



A near-complete and time-calibrated phylogeny of the Old World flycatchers, robins and chats (Aves, Muscicapidae)

Min Zhao^a, J. Gordon Burleigh^a, Urban Olsson^{b,c}, Per Alström^{d,e,*}, Rebecca T. Kimball^{a,*}

^a Department of Biology, University of Florida, Gainesville, FL 32611, USA

^b Department of Biology and Environmental Sciences, Systematics and Biodiversity, University of Gothenburg, Box 463, SE 405 30 Gothenburg, Sweden

^c Gothenburg Global Biodiversity Centre, Box 461, SE-405 30 Gothenburg, Sweden

^d Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

^e Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:

Phylogenetics
Systematics
Muscicapids
Super-matrix
Multi-locus
Divergence time

ABSTRACT

The Old World flycatchers, robins and chats (Aves, Muscicapidae) are a diverse songbird family with over three hundred species. Despite continuous efforts over the past two decades, there is still no comprehensive and well-resolved species-level phylogeny for Muscicapidae. Here we present a supermatrix phylogeny that includes all 50 currently recognized genera and ca. 92% of all the species, built using data from up to 15 mitochondrial and 13 nuclear loci. In addition to assembling nucleotide sequences available in public databases, we also extracted sequences from the genome assemblies and raw sequencing reads from GenBank and included a few unpublished sequences. Our analyses resolved the phylogenetic position for several previously unsampled taxa, for example, the Grand Comoro Flycatcher *Humblotia flavirostris*, the Collared Palm Thrush *Cichladusa arquata*, and the Taiwan Whistling-Thrush *Myophonus insularis*, etc. We also provide taxonomic recommendations for genera that exhibit paraphyly or polyphyly. Our results suggest that Muscicapidae diverged from Turdidae (thrushes and allies) in the early Miocene, and the most recent common ancestors for the four subfamilies (Muscicapinae, Niltavinae, Cossyphinae and Saxicolinae) all arose around the middle Miocene.

1. Introduction

Muscicapidae, the chats, robins and Old World flycatchers, is one of the largest bird families in the world, with 49 – 57 different genera proposed and 303 – 343 recognized species, depending on the taxonomy (e.g., Dickinson and Christidis, 2014; Clements et al., 2021; and Gill et al., 2022). It is widely distributed in various habitats across the entire Old World and adjacent Australasia and Nearctic (Winkler et al., 2020). Birds of this family exhibit great diversity in morphology, behaviors, vocalizations, and life history, which makes Muscicapidae a great study group to address various questions on evolution, diversity and biogeography. However, for very similar reasons, it also makes resolving the phylogeny for this species-rich family challenging.

Numerous studies in the past two decades sought to infer the phylogenetic relationships for subsets of the Muscicapidae family (e.g., Roy et al., 2001; Beresford, 2003; Cibois and Cracraft, 2004; Outlaw and Voelker, 2006; Outlaw et al., 2007; Illera et al., 2008; Lim et al., 2010;

Outlaw et al., 2010; Sangster et al., 2010; Zuccon and Ericson, 2010a; Zuccon and Ericson, 2010b; Aliabadian et al., 2012; Voelker et al., 2016, 2012; Barve and Mason, 2015; Moyle et al., 2015; Hooper et al., 2016; Zhang et al., 2016; Zhao et al., 2017; Alström et al., 2018; Fjeldså et al., 2020; Ng et al., 2022; Wei et al., 2022). Molecular phylogenetic studies revealed that the so called “flycatchers”, “robins”, and “chats” are not reciprocally monophyletic groups. Also, some species that were historically included in other bird families are more closely related to species in Muscicapidae. For example, *Monticola* (rock thrushes), *Myophonus* (whistling thrushes), *Heinrichia* (Great Shortwing), *Alethe* (alethes), and *Brachypteryx* (shortwings) were formerly in Turdidae (thrushes and allies) (Voelker and Spellman, 2004; Sangster et al., 2010), and Bagobo Robin *Leonardina woodi* was previously in Timaliidae (Old World babblers) (Oliveros et al., 2012). On the other hand, *Cochoa* (cochoas) was transferred from Muscicapidae to Turdidae (Sangster et al., 2010), and more recently, the monotypic genus *Pinarornis* (Boulder Chat *P. plumosus*) was placed within Turdidae based on unpublished

* Corresponding authors at: Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden (P. Alström). Department of Biology, University of Florida, Gainesville, FL 32611, USA (R.T. Kimball).

E-mail addresses: per.alstrom@ebc.uu.se (P. Alström), rkimball@ufl.edu (R.T. Kimball).

<https://doi.org/10.1016/j.ympev.2022.107646>

Received 6 June 2022; Received in revised form 8 October 2022; Accepted 12 October 2022

Available online 17 October 2022

1055-7903/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

molecular data (Fjelds  et al., 2020). Many of these original placements were based on evidence of shared morphological or behavioral traits, which may be misleading due to convergent evolution.

Two of the latest efforts to build a large-scale species-level phylogeny for Muscicapidae, Sangster et al. (2010) and Zuccon and Ericson (2010a), laid the foundation for the classification of Muscicapidae and resolved many taxonomic uncertainties, although their taxon sampling was limited and with many deep nodes having poor support. Barve and Mason (2015) assembled another large-scale phylogeny for Muscicapidae using a supermatrix with both increased taxon sampling (252 taxa) and molecular data (11 molecular markers). This study provides support for many lineages delineated in previous studies as well as provides some new insights; however, it lacks comprehensive taxonomic reconciliations, assessments of the phylogenetic relationships, or suggestions for taxonomic change.

The goal of this study is to reconstruct a comprehensive species-level phylogeny for Muscicapidae with more extensive taxon and locus sampling than previous studies, which incorporates and synthesizes molecular data from numerous resources (i.e., nucleotide databases, genome assemblies, raw sequencing reads, and our own unpublished Sanger sequencing data). The resulting phylogenetic tree includes 301 species (ca. 92% species coverage based on Clements/eBird Checklist of Birds of the World v.2021) from all currently recognized genera and provides a robust basis for the taxonomic classification of Muscicapidae as well as a resource for comparative studies.

2. Methods

2.1. Taxon name reconciliation

We downloaded all the sequences from the NCBI nucleotide database (as of Nov 10, 2020) for Muscicapidae and Turdidae, since many muscicapids are placed under Turdidae in GenBank. As GenBank taxa names do not follow a specific standardized avian taxonomy, we reconciled the GenBank names to match the Clements/eBird Checklist of Birds of the World v.2021, which includes 49 genera and 327 species within Muscicapidae (Clements et al., 2021), based on information from Avibase, online searches, and results presented in Hosner et al. (2022). We standardized the names based on the Clements/eBird checklist because it is updated regularly and frequently and intrinsically linked to trait databases such as Birds of the World (<https://birdsoftheworld.org/>) and Macaulay Library (<https://macaulaylibrary.org/>), which facilitates downstream phylogenetic comparative analyses. We also implemented taxonomic updates suggested by IOC World Bird List v12.1 (Gill et al., 2022) that are shown in Table 1. Collectively, this provided us with a target set consisting of 326 species from 50 genera.

2.2. Molecular data collection

After taxon name reconciliation, we performed an all by all BLASTN search on the downloaded nucleotide sequences (Camacho et al., 2009).

For the BLASTN search, each accession from a nuclear sequence was transformed into a single FASTA file, and for the mitochondria, each individual coding or rRNA region was made into a FASTA file. Any pair of sequences that had a maximum e-value of $1.0e-5$ and the pairwise alignment length covering at least 50% of the length of both sequences were considered homologous. We then performed single linkage clustering to obtain all groups of sequences (representing homologous loci) that were linked by significant BLAST hits. If a cluster contained multiple sequences from a species, we retained only a single, longest sequence. We manually searched GenBank nucleotide database for recently published sequences or sequences that may have been missed by the clustering (as of July 2022).

As noted by Hosner et al. (2022), sequences from closely related species may still be placed under the previous name in GenBank if they were split recently. Therefore, we also searched associated literature and museum vouchers for species with recent splits proposed. In this way, we identified more sequences, for example, *Ficedula luzoniensis* from *Ficedula hyperythra* (Moyle et al., 2015), *Ficedula riedeli* from *Ficedula dumetoria* (Outlaw and Voelker, 2006), *Calliope tschebaiewi* from *Calliope pectoralis* (Liu et al., 2016), *Cyornis pelingensis* from *Cyornis colonus* (Garg et al., 2018), and *Copsychus superciliosus* from *Copsychus luzoniensis* (Lim et al., 2010). These sequences were added into the supermatrix to represent their corresponding species.

To further enrich the data matrix, we extracted molecular markers from genome assemblies and raw sequencing reads (Table 2). We extracted nuclear loci from the six muscicapids that have genome assemblies available on GenBank (i.e., *Erithacus rubecula*, *Saxicola maurus*, *Cercotrichas coryphoeus*, *Oenanthe oenanthe*, *Copsychus sechellarum*, and *Ficedula albicollis*) using the Extract seq pipeline (https://github.com/aa-kanksha12/Extract_seq) which incorporates the nhmmer tool (Wheeler and Eddy, 2013) to improve the detection of remote homologs for standard BLASTN (Camacho et al., 2009). We used MitoFinder (Allio et al., 2020) to extract mitochondrial regions from 27 mitochondrial genomes, as well as from the UCE target enrichment sequencing reads for *Muscicapa striata* that are available in the NCBI SRA database. The mitochondrial genomes for *Muscicapa griseisticta* (NC_045181) and *Cyanoptila cyanomelana* (NC_015232) were misidentified according to Sangster and Luksenburg (2021), therefore we did not include them in our data matrix. We also created a reference fasta file using the nuclear loci from *Erithacus rubecula* and mapped whole genome sequencing reads in the NCBI SRA database for 20 muscicapids, including *Cichladusa arquata*, which had no data available in the NCBI nucleotide database (as of July 2022). We applied the same process as described above to collect sequence data for three Turdidae species as the outgroup (*Catharus fuscescens*, *Sialia sialis*, and *Turdus rufiventris*), all with genome assembly data available on GenBank.

We obtained additional cytochrome c oxidase subunit 1 (COI) sequences from the Barcode of Life Data (BOLD) System v4. Lastly, we included five unpublished sequences for *Myophonus insularis* [cytochrome *b* (Cytb), myoglobin intron 2 (Myo), and ornithine

Table 1
Taxonomic updates based on two popular standardized avian taxonomies.

| Taxon in Clements/eBird 2021 | Taxon in IOC World Bird List 12.1 | Name used in this study | Reference |
|---|---------------------------------------|--------------------------------|---|
| <i>Cyornis concretus</i> | <i>Leucoptilon concretum</i> | <i>Leucoptilon concretum</i> | (Sangster et al., 2021) |
| <i>Brachypteryx stellata</i> | <i>Heteroxenicus stellatus</i> | <i>Heteroxenicus stellatus</i> | (Rasmussen and Anderton 2005; Price et al., 2014) |
| <i>Thamnolaea semirufa</i> | <i>Monticola semirufus</i> | <i>Monticola semirufus</i> | (Zuccon and Ericson 2010b) |
| <i>Pinarornis plumosus</i> (Muscicapidae) | <i>Pinarornis plumosus</i> (Turdidae) | Removed | (Fjelds  et al., 2020) |

Table 2
Data sources of the molecular markers used. For details of data accessions, see Supplementary Table S1.

| Data Source | Species sampled | Unique species | Loci | Sequences included |
|------------------------------------|-----------------|----------------|------|--------------------|
| GenBank Nucleotide | 295 | 295 | 28 | 2053 |
| GenBank Genome Assemblies | 6 | 0 | 13 | 47 |
| GenBank Mitochondrial Genomes | 27 | 0 | 15 | 133 |
| GenBank SRA | 21 | 1 | 28 | 203 |
| Barcode of Life Data System | 19 | 3 | 1 | 19 |
| Unpublished Sanger sequencing data | 3 | 2 | 4 | 5 |
| Summary | 301 | 301 | 28 | 2460 |

decarboxylase introns 6–7 (ODC)], *Cyornis banyumas* (Cytb), and *Ficedula albicilla* [glyceraldehyde-3-phosphate dehydrogenase intron 11 (GAPDH)]. Genomic DNA from blood or muscle was extracted using the Qiagen DNeasy Blood & Tissue Kit. PCR amplifications were performed using the following cycling conditions: 1) for Cytb, 5 min at 95 °C, 40 cycles of 40 sec at 95 °C, 1 min at 45 °C, and 2 min at 72 °C, and a final extension of 6 min at 72 °C, with primers L14995 and H16065 (Helbig et al., 1995); 2) for GAPDH, 5 min at 95 °C, 40 cycles of 40 sec at 95 °C, 40 sec at 58 °C, and 1 min at 72 °C, and a final extension of 6 min at 72 °C, with primers G3P13b and G3P14b (Fjeldsø et al., 2003); 3) for ODC, similar as GAPDH, except using an annealing temperature of 59 °C, with primers OD6 and OD8r (Friesen et al., 1999); 4) for Myo, a nested PCR using primers Myo2 and Myo3 (Slade et al., 1993) in the first step (5 min at 95 °C; followed by 20 cycles of 40 s at 95 °C, 40 s at 62 °C, and 1 min at 72 °C; terminated by 8 min at 72 °C), and primers Myo 2 and 3F (Heslewood et al., 1998) in the second amplification. All genes were sequenced with the respective PCR primers, except for Cytb which was sequenced with H15298 and L15722 (Helbig et al., 1995), and L15152 (5'-GTC CAA TTC GGC TGA CTA ATT CGC AAC CTA CAC GCA AAC GG-3'). For details of data accessions, see [Supplementary Table S1](#).

2.3. Data verification

We used MAFFT 7.407 (Katoh and Standley, 2013) to align each locus individually, and we manually checked alignments using MEGA X (Kumar et al., 2018). We then estimated gene trees in IQ-TREE 2.1.2 (Minh et al., 2020) with 1000 ultrafast bootstrap replicates (–ufboot 1000) by considering all standard substitution models and allowing for invariable sites and discrete Gamma model for rate heterogeneity (–m TEST). Gene trees were visualized with FigTree 1.4.4 (Rambaut, 2018).

We examined each gene tree for exceptionally long branches and any species that appeared far apart from its congeners or in an unexpected position. This could be caused by sequences that were too short, of poor quality, or mislabeled. We removed the problematic sequences and replaced them with alternate sequences if available. For example, in the initial COI gene tree, *Cyornis oscillans* appeared close to the outgroup (Turdidae) instead of in the genus *Cyornis* or subfamily Niltavinae. The COI sequence in the BOLD System that was used to build the gene tree (id: BBIND665-08) was BLASTed and turned out to be most similar to *Pachycephala homeyeri* (Pachycephalidae). As there was no other molecular data available for this species, it was removed from the data matrix. A second example of mislabeling is that both the Cytb and NADH dehydrogenase subunit 2 (ND2) sequences of *Phoenicurus fuliginosus* (GenBank accessions: KJ024172 & KJ024236), likely belong to *Niltava macgrigoriae*. Thus, they were replaced with correctly labeled sequences. Additionally, we followed the findings in Moyle et al. (2014) and removed/replaced problematic sequences for *Ficedula bonthaina*,

F. buruensis and *F. subrubra*. Only the loci that had data available for a minimum of 30 taxa were kept for further phylogenetic analyses (Table 3).

2.4. Phylogenetic analyses

After the process of data verification, we re-aligned the sequences and re-built gene trees as described above. The concatenated alignment and NEXUS data block were built using phyutility 2.7.1 (Smith and Dunn, 2008). Mitochondrial protein coding genes were partitioned by codon positions (i.e., 1st, 2nd, and 3rd codon positions). We then performed a partitioned analysis for the concatenated data in IQ-TREE (Minh et al., 2020), which infers phylogenetic trees by maximum likelihood and uses a greedy strategy to find the partitioning scheme with best model-fit (–m TESTMERGE). Tree estimation was run with 1000 ultrafast bootstrap replicates (–ufboot 1000) and edge-proportional branch lengths between partitions but allowing each partition to have its own substitution rate (–p). We also ran a partitioned analysis for mitochondrial and nuclear loci respectively using IQ-TREE with the same parameter settings as described above.

We used IQ-TREE to run model testing (–m TESTMERGEONLY –mset mrbayes) similar to PartitionFinder (Lanfear et al., 2012). The best-fitting partitions and models were then used for a partitioned analysis with Bayesian inference in MrBayes 3.2.6 (Ronquist and Huelsenbeck, 2003) on the CIPRES Science Gateway (Miller et al., 2010). MCMC chains were run for 50 million generations (with four chains, chain temperature of 0.2, and two runs) and sampled every thousand generations. The first 20% of trees were discarded before summarizing the trees to a 50-majority-rule consensus tree (contype = Halfcompat). The mixing of chains was checked in Tracer 1.7.1 (Rambaut et al., 2018) to ensure all parameters had effective sample sizes (ESSs) > 200. Separate analyses were performed for all loci, mitochondrial loci only, and nuclear loci only.

2.5. Divergence time estimation

To estimate divergence times for Muscicapidae, we used TreePL under penalized likelihood (Smith and O'Meara, 2012), which allows for different rates across branches but penalizes rate differences over the tree with a rate smoothing parameter. Given that the ML tree based on all loci overall yielded better estimates of the topology than the Bayesian tree, we used the rooted ML tree inferred from all loci as the phylogeny for TreePL time calibration. A secondary calibration point for the split between Muscicapidae and Turdidae [min = 13.5 million years ago (Mya), max = 22 Mya] from Oliveros et al. (2019) was applied. We used the *thorough* option, which allowed the program to continue to iterate until convergence, and the *prime* option, which tested different

Table 3

Molecular markers included in the super-matrix. Taxa number refers to the number of Muscicapidae species with sequences from each locus. Locus length shows the length of final alignment.

| Locus (mitochondrial) | Taxa | Length | Locus (nuclear) | Taxa | Length |
|--|------|--------|--|------|--------|
| COI (cytochrome c oxidase subunit 1) | 162 | 1563 | ACO1 (aconitase 1 intron 9) | 42 | 1052 |
| COII (cytochrome c oxidase subunit 2) | 31 | 684 | BRM (brama protein intron 15 and partial cds) | 41 | 378 |
| CO3 (cytochrome c oxidase subunit 3) | 31 | 784 | CHDZ (chromosome Z chromo-helicase-DNA binding protein intron A) | 45 | 655 |
| Cytb (cytochrome b) | 272 | 1178 | FGB (beta-fibrinogen intron 5 and partial cds) | 96 | 575 |
| ND1 (NADH dehydrogenase subunit 1) | 30 | 978 | GAPDH (glyceraldehyde-3-phosphate dehydrogenase intron 11) | 135 | 429 |
| ND2 (NADH dehydrogenase subunit 2) | 261 | 1041 | LDH (lactate dehydrogenase) | 149 | 690 |
| ND3 (NADH dehydrogenase subunit 3) | 108 | 412 | MUSK (muscle skeletal receptor tyrosine kinase intron 3 and partial cds) | 61 | 615 |
| ND4 (NADH dehydrogenase subunit 4) | 31 | 1378 | MYO (myoglobin intron 2 and partial cds) | 235 | 748 |
| ND4L (NADH dehydrogenase subunit 4L) | 31 | 297 | ODC (ornithine decarboxylase introns 6–7 and partial cds) | 212 | 831 |
| ND5 (NADH dehydrogenase subunit 5) | 38 | 1821 | PEPCK (phosphoenolpyruvate carboxykinase intron 9 and partial cds) | 91 | 623 |
| ND6 (NADH dehydrogenase subunit 6) | 33 | 519 | RAG1 (recombination-activating protein 1 partial cds) | 52 | 2873 |
| ATP6 (ATP synthase membrane subunit 6) | 38 | 684 | TGFB2 (transforming growth factor beta-2 intron 5 and partial cds) | 65 | 609 |
| ATP8 (ATP synthase membrane subunit 8) | 34 | 168 | Z185 (anonymous nuclear locus Z-185 genomic sequence) | 39 | 898 |
| 12Srrn (12S ribosomal RNA) | 33 | 995 | | | |
| 16Srrn (16S ribosomal RNA) | 64 | 1620 | | | |

optimization possibilities. The optimal rate smoothing parameter was identified through cross-validation that tested seven values (0.001, 0.01, 0.1, 1, 10, 100, 1000); 0.001 was the optimal smoothing parameter that we used in this analysis.

We also repeated the above analysis twice using different calibration points: 1) the most recent common ancestor (MRCA) of the three out-group species, *Catharus fuscescens*, *Sialia sialis*, and *Turdus rufiventris* (min = 15.15 Mya, max = 20.75 Mya, optimal smoothing parameter = 1000), which corresponds to the crown age for Turdidae estimated in [Selvatti et al. \(2015\)](#); and 2) the crown age of Muscicapidae (min = 14.09 Mya, max = 19.10 Mya, optimal smoothing parameter = 10) from [Selvatti et al. \(2015\)](#).

3. Results

3.1. Taxon sampling and genetic data

We assembled genetic data for 301 Muscicapidae species, covering

all genera targeted and 92% of recognized species (out of 326 species and 50 genera). The super-matrix included 2460 sequences from a total of 28 molecular markers (15 mitochondrial and 13 nuclear; [Table 3](#)) with ca. 30% data coverage and an average of eight loci for each sampled species. Among them, 263 species were sampled with both mitochondrial and nuclear loci, 37 species with only mitochondrial loci, and one species with only nuclear loci. The resulting concatenated dataset contained 25,098 characters.

3.2. Phylogenetic relationships

Phylogenetic analyses using IQ-TREE (ML tree, [Fig. 1](#)) and MrBayes (the 50-majority-rule consensus tree, hereafter the Bayesian tree, [Fig. 2](#)) based on all taxa and all loci overall yielded similar topologies. Both trees showed four major lineages with strong support (> 95%), which corresponded to the four recognized subfamilies: Muscicapinae ([Figs. 1 & 2](#), clade A), Niltavinae (clade B), Cossyphinae (clade C) and Saxicolinae (clade D). 86% of the nodes in the ML tree had bootstrap



Fig. 1. Maximum likelihood tree estimated by IQ-TREE using a partitioned dataset containing up to 15 mitochondrial and 13 nuclear loci from 301 Muscicapidae and three Turdidae species. Clades A – C (top panel) and clade D (bottom panel) represent the four subfamilies. A1 represents the tribe Muscicapini, and A2 represents the tribe Copsychini. Values at nodes show bootstrap support; 100% bootstrap support is indicated with an asterisk. Red branches denote lineages that are in different places than in the Bayesian tree in [Fig. 2](#). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

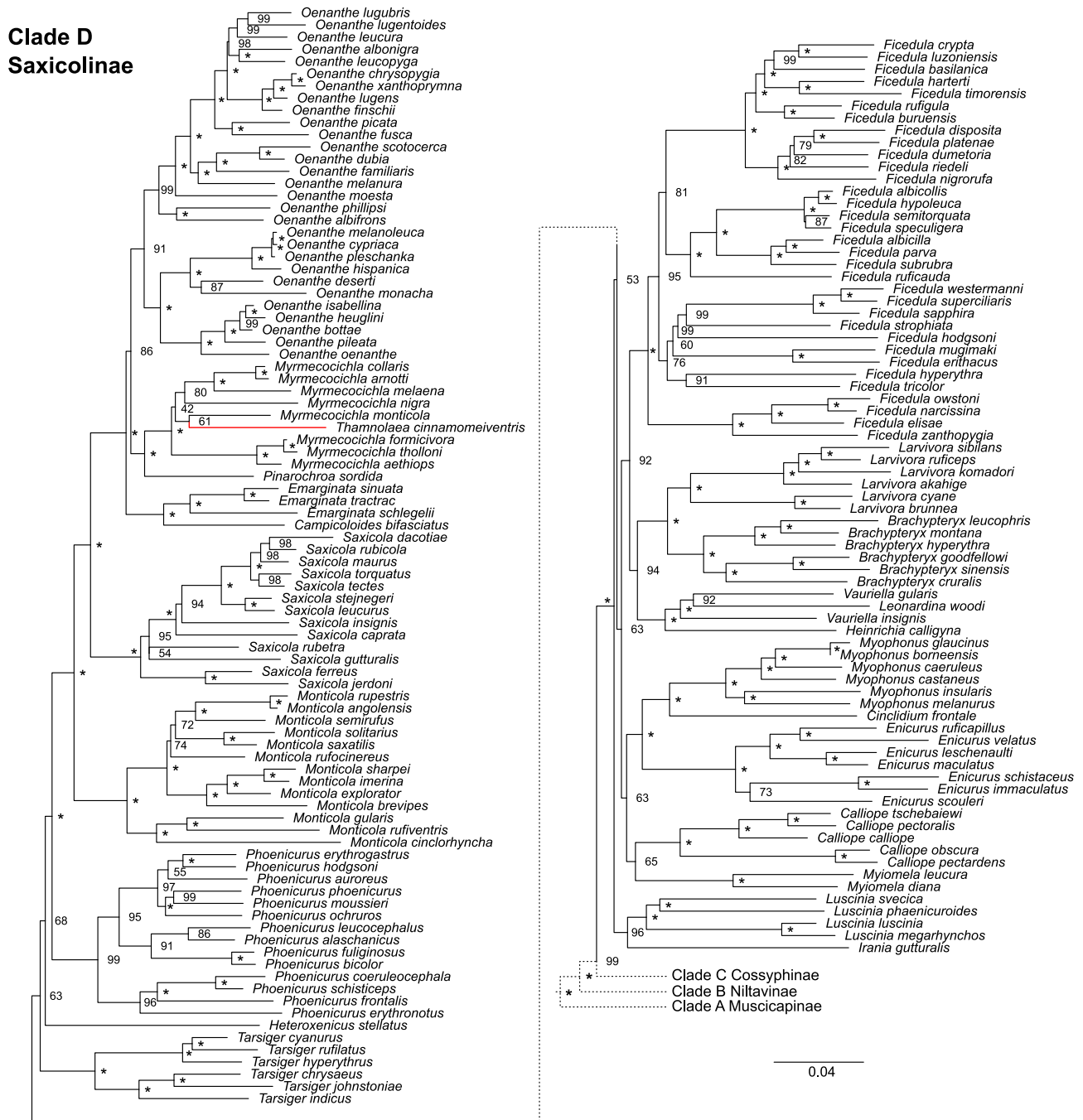


Fig. 1. (continued).

support $\geq 90\%$, and only 6.6% of the nodes had bootstrap support lower than 75%. The Bayesian tree was less resolved for some of the intra-generic relationships (e.g., *Niltava*, *Cyornis* and *Ficedula*) and failed to resolve some of the deep nodes in Niltavinae and Cossyphinae (Fig. 2); however, it agreed with the ML tree for most intergeneric relationships. The two tribes within Muscipapinae, Muscipapini (A1) and Copsychini (A2), are better supported by the ML tree based on all loci [bootstrap (BS) = (100, 90)] than the Bayesian tree [posterior probability (PP) = (1.00, 0.76)]. Within Saxicolinae, some of the deep nodes were connected by short internal branches, which were poorly supported in both ML tree and Bayesian trees (Figs. 1 & 2); however, the shallower relationships were more strongly supported by both analyses.

Mitochondrial and nuclear trees from both ML and Bayesian analyses all recovered four major clades (Supplementary Figs. S1–S4), however,

they differed in the placement of Niltavinae and Cossyphinae. Both nuclear trees supported that Niltavinae is sister to Saxicolinae and Cossyphinae (BS = 100; PP = 1.00); the same topology was recovered in our trees based on all loci, as well as in previous studies (e.g., Sangster et al., 2010; Zuccon and Ericson, 2010a; Barve and Mason, 2015). In contrast, in the mitochondrial trees, Niltavinae was found to be sister to Saxicolinae with only moderate support (BS = 78; PP = 0.72). For intergeneric relationships within each subfamily, mitochondrial and nuclear trees largely agreed with each other, but did exhibit a few topological discordances. For example, the nuclear trees supported that *Alethe* is sister to the rest of the Muscipapinae species, rendering the tribe Copsychini (A2) paraphyletic. The mitochondrial trees, on the other hand, placed *Alethe* within Copsychini.

Most of the taxa that have not been included in previous

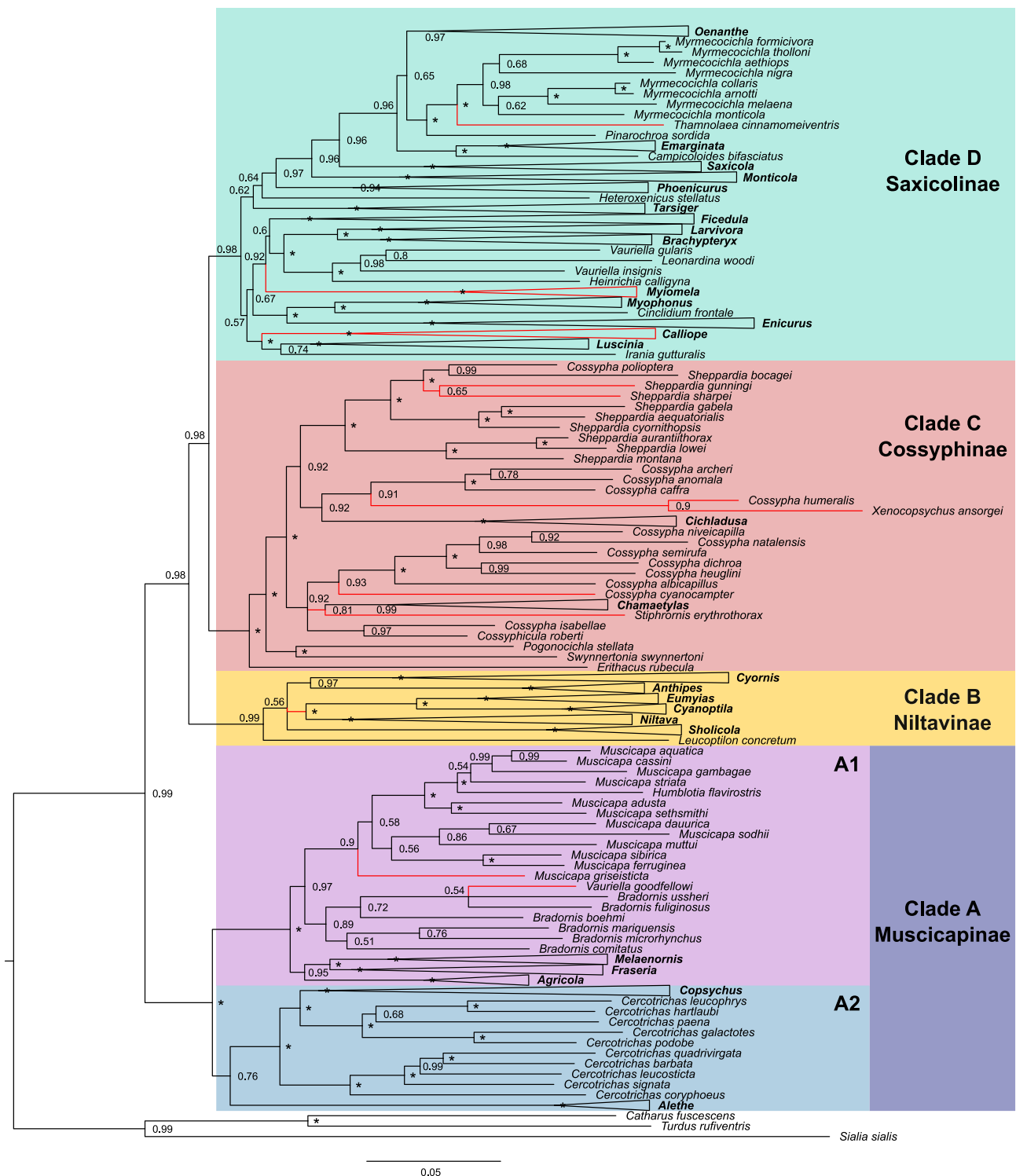


Fig. 2. The 50-majority-rule consensus tree reconstructed using Bayesian inference based on a partitioned dataset containing up to 15 mitochondrial and 13 nuclear loci for 301 Muscipidae species and three Turdidae species as outgroup. Genera that form reciprocally monophyletic clades have been collapsed and are in bold. Values at nodes show posterior probabilities; full support is indicated with an asterisk. Red branches denote lineages that are in different places than in the ML tree in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

phylogenetic trees were placed in their corresponding genera as defined by Clements et al. (2021). Twenty-seven of the 34 polytypic genera were monophyletic in both ML and Bayesian trees based on all loci with high support (> 90%); however, *Muscicapinae*, *Bradornis*, *Cercotrichas*, *Vauriella*, *Sheppardia*, *Cossypha* and *Myrmecocichla* were not reciprocally monophyletic (Figs. 1 & 2), with species assigned to *Vauriella* falling into

different subfamilies, which will be discussed in detail below. All original tree files are available in [supplementary data](#). Gene trees estimated by IQ-TREE are presented in [Supplementary Figs. S5–S32](#).

3.3. Divergence time estimation

Three TreePL analyses overall provided similar estimates of the divergence time across Muscicapidae (Table 4). Muscicapidae was estimated to diverge from Turdidae in the early Miocene and the most recent common ancestors for the four subfamilies all arose around middle Miocene. The TreePL tree with a crown age for Muscicapidae estimated to be 17.43 Mya (Fig. 3; Supplementary Fig. S33) has divergence times most similar to those in Selvatti et al. (2015), so its dates were used in Discussion below.

4. Discussion

This study provides the most comprehensive species-level time-calibrated phylogeny for Muscicapidae. Combining molecular data collected from various sources, most of the nodes were strongly supported by the phylogenetic analyses. For taxa that exhibited conflicts in topology among analyses, we discuss in detail below and propose taxonomic changes (Table 5). We also reconciled the taxon names in IOC World Bird List 12.1 that differ from those in Clements/eBird 2021 checklist (Table 6).

4.1. Muscicapinae

Clade A1 – paraphyly of *Muscicapa* flycatchers. The inclusion of the Grand Comoro Flycatcher *Humblotia flavirostris* renders *Muscicapa* paraphyletic. Although only *Cytb* was sampled in the data matrix for this species (GenBank accession: MH307512), concatenated (ML & Bayesian) and single-locus analyses all strongly support that *H. flavirostris* is nested within *Muscicapa* and is sister to the clade containing *M. aquatica*, *M. cassini*, *M. gambagae* and *M. striata* (BS = 97; PP = 1.00; BS = 100 in *Cytb* gene tree). This sample of *H. flavirostris* was collected from Grande Comore, where there are currently no *Muscicapa* flycatchers according to eBird sightings (<https://ebird.org/region/KM-G?yr=all>). As this species is sufficiently distinct from *Muscicapa* flycatchers in morphology, we conclude that it is not a mislabeled *Muscicapa* species. Based on our results, *Humblotia* Milne-Edwards and Oustalet, 1885 should be subsumed into *Muscicapa* Brisson, 1760. Therefore, we propose to rename *Humblotia flavirostris* as *Muscicapa flavirostris*. This is the first time that this species has been included in a comprehensive phylogenetic study.

Clade A1 – polyphyly of *Vauriella* jungle-flycatchers. The Mindanao Jungle Flycatcher *V. goodfellowi* is recovered in the Muscicapinae clade as most closely related to Muscicapini flycatchers, whereas the rest of the sampled *Vauriella* species (*V. gularis* and *V. insignis*) unexpectedly formed a well-supported clade with the Bagobo Robin *Leonardina woodi* (BS = 100; PP = 0.98) in the Saxicolinae clade, rendering *Vauriella* polyphyletic. The ML tree based on all loci placed *V. goodfellowi* as part of Muscicapini (BS = 100), whereas the Bayesian tree placed it within

Bradornis, forming a polytomy with *B. ussheri* and *B. fuliginosus* (PP = 0.54). Two previous studies also reported that *V. goodfellowi* appeared more closely related to the African flycatchers (Cibois and Cracraft, 2004; Barve and Mason, 2015), but both these were based on the same RAG1 sequence (GenBank accession: AY307211). There is now also a mitochondrial COI sequence of *V. goodfellowi* available on GenBank (KC354954), and the COI gene tree supports a sister relationship between *V. goodfellowi* and all *Muscicapa* flycatchers (BS = 99; Supplementary Fig. S9), while the RAG1 gene tree supports that *V. goodfellowi* is sister to the rest of Muscicapini (BS = 100; Supplementary Fig. S30), as in our ML tree. These two sequences were collected from different individuals and sequenced by different institutions (DNA Barcoding of Philippine Birds project, 2011, voucher MGP3; Field Museum, 1992, voucher FMNH 357498), therefore diminishing the possibility that this “misplacement” was solely due to mislabeling. Although the exact position of *V. goodfellowi* is not fully resolved, it is clear that it should be placed within Muscicapini instead of in Saxicolinae. Since *V. insignis* is the type species of *Vauriella* Wolters, 1980, a new generic name is warranted for *V. goodfellowi*, as well as phylogenetic analysis of the unsampled *V. albigularis*, which is provisionally assumed to be closely related to the two other current *Vauriella* species (see Clade D, below).

Clade A1 – taxonomic notes for *Bradornis*. The Bayesian trees based on all loci (Fig. 2) and mitochondrial loci (Supplementary Fig. S3) placed *Vauriella goodfellowi* within *Bradornis*, rendering it paraphyletic. However, both concatenated nuclear trees recovered *Bradornis* as a monophyletic clade (BS = 96; PP = 0.98; Supplementary Figs. S2 & S4), as well as the ML tree based on all loci (BS = 93; Fig. 1), with an estimated crown age at 10.26 Mya. Considering that the position of *V. goodfellowi* within Muscicapini was unstable, we also included a Bayesian inference based on all loci concatenated for just clade A1 but excluding *V. goodfellowi*. We found that the six sampled *Bradornis* species formed a strongly supported clade (PP = 1.00; Supplementary Fig. S34). Despite the enigma of *V. goodfellowi*, it is appropriate to continue using *Bradornis* as the valid genus name for *B. ussheri*, *B. fuliginosus*, *B. boehmi*, *B. mariquensis*, *B. microrhynchus* and *B. comitatus*.

Clade A2 – paraphyly of *Cercotrichas* scrub-robins. Our results showed that *Cercotrichas* was paraphyletic, as five of the 10 currently recognized species form a strongly supported clade that is sister to *Copsychus* magpie-robins while the remaining five are sister to that *Copsychus* plus *Cercotrichas* clade. The subdivision of *Cercotrichas* in our study agrees with the findings in recent phylogenetic work on this group (Sangster et al., 2010; Voelker et al., 2014). As suggested by Sangster et al. (2010), *C. podobe*, *C. galactotes*, *C. paena*, *C. hartlaubi* and *C. leucophrys* should retain the name *Cercotrichas* as *C. podobe* is the type species for *Cercotrichas*. Since *C. signata* is the type species for *Tychaodon* Richmond, 1917, *C. coryphoeus*, *C. signata*, *C. leucosticta*, *C. barbata* and *C. quadrivirgata* should be placed in *Tychaodon*, with *Salsicola* Oatley, 2004 as synonym.

4.2. Cossyphinae

Clade C – paraphyly of *Sheppardia* akalats. The clade comprising the *Sheppardia* akalats and the Gray-winged Robin-Chat *Cossypha polioptera* was strongly supported by both ML and Bayesian trees (BS = 100; PP = 0.99). All trees based on concatenation support a sister relationship between *C. polioptera* and *S. bocagei*, except for the concatenated nuclear trees (Supplementary Figs. S2 & S4) as *S. bocagei* lacks nuclear data. In agreement to several previous studies (Sangster et al., 2010; Voelker et al., 2010; del Hoyo et al., 2016; Fjeldsø and Bowie, 2021), our results support the placement of *C. polioptera* in *Sheppardia*.

Clade C – polyphyly of *Cossypha* robin-chats. The phylogenetic relationships among these African robin-chats remain highly inconsistent across different studies, likely due to incomplete taxon sampling (Zuccon, 2011). Our study sampled 13 *Cossypha* species (all currently recognized species except for *C. heinrichi*), which were placed in five

Table 4

Divergence time (million years ago) of major nodes estimated by TreePL using three different calibration points for our study and results from two previous studies. An asterisk represents where we put the secondary calibration point in the analysis. Dates in bold were used in the Discussion.

| | This study | | | Selvatti et al. (2015) | Oliveros et al. (2019) |
|-------------------------|---------------|--------|--------|------------------------|------------------------|
| Muscicapidae | 17.43 | 18.66 | 19.10* | 16.35 [14.09, 19.10] | NA |
| Muscicapinae | 14.76 | 15.32 | 16.13 | NA | NA |
| Niltavinae | 14.00 | 14.95 | 15.63 | NA | NA |
| Cossyphinae | 13.82 | 14.57 | 15.16 | NA | NA |
| Saxicolinae | 14.19 | 14.98 | 15.70 | NA | NA |
| Turdidae | 18.14 | 20.75* | 21.37 | 17.79 [15.15, 20.75] | NA |
| Muscicapidae + Turdidae | 21.96* | 24.87 | 25.62 | 21.02 [18.42, 24.13] | 17.75 [13.50, 22.00] |

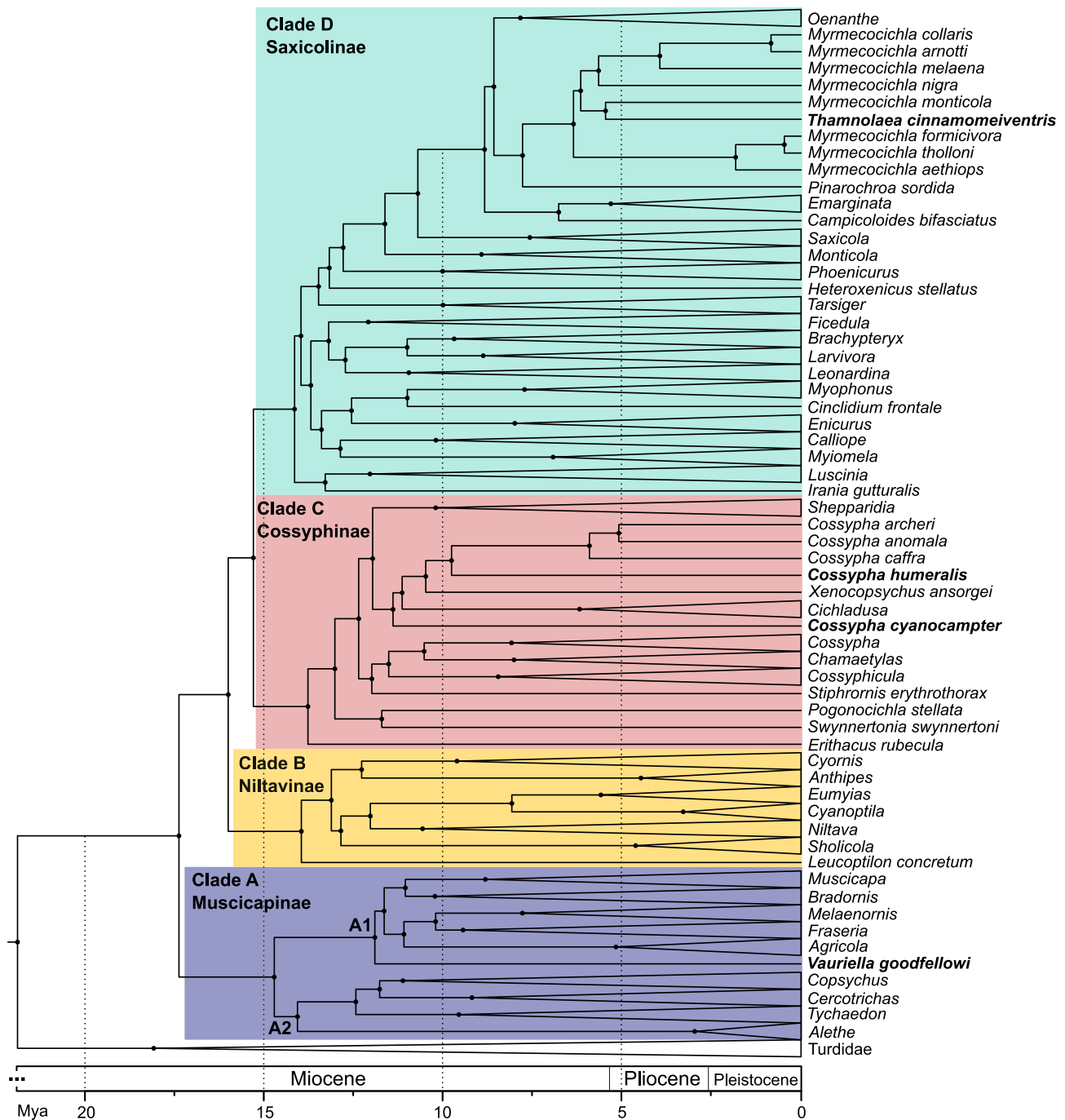


Fig. 3. Divergence time in million years ago (Mya) estimated by TreePL using the topology of the ML tree based on all loci. A secondary calibration point for the split between Muscicapidae and Turdidae (min = 13.5 Mya, max = 22 Mya) from [Oliveros et al. \(2019\)](#) was applied. Dotted vertical lines were drawn every-five million years. Genera that form reciprocally monophyletic clades have been collapsed. Tip labels in bold are the taxa that were placed in different positions by the ML and Bayesian analyses. The tree with all collapsed clades expanded is available in [Supplementary Fig. S33](#).

different positions in the ML tree ([Fig. 1](#)): (1) *C. polioptera* within *Sheppardia* (see above); (2) *C. isabellae* was sister to the monotypic *Cossyphicula* (BS = 100; PP = 0.97); (3) *C. archeri*, *C. anomala*, *C. caffra* and *C. humeralis* formed a strongly supported clade (BS = 99) that was sister to the monotypic Angola Cave-Chat *Xenocopsychus ansorgei* (BS = 94), although *C. humeralis* showed relatively deep divergence compared to the other three *Cossypha* species; (4) *C. cyanocampter* was sister to the clade containing the previous four *Cossypha* species, *Xenocopsychus* and *Cichladusa* palm-thrushes, with moderate support (BS = 78); (5) the remaining six *Cossypha* species formed a clade with strong support (BS = 100; PP = 1.00) that was sister to *Chamaetylas alethes* (BS = 93).

The Bayesian analysis generally agreed with the above placements, except that *C. cyanocampter* was sister to the clade containing six *Cossypha* species (PP = 0.93; [Fig. 2](#)), and that *C. humeralis* was sister to *Xenocopsychus* (PP = 0.90). Given that the branch uniting *C. humeralis* and *Xenocopsychus* appears to be exceptionally long in the Bayesian tree (two long branches in the ML tree), the sister relationship between these two species likely resulted from long-branch attraction ([Bergsten, 2005](#)). Both ML and Bayesian mitochondrial trees supported a sister relationship between *C. cyanocampter* and the six *Cossypha* species with strong support (BS = 99; PP = 0.94; [Supplementary Figs. S1 & S3](#)). However, the two nuclear trees placed *C. cyanocampter* within this clade as sister to

Table 5

Recommended taxonomic changes given results in our study and previous work. For taxon names in IOC World Bird List 12.1 that differ from those in Clements/eBird 2021 checklist and from the results shown in present study, see Table 6 for reconciliations.

| Name in Clements/eBird 2021 and IOC World Bird List 12.1 | Recommended name given results in our study and previous work |
|--|---|
| <i>Humboldtia flaviviridis</i> | <i>Muscicapa flaviviridis</i> |
| <i>Cercotrichas coryphoeus</i> | <i>Tychaemon coryphoeus</i> |
| <i>Cercotrichas signata</i> | <i>Tychaemon signata</i> |
| <i>Cercotrichas leucosticta</i> | <i>Tychaemon leucosticta</i> |
| <i>Cercotrichas barbata</i> | <i>Tychaemon barbata</i> |
| <i>Cercotrichas quadrivirgata</i> | <i>Tychaemon quadrivirgata</i> |
| <i>Cossypha polioptera</i> | <i>Sheppardia polioptera</i> |
| <i>Cossypha isabellae</i> | <i>Cossyphicula isabellae</i> |
| <i>Cossypha anomala</i> | <i>Caffronis anomala</i> |
| <i>Cossypha archeri</i> | <i>Caffronis archeri</i> |
| <i>Cossypha caffra</i> | <i>Caffronis caffer</i> |
| <i>Cossypha humeralis</i> | <i>Bessonornis humeralis</i> * |
| <i>Vauriella gularis</i> | <i>Leonardina gularis</i> |
| <i>Vauriella insignis</i> | <i>Leonardina insignis</i> |
| <i>Vauriella albigularis</i> | <i>Leonardina albigularis</i> ** |
| <i>Heinrichia calligyna</i> | <i>Leonardina calligyna</i> |

* See alternative suggestion in Discussion.

** Provisionally assumed to be closely related to *L. gularis* and *L. insignis*, pending further phylogenetic analysis.

Table 6

Recommended taxonomic changes for taxon names in IOC World Bird List 12.1 which differ from those in Clements/eBird 2021 checklist.

| Taxon name in IOC 12.1 | Taxon name in Clements/eBird 2021 | Recommend name given our results |
|----------------------------------|-----------------------------------|----------------------------------|
| <i>Empidonax semipartitus</i> | <i>Melaenornis semipartitus</i> | <i>Melaenornis semipartitus</i> |
| <i>Melaenornis infuscatus</i> | <i>Agricola infuscatus</i> | <i>Agricola infuscatus</i> |
| <i>Melaenornis mariquensis</i> | <i>Bradornis mariquensis</i> | <i>Bradornis mariquensis</i> |
| <i>Melaenornis microrhynchus</i> | <i>Bradornis microrhynchus</i> | <i>Bradornis microrhynchus</i> |
| <i>Melaenornis pallidus</i> | <i>Agricola pallidus</i> | <i>Agricola pallidus</i> |
| <i>Muscicapa boehmi</i> | <i>Bradornis boehmi</i> | <i>Bradornis boehmi</i> |
| <i>Muscicapa caerulea</i> | <i>Fraseria caerulea</i> | <i>Fraseria caerulea</i> |
| <i>Muscicapa comitata</i> | <i>Bradornis comitata</i> | <i>Bradornis comitata</i> |
| <i>Muscicapa infusca</i> | <i>Bradornis fuliginosus</i> | <i>Bradornis fuliginosus</i> |
| <i>Muscicapa lendu</i> | <i>Fraseria lendu</i> | <i>Fraseria lendu</i> |
| <i>Muscicapa olivacea</i> | <i>Fraseria olivacea</i> | <i>Fraseria olivacea</i> |
| <i>Muscicapa tessmanni</i> | <i>Fraseria tessmanni</i> | <i>Fraseria tessmanni</i> |
| <i>Muscicapa ussheri</i> | <i>Bradornis ussheri</i> | <i>Bradornis ussheri</i> |
| <i>Myioparus griseigularis</i> | <i>Fraseria griseigularis</i> | <i>Fraseria griseigularis</i> |
| <i>Myioparus plumbeus</i> | <i>Fraseria plumbea</i> | <i>Fraseria plumbea</i> |
| <i>Namibornis herero</i> | <i>Melaenornis herero</i> | <i>Melaenornis herero</i> |

C. niveicapilla (BS = 98; PP = 0.86; Supplementary Figs. S2 and S4), similarly in Beresford (2003), but not in Sangster et al. (2010). It is possible that *C. cyanocampter* might be more closely related to the clade containing the six *Cossypha* species than to the other *Cossypha* robin-chats. However, due to the disagreement among mitochondrial, nuclear and all loci concatenated trees, the exact position of *C. cyanocampter* still remains unresolved.

Our concatenated trees all recovered *Cossypha isabellae* and *Cossyphicula roberti* as sister taxa with strong support (BS = 100; PP = 0.97). *Cossypha isabellae* was sampled for Cytb and RAG1, and *Cossyphicula* only for Cytb. The RAG1 gene tree (Supplementary Fig. S30) puts *C. isabellae* outside the clade containing *Cossypha* and *Chamaetylas* (BS = 94), whereas the Cytb gene tree supports a sister relationship for *C. isabellae* and *Cossyphicula* (BS = 99). We tentatively propose to place *C. isabellae* in *Cossyphicula*. Future studies with increased molecular

sampling should better evaluate their relationship.

Our results suggest that *C. caffra*, *C. archeri*, *C. anomala* and *C. humeralis* should have a separate generic status apart from the rest of the *Cossypha* robin-chats. Zuccon (2011) pointed out that the name *Caffronis* Roberts, 1922 (type species *C. caffra*) is available for *C. caffra*, *C. archeri*, and *C. anomala*, which in our analyses form a well-supported clade (BS = 100; PP = 1.00) with an estimated most recent common ancestor at 5.93 Mya. There is strong (BS = 99) or fairly strong (PP = 0.91) support for inclusion of *C. humeralis* in the same clade as the three previous species, although it is more anciently diverged from these, with an estimated divergence time of 9.79 Mya. If these four species are placed in the same genus, the name *Bessonornis* (Smith, 1844) (the corrected spelling for *Dessonornis* A. Smith, 1836) has priority over *Caffronis* Roberts, 1922. Since *Xenocopsychus ansorgei* is supported as either sister to the four others (BS = 94) or as sister to *C. humeralis* (PP = 0.90), this species is also appropriately placed in the same genus, which in that case would be *Bessonornis* A. Smith, 1836, which has priority over *Xenocopsychus* Hartert, 1907. An alternative would be to recognize three genera, *Caffronis* (including *C. caffer*, *C. archeri* and *C. anomala*), *Bessonornis* (monotypic: *B. humeralis*) and *Xenocopsychus* (monotypic: *X. ansorgei*).

Clade C – additional taxonomic notes. The ML and Bayesian trees based on all loci (Figs. 1 & 2) strongly supported *Pogonocichla stellata* and *Swynnertonina swynnertonii* as sister taxa (BS = 96; PP = 1.00). This sister relationship was also reported in Sangster et al. (2010). Our results suggest that *Swynnertonina* Roberts, 1922 could be subsumed in *Pogonocichla* Cabanis, 1847, and *S. swynnertonii* renamed as *Pogonocichla swynnertonii*. However, in view of the deep divergence between these (11.75 Mya), continued treatment as monotypic genera is equally valid.

4.3. Saxicolinae

Although most of the relationships were well resolved and strongly supported by the concatenated trees, a few deep nodes in Saxicolinae still remained poorly supported, especially for the nodes uniting *Ficedula* flycatchers and the forest understory birds that are now mainly distributed in subtropical and tropical Asia (e.g., *Enicurus*, *Myophonus*, *Cinclidium*, *Larvivora*, *Brachypteryx*, *Myiomela*, *Calliope*, etc.). These groups may have evolved through rapid radiations. Therefore, they have not had the time to accumulate sufficient substitutions, which makes fully resolving their phylogenetic relationships challenging.

Clade D – paraphyly of *Myrmecocichla* chats. In the ML tree based on all loci concatenated (Fig. 1), the Mocking Cliff-Chat *Thamnolaea cinnamomeiventris* is nested within *Myrmecocichla* and sister to *M. monticola*, rendering *Myrmecocichla* paraphyletic, although the sister relationship between *T. cinnamomeiventris* and *M. monticola* was poorly supported (BS = 61). Sangster et al. (2010) and Barve and Mason (2015) also reported *T. cinnamomeiventris* and *M. monticola* as sister taxa. However, the mitochondrial ML tree (Supplementary Fig. S1) and the Bayesian tree based on all loci (Fig. 2) strongly support a sister relationship between *T. cinnamomeiventris* and all *Myrmecocichla* chats (BS = 100; PP = 1.00), as well as in some previous studies (Aliabadian et al., 2012; Voelker et al., 2012; Alaei Kakhki et al., 2016). In our nuclear concatenated trees (Supplementary Figs. S2 & S4), *T. cinnamomeiventris* is part of an unresolved polytomy with *Myrmecocichla* chats and the Moorland Chat *Pinarochroa sordida*. Collectively, this suggests that *Thamnolaea* is more likely to be sister to all *Myrmecocichla* chats. Future phylogenetic work with increased molecular sampling for *Thamnolaea* and *Myrmecocichla* chats as well as increased taxon sampling within the polytypic *T. cinnamomeiventris* may be warranted.

Clade D – polyphyly of *Vauriella* jungle-flycatchers. Both ML and Bayesian trees (Figs. 1 & 2) found that *Leonardina woodi* is nested with two current members of *Vauriella*: *V. gularis* and *V. insignis* (BS = 100; PP = 0.98) and is sister to *V. gularis* (BS = 92; PP = 0.80). This has also been reported by Oliveros et al. (2012) and Fjeldsø et al. (2020). *Heinrichia calligyna* is strongly supported as sister to these (BS = 100; PP = 1.00),

with relatively shallow divergence (MRCA at 10.99 Mya) from the previous three species (MRCA at 9.99 Mya), and based on this it seems reasonable to include all of these in a single genus. One option is to combine *V. gularis*, *V. insignis*, and tentatively *V. albigularis*, with *H. calligyna* and *L. woodi* in *Leonardina* Mearns, 1905, by priority (*Heinrichia* was introduced by Stresemann 1931 and *Vauriella* by Wolters 1980).

4.4. Taxa in need of further phylogenetic study

Based on the Clements/eBird 2021 checklist (excluding *Pinarornis plumosus*), there are 25 unsampled Muscicapidae species in our study (Supplementary Table S2). Most of them either occur only on islands or are under threat of extinction. About 30% of the missing taxa are in the genus *Cyornis*, including the critically endangered island-dwelling *C. ruckii*. Advances in next-generation sequencing technologies (such as the probe capture for target enrichment) have enabled us to obtain substantial amount of molecular data from toepad samples in museum collections, which could be a promising future step for sampling these difficult species.

Additionally, 17 subspecies in Clements/eBird 2021 checklist, for example *Larvivora komadori namiyei* (*L. namiyei*) and *Thamnolaea cinnamomeiventris coronata* (*T. coronata*), have been elevated to the species level in IOC World Bird List 12.1 (Supplementary Table S3). Although these taxa were not included in our study, future studies can easily incorporate them, since most of them have molecular data available in GenBank. Careful curation of taxon name, sampling locality and associated literature may be required to accurately assign additional GenBank sequences to the correct taxa in cases of taxonomic change (see Hosner et al., 2022).

5. Conclusion

Our study provides the most comprehensive species-level phylogeny for Muscicapidae, covering 92% of the species from all recognized genera. Most of the nodes are well resolved and strongly supported. For genera that appeared to be non-monophyletic in our analyses, we proposed taxonomic revisions based on results both from our study and previous work. Given the number of taxonomic changes already observed, it is likely that more complete taxon sampling, including expanded sampling from subspecies and distinct populations, may further clarify patterns of diversification within the Muscicapidae. In addition, use of next-generation sequencing approaches may be particularly helpful to resolve those nodes with limited support, particularly within the Saxicolinae.

CRedit authorship contribution statement

Min Zhao: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **J. Gordon Burleigh:** Methodology, Software, Investigation, Data curation, Writing – review & editing. **Urban Olsson:** Investigation, Resources, Data curation, Writing – review & editing. **Per Alström:** Conceptualization, Resources, Data curation, Writing – review & editing, Funding acquisition. **Rebecca T. Kimball:** Conceptualization, Methodology, Writing – review & editing, Supervision, 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.

Acknowledgements

We are grateful to Edward Braun, Akanksha Pandey and Sarah Kurtis for their support in sequence data extraction. We thank Sahas Barve and Nick Mason for sharing the tree file from their 2015 study with us. We are grateful to George Sangster for his valuable comments on this manuscript. We thank photographers Yan Shen, Zak Pohlen, Toni Ferrer, Mike Rose, Roger Clark and Ed Thomas for their photos that the bird illustrations in the graphical abstract were painted from (detailed photo credits in Supplementary Table S4). P.A. was supported by the Swedish Research Council (grant No. 2019-04486). R.T.K. was supported by the United States National Science Foundation (DEB 1655683). We are grateful to the two anonymous reviewers who provided comments that improved the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2022.107646>.

References

- Alaee Kakhki, N., Aliabadian, M., Schweizer, M., 2016. Out of Africa: Biogeographic history of the open-habitat chats (Aves, Muscicapidae: Saxicolinae) across arid areas of the Old World. *Zool. Scr.* 45, 237–251.
- Aliabadian, M., Kaboli, M., Förschler, M.I., Nijman, V., Chamani, A., Tillier, A., Prodon, R., Pasquet, E., Ericson, P.G.P., Zuccon, D., 2012. Convergent evolution of morphological and ecological traits in the open-habitat chat complex (Aves, Muscicapidae: Saxicolinae). *Mol. Phylogenet. Evol.* 65, 35–45.
- Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdoci, F., Nabholz, B., Delsuc, F., 2020. MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Mol. Ecol. Resour.* 20, 892–905.
- Alström, P., Rasmussen, P.C., Xia, C., Gelang, M., Liu, Y., Chen, G., Zhao, M., Hao, Y., Zhao, C., Zhao, J., 2018. Taxonomy of the White-browed Shortwing (*Brachypteryx montana*) complex on mainland Asia and Taiwan: an integrative approach supports recognition of three instead of one species. *Avian Res.* 9, 1–13.
- Barve, S., Mason, N.A., 2015. Interspecific competition affects evolutionary links between cavity nesting, migration and clutch size in Old World flycatchers (Muscicapidae). *Ibis (Lond. 1859)* 157, 299–311.
- Beresford, P., 2003. Molecular systematics of *Alethe*, *Sheppardia* and some other African robins (Muscipapidae). *Ostrich* 74, 58–73.
- Bergsten, J., 2005. A review of long-branch attraction. *Cladistics* 21, 163–193.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. *BMC Bioinf.* 10, 1–9.
- Cibois, A., Cracraft, J., 2004. Assessing the passerine “Tapestry”: phylogenetic relationships of the Muscipoidea inferred from nuclear DNA sequences. *Mol. Phylogenet. Evol.* 32, 264–273.
- Clements, J., Schulenberg, T., Iliff, M., Billerman, S., Fredericks, T., Gerbracht, J.A., Lepage, D., Sullivan, B., Wood, C., 2021. The eBird/Clements checklist of birds of the world: v2021. Downloaded from <https://www.birds.cornell.edu/clementschecklist/download/>.
- del Hoyo, J., Collar, N.J., Christie, D.A., Elliott, A., Fishpool, L.D.C., Boesman, P., Kirwan, G.M., 2016. HBW and BirdLife International Illustrated Checklist of the Birds of the World. Volume 2: Passerines. Lynx Edicions and BirdLife International, Barcelona, Spain and Cambridge, UK.
- Dickinson, E., Christidis, L., 2014. The Howard and Moore Complete Checklist of the Birds of the World.
- Fjeldså, J., Bowie, R., 2021. Evolutionary and ecological explanations for the elevational flexibility of several East African bird species complexes. *Front. Ecol. Evol.* 885.
- Fjeldså, J., Zuccon, D., Irestedt, M., Johansson, U.S., Ericson, P.G.P., 2003. *Sapayoa aenigma*: a New World representative of ‘Old World’ suboscines. *Proc. R. Soc. London. Ser. B Biol. Sci.* 270, S238–S241.
- Fjeldså, J., Alström, P., Olsson, U., 2020. Superfamily Muscipoidea: mockingbirds, starlings, chats, Old World flycatchers, thrushes and allies. In: Fjeldså, J., Ericson, P.G.P., Christidis, L. (Eds.), *The Largest Avian Radiation: The Evolution of Perching Birds, or the Order Passeriformes*. Lynx Edicions, Barcelona.
- Friesen, V.L., Congdon, B.C., Kidd, M.G., Birt, T.P., 1999. Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Mol. Ecol.* 8, 2147–2149.
- Garg, K.M., Chattopadhyay, B., Wilton, P.R., Prawiradilaga, D.M., Rheindt, F.E., 2018. Pleistocene land bridges act as semipermeable agents of avian gene flow in Wallacea. *Mol. Phylogenet. Evol.* 125, 196–203.
- Gill, F., Donsker, D., Rasmussen, P., 2022. IOC world bird list v 12.1. IOC.
- Helbig, A.J., Seibold, I., Martens, J., Wink, M., 1995. Genetic differentiation and phylogenetic relationships of Bonelli’s Warbler *Phylloscopus bonelli* and Green Warbler *P. nitidus*. *J. Avian Biol.* 139–153.
- Heslewood, M.M., Elphinstone, M.S., Tidemann, S.C., Baverstock, P.R., 1998. Myoglobin intron variation in the Gouldian Finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoresis. *Electrophoresis* 19, 142–151.

- Hooper, D.M., Olsson, U., Alström, P., 2016. The Rusty-tailed Flycatcher (*Muscicapa ruficauda*; Aves: Muscipidae) is a member of the genus *Ficedula*. *Mol. Phylogenet. Evol.* 102, 56–61.
- Hosner, P.A., Zhao, M., Kimball, R.T., Braun, E.L., Burleigh, J.G., 2022. Updating splits, lumps, and shuffles: Reconciling GenBank names with standardized avian taxonomies. *Ornithology* 139 (4), 1–15.
- Illera, J.C., Richardson, D.S., Helm, B., Atienza, J.C., Emerson, B.C., 2008. Phylogenetic relationships, biogeography and speciation in the avian genus *Saxicola*. *Mol. Phylogenet. Evol.* 48, 1145–1154.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547.
- Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701.
- Lim, H.C., Zou, F., Taylor, S.S., Marks, B.D., Moyle, R.G., Voelker, G., Sheldon, F.H., 2010. Phylogeny of magpie-robins and shamas (Aves: Turdidae: *Copsychus* and *Trichixos*): implications for island biogeography in Southeast Asia. *J. Biogeogr.* 37, 1894–1906.
- Liu, Y., Chen, G., Huang, Q., Jia, C., Carey, G., Leader, P., Li, Y., Zou, F., Yang, X., Olsson, U., 2016. Species delimitation of the white-tailed rubythroat *Calliope pectoralis* complex (Aves, Muscipidae) using an integrative taxonomic approach. *J. Avian Biol.* 47, 899–910.
- Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Gatew. Comput. Environ. Work.* 1–8.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von 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.
- Moyle, R.G., Hosner, P.A., Jones, A.W., Outlaw, D.C., 2015. Phylogeny and biogeography of *Ficedula* flycatchers (Aves: Muscipidae): Novel results from fresh source material. *Mol. Phylogenet. Evol.* 82, 87–94.
- Ng, E.Y.X., Li, S., Zhang, D., Garg, K.M., Song, G., Martinez, J., Manh, H.L., Tu, V.T., Fuchs, J., Dong, L., Olsson, U., Huang, Y., Alström, P., Rheindt, F.E., Lei, F., 2022. Genome-wide SNPs confirm plumage polymorphism and hybridization within a *Cyornis* flycatcher species complex. *Scr. Zool.* <https://doi.org/10.1111/zsc.12568>.
- Oatley, T.B., 2004. A re-appraisal of the taxonomic status of the Karoo Scrub-robin. *Ostrich J. Afr. Ornithol.* 75, 156–158.
- Oliveros, C.H., Reddy, S., Moyle, R.G., 2012. The phylogenetic position of some Philippine “babbler” spans the muscipoid and sylvioid bird radiations. *Mol. Phylogenet. Evol.* 65, 799–804.
- 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, 7916–7925.
- Outlaw, D.C., Voelker, G., 2006. Systematics of *Ficedula* flycatchers (Muscipidae): A molecular reassessment of a taxonomic enigma. *Mol. Phylogenet. Evol.* 41, 118–126.
- Outlaw, R.K., Voelker, G., Outlaw, D.C., 2007. Molecular systematics and historical biogeography of the rock-thrushes (Muscipidae: *Monticola*). *Auk* 124, 561–577.
- Outlaw, R.K., Voelker, G., Bowie, R.C.K., 2010. Shall we chat? Evolutionary relationships in the genus *Cercomela* (Muscipidae) and its relation to *Oenanthe* reveals extensive polyphyly among chats distributed in Africa, India and the Palearctic. *Mol. Phylogenet. Evol.* 55, 284–292.
- Price, T.D., Hooper, D.M., Buchanan, C.D., Johansson, U.S., Tietze, D.T., Alström, P., Olsson, U., Ghosh-Harihar, M., Ishtiaq, F., Gupta, S.K., 2014. Niche filling slows the diversification of Himalayan songbirds. *Nature* 509, 222–225.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901.
- Rambaut A. 2018. FigTree version 1.4.4. Available at: <http://tree.bio.ed.ac.uk/software/figtree/>.
- Rasmussen P.C., Anderton J.C. 2005. Birds of south Asia: the Ripley guide.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Roy, M.S., Spomer, R., Fjeldsø, J., 2001. Molecular systematics and evolutionary history of akalats (genus *Sheppardia*): a pre-Pleistocene radiation in a group of African forest birds. *Mol. Phylogenet. Evol.* 18, 74–83.
- Sangster, G., Alström, P., Forsmark, E., Olsson, U., 2010. Multi-locus phylogenetic analysis of Old World chats and flycatchers reveals extensive paraphyly at family, subfamily and genus level (Aves: Muscipidae). *Mol. Phylogenet. Evol.* 57, 380–392.
- Sangster, G., Alström, P., Gaudin, J., Olsson, U., 2021. A new genus for the White-tailed Flycatcher *Cyornis concretus* (Aves: Muscipidae). *Zootaxa* 5072, 599–600.
- Sangster, G., Luksenburg, J.A., 2021. Sharp increase of problematic mitogenomes of birds: Causes, consequences, and remedies. *Genome Biol. Evol.* 13, evab210.
- Selvatti, A.P., Gonzaga, L.P., de Moraes, R.C.A., 2015. A Paleogene origin for crown passerines and the diversification of the Oscines in the New World. *Mol. Phylogenet. Evol.* 88, 1–15.
- Slade, R.W., Moritz, C., Heideman, A., Hale, P.T., 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. *Mol. Ecol.* 2, 359–373.
- Smith, S.A., Dunn, C.W., 2008. Phylutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* 24, 715–716.
- Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* 28, 2689–2690.
- Smith, A., 1844. Illustrations of the Zoology of South Africa; Consisting Chiefly of Figs and Descriptions of the Objects of Natural History Collected During an Expedition into the Interior of South Africa, in the Years 1834, 1835, and 1836. Volume II, pl. 48. London, Smith, Elder and Co, pp. 1838–1849.
- Voelker, G., Outlaw, R.K., Bowie, R.C.K., 2010. Pliocene forest dynamics as a primary driver of African bird speciation. *Glob. Ecol. Biogeogr.* 19, 111–121.
- Voelker, G., Bowie, R.C.K., Wilson, B., Anderson, C., 2012. Phylogenetic relationships and speciation patterns in an African savanna dwelling bird genus (*Myrmecocichla*). *Biol. J. Linn. Soc.* 106, 180–190.
- Voelker, G., Peñalba, J.V., Huntley, J.W., Bowie, R.C.K., 2014. Diversification in an Afro-Asian songbird clade (*Erythropgia-Copsychus*) reveals founder-event speciation via trans-oceanic dispersals and a southern to northern colonization pattern in Africa. *Mol. Phylogenet. Evol.* 73, 97–105.
- Voelker, G., Huntley, J.W., Peñalba, J.V., Bowie, R.C.K., 2016. Resolving taxonomic uncertainty and historical biogeographic patterns in *Muscicapa* flycatchers and their allies. *Mol. Phylogenet. Evol.* 94, 618–625.
- Voelker, G., Spellman, G.M., 2004. Nuclear and mitochondrial DNA evidence of polyphyly in the avian superfamily Muscipoidae. *Mol. Phylogenet. Evol.* 30, 386–394.
- Wei, C., Sangster, G., Olsson, U., Rasmussen, P.C., Svensson, L., Yao, C., Carey, G.J., Leader, P.J., Zhang, R., Chen, G., Song, G., Lei, F., Wilcove, D.S., Alström, P., Liu, Y., 2022. Cryptic species in a colorful genus: integrative taxonomy of the bush robins (Aves, Muscipidae, *Tarsiger*) suggests two overlooked species. *Mol. Phylogenet. Evol.* 175, 107580.
- Wheeler, T.J., Eddy, S.R., 2013. nhmmer: DNA homology search with profile HMMs. *Bioinformatics* 29, 2487–2489.
- Winkler, D.W., Billerman, S.M., Lovette, I.J., 2020. Old World Flycatchers (Muscipidae), version 1.0. In: Billerman, S.M., Keeney, B.K., Rodewald, P.G., Schulenberg, T.S. (Eds.), *Birds of the World*. Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.muscic3.01>.
- Zhang, Z., Wang, X., Huang, Y., Olsson, U., Martinez, J., Alström, P., Lei, F., 2016. Unexpected divergence and lack of divergence revealed in continental Asian *Cyornis* flycatchers (Aves: Muscipidae). *Mol. Phylogenet. Evol.* 94, 232–241.
- Zhao, M., Alström, P., Hu, R., Zhao, C., Hao, Y., Lei, F., Qu, Y., 2017. Phylogenetic relationships, song and distribution of the endangered Rufous-headed Robin *Larvivora ruficeps*. *Ibis* 159, 204–216.
- Zuccon, D., 2011. Taxonomic notes on some Muscipidae. *Br. Ornithol. Club* 131, 196–199.
- Zuccon, D., Ericson, P.G.P., 2010a. A multi-gene phylogeny disentangles the chat-flycatcher complex (Aves: Muscipidae). *Zool. Scr.* 39, 213–224.
- Zuccon, D., Ericson, P.G.P., 2010b. The *Monticola* rock-thrushes: Phylogeny and biogeography revisited. *Mol. Phylogenet. Evol.* 55, 901–910.