Congruence and conflict in the higher-level phylogenetics of squamate reptiles: an expanded phylogenomic perspective

Authors and Affiliations

Sonal Singhal[1,2,3*], Timothy J. Colston[4,5], Maggie R. Grundler[1,2,6], Stephen A. Smith[1], Gabriel C. Costa[7], Guarino R. Colli[8], Craig Moritz[9], R. Alexander Pyron[4], Daniel L. Rabosky[1,2]

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109

[2]Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109

[3]Department of Biology, CSU Dominguez Hills, Carson, California 90747

[4]Department of Biological Sciences, The George Washington University, Washington D.C. 20052

[5]Department of Biological Science, Florida State University, Tallahassee, Florida 32306

[6]Department of Environmental Science, Policy, & Management, University of California Berkeley, Berkeley, CA 94720

[7]Department of Biology and Environmental Sciences, Auburn University at Montgomery, Montgomery, Alabama

[8]Departamento de Zoologia, Universidade de Brasília, Brasília, DF, Brazil

[9]Division of Ecology and Evolution, Research School of Biology, and Centre for Biodiversity Analysis, The Australian National University, 46 Sullivans Creek Road, Acton, ACT, 2601, Australia

* corresponding author: sonal.singhal1@gmail.com

Keywords

target capture, molecular evolution, gene tree conflict, anchored hybrid enrichment (AHE), ultraconserved elements (UCE), phylogenomic concordance

Abstract

Genome-scale data have the potential to clarify phylogenetic relationships across the tree of life, but have also revealed extensive gene tree conflict. This seeming paradox, whereby larger datasets both increase statistical confidence and uncover significant discordance, suggests that understanding sources of conflict is important for accurate reconstruction of evolutionary history. We explore this paradox in squamate reptiles, the vertebrate clade comprising lizards, snakes, and amphisbaenians. We collected an average of 5103 loci for 91 species of squamates that span higher-level diversity within the clade, which we augmented with publicly available sequences for an additional 17 taxa. Using a locus-by-locus approach, we evaluated support for alternative topologies at 17 contentious nodes in the phylogeny. We identified shared properties of conflicting loci, finding that rate and compositional heterogeneity drives discordance between gene trees and species tree and that conflicting loci rarely overlap across contentious nodes. Finally, by comparing our tests of nodal conflict to previous phylogenomic studies, we confidently resolve nine of the 17 problematic nodes. We suggest this locus-by-locus and node-by-node approach can be used to build consensus on which topological resolutions remain uncertain in phylogenomic studies of other contentious groups.

Introduction

Phylogenomic analyses face several major challenges. Because large datasets are used to generate these trees, many nodes in a tree often have strong statistical support (Rokas and Carroll 2006), whether measured by bootstrap or posterior probability metrics. However, this support is somewhat illusory, because alternative datasets and inference methods can yield strongly discordant results. Notable examples include the placement of ctenophores within animals (Pisani et al. 2015; Whelan et al. 2015, 2017) and relationships among bird families (Jarvis et al. 2014; Prum et al. 2015). In both cases, phylogenies were inferred with millions of sites, and most nodes in a given analysis were strongly statistically supported. Yet, some of these strongly supported nodes conflict with each other across datasets and analytical methods, suggesting that these estimates of statistical support might be inflated for some nodes (Cummings et al. 2003; Jeffroy et al. 2006). In addition – and somewhat paradoxically – phylogenomic datasets sometimes fail to provide additional resolution for some contentious nodes, despite massive amounts of data (Philippe et al. 2011).

To address these challenges, we can instead interrogate support for alternative phylogenetic hypotheses using a locus-by-locus approach (Brown and Thomson 2016; Arcila et al. 2017; Shen et al. 2017; Smith et al. 2020; Walker et al. 2018). Given the assumption of a single underlying species tree, this approach explicitly measures levels of conflict among gene trees and attempts to determine its potential causes. Researchers can then filter loci or use more sophisticated analytical methods (i.e., modeling introgression across tips, Wen et al. 2018) to better resolve nodes with high levels of conflict.

Conflict among gene trees can result from both biological processes and methodological issues. With respect to biology, certain evolutionary histories can increase gene tree conflict,

including introgression among lineages, large or structured ancestral populations, and periods of rapid speciation (Maddison 1997; Degnan and Rosenberg 2006; Edwards 2009). Gene-tree conflict can also arise if gene trees were estimated incorrectly due to methodological issues such as undetected paralogy, model violation, or low information-content. Identifying and removing sources of gene tree estimation error can generate better-resolved phylogenies (Jeffroy et al. 2006; Salichos and Rokas 2013; Doyle et al. 2015). However, such filtering approaches cannot ameliorate gene tree incongruence that results from biological processes (but see Knowles et al. 2018). Instead, we must evaluate what these conflicts tell us about our confidence in a given node as well as the processes that have led to conflict in the first place.

Here, we apply a locus-by-locus approach to understand gene tree conflict in Squamata, the vertebrate clade comprising lizards, snakes and amphisbaenians. This clade includes over 10,000 species and exhibits striking instances of evolutionary convergence, with multiple independent origins of viviparity, parthenogenesis, limblessness, sex chromosomes, and venom production (Uetz and Stylianou 2018, Fry et al. 2006; Brandley et al. 2008; Kearney et al. 2009; Pyron and Burbrink 2014; Gamble et al. 2015). This group has been subject to three recent, wide-ranging phylogenomic studies (Burbrink et al. 2020; Streicher and Wiens 2016, 2017), all of which clarified key relationships among clades and identified topological relationships that remain uncertain. Building on these studies, we provide a consensus view on higher-level squamate phylogenetics by assessing conflict and congruence across thousands of independent loci, conducting targeted tests of support across high-conflict nodes, and identifying the shared properties of conflicting loci. In doing so, we show how our locus-by-locus and node-by-node approach can help focus attention on which phylogenetic relationships remain uncertain.

Methods

Sampling, Data Acquisition, and Data Processing

We used both newly-collected and previously-published genome-wide sequence data to infer a family-level phylogeny for squamate reptiles. We sequenced 92 target samples, prioritizing samples that were linked to vouchered museum specimens. We addressed key gaps in our phylogenetic sampling by further including 17 samples from previously-published phylogenomic studies (Leaché et al. 2015, Streicher et al. 2016, Streicher and Wiens 2016, Streicher and Wiens 2017). Where possible, we downloaded the raw sequence data associated with these samples and processed them similarly to newly collected data. In total, we included 109 samples across 108 species, spanning 58 of the 67 squamate families (Table S1). Most families are represented by two species that span the phylogenetic breadth of the family. Our taxonomy follows Uetz *et al.* 2018.

We used a target capture approach to sequence 5,462 phylogenomic loci per newly-collected sample (SqCL marker set; Singhal et al. 2017). This marker set consists of three loci types, all commonly used in vertebrate phylogenomics: 372 anchored hybrid enrichment loci (AHE; Lemmon et al. 2012), 5052 ultraconserved elements (UCE; Faircloth et al. 2012b), and 38 single-copy nuclear genes (Wiens et al. 2012). The AHE and nuclear genes are conserved exons, whereas UCEs are non-exonic conserved loci. To generate these data, we first extracted DNA from either tail or liver tissue using a high-salt or phenol-chloroform DNA extraction (Aljanabi and Martinez 1997). Following Illumina protocols, the commercial services Rapid Genomics (Gainesville, FL, USA) and Arbor Biosciences (Ann Arbor, MI, USA) then prepared dual-barcoded genomic libraries from ~1.0 ng of sheared DNA. Libraries were pooled in sets of eight; pooled libraries were then used as template for standard capture reactions following the MyBaits

v3 Protocol (Arbor Biosciences; Ann Arbor, MI, USA). Following capture, libraries were pooled further and 100 libraries were sequenced per one lane of 125PE reads with the Illumina HiSeq 4000 at the University of Michigan Sequencing Core (Ann Arbor, MI, USA) and at HudsonAlpha (Huntsville, AL, USA).

We processed sequenced reads as follows; full details are available at Singhal et al. (2017). Following demultiplexing, we removed adaptor sequence using Trimmomatic v0.36 and merged overlapping reads with PEAR v0.9.6 (Zhang et al. 2013; Bolger et al. 2014). We used Trinity v2.3.2 to assemble reads and blat v36x1 to annotate assemblies (Kent 2002; Grabherr et al. 2011). To call variants per individual, we aligned trimmed reads using bwa v0.7.17 and called genotypes using GATK v3.4 (Li 2013, Van der Auwera et al. 2013). For use as outgroups, we used BLAST v2.2.29 and samtools v1.3 to extract our target loci from the human (hg38), chicken (galGal2), turtle (chrPic1), zebra finch (taeGut2), and alligator (allMis1) reference genomes (Altschul et al. 1997; Li et al. 2009).

Phylogenetic Inference

We inferred a phylogeny across species using both a coalescent-based approach (ASTRAL-III v5.5.9; Zhang et al. 2018) and concatenated approach (ExaML v3.0.19; Kozlov et al. 2015). First, we generated locus-specific alignments using mafft v7.294 (Katoh and Standley 2013). We removed any alignments that sampled <5% of individuals and then trimmed the remaining alignments to remove any individual sequences that were <300 bp and any sites that were >70% missing.

To generate a coalescent-based tree, we used RAxML v8.2.8 under the rapid hill-climbing algorithm to infer a gene tree for each locus under the GTRGAMMA model

(Stamatakis 2014). To evaluate support for each gene tree, we calculated Shimodaira–Hasegawa (SH)-like values per node. We then collapsed all gene tree nodes with <10 SH-like support, resulting in an average of 9% of nodes collapsed. We used ASTRAL-III to infer a phylogeny across these gene trees.

To infer a concatenated phylogeny, we used ExaML under the CAT model. We generated 100 bootstraps by randomly subsampling 5% of the loci in the original alignment and then inferring topology with ExaML. Because bootstrapping values were uniformly high even with a small subsample and because this subsampling strategy was computationally efficient, we did not explore alternative subsampling strategies.

We then inferred both a concatenated and coalescent phylogeny using an AHE-only or UCE-only alignment, because marker type has been shown to affect phylogenetic inference (Jarvis et al. 2014; Reddy et al. 2017). We did not analyze an alignment of traditional phylogenetic genes only due to its small sample size. Then, we identified nodes that differed among inferred trees using phyparts v0.0.1 (Smith et al. 2015). phyparts identifies concordant nodes as those that share the same set of descendants; all other nodes are discordant.

Finally, a major source of gene tree conflict can be topologies that fall into the anomaly zone, the parameter space in which gene trees are more likely to be discordant with the species tree than concordant (Degnan and Rosenberg 2006). Using scripts provided by Linkem et al. (2016), we calculated the limit of the anomaly zone for each pair of parent-child internodes (equation 4 in Degnan and Rosenberg 2006). If the descendant internal branch is shorter than the limit, this branch falls into the anomaly zone. We calculated internal branch lengths in coalescent units based on the ASTRAL-III tree.

Testing Phylogenetic Conflicts

We identified uncertain nodes in the family-level phylogeny for subsequent interrogation using several approaches. First, we identified nodes that have been resolved inconsistently across different studies (Wiens et al. 2012; Pyron et al. 2013; Streicher and Wiens 2017). In addition, we considered nodes that have been historically contentious, such as the placement of Iguania (as summarized in Losos et al. 2012). Second, we identified nodes that conflicted across the phylogenies inferred in this study (Fig. 1, Fig. S1). Third, we identified common conflicting topologies across gene trees. To do so, we used bp to compare rooted gene trees to the concatenated phylogeny (Smith et al. 2020). For every node, bp outputs all conflicting topologies found in the gene trees, ranked by frequency. We then manually reviewed this output to both identify high conflict nodes and their alternate topological resolutions. Through these three approaches, we selected 17 relationships for further investigation; each had two to four alternate topological resolutions (see Table 1).

We used two complementary approaches to evaluate support for alternative topological resolutions across our 17 putatively uncertain nodes. First, we measured levels of gene tree conflict using bp. For a given node, if the gene tree and species tree have different descendants, bp will classify the gene tree as conflicting. We measured conflict using gene trees that were outgroup rooted and for which all nodes with <80 SH-like values were collapsed. Second, we measured the difference in log-likelihoods for a given locus across all alternate topologies, as introduced by Smith et al. (2020). Per node and locus, we calculated the log-likelihood under each alternate topology by specifying these topologies as constraints in RAxML. We then collated all likelihoods across all topological resolutions and took the difference between the two largest likelihoods as D_{LNL}. D_{LNL} is thus an estimate of the extent to which a particular

topological resolution is favored over the next-best topological resolution for a given locus and node. Then, per topology, we summed D_{LNL} values across the loci that best supported that topological resolution. The summed D_{LNL} thus tells us the total weight of evidence favoring the focal topology; this metric quantifies how strongly (summed D_{LNL} large) or weakly (summed D_{LNL} small) a set of loci favors a particular topology. Similar to other measures of nodal support based on likelihood (e.g., Shen et al. 2017), the D_{LNL} approach does not account for how demographic parameters affect the likelihood of a gene tree given a species tree and thus might fail in situations like the anomaly zone (Degnan and Rosenberg 2006).

Shared properties of conflicting loci

The properties of a given locus affect phylogenetic inference and thus levels of gene tree conflict (Jeffroy et al. 2006). Accordingly, we calculated 14 summary statistics that characterized the loci's overall data quality and patterns of molecular evolution (Table 2). We measured levels of missing data (missingness and occupancy), informativeness (locus length, total tree length, average SH-like value, and two metrics related to phylogenetic informativeness [PI]), heterogeneity (nucleotide compositional heterogeneity, root-tip variance, and residuals of root-tip length against root-tip node depth), quality (heterozygosity, number of long branches), GC content, and saturation C value (Kück and Struck 2014, Townsend 2007). To calculate phylogenetic informativeness, we calibrated the concatenated phylogeny using treePL (Smith and O'Meara 2012) and fossil and secondary calibrations from Irisarri et al. 2017 and then estimated PI using TAPIR (Faircloth et al. 2012a).

To determine what shared properties of loci might drive conflict, we conducted five analyses. Across all these analyses, we used the D_{LNL} results to categorize loci as either

conflicting or supporting. First, per metric and per putatively contentious relationship, we calculated the mean difference between loci that supported the most-preferred topology vs. those that conflicted. We then generated 1000 non-parametric bootstraps and calculated the difference for each of these scrambled datasets. We calculated significance as the number of bootstraps in which the absolute difference was greater than the observed difference. Second, we determined which locus-level properties might explain the level of conflict between the gene tree and the species tree. Here, we measured the level of conflict as the difference in log-likelihoods of an unconstrained gene tree vs. one constrained to the concatenated species tree. Before conducting correlations, we took the residuals of all metrics and log-likelihoods against 'tree length'. Third, we correlated patterns of D_{LNL} values across all pairwise combinations of our 17 putatively contentious nodes. Fourth, we determined if the identity of conflicting loci overlap more across topological resolutions than would be expected by random chance. To calculate the percent overlap expected under random, we scrambled the identity of conflicting vs. supporting loci for each comparison, keeping proportions constant, and then measured percent overlap across 100 bootstraps.

Finally, fifth, patterns of molecular evolution can vary across locus types. For example, UCEs contain a central conserved region and more quickly evolving flanking regions, whereas AHE exons exhibit modest levels of conservation across their entire region (Faircloth et al. 2012b; Lemmon et al. 2012, Singhal et al. 2017). To determine if locus type might affect our phylogenetic inference, we compared our locus-level metrics across all three locus types and repeated the D_{LNL} analyses for both AHE- and UCE-only data sets. We did not conduct D_{LNL} analyses with traditional phylogenetic genes because of the small sample size.

Data analysis and visualization

All code used to process genomic data and analyze data is available at <GitHubLink>. We used python v3, R v3.3.3, ape, phangorn, phytools, and cowplot to process and visualize these data (Paradis et al. 2004; Schliep 2010; Revell 2012; Wilke 2016).

Results

Phylogenetic inference

Our target capture approach was highly effective; we collected an average of 4.5 Mb of sequence across 5103 loci across our 92 individuals (Table S1). Per locus, average completeness across individuals was 92%. Our newly-generated data were of higher quality – higher coverage (mean 80×) and longer loci (mean 880 bp) – than previously-published data, likely because of greater high sequencing effort (Fig. S3).

Using these data, we inferred both coalescent-based and concatenated trees. The two trees were largely concordant but differed at several nodes, particularly with respect to family-level relationships within Iguania (Fig. 1). Given that the two trees are fairly similar and mainly disagree at known discordant nodes, we focus further analyses and discussion on the concatenated phylogeny.

The concatenated phylogeny was largely concordant with previous squamate phylogenies, whether these phylogenies were inferred with a few loci or with phylogenomic datasets (Burbrink et al. 2020; Wiens et al. 2012; Streicher and Wiens 2017). However, some inferred relationships differed. For example, in the concatenated topology, Dibamidae is sister to all non-gecko squamates (as in Townsend et al. 2004), whereas other studies have found it sister to all squamates (Pyron et al. 2013; Streicher and Wiens 2017) or sister to Gekkota (Burbrink et

al. 2020; Wiens et al. 2012; Reeder et al. 2015). Other conflicts emerged by comparing phylogenies inferred using different marker sets and different analytical methods (Fig. 1 and S1). For example, the position of Eublepharidae differs in trees inferred with AHE vs. UCE loci (Fig. S1), as seen in other studies (Townsend et al. 2004; Wiens et al. 2012; Pyron and Burbrink 2014; Reeder et al. 2015).

Testing Phylogenetic Conflicts

To more systematically evaluate conflict, we compared gene tree and species tree topologies to determine the number of gene trees that conflict at each node. Levels of support and conflict varied considerably both across clades and across clade depth (Fig. 2, Fig. S2). Although our within-family sampling was limited, monophyly of families was well-supported by the majority of gene trees (average support = 71%; Fig. S2). However, for relationships deeper than family-level, gene tree support averaged 40%. Conflict was particularly common among early branching relationships in Serpentes and Iguania; many of these branches fall into anomaly zones (Fig. S4). In fact, conflict was so rampant within Iguania that we could not identify alternate topological resolutions to test (see also Burbrink et al. 2020). Conflict was high even across nodes that had high statistical support as measured by bootstrap and local posterior probability (Fig. S5).

We then identified 17 putatively contentious nodes and used a summed log-likelihood approach to evaluate support for alternate topological resolutions at each node (Table 1). Most loci had very low D_{LNL} values (median D_{LNL} = 1.66; Fig. 3), indicating that they did not strongly distinguish amongst alternate topologies. Nonetheless, the summed D_{LNL} approach strongly resolved several uncertain nodes (Table 1, Table S2), including the historically contentious

placement of Iguania (Losos et al. 2012, see also Burbrink et al. 2020). For 10 of the 17 nodes, comparing summed D_{LNL} across topologies provided strong support for one resolution among others (Table 1). Here, we interpret a given topology as "strongly supported" when the top resolution has a summed D_{LNL} at least 50% greater than the next best topological resolution. The summed D_{LNL} approach supported the same topology found in the concatenated tree for 11 out of 17 tested nodes (Fig. 1; Table 1). Of the remaining six nodes, four of them (position of Bolyeriidae, Eublepharidae, Rhineuridae, and Xenosauridae) had fairly equivocal support across alternate topologies – i.e., alternate topological resolutions had very similar summed D_{LNL} values.

Shared properties of conflicting loci

We tested if locus-specific patterns of data quality and molecular evolution could possibly be driving conflict at nodes using five approaches. First, we compared how loci properties differed between loci that supported the preferred vs. alternate topologies. In general, supporting vs. conflicting loci were similar across most metrics, even when these differences were significant (Table 2, Table S3). In the cases where metrics differed significantly across loci supporting different topologies, typically higher quality loci – i.e., loci with less missingness, less heterogeneity – supported the preferred topology (Table 2, Table S3). Exceptions included the placement of Dibamidae, Gymnophthalmidae, and Xenosauridae, in which the best-supported topology was supported by a biased subset of lower-quality loci.

Second, we calculated the correlation between locus summary statistics and the adequacy of the concatenated topology for individual loci, finding that loci with increased compositional

298 heterogeneity and greater root-tip variance (indicative of heterotachy) showed the greatest
299 differences in likelihood (Fig. 4).
300 Third, we compared patterns of D_{LNL} values across topological tests. In general,
301 correlations in D_{LNL} values across different tests were weak; the average correlation was r =
302 0.175 (Fig. 5A). All correlations > 0.5 were between topological tests within snakes – e.g., the
303 correlation in D_{LNL} values between "position of Cylindrophiidae & Uropeltidae" and "position of
304 Anomalepididae".
305 Fourth, we determined if the identity of conflicting loci overlap more across topological
306 resolutions than would be expected by chance, finding no more or less overlap than expected
307 under random (Fig. 5B). Together, this result and the D_{LNL} correlations suggest little consistency
308 in which loci conflict across different nodes.
309 Fifth, we repeated the summed D_{LNL} tests with AHE loci only, finding patterns in
310 agreement with the full data set at 11 of the 17 contentious nodes (Table S4). Of the remaining
311 six, the summed D_{LNL} values across topological resolutions were similar, suggesting that the D_{LNL}
312 test was inconclusive. Finally, the AHE markers generally exhibited less conflict with the species
313 tree, had less missing data and were more informative, and showed less evidence of
314 heterogeneity (Fig. 6, S6).
315
316 **Discussion**
317 *Squamate phylogenomics*
318 Our 5,343-locus phylogeny captures 86% of the family-level diversity in squamate
319 reptiles and recapitulates many of the same relationships identified by studies with more taxa and
320 fewer loci (Pyron et al. 2013; Tonini et al. 2016) and similar phylogenomic datasets (Burbrink et

al. 2020; Streicher and Wiens 2017). Many of the differences between our tree and previously-published trees – for example, relationships among gecko families, placement of Xenosauridae, placement of Dibamidae – have shown instability across studies that either sample different loci and taxa and / or use different analytical methods. We replicate this pattern of discordance in our study, finding topological differences across trees inferred using concatenated versus coalescent-based methods (Fig. 1), as well as for UCE versus AHE loci only (Fig. S1). Given that levels of gene tree conflict are high for most nodes in the phylogeny (Fig. 2, S2), this discordance across datasets and studies is perhaps unsurprising.

We explored 17 putatively contentious nodes in detail. Some of these are nodes that have low statistical support, some are nodes that have alternate topologies depending on the dataset and analytical method used, and others show extensive gene tree conflict. By comparing summed D_{LNL} values across topological resolutions, we could strongly resolve 10 of these 17 nodes, (Table 1). However, although the placements of Dibamidae and Gymnophthalmidae were strongly resolved, they should remain open questions. For both, support for the preferred topology is partially driven by markers with greater data missingness and more heterogeneity, and for Dibamidae, relatively few markers were sampled.

Further, eight of these ten strongly supported topological resolutions were also recovered in the concatenated phylogeny. The exceptions are the placements of Dibamidae and Anomalepididae. Anomalepididae, along with Leptotyphlopidae and Typhlopidae, constitutes the blind snakes, a group of fossorial snakes with reduced eyes. Most phylogenetic studies have placed Anomalepididae as either sister to all snakes or sister to all non-blind snakes (Streicher and Wiens 2016). In all inferred phylogenies (Fig. 1, S1, S2), we recover Anomalepididae as sister to all non-blind snakes, which would suggest the ancestor of all snakes likely resembled

blind-snakes (Bellairs and Underwood 1951). In contrast, our D_{LNL} results recover Anomalepididae as sister to other blind snakes (Table 1), as found in phylogenetic studies that consider morphological data (Hsiang et al. 2015). However, our D_{LNL} analysis based solely on AHE loci weakly supports Anomalepididae as sister to all non-blind snakes (Table S4). Supporting versus conflicting loci for Anomalepididae are similar across all measured metrics (Table S3); thus, this discrepancy between our topologies and D_{LNL} results might result from variance at some other unmeasured metric of the sampled loci (e.g., gappiness of alignment).

Comparing our results to other phylogenomic analyses (Burbrink et al. 2020, Streicher and Wiens 2016 & 2017), we can build consensus on which relationships in the squamate phylogeny remain uncertain. These three studies and ours employ different sampling, similar marker sets (either AHEs or UCEs or both), and different approaches to inferring nodal support (bootstrap, local posterior probability, or locus-by-locus approaches). Thus, they can be regarded as semi-independent studies. Summarizing these studies suggests that nine of the 17 putatively contentious nodes in Squamata have been resolved (Table 3). Most notable among these eight nodes is the placement of Iguania, which has been historically contentious (Losos et al. 2012). Further, like Burbrink et al. (2020), we find no evidence that biased loci drive the placement of Iguania (Table S3), as has been suggested in previous analyses (Gauthier et al. 2012, Koch et al. 2018). A few nodes – e.g., the placement of Dibamidae, the position of Eublepharidae – remain uncertain and also have low statistical support across studies (Table 3). However, we also identify a few nodes – e.g., position of Bolyeridae, position of Cylindrophiidae & Uropeltidae, position of Xenosauridae – which both have strong statistical support and conflicting topologies across previous studies. Our D_{LNL} analysis identified these nodes as having ambiguous support, even when traditional measures of support failed to capture this ambiguity. These results suggest

the power of locus-by-locus approaches to identify contentious nodes in phylogenies. Below, we explore potential causes for this conflict at these contentious nodes.

Sources of conflict

Biological sources of conflict

Gene tree conflict can arise from multiple biological sources – incomplete lineage sorting, introgression, gene duplication, or varying selective or recombination regimes across loci (Degnan and Rosenberg 2006, Maddison 1997, Duchêne et al. 2018). Of these sources of conflict, incomplete lineage sorting – particularly as it arises during rapid radiation (e.g., Cloutier et al. 2019) – most likely affects our dataset. Many of the internode distances within snakes and iguanids are very short (Fig. 1), which could reflect rapid radiations in these clades. Accordingly, we tested if any branches in our tree are in anomaly zones (Degnan and Rosenberg 2006). We found that relationships within Iguania and within the clade spanning Boidae to Pythonidae in Serpentes are in anomaly zones (Fig. S4). Both 'position of Bolyeridae' and 'position of Cylindrophiidae & Uropeltidae' within the anomaly zone in Serpentes (see Table 1), which limits our ability to interrogate these nodes using likelihood-based tests. Nonetheless, our tests of these nodes were inconclusive (Table 3). In such cases where poor resolution is driven by biological processes, phylogenetic uncertainty cannot be simply addressed through better sampling, and these relationships are likely to persist as unresolvable.

Gene tree estimation error: uninformative loci

If loci have low information content, then some nodes in the inferred gene tree can be essentially resolved randomly. This leads to extensive gene tree conflict, although this conflict does not

necessarily impact the reliability of species tree inference (Lanier et al. 2014, Blom et al. 2016). To test if uninformative loci are driving conflict, we measured locus properties that reflect information content, including SH values, tree length, locus length, and phylogenetic informativeness. Generally, we found loci with greater informativeness (greater locus length, higher SH, greater tree length) had higher concordance with our species tree (Fig. 4, see also Burbrink et al. 2020), though results across phylogenetic informativeness were mixed. In our dataset, more than 68% of our loci reached their maximum phylogenetic informativeness >100 million years ago (Fig. S7). Most of our loci should thus have adequate power to inform deeper relationships in squamates, such as family-level relationships within Iguania, many of which formed ~80 - 100 million years ago. Yet, most loci exhibit only minimal differences in log-likelihoods across competing relationships (Fig. 3), suggesting these loci might be weakly informative about these deeper nodes. Indeed, on average, only 2424 of the 5354 loci sampled offered strong support for one relationship over another ($D_{LNL} > 2$). Possibly, loci with greater information content – perhaps ones that are longer or that evolve more quickly – might be more variable in their relative likelihoods across these relationships. However, perhaps because of the low correlation of loci D_{LNL} values across nodes (Fig. 5A), we found no relationship between a locus's average D_{LNL} and our measures of loci informativeness.

Gene tree estimation error: model violation

Model violation is an important source of gene tree estimation error. We quantified several metrics of loci and their inferred trees that suggest the potential for model violation. For example, high root-tip variance might reflect rate heterogeneity across lineages, high compositional heterogeneity might reflect biased mutational process, high GC might reflect high

recombination rates (Romiguier et al. 2016), and high saturation c-values might reflect multiple mutations to the same position. In our pipeline, we implemented fairly simple models of sequence and tree evolution. Particularly for UCEs – in which there is marked spatial heterogeneity in rates of evolution across the locus – these models might be too simple which could then lead to gene tree estimation error (but see Abadi et al. 2019). Such model violation might partially explain why loci with greater rate and compositional heterogeneity showed the greatest difference between unconstrained and constrained gene tree likelihoods (Fig. 4), why many of the loci supporting alternate, less-supported topologies exhibited higher rates of rate and compositional heterogeneity (Table S3), and why AHE gene trees showed better fit to the species tree than UCE gene trees (Fig. 6).

Gene tree estimation error: poor data quality

In a phylogenomic pipeline, data quality issues can arise across multiple steps, including poor sequencing quality, mis-assemblies, and mistaken ortholog identification. These technical issues can result in messy alignments, which could include poorly aligned regions or regions with high missingness. The gene trees inferred from these alignments might then have inaccurate topologies (Wong et al. 2008) or have longer branch lengths, more branch outliers, or show higher levels of root-tip variance. Together, these sources of error can create gene tree conflict even if they do not necessarily impact species tree inference (Nute et al. 2018). We attempted to mitigate some of these quality issues by trimming alignments and requiring strict orthology identification. Yet, we still see evidence for variance across all these metrics of locus and tree quality (Fig. S7). In particular, loci with high levels of missingness and greater number of branch outliers exhibit bigger log-likelihood differences in unconstrained topologies vs. topologies

436 constrained to the species tree (Fig. 4), and we found conflicting loci were more likely to have
437 greater missingness (Table S3). Emerging tools like SpruceUp and TreeShrink (Borowiec 2019,
438 Mai and Mirarab 2018) automatically profile alignments and inferred trees, offering a promising
439 way to identify and remove low-quality samples and loci that can increase gene tree conflict.
440
441 **Comparisons across marker types**
442 Other analyses have found the type of marker – e.g., intron versus exon – can influence
443 phylogenetic inference (Jarvis et al. 2014; Reddy et al. 2017). In this study, we sequenced three
444 markers types, which are relatively similar. These markers all have relatively slow evolutionary
445 rates (Faircloth et al. 2012b; Lemmon et al. 2012), and they almost certainly evolved under a
446 history of purifying selection (Katzman et al. 2007). Despite these similarities, AHE markers
447 have less missing data, exhibit less heterogeneity, and are more informative than UCE markers
448 or genes (Fig. 6B). These locus-level properties reduce discordance between gene trees and the
449 species tree (Fig. 5). Consequently, AHE markers show smaller differences in log-likelihoods
450 between their unconstrained topologies and the topologies constrained to the species tree (Fig.
451 6A). Despite the differences in quality across marker types, an AHE-only D_{LNL} analysis returned
452 a concordantly strong resolution for nine of the ten contentious nodes resolved strongly by the
453 full dataset (Table S4).
454
455 *Phylogenomics and phylogenetic conflict*
456 In many phylogenomic studies, independent analyses of the same clade often return trees
457 that conflict with one another yet have high statistical support (e.g., Pisani et al. 2015; Whelan et
458 al. 2015, 2017). Here, we recapitulate this finding; our inferred trees have nodes that conflict

459 with those found in three other squamate phylogenomic studies (Table 3, Burbrink et al. 2020,
460 Streicher and Wiens 2016 & 2017). Several of these conflicting nodes have strong statistical
461 support, but our D_{LNL} analysis identifies these nodes as remaining uncertain – thus showing the
462 power of a locus-by-locus and node-by-node approach.
463 Further, although our trees conflict, assessments of which nodes are uncertain – and
464 which nodes remain uncertain – are robust across the locus-by-locus and node-by-node analysis
465 we conducted and one conducted by Burbrink et al. 2020. We independently designed different
466 studies to address the same question, using different marker and taxon sets and different gene-
467 wise analyses to assess support and conflict. Yet, both Burbrink et al. 2020 and our study found
468 the same pattern across the two nodes we both tested; both studies strongly supported a nested
469 relationship for Iguania and showed uncertainty in the placement of Dibamidae. Although the
470 number of shared comparisons is small, this concordance suggests this locus-by-locus and node-
471 by-node approach provides better insights into levels of support for particular topological
472 resolutions (Shen et al. 2017; Smith et al. 2020; Walker et al. 2018), relative to traditional
473 measures that use clade posterior probabilities or bootstrap proportions (see also Fig. S5). Thus,
474 this general approach of interrogating nodes might help build consensus across different
475 phylogenomic studies on which nodes are resolved and which remain uncertain. Based on this
476 consensus, future researchers could then target uncertain relationships with different locus or
477 taxon sampling or improved analytical methods.
478 Other studies have argued to filter loci to ameliorate gene tree conflict (Jeffroy et al.
479 2006; Doyle et al. 2015; Whelan et al. 2015), specifically removing loci with low information
480 content. Particularly for coalescent-based methods, less-informative loci tend to lead to less
481 accurate gene trees, which could lead to inaccurate species trees (Gatesy and Springer 2014 but

see Blom et al. 2016). Removing such loci often results in better resolved species trees. Our results suggest, however, that supporting versus conflicting loci do not dramatically differ in information content (Fig. 4, Table S3), suggesting low information content might simply increase noise rather than introducing bias.

Employing a locus-by-locus approach sidesteps this debate. Instead of removing less-informative loci, we quantified how much support a given locus has for a particular topology relative to others. Most loci show only minimal differences in likelihoods across different constrained topologies (Fig. 3), which accords with a more general finding that only a small proportion of sequenced loci can drive overall phylogenetic patterns (Brown and Thomson 2016; Shen et al. 2017). Further, different loci have power to resolve nodes in different parts of the phylogenetic tree. For example, we see little correlation in D_{LNL} values across loci for different tested relationships, even across adjacent nodes or nodes with similar splitting times (Fig. 5A). Filtering loci on general informativeness risks removing loci that might inform specific relationships (Chen et al. 2015, Dornburg et al. 2019, Smith et al. 2020). Instead, the pipeline used here, where we ensure that biased loci are not driving topological resolutions, provides an alternative approach (Table S3). Better identification and then removal of loci with poor data quality – e.g., mistaken orthology assignment, chimeric assemblies – from large phylogenomic datasets could further strengthen this approach.

Finally, traditionally, downstream phylogenetic analyses such as ancestral state reconstruction have incorporated uncertainty in topologies by sampling across bootstrapped trees or a posterior distribution. But, when inferred from phylogenomic data, bootstrap trees and posterior distributions often fail to properly capture the uncertainty inherent in evolutionary relationships (e.g., Arcila et al. 2017; Smith et al. 2020). A potential solution is to conduct

505 comparative analyses across gene trees, particularly in cases where gene tree conflict is driven by
506 biological processes (Hahn and Nakleh 2016). An additional solution might be to develop new
507 approaches for translating these alternative measures of nodal support (e.g., number of gene trees
508 supporting a given node, summed log-likelihoods) into uncertainty metrics that can then be
509 properly modeled in comparative analyses. As we collect larger and larger phylogenomic
510 datasets, such advances, along with improved methods for inferring and modeling sources of
511 conflict, will allow us to both better generate robust phylogenies and to use these phylogenies to
512 understand the evolution of life's diversity.
513

514 **Acknowledgements**

515 We gratefully acknowledge funding from: a fellowship from the David and Lucile Packard
516 Foundation to DLR, NSF 1754398 to DLR, NSF DEB-1441719 to RAP, NSF DEB-1519732 to
517 SS, CSU Dominguez Hills RSCA & EFA to SS, CAPES, CNPq, FAPDF to GRC. For technical
518 & logistical support, we thank Alison Devault & Jake Enk at Arbor Biosciences, Robbin Murrell,
519 Raquel Rivadeneira, and the staff of University of Michigan ARC TS Flux. For helpful
520 discussions, we thank Mike Harvey and Joseph Walker.
521

522 **Data availability**

523 • All scripts used in processing and analyzing the data: <GITHUBLINK>
524 • Short-read sequencing data: <SRALINK>
525 • Locus alignments and tree topologies: <DRYADLINK>
526 • Many sequenced individuals are accessioned as museum specimens (Table S1)
527

528 **Literature Cited**

529 Abadi S., Azouri D., Pupko T., Mayrose I. 2019. Model selection may not be a mandatory step
530 for phylogeny reconstruction. Nat. Commun. 10:934.
531 Aljanabi S.M., Martinez I. 1997. Universal and rapid salt-extraction of high quality genomic
532 DNA for PCR-based techniques. Nucleic Acids Res. 25:4692–4693.
533 Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. 1997.
534 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.
535 Nucleic Acids Res. 25:3389–3402.
536 Arcila D., Ortí G., Vari R., Armbruster J.W., Stiassny M.L.J., Ko K.D., Sabaj M.H., Lundberg J.,
537 Revell L.J., Betancur-R R. 2017. Genome-wide interrogation advances resolution of
538 recalcitrant groups in the tree of life. Nat. Ecol. Evol. 1:20.
539 Bellairs A. d'A, Underwood G. 1951. The origin of snakes. Biol. Rev. 26:193–237.
540 Blom M.P.K., Bragg J.G., Potter S., Moritz C. 2016. Accounting for uncertainty in gene tree
541 estimation: summary-coalescent species tree inference in a challenging radiation of
542 Australian lizards. Syst. Biol. 66:352–366.
543 Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence
544 data. Bioinformatics. 30:2114–2120.
545 Borowiec, M., 2019. Spruceup: fast and flexible identification, visualization, and removal of
546 outliers from large multiple sequence alignments. J. Open Source Softw. 4:1635.
547 Brandley M.C., Huelsenbeck J.P., Wiens J.J. 2008. Rates and patterns in the evolution of snake-
548 like body form in squamate reptiles: evidence for repeated re-evolution of lost digits and
549 long-term persistence of intermediate body forms. Evol. Int. J. Org. Evol. 62:2042–2064.
550 Brown J.M., Thomson R.C. 2016. Bayes factors unmask highly variable information content,

551 bias, and extreme influence in phylogenomic analyses. Syst. Biol. 66:517–530.
552 Burbrink F.T., Grazziotin F.G., Pyron R.A., Cundall D., Donnellan S., Irish F., Keogh J.S.,
553 Kraus F., Murphy R.W., Noonan B., Raxworthy C.J., Ruane S., Lemmon A.R., Lemmon
554 E.M., Zaher H. 2020. Interrogating genomic-scale data for Squamata (lizards, snakes, and
555 amphisbaenians) shows no support for key traditional morphological relationships. Syst.
556 Biol. 69:502–520.
557 Chen, M.Y., Liang, D. and Zhang, P., 2015. Selecting question-specific genes to reduce
558 incongruence in phylogenomics: a case study of jawed vertebrate backbone phylogeny.
559 Syst. Biol. 64:1104–1120.
560 Cloutier, A., Sackton, T.B., Grayson, P., Clamp, M., Baker, A.J. and Edwards, S.V. 2019.
561 Whole-genome analyses resolve the phylogeny of flightless birds (Palaeognathae) in the
562 presence of an empirical anomaly zone. Syst. Biol. 68:937–955.
563 Cummings M.P., Handley S.A., Myers D.S., Reed D.L., Rokas A., Winka K. 2003. Comparing
564 bootstrap and posterior probability values in the four-taxon case. Syst. Biol. 52:477–487.
565 Degnan J.H., Rosenberg N.A. 2006. Discordance of species trees with their most likely gene
566 trees. PLoS Genet. 2:e68.
567 Dornburg, A., Su, Z. and Townsend, J.P. 2019. Optimal rates for phylogenetic inference and
568 experimental Design in the era of genome-scale data sets. Syst. Biol.68:145–156.
569 Doyle V.P., Young R.E., Naylor G.J.P., Brown J.M. 2015. Can we identify genes with increased
570 phylogenetic reliability? Syst. Biol. 64:824–837.
571 Duchêne, D.A., Bragg, J.G., Duchêne, S., Neaves, L.E., Potter, S., Moritz, C., Johnson, R.N.,
572 Ho, S.Y. and Eldridge, M.D. 2018. Analysis of phylogenomic tree space resolves
573 relationships among marsupial families. Syst. Biol. 67:400–412.

574 Edwards S. V. 2009. Is a new and general theory of molecular systematics emerging? Evol.
575 63:1–19.
576 Faircloth B.C., Chang J., Alfaro M.E. 2012a. TAPIR enables high-throughput estimation and
577 comparison of phylogenetic informativeness using locus-specific substitution models. arXiv
578 Prepr. arXiv1202.1215.
579 Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C.
580 2012b. Ultraconserved elements anchor thousands of genetic markers spanning multiple
581 evolutionary timescales. Syst. Biol. 61:717–726.
582 Fry B.G., Vidal N., Norman J.A., Vonk F.J., Scheib H., Ramjan S.F.R., Kuruppu S., Fung K.,
583 Hedges S.B., Richardson M.K. 2006. Early evolution of the venom system in lizards and
584 snakes. Nature. 439:584–588.
585 Gamble T., Coryell J., Ezaz T., Lynch J., Scantlebury D.P., Zarkower D. 2015. Restriction site-
586 associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions
587 among gecko sex-determining systems. Mol. Biol. Evol. 32:1296–1309.
588 Gatesy J., Springer M.S. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees,
589 bypassed hidden support, and the coalescence/concatalescence conundrum. Mol.
590 Phylogenet. Evol. 80:231–266.
591 Gauthier J.A., Kearney M., Maisano J.A., Rieppel O., Behlke A.D.B. 2012. Assembling the
592 squamate tree of life: perspectives from the phenotype and the fossil record. Bull. Peabody
593 Museum Nat. Hist. 53:3–309.
594 Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan
595 L., Raychowdhury R., Zeng Q. 2011. Trinity: reconstructing a full-length transcriptome
596 without a genome from RNA-Seq data. Nat. Biotechnol. 29:644–652.

597 Hahn, M.W. and Nakhleh, L., 2016. Irrational exuberance for resolved species trees. Evolution,
598 70(1), pp.7–17.
599 Hsiang A.Y., Field D.J., Webster T.H., Behlke A.D.B., Davis M.B., Racicot R.A., Gauthier J.A.
600 2015. The origin of snakes: revealing the ecology, behavior, and evolutionary history of
601 early snakes using genomics, phenomics, and the fossil record. BMC Evol. Biol. 15:87.
602 Irisarri I., Baurain D., Brinkmann H., Delsuc F., Sire J.-Y., Kupfer A., Petersen J., Jarek M.,
603 Meyer A., Vences M. 2017. Phylotranscriptomic consolidation of the jawed vertebrate
604 timetree. Nat. Ecol. Evol. 1:1370.
605 Jarvis E.D., Mirarab S., Aberer A.J., Li B., Houde P., Li C., Ho S.Y.W., Faircloth B.C., Nabholz
606 B., Howard J.T. 2014. Whole-genome analyses resolve early branches in the tree of life of
607 modern birds. Science 346:1320–1331.
608 Jeffroy O., Brinkmann H., Delsuc F., Philippe H. 2006. Phylogenomics: the beginning of
609 incongruence? Trends Genet. 22:225–231.
610 Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7:
611 improvements in performance and usability. Mol. Biol. Evol. 30:772–780.
612 Katzman S., Kern A.D., Bejerano G., Fewell G., Fulton L., Wilson R.K., Salama S.R., Haussler
613 D. 2007. Human genome ultraconserved elements are ultraselected. Science. 317:915.
614 Kearney M., Fujita M.K., Ridenour J. 2009. Lost sex in the reptiles: constraints and correlations.
615 Lost Sex. Springer. p. 447–474.
616 Kent W.J. 2002. BLAT—the BLAST-like alignment tool. Genome Res. 12:656–664.
617 Knowles L.L., Huang H., Sukumaran J., Smith S.A. 2018. A matter of phylogenetic scale:
618 Distinguishing incomplete lineage sorting from lateral gene transfer as the cause of gene
619 tree discord in recent versus deep diversification histories. Am. J. Bot. 105:376–384.

620 Koch N.M., Gauthier J.A. 2018. Noise and biases in genomic data may underlie radically
621 different hypotheses for the position of Iguania within Squamata. PLoS One. 13:e0202729.
622 Kozlov A.M., Aberer A.J., Stamatakis A. 2015. ExaML version 3: a tool for phylogenomic
623 analyses on supercomputers. Bioinformatics. 31:2577–2579.
624 Kück P., Struck T.H. 2014. BaCoCa--A heuristic software tool for the parallel assessment of
625 sequence biases in hundreds of gene and taxon partitions. Mol. Phylogenet. Evol. 70:94–98.
626 Lanier H.C. Huang H. Knowles L.L. 2014. How low can you go? The effects of mutation rate on
627 the accuracy of species-tree estimation. Mol. Phylogenet. Evol. 70:112–119.
628 Leache, A.D., Chavez, A.S., Jones, L.N., Grummer, J.A., Gottscho, A.D. and Linkem, C.W.
629 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture
630 versus restriction site associated DNA sequencing. Genome Biol. Evol. 7:706–719.
631 Lemmon A.R., Emme S.A., Lemmon E.M. 2012. Anchored hybrid enrichment for massively
632 high-throughput phylogenomics. Syst. Biol. 61:727–744.
633 Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin
634 R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25:2078–
635 2079.
636 Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
637 arXiv preprint arXiv:1303.3997.
638 Linkem, C.W., Minin, V.N. and Leaché, A.D. 2016. Detecting the anomaly zone in species trees
639 and evidence for a misleading signal in higher-level skink phylogeny (Squamata:
640 Scincidae). Syst. Biol. 65: 465–477.
641 Losos J.B., Hillis D.M., Greene H.W. 2012. Who speaks with a forked tongue? Science.
642 338:1428–1429.

643 Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.
644 Mai, U. and Mirarab, S. 2018. TreeShrink: fast and accurate detection of outlier long branches in
645 collections of phylogenetic trees. BMC Genomics. 19:272.
646 Nute, M., Chou, J., Molloy, E.K. and Warnow, T. 2018. The performance of coalescent-based
647 species tree estimation methods under models of missing data. BMC Genomics. 19:286.
648 Paradis E., Claude J., Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R
649 language. Bioinformatics. 20:289–290.
650 Philippe H., Brinkmann H., Lavrov D. V, Littlewood D.T.J., Manuel M., Wörheide G., Baurain
651 D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough.
652 PLoS Biol. 9:e1000602.
653 Pisani D., Pett W., Dohrmann M., Feuda R., Rota-Stabelli O., Philippe H., Lartillot N., Wörheide
654 G. 2015. Genomic data do not support comb jellies as the sister group to all other animals.
655 Proc. Natl. Acad. Sci. 112:15402–15407.
656 Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R.
657 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA
658 sequencing. Nature. 526:569–573.
659 Pyron R.A., Burbrink F.T. 2014. Early origin of viviparity and multiple reversions to oviparity in
660 squamate reptiles. Ecol. Lett. 17:13–21.
661 Pyron R.A., Burbrink F.T., Wiens J.J. 2013. A phylogeny and revised classification of Squamata,
662 including 4161 species of lizards and snakes. BMC Evol. Biol. 13:93.
663 Reddy S., Kimball R.T., Pandey A., Hosner P.A., Braun M.J., Hackett S.J., Han K.-L.,
664 Harshman J., Huddleston C.J., Kingston S. 2017. Why do phylogenomic data sets yield
665 conflicting trees? Data type influences the avian tree of life more than taxon sampling. Syst.

Biol. 66:857–879.

Reeder T.W., Townsend T.M., Mulcahy D.G., Noonan B.P., Wood Jr P.L., Sites Jr J.W., Wiens J.J. 2015. Integrated analyses resolve conflicts over squamate reptile phylogeny and reveal unexpected placements for fossil taxa. PLoS One. 10:e0118199.

Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.

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

Romiguier, J., Cameron, S.A., Woodard, S.H., Fischman, B.J., Keller, L. and Praz, C.J., 2016. Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. Mol. Biol. Evol. 33: 670–678.

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

Schliep K.P. 2010. phangorn: phylogenetic analysis in R. Bioinformatics. 27:592–593.

Shen X.-X., Hittinger C.T., Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. 1:126.

Singhal S., Grundler M., Colli G., Rabosky D.L. 2017. Squamate Conserved Loci (Sq CL): A unified set of conserved loci for phylogenomics and population genetics of squamate reptiles. Mol. Ecol. Resour. 17:e12–e24.

Smith S.A., Moore M.J., Brown J.W., Yang Y. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15:150.

Smith S.A., O'Meara B.C. 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics. 28:2689–2690.

689 Smith S.A., Walker J.F., Brown J., Walker-Hale N. 2020. Nested phylogenetic conflicts,
690 combinability, and deep phylogenomics in plants. Syst. Biol. 69:579–592.
691 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
692 large phylogenies. Bioinformatics. 30:1312–1313.
693 Streicher, J.W., Schulte, J.A. and Wiens, J.J., 2016. How should genes and taxa be sampled for
694 phylogenomic analyses with missing data? An empirical study in iguanian lizards. Syst.
695 Biol. 65:128–145.
696 Streicher J.W., Wiens J.J. 2016. Phylogenomic analyses reveal novel relationships among snake
697 families. Mol. Phylogenet. Evol. 100:160–169.
698 Streicher J.W., Wiens J.J. 2017. Phylogenomic analyses of more than 4000 nuclear loci resolve
699 the origin of snakes among lizard families. Biol. Lett. 13:20170393.
700 Tonini J.F.R., Beard K.H., Ferreira R.B., Jetz W., Pyron R.A. 2016. Fully-sampled phylogenies
701 of squamates reveal evolutionary patterns in threat status. Biol. Conserv. 204:23–31.
702 Townsend J.P. 2007. Profiling phylogenetic informativeness. Syst. Biol. 56:222–231.
703 Townsend T.M., Larson A., Louis E., Macey J.R. 2004. Molecular phylogenetics of Squamata:
704 the position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree.
705 Syst. Biol. 53:735–757.
706 Uetz P., Stylianou A. 2018. The original descriptions of reptiles and their subspecies. Zootaxa.
707 4375:257–264.
708 Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine,
709 A., Jordan, T., Shakir, K., Roazen, D., Thibault, J. and Banks, E. 2013. From FastQ data to
710 high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr.
711 Protoc. Bioinformatics. 43:11.

712 Walker J.F., Brown J.W., Smith S.A. 2018. Analyzing contentious relationships and outlier
713 genes in phylogenomics. Syst. Biol. 67:916–924.
714 Wen D., Yu Y., Zhu J., Nakhleh L. 2018. Inferring phylogenetic networks using PhyloNet. Syst.
715 Biol. 67:735–740.
716 Whelan N. V, Kocot K.M., Moroz L.L., Halanych K.M. 2015. Error, signal, and the placement
717 of Ctenophora sister to all other animals. Proc. Natl. Acad. Sci. 112:5773–5778.
718 Whelan N. V, Kocot K.M., Moroz T.P., Mukherjee K., Williams P., Paulay G., Moroz L.L.,
719 Halanych K.M. 2017. Ctenophore relationships and their placement as the sister group to all
720 other animals. Nat. Ecol. Evol. 1:1737.
721 Wiens J.J., Hutter C.R., Mulcahy D.G., Noonan B.P., Townsend T.M., Sites Jr J.W., Reeder
722 T.W. 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive
723 sampling of genes and species. Biol. Lett. 8:1043–1046.
724 Wilke C.O. 2016. Cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. 2016.
725 URL https//CRAN. R-project. org/package= cowplot. R Packag. version 0.7. 0.[p 287].
726 Wong, K.M., Suchard, M.A. and Huelsenbeck, J.P. 2008. Alignment uncertainty and genomic
727 analysis. Science. 319: 473–476.
728 Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: polynomial time species tree
729 reconstruction from partially resolved gene trees. BMC Bioinformatics. 19:153.
730 Zhang J., Kobert K., Flouri T., Stamatakis A. 2013. PEAR: a fast and accurate Illumina Paired-
731 End reAd mergeR. Bioinformatics. 30:614–620.
732
733
734

FIGURES

Figure 1: Concatenated phylogeny inferred using ExaML. Branch colors denote major squamate clades, and each clade is depicted by a representative taxon (all photographs courtesy of author TJC and Pascal Title). Nodes marked by black circles have high statistical support (bootstrap > 95) but conflict between the concatenated and coalescent-based inferred tree (Fig. S2); nodes in gray conflict and have low statistical support; nodes in white are congruent but have low statistical support. Many conflicting nodes have high statistical support.

Figure 2: ExaML-inferred tree with levels of conflict shown at each node. Pie proportions represent the number of gene trees that either support a node, support the most common conflicting relationship, support other less common conflicting relationships, or are non-informative. Branches in gene trees with <80 SH-like support were collapsed prior to analysis. Node labels mark putatively contentious nodes; labels follow Table 1. Many nodes exhibit high levels of gene tree conflict.

Figure 3: D_{LNL} values across all loci for each of the seventeen putative conflicts investigated. Per locus, D_{LNL} values are measured as the difference in log-likelihoods between the two best-supported topological resolutions with respect to a focal relationship (e.g., Anniellidae). Loci are categorized by whether they support the best-supported topology (see Table 1) or not; the dotted line is where $D_{LNL} > 2$. Most loci had fairly small D_{LNL} values, suggesting they do not strongly support any given topological resolution.

Figure 4: Correlation between locus summary statistics (Table 2) and the level of conflict between gene trees as species trees across marker types. Conflict level was measured difference in log-likelihoods of an unconstrained gene tree vs. one constrained to the concatenated species tree. Larger values suggest greater conflict. Linear model fit shown for significant correlations as measured by Spearman's correlation and shown for visualization only. The strongest absolute correlations are for compositional heterogeneity (Spearman's $\rho = 0.19$; $p = 3.0e\text{-}49$) and root-tip variance (Spearman's $\rho = 0.18$; $p = 1.3e\text{-}39$). These results suggest that loci with greater compositional heterogeneity or greater rate variation across the tree are more likely to differ from the concatenated topology.

Figure 5: (A) Correlation in locus D_{LNL} values across different topological tests. The mean correlation in D_{LNL} is $r = 0.175$; the few correlations >0.5 all stem from topological comparisons within snakes. (B) Percent overlap in conflicting loci across different topological tests, shown as the mean deviation from percent overlap of 100 random bootstraps. Values > 0 indicate greater overlap than expected by random. Together, these results suggest that there is little to modest consistency in which loci conflict across different nodes.

Figure 6: Comparative performance across the three marker types used in this study: Anchored Hybrid Enrichment (AHE) markers, standard phylogenetic genes, and ultraconserved elements (UCEs). (A) The level of conflict between gene trees as species trees across marker types. Conflict level was measured difference in log-likelihoods of an unconstrained gene tree vs. one constrained to the concatenated species tree. These results suggest that AHE loci better fit the concatenated tree. (B) Differences in locus quality metrics across marker types. In general, AHE

780 markers showed evidence of being higher quality (i.e., they had lower levels of missingness) and
781 more informative (i.e., trees inferred with AHE markers had higher nodal support as measured
782 by Shimodaira–Hasegawa (SH)-like support).