Syst. Biol. 0(0):1–14, 2022

https://doi.org/10.1093/sysbio/syac055
Advance Access Publication 02 August 2022

Phylogenomic Analysis of the Parrots of the World Distinguishes Artifactual from Biological Sources of Gene Tree Discordance

BRIAN TILSTON SMITH[1,*], JON MERWIN[2,3], KAIYA L. PROVOST[4], GREGORY THOM[5,6], ROBB T. BRUMFIELD[5,6], MATEUS FERREIRA[7], WILLIAM M. MAUCK III[1], ROBERT G. MOYLE[8], TIMOTHY F. WRIGHT[9], AND LEO JOSEPH[10]

[1]*Department of Ornithology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA;* [2]*Department of Ornithology, Academy of Natural Sciences of Drexel University, 1900 Benjamin Franklin Parkway, Philadelphia, PA 19103, USA;* [3]*Department of Biodiversity, Earth, and Environmental Science, Drexel University, 3245 Chestnut St, Philadelphia, PA 19104, USA;* [4]*Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, 318 W. 12th Avenue, Columbus, OH 43210, USA;* [5]*Museum of Natural Science, Louisiana State University, 119 Foster Hall, Baton Rouge, LA 70803, USA;* [6]*Department of Biological Sciences, Louisiana State University, 202 Life Science Bldg, Baton Rouge, LA 70803, USA;* [7]*Centro de Estudos da Biodiversidade, Universidade Federal de Roraima, Av. Cap. Ene Garcez, 2413, Boa Vista, RR, Brazil;* [8]*Department of Ecology and Evolutionary Biology and Biodiversity Institute, University of Kansas, 1345 Jayhawk Blvd., Lawrence, KS 66045, USA;* [9]*Department of Biology, New Mexico State University, MSC 3AF, Las Cruces, NM 88003, USA and* [10]*Australian National Wildlife Collection, National Research Collections Australia, CSIRO, GPO Box 1700, Canberra, ACT 2601, Australia*

**Correspondence to be sent to: Department of Ornithology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA;*
Email: briantilstonsmith@gmail.com

Received 26 April 2021; reviews returned 22 February 2022; accepted 22 July 2022
Associate Editor: Bryan Carstens

Abstract.—Gene tree discordance is expected in phylogenomic trees and biological processes are often invoked to explain it. However, heterogeneous levels of phylogenetic signal among individuals within data sets may cause artifactual sources of topological discordance. We examined how the information content in tips and subclades impacts topological discordance in the parrots (Order: Psittaciformes), a diverse and highly threatened clade of nearly 400 species. Using ultraconserved elements from 96% of the clade's species-level diversity, we estimated concatenated and species trees for 382 ingroup taxa. We found that discordance among tree topologies was most common at nodes dating between the late Miocene and Pliocene, and often at the taxonomic level of the genus. Accordingly, we used two metrics to characterize information content in tips and assess the degree to which conflict between trees was being driven by lower-quality samples. Most instances of topological conflict and nonmonophyletic genera in the species tree could be objectively identified using these metrics. For subclades still discordant after tip-based filtering, we used a machine learning approach to determine whether phylogenetic signal or noise was the more important predictor of metrics supporting the alternative topologies. We found that when signal favored one of the topologies, the noise was the most important variable in poorly performing models that favored the alternative topology. In sum, we show that artifactual sources of gene tree discordance, which are likely a common phenomenon in many data sets, can be distinguished from biological sources by quantifying the information content in each tip and modeling which factors support each topology. [Historical DNA; machine learning; museomics; Psittaciformes; species tree.]

Gene tree discordance is a common feature of phylogenomic studies (Smith et al. 2015; Sharma et al. 2014; Burbrink et al. 2020; Hime et al. 2021; Morales-Briones et al. 2021). The conflict between gene and species trees has been attributed to biological processes and artifacts related to gene tree estimation (Gatesy and Springer 2014). Ancient and contemporary introgression (Eaton et al. 2015; Li et al. 2016; Thom et al. 2018), horizontal gene transfer (Galtier and Daubin 2008), and incomplete lineage sorting (Mirarab et al. 2016) are among the most invoked biological processes to explain gene tree discordance. Although introgression is regularly modeled in small to moderately sized data sets (e.g., Burbrink and Gehara 2018), incomplete lineage sorting can be accounted for even in large phylogenies using approaches that summarize gene trees (Zhang et al. 2018). Missing data, biases introduced during bioinformatic processing and sequence alignment, and errors associated with gene tree estimation are the most common artifacts that can impact the reliability of gene tree topologies (Springer and Gatesy 2016). An assessment of data quality issues is less constrained by phylogeny size because a complex biological process does not need to be modeled during phylogenetic inference. This combination of evolutionary processes, computational limits, and underlying data quality issues makes phylogenomic inference challenging.

Adding to this complexity is that phylogenetic information content varies among loci and individuals. This heterogeneity can obscure phylogenetic relationships and produce discordance among concatenated and species trees. For example, sequence capture of loci from historical museum specimens typically yields short loci with low variation (e.g., McCormack et al. 2016), which subsequently impacts species tree estimation (Hosner et al. 2016; Moyle et al. 2016) and can yield questionable phylogenetic relationships (Smith et al. 2020). Although sequence data from historical samples are expected to have lower quality than more recent material, other characteristics of sequence data could lead to spurious phylogenetic signal (i.e., phylogenetic noise) in any locus or individual. Genomic regions with high GC richness can favor alternative topologies, presumably because these loci are subject to higher meiotic recombination (Bossert et al. 2017). Other factors such as genomic library preparation, sequencing, or bioinformatic processing

can cause the information content in individuals to vary. These sources of heterogeneity can lead to gene tree estimation error and cause artifactual sources of phylogenetic discordance. For these reasons, gene interrogation approaches that identify outlier individuals are important for characterizing the information content underlying gene tree discordance. These approaches can help determine whether biological or artifactual factors are the cause.

One way to address the complexities associated with the multiple dimensions of phylogenetic inference is through machine learning algorithms. This class of methods has been used to infer topologies from multiple sequence alignments (Suvorov et al. 2020; Zou et al. 2020), distinguish between bifurcating or reticulating topologies (Burbrink and Gehara 2018), determine whether maximum likelihood or maximum parsimony are more appropriate for a particular data set (Leuchtenberger et al. 2020), and interrogate genes in phylogenomic studies (Burbrink et al. 2020). The applications of machine learning to gene interrogation methods are particularly interesting because the approach can build on previous work showing that single or few loci can have a disproportionate impact on a particular topology (Arcila et al. 2017; Shen et al. 2017; Brown and Thomson 2017; Walker et al. 2018; Gatesy et al. 2019). For example, machine learning has been used to understand why certain loci fail to support previously proposed relationships (Burbrink et al. 2020), or why topologies with and without missing data vary (Smith 2020). Understanding the association between a phylogenetic hypothesis and the signal driving topological relationships will help clarify the causes of discordance between trees.

In this study, we use tip-based filtering and machine learning to examine the impact of heterogeneous data quality among terminals and subclades in establishing evolutionary relationships in parrots (Order: Psittaciformes). The species of this highly diverse pantropical clade of birds are known for their colorful plumage (Merwin et al. 2020), intelligence (Pepperberg 2009), and highly threatened and endangered status (Snyder and McGowan 2000). Because of their conservation status and relatively high species diversity, many species in this clade are not represented by modern genetic samples in natural history collections and so must be studied through historical museum specimens. This limitation should lead to a sequence alignment with varying levels of data completeness among individuals. Despite the challenges in the availability of samples, major gains in determining phylogenetic relationships in parrots have been made through studies focusing on particular genera (e.g., Ribas and Miyaki 2004; Ribas et al. 2005, 2007a, 2009; Joseph et al. 2011; Kirchman et al. 2012; Smith et al. 2013; Braun et al. 2019) or species-complexes (e.g., Ribas et al. 2006, 2007b; Joseph et al. 2008; Braun et al. 2017), and higher-level relationships (Tavares et al. 2006; Wright et al. 2008; White et al. 2011; Schweizer et al. 2012, Schweizer et al. 2014; Schweizer et al. 2015; Joseph et al. 2020; see Provost et al. 2018 for a review at all levels). Several synthesis studies using supermatrix (Burleigh et al. 2015; Provost et al. 2018) and supertree (Burgio et al. 2019) approaches from legacy markers have provided a more complete overview of relationships across the order. Published multilocus studies have verified the support for most higher-level groups, but incomplete sampling and/or limited molecular data have caused uncertainty in the placement of some groups (e.g., Platycercinae) and the monophyly of particular genera (e.g., *Polytelis*; Provost et al. 2018). Taxonomic relationships at lower ranks (genera, species) have been difficult to resolve because of plumage convergence (e.g., Merwin et al. 2020), a high frequency of derived traits (e.g., Provost et al. 2018), and issues of lumping versus splitting genera (e.g., Joseph et al. 2020). A phylogenomic approach offers an opportunity to resolve recalcitrant nodes and clarify relationships among genera and species.

We used genome-wide markers and nearly complete species-level sampling to produce a phylogenomic tree for parrots. Our aim was to assess evolutionary relationships and the sensitivity of those relationships to the information content available in the DNA sequence alignments (Fig. 1). We estimated trees using concatenated and multispecies coalescent methods because topological relationships are expected to vary between these approaches (Maddison 1997). To test whether topological discordance was an artifact of species in the alignment with low data completeness, we employed a two-step procedure. First, we used two metrics to identify taxa with the lowest phylogenetic information to filter those samples from the alignment. Second, we applied predictive modeling for subclades that remained discordant after tip-based filtering. This was done to characterize the association between summary statistics and their relationship to alternative topologies. From these analyses, we demonstrated how phylogenetic signal can be distinguished from noise to interpret sources of conflict among phylogenies.

Materials and Methods

We sampled ingroup taxa representing 96% of species diversity in the Psittaciformes (382/398 recognized species), including 58 species of lorikeets (Smith et al. 2020), the extinct *Conuropsis carolinensis* (Smith et al. 2021), the endangered *Amazona vittata* (Oleksyk et al. 2012) and three outgroup taxa (*Caracara plancus cheriway*, *Calyptomena viridis*, *Icterus cucullatus*; Oliveros et al. 2019). Specimen details and locality information are available in Supplementary Table S1 available on Dryad at http://dx.doi.org/10.5061/dryad.b5mkkwhfm. The unsampled species include both extinct (*Nestor productus*, *Cyanoramphus ulietanus*, *Cyanoramphus zealandicus*, *Psittacula wardi*, *Psittacula exsul*, *Mascarinus mascarin*, *Vini diadema*, and *Ara tricolor*) and extant (*Pionus menstruus*, *Tanygnathus gramineus*, *Psephotellus chrysopterygius*, *Loriculus camiguinensis*, *Loriculus flosculus*, *Aratinga*

MANUSCRIPT CATEGORY: Systematic Biology

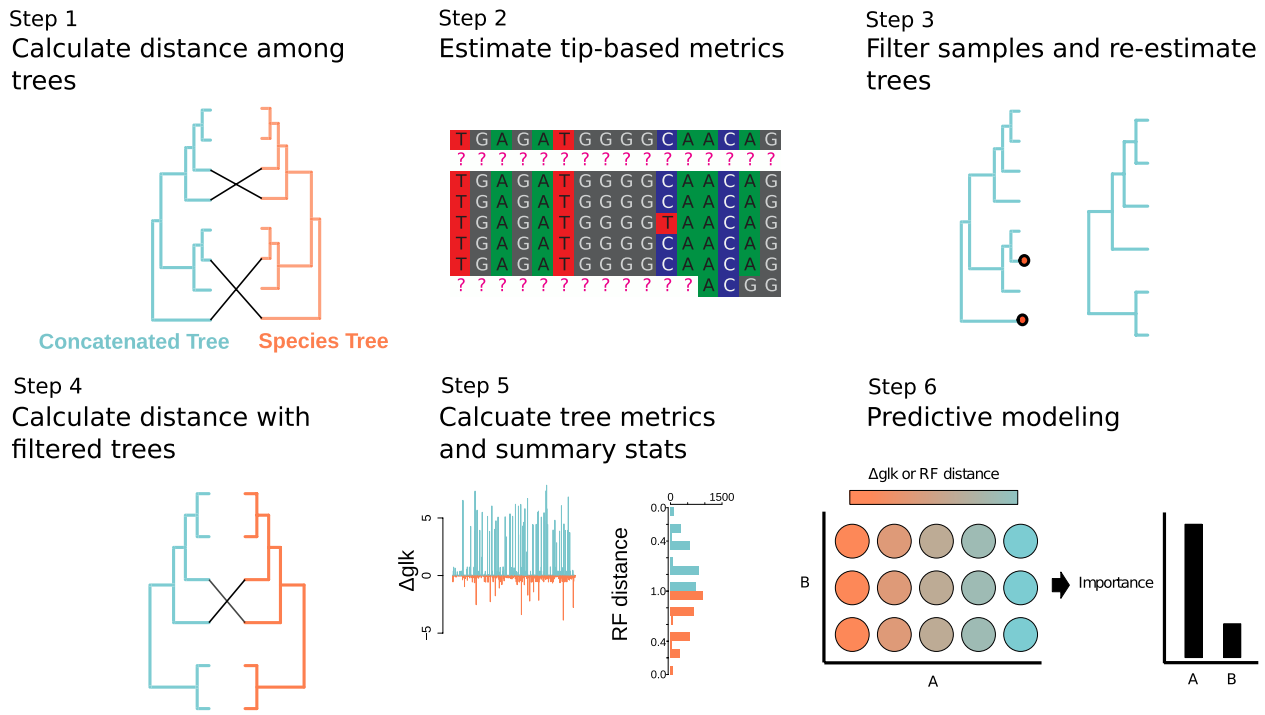

FIGURE 1. Flowchart for evaluating the impact of artifactual sources of topological discordance. Step 1: estimate concatenated (aqua on left) and species (orange on right) trees, and then quantify differences among the topologies. Step 2: estimate tip-based metrics that reflect the amount of phylogenetic signal in each species. Step 3: using the tip-based metrics to filter individuals above threshold values. Step 4: calculate how different the concatenated and species tree topologies are after filtering. Step 5: for subclades, calculate gene-wise log-likelihood scores (Δglk) for each topology and Robinson–Foulds (RF) distances. Step 6: use machine learning and alignment summary statistics to predict Δglk and the RF distance between each gene tree to the species tree.

maculata, *Psittacara wagleri*, *Pyrrhura pfrimeri*, and *Poicephalus crassus*) species. In comparison to the previously largest published parrot phylogeny (Provost et al. 2018; excluding Jetz et al. 2012 and Burgio et al. 2019, which included species without genetic data), we added 82 taxa. Our sampling included several recent species splits in *Amazona*, *Pyrrhura*, and *Psittacara*. This data set also included more comprehensive sampling in the lorikeets (Smith et al. 2020), and from groups that have not been the focus of phylogenetic studies, such as *Loriculus* and *Touit*, and the species *Ognorhynchus icterotis* in an unsampled monotypic genus.

For species with tissue samples, we extracted total genomic DNA from genetic resources using QIAamp DNeasy extraction kits (Qiagen, Valencia, CA). For species without modern genetic samples, we sampled skin from the toes of museum specimens and applied a modified DNeasy extraction protocol that used QIAquick PCR filter columns that size selected for smaller fragments of DNA. The modified protocol also included washing the sample with H_2O and EtOH prior to extracting as well as extra time for digestion. DNA extraction from historical samples was done in a dedicated lab for working on degraded samples to reduce contamination risk. We quantified DNA extracts using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Library preparation of UCEs, enrichment, and Illumina sequencing were performed by RAPiD Genomics (Gainesville, FL). The Tetrapod UCE 5K probe set was used to enrich 5060 UCE loci (Faircloth et al. 2012). Sequencing was done on an Illumina HiSeq 3000 PE 150 and HiSeq 10×. FASTQ files are available on the Short Read Archive (SRA numbers are available in Supplementary Table S1 available on Dryad).

We used a modified data-processing pipeline that incorporated PHYLUCE v. 1.6.8 (Faircloth 2016), a software package developed for analyzing UCE data and a variant calling pipeline (Harvey et al. 2016). Low-quality bases and adaptor sequences were trimmed from demultiplexed FASTQ files using illumiprocessor v1 (Faircloth 2013; Bolger et al. 2014). Next, we chose a single individual per genus and produced *de novo* contigs with ABySS v.1.5.2 (Simpson et al. 2009). We explored using several assemblers, and although ABySS produced shorter contigs, preliminary results found that these produced more stable phylogenetic relationships among modern and historical samples, and less branch length heterogeneity. We mapped contigs to UCE probes and generated an index for the genus-specific reference sequence and independently mapped reads from each of the species belonging to the same genus using BWA v0.7.13-r1126 (Li and Durbin 2009). SAM files produced from the BWA mapping were converted to BAM files and sorted with SAMtools v. 1.10 (Li et al. 2009). Then, we used the mpileup function in SAMtools v. 1.10 (-C 30; -Q 20) to produce a VCF file,

bcftools v. 1.12 (Danecek et al. 2021), and vcfutils to call variant sites and filter sites with <5× coverage per SNP and a quality score < 20, and seqtk to convert FASTQ files to FASTA. Using the PHYLUCE default settings, MAFFT v. 7.455 (Katoh and Standley 2013) aligned loci that were a minimum of 100 base pairs (bp). Heterozygous sites were labeled with IUPAC ambiguity codes. The final concatenated alignment retained only loci for which 75% of the samples were present.

Phylogenomic Analyses

We estimated phylogenies using concatenated and coalescent approaches to assess the stability of relationships and the extent of phylogenetic discordance. To infer phylogenomic relationships from the concatenated UCE alignment, we used IQ-TREE2 v. 2.1.3 (Minh et al. 2020) with 1000 rapid bootstraps, and the software's ModelFinder implementation (Kalyaanamoorthy et al. 2017) to select the best-fit substitution model for each gene partition (Chernomor et al. 2016). We also generated gene trees from locus alignments where heterozygous sites had an ambiguity code in IQ-TREE2, and we collapsed all nodes in the gene trees with 0% SH-like aLRT support using newick utilities (Junier and Zdobnov 2010) following Simmons and Kessenich (2020) and Simmons and Gatesy (2021). These gene trees were used for species tree estimation in ASTRAL-III v. 5.7.3 where support for nodes was determined by local posterior probabilities (Zhang et al. 2017). Finally, we used an approximately unbiased topology test (AU test; Shimodaira 2002) to determine whether the concatenated and species tree topology were significantly different than each gene tree topology for each subclade with 10,000 multiscale bootstrap replicates in IQ-TREE2 v. 2.1.3 (Minh et al. 2020).

Detecting Outlier Samples

Poorly resolved gene trees have been shown to decrease the performance of species tree estimation using gene tree summary methods such as ASTRAL (Zhang et al. 2018). Typically, low-quality gene trees are caused by loci with overall low phylogenetic signal or by individuals with high proportions of missing data (Sayyari et al. 2017). We determined how the species tree topology changed by filtering individuals based on two metrics that reflect data quality. The number of parsimony informative sites (PIS) is a commonly used metric that reflects the information content of a locus. However, the total PIS varies among loci because of differences in sequence length and mutation rate. To scale PIS per individual by the average PIS per locus, we estimated the difference between the overall mean PIS per locus and the per sample mean PIS per locus. We refer to this metric as the deviation in PIS (dPIS). The per species mean of this metric is similar to the total PIS per individual, which is useful for identifying low information content samples. The per locus value of dPIS can indicate which individuals have fewer or more PIS than expected for a given locus; this can be useful for filtering individuals from specific gene trees. To calculate the PIS per individual per locus, we converted the FASTA alignment for each locus into a VCF file using the program SNP-SITES v. 2.5.1 (Page et al. 2016). We used VCFTOOLS v0.1.15 (Danecek et al. 2021) to calculate the PIS by first filtering out all sites with minor allele count < 4 and, and then calculating the number of homozygous sites for the alternative allele to avoid counting heterozygous sites as PIS, using the most distant outgroup (*C. cheriway*) as the reference. The second metric was the number of missing loci for each individual, which was calculated using the summary statistics R script from Burbrink et al. (2020). For both metrics, we estimated the 90th, 80th, and 70th percentiles as filtering thresholds: 1) dPIS: > 7.43 (90th percentile), > 5.30 (80%), and > 1.31 (70%) and 2) number of missing loci per taxon: > 358.6 (90%), > 280.4 (80%), and > 169.8 (70%). From these filtered alignments, we followed the methods listed above for estimating concatenated and species trees. We estimated the generalized Robinson–Foulds distances (Smith 2020) between the full and filtered trees. Trees were plotted in R v. 3.6.3 using ape v. 5.4-1 (Paradis and Schliep 2019), ggtree v. 2.04 (Yu et al. 2017), phytools v. 0.7-70 (Revell 2012), diversitree v. 0.9-14 (FitzJohn 2012), geiger v. 2.0.7 (Pennell et al. 2014), and RColorBrewer v. 1.1-2 (Neuwirth and Neuwirth 2011).

Quantifying Topological Discordance in Subclades

The relationship between locus properties and topological discordance can be modeled to distinguish between loci with signal versus those with noise. We focused on subclades, as opposed to the full tree, to capture more nuanced patterns within clades. From our filtered trees (threshold dPIS > 1.31; 70th percentile), we sampled subclades that showed either: 1) topological differences between the concatenated and species trees ($n = 21$), 2) differences in the phylogenomic trees (concatenated and species tree) from previously published relationships ($n = 1$), or 3) nonmonophyletic genera in both trees ($n = 2$), which were compared to a topology where monophyly was enforced. We subsampled the number of taxa in each subclade, when appropriate, to retain the minimal number of taxa that showed the alternative relationships among topologies. Next, we used per-locus alignment summary statistics as independent variables in our models (alignment length, undetermined characters, number of variable sites, PIS, GC content, frequency of gaps, segregating sites with gaps, and number of individuals) which we estimated in AMAS (Borowiec 2016) and the summary statistics R script from Burbrink et al. (2020). These statistics provided an overview of the phylogenetic signal and potential biases within each locus. We also included which chromosome each locus was on as an independent variable by mapping the UCE probes to the *Melopsittacus undulatus* genome (Ganapathy et al. 2014) in PHYLUCE. This variable was included to determine

whether loci on particular chromosomes favored a topology. We estimated two dependent variables: 1) the normalized Robinson–Foulds distance (Robinson and Foulds 1981), or RF distance, a metric that captures the distance between the phylogenomic tree and each gene tree and 2) the gene-wise log-likelihood for each topology (Shen et al. 2017; Walker et al. 2018). We estimated normalized RF distances between every gene tree and the species/alternative and concatenated trees estimated with the same filtered data sets in the R packages ape v. 5.4 (Paradis and Schliep 2019), phytools v. 0.7-47 (Revell 2012), and phangorn v. 2.5.5 (Schliep 2011). Using RAxML (Stamatakis 2014), site-wise log-likelihoods were calculated for the concatenated (T_1) and species/alternative tree (T_2) topologies for each subclade, and gene-wise log-likelihoods were the sum of site-wise log-likelihoods for each gene. The change in gene-wise log-likelihoods were estimated for each gene partition (hereafter Δglk = T_1 gene log-likelihood - T_2 gene log-likelihood). Positive Δglk scores supported T_1 and negative scores supported T_2.

We used the K-nearest neighbor (KNN) classifier in the machine learning R package caret v. 6.0.79 (Kuhn 2008) to determine which summary statistics were most important in predicting: 1) Δglk scores and 2) the species-gene tree RF distance in each of the 24 subclades. We used KNN instead of a neural net because it could handle chromosome location as a single variable, and the results were qualitatively similar with either algorithm. The input data were centered and scaled, and the percentage of training versus test data was set to 70% and 30%, respectively. We produced 100 training/test data sets, and reported the mean: R^2, root mean square error, and variable importance. To obtain the directionality of relationships for the most important variables in the model, we estimated the correlation coefficient in R. We used KNN modeling to determine whether topological discordance, between species and concatenated trees, was caused by phylogenetic signal or noise. It is expected that multilocus data sets will harbor both informative and noninformative loci, but the overall ratio of signal-noise can be captured by the KNN modeling. For example, if a metric that reflects phylogenetic signal (e.g., PIS) was an important variable that had a positive correlation with Δglk, it would indicate that the topology in the concatenated tree was being driven by phylogenetic information content. Alternatively, if models where statistics that may capture noise in the data (e.g., GC content; undetermined characters) are found to be the most important variable, that would indicate the topology they were correlated with was likely biased by low information content. For subclades where the models were a poor fit to the data (i.e., low R^2 values), the alternative topologies could not be predicted by the summary statistics.

Time-Calibrating the Parrot Phylogeny

To provide a temporal perspective on topological discordance in the Psittaciformes, we estimated a time-calibrated tree for the clade. We used a penalized likelihood method to estimate divergence dates to accommodate the large number of species and characters in our data set. We dated the complete concatenated tree using treePL (Smith and O'Meara 2012) and calibrated nodes with the fossil ages of the following extinct taxa: *Eozygodactylus americanus* (Weidig 2010; minimum age: 51.81 Ma; maximum age: 66.5 Ma) for calibrating the split between Psittacopasserae (Psittaciformes and Passeriformes; *Nelepsittacus minimus* (Worthy et al. 2011; minimum age: 15.9 Ma; maximum age: 66.5 Ma) for the split between *Strigops* and *Nestor*; an unidentified member of the extant genus *Cacatua* (Boles 1993; minimum age: 11.608 Ma; maximum age: 23.03 Ma) for the split between *Eolophus* and *Callocephalon*; and Suboscines indet. (Mayr and Manegold 2006; minimum age: 27.25 Ma; maximum age: 56.0 Ma) for the split between oscine and suboscine passerines. To specify an upper bound on the tree root, we used an external calibration from Jarvis et al. (2014) for the split between Falconiformes and Psittacopasserae (minimum age: 57; maximum age: 62 Ma). Justification for these fossil calibrations and their ages, except for *Cacatua*, are discussed in detail in Kimball et al. (2019) and Oliveros et al. (2019). In treePL, we estimated the optimal parameter settings using the prime and thorough options and randomly sampled during the cross-validation over a range of smoothing parameters ($1\times10^{-7} - 1\times10^{4}$) for 10 iterations and ran the analysis on 100 bootstrapped trees.

RESULTS

Data Characteristics

We obtained 3242 loci from 385 taxa totaling 1,285,685 bp per taxon consisting of 132,895 PIS, 84,944 singleton variable sites, and 941,672 constant sites. The average number of species per locus was 340 (range = 3–382); locus length average 352 bp (range = 100–1572 bp), and GC content was 0.40 (range = 0.39–0.45). The metrics dPIS and number of missing loci were partially correlated (0.66), but different sets of individuals were identified in the threshold categories (Supplementary Table S2 available on Dryad). Both metrics were plotted on the concatenated tree to show their distribution across taxonomic groups (Supplementary Fig. S1 available on Dryad). The number of missing loci per taxon ranged from 18 to 3086 with a mean ± SD of 198 ± 327. Based on our filtering threshold categories, 39 individuals had >358.6 missing loci (samples from fresh tissues: n = 21; museum skins: n = 18), 76 had >280.4 missing loci (fresh tissues: n = 37; museum skins: n = 39), and 102 had >169.8 missing loci (fresh tissues: n = 61; museum skins: n = 54; Supplementary Figs. S2 and S3 available on Dryad; Supplementary Table S2 available on Dryad). The mean dPIS across species was 0.02 (SD = 5.52) with a range of -4.85 to 21.65. When the dPIS was greater than zero, the individual had fewer mean PIS per locus than the average across species. Based on filtering threshold categories, 42 individuals had a dPIS > 7.43 (fresh tissues:

MANUSCRIPT CATEGORY: Systematic Biology

$n=7$; museum skins: $n=35$), 77 had >5.30 (fresh tissues: $n=31$; museum skins: $n=46$), and 116 had >1.31 (fresh tissues: $n=40$; museum skins: $n=76$; Supplementary Figs. S2 and S3 available on Dryad; Supplementary Table S2 available on Dryad). A total of 211/269 taxa were identified in the 70th percentile lists for both metrics (Supplementary Table S2 available on Dryad).

Phylogenetic Discordance and Tip-based Filtering

Most nodes in the concatenated tree had 100% bootstrap support, whereas support values in the species tree were more varied but also high (Supplementary Figs. S4 and S5 available on Dryad). Higher-level phylogenetic relationships were similar to previous phylogenetic studies, with some notable exceptions. In the concatenated tree, relationships among three superfamilies were as expected based on previous work, showing Strigopoidea from New Zealand as sister to a clade containing Australasian Cacatuoidea and pantropical Psittacoidea. Within Psittacoidea, Psittacidae was sister to a clade containing Psittrichasidae and Psittaculidae. The species tree, however, showed extensive discordance with the concatenated tree (Supplementary Figs. S4 and S5 available on Dryad). For example, Strigopoidea consists of two low-diversity families, Strigopidae and Nestoridae. In the concatenated tree, the relationships between *Strigops* and *Nestor* and between the two extant *Nestor* species (*Nestor meridionalis* and *Nestor notabilis*) were highly supported. In contrast, in the species tree, Strigopoidea was paraphyletic because Strigopidae and then Nestoridae were on successive branches separated by a short branch with 100% support. A similar pattern was recovered for Psittrichasidae (Coracopseinae and Psittrichasinae), which was monophyletic only in the concatenated tree. In the species tree, Coracopseinae and then Psittrichasiinae were on successive branches that were short but highly supported (100% and 99%, respectively). Higher-level relationships within Cacatuidae, Psittacidae, and Psittaculidae were similar between the species and concatenated trees, with a few exceptions. The Afrotropical Psittacinae was sister to the New World Arinae, and the tribes in Arinae had the same relationships in the species and concatenated trees. One notable difference was that the support for Androglossini of the Neotropics as a clade was lower in the species tree than the concatenated tree (BS = 75% vs. 100%, respectively). Agapornithinae was sister to the Loriini in the concatenated tree (BS = 100%) and sister to the Platycercinae in the species tree (BS = 63). The genera *Amazona*, *Anodorhynchus*, *Aratinga*, *Bolborhynchus*, *Cacatua*, *Cyanoramphus*, *Cyclopsitta*, *Eos*, *Nannopsittaca*, *Polytelis*, *Psilopsiagon*, *Psittacula*, *Saudareos*, *Tanygnathus*, *Trichoglossus*, and *Vini* were not monophyletic in the species tree (Supplementary Figs. S5 and S6 available on Dryad). Many of these cases of nonmonophyly were unexpected based on the known phylogenetics and taxonomy of these groups. *Psittacula*, *Nannopsittaca*, *Bolborhynchus*, *Psilopsiagon*, and *Cyclopsitta* were not monophyletic in either tree (Supplementary Figs. S4–S6 available on Dryad).

We filtered individuals from the alignment using the metrics dPIS and number of missing loci and found that many of the cases of nonmonophyly of genera were driven by low-quality samples (e.g., *Amazona*, *Anodorhynchus*, or *Aratinga;* Supplementary Figs. S4–S18 available on Dryad). Tip-based filtering reduced RF distances among the concatenated and species trees from 0.27 (full concatenated vs. full species trees) to 0.08 (dPIS $>$ 1.31 concatenated tree vs. dPIS $>$ 1.31 species tree; Supplementary Fig. S19 available on Dryad). RF distances between the full concatenated and the filtered species tree topology were lower when using the dPIS filter ($>7.43=0.18$; $>5.3=0.15$; $>1.31=0.08$) than the number of missing loci filter ($>358.6=0.24$; $>280.4=0.22$; $>169.8=0.19$). dPIS ($>7.32=3$ cases; $>5.3=2$; $>1.31=1$) was more efficient than number of missing loci ($>358.6=9$; $>280.4=8$; $>169.8=4$) at identifying samples that produced unexpected cases of nonmonophyly in the species tree based on taxonomy (Supplementary Fig. S6 available on Dryad). Tip filtering by the two metrics identified the most data-poor samples. It showed that in the species tree, most instances of nonmonophyletic genera were due to data quality, although discordance between the two trees persisted after filtering.

Discordant Subclades

To further assess drivers of topological conflict, we focused on three categories: 1) subclades ($n=21$) that were discordant between the concatenated and species tree after the tip-based filtering, 2) one case in which the phylogenomic tree differed from previously published phylogenetic relationships (*Nymphicus*; White et al. 2011), and 3) two cases in which nonmonophyletic genera (*Psittacula* and *Cyclopsitta-Psittaculirostris*) were found in both phylogenomic trees and monophyly was enforced in the alternative topology (Fig. 3; Supplementary Figs. S20–S26 available on Dryad). We split several genera into smaller subclades to decrease the number of relationships that differed among trees. For example, there were three clades in *Cacatua*, one composed of species with dark bills (*Cacatua* A) and another with lighter colored bills (*Cacatua* B), and a separate subclade where the placement of *C. haematuropygia* (*Cacatua* C) differed. *Pyrrhura* was also split into A and B clades. In some subclades (*Agapornis*, *Eupsittula*, *Eupsittula-Psittacara*, *Micropsitta*, *Poicephalus*, *Pyrilia*, *Pyrrhura* A, and *Pyrrhura* B), relationships differed between the two topologies, but the relationships were poorly supported in one tree. There was also one subclade where the placement of the genera *Eupsittula* and *Psittacara* varied among trees, lower support being in the species tree (BS 80%). Most subclades had discordance within genera (e.g., *Ara*, *Cacatua* A–C, *Pionus*, *Primolius*, *Psittacula*, *Touit*, and *Psittacara*). In the species tree, *Polytelis alexandrae* was sister to *Aprosmictus* (BS = 100%) whereas in the concatenated tree *P. alexandrae* was sister to the other two taxa in *Polytelis* (BS = 69%). Comparisons of these

MANUSCRIPT CATEGORY: Systematic Biology

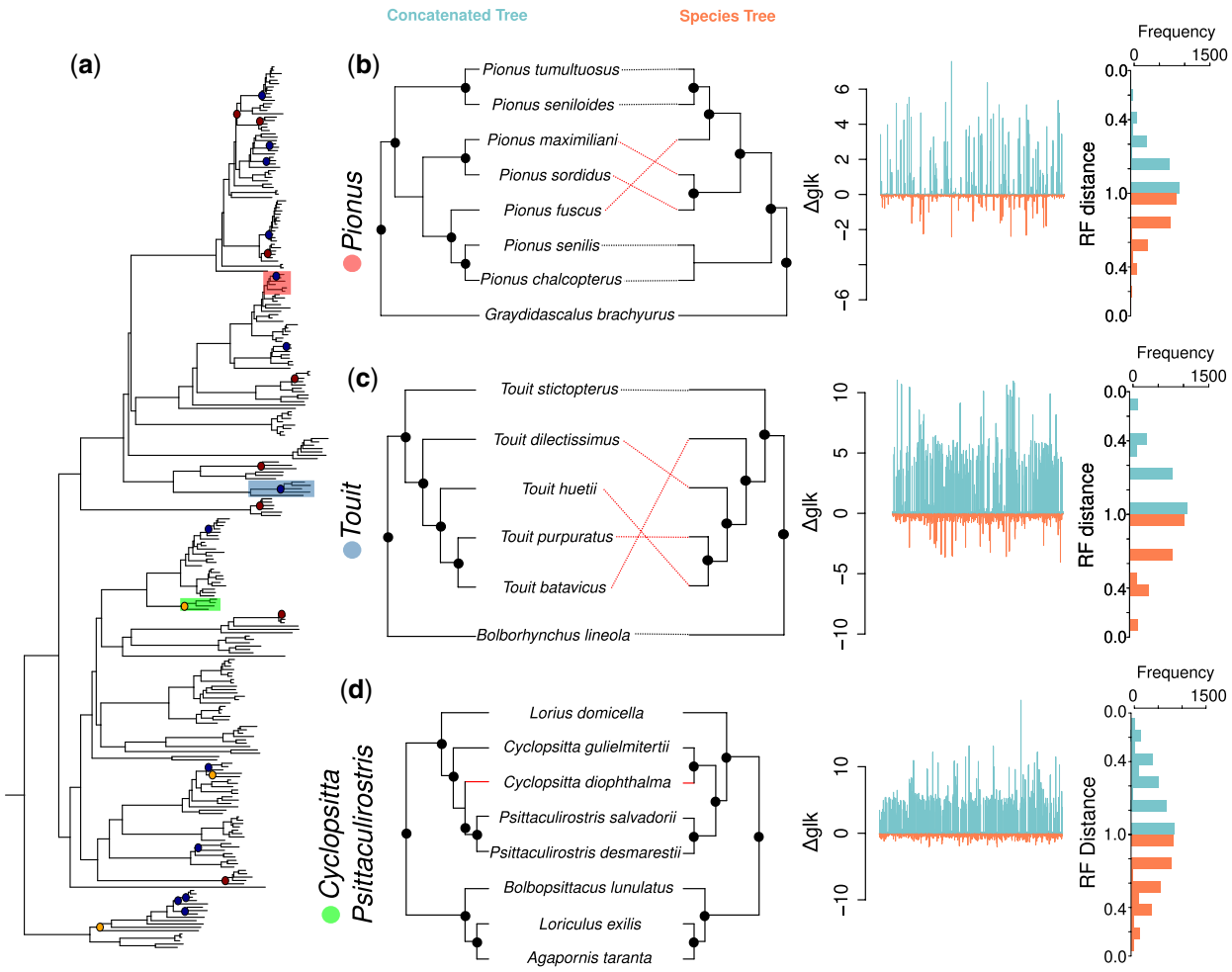

FIGURE 2. Exemplar subclades with alternative topologies between the concatenated (left) and species (right) trees with plots showing the change in gene-wise likelihood scores for the two topologies, and the distance among gene trees to each of the topologies. a) Shown is the filtered concatenated tree (dPIS < 1.31) with exemplar subclades highlighted, and the 24 subclades denoted with colored circles. The colored nodes reflect different conditions. The different relationships between the concatenated and species tree had support values >95% (blue), <95% (red), or where an alternative relationship was forced (orange). b) *Pionus* (red) was an example where summary statistics were a good predictor of the Robinson–Foulds distance (RF distance) between gene trees and the species tree. c) *Touit* (light blue) was an example where summary statistics were a good predictor of the gene-wise Δ log-likelihood (Δglk). d) *Cyclopsitta-Psittaculirostris* (bright green) was an example where both phylogenomic trees agreed, but we forced monophyly. b–d) Shown for each subclade are the branches of conflicting topologies that differ for the concatenated (aqua) and species tree (orange). Nodes with support values ≥95% are denoted with circles. The trees were filtered by excluding all tips with dPIS values >1.31. To the right are the Δglk scores for each gene partition across each subclade's alignment, where positive scores are of genes that support the concatenated tree (aqua) and negative scores favor the species tree (orange). Far right are histograms of the RF distances for gene trees versus species trees (orange), and gene trees versus the concatenated tree (aqua). Plots for the other subclades are available in Supplementary Figures S20–S26 available on Dryad.

topologies with prior Sanger-based topologies (e.g., Provost et al. 2018) resulted in four categories of congruence. The Sanger tree was congruent with either the concatenated or species trees, or was incongruent with either phylogenomic tree, or incomplete sampling prevented the comparison.

There was a wide variation in the number of loci supporting a topology across subclades. Using an AU test, we found that the percentage of gene tree topologies that were not statistically different ($\alpha = 0.05$) from either the concatenated or species tree topology ranged from 36% (*Ara*) to 86% (*Primolius*) (Supplementary Table S3 available on Dryad). There were a low number of loci supporting either topology detected in the gene-wise likelihood analyses. Across subclades, 1−26% of gene trees had Δglk scores > 2, which favored the concatenated topology, and no more than 2% of gene trees supported the species or alternative tree (Δglk scores <-2). RF distances between the species tree/concatenated tree and gene trees also showed high variability (Supplementary Figs. S20–S26 available on Dryad).

Predictive Modeling

Model performance varied across the 24 subclades (Fig. 2a) and the two dependent variables that were

run independently (Supplementary Table S4 available on Dryad). We found two general patterns across the modeling results where the data were a good fit in predicting: 1) Δglk scores favoring the concatenated tree (Fig. 2b; e.g., *Pionus*) or 2) RF distance between gene trees and the species (Fig. 2c; e.g., *Touit*) or alternative tree (Fig. 2d; *Cyclopsitta-Psittaculirostris*). The distributions of Δglk scores were heterogeneous across alignments (e.g., Fig. 2), but more loci favored the concatenated tree versus the species-tree topology for all subclades.

Predictive modeling revealed that the RF distances were an overall better fit to the data than the Δglk, but the pattern varied across subclades. The range in R^2 values varied across subclades for the Δglk scores (0.00–0.21) and RF distances (0.00–0.48). The relative importance of variables was summarized in a heatmap (Fig. 3). The best performing models had PIS as the most important variable Δglk: $N = 5$; R^2 range = 0.05–0.21; RF distance: $N = 12$; R^2 range = 0.17–0.48), and nearly all runs had PIS as one of the top three most important variables. PIS was positively correlated with Δglk and negatively correlated with RF distances (gene trees s. species trees). The next most important variable was GC content, and for the cases in which GC content was the most important variable, the R^2 values were lower (Δglk: $N = 15$; R^2 range = 0.00–0.10; RF distance: $N = 9$; R^2 range = 0.00–0.12). The frequency of gaps and segregating sites with gaps were in the top four most important variables in most models. Percentage of missing data, alignment length, chromosome, and number of individuals were never in the top three most important variables for any model. For 12 of the subclades (*Agapornis*, *Ara*, *Cacatua* A, *Cyclopsitta-Psittaculirostris*, *Eupsittula-Psittacara*, *Nymphicus*, *Pionus*, *Polytelis*, *Psilopsiagon-Bolborhynchus-Nannopsittaca*, *Psittacula*, *Psittacula-Tanygnathus*, and *Pyrilia*), RF distances between the gene trees and species tree were best explained by PIS. In contrast, the best performing models for Δglk were in *Amazona*, *Micropsitta*, *Psittacara*, *Pyrrhura* B, and *Vini*, indicating that the support for the concatenated tree over species tree was driven by phylogenetic signal. The three subclades in which the phylogenomic trees were concordant but an alternative topology was assessed (*Cyclopsitta-Psittaculirostris*, *Nymphicus*, and *Psittacula*) had well-fit models for RF

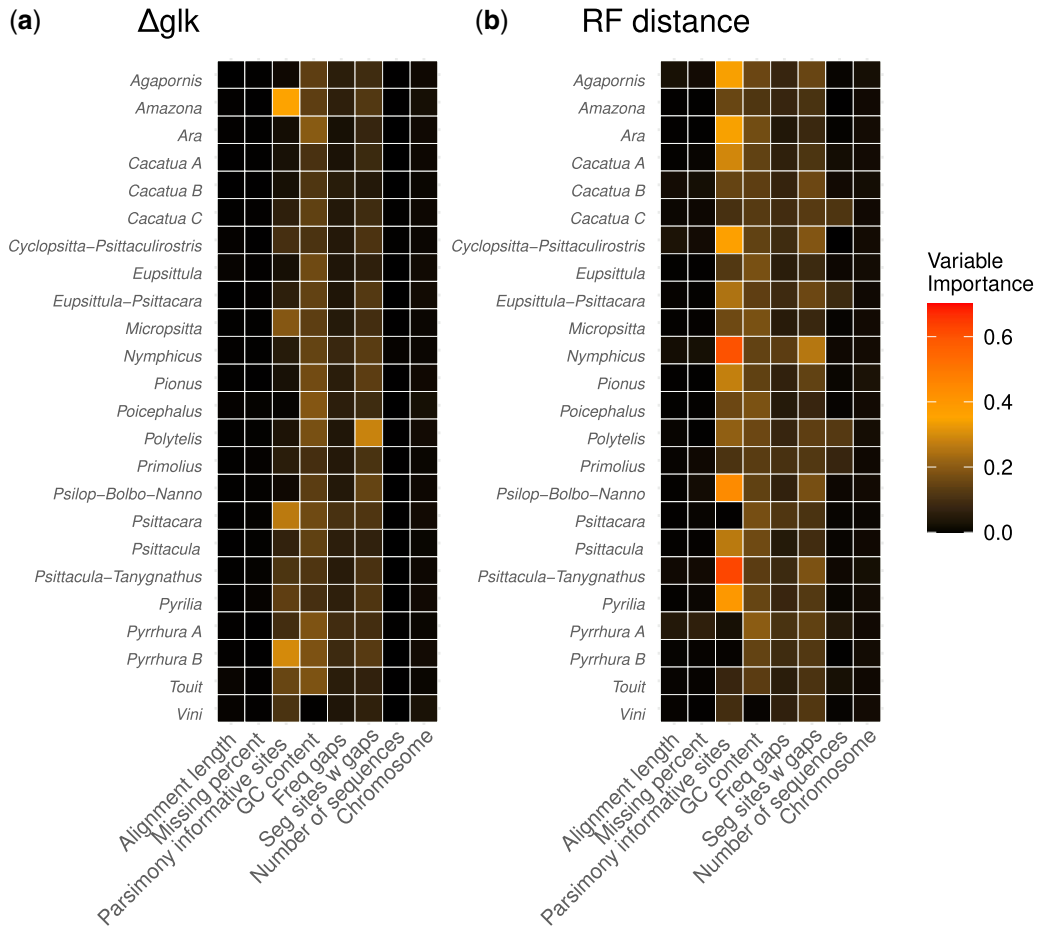

FIGURE 3. Heatmap showing variable importance for the KNN model for a) concatenated-species tree gene-wise Δ log-likelihood (Δglk) scores and b) Robinson–Foulds distance (RF distance) from gene trees to the species tree. On the vertical axes are each of the 24 subclades, which were independently analyzed, and on the horizontal axes are summary statistics used for the KNN classification. Warmer colors denote more important variables.

MANUSCRIPT CATEGORY: Systematic Biology

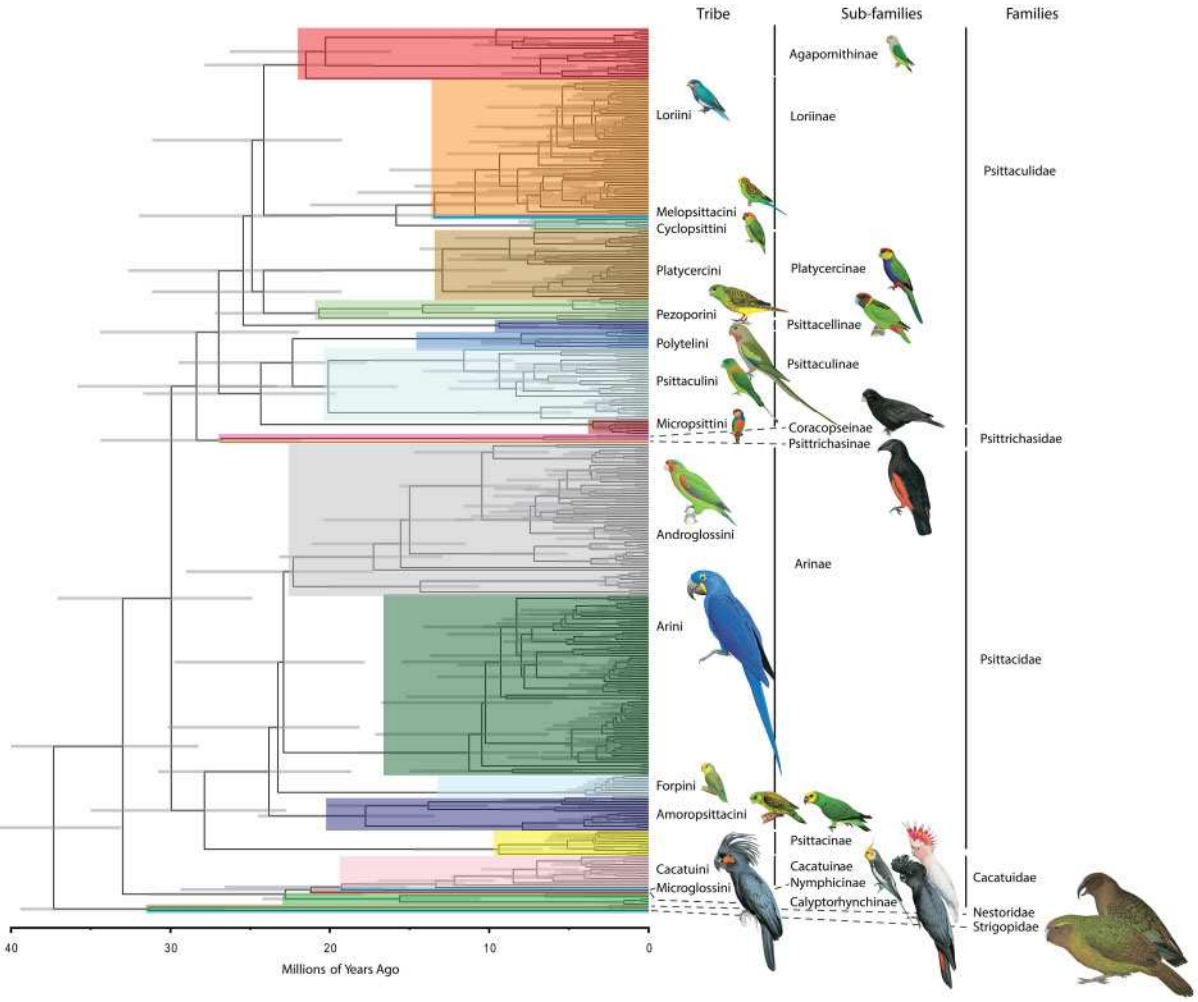

FIGURE 4. Time-calibrated tree of the Psittaciformes. Colored clades represent the lowest taxonomic rank from Family to Tribe in each lineage. Bars on nodes are from 100 rapid bootstrap trees. Included is an exemplar from each taxonomic group. Bird illustrations reproduced by permission of Lynx Edicions under an Open Access Creative Commons License.

distance but not Δglk. Interestingly, for these thre subclades, there were few genes (*Cyclopsitta-Psittaculirostris*: $n=1$; *Nymphicus*: $n=3$; and *Psittacula*: $n=17$) based on Δglk scores (−2) that supported the alternative tree.

Temporal Divergences

For node ages, we report estimates from the maximum likelihood concatenated tree and the range from 100 rapid bootstrapped trees (Fig. 4). A time-calibrated tree with tips ($n=384$) is provided in Supplementary Figure S28 available on Dryad. *Pyrrhura snethlageae*, now known as *Pyrrhura amazonum pallescens*, was removed because it was a low-quality sample with an unexpected position within *Pyrrhura*. Strigopoidea diverged from Cacatuoidea/Psittacoidea at 37.45 Ma (33.03–43.81 Ma), and Cacatuoidea and Psittacoidea split at 33.15 Ma (28.27–40.11 Ma). The six families diverged during a similar timeframe in the Oligocene. Strigopidae and Nestoridae diverged 31.45 Ma (26.16–39.67). Psittacidae had a crown age of 28.13 Ma (22.81–35.17). Psittrichasiidae and Psittaculidae diverged 28.51 Ma (23.27–35.89). The 12 subfamilies originated throughout the Miocene. The 14 lower diversity tribes (e.g., Melopsittacini and Microglossini) and the species-rich ones (e.g., Arini and Loriini) diversified from the early Miocene to Pliocene. The crown ages of the tribes ranged from the early Miocene (e.g., Androglossini) to Pliocene (e.g., Forpini). The oldest species divergences are for monotypic lineages (e.g., *Psittrichas fulgidus* split from *Coracopsis* 26.94 Ma (21.86–34.56)) and the youngest divergences are <1 Ma at the species level (e.g., *Rhynchopsitta pachyrhyncha* and *R. terrisi* split at 0.39 Ma (0.2–0.74)). The majority of the 21 subclades that show phylogenetic discordance involved divergences that occurred within the last 10 million years.

DISCUSSION

We showed that by identifying low-quality samples with simple metrics and applying a predictive modeling approach to evaluate the support for a topology, we

could distinguish artifactual from biological sources of topological discordance in the phylogeny of parrots. In phylogenies inferred from genome-wide markers and 96% of the species diversity in the clade, we recovered extensive discordance among the concatenated and species tree topologies. In contrast to previous studies showing ancient phylogenetic discordance in other systems (e.g., Burbrink et al. 2020; Hime et al. 2021; Morales-Briones et al. 2021), we found that the deeper nodes in the parrot tree, which date to the Eocene through to the middle Miocene, were congruent among the species and concatenated analyses. Topological conflict, however, was most evident in portions of the tree that were previously poorly understood (i.e., within genera). By using simple, tip-based metrics, we differentiated among samples having low versus higher phylogenetic information content. Likewise, by filtering these samples at different thresholds, we resolved the bulk of the discordance, particularly cases of unexpected nonmonophyly in the species tree based on taxonomy. Support for alternative topologies in individual subclades was further determined using tree-based metrics and alignment summary statistics. For subclades where the models performed well, PIS was the most important variable in predicting gene-wise log-likelihood scores. A similar pattern was found in models for a different subset of subclades where PIS was most important in predicting the distances between gene trees and the species tree. That, in turn, indicates phylogenetic information was driving the species tree topology in those particular subclades. Collectively, this shows that neither the concatenated nor the species tree can be treated as the preferred tree when alignments consist of varying levels of data completeness.

Detecting Artifactual Signal

Our approach differs from previously applied gene interrogation methods in that it consisted of multiple steps that filtered low-quality samples, modeled gene characteristics supporting alternative trees, and examined subclades versus the whole tree. Excluding loci based on their phylogenetic signal (Gatesy et al. 2019), amount of missing data (Hosner et al. 2016), or evolutionary rate and phylogenetic usefulness (Mongiardino Koch 2021) have been used to decrease gene tree estimation error prior to species tree estimation (Simmons et al. 2016; Molloy and Warnow 2018). Our results indicate that a similar result is achievable by removing lower-quality samples from the alignment. We built on work showing that filtering of fragmentary sequences using missing data thresholds reduced gene tree estimation error (Sayyari et al. 2017). We showed that excluding species with fewer PIS than expected per locus was more effective at reducing RF distances among trees than the number of missing genes per tip. Previous studies have examined the correlation between alignment statistics and used tree-based metrics to identify the loci that may be most suitable for phylogenomic inference (Shen et al. 2016), or to quantify the properties of loci that support conflicting nodes (Burbrink et al. 2020; Singhal et al. 2021). In our study, we developed a framework that quantitatively identified which samples would most likely produce artifactual discordance in the species tree and modeled whether signal or noise was driving a conflicting topology.

We focused on subclades that largely represented genera or complexes of genera because these were the sources of most topological conflict. By examining discordance within subclades, our modeling clarified whether the phylogenetic signal favored one of the topologies, a result that would have been masked if we had examined the entire tree. Most studies that employ phylogenomic subsampling on gene interrogation do so on the full alignment and tree. These global comparisons of phylogenetic discordance across an entire tree may overlook more nuanced patterns within subclades if data quality issues bias relationships in only portions of a tree. An alternative approach would be to apply hierarchical sampling and examine nested subclades. This could be done independently or in conjunction with previously implemented gene interrogation approaches on an entire tree and alignment.

Although some samples may have too little information for species tree analyses, as evident in our study and previous work (e.g., Hosner et al. 2016), they can be of value in concatenated approaches. Much of the debate over using supermatrix versus coalescent approaches to infer phylogenies has centered on accounting for the independent evolutionary histories of gene partitions versus optimizing the signal contained in DNA alignments (Edwards et al. 2016; Springer and Gatesy 2016). Our results favor a middle ground for tree inference from alignments with highly heterogeneous data completeness. We found that concatenation was more effective at placing samples with limited data on the tree, as is evident in samples of these taxa always being placed in their appropriate genera. Because the distribution of information content among samples will be data set-specific, our general recommendation for filtering is to use thresholds that identify problematic samples and/or loci that will bias species tree estimation while not over-filtering individuals with adequate DNA sequence. The potential cost of excluding loci is that sites that resolve relationships in other portions of the tree may be dropped, which may be problematic in alignments where relationships are driven by few sites. In contrast, the more targeted approach that we applied may remove taxa that could be critical to project goals or have an accurate and stable placement in the phylogeny despite having low-quality data. Irrespective of which approach is implemented, artifactual sources of phylogenetic discordance may remain.

Collectively, our results show that neither the concatenated nor the species tree is likely to be completely correct. For several subclades, we found that when PIS favored one of the topologies, GC content was the most important variable in poorly performing models that

favored the alternative topology, and *vice versa*. The identification of GC content as an important variable likely reflects loci with shorter flanking sequences (and thus lower information content). This importance was driven by the highly variable GC content of the conserved UCE core (range: 20–100%). For the subclades where the modeling found phylogenetic signal supporting the concatenated tree, there were no existing phylogenies to compare. In *Psittacara*, where topologies differed because of the placement of taxa such as *P. brevipes*, the concatenated tree had the taxon nested within the geographic and phenotypic species complex of *P. holochlorus*, whereas in the species tree the taxon was outside of that complex. It is more likely that the species tree topology is aberrant for that subclade because PIS predicted the Δglk scores, and there were 303 to 44 loci favoring the concatenated over the species tree, respectively. *Nymphicus* had one of the best performing models in predicting RF distance between the species and gene trees. For this subclade, both phylogenomic trees agreed on its position and we tested an alternative placement based on a previous study (White et al. 2011). However, only three loci were found to support the previous studies' topology using gene-wise likelihood scores $\Delta\text{glk} < -2$) indicating strong support for our phylogenomic tree. A similar pattern was observed in the two subclades (*Psittacula* and *Cyclopsitta-Psittaculirostris*) where we enforced monophyly; only the RF distance model performed well, but there were few genes supporting the alternative topology. A range of other scenarios including sampling differences in *Touit*, *Cacatua* A, and *Bolborhynchus-Psilopsiagon-Nannopsittaca* made phylogenetic discordance in these clades incomparable to previous trees. In some cases, such as in *Pionus* (Ribas et al. 2007a) and *Pyrilia* (Eberhard and Bermingham 2005; Ribas et al. 2005), both phylogenomic trees differed from previously published trees using Sanger data. The modeling also gave insight into previously ambiguous results. In *Polytelis*, where we have provided the first complete sampling of the genus, the still surprising though previously observed nonmonophyly (Provost et al. 2018) recovered was due to the species tree finding *P. alexandrae* strongly supported as sister to the morphologically distinct genus *Aprosmictus*. Because the modeling indicates that loci with higher PIS favor the species tree topology, this relationship may represent introgression as old as 6 Ma. We are pursuing this in a separate study.

Biogeographic and Phenotypic Considerations for Evaluating Trees

Interpreting phylogenetic discordance in the context of biogeographic and phenotypic evolution provides a means of ground-truthing phylogenetic hypotheses. Several cases of phylogenetic discordance in our study involve clades distributed over tectonically dynamic regions such as New Guinea, Melanesia, Indonesia, and the Philippines, where present-day geography may be a misleading indicator of deeper history (e.g., Oliver et al. 2020). For example, *Micropsitta* (of New Guinea and surrounding archipelagos) and *Loriculus* (largely distributed from southeast Asia through New Guinea) have nodes that differ among concatenated and species trees and date to a similar period during the Late Pliocene–Early Pleistocene. Dynamic Earth history events such as tectonics and rising and falling sea levels that connect and disconnect land masses may increase gene tree discordance because of rapid divergence events and post-divergence gene flow (e.g., Peñalba et al. 2019). Phylogenetic patterns of plumage color may also offer insight into alternative topologies. *Ara*, the large-bodied Neotropical macaws, contains three general color phenotypes: consisting predominantly of green, or red, or blue and yellow species. The phylogenetic distribution of colors in *Ara* is more concordant in the concatenated tree, which showed the red species and the blue/yellow species each being monophyletic. While noting that reliance on phenotype to inform taxonomic relationships has been shown to be misleading, particularly in parrots (Joseph et al. 2020; Merwin et al. 2020), we find that color evolution in these macaws on the concatenated tree would involve single evolutionary origins of red and blue/yellow clades/species, whereas the species tree would entail multiple gains and losses of each of the three color states. The more parsimonious scenario is also preferred because the evolution of an ancestral green lineage to blue may only require a single base-pair change as observed in *M. undulatus* (Cooke et al. 2017) and is a phylogenetic pattern of color evolution observed in the lorikeets from Oceania (Merwin et al. 2020).

Future Directions

Even as the techniques used to produce genomic data from museum specimens and degraded samples improve (Raxworthy and Smith 2021), distinguishing the ratio of phylogenetic signal to noise will remain an important feature of interpreting gene tree discordance because alignment completeness can vary among any sample types. We used our approach to examine data quality, but the method could be further extended to evaluate the role of biological factors such as gene flow and variation in recombination rate that are known to cause phylogenetic discordance (e.g., Thom et al. 2018; Li et al. 2019). As phylogenomic studies move towards the use of whole genomes, there will be an increased need to identify portions of the genome that best capture evolutionary history (Thom et al. 2021) and new measures to assess support for relationships (Thomson and Brown 2022). While larger genomic data sets provide a more detailed view of evolutionary history, accounting for the sources of gene tree heterogeneity is a pressing challenge in phylogenomics.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.b5mkkwhfm.

ACKNOWLEDGMENTS

We thank the following institutions (acronyms as in Supplementary Table S1 available on Dryad) and people for providing samples: ANSP (N. Rice, J. Weckstein), AMNH (P. Sweet, T. Trombone), CSIRO/ANWC (R. Palmer, C. Wilson, R. Schodde), CM (S. Rogers), CUMV (C. Dardia), DMNH (J. Woods), FMNH (B. Marks, J. Bates, S. Hackett), KUNHM (M. Robbins), LACM (K. Garrett), LSUMZ (D. Dittman, S. Cardiff, F. Sheldon, B. Faircloth), MCZ (J. Trimble, S. Edwards), MSB (A. Johnson, C. Witt, M. Andersen), ZMUC (P. Hosner, K. Jønsson), NMS (R. McGowan), ROM (S. Claramunt, M. Peck), SDNHM (P. Unitt), SDSU (K. Burns), UFMNH (A. Kratter, D. Steadman), USNM (B. Schmidt, H. James, G. Graves), UWBM (S. Birks, R. Faucett, J. Klicka), WAM (R. Bray, R. Johnstone), WFVZ (R. Corado), LACM (K. Garrett, A. Shultz), YPM (K. Zyskowski, R. Prum), G. Amato, and WCS. We thank F. Burbrink for assistance and feedback, and we acknowledge B. Carstens, K. Halanych, Ed Braun, and anonymous reviewers for their suggestions.

FUNDING

This work was funded by the National Science Foundation awards [DEB-1655736 and DBI-2029955 to B.T.S.] and [DEB-1557053 to R.G.M.].

REFERENCES

Arcila D., Ortí G., Vari R., Armbruster J.W., Stiassny M.L., Ko K.D., Sabaj M.H., Lundberg J., Revell L.J., Betancur-R R. 2017. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. Nat. Ecol. Evol. 1:20.

Boles W.E. 1993. A new cockatoo (Psittaciformes: Cacatuidae) from the Tertiary of Riversleigh, northwestern Queensland, and an evaluation of rostral characters in the systematics of parrots. Ibis 135(1):8–18.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Borowiec M.L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4:e1660.

Bossert S., Murray E.A., Blaimer B.B., Danforth B.N. 2017. The impact of GC bias on phylogenetic accuracy using targeted enrichment phylogenomic data. Mol. Phylogenet. Evol. 111:149–157.

Braun M.P., Datzmann T., Arndt T., Reinschmidt M., Schnitker H., Bahr N., Sauer-Gürth H., Wink M. 2019. A molecular phylogeny of the genus Psittacula sensu lato (Aves: Psittaciformes: Psittacidae: Psittacula, Psittinus, Tanygnathus, †Mascarinus) with taxonomic implications. Zootaxa 4563:547–562.

Braun M.P., Reinschmidt M., Datzmann T., Waugh D., Zamora R., Häbich A., Neves L., Gerlach H., Arndt T., Mettke-Hofmann C., Sauer-Gürth H., Wink M. 2017. Influences of oceanic islands and the Pleistocene on the biogeography and evolution of two groups of Australasian parrots (Aves: Psittaciformes: Eclectus roratus, Trichoglossus haematodus complex). Rapid evolution and implications for taxonomy and conservation. Eur. J. Ecol. 3(2):47–66.

Brown J.M., Thomson R.C. 2017. Bayes factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. Syst. Biol. 66:517–530.

Burbrink F.T., Gehara, M. 2018. The biogeography of deep time phylogenetic reticulation. Syst. Biol. 67(5):743–755.

Burbrink F.T., Grazziotin F.G., Pyron R.A., Cundall D., Donnellan S., Irish F., Keogh J.S., Kraus F., Murphy R.W., Noonan B., Raxworthy C.J. 2020. Interrogating genomic-scale data for Squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. Syst. Biol. 69(3):502–520.

Burgio K.R., Davis K.E., Dreiss L.M., Cisneros L.M., Klingbeil B.T., Presley S.J., Willig M.R. 2019. Phylogenetic supertree and functional trait database for all extant parrots. Data Brief. 24:103882.

Burleigh J.G., Kimball R.T., Braun E.L. 2015. Building the avian tree of life using a large-scale, sparse supermatrix. Mol. Phylogenet. Evol. 84: 53–63.

Chernomor O., von Haeseler A., Minh B.Q. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst. Biol. 65:997–1008.

Cooke T.F., Fischer C.R., Wu P., Jiang T.X., Xie K.T., Kuo J., Doctorov E., Zehnder A., Khosla C., Chuong C.M., Bustamante C.D. 2017. Genetic mapping and biochemical basis of yellow feather pigmentation in budgerigars. Cell 171(2):427–439.

Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., McCarthy S.A., Davies R.M., Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10(2):giab008.

Eaton D.A., Hipp A.L., González-Rodríguez A., Cavender-Bares J. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution 69(10):2587–2601.

Eberhard J.R., Bermingham E. 2005. Phylogeny and comparative biogeography of Pionopsitta parrots and Pteroglossus toucans. Mol. Phylogenet. Evol. 36(2):288–304.

Edwards S.V., Xi Z., Janke A., Faircloth B.C., McCormack J.E., Glenn T.C., Zhong B., Wu S., Lemmon E.M., Lemmon A.R., Leaché A.D. 2016. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. Mol. Phylogenet. Evol. 94:447–462.

Faircloth B.C. 2013. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. Available at: http://dx.doi.org/10.6079/J9ILL.

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788.

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

FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3(6):1084–1092.

Galtier N., Daubin V. 2008. Dealing with incongruence in phylogenomic analyses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363(1512):4023–4029.

Ganapathy G., Howard J.T., Ward J.M., Li J., Li B., Li Y., Xiong Y., Zhang Y., Zhou S., Schwartz D.C., Schatz M. 2014. High-coverage sequencing and annotated assemblies of the budgerigar genome. GigaScience 3(1):2047–217X–3–11.

Gatesy J., Sloan D.B., Warren J.M., Baker R.H., Simmons M.P., Springer M.S. 2019. Partitioned coalescence support reveals biases in species-tree methods and detects gene trees that determine phylogenomic conflicts. Mol. Phylogenet. Evol. 139:106539.

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

Harvey M.G., Smith B.T., Glenn T.C., Faircloth B.C., Brumfield R.T. 2016. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. Syst. Biol. 65:910–924.

Hime P.M., Lemmon A.R., Lemmon E.C.M., Prendini E., Brown J.M., Thomson R.C., Kratovil J.D., Noonan B.P., Pyron R.A., Peloso P.L., Kortyna M.L. 2021. Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. Syst. Biol. 70(1):49–66.

Hosner P.A., Faircloth B.C., Glenn T.C., Braun E.L., Kimball R.T. 2016. Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Mol. Biol. Evol. 33(4):1110–1125.

Jarvis E.D., Mirarab S., Aberer A.J., Li B., Houde P., Li C., Ho S.Y., Faircloth B.C., Nabholz B., Howard J.T., Suh A., Weber C.C., da Fonseca R.R., Li J., Zhang F., Li H., Zhou L., Narula N., Liu L., Ganapathy G., Boussau B., Bayzid M.S., Zavidovych V., Subramanian S., Gabaldon T., Capella-Gutierrez S., Huerta-Cepas J.,

Rekepalli B., Munch K., Schierup M., Lindow B., Warren W.C., Ray D., Green R.E., Bruford M.W., Zhan X., Dixon A., Li S., Li N., Huang Y., Derryberry E.P., Bertelsen M.F., Sheldon F.H., Brumfield R.T., Mello C.V., Lovell P.V., Wirthlin M., Schneider M.P., Prosdocimi F., Samaniego J.A., Vargas Velazquez A.M., Alfaro-Nunez A., Campos P.F., Petersen B., Sicheritz-Ponten T., Pas A., Bailey T., Scofield P., Bunce M., Lambert D.M., Zhou Q., Perelman P., Driskell A.C., Shapiro B., Xiong Z., Zeng Y., Liu S., Li Z., Liu B., Wu K., Xiao J., Yinqi X., Zheng Q., Zhang Y., Yang H., Wang J., Smeds L., Rheindt F.E., Braun M., Fjeldsa J., Orlando L., Barker F.K., Jonsson K.A., Johnson W., Koepfli K.P., O'Brien S., Haussler D., Ryder O.A., Rahbek C., Willerslev E., Graves G.R., Glenn T.C., McCormack J., Burt D., Ellegren H., Alstrom P., Edwards S.V., Stamatakis A., Mindell D.P., Cracraft J., Braun E.L., Warnow T., Jun W., Gilbert M.T., Zhang G. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346:1320–1331.

Jetz W., Thomas G.H., Joy J.B., Hartmann K., Mooers A.O. 2012. The global diversity of birds in space and time. Nature 491(7424):444–448.

Joseph L., Dolman G., Saint K., Donnellan S., Berg M., Bennett A. 2008. Where and when does a ring start and end? Testing the ring species hypothesis in a species complex of Australian parrots. Proc. Biol. Sci. 275:2431–2440.

Joseph L., Merwin J., Smith B.T. 2020. Improved systematics of lorikeets reflects their evolutionary history and frames conservation priorities. Emu 120:201–215.

Joseph L., Toon A., Schirtzinger E., Wright T. 2011. Molecular systematics of two enigmatic genera *Psittacella* and *Pezoporus* illuminate the ecological radiation of Australo-Papuan parrots. Mol. Phylogenet. Evol. 59:675–684.

Junier T., Zdobnov E.M. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. Bioinformatics 26:1669–1670.

Kalyaanamoorthy S., Minh B.Q., Wong T.K., von Haeseler A., Jermiin L.S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14(6):587.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol 30(4):772–780.

Kimball R.T., Oliveros C.H., Wang N., White N.D., Barker F.K., Field D.J., Ksepka D.T., Chesser R.T., Moyle R.G., Braun M.J., Brumfield R.T. 2019. A phylogenomic supertree of birds. Diversity 11(7):109.

Kirchman J.J., Schirtzinger E.E., Wright T.F. 2012. Phylogenetic relationships of the extinct Carolina Parakeet (*Conuropsis carolinensis*) inferred from DNA sequence data. Auk 129:197–204.

Kuhn M. 2008. Caret package. J. Stat. Softw. 28:1–26.

Leuchtenberger A.F., Crotty S.M., Drucks T., Schmidt H.A., Burgstaller-Muehlbacher S., von Haeseler A. 2020. Distinguishing Felsenstein zone from Farris zone using neural networks. Mol. Biol. Evol. 37(12):3632–3641.

Li G., Davis B.W., Eizirik E., Murphy W.J. 2016. Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). Genome Res. 26(1):1–11.

Li G., Figueiró H.V., Eizirik E., Murphy W.J. 2019. Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat species. Mol. Biol. Evol. 36(10):2111–2126.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–60.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25:2078–2079.

Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46(3):523–536.

Mayr G., Manegold A. 2006. New specimens of the earliest European passeriform bird. Acta Palaeontol. Pol. 51:315–323.

McCormack J.E., Tsai W.L., Faircloth B.C. 2016. Sequence capture of ultraconserved elements from bird museum specimens. Mol. Ecol. Resour. 16(5):1189–1203.

Merwin J.T., Seeholzer G.F., Smith B.T. 2020. Macroevolutionary bursts and constraints generate a rainbow in a clade of tropical birds. BMC Evol. Biol. 20:32.

Minh B.Q., Schmidt H.A., Cernomor O., Scrempg D., Woodhams M.D., Haeseler A., Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37:1530–1534.

Mirarab S., Bayzid M.S., Warnow T. 2016. Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. Syst. Biol. 65(3):366–380.

Molloy E.K., Warnow T. 2018. To include or not to include: the impact of gene filtering on species tree estimation methods. Syst. Biol. 67(2):285–303.

Mongiardino Koch, N. 2021. Phylogenomic subsampling and the search for phylogenetically reliable loci. Mol. Biol. Evol. 38(9):4025–4038.

Morales-Briones D.F., Kadereit G., Tefarikis D.T., Moore M.J., Smith S.A., Brockington S.F., Timoneda A., Yim W.C., Cushman J.C., Yang Y. 2021. Disentangling sources of gene tree discordance in phylogenomic data sets: testing ancient hybridizations in Amaranthaceae sl. Syst. Biol. 70(2):219–235.

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

Neuwirth E., Neuwirth M.E. 2011. Package 'RColorBrewer'. CRAN 2011–06-17 08: 34: 00. Apache License 2.0.

Oleksyk T.K., Pombert J.F., Siu D., Mazo-Vargas A., Ramos B., Guiblet W., Afanador Y., Ruiz-Rodriguez C.T., Nickerson M.L., Logue D.M., Dean M. 2012. A locally funded Puerto Rican parrot (*Amazona vittata*) genome sequencing project increases avian data and advances young researcher education. GigaScience 1(1):2047–217X-1–14.

Oliver P.M., Heiniger H., Hugall A.F., Joseph L., Mitchell K.J. 2020. Oligocene divergence of frogmouth birds (Podargidae) across Wallace's Line. Biol. Lett. 16(5):20200040.

Oliveros C.H., Field D.J., Ksepka D.T., Barker F.K., Aleixo A., Andersen M.J., Alström P., Benz B.W., Braun E.L., Bravo G.A., Brumfield R.T., Chesser R.T., Claramunt S., Cracraft J., Cuervo A.M., Derryberry E.P., Glenn T.C., Harvey M.G., Hosner P.A., Joseph L., Kimball R.T., Mack A.L., Miskelly C.M., Peterson A.T., Robbins M.B., Sheldon F.H., Silveira L.F., Smith B.T., White N.D., Moyle R.G., Faircloth B.C. 2019. Earth history and the passerine superradiation. Proc. Natl. Acad. Sci. USA 116:7916–7925.

Page A.J., Taylor B., Delaney A.J., Soares J., Seemann T., Keane J.A., Harris S.R. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb. Genom. 2(4):e000056.

Paradis E., Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35(3):526–528.

Peñalba J.V., Joseph L., Moritz C. 2019. Current geography masks dynamic history of gene flow during speciation in northern Australian birds. Mol. Ecol. 28(3):630–643.

Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro M.E., Harmon L.J. 2014. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics 30(15):2216–2218.

Pepperberg I.M. 2009. The Alex studies: cognitive and communicative abilities of grey parrots. Cambridge (MA): Harvard University Press.

Provost K., Joseph L., Smith B.T. 2018. Resolving a phylogenetic hypothesis for parrots: implications from systematics to conservation. Emu 118:7–21.

R Core Team. 2019. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: URL http://www.R-project.org/.

Raxworthy C.J., Smith B.T. 2021. Mining museums for historical DNA: advances and challenges in museomics. Trends Ecol. Evol. 36(11):1049–1060.

Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3(2):217–223.

Ribas C.C., Gaban-Lima R., Miyaki C.Y., Cracraft J. 2005. Historical biogeography and diversification within the Neotropical parrot genus *Pionopsitta* (Aves: Psittacidae). J. Biogeogr. 32:1409–1427.

Ribas C.C., Joseph L., Miyaki C.Y. 2006. Molecular systematics and patterns of diversification in *Pyrrhura* (Psittacidae), with special reference to the *picta-leucotis* complex. Auk 123:660–680.

Ribas C.C., Miyaki C.Y. 2004. Molecular systematics in Aratinga parakeets: species limits and historical biogeography in the '*solstitialis*' group, and the systematic position of *Nandayus nenday*. Mol. Phylogenet. Evol. 30:663–675.
Ribas C.C., Miyaki C.Y., Cracraft J. 2009. Phylogenetic relationships, diversification and biogeography in Neotropical *Brotogeris* parakeets. J. Biogeogr. 36:1712–1729.
Ribas C.C., Moyle R.G., Miyaki C.Y., Cracraft J. 2007a. The assembly of montane biotas: linking Andean tectonics and climatic oscillations to independent regimes of diversification in *Pionus* parrots. Proc. Biol. Sci. 274:2399–2408.
Ribas C.C., Tavares E.S., Yoshihara C., Miyaki C.Y. 2007b. Phylogeny and biogeography of Yellow-headed and Blue-fronted Parrots (*Amazona ochrocephala* and *Amazona aestiva*) with special reference to the South American taxa. Ibis 149:564–574.
Robinson D.F., Foulds L.R. 1981. Comparison of phylogenetic trees. Math. Biosci. 53(1–2):131–147.
Sayyari E., Whitfield J.B., Mirarab S. 2017. Fragmentary gene sequences negatively impact gene tree and species tree reconstruction. Mol. Biol. Evol. 34(12):279–3291.
Schliep K.P. 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27(4):592–593.
Schweizer M., Güntert M., Hertwig S.T. 2012. Out of the Bassian province: historical biogeography of the Australasian platycercine parrots (Aves, Psittaciformes). Zool. Scr. 42:13–27.
Schweizer M., Hertwig S.T., Seehausen O. 2014. Diversity versus disparity and the role of ecological opportunity in a continental bird radiation. J. Biogeogr. 41:1301–1312.
Schweizer M., Wright T.F., Peñalba J.V., Schirtzinger E.E., Joseph L. 2015. Molecular phylogenetics suggests a New Guinean origin and frequent episodes of founder-event speciation in the nectarivorous lories and lorikeets (Aves: Psittaciformes). Mol. Phylogenet. Evol. 90:34–48.
Sharma P.P., Kaluziak S.T., Pérez-Porro A.R., González V.L., Hormiga G., Wheeler W.C., Giribet G. 2014. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. Mol. Biol. Evol. 31(11):2963–2984.
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(5):1–10.
Shen X.X., Salichos L., Rokas A. 2016. A genome-scale investigation of how sequence, function, and tree-based gene properties influence phylogenetic inference. Genome Biol. Evol. 8(8):2565–2580.
Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51(3):492–508.
Simmons M.P., Gatesy J. 2021. Collapsing dubiously resolved gene-tree branches in phylogenomic coalescent analyses. Mol. Phylogenet. Evol. 158:107092.
Simmons M.P., Kessenich J. 2020. Divergence and support among slightly suboptimal likelihood gene trees. Cladistics 36:322–340.
Simmons M.P., Sloan D.B., Gatesy J. 2016. The effects of subsampling gene trees on coalescent methods applied to ancient divergences. Mol. Phylogenet. Evol. 97:76–89.
Simpson J.T., Wong K., Jackman S.D., Schein J.E., Jones S.J., Birol I. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19(6):1117–1123.
Singhal S., Colston T.J., Grundler M.R., Smith S.A., Costa G.C., Colli G.R., Moritz C., Pyron R.A., Rabosky D.L. 2021. Congruence and conflict in the higher-level phylogenetics of squamate reptiles: an expanded phylogenomic perspective. Syst. Biol. 70(3):542–557.
Smith B.T., Gehara M., Harvey M.G. 2021. The demography of extinction in eastern North American birds. Proc. Biol. Sci. 288(1944):20201945.
Smith B.T., Mauck III W.M., Benz B.W., Andersen M.J. 2020. Uneven missing data skew phylogenomic relationships within the lories and lorikeets. Genome Biol. Evol. 12(7):1131–1147.
Smith B.T., Ribas C.C., Whitney B.M., Hernández-Baños B.E., Klicka J. 2013. Identifying biases at different spatial and temporal scales of diversification: a case study in the Neotropical parrotlet genus *Forpus*. Mol. Ecol. 22:483–494.
Smith M.R. 2020. Information theoretic generalized Robinson–Foulds metrics for comparing phylogenetic trees. Bioinformatics 36(20):5007–5013.
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(1):150.
Smith S.A., O'Meara B.C. 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28:2689–2690.
Snyder N.F., McGowan P., editors. 2000. Parrots: status survey and conservation action plan 2000–2004. IUCN.
Springer M.S., Gatesy J. 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94:1–33.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.
Suvorov A., Hochuli J., Schrider D.R. 2020. Accurate inference of tree topologies from multiple sequence alignments using deep learning. Syst. Biol. 69(2):221–233.
Tavares E.S., Baker A.J., Pereira S.L., Miyaki C.Y. 2006. Phylogenetic relationships and historical biogeography of Neotropical parrots (Psittaciformes: Psittacidae: Arini) inferred from mitochondrial and nuclear DNA sequences. Syst. Biol. 55:454–70.
Thom G., Amaral F.R.D., Hickerson M.J., Aleixo A., Araujo-Silva L.E., Ribas C.C., Choueri E., Miyaki C.Y. 2018. Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species. Syst. Biol. 67(4):700–718.
Thom G., Moreira L.R., Batista R., Gehara M., Aleixo A., Smith B.T. 2021. Genomic architecture controls spatial structuring in Amazonian birds. bioRxiv, doi: 10.1101/2021.12.01.470789.
Thomson R.C., Brown J.M. 2022. On the need for new measures of phylogenomic support. Syst. Biol. 71:917–920.
Yu G., Smith D.K., Zhu H., Guan Y., Lam T.T.Y. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8(1):28–36.
Walker J.F., Brown J.W., Smith S.A. 2018. Analyzing contentious relationships and outlier genes in phylogenomics. Syst. Biol. 67(5):916–924.
Weidig I. 2010. New birds from the Lower Eocene Green River Formation, North America. In: Boles W.E., Worthy, T.H., editors. Proceedings of the VII International Meeting of the Society of Avian Paleontology and Evolution. Rec Aust Museum, vol. 62. p. 29–44.
White N.E., Phillips M.J., Gilbert T., Alfaro-Núñez A., Willerslev E., Mawson P.R., Spencer P.B.S., Bunce M. 2011. The evolutionary history of cockatoos (Aves: Psittaciformes: Cacatuidae). Mol. Phylogenet. Evol. 59:615–622.
Worthy T.H., Tennyson A.J.D., Scofield R.P. 2011. An early Miocene diversity of parrots (Aves, Strigopidae, Nestorinae) from New Zealand. J. Vertebr. Paleontol. 31:1102–1116.
Wright T.F., Schirtzinger E.E., Matsumoto T., Eberhard J.R., Graves G.R., Sanchez J.J., Capelli S., Muller H., Scharpegge J., Chambers G.K., Fleischer R.C. 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the cretaceous. Mol. Biol. Evol. 25:2141–2156.
Zhang C., Sayyari E., Mirarab S. 2017. ASTRAL-III: increased scalability and impacts of contracting low support branches. In: RECOMB international workshop on comparative genomics. Cham: Springer. p. 53–75.
Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19(6):15–30.
Zou Z., Zhang H., Guan Y., Zhang J. 2020. Deep residual neural networks resolve quartet molecular phylogenies. Mol. Biol. Evol. 37(5):1495–1507.