Molecular Phylogenetics and Evolution 155 (2021) 107013

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogenomics of manakins (Aves: Pipridae) using alternative locus filtering strategies based on informativeness

Rafael N. Leite [a,*], Rebecca T. Kimball [b], Edward L. Braun [b], Elizabeth P. Derryberry [c], Peter A. Hosner [d], Graham E. Derryberry [e], Marina Anciães [f], Jessica S. McKay [g], Alexandre Aleixo [h,i], Camila C. Ribas [f], Robb T. Brumfield [e], Joel Cracraft [g]

[a] Department of Genetics and Evolutionary Biology, University of São Paulo, São Paulo, SP, Brazil
[b] Department of Biology, University of Florida, Gainesville, FL, USA
[c] Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA
[d] Natural History Museum of Denmark and Center for Macroecology, Evolution and Climate, University of Copenhagen, Copenhagen, Denmark
[e] Museum of Natural Science and Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA
[f] Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil
[g] Department of Ornithology, American Museum of Natural History, New York, NY, USA
[h] Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland
[i] Department of Zoology, Museu Paraense Emílio Goeldi, Belém, PA, Brazil

ARTICLE INFO

Keywords:
Ultraconserved elements
Exon probes
Informative loci
Concatenation
Multispecies coalescent
Tyrannides

ABSTRACT

Target capture sequencing effectively generates molecular marker arrays useful for molecular systematics. These extensive data sets are advantageous where previous studies using a few loci have failed to resolve relationships confidently. Moreover, target capture is well-suited to fragmented source DNA, allowing data collection from species that lack fresh tissues. Herein we use target capture to generate data for a phylogeny of the avian family Pipridae (manakins), a group that has been the subject of many behavioral and ecological studies. Most manakin species feature lek mating systems, where males exhibit complex behavioral displays including mechanical and vocal sounds, coordinated movements of multiple males, and high speed movements. We analyzed thousands of ultraconserved element (UCE) loci along with a smaller number of coding exons and their flanking regions from all but one species of Pipridae. We examined three different methods of phylogenetic estimation (concatenation and two multispecies coalescent methods). Phylogenetic inferences using UCE data yielded strongly supported estimates of phylogeny regardless of analytical method. Exon probes had limited capability to capture sequence data and resulted in phylogeny estimates with reduced support and modest topological differences relative to the UCE trees, although these conflicts had limited support. Two genera were paraphyletic among all analyses and data sets, with *Antilophia* nested within *Chiroxiphia* and *Tyranneutes* nested within *Neopelma*. The *Chiroxiphia–Antilophia* clade was an exception to the generally high support we observed; the topology of this clade differed among analyses, even those based on UCE data. To further explore relationships within this group, we employed two filtering strategies to remove low-information loci. Those analyses resulted in distinct topologies, suggesting that the relationships we identified within *Chiroxiphia–Antilophia* should be interpreted with caution. Despite the existence of a few continuing uncertainties, our analyses resulted in a robust phylogenetic hypothesis of the family Pipridae that provides a comparative framework for future ecomorphological and behavioral studies.

* Corresponding author at: Programa de Pós-Graduação em Ecologia, Instituto Nacional de Pesquisas da Amazônia, Avenida André Araújo 2936, Petrópolis, Manaus, AM 69.067-375, Brazil.
E-mail addresses: rnleite@gmail.com (R.N. Leite), rkimball@ufl.edu (R.T. Kimball), ebraun68@ufl.edu (E.L. Braun), ederryb1@utk.edu (E.P. Derryberry), peter.hosner@snm.ku.dk (P.A. Hosner), alexandre.aleixo@helsinki.fi (A. Aleixo), robb@lsu.edu (R.T. Brumfield), jlc@amnh.org (J. Cracraft).

https://doi.org/10.1016/j.ympev.2020.107013
Received 13 August 2019; Received in revised form 7 November 2020; Accepted 11 November 2020
Available online 17 November 2020

1. Introduction

The field of phylogenetics has significantly advanced due to multilocus inferences using next-generation sequencing (NGS) technologies (McCormack and Faircloth, 2013). Massive-parallel sequencing can be employed via target capture protocols (Mamanova et al., 2010) to generate new data for hundreds or thousands of unlinked loci (McCormack et al., 2013). This increase in the number of independent markers suitable for phylogenetic studies has revolutionized molecular systematics and the tree of life (e.g., Kimball et al., 2019; Prum et al., 2015). A commonly used class of nuclear markers is ultraconserved elements, or UCEs (Faircloth et al., 2012). UCEs constitute highly conserved orthologous segments found across the genome of distinct vertebrates (Bejerano et al., 2004), characterized by flanking regions with more variable sites that can be used to investigate historical relationships at deep and shallow taxonomic levels (Faircloth et al., 2012).

Many phylogenomic studies have used concatenation, in which sequences of all genes are combined for each taxon and analyzed as single sequences in a supermatrix with all taxa. Analyses of concatenated data are computationally efficient (e.g., RAxML; Stamatakis, 2014) and provide intuitive measures of branch lengths. However, concatenation has received criticism (see Braun et al., 2019; Edwards et al., 2016 for review) because it assumes all genes share the same underlying history; this is not expected to be the case due to phenomena such as incomplete lineage sorting (ILS). Indeed, ILS has two important implications for phylogenetic analyses. First, the potential for individual gene trees to conflict with the species tree suggests that it may be necessary to collect data for a large number of loci to obtain an accurate estimate of the species tree. Second, it is possible that the most probable gene trees differ from the species tree when the internal branches of a phylogeny are specially short relative to effective population size; this occurs in a part of parameter space called the anomaly zone (Degnan and Rosenberg, 2006). Under this circumstance, maximum likelihood (ML) analyses of concatenated data may yield high support for an incorrect topology (Kubatko and Degnan, 2007; Mendes and Hahn, 2018; Roch and Steel 2015).

Although theory indicates that methods of phylogenetic inference that account for ILS given the multispecies coalescent (MSC) are consistent estimators of the species tree, the performance of MSC methods in empirical settings, particularly "summary" methods that rely on analyses of individual gene trees, has been the subject of intense debate (Gatesy and Springer, 2014; Meiklejohn et al., 2016; Patel et al., 2013; Song et al., 2012; Springer and Gatesy, 2016). Additionally, some MSC methods circumvent gene tree estimation by extracting phylogenetic signal directly from site patterns in a sequence data matrix (Chifman and Kubatko, 2014; Chou et al., 2015). Despite the debate, simulation and empirical studies find similar relationships for the majority of nodes under both frameworks (Chen et al., 2015; Hosner et al., 2016; Pyron et al., 2014; Tonini et al., 2015), though analyses using multiple approaches can be important to identify relationships that might need further study.

Manakins constitute a family (Pipridae) of small suboscine passerine birds characterized by a number of unique behaviors and morphological features (Kirwan and Green, 2012). Piprids have their greatest diversity in lowland Neotropical humid forests, but some taxa occur in dry woodlands, along riparian forests, and in montane forests (Anciães and Peterson, 2009; Kirwan and Green, 2012). Most species have strong sexually dimorphic plumage and elaborate lekking courtship rituals that may include displays involving coordinated movement between multiple males, mechanical and vocal sounds, and high speed movements. Thus, manakins have been the focus of many behavioral studies.

The Pipridae family is a well-supported clade including 53 named species that have been divided into 17 genera (Gill and Donsker, 2018). Early phylogenetic hypotheses were based on syringeal morphology, lek-display behavior and sexual plumage traits (Prum, 1990, 1992, 1994, 1997). These analyses suggested the "tyrant" manakins *Neopelma* and *Tyranneutes* should not be members of Pipridae. Subsequent molecular phylogenetic studies using small numbers of loci have identified a well-supported Pipridae that includes the two genera of tyrant manakins (Barber et al., 2007; Chesser, 2004; McKay et al., 2010; Ohlson et al., 2008; Ohlson et al., 2013a; Ohlson et al., 2013b; Tello et al., 2009). Although none of these works includes complete taxon sampling, the studies highlight genera that were not monophyletic, which has led to new genera and changed generic circumscription within the family. Two recent studies, more narrowly focused on specific clades, further suggest two additional genera that are not monophyletic: *Neopelma* (Capurucho et al., 2018) and *Chiroxiphia* (Silva et al., 2018). However, these former assessments used a limited number of markers, relationships among a number of key taxa remain unclear, and many species are still not included in any molecular phylogeny. This absence of a robust phylogenetic hypothesis for the family limits comparative studies that might leverage our current knowledge on the ecology, behavior and traits in this group.

Our goal is to advance a robust phylogenetic hypothesis for Pipridae that can be used for future macroevolutionary studies on these fascinating neotropical birds. To fully understand relationships within the family, we generated data for thousands of UCE loci and employed several different analytical approaches: standard concatenation and two MSC methods (one summary method and one the site-pattern method). To further explore relationships not strongly supported in our initial analyses, we also used locus filtering to remove loci with low phylogenetic information (Chen et al., 2015; Molloy and Warnow, 2018). We filtered these data because it has been suggested that low-information loci may render inaccurate estimates, and may compromise species tree analyses especially for summary MSC methods (Meiklejohn et al., 2016; Xi et al., 2015). We expected to observe improved congruence for recalcitrant relationships if the locus filtering methods we used were able provide robust phylogenetic inferences across analyses. We suggest that those relationships for which contrasting results emerge from alternative approaches of species tree estimation and locus filtering should be interpreted with caution and be the focus of future studies. Notwithstanding a few uncertain relationships for difficult nodes of the Pipridae phylogeny, the majority of relationships received strong support among UCE analyses and most were congruent when data from exonic regions were analyzed.

2. Materials and methods

2.1. Taxon sampling

We obtained samples for a total of 51 taxa within the family Pipridae (Tello et al., 2009), including almost all currently recognized species except for *Neopelma aurifrons* and the newly described species *Machaeropterus eckelberryi* (Lane et al., 2017). We also sampled three additional taxa (*Pyroderus scutatus*, *Onychorhynchus coronatus*, *Pachyramphus minor*) as representative genera of closely related families (Cotingidae, Tyrannidae and Tityridae, respectively). Most samples came from freshly preserved tissue or blood, but we also successfully sequenced two samples from museum specimens. Voucher numbers and institutions are listed in the Supplementary material (Table A.1).

2.2. Library preparation, target enrichment and sequencing

We extracted total DNA from samples using Qiagen DNeasy Blood & Tissue Kits (Qiagen, Valencia, CA, USA). Sequence data were obtained by RAPiD Genomics (Gainesville, FL, USA) following methods detailed in Faircloth et al. (2012) with minor modifications. Briefly, the sequence-capture workflow involved preparation of Illumina TruSeq libraries using the manufacturer's protocols (Illumina Inc., San Diego, CA, USA) and primers with custom index tags for multiplexing. We enriched each library using a set of 4,715 custom probes (MYbaits, MYcroarray, Ann Arbor, MI, USA) targeting 49 exons plus 2,320 UCE

loci with 100-nt paired-end reads sequenced on an Illumina HiSeq 2500 (Harvey et al., 2017). Raw sequence data are archived on NCBI databases under BioProject Accession PRJNA655842.

2.3. Bioinformatic preprocessing

After massive parallel sequencing, we de-multiplexed the raw reads in fastq format and removed adapter contamination and low-quality bases from reads using Illumiprocessor (Faircloth, 2013) as a parallel wrapper for Trimmomatic (Bolger et al., 2014). We processed the cleaned read-data following standard bioinformatic pipelines implemented in Phyluce (Faircloth, 2016). We assembled the contigs using Trinity r2013-02-25 (Grabherr et al., 2011), then extracted sequences from those contigs matching targeted loci (UCE or exon probes), and discarded as putative duplicates the same contigs matching probes designed for multiple loci or multiple contigs matching probes for the same locus. We performed sequence alignments in parallel across all loci using MAFFT (Katoh and Standley, 2013) with the default edge-trimming settings of Phyluce.

2.4. Data sets

We assembled a total of seven different data sets (available from Zenodo repository at doi: https://doi.org//10.5281/zenodo.4118662). Our three concatenated data sets (Table 1) included: 1) all UCE loci in which at least 75% of species were sampled; 2) all UCE loci in which at least 95% of species were sampled; and 3) the exon loci (which also included flanking non-coding sequences). The exon loci provided an independent estimate of the manakin phylogeny using a different data type, while the two UCE alignments provided one data set with little missing data (95%) and one with more loci but also more missing data (75%).

For the MSC analyses, we focused just on UCEs as relatively few exons were assembled. To ensure we included an outgroup in all gene trees, we retained only those UCE loci that included the outgroup *Pyroderus scutatus* (though many of these loci also included the other outgroup taxa as well, and all outgroups were retained when present). For these filtered data sets, we used two different strategies to further filter low information loci prior to analyses based on parsimony informative sites (Table 2). While longer loci tend to have slightly more informative sites (see Table 2), our goal was to identify loci with more information (regardless of length). To do this, we employed two basic strategies: one "inclusive" filtering that identified the loci to retain based on the number of parsimony informative sites in the data set as a whole; and a second "clade-specific" filtering that used the number of parsimony informative sites in a specific clade (comprising *Chiroxiphia* and *Antilophia* species) to identify the loci to retain with the potential to be the most informative for resolving relationships in a problematic group (see Section 3.2.3 for justification). The first inclusive filtering strategy resulted in the following data sets: 4) all informative loci (those with at least one parsimony informative site); and 5) the 25% most informative loci. In our second clade-specific filtering strategy, we calculated the number of parsimony informative sites in alignments that just included the *Chiroxiphia* and *Antilophia* taxa, and then applied that information to alignments containing all taxa to generate two additional data sets: 6) *Chiroxiphia–Antilophia* all informative loci (those with at least one parsimony informative site in the *Chiroxiphia–Antilophia* clade); and 7) the *Chiroxiphia–Antilophia* 25% most informative loci (those 25% most informative for the *Chiroxiphia–Antilophia* clade).

Since UCEs had a better capture rate than the exon loci, we restricted our comparisons of concatenation and MSC methods (with locus filtering) to the UCEs. We used Phyluce (Faircloth, 2016) for data management, including alignment filtering and to compute numbers of parsimony informative sites.

2.5. Phylogenomic analyses

2.5.1. Data partitioning

The concatenated 75% and 95% complete UCE matrices and the exon matrix were each used as input to select the best partitioning scheme for each data set in PartitionFinder 2 (Lanfear et al., 2017). The exon loci included coding regions targeted by the probes as well as flanking intronic and untranslated region (UTR) sequences. We defined separate data blocks within each exon locus based on the three codon positions for coding regions and on intron or UTR for the associated non-coding regions. UCE data blocks were defined by locus. We applied the relaxed hierarchical clustering algorithm (Lanfear et al., 2014) using default weights and percentage of schemes analyzed, with the maximum number of subsets set to 100, and estimated a maximum parsimony starting tree and unlinked branch lengths in RAxML (Stamatakis, 2014) using the general time reversible (GTR) model with gamma distribution for rate heterogeneity (+G). We used the Bayesian information criterion (BIC) to select the best partitioning scheme among three model options (GTR; GTR+G; or GTR+G+I, with a proportion of invariable sites) available under these settings in PartitionFinder 2.

2.5.2. Concatenated analyses

We performed standard concatenated analyses using the first three data sets (those without filtering loci by parsimony informativeness): the 75% and 95% UCE data sets, and the exon data set (data sets 1–3 above). We conducted ML inferences obtained by concatenation of unpartitioned and partitioned data sets using RAxML under the GTR+G model, with *Pyroderus scutatus* as the outgroup and 20 initial random trees. We assessed nodal support via the *autoMRE* option to generate bootstrap replicates until convergence was reached and to draw bipartitions onto the best-scoring ML tree.

2.5.3. MSC analyses

We also estimated ML gene trees and 100 bootstrapped gene tree replicates for each locus under these settings in RAxML. These estimated gene trees (available at doi: https://doi.org//10.5281/zenodo.4118662) were used as input for the MSC gene tree reconciliation program ASTRAL-II (Mirarab and Warnow, 2015). We assessed branch support in two different ways. First, we conducted 100 bootstrap replicates resampling by locus and by site (Seo, 2008), and computed a greedy consensus tree from bootstrapped species trees. Second, we used the local posterior probabilities of branch support based on quartet frequencies (Sayyari and Mirarab, 2016). We ran ASTRAL on each of the four filtered data sets (data sets 4–7).

We also evaluated an MSC approach that takes input directly from

Table 1
Summary statistics for UCE and exon data sets used in standard concatenated analyses.

Data set	Data set number	Total number of loci	Average locus length	Total number of parsimony informative sites	Average number of parsimony informative sites per locus
UCE 75% complete loci	1	2,237	639	63,741	28
UCE 95% complete loci	2	1,796	653	52,642	29
Exon loci	3	36	955	1,875	52

Table 2
Summary statistics comparing UCE data sets under different inclusive and clade-specific filtering schemes. UCEs were filtered based on the number of parsimony informative sites for all taxa and for clade-specific taxa. Numbers within each scheme separated by a slash symbol correspond to values calculated for the entire alignments (before slash) or alignments including only *Chiroxiphia–Antilophia* taxa (after slash), respectively.

Filtering scheme	Data set number	Total number of loci	Average locus length	Total number of parsimony informative sites	Average number of parsimony informative sites per locus
All informative loci	4	2,062	640.1	59,9803,966	29.1/1.9
25% most informative loci	5	520	665.1	28,532/1,888	54.9/3.6
Chiroxiphia–Antilophia All informative loci	6	1,516	644.1	50,774/3,966	33.5/2.6
Chiroxiphia–Antilophia 25% most informative loci	7	600	655.7	26,298/2,691	43.8/4.5

the concatenated sequence data (SVDquartets). SVDquartets (Chifman and Kubatko, 2014, 2015) computes singular value decomposition scores to infer relationships among quartets of taxa and then estimates the species tree by assembling the collection of quartet splits. SVDquartets analyses were implemented in PAUP* (Swofford, 2017) using 100,000 random quartets and we computed a 50% majority-rule consensus tree from 100 bootstrap replicates as measure of uncertainty. This was also run on each of the four filtered data sets.

To allow a direct comparison between concatenation and MSC analyses, each of the four filtered data sets were also concatenated and an unpartitioned analysis in RAxML was performed using the same settings as above.

3. Results

3.1. Sequence data

After we trimmed the raw data for adapter contamination and low-quality bases we obtained an average of 5,671,732 sequence reads per taxon (95% confidence interval [CI]: ±952298), with an average length of 97.4 base pairs (bp) (95% CI: ±2.1) (Table A2). The cleaned reads were assembled into an average of 11,438 contigs (95% CI: ±6753), with an average length of 492 bp (95% CI: ±29) and an average sequencing coverage of 31× (95% CI: ±9) (Table A3). The UCE data sets of 75% and 95% completeness contained an average of 52 and 53 taxa (out of 54 taxa) per locus, respectively, and the best partition schemes included 14 and 12 subsets, respectively. The UCE data comprised 2,314 (99.7%) of the 2,320 UCE loci targeted by probes.

Sequence data captured using the exon probes averaged 28 taxa per locus (ranging from 10 to 43 taxa). The "exon" data set comprised 36 (73.5%) of the 49 loci targeted by probes and contained the coding regions targeted by probes along with flanking non-coding regions; for simplicity we refer to these regions as "exon loci" since their sequencing reflects the use of exon probes. We note that the PSMA2 locus was not assembled as a single contig, instead it was captured as two non-contiguous segments. The aligned exon data set included 15,668 coding sites (45.6%), 16,256 intron sites (47.3%), and 2,227 UTR sites (6.5%). The best partition scheme selected for the exon data set included five subsets.

3.2. Phylogenomics

3.2.1. Concatenation

The topologies obtained by ML analyses of the concatenated data sets of UCE loci with 75% and 95% completeness using both unpartitioned and partitioned analyses were completely congruent among all four inferences (Fig. 1). For higher-level relationships, our results included the nominal subfamilies Neopelminae and Piprinae as clades A and B, respectively (Fig. 1) with high support. In addition, our results suggest sub-clades B1 (*Ilicura*, *Masius*, *Corapipo*, *Chiroxiphia* and *Antilophia*) and B2 (*Xenopipo*, *Chloropipo*, *Cryptopipo*, *Lepidothrix*, *Heterocercus*, *Manacus*, *Pipra*, *Machaeropterus*, *Pseudopipra* and *Ceratopipra*) within the Piprinae. Although most genera were monophyletic, two genera were not: *Tyranneutes* nested within *Neopelma*, and *Antilophia* nested within *Chiroxiphia*. Most nodes on the phylogeny had 100% bootstrap support; however, a few relationships did not receive full support in the ML analyses and their bootstrap values varied according to the amount of taxon completeness, number of loci and data partitioning. Partitioned ML analyses produced overall higher bootstrap support values using the 75% complete data set, though this was not true for the 95% complete matrix. The lowest support was for relationships within the *Chiroxiphia–Antilophia* clade, as well as within *Pipra*.

The smaller exon data set showed some interrelationships with moderate to high bootstrap support (i.e., ≥70% bootstrap support) in unpartitioned and partitioned ML analyses, but several nodes had low bootstrap support, particularly among many of the genera (Fig. 2). This likely reflected both poor capture efficiency for the exon probes (see above) and the more limited size of the exon data set. Moreover, unpartitioned versus partitioned inferences differed only in the placement of *Chiroxiphia caudata* and *C. pareola* as well as of *Lepidothrix isidorei* and *L. coeruleocapilla*. Nevertheless, for the nodes that showed high support values (i.e., ≥95% bootstrap support) in both analyses of the exon data, the only relationship that conflicted with the UCE results (including those topologies estimated under different filtering schemes, except when the 25% most informative loci were used with ASTRAL; see Section 3.2.2) was in the *Lepidothrix iris*, *L. nattereri*, *L. vilasboasi* clade (see Discussion).

3.2.2. Coalescent-based species trees

Estimates of the species tree obtained using MSC methods were largely congruent with concatenation results. Nodes with 100% bootstrap in the ML concatenated trees had strong support in the ASTRAL and SVDquartets species trees. At the same time, those nodes with lower support in concatenated analyses varied in topology and/or were poorly supported in coalescent trees (Fig. 3).

In general, posterior probabilities from ASTRAL were lower when analyzing the data set including just the 25% most informative loci (data set 5) relative to analyzing all informative loci (data set 4; Fig. 3a). For ASTRAL, topologies were identical, and support was typically high for deep-branching relationships. However there were some topological differences among taxa within genera (*Neopelma*, *Chiroxiphia*, *Lepidothrix* and *Heterocercus*). Within the *Chiroxiphia–Antilophia* clade there was very low support for some relationships, as well as topological differences among analyses of the different data sets.

The ASTRAL species tree estimated from the ML trees differed from that estimated from bootstrap consensus trees in the placement of several taxa (Figs. 3a, A1 and A2), particularly when more loci were included in the data set. For instance, in the ASTRAL bootstrap tree based on all informative loci, *Machaeropterus* was not monophyletic as *Machaeropterus regulus* was sister to a large clade that included other *Machaeropterus* as well as other genera, and *L. isidorei* was sister to the other *Lepidothrix*. However, in the ASTRAL bootstrap tree estimated from the 25% most informative loci, *Machaeropterus* was monophyletic and *L. serena* + *L. suavissima* was sister to the remaining *Lepidothrix*, both

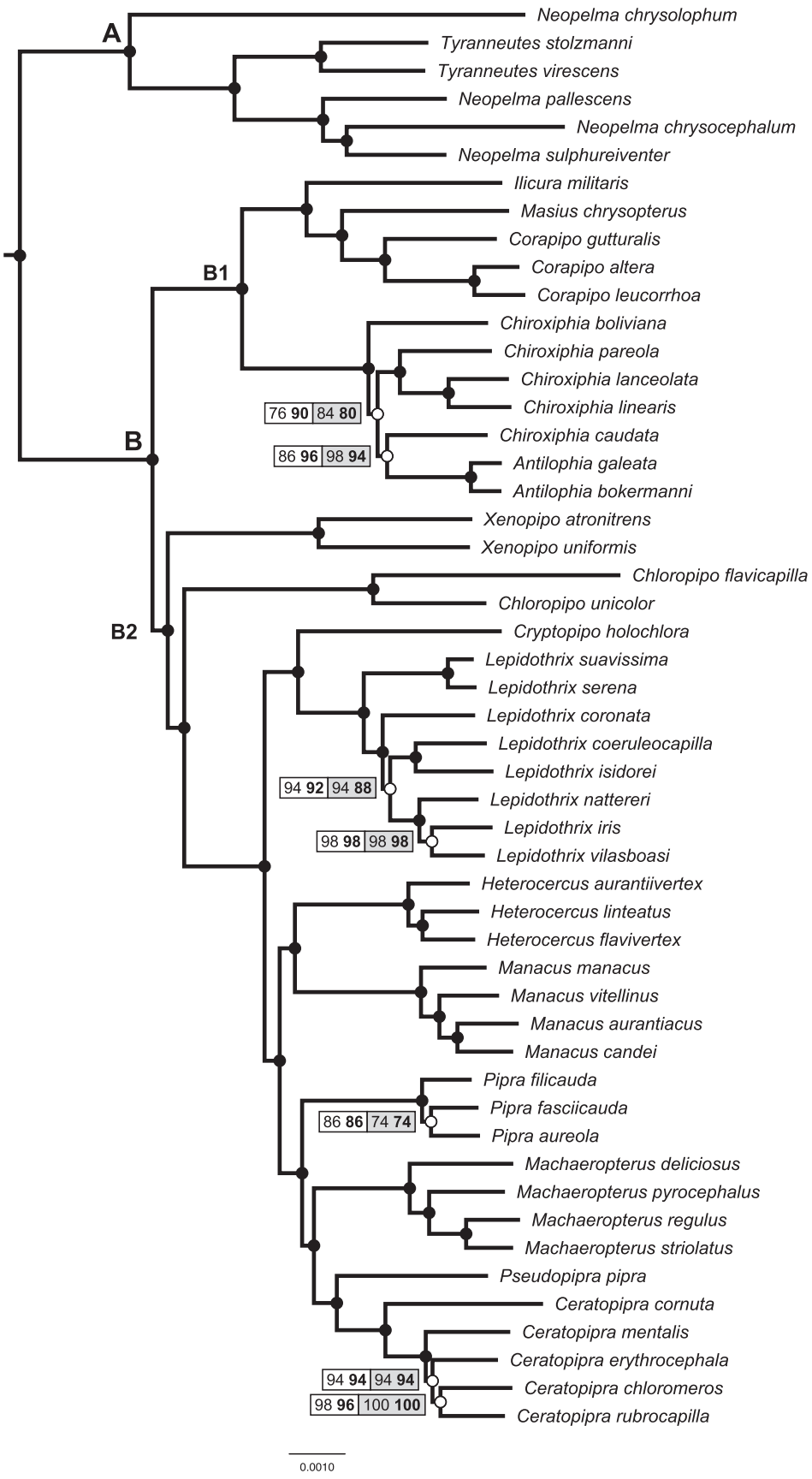

Fig. 1. Maximum likelihood tree obtained using the concatenated UCE data sets. Numbers inside boxes correspond to nodal support values for unpartitioned and partitioned (in bold) inferences of the 75% (white) and 95% (gray) complete data sets (data sets 1 and 2); dark circles indicate 100% bootstrap support in all analyses. Subfamily ranks for Neopelminae (A) and Piprinae (B) follow the South American Classification Committee SACC591 (Remsen et al., 2018), and the proposed tribes Ilicurini (B1) and Piprini (B2) are based on a classification scheme modified from Ohlson et al. (2013a).

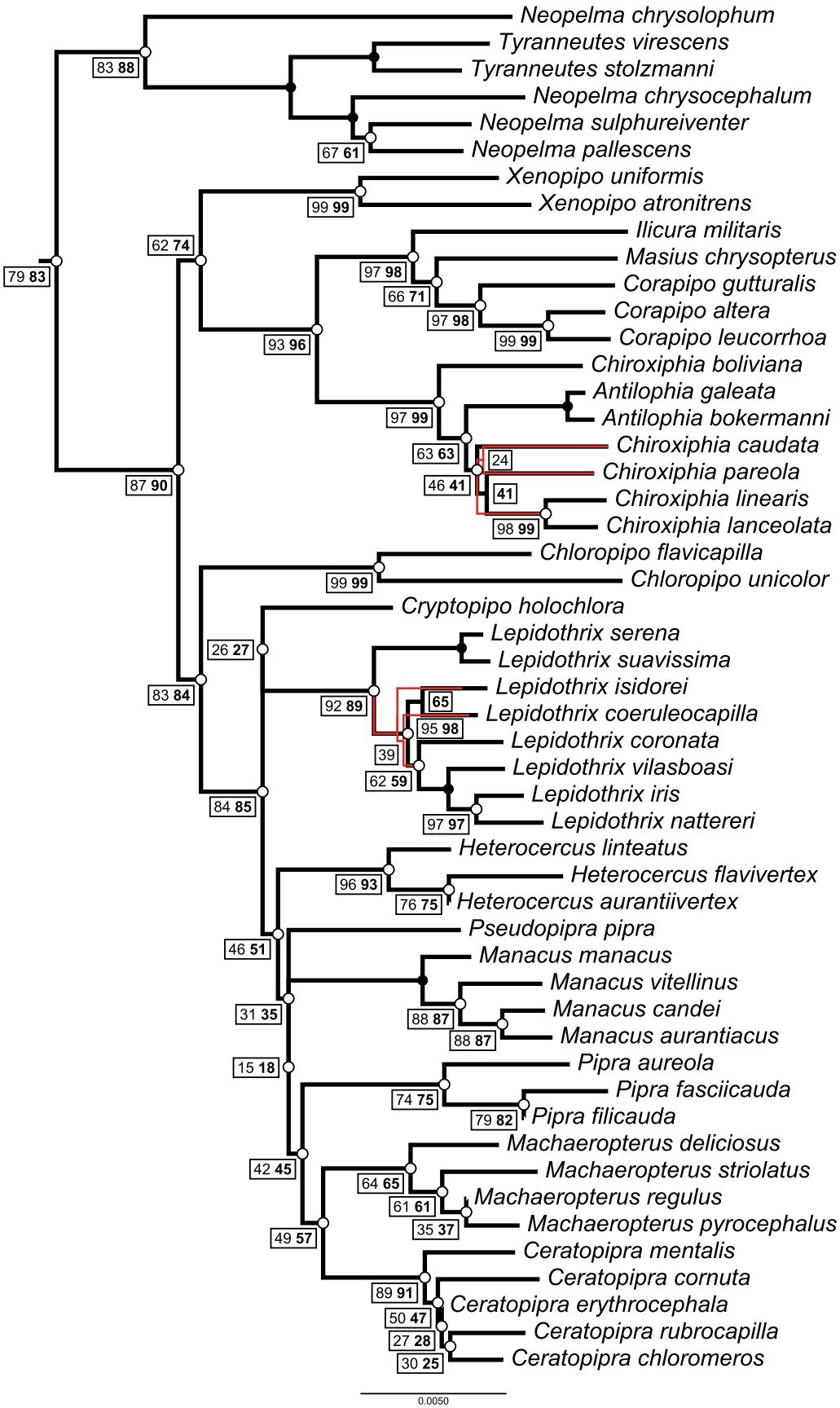

Fig. 2. Maximum likelihood tree obtained using the concatenated exon data set (data set 3) with 36 loci. Numbers inside boxes correspond to nodal support values for unpartitioned and partitioned (in bold) inferences; dark circles indicate 100% bootstrap support in both analyses. Topological discordances between partition schemes are overlaid, with respective support values in separate boxes and the unpartitioned estimate depicted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

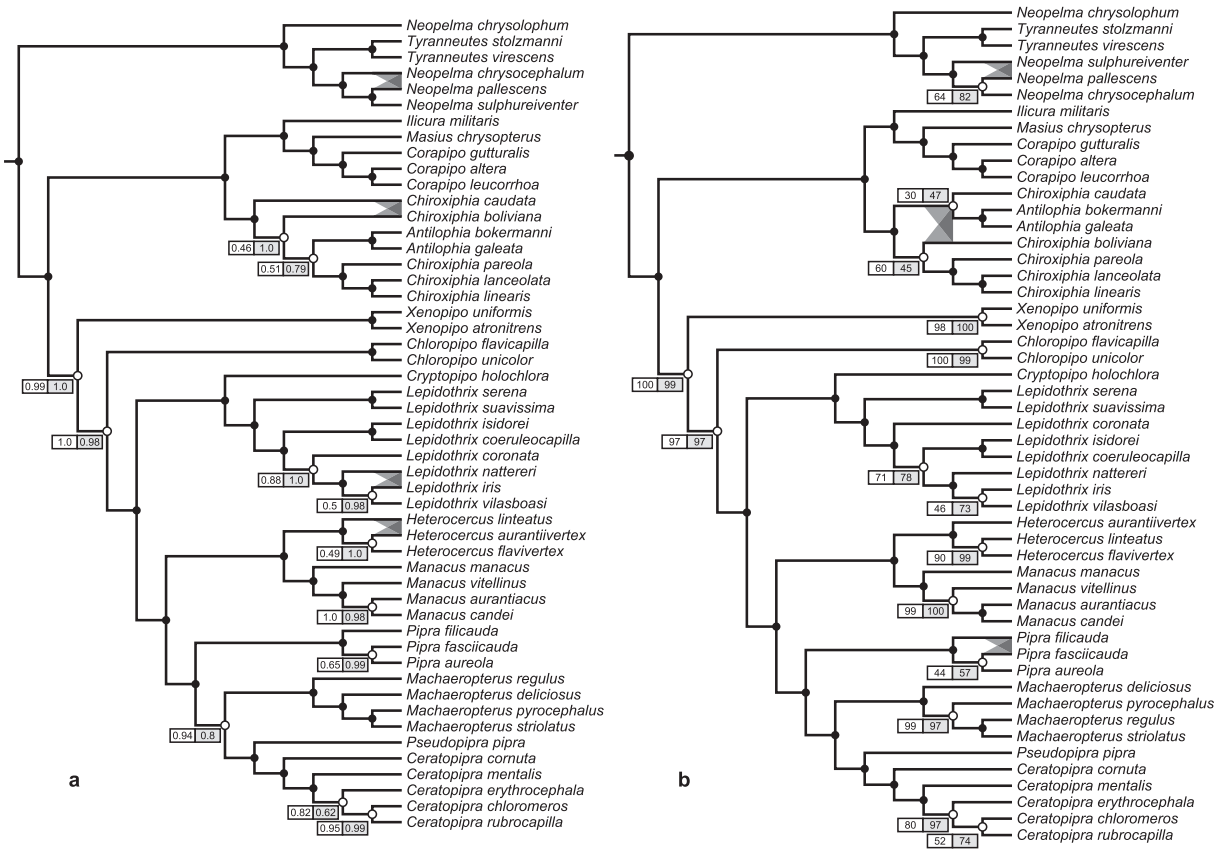

Fig. 3. Estimates of the species tree based on (a) ASTRAL using optimal trees and (b) SVDquartets trees. Numbers inside boxes are nodal support values of the posterior probabilities from quartet frequencies (ASTRAL) and bootstrap replicates (SVDquartets) inferred using UCE data sets with at least one informative site per locus (data set 4, gray) and the 25% most informative loci (data set 5, white); dark circles indicate full support in both analyses. The depicted topologies were obtained using data set 4 (i.e., all informative loci) and shaded zones represent taxa with relationships that conflict with those found in analyses of data set 5 (i.e., the 25% most informative loci).

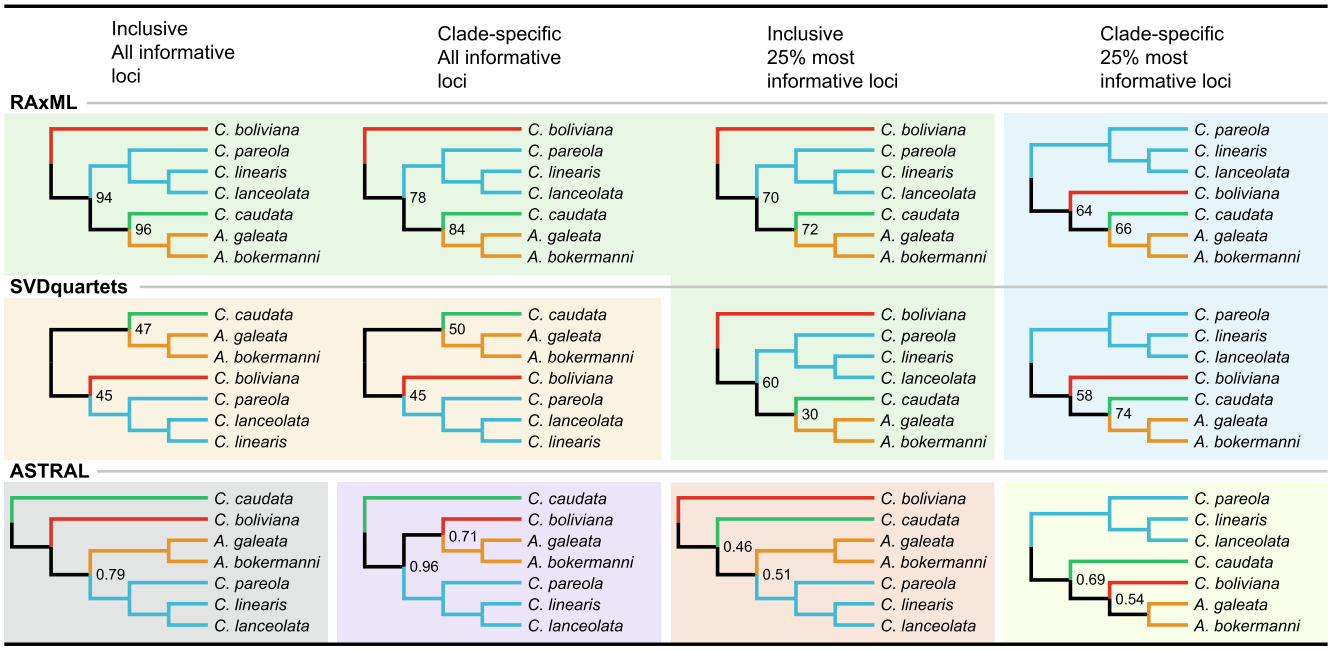

Fig. 4. UCE data sets filtered by the number of parsimony-informative sites per locus. Nodal values represent bootstrap support (RAxML and SVDquartets trees) and posterior probabilities from quartet frequencies (ASTRAL using optimal trees); support values of 100% or 1.0 were omitted. Colored shades indicate alternative topologies. Only the ingroup taxa are shown; the position of the root for each tree reflects the inclusion of the complete set of taxa included in all other analyses.

consistent with analyses of the ASTRAL optimal tree (Fig. 3a) and the concatenated tree (Fig. 2).

The species trees inferred using SVDquartets were more similar to the topology of concatenated trees than to ASTRAL trees, and the early diverging nodes of the SVDquartets trees likewise had overall strong support (Figs. 3b and A3). Yet, phylogenetic relationships within *Neopelma*, *Chiroxiphia* and *Pipra* contained areas of disagreement among the data sets with different filtering schemes.

3.2.3. Topology of the Chiroxiphia–Antilophia clade

Topologies within the *Chiroxiphia/Antilophia* clade were especially variable among our previous analyses (e.g., Figs. 1 and 3), so we conducted additional analyses using data sets 6 and 7 that included loci more likely to resolve these relationships (filtered based on parsimony informative sites just among these taxa). For the *Chiroxiphia–Antilophia* clade, the average number of parsimony-informative sites under the clade-specific filtering strategy was higher than the number of informative sites calculated using all taxa for both all informative loci (2.6 versus 1.9 parsimony-informative sites, respectively) and the 25% most informative loci (4.5 versus 3.6 parsimony-informative sites, respectively; Table 2).

Across all analyses there was high support (100% bootstrap support or posterior probability of 1.0) for uniting the two species of *Antilophia* as a clade, and *Chiroxiphia pareola*, *C. linearis*, and *C. lanceolata* as a second clade. However, relationships among these two clades, *C. boliviana* and *C. caudata* were highly variable (Fig. 4). Of the 15 possible topologies that could arise from our rooted tree among these two well-supported clades and the two other taxa (i.e., essentially a rooted 4-taxon tree), our analyses found seven distinct topologies (represented by the different color shades in Fig. 4). While some topologies were identified in multiple analyses (e.g., three of four concatenated analyses were topologically identical), ASTRAL estimated different topologies with each data set. The ASTRAL bootstrap consensus trees (Fig. A1) using both all informative and the 25% most informative loci were identical to the ASTRAL "Inclusive All informative loci" tree, even though the ASTRAL optimal tree using 25% most informative loci differed (Fig. 4).

In general, a decrease in total nodal support (as measured by the average sum of bootstrap values or posterior probabilities for the clade) was observed for the relationships among *Chiroxiphia–Antilophia* taxa inferred from data sets with fewer loci (Figs. 4 and A5), despite more information per locus (Table 2). For instance, compare the results of the concatenated and ASTRAL analyses using the 25% most informative loci to the results including all those loci with at least one informative site per locus (Fig. 4, also note horizontal and diagonal arrows in Fig. A5). This tendency was also seen in the comparison between inclusive and clade-specific filtering schemes, although the magnitude of change in support difference was overall smaller (Fig. 4; vertical arrows in Fig. A5). However, we detected some disparities in this general pattern in relation to the coalescent-based estimates of the ASTRAL optimal and SVDquartets trees using the 25% most informative loci combined with the clade-specific filtering (Fig. 4; red arrows in Fig. A5).

4. Discussion

This study provides the best-supported tree to date for the Pipridae. Our results were largely congruent across analyses, and led to a robust hypothesis about the phylogenetic relationships of manakins. Although the exon data set had the potential to provide information from a different type of marker than UCEs, the low capture efficiency for these regions resulted in a poorly supported tree. However, most well-supported nodes in the exon tree were the same as those identified using the UCE data. Even using UCEs, there were some nodes that lacked 100% bootstrap support (or posteriors of 1.0) in some inferences; though with the exception of the *Chiroxiphia–Antilophia* clade, most nodes were congruent across analyses. Overall, in spite of some continuing uncertainties, the phylogenetic hypothesis advanced herein, including all but one species, will provide a firmer comparative context for future ecomorphological and behavioral studies.

4.1. Systematic considerations

For higher-level relationships, our results agreed with other molecular studies (Barber et al., 2007; Chesser, 2004; McKay et al., 2010; Ohlson et al., 2008; Ohlson et al., 2013a; Ohlson et al., 2013b; Tello et al., 2009) in that the sexually monomorphic genera *Neopelma* and *Tyranneutes* form a clade that is sister to all other manakin genera, which have typical plumage dichromatism (the "core" manakins), in contrast to earlier morpho-behavorial data that had suggested otherwise (e.g., Prum, 1992). Within the core manakins (clade B, Fig. 1), previous studies have also supported their separation in two groups (Ohlson et al., 2013a; Tello et al., 2009), though these studies assigned *Xenopipo* to different groups. Our results are in agreement with Ohlson et al. (2013a) in placing *Xenopipo* as sister to the remaining taxa in our sub-clade B2 rather than in sub-clade B1. However, we differed from Ohlson et al. (2013a) as we found strong support for placing *Chloropipo* within sub-clade B2, rather than as sister to the taxa of sub-clade B1 (Fig. 1). Within these major groups, relationships among genera were congruent among our analyses.

Our results also agreed with generic reassignments suggested by Ohlson et al. (2013a), the most recent taxonomic treatment of this group. However, we found substantial differences from Ohlson et al. (2013a) in species relationships within various genera, including *Manacus*, *Machaeropterus*, and *Ceratopipra*. It also became evident from our results that the available taxonomy awaits revision of two paraphyletic genera, whose monophyly have already been questioned using smaller numbers of loci (Capurucho et al., 2018; Silva et al., 2018). Our results support the recent study on *Neopelma* and *Tyranneutes* (Capurucho et al., 2018) in finding that *Tyranneutes* nests within *Neopelma*. The divergence of *Neopelma chrysolophum* from the other members of *Neopelma* and *Tyranneutes* is quite deep (Figs. 1 and 2), suggesting that it might be appropriate to transfer that taxon to a new genus. Our results also support the inclusion of *Antilophia* within *Chiroxiphia*, though more loci with greater information content will likely be required to fully understand relationships within this clade.

There was one strongly supported conflicting node between the UCE and exon trees, which was the clade that comprised *Lepidothrix iris*, *L. nattereri* and *L. vilasboasi*. Molecular studies of this clade (Barrera-Guzmán et al., 2018; Dias et al., 2018) already showed that *L. vilasboasi* is likely a hybrid species derived from the *L. iris* and *L. nattereri* lineages.

Thus, the conflict we observed among these species (Fig. 1 versus Fig. 2) is likely due to our different data sets containing more of one parental species versus the other (see also differences in Figs. 3 and A1). Whether other species of manakin might be of hybrid origin will require more extensive genomic and population sampling.

4.2. Conflicts in phylogenomic analyses

Like many other phylogenomic studies, we found topological differences among our analyses (e.g., Jarvis et al., 2014; Hosner et al., 2016; Meiklejohn et al., 2016). One source of difference could be the failure of concatenation to estimate the underlying species tree. However, if biased estimation of the species tree due to ILS is present, the MSC methods should yield congruent topologies, and concatenation should yield a distinct topology. However, for some relationships we found differences among all of our methods, including within and between the two MSC approaches (e.g., Hosner et al., 2016; Meiklejohn et al., 2016). While both MSC methods are consistent given the multispecies coalescent, they make different assumptions. SVDquartets estimates the species tree directly from site patterns in the aligned sequences, considering the mutational process as a source of variability (Chifman and Kubatko, 2014). In contrast, summary methods such as

ASTRAL account for the MSC but rely on estimated gene trees. In cases in which individual loci typically have relatively little phylogenetic information, like UCEs, it may be difficult to estimate the topology, branch lengths and substitution model for each locus, so gene tree estimation error is expected to be relatively high (e.g., Meiklejohn et al., 2016) and can lead to inaccurate estimates of the species tree.

One approach to minimize the problem of gene tree estimation error is to focus on gene trees that are likely to be more accurate, though there has been substantial debate regarding the value of excluding subsets of the genome in phylogenomic analyses (i.e., locus filtering). Using simulations, Molloy and Warnow (2018) showed that removing loci based on proxies for gene-tree estimation error did not improve results from RAxML or SVDquartets (both of which involve concatenating loci as the input format), but it could improve gene tree reconciliation methods (e.g., ASTRAL) when levels of ILS were low to moderate. Like some prior studies (Hosner et al., 2016; Meiklejohn et al., 2016), we used the number of parsimony informative sites as a proxy for locus informativeness, though there are a variety of other approaches to identify informative loci or address problematic relationships (e.g., Arcila et al., 2017; Chen et al., 2015; Dornburg et al., 2016; Dornburg et al., 2019; Salichos and Rokas, 2013). Molloy and Warnow (2018) highlighted recent empirical papers that examined the impact of locus filtering using various proxies for gene tree estimation error (such as parsimony informative sites) and noted that the recommendations based on those empirical studies were at least somewhat contradictory. While the specific approach that may be robust for a given empirical data set might vary, we argue that an advantage of employing at least one of these strategies is that it can at least highlight whether a relationship seems more robust (e.g., Meiklejohn et al., 2016) or whether it appears very unstable and should be treated with caution (e.g., as we observed for the *Chiroxiphia–Antilophia* clade).

All of these results emphasize the caution with which systematists should approach analyses of NGS sequence data when challenging nodes are examined. This is not surprising considering that analyses of whole-genome data have been unable to resolve some recalcitrant nodes at the base of Neoaves (Jarvis et al., 2014), and that the results of some analyses appear to depend upon specific types of loci that are analyzed (Braun et al., 2019; Jarvis et al., 2014; Reddy et al., 2017). Despite these challenges, phylogenomic analyses often yield trees in which most nodes are both well supported and insensitive to analytical methodology (e.g., Hosner et al., 2016; Moyle et al., 2016; Oliveros et al., 2019). That was certainly true in this study; the backbone for manakin relationships was strongly supported and estimates of phylogeny based on concatenation, ASTRAL, and SVDquartets were congruent with just a few exceptions.

CRediT authorship contribution statement

Rafael N. Leite: Conceptualization, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Visualization. **Rebecca T. Kimball:** Conceptualization, Resources, Writing - original draft, Writing - review & editing. **Edward L. Braun:** Conceptualization, Data curation, Writing - original draft, Writing - review & editing. **Elizabeth P. Derryberry:** Resources, Funding acquisition. **Peter A. Hosner:** Writing - review & editing. **Graham E. Derryberry:** Data curation. **Marina Anciães:** Resources, Writing - review & editing. **Jessica S. McKay:** Data curation. **Alexandre Aleixo:** Resources, Writing - review & editing. **Camila C. Ribas:** Resources, Writing - review & editing. **Robb T. Brumfield:** Resources, Writing - review & editing, Funding acquisition. **Joel Cracraft:** Resources, Writing - review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgments

This work was funded by research grants from NSF (DEB 1146265 to R.T.B., DEB 1146423 to E.P.D., and DEB 1146248 and DEB 1241066 to J.C.) and from FAPESP (2012/50260-6 to Lucia G. Lohmann); R.N.L. was supported by FAPESP (2016/08356-7) and CNPq (457267/2014-3; 409430/2017-0; 150128/2018-5). R.T.K. and E.L.B. received support from NSF (DEB 1655683). P.A.H. acknowledges the Danish National Research Foundation for support to the Center for Macroecology, Evolution and Climate (DNRF96). This work was also facilitated by the Manakin Genomics RCN with support from NSF (DEB 1457541) to Drs. Bette Loiselle, Emily DuVal, Christopher Balakrishnan, Michael Braun and Alice Boyle. We are thankful to scientific collections and museums (American Museum of Natural History, Burke Museum, Field Museum of Natural History, Florida Museum of Natural History, Instituto Nacional de Pesquisas da Amazônia, Louisiana State University Museum of Natural Science, Museu Paraense Emílio Goeldi, and Universidad Nacional de Colombia) and to the collectors and curatorial personnel who provided samples for this study. R.N.L. is thankful to Drs. Fernanda Werneck and Izeni Farias for providing logistic support and to Dr. Diogo Meyer for facilitating this research. We also thank two anonymous reviewers whose comments greatly improved this manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2020.107013.

References

Anciães, M., Peterson, A.T., 2009. Ecological niches and their evolution among Neotropical manakins (Aves: Pipridae). J. Avian Biol. 40, 591–604.
Arcila, D., Ortí, G., Vari, R., Armbruster, J.W., Stiassny, M.L.J., Ko, K.D., Sabaj, M.H., Lundberg, J., Revell, L.J., Bentacur-R, R., 2017. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. Nat. Ecol. Evol. 1, 0020.
Barber, B.R., Rice, N.H., Klicka, J., 2007. Systematics and evolution in the Tityrinae (Passeriformes: Tyrannoidea). Auk 124, 1317–1329.
Barrera-Guzmán, A.O., Aleixo, A., Shawkey, M.D., Weir, J.T., 2018. Hybrid speciation leads to novel male secondary sexual ornamentation of an Amazonian bird. Proc. Natl. Acad. Sci. 115, E218.
Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., Haussler, D., 2004. Ultraconserved Elements in the Human Genome. Science 304, 1321.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.
Braun, E.L., Cracraft, J., Houde, P., 2019. Resolving the avian tree of life from top to bottom: the promise and potential boundaries of the phylogenomic era. In: Kraus, R.H.S. (Ed.), Avian Genomics in Ecology and Evolution. Springer, Cham, pp. 151–210.
Capurucho, J.M.G., Ashley, M.V., Ribas, C.C., Bates, J.M., 2018. Connecting Amazonian, Cerrado, and Atlantic forest histories: Paraphyly, old divergences, and modern population dynamics in tyrant-manakins (Neopelma/Tyranneutes, Aves: Pipridae). Mol. Phylogenet. Evol. 127, 696–705.
Chen, M.-Y., Liang, D., Zhang, P., 2015. Selecting Question-Specific Genes to Reduce Incongruence in Phylogenomics: A Case Study of Jawed Vertebrate Backbone Phylogeny. Syst. Biol. 64, 1104–1120.
Chesser, R.T., 2004. Molecular systematics of New World suboscine birds. Mol. Phylogenet. Evol. 32, 11–24.
Chifman, J., Kubatko, L., 2014. Quartet Inference from SNP Data Under the Coalescent Model. Bioinformatics 30, 3317–3324.
Chifman, J., Kubatko, L., 2015. Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. J. Theor. Biol. 374, 35–47.
Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., Warnow, T., 2015. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. BMC Genomics 16, S2.
Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2, 762–768.
Dias, C., Lima, K.A., Araripe, J., Aleixo, A., Vallinoto, M., Sampaio, I., Scheneider, H., Rêgo, P.S., 2018. Mitochondrial introgression obscures phylogenetic relationships among manakins of the genus Lepidothrix (Aves: Pipridae). Mol. Phylogenet. Evol. 126, 314–320.
Dornburg, A., Fisk, J.N., Tamagnan, J., Townsend, J.P., 2016. PhyInformR: phylogenetic experimental design and phylogenomic data exploration in R. BMC Evol. Biol. 16, 262.
Dornburg, A., Su, Z., Townsend, J.P., 2019. Optimal rates for phylogenetic inference and experimental design in the era of genome-scale data sets. Syst. Biol. 68, 145–156.

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., Liu, L., Davis, C.C., 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. Evol. 94 (Part A), 447–462.

Faircloth, B.C., 2013. Illumiprocessor: a Trimmomatic wrapper for parallel adapter and quality trimming.

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.

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.

Gill, F., Donsker, D., 2018. IOC World Bird List (version 8.2).

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotech. 29, 644–652.

Harvey, M.G., Aleixo, A., Ribas, C.C., Brumfield, R.T., 2017. Habitat association predicts genetic diversity and population divergence in Amazonian birds. Am. Nat. 190, 631–648.

Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T., 2016. Avoiding Missing Data Biases in Phylogenomic Inference: An Empirical Study in the Landfowl (Aves: Galliformes). Mol. Biol. Evol. 33, 1110–1125.

Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y.W., 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., Gabaldón, T., Capella-Gutiérrez, 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.C., Prosdocimi, F., Samaniego, J.A., Velazquez, A.M.V., Alfaro-Núñez, 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., Jønsson, 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., Alström, P., Edwards, S.V., Stamatakis, A., Mindell, D.P., Cracraft, J., Braun, E.L., Warnow, T., Jun, W., Gilbert, M.T.P., Zhang, G., 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346, 1320–1331.

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30, 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., Faircloth, B.C., Smith, B. T., Braun, E.L., 2019. A phylogenomic supertree of birds. Diversity 11, 109.

Kirwan, G.M., Green, G., 2012. Cotingas and Manakins. Princeton University Press, Princeton, New Jersey.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. Syst. Biol. 56, 17–24.

Lane, D.F., Kratter, A.W., O'Neill, J.P., 2017. A new species of manakin (Aves: Pipridae; Machaeropterus) from Peru with a taxonomic reassessment of the striped manakin (*M. regulus*) complex. Zootaxa 4320, 379–390.

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

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol. Biol. Evol. 34, 772–773.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., Turner, D.J., 2010. Target-enrichment strategies for next-generation sequencing. Nat. Methods 7, 111.

McCormack, J.E., Faircloth, B.C., 2013. Next-generation phylogenetics takes root. Mol. Ecol. 22, 19–21.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A Phylogeny of Birds Based on Over 1,500 Loci Collected by Target Enrichment and High-Throughput Sequencing. PLoS ONE 8.

McKay, B.D., Barker, F.K., Mays Jr, H.L., Doucet, S.M., Hill, G.E., 2010. A molecular phylogenetic hypothesis for the manakins (Aves: Pipridae). Mol. Phylogenet. Evol. 55, 733–737.

Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Kimball, R.T., Braun, E.L., 2016. Analysis of a Rapid Evolutionary Radiation Using Ultraconserved Elements: Evidence for a Bias in Some Multispecies Coalescent Methods. Syst. Biol. 65, 612–627.

Mendes, F.K., Hahn, M.W., 2018. Why concatenation fails near the anomaly zone. Syst. Biol. 67, 158–169.

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

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, 285–303.

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.

Ohlson, J., Fjeldså, J., Ericson, P.G.P., 2008. Tyrant flycatchers coming out in the open: phylogeny and ecological radiation of Tyrannidae (Aves, Passeriformes). Zoolog. Scr. 37, 315–335.

Ohlson, J.I., Fjeldså, J., Ericson, P.G.P., 2013a. Molecular phylogeny of the manakins (Aves: Passeriformes: Pipridae), with a new classification and the description of a new genus. Mol. Phylogenet. Evol. 69, 796–804.

Ohlson, J.I., Irestedt, M., Ericson, P.G.P., Fjeldså, J.O.N., 2013b. Phylogeny and classification of the New World suboscines (Aves, Passeriformes). Zootaxa 3613, 1–35.

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., Braun, M.J., 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. 116, 7916.

Patel, S., Kimball, R.T., Braun, E.L., 2013. Error in phylogenetic estimation for bushes in the tree of life. J. Phylogenet. Evol. Biol. 1, 110.

Prum, R.O., 1990. Phylogenetic Analysis of the Evolution of Display Behavior in the Neotropical Manakins (Aves: Pipridae). Ethology 84, 202–231.

Prum, R.O., 1992. Syringeal Morphology, Phylogeny, and Evolution of the Neotropical Manakins (Aves: Pipridae). Am. Mus. Novit. 3043, 65 pp.

Prum, R.O., 1994. Phylogenetic analysis of the evolution of alternative social behavior in the manakins (Aves: Pipridae). Evolution 48, 1657–1675.

Prum, R.O., 1997. Phylogenetic Tests of Alternative Intersexual Selection Mechanisms: Trait Macroevolution in a Polygynous Clade (Aves: Pipridae). Am. Nat. 149, 668–692.

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

Pyron, R.A., Hendry, C.R., Chou, V.M., Lemmon, E.M., Lemmon, A.R., Burbrink, F.T., 2014. Effectiveness of phylogenomic data and coalescent species-tree methods for resolving difficult nodes in the phylogeny of advanced snakes (Serpentes: Caenophidia). Mol. Phylogenet. Evol. 81, 221–231.

Reddy, S., Kimball, R.T., Pandey, A., Hosner, P.A., Braun, M.J., Hackett, S.J., Han, K.-L., Harshman, J., Huddleston, C.J., Kingston, S., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Witt, C.C., Yuri, T., Braun, E.L., 2017. Why Do Phylogenomic Data Sets Yield Conflicting Trees? Data Type Influences the Avian Tree of Life more than Taxon Sampling. Syst. Biol. 66, 857–879.

Remsen, J.V., Jr, Areta, J.I., Cadena, C.D., Claramunt, S., Jaramillo, A., Pacheco, J.F., Robbins, M.B., Stiles, F.G., Stoz, D.F., Zimmer, K.J., 2018. A classification of the bird species of South America. Version: 8 May 2019. American Ornithologists' Union http://www.museum.lsu.edu/~Remsen/SACCBaseline.htm.

Roch, S., Steel, M., 2015. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theor. Popul. Biol. 100, 56–62.

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

Sayyari, E., Mirarab, S., 2016. Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. Mol. Biol. Evol. 33, 1654–1668.

Seo, T.-K., 2008. Calculating Bootstrap Probabilities of Phylogeny Using Multilocus Sequence Data. Mol. Biol. Evol. 25, 960–971.

Silva, S.M., Agne, C.E., Aleixo, A., Bonatto, S.L., 2018. Phylogeny and systematics of *Chiroxiphia* and *Antilophia* manakins (Aves, Pipridae). Mol. Phylogenet. Evol. 127, 706–711.

Song, S., Liu, L., Edwards, S.V., Wu, S., 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proc. Natl. Acad. Sci. 109, 14942–14947.

Springer, M.S., Gatesy, J., 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94 (Part A), 1–33.

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

Swofford, D.L., 2017. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Tello, J.G., Moyle, R.G., Marchese, D.J., Cracraft, J., 2009. Phylogeny and phylogenetic classification of the tyrant flycatchers, cotingas, manakins, and their allies (Aves: Tyrannides). Cladistics 25, 429–467.

Tonini, J., Moore, A., Stern, D., Shcheglovitova, M., Ortí, G., 2015. Concatenation and Species Tree Methods Exhibit Statistically Indistinguishable Accuracy under a Range of Simulated Conditions. PLoS. Currents 7 ecurrents. tol.34260cc27551a34527b34124ec34265f36334b34266be.

Xi, Z., Liu, L., Davis, C.C., 2015. Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. Mol. Phylogenet. Evol. 92, 63–71.