Diversification history in the *Dendrocincla fuliginosa* complex (Aves: Dendrocolaptidae): insights from broad geographic sampling

Eduardo D. Schultz, Jorge Pérez-Emán, Alexandre Aleixo, Cristina Y. Miyaki, Robb T. Brumfield, Joel Cracraft, Camila C. Ribas

PII: S1055-7903(18)30787-5
DOI: https://doi.org/10.1016/j.ympev.2019.106581
Reference: YMPEV 106581

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 8 December 2018
Revised Date: 17 July 2019
Accepted Date: 12 August 2019


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier Inc.
Diversification history in the *Dendrocincla fuliginosa* complex (Aves: Dendrocolaptidae): insights from broad geographic sampling.

Eduardo D. Schultz\(^a\)*, Jorge Pérez-Emán\(^b\)\(^c\), Alexandre Aleixo\(^d\)\(^e\)\(^l\), Cristina Y. Miyaki\(^f\), Robb T. Brumfield\(^g\)\(^h\), Joel Cracraft\(^i\), Camila C. Ribas\(^j\)

\(^a\) Programa de Pós Graduação em Biologia (Ecologia), Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2936 Manaus, AM, Brazil.
\(^b\) Instituto de Zoología y Ecología Tropical, Universidad Central de Venezuela, Caracas 1041-A, Venezuela.
\(^c\) Colección Ornitológica Phelps, Caracas 1010-A, Venezuela.
\(^d\) Museu Paraense Emílio Goeldi, Coordenação de Zoologia, Caixa Postal 399, 66040-170 Belém, Pará, Brazil.
\(^e\) Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland.
\(^f\) Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil.
\(^g\) Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803.
\(^h\) Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803.
\(^i\) Department of Ornithology, American Museum of Natural History, New York, NY, USA.
\(^j\) Instituto Nacional de Pesquisas da Amazônia, Coordenação de Biodiversidade, Av. André Araújo, 2936, Manaus, Amazonas, Brazil.

* Corresponding author: edsbio@gmail.com
\(^l\) Present address: Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland.
Diversification history in the *Dendrocincla fuliginosa* complex (Aves: Dendrocolaptidae): insights from broad geographic sampling.

**ABSTRACT**

*Dendrocincla* woodcreepers are ant-following birds widespread throughout tropical America. Species in the genus are widely distributed and show little phenotypic variation. Notwithstanding, several subspecies have been described, but the validity of some of these taxa and the boundaries among them have been discussed for decades. Recent genetic evidence based on limited sampling has pointed to the paraphyly of *D. fuliginosa*, showing that its subspecies constitute a complex that also includes *D. anabatina* and *D. turdina*. In this study we sequenced nuclear and mitochondrial markers for over two hundred individuals belonging to the *D. fuliginosa* complex to recover phylogenetic relationships, describe intraspecific genetic diversity and provide historical biogeographic scenarios of diversification. Our results corroborate the paraphyly of *D. fuliginosa*, with *D. turdina* and *D. anabatina* nested within its recognized subspecies. Recovered genetic lineages roughly match the distributions of described subspecies and congruence among phylogenetic structure, phenotypic diagnosis and distribution limits were used to discuss current systematics and taxonomy within the complex, with special attention to Northern South America. Our data suggest the origin of the complex in western Amazonia, associated with the establishment of upland forests in the area during the early Pliocene. Paleoclimatic cycles and river rearrangements during the Pleistocene could have, at different times, both facilitated dispersal across large Amazonian rivers and the Andes and isolated populations, likely playing an important role in differentiation of extant species. Previously described hybridization in the headwaters of the Tapajós river represents a secondary contact of non-sister lineages that cannot be used to test the role of the river as primary source of diversification. Based on comparisons of *D. fuliginosa* with closely related understory upland forest taxa, we suggest that differential habitat use could influence diversification processes in a historically changing landscape, and should be considered for proposing general mechanisms of diversification.

**1. INTRODUCTION**

The genus *Dendrocincla* is composed of mid-sized woodcreepers that inhabit humid forests and woodlands in Central and South America, all but one (*D. tyrannina*) in the lowlands (Ridgely & Tudor, 2009). Usually associated with swarms of army ants, species vary from facultative to obligate ant-followers, feeding mainly on arthropods flushed by the ants (Willis & Oniki, 1978) but also by other animals such as monkeys (Boinski & Scott; 1988). The taxa in the genus are mostly cryptic, and species limits and taxonomic affinities have been debated for over a century (Oberholser, 1904; Zimmer, 1934; Todd, 1948; Weir & Price, 2011; Marantz et al., 2018a). Currently, six species and 29 subspecies are recognized within *Dendrocincla* but the boundaries and relationships among lineages remain problematic (Weir & Price, 2011; Gill & Donsker 2016). A molecular study of the genus pointed to the paraphyly of *D. fuliginosa*, with its subspecies forming a species complex with *D. anabatina*, from Central America, and *D. turdina*, from the Atlantic Forest (Weir & Price, 2011). Despite the valuable insights in the systematics of the genus, only some of the described subspecies and a few individuals per taxon were sampled, with some of the absences including recurrently underrepresented (in Neotropical bird studies) taxa distributed in northern South America, precluding the understanding of the evolutionary history of the complex and its finer phylogeographic structure.

In the highly diverse Neotropics, broad sampling has repeatedly revealed that phylogenetic diversity in birds is larger than recognized by traditional taxonomy (d’Horta et al., 2013; Thom & Aleixo, 2015; Ferreira et al., 2017; Schultz et al., 2017). Accurate recognition of genetic diversity is crucial for tracing conservation priorities (Laikre et al., 2010), understanding phylogenetic relationships (Avendaño et al., 2017, Pérez-Emán et al., 2018) and making inferences about the evolutionary history of lineages and, consequently, the landscape they inhabit (Knowles, 2009).
The Neotropical region is currently composed of complex topographical elements that shape the landscape and constrain species distributions. Features like the Andes, large Amazonian rivers, the Panamanian Isthmus and open savannas are long known to isolate populations of forest species, particularly birds (Wallace, 1852; Snethlage, 1913; Chapman 1917). Accordingly, comparative analysis of distribution patterns of upland forest birds has revealed areas of endemism (Haffer 1974, Cracraft, 1985). The observation of such congruent patterns of distribution has shaped the search for general explanations and for the main mechanisms responsible for the origin of the large Neotropical species diversity (e.g. Haffer, 1969; Ribas et al., 2012; Smith et al., 2014).

*Dendrocincla*, particularly the *D. fuliginosa* complex, has played a major role in these discussions and has been used to test Neotropical biogeographic hypotheses regarding the effectiveness of landscape features as barriers to gene flow and their role in diversification of forest birds (e.g. Burney & Brumfield, 2009; Weir et al., 2009; Weir & Price, 2011; Milá et al., 2012; Batalha-Filho et al., 2013; Harvey et al., 2014; Smith et al., 2014; Weir et al., 2015; Naka & Brumfield, 2018). *Dendrocincla* species are particularly informative for that purpose because they occur in most forested regions from southern Mexico to northern Argentina, with several intraspecific lineages either distributed across those landscapes features or limited by them (Weir & Price, 2011). However, within Amazonia, the distribution limits of *D. fuliginosa* subspecies are not clear, with some lineages trespassing large rivers known to limit distributions of several understory bird taxa, such as the Solimões, Negro and Madeira (Marantz et al., 2018a), and examples of hybridization of two subspecies in the headwaters of the Tapajós and Xingu rivers (Weir et al., 2015).

Sampling as extensively as possible across geographical ranges and taxonomic units is essential in phylogeographic studies (Hawkins, 2006; Buckley, 2009). Consequently, here we analyze mitochondrial and nuclear sequences from over two hundred individuals of the *Dendrocincla fuliginosa* complex. This is the largest dataset focused on the systematics and phylogeography of the group to date, comprising samples broadly distributed across the species’ ranges and most described subspecies. Based on this thorough sampling, we aim to: (i) understand the spatial distribution of the genetic diversity and population structure within the *D. fuliginosa* complex and the relationships among phylogenetic lineages; (ii) contrast molecular data with available information on distribution and phenotypic variation to evaluate current taxonomy; and (iii) infer the biogeographical history of the group and its relationship to Neotropical landscape evolution.

2.MATERIAL AND METHODS

2.1. Taxon sampling and molecular data acquisition

Considering previous evidence for complex relationships among described species and subspecies in the genus (Weir & Price, 2011, Marantz et al., 2018a), we sampled 218 individuals in the *D. fuliginosa* complex (*D. fuliginosa*, *D. anabatina* and *D. turdina*). We focused on thoroughly sampling the distributions of all named taxa (including subspecies) in the complex (Table A1), with the exception of two subspecies of *D. anabatina* (*D. a. typhla* and *D. a. saturata*). One individual of *D. homochroa* was included as outgroup (Weir & Price, 2011). Two mitochondrial (NADH dehydrogenase subunit 2 [ND2] and cytochrome b [cytb]) and one nuclear (b-fibrinogen intron 7 [Fib7]) marker were sequenced. Collection localities were gathered from source museum collections (Table A1, Fig. 1). Additionally, 28 cytb sequences from Genbank were added to fill sampling gaps. Details on laboratory procedures and sampling are in Appendix A. Sequences were assembled, edited and aligned in Geneious (Kearse et al., 2012). Alleles from heterozygous individuals for the Fib7 nuclear gene region were inferred through the Bayesian algorithm implemented in PHASE v2.1.1 (Stephens et al., 2001), using SeqPHASE (Flot, 2010) to format the input files.

2.2. Phylogenetic analyses and genetic structure of separate loci
To access the phylogenetic relationships of the *D. fuliginosa* complex we analyzed the concatenated mtDNA and the Fib7 datasets separately. The mitochondrial dataset was analyzed under Bayesian Inference (BI) and Maximum Likelihood (ML) approaches in MrBayes 3.2.3 (Ronquist et al., 2012) and RAxML v.8.2 (Stamatakis, 2014), respectively, with data partitioned by gene. For the BI analyses, the best-fitting model of molecular evolution for each marker was estimated in jModelTest 2.1.7 (Darriba et al., 2012) using the Bayesian Information Criterion (BIC). For each dataset, three parallel runs were performed with four chains each for 50 million generations, parameters and trees being sampled every 50000 generations. The first 25% of trees were discarded as burnin. Convergence was visually accessed using Tracer v1.6 (Rambaut et al., 2015). For the ML analysis, a GTR + Gamma model was used, with 100 independent searches. Nodal support was accessed through 1000 non-parametric bootstrap pseudoreplicates. Fib7 sequences were only analyzed under the Bayesian Inference approach using MrBayes and the same protocols as described above. MrBayes analyses were conducted on the Cipres Gateway (Miller et al., 2010).

Haplotype networks were built using ‘haploNet’ function in the R package ‘pegas’ (Paradis, 2010; R Development Core Team, 2015) for all three loci sequenced focused on the mitochondrial lineages recovered in the phylogenetic analyses and on described subspecies. Mean pairwise p-distances within and among clades recovered by the phylogenetic analyses were calculated for each mtDNA marker in MEGA 7.0 (Kumar et al., 2016). Finally, well supported clades and the geographic distribution of our samples were contrasted with the distribution of described taxa in the complex (species and subspecies).

2.3. Molecular species delimitation, species tree and divergence dating

Multilocus coalescent approaches were applied to estimate genetic structure, phylogenetic relationships and divergence dating. First, well supported clades in the mtDNA analyses were tested as independent evolutionary units using a multilocus approach in BEAST2 package STACEY (Bouckaert et al., 2014; Jones, 2017). This analysis requires no guide tree, avoiding potential bias from *a priori* phylogenetic assumptions (Leaché & Fujita, 2010). Independent lineages supported by both mtDNA and STACEY analyses were used as independent evolutionary units for the downstream analyses. Two different approaches were applied to estimate species trees and divergence dating among recognized lineages in *BEAST* (Heled & Drummond, 2010). First, a multilocus analysis were ran including all three loci sequenced. Additionally, as a single nuclear marker with low phylogenetic information (Fig. B1) might be problematic for coalescent analyses (Xi et al., 2015), a species tree based only on mtDNA data was generated. An additional species tree with the mitochondrial markers was generated including *D. homochroa* as an outgroup to date the origin of the *D. fuliginosa* lineage. A Yule prior was used on all tree inferences, with substitution models following jModeltest’s results. For each marker, tests for clock-like evolutionary models were performed in MEGA and strict clocks were rejected for all three markers. Therefore, relaxed clocks were used, linking the mitochondrial markers using a substitution rate of 0.0105 (SD = 0.0034) substitution/site/million years (Weir and Schluter, 2008) with a normal distribution. Tests unlinking the mitochondrial clocks using two different rates (cytb mean rate = 0.0105 from Weir and Schluter (2008), ND2 mean rate = 0.0123 from Smith & Klicka (2013)), and a test using just one mean rate for mtDNA = 0.0114 from Weir & Price (2011) were also performed to contrast with results using the first approach. For each STACEY and *BEAST* analyses, three independent runs of 5x10⁷ generations were performed and results were combined using logcombiner (Rambaut & Drummond, 2014), removing the first 10% of samples from each run as burnin. *BEAST* analyses were conducted on the Cipres Gateway (Miller et al., 2010).

2.4. Biogeographic reconstruction

Ancestral area reconstructions were performed in the R package BioGeoBEARS (Matzke, 2014). Due to the uncertainties associated with weakly supported nodes, the reconstructions were made
for the mtDNA tree with \textit{D. f. atrirostris} as sister to the Atlantic Forest taxa; for the mtDNA tree with \textit{D. f. atrirostris} sister to the Northern South American taxa; and for the multilocus tree. For each analysis, twelve different models were tested: DEC, DIVALIKE, BAYAREALIKE and modifications of each to include founder-event speciation (“+J”) and to remove null range from space state, improving estimation of local extinctions (“**” models; Massana \textit{et al.}, 2015). Afterwards, as an additional approach to deal with uncertainties in the phylogenetic analyses, for each tree the most likely model identified was ran again including 100 random trees from its respective *BEAST analysis output.

Considering the distribution of the lineages (Fig. 1) and the Neotropical Areas of Endemism for birds (Cracraft, 1985), nine areas were used in the analysis: Central America (CAM), Chocó (CHO), Northern South America (NSA), Napo (NAP), Inambari (INA), Rondonia (RON), Guiana (GUI), Para (PAR) and Atlantic Forest (AFO). The maximum value of the “e” parameter, representing extinction events, was raised to 100 as preliminary runs hit the default maximum in some models. To improve the geographical information in the analysis, an adjacency matrix was created wherein areas that cannot be connected without trespassing another area cannot be joined as an ancestral area. This approach adds a geographical component to the analysis and avoids unrealistic ancestral distributions, as AFO/CAM, for example. The maximum range for ancestral states was set to three.

2.5. Demography of Amazon and Atlantic forest lineages

Due to their wide distribution throughout Amazon and Atlantic forests, \textit{Dendrocincla} lineages are potentially good models for testing hypotheses about forest retractions and expansions during Pleistocene climatic cycles (Haffer 1969). To examine whether current distribution patterns could be due to recent expansion of populations and if patterns of demographic change varied in different populations, we generated Extended Bayesian Skyline Plots (EBSP; Heled \& Drummond, 2008). For each lineage, three runs of $5 \times 10^7$ generations were performed, and the results were combined using a 10% burnin from each run. Substitution models and dating configurations were set as in the *BEAST analysis and analyses were conducted on the Cipres Gateway (Miller \textit{et al.}, 2010).

3. RESULTS

3.1. Phylogenetic analyses and genetic structure based on independent loci

BI and ML trees based on the mitochondrial markers recovered the same topology (Fig. 1). \textit{D. fuliginosa} as currently recognized is not monophyletic, as both \textit{D. anabatina} and \textit{D. turdina} are embedded within the \textit{D. fuliginosa} clade. The analyses identified a phylogenetic structure including several lineages, being eight of them with maximum support and two with lower support (Fig. 1), most of which corresponding to described subspecies based on their reported geographic ranges (Marantz \textit{et al.}, 2018a). However, three main instances of discordance between genetic evidence and current taxonomy were found. First, although several subspecies occur in northern Colombia and Venezuela, our genetic evidence suggests those subspecies could be included into two main taxa, \textit{D. f. meruloides} and \textit{D. f. ridgwayi} (Fig. 2). Second, two lineages were recognized within \textit{D. f. atristrostris} (atristrostris 1 and 2) in south-central Amazonia, (Figs. 1, B2). Third, a single lineage was found throughout the described distribution of \textit{D. f. neglecta} and \textit{D. f. phaeochroa}, including individuals from Northern and Western Amazonia. We were also unable to identify an independent lineage previously described as \textit{D. f. trumaii} (Sick, 1950), from the upper Xingu River, despite dense sampling in the area (Figs. 1, B2).

The Fib7 tree was poorly resolved and recovered only the Atlantic Forest lineages with high support (Fig. B1), but including one \textit{D. f. atristrostris} individual (P198). Concordantly, mitochondrial haplotype networks revealed unique haplotypes for most taxa as opposed to the Fib7 network which showed several haplotypes shared among most taxa (Fig. 3). However, atristrostris and turdina only shared Fib7 haplotypes between each other.
Mean p-distances among lineages ranged from 0.8% to 7.2% for ND2 and 0.6% to 6.3% for cytb. Distances among lineages were mostly larger than 5%, with the lowest value found between \textit{D. f. rufoolivacea} and \textit{D. f. fuliginosa} (0.8% for ND2 and 0.6% for cytb). Distances within lineages ranged from 0.1% to 0.8% and 0.1% to 0.4% for ND2 and cytb, respectively (Table B1).

### 3.2. Multilocus species delimitation, phylogenetics and divergence dating

In approximately 79% of the runs, multilocus clustering in STACEY distinguished all 10 lineages recovered by the mtDNA analysis. The remaining runs recovered different groupings of taxa, all of them with very low probabilities. Therefore, all 10 lineages were considered as independent units for the species trees, molecular dating, and demographic analyses (Table B2).

The multilocus species-tree recovered two main groups (Fig. 4). One of them comprised the Atlantic Forest taxa \textit{D. turdina} and \textit{D. f. taunayi}, potentially related to southwestern Amazonian taxa (\textit{D. f. atrirostris 1} and 2). The other main group, with high support, included Trans-Andean, Northern South American and Northern/Eastern Amazonian taxa. Within this clade, the low support of two internal nodes made the phylogenetic position of included taxa unclear, particularly Central America’s \textit{D. anabatina} and the Chocoan \textit{D. f. ridgwayi}. Eastern Amazonian \textit{D. f. fuliginosa} and \textit{D. f. rufoolivacea} appeared as sister lineages with high support. Similarly, \textit{D. f. phaeochroa} was recovered as being closely related to \textit{D. f. meruloides}. The multilocus tree showed some topological similarities to the mitochondrial tree but nodes had lower support and differed in the position of Trans Andean lineages (\textit{D. anabatina} and \textit{D. f. ridgwayi}) as well as \textit{D. f. fuliginosa} and \textit{D. f. rufoolivacea} (Fig. 4).

Multilocus and mitochondrial species trees dated the first split within the complex in the Pliocene, around 3.33 Ma (2.52-4.06 Ma 95%HPD) and 3.87 Ma (2.99-4.77 Ma 95%HPD), respectively (Fig. 4). Species trees unlinking mitochondrial markers or using the average mtDNA rate showed similar results; however, these analyses produced higher confidence intervals that encompassed the ones from linked clocks for most nodes. For example, mitochondrial trees using the two alternative rates recovered the same first split at 4.69 Ma (2.91-8.34 Ma 95%HPD) and 3.58 Ma (2.25-5.71 Ma 95%HPD). Consequently, we used the first results for biogeographic interpretations. The species tree including the outgroup estimated the origin of the \textit{D. fuliginosa} lineage during the Pliocene (4.27 Ma, 3.26-5.52 95%HPD), when it split from Andean \textit{D. homochroa} (Fig. B2).

### 3.3. Biogeographic reconstruction

The best-fitting model in all BioGeoBEARS analyses was DIVALIKE (Table B3). In a general way, all reconstructions tell a similar story. First of all, they recover the most likely origin of the complex in western Amazonia, from where it dispersed northwards and southwards throughout the Neotropical forests. To the south, a dispersal event from Western Amazonia in the Pliocene resulted in the colonization of the Atlantic forest. Within Amazonia, the first differentiation resulted in lineages at north and south of the region. However, the order that these two events took place is not clear with our data. In the north of South America, two dispersals events happened across the Andes, the first one originating Central American \textit{D. anabatina} and the second originating \textit{D. f. ridgwayi}, in multilocus tree, or the ancestor of \textit{D. f. ridgwayi}/\textit{D. f. meruloides}, in the mtDNA trees. Therefore, the origin of the NSA lineage \textit{D. f. meruloides} could not be recovered and could be related either to northern Amazonia or to the west of the Andes. Finally, the colonization of southeastern Amazonia constitutes a recent event, resulting from dispersal from the north across the lower Amazon river.

### 3.4. Demography of Amazon and Atlantic forest lineages

Among Amazonian lineages, there was a strong signal of demographic expansion for both northeastern (\textit{D. f. fuliginosa}) and southwestern (\textit{D. f. atrirostris 1}) lineages around 40-20 kyr (Fig. 6). For those lineages the 95% HPD estimate of population change over time did not include zero,
statistically rejecting stability. When the 95% HPD includes zero, stability cannot be rejected, but means in the population size change parameter higher than one (as shown in most lineages) might also indicate slight population expansions. Means lower than one in *D. f. rufoolivacea* and *D. f. taunayi* suggest population size stability.

4. DISCUSSION

4.1. Genetic Diversity and Systematics

Geographical variation in phenotypic characters of *D. fuliginosa* has resulted in the recognition of 12 subspecies. These subspecies have been included in three groups: *meruloides*-group (*ridgwayi, lafresnayeii, barinensis, meruloides, deltana, phaeochroa, neglecta*), *fuliginosa*-group (*fuliginosa, rufoolivacea, atrirostris, trumaii*) and *taunayi* (Todd, 1948; Marantz *et al.*, 2018a). Our study partially supports this classification. The *meruloides* group was well supported in our mtDNA tree, but some of the taxa within it were not recovered as monophyletic. None of our results supported the *fuliginosa* group as suggested above, as *atrirostris* appeared mostly as sister to the Atlantic forest lineages, but with low support. Finally, *D. f. taunayi* was sister to *D. turdina*. The paraphyly of *D. fuliginosa* based on this last result and the inclusion of *D. anabatina* within this complex is in agreement with previous studies on the genus (Weir & Price, 2011).

A close relationship of *D. f. taunayi* to *D. turdina* has been suggested since its description, when both taxa were recognized as subspecies of *D. fuliginosa* (Pinto, 1947; Zimmer & Mayr, 1943; Peters, 1951). With the elevation of *D. turdina* to species, the taxonomic placement of *taunayi* remained uncertain (Piacentini *et al.*, 2015, Gill & Donsker, 2017). Our results, in agreement with Weir & Price (2011), found large support for the sister taxon relationship of *D. f. taunayi* and *D. turdina*; similarly, we found that the clade formed by these taxa diverged from all other lineages around 1.3 Ma, having unique haplotypes, even in slow-evolving Fib7 (Fig. 3). Hence, our data, coupled with known morphological and vocal differences, support the treatment of *D. taunayi* as a species apart from *D. fuliginosa* (Pinto, 1939; Zimmer & Mayr, 1943; Marantz *et al.*, 2018a).

A monophyletic *fuliginosa* group of subspecies was not supported by our results. The only clear phylogenetic relationship within this group was between *D. f. fuliginosa* and *D. f. rufoolivacea*, which formed a well-supported eastern Amazonian clade distributed north and south of the Amazon river, respectively. Genetic divergence between them was very low (<1% for both mtDNA markers). Moreover, one individual (FMNH391299) collected north of the Amazon river grouped with *D. f. rufoolivacea*, indicating the possibility of ongoing introgression across the lower Amazonas, although this pattern can also be due to ancestral polymorphism. Phenotypically, those taxa are also very similar and authors diverge regarding their morphological distinction (Cory & Hellmayr, 1925; Zimmer, 1934; Todd, 1948). A more detailed morphological and genetic study is required to investigate their diagnosis and the occurrence of gene flow among those lineages. The phylogenetic placement of the other subspecies in this group, *atrirostris*, could not be reliably estimated (comparing MrBayes tree vs. both species trees). Species trees (both multilocus and mtDNA) placed *D. f. atrirostris* as potential sister to Atlantic Forest lineages contradicting its position within the *fuliginosa* group and a previous phylogeny of the genus (Weir & Price, 2011), characterized also by low supported nodes associated to *atrirostris* phylogenetic position. Similarities in vocalizations and plumage patterns between *atrirostris* and *turdina* have already been described and provide additional evidence for their potential close relationship (Oberholser, 1904; Cory and Hellmayr, 1925; Willis, 1992, Marantz *et al.*, 2018a), which should be addressed more thoroughly with genomic data. Confirmation of this phylogenetic pattern would provide with another example of southwestern Amazonian bird taxa that are closely related to taxa from southern Atlantic Forest (Batalha-Filho *et al.*, 2013). Our results also recovered, within *D. f. atrirostris*, two monophyletic and seemingly allopatric lineages, which were previously undescribed. Interestingly, Todd (1948) mentioned that birds from Bolivia and the west bank of the Tapajós river might represent different taxa but did not give details about this; our data show that birds from each of
these areas belong, respectively, to the *D. f. atrirostris* 2 and *D. f. atrirostris* 1 lineages. We did not find evidence for the presence of *D. fuliginosa trumaii*, a taxon described by Sick (1950) from the confluence of Culune and Xingu rivers in Mato Grosso state in Brazil. Further references to this subspecies seem to be based solely on its supposed range (Novaes & Lima, 1996; Pacheco & Olmos, 2005; Somenzari et al., 2011). Here, we included specimens (Table A1: FLJA11, MTMA 3, 5, 11, 12, 14, 15, 18 and 59) collected very close (ca 85 km) to the type locality (Fig. B3) and showed that they belong to the *D. f. atrirostris* 1 clade, suggesting that *D. f. trumaii* is not a distinct taxon.

The most geographically differentiated group within *D. fuliginosa*, the *meruloides* group, with its seven recognized subspecies, proved to be monophyletic but the status of those subspecies need to be evaluated more thoroughly. In Northwestern Amazonia, our results identified a single lineage in the described distribution of *D. f. neglecta* and *D. f. phaeochroa*, with no clear separation of haplotypes assigned to each subspecies. Slight morphological variation has been attributed to these subspecies, and their distinctiveness has long been debated (Zimmer, 1934; Todd, 1948; Peters, 1951; Marantz et al., 2018a). The lack of genetic differentiation might support merging *neglecta* (Todd, 1948) with *phaeochroa* (Berlepsch & Hartert, 1902). The other five subspecies represent the large phenotypic diversity described for *D. fuliginosa* in Northern Colombia and Venezuela (Peters, 1951). Here, we present the first extensive genetic sampling of the clade in the region and our results point to a new interesting perspective. In our analyses, samples from the described distribution of subspecies *meruloides*, *deltana*, *barinensis*, *lafresnayei* and *ridgwayi* were grouped into two main lineages treated here as *D. f. meruloides* and *D. f. ridgwayi*. In this context, the distribution of *D. f. lafresnayei* deserves some discussion. This taxon is supposed to occur in N & E Colombia (Santa Marta region, Magdalena and Cauca Valleys) and adjacent NW Venezuela (NW Zulia, Merida, Táchira) (Marantz et al., 2018a). However, and somehow confusingly, *lafresnayei* distribution in Venezuela includes the northwestern area of Zulia (in Perijá) and western Táchira (San Juan de Colón, within the eastern Maracaibo Lake basin), suggesting this taxon potentially co-occurs with *meruloides*, which is supposed to be distributed in the southern (and likely eastern) shore of Maracaibo Lake (Cory & Hellmayr, 1925; Peters, 1951). Our sampling shows that individuals from Magdalena and Cauca Valleys, as well as from the Colombian side of northern Perijá, grouped within *D. f. ridgwayi*, while samples from the southeastern region of the Maracaibo Lake basin were nested within *D. f. meruloides*, suggesting a polyphyletic *D. f. lafresnayei*. This result highlights the need for more work in the region to better understand the geographical distribution of these taxa.

4.2. Species delimitation and taxonomy

There is a vast discussion in the literature about how to delimit species (de Queiroz, 2007). In birds, both morphological and genetic evidence suggest that species diversity is underestimated by current taxonomy by at least a factor of two (Barrowclough et al., 2016). Over the last years the development of probabilistic models and coalescent-based species delimitation methods have provided statistical frameworks to test alternative taxonomic hypotheses (Fujita et al., 2012), but these should also be taken with care (Sukumaran & Knowles, 2017).

In the Neotropics, many widespread bird species are composed of several diagnosable allopatric lineages divided by major geographical features (Cracraft, 1985). Even in occasional contact zones, as in rivers’ headwaters, hybridization does not seem to be capable of preventing genetic and geographic differentiation (Weir et al., 2015; Pulido-Santacruz et al., 2018). Here we recognize genetic lineages within *D. fuliginosa* by different molecular analyses, including mtDNA phylogenies and multispecies coalescent delimitation. Despite some of these lineages trespass putative barriers as large Amazonian rivers and could be in contact at interfluvial regions, they represent reciprocally monophyletic taxa that generally correspond to described subspecies that have been known for over a century (Marantz et al., 2018a). Such agreement among genetic structure, geographic distribution and previous knowledge on phenotypic differentiation (Cory & Hellmayr, 1925; Zimmer 1934; Todd, 1948; Peters, 1951; Marantz
et al., 2018a) suggest many of the subspecies within the *D. fuliginosa* complex might deserve species status. We propose that thorough studies on species limits need to be done in this complex focusing on the following potential species: *D. atrirostris*, *D. fuliginosa* (including *D. f. rufoolivacea*) (see below for the discussion about hybridization between these two taxa), *D. meruloides* (including *D. f. barinensis*, *D. f. deltana*, and *D. f. lafresnayei* east of Serranía del Perijá), *D. phaeochroa* (including *D. f. neglecta*), *D. ridgwayi* (including *D. f. lafresnayei* west of Serranía del Perijá and acknowledging that if future studies confirm this hypothesis the seemingly older name *D. lafresnayei* (Peters, 1951) would probably be more adequate for such taxon), *D. taunayi* and *D. turdina*. The apparently cryptic nature of phenotypic variation in both *atrirostris* lineages suggests keeping them as a single taxa, at least until more detailed analyses could be carried out. Similarly, the status of *fuliginosa/rufoolivacea* and *phaeochroa/neglecta* must be reviewed more carefully and a more detailed phenotypic and genomic comparison could clarify the taxonomic status of those taxa.

4.3. Diversification and biogeography

The overall biogeographic history of the complex indicates an ancestral distribution in Western Amazonia, with posterior expansion to forests throughout the Neotropics. This scenario matches biogeographic histories proposed for other Neotropical understory forest birds (Thom & Aleixo, 2015; Ferreira et al., 2016; Schultz et al., 2017; Ribas et al., 2018) and reinforces the importance of western Amazonia as a stable area and a potential source of Amazonian diversity (Silva et al. 2019). During the early Pliocene, South America was experiencing landscape changes due to the rearrangement of the drainage network and establishment of extensive non-flooded upland forests in the west, with consequences for the distribution and connectivity of the biota (Hoorn et al., 2010, Latrubesse et al., 2010, Nogueira et al., 2013, Pupim et al. 2019). The establishment of upland *terra firme* forests in western Amazonia, previously occupied by wetlands, resulted in connections among upland forest populations from the Andean foothills, Brazilian and Guianan shields (Bates 2001; Pupim et al. 2019; Silva et al. 2019). The sister relationship to Trans-Andean *D. homochroa* (Weir & Price, 2011) supports this scenario of origin of the *D. fuliginosa* complex at the Andean foothills.

Although the low support for the node representing the position of *D. f. atrirostris* veils the order of events associated to the differentiation of lineages both in Amazonia and Atlantic Forest, all of our reconstructions point to a dispersal event from southwestern Amazonia to the Atlantic forest (originating the ancestor of *D. turdina* and *D. f. taunayi*). The possible connection between southwestern Amazonia and southern Atlantic Forest that could have facilitated that dispersal is well documented for several taxa and is suggested as the main historical bridge between the two biomes (Por, 1992). This connection is postulated to have occurred several times since the final stages of Andean uplift and is associated with forest expansion and retractions caused by climatic fluctuations (Ledo & Colli, 2017). Despite the suggestion of a Miocene divergence of birds from SW Amazon and SE Atlantic Forests (Batalha-Filho et al., 2013), our data suggest disconnection among lineages from those regions at a more recent time, during the Pliocene, in agreement with other recent studies (d’Horta et al., 2013; Lavinia et al., 2015; Ferreira et al., 2016). Genomic data could be useful to assess the support for this taxon relationship (*atrirostris* sister to a clade formed by *taunayi* and *turdina*) and, consequently, strongly contribute to the history of connectivity among these areas.

In the Atlantic Forest, our divergence estimates suggest that the southern and northern *Dendrocincla* lineages (*turdina* and *taunayi*) diverged more than 1 Ma. The demographic histories of those lineages are contrasting, with indications of constant population size in *D. f. taunayi* and expansion in *D. turdina*. However, the signal of population expansion is subtle in *D. turdina*, perhaps due to limited sampling, and a more extensive sampling could confirm this pattern. Nonetheless, this finding is in agreement with evidences of climatic stability in northern and instability in southern Atlantic Forest (Carnaval et al., 2009).
Within the Madeira-Tapajós interfluvium, lineages of *D. fuliginosa* show an intricate pattern of distributions, with *D. f. atrirostris* seemingly restricted to the northern portion and *D. f. atrirostris* I widely distributed in the southern portion and across the headwaters of those rivers. Moreover, two individuals (MPDS748, MPDS616) belonging to the *D. phaeochroa* clade occur on the right bank of the Madeira river, between the rivers Ji-Paraná and Aripuanã (Fig. B3). This region is known for having complex distributional patterns for several bird taxa (Fernandes, 2013). Like in *Dendrocincla*, several groups have young lineages aging no more than 1.5 Ma, isolated within the Madeira/Tapajós interfluve (microendemisms, Fernandes *et al.*, 2012, 2013, 2014; Thom & Aleixo, 2015; Ferreira *et al.*, 2017; Ribas *et al.*, 2018). Geological evidence suggests that changes in the courses of the Madeira river and its tributaries occurred during the Pleistocene (Latrubesse, 2002; Hayakawa & Rossetti, 2015), and may have been responsible for these intricate distribution patterns.

The differentiation between *D. f. fuliginosa* and *D. f. rufoolivacea*, currently separated by the lower Amazon River, dates to a few hundred thousand years, suggesting an unusual pattern of recent dispersal across this barrier. The fact that we have found one *rufoolivacea* individual north of the Amazon River, its occurrence on Marajó Island, and its use of flooded forests (várzea) and tidal-swamp forests in Eastern Amazonia (Oniki, 1972), suggest dispersal through the mouth of Amazon and possible ongoing gene flow across this partially permeable barrier. Considering the shallow divergence among these lineages, shared haplotypes could also be caused by incomplete lineage sorting.

Several *Dendrocincla* taxa have their distributions restricted to either side of the Andes, making them interesting models to study divergence across this range. As suggested by Weir & Price (2011), our results show asynchronous diversification across the Andes, pointing to independent events, with repeated colonization of Trans-Andean regions from western Amazonia. While the split between *D. homochroa* and the *D. fuliginosa* complex at the early Pliocene could be related to a vicariant event caused by the final uplift of the Andes (Weir & Price, 2011), the following events of cross-Andean diversification occurred at a time when the Andes had reached altitudes much similar to current ones (Horton *et al.*, 2010; Baker *et al.*, 2014). However, Pleistocene climatic fluctuations resulted in vegetation changes in the Andes and adjacent areas that affected the connectivity or lack of it of both lowland and highland forests (Flantua & Hooghiemstra, 2018). Low passes through the Andes and through Caribbean lowlands, currently unsuitable to dispersal of lowland forest birds, may have acted as corridors under different climate conditions and facilitated the dispersal of *D. fuliginosa* ancestors across the Andes (Haffer 1967; Cadena *et al.*, 2016). The first cross Andean dispersal resulted in the colonization of Central America by the ancestor of *D. anabatina*. All estimated species trees suggested a dispersal mean age younger than 2.7 Ma, at a time when the land bridge between South and Central America was already complete, according to both early and late hypotheses of Panamanian Isthmus formation (Bacon *et al.*, 2015). Afterwards, because of disagreement in the position of some nodes in the northern clade, our reconstructions suggest two different hypotheses for the diversification of lineages in Northern South America (NSA) and Chocó. The mitochondrial tree points to two consecutive diversification events across the Andes. First, ca 1.26 Ma (1.32-3.58 Ma 95%HPD), separating Western Amazonian *D. f. phaeochroa* from the ancestor of *D. f. ridgwayi/meruloides* and subsequently, almost one million years ago, separating the latter taxa at east and west of the Eastern Colombian Cordillera. On the other hand, the multilocus tree suggests a single cross-Andean diversification event, but resulting from a seemingly less likely connection of Chocó and eastern Amazonia. In this scenario, *D. f. meruloides* diverged from *D. f. phaeochroa* ca. 0.25 Ma (0.14-0.45 Ma 95% HPD).

### 4.4. Importance of sampling in widespread species

The first studies in molecular phylogenies and biogeography of Neotropical birds were usually based on one or a few individuals per species or subspecies. These studies were essential to reveal general patterns of relationship and to explore ideas about the evolutionary history of the region.
However, here we show that in a widely distributed clade such as the *D. fuliginosa* complex, thoroughly sampling of individuals across the whole distribution of the taxa provides new relevant discussions in a finer scale.

Regarding taxonomy and lineage distributions we show that in northern Venezuela, in spite of several described subspecies, they do not match genetic lineages suggesting that populations ranging from the eastern side of Maracaibo Lake to the Orinoco delta might not represent multiple subspecies but geographical variation associated to *D. f. meruloides*. However, such lack of phylogenetic structure could also be the result of a process of rapid and recent differentiation of these populations. To the west of that range, what has been called of *D. f. lafresnayei* is actually co-specific with *D. f. ridgwayi*. Also, we show that *phaeochroa* and *neglecta* cannot be distinguished genetically and should be treated as a single taxon. Finally, we show that, according to our data, the taxa *trumaii* is not valid and individuals sequenced from the type locality are actually *atrirostris*. These findings provide a strong rationale for not only studying potential distributional limits of each of these taxa but also the need to evaluate the phenotypic variation of these taxa to discern the merits of dividing such variation in the currently known subspecies or explaining it as a result of evolutionary processes such as rapid recent differentiation.

Our results also give a new interesting perspective on the diversification of *D. fuliginosa* lineages in southern Amazonia. In the headwaters of the Tapajós and Teles-Pires rivers, Weir et al. (2015) found hybrids between several bird lineages, including *D. f. atrirostris* and *D. f. rufoolivacea*. Our phylogenetic hypothesis indicates that *atritrostris* and *rufoolivacea* are not each other closest relatives within the *D. fuliginosa* complex, with *atrirostris* probably more related to the Atlantic forest taxa, and *rufoolivacea* having reached SE Amazonia recently, potentially coming from the north. Consequently, their current parapatric distribution should be secondary as a result of changes in their distributions, and not a result of vicariance events related to the evolution of the Tapajós drainage itself or to climatic fluctuations, as suggested by Weir et al. (2015). In fact, genetic divergence between both taxa cannot be used to provide historical scenarios of vicariance for these taxa as suggested by these authors. An additional information from our study is that, while most Amazonian lineages in the complex seem to have kept relatively stable demographic sizes, the *atritrostris* lineage shows a strong signal of population expansion around 30-20 kyr, which could explain the trespassing of putative barriers and the interfluvial co-occurrence with both *D. f. rufoolivacea* and *D. f. phaeochroa*. This demographic expansion could be a result of recent forest expansion at the southern edge of Amazonia, agreeing with paleoclimatic reconstructions (Wang et al., 2017) and ultimately conforming to the basic prediction of the refuge theory (Haffer 1969).

4.5. The role of ecology in the diversification of Neotropical forest birds

Recent studies have shown that pairs of taxa currently separated by the same barriers may have temporally discordant divergence times (Smith et al., 2014; Naka and Brumfield 2018). A current barrier such as a river or a mountain could have acted as a primary source of diversification at its formation, but could also serve as a secondary barrier, avoiding contact of populations that differentiated elsewhere (Naka & Brumfield, 2018). Additionally, ecological traits like habitat affinities significantly influence connectivity and diversification in Neotropical bird populations (Burney & Brumfield, 2009; Harvey et al., 2017).

Throughout the Amazon Basin, *D. fuliginosa* occurs in sympatry with *D. merula*, a taxon restricted to the region. However, while *D. merula* favors the understory of mature forests, *D. fuliginosa* is also common in second growth and forest edges (Cintra et al., 2006; Rodrigues et al., 2016; Marantz et al., 2018a; Marantz et al., 2018b). Accordingly, both taxa have contrasting biogeographical patterns within Amazonia. Some Amazonian lineages of *D. fuliginosa* have experienced recent population expansion and distributional shifts, with several extant taxa aging a few hundred thousand years and distributions trespassing rivers known as important biogeographic barriers.
to several other bird taxa. In contrast, despite the lack of a detailed phylogeography, *D. merula* seems to have an older history within the Amazon basin and the subspecies distribution is more clearly delimited by large rivers (Weir & Price, 2011; Marantz et al., 2018b). A comparable pattern is found in *Automolus* foliage-gleaners, with taxa less restricted to the interior of upland forest (*A. ochrolaemus* and *A. subulatus*) widespread in Neotropical forests and showing younger diversification across the Amazon Basin than Amazonian endemic *A. infuscatus* (Schultz et al., 2017). As in *Automolus*, the history told by *Dendrocincla* suggests that taxa more specialized in the interior of mature Amazonian forests have occupied the forest for longer and their movements are more restricted by discontinuities in forest cover, while species adapted to edges and second growth represent a more recent occupation. The use of forest edges and second growth itself may be a response to the colonization of a region already occupied by a congener specialized in upland forests (Schultz et al., 2017). Therefore, lineages more ecologically restricted to the understory of mature upland forest may have been more affected by paleoenvironmental changes than younger and more generalist ones. These ecological differences would result in discordant responses to the same historical events and thus to current barriers (Naka & Brumfield 2018).

Albeit simple models cannot explain the diversification history of the complex Neotropical biota, considering how species occupy the landscape can give important clues to clarify incongruences in biogeographic histories across the landscape. Comparative phylogeographies of sympatric, closely-related taxa, will help understand the coexistence of congeners in Amazonia and how different ecological affinities relate to distinct biogeographic histories and seemingly idiosyncratic responses to the same barriers.

**FUNDING**

This study was co-funded by FAPESP (2009/12989-1, BIOTA 2013/50297-0), NSF (DOB 1343578) and NASA. This work was developed in the Research Center on Biodiversity and Computing (BioComp) of the Universidade de São Paulo (USP), supported by the USP Provost’s Office for Research. EDS is supported by FAPEAM (Doctoral scholarship). AA, CCR, and CYM are supported by CNPq research productivity fellowships (respectively 310880/2012-2, 458311/2013-8, 303713/2015-1).

**ACKNOWLEDGEMENTS**

We thank the curator and curatorial assistants of the American Museum of Natural History (AMNH), New York, USA; Academy of Natural Sciences of Drexel University, Philadelphia, USA (ANSP); Colección Ornitológica Phelps (COP), Venezuela; Field Museum of Natural History, Chicago, USA (FMNH); Instituto Alexander von Humbold (IAvH), Colombia; Instituto de Ciencias Naturales, Universidad Nacional de Colombia (ICN), Colombia; Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil (INPA); Laboratório de Genética e Evolução Molecular de Aves, Universidade de São Paulo, Brazil (LGEMA); Lousiana State University Museum of Natural Science (LSUMZ), Baton Rouge, USA; Marjorie Barrick Museum (MBM), USA, and Museu Paraense Emílio Goeldi, Belém, Brazil (MPEG) for providing tissue loans.

**REFERENCES**


Haffer, J. (1974). Avian speciation in tropical South America, with a systematic survey of the toucans (Ramphastidae) and jacamars (Galbulidae).


Highlights

- *D. fuliginosa* is paraphyletic in relation to *D. anabatina* and *D. turdina*
- Genetic diversity significantly matches intraspecific phenotypic diversity
- Subspecies from Northern South America consist of two main taxa
- The origin of the complex took place in Western Amazonia in the Pliocene
- Habitat affinities shaped a different evolutionary history from an amazonian congener