset. Gene concordance (gCF) did not significantly improve when data was filtered by taxon completeness but showed marginal improvements when filtered by PIS. Site concordance (sCF) was unaffected by data filtering.

Table S1. List of samples used in the study and their associated metadata. KU = University of Kansas; LSUHC = La Sierra University Herpetological Collection (Riverside, California); FMNH = The Field Museum of Natural History (Chicago); AMH = Australian Museum (Sydney); CAS = California Academy of Sciences; MVZ = Museum of Vertebrate Zoology (University of California, Berkeley); UNS = University of New South Wales.

Species	**Voucher**	**Locality**
Outgroup		
Arthroleptis variabilis	KU 341900	Cameroon: Korup National Park, Chimpanzee Camp
Cornufer guentheri	KU 341235	Solomon Islands: Koliai Village, Taloven Camp
Boophis tephraeomystax	KU 347343	Madagascar: Ranomanfana National Park
Mantidactylus melanopleura	KU 347429	Madagascar: Marojejy National Park
Scaphiophryne marmorata	KU 343458	Madagascar: Andasibe Village
Ingroup		
Chiromantis doriae	FMNH 255213	Laos: Huaphahn province
Chiromantis doriae	JJLR 451 (AMH)	Vietnam: Pu Hoat Proposed Nature Reserve
Chiromantis nongkhorensis	AMH 48646	Cambodia: Kratie Province
Chiromantis nongkhorensis	FMNH 263098	Cambodia: Mondolkiri Province
Chiromantis rufescens	CAS 258328	Gabon: Ogooue-Lolo Province
Feihyla palpebralis	FMNH 254425	Vietnam: Gia-Lai province, Ankhe district
Feihyla vittatus	KU 328224	Thailand: Nakhon Ratchasima
Feihyla vittatus	AMH 48555	Cambodia: Kratie Province

Feihyla vittatus	KU 328245	Thailand: Khao Luang
Gracixalus carinensis	FMNH 271715	Thailand: Chiangmai, Amphoe Chom Thong
Gracixalus jinxuensis	JJLR 414 (AMH)	Vietnam: Pu Hoat Proposed Nature Reserve
Gracixalus lumarius	UNS 00340	Vietnam: Ngoc Linh Nature Reserve
Kurixalus verrucosus	FMNH 223856	Vietnam: Vinh Phuc
Nyctixalus pictus	MVZ 239460	Indonesia: Bengkulu
Nyctixalus pictus	CAS 247476	Myanmar: Tanintharyi Division
Nyctixalus spinosus	KU 337973	Philippines: Eastern Samar
Philautus abditus	AMH 48597	Vietnam: Quang Nam Province
Philautus aurantium	FMNH 233225	Malaysia: Sabah, Sipitang District
Philautus aurifasciatus	MVZ 253933	Indonesia: West Java
Philautus poecilus	ACD 489 (Tissue at KU)	Philippines: Mindanao
Polypedates leucomystax	KU 327999	Thailand: Khao Luang
Polypedates otilophus	FMNH 230836	Malaysia: Sabah
Raorchestes parvulus	AMH 48886	Vietnam: Pu Mat National Park
Raorchestes longchuanensis	CAS 241491	Myanmar: Kachin State
Rhacophorus achantharrhena	MVZ 272105	Indonesia: West Sumatra
Rhacophorus baluensis	FMNH 235957	Malaysia: Sabah
Rhacophorus barisani	MVZ 272175	Indonesia: West Sumatra
Rhacophorus bimaculatus	KU 307676	Philippines: Quezon Province
Rhacophorus bipunctatus	CAS 221351	Myanmar: Kachin State
Rhacophorus bipunctatus	FMNH 253123	Vietnam: Nghe An, Con Cuong District
Rhacophorus cf. pardalis	FMNH 235741	Malaysia: Sarawak

Rhacophorus cyanopunctatus	FMNH 267835	Indonesia: West Sumatra
Rhacophorus edentulus	MVZ 256034	Indonesia: Central Sulawesi
Rhacophorus nigropalmatus	FMNH 230901	Malaysia: Sarawak
Rhacophorus pardalis	FMNH 269175	Malaysia: Sarawak
Rhacophorus pardalis	MVZ 272185	Indonesia: West Sumatra
Rhacophorus pardalis	LSUHC 3998	Malaysia: Selangor, Kepong
Rhacophorus rhodopus	CAS 221344	China: Yunnan
Rhacophorus vampyrus	AMS R173126	Vietnam: Khanh Hoa Province
Theloderma asperum	CAS 241559	Myanmar: Shan State
Theloderma corticale	MVZ 223905	Vietnam: Vinh Phuc
Theloderma licin	MVZ 272184	Indonesia: West Sumatra
Theloderma petilus	FMNH 257902	Laos: Phongsaly Province, Phongsaly District
Theloderma stellatum	FMNH 261896	Cambodia: Mondolkiri Province, Keo Seima District
Theloderma stellatum	FMNH 267765	Cambodia: Koh Kong Province

Table S2. Results of saturation tests at the first, second, and third codon positions. OTU=operational taxonomic unit; I_{ss} = Index of substitution saturation [19]; cSym=symmetrical topology; Asym=asymmetrical topology.

Num. OTU	I_{ss}	I_{ss}.cSym	P-value	I_{ss}.cAsym	P-value
First codon:					
4	0.074	0.850	0.0000	0.840	0.0000
8	0.084	0.845	0.0000	0.764	0.0000
16	0.087	0.830	0.0000	0.676	0.0000
32	0.092	0.810	0.0000	0.561	0.0000
Second codon:					
4	0.050	0.850	0.0000	0.840	0.0000
8	0.053	0.845	0.0000	0.764	0.0000
16	0.056	0.830	0.0000	0.676	0.0000
32	0.059	0.810	0.0000	0.561	0.0000
Third codon:					
4	0.229	0.85	0.0000	0.84	0.0000
8	0.219	0.845	0.0000	0.764	0.0000
16	0.247	0.83	0.0000	0.676	0.0000
32	0.258	0.81	0.0000	0.561	0.0000

Iss < Iss.c + P < 0.05 = Little saturation

Iss < Iss.c + P > 0.05 = Substantial saturation

Iss > Iss.c + P < 0.05 = Useless sequences

Iss > Iss.c + P > 0.05 = Very poor for phylogenetics

Table S3. Topology tests of all consensus topologies (T1–T5) against the All-combined alignment.

Plus signs (+) denote the 95% confidence sets, while minus signs (-) denote significant exclusion.

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	p-WKH	p-WSH	c-ELW	p-AU
T1	-64605050.61	0	1 (+)	1 (+)	1 (+)	1 (+)	1 (+)	1 (+)	1 (+)
T2	-64607213.2	2162.597	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)
T3	-64607346.02	2295.413	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0.0187 (-)
T4	-64612663.06	7612.449	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)
T5	-64622696.99	17646.386	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0.0052 (-)

deltaL: logL difference from the maximal loglikelihood in the set.

bp-RELL: bootstrap proportion using RELL method.

p-KH : *p*-value of one sided Kishino-Hasegawa test.

p-SH : *p*-value of Shimodaira-Hasegawa test.

c-ELW: Expected Likelihood Weight.

p-AU: *p*-value of approximately unbiased (AU) test.

Table S4. Resulting *p*-values from the chi-square test of equal gene (gEF) and site frequencies (sEF) at each focal node. P > 0.05 rejects the null hypothesis of equal frequencies, indicating that ILS is likely to be present around that node.

Dataset	gEF (N1 | N2 | N3)	sEF (N1 | N2 | N3)
All-combined	0.0097 | 0.0000 | 0.0000	0.6409 | 0.0000 | 0.7652
Exon-unfiltered	0.5778 | 0.0004 | 0.0000	0.4586 | 0.0000 | 0.6778
Exon-mis75	0.3474 | 0.0011 | 0.0000	0.9704 | 0.0000 | 0.0000
Exon-pis50	0.7263 | 0.0015 | 0.0000	0.4780 | 0.0000 | 0.2006
Intron-unfiltered	0.0033 | 0.0000 | 0.0002	0.6701 | 0.0317 | 0.2824
Intron-mis75	0.0141 | 0.0000 | 0.0047	0.1014 | 0.3296 | 0.8694
Intron-PIS50	0.0393 | 0.0000 | 0.0251	0.1465 | 0.0004 | 0.3646
UCE-unfiltered	0.9037 | 0.0066 | 0.0868	0.7820 | 0.6203 | 0.6506
UCE-mis75	0.9004 | 0.0789 | 0.3736	0.6945 | 0.7200 | 0.7239
UCE-pis50	0.6941 | 0.0796 | 0.8230	0.8657 | 0.8066 | 0.9004
Legacy	0.0325 | 0.0455 | 0.5270	0.2145 | 0.9046 | 0.6907

Table S5. Results of the paralogy detection analysis showing Mean/Min/Max paralogs of the species in each genus. MajorityClade = 50% or greater of the species in the clade share the same paralog. SingleSpecies = number of paralogs found in a single species only.

Genus	MeanParalogs	MinParalogs	MaxParalogs	MajorityClade	SingleSpecies
Chiromantis	338.5	124	635	6	96.5
Feihyla	269	269	269	1	236
Gracixalus	212.3	177	276	2	76.3
Kurixalus	329	204	454	13	198.5
Nyctixalus	245.6	147	330	6	62
Philautus	174.2	134	215	29	52.4
Polypedates	548.5	212	885	16	396.5
Raorchestes	248	248	248	16	232
Rhacophorus	328.9	190	641	4	74.9
Theloderma	232.5	152	303	3	43.7

Literature Cited

1. Hutter CR, Cobb KA, Portik DM, Travers SL, Wood PL, Brown RM. 2019 FrogCap: A modular sequence capture probe set for phylogenomics and population genetics for all frogs, assessed across multiple phylogenetic scales. *bioRxiv* **825307**. (doi:10.1101/825307)

2. Kent WJ. 2002 BLAT — The BLAST -Like Alignment Tool. *Genome Res.* **12**, 656–664. (doi:10.1101/gr.229202.)

3. Slater G, Birney E. 2005 Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31. (doi:10.1186/1471-2105-6-31)

4. Katoh K, Standley DM. 2013 MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. (doi:10.1093/molbev/mst010)

5. Alexander AM, Su YC, Oliveros CH, Olson K V., Travers SL, Brown RM. 2017 Genomic data reveals potential for hybridization, introgression, and incomplete lineage sorting to confound phylogenetic relationships in an adaptive radiation of narrow-mouth frogs. *Evolution (N. Y).* **71**, 475–488. (doi:10.1111/evo.13133)

6. Chen S, Zhou Y, Chen Y, Gu J. 2018 Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890. (doi:10.1093/bioinformatics/bty560)

7. Bushnell B, Rood J, Singer E. 2017 BBMerge – Accurate paired shotgun read merging via overlap. *PLoS One* **12**, 1–15. (doi:10.1371/journal.pone.0185056)

8. Bankevich A *et al.* 2012 SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477. (doi:10.1089/cmb.2012.0021)

9. Nikolenko SI, Alekseyev M. 2013 BAYESHAMMER: Bayesian subclustering for error correction in single cell sequencing. *BMC Genomics* **14**, S7.

10. Safonova Y, Bankevich A, Pevzner PA. 2015 DipSPAdes: Assembler for Highly Polymorphic Diploid Genomes. *J. Comput. Biol.* **22**, 528–545. (doi:10.1089/cmb.2014.0153)

11. Linkem CW, Minin VN, Leaché AD. 2016 Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Syst. Biol.* **65**, 465–477. (doi:10.1093/sysbio/syw001)

12. Degnan JH, Rosenberg N a. 2009 Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340. (doi:10.1016/j.tree.2009.01.009)

13. Rosenberg NA. 2013 Discordance of species trees with their most likely gene trees: a unifying principle. *Mol. Biol. Evol.* **30**, 2709–2713. (doi:10.1093/molbev/mst160)

14. Shimodaira H. 2002 An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508. (doi:10.1080/10635150290069913)

15. Kishino H, Miyata T, Hasegawa M. 1990 Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **31**, 151–160. (doi:10.1007/BF02109483)

16. Kishino H, Hasegawa M. 1989 Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *J. Mol. Evol.* **29**, 170–179. (doi:10.1007/BF02100115)

17. Shimodaira H, Hasegawa M. 1999 Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116. (doi:10.1016/j.arth.2012.06.015)

18. Strimmer K, Rambaut A. 2002 Inferring confidence sets of possibly misspecified gene trees. *Proc. R. Soc. B Biol. Sci.* **269**, 137–142. (doi:10.1098/rspb.2001.1862)

19. Xia X, Xie Z, Salemi M, Chen L, Wang Y. 2003 An index of substitution saturation and its application. *Mol. Phylogenet. Evol.* **26**, 1–7. (doi:10.1016/S1055-7903(02)00326-3)