Gregory Thom

Filogeografia comparada de aves de várzea baseada em sequências de elementos ultra conservados: reconstruindo padrões biogeográficos da Amazônia

Comparative phylogeography of floodplain specialist birds based on sequences of ultra conserved elements: inferring Amazonian biogeographic patterns

São Paulo
2018

Gregory Thom

Filogeografia comparada de aves de várzea baseada em sequências de elementos ultra conservados: reconstruindo padrões biogeográficos da Amazônia

Comparative phylogeography of floodplain specialist birds based on sequences of ultra conserved elements: inferring Amazonian biogeographic patterns

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Doutor em Ciências Biológicas, na Área de Biologia – Genética, Departamento de Genética e Biologia Evolutiva

Orientador(a): Cristina Yumi Miyaki

São Paulo

2018

	Thom, Gregory Filogeografia comparada de aves de várzea baseada em sequências de elementos ultra conservados: reconstruindo padrões biogeográficos da Amazônia 232 páginas Tese (Doutorado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Genética e Biologia Evolutiva. 1. Diversificação 2. Biodiversidade 3. Demografia histórica 4. Região Neotropical I. Universidade de São Paulo. Instituto de Biociências. Departamento de Genética e Biologia Evolutiva

Comissão Julgadora:

_____ _____

Prof.(a) Dr.(a) Prof.(a) Dr.(a)

_____ _____

Prof.(a) Dr.(a) Prof.(a) Dr.(a)

_____

Profa. Dra. Cristina Yumi Miyaki

Orientadora

“The Universe is under no obligation
to make sense to you”
Neil deGrasse Tyson

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Prof. Cristina Yumi Miyaki for her guidance and continuous support of my Ph.D. study, allowing me to grow as a research scientist.

My sincere thanks also to Dr. Michael Hickerson, who provided me the opportunity to join the Hickerlab team as a visiting student at the City Colledge of New York. This was an outstanding experience, for both my scientific and personal growth. Additionally, I would like to thank all members of the Hickerlab and Carnaval lab for the continuous support in the laboratory and for making the internship so enjoyable. Special thanks to Isaac Overcast, Alexander Xue, Ivan Prates, Vanessa Bonatti, Pedro Taucce, Laura Bertolla and Roberta Damasceno not only for the scientific help but also for their invaluable friendship.

I wish to express my sincerest gratitude and warm appreciation to all my fellow labmates who came and went over the years. In particular, I am grateful to Claydson Assis, Andressa Nuss, Manuelita Sotelo Munoz, Rafaella Monteiro for all the great moments in and outside the lab.

I also would like to express my gratitude to Dr. Alexandre Aleixo for his endless motivation, immense knowledge and advices, both on professional and life matters.

I would especially like to thank Dr. Fabio Raposo do Amaral for the friendship and for sharing his knowledge that greatly improved this study.

Last but not least, I would like to express my gratitude to my beloved wife that in every moment stood by me, sharing her motivation, mutual love for science and by showing day after day that is possible to develop two independent academic careers without being away from each other.

This work would not have been possible without the financial support of CAPES, FAPESP (FAPESP grants 2014/00113-2, 2015/12551-7, BIOTA 2012/50260-6 and 2013/50297-0), NSF (DOB 1343578), NASA, CNPq (310593/2009-3, 574008/2008-0, 563236/2010-8, 471342/ 2011-4) and BioComp.

Contents

Resumo

Nesta Tese foram inferidos os processos históricos que podem ter influenciado a evolução da comunidade de Aves restritas às planícies alagadas amazônicas baseado na análise de seus padrões de diversificação associados a dados geológicos e climáticos. Foram selecionados três complexos de espécies amplamente distribuídos pela bacia Amazônica e intimamente associados às florestas de beira de rio, *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus* e *Myrmotherula assimilis*. Para acessar a diversidade genética e realizar análises filogenéticas e demográficas das espécies foi realizada a captura e sequenciamento de ~2.300 Elementos Ultra Conservados. A Tese foi dividida em três capítulos que abordam aspectos distintos sobre a diversificação desses organismos. No Capítulo 1, *M. lugubris* foi estudado com o intuito de explorar os efeitos de processos microevolutivos, em particular fluxo gênico e sorteamento incompleto de linhagens, na reconstrução filogenética das populações deste complexo. Esse estudo demonstrou a problemática em se estimar árvores de espécies em cenários recentes de diversificação utilizando métodos que não acomodam fluxo gênico como um parâmetro, e explorou o histórico evolutivo desse complexo de espécies, mostrando a presença de fluxo gênico entre populações não irmãs e a presença de uma zona de hibridação com um potencial cenário de heterose. No Capítulo 2 foi realizado um estudo filogeográfico comparado dos três complexos de espécies que sugeriu que ciclos climáticos ocorridos ao longo do Pleistoceno, que parecem ter alterado o padrão de sedimentação e formação de florestas de beira de rio, tiveram importante papel na diversificação deste conjunto de organismos, promovendo períodos de alopatria seguidos de contato secundário. No terceiro Capítulo da Tese exploramos como a atual diversidade genética de populações restritas ao Rio Solimões está distribuída no espaço e se alterações demográficas históricas poderiam estar relacionadas à expansão nas distribuições geográficas. Nesse estudo encontramos sinal de expansão geográfica para *T. cryptoleucus*, mas não para *M. lugubris* e *M. assimilis*, sugerindo que em uma escala intrapopulacional variáveis ecológicas intrínsecas aos táxons estudados devem promover padrões distintos de diversidade genética. Ainda nesse estudo foi encontrado que a atual diversidade genética desses complexos espécies está distribuída de forma heterogênea na paisagem. Os dados apresentados nesta Tese permitiram testar de forma inédita hipóteses de diversificação para a comunidade de Aves que ocorre ao longo de toda a bacia Amazônica, sugerindo que organismos de ambientes pouco explorados por estudos filogeográficos podem revelar novas facetas sobre a diversificação da elevada diversidade Amazônica.

Abstract

This Ph.D. Dissertation inferred the historical processes that seem to have built the avian community assemblage restricted to the Amazonian floodplains based on their patterns of diversification and geological and climatic data. We analyzed three species complexes widely distributed over the Amazon Basin and specialists of river edge forests, *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis*. In order to access their genetic diversity and perform phylogenetic and demographic analyses, we captured and sequenced ~2,300 Ultra Conserved Elements. This Dissertation was subdivided into three chapters that discuss distinct diversification aspects of these taxa. In Chapter 1, we explored the effects of microevolutionary processes in *M. lugubris*, especially gene flow and incomplete lineage sorting in phylogenetic reconstructions of its populations. We demonstrated the potential bias of estimating species tree without accommodating gene flow in recent scenarios of divergence. Additionally, we explored the evolutionary history of this lineage, supporting the presence of gene flow between nonsister populations and a hybrid zone with potential heterosis. In Chapter 2 we performed a comparative phylogeographic study including all three species complexes. The results indicated that climatic oscillations during the Quaternary that altered the pattern of sedimentation and formation of river edge forests seem to have promoted cyclical periods of allopatry and secondary contact. In the third chapter we explored how the genetic diversity of populations restricted to the Solimões river is currently distributed in space and if historical demographic alterations could be related to geographic range expansions. A signal of range expansion was detected only for *T. cryptoleucus* but not for *M. lugubris* and *M. assimilis,* suggesting that at the intra-populational scale species-specific ecological variables may promote distinct patterns of genetic diversity. However, despite the absence of a shared pattern of range expansion the genetic diversity of each of the three taxa is heterogeneously distributed in the landscape. The data presented in this Dissertation allowed an unprecedented test of diversification hypotheses for the bird community that occurs throughout the Amazonian floodplains, suggesting that organisms from environments poorly explored by phylogeographic studies may reveal new facets about the diversification of the Amazonian diversity.

Introduction

Wild populations are subjected to changes in habitat distribution such as expansions, connections among previously isolated areas, contractions and fragmentations. From an evolutionary perspective, these fluctuations can affect patterns of genetic diversity, population structure and gene flow (Avise 2000). Diversification patterns are influenced by historical events such as climatic and geomorphological changes, current landscape conditions including anthropogenic alterations and/or ecological conditions and intrinsic characteristics of the organisms under study. Hence, understanding the mechanisms of speciation and diversification can assist strategies for biodiversity conservation and to predict the effects of future climatic alteration on natural communities (Moritz and Faith 1998; Moritz 2002; Taberlet and Cheddadi 2002; Williams et al. 2007; Brown et al. 2016; Prates et al. 2016).

Phylogeography aims to understand the principles and processes behind the current geographic distribution and demography of intra-specific lineages or closely related species based on the spatial distribution of genealogies (Avise et al. 1987; Avise 2000). Thus, it can help to reveal mechanisms of speciation and diversification. Historical alteration in climate and landscape may affect genealogical patterns of an entire biological community, and if a process is significant enough, different organisms should present concordant genealogies (Avise et al. 1987; Avise 2000; Soltis et al. 2006; Hickerson et al. 2010). However, variation in the capacity of individuals to respond to changing conditions may increase variation among diversification processes (Smith et al. 2014). Avise (2000) described four levels of genealogical concordance in phylogeography: 1) concordance among sites within a locus indicating genetic structure; 2) concordance among multiple loci providing genomic evidence for a specific pattern; 3) concordance among multiple co-distributed lineages that indicates a similar pattern in the geographic distribution of the genetic structure providing evidence of a shared diversification process; 4) correlation between a genetic pattern shared by various taxa and previously described biogeographic information supporting historical factors that shaped the distribution of genetic lineages.

Initially, the molecular marker mostly used in phylogeographic studies of animals was the mitochondrial DNA (mtDNA) given its intrinsic characteristics such as the apparent lack of recombination, putative neutrality and haploid matrilineal inheritance with smaller effective population size when compared to autosomal loci leading to more pronounced effects of genetic drift and lineage sorting (Avise 2000). However, the use of a single locus to infer concordance among distinct levels of phylogeography can be compromised as explained as follows (Edwards & Beerli 2002; Edwards 2009; Brito & Edwards 2009). The theoretical basis of phylogeography is anchored on population genetics and the coalescent theory, which mathematically describes the probability of a given group of alleles to coalesce into a common ancestral over the generations backwards in time (Felsenstein 2008). With the continuous development of the coalescent theory, it became possible, based only on sampled alleles, to model virtually any complex diversification scenario and estimate demographic parameters including current and historical population sizes, divergence times and migration rates (Wakeley 2009). However, single locus mtDNA data can dramatically bias and limit demographic parameter estimation using the coalescent (Edwards 2009). For example, analyses performed with a single locus or with multiple concatenated loci can produce statistically supported topologies that are not in agreement with the evolutionary history of a species, especially in recent scenarios of diversification (Kubatko & Degnan 2007; Degnan & Rosenberg 2009). Additionally, it can generate overestimated divergence times since the coalescence process of a gene to a common ancestor usually occurs after the true cladogenetic event backwards in time (Maddison & Knowles 2006; Carstens et al. 2007; Edwards et al. 2007; Heled & Drummond 2010).

The main causes of incongruences among genes' history and species' history are related to the fact that genes can be independently affected by distinct evolutionary processes such as horizontal gene transfer, duplication/deletion, incomplete lineage sorting (ILS), selection and hybridization (Degnan & Rosenberg 2009; Huang & Knowles 2009). Recent studies have been

showing that when natural variation among independent loci is not taken into account statistical power of demographic parameters is reduced and phylogenies with taxa that are not reciprocally monophyletic are hardly resolved (Carstens & Knowles 2007). These reasons resulted in a fast transition from the single locus mtDNA to multilocus data sets (McCormack et al. 2013).

Another important shift in phylogeography is related to the use of coalescent models for statistically test competing hypotheses and estimate demographic parameters, known as statistical phylogeography (Knowles & Maddison 2002; Hickerson et al. 2010). In its initial phase phylogeographic inferences were based on quantitative interpretations such as the shape of a topology, the presence of geographic structured clades, the time of divergence among branches or specific summary statistics that described the data (Slatkin & Maddison 1989; Tajima 1989; Excoffier et al. 1992; Fu & Li 1993). In statistical phylogeography, distinct demographic hypotheses/models are evaluated under Bayesian or likelihood-based approaches and the probabilities of fitting the observed data of each model are compared (Knowles 2009). Under this procedure, models are never treated as the complete history of a taxon but it tests important parameters that describe alternative demographic histories (Anderson 2007; Wakeley 2004). A commonly used approach in statistical phylogeography is the Approximate Bayesian Computation (ABC), which provides a rigorous statistical method to test for competing *a priori* hypotheses, and its use has grown considerably in recent years (Beaumont et al. 2002; Beaumont et al. 2010; Bertorele et al. 2010; Huang et al. 2011). This Bayesian approach estimates the posterior distribution of model parameters by replacing the likelihood computation by a measure of similarity between observed and simulated data, reducing the complexity of the data to summary statistics, allowing to test complex and more realistic diversification scenarios and also to use large datasets (Beaumont et al. 2010).

The demand for larger multi-loci data sets in order to obtain more robust parameter estimation has encouraged the development and application of new DNA sequencing methods so-

called next-generation sequencing (NGS), which is less expensive and laborious than Sanger sequencing that requires isolated amplification and sequencing of each marker for each individual. In NGS platforms it is possible to combine multiple individuals in a massively parallel sequencing of thousands of loci that are later processed with bioinformatic pipelines (McCormack et al. 2013). While whole-genome sequencing techniques are still restrictive for most phylogeographic studies, due to the elevated costs and analytical complexity (McCormack et al. 2013), reduced genomic representation techniques are quickly becoming standard methods in the field (Lerner & Fleischer 2010; McCormack et al. 2013). Despite the quick development and relatively large number of new sequencing methods, there are limitations to be taken into account (Rokas & Abbot 2009; Lerner & Fleischer 2010; Lemmon & Lemmon 2012; McCormack et al. 2012; McCormack et al. 2013). Among the most common limitations are: 1) in general phylogeographic studies are focused on non model organisms (without reference genomes), which makes it difficult to establish homology among loci from different individuals; 2) the need for large numbers of individuals per population, producing intractable data sets for most of the available full-likelihood methods; 3) lack of consensus between protocols for genomic library preparation; and 4) arbitrary threshold for data filtering parameters in bioinformatics pipelines (McCormack et al. 2012, 2013; Harvey et al. 2016).

Among subgenomic sequencing methods currently applied to phylogeography, sequence capture or target enrichment methods have overcome some limitations presented above. This procedure involves the capture of conserved genomic regions by using DNA or RNA probes that hybridize with target DNA enabling the retention of captured fragments that can be sequenced by NGS platforms (Gnirke et al. 2009). This method uses probes that are analogous to primers used in PCR for Sanger sequencing that anneal to specific conserved regions in the DNA of the organisms being studied (eg. mammals and birds; McCormack et al. 2012). When compared to other NGS methodologies, as restriction digest sequencing (RADseq; Van Tassel et al. 2008), sequence capture does not produce a random genomic representation that demands reference genomes. Among the

benefits of sequence capture are: 1) easy selection of locus size; 2) probes are usually widely distributed along the genome; 3) probes usually target regions with low copy number reducing paralogy; 4) it is possible to obtain larger loci by using interconnected probes increasing the information content of each locus and gene tree resolution (McCormack et al. 2013).

Among the targeted genomic regions by sequence capture approaches, ultra conserved elements (UCEs) have shown promising results in distinct timescales (Faircloth et al. 2012; Smith et al. 2014; Manthey et al. 2015; Harvey et al. 2016). UCEs are conserved genomic regions between 50 bp to 200 bp shared among highly divergent lineages (Bejerano et al. 2004). Initially described in the alignment of human and mouse genomes, UCEs were then identified in several groups of vertebrates (Stephen et al. 2008; Janes et al. 2011; Crawford et al. 2012), invertebrates and fungi (Siepel et al. 2005). The functions of UCEs are still poorly understood, but so far studies suggest that part of these markers are related to non coding regions associated to gene regulation and early-stage development (Bejerano et al. 2004; Woolfe et al. 2005; Ahituv et al. 2007; Warnefors et al. 2016; Terracciano et al. 2017). The advantages of using sequence capture of UCEs compared to other markers and sequencing techniques are: 1) reduced overlap with duplicated areas of the genome (Derti et al. 2006); 2) flanking regions progressively less conserved, enabling phylogenetic and phylogeographic studies in distinct time scales (Faircloth et al. 2012); 3) probes are used as reference enabling straightforward identification of orthologous contigs. Even though most of the studies based on UCEs focus on deep phylogenies (classes, orders, families, and genus), recent studies reported well supported intraspecific relationships and demographic parameters (Smith et al. 2013; Manthey et al. 2015; Harvey et al. 2016). Smith et al. (2013) tested the applicability of UCEs in shallow diversification scenarios in five bird lineages with populations occurring in four Neotropical areas of endemism (Central America, Choco, Napo, and south of Amazon River). The results, revealed a similar diversification pattern to those obtained with mtDNA but with more recent estimated divergence times and a drastic reduction in the posterior distribution of

demographic parameters, suggesting a more precise estimation. The results obtained by Smith et al. (2013) indicated that UCEs can be successfully used in phylogeographic studies including demographic parameter estimation increasing the resolution obtained with mtDNA.

Amazonian biogeography

One of the greatest challenges in biogeography is to understand how the high biodiversity of tropical environments was generated and how historical processes and ecological traits acted enabling a large number of species to co-exist (Wiens et al. 2011). Amongst tropical regions, the Amazon (*sensu lato*) stands out as the largest fluvial basin with more than 5.5 million km^2, around 40% of the remaining tropical forests in the world, as well as one of the greatest biodiversity in the planet (Silva et al. 2005; Irion & Kalliola 2010). Despite the fact that new species are still being described, the Amazon basin harbors at least 7,500 species of Lepidoptera, 1,500 species of birds and around 112,000 species of trees (Hubbell et al. 2008; Hoorn & Wesselingh 2010). The megadiversity of the Amazon fauna has led to the formulation of different biogeographic hypotheses and speciation models (Antonelli et al. 2010), mostly based on shared species distribution patterns and phylogeographic data correlated to historical events (Wallace 1852; Haffer 1969, 2001; Moritz et al. 2000; Ribas et al. 2011). Since the beginning of the 19^{th} century, the high diversity associated with complex geographic distributions instigated naturalists and researchers to propose hypotheses that explain the formation of the current Amazonian biodiversity (Wallace 1852; Haffer 1969, 2001). These hypotheses were mostly based on congruent biogeographic patterns, suggesting that certain common cladogenetic events were determinant for the formation of the present diversity (Haffer 2001). The application of molecular biology tools in phylogenetic and population genetics of Amazonian taxa has allowed testing some of these hypotheses, including Amazonian birds (Aleixo 2004; Fernandes et al. 2012, 2014; d'Horta 2013 et al.; Capurucho et al. 2013). However, the lack of hierarchical division in spatial and temporal assumptions does not

allow mutual exclusiveness among most hypotheses, blurring the effects of particular historical events on the formation of the current genetic diversity (Patton & Silva 1998; Hall & Harvey 2002). Additionally, intrinsic characteristics of each organism studied such as ecology and the evolutionary origin tend to produce distinct dispersal probabilities over physical barriers, suggesting an overall complex scenario that cannot be explained by a single generalized hypothesis (Burney & Brumfield 2009; Smith et al. 2014).

Among the most discussed hypotheses for the diversification of terrestrial vertebrates in the Amazon, past forest refugia have long been adopted to explain the origin of the Amazon diversity (refugia hypothesis; Haffer 1969). Evidence supporting this hypothesis have been observed in the Northern Hemisphere (Lessa et al. 2003; Hewitt 2000) and the Neotropical region, in the Atlantic Forest (Carnaval et al., 2009). Although some studies on Amazon birds reported demographic expansions for some populations (Aleixo 2004; Fernandes et al., 2012), a direct relationship between these forest range oscillations and cladogenetic events were not observed, refuting the main assumption of this hypothesis. Phylogeographic studies on Amazonian birds and mammals that refute the refugia hypothesis usually assume that large rivers are effective barriers to gene flow between populations of opposite margins (riverine hypothesis; Wallace, 1854), suggesting that the formation of the current Amazon basin produced the observed distribution patterns (Fernandes et al. 2012; Ribas et al. 2011; Boubli et al. 2015; Thom & Aleixo 2015). Despite the intense debate around the refugia and riverine hypotheses, it is common sense that several other processes may have operated on the formation of the current Amazon diversity (Bush 1994; Haffer 2001; Leite & Rogers 2013).

The increase in the number of phylogeographic studies, as well as the reconstruction of paleobiogeographic models (Aleixo 2004; Solomon et al. 2008; Patel et al. 2011; Weir & Price 2011; Ribas et al. 2011), have shed light on the formation history of the Amazon basin. However, this understanding is still sparse when compared to the high and complex biological diversity in this

region. Nevertheless, this complexity in patterns and processes tends to increase since, in its vast majority, phylogeographic studies have been based on organisms restricted to the non-flooded ombrophilous forests (upland forests), the most abundant and diverse environment in the biome (Silva et al. 2005; Capurucho et al.2013).

Biogeography of floodplain forest species

Studies of upland species of vertebrates, mainly birds and mammals, set the basis for the current knowledge on Amazonian biogeography (Haffer 1969; Cracraft 1985; Ribas et al. 2011; D'Horta et al. 2013). However, organisms occurring in other kinds of environments such as open areas (eg. meadows and campinaranas) and floodplains have rarely been subjects of phylogeographic studies. These organisms tend to respond differently to the same historical events or could be affected by processes that did not act on upland forest species (Capurucho et al. 2013).

The Amazonian floodplains house the most diverse and largest flooded forests in the world, covering more than 300,000 km^2, with high levels of species endemicity - approximately ten percent of tree species and 15 percent of non aquatic bird species are endemic (Remsen & Parcker 1983; Wittmann et al. 2006; Junk et al. 2011; Wittmann et al. 2012). Despite its large diversity and relatively easy access, endemic terrestrial species are poorly represented in phylogeographic studies and the potential diversification processes shaping the current genetic diversity are poorly understood resulting in unknown levels of cryptic diversity (Aleixo 2006; Cadena et al. 2011; O'Neil et al. 2011; Harvey et al. 2017). This is critical since this is a highly threatened environment due to habitat conversion for agriculture and energy production by hydroelectric power plants (Albernaz 2011; Latrubesse et al. 2017) with several endangered endemic species (BirdLife international 2016).

The linear connectivity of floodplains and other flooded habitats along Amazonian rivers potentially allows adapted species to disperse freely, forming panmictic populations (Aleixo et al.,

2006). Annual fluctuations in the level of the rivers in Amazonian floodplains resulted in cycles characterized by intense sediment and nutrient transportation resulting in constant changes in river channels and river-created environments (Wittmann et al. 2006; Junk et al. 2011). Thus the dynamics of this environment may lead to selection for better dispersion abilities (Remsen & Parker et al., 1983). For these reasons, it is more likely that taxa adapted to floodplains do not have the same population structure as upland species, whose distributions are generally subdivided by large rivers (Ribas et al. 2011; Smith et al., 2014). The few phylogeographic studies of terrestrial vertebrates occupying flooded environments supported this scenario of lack of genetic structure between populations distributed throughout the entire Amazon basin with lower levels of genetic diversity when compared to upland forest species (Aleixo et al. 2006; Cadena et al. 2011; Harvey et al. 2017). However, when Cohn-Haft et al. (2007) characterized the geographical distribution of *várzea* bird, they described three zones of endemism for the Amazonian floodplains, suggesting that the region separating the Solimões and Amazonas rivers is a suture zone with a gradual transition between distinct communities. Similarly, community turnovers based on alpha diversity were described in several taxonomic groups including fishes, spiders, and plants, suggesting a shared pattern of diversification that so far was not properly explored with molecular tools (Hubert & Renno 2006; Albernaz et al. 2011; Vinticinque et al. 2007).

Choueri et al. (2017), while describing the landscape genetics of four antbird species along the Negro river archipelagos, included individual samples from other large Amazonian rivers (Madeira, Solimões, and Amazonas) and found a considerable level of divergence between samples restricted to distinct rivers. Despite the low number of samples used by Choueri et al. (2017), this geographically structured pattern suggested that historical events might have affected the diversification of distinct lineages of these lowland bird species. Thus, comparative phylogeographic studies focusing on taxa restricted to Amazonian floodplains with phenotypic structure along the main Amazonian tributaries (as the ones used by Cohn-Haft et al. [2007] and

Choueri et al. [2017]) can reveal new scenarios and drivers for the diversification of the large Amazon biodiversity.

Study groups

In the present study, we focus on the patterns and processes related to the diversification of birds restricted to Amazonian floodplains, specifically Antbirds (Thamnophilidae - Aves). South America concentrates the highest diversity of birds, with more than 2,700 species (del Hoyo et al. 2018) and the Amazon basin concentrates more than half of this number (Hubbel et al. 2008; Hoorn & Wesselingh 2010). Antbirds are a Neotropical radiation with approximately 241 species (Zimmer & Isler 2017), considered as good models for phylogeographic studies in the Amazon due the high diversity in the biome - more than 40 species can be found syntopically -, intimate relationship with specific environments, and high intraspecific genetic structure possibly related to poor dispersal abilities (Thom & Aleixo 2015; Fernandes et al. 2012; Fernandes et al. 2014; Zimmer & Isler 2017). Here we selected three lineages (species complexes) of the family Thamnophilidae, intimately related to river-created environments widely distributed over the main Amazonian rivers such as Amazon, Solimões, Madeira, Negro, Branco and Tapajós rivers. The selected species complexes and their distribution areas are described below (Figure 1).

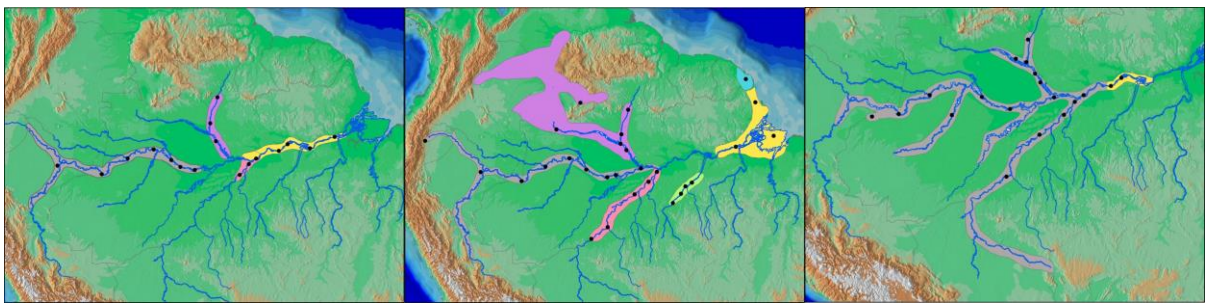

Figure 1: Geographic distribution of the three taxa studied. Left - *Myrmoborus lugubris* (gray - *M. l. berlepschi*; purple - *M. l. stictopterus*; pink - *M. l. femininus*; yellow - *M. l. lugubris*). Center - *Thamnophilus nigrocinereus/cryptoleucus* (gray - *T. cryptoleucus*; purple - *T. n. cinereoniger*; pink - *T. n. tschudii*; green - *T. n. huberi*; yellow - *T. n. nigrocinereus*; blue - *T. n. kulczynskii*). Right - *Myrmotherula assimilis* (grey - *M. a. assimilis*; yellow - *M. a. transamazonica*).

Myrmoborus lugubris (Cabanis, 1847)

Occurs in the understory of river-created environments in *várzea* forests, mainly in islands of whitewater rivers (Amazon, Branco, and Solimões rivers) usually in marshy forests dominated by *Cecropia* sp. and *Heliconia* sp. (Zimmer & Isler 2017). Four subspecies are recognized (Figure 1): *M. l. lugubris* occurs in the Amazon River west of the Madeira river; *M. l. femininus* is restricted to the lower course of the Madeira river; *M. l. stictopterus* is distributed in the Negro and Branco rivers; *M. l. berlepschi* is restricted to the Solimões Basin. The taxa *M. l. stictopterus* and *M. l. female*, resemble intermediate phenotypes between the nominal form and *M. l. berlepschi*, located in the distal portions of the geographic distribution (Haffer & Fitzpatrick 1985). Thus the morphological structure of the complex could be related to a clinal variation with a continuous transition between taxa (Zimmer & Isler 2017). The sister species of *M. lugubris* complex is *M. leucophrys* (Bravo 2012).

Thamnophilus nigrocinereus Sclater, 1855

Distributed over the entire Amazon river and some large tributary rivers such as Negro (Branco), Madeira, Tapajós, Xingu, and Tocantins rivers (Figure 1). It occurs in the understory and medium strata of floodplain forests, gallery forests, mangroves and savanna shrubs, always close to large rivers. However, in most of the distribution, it prefers islands and borders of *igapós* forest (Zimmer & Isler 2017). This species complex is composed of two species, *T. cryptoleucus*, which is monotypic, and *T. nigrocinereus*, which groups five subspecies. *T. n. nigrocinereus* is distributed in the lower Amazon river; *T. n. kulczynskii* occurs in French Guiana and northeast Amapá state in Brazil mainly in mangroves; *T. n. huberi* is restricted to the medium and upper Tapajos river; *T. n. cinereoniger* occurs in the Negro and Branco rivers as well as part of the Orinoco basin; *T. n. tschudii* is endemic to the lower and middle courses of the Madeira river. Zimmer & Isler (2017)

reported that the morphological and vocal variability within *T. nigrocinereus* suggests that a taxonomic revision is needed as some of these taxa should be recognized as full species. Brumfield & Edwards (2007) supported the reciprocal monophyly of *T. nigrocinereus* and *T. cryptoleucus*, with low genetic distance in mitochondrial markers (ND2: 1.1% and cyt b: 1.3%), proposing a recent scenario of diversification or introgression, since these species occur in sympatry at the confluence of the Madeira and Solimões rivers. These authors did not recover with high statistical support the sister group of *T. nigrocinereus* complex, but this complex is in a clade composed by *T. punctatus, T. stictocephalus, T. caerulescens, T. unicolor, T. aroyae*, and *T. aethiops.*

Myrmotherula assimilis Pelzeln, 1868

It is distributed over the Solimões, Madeira, Purus, Jurua, Negro, and upper Amazon rivers, occupying a wider range of environments than the two previous study groups, including areas of tall *várzea* and *igapó* forests (Figure 1). Two subspecies are recognized: *M. a. assimilis* occurs along most part of the distribution, and *M. a. transamazonica* is restricted to the Amazon River in Pará state, Brazil (Zimmer & Isler 2017). The sister species is *Myrmotherula menetriesii* (Bravo et al. 2012).

Objectives

The aim of this doctoral Dissertation was to produce new evidence about the processes shaping the current genetic diversity of Amazonian floodplain communities to contribute with knowledge on the evolution of this ecosystem. We analyzed the molecular systematics, phylogeographic structure, and historical demography of endemic Amazonian floodplain forest birds to understand the diversification patterns of this group of organisms. The specific objectives are presented in each chapter of the Dissertation.

Bibliography

Ahituv N., Zhu Y., Visel A., Holt A., Afzal V., Pennacchio L.A., Rubin E.M. 2007. Deletion of ultraconserved elements yields viable mice. PLoS Biology. 5.

Albernaz A.L., Pressey R.L., Costa L.R., Moreira M.P., Ramos J.F., Assunção P.A., Franciscon C.H. 2011. Tree species compositional change and conservation implications in the white-water flooded forests of the Brazilian Amazon. Journal of Biogeography. 39:869–883.

Aleixo A. 2004. Historical diversification of a terra-firme forest bird superspecies: a phylogeographic perspective on the role of different hypotheses of Amazonian diversification. Evolution. 58:1303.

Aleixo A. 2006. Historical diversification of floodplain forest specialist species in the Amazon: a case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). Biological Journal of the Linnean Society. 89:383–395.

Anderson D.R. 2008. Model-based inference in the life sciences: a primer on evidence. Springer, New York.

Antonelli A., Quijada-Mascareñas A., Crawford A. J., Bates J. M., Velazco P. M., Wüster W. 2010. Molecular studies and phylogeography of Amazonian tetrapods and their relation to geological and climatic models. Pp. 387-404. In: Hoorn C., Wesselingh E. P. (Eds.) Amazonia, Landscape, and Species Evolution: A Look into the Past. Wiley-Blackwell, London.

Beaumont M.A., Zhang W., Balding D.J. 2002. Approximate Bayesian computation in population genetics. Genetics 162: 2025–2035.

Beaumont M.A. 2010. Approximate Bayesian computation in evolution and ecology. Annual Review of Ecology, Evolution, and Systematics. 41:379–406.

Beaumont M.A., Nielsen R., Robert C., Hey J., Gaggiotti O., Knowles L., Estoup A., Panchal M.,

Corander J., Hickerson M., Sisson S.A., Fagundes N., Chikhi L., Beerli P., Vitalis R., Cornuet J.-M., Huelsenbeck J., Foll M., Yang Z., Rousset F., Balding D., Excoffier L. 2010. In defense of model-based inference in phylogeography. Molecular Ecology. 19:436–446.

Bejerano G. 2004. Ultraconserved elements in the human genome. Science. 304:1321–1325.

Bertorelle G., Benazzo A., Mona S. 2010. ABC as a flexible framework to estimate demography over space and time: some cons, many pros. Molecular Ecology. 19:2609–2625.

BirdLife International. 2016. The IUCN Red List of Threatened Species 2016: e.T22680736A92875297.http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22680736A92875297.en. Downloaded on 20 November 2017.

Boubli J.P., Ribas C., Alfaro J.W.L., Alfaro M.E., Silva M.N.F.D., Pinho G.M., Farias I.P. 2015. Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. Molecular Phylogenetics and Evolution. 82:400–412.

Bravo G.A. 2012. Phenotypic and niche evolution in the antbirds (Aves, Thamnophilidae). Ph.D. Dissertation, Louisiana State University, Baton Rouge.

Brown J.L., Weber J.J., Alvarado-Serrano D.F., Hickerson M.J., Franks S.J., Carnaval A.C. 2016. Predicting the genetic consequences of future climate change: The power of coupling spatial demography, the coalescent, and historical landscape changes. American Journal of Botany. 103:153–163.

Brumfield R.T., Edwards S.V. 2007. Evolution into and out of the Andes: a Bayesian analysis of historical diversification in *Thamnophilus* antshrikes. Evolution. 61:346–367.

Burney, Brumfield. 2009. Ecology predicts levels of genetic differentiation in neotropical birds. The American Naturalist. 174:358.

Bush M.B. 1994. Amazonian speciation: a necessarily complex model. Journal of Biogeography. 21:5-17.

Cadena C.D., Gutiérrez-Pinto N., Dávila N., Chesser R.T. 2011. No population genetic structure in a widespread aquatic songbird from the Neotropics. Molecular Phylogenetics and Evolution. 58:540–545.

Capurucho J.M.G., Cornelius C., Borges S.H., Cohn-Haft M., Aleixo A., Metzger J.P., Ribas C.C. 2013. Combining phylogeography and landscape genetics of *Xenopipo atronitens* (Aves: Pipridae), a white sand campina specialist, to understand Pleistocene landscape evolution in Amazonia. Biological Journal of the Linnean Society. 110:60–76.

Carnaval A.C., Hickerson M.J., Haddad C.F.B., Rodrigues M.T., Moritz C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science. 323:785–789.

Carstens B.C., Knowles L.L., Collins T. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. Systematic Biology. 56:400–411.

Choueri É.L., Gubili C., Borges S.H., Thom G., Sawakuchi A.O., Soares E.A.A., Ribas C.C. 2017. Phylogeography and population dynamics of antbirds (Thamnophilidae) from Amazonian fluvial islands. Journal of Biogeography. 44:2284–2294.

Cohn-Haft M., Naka L.N., Fernandes A.M. 2007. Padrões de distribuição da avifauna da várzea dos rios Solimões-Amazonas. Pp. 287-324. In: Albernaz A.L. (Eds). Conservação da Várzea, Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ ProVárzea/INPA, Manaus.

Cracraft J. 1985. Historical biogeography and patterns of differentiation within the South American avifauna: areas of endemism. Ornithological Monographs.:49–84.

Crawford N.G., Faircloth B.C., Mccormack J.E., Brumfield R.T., Winker K., Glenn T.C. 2012.

More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. Biology Letters. 8:783–786.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference, and the multispecies coalescent. Trends in Ecology and Evolution. 24:332–340.

Derti A., Roth F.P., Church G.M., Wu C.-T. 2006. Mammalian ultraconserved elements are strongly depleted among segmental duplications and copy number variants. Nature Genetics. 38:1216–1220.

D'Horta F.M., Cuervo A.M., Ribas C.C., Brumfield R.T., Miyaki C.Y. 2012. Phylogeny and comparative phylogeography of *Sclerurus* (Aves: Furnariidae) reveal constant and cryptic diversification in an old radiation of rainforest understorey specialists. Journal of Biogeography. 40:37–49.

Edwards S.V., Liu L., Pearl D.K. 2007. High-resolution species trees without concatenation. Proceedings of the National Academy of Sciences of the USA. 104:5936–5941.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

Edwards S.V., Beerli P. 2000. Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution. 54:1839.

Excoffier L., Smouse P.E., Quattro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

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

Felsenstein J. 2008. Inferring Phylogenies. Sinauer Associates, Inc., Sunderland.

Fernandes A.M., Wink M., Sardelli C.H., Aleixo A. 2014. Multiple speciations across the Andes and throughout Amazonia: the case of the spot-backed antbird species complex (*Hylophylax naevius/Hylophylax naevioides*). Journal of Biogeography. 41:1094–1104.

Fernandes A.M., Wink M., Aleixo A. 2012. Phylogeography of the chestnut-tailed antbird (*Myrmeciza hemimelaena*) clarifies the role of rivers in Amazonian biogeography. Journal of Biogeography. 39:1524–1535.

Gnirke A., Melnikov A., Maguire J., Rogov P., Leproust E.M., Brockman W., Fennell T., Giannoukos G., Fisher S., Russ C., Gabriel S., Jaffe D.B., Lander E.S., Nusbaum C. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nature Biotechnology. 27:182–189.

Haffer J. 1969. Speciation in Amazonian forest birds. Science. 165:131–137.

Haffer J., Fitzpatrick J.W. 1985. Geographic variation in some Amazonian forest birds. Ornithological Monographs. 36:147–168.

Haffer J. 2001. Hypotheses to explain the origin of species in Amazonia. Pp. 45-118. In: Viera I., D'Incao M.A., Silva J.M.C., Oren D. (Eds.) Diversidade Biológica e Cultural da Amazônia. Museu Paraense Emilio Goeldi, Belém.

Hall J.P.W., Harvey D.J. 2002. The phylogeography of Amazonia revisited: new evidence from riodinid butterflies. Evolution. 56:1489.

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

Harvey M., Aleixo A., Ribas C.C., Brumfield R.T. 2016. Habitat preference predicts genetic diversity and population divergence in Amazonian birds. The American Naturalist 190: 631-648

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

Heled J., Drummond A.J. 2009. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution. 27:570–580.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. Nature. 405:907–913.

Hickerson M., Carstens B., Cavender-Bares J., Crandall K., Graham C., Johnson J., Rissler L., Victoriano P., Yoder A. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution. 54:291–301.

Hoorn C., Wesselingh F.P. 2010. Amazonia landscape and species evolution: a look into the past. Wiley-Blackwell, London.

Huang H., Knowles L.L. 2009. What is the danger of the anomaly zone for empirical phylogenetics? Systematic Biology. 58:527–536.

Huang W., Takebayashi N., Qi Y., Hickerson M.J. 2011. MTML-msBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. BMC Bioinformatics. 12:1.

Hubbell S.P., He F., Condit R., Borda-De-Agua L., Kellner J., Steege H.T. 2008. How many tree species are there in the Amazon and how many of them will go extinct? Proceedings of the National Academy of Sciences of the USA. 105:11498–11504.

Hubert N., Renno J.-F. 2006. Historical biogeography of South American freshwater fishes. Journal of Biogeography. 33:1414–1436.

Irion G., Kalliola R. 2010. Long-term landscape development process in Amazônia. Pp. 185-197. In: Hoorn, C., Wesselingh, E.P. (Eds.) Amazonia, Landscape, and Species Evolution: A Look

into the Past. Wiley-Blackwell, London.

Janes D.E., Chapus C., Gondo Y., Clayton D.F., Sinha S., Blatti C.A., Organ C.L., Fujita M.K., Balakrishnan C.N., Edwards S.V. 2010. Reptiles and mammals have differentially retained long conserved noncoding sequences from the amniote ancestor. Genome Biology and Evolution. 3:102–113.

Junk W.J. 2011. Amazonian Floodplain Forests: Ecophysiology, Biodiversity, and Sustainable Management. Springer, Berlin.

Junk W.J., Piedade M.T.F., Schöngart J., Cohn-Haft M., Adeney J.M., Wittmann F. 2011. A classification of major naturally-occurring Amazonian lowland wetlands. Wetlands. 31:623–640.

Knowles L.L. 2009. Statistical phylogeography. Annual Review of Ecology, Evolution, and Systematics. 40:593–612.

Knowles L.L., Maddison W.P. 2002. Statistical phylogeography. Molecular Ecology. 11:2623–2635.

Kubatko L.S., Degnan J.H., Collins T. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology. 56:17–24.

Latrubesse E.M., Arima E.Y., Dunne T., Park E., Baker V.R., D'Horta F.M., Wight C., Wittmann F., Zuanon J., Baker P.A., Ribas C.C., Norgaard R.B., Filizola N., Ansar A., Flyvbjerg B., Stevaux J.C. 2017. Damming the rivers of the Amazon basin. Nature. 546:363–369.

Leite R.N., Rogers D.S. 2013. Revisiting Amazonian phylogeography: insights into diversification hypotheses and novel perspectives. Organisms, Diversity and Evolution. 13:639–664.

Lemmon A.R., Lemmon E.M. 2012. High-throughput identification of informative nuclear loci for shallow-scale phylogenetics and phylogeography. Systematic Biology. 61:745–761.

Lerner H., Fleischer R. 2010. Prospects for the use of next-generation sequencing methods in

ornithology. The Auk. 127:4–15.

Lessa E.P., Cook J.A., Patton J.L. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. Proceedings of the National Academy of Sciences of the USA. 100:10331–10334.

Fu Y.X., Li W.H. 1993. Statistical tests of neutrality of mutations. Genetics 133: 693–709.

Maddison W.P., Knowles L.L., Collins T. 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology. 55:21–30.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (aves, genus: *Piranga*). Systematic Biology. 65:640–650.

Mccormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., Brumfield R.T. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution. 66:526–538.

Mccormack J.E., Maley J.M., Hird S.M., Derryberry E.P., Graves G.R., Brumfield R.T. 2012. Next-generation sequencing reveals phylogeographic structure and a species tree for recent bird divergences. Molecular Phylogenetics and Evolution. 62:397–406.

Moritz C., Patton J.L., Schneider C.J., Smith T.B. 2000. Diversification of rainforest faunas: an integrated molecular approach. Annual Review of Ecology and Systematics. 31:533–563.

Moritz C., Faith D.P. 1998. Comparative phylogeography and the identification of genetically divergent areas for conservation. Molecular Ecology. 7:419–429.

Moritz C., Funk V., Sakai A.K. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology. 51:238–254.

Nei M., Li W.H. 1979. Mathematical model for studying genetic variation in terms of restriction

endonucleases. Proceedings of the National Academy of Sciences of the USA. 76:5269–5273.

O'Neill J.P., Lane D.F., Naka L.N. 2011. A cryptic new species of thrush (Turdidae: T*urdus*) from western Amazonia. The Condor. 113:869–880.

Patel S., Weckstein J.D., Patané J.S., Bates J.M., Aleixo A. 2011. Temporal and spatial diversification of *Pteroglossus* araçaris (AVES: Ramphastidae) in the Neotropics: Constant rate of diversification does not support an increase in radiation during the Pleistocene. Molecular Phylogenetics and Evolution. 58:105–115.

Patton J.L., da Silva M.N.F. 1998. Rivers, refuges, and ridges. The geography of speciation of Amazonian mammals. Pp. 202-213 In D.J. Howard, S.H. Berlocher (Eds.) Endless Forms: Species and Speciation. Oxford University Press, New York.

Prates I., Xue A.T., Brown J.L., Alvarado-Serrano D.F., Rodrigues M.T., Hickerson M.J., Carnaval A.C. 2016. Inferring responses to climate dynamics from historical demography in neotropical forest lizards. Proceedings of the National Academy of Sciences of the USA. 113:7978–7985.

Remsen J.V., Iii T.A.P. 1983. Contribution of river-created habitats to bird species richness in Amazonia. Biotropica. 15:223.

Ribas C.C., Aleixo A., Nogueira A.C.R., Miyaki C.Y., Cracraft J. 2011. A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. Proceedings of the Royal Society B: Biological Sciences. 279:681–689.

Rokas A., Abbot P. 2009. Harnessing genomics for evolutionary insights. Trends in Ecology and Evolution. 24:192–200.

Siepel A. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Research. 15:1034–1050.

Silva J.M.C.D., Rylands A.B., Fonseca G.A.B.D. 2005. The fate of the Amazonian areas of

endemism. Conservation Biology. 19:689–694.

Slatkin M., Maddison W.P. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123: 603–613.

Smith B.T., Harvey M.G., Faircloth B.C., Glenn T.C., Brumfield R.T. 2013. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Systematic Biology. 63:83–95.

Smith B.T., Mccormack J.E., Cuervo A.M., Hickerson M.J., Aleixo A., Cadena C.D., Pérez-Emán J., Burney C.W., Xie X., Harvey M.G., Faircloth B.C., Glenn T.C., Derryberry E.P., Prejean J., Fields S., Brumfield R.T. 2014. The drivers of tropical speciation. Nature. 515:406–409.

Solomon S.E., Bacci M., Martins J., Vinha G.G., Mueller U.G. 2008. Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* spp.) Provide new insight into the origins of Amazonian diversity. PLoS ONE. 3:e2738.

Soltis D.E., Morris A.B., Mclachlan J.S., Manos P.S., Soltis P.S. 2006. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology. 15:4261–4293.

Stephen S., Pheasant M., Makunin I.V., Mattick J.S. 2008. Large-scale appearance of ultraconserved elements in tetrapod genomes and slowdown of the molecular clock. Molecular Biology and Evolution. 25:402–408.

Taberlet P. 2002. Ecology: enhanced: quaternary refugia and persistence of biodiversity. Science. 297:2009–2010.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Tassell C.P.V., Smith T.P.L., Matukumalli L.K., Taylor J.F., Schnabel R.D., Lawley C.T., Haudenschild C.D., Moore S.S., Warren W.C., Sonstegard T.S. 2008. SNP discovery and allele

frequency estimation by deep sequencing of reduced representation libraries. Nature Methods. 5:247–252.

Terracciano D., Terreri S., Nigris F.D., Costa V., Calin G.A., Cimmino A. 2017. The role of a new class of long noncoding RNAs transcribed from ultraconserved regions in cancer. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 1868:449–455.

Thom G., Aleixo A. 2015. Cryptic speciation in the white-shouldered antshrike (*Thamnophilus aethiops*, Aves – Thamnophilidae): The tale of a transcontinental radiation across rivers in lowland Amazonia and the northeastern Atlantic Forest. Molecular Phylogenetics and Evolution. 82:95–110.

Venticinque E.M., Rego F.N.A.A., Brescovit A.D., Rheims C.A., Ruiz G.R.S. 2007. A araneofauna (Arachnida, Araneae) das várzeas do Rio Amazonas: padrões de distribuicao e estado do conhecimento atual. Pp. 179-198. In: Albernaz A.L. (Eds). Conservação da Várzea, Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ ProVárzea/INPA, Manaus.

Wakeley J. 2004. Recent trends in population genetics: more data! More math! Simple models? Journal of Heredity. 95:397–405.

Wakeley J. 2009. Coalescent Theory: an Introduction. Roberts & Co. Publishers, Greenwood Village.

Wallace A.R. 1854. On the monkeys of the Amazon. Journal of Natural History Series 2. 14: 451–454.

Warnefors M., Hartmann B., Thomsen S., Alonso C.R. 2016. Combinatorial gene regulatory functions underlie ultraconserved elements in *Drosophila*. Molecular Biology and Evolution. 33:2294–2306.

Weir J.T., Price M. 2011. Andean uplift promotes lowland speciation through vicariance and

dispersal in *Dendrocincla* woodcreepers. Molecular Ecology. 20:4550–4563.

Wiens J.J., Pyron R.A., Moen D.S. 2011. Phylogenetic origins of local-scale diversity patterns and the causes of Amazonian megadiversity. Ecology Letters. 14:643–652.

Williams J.W., Jackson S.T., Kutzbach J.E. 2007. Projected distributions of novel and disappearing climates by 2100 AD. Proceedings of the National Academy of Sciences of the USA. 104:5738–5742.

Wittmann F., Householder E., Piedade M.T.F., Assis R.L.D., Schöngart J., Parolin P., Junk W.J. 2012. Habitat specifity, endemism and the neotropical distribution of Amazonian white-water floodplain trees. Ecography. 36:690–707.

Wittmann F., Schongart J., Montero J.C., Motzer T., Junk W.J., Piedade M.T.F., Queiroz H.L., Worbes M. 2006. Tree species composition and diversity gradients in white-water forests across the Amazon Basin. Journal of Biogeography. 33:1334–1347.

Woolfe A., Goodson M., Goode D.K., Snell P., Mcewen G.K., Vavouri T., Smith S.F., North P., Callaway H., Kelly K., Walter K., Abnizova I., Gilks W., Edwards Y.J.K., Cooke J.E., Elgar G. 2004. Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biology. 3.

Zimmer K., Isler M.L. 2017. Typical Antbirds (Thamnophilidae). In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A. & de Juana, E. (Eds.). Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona. (retrieved from https://www.hbw.com/node/52291 on 20 November 2017).

Chapter 1

Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species

Manuscript accepted for publication in the Journal *Systematic Biology*.

DOI:10.1093/sysbio/syy004/4828098

Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species

Gregory Thom[1]*; Fabio R. Amaral[2]; Michael J. Hickerson[3]; Alexandre Aleixo[4]; Lucas E. Araujo-Silva[4]; Camila C. Ribas[5]; Erik Choueri[5]; Cristina Y. Miyaki[1]

1 Departamento de Genética e Biologia Evolutiva, Universidade de São Paulo, Rua do Matão, 277, Cidade Universitária, São Paulo, SP, 05508-090, Brazil.

2 Universidade Federal de São Paulo, Instituto de Ciências Ambientais, Químicas e Farmacêuticas, Laboratório de Genética Evolutiva, Rua Professor Artur Riedel, 275, Diadema, SP , 09972–270, Brazil.

3 Department of Biology, Marshak Science Building, City College of New York, New York, NY, USA.

4 Museu Paraense Emílio Goeldi (MPEG), Caixa Postal 399, Belém, PA 66040-170, Brazil.

5 Instituto Nacional de Pesquisas da Amazônia (INPA),Av. André Araújo 2936, Manaus, AM 69060-001, Brazil.

* Corresponding Author: Departamento de Genética e Biologia Evolutiva, Universidade de São Paulo, Rua do Matão, 277, Cidade Universitária, São Paulo, SP, 05508-090, Brazil; Phone: +55 (11) 3091-7582; E-mail: biogrego@yahoo.com.br

Abstract

Before populations become independent evolutionary lineages, the effects of micro evolutionary processes tend to generate complex scenarios of diversification that may affect phylogenetic reconstruction. Not accounting for gene flow in species tree estimates can directly impact topology, effective population sizes and branch lengths and the resulting estimation errors are still poorly understood in wild populations. In this study, we used an integrative approach, including sequence capture of Ultra Conserved Elements (UCEs), mtDNA Sanger sequencing and morphological data to investigate species limits and phylogenetic relationships in face of gene flow in an Amazonian endemic species (*Myrmoborus lugubris*: Aves). We used commonly implemented species tree and model-based approaches to understand the potential effects of gene flow in phylogenetic reconstructions. The genetic structure observed was congruent with the four recognized subspecies of *M. lugubris.* Morphological and UCEs data supported the presence of a wide hybrid zone between *M. l. femininus* from the Madeira river and *M. l. lugubris* from the Middle and lower Amazon river, which were recovered as sister taxa by species tree methods. When fitting gene flow into simulated demographic models with different topologies, the best-fit model indicated these two taxa as nonsister lineages, a finding that is in agreement with the results of mitochondrial and morphological analyses. Our results demonstrated that failing to account for gene flow when estimating phylogenies at shallow divergence levels can generate topological uncertainty, which can nevertheless be statistically well supported, and that model testing approaches using simulated data can be useful tools to test alternative phylogenetic hypotheses.

Keywords: Multispecies coalescent; hybridization; *Myrmoborus lugubris*; Antbird; migration

Introduction

In an evolutionary perspective the speciation process is concluded when genetic differentiation promotes complete reproductive isolation between diverging populations (Harrison 1990). However, before populations become independent evolutionary lineages, the effects of micro evolutionary processes tend to generate complex scenarios of diversification that may affect phylogenetic reconstruction (Hickerson et al. 2006; Leaché et al. 2014; Poelstra et al. 2014; Mallet et al. 2016; Nater et al. 2015; Edwards et al. 2016; Barley et al. 2017).

Incomplete lineage sorting (ILS) and gene flow between taxa are commonly reported in studies involving recent diversification scenarios in the gray zone of phylogeography (Edwards et al. 2016; Meyer et al. 2016; Oswald et al. 2017). Around 10% of animal and 25% of plant taxa exchange genes with other taxa (Mallet 2005, 2007). In phylogenetic reconstructions the effects of ILS and gene flow can be misleading, since both tend to produce inconsistencies between gene trees and the species tree and the outcomes of each process are often difficult to be distinguished and modeled by the coalescent process (Kubatko & Degnan 2007; Kubatko 2009).

The multi-species coalescent (MSC) methods have become one of the most used approaches to perform species tree estimation by using a statistically standardized methodology to test alternative hypotheses of isolation (Liu et al. 2009a; Leaché and Fujita 2010; Bryant et al. 2012; Leaché et al. 2014). Such methods can incorporate the uncertainty of gene trees given the coalescent stochasticity as well as parameters (e.g. effective population sizes and times of divergence) during the estimation of the species tree (Yang and Rannala 2010; Zhang et al. 2014). However, gene flow is not taken into account by most of the species tree methods available (Liu et al. 2009a; Liu and Yu 2011; Bryant et al. 2012; Sousa and Hey 2013; Mirarab et al. 2014), whereby ILS is assumed as the only source of gene tree discordance in a given dataset.

The presence of gene flow between taxa violates the assumption of bifurcating branches

arising from a common ancestor, and not properly accounting for it in species tree estimates can directly impact parameter inferences such as topology, effective population sizes and branch lengths (Eckert & Carstens 2008; Leaché et al. 2014; Solís-Lemus et al. 2016). Notwithstanding the fact that some species tree methods have proven to be robust in the presence of low levels of gene flow, such as ASTRAL and NJst (Solís-Lemus et al. 2016), the explicit accommodation of migration in phylogenetic estimations is critical, specially due to the large amount of recent evidence for reticulated evolution (Lamichhaney et al. 2015; Mallet et al. 2016; Edwards et al. 2016), including cases of speciation with gene flow (Feder et al. 2012), ephemeral speciation (Rosenblum et al. 2012) and hybrid speciation (Mallet 2007).

Temporally dynamic environments, such as areas affected by climatic oscillations during the Quaternary when patches of specific habitats were connected and disconnected as predicted by the refugia hypothesis (Haffer 1969), or riverine environments constantly changed by river dynamics (Bagley et al. 2011), are prone to produce complex scenarios of diversification. The combination of recent history of divergence and dynamic secondary contact between populations can potentially result on paraphyly and introgression of diverging lineages (Hewitt 2004). This scenario demands the application of methods that explicitly model ILS and gene flow to explore phylogenetic relationships among populations or species (Nater et al. 2015; Morales et. al. 2017)

The Amazon Basin comprises the largest fluvial system in the world and its floodplains extend for over 550,000 km^2, with the major tributaries representing approximately 8% of the area of this biome (Wittmann et al. 2010; Junk et al. 2011). Floodplain environments in the Amazon are strongly influenced by annual flooding cycles and were historically shaped by paleoclimatic changes affecting the river dynamics, yielding the formation of a plethora of specific and ephemeral habitats heterogeneously distributed over the basin (Irion et al. 2009; Junk et al. 2011; Wittmann et al. 2012; Rossetti et al. 2015). As a result, a high level of species endemicity is reported, such as ten % of tree species and 15 % of non-aquatic bird species

(Remsen & Parcker 1983; Wittmann et al 2012). The high levels of endemicity contrast with the presence of widely distributed species with low levels of genetic diversity suggesting that the linear distribution and connectivity of floodplains enable high levels of gene flow or that the ephemeral environments could be prone to high extinction rates (Aleixo 2006; Cadena et al. 2011; but see Choueri et al 2017).

The Ash-breasted Antbird (*Myrmoborus lugubris*) inhabits the floodplain forests of large Amazonian rivers, being restricted to a narrow strip of river edge forests, mainly found on river-created islands (Zimmer & Isler 2003). *M. lugubris* as currently recognized, is a single species with four subspecies that are recognized based on minor to moderate plumage differences: *M. l. lugubris* occurs in the east of the Madeira river to the Amazon river mouth; *M. l. berlepschi* is restricted to the Solimões river*; M. l. femininus* is restricted to the Madeira river; and *M. l. stictopterus* occurs in the Negro river (Fig. 1a). Despite the apparent continuity of river created environments along the entire extension of the main channel of the Amazon river, there is no evidence – historical or recent records - that any taxa of this group occurs in the lower portion of the Solimões and Negro rivers (Figure 1a; Petermann 1997; Cohn-Haft et al 2007). It is worth mentioning that these are thoroughly surveyed areas by ornithologists and birdwatchers. Thus, there is strong evidence that *M. l. berlepschi* and *M. l. stictopterus* are allopatric taxa. Haffer and Fitzpatrick (1985) suggested that the central Amazonian taxa (*M. l. femininus* and *M. l. stictopterus*) could be intermediate forms between the two extreme plumage types (*M. l. berlepschi* and *M. l. lugubris*), as a result of secondary contact and intergradation between the western black-faced (*M. l. berlepschi*) and eastern reddish-faced (*M. l. lugubris*) taxa. However, no study has been conducted so far to describe the relationships among these taxa nor whether phenotypic variation mirrors neutral genetic structure. Thus, *M. lugubris* is an interesting model to test the robustness of distinct phylogenetic inference methods in face of a complex scenario of diversification that may present ILS and gene flow.

To test this, we used a model-based integrative approach, that combined sequence capture of Ultra Conserved Elements (UCEs), mtDNA Sanger sequencing and morphological data to investigate the phylogenetic relationships in the Amazon endemic Ash-breasted Antbird complex. We applied empirical and model-based approaches to explore hypotheses regarding species tree estimation. In addition, we tested the robustness of typically used species tree methods in the presence of gene flow between non-sister taxa. Our results indicated that species tree methods, including full likelihood and summary-based approaches, which do not account for migration, supported distinct topologies or did not have phylogenetic resolution. Our results supported that phylogenetic network methods and model-testing approaches based on simulation with thousands of SNPs agreed with the mtDNA phylogeny and phenotypic analyses. Also, diversification occurred around the Middle and Late Pleistocene and secondary contact between nonsister species seem to be related to the dynamics of Amazon floodplains.

Material and Methods

Morphological analyses

We examined 160 skins belonging to all taxa of *Myrmoborus lugubris* (*M. l. lugubris, M. l. berlepschi, M. l. femininus,* and *M. l. stictopterus*) housed at various institutions (Table s1; doi:10.5061/dryad.13b88). The phenotypic variation among *M. lugubris* subspecies is a case of heterogynism (Hellmayr 1929), in which female plumage from different taxa differ more conspicuously than that of males. In general, the front and sides of the head are light ferruginous in females from the lower Amazon river (*M. l. lugubris*), whereas those body parts are black or blackish in western Amazonian subspecies (*M. l. berlepschii, M. l. femininus, M. l. stictopterus*; Zimmer & Isler 2003). Plumage variation was described observing discrete characters over the distribution of the species and comparing with the original description of each taxon. We used Smithe (1975, 1981) color

catalog to describe color variation and plumage tones. Phenotypically intermediate individuals between *M. l. lugubris* and *M. l. femininus* included females from a contact zone between these taxa with a mixed combination of plumage characters diagnostic of each taxa, whereas males from the same localities were assigned by default as intermediates as well. Measurements of the following morphometric characters were taken by G.T. to the nearest 0.1 mm with an Mitutoyo electronic caliper: wing length (from the wrist to the tip of the longest primary), tail length (from the pygostyle insertion to the tip of the longest rectrices), tarsus length, culmen, bill length from the distal points of the nostril to the tip of the bill, bill depth and width at the distal point of the nostril. Wing and tarsus measurements were always taken from the specimens' right side. The normality of the data was assessed with Kolmogorov-Smirnov tests. We used discriminant function analysis (DFA) to test for differences in the morphometric space among recognized taxa of *M. lugubris*. Missing morphometric values for any given character were replaced with the corresponding average value obtained for the taxon to which the individual was assigned and did not exceeded > 5% of the total number of measurements. Intermediate individuals (based on plumage patterns) between *M. l. lugubris* and *M. l. femininus* were considered as a distinct group. All statistical tests were performed with STATISTICA version 8.0. We adopted a statistical significance of $P \leq 0.05$.

mtDNA – DNA extraction, sequencing and data analyses

A total of 198 individuals from 25 localities were sampled throughout the entire distribution of *M. lugubris* (Table s1; Fig. 1a). Subspecific identification was based on the plumage of vouchered individuals. We used *M. leucophrys, M. myotherinus* and *Percnostola lophotes* as outgroups due their close phylogenetic relationship to *M. lugubris* (Isler et al. 2013). Total DNA was extracted from approximately 20 mg of muscle tissue following a standard phenol/chloroform protocol (Sambrook et al. 1989). We amplified the

mitochondrial cytochrome b gene (cyt b, ~1040 bp; Primers L14841: CCATCCAACATCTCAGCATGATGAAA and H16065: AACTGCAGTCATCTCCGGTTTACAAGAC) for all sampled individuals. Polymerase chain reaction (PCR) was performed in 25 µl of final volume, and approximately 50 ng of genomic DNA, 1.5 mM to 2.5 mM of $MgCl_2$, 200 mM of dNTPs and 0.1 U of *Taq* DNA polymerase (Promega, Madison, WI, USA). An initial denaturation step was performed at 94°C for five minutes, followed by 35 cycles of: 1) 94°C for one minute; 2) 50°C for 1minute and 3) 72°C for one minute. The final extension was at 72°C for five minutes. PCR products were purified with Exo-Fap enzymes. Sequencing was carried out on an ABI 3130 or 3730 automated capillary sequencer (Applied Biosystems, Foster City, California, USA) with the ABI Prism Big Dye terminator Kit. Both DNA strands were sequenced for all samples. The DNA sequences were edited and aligned in CodonCode aligner 3.7.1 (CodonCode Inc.).

Phylogenetic analyses were performed using Bayesian inference (BI) in MrBAYES 3.1.2 (Ronquist et al. 2012). The best-fit evolutionary model was selected in JMODELTEST 0.1 .1 (Posada 2008), using the Bayesian information criterion (BIC). BI analysis was performed by two independent runs of 10 millions generations, each with four Markov chains. The parameters of the chains were sampled every 1,000 generations and the first 1,000 trees were discarded as burn-in. We used random seed for starting tree and default priors and initial searching values. The posterior probability for each estimated node was obtained through a majority rule consensus of the remaining MCMC samples.

We performed a Bayesian population (k) clustering analyses without prior information of the sampling locations and the number of taxa in BAPS 6.0 (Corander and Marttinen 2006), using the option “clustering with linked loci”, which accounts for dependence between sites within the aligned sequences. We tested multiple k values (1-8) and performed ten independent runs in order to assess the consistency of the results.

Sequence capture of Ultra Conserved Elements (UCEs)

We selected 22 tissue samples of *M. lugubris* representing all four taxa, from 13 localities covering the entire geographic distribution of the species for capture and sequencing Ultra Conserved Elements (UCEs). Given the close phylogenetic relationship to *M. lugubris*, one sample of *M. leucophrys* and one sample of *M. myotherinus* were used as outgroups (Isler et al. 2013). Genomic DNA of all samples were extracted with QIAGEN DNeasy tissue and Blood kit (Valencia, CA) according to the manufacturer's protocol with a minimum amount of 2 µg of DNA eluted in ~50 µl of AE buffer. All samples were treated with QIAGEN RNAse during extraction.

Library preparation and sequencing of UCEs were performed by RAPiD Genomics (Gainsville, FL, USA), following most of the protocol described by Faircloth et al. (2012). Modifications of the original protocol concern the use of a probe set containing 2312 probes targeting UCEs (ultraconserved.org) plus 97 probes targeting exons typically used in avian phylogenetic studies (Hackett et al. 2008; Kimball et al. 2009), and sequencing of 150 bp paired-ends in Illumina Hiseq 2500. Samples were sequenced in a multiplexed batch of 96 samples, including individuals not analyzed in the present study.

Raw read files were separated per individual using Illumina's Casava software and read quality of each individual was evaluated in FastQC 0.11.4 (Andrews 2014). We filtered raw reads for each individual using Illumiprocessor which trims off adapter sequences and excludes low-quality reads (Faircloth et al. 2012). We conducted de novo assembly across all samples using Trinity 2.4 (Grabherr et al. 2011). Contigs were aligned to UCE probe reference sequences. Contigs that did not align or that matched multiple UCEs were removed with LASTZ using the "match_contigs_to_probes.py" script, available in PHYLUCE 1.4 package (Faircloth et al. 2012; https://github.com/faircloth-lab/phyluce). After generating a SQL database of matches for all individuals, we generated alignments in

fasta format and trimmed off long ragged-ends. A matrix of concatenated sequences with 15% of missing data was used for the phylogenetic analyses. For the remaining analyses we generated a sequence matrix with all *M. lugubris* individuals without filtering for missing data nor trimming, which was used as reference for SNP calling (see below). The longest sequence without indels per loci was selected as reference. Sequence alignments for each locus were performed using MAFFT (Katoh and Standley 2013). Finally, additional scripts available in the PHYLUCE package were used to obtain summary statistics of all loci and to convert file formats for phylogenetic analyses.

We aligned the cleaned and synchronized reads of each individual with the generated reference using BWA (Li and Durbin 2009) and called SNPs using the Genome Analyses Tool Kit (GATK; McKenna et al. 2010), hard-masking low-quality bases (< Q30) and keeping sites with a minimum read depth of > 8. In order to obtain an unlinked SNP matrix, we randomly selected one SNP per locus, excluding sites with missing data. We also blasted (Blast+) the reference sequence used to call the SNPs against the zebra finch genome (available at https://www.ncbi.nlm.nih.gov/books/NBK279690/) to select for autosomal loci (excluding UCEs in sexual chromosomes). We used VCFTOOLS (Danecek et al. 2011) and custom scripts to generate input files for the programs described below. Finally, to obtain phased alleles for each locus we used the function "ReadBackedPhasing" of GATK (McKenna et al. 2010) on the final VCF with a phasing quality threshold of 20.

Phased alleles were incorporated to the reference sequences using the "add_phased_snps_to_seqs_filter.py" from the seqcap_pop package (https://github.com/mgharvey/seqcap_pop; Harvey et al. 2016). Sequence alignments for each locus were performed using MAFFT (Katoh and Standley 2013). No threshold for missing data was applied.

UCEs: Genetic structure

We used sNMF (Frichot et al. 2014) to infer the best-fit number of ancestral populations (k) within *M. lugubris* and to assign individuals to populations. The sNMF algorithm applies a sparse non-negative matrix factorization to compute least-squares estimates of ancestry coefficients. When compared to more commonly used programs, such as STRUCTURE (Pritchard et al. 2000) and ADMIXTURE (Alexander et al. 2009), sNMF can analyze large bi-allelic datasets more efficiently without loss of accuracy. Also, it seems to perform well even in scenarios of departure from the Hardy Weinberg and linkage equilibrium assumptions (Frichot et al. 2014; Harris et al. 2015).

We tested multiple k values (1-8), with 100 replicate runs for each k value. The robustness of the results was assessed by testing four values of the alpha regularization parameter (1, 10, 100, 1000). Additionally, we used the k-means (find.clusters) algorithm available in the package ADEGENET 2.0 (Jombart and Ahmed 2011) to test for the number of genetic clusters in the sample. K-means reduces the number of variables (in our case, SNPs) through a principal component analyses (PCA), maximizes the variation between groups and indicates the optimal number of groups by comparing different clustering solutions with Bayesian Information Criterion (BIC). In order to avoid information loss, all components of the PCA were retained. To check the concordance between both clustering methods we plotted the sNMF results as a function of the k-means classification. Clusters obtained in k-means analysis were graphically described using a discriminant analysis of principal components (DAPC; Jombart et al. 2010) available in ADEGENET 2.0 (Jombart and Ahmed 2011).

UCEs: Phylogenetic estimations and species tree

We analyzed the concatenated loci data set using two phylogenetic methods. First we performed a maximum likelihood (ML) search in RAxML v8.1.18 (Stamatakis 2014) with the GTR+G model of sequence evolution. Node support values were accessed based on

1000 rapid bootstrap replicates. Second, we implemented a Bayesian inference (BI) analysis in MrBayes v3.2 (Ronquist et al. 2012). We performed two independent runs of 2 million generations, each with 4 chains, sampling every 1000 generations, excluding the first 15% of generations as burn-in. Best-fit models of sequence evolution were determined in CloudForest (github.com/ngcrawford/CloudForest). We used random seed for starting tree and default priors and initial search values.

We also used species tree (ST) methods under the multi-species coalescent (MSC) framework. To test the robustness of alternative ST methods in face of potential gene flow among *M. lugubris* taxa, we implemented three methods that sample individual gene trees to infer the ST under the coalescent: (1) Maximum pseudo-likelihood to estimating ST (MP-EST; Liu et al. 2010); (2) ST from average ranks of coalescence (STAR; Liu et al. 2009b); and (3) Neighbor joining ST (NJst; Liu and Yu 2011), using the STRAW online platform (Shaw et al. 2013; http://bioinformatics.publichealth.uga.edu/SpeciesTreeAnalysis/index.php). We identified the best-fit molecular substitution model and estimated ML gene trees and 100 bootstrap replicates for each gene tree based on sequences of phased alleles with Cloudforest, which runs the Phyml (Guindon et al. 2010) in background (github.com/ngcrawford/CloudForest). The node support values were obtained by calculating an extended majority-rule consensus tree for the 100 ST obtained by each of the three methods. Despite the potential ascertainment bias introduced by sub sampling loci (Knowles 2010), previous studies suggested that informative loci with limited variation - as typically observed in UCEs datasets - can negatively affect results of summary-based ST methods, given the assumption of bifurcating trees (Lanier et al. 2014; Hosner et al. 2015; Manthey et al. 2016). Thus for these analyses we selected loci with more than five informative sites that represent approximately 25% of the most variable loci, which is assumed to be a good threshold following Hosner et al. (2015). Given that

the inclusion of hybrid individuals explicitly violates the MSC model, for all three methods we performed analyses with and without individuals that are morphological intermediates between *M. l. femininus* and *M. l. lugubris*.

Comparatively, we implemented the full likelihood MSC model of Bayesian Phylogenetics and Phylogeography (BP&P; Yang & Rannala 2010; Rannala & Yang 2013; Yang & Rannala 2014) to estimate ST and the probability of speciation at each node, testing our hypothesized taxa, with and without intermediate individuals between *M. l. femininus* and *M. l. lugubris*. In BP&P we ran 2x10^6 generations sampling every 5 generations with a burn-in of 10 %. To estimate divergence time (Tau) and population size (Theta) we set a gamma distribution [G(α,β)] as G(2,1000), with the prior mean = α/β; and prior variance = α/β^2. All analyses were performed twice to check for consistency between independent runs. As for the summary-based ST methods, we used the data set containing loci with five or more informative sites.

Additionally, we applied the MSC method implemented in SNAPP (Bryant et al. 2012) that is executable in BEAST v.2 (Bouckaert et al. 2014). SNAPP infers a likelihood ST using allele frequency of unlinked SNPs bypassing the need to integrate the probabilities of gene trees as a function of a given ST. We used gamma rate priors for alpha and beta parameters, with all other priors with default values. Two replicates of 2,5 million MCMC generations were run with 100,000 burn-in iterations. Estimated parameters were sampled every 500 generations. Burn-in values for the MCMC chains were accessed in Tracer v1.4 (Rambaut and Drummond 2007). Finally, using SNAPP we implemented the Bayes Factor Delimitation (BFD) of species to test which arrangement of individuals as potential species had the best Bayes factor value based on Marginal Likelihoods (Leaché et al. 2014). Models tested the potential effects of gene flow in ST analyses and the results obtained based on mtDNA (Tables 1, 2). We ran 50 steps of 100,000 MCMC generations with a pre burn-in of 10,000 MCMC generations for each arrangement of individuals. For both methods, SNAPP and BFD, we performed analyses with and without intermediate individuals between

M. l. femininus and *M. l. lugubris*. To visualize the posterior distribution of sampled species trees we used DENSITREE v2.2 (Bouckaert 2010).

UCEs: gene flow and demographic models

In order to account for gene flow while inferring the relationships among *M. lugubris* taxa we used TREEMIX 1.12 (Pickrell and Pritchard 2012). This method infers patterns of population splitting and mixing accessing the covariance structure of allele frequencies between populations and performing a Gaussian approximation for genetic drift, producing a ML graph linking sampled populations to their most common ancestor, improving the fit of the inferred model given the observed data by enabling migration edges. We performed the analyses varying from zero to four migration edges assessing the significance of each added migration edge and the residue covariance matrix to access the fit of the model to the data. Additionally, the $f3$ statistic was calculated using the three-pop test for admixture, available in the TREEMIX package (Reich et al. 2009). This test detects correlations in allele frequencies that are not compatible with population evolution following a bifurcating tree (Reich et al. 2009). All possible clusters of populations were tested.

Additionally we estimated phylogenetic networks including branch length and inheritance probabilities with the software PhyloNet 3.6.1 (Than et al. 2008). Given the relative large number of terminal taxa and number of sampled individuals we implemented the Maximum Pseudo Likelihood algorithm (Yu & Nakhleh 2015), "InferNetwork_MPL", allowing for none and two reticulations and performing 1,000 independent searches to avoid sampling in local optimums, with the other parameters set as default. For this analysis we used the bootstrapped (100 replicates) gene trees of loci with at least five informative sites. To bypass uncertain nodes in gene trees, we applied a bootstrap support threshold of 75 in all PhyloNet analyses with the -b flag.

In order to compare the topologies obtained based on mtDNA and ST/TREEMIX/PhyloNet

analyses based on UCEs and to infer the potential effects of gene flow in the diversification of *M. lugubris*, we implemented a coalescent model-based approach using Fastsimcoal2 (Excoffier et al. 2013). Fastsimcoal2 estimates the composite likelihood of a specific scenario of diversification given the observed data as well as population genetic parameters such as divergence time, effective population size and gene flow using the joint site frequency spectrum (jSFS) as input, to summarize the complexity of the data. The jSFS for each pairwise population (n=10; 5 groups) were generated in ∂a∂I 1.7 (Gutenkunst et al. 2009). For parameter estimation and model comparison we ran 50 replicates per model retaining the parameters that maximized the composite likelihood across all iterations. Parameter optimization was performed through 50 cycles of the Brent algorithm and the composite likelihood calculated using 100,000 simulations per replicate. Using the replicate that optimized the composite likelihood for each tested model we calculated the Akaike Information Criterion (AIC), AIC=2k-2ln(L), where k is the number of parameters estimated in the model and L the composite likelihood value. To obtain confidence intervals we performed 100 parametric bootstraps using the parameters of the best model to simulate 100 new sets of jSFS and re-estimated these parameters. We performed 30 independent runs for each new simulated data set running 50 cycles of the Brent algorithm and 100,000 simulations. We applied a mutation rate of 2.5 X 10^-9 substitution per site per generation (Nadachowska-Brzyska et al. 2015) and assumed a generation time of 2.33 year (Maldonado-Coelho 2012).

Results

Morphology

Plumage variation (see supplementary information [SI] 1) was congruent with the original taxonomic descriptions of each taxon, with subtle differences regarding population polymorphisms detected due to the large number of specimens analyzed in this study. In general, male plumage was conserved along the entire distribution of *M. lugubris*, despite a

tendency for darker neutral gray individuals in *M. l. berlepschi* (Fig. s1). The plumage of females was more variable and geographically diagnosable than that of males, enabling the distinction of four groups congruent with the current taxonomy (see Material and Methods; Figs. s2, s3; SI1). In general, the front and sides of the head were light ferruginous in females from the lower Amazon river (*M. l. lugubris*), whereas those body parts were black or blackish in western Amazonian subspecies (*M. l. berlepschii, M. l. femininus, M. l. stictopterus*; Zimmer & Isler 2003). Besides, females of *M. l. lugubris* from Almerim (extreme of its distribution) presented throat and central portions of the chest pure white. *M. l. femininus* females from Borba and Autazes (extreme of its distribution) showed white throat with small and subtle black spots (scale) and a subtle black spotted collar delimiting the throat and chest. Females of *M. l. berlepschi* had white throat with small and subtle black spots (scale) as observed in *M. l. femininus* and an evident collar of black spots with variable intensity delimiting the throat and chest. Finally, *M. l. stictopterus* females were very similar to *M. l. femininus* and diagnosable only by the color of the facial mask, which was larger and pure black without signs of ferruginous feathers. Despite the clear diagnosis of the females of *M. l. lugubris* and *M. l. femininus* in the extreme of their distributions (Almeirim for *M. l. lugubris* and Borba for *M. l. femininus*, localities 18 and 11 in Fig. 1a, respectively), these diagnostic characters were blurred in geographically intermediate localities forming a gradual transition between these two forms (Fig. 2).

We detected significant ($P \leq 0.05$) differences between sexes in bill width and wing length in all four taxa of *M. lugubris*. To extract the variation in the data not explained by sexual dimorphism, we performed a linear regression of each character measured as a function of sex, retaining the residues of these analyses. The DFA resulted in significant differences in the morphometric space among *M. lugubris* taxa, with the first two canonical discriminant variables accounting for 95% of the total variation (Fig. s4). The

variables that most contributed to this differentiation were tail length, wing length, and exposed culmem (Wilks's lambda = 0.1149, $P > 0.0001$; Fig. s4; Table 3). The classification matrix revealed that 68% (11 specimens) of *M. l. lugubris* individuals were correctly identified, with all five misclassified specimens placed in the morphologically intermediate group (hereafter named as intermediate group). A similar pattern was observed in *M. l. femininus*, with 80% of the individuals (24 specimens) correctly classified and six specimens placed in the intermediate group (Tables 1, 2). Among the *M. l. stictopterus* individuals, 75% (6 specimens) were correctly classified and two were misclassified as belonging to the intermediate group. All the *M. l. berlepschi* individuals (100%; 48 specimens) were correctly classified. The intermediate group (53 specimens) had the second highest correct classification index (88.67%), with 47 individuals correctly identified and only six individuals misclassified in other groups (Table 3). Thus, individuals from the intermediate group that were inferred to be hybrids between *M. l. lugubris* and *M. l. femininus* (see below) can be diagnosed by morphometric characters. The geographic distribution (from west to east) of the first canonical axis (that explains 85.4% of the total variation and expresses the size of individuals) and the three measured characters that most contributed to the observed pattern (wing length, tail length, and exposed culmem), revealed that intermediate individuals in plumage are larger than individual with "pure" plumage patterns (Figs. 2, s5). Additionally, most of the morphometric variables of the intermediate group was statistically different (Mann-Whitney test) from those of "pure" *M. l. femininus* and *M. l. lugubris* (Table s3). Considering only the plumage pattern of non-introgressed individuals, *M. l. lugubris* had significantly larger wings, tail and bill length than any of the other taxa. The opposite occured with *M. l. berlepschi*, which was significantly smaller for the same characters than the other taxa, with *M. l. stictopterus* and *M. l. femininus* widely overlapping in measurements of these characters (Tables s2, s3).

mtDNA phylogeography

Mitochondrial phylogenetic analyses supported three main reciprocally monophyletic groups within *M. lugubris*, which were congruent with three described subspecies that present distinct plumage. However, it showed paraphyly between *M. l. stictopterus* and *M. l. femininus* (Fig. 3a), which could be distinguished by plumage characters. The first split supported the separation of *M. l. berlepschi* from the ancestral population of the remaining taxa, whereas a second event split *M. l. lugubris* from *M. l. femininus* and *M. l. stictopterus.* Within *M. l. berlepschi* an internal clade with some individuals from localities 1 and 2 was observed. Within *M. l. lugubris*, two groups had high support partially clustering individuals from westernmost (locality 14) and easternmost (localities 17 and 18) localities. The paraphyly between *M. l. stictopterus* and *M. l. femininus* was due to two individuals from locality 9 (*M. l. stictopterus*) that were more closely related to *M. l. femininus* individuals than to the reciprocally monophyletic clade formed by the remaining individuals of *M. l. stictopterus*. Clustering analyses considering linked sites performed in BAPS (Corander and Marttinen 2006) supported K=5 as the best model (likelihood = -699.0878; probability for 5 populations = 0.9999), clustering individuals congruently with the phylogenetic analyses and splitting *M. l. berlepschi* in two clusters with partial geographic structure (Fig. 3b). The gradual transition between the phenotypes of *M. l. femininus* and *M. l. lugubris* detected in morphometric and plumage analyses was not observed in BAPS results, suggesting an abrupt transition between these two taxa. Only specimens from Parintins (locality n° 14 in Fig. 1a, Fig. 3b) included individuals of the two clades occurring in sympatry. However the individual of locality 15, eastwards from Parintins, clustered with *M. l. femininus* individuals.

UCEs – Summary information and genetic structure

Data processing in PHYLUCE produced a concatenated matrix with 15% missing data with

2,151 loci. The average locus size was 549 bp ranging from 259 to 763 bp. For the SNP calling procedure we were able to obtain a reference for 2,378 loci without cleaning for missing or trimming long edges. Data filtering for loci with five or more informative sites between all individuals of *M. lugubris* produced a data set with 473 loci with average size of 813 bp ranging from 330 (ARNTL-exon13) to 2092 bp (RAG2) with an average of 6.57 informative sites per locus (range of 5 to 21 informative sites per locus). The selection of a single variant per locus without missing data resulted in a matrix with 1,664 SNPs. Overall mean sequencing coverage of SNPs among individuals was 29.3 X varying from 20.9 to 46.3 X.

The population structure obtained in sNMF was concordant with the plumage variation, with the best value of cross entropy (alpha parameter = 10; cross-entropy = 0.513584) supporting the presence of four ancestral populations (K=4) that matched the geographical distributions of *M. l. berlepschi, M. l. stictopterus, M. l. femininus* and *M. l. lugubris* (Fig. 1b). The pattern of individuals with intermediate morphology between *M. l. femininus* and *M. l. lugubris* was congruent with the result that individuals of localities 13 (Itacoatiara) and 14 (Parintins) presented shared ancestry (Figs. 1a, b). The k-means and DAPC analyses recovered the same number of clusters as sNMF (K=4; Fig. s6). The only discordance between these two analyses was in the classification of the morphologically intermediate individuals from Parintins (locality 14 in Fig. 1a) as the sNMF analyses showed a higher coefficient of ancestry with *M. l. lugubris* while in the classification matrix of the DAPC analysis they were classified as *M. l. femininus*. The second best model estimated in sNMF supported the presence of three ancestral populations (K=3), placing *M. l. femininus* individuals as intermediates between *M. l. lugubris* and *M. l. stictopterus* (alpha parameter = 10; cross-entropy = 0.521805; Fig. 1c).

UCEs – Phylogenetic estimates, species tree and gene flow

The RAxML and MrBayes phylogenies based on concatenated loci supported similar topologies that suggested the monophyly of *M. lugubris* and *M. leucophrys* as its sister

species, and revealed at least three main clades within *M. lugubris* (Fig. s7). These relationships were partially congruent with currently recognized subspecies. The first split within *M. lugubris* separated *M. l. berlepschi*, which occurs in western Amazon (Solimões river) from the remaining populations. The second split separated *M. l. stictopterus* (Negro - Branco rivers) from a clade including individuals assigned to *M. l. femininus* and *M. l. lugubris* (Madeira and Amazon rivers). This latter clade was not recovered by mtDNA results (Fig. 3). Furthermore, internal relationships within this clade were generally poorly supported, with two exceptions: 1) easternmost individuals of *M. l. lugubris* (localities 17 and 18; Fig. 1a) clustered in a well supported clade; and 2) some specimens of *M. l. femininus* and morphological intermediates from the westernmost part of *M. l. lugubris* range (localities 11 and 13; Fig. 1a) also grouped in a well-supported clade. The position of individuals from localities 12 and 14 is congruent with a steeping stone pattern between these two groups, which is in agreement with the presence of intermediate plumage patterns in these localities (Figure s2). Assuming that the sNMF clusters (K=4) as taxa, the species tree estimation methods (Njst, Star and MP-EST) recovered topologies with all nodes having bootstrap support < 0.75, except for the node clustering all individuals of *M. lugubris* (bootstrap = 100; Fig. 4a). When we considered the intermediate individuals as a separate lineage, in all three methods these intermediate individuals formed a clade with *M. l. lugubris* (bootstrap > 99; Fig. 4). Only in Njst the relationship of *M. l. femininus* with the intermediate individuals and *M. l. lugubris* was marginally supported (bootstrap = 73). The full likelihood approach of BP&P produced a totally resolved species tree with maximum posterior probability for all taxa, with and without intermediate individuals (Fig. 4). This topology showed a first split between *M. l. berlepschi* and eastern Amazon taxa, a second split between *M. l. stictopterus* and the remaining individuals, a third split separating *M. l. femninius*, and a fourth split (only in the analyses with five taxa) between

the intermediate individuals and *M. l. lugubris* (Fig. 4).

The species trees estimated with SNAPP, assuming the same set of groups (four and five taxa), were identical with the trees obtained with BP&P. The posterior distribution of sampled species tree in DensiTree suggested a well-resolved relationship among the recognized populations with posterior probability of 1.0 for all estimated nodes, except for the relationship between *M. l. lugubris* and *M. l. femininus* (PP=0.98) in the four taxa analysis (Fig. 5). The Bayes Factor Delimitation (BFD) results for alternative assignments of individuals to taxa supported with the highest marginal likelihood and a high value of Bayes factor (four taxa analysis, Marginal likelihood = - 1770.50, ln(BF) > 5; five taxa analysis, Marginal likelihood = -22095.83, ln(BF) > 5) indicated the same groups used for the SNAPP analyses as the best arrangement of individuals as independent evolutionary lineages (Tables 1, 2).

The TREEMIX topology without migration edges was identical to the SNAPP topology assuming five taxa (Fig. s8; likelihood = 131.558). However, the model that best fit the data supported the presence of two migration edges (Fig. 6; likelihood = 135.843). The first took place between *M. l. lugubris* and *M. leucophrys* (the sister taxon to the polytypic *M. lugubris*), which was not significant (Jacknife estimate *P*-value = 0.1606), and the second between the intermediate individuals and *M. l. femininus*, which was statistically significant (*P*-value = 0.0006). This latter relationship was the only admixture event supported by the *f*3 statistic, with a highly negative *Z-score* (*M. l. femininus; M. l. stictopterus*, morphologically intermediate individuals; *Zscore* = -6.87209). The presence of this significant migration edge affected the inferred tree topologies, suggesting a sister relationship between *M. l. femininus* and *M. l. stictopterus* (Fig. 6) as observed in the mtDNA results.

The Maximum Pseudo Likelihood algorithm of PhyloNet with introgression edges set to zero recovered similar relationships obtained by BP&P and SNAPP, clustering *M. l. femininus, M. l. lugubris* and intermediate individuals as a clade (likelihood (ln(lik)= -

739516.4899; Fig. 7). The PhyloNet runs allowing for 2 introgression edges supported a topology similar to those obtained by TREEMIX and mtDNA, with *M. l. femininus* and *M. l. stictopterus* as sister taxa in the two networks with the highest likelihood (Fig. 7). The inheritance probability of the two best trees were higher for *M. l. lugubris* (γ=0.70 and γ=0.97 in the first edge and γ=0.97 and γ=0.80 in the second edge for the first and second best runs, respectively) than for *M. l. stictopterus/M. l. femininus* (γ=0.30 and γ=0.03 in the first edge and γ=0.03 and γ=0.20 in the second edge for the first and second best runs, respectively) with a higher contribution of *M. l. stictopterus* than *M. l. femininus* in both estimations. The three remaining runs supported the same relationship with zero migration edges.

To assess the potential impacts of gene flow in species tree estimation, we used Fastsimcoal2 (Excoffier et al. 2013) to test the fit of demographic models concerning the different topologies with and without post-divergence gene flow. We tested four demographic models (Fig. 8): A) based on the SNAPP topology, with no gene flow among lineages; B) based on the mtDNA/TREEMIX topologies, also with no gene flow among lineages and considering the intermediate individuals used in SNAPP as a distinct lineage sister to *M. l. lugubris*; C) based on the BP&P/SNAPP topologies, but considering the intermediate individuals as a hybrid population between *M. l. femininus* and *M. l. lugubris* with gene flow among *M. l. femininus*, intermediate individuals and *M. l. lugubris*; D) based on the mtDNA/TREEMIX topology, with gene flow among *M. l. femininus*, intermediate individuals and *M. l. lugubris*, and assuming that the intermediate individuals form a hybrid population between *M. l. femininus* and *M. l. lugubris*. The results supported a maximum relative contribution, based on AIC values (1.0), for model D which assumed the mtDNA/TREEMIX topology and a hybrid population between *M. l. femininus* and *M. l. lugubris* (Tables 4, 5; Fig. 8d). Parameter estimation and confidence intervals obtained through bootstrap replicates supported relatively large current effective population sizes

for the four *M. lugubris* taxa but a reduced effective size of hybrids (Table 4). The diversification time estimates supported the split of *M. l. berlepschi* from the remaining taxa at approximately 1.5 Ma, the split between *M. l. lugubris* from *M. l. femininus* and *M. l. stictopterus* at approximately 1.4 Ma, the split between *M. l femininus* and *M. l. stictopterus* at approximately 0.35 Ma and the beginning of the introgression between *M. l. lugubris* and *M. l. femininus* at approximately 0.14 Ma (Table 4). Gene flow estimates supported an asymmetric pattern and relative high number of migrants per generation from *M. l. femininus* into the morphologically intermediate population (4.90 – 95% CI = 4.59-5.01) and from *M. l. lugubris* into *M. l. femininus* (0.91 – 95%CI = 0.97-1.32; Table 4).

Discussion

Phylogenetic inference and species tree discordances

Our study indicated that failing to account for gene flow in species tree estimation can affect topology in cases of recent diversification. In addition, our results showed that the use of phenotypic information can help detect the effects of introgression in wild populations and in the interpretation of species tree (ST) analyses (Edwards and Knowles 2014).

The ST methods applied here have been widely used to reconstruct shallow diversification scenarios that could be affected by gene flow between populations assumed to be isolated. In these cases the implicit assumption that ILS solely generated the observed genealogical patterns is restrictive and can bias topology, branch length and effective population size estimations if considerable levels of gene flow are present (Leaché et al. 2014; Solís-Lemus et al. 2016). Here we detected admixture among non-sister lineages, which would have exacerbated this bias and increased the complexity underlying the evolutionary relationship among lineages. Summary-based ST methods that use individual gene trees as inputs have shown accurate results under low levels of gene flow, although higher levels of gene flow increase gene tree heterogeneity promoting loss

of resolution and weak node support in bootstrap replicates (Liu et al. 2009a; Solís-Lemus et al. 2016). Solís-Lemus et al. (2016) reported anomalous results based on concatenated data sets and using two summary-based ST methods (STAR and Njst) when more than 30% ($\gamma > 0.3$) of loci were inherited through reticulation, even with high numbers of markers such as 10,000 simulated genes. The strong effects of gene flow reported by Solís-Lemus et al. (2016) is in agreement with our results that showed inconsistencies among results of ST methods that do not account for gene flow (Fig. 4, 5) and those of PhyloNet, whose model with the highest likelihood supported values higher than 0.3 (Fig. 7).

On the other hand, little is known about how ST methods such as SNAPP are affected by high levels of gene flow. Although Rheindt et al. (2014) suggested that SNAPP is robust with limited gene flow between populations of *Zimmerius* flycatchers, here we obtained full statistical support for topologies indicating to a distinct relationship between populations that was not consistent with results derived from our mtDNA phylogeny, phenotypic data, TREEMIX, PhyloNet as well as the best-fit model recovered with fastsimcoal2. The SNAPP algorithm uses SNP allele frequencies assuming that the probability of an allele frequency at a given locus is correlated with the probability of a site in a gene tree multiplied by the gene tree probability given a ST (Bryant et al. 2012; Leaché et al. 2014). Thus it is possible that a hybrid population, with a unique pattern of allele frequencies (intermediary allele frequency), tends to be weighted as an independent lineage by SNAPP and BFD methods. The results presented here are in agreement with Sukumaran and Knowles (2017) that suggested that MSC methods can produce misleading results when used to directly infer species limits in face of relative high gene flow but are useful tools to estimate phylogenetic relationships and genetic structure if their main assumptions are preserved.

The phylogenetic estimate obtained based on mtDNA and the model-based approach implemented on fastsimcoal2 are partially congruent with the morphological data. The

most readily diagnosable taxon is *M. l. berlepschi* given its significantly smaller size and its characteristic female plumage, and it is also the first population to diverge within *M. lugubris*. The second split within *M. lugubris* separated the taxa of central Amazonia (*M. l. femininus* and *M. l. stictopterus*) from *M. l. lugubris* of the lower Amazon river, which is concordant with the black-faced pattern and the overlap in morphometric variables observed in *M. l. femininus* and *M. l. stictopterus*. However, the pattern of gradual transition in plumage characters, size and the coefficient of ancestry of UCEs observed between *M. l. femininus* and *M. l. lugubris* was not observed in mtDNA, which supported an abrupt transition between these taxa with only Parintins (locality n° 14 in Fig. 1a) presenting individuals from both clades (Fig. 3). The absence of mtDNA introgression could be explained by the matrilineal inheritance without recombination of this marker (Carling and Brumfield 2008) in audition to a recent secondary contact, as suggested by fastsimcoal2 (Table 4) or even male biased dispersion. However, given the biology of antbirds, including territorial defense by both sexes (Zimmer & Isler 2003), the latter hypothesis seems to be less likely.

Phenotypes and hybridization

In scenarios of recent diversification, the use of multiple types of evidence - phenotypic data such as plumage, morphometric, vocal, ecological and physiological traits - can provide more accurate estimates of and a better understanding of demographic processes than genetic data alone (Edwards and Knowles 2014; Solís-Lemus et al. 2015; Zamudio et al. 2016). The results obtained here revealed high levels of introgression between *M. l. lugubris* and *M. l. femininus* (Table 4; Figs. 1b, c, 2, s2, s4), with a wide hybrid zone at least 500 km long and with a clear pattern of transition in female plumage and morphometric measurements (tail length, wing length and exposed culmem length). The hypothesis of Haffer and Fitzpatrick (1985) that the central Amazonian taxa could be

intermediate between the two extreme plumage forms as a result of secondary intergradation of *M. l. berlepschi* and *M. l. lugubris* was not corroborated, since we observed secondary intergradation only between *M. l. femininus* and *M. l. lugubris* and a clear diagnosis of *M. l. berlepschi* that seems restricted to the Solimões river. Size variation over the geographic distribution of *M. lugubris* was first noted by Hellmayr (1910) who, when describing *M. l. femininus* and *M. l. berlepschi*, noticed the smaller size of western populations. The significantly larger size of hybrid individuals from localities 14 to 17 (Figs. 1a, 2) in most of the morphometric characters measured (Table s3) suggest that the secondary contact and introgression between *M. l. lugubris* and *M. l. femininus* could result on heterosis. However, despite the general larger size of the morphologically intermediate population, more detailed studies must be performed to better understand the potential effects of hybridization on the fitness of these individuals compared to that of the parental forms.

Diversification of the Ash-breasted Antbird over the Amazonian floodplains

Large Amazonian rivers have been historically recognized as effective barriers to gene flow in upland forest (*terra-firme*) species, delimiting their geographic distributions (Cracraft 1985; Antonelli et al. 2010; Smith et al. 2014). However, this well documented pattern is not expected to be observed in organisms occurring along riverine environments. Phylogeographic studies of non-aquatic bird species associated to floodplains did not find population structure throughout the Amazon basin, suggesting that the linear distribution and connectivity of floodplains enable high levels of gene flow (Aleixo 2006; Cadena et al. 2011; but see Choueri et al 2017). Furthermore, species occurring in these environments are expected to be good dispersers due to the constant effect of flooding cycles and sedimentation dynamics, which affect the distribution of river created environments (Aleixo 2006). Against these patterns and assumptions, our study supported the presence of four structured

populations within the ash-breasted antbird complex geographically associated with main Amazonian rivers (Solimões, Madeira, Negro and Amazonas). This hitherto uncommon pattern suggests that despite the apparent absence of modern physical barriers, historical processes could have promoted opportunities of isolation along floodplain forests, mainly in organisms such as *M. lugubris*, which is restricted to specific environments of river edge forests and islands (Rosenberg 1990).

The parameter estimates obtained with fastsimcoal2 supported a recent scenario of diversification starting at the mid-Pleistocene with the split between *M. l. berlepschi* and the remaining taxa followed by an almost simultaneous diversification of *M. l. lugubris* and the two taxa of central Amazonia (Table 4). These results indicate a scenario where paleoclimatic fluctuations during the mid and late Pleistocene potentially interrupted the distribution and availability of river-created environments (Latrubesse and Franzinelli 2005; Irion et al. 2009; Rossetti et al. 2015). Similar results were obtained by Choueri et al. (2017) studying landscape genetics of four antbird species - including *M. lugubris* – along the Negro river archipelagos, suggesting that populations restricted to this river diverged from the Amazon and Solimões populations during the late Pleistocene. During this time span, local river channel incisions occurred at the main course of the Solimões river and tributaries, with floodplain environments being replaced by upland forest habitats (Nogueira et al. 2013; Rossetti et al. 2015). This process could have fragmented riverine environments, specially in central Amazon where the sedimentary basin of the Solimões runs through hard rock substrates of the Amazon Craton (Choueri et al. 2017).

River incision cyclically alternated with sediment accumulation periods with the expansion of floodplains environments and potentially connection of previously isolated areas. Although these cycles could be linked to sea level fluctuations related to the intensification of glacial/interglacial periods during the mid-Pleistocene in the lower Amazon river (Irion et al. 2009; Bertani et al. 2015), areas from the central and western

Amazon Basin are more likely to be influenced by paleoclimatic and tectonic processes occurring in source areas of sediments, increasing water discharge and sediment load (Choueri et al. 2017). This intense dynamic cycle of sediment accumulation and erosion could have originated the hybrid zone between *M. l. lugubris* and *M. l. femininus* around 0.14 Ma. Additionally, during glacial cycles, large lakes (the so called Ria lakes) were formed in the lower courses of rivers that carry low amounts of sediments, such as the Negro and Tapajós rivers (Irion et al. 2009; Bertani et al. 2015). The Ria lakes formation produced large areas without islands and with poorly developed floodplains on river banks (Franzinelli and Igreja 2002; Irion et al. 2009), potentially acting as barrier for *M. lugubris* taxa, explaining the possible absence of *M. l. stictopterus* in the lower course of the Negro river. However, despite the clear pattern of genetic structure and the recent diversification scenario reported here, comparative phylogeographic studies including large sample sizes and estimates of population size changes over time, have the potential to reveal a much more refined pattern of diversification of organisms associated with Amazonian floodplain forests.

The first diversification event observed in *M. lugubris* separating *M. l. berlepschi* of the upper Amazon (Solimões) river from the eastern taxa of the Negro, Madeira and lower Amazon rivers is in accordance with studies reporting abrupt transitions in community composition of trees, spiders, fishes and birds (Hubert and Renno 2006; Albernaz et al. 2012; Cohn-Haft et al 2007), as well as sedimentation patterns and river dynamics (Mertes et al. 1996), along a strong longitudinal biogeographic gradient encompassing floodplain forests of the entire course of the Amazon river and its main tributaries. However, few phylogeographic studies on populations restricted to the floodplains have been conducted so far to evaluate whether this biogeographic pattern is mirrored by the genetic structure of populations along this region (Choueri et al. 2017). Farias and Hrbek (2008), based on mtDNA data, recovered reciprocally monophyletic groups in discus fishes of the genus *Symphysodon* from the

Solimões and Amazon rivers, suggesting that this diversification process was mediated by the breach of Purus Arch during the Pliocene and by the different chemical composition of these rivers. Despite the similarities in the geographic distributions of the taxa studied here and the ones of the *Symphysodon* clades identified by Farias and Hrbek (2008), the more recent divergence times reported in *M. lugubris* suggests that distinct historical processes shaped the current pattern of genetic diversity in these two groups.

Another potentially important process of diversification that we cannot rule out as an explanation for the results we obtained, is the presence of ecological gradients with distinct selective pressures in specific regions of the Amazonian floodplains. For example, the transitions between white water rivers (with high sediment load) and black water rivers (with low sediment load) delimit distinct vegetation types and affect the pattern of genetic diversity in several species of fishes (Cooke et al. 2014).

Species limits and taxonomy

Here we showed significant levels of differentiation among subspecies of *M. lugubris* in phenotypic and genetic characters which, when interpreted together, allow for a redefinition of species limits within this complex. The clear diagnosis of *M. l. berlepschi* by all morphological and genetic analyses, including the absence of gene flow between this taxon and the remaining ones grouped under *M. lugubris* provide the basis for considering it a separate species under distinct species concepts including the Biological (BSC; Mayr 1942; Queiroz 2007) and the Phylogenetic ones (PSC; Cracraft 1983; Queiroz 2007). However the observed pattern does not guarantee total reproductive isolation given their apparent allopatric distribution. The high support obtained for the reciprocal monophyly of *M. l. stictopterus* by most phylogenetic estimates based on UCEs, was consistent with its apparent allopatric distribution and differences in the coloration of the face (pure black in *M. l. stictopterus* without sparse reddish feather as in *M. l. femininus*) thereby supporting a clear diagnosis for these taxa that could be

recognized as valid species under distinct species concepts such as the PSC (Queiroz 2007).

The hybridization pattern observed between *M. l. lugubris* and *M. l. femininus* indicates a complex scenario with multiple taxonomic interpretations including: 1) the clearly diagnosable parental forms and the heterosis pattern among hybrids suggest secondary contact and post zygotic incompatibilities between these taxa that could be recognized as valid species under the PSC (Queiroz 2007); 2) the gene flow pattern with a hybrid zone geographically as large as the distribution of the parental forms could suggest an ephemeral speciation with fusion of nonsister lineages not supporting species status for these taxa. Thus, studies with denser sampling over the hybrid zone to better explore the clinal variation, as well as studies focusing on hybrid fitness, must be performed to better explore the genetic cohesion of these two taxa.

Considerations when accounting for gene flow in phylogenetic reconstructions

The importance of MSC methods arose with the assumption that ILS can produce discordance among gene trees, which cannot be modeled with concatenated markers. Moreover, methods that explicitly incorporate gene flow showed that ILS is not the only source of inconsistencies among gene trees. The increasing amount of evidence for reticulated evolution reinforces this demand (Mallet et al. 2016; Edwards et al. 2016). Phylogenetic inferences are the starting analyses of phylogeographic studies that require a strong phylogenetic hypothesis to explore complex demographic parameters. Commonly used methods to estimate demographic parameters such as Ima2 (Hey & Nielsen 2007) and G-Phocs (Gronau et al. 2011) demand a priori topology for the estimation of demographic parameters, including effective population size, gene flow and divergence times. Under this perspective, network estimation methods that explicitly accept gene flow into species tree estimation can be important tools to understand complex patterns of diversification among relatively large species or population complexes (Wen et al. 2016). Similarly, model-based approaches such as the composite likelihood based method implemented in fastsimcoal2 or Approximate Bayesian Computation can provide robust parameter estimations to test competing phylogenetic

hypotheses (Robinson et al. 2014; Martin et al. 2015; Jackson et al. 2017). However, these methods do not allow for exhaustive tree search, meaning that robust alternative hypotheses need to be constructed based on complementary analyses (Nater et al. 2015).

In recent years, a relatively large number of analytical methods were proposed to accommodate both ILS and gene flow in phylogenetic inferences (Than et al. 2008; Kubatko 2009; Jackson et al. 2017; Solís-Lemus & Ané 2016). As examples, there are fast and reliable algorithms such as the Maximum Pseudo Likelihood of PhyloNet (Yu & Nakhleh 2015) and SNaQ (Solís-Lemus & Ané 2016) and a model selection framework that allows to test the fit of hundreds of models to varying topologies and gene flow direction as implemented in the software PHRAPL (Jackson et al. 2017). Thereby, given the relative large availability of methods and the results presented here and elsewhere (Morales et al. 2017; Meyer et al. 2016), studies involving phylogenetic estimations of recently diverged taxa in the gray zone of phylogeography must consider the effects of gene flow by applying such alternative methods. Additionally, the application of these alternative methods can help to detect errors in phylogenetic estimations (Knowles & Kubatko 2010) and incongruent topologies can be later on explored by model testing approaches.

Funding

This study was co-funded by FAPESP (BIOTA, 2012/50260-6, 2013/50297-0), NSF (DOB 1343578), NASA, CNPq (310593/2009-3, 574008/2008-0, 563236/2010-8, 471342/ 2011-4), and FAPESPA (ICAAF 023/2011). GT's Ph.D scholarships were granted by CAPES and then FAPESP (2014/00113-2, 2015/12551-7). AA, CCR, and CYM are supported by CNPq research productivity fellowships (310880/2012-2, 458311/2013-8, 303713/2015-1). FA was funded by FAPESP (2011/50143-7, 2011/23155-4).

Acknowledgements

We thank the curator and curatorial assistants of the Instituto Nacional de Pesquisas da

Amazônia (INPA); and Museu Paraense Emílio Goeldi (MPEG) for allowing us to analyze study specimens under their care. We thank members of the Hickerson lab, more specifically Isaac Overcast and Alexander Xue for the discussion related to the use of model-based approaches. 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. We thank the three anonymous reviewers and the associate editor for the suggestions and comments that greatly improved this publication

Literature Cited

Andrews S. 2014. FastQC: A quality control tool for high-throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ accessed March 9, 2017.

Antonelli A., Quijada-Mascareñas A., Crawford A.J., Bates J.M., Velazco P.M., Wüster W. 2010. Molecular studies and phylogeography of Amazonian tetrapods and their relation to geological and climatic models. in: Hoorn, C. & Wesselingh, F. Editors. Amazonia: Landscape and species evolution, 1st ed. Oxford: Wiley-Blackwell. p. 386–404.

Albernaz A.L., Pressey R.L., Costa L.R.F., Moreira M.P., Ramos J.F., Assunção P. A., Franciscon C.H. 2012. Tree species compositional change and conservation implications in the white-water flooded forests of the Brazilian Amazon. J. Biogeogr. 39:869–883.

Aleixo A. 2006. Historical diversification of floodplain forest specialist species in the Amazon: A case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). Biol. J. Linn. Soc. 89:383–395.

Alexander D.H., Novembre J., Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19:1655–1664.

Bagley J.C., Mayden R.L., Roe K.J., Holznagel W., Harris P.M. 2011. Congeneric phylogeographical sampling reveals polyphyly and novel biodiversity within black basses (Centrarchidae: Micropterus). Biol. J. Linn. Soc. 104:346–363.

Bapteste E., van Iersel L., Janke A., Kelchner S., Kelk S., McInerney J.O., Morrison D.A.,

Nakhleh L., Steel M., Stougie L., Whitfield J. 2013. Networks: Expanding evolutionary thinking. Trends Genet. 29:439–441.

Barley A.J., Brown J.M., Thomson R.C. 2017. Impact of model violation on the inference of species boundaries under the multispecies coalescent. Syst. Biol. 0:1-17.

Bertani T.C., Rossetti D.F., Hayakawa E.H., Cohen, M.C.L. 2015. Understanding Amazonian fluvial rias based on a Late Pleistocene-Holocene analog. Earth Surf. Process. Landf. 40:285-292.

Bouckaert R., Heled J., Khnert D., Vaughan T., Wu C.H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: A software platform for Bayesian Evolutionary Analysis. PLoS Comput. Biol. 10:1–6.

Bouckaert R.R. 2010. DensiTree: Making sense of sets of phylogenetic trees. Bioinformatics. 26:1372–1373.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N. A., Roychoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Cadena C.D., Gutiérrez-Pinto N., Dávila N., Chesser R.T. 2011. No population genetic structure in a widespread aquatic songbird from the Neotropics. Mol. Phylogenet. Evol. 58:540–545.

Carling M.D., Brumfield R.T. 2008. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. Evolution. 62:2600–2615.

Choueri E., Gubill C., Borges S., Thom G., Sawakuchi A., Soares E., Ribas C. 2017. Phylogeography and population dynamics of Antbirds (Thamnophilidae from Amazonian fluvial islands. J. Biogeogr. 10.1111/jbi.13042

Cohn-Haft M., Naka L.N. F.A.M. 2007. Padrões de distribuição da avifauna da várzea dos

rios Solimões- Amazonas. In: Albernaz A.L., editor. Conservação da Várzea, Identificação e Caracterização de Regiões Biogeográficas. Manaus: IBAMA/ ProVárzea/INPA. p. 287–324.

Cooke G.M., Landguth E.L., Beheregaray L.B. 2014. Riverscape genetics identifies replicated ecological divergence across an Amazonian ecotone. Evolution 68:1947-1960.

Corander J., Marttinen P. 2006. Bayesian identification of admixture events using multilocus molecular markers. Mol. Ecol. 15:2833–43.

Cracraft J. 1983. Species concept and speciation analysis. Curr. Ornithol. 1:159-187 Cracraft J. 1985. Historical biogeography and patterns of differentiation within the

South Amercian avifauna: areas of endemism. Ornithol. Monogr. 36:49–84. Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker

R.E., Lunter G., Marth G.T., Sherry S.T., McVean G., Durbin R. 2011. The variant call format and VCFtools. Bioinformatics. 27:2156–2158.

Edwards D.L., Knowles L.L. 2014. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? Proc. R. Soc. B. 281:20132765.

Edwards S. V., Potter S., Schmitt C.J., Bragg J.G., Moritz C. 2016. Reticulation, divergence, and the phylogeography–phylogenetics continuum. Proc. Natl. Acad. Sci. USA.113:8025–8032.

Eckert A., Carstens B. C. 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. Mol. Phylogenet. Evol. 49:832-842.

Excoffier L., Dupanloup I., Huerta-Sanchez E., Sousa V.C., Foll M. 2013. Robust demographic inference from genomic and SNP data. PLoS Genet. 9:e1003905.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61:717–26.

Farias I.P., Hrbek T. 2008. Patterns of diversification in the discus fishes (*Symphysodon*

spp. Cichlidae) of the Amazon basin. Mol. Phylogenet. Evol. 49:32–43.

Feder J.L., Egan S.P., Nosil P. 2012. The genomics of speciation-with-gene-flow. Trends Genet. 28:342–350.

Franzinelli E., Igreja H. 2002. Modern sedimentation in the lower Negro River, Amazonas State, Brazil. Geomorphology. 44:259–271.

Frichot E., Mathieu F., Trouillon T., Bouchard G., François O. 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics. 196:973–983.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnol. 29:644–52.

Gronau I., Hubisz M. J., Gulko B., Danko C. G., Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nature Genet. 43:1031-1034.

Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59:307–321.

Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. 5:1–11.

Hackett S.J., Kimball R.T., Reddy S., Bowie R.C.K., Braun E.L., Braun M.J., Chojnowski J.L., Cox W.A., Han K.-L., Harshman J., Huddleston C.J., Marks B.D., Miglia K.J., Moore W.S., Sheldon F.H., Steadman D.W., Witt C.C., Yuri T. 2008. A Phylogenomic Study of Birds Reveals Their Evolutionary History. Science. 320:1763–1768.

Haffer J. 1969. Speciation in Amazonian forest birds. Science 165:131-137.

Haffer J., Fitzpatrick J.W. 1985. Geographic variation in some Amazonian forest birds.

Ornithol. Monogr. 36:147–168.

Harris S.E., O'Neill R.J., Munshi-South J. 2015. Transcriptome resources for the white-footed mouse (Peromyscus leucopus): new genomic tools for investigating ecologically divergent urban and rural populations. Mol. Ecol. Resour. 15:382-394.

Harrison R.G. 1990. Hybrid zones: windows on evolutionary process. Oxf. Surv. Evol. Biol. 7:69–128

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

Hellmayr C.E. 1910. The birds of the Rio Madeira. Novit. Zoolog. 17:257-428. Hellmayr C.E. 1929. A contribution to the ornithology of Northeastern Brazil. Field Mus. Nat. His. Zoo. Ser. 12:235500.

Hewitt G.M. 2004. Genetic consequences of climatic changes in the Quaternary. Phil. Trans. R. Soc. Lond. B 359:183–195.

Hey J., Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. USA. 104:2785–2790.

Hickerson M.J., Meyer C.P., Moritz C. 2006. DNA barcoding will often fail to discover new animal species over broad parameter space. Syst. Biol. 55:729-739.

Hoorn C., Wesselingh F.P., ter Steege H., Bermudez M. A, Mora A, Sevink J., Sanmartín I., Sanchez-Meseguer A, Anderson C.L., Figueiredo J.P., Jaramillo C., Riff D., Negri F.R., Hooghiemstra H., Lundberg J., Stadler T., Särkinen T., Antonelli A. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science. 330:927–931.

Hosner P., Faircloth B.C., Glenn T.C., Braun E.L., Kimball R.T. 2015. Avoiding missing data biases in phylogenomic inference: An empirical study in the landfowl (Aves:

Galliformes). Mol. Biol. Evol. 33:1110-1125.

del Hoyo J., Elliott A., Sargatal J., Christie D.A., de Juana E. 2017. Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona. (retrieved from http://www.hbw.com/ on 01/10/2018).

Hubert N., Renno J.F. 2006. Historical biogeography of South American freshwater fishes. J. Biogeogr. 33:1414–1436.

Irion G., Müller J., Morais J.O., Keim G., de Mello J.N., Junk W.J. 2009. The impact of Quaternary sea level changes on the evolution of the Amazonian lowland. Hydrol. Process. 23:3168–3172.

Isler M.L., Bravo G.A., Brumfield R.T. 2013. Taxonomic revision of *Myrmeciza* (Aves: Passeriformes: Thamnophilidae) into 12 genera based on phylogenetic, morphological, behavioral, and ecological data. Zootaxa. 3717:469–497.

Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species delimitation with gene flow. Syst. Biol. 66:799-812

Jombart T., Ahmed I. 2011. Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics. 27:3070–3071.

Jombart T., Devillard S., Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11:94.

Junk W.J., Piedade M.T.F., Schöngart J., Cohn-Haft M., Adeney J.M., Wittmann F. 2011. A Classification of Major Naturally-Occurring Amazonian Lowland Wetlands. Wetlands. 31:623–640.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software versionImprovements in performance and usability. Mol. Biol. Evol. 30:772–780. Kimball R.T., Braun E.L., Barker F.K., Bowie R.C., Braun M.J., Chojnowski J.L.,

Hackett S.J., Han K.-L., Harshman J., Heimer-Torres V. 2009. A well-tested set of primers to amplify regions spread across the avian genome. Mol. Phyl. Evol. 50:654–660.

Knowles L.L., Kubatko L.S. eds. 2010. Estimating Species Trees: Practical and Theoretical Aspects. Wiley Blackwell, Hoboken.

Kubatko L.S. 2009. Identifying hybridization events in the presence of coalescence via model selection. Syst. Biol. 58:478–488.

Kubatko L.S., Degnan J.H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56:17–24.

Lamichhaney S., Berglund J., Almén M.S., Maqbool K., Grabherr M., Martinez-Barrio A., Promerová M., Rubin C.J., Wang C., Zamani N., Grant B.R., Grant P.R., Webster M.T., Andersson L. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature. 518:371-375.

Lanier H.C., Huang H., Knowles L.L. 2014. How low can you go? The effects of mutation rate on the accuracy of species-tree estimation. Mol. Phylogenet. Evol. 70:112–119.

Latrubesse E.M., Franzinelli E. 2005. The late Quaternary evolution of the Negro River, Amazon, Brazil: Implications for island and floodplain formation in large anabranching tropical systems. Geomorphology. 70:372–397.

Leaché A.D., Fujita M.K., Minin V.N., Bouckaert R.R. 2014. Species delimitation using genome-wide SNP data. Syst. Biol. 63:534–542.

Leaché A.D., Fujita M.K. 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. Biol. Sci. 277:3071–3077.

Leaché A.D., Harris R.B., Rannala B., Yang Z. 2014. The influence of gene flow on species tree estimation: A simulation study. Syst. Biol. 63:17–30.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25:1754–1760.

Liu L., Yu L., Edwards S. V. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol. Biol. 10:302.

Liu L., Yu L., Kubatko L., Pearl D.K., Edwards S. V. 2009a. Coalescent methods for

estimating phylogenetic trees. Mol. Phylogenet. Evol. 53:320–328.

Liu L., Yu L., Pearl D.K., Edwards S. V. 2009b. Estimating species phylogenies using coalescence times among sequences. Syst. Biol. 58:468–77.

Liu L., Yu L. 2011. Estimating species trees from unrooted gene trees. Syst. Biol. 60:661–667.

Maldonado-Coelho M. 2012. Climatic oscillations shape the phylogeographical structure of Atlantic Forest fire-eyes (Aves: Thamnophilidae). Biol. J. Linn. Soc. 105:900–924.

Mallet J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20:29– 37.

Mallet J. 2007. Hybrid speciation. Nature. 446:279–283.

Mallet J., Besansky N., Hahn M.W. 2016. How reticulated are species? BioEssays.

38:140–149.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: A test in cardinalid tanagers (Aves, Genus: *Piranga*). Syst. Biol. 65:640–650.

Martin S.H., Eriksson A., Kozak, K.M., Manica A., Jiggins C.D. 2015. Speciation in *Heliconius* butterflies: minimal contact followed by millions of generations of hybridization. BioRxiv, doi: https://doi.org/10.1101/015800.

Mayr E. 1942. Systematics and the origin of species. Columbia University Press, New York.

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

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo M.A. 2010. The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303.

Mertes L. A. K., Dunne T., Martinelli L. A. 1996. Channel-floodplain geomorphology along the Solimões-Amazon River, Brazil. Geol. Soc. Am. Bull. 108:1089–1107.

Meyer B.S., Matschiner M., Salzburger W. 2016. Disentangling incomplete lineage sorting and introgression to refine species-tree estimates for lake Tanganyika cichlid fishes. Syst. Biol. 66:531-550.

Mirarab S., Reaz R., Bayzid M.S., Zimmermann T., S. Swenson M., Warnow T. 2014. ASTRAL: Genome-scale coalescent-based species tree estimation. Bioinformatics. 30:541–548.

Morales A.E., Jackson N. ., Dewey T.A., O'Meara B., Carstens B.C. 2017. Speciation with gene flow in North American *Myotis* bats. Syst. Biol. 66:440-452.

Nadachowska-Brzyska K., Li C., Smeds L., Zhang G., Ellegren H. 2015. Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. Curr. Biol. 25:1375–1380.

Nakhleh L. 2013. Computational approaches to species phylogeny inference and gene tree reconciliation. Trends Ecol. Evol. 28:719–728.

Nater A., Burri R., Kawakami T., Smeds L., Ellegren H. 2015. Resolving evolutionary relationships in closely related species with whole-genome sequencing data. Syst. Biol. 46:1–54.

Nogueira A.C.R., Silveira R., Guimarães J.T.F. 2013. Neogene–Quaternary sedimentary and paleovegetation history of the eastern Solimões basin, central Amazon region. J. South Amer. Earth Sciences. 46: 89–99.

Oswald J.A., Overcast I., Mauck, W.M.I., Andersen M.J., Smith B.T. 2017. Isolation with asymmetric gene flow during the nonsynchronous divergence of dry forest birds. Mol. Ecol. 26:1386–1400.

Petermann P. 1997. The birds. In: Junk W.J., editor. The Central Amazon Floodplain: Ecology of a Pulsing System. Berlin: Springer Verlag. p. 419-452.

Pickrell J.K., Pritchard J.K. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 8:e1002967.

Poelstra J.W., Vijay N., Bossu C.M., Lantz H., Ryll B., Müller I., Baglione V., Unneberg P., Wikelski M., Grabherr M.G., Wolf J.B.W. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. Science. 344:1410–4.

Posada D. 2008. jModelTest: Phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.

Potter S., Bragg J.G., Peter B.M., Bi K., Moritz C. 2016. Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. Mol. Ecol. 25:1367–1380.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Queiroz K. De. 2007. Species concepts and species delimitation. Syst. Bot. 56:879–886.

Rambaut A., Drummond A.J. 2007. Tracer v1.4, http://tree.bio.ed.ac.uk/software/tracer/. Accessed March 9, 2017.

Remsen J.V., Parcker III T.A. 1983. Contribution of river-Created Habitats to Bird Species Richness in Amazonia. Biotropica. 15:223.

Rannala B., Yang Z. 2013. Improved reversible jump algorithms for Bayesian species delimitation. Genetics. 194:245–253.

Reich D., Thangaraj K., Patterson N., Price A.L., Singh L. 2009. Reconstructing Indian population history. Nature. 461:489–94.

Rheindt F.E., Fujita M.K., Wilton P.R., Edwards S.V. 2014. Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): population genetic and phylogenetic inferences from genome-wide SNPs. Syst. Biol. 63:134-152.

Robinson J.D., Bunnefeld L., Hearn J., Stone G.N., Hickerson M.J. 2014. ABC inference of multi-population divergence with admixture from unphased population genomic data. Mol.

Ecol. 23:4458–4471.

Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–542.

Rosenberg G.H. 1990. Habitat specialization and foraging behavior by birds of Amazonian river islands in northeastern Peru. Condor. 92:427.

Rosenblum E.B., Sarver B.A.J., Brown J.W., Des Roches S., Hardwick K.M., Hether T.D., Eastman J.M., Pennell M.W., Harmon L.J. 2012. Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. Evol. Biol. 39:255–261.

Rossetti D.F., Cohen M.C.L., Tatumi S.H., Sawakuchi A.O., Cremon E.H., Mittani J.C.R., Bertani T.C., Munita C.J.A.S., Tudela D.R.G., Yee M., Moya G. 2015. Mid-Late Pleistocene OSL chronology in western Amazonia and implications for the transcontinental Amazon pathway. Sediment. Geol. 330: 1-15.

Sambrook J., Fritsch E.F., Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd ed. New York: Cold Spring Harbor Laboratory Press.

Shaw T.I., Ruan Z., Glenn T.C., Liu L. 2013. STRAW: Species TRee Analysis Web server. Nucleic Acids Res. 41:W238–W241.

Smith B.T., McCormack J.E., Cuervo A.M., Hickerson M.J., Aleixo A., Cadena C.D., Pérez-Emán J., Burney C.W., Xie X., Harvey M.G., Faircloth B.C., Glenn T.C., Derryberry E.P., Prejean J., Fields S., Brumfield R.T. 2014. The drivers of tropical speciation. Nature. 515:406–409.

Smithe F.B. 1975. Naturalist's Color Guide. American Museum of Natural History, New York.

Smithe F.B. 1981. Naturalist's Color Guide. Part III. American Museum of Natural History, New York.

Solís-Lemus C., Knowles L.L., Ané C. 2015. Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution. 69:492–507.

Solís-Lemus C., Yang M., Ané C. 2016. Inconsistency of species-tree methods under gene flow. Syst. Biol. 65:843-851.

Sousa V., Hey J. 2013. Understanding the origin of species with genome-scale data: modeling gene-flow. Nature Rev. Genet. 14:404-414.

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

Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. Proc. Natl. Acad. Sci. USA. 114:1607-1612.

Than C., Ruths D., Nakhleh L. 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinformatics. 9: 322.

Wen D., Yu Y., Hahn M.W., Nakhleh L. 2016. Reticulate evolutionary history and extensive introgression in mosquito species revealed by phylogenetic network analysis. Mol. Ecol. 25:2361–2372.

Wittmann F., Schongart J., Junk W. J. 2010. Phytogeography, species diversity, community structure and dynamics of central Amazonian floodplais forests. In: Junk W. J., Piedade M. T. F., Wittmann F., Schongart J., Parolin P., editors. Amazonian floodplain forests: Ecophysiology, biodiversity and sustainable management. NewYork, NY: Springer. p. 61–102.

Wittmann F., Householder E., Piedade M.T.F., Assis R.L.D., Schöngart J., Parolin P., Junk W.J. 2012. Habitat specifity, endemism and the neotropical distribution of Amazonian white-water floodplain trees. Ecography. 36:690–707.

Yang Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. USA. 107:9264–9.

Yang Z. Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol. Biol. Evol. 31:3125–3135.

Yu Y., Nakhleh L. 2015. A maximum pseudo-likelihood approach for phylogenetic networks. BMC Genomics. 16:S10.

Zamudio K.R., Bell R.C., Mason N.A. 2016. Phenotypes in phylogeography: Species ' traits, environmental variation , and vertebrate diversification. Proc. Natl. Acad. Sci. USA. 113:1–8.

Zhang C., Rannala B., Yang Z. 2014. Bayesian species delimitation can be robust to guide-tree inference errors. Syst. Biol. 63:993–1004.

Zimmer K.J., Isler M.L. 2003. Family Thamnophilidae (typical antbirds). In: del Hoyo J., Elliot A., Christie D. A., editors. Handbook of the Birds of the World. Volume 8. Broadbills to Tapaculos. Barcelona: Lynx Edicións. p. 448–681.

Figure captions

Figure 1: **a**) Map of sampling localities (black circles, see table s1) of the *Myrmoborus lugubris* species complex and numbers of samples used for mtDNA, sequence capture of Ultra Conserved Elements and morphological analyses, respectively. Colors represent the geographic distribution of the four subspecies: gray – *M. l. berlepschi*, purple – *M. l. stictopterus*, pink – *M. l. femininus* and yellow – *M. l. lugubris*; **b**) Model with the best value of cross-entropy (K=4) for the population genetic structure inferred in sNMF based on 1664 SNPs. Numbers in the bars correspond to the localities in the map; **c**) Model with the second best value of cross-entropy (K=3) for the population genetic structure inferred in sNMF. Ilustrations retrieved from del Hoyo et al. 2017 (Handbook of the Birds of the World Alive).

Figure 2: **a)**: Score values of the first canonical axis of the discriminant functional analysis performed on specimens of the four *Myrmoborus lugubris* subspecies and morphologically intermediate individuals between *M. l. lugubris* and *M l. femininus* based on measurements of seven morphometric characters plotted as a function of the 18 localities sampled longitudinally and displayed from west to east (see locality numbers in Fig. 1a); **b)**: Picture of the left side of the head of females collected between localities 11 and 18 (see Fig. 1a) representing the cline in plumage between *M. l. lugubris* and *M l. femininus.*

Figure 3: **a)** Bayesian phylogenetic inference based on mtDNA (cyt b, 1040 bp). * posterior probability >0.95; **b)** Population genetic structure inferred with BAPS (K=5) using the mtDNA data (cyt b, 1040 bp). Note that *M. l. berlepschi* individuals were split into two groups (light and dark gray). Numbers represent localities and colors, taxa in figure 1a.

Figure 4: Species tree topologies based on 472 loci with >5 informative site and respective

bootstrap nodal support obtained with Mpest, Njst and STAR species tree methods, and posterior probability for the Bayesian species delimitation analyses on BP&P. **a)** Analyses without intermediate individuals between *Myrmoborus lugubris lugubris* and *M. l. femininus* **b)** Analyses assuming intermediate individuals between *M. l. lugubris* and *M. l. femininus* as a distinct taxa. le = *M. leucophrys*; be = *M. l. berlepschi;* st = *M. l. stictopterus;* fe = *M. l. femininus;* lu = *M. l. lugubris;* hy = Intermediate individuals between *M. l. lugubris* and *M. l. femininus.*

Figure 5: Overlapped species tree topologies obtained with SNAPP based on the complete SNP matrix (1,664 SNPs). **a)** Analysis without intermediate individuals between *Myrmoborus lugubris lugubris* and *M. l.*; **b)** Analysis including intermediate individuals between *M. l. lugubris* and *M. l. femininus* as a distinct taxon. be = *M. l. berlepschi;* st = *M. l. stictopterus;* fe = *M. l. femininus;* lu = *M. l. lugubris;* hy = intermediate individuals between *M. l. lugubris* and *M. l. femininus.*

Figure 6: Population relationships and migration edges inferred by TREEMIX. Color-scale indicates the weight of migration edges. Drift parameter is a relative temporal measure and the scale bar indicates 10 times the average standard error of the relatedness among populations based on the variance-covariance matrix of allele frequencies; be = *M. l. berlepschi;* st = *M. l. stictopterus;* fe = *M. l. femininus;* lu = *M. l. lugubris;* hy = intermediate individuals between *M. l. lugubris* and *M. l. femininus; le = M. leucophrys* (outgroup).

Figure 7: Species tree (top left; no introgression allowed) and the five best phylogenetic networks inferred with the Maximum Pseudo Likelihood algorithm of Phylonet allowing for two introgression events. In red, lines represent the introgression edges and number, the inheritance probabilities of each edge. Numbers in black represent branch length. be = *M. l.*

berlepschi; st = *M. l. stictopterus*; fe = *M. l. femininus*; lu = *M. l. lugubris*; hy = Intermediate individuals between *M. l. lugubris* and *M. l. femininus.*

Figure 8: The four demographic models tested in Fastsimcoal2. Model parameters are shown in Table 3, and include divergence times (Tdiv), current and ancestral effective population sizes (Ne) and migration (horizontal arrows).

Table 1: Bayes Factor Delimitation (BFD) results for alternative arrangements of individuals for the four *Myrmoborus lugubris* subspecies.

Model	Rank	ML	BF	Ln(BF)
(be),(st),(fe),(lu)	1	-1770.50	***	***
(be),(st),(fe,lu)	3	-18069.44	735.861380	6.60
(be),(st,fe),(lu)	2	-18044.73	686.447462	6.53

Clusters of individuals are represented between parentheses; be = *M. l. berlepschi;* st = *M. l. stictopterus;* fe = *M. l. femininus;* lu = *M. l. lugubris.*

Table 2: Bayes Factor Delimitation (BFD) results for alternative arrangements of individuals for the four *Myrmoborus lugubris* subspecies and morphologically intermediate individuals between *M. l. lugubris* and *M l. femininus*.

Model	Rank	ML	BF	Ln(BF)
(be),(st),(fe),(hy),(lu)	1	-22095.83	***	***
(be),(st),(fe,hy,lu)	4	-22483.43	775.18983	6.65
(be),(st,fe),(hy),(lu)	5	-22484.03	776.394753	6.65
(be,(st,fe),(hy,lu)	6	-22665.59	1139.50468	7.04
(be),(st,fe,hy),(lu)	7	-22689.10	1186.54008	7.08
(be),(st),(fe),(hy,lu)	2	-22281.99	372.312179	5.92
(be),(st),(fe,hy),(lu)	3	-22352.77	513.873197	6,24

Clusters of individuals are represented between parenthesis; be = *M. l. berlepschi;* st = *M. l. stictopterus;* fe = *M. l. femininus;* lu = *M. l. lugubris;* hy = Intermediate individuals between *M. l. lugubris* and *M. l. femininus.*

Table 3: Percentage (% Correct) and number of individuals correctly identified by a discriminant function analysis based on measurements of seven morphometric characters expressing the individual size of *Myrmoborus lugubris* taxa.

Taxon	**% Correct**	***M. l. berlepschi***	***M. l. femininus***	***M. l. lugubris***	***M. l. lugubris/femininus***	***M. l. stictopterus***
M. l. berlepschi (48)	100.0	48	0	0	0	0
M. l. femininus (30)	80.0	0	24	0	6	0
M. l. lugubris (16)	68.7	0	1	11	4	0
M. l. lugubris / femininus (53)	88.7	0	3	1	47	2
M. l. stictopterus (8)	75.0	0	0	0	2	6
Total	87.7	48	28	12	59	8

Numbers in parenthesis represent the total number of individuals analyzed per taxon.

Table 4: Parameters estimated in Fastsimcoal 2 under four demographic models for the diversification of *Myrmoborus lugubris*. Models are described in Figure 8.

Model	**N1**	**N2**	**N3**	**N4**	**N5**	**Na23**	**Na34**	**Na234**	**Na**	**NM43**
A	1101907	389003	663662	356988	145770	***	1878408	940731	112603	***
B	500690	714630	1110812	654584	230709	148122	***	1005070	1956792	***
C	1074254	645673	1085734	66633	208507	***	1570047	1964183	123272	0.00
D	1075223	731159	1133674	66832	590531	1556677	***	2175098	2160755	0.02
Lo 95%	896418	661016	1230587	61972	456088	1494866	***	1984074	2137099	0.02
Up 95%	962213	741371	1368173	68617	457183	1574056	***	2149316	2148960	0.07

Model	**NM34**	**NM54**	**NM45**	**NM53**	**NM35**	**TDIV1**	**TDIV2**	**TDIV3**	**TDIV4**	**Lhood**
A	***	***	***	***	***	1253010	266216.5	290599.9	1603298.6	-19813.85
B	***	***	***	***	***	443909.3	455759.6	477831.7	494954.9	-19830.184
C	4.65	0.03	0.00	0.02	0.23	169127.7	485341.3	505777.8	1484885.7	-19679.779
D	4.90	0.02	0.01	0.91	0.01	148078.5	350702.2	1486209.1	1508230	-19554.501
Lo 95%	4.59	0.01	0.01	0.97	0.02	125994.7	303123.7	1136881.6	1171892.1	***
Up 95%	5.01	0.03	0.02	1.32	0.07	148453.6	347698.9	1198295.7	1222050	***

Estimated parameters are effective population sizes (Ne) of current and ancestral populations in haploid individuals (N1 – Ne *M. l. berlepschi*; N2 – Ne *M. l. stictopterus*; N3 – Ne *M. l. femininus*; N4 – Ne morphologically intermediate group between *M. l. lugubris* and *M. l. femininus*; N5 – Ne *M. l. lugubris*; Na23 – ancestral of *M. l. femininus* and *M. l. stictopterus;* Na34 – ancestral Ne of *M. l. femininus M. l. lugubris;* Na234 – ancestral Ne of *M. l. stictopterus, M. l. femininus* and *M. l. lugubris;* Na – ancestral Ne of *M. lugubris*), number of migrants per generation from population x to y forward in time (NMxy), divergence times backwards in time as in figure 8 (TDIV), lower and upper limits of the 95% confidence interval of maximum likelihood parameter estimates based on 100 parametric bootstraps (Lo 95% and Up 95%).

Table 5: Maximum likelihood (Max ln(L)) obtained for the four demographic models simulated in Fastsimcoal2.

Model	Max ln(L)	2*par	AIC	ΔAIC	AIC weight
A	-19814	18	91264.2	1182.3	2E-257
B	-19830	18	91339.4	1257.6	8E-274
C	-19680	30	90658.7	576.9	5E-126
D	-19555	30	90081.8	0	1

Number of parameters (par); Akaike Information Criterion (AIC); differences in AIC relative to the best-fit model (ΔAIC); relative Akaike's weight of evidence based on the maximum-likelihood estimates (AIC weight).

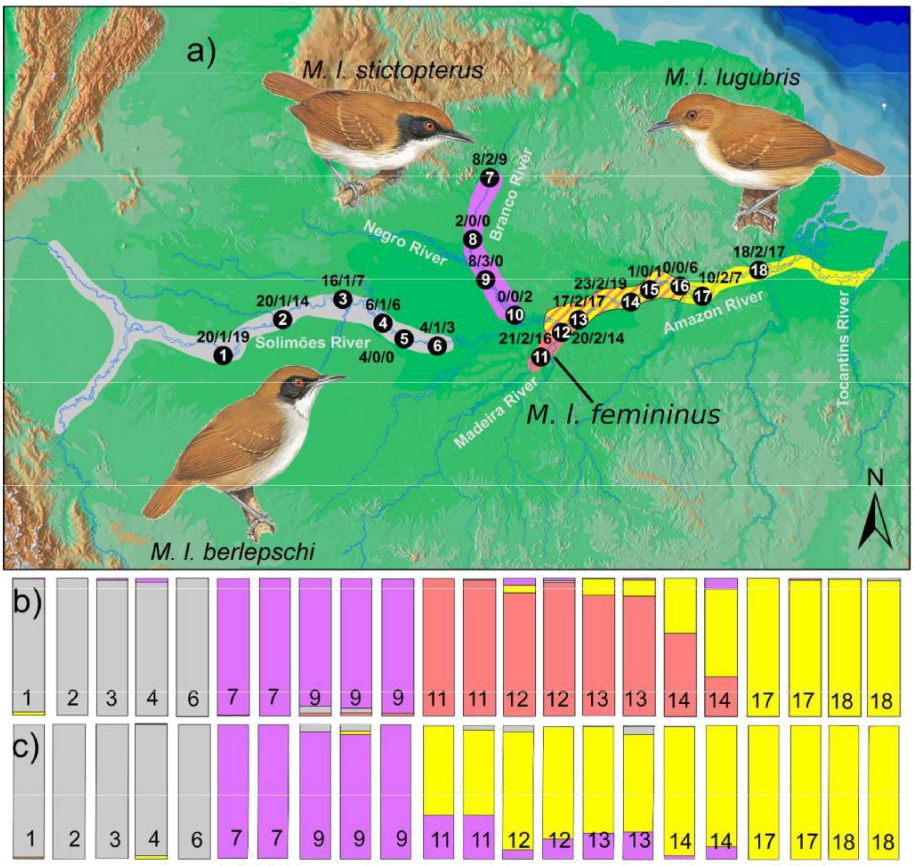

Figure 1: a) Map of sampling localities (black circles, see table s1) of the *Myrmoborus lugubris* species complex and numbers of samples used for mtDNA, sequence capture of Ultra Conserved Elements and morphological analyses, respectively. Colors represent the geographic distribution of the four subspecies: gray – *M. l. berlepschi*, purple – *M. l. stictopterus*, pink – *M. l. femininus* and yellow – *M. l. lugubris*; b) Model with the best value of cross-entropy (K=4) for the population genetic structure inferred in sNMF based on 1664 SNPs. Numbers in the bars correspond to the localities in the map; c) Model with the second best value of cross-entropy (K=3) for the population genetic structure inferred in sNMF. Ilustrations retrieved from del Hoyo et al. 2017 (Handbook of the Birds of the World Alive).

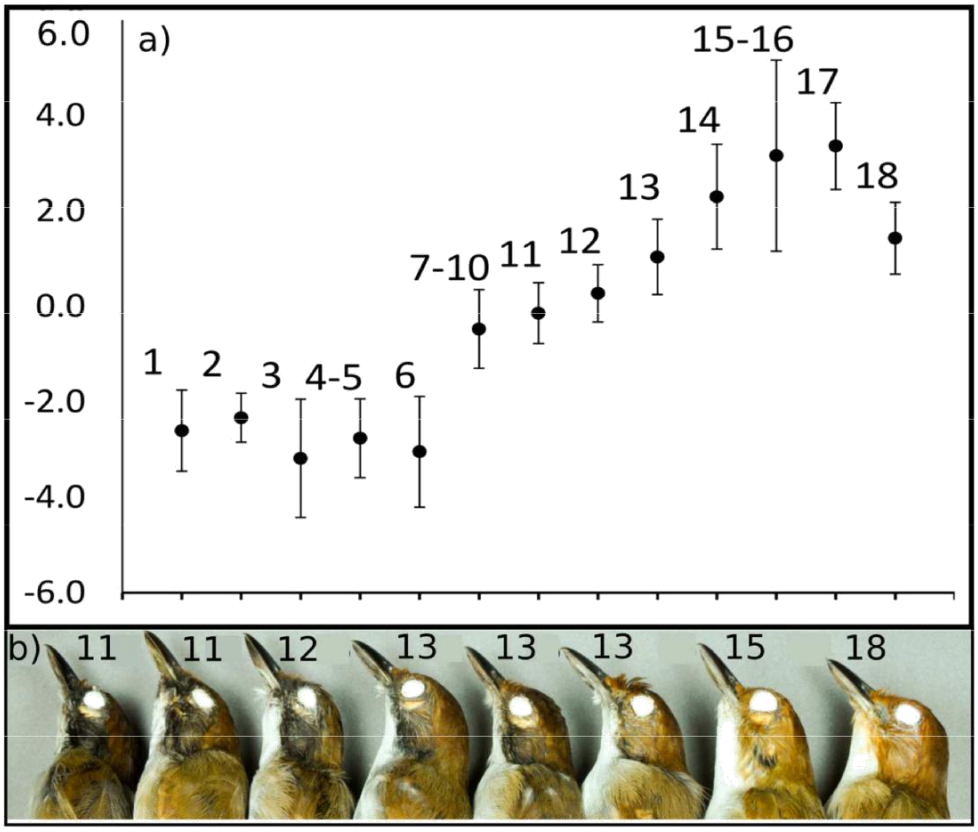

Figure 2: a): Score values of the first canonical axis of the discriminant functional analysis performed on specimens of the four Myrmoborus lugubris subspecies and morphologically intermediate individuals between *M. l. lugubris* and *M l. femininus* based on measurements of seven morphometric characters plotted as a function of the 18 localities sampled longitudinally and displayed from west to east (see locality numbers in Fig. 1a); b): Picture of the left side of the head of females collected between localities 11 and 18 (see Fig. 1a) representing the cline in plumage between *M. l. lugubris* and *M l. femininus*.

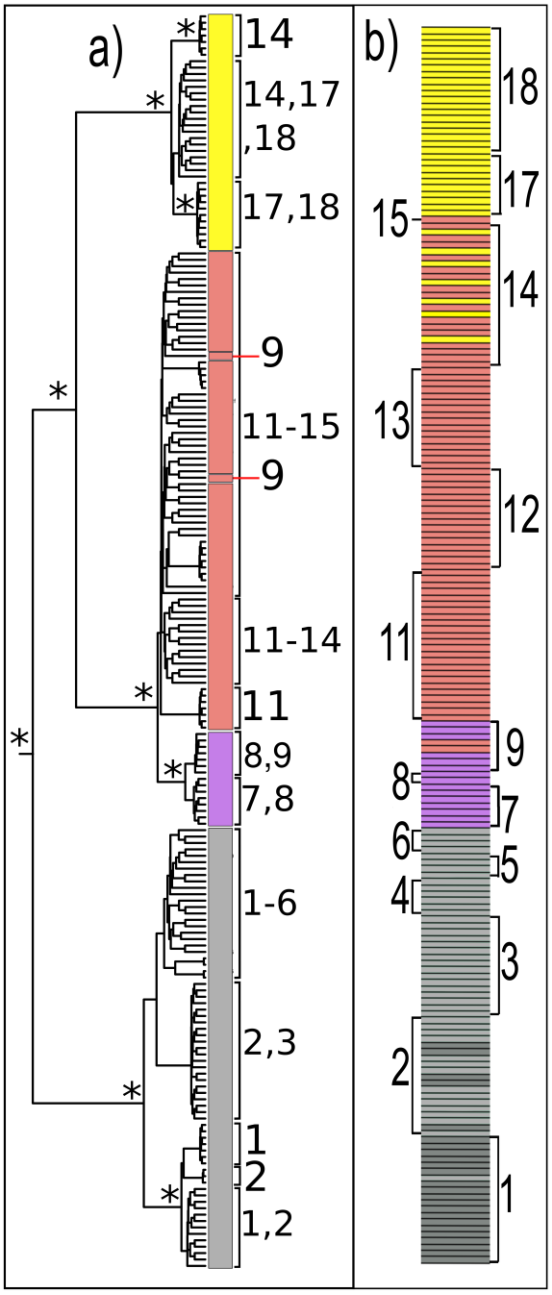

Figure 3: a) Bayesian phylogenetic inference based on mtDNA (cyt b, 1040 bp). * posterior probability >0.95; b) Population genetic structure inferred with BAPS (k=5) using the mtDNA data (cyt b, 1040 bp). Note that *M. l. berlepschi* individuals were split into two groups (light and dark gray). Numbers represent localities and colors, taxa in figure 1a.

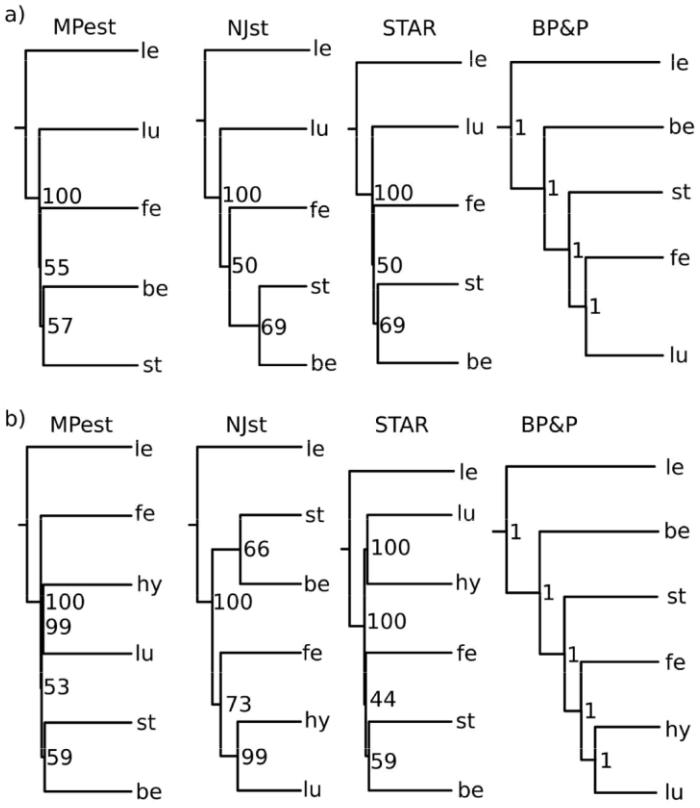

Figure 4: Species tree topologies based on 472 loci with >5 informative site and respective bootstrap nodal support obtained with Mpest, Njst and STAR species tree methods, and posterior probability for the Bayesian species delimitation analyses on BP&P. a) Analyses without intermediate individuals between *Myrmoborus lugubris lugubris* and *M. l. femininus* b) Analyses assuming intermediate individuals between *M. l. lugubris* and *M. l. femininus* as a distinct taxa. le = *M. leucophrys*; be = *M. l. berlepschi*; st = *M. l. stictopterus*; fe = *M. l. femininus*; lu = *M. l. lugubris*; hy = Intermediate individuals between *M. l. lugubris* and *M. l. femininus*.

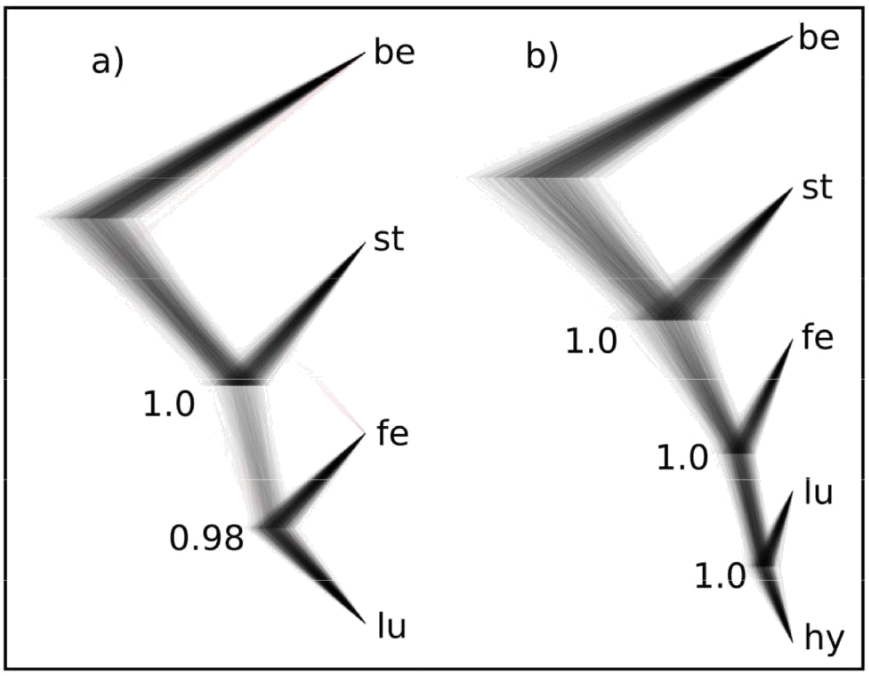

Figure 5: Overlapped species tree topologies obtained with SNAPP based on the complete SNP matrix (1,664 SNPs). a) Analysis without intermediate individuals between *Myrmoborus lugubris lugubris* and *M. l. femininus*; b) Analysis including intermediate individuals between *M. l. lugubris* and *M. l. femininus* as a distinct taxon. be = *M. l. berlepschi*; st = *M. l. stictopterus*; fe = *M. l. femininus*; lu = *M. l. lugubris*; hy = intermediate individuals between *M. l. lugubris* and *M. l. femininus*.

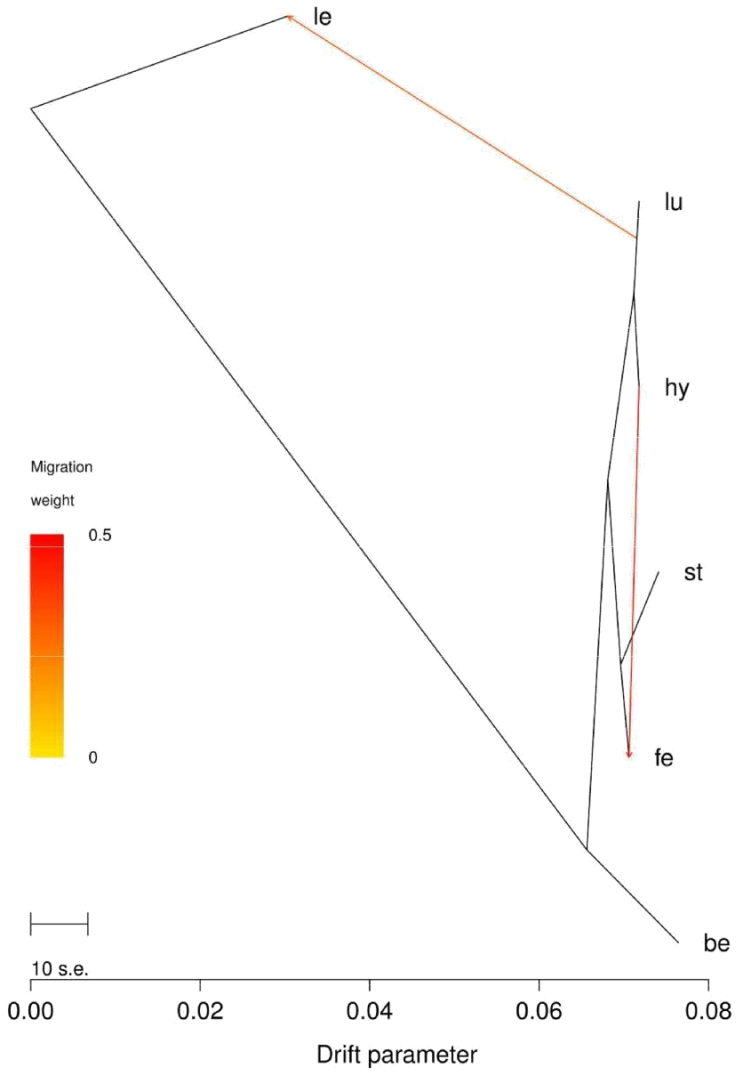

Figure 6: Population relationships and migration edges inferred by TREEMIX. Color-scale indicates the weight of migration edges. Drift parameter is a relative temporal measure and the scale bar indicates 10 times the average standard error of the relatedness among populations based on the variance-covariance matrix of allele frequencies; be = *M. l. berlepschi*; st = M. *l. stictopterus*; fe = *M. l. femininus*; lu = *M. l. lugubris*; hy = intermediate individuals between *M. l. lugubris* and *M. l. femininus*; le = *M. leucophrys* (outgroup).

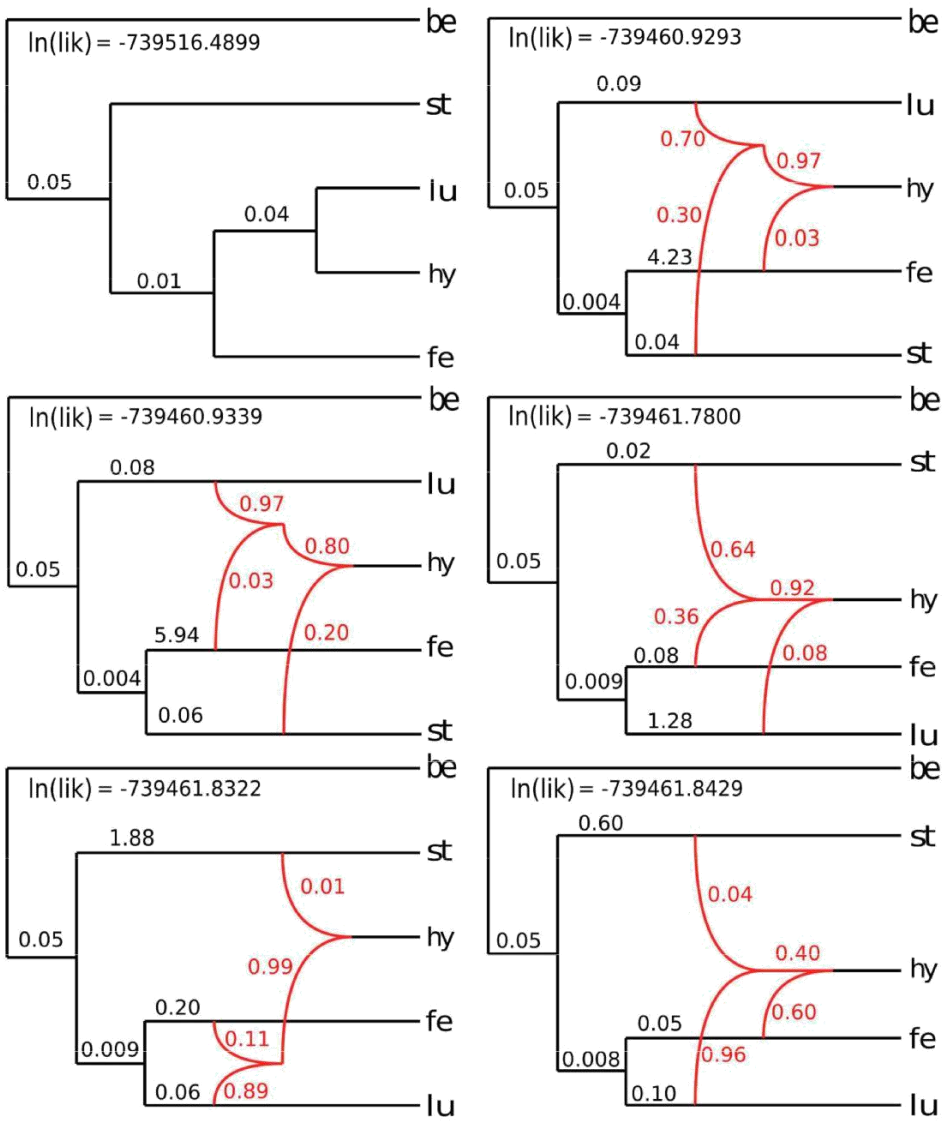

Figure 7: Species tree (top left; no introgression allowed) and the five best phylogenetic networks inferred with the Maximum Pseudo Likelihood algorithm of Phylonet allowing for two introgression events. In red, lines represent the introgression edges and number, the inheritance probabilities of each edge. Numbers in black represent branch length. be = *M. l. berlepschi*; st = *M. l. stictopterus*; fe = *M. l. femininus*; lu = *M. l. lugubris*; hy = Intermediate individuals between *M. l. lugubris* and *M. l. femininus*.

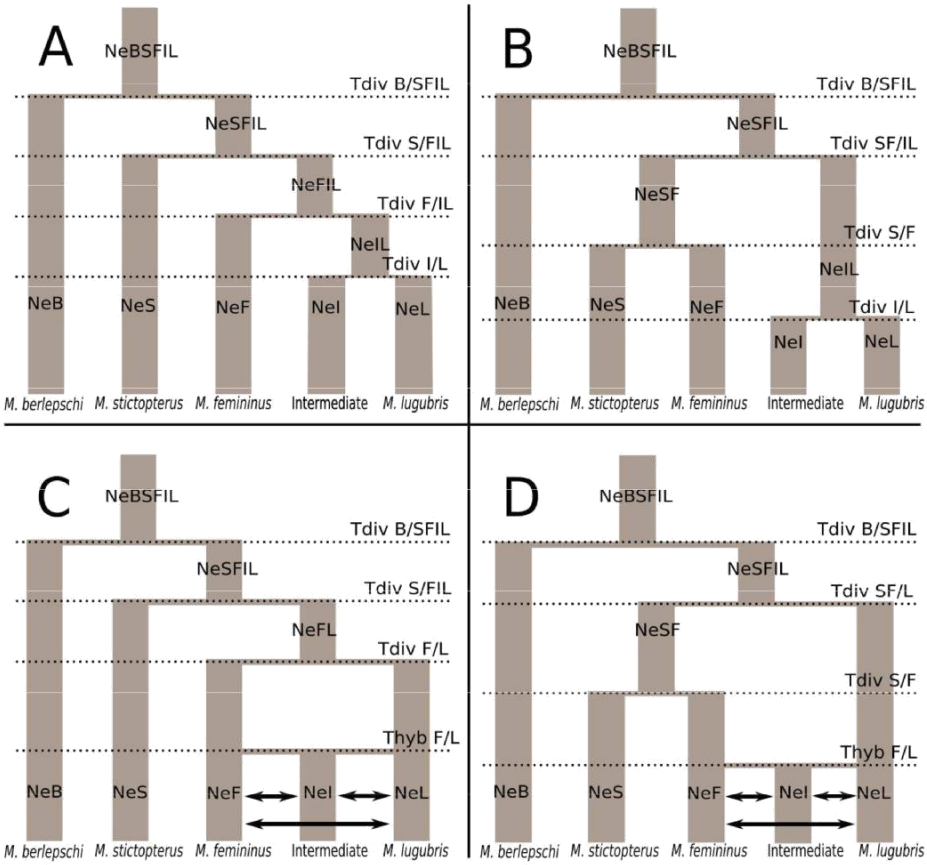

Figure 8: The four demographic models tested in Fastsimcoal2. Model parameters are shown in Table 3, and include divergence times (Tdiv), current and ancestral effective population sizes (Ne) and migration (horizontal arrows).

Supplementary Material

Supplementary material, including data files and/or online-only appendices, can be found in the Dryad data repository (doi: 10.5061/dryad.13b88).

Figure captions

Figure s1: Ventral, lateral and dorsal views of males of *Myrmoborus lugubris* representing the four subspecies. Two individuals of each taxon are displayed side by side from left to right: *M. l. lugubris, M. l. femininus, M. l. stictopterus* and *M. l. berlepschi.*

Figure s2: Lateral view of female heads of *Myrmoborus lugubris femininus* (localities 11 and 12), morphologically intermediate individuals (localities 13 – 17) and *M. l. lugubris* (locality 18), illustrating the gradual transition between these two taxa. Numbers in photos are the corresponding voucher numbers of the Museu Paraense Emilio Goeldi (MPEG).

Figure s3: Ventral, lateral, frontal head and dorsal views of females of *Myrmoborus lugubris* representing the four subspecies. Two individuals of each taxon are displayed side by side from left to right: *M. l. lugubris, M. l. femininus, M. l. stictopterus* and *M. l. berlepschi.*

Figure s4: Graphic representation of the first and second canonical axes of the discriminant functional analysis performed on specimens of the four *Myrmoborus lugubris* subspecies (gray – *M. l. berlepschi*, purple – *M. l. stictopterus*, pink – *M. l. femininus* and yellow – *M. l. lugubris*) and morphologically intermediate individuals between *M. l. lugubris* and *M l. femininus* (black triangles) based on measurements of seven morphometric characters (see Tables 1 and s2 for details).

Figure s5: **C1**: Scores of the first canonical axe of the discriminant functional analysis performed

on specimens of the four *Myrmoborus lugubris* subspecies and morphologically intermediate individuals between *M. l. lugubris* and *M l. femininus* based on measurements of seven morphometric characters and the three characters that most contributed to the observed variation in the first canonical axe (**EC** – exposed culmem; **CA1** – wing length; **CC1** – tail length) plotted in function of the 18 localities sampled longitudinally displayed from west to east (see locality numbers in Fig. 1a).

Figure s6: Density plot of the first discriminant function of the Discriminant Analysis of Principal Components. Yellow – *Myrmoborus l. lugubris;* Pink – *M. l. femininus;* Purple – *M. l. stictopterus;* Gray – *M. l. berlepschi.*

Figure s7: **a)** Maximum-likelihood (*) and Bayesian (°) phylogenetic inference of the complete concatenated UCE data set (751 loci). Node support values are based on 1000 quick bootstrap replicates (> 0.75) and posterior probability (>0.95); **b)** Model with the best value of cross-entropy (K=4) for the population genetic structure inferred in sNMF. Numbers next to the bars correspond to the localities in the map (Figure 1a).

Figure s8: Population relationships of *Myrmoborus lugubris* inferred by TREEMIX assuming no migration edges. Drift parameter is a relative temporal measure and the scale bar indicates 10 times the average standard error of the relatedness among populations based on the variance-covariance matrix of allele frequencies. le – *Myrmoborus leucophrys* (outgroup); be – *M. l. berlepschi*; st – *M. l. stictopterus*; fe – *M. l. femininus*; hy – morphologically intermediate individuals between *M. l. femininus* and *M. l. lugubris*; lu – *M. l. lugubris*.

Table s1: Specimens of *Myrmoborus lugubris* and outgroups (*M. leucophrys*, *M. myotherinus*) used in this study. Locality number as in Fig. 1a (Code); Institutions were samples are deposited (Inst.; MPEG: Museu Paraense Emílio Goeldi; INPA: Instituto Nacional de Pesquisas da Amazônia; LGEMA: Laboratório de Genética e Evolução Molecular de Aves); voucher number (N°); sex; taxon (species or subspecies following the taxonomy in Zimmer & Isler, 2003); study skins (1) used for plumage and morphometric analyses (Skin); samples used to sequence mtDNA (1) and Ultra Conserved Elements (*, DNA); Country; State; Locality; Latitude; Longitude (as in Fig. 1a).

Code	**Inst.**	**N°**	**Sex**	**Taxon**	**Skin**	**DNA**	**Country**	**State**	**Locality**	**Latitude**	**Longitude**
1	MPEG	78759	F	*M. l. berlepschi*	1	1*	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78760	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78762	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78764	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78767	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78768	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78769	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78770	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78771	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78772	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78773	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
1	MPEG	78774	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78791	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78792	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78793	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78794	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78795	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78803	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	78809	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
1	MPEG	A010932	M	*M. l. berlepschi*		1	Brazil	Amazonas	Tabatinga, Ilha Aramaçá	04°19'27,7"S	69°53'10,2"W
2	MPEG	78825	M	*M. l. berlepschi*	1	1*	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78826	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78827	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78828	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78829	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto	03°08'52,7"S	67°57'35,1"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
									América		
2	MPEG	A010937	M	*M. l. berlepschi*		1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78830	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	A010940	M	*M. l. berlepschi*		1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78836	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78837	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	A010941	M	*M. l. berlepschi*		1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78838	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78839	M	*M. l. berlepschi*		1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78840	M	*M. l. berlepschi*		1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto	03°08'52,7"S	67°57'35,1"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
									América		
2	MPEG	78841	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78842	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78843	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78856	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
2	MPEG	78857	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Santo Antônio do Içá, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
3	MPEG	A010955	M	*M. l. berlepschi*		1*	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010956	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	78895	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	78896	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	78897	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010957	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010958	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
3	MPEG	78898	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	78899	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	78900	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	78901	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010966	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010967	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010968	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010969	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
3	MPEG	A010970	M	*M. l. berlepschi*		1	Brazil	Amazonas	Fonte Boa, Ilha Nova Esperança	02°28'58,2"S	66°00'38,4"W
4	MPEG	78941	M	*M. l. berlepschi*	1	1*	Brazil	Amazonas	Tefé, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
4	MPEG	78942	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tefé, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
4	MPEG	78943	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tefé, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
4	MPEG	78944	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tefé, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
4	MPEG	78945	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tefé, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
4	MPEG	78946	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Tefé, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
6	MPEG	78954	F	*M. l. berlepschi*	1	1*	Brazil	Amazonas	Coari, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
6	MPEG	78960	M	*M. l. berlepschi*		1	Brazil	Amazonas	Coari, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
6	MPEG	78961	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Coari, Ilha da Botija	04°00'49,5"S	62°57'00,5"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
6	MPEG	78962	M	*M. l. berlepschi*	1	1	Brazil	Amazonas	Coari, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
11	MPEG	79060	F	*M. l. berlepschi*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79061	M	*M. l. berlepschi*		1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79062	F	*M. l. berlepschi*	1	1*	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79063	?	*M. l. berlepschi*		1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	A010995	?	*M. l. berlepschi*		1*	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	A010996	F	*M. l. femininus*		1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79064	F	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79065	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79066	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79067	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79068	F	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79069	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79070	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79071	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79072	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79073	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79074	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
11	MPEG	79075	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79076	M	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
11	MPEG	79092	F	*M. l. femininus*	1	1	Brazil	Amazonas	Borba, Ilha de Borba	04°22'31,7"S	59°35'49,5"W
12	MPEG	79112	M	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79113	F	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	A11005	M	*M. l. femininus*		1*	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79114	M	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79115	M	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79116	M	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79117	F	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79118	M	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79119	F	*M. l. femininus*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79120	M	*M. l. femininus / lugubris*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	A11006	F	*M. l. femininus/lugubris*		1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79135	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79136	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79137	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
12	MPEG	79138	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
12	MPEG	79139	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°34'17,3"S	58°55'22,2"W
13	MPEG	79140	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Risco	03°09'31,6"S	58°22'13,0"W
13	MPEG	79141	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Risco	03°09'31,6"S	58°22'13,0"W
13	MPEG	79142	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Risco	03°09'31,6"S	58°22'13,0"W
13	MPEG	79143	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Risco	03°09'31,6"S	58°22'13,0"W
13	MPEG	79144	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Risco	03°09'31,6"S	58°22'13,0"W
13	MPEG	79167	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79168	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79169	F	*M. l. femininus/lugubris*	1	1*	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79170	F	*M. l. femininus/lugubris*	1	1*	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79171	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79172	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79173	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79174	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79175	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79176	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79177	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W
13	MPEG	79178	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Itacoatiara, Ilha do Boi	03°11'4,3"S	58°15'52,1"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
14	MPEG	79181	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79182	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79183	F	*M. l. femininus/lugubris*	1	1*	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79184	F	*M. l. femininus/lugubris*	1	1*	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79185	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79186	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79187	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	A011019	F	*M. l. femininus/lugubris*		1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79188	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79189	F	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	A011020	M	*M. l. femininus/lugubris*		1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79190	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79191	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	A011021	M	*M. l. femininus/lugubris*		1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79192	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79193	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79194	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79195	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
14	MPEG	A011022	M	*M. l. femininus/lugubris*		1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79196	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79197	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79198	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
14	MPEG	79199	M	*M. l. femininus/lugubris*	1	1	Brazil	Amazonas	Parintins, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
17	MPEG	79225	M	*M. l. femininus/lugubris*	1	1*	Brazil	Pará	Monte Alegre, Ilha Cacoal Grande	02°23'17,2"S	54°21'41,9"W
17	MPEG	79226	M	*M. l. femininus/lugubris*	1	1*	Brazil	Pará	Monte Alegre, Ilha Cacoal Grande	02°23'17,2"S	54°21'41,9"W
17	MPEG	79227	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Monte Alegre, Ilha Cacoal Grande	02°23'17,2"S	54°21'41,9"W
17	MPEG	79228	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Monte Alegre, Ilha Cacoal Grande	02°23'17,2"S	54°21'41,9"W
17	MPEG	79229	F	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Monte Alegre, Ilha Cacoal Grande	02°23'17,2"S	54°21'41,9"W
17	MPEG	79230	F	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Monte Alegre, Ilha Cacoal Grande	02°23'17,2"S	54°21'41,9"W
17	MPEG	79252	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Monte Alegre, Rio Amazonas	02°25'25,6"S	54°29'54,5"W
17	MPEG	A011028	M	*M. l. femininus/lugubris*		1	Brazil	Pará	Monte Alegre, Rio Amazonas	02°25'25,6"S	54°29'54,5"W
17	MPEG	A011029	F	*M. l. femininus/lugubris*		1	Brazil	Pará	Monte Alegre, Rio Amazonas	02°25'25,6"S	54°29'54,5"W
17	MPEG	A011030	F	*M. l. femininus/lugubris*		1	Brazil	Pará	Monte Alegre, Rio Amazonas	02°25'25,6"S	54°29'54,5"W
18	MPEG	79254	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79255	F	*M. l. femininus/lugubris*	1	1*	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79256	M	*M. l. femininus/lugubris*	1	1*	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
18	MPEG	79257	F	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79258	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79259	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79260	F	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79261	F	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79262	F	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79263	M	*M. l. femininus/lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79264	F	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79265	M	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79266	F	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79267	M	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79294	F	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79295	M	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	79296	M	*M. l. lugubris*	1	1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
18	MPEG	A011037	M	*M. l. lugubris*		1	Brazil	Pará	Almeirim, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
2	INPA	A015	?	*M. l. lugubris*		1	Brazil	Amazonas	Santo Antônio do Iça, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
9	INPA	A10504	F	*M. l. lugubris*		1	Brazil	Amazonas	Parna Jaú, Ilha das Onças	1°50'37,8"S	61°22'48,2"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
9	INPA	A10505	F	*M. l. lugubris*		1*	Brazil	Amazonas	Parna Jaú, Ilha das Onças	1°50'37,8"S	61°22'48,2"W
9	INPA	A10508	F	*M. l. lugubris*		1	Brazil	Amazonas	Parna Jaú, Ilha Jussara	1°52'32,4"S	61°22'1,7"W
9	INPA	A15803	M	*M. l. lugubris*		1	Brazil	Amazonas	Parna Jaú, Ilha das Onças	1°48'37"S	61°23'42"W
9	INPA	A15804	M	*M. l. lugubris*		1	Brazil	Amazonas	Parna Jaú, Ilha das Onças	1°48'37"S	61°23'42"W
7	INPA	A2191	F	*M. l. lugubris*		1*	Brazil	Roraima	Caracaraí, Ilha do Palhal	1°14'25"N	61°18'51"W
7	INPA	A2193	M	*M. l. lugubris*		1*	Brazil	Roraima	Caracaraí, Ilha do Palhal	1°14'25"N	61°18'51"W
9	INPA	A3095	F	*M. l. lugubris*		1*	Brazil	Amazonas	Novo Airão, Ilhas do Rio Negro	1°51'56,7"S	61°22'20,8"W
9	INPA	A3100	M	*M. l. lugubris*		1	Brazil	Amazonas	Novo Airão, Ilhas do Rio Negro	1°51'56,7"S	61°22'20,8"W
9	INPA	A3103	M	*M. l. lugubris*		1*	Brazil	Amazonas	Novo Airão, Ilhas do Rio Negro	1°51'56,7"S	61°22'20,8"W
8	INPA	A8376	M	*M. l. lugubris*		1	Brazil	Roraima	Santa Maria do Boiaçu, Ilha no Rio Branco	0°31'48"S	61°47'57"W
8	INPA	A8379	F	*M. l. lugubris*		1	Brazil	Roraima	Santa Maria do Boiaçu, Ilha no Rio Branco	0°31'48"S	61°47'57"W
5	INPA	A936	?	*M. l. lugubris*		1	Brazil	Amazonas	Tefé, Ilha do Barbado	3°46'51,8"S	64°1'31,3"W
5	INPA	A960	?	*M. l. lugubris*		1	Brazil	Amazonas	Tefé, Ilha do Barbado	3°46'51,8"S	64°1'31,3"W
5	INPA	A966	?	*M. l. lugubris*		1	Brazil	Amazonas	Resex Catuá-Ipixuna, Ilha do Ipixuna	3°50'45"S	63°49'48,2"W
5	INPA	A971	?	*M. l. lugubris*		1	Brazil	Amazonas	Resex Catuá-Ipixuna, Ilha do	3°50'45"S	63°49'48,2"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
									Ipixuna		
7	MPEG	56402	M	*M. l. lugubris*	1	1	Brazil	Roraima	Caracaraí, PARNA Viruá - Ilha Aliança	01º28'N	61º15'W
7	MPEG	56403	M	*M. l. lugubris*		1	Brazil	Roraima	Caracaraí, PARNA Viruá - Ilha Inajatuba	01º25'N	61º16'W
7	MPEG	56404	M	*M. l. lugubris*	1	1	Brazil	Roraima	Caracaraí, PARNA Viruá - Ilha Inajatuba	01º25'N	61º16'W
7	MPEG	56405	M	*M. l. stictopterus*	1	1	Brazil	Roraima	Caracaraí, PARNA Viruá - Ilha Inajatuba	01º25'N	61º16'W
7	MPEG	56406	F	*M. l. stictopterus*	1	1	Brazil	Roraima	Caracaraí, PARNA Viruá - Ilha Inajatuba	01º25'N	61º16'W
7	MPEG	56828	M	*M. l. stictopterus*	1	1	Brazil	Roraima	Caracaraí, PARNA Viruá - Ilha Inajatuba	01º25'N	61º16'W
15	MPEG	65253	M	*M. l. stictopterus*	1	1	Brazil	Pará	Juruti, Lago Santana	02°05'01,1"S	59°45'06,8"W
11	MPEG	73716	M	*M. l. stictopterus*		1	Brazil	Amazonas	Borba, Auarazinho	04°22'57,6"S	59°45'06,8"W
12	MPEG	73799	M	*M. l. stictopterus*		1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°35'31,2"S	58°56'35,6"W
12	MPEG	73802	?	*M. l. stictopterus*		1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°35'31,2"S	58°56'35,6"W
12	MPEG	73846	M	*M. l. stictopterus*		1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°35'31,2"S	58°56'35,6"W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
12	MPEG	73847	F	*M. l. stictopterus*		1	Brazil	Amazonas	Autazes, Ilha Caiçara	03°35'31,2"S	58°56'35,6"W
11	MPEG	35520	M	*M. l. stictopterus*	1		Brazil	Amazonas	Borba, Foz do Igarapé Jaraqui	04o06'S	59o58'W
11	MPEG	35526	M	*M. l. stictopterus*	1		Brazil	Amazonas	Borba, Foz do Igarapé Jaraqui	04o06'S	59o58'W
11	MPEG	35527	F	*M. l. stictopterus*	1		Brazil	Amazonas	Borba, Foz do Igarapé Jaraqui	04o06'S	59o58'W
10	MPEG	43843	M	*M. l. stictopterus*	1		Brazil	Amazonas	Rio Negro, Arquipélago das Anavilhanas	03o00'S	60o26'W
10	MPEG	43844	F	*M. l. stictopterus*	1		Brazil	Amazonas	Rio Negro, Arquipélago das Anavilhanas	03o00'S	60o26'W
16	MPEG	56159	M	*M. l. stictopterus*	1		Brazil	Pará	Curuá, Pé na Cova	2o03'08"S	55o04'43"W
16	MPEG	56160	M	*M. l. stictopterus*	1		Brazil	Pará	Curuá, Pé na Cova	2o03'08"S	55o04'43"W
16	MPEG	56161	M	*M. l. stictopterus*	1		Brazil	Pará	Curuá, Pé na Cova	2o03'08"S	55o04'43"W
16	MPEG	56162	F	*M. l. stictopterus*	1		Brazil	Pará	Curuá, Pé na Cova	2o03'08"S	55o04'43"W
16	MPEG	56163	F	*M. l. stictopterus*	1		Brazil	Pará	Curuá, Pé na Cova	2o03'08"S	55o04'43"W
16	MPEG	56164	F	*M. l. stictopterus*	1		Brazil	Pará	Curuá, Pé na Cova	2o03'08"S	55o04'43"W
7	MPEG	56401	M	*M. l. stictopterus*	1		Brazil	Roraima	Caracaraí, Parna Viruá, Ilha Aliança	1o28'N	61o15'W
7	MPEG	56704	F	*M. l. stictopterus*	1		Brazil	Roraima	Caracaraí, Parna Viruá, Ilha Aliança	1o28'N	61o15'W

Code	Inst.	N°	Sex	Taxon	Skin	DNA	Country	State	Locality	Latitude	Longitude
7	MPEG	56705	M	*M. l. stictopterus*	1		Brazil	Roraima	Caracaraí, Parna Viruá, Ilha Aliança	1o28'N	61o15'W
7	MPEG	56706	M	*M. l. stictopterus*	1		Brazil	Roraima	Caracaraí, Parna Viruá, Ilha Aliança	1o28'N	61o15'W
	LGEMA	17438	M	*M. leucophrys*		1*	Brazil	Amazonas	Codajas, Ilha do Camaleao	3o50'36"S	62o14'29"S
	LGEMA	12450		*M. myotherinus*		1*	Brazil	Pará	Jacareacanga, Rio Teles Pires	9o31'08"S	56o77'80"W

Table s2: Mean values (in millimeters) and size range (in parenthesis) of selected morphometric variables of the five recognized groups of *Myrmoborus lugubris*.

Taxon	Sex	Sample size	Bill depth	Bill width	Culmem	Bill length	Wing	Tail	Tarsus
M. l. berlepschi	Male	38	5.63 (5.04-6.05)	5.02 (4.69-5.53)	19.16 (17.65-20.89	11.42 (10.02-12.62)	64.09 (59.86-67.35)	38.11 (35.83-39.99)	26.53 (24.06-28.46)
	Female	11	5.57 (5.30-6.02)	4.88 (4.68-5.20)	18.64 (16.93-19.93)	11.19 (10.65-12.26)	62.02 (59.26-64.91)	37.52 (34.8-39.31)	26.27 (25.37-27.49)
M. l. femininus	Male	21	5.89 (5.13-6.58)	5.18 (4.58-5.63)	19.89 (18.8-21.1)	11.78 (11.14-12.7)	67.60 (65.14-72.88)	42.6 (40.30-45.46)	27.56 (26.30-29.75)
	Female	12	5.89 (5.43-6.36)	5.04 (4.54-5.47)	19.84 (18.44-21.54)	11.61 (10.94-12.50)	65.67 (63.63-68.05)	41.76 (39.89-43.05)	26.96 (25.5-28.13)
M. l. lugubris	Male	9	6.03 (5.82-6.26)	5.37 (5.02-5.64)	20.25 (19.46-20.88)	11.81 (11.00-12.27)	69.91 (68.45-71.34)	45.64 (43.40-46.8)	27.52 (26.37-28.72)
	Female	7	5.94 (5.41-6.11)	5.19 (4.8-5.55)	20.12 (19.62-21.14)	12.47 (11.65-15.5)	65.82 (63.73-68.33)	44.50 (43.64-45.9)	26.91 (25.88-27.68)
M. l. lugubris/femininus	Male	30	5.80 (5.22-6.36)	5.10 (4.71-5.81)	20.85 (18.19-21.76)	12.17 (10.55-12.91)	69.68 (65.14-74.09)	44.94 (40.42-49.14)	27.64 (25.61-28.88)
	Female	21	5.86 (5.28-6.52)	5.10 (4.71-5.7)	20.98 (19.74-22.36)	12.18 (11.02-13.07)	68.37 (64.02-71.99)	43.47 (40.42-48.02)	27.31 (24.9-28.6)
M. l. stictopterus	Male	6	5.76 (5.52-6.0)	5.39 (5.01-5.86)	20.26 (19.21-20.84)	11.70 (11.03-11.85)	66.39 (64.25-68.14)	41.64 (40.55-43.5)	26.51 (26.97-26.13)
	Female	2	5.47 (5.45-5.49)	5.05 (4.67-5.43)	20.4 (19.89-20.9)	11.61 (11.62-11.6)	64.99 (64.3-65.67)	42.73 (41.65-43.8)	25.97 (25.57-26.36)

Table s3: Morphometric characters (see Tables 1 and s2 for details) that showed significant Mann-Whitney test values between the four *Myrmoborus lugubris* subspecies and morphologically intermediate individuals between *M. l. lugubris* and *M l. femininus*. 1 – bill depth; 2 – bill width; 3 – exposed culmem; 4 - bill length; 5 - wing length; 6 - tail length; 7 - tarsus length. Accepted significance $P< 0.05$.

Taxon	***M. l. berlepschi***	***M. l. stictopterus***	***M. l. femininus***	***M. l. lugubris/femininus***
M. l. stictopterus	2-6	***	***	***
M. l. femininus	1-7	3-7	***	***
M. l. lugubris/femininus	1-7	5,7	3-6	***
M. l. lugubris	1-7	1,6,7	1,2,4,6	1-5

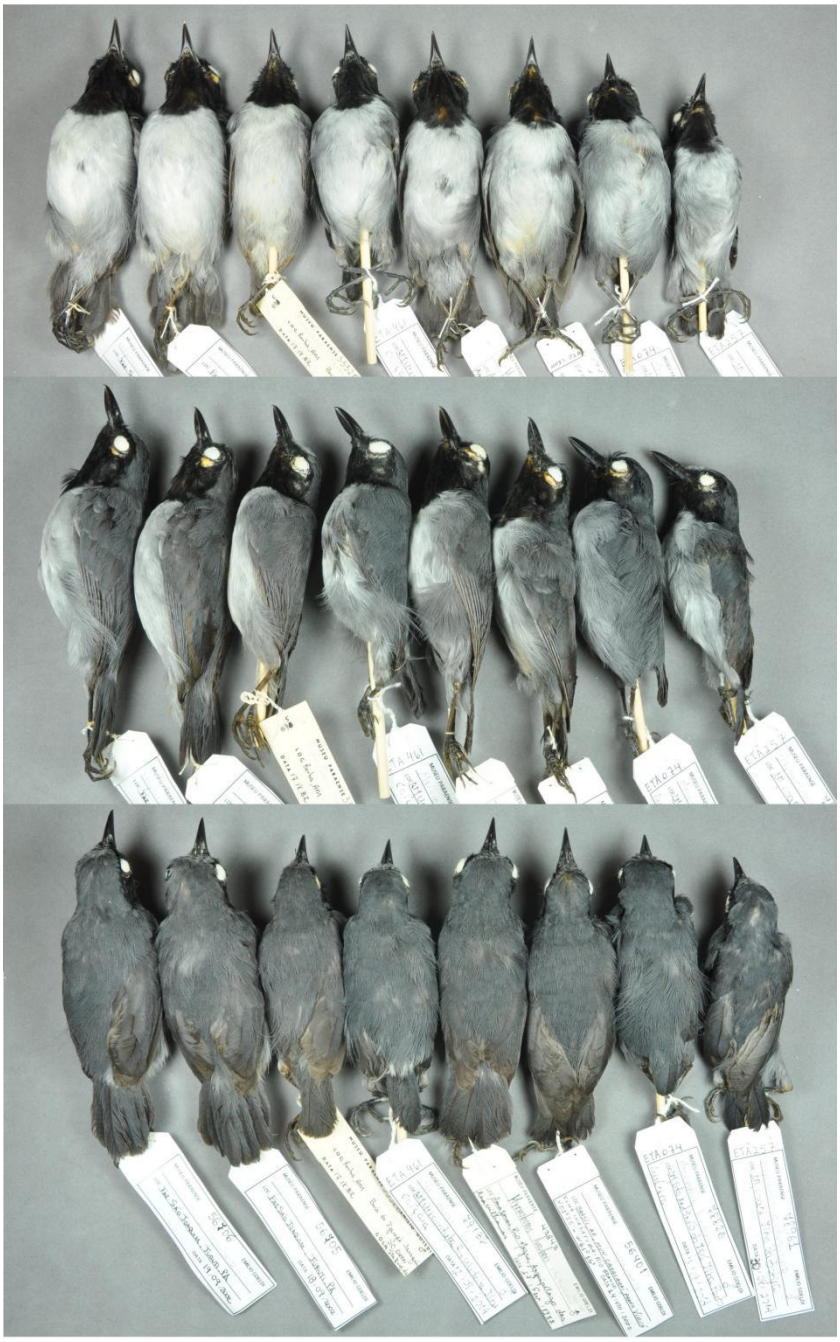

Figure s1

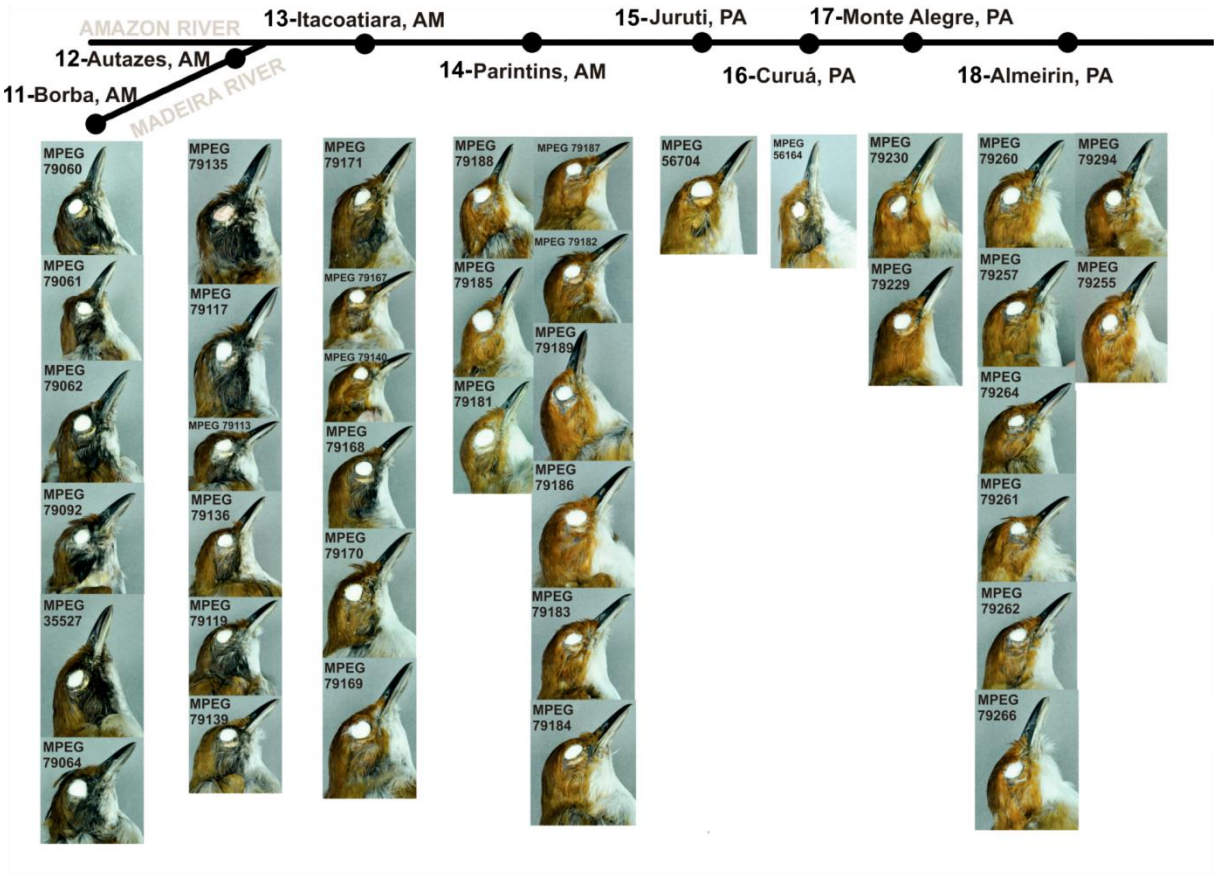
AMAZON RIVER
MADEIRA RIVER
11-Borba, AM
12-Autazes, AM
13-Itacoatiara, AM
14-Parintins, AM
15-Juruti, PA
16-Curuá, PA
17-Monte Alegre, PA
18-Almeirin, PA
MPEG 79060
MPEG 79061
MPEG 79062
MPEG 79092
MPEG 35527
MPEG 79064
MPEG 79135
MPEG 79117
MPEG 79113
MPEG 79136
MPEG 79119
MPEG 79139
MPEG 79171
MPEG 79167
MPEG 79140
MPEG 79168
MPEG 79170
MPEG 79169
MPEG 79188
MPEG 79187
MPEG 79185
MPEG 79182
MPEG 79189
MPEG 79181
MPEG 79186
MPEG 79183
MPEG 79184
MPEG 56704
MPEG 56164
MPEG 79230
MPEG 79229
MPEG 79260
MPEG 79294
MPEG 79257
MPEG 79255
MPEG 79264
MPEG 79261
MPEG 79262
MPEG 79266

Figure s2

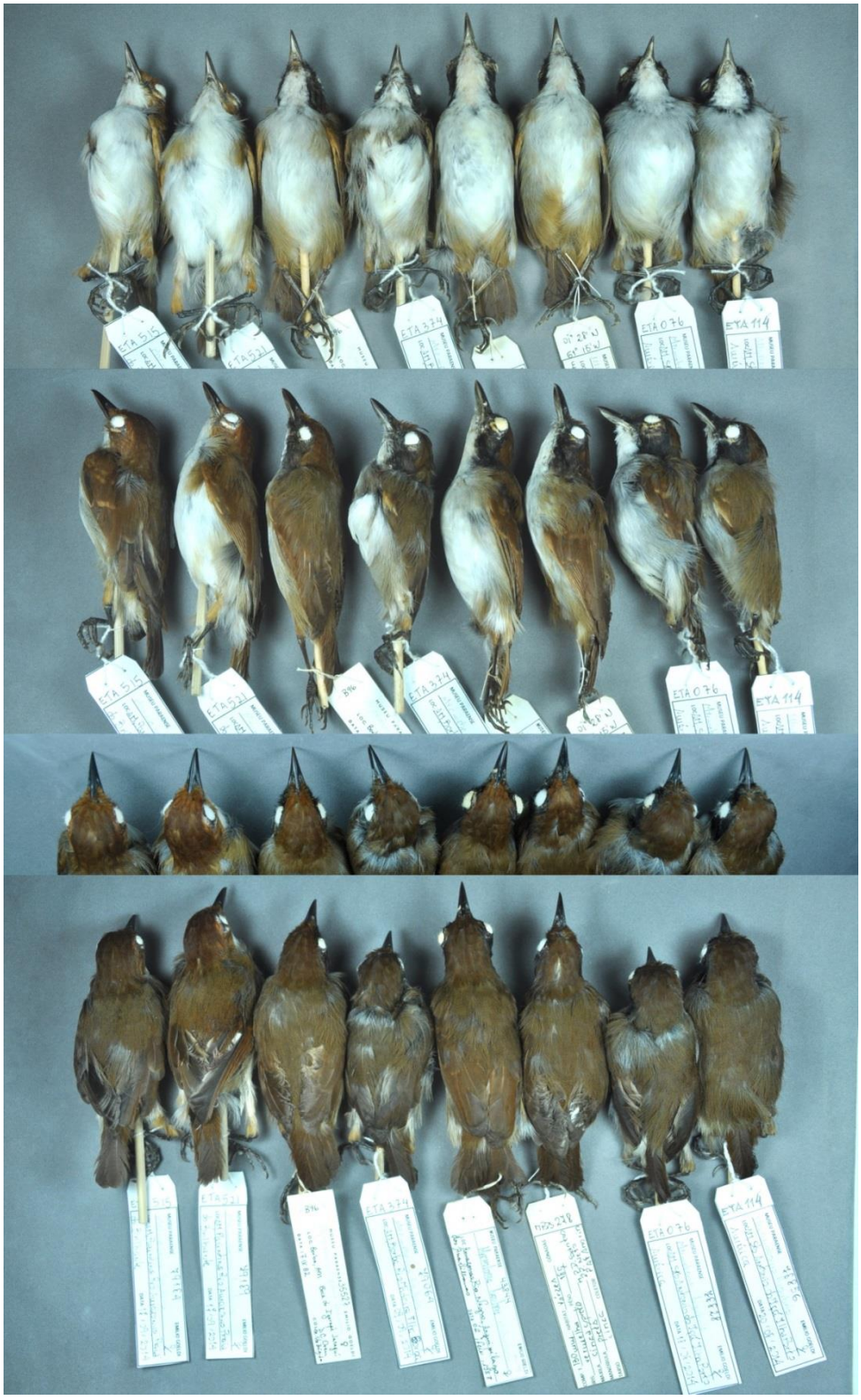

Figure s3

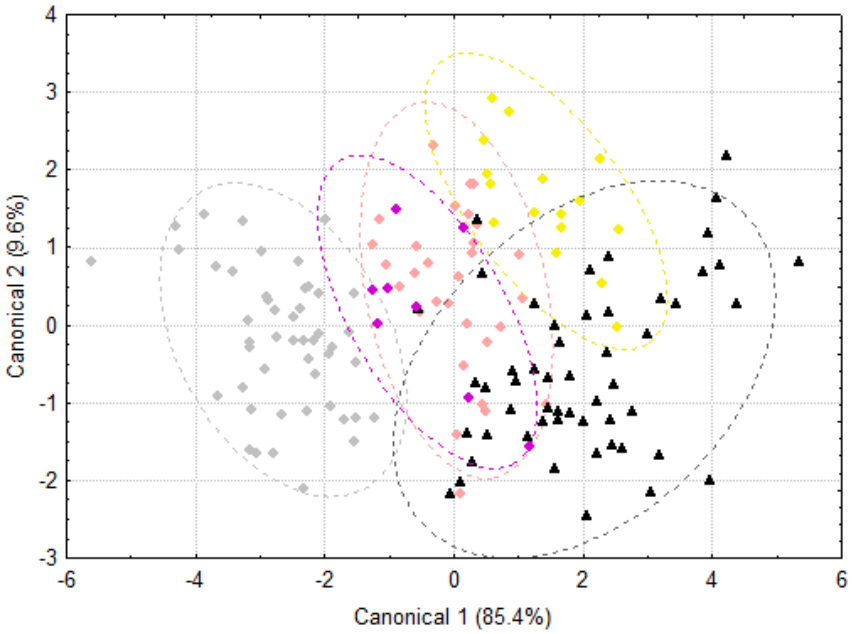

Figure s4

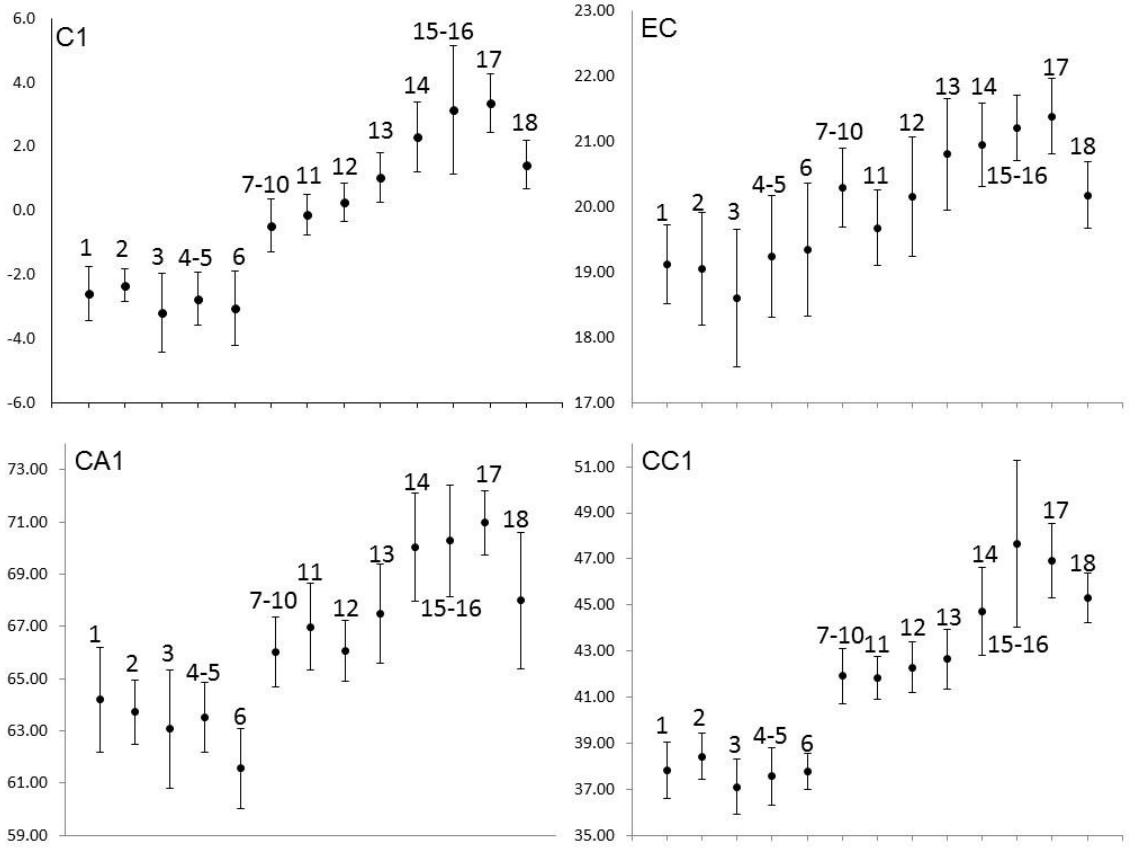
C1
6.0
4.0
2.0
0.0
-2.0
-4.0
-6.0
1
2
3
4-5
6
7-10
11
12
13
14
15-16
17
18
EC
23.00
22.00
21.00
20.00
19.00
18.00
17.00
1
2
3
4-5
6
7-10
11
12
13
14
15-16
17
18
CA1
73.00
71.00
69.00
67.00
65.00
63.00
61.00
59.00
1
2
3
4-5
6
7-10
11
12
13
14
15-16
17
18
CC1
51.00
49.00
47.00
45.00
43.00
41.00
39.00
37.00
35.00
1
2
3
4-5
6
7-10
11
12
13
14
15-16
17
18

Figure s5

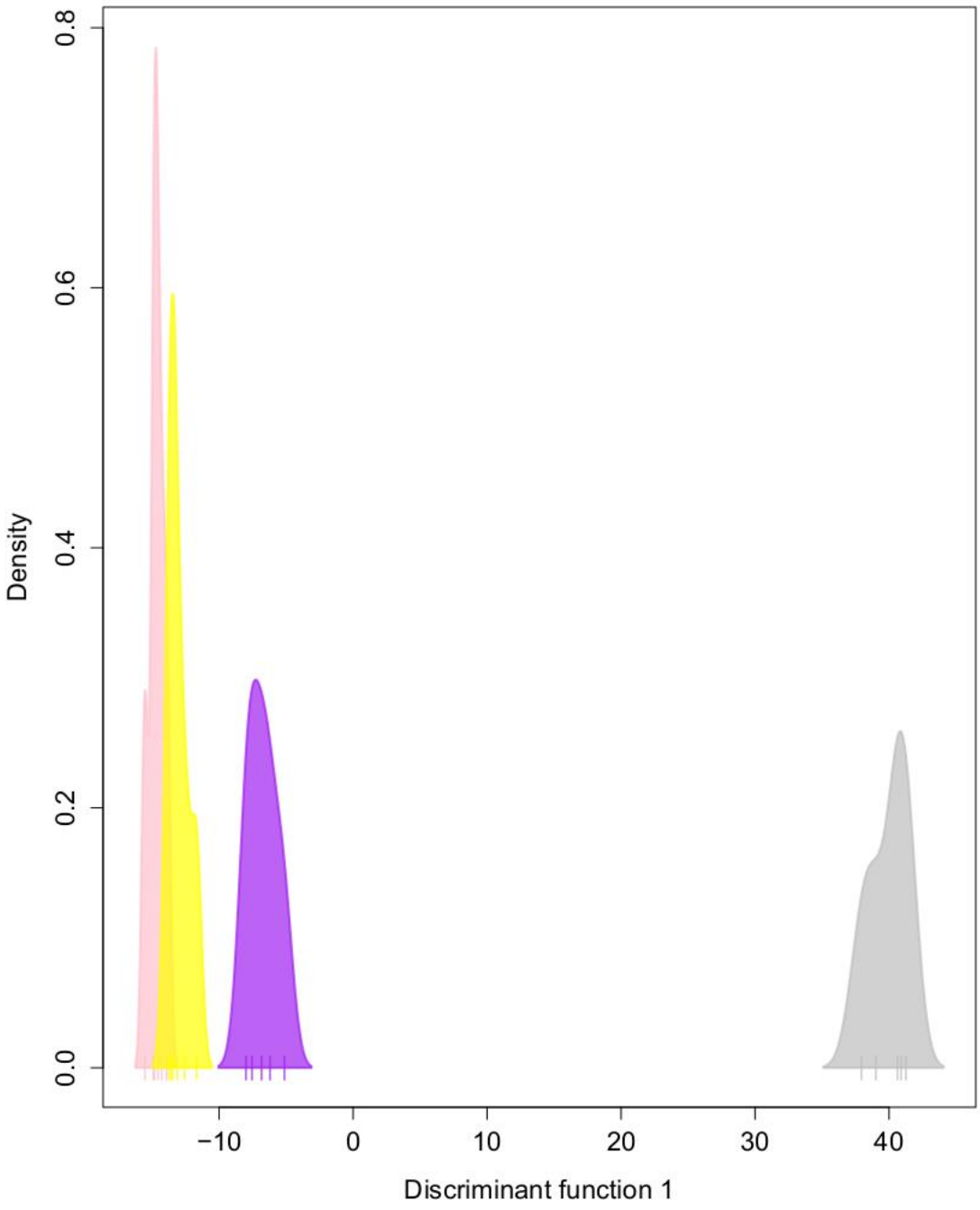

Figure s6

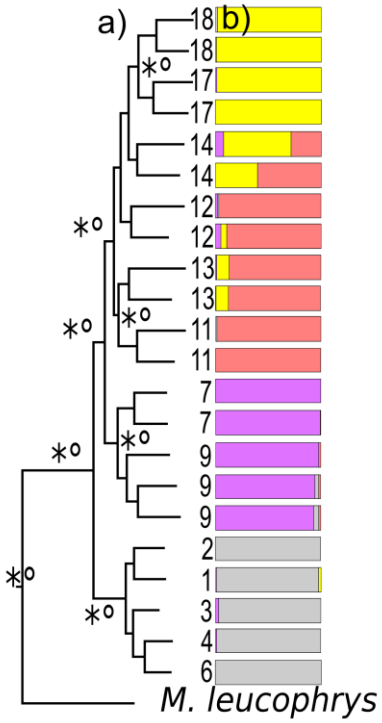

Figure s7

Supplementary information 1

Morphology – plumage patterns

Plumage variation was congruent with the original taxonomic descriptions of each taxa, with subtle differences regarding population polymorphism due to the larger sample size analyzed in this study. Male plumage is conserved over the entire distribution of *M. lugubris*, despite a tendency for darker neutral gray individuals in *M. l. berlepschi* (Fig. s1). The presence of whitish pale neutral gray spots in the edges of the upper and middle wing coverts is highly variable among individuals of *M. l lugubris*, *M. l. femininus* and *M. l. stictopterus*, with a prevalence of 45% in these taxa, but is absent in *M. l. berlepschi*.

The plumage of females is more variable and geographically diagnosable than that of males, enabling the distinction of four groups congruent with current taxonomy (Fig. s2, s3). Females of *M. l. lugubris* (Cabanis, 1847) from Almerim (locality 18 in Fig. 1a; Fig. s2) and Monte Alegre (locality 17 in Fig. 1a; Fig. s2) are diagnosable from the remaining taxa by: (1) throat and central portions of the chest pure white, without black spots or a grayish band separating the throat from the abdomen; (2) sides of the chest varying from Raw Umber to Tawny Olive and usually Robin flanks; (3) abrupt transition to more rufous tones from the back (Rousset) to the crown (cinnamon-brown) and forehead (cinnamon-Rufous); (4) sides of the head Cinnamon-Rufous without signs of black feathers; (5) wings coverts are Rousset with larger and more conspicuous apical spots (more ferruginous than cinnamon) than in other taxa.

M. l. femininus females from Borba (locality 11 in Fig. 1a; Fig. s2) and Autazes

(locality 12 in Fig. 1a; Fig. s2) are distinguishable from other taxa by: (1) white throat with small and subtle black spots (scale) and a subtle black spotted collar delimiting the throat and chest; (2) whitish gray band on chest right after the collar; (3) sides of the chest and flanks more olivaceus than in *M. l. lugubris* individuals; (4) the transition between the back (less ferruginous than Rousset), crown (Prout's Brown) and forehead (Tawny) is less abrupt than in *M. l. lugubris*; (5) The sides of the face including lores, eye ring and auricular areas are black, faded by some tawny feathers producing a pattern on of rufescent shaft lines (Hellmayr, 1910).

Despite the clear diagnosis of the females of *M. l. lugubris* and *M. l. femininus* in the extreme of their distributions (Almeirim for *M. l. lugubris* and Borba for *M. l. femininus*, localities 18 and 11 in Fig. 1a, respectively), these diagnostic characters are blurred in geographically intermediate localities forming a gradual transition between these two forms (Fig. 2, s2). At Itacoatiara (locality 13 in Fig. 1a; s2), despite a prevalence of the *M. l. femininus* plumage pattern there is a noticed increase in the amount of ferruginous feathers on the sides of the head, with some individuals presenting half of the face pure ferruginous (Fig. 2, s3). The opposite is observed in Parintins (locality 14 in Fig. 1a; Fig. s2), where despite the prevalence of typical *M. l. lugubris* plumage pattern, some individuals have black feathers in the auricular region (Fig. s2). Still the individual from Curuá (locality 16 in Fig. 1a; Fig. s2, MPEG 56164) which is 530 km from the closest locality with a pure *M. l. femininus* plumage pattern (Autazes) and 147 km from the closest locality with a pure *M. l. lugubris* plumage pattern (Monte Alegre), has the typical black face pattern of *M. l. femininus* and the crown (cinnamon-

brown) and forehead (cinnamon-Rufous) typical of *M. l. lugubris*. Localities with morphologically intermediate individuals are distributed in an area of at least 500 kilometers (Fig. 1a, s3).

M. l. berlepschi is diagnosable by the following characteristics: (1) white throat with small and subtle black spots (scale) as observed in *M. l. femininus* and an evident collar of black spots with variable intensity delimiting the throat and chest; (2) evident light neutral gray band after the black spotted collar; (3) soft transition between the back (Rousset) and crown (Prout's Brown) without color differences between the crown and the forehead; (4) sides of the head black without signs of ferruginous feathers.

M. l. stictopterus females are very similar to *M. l. femininus* and diagnosable only by the color of the facial mask which is larger and pure black without signs of ferruginous feathers.

Chapter 2

New tales for the Amazonian biogeography: Overlooked floodplain forest birds support climatic oscillation in the Pleistocene as driver of speciation in the Amazon

Abstract

The large Amazonian diversity has been instigating naturalists and researchers to develop diversification hypotheses seeking to understand the historical processes shaping the observed patterns. However, most of the studies on Amazonian biogeography have been conducted on taxa restricted to the upland forest, while other kinds of habitats such as floodplains are still poorly explored. The Amazonian floodplains house the most diverse and largest flooded forests in the world, with high levels of species endemicity. The potential diversification processes shaping the current genetic diversity are still poorly characterized and levels of cryptic diversity are unknown. In the present study, we conducted a comparative phylogenomic analysis of floodplain forest birds testing diversification hypotheses regarding the potential effects of ecological traits and historical environmental changes. We selected three widely distributed species complexes of antbirds (*Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis*) and obtained more than 1,400 SNPs from Ultra Conserved Elements of each of the species complexes. We implemented a comprehensive set of phylogenetic and demographic analyses including model-based approaches to test alternative hypotheses. Our findings supported a similar pattern of diversification between species complexes including time of divergence, demographic history, and current gene flow among populations. Thus, our results supported historical processes related to climatic oscillations that affected the water levels of the main Amazonian rivers during the mid and late Pleistocene as possible drivers of these genomic patterns.

Introduction

The drivers for the diversification of the large Amazonian diversity have been intriguing naturalists and researchers since the 17th century (Wallace 1852; Haffer 1969; Smith et al. 2014). The correlation between patterns of phenotypic distribution and environmental history including geological and paleoclimatic data were the starting point for the development of diversification hypotheses seeking to understand the historical process shaping the observed patterns (Haffer 2008; Antonelli et al. 2010; Leite & Rogers 2013). Phylogenetic and phylogeographic data allowed to directly study diversification scenarios testing assumptions related to specific hypotheses of diversification (Avise 1987; Knowles & Maddison 2002; Knowles 2009). However, most of these assumptions do not present a clear hierarchical separation in space and time, i.e., some assumptions are not mutually exclusive; thus blurring the effects of particular historical events in the history of the current genetic diversity (Patton & Silva 1998; Hall & Harvey 2002). Additionally intrinsic characteristics of each studied organism such as ecology and the evolutionary origin tend to produce distinct dispersal probabilities over physical barriers, suggesting an overall complex scenario that can not be explained by a single generalized hypothesis (Burney & Brumfield 2009; Smith et al. 2014). Nevertheless, this complexity in patterns and processes tends to increase since, in its vast majority, phylogeographic studies of Amazonian taxa have been based on organisms restricted to the nonflooded ombrophilous forests (upland forests), the most abundant and diverse environment in the biome.

The Amazon, however, is an intricate mosaic with several highly distinct environments with endemic fauna and flora, such as the floodplain forest habitats of "várzea" and "igapó" (Junk et al. 2011; Junk et al. 2013). Organisms restricted to these environments tend to respond differently when affected by the same historical events or could be affected by processes that did not affect upland forest species (Capurucho et al 2013). For example, while rivers are considered as barriers to gene flow for upland forest species, they act as corridors for floodplain forest organisms (Aleixo

2006) and in this sense areas with reduced gene flow among populations of floodplain forest species can help understand potential routes of dispersal along rivers by upland forest species.

The Amazonian floodplains house the most diverse and largest flooded forests in the world, covering more than 300,000 km^2 and with high levels of species endemicity - approximately 10% of tree species and 15% of non-aquatic bird species are endemic (Remsen & Parcker 1983; Wittmann et al. 2010; Junk et al. 2011; Wittmann et al. 2013). Unfortunately, phylogeographic studies of endemic terrestrial species are not yet representative enough and consequently, the potential diversification processes shaping the current genetic diversity are poorly characterized and levels of cryptic diversity are unknown (Aleixo 2006; Cadena et al. 2011; O'Neil et al. 2011; Harvey et al. 2017). This is critical since many areas within the Amazon are being used for agriculture and energy production by hydroelectric power plants (Albernaz 2011; Latrubesse et al. 2017). And this pressure is affecting several populations of endangered endemic bird species (BirdLife international 2016).

A few phylogeographic studies conducted on floodplain terrestrial species supported a scenario of widely distributed populations with low levels of genetic diversity, without genetic structure over the entire basin (Aleixo 2006; Cadena 2011; Harvey et al. 2017). These results reinforced assumptions for the diversification of floodplain organisms that suggested that the linear connectivity of flooded habitats enables high levels of gene flow over the entire distribution of these taxa that, in turn, are expected to be good dispersers due the high dynamic of these environments related to the intense effects of the cyclic flooding seasons (Remsen & Parcker 1983). However, despite the apparent linear connectivity of river-created environments, community turnover analyses based on alpha diversity of several taxonomic groups including birds, spiders and plants, suggested a common pattern of diversification that has not been properly explored with genetic data (Hubert & Renno 2006; Albernaz et al. 2011; Vinticinque et al. 2007; Cohn-Haft et al. 2007). Besides, recent studies on bird species restricted to river-created environments supported the

presence of genetic clusters associated to the main tributaries of the Amazon basin matching the community turnovers described above (Choueri et al. 2017; Thom et al. 2018).

From a phylogeographic and population genetics perspective, processes capable to affect the connectivity of the floodplains and to result in distinct phenotypes geographically structured could be related to 1) isolation by distance and river architecture – the linear and extensive distribution of floodplain habitats could potentially affect the pattern of gene flow between populations, forming gradients of genetic diversity correlated with the geographic distance that can promote genetic differentiation (Thomaz et al. 2016). In this scenario it is expected to observe a continuous distribution of the species along its entire geographic range and a gradual transition in patterns of genetic diversity and genetic structure matching the expectations of an isolation by distance model; 2) ecological gradients – spatially segregated habitats could generate local adaptation reducing gene flow among populations occurring in distinct environments. For example, chemical features of the rivers, as the amount of sediments, result in whitewater rivers (sediment-rich) and black/clear water rivers (sediment-poor) that can produce abrupt transitions that are known to affect the pattern of genetic diversity and allele frequencies in species of fishes locally adapted to these environments (Beheregaray et al. 2015). In this scenario it is expected that the species presents a continuous distribution over its entire geographic range with gradual or abrupt transition in patterns of genetic diversity and genetic structure due to reduced but continuous gene flow matching areas of ecological gradients, such as different types of flooded forests (i.e. “várzea” vs. “igapó”); 3) landscape changes related to the evolution of the current Amazon river - distinct paleogeographic hypotheses postulate that the formation of the current Amazon basin could have started at the Miocene (Hoorn et al. 2010), Pliocene (Campbell et al. 2006) or early Pleistocene (Rossetti et al. 2015) with the connection of the sedimentary basin of the Pebas system of western Amazonia with the fluvial system of the Brazillian and Guyana shields from eastern Amazonia. This process of continuous structuring of the main Amazonian rivers could potentially have enabled dispersal and

colonization events of new environments and subsequent diversification (Ribas et al. 2012; Thom & Aleixo 2015). In this scenario it is expected to observe a continuous or disrupted distribution range, divergence times overlapping with geologically more dynamic periods and spatial congruence with known geological features, as structural arches and river capture events as well as no or reduced levels of gene flow due to allopatry; 4) climatic oscillations during the Quaternary affecting river levels - climatic cycles increasing in intensity in the last 0.8 million years (Myr) affected global sea level and rainfall patterns over the Amazon producing periods of lower and higher river levels (Irion et al. 2009; Choueri et al. 2017). Drastic reductions on the average sea and river levels tend to increase the angle of the drainage, water flow speed, and channels' incision, narrowing the distribution of floodplain environments and potentially interrupting the distribution of river-created environments by vanishing areas with sediments. In this scenario it is expected to observe continuously or disrupted distribution range, with recently diverged (<0.8 Myr) populations, high levels of gene flow due to secondary contact and synchronic expansions of effective population sizes given the current interglacial cycle with more widespread floodplains. These four processes are not mutually exclusive hypotheses and do not represent all potential processes capable to affect the diversification of floodplain species. However, they represent a wide variety of scenarios that were poorly explored so far for Amazonian terrestrial floodplain forest organisms.

A recent molecular study on the Ash-breasted Antbird species complex (*Myrmoborus lugubris*) - a river edge forest specialist - supported the presence of distinct genetic clusters associated to the large Amazon tributaries and a recent and complex scenario of diversification related to cyclical river level fluctuations related to climatic oscillations during the Pleistocene (Thom et al. 2018). The unexpected results obtained in this study were based on a single species complex and thus, does not allow to draw more general conclusions about the diversification of the floodplain bird community (but see Choueri et al. 2017; Harvey et al. 2017).

In the present study, we conducted a comparative phylogenomic analysis of floodplain forest

birds testing diversification hypothesis regarding the potential effects of environmental changes. We selected widely distributed species complexes of Antbirds comprising 12 subspecific taxa with similar patterns of geographic distribution. Our findings supported a similar pattern of diversification between species complexes including time of divergence, demographic history, and current gene flow among populations, supporting historical processes related to climatic oscillations during the mid and late Pleistocene that affected the levels of the main Amazon rivers as possible drivers of these genomic patterns

Material and Methods

Study Groups and Floodplain Environments

Antbirds (Thamnophilidae - Aves) are a Neotropical radiation with approximately 241 species (Zimmer & Isler 2017). They are considered good models for phylogeographic studies in the Amazon given their high diversity in the biome - more than 40 species can be found syntopically -, intimate relationship with specific environments, and high intraspecific genetic structure possibly related to poor dispersal abilities (Thom & Aleixo 2015; Fernandes et al. 2012; Fernandes et al. 2014; Zimmer & Isler 2017). Here we selected three antbird species complexes, *Myrmoborus lugubris, Thamnophilus nigrocinereus/T. cryptoleucus,* and *Myrmotherula assimilis,* distributed along floodplain forests of large Amazonian rivers and mainly restricted to river edge forests (Figure 1; Remsen & Parker 1983; Rosemberg 1990; Zimmer & Isler 2017). These river edge forests are "river-created" environments shaped by the intense effects of sedimentation related to the flooding system of large Amazonian rivers. These ephemeral environments are characterized by an early successional stage vegetation spread over a narrow band of margins and islands, usually flooded for more than three months a year (Remsen & Parker 1983; Junk et al. 2011). This habitat is dominated by widespread plant genera such as *Cecropia, Ochrona, Erythrina, Heliconia*, and *Gynerium,* occasionally forming monospecific stands (Junk et al. 2011). In this environment,

islands are frequently affected by river dynamics being constantly eroded at their upstream side and sediment is deposited at their downstream portion, forming a successional gradient (downstream portions tend to be younger), and a potential movement of islands downstream wards. Given this intimate relationship between the river edge forest and specific characteristics of the rivers, we expect that species restricted to this environment tend to be affected by major historical events capable of changing the sedimentation dynamics, such as formation/structuring of the current Amazonian basin and drastic river level fluctuations.

Although river edge forests are widespread over white-water rivers, some populations of the selected taxa also occur on islands of sediment-poor rivers such as Negro and Tapajós, which maintain a distinct vegetation type ("igapó") with similar dynamics related to the flooding system of the Amazonian basin (Junk et al. 2011). However, distinct ecological characteristics between these environments could potentially produce distinct selective pressures increasing diversification rates and reducing gene flow between populations.

Sequence Capture of Ultra Conserved Elements (UCEs) and Bioinformatics

We selected 80, 56 and 38 samples of *M. lugubris*, *T. nigrocinereus/T. cryptoleucus* and *M. assimilis*, respectively. They represent all the twelve described subspecies, covering most of the geographic distribution of these taxa (Table S1; Figure 1). To obtain genome-wide data we used sequence capture of Ultra Conserved Elements (UCEs; Faircloth et al. 2012). Initially, genomic DNA was extracted from blood or muscle tissue using the QIAGEN DNeasy tissue and Blood kit (Valencia, CA) and QIAGEN RNAse, targeting a minimum amount of 2 µg of DNA eluted in ~50 µl of AE buffer. DNA quantification was performed in Qubit 2.0 fluorometer. Library preparation and sequencing of UCEs were outsourced in RAPiD Genomics (Gainsville, FL, USA), following the protocol by Faircloth et. al (2012). Modifications in the original protocol included the use of a probe set targeting 2,312 UCEs (ultraconserved.org) and 97 additional probes targeting exons

typically used in avian phylogenetic studies (Hackett et al. 2008; Kimball et al. 2009), and the use of 150 bp paired-end Illumina Hiseq 2500 sequencing. Samples were sequenced in a multiplexed batch of 96 samples.

Raw read quality was evaluated per individual using FastQC 0.11.4 (Andrews 2014). We used Illumiprocessor to trim off adapter sequences and exclude low quality reads. To assemble *de novo* reads into contigs, we used Trinity 2.4 (Grabherr et al. 2011). The following two steps were performed using PHYLUCE 1.4 (Faircloth et al. 2012; https://github.com/faircloth-lab/phyluce): 1) obtained contigs were matched to UCEs probes with LASTZ using the "match_contigs_to_probes.py". The un-aligned contigs and contigs matching multiple UCE loci were removed; 2) SQL database of matches for all individuals was used to generate alignments in fasta format, allowing for missing individuals and without trimming long ragged-ends with MAFFT (Katoh et al. 2013). Additionally, we trimmed long ragged-ends applying a threshold of 50% missing sequences among individuals with TrimAl (Capella-Gutierrez et al. 2009).

We selected the longest sequence without indels among all individuals per species complex for each locus as a reference for the SNP calling procedure. Initially, we aligned the cleaned and synchronized reads of each individual with the produced reference using BWA (Li & Durbin 2009), enabling up to 4 mismatches per read. We converted the obtained sam files to bam format with Samtools (Li et al. 2009), trimmed read ends covering regions beyond the reference sequence (CleanSam.jar), reassigned reads to groups (AddOrReplaceReadGroups.jar), marked duplicated reads for posterior exclusion (MarkDuplicates.jar), and merged bam files per individual in a single bam file (MergeSamFiles.jar) using Picard (Broad Institute, Cambridge, MA; http://broadinstitute.github.io/picard/). All reads and indels from all individuals from each species complex were realigned using the RealignerTargetCreator and IndelRealigner from the Genome Analyses Tool Kit (GATK; McKenna et al. 2010). Finally, SNPs were called, hard-masking low-quality bases (< Q30) with UnifiedGenotyper and VariantAnotator and indels were masked using

VariantFiltration in GATK. For each species complex, we obtained an unphased VCF file containing all variant sites among individuals. The raw VCF files were filtered for a minimum read depth of > 8 using VCFTOOLS (Danecek et al. 2011). We generated two distinct data sets that were used in further analyses: 1) using VCFTOOLS we randomly selected one SNP per locus, excluding sites with missing data, resulting in a complete SNP matrix; 2) we conducted a read-backed gametic phasing (ReadBackedPhasing) in GATK and incorporated the called SNPs for each individual in the reference sequences using "add_phased_snps_to_seqs_filter.py" from the seqcap_pop package (https://github.com/mgharvey/seqcap_pop; Harvey et al. 2016) in order to obtain full sequences for each locus. Heterozygote sites that did not reach the minimum phasing quality (>20.0) were replaced with the IUPAC ambiguity code. Obtained sequences were aligned with MAFFT (Katoh et al. 2013).

Genetic Structure

To infer the genetic structure we used sNMF (Frichot et al. 2014) to test the best-fit number of ancestral populations (K) for each species complex and to cluster individuals to populations by applying sparse non-negative matrix factorization to compute least-square estimates of ancestry coefficients. When compared to widely used software, such as STRUCTURE (Pritchard et al. 2000) and ADMIXTURE (Alexander et al. 2009), sNMF can process large bi-allelic datasets more efficiently with good accuracy, even in scenarios of departure from the Hardy Weinberg and linkage equilibria (Frichot et al. 2014; Harris et al. 2015). We tested K values ranging from 1 to 8, running 100 replicates for each K value. Four values of alpha regularization parameter (1, 10, 100, 1000) were used in order to test the robustness of the results. Similarly, we implemented the k-means (find.clusters) of ADEGENET 2.0 (Jombart et al. 2011) to infer the number of genetic clusters in the sample. k-means reduces the complexity of the data using a principal component analysis (PCA), obtaining the optimal number of groups by comparing different clustering solutions

with Bayesian Information Criterion (BIC). All components of the initial PCA analysis were retained. To check the concordance between both clustering methods we plotted the sNMF results in function of the k-means classification. Finally, we used a discriminant analysis of principal components (DAPC) in ADEGENET 2.0 to graphically describe the results obtained.

In order to test if the observed genetic diversity among populations of each species complex deviates from an isolation by distance scenario and identify corridors or barriers to gene flow, we used EEMS (Petkova et al. 2015). EEMS maps genetic differentiation by estimating effective migration surface among demes based on a spatially explicit approach, integrating all possible routes of migration between two demes using a Markov Chain Monte Carlo (MCMC) approach to estimate demographic parameters by sampling from their posterior distributions given the observed genetic dissimilarity between individuals. Euclidian genetic dissimilarity matrices between individuals for each species complex were generated using the SNP datasets in ADEGENET (Jombart et al. 2011). Habitat polygons were produced based on the geographic distribution of each species complex and 300 demes were distributed over the habitat area. Each MCMC run was performed for $30x10^6$ generations with the first $5x10^6$ generations excluded as burn-in. Maps were generated using additional features of the EEMS R package.

Phylogenetic Relationships, Species Tree, and Gene Flow

We inferred phylogenies for each species complex using the coalescent-based method implemented in SNAPP (Bryant et al. 2012). SNAPP infers the likelihood of a given species tree using allele frequency of unlinked SNPs bypassing the need to integrate the probabilities of gene trees in function of a given species tree. We used gamma rate priors for alpha and beta parameters, with all other priors with default values. Two replicates of 2.5 million MCMC generations were run with 100,000 burn-in iterations for each species complex. Estimated parameters were sampled every 500 generations. Burn-in values for the MCMC chains were assessed in Tracer (Rambaut &

Drummond 2007). To visualize the posterior distribution of sampled species trees we used DENSITREE v2.2 (Bouckaert et al. 2010).

In order to estimate the relationship graph among populations of each species complex accounting for gene flow, we used TREEMIX 1.12 (Pickrell & Pritchard 2012). This method performs a Gaussian approximation of genetic drift estimating the patterns of population splitting and mixing, accessing the covariance structure of allele frequencies between populations to their most common ancestor. We ran TREEMIX varying from zero to six migration edges for each species complex. The best model was selected observing the significance of migration edges and the residue covariance matrix. We subdivided widely distributed populations by clustering individuals from close localities. Individuals with shared coefficient of ancestry between two populations in sNMF were clustered as distinct groups (potentially introgressed individuals). Additionally, using the TREEMIX package, we implemented the $f3$ statistics (Reich et al. 2009) using the three-pop test for admixture. This test detects correlations in allele frequencies that are not compatible with populations fitting a bifurcating tree. All possible clusters of groups were tested. We assumed as groups the genetic clusters inferred by sNMF. Individuals with shared coefficient of ancestry close to 0.5 (potentially introgressed) were assumed as distinct groups.

Population Demography and Synchronicity of Size Change Among Populations

Since demographic analyses can potentially be affected by distinct bioinformatic pipelines and intrinsic characteristics of the methods (Harvey et al. 2016; Oswald et al. 2017), we estimated demographic parameters using two distinct and complementary approaches. First we used the Generalized Phylogenetic Coalescent Sampler (G-Phocs; Gronau et al 2011) to estimate mutational-scaled (μ) divergence times given in τ ($\tau=T\mu/g$, where T is the absolute divergence time in years, g is the average generation time), effective population sizes based on θ ($\theta=4Ne\mu$, where Ne is the absolute effective population size in number of individuals) and gene flow between populations

measured as migration bands - converted in migrants per generation following $Msx=mst.\theta t$, where *Mst* is the number of migrants per generation from *s* (source) to *t* (target) forward in time, *mst* is the output migration rate and θt is the population mutation rate of the target population (Figure 2). G-Phocs is conditioned to a population phylogeny and a set of parameter priors, using a full probabilistic model of coalescent with migration in a Markov Chain Monte Carlo (MCMC) sampling strategy. G-phocs use as input unphased loci sequences, integrating the likelihood computation over all possible gametic phases. Preliminary MCMC runs were conducted for all three species complexes using default priors defined as a product of a gamma distribution, with $\alpha = 1.0$ and $\beta = 10{,}000$ for τ and θ and $\alpha = 0.002$ and $\beta = 0.00001$ for mutation-scaled migration rates. Each run included 300,000 iterations, sampling every five iterations, and the first 100,000 iterations were conservatively excluded as burn-in. Final runs were performed with $\alpha = 1.0$ and $\beta = 5{,}000$ for τ and θ and $\alpha = 1.2$ and $\beta = 0.01$ for mutation-scaled migration rates, and 500,000 iterations were performed, sampling every five. The fine-tune parameters were automatically estimated during the first 10,000 burn-in iterations. Migration bands were assumed to have significant levels of gene flow if the 95% Bayesian confidence interval of the migration rate did not overlap zero. We assumed an average mutation rate of 2.5×10^{-9} (Nadachowska-Brzyska et al. 2015). In order to reduce the computational time, we sub-sampled between five to 10 individuals per population. We assumed a generation time of one year (see Discussion).

Comparatively, we implemented a model-based approach in Fastsimcoal2 (Excoffier et al. 2013; FSC) in order to explicitly test distinct demographic scenarios and estimate demographic parameters to compare with the G-phocs results. FSC estimates the composite likelihood of a specific scenario of diversification comparing simulated and observed data as well as estimates population genetic parameters such as divergence time, effective population size and gene flow using the joint site frequency spectrum (jSFS) as summary statistics. We converted the VCF files to jSFS in ∂a∂I 1.7 (Gutenkunst et al. 2009). We tested a total of five demographic models per species

complex concerning the presence of gene flow and population size changes after divergence. These models consider the assumptions of the hypotheses of diversification described in the Introduction (Figure 2). The search ranges for the model parameters are shown in Table S2. Since G-phocs is known to produce robust estimates of divergence times even under gene flow, in order to constrain the likelihood search of FSC to meaningful intervals we fixed the first divergence event (forward in time) for each species complex using the mean of the posterior distribution estimated by G-phocs. This procedure allowed us to avoid the need to calculate the proportion of monomorphic sites, using only randomly selected SNPs for the analyses. We performed the composite likelihood search for each model using 100 independent runs keeping the parameters that maximized the composite likelihood across all iterations. For each independent run, parameter optimization was performed with a minimum of 25 and a maximum of 50 cycles of Brent algorithm (ECM cycle) with 100,000 simulations per replicate.

In order to select the best-fit model to the observed data we implemented an information theory procedure based on the Akaike information criterion (AIC=2k-2ln(L), where k is the number of parameters estimated in the model and L the composite likelihood value), and calculated the ΔAIC, relative likelihood, and AIC weights for each model. To obtain 95% confidence intervals for the parameters of the best-fit model we performed 50 parametric bootstraps.

Additionally to the population size changes tested in the FSC multi-population models we tested individual populations identified with sNMF for size changes and the degree of synchronicity of these changes. By analyzing each genetic cluster separately it is possible to optimize the number of selected SNPs for each population. Widely distributed populations were split into more restricted areas (Table S3). As in this analysis the same number of individuals per population is mandatory we used six diploid individuals for each population. These individuals presented no more than 30% of shared ancestry based on sNMF results. In a preliminary analysis, we simulated data under three demographic syndromes for each population with FSC: 1) instantaneous population expansion; 2)

constant size; and 3) instantaneous population contraction. We performed 400,000 coalescent simulations per demographic syndrome applying an Approximate Bayesian Computation procedure to compare the observed folded SFSs with simulations that approximate the data. Given that all the populations presented relatively similar estimated values of effective population size and time of populations size changes in the multi-populations models performed with FSC, we set wide uniform priors for current effective population size (Ne; 10,000-1,000,000), time of instantaneous population size (τ; 1,000-150,00), and proportion of the ancestral population size compared to the current one ($\mathcal{E}$; expansion model: 0.01-0.2, constant size model: 1 and contraction: 5-100), and used the same set of simulations for all 15 populations. ABC model selection was performed with abc R package (Csilléry et al. 2012), using the postpr() function to select the 0.125% simulations (1,500 simulations) with the shortest Euclidean distance between simulated and observed SFS and abc() function to obtain a posterior distribution of the estimated parameters. We assumed the best-fit demographic syndrome for each population as the one with the highest posterior distribution. In order to observe if simulated models could be distinguished based on the SFS as a summary statistic and to check if the best-fit syndrome could generate the observed data, we conducted a PCA analysis.

In order to estimate the coefficient of synchronicity in population size changes, populations with the same best-fit syndrome were grouped by converting the observed SFS of each population into an aggregated SFS (aSFS; Xue & Hickerson 2015). We then implemented a hierarchical ABC (hABC) approach in the Multi-Dice R package (Xue & Hickerson 2017), comparing the observed data with simulated hierarchical demographic models with different degrees of temporal synchronicity in instantaneous population size change performed in Fastsimcoal2. This procedure allows us to obtain the proportion of populations with synchronic demography, expressed as ζ, the mean and variance of the time for a demographic change across populations, and the timing of synchronous population size change (Xue & Hicherson 2015). We performed 1 million simulations,

with the same uniform prior distributions implemented in the demographic syndrome test. The hyper-prior for ζ followed a uniform discrete distribution from 0 (completely asynchronous) to 1 (all populations in demographic synchrony) and each discrete value was treated as a distinct model. We tested for a single synchronic pulse (Ψ = {0,1}). Other uniform parameters priors were: tau.shared.prior ~U{1k, 300k}; epsilon.idio.prior ~U{0.01, 0.2}; NE.idio.prior ~U{10k, 1000k}; tau.idio.prior ~U{1k, 3000k}. We enabled a time buffer of 10k generations to ensure demographic asynchrony between populations, increasing the possibility to identify ζ by reducing the temporal proximity of asynchronous responses. After performing simulations, the hABC hyper-parameter inference followed the same steps of the demographic syndrome selection. We retained the 1,500 simulations with shortest Euclidian distance from the observed aSFS to estimate the posterior distribution of models and parameters. Lastly, we performed 50 "leave one out" cross-validations to access hABC performance (accuracy and bias) using the aSFS. For each cross-validation, one simulation was selected as a pseudo-observed dataset (POD) using the remaining simulations to repeat the hABC estimations using the functions cv4postpr() and cv4abc() from abc R package for ζ and parameter summaries, respectively. For each POD we calculated Pearson's correlation and the root mean square error (rmse) between the observed and estimated values.

Results

Data processing produced raw VCF files with 19,622, 31,453, and 23,569 variant sites for *M. lugubris, T. nigrocinereus/T. cryptoleucus,* and *M. assimilis,* respectively (Table S4). The average site depth was 18.86 (SD = 18.50), 20.41 (SD = 20.92), and 20.58 (SD = 20.86) for *M. lugubris, T. nigrocinereus/T. cryptoleucus*, and *M. assimilis*, respectively. After randomly selecting one SNP per UCE loci we obtained complete matrixes with 1,521, 1,588 and 1,408 SNPs for *M. lugubris, T. nigrocinereus/T. cryptoleucus*, and *M. assimilis*, respectively. After incorporating all variants to the reference sequences we obtained 2,144, 2,241, and 2,251 UCE loci for *M. lugubris,*

T. nigrocinereus/T. cryptoleucus, and *M. assimilis*, respectively. The average number of variable sites per loci was 3.13 (SD = 1.86), 4.83 (SD = 2.93), and 5.49 (SD = 3.66) for *M. lugubris, T. nigrocinereus/T. cryptoleucus*, and *M. assimilis*, respectively (Figure S1). One individual of *M. assimilis* was excluded from the dataset due to sequencing failure. One individual of *T. nigrocinereus* (A3528; cytochrome B identity of 99% with *T. aethiops*), and one of *M. assimilis* (A10383; cytochrome B identity of 99% with *M. menetriesii*) were excluded due to potential misidentification after blasting the mitochondrial DNA (obtained as sequencing off-target) with the NCBI platform (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

The number of genetic clusters obtained with sNMF is partially concordant with the phenotypic variation associated with each subspecies (Figure 1). For *M. lugubris* the best value of masked cross-entropy was achieved with K=4 and alpha = 10 (masked cross-entropy = 0.209927; Figure S2) with different populations distributed at 1) Solimões river; 2) Negro river; 3) Madeira river; and 4) Amazon river + east of Madeira river (Figure 1). For *T. nigrocinereus/T. cryptoleucus* the best values were K=5 and alpha = 10 (masked cross-entropy = 0.206068; Figure S3) with the following populations: 1) Solimões + west of Madeira rivers; 2) Madeira river; 3) Negro river; 4) Amazon + east of Tapajós rivers; and 5) Tapajós river (Figure 1). For *M. assimilis* the best values were K=3 and alpha = 10 (masked cross-entropy = 0.3625; Figure S4) with the populations geographically structured as 1) Solimões + west of Madeira rivers; 2) Madeira river; and 3) Negro + Amazon rivers (Figure 1). The amount of admixture between populations varied among species complex, with individuals from the central portion of Amazonia (where the Negro, Madeira, and Solimões rivers meet to form the Amazon river) showing a considerable higher shared coefficient of ancestry (Figure 1). This region is where different populations present their distribution limits. For *M. lugubris* and *M. assimilis* the individuals with more admixture have shared ancestry of Madeira and Amazonas rivers populations; while those of *T. nigrocinereus/T. cryptoleucus*, of Solimões and Madeira rivers populations.

The number and geographic structure of genetic clusters inferred by k-means (Figures 1, S5-S13) showed similar results to those by sNMF, except for *M. assimilis,* for which the model with the lowest BIC supported K=2, with the population of Madeira river merged with Negro/Amazonas rivers populations (Figures S7, S10, S13). The EEMS plots for the three species complexes (Figures 3, S14) supported a considerable reduction of effective migration at the central portion of the Amazonian basin (where differentiated populations get into contact) compared to the values observed within each of these populations. This suggests that the genetic structure observed in sNMF and k-means cannot be explained solely by a scenario of isolation by distance. Similarly, the highest estimated effective diversity parameter values were observed in the areas of transition between populations (central portion of the Amazonian basin; Figure S15), which reflects the genetic dissimilarity between individuals, indicating a secondary contact scenario with low levels of gene flow when compared to the within population gene flow.

Species tree estimated with SNAPP supported distinct splitting patterns for the three species complexes (Figure 1). In *M. lugubris* the first diversification event was between the Solimões population and the eastern Amazonia populations (posterior probability, PP = 1.0), followed by the split of the Negro river population from those of Madeira and Amazonas rivers (PP = 1.0), and a subsequent divergence between Madeira and Amazonas rivers populations (PP = 0.93). The sampled topologies that were incongruent with the consensus tree supported the Negro and Madeira rivers populations as a clade with Amazonas river population as a sister group (Figure 1). For *T. nigrocinereus/T. cryptoleucus* the consensus species tree supported a sister relationship between Solimões and Madeira rivers populations (PP = 1.0) and between Tapajós and Amazonas rivers populations (PP = 1.0). However, the relationship of the Negro river population and these two groups was unclear (PP = 0.93). Overlapped sampled topologies in densitree indicated three distinct scenarios: 1) Negro river as sister to Solimões and Madeira rivers populations; 2) Negro river as sister to Tapajós and Madeira rivers populations; and 3) a simultaneous split of the three groups. In

M. assimilis the consensus topology suggested with maximum statistical support the sister relationship between Madeira and Negro + Amazonas rivers populations with the Solimões river population as sister to this group.

Populations relationship graphs estimated with TREEMIX had similar topologies as of those from SNAPP analyses and indicated the presence of significant migration edges among populations in all three species complexes (Figure 4). In *M. lugubris*, three migration edges were significant (jack-knife p-value <0.05). In *T. nigrocinereus/T. cryptoleucus* two migration edges, both between Solimões and Madeira subgroups were significant. In *M. assimilis* five significant migration edges were inferred and were in general related to the Negro+Amazonas and Solimões populations (Figure 4). The *f3* statistics partially corroborated the TREEMIX results. For *M. lugubris* the only significant estimate value supported the individuals with a shared coefficient of ancestry between Madeira and Amazonas populations in sNMF as an introgressed population between Negro and Amazonas populations (*f3* = -0.0007; SD = 0.0005; Z-score = -6.39102). In *T. nigrocinereus/T. cryptoleucus* the only significant value supported that the individuals with a shared coefficient of ancestry between Solimões and Madeira populations in sNMF are an introgressed population between these groups (*f3* = -0.0061; SD=0.0016; Z-score = -7.3499). In *M. assimilis*, despite the five significant migration edges inferred with TREEMIX representing 13.4 % of the total variation in the data set, no estimated value was significant using threepop test (Table S5).

The presence of moderate to high levels of gene flow between nonsister populations can directly affect topologies (Leache et al 2014; Thom et al 2018). Given the inferred gene-flow pattern revealed by TREEMIX in *M. lugubris* and *M. assimilis* and the low node support in *T. nigrocinereus/T. cryptoleucus* inferred by SNAPP, we performed a topology test using Fastsimcoal2 assuming the presence of gene flow before testing the demographic models and estimating parameters. For *M. lugubris* we tested two distinct models varying the topology in the presence of constant gene flow since divergence (Figure 2c): 1) consensus tree obtained with

SNAPP, (Solimões,(Negro,(Madeira,Amazonas))); and 2) alternative topology obtained with SNAPP clustering the populations that exchange alleles as identified by TREEMIX (Solimões,(Amazonas,(Negro, Madeira))). The model with the best likelihood supported the alternative topology with Madeira and Negro as sister populations with maximum relative weight based on AIC (ln(L) = -14905.47; Table S6). As previously reported by Thom et al (2018), the high amount of current gene flow, including the presence of a wide hybrid zone between Madeira and Amazonas populations, is potentially affecting phylogenetic estimations that do not account for gene flow, such as SNAPP. For *T. nigrocinereus/T. cryptoleucus* we tested two models concerning the relationship of Negro river population with the other two groups (Figure 2d): 1) consensus tree obtained with SNAPP ((Solimões, Madeira),(Negro,(Tapajós,Amazonas))); and 2) alternative topology obtained with SNAPP (((Solimões, Madeira),Negro),(Tapajós,Amazonas)). The model with the best likelihood supported with maximum relative weight the consensus topology of SNAPP with Negro as sister population of Tapajós and Amazonas clade (ln(L) = -17777.802; Table S6). In *M. assimilis* we tested three possible topologies (Figure 2e): 1) consensus tree obtained with SNAPP, (Solimões,(Madeira,Amazonas/Negro)); 2) (Amazonas/Negro,(Solimões,Madeira)); and 3) (Solimões,(Madeira, Amazonas/Negro)). The model with the best likelihood supported with 68.3 % of relative weight the consensus topology of SNAPP (ln(L) = -3345,271; Table S6).

Demographic Modeling

The demographic parameter values (mean and 95% HPD) estimated with the isolation-migration model of G-phocs were highly consistent among preliminary and final runs (results not shown), even despite the low Effective Sample Size (ESS) and high Auto Correlation Time (ACT) of some parameters that suggested a weak signal in the data for specific parameters (Tables 1-3). Preliminary runs suggested that 100,000 generations in the MCMC were a conservative burn-in, allowing all parameter values of primary interest to converge. Distinct gamma priors used in

preliminary runs lead to very similar estimates and were not sufficiently diffuse to strongly influence the posterior distributions. For *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* populations the current effective population size (Ne) spanned from 34,180 (95% HPD 31,600-36,590) to 138,480 (95% HPD 129,180-147,840) diploid individuals (Tables 1, 2). For *M. assimilis* the Ne of Solimões (83,550; 95% HPD 72,270-94,270) and Amazonas/Negro populations (117,760; 95% HPD 107,190-128,380) were similar to values obtained for the other species, but the Madeira population was larger (480,860 95% HPD; 419,190-546,020; Table 3). Mean divergence times ranged from 25,320 to 225,960 years ago (ya), with the last divergence events for *M. lugubris, T. nigrocinereus/T. cryptoleucus* and *M. assimilis* occurring at 25,360 ya (95% HPD 22,280-28,600), 58,800 ya (95% HPD 53,920-63,320) and 38,200 ya (95% HPD 35,120-41,320), respectively. Migration bands varied considerably between pairs of populations (Tables 1-3; pairwise comparisons overlapping 0 migration are not shown). For *M. lugubris* the highest migration values were observed between Madeira and Amazonas populations (M→A = 1.47 and A→M = 2.01 migrants per generation; 95% HPD 1.40-1.80 and 1.57-2.46, respectively). The remaining pairwise comparisons overlapped 0 or were lower than 0.6 migrants per generation (Table 1). In *T. nigrocinereus/T. cryptoleucus* all pairwise migration bands were lower than 1 migrant per generation. The migration bands with the highest values were from Solimões to Madeira (S→M = 0.51 migrants per generation; 95% HPD 0.46-0.57; Table 2) and between Madeira and Amazonas (M→A = 0.42 and A→M = 0.54 migrants per generation; 95% HPD 0.32-0.53 and 0.44-0.64, respectively). For *M. assimilis* migration bands were very high and asymmetric between Solimões and Madeira (S→M = 0.58 and M→S= 8.00 migrants per generation; 95% HPD 0.40-0.76 and 3.29-12.93, respectively) and between Solimões and Amazonas/Negro (S→A/N = 1.05 and A/N→S = 3.69 migrants per generation; 95% HPD 0.68-1.42 and 2.69-4.71, respectively; Table 3).

The FSC results supported demographic models including the presence of gene flow for all taxa (Tables 4-7). For *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* the best-fit demographic

model supported, with maximum relative contribution on AIC, isolation with gene flow with all divergence times fixed with G-phocs estimated values and instantaneous population size changes after divergence (Figure 2c, 2d; Table 4). For *M. assimilis* the best-fit demographic model supported, with 0.97 of relative contribution on AIC, isolation with gene flow without population size change (Figure 2e; Table 4). Parameter estimations based on the best-fit models for Ne were relatively congruent with G-phocs results, being most of the values no more than 2 fold different (Tables 1-3, 5-7). Divergence time between Amazonas/Negro and Madeira population for *M. assimilis* was estimated at 166,612 years ago (ya) with relatively high gene flow between these populations (> 3 individuals per generation), which was not detected in the G-phocs model. Migration values were usually higher in FSC best-fit model than those estimated by G-phocs (Tables 1-3, 5-7).

For the demographic syndromes tested with Multi-Dice, we considered 15 populations, 12 recognized with sNMF (Figure 1) plus three that were split from the widely distributed population from Solimões of *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* and the Amazonas/Negro population (Negro and Amazonas regions were split) of *M. assimilis*. The ABC procedure supported 10 populations with instantaneous population expansion as the best-fit model (PP > 0.99). Population size change intensity varied from ~5 to ~11 fold (based on mean PP), occurring from ~8,000 to ~110,000 ya (Table S3). The models tested were clearly diagnosed in the PCA analysis (Figure S16). Regarding the remaining five populations, for three of them (Negro of *M. lugubris*; Negro and Tapajos of *T. nigrocinereus/T. cryptoleucus*) the best-fit model supported populations with stable sizes, and for the *T. nigrocinereus/T. cryptoleucus* Madeira population and *M. lugubris* Solimões 2 population, a signal of instantaneous contraction with an intensity of ~5 fold dated at ~27,000 ya and ~133,000 ya, respectively (Table S3). The results obtained of demographic changes analyzing models of individual populations were partially congruent with the multi-population models for *M. lugubris* and *T. nigrocinereus/T. cryptoleucus*, as only two populations had distinct

estimated syndromes: 1) Madeira of *M. lugubris* presented an ancestral population size approximately ten fold smaller (population expansion) in the single population model and five fold (subtle population contraction) larger in the multi-population model; and 2) Amazonas population of *T. nigrocinereus/T. cryptoleucus* showed an ancestral population size ten fold smaller (population expansion) in the single population model and 4.5 fold larger (subtle population contraction) in the multi-population model. For *M. assimilis*, despite the best multi-population model supported no population size change, all four single population best models supported relative subtle population expansion (Table S3).

Given the high proportion of expanding populations (ten out of 15) we applied the co-demography hABC approach of Multi-Dice on these ten groups. A synchronic pulse of expansion was detected including almost all of the 10 tested populations (mean $\zeta = 0{,}9301$, with a mean cross-validation correlation, r = 0.8356; Figures S18, S19; Tables S7, S8) occurring at ~ 0.2 Myr (r = 0.7325; 95% HPD = 0.084, 0.295 Myr) with a mean dispersion index = 72,010 (r = 0.707), suggesting a temporally wide but strong synchronic pulse that occurred during the Quaternary climatic cycles. It is noteworthy that the estimated time for the synchronic expansion pulse of Multi-Dice is older than any estimated population expansion time obtained with the single population ABC analyses.

Discussion

Here we used genome-wide variation to model the diversification and demographic history of three antbird species complexes restricted to the Amazonian floodplain forest, supporting a shared pattern of genetic structure, gene flow, and population size changes in these taxa. This is one of the first studies to report a pattern of population structure for terrestrial species associated to the main Amazonian rivers (Solimões, Negro, Madeira, Amazon, and Tapajós) and the first to perform a comparative approach to explore the diversification of this group of organisms (Choueri et al. 2017;

Thom et al. 2018). In contrast to results observed in upland forest organisms - mostly birds and mammals (Ribas et al. 2012; Boubli et al. 2015) - it is not possible to observe any sign of differentiation between opposite margins or between island and margins, but a shared pattern of genetic structure associated to the main Amazonian sub-basins.

We found that the central portion of the Amazon basin, more specifically the confluence of Solimões, Negro, and Madeira rivers represent a suture zone delimiting the distribution of multiple populations with distinct levels of secondary contact and gene flow (Figure 1; Tables 1-3). Previous studies of floodplain bird species supported widely distributed populations without genetic structure over the entire Amazon basin (Aleixo 2006; Cadena et al. 2011; but see Choueri et al. 2017; Thom et al. 2018). These contrasting results could be related to habitat specificity of the species used in these studies, as *Xiphorhynchus obsoletus, Dendroplex kieneri*, and *Cacicus icterocephalus* are more generalist (they occur in the majority of flooded environments, including open vegetation areas) than the taxa analyzed here. Concordantly to the results obtained here, transitions in the floodplain community composition based on alpha diversity between the western (Solimões river) and eastern (Amazon river) Amazonian basin were reported for birds, spiders, fishes, and trees (Cohn-Haft et al. 2007; Venticinque et al. 2007; Farias & Hrbek 2008; Albernaz et al. 2011). This could suggest that the shared location of this biotic transition indicates a shared mechanism of diversification. Unfortunately, these studies were based on distributional data only and it is not known if the temporal pattern is also shared among them. The population divergence pattern we observed indicated that the populations from Solimões river were the first to diverge in *M. lugubris* and *M. assimilis* while in *T. nigrocinereus/T. cryptoleucus* there was a simultaneous divergence of the Solimões and Madeira populations (Tables 1-3). The usual low dispersal capacity of antbirds (Zimmer & Isler 2017) associated to an intimate relationship with river edge forests and islands suggests that isolation and connection of populations is linked to the high sedimentation dynamics of rivers (Remsen & Parker 1983; Rosemberg 1990; Choueri et al. 2017). The ephemeral nature of

river edge forests is congruent with the recent (Mid and Late Pleistocene) scenario of diversification inferred here (Tables 1-3). Given the lack of knowledge about the generation time of the taxa analyzed here, we assumed one year as the generation time to estimate divergence times. However, this value could be underestimating the true splitting dates. Given the absence of specific ecological information about Neotropical passerine birds distinct studies assumed a wide range of values for generation time of antbirds, including one year (Oswald et al. 2017), 2.33 years (Maldonado-Coelho et al. 2012), and 5.1 years (BirdLife International 2017). Hence, if the generation time of the species studied is between one and five years, the diversification process of the three species occurred in the last 1.2 million years (Myr), with the oldest divergence events occurring at approximately 0.56, 1, and 1.25 Myr ago for *M. lugubris, T. nigrocinereus/T. cryptoleucus*, and *M. assimilis*, respectively (assuming 5.1 years as generation time and the upper limit of the 95% HPD of G-phocs).

Another important point is that all three species complexes studied diverged from their sister taxa at least 5 Myr ago (Bravo 2012). Therefore there is a time window of several million years that we do not have information about the coalescent process of these populations, indicating high levels of population extinctions or less opportunity for speciation when compared to other Neotropical environments (Weir & Schluter 2008). However, the latter hypothesis is less likely given the complex and dynamic formation history of the Amazonian floodplains (Rossetti et al. 2015). These observations are in agreement with previous studies suggesting low genetic diversity for widespread species (Aleixo et al. 2006).

Another factor that could have affected the diversification pattern of the studied species is interspecific competition, as it is noteworthy that along most of their distribution these species only occur on islands even where river edge forests are available on river margins (Rosemberg 1990; but see Armacost Jr. & Capparella 2012). The three studied species complexes have closely related species that are potential ecological competitors that occupy upland forest and floodplain environments on the river margins (*M. leucophrys, M. myotherinus*, and *M. melanurus* for *M.*

lugubris; *M. menetriessi* and *M. axillaris* for *M. assimilis*; and *T. amaznonicus*, *T. aethiops*, and *Sakesphorus luctuosu*s for *T. nigrocinereus/T. cryptoleucus*). Although there is a general lack of knowledge about interspecific interactions of Amazonian birds, playback experiments conducted over a successional gradient in an Amazonian white-water river in Peru supported strong interspecific aggression without territorial overlap between several species of birds, including *M. leucophrys* and *M. myotherinus* (Robinson & Terborgh 1995). These results could indicate interspecific exclusion in upland and floodplain habitat transitions, which could nevertheless promote ecological speciation and habitat specialization leading to higher rates of extinction or speciation (Rabosky 2013). River islands usually submerge during the flooding season and species that cannot disperse horizontally moving from the islands to other environments have to have the ability to disperse vertically to upper forest strata, which could be a limitation for understory species. Hence interspecific competition could have resulted in habitat specialization to islands in the case of the species complexes studied here. However, experiments must be performed to test the effects of interspecific aggression and habitat segregation on floodplain specialists.

Concordant Demography and Diversification of Amazonian Floodplain Species

We found that despite intrinsic ecological characteristics of each species studied their diversification patterns are similar (Figure 1; Tables 1-3, 5-7). Despite the potential strong effects of isolation by distance reported for taxa with linear distributions (Thomaz et al. 2016), our results point to a scenario that cannot be explained solely by this process. Similarly, given the allopatric distribution of some populations, the demographic signature of expansion in most of the populations occurring in white-water rivers and the variable levels of gene flow in each species complex (Figure 1; Tables 1-3, 5-7), ecological gradients solely seem not to result on the observed pattern.

Although the posterior distributions of divergence times estimated with G-phocs for the

three taxa were extremely narrow (see Full vs. Composite Likelihood; Tables 1-3) they overlapped or were similar, suggesting that these divergences could have been generated by the same process of diversification. For example the first and second divergence in *T. nigrocinereus/T. cryptoleucus* (~0.18 Mya) occurred close to the first divergence event in *M. assimilis* (~0.225 Mya), both related to the split of western and eastern Amazonian populations. Around 0.06 and 0.1 Mya populations of *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* from Solimões, Negro, lower Amazon, and Tapajós rivers diverged (Tables 1-3). Between 0.02 and 0.04 Mya, Madeira populations of *M. assimilis* and *M. lugubris* diverged (Tables 1-3). Similarities were also observed in the patterns of gene flow, with populations from white-water rivers (Solimões, Madeira, and Amazonas) sharing higher levels of migrants per generation than populations from black and clear-water rivers (Tables 1-3, 5-7). Concordantly, size changes were detected only in populations occurring at white-water rivers, mostly subtle expansions (Table S3), suggesting that the higher dynamics of sediment-rich environments seem to have produced more unstable environments (Junk et al. 2013). Choueri et al. (2017), studying landscape genetics of island specialist birds of the Negro river, reported for four antbird species - including the three species studied here - subtle population expansions. Additionally, the authors suggested that the river-created environments possibly suffered rearrangements and interruption in the connectivity of habitats but not drastic reductions or expansions as observed for upland forest species (Aleixo 2006; Fernandes et al. 2012; Ribas et al. 2012; Thom & Aleixo 2015). Our results support this conclusion, mainly for clear and black-water rivers such as Negro and Tapajós.

The period for the initial diversification of the three taxa studied here took place during the Mid-Late Pleistocene, after the major reorganization of Amazonian rivers and changes in the sedimentation dynamics following the Içá Formation deposition (Rossetti et al. 2015). The formation and structuring of the current transcontinental Amazon river occurred through the connection of a western lacustrine system of what is today the Solimões river with the fluvial

system of the eastern Amazon through the breaching of the Purus Arch (Rossetti et al. 2015). However the time when this process occurred is under intense debate, with estimated dates ranging from the Miocene (~10-12 Mya) based on stratigraphic information and increase of sedimentation rate in the Amazon fan (Hoorn et al. 1995; Gorini et al. 2013; Figueiredo et al. 2009, 2010), to the Plio-Pleistocene (Campbell et al. 2006; Nogueira et al. 2013), and more recently, to Mid-Late Pleistocene (Rossetti et al. 2015). Our results corroborate a more recent formation of the current Amazon river since the connection among the Solimões and Amazon systems could have enabled dispersal and colonization events to new environments – mainly from western (sediment-rich system) to eastern Amazon - which is coincident with our results that indicate that the more ancient splits separated the Solimões populations. With the rearrangement of the drainage system after the Mid-Pleistocene, widespread floodplains mainly in western Amazon were converted into upland environments following river channel incisions – reducing the distribution and fragmenting floodplain environments - potentially affecting the establishment of river edge forests and islands (Rossetti et al. 2015; Choueri et al. 2017). After this process, intensification of Milankovitch cycles about 0.8 Mya yielded pronounced alterations in the levels of Amazonian rivers mainly due to global sea level fluctuations and alteration of the rainfall pattern in western Amazon and Andean foothills, modulating the pattern of sedimentation over the Amazon basin and potentially producing expansion of the floodplains following aggradation or contraction and channel incision (Baker et al., 2001; Cheng et al., 2013; Govin et al., 2014; Zhang et al., 2016). As stated by Iron et al. (2009), with a reduction of more than 100 meters in the global sea level during the Last Glacial Maximum (LGM; 0.02 Mya) part of the Amazon floodplains experienced an increase in the slope of the rivers, resulting in a much higher rate of erosion and incision in the main channels, contracting the distribution of floodplains potentially isolating portions of specific environments. However, the effects of sea level fluctuation in water discharge and sediment yield tend to be minor or absent in areas farther from the coast, such as Negro and Solimões rivers (>1,000 km from the coast) which

tend to be more affected by tectonics or climate variations in the headwaters that are the source of the sediments (Schumm 1993; Archer 2005; Choueri et al. 2017). Despite the lack of geological knowledge on which areas in the Amazon basin are more prone to be affected by climatic cycles, our data suggested that the central portion of the basin, where the sedimentary system of the Solimões river flows into the hard rock bed of the Amazon craton seems to have faced critical alterations in the landscape. Habitat changes in this region seem to have simultaneously isolated upstream populations at the Solimões, Negro, and Madeira rivers from those of the lower course of the Amazon river. This matches the genetic structure we observed (Figure 1) and overlapped or has similar posterior distributions of divergence times (Tables 1-3). Additionally, the high synchronicity in time of the demographic expansions supported a scenario of white-water river populations being simultaneously affected (Tables 7, 8), potentially responding to habitat contraction in the central portion of the basin. Landscape genetics studies focusing on species restricted to river edge forests that applied spatial explicit methods to estimate geographic expansion could produce new evidence supporting this hypothesis (Peter & Slatkin 2015; He et al. 2017).

Another important point is related to the high dynamics of channel incision promoting the erosion of the lower course of rivers such as Negro and Tapajós, which during periods of sediment accumulation cannot fill the eroded area due to the low amount of sediments transported by these rivers. As a consequence, large lakes - so-called Ria Lakes - were formed producing large areas without islands and with poorly developed floodplains on the margins (Franzinelli and Igreja 2002; Irion et al. 2009; Bertani et al. 2015). The presence of current Ria lakes in the lower course of Negro and Tapajós rivers is congruent with results obtained here of populations restricted to these rivers being recently isolated (except for *M. assimilis*) and with low levels of gene flow with other populations (Tables 1-3). Additionally, there is no evidence suggesting the presence of *M. lugubris* and *T. nigrocinereus* in the lower course of these rivers (Choueri et al. 2017; Thom et al. 2018).

Full vs. Composite Likelihood

Here we implemented two distinct coalescent-based approaches, G-phocs and FSC/Multi-Dice, widely used for modeling demographic histories. G-phocs is a full likelihood method that uses as input loci sequences containing all the variable sites (Gronau et al. 2011). FSC approximates the likelihood of a model based on simulations, reducing the complexity of genomic data by using the site frequency spectrum as summary statistic calculated from a single SNP per locus (Excoffier et al. 2013). G-phocs demands previous information about the phylogenetic relationships between populations and does not allow for testing the best-fit of distinct demographic models, acting as a parameter estimation method. On the other hand, FSC is more flexible, enabling to draw distinct demographic scenarios and approximating the likelihood of each model given the data, testing distinct hypotheses of diversification and parameters. Despite the intrinsic differences of these approaches we obtained consistent parameter estimates for the best-fit model for *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* as obtained in FSC with all divergence times fixed with G-phocs estimates. Besides, among the 21 estimated population size parameters, including current and ancestral populations, only four differed by more than twice comparing both methods (Tables 1-3, 5-7). The major differences in parameter estimations between these methods were related to gene flow, as in general FSC resulted in higher values than G-phocs (Tables 1-3, 5-7). This was possibly caused by the fact that FSC models force migration to stop before any instantaneous population size changes backwards in time, while G-phocs assumes an isolation-migration model. Hence a shorter time interval of gene flow in FSC models could overestimate migration values when compared to G-phocs. However, the proportion of migrants per generation between a given pair of populations when compared to other pairwise estimates is consistent between these approaches. Gronau et al. (2011) suggest that for the recent diversification of human populations, robust quantification of gene flow is difficult to achieve, however it does not appear to bias estimates of divergence times, while in scenarios with strong signal of gene flow and well-differentiated population, fairly

complex scenarios involving multiple migration bands can be recovered. In our study, the relatively well-differentiated populations and the similar proportions of pairwise gene flow between populations inferred with FSC indicated that migration estimates were fairly robust (Tables 1-3, 5-7). Similarly, Thom et al. (2018) supported the presence of a wide hybrid zone between Madeira and Amazonas populations of *M. lugubris* based on phenotypic characters, which matched the higher migration bands inferred by G-phocs (Table 1). Another noteworthy result of G-phocs was the excessively narrow range of 95% HPD (Tables 1-3). Narrow ranges of HPD can underestimate the uncertainty of parameter estimation and are potentially related to poor mixing of the MCMC with a reduced exploration of the likelihood surface (Gronau et al. 2011).

We observed that the results obtained with the FSC multi-population models and with the hABC approach implemented in Multi-Dice (which uses FSC as simulator and analyzes each population separately) were concordant. Both detected demographic expansions for most of the *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* populations but not for *M. assimilis* populations (Tables 5-7, s3). However, the magnitude of the instantaneous population size change inferred by FSC multi-population models was always smaller - in general, less than five times the size of the current population – than those inferred by the Multi-Dice approach (Tables 5-7, s3). This difference could be due to the inclusion of possibly more informative genetic information in the Multi-Dice analysis, given that a new set of SNPs was called for each population - which could increase the resolution of the analysis - as well as by the fact that Multi-Dice single population analysis did not account for recent gene flow from other populations that could mimic a population expansion event. Hence, our comparative pipeline of demographic methods including empirical and model-based approaches enabled us to observe potential trade-offs, limitations and concordances between the methods, which is highly appreciated when inferring historical processes (Oswald et al. 2016).

Implications for Studying Diversification in the Amazon Floodplains

In general phylogeographic studies of upland forest birds support the formation and structuring of large Amazonian rivers as the main cause of diversification, refuting the refugia hypothesis given the lack or low intensity of demographic changes (Fernandes et al. 2012; Ribas et al. 2012; Fernandes et al. 2014; Thom & Aleixo 2015; Ferreira et al. 2017). Our data suggested that climatic oscillations drove speciation on floodplain Amazonian taxa by an analogous scenario modeled by the refugia hypothesis: the contraction of river edge forest into isolated blocks restricted to the main Amazonian rivers.

The study of organisms occurring in historically sub-sampled environments by phylogeographic approaches can bring new insights to interpret the diversification of the Amazon as a whole. Here we support that the Amazonian floodplains are highly dynamic and constantly affected by paleoclimatic oscillations, mainly in the confluence of Solimões, Negro, and Madeira rivers, potentially isolating populations restricted to these rivers. In this sense, periods when floodplain distribution was contracted potentially allowed gene flow and dispersion over rivers by upland forest species. For example, the central portion of the basin - western Negro river - is the only region where *Myrmoborus leucophrys* occurs northwards of the Solimões river. Similarly, species mainly distributed in the north of the Amazon river, such as *Leucopternis melanops* and *Terenura spodioptila*, also present populations in the eastern portion of the Madeira river in the south margin of the Amazon River (del Hoyo et al. 2018). In the lower Tapajós river, *Willisornis nigrigula* populations are found in opposite margins, but not in the middle or upper courses of the river, suggesting that the current Ria lake was narrower in the past (del Hoyo et al. 2018). These examples could be good models to test dispersion of upland forest species over floodplain environments.

Taxonomy, Cryptic Speciation and Conservation of Floodplain Species

Although taxonomy is not the central scope of our study, the genetic structure observed matched diagnosable phenotypes. This suggests that at least for *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* their subspecies could be recognized as full species based on distinct species concepts (Queiroz et al. 2007). Thom et al. (2018) inferred the phylogenetic relationships of *M. lugubris* populations based on molecular and phenotypic data and found a complex scenario of diversification including a wide hybrid zone between nonsister lineages, suggesting the recognition of all four taxa as full species under the phylogenetic species concept (Cracraft 1983; Queiroz et al. 2007). A similar pattern was observed for *T. nigrocinereus/T. cryptoleucus* as all taxa present phenotypic diagnosys that match genetic clusters, except *T. n. kulczynskii* that, despite the unique plumage, clusterd with *T. n. nigrocinereus.* Another noteworthy result was the paraphyly of *T. nigrocinereus*, given that phylogenetic estimations (Figure 1) supported the sister relationship between *T. n. tschudi* and *T. cryptoleucus*. In *M. assimilis* two subspecies are currently recognized based on slight plumage variation (Zimmer & Isler 2003): *M. a. assimilis* (found along almost all the distribution of the species) and *M. a. transamazonica* (restricted to the eastern part of the distribution on the Amazon River in Pará state). Our results did not match these taxa as the Amazon river individuals clustered with Negro river individuals (Figure 1). Additionally, we identified a cryptic population restricted to the Madeira river (Figure 1). Thus, it would be important to test the species status of these populations using similar approaches as described by Thom et al. (2018).

Populations restricted to river edge forests are expected to have lower population sizes than upland forest species, given the linear and narrow distribution restricted to the large Amazonian rivers (Aleixo 2006). The association of these species to secondary and early stage regenerated forests potentially allows the use of areas near to non-extensive agricultural activities (Armacost Jr. & Capparella 2012). However, extensive livestock production such as cattle and water buffalo, as well as dam construction could highly affect these species (Barten et al. 2004; Albernaz et al. 2012). This latter disturbance mainly affects populations occurring on rivers running over the Amazon

craton, such as Tapajós river, with high potential for hydroelectric energy production (Latrubesse et al. 2017). Damming rivers for hydroelectric power plants can drastically change sedimentation dynamics and the lakes formed by the dams can vanish sediment islands and floodplains, also acting as barriers for river edge specialists as is currently observed in the Ria Lakes (Latrubesse et al. 2017). The Tapajós populations of *T. nigrocinereus* match the subspecies *T. n. huberi*, a phenotypically diagnosable taxon endemic to the middle and upper course of the Tapajós river (Zimmer& Isler 2003). The effective population size of *T. n. huberi* was estimated to be less than 40,000 individuals (Table 2, 6), the smallest within *T. nigrocinereus*. Thus, given the large number of hydroelectric power plants planned to be constructed along the Tapajós river (Latrubesse et al. 2017), conservation strategies have to be designed in order to prevent future threats and local extinctions of this and other floodplain forest species.

References

Aleixo A. 2006. Historical diversification of floodplain forest specialist species in the Amazon: A case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). Biol. J. Linn. Soc. 89:383–395.

Alexander D.H., Novembre J., Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research. 19:1655–1664.

Andrews S. 2014. FastQC: A quality control tool for high-throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ Accessed March 9, 2017.

Antonelli A., Quijada-Mascareñas A., Crawford A.J., Bates J.M., Velazco P.M., Wüster W. 2010. Molecular studies and phylogeography of Amazonian tetrapods and their relation to geological and climatic models. Pp. 386–404 In: Hoorn C. & Wesselingh F. (eds.) Amazonia: Landscape and Species Evolution, 1st ed. Wiley-Blackwell, Oxford.

Archer A.W. 2005. Review of Amazonian depositional systems. Pp. 17-39 In: Blum M.D., Marriott

S. B., Leclair S. F. (eds), Fluvial Sedimentology VII. Blackwell Publishing Ltd, Oxford.

Baker P.A., Rigsby C.A., Seltzer G.O., Fritz S.C., Lowenstein T.K., Bacher N.P., Veliz C. 2001. Tropical climate changes at millennial and orbital timescales on the Bolivian Altiplano. Nature. 409:698–701.

Barthem R.B. 2004. Aquatic Biota. Pp. 62-79 In: Capobianco J.P.R., Verıíssimo A., Moreira A., Sawyer D., dos Santos I., Pinto L.P. (eds) Biodiversity in the Brazilian Amazon: Assessment and Priority Actions for Conservation, Sustainable use and Benefit Sharing. Estação Liberdade & Instituto Socioambiental, São Paulo.

Beheregaray L.B., Cooke G.M., Chao N.L., Landguth E.L. 2015. Ecological speciation in the tropics: insights from comparative genetic studies in Amazonia. Frontiers in Genetics. 5.

Bertani T.C., Rossetti D.F., Hayakawa E.H., Cohen M.C.L. 2014. Understanding Amazonian fluvial rias based on a Late Pleistocene-Holocene analog. Earth Surface Processes and Landforms. 40:285–292.

BirdLife International. 2016. The IUCN Red List of Threatened Species 2016: e.T22680736A92875297.http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22680736A92875297.en. Downloaded on 20 November 2017.

Boubli J.P., Ribas C., Alfaro J.W.L., Alfaro M.E., Silva M.N.F.D., Pinho G.M., Farias I.P. 2015. Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. Molecular Phylogenetics and Evolution. 82:400–412.

Bouckaert R.R. 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics. 26:1372–1373.

Bravo G.A. 2012. Phenotypic and niche evolution in the antbirds (Aves, Thamnophilidae). Ph.D. Dissertation, Louisiana State University, Baton Rouge.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N.A., Roychoudhury A. 2012. Inferring species

trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution. 29:1917–1932.

Burney C.W., Brumfield R.T. 2009. Ecology predicts levels of genetic differentiation in Neotropical birds. The American Naturalist. 174:358–368.

Cadena C.D., Gutiérrez-Pinto N., Dávila N., Chesser R.T. 2011. No population genetic structure in a widespread aquatic songbird from the Neotropics. Molecular Phylogenetics and Evolution. 58:540–545.

Campbell K.E., Frailey C.D., Romero-Pittman L. 2006. The Pan-Amazonian Ucayali Peneplain, late Neogene sedimentation in Amazonia, and the birth of the modern Amazon river system. Palaeogeography, Palaeoclimatology, Palaeoecology. 239:166–219.

Capella-Gutierrez S., Silla-Martinez J.M., Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25:1972–1973.

Capurucho J.M.G., Cornelius C., Borges S.H., Cohn-Haft M., Aleixo A., Metzger J.P., Ribas C.C. 2013. Combining phylogeography and landscape genetics of *Xenopipo atronitens* (Aves: Pipridae), a white sand campinas specialist, to understand Pleistocene landscape evolution in Amazonia. Biological Journal of the Linnean Society. 110:60–76.

Cheng H., Sinha A., Cruz F.W., Wang X., Edwards R.L., D'Horta F.M., Ribas C.C., Vuille M., Stott L.D., Auler A.S. 2013. Climate change patterns in Amazonia and biodiversity. Nature Communications. 4:1411.

Choueri É.L., Gubili C., Borges S.H., Thom G., Sawakuchi A.O., Soares E.A.A., Ribas C.C. 2017. Phylogeography and population dynamics of antbirds (Thamnophilidae) from Amazonian fluvial islands. Journal of Biogeography. 44:2284–2294.

Cohn-Haft M., Naka L.N., Fernandes A.M. 2007. Padrões de distribuição da avifauna da várzea dos rios Solimões-Amazonas. Pp 287-324 In: Albernaz A.L. (eds). Conservação da Várzea: Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ProVárzea, Manaus.

Cracraft J. 1983. Species concepts and speciation analysis. Current Ornithology. 1:159–187.

Csilléry K., François O., Blum M.G.B. 2012. abc: an R package for approximate Bayesian computation (ABC). Methods in Ecology and Evolution. 3:475–479.

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., Depristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T., Mcvean G., Durbin R. 2011. The variant call format and VCFtools. Bioinformatics. 27:2156–2158.

Excoffier L., Dupanloup I., Huerta-Sánchez E., Sousa V.C., Foll M. 2013. Robust demographic inference from genomic and SNP data. PLoS Genetics. 9.

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

Farias I.P., Hrbek T. 2008. Patterns of diversification in the discus fishes (*Symphysodon* spp. Cichlidae) of the Amazon basin. Molecular Phylogenetics and Evolution. 49:32–43.

Fernandes A.M., Wink M., Sardelli C.H., Aleixo A. 2014. Multiple speciation across the Andes and throughout Amazonia: the case of the spot-backed antbird species complex (*Hylophylax naevius*/*Hylophylax naevioides*). Journal of Biogeography. 41:1094–1104.

Fernandes A.M., Wink M., Aleixo A. 2012. Phylogeography of the chestnut-tailed antbird (*Myrmeciza hemimelaena*) clarifies the role of rivers in Amazonian biogeography. Journal of Biogeography. 39:1524–1535.

Franzinelli E., Igreja H. 2002. Modern sedimentation in the Lower Negro river, Amazonas state, Brazil. Geomorphology. 44:259–271.

Frichot E., Mathieu F., Trouillon T., Bouchard G., François O. 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics. 196:973–983.

Govin A., Chiessi C.M., Zabel M., Sawakuchi A.O., Heslop D., Hörner T., Zhang Y., Mulitza S. 2013. Terrigenous input off northern South America driven by changes in Amazonian

climate and the North Brazil current retroflection during the last 250 ka. Climate of the Past Discussions. 9:5855–5898.

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

Gronau I., Hubisz M.J., Gulko B., Danko C.G., Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nature Genetics. 43:1031–1034.

Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genetics. 5.

Hackett S.J., Kimball R.T., Reddy S., Bowie R.C.K., Braun E.L., Braun M.J., Chojnowski J.L., Cox W.A., Han K.-L., Harshman J., Huddleston C.J., Marks B.D., Miglia K.J., Moore W.S., Sheldon F.H., Steadman D.W., Witt C.C., Yuri T. 2008. A phylogenomic study of birds reveals their evolutionary history. Science. 320:1763–1768.

Haffer J. 1969. Speciation in Amazonian forest birds. Science. 165:131–137.

Haffer J. 1997. Biodiversity and Conservation. Biodiversity and Conservation. 6:451–476.

Hall J.P.W., Harvey D.J. 2002. The Phylogeography of Amazonia revisited: new evidence from Riodinid butterflies. Evolution. 56:1489.

Harris S.E., O'Neill R.J., Munshi-South J. 2014. Transcriptome resources for the white-footed mouse (*Peromyscus leucopus*): new genomic tools for investigating ecologically divergent urban and rural populations. Molecular Ecology Resources. 15:382–394.

Harvey M.G., Aleixo A., Ribas C.C., Brumfield R.T. 2017. Habitat association predicts genetic diversity and population divergence in Amazonian birds. The American Naturalist.

190:631–648.

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

He Q., Prado J.R., Knowles L.L. 2017. Inferring the geographic origin of a range expansion: Latitudinal and longitudinal coordinates inferred from genomic data in an ABC framework with the program x-origin. Molecular Ecology. 26:6908-6920.

Albert J.S., Reis R.E.dos. 2011. Historical biogeography of neotropical freshwater fishes. University of California Press.

Hoorn C., Wesselingh F.P., Steege H.T., Bermudez M.A., Mora A., Sevink J., Sanmartin I., Sanchez-Meseguer A., Anderson C.L., Figueiredo J.P., Jaramillo C., Riff D., Negri F.R., Hooghiemstra H., Lundberg J., Stadler T., Sarkinen T., Antonelli A. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science. 330:927–931.

Howard D.J., Berlocher S.H. 1999. Endless forms: Species and speciation. Oxford University Press Inc., New York.

Hubert N., Renno J.-F. 2006. Historical biogeography of South American freshwater fishes. Journal of Biogeography. 33:1414–1436.

Irion G., Müller J., Morais J.O., Keim G., Mello J.N.D., Junk W.J. 2009. The impact of Quaternary sea level changes on the evolution of the Amazonian lowland. Hydrological Processes. 23:3168–3172.

Jombart T., Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 27:3070–3071.

Junk W.J., Piedade M.T.F., Lourival R., Wittmann F., Kandus P., Lacerda L.D., Bozelli R.L., Esteves F.A., Cunha C.N.D., Maltchik L., Schöngart J., Schaeffer-Novelli Y., Agostinho

A.A. 2013. Brazilian wetlands: their definition, delineation, and classification for research, sustainable management, and protection. Aquatic Conservation: Marine and Freshwater Ecosystems. 24:5–22.

Junk W.J., Piedade M.T.F., Schöngart J., Cohn-Haft M., Adeney J.M., Wittmann F. 2011. A classification of major naturally-occurring Amazonian lowland wetlands. Wetlands. 31:623–640.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 30:772–780.

Kimball R.T., Braun E.L., Barker F.K., Bowie R.C., Braun M.J., Chojnowski J.L., Hackett S.J., Han K.-L., Harshman J., Heimer-Torres V. 2009. A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolution. 50:654–660.

Latrubesse E.M., Arima E.Y., Dunne T., Park E., Baker V.R., D'Horta F.M., Wight C., Wittmann F., Zuanon J., Baker P.A.(Burney and Brumfield 2007; Harvey et al. 2017; Ribas C.C., Norgaard R.B., Filizola N., Ansar A., Flyvbjerg B., Stevaux J.C. 2017. Damming the rivers of the Amazon basin. Nature. 546:363–369.

Leaché A.D., Harris R.B., Rannala B., Yang Z. 2013. The influence of gene flow on species tree estimation: a simulation study. Systematic Biology. 63:17–30.

Leite R.N., Rogers D.S. 2013. Revisiting Amazonian phylogeography: insights into diversification hypotheses and novel perspectives. Organisms Diversity and Evolution. 13:639–664.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25:1754–1760.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079.

Maldonado-Coelho M. 2012. Climatic oscillations shape the phylogeographical structure of Atlantic

Forest fire-eyes (Aves: Thamnophilidae). Biol. J. Linn. Soc. 105:900–924.

Mckenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., Depristo M.A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research. 20:1297–1303.

Nadachowska-Brzyska K., Li C., Smeds L., Zhang G., Ellegren H. 2015. Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. Current Biology. 25:1375–1380.

Oswald J.A., Overcast I., Mauck W.M., Andersen M.J., Smith B.T. 2017. Isolation with asymmetric gene flow during the nonsynchronous divergence of dry forest birds. Molecular Ecology. 26:1386–1400.

O'Neill J.P., Lane D.F., Naka L.N. 2011. A cryptic new species of thrush (Turdidae: *Turdus*) from western Amazonia. The Condor. 113:869–880.

Parolin P., Wittmann F. 2010. Tree phenology in Amazonian floodplain forests. Ecological Studies Amazonian Floodplain Forests.:105–126.

Peter B.M., Slatkin M. 2015. The effective founder effect in a spatially expanding population. Evolution. 69:721–734.

Petkova D., Novembre J., Stephens M. 2015. Visualizing spatial population structure with estimated effective migration surfaces. Nature Genetics. 48:94–100.

Pickrell J.K., Pritchard J.K. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genetics. 8.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Queiroz K.D. 2007. Species concepts and species delimitation. Systematic Biology. 56:879–886.

Rabosky D.L. 2013. Diversity-dependence, ecological speciation, and the role of competition in

macroevolution. Annual Review of Ecology, Evolution, and Systematics. 44:481–502.

Rambaut A., Drummond A.J. 2007. Tracer v1.4, http://tree.bio.ed.ac.uk/software/ tracer/. Accessed March 9, 2017.

Reich D., Thangaraj K., Patterson N., Price A.L., Singh L. 2009. Reconstructing Indian population history. Nature. 461:489–494.

Remsen J.V., Iii T.A.P. 1983. Contribution of river-created habitats to bird species richness in Amazonia. Biotropica. 15:223.

Ribas C.C., Aleixo A., Nogueira A.C.R., Miyaki C.Y., Cracraft J. 2011. A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. Proceedings of the Royal Society B: Biological Sciences. 279:681–689.

Robinson S.K., Terborgh J. 1995. Interspecific aggression and habitat selection by Amazonian birds. The Journal of Animal Ecology. 64:1.

Rosenberg G.H. 1990. Habitat specialization and foraging behavior by birds of Amazonian river islands in Northeastern Peru. The Condor. 92:427.

Rossetti D.F., Cohen M.C., Tatumi S.H., Sawakuchi A.O., Cremon É.H., Mittani J.C., Bertani T.C., Munita C.J., Tudela D.R., Yee M., Moya G. 2015. Mid-Late Pleistocene OSL chronology in western Amazonia and implications for the transcontinental Amazon pathway. Sedimentary Geology. 330:1–15.

Schumm S.A. 1993. River response to base level change: implications for sequence stratigraphy. The Journal of Geology. 101:279–294.

Smith B.T., Mccormack J.E., Cuervo A.M., Hickerson M.J., Aleixo A., Cadena C.D., Pérez-Emán J., Burney C.W., Xie X., Harvey M.G., Faircloth B.C., Glenn T.C., Derryberry E.P., Prejean J., Fields S., Brumfield R.T. 2014. The drivers of tropical speciation. Nature. 515:406–409.

Thom G., Aleixo A. 2015. Cryptic speciation in the white-shouldered antshrike (*Thamnophilus aethiops*, Aves – Thamnophilidae): The tale of a transcontinental radiation across rivers in

lowland Amazonia and the northeastern Atlantic Forest. Molecular Phylogenetics and Evolution. 82:95–110.

Thom G., Amaral F.R., Hickerson M.J., Aleixo A., Araujo-Silva L.E., Ribas C.C., Choueri E., Miyaki C.Y. 2018. Phenotypic and genetic structure support gene flow generating gene Tree discordances in an Amazonian floodplain endemic species. Systematic Biology. https://doi.org/10.1093/sysbio/syy004.

Thomaz A.T., Christie M.R., Knowles L.L. 2016. The architecture of river networks can drive the evolutionary dynamics of aquatic populations. Evolution. 70:731–739.

Rosenberg G.H. 1990. Use of mainland habitats by supposed river-island obligate birds along the Amazon river in Peru. The Condor. 114:56–61.

Venticinque E.M., Rego F.N.A.A., Brescovit A.D., Rheims C.A., Ruiz G.R.S. 2007. A araneofauna (Arachnida, Araneae) das varzeas do Rio Amazonas: padrões de distribuicao e estado do conhecimento atual. Pp 179-198 In: Albernaz A.L. (eds). Conservação da Várzea: Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ProVárzea, Manaus.

Wallace A.R. 1854. On the monkeys of the Amazon. Journal of Natural History Series 2. 14:451–454.

Weir J.T., Schluter D. 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science. 315:1574–1576.

Wittmann F., Schongart J., Junk W. J. 2010. Phytogeography, species diversity, community structure and dynamics of central Amazonian floodplains forests. Pp. 61-102 In: Junk W. J., Piedade M. T. F., Wittmann F., Schongart J., Parolin P. (eds), Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management. Springer, New York.

Wittmann F., Householder E., Piedade M.T.F., Assis R.L.D., Schöngart J., Parolin P., Junk W.J. 2012. Habitat specifity, endemism and the neotropical distribution of Amazonian white-water floodplain trees. Ecography. 36:690–707.

Xue A.T., Hickerson M.J. 2015. The aggregate site frequency spectrum for comparative population genomic inference. Molecular Ecology. 24:6223–6240.

Xue A.T., Hickerson M.J. 2017. Multi-dice: r package for comparative population genomic inference under hierarchical co-demographic models of independent single-population size changes. Molecular Ecology Resources. 17:e212-e224

Zhang Y., Zhang X., Chiessi C.M., Mulitza S., Zhang X., Lohmann G., Prange M., Behling H., Zabel M., Govin A., Sawakuchi A.O., Cruz F.W., Wefer G. 2016. Equatorial Pacific forcing of western Amazonian precipitation during Heinrich Stadial 1. Scientific Reports. 6:35866.

Zimmer K., Isler M.L. 2017. Typical antbirds (Thamnophilidae). In: del Hoyo J., Elliott A., Sargatal J., Christie D.A., de Juana E. (eds.). Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona (retrieved from https://www.hbw.com/node/52291 on 20 November 2017).

Table 1: Demographic parameters estimated for *Myrmoborus lugubris* obtained with G-phocs.

Parameters	**Ne solimoes**	**Ne negro**	**Ne madeira**	**Ne amazonas**	**Ne negmad**	**Ne amanegmad**	**Ne ancestral**	**Tdiv negmad**
mean	44500	62330	57480	44090	119260	200040	79600	25360
median	44490	62300	57390	44080	119030	199030	79600	25320
95% HPD Lower	41450	54800	51100	40250	96450	163660	74600	22280
95% HPD Upper	47510	69550	64580	47990	140420	238290	84530	28600
ACT	482.0	2807.3	1422.3	839.1	3478.4	530.9	937.1	5992.4
ESS	518.7	89.1	175.8	297.9	71.9	470.9	266.8	41.7
Parameters	**Tdiv amanegmad**	**Tdiv ancestral**	**m_sol→mad**	**m_mad→sol**	**m_sol→neg**	**m_neg→sol**	**m_mad→ama**	**m_ama→mad**
mean	62080	107880	0.42	0.39	0.53	0.32	1.47	2.01
median	61880	107760	0.42	0.38	0.53	0.32	1.41	2.00
95% HPD Lower	53840	100280	0.28	0.23	0.39	0.16	1.40	1.57
95% HPD Upper	70520	115640	0.56	0.56	0.68	0.49	1.81	2.46
ACT	4411.8	2357.8	2825.6	2618.6	1335.8	3234.0	3755.1	1292.1
ESS	56.7	106.0	88.5	95.5	187.2	77.3	66.6	193.5

Ne – Effective population size number in diploid individuals; negmad – ancestral population of Negro and Madeira populations; amanegmad – ancestral population of Amazonas, Negro, and Madeira populations; ancestral – ancestral population of the group; Tdiv – divergence time in years; m_x→y – estimated migration band (posterior distribution not overlapping zero) from x to y forward in time in number of individuals per generation; sol – Solimões population; mad – Madeira population; neg – Negro population; ama – Amazonas population.

Table 2: Demographic parameters estimated for *Thamnophilus nigrocinereus/T. cryptoleucus* obtained with G-phocs.

Parameters	**Ne solimoes**	**Ne negro**	**Ne madeira**	**Ne amazonas**	**Ne tapajos**	**Ne solmad**	**Ne amatap**	**Ne amatapneg**
mean	82340	84050	55410	138480	34180	54190	215350	477130
median	82330	84020	55390	138450	34190	46580	214210	476080
95% HPD Lower	78930	79330	51740	129180	31600	2340	146010	424960
95% HPD Upper	85920	88900	59080	147840	36590	126380	287940	530780
ACT	555.4	697.8	972.9	2872.3	4782.3	2630.8	3165.3	68108
ESS	810.2	644.8	462.5	156.6	94.1	171.1	142.1	660.1
Parameters	**Ne ancestral**	**Tdiv solmad**	**Tdiv_amatap**	**Tdiv amatapneg**	**Tdiv ancestral**	**m_sol→mad**	**m_mad→sol**	**m_sol→neg**
mean	117800	189320	58800	79960	189640	0.51	0.03	0.31
median	117780	189280	58880	79880	189680	0.51	0.03	0.31
95% HPD Lower	112150	182360	53920	74760	182600	0.46	0.01	0.26
95% HPD Upper	123440	196120	63320	85280	196400	0.57	0.05	0.36
ACT	1395.1	506.4	5957.2	1277.5	2519.1	662.7	4601.7	651
ESS	322.5	789.7	67.1	313.1	158.7	603.5	89.9	614.3
Parameters	**m_neg→sol**	**m_neg→mad**	**m_mad→neg**	**m_mad→ama**	**m_ama→mad**			
mean	0.32	0.31	0.13	0.42	0.54			
median	0.32	0.31	0.12	0.42	0.54			
95% HPD Lower	0.26	0.23	0.08	0.32	0.44			
95% HPD Upper	0.37	0.38	0.18	0.53	0.64			
ACT	2131.3	1384.1	1748.9	2260.3	1833.1			
ESS	196.9	288.9	228.7	176.9	218.2			

Ne – Effective population size number in diploid individuals; solmad – ancestral population of Solimões and Madeira populations; amatap - ancestral population of Amazonas and Tapajós populations; amatapneg – ancestral population of Amazonas, Tapajós, and Negro populations; ancestral – ancestral population of the group; Tdiv – divergence time in years; m_x→y – estimated migration band (posterior distribution not overlapping zero) from x to y forward in time in number of individuals per generation; sol – Solimões population; mad – Madeira population; neg – Negro population; ama – Amazonas population; tap – Tapajós population.

Table 3: Demographic parameters estimated for *Myrmotherula assimilis* obtained with G-phocs.

Parameters	**Ne solimoes**	**Ne negro/amazonas**	**Ne madeira**	**Ne negamamad**	**Ne ancestral**	**Tdiv negamamad**	**Tdiv amanegmad**	**m_sol→mad**
mean	83550	117760	480860	276090	81610	38200	225960	0.58
median	83460	117620	479430	275970	81680	38160	225760	0.57
95% HPD Lower	72270	107190	419190	258060	73370	35120	207960	0.40
95% HPD Upper	94270	128380	546020	293830	89580	41320	244720	0.76
ACT	5552.6	2460.8	607.4	1215.6	3179.5	1543.5	4354.5	6386.6
ESS	72.0	162.6	658.4	329.1	125.8	259.2	91.9	62.6
Parameters	**m_mad→sol**	**m_sol→negama**	**m_negama→sol**					
mean	8.00	1.05	3.69					
median	7.91	1.05	3.67					
95% HPD Lower	3.29	0.68	2.69					
95% HPD Upper	12.93	1.42	4.71					
ACT	4110.6	6377.9	4760.1					
ESS	97.3	62.7	84.1					

Ne – Effective population size number in diploid individuals; negamamad – ancestral population of Negro, Amazonas, and Madeira populations; ancestral – ancestral population of the group; Tdiv – divergence time in years; m_x→y – estimated migration band (posterior distribution not overlapping zero) from x to y forward in time in number of individuals per generation; sol – Solimões population; mad – Madeira population; negama – Negro/Amazonas population.

Table 4: Composite likelihood (Max ln(L)), Akaike information criterion (AIC) and relative contribution (Weights) for each of the demographic models tested for *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis*. In bold: best model.

Taxon	Model	Max ln(L)	nparam	AIC	Relative Weights
M. lugubris	0mig_0siz	-15003.699	9	69112.59	2.72E-201
	0mig_siz	-14986.31	17	69048.51	2.23E-187
	mig_0siz	-14880.81	21	68570.66	1.29E-083
	mig-siz	-14804.56	29	68235.52	7.72E-011
	mig_siz_All_G-phocs	**-14795.316**	**27**	**68188.95**	**1.00E+000**
T. nigrocinereus/ T. cryptoleucus	0mig_0siz	-17880.913	12	82368.65	9.37E-239
	0mig_siz	-17790.154	22	81970.69	2.44E-152
	mig_0siz	-17777.802	26	81921.80	1.00E-141
	mig-siz	-17672.851	36	81458.49	4.08E-41
	mig_siz_All_G-phocs	**-17633.765**	**33**	**81272.49**	**1.00E+000**
M. assimilis	0mig_0siz	-3415.642	6	15741.61	5.25E-66
	0mig_siz	-3398.839	12	15676.23	8.27E-52
	mig_0siz	**-3347.767**	**12**	**15441.04**	**9.76E-01**
	mig-siz	-3346.776	18	15448.47	2.37E-02
	mig_siz_All_G-phocs	-3352.539	16	15471.01	3.02E-07

0mig_0siz – demographic model with both migration between populations and population size changes fixed in zero; 0mig_siz - demographic model with migration between populations fixed to zero but with population size changes; mig_0siz – demographic model with migration between populations but population size changes fixed to zero; mig_siz – demographic model with migration between populations and population size changes; mig_siz_ALL_G-phocs – demographic model with migration between populations and population size changes with all divergence times fixed with G-phocs mean estimations (tables 1-3); 2*nparam – number of parameters in the models.

Table 5: Parameter values of simulated models in Fastsimcoal 2 with the best composite likelihood for *Myrmoborus lugubris*. Models are described in Figure 2. In bold: best model.

models	**Ne Solimoes**	**Ne Negro**	**Ne Madeira**	**Ne Amazonas**	**Ne n-m**	**Ne n-m-a**	**Ne anc**	**Ms→n**	**Mn→s**	**Ms→m**
0mig_0siz	600341	34294	46334	67628	1302674	749677	138407	***	***	***
0mig_siz	260767	271458	83113	331515	446847	383069	248803	***	***	***
mig_0siz	172387	39433	35312	34550	851096	533331	171403	0.22	0.62	0.01
mig_siz	313537	36900	26456	60856	624787	531157	285615	0.03	0.60	0.02
mig_siz_Gph	**577831**	**187753**	**41753**	**64241**	**592914**	**402991**	**88469**	**0.04**	**0.50**	**0.02**
models	**Mm→s**	**Ms→a**	**Ma→s**	**Mn→m**	**Mm→n**	**Mn→a**	**Ma→n**	**Mm→a**	**Ma→m**	**TDIVn-m**
0mig_0siz	***	***	***	***	***	***	***	***	***	6079
0mig_siz	***	***	***	***	***	***	***	***	***	76601
mig_0siz	0.17	0.00	0.01	0.04	0.01	0.05	0.03	0.33	0.42	6037
mig_siz	0.00	0.64	0.07	2.13	0.00	0.44	0.02	0.72	0.82	20945
mig_siz_Gph	**0.02**	**1.56**	**0.07**	**1.67**	**0.00**	**1.88**	**0.00**	**0.10**	**5.77**	*******
models	**TDIVn-m-a**	**Epsilon s**	**Epsilon n**	**Epsilon m**	**Epsilon a**	**Tau s**	**Tau n**	**Tau m**	**Tau a**	
0mig_0siz	12320	***	***	***	***	***	***	***	***	
0mig_siz	84230	5.92	4.17	9.86	2.84	30108	39486	3973	59768	
mig_0siz	8351	***	***	***	***	***	***	***	***	
mig_siz	21352	0.23	29.59	0.52	0.34	79055	3793	4824	6780	
mig_siz_Gph	*******	**0.22**	**3.96**	**5.50**	**0.36**	**20203**	**21632**	**20255**	**20025**	
mig_siz_Gph	*******	**0.22**	**3.96**	**5.50**	**0.36**	**20203**	**21632**	**20255**	**20025**	

0mig_0siz – demographic model with both migration between populations and population size changes fixed in zero; 0mig_siz - demographic model with migration between populations fixed to zero but with population size changes; mig_0siz – demographic model with migration between populations but population size changes fixed to zero; mig_siz – demographic model with migration between populations and population size changes; mig_siz_ALL_G-phocs – demographic model with migration between populations and population size changes with all divergence times fixed with G-

phocs mean estimations (tables 1-3); Ne – Effective population size number in diploid individuals ; n-m – ancestral population of Negro and Madeira populations; n-m-a – ancestral population of Amazonas, Negro, and Madeira populations; anc – ancestral population of the group; Mx→y – estimated migration band from x to y forward in time in number of individuals per generation; Tdiv – divergence time in years; s – Solimões population; m – Madeira population; n – Negro population; a – Amazonas population; Epsilon – proportional size of the ancestral population in relation to the current one when size change was allowed; Tau – time of population size change.

Table 6: Parameter values of simulated models in Fastsimcoal 2 with the best composite likelihood for *Thamnophilus nigrocinereus/T. cryptoleucus*. Models are described in Figure 2. In bold: best model.

models	**Ne Solimoes**	**Ne Negro**	**Ne Madeira**	**Ne Amazonas**	**Ne Tapajos**	**Ne s-m**	**Ne a-t**	**Ne a-t-n**	**Ne anc**	**Ms→n**
0mig_0siz	1100085	58877	470456	58173	15416	432143	551000	597060	547722	***
0mig_siz	222213	218433	170016	152138	146217	557512	1182411	748574	431834	***
mig_0siz	1694375	302923	475169	292775	59937	840392	1725742	1617784	1508096	1.44
mig_siz	197213	113262	308502	71154	28175	217254	606171	437210	406848	1.01
mig_siz_Gph	**231432**	**196767**	**127012**	**100777**	**116840**	**311298**	**447604**	**211583**	**356013**	**1.36**
models	**Mn→s**	**Ms→a**	**Ma→s**	**Ms→m**	**Mm→s**	**Mn→a**	**Ma→n**	**Mn→m**	**Mm→n**	**Ma→m**
0mig_0siz	***	***	***	***	***	***	***	***	***	***
0mig_siz	***	***	***	***	***	***	***	***	***	***
mig_0siz	0.00	0.09	0.02	1.50	0.01	0.11	0.07	0.02	0.52	0.01
mig_siz	0.61	0.12	1.09	1.41	1.98	0.00	0.72	0.02	0.20	0.19
mig_siz_Gph	**0.52**	**0.06**	**1.15**	**0.30**	**1.04**	**0.04**	**0.57**	**0.22**	**0.08**	**0.01**
models	**Mm→a**	**Ma→t**	**Mt→a**	**TDIVs-m**	**TDIVa-t**	**TDIVa-t-n**	**Epsilon s**	**Epsilonn**	**Epsilon m**	**Epsilon a**
0mig_0siz	***	***	***	142235	11256	17301	***	***	***	***
0mig_siz	***	***	***	88826	74214	79850	8.53	1.73	6.92	2.99
mig_0siz	0.03	0.32	0.00	143073	42739	128754	***	***	***	***
mig_siz	0.01	0.02	0.01	138753	11452	38726	0.55	0.30	2.38	2.49
mig_siz_Gph	**0.00**	**0.26**	**0.01**	*******	*******	*******	**0.28**	**1.09**	**6.25**	**4.43**
models	**Epsilon t**	**Tau s**	**Tau n**	**Tau m**	**Tau a**	**Tau t**				
0mig_0siz	***	***	***	***	***	***				
0mig_siz	0.51	5172	53627	38036	1586	23424				
mig_0siz	***	***	***	***	***	***				
mig_siz	0.23	23604	17523	135141	8104	7840				

mig_siz_Gph	**0.35**	**79733**	**60520**	**19378**	**11112**	**13927**

0mig_0siz – demographic model with both migration between populations and population size changes fixed in zero; 0mig_siz - demographic model with migration between populations fixed to zero but with population size changes; mig_0siz – demographic model with migration between populations but population size changes fixed to zero; mig_siz – demographic model with migration between populations and population size changes; mig_siz_ALL_G-phocs – demographic model with migration between populations and population size changes with all divergence times fixed with G-phocs mean estimations (tables 1-3); Ne – Effective population size number in diploid individuals ; s-m – ancestral population of Solimoes and Madeira populations; a-t – ancestral population of Amazonas and Tapajos populations; a-t-n – ancestral population of Amazonas, Tapajos and Negro populations; anc – ancestral population of the group; Mx→y – estimated migration band from x to y forward in time in number of individuals per generation; Tdiv – divergence time in years; s – Solimoes population; m – Madeira population; n – Negro population; a – Amazonas population; t – Tapajos population; Epsilon – proportional size of the ancestral population in relation to the current, when size change was allowed; Tau – time of population size change.

Table 7: Parameter values of simulated models in Fastsimcoal 2 with the best composite likelihood for *Myrmotherula assimilis*. Models are described in Figure 2. In bold: best model.

models	**Ne Solimoes**	**Ne Amazonas/Negro**	**Ne Madeira**	**Ne m-a/n**	**Ne anc**	**Mm→s**	**Ms→m**	**Mm→a/n**	**Ma/n→m**	**Ms→a/n**
0mig_0siz	1504511	3350761	2195186	872433	909417	***	***	***	***	***
0mig_siz	337653	619961	419983	894083	29444	***	***	***	***	***
mig_0siz	358432	525318	765488	195303	249877	0.35	0.20	3.05	3.94	2.37
mig-siz	270456	442542	606222	559583	190556	0.70	0.01	2.00	5.71	1.79
mig_siz_Gph	255566	446985	760556	555334	42213	1.23	1.07	1.58	0.10	3.23
models	**Ma/n→s**	**TDIVa/n-m**	**Epsilon s**	**Epsilon a/n**	**Epsilon m**	**Tau s**	**Tau a/n**	**Tau m**		
0mig_0siz	***	147111	***	***	***	***	***	***		
0mig_siz	***	32156	9.13	6.12	2.83	19168	19276	7938		
mig_0siz	3.07	166612	***	***	***	***	***	***		
mig-siz	3.63	219653	0.93	0.23	1.95	160324	133830	136006		
mig_siz_Gph	1.75	***	8.74	2.03	5.91	30271	34111	29936		

0mig_0siz – demographic model with both migration between populations and population size changes fixed in zero; 0mig_siz - demographic model with migration between populations fixed to zero but with population size changes; mig_0siz – demographic model with migration between populations but population size changes fixed to zero; mig_siz – demographic model with migration between populations and population size changes; mig_siz_ALL_G-phocs – demographic model with migration between populations and population size changes with all divergence times fixed with G-phocs mean estimations (tables 1-3); Ne – Effective population size number in diploid individuals ; m-a/n – ancestral population of Madeira and Amazonas/Negro populations; anc – ancestral population of the group; Mx→y – estimated migration band from x to y forward in time in number of individuals per generation; Tdiv – divergence time in years; s – Solimoes population; m – Madeira population; a/n – Amazonas/Negro population; Epsilon – proportional size of the ancestral population in relation to the current, when size change was allowed; Tau – time of population size change.

Table S1. Individuals sampled of *Myrmoborus lugubris*, *Myrmotherula assimilis*, and *Thamnophilus nigrocinereus/T. cryptoleucus*. **n° -** number of sampled localities in Figure 1.

n°	Taxon	Population	Field Id.	Institution	Voucher	Sex	Locality	Latitude	Longitude
2	*M. l. berleschi*	Solimoes	ETA 001	MPEG	T22430	female	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. l. berleschi*	Solimoes	ETA 002	MPEG	T22431	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. l. berleschi*	Solimoes	ETA 004	MPEG	T22433	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. l. berleschi*	Solimoes	ETA 009	MPEG	T22438	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. l. berleschi*	Solimoes	ETA 052	MPEG	T22480	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. l. berleschi*	Solimoes	ETA 053	MPEG	T22481	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
3	*M. l. berleschi*	Solimoes	ETA 073	MPEG	T22496	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. l. berleschi*	Solimoes	ETA 074	MPEG	T22497	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. l. berleschi*	Solimoes	ETA 093	MPEG	T22511	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. l. berleschi*	Solimoes	ETA 094	MPEG	T22512	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. l. berleschi*	Solimoes	ETA 096	MPEG	T22514	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. l. berleschi*	Solimoes	ETA 114	MPEG	T22527	female	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. l. berleschi*	Solimoes	ETA 115	MPEG	T22528	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
5	*M. l. berleschi*	Solimoes	ETA 162	MPEG	T23519	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W
5	*M. l. berleschi*	Solimoes	ETA 163	MPEG	T23520	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W
5	*M. l. berleschi*	Solimoes	ETA 164	MPEG	T22566	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W
5	*M. l. berleschi*	Solimoes	ETA 165	MPEG	T22567	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W
5	*M. l. berleschi*	Solimoes	ETA 166	MPEG	T22568	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W
5	*M. l. berleschi*	Solimoes	ETA 194	MPEG	T23532	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W
5	*M. l. berleschi*	Solimoes	ETA 196	MPEG	T23534	male	AM, Fonte Boa, Rio Solimões, Ilha Nova Esperança	02°28'58.2"S	66°00'38.4"W

6	*M. l. berleschi*	Solimoes	ETA 233	MPEG	T22612	male	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57.2"S	64°42'25.9"W
6	*M. l. berleschi*	Solimoes	ETA 234	MPEG	T22613	male	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57.2"S	64°42'25.9"W
6	*M. l. berleschi*	Solimoes	ETA 235	MPEG	T22614	male	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57.2"S	64°42'25.9"W
6	*M. l. berleschi*	Solimoes	ETA 236	MPEG	T22615	male	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57.2"S	64°42'25.9"W
6	*M. l. berleschi*	Solimoes	ETA 237	MPEG	T22616	male	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57.2"S	64°42'25.9"W
6	*M. l. berleschi*	Solimoes	ETA 238	MPEG	T22617	male	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57.2"S	64°42'25.9"W
7	*M. l. berleschi*	Solimoes	TVVC 25	INPA	A936	unknown	AM, Tefe, Ilha do Barbado	3°46'51.8"S	64°1'31.3"W
8	*M. l. berleschi*	Solimoes	ETA 246	MPEG	T22625	female	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
8	*M. l. berleschi*	Solimoes	ETA 256	MPEG	T22631	male	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
8	*M. l. berleschi*	Solimoes	ETA 257	MPEG	T22632	male	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
8	*M. l. berleschi*	Solimoes	ETA 258	MPEG	T22633	male	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
13	*M. l. femininus*	Madeira	ETA 368	MPEG	T22643	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. l. femininus*	Madeira	ETA 369	MPEG	T22750	male	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. l. femininus*	Madeira	ETA 370	MPEG	T22751	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. l. femininus*	Madeira	ETA 371	MPEG	T22752	unknown	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. l. femininus*	Madeira	ETA 372	MPEG	T22753	unknown	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. l. femininus*	Madeira	ETA 373	MPEG	T23560	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. l. femininus*	Madeira	ETA 385	MPEG	T22765	male	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
14	*M. l. femininus*	Madeira	ETA 429	MPEG	T22688	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*M. l. femininus*	Madeira	ETA 430	MPEG	T22689	female	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*M. l. femininus*	Madeira	ETA 431	MPEG	T23569	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*M. l. femininus*	Madeira	ETA 432	MPEG	T22690	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*M. l. femininus*	Madeira	ETA 433	MPEG	T22786	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W

14	*M. l. femininus*	Madeira	ETA 434	MPEG	T22787	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 465	MPEG	T22812	male	AM, Itacoatiara, Rio Amazonas, Ilha do Risco	03°09'31.6"S	58°22'13.0"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 497	MPEG	T22838	female	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 498	MPEG	T22839	female	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 499	MPEG	T22840	female	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 500	MPEG	T22841	female	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 501	MPEG	T22842	female	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 502	MPEG	T22843	male	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
20	*M. l. femininusXlugubris*	Madeira/Amazonas	ETA 508	MPEG	T22849	male	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
21	*M. l. lugibris*	Amazonas	ETA 512	MPEG	T22852	female	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
21	*M. l. lugibris*	Amazonas	ETA 513	MPEG	T22853	female	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
21	*M. l. lugibris*	Amazonas	ETA 514	MPEG	T22854	female	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
21	*M. l. lugibris*	Amazonas	ETA 515	MPEG	T22855	female	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
21	*M. l. lugibris*	Amazonas	ETA 516	MPEG	T22856	female	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
21	*M. l. lugibris*	Amazonas	ETA 534	MPEG	T22870	male	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
26	*M. l. lugibris*	Amazonas	ETA 563	MPEG	T22896	male	PA, Monte Alegre, Ilha Cacoal Grande	02°23'17.2"S	54°21'41.9"W
26	*M. l. lugibris*	Amazonas	ETA 564	MPEG	T22897	male	PA, Monte Alegre, Ilha Cacoal Grande	02°23'17.2"S	54°21'41.9"W
26	*M. l. lugibris*	Amazonas	ETA 565	MPEG	T22898	male	PA, Monte Alegre, Ilha Cacoal Grande	02°23'17.2"S	54°21'41.9"W
26	*M. l. lugibris*	Amazonas	ETA 566	MPEG	T22899	male	PA, Monte Alegre, Ilha Cacoal Grande	02°23'17.2"S	54°21'41.9"W
26	*M. l. lugibris*	Amazonas	ETA 567	MPEG	T22900	female	PA, Monte Alegre, Ilha Cacoal Grande	02°23'17.2"S	54°21'41.9"W
27	*M. l. lugibris*	Amazonas	ETA 597	MPEG	T22925	male	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*M. l. lugibris*	Amazonas	ETA 598	MPEG	T22926	female	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*M. l. lugibris*	Amazonas	ETA 599	MPEG	T22927	male	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W

27	*M. l. lugibris*	Amazonas	ETA 600	MPEG	T22928	female	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*M. l. lugibris*	Amazonas	ETA 601	MPEG	T22929	male	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*M. l. lugibris*	Amazonas	ETA 645	MPEG	T22967	male	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*M. l. lugibris*	Amazonas	ETA 646	MPEG	T23601	male	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
16	*M. l. stictopterus*	Negro	CFV 65	INPA	A2191	female	RR, Caracaraí, Vista Alegre, Ilha do Palhal	1°14'25"N	61°18'51"W
16	*M. l. stictopterus*	Negro	AMFP 126	INPA	A2193	male	RR, Caracaraí, Vista Alegre, Ilha do Palhal	1°14'25"N	61°18'51"W
17	*M. l. stictopterus*	Negro	GRL 1237	INPA	A8376	male	RR, Ilha no Rio Branco, próximo a Santa Maria do Boiaçu, ca 280 km Caracaraí	0°31'48"S	61°47'57"W
17	*M. l. stictopterus*	Negro	GRL 1240	INPA	A8379	female	RR, Ilha no Rio Branco, próximo a Santa Maria do Boiaçu, ca 280 km Caracaraí	0°31'48"S	61°47'57"W
19	*M. l. stictopterus*	Negro	SHB 308	INPA	A3095	female	AM, Novo Airão, Ilhas do Rio Negro em frente à boca do Rio Jaú, Ilha Jussara	1°51'56.7"S	61°22'20.8"W
19	*M. l. stictopterus*	Negro	SHB 293	INPA	A3103	male	AM, Novo Airão, Ilhas do Rio Negro em frente à boca do Rio Jaú, Ilha Jussara	1°51'56.7"S	61°22'20.8"W
19	*M. l. stictopterus*	Negro	GRL 1426	INPA	A10505	female	AM, Parna Jaú, Rio Negro, Ilha das Onças	1°50'37.8"S	61°22'48.2"W
19	*M. l. stictopterus*	Negro	GRL 1429	INPA	A10508	female	AM, Parna Jaú, Rio Negro, Ilha das Onças	1°52'32.4"S	61°22'1.7"W
19	*M. l. stictopterus*	Negro	SHBs 54	INPA	A15803	male	AM, Parna Jaú, Rio Negro, Ilha das Onças	1°48'37"S	61°23'42"W
19	*M. l. stictopterus*	Negro	SHBs 56	INPA	A15805	female	AM, Parna Jaú, Rio Negro, Ilha das Onças	1°48'37"S	61°23'42"W
1	*M. a. assimilis*	Solimoes	AMA 427	MPEG	72924	female	AM, Estirão do Equador, Atalaia do Norte	04°29'12.1"S	71°33'02.3"W
1	*M. a. assimilis*	Solimoes	AMA 475	MPEG	72972	unknown	AM, Estirão do Equador, Atalaia do Norte	04°29'12.1"S	71°33'02.3"W
2	*M. a. assimilis*	Solimoes	ETA 046	MPEG	T22475	female	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. a. assimilis*	Solimoes	ETA 051	MPEG	T23495	female	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*M. a. assimilis*	Solimoes	ETA 054	MPEG	T23497	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
3	*M. a. assimilis*	Solimoes	ETA 081	MPEG	T22503	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*M. a. assimilis*	Solimoes	ETA 097	MPEG	T23506	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W

3	*M. a. assimilis*	Solimoes	ETA 098	MPEG	T23507	male	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
4	*M. a. assimilis*	Solimoes	JUT269	MPEG	T23446	female	AM, Jutaí, ESEC Jutaí/Solimões, Pati	03°13'29.9"S	67°26'10"W
6	*M. a. assimilis*	Solimoes	ETA 212	MPEG	T22595	male	AM, Tefé, Rio Solimões, Ilha 1	03°22'46.0"S	64°38'27.8"W
8	*M. a. assimilis*	Solimoes	ETA 254	MPEG	T22630	male	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
8	*M. a. assimilis*	Solimoes	ETA 269	MPEG	T23548	female	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
10	*M. a. assimilis*	Solimoes/Amazonas/ Negro	ETA 337	MPEG	T22719	female	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
10	*M. a. assimilis*	Solimoes/Amazonas/ Negro	ETA 338	MPEG	T22720	female	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
10	*M. a. assimilis*	Solimoes/Amazonas/ Negro	ETA 339	MPEG	T22721	male	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
10	*M. a. assimilis*	Solimoes/Amazonas/ Negro	ETA 353	MPEG	T22735	male	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
11	*M. a. assimilis*	Madeira	MAD 057	MPEG	T15881	female	RO, Porto Velho, Conceição do Garcia, margem leste Rio Madeira	8°13'26"S	63°10'54.6"W
11	*M. a. assimilis*	Madeira	MAD 060	MPEG	T15884	male	RO, Porto Velho, Conceição do Garcia, margem leste Rio Madeira	8°13'26"S	63°10'54.6"W
11	*M. a. assimilis*	Madeira	MSF408	MPEG	T22145	unknown	RO, São Carlos, Rio Jamari	8°56'12"S	63°21'05"W
12	*M. a. assimilis*	Madeira	MAD 303	MPEG	T16124	male	AM, Mancicoré, Ribamar, margem leste Rio Madeira	05°36'46.5"S	61°07'17.9"W
12	*M. a. assimilis*	Madeira	MAD 337	MPEG	T16156	male	AM, Mancicoré, Ribamar, margem leste Rio Madeira	05°36'46.5"S	61°07'17.9"W
13	*M. a. assimilis*	Madeira/Amazonas/ Negro	ETA 399	MPEG	T23561	male	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. a. assimilis*	Madeira/Amazonas/ Negro	ETA 400	MPEG	T23562	male	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. a. assimilis*	Madeira/Amazonas/ Negro	ETA 401	MPEG	T23563	male	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. a. assimilis*	Madeira/Amazonas/ Negro	ETA 402	MPEG	T22779	male	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*M. a. assimilis*	Madeira/Amazonas/ Negro	ETA 403	MPEG	T23565	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W

14	*M. a. assimilis*	Amazonas/Negro	ETA 440	MPEG	T23571	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*M. a. assimilis*	Amazonas/Negro	ETA 441	MPEG	T23572	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*M. a. assimilis*	Amazonas/Negro	ETA 442	MPEG	T23573	female	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
18	*M. a. assimilis*	Amazonas/Negro	LNN	INPA	A8418	male	AM, Margem esquerda do Rio Branco, ~ 50 km N da confluência com o Rio Negro	1°05'55"S	61°56'35"W
19	*M. a. assimilis*	Amazonas/Negro	SHBs	INPA	A15788	male	AM, Ilha do Caroçal, Rio Negro, Parque Nacional do Jaú	1°51'18"S	61°26'18"W
20	*M. a. assimilis*	Amazonas/Negro	ETA 490	MPEG	T23576	male	AM, Itacoatiara, Rio Amazonas, Ilha do Risco	03°09'31.6"S	58°22'13.0"W
20	*M. a. assimilis*	Amazonas/Negro	ETA 491	MPEG	T23577	female	AM, Itacoatiara, Rio Amazonas, Ilha do Risco	03°09'31.6"S	58°22'13.0"W
20	*M. a. assimilis*	Amazonas/Negro	ETA 509	MPEG	T22850	male	AM, Itacoatiara, Rio Amazonas, Ilha do Boi	03°11'4.3"S	58°15'52.1"W
21	*M. a. assimilis*	Amazonas/Negro	ETA 559	MPEG	T23587	male	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
21	*M. a. assimilis*	Amazonas/Negro	ETA 560	MPEG	T23588	female	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47.5"S	56°40'45.0"W
22	*M. a. transamazonica*	Amazonas/Negro	JUR 011	MPEG	T6623	unknown	PA, Juruti, Lago Santana	02°05'01.1"S	56°01'30.9"W
22	*M. a. transamazonica*	Amazonas/Negro	56696	MPEG	T2352	male	PA, Juruti, Fazenda São Joaquim	02°05'01.1"S	56°01'30.9"W
2	*T. cryptoleucus*	Solimoes	ETA 003	MPEG	T22432	female	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*T. cryptoleucus*	Solimoes	ETA 005	MPEG	T22434	cho	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*T. cryptoleucus*	Solimoes	ETA 007	MPEG	T22436	female	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
2	*T. cryptoleucus*	Solimoes	ETA 008	MPEG	T22437	male	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27.7"S	69°53'10.2"W
3	*T. cryptoleucus*	Solimoes	ETA 062	MPEG	T23498	female	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
3	*T. cryptoleucus*	Solimoes	ETA 063	MPEG	T22488	male	AM, Santo Antônio do Içá, Ilha Porto América	03°08'52.7"S	67°57'35.1"W
5	*T. cryptoleucus*	Solimoes	ETA 174	MPEG	T22574	female	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58.2"S	66°00'38.4"W
5	*T. cryptoleucus*	Solimoes	ETA 175	MPEG	T22575	male	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58.2"S	66°00'38.4"W
5	*T. cryptoleucus*	Solimoes	ETA 176	MPEG	T22576	female	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58.2"S	66°00'38.4"W
6	*T. cryptoleucus*	Solimoes	ETA 201	MPEG	T22585	female	AM, Tefé, Rio Solimões, Ilha 1	03°22'46.0"S	64°38'27.8"W

6	*T. cryptoleucus*	Solimoes	ETA 202	MPEG	T22586	male	AM, Tefé, Rio Solimões, Ilha 1	03°22'46.0"S	64°38'27.8"W
8	*T. cryptoleucus*	Solimoes	ETA 240	MPEG	T22619	female	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
8	*T. cryptoleucus*	Solimoes	ETA 242	MPEG	T22621	male	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
8	*T. cryptoleucus*	Solimoes	ETA 243	MPEG	T22622	male	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49.5"S	62°57'00.5"W
9	*T. cryptoleucus*	Solimoes	ETA 289	MPEG	T22661	male	AM, Codajás, Rio Solimões, ilha do camaleão	03°50'36.4"S	62°14'29.7"W
9	*T. cryptoleucus*	Solimoes	ETA 290	MPEG	T22662	male	AM, Codajás, Rio Solimões, ilha do camaleão	03°50'36.4"S	62°14'29.7"W
10	*T. cryptoleucus*	Solimoes	ETA 331	MPEG	T22713	female	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
10	*T. cryptoleucus*	Solimoes	ETA 332	MPEG	T22714	male	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
10	*T. cryptoleucus*	Solimoes	ETA 334	MPEG	T22716	male	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
10	*T. cryptoleucus*	Solimoes	ETA 335	MPEG	T22717	female	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15.9"S	60°12'54.1"W
14	*T. cryptoleucus*	Solimoes/Madeira	ETA 418	MPEG	T22677	male	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*T. cryptoleucus*	Solimoes/Madeira	ETA 419	MPEG	T22678	female	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
14	*T. cryptoleucus*	Solimoes/Madeira	ETA 420	MPEG	T22679	female	AM, Autazes, Uricurituba, Rio Madeira, Ilha Caiçara	03°34'17.3"S	58°55'22.2"W
15	*T. n. cinereoniger*	Negro	SGC 200	MPEG	T20464	female	AM, São Gabriel da Cachoeira, Ilha da Brigada	0°08'S	67°04'W
15	*T. n. cinereoniger*	Negro	SGC 515	MPEG	T20770	male	AM, São Gabriel da Cachoeira, Ilha da Brigada	0°08'S	67°04'W
16	*T. n. cinereoniger*	Negro	GRL17	INPA	A2165	male	RR, Caracaraí, Vista Alegre, margem esquerda do Rio Branco	1°28'2.2"N	61°15'12.8"W
16	*T. n. cinereoniger*	Negro	GRL1149	INPA	A8278	male	RR, Ilha Ajarani no Rio Branco, ca 30 km S Caracaraí	1°31'52"N	61°14'44"W
18	*T. n. cinereoniger*	Negro	GRL1261	INPA	A8394	female	RR, Margem direita do Rio Branco a ~ 10 km da confluência com o Rio Negro	1°17'43"S	61°51'45"W
19	*T. n. cinereoniger*	Negro	SHB360	INPA	A3093	male	AM, Novo Airão, Ilhas do Rio Negro	1°48'41.3"S	61°23'35.6"W
19	*T. n. cinereoniger*	Negro	SHB291	INPA	A3099	male	AM, Novo Airão, Ilhas do Rio Negro em frente à boca do Rio Jaú, Ilha Jussara	1°51'56.7"S	61°22'20.8"W
19	*T. n. cinereoniger*	Negro	SHB297	INPA	A3101	female	AM, Novo Airão, Ilhas do Rio Negro em frente à boca do Rio Jaú, Ilha Jussara	1°51'56.7"S	61°22'20.8"W

19	*T. n. cinereoniger*	Negro	SHBs50	INPA	A15799	female	AM, Ilha das Onças, Rio Negro, Parque Nacional do Jaú	1°48'37"S	61°23'42"W
19	*T. n. cinereoniger*	Negro	SHBs51	INPA	A15800	female	AM, Ilha das Onças, Rio Negro, Parque Nacional do Jaú	1°48'37"S	61°23'42"W
23	*T. n. huberi*	Tapajos	JAT(A) 026	MPEG	T18623	female	PA, Jacareacanga, margem esquerda Rio Tapajós, Vila Mamãe-anã	05°45'31.5"S	57°24.'45.85"W
24	*T. n. huberi*	Tapajos	HSLT217	INPA	A12364	male	PA, PA: Ilha do Rio Tapajós, ca. 150 km SW Itaituba, "I. do Rato", mod. Ilha Rato	5°21'10.5"S	56°57'52.67"W
25	*T. n. huberi*	Tapajos	A9996	INPA	A9996	male	PA, Itaituba, ilha fluvial no Rio Tapajós	4°27'S	56°5'W
25	*T. n. huberi*	Tapajos	A10000	INPA	A10000	male	PA, Itaituba, ilha fluvial no Rio Tapajós	4°27'S	56°5'W
25	*T. n. huberi*	Tapajos	A10843	INPA	A10843	male	PA, Itaituba, Rio Tapajós, ilha fluvial Vila Rayol	4°27'S	56°5'W
25	*T. n. huberi*	Tapajos	A10844	INPA	A10844	male	PA, Itaituba, Rio Tapajós, ilha fluvial Vila Rayol	4°27'S	56°5'W
25	*T. n. huberi*	Tapajos	A10846	INPA	A10846	female	PA, Itaituba, Rio Tapajós, ilha fluvial Vila Rayol	4°27'S	56°5'W
30	*T. n. kulcynskii*	Amazonas	OI001	MPEG	T24378	male	AP, Oiapoque, Rio Oiapoque, ilha	04°01'36.3"S	51°41'59.9"W
30	*T. n. kulcynskii*	Amazonas	OI002	MPEG	T24379	male	AP, Oiapoque, Rio Oiapoque, ilha	04°01'36.3"S	51°41'59.9"W
30	*T. n. kulcynskii*	Amazonas	OI004	MPEG	T24381	male	AP, Oiapoque, Rio Oiapoque, ilha	04°01'36.3"S	51°41'59.9"W
27	*T. n. nigrocinereus*	Amazonas	ETA 611	MPEG	T23595	female	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*T. n. nigrocinereus*	Amazonas	ETA 612	MPEG	T23596	female	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
27	*T. n. nigrocinereus*	Amazonas	ETA 613	MPEG	T22939	female	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15.2"S	52°31'25.9"W
28	*T. n. nigrocinereus*	Amazonas	ETA671	MPEG	xxx	unknown	PA, Cameta ou rio PARA	1°41'13"S	49°17'57"W
28	*T. n. nigrocinereus*	Amazonas	ETA673	MPEG	xxx	unknown	PA, Cameta ou rio PARA	1°41'13"S	49°17'57"W
29	*T. n. nigrocinereus*	Amazonas	MAYA 003	MPEG	T4156	female	PA, Marajó, Anajás, Mocoons	00°39'06.5"S	49°27'12.8"W
29	*T. n. nigrocinereus*	Amazonas	MAYA 019	MPEG	T4171	male	PA, Marajó, Anajás, Mocoons	00°39'06.5"S	49°27'12.8"W
12	*T. n. tschudii*	Madeira	CHS 29	INPA	A295	female	AM, Manicoré, Ilha do Uruá, Rio Madeira, ca 80 km NE	5°21'37"S	60°43'51"W
12	*T. n. tschudii*	Madeira	CHS30	INPA	A296	female	AM, Manicoré, Ilha do Uruá, Rio Madeira, ca 80 km NE	5°21'37"S	60°43'51"W
12	*T. n. tschudii*	Madeira	CHS31	INPA	A297	male	AM, Manicoré, Ilha do Uruá, Rio Madeira, ca 80 km NE	5°21'37"S	60°43'51"W

13	*T. n. tschudii*	Madeira	ETA 389	MPEG	T22769	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*T. n. tschudii*	Madeira	ETA 390	MPEG	T22770	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W
13	*T. n. tschudii*	Madeira	ETA 391	MPEG	T22771	female	AM, Borba, Rio Madeira, Ilha de Borba	04°22'31.7"S	59°35'49.5"W

Table S2: Initial search range of Fastsimcoal2 models tested.

Taxon	Model	Ne current pops	Ne ancestral pops	Migration	TDIVs	Epsilon	Tau
Myrmoborus lugubris	0mig_0siz	{20k, 250K}	{50k, 500K}	***	{5k, ~108K}	***	***
	0mig_siz	{20k, 250K}	{50k, 500K}	***	{5k, ~108K}	{0.02, 100}	{1k, <TDIV}
	mig_0siz	{20k, 250K}	{50k, 500K}	{1e-3, 2}	{5k, ~108K}	***	***
	mig_siz	{20k, 250K}	{50k, 500K}	{1e-3, 2}	{5k, ~108K}	{0.02, 100}	{1k, <TDIV}
	mig_siz_All_G-phocs	{20k, 250K}	{50k, 500K}	{1e-3, 2}	***	{0.02, 100}	{1k, <TDIV}
	Model	**Ne current pops**	**Ne ancestral pops**	**Migration**	**TDIVs**	**Epsilon**	**Tau**
Thamnophilus nigrocinereus/ T. cryptoleucus	0mig_0siz	{15k, 150k}	{10k, 500K}	***	{1K, ~145k}	***	***
	0mig_siz	{15k, 150k}	{10k, 500K}	***	{1K, ~145k}	{0.02, 100}	{1k, <TDIV}
	mig_0siz	{15k, 150k}	{10k, 500K}	{1e-3, 2}	{1K, ~145k}	***	***
	mig_siz	{15k, 150k}	{10k, 500K}	{1e-3, 2}	{1K, ~145k}	{0.02, 100}	{1k, <TDIV}
	mig_siz_All_G-phocs	{15k, 150k}	{10k, 500K}	{1e-3, 2}	***	{0.02, 100}	{1k, <TDIV}
	Model	**Ne current pops**	**Ne ancestral pops**	**Migration**	**TDIVs**	**Epsilon**	**Tau**
Myrmotherula assimilis	0mig_0siz	{5k, 500k}	{10k, 800k}	***	{10K, ~226K}	***	***
	0mig_siz	{5k, 500k}	{10k, 800k}	***	{10K, ~226K}	{0.02, 100}	{1k, <TDIV}
	mig_0siz	{5k, 500k}	{10k, 800k}	{1e-3, 2}	{10K, ~226K}	***	***
	mig-siz	{5k, 500k}	{10k, 800k}	{1e-3, 2}	{10K, ~226K}	{0.02, 100}	{1k, <TDIV}
	mig_siz_All_G-phocs	{5k, 500k}	{10k, 800k}	{1e-3, 2}	***	{0.02, 100}	{1k, <TDIV}

0mig_0siz – demographic model with both migration between populations and population size changes fixed in zero; 0mig_siz - demographic model with migration between populations fixed to zero but with population size changes; mig_0siz – demographic model with migration between populations but population size changes fixed to zero; mig_siz – demographic model with migration between populations and population size changes; mig_siz_ALL_G-phocs – demographic model with migration between populations and population size changes with all divergence times fixed with G-phocs mean estimations (tables 1-3); Ne current pops – Effective population size of diploid individuals of current populations (Solimões, Negro, Madeira, Tapajós, and Amazonas); Ne ancestral pops - Effective population size of diploid individuals of ancestral populations; TDIVs – Divergence times bounded by the oldest split estimated by G-phocs; Epsilon - proportional size of ancestral populations before population size changes; Tau – Time of instantaneous population size change.

Table S3: Model selection and parameter estimation of Multi-Dice single-population models for *Myrmotherula assimilis*, *Myrmoborus lugubris*, and *T. nigrocinereus/T. cryptoleucus.*

Species	**populations**	**SNPs**	**Model/Posterior probability**			**Parameter estimation**		
			Expansion	Stable	Contraction	Ne (95% HPD)	Epsilon (95% HPD)	τ (95% HPD)
M. assimilis	a_amazonas	1497	1	0	0	95483 (53414-179220)	0.175 (0.0159-0.1962)	140831 (56065-148440)
	a_madeira	1664	1	0	0	734888 (152587-960562)	0.1607 (0.1309-0.1892)	132815 (25391-147514)
	a_negro	1343	1	0	0	195406 (60796-311281)	0.0712 (0.0148-0.1954)	138593 (36786-148049)
	a_solimoes	1392	1	0	0	349423 (100621-859164)	0.1968 (0.1404-0.1997)	135471 (27124-148210)
M. lugubris	l_amazonas	1005	0.99	0.001	0.0027	922256 (186568-988659)	0.1697 (0.0430-0.1979)	4304 (1404-20906)
	l_madeira	1087	1	0	0	187634 (61711-318840)	0.0316 (0.0142-0.1946)	138877 (36309-148045)
	l_negro	1124	0.2952	0.6772	0.0277	57189 (15346-963449)	1.083 (0.0767-5.2698)	4809 (1394-145311)
	l_solimoes1	844	1	0	0	850431 (165382-981202)	0.1964 (0.1732-0.1996	104989 (23722-145586)
	l_solimoes2	582	0.08	0.0383	0.8807	39097 (14333-857116)	5.7851 (0.1037-32.5733)	133579 (3356-148050)
T. nigrocinereus/ T. cryptoleucus	n_amazonas	1790	1	0	0	161359 (60265-298735)	0.0679 (0.0147-0.1953)	139663 (38480-148049)
	n_madeira	1144	0	0	1	455026 (29251-491711)	5.7545 (5.0633-50.7614)	27009 (6445-131378)
	n_negro	1602	0.174	0.7168	0.1092	67434 (14148-482841)	0.8921 (0.0457-17.8587)	134655 (2934-147283)
	n_solimoes1	935	1	0	0	659007 (145679-967496)	0.1969 (0.1709-0.1997)	122941 (24706-147450)
	n_solimoes2	935	1	0	0	924572 (160641-988182)	0.1967 (0.1082-0.1996)	16532 (4630-127589)
	n_tapajos	740	0.0005	0.6975	0.302	33718 (13206-480791)	0.9209 (1-16.9499)	131934 (5268-147100)

SNPs – number of SNPs recovered for each population; Ne – Effective population size in haploid individuals; Epsilon - proportional size of ancestral population; τ – Time of instantaneous population size change.

Table S4: Absolute number of SNPs recovered for *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis.*

Species	n SNPs raw VCF	n SNPs depth > 6	Mean depth (SD)	n SNPs (1 SNP/UCE) no missing data
M. lugubris	19622	12332	18.86 (18.50)	1521
T. nigrocinereus/ T. cryptoleucus	31453	20889	20.41 (20.92)	1588
M. assimilis	23569	16021	20.58 (20.86)	1408

n SNPs raw VCF – number of SNPs in the raw VCF; n SNPs depth > 6 – number of SNPs with depth higher than 6; n SNPs (1 SNP/UCE) no missing data – total number of SNPs after randomly selecting one SNP per locus.

Table S5: Percentage of explained variance of three Treemix models for *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis*

Species	0 migration (% explained variance)	migration (% explained variance)	Edges with jack-knife p-value < 0.05
M. lugubris	0.991	0.9973	3
T. nigrocinereus/T. cryptoleucus	0.9826	0.9982	2
M. assimilis	0.8628	0.9968	5

Table S6: Topologies tested for *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis* in Fastsimcoal2

Species	Model	Max ln(L)	param	AIC	Relative Weights
M. lugubris	(A,(N,M))	-14905.47	21	68684.22	1.00E+00
	(N,(A,M))	-14911.62	21	68712.56	7.01E-07
T. nigrocinereus/ T. cryptoleucus	((S,M),(N,(A,T)))	-17777.802	26	81921.80	1.00E+00
	(((S,M),N),(A,T))	-17786.125	26	81960.13	4.75E-09
M. assimilis	(S,(M,A))	-3345.271	13	15431.54	6.83E-01
	(M,(S,A))	-3345.713	13	15433.58	2.47E-01
	(A,(S,M))	-3346.259	13	15436.09	7.02E-02

Max ln(L) - Composite likelihood; AIC - Akaike information criterion; Relative Weights - relative contribution for each of the topologies tested for *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis*.

Table S7: Posterior estimates obtained with the hABC procedure implemented in Multi-Dice using aSFS.

hABC parameter estimation	**ζ**	**Ω**	**Synchronous Pulse Time (years)**	**Mean Expansion Time (years)**
2.5% posterior density	0.70	0.00	84,569	111,181
mean	0.93	72,010	213,678	241,857
mode	1.00	1,101	276,205	272,128
97.5% posterior density	1.00	476,993	295,558	384,028

ζ - proportion of expanding population in the synchronic pulse; Ω – dispersion index of expansion times (variance over the mean value).

Table S8: Results of the "leave one out" cross-validations obtained with the hABC procedure implemented in Multi-Dice using the aSFS.

Estimate	**ζ**	**Ω**	**Synchronous Pulse Time**	**mean expansion Time**
Mean correlation	0.8356	0.7070	0.7325	0.8744
Mean estimate root mean squared error	0.1540	333417	51671	197,256
Median correlation	0.9046	0.5392	0.8377	0.8850
Median estimate root mean squared error	0.1256	443401	49592	205319

ζ - proportion of expanding population in the synchronic pulse; Ω – dispersion index of expansion times (variance over the mean value).

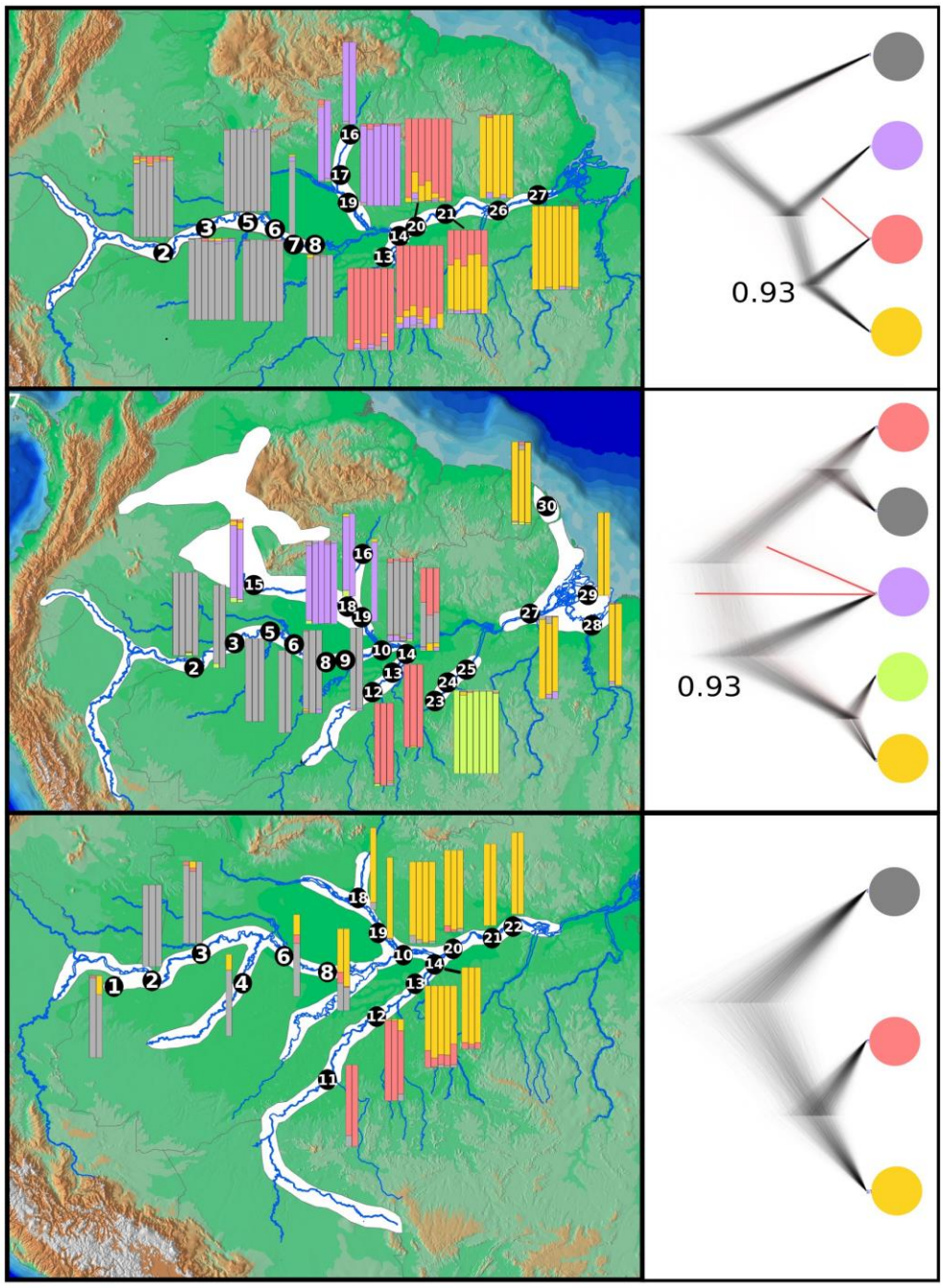

Figure 1: Geographic distribution of the three species complexes studied (numbers are localities in Table S1), population structure inferred with sNMF (bars), and SNAPP species tree (numbers are posterior probabilities < 0.95, all other nodes presented maximum posterior probabilities). Top – *Myrmoborus lugubris;* center – *Thmanophilus nigrocinereus/T. cryptoleucus*; bottom – *Myrmotherula assimilis.* Gray – Solimões population; Pink – Madeira population; Green – Tapajós population; Purple – Negro population; Yellow – Amazonas population (Amazonas+Negro for *M. assimilis*).

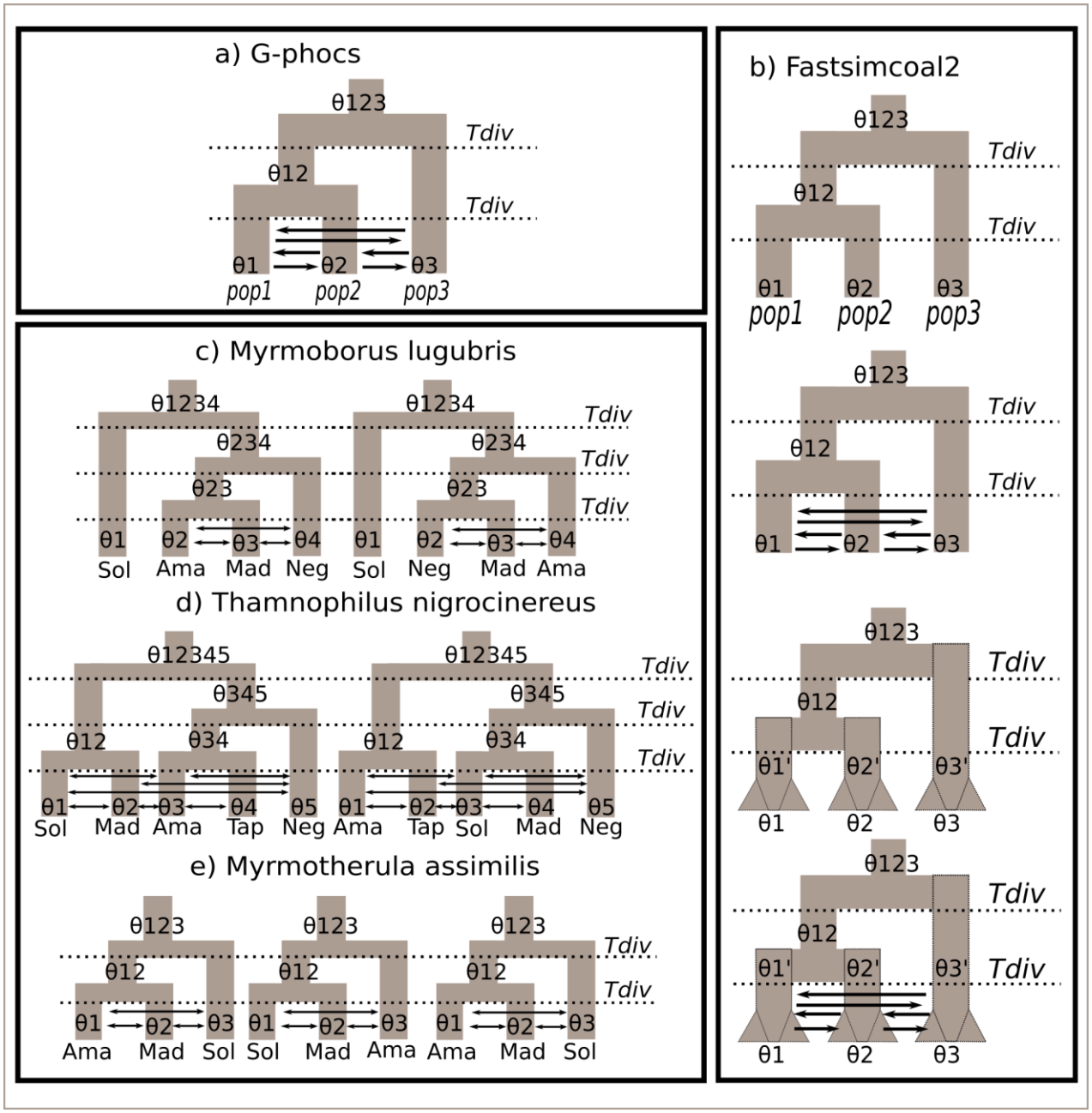

Figure 2: Example of the demographic model applied in G-phocs (a); Demographic models simulated in Fastsimcoal2 (b); Topologies tested with Fastsimcoal2 for *Myrmoborus lugubris* (c), *Thamnophilus nigrocinereus/T. cryptoleucus* (d) and *Myrmotherula assimilis* (e). θ – Effective population size; Tdiv – divergence time; Sol – Solimões population; Mad – Madeira population; Neg – Negro population; Ama – Amazonas population; Tap – Tapajós population.

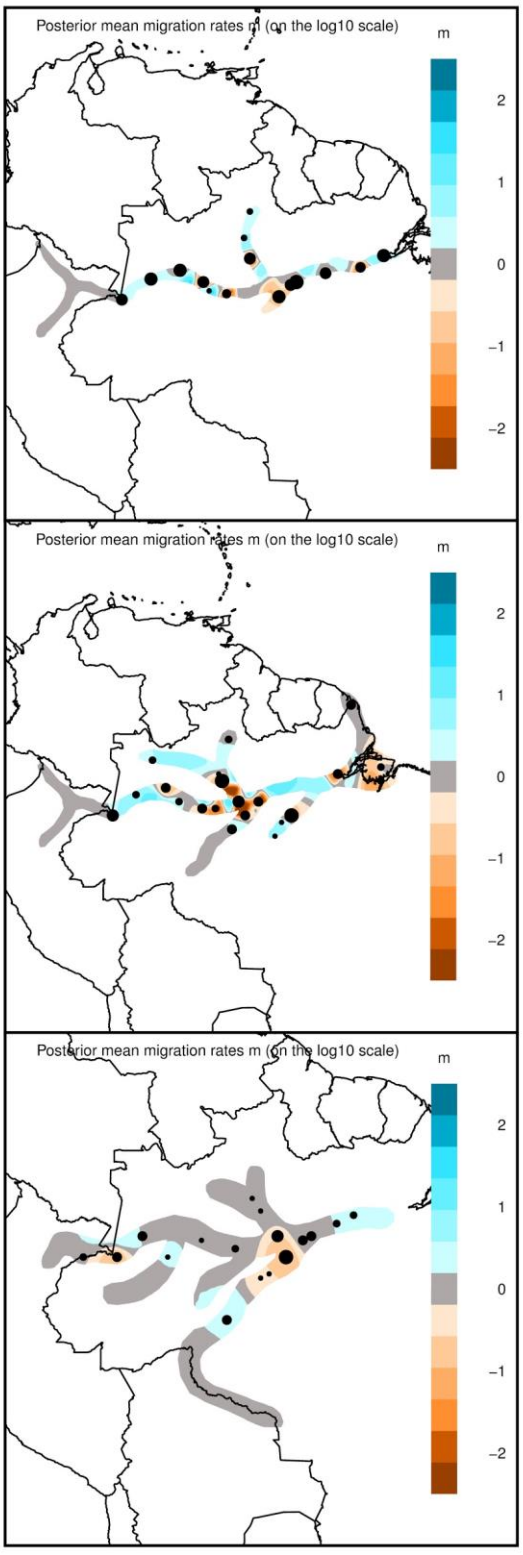

Figure 3: Estimated effective migration surface (EEMS) for *Myrmoborus lugubris* (top), *Thamnophilus nigrocinereus/T. cryptoleucus* (center) and *Myrmotherula assimilis* (bottom). Color bar shows the effective migration rate on a log10 scale relative to the average migration rate over the entire range of the corresponding species

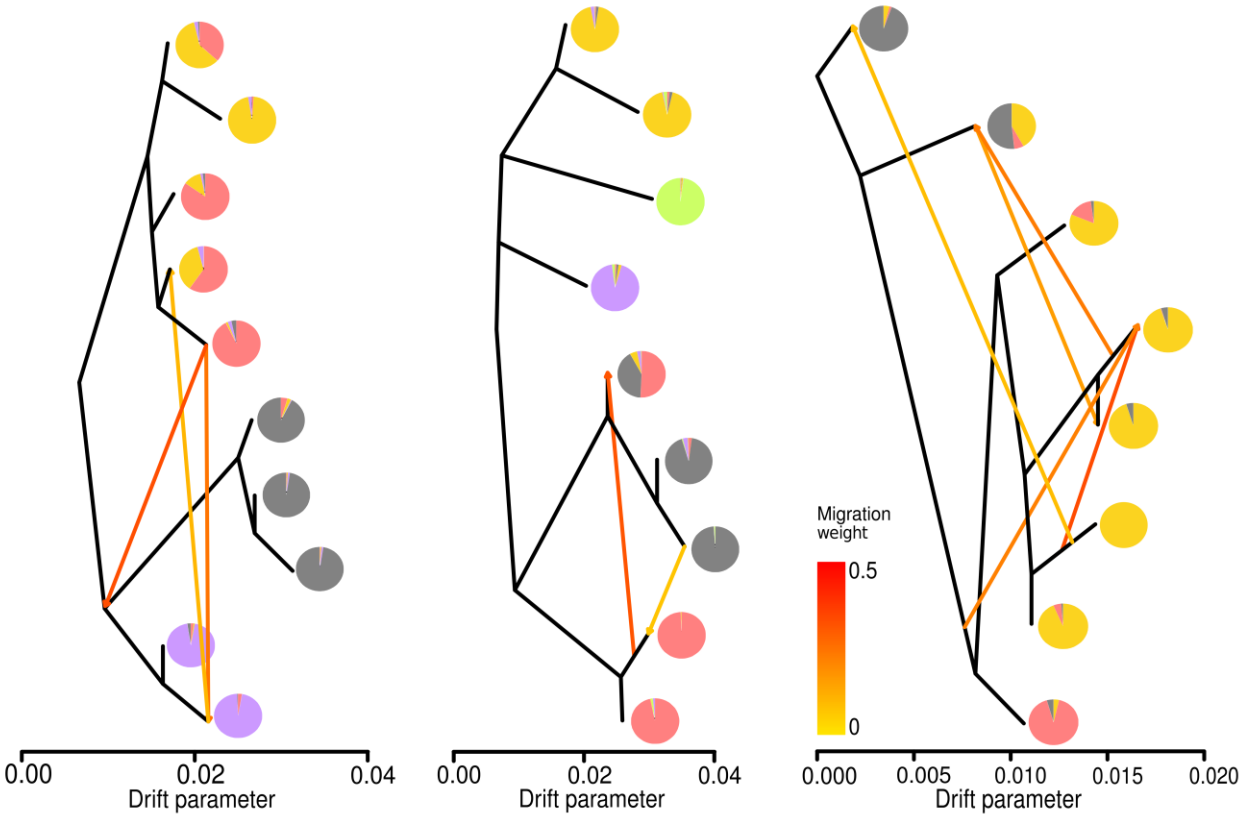

Figure 4: Population relationship graphs for *Myrmoborus lugubris* (left), *Thamnophilus nigrocinereus/T. cryptoleucus* (center), and *Myrmotherula assimilis* (right) obtained with Treemix. Colors represent populations in Figure 1; pie charts represent the coefficient of ancestrality obtained with sNMF.

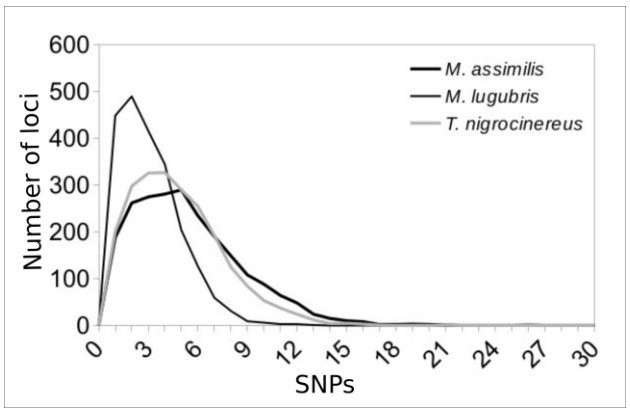

Figure S1: Distribution of number SNPs per loci.

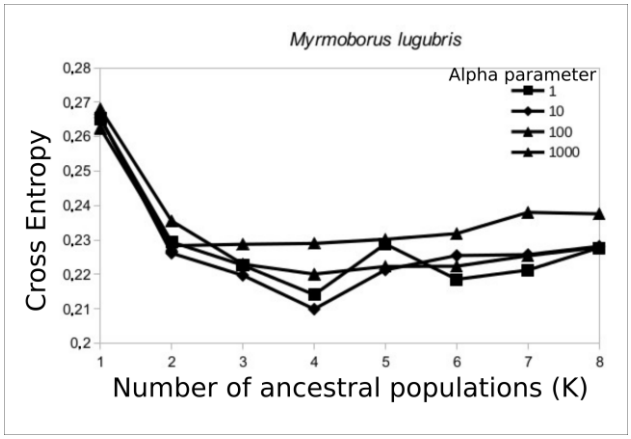

Figure S2: Masked cross-entropy values estimated in sNMF with distinct K and four different alpha regularization parameter values for *Myrmoborus lugubris*.

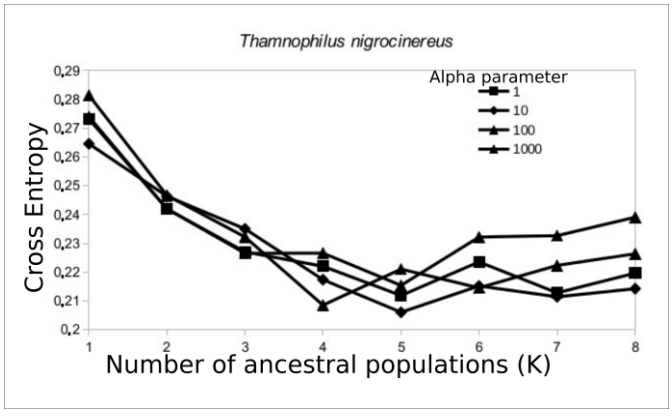

Figure S3: Masked cross-entropy values estimated in sNMF with distinct K and four different alpha regularization parameter values for *Thamnophilus nigrocinereus/T. cryptoleucus*.

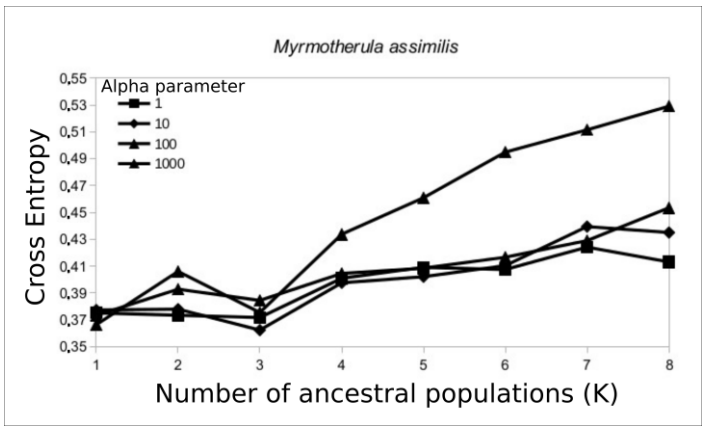

Figure S4: Masked cross-entropy values estimated in sNMF with distinct K and four different alpha regularization parameter values for *Myrmotherula assimilis.*

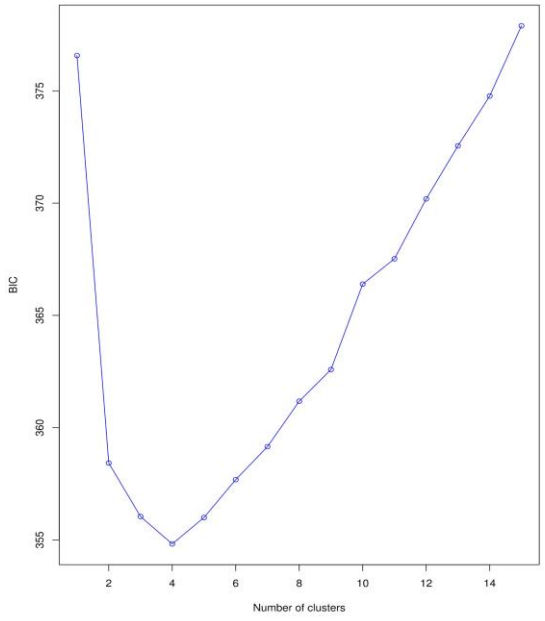

Figure S5: k-means results for *Myrmoborus lugubris.* Bayesian information criterion values (BIC) for distinct clusters of individuals (Number of clusters).

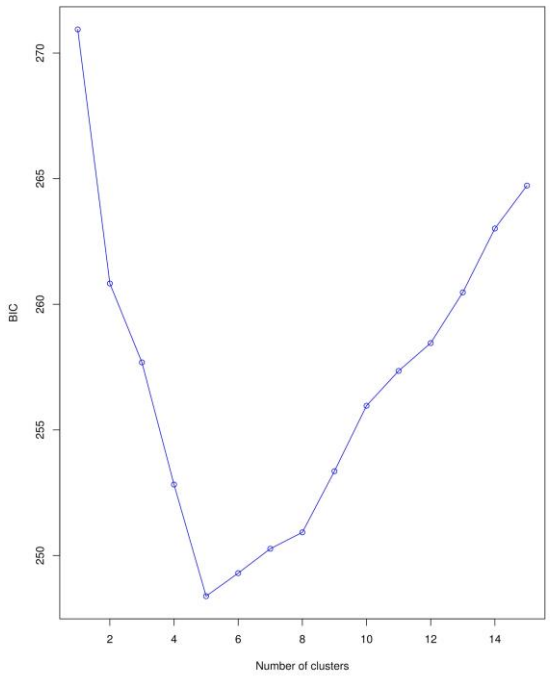

Figure S6: k-means results for *Thamnophilus nigrocinereus/T. cryptoleucus*. Bayesian information criterion values (BIC) for distinct clusters of individuals (Number of clusters).

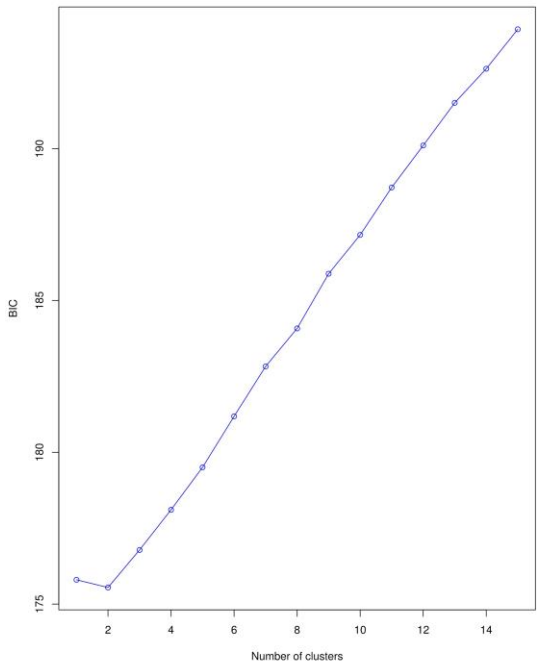

Figure S7: k-means results for *Myrmotherula assimilis*. Bayesian information criterion values (BIC) for distinct clusters of individuals (Number of clusters).

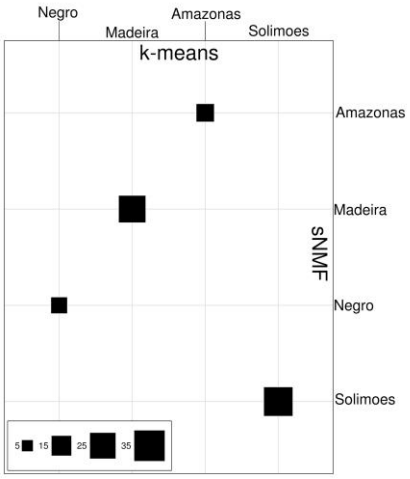

Figure S8: Classification matrix of the discriminant analyses of principal components versus sNMF groups for *Myrmoborus lugubris.* Square size represents number of individuals.

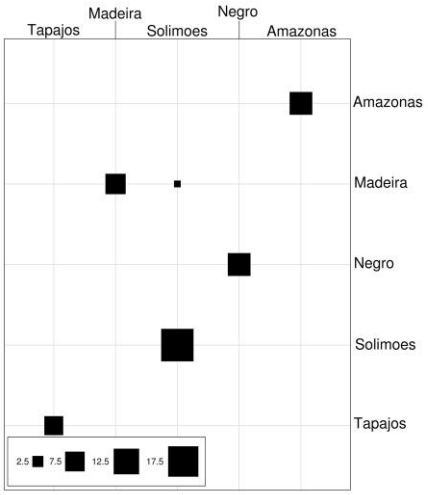

Figure S9: Classification matrix of the discriminant analyses of principal components versus sNMF groups for *Thamnophilus nigrocinereus/T. cryptoleucus.* Square size represents number of individuals.

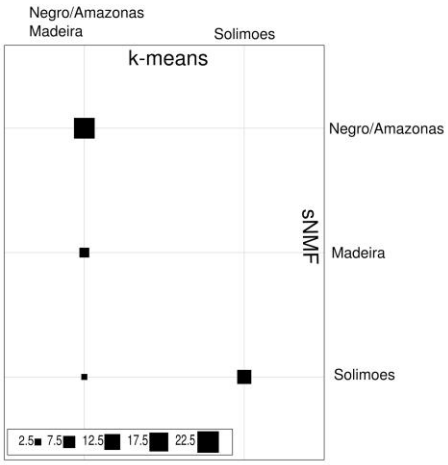

Figure S10: Classification matrix of the discriminant analyses of principal components versus sNMF groups for *Myrmotherula assimilis.* Square size represents number of individuals.

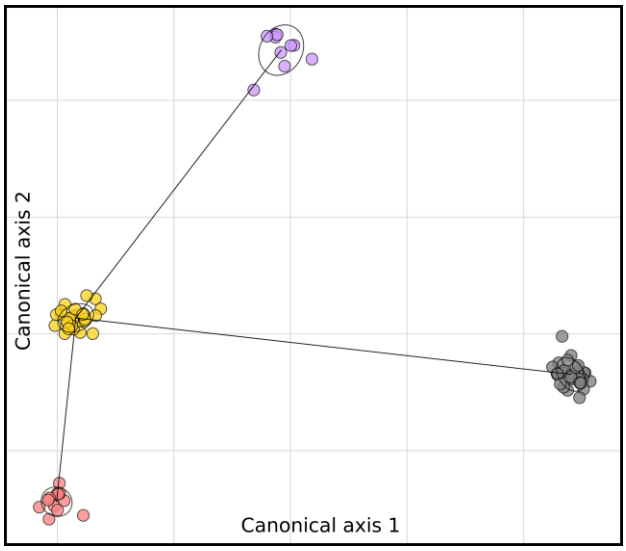

Figure S11: Graphical visualization of the first (x-axis) versus the second (y-axis) canonical axes obtained through a discriminant analysis of principal components for *Myrmoborus lugubris*

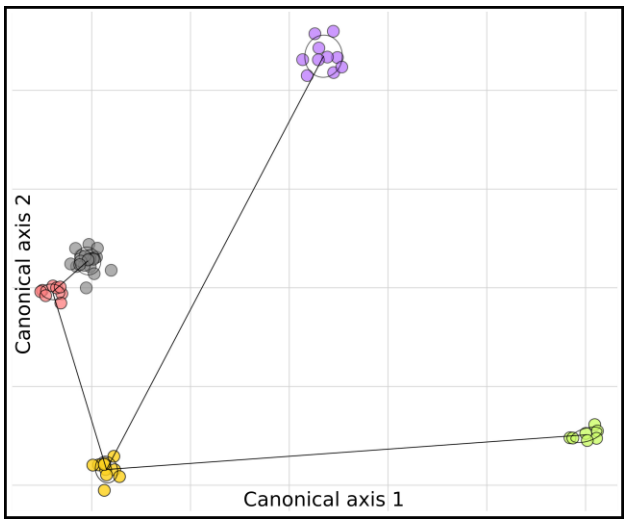

Figure S12: Graphical visualization of the first (x-axis) versus the second (y-axis) canonical axes obtained through a discriminant analysis of principal components for *Thamnophilus nigrocinereus/T. cryptoleucus.*

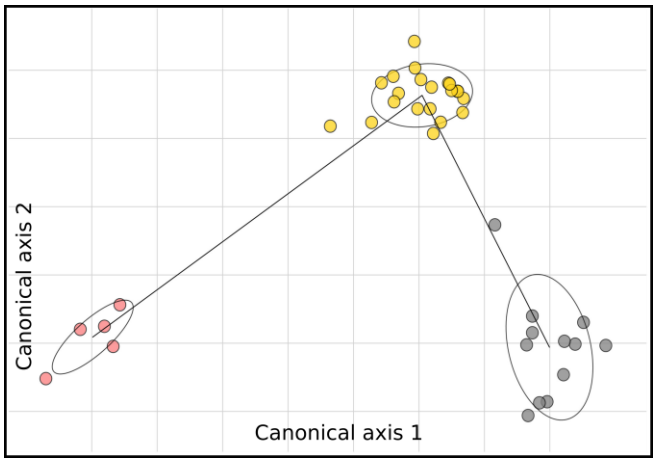

Figure S13: Graphical visualization of the first (x-axis) versus the second (y-axis) canonical axes obtained through a discriminant analysis of principal components for *Myrmotherula assimilis.*

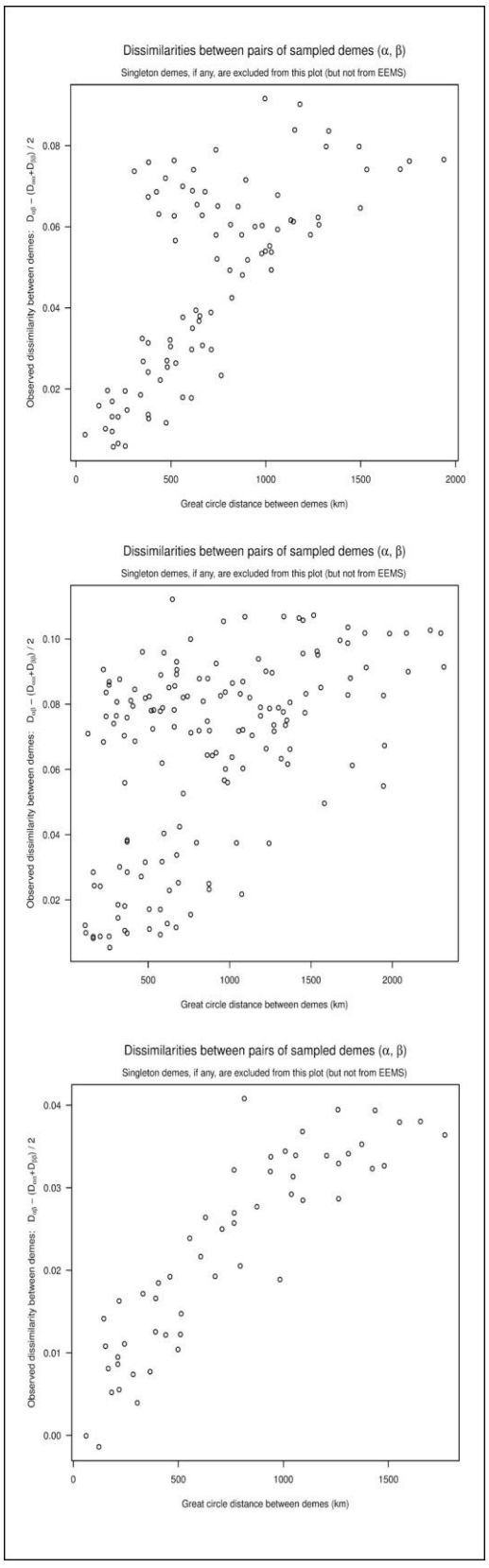

Figure S14: Genetic dissimilarity between geographical demes (y-axis) versus geographic distance between demes obtained in EEMS for *Myrmoborus lugubris* (top), *Thamnophilus nigrocinereus/T. cryptoleucus* (center), and *Myrmotherula assimilis* (bottom).

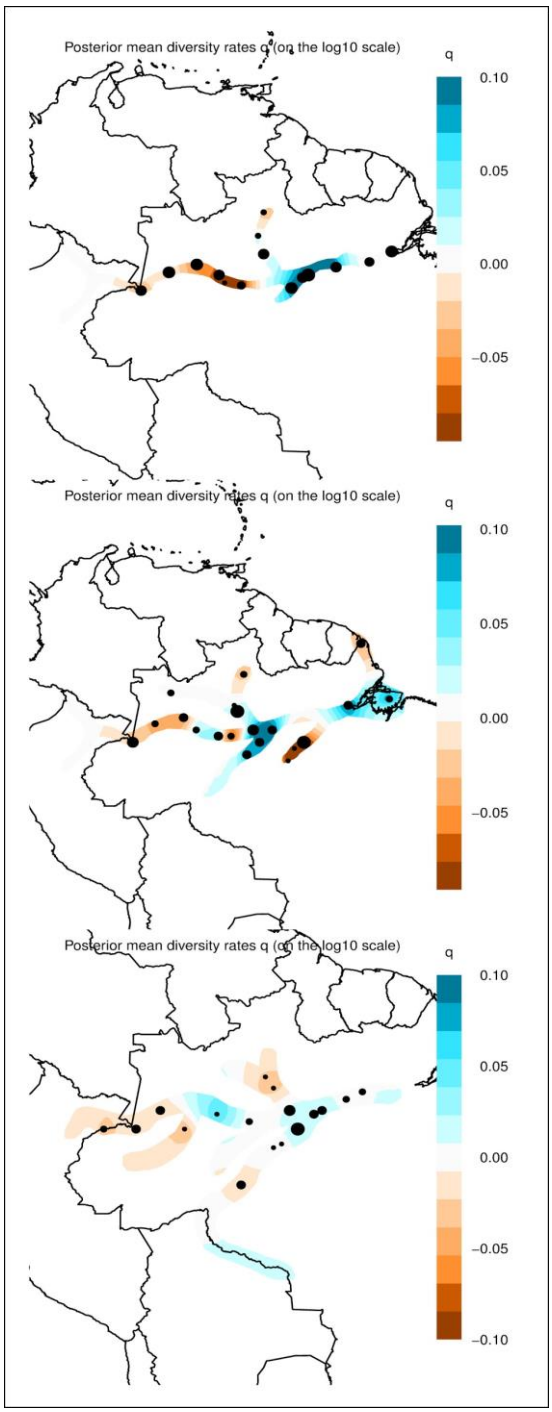

Figure S15: Estimated effective diversity surface in EEMS for *Myrmoborus lugubris* (top), *Thamnophilus nigrocinereus/T. cryptoleucus* (center), and *Myrmotherula assimilis* (bottom). Color bar shows the effective diversity on a log10 scale relative to the average diversity over the entire range of the corresponding species.

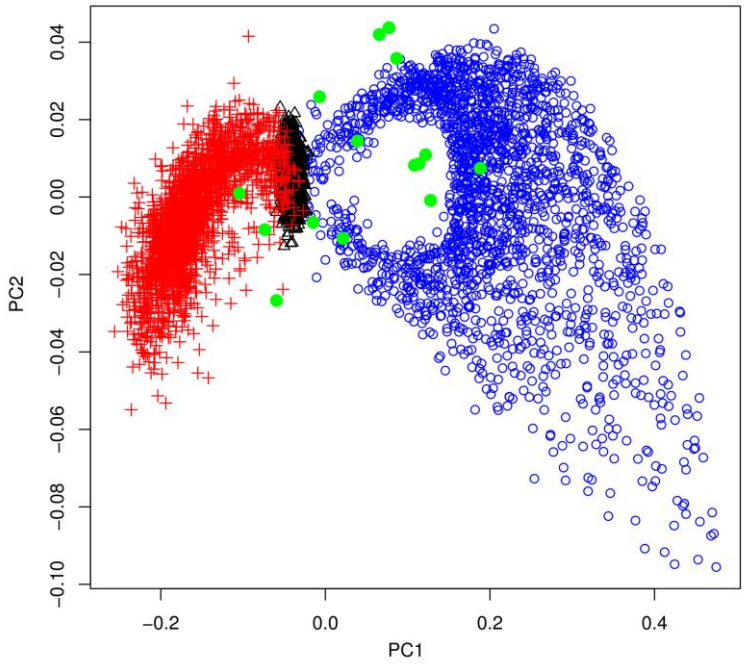

Figure S16: Principal component analyses of the simulated single-pop models with Multi-Dice. Red: instantaneous populational contraction; Black: Populational constant size; Blue: instantaneous population expansion; Green: observed data for the 15 studied population.

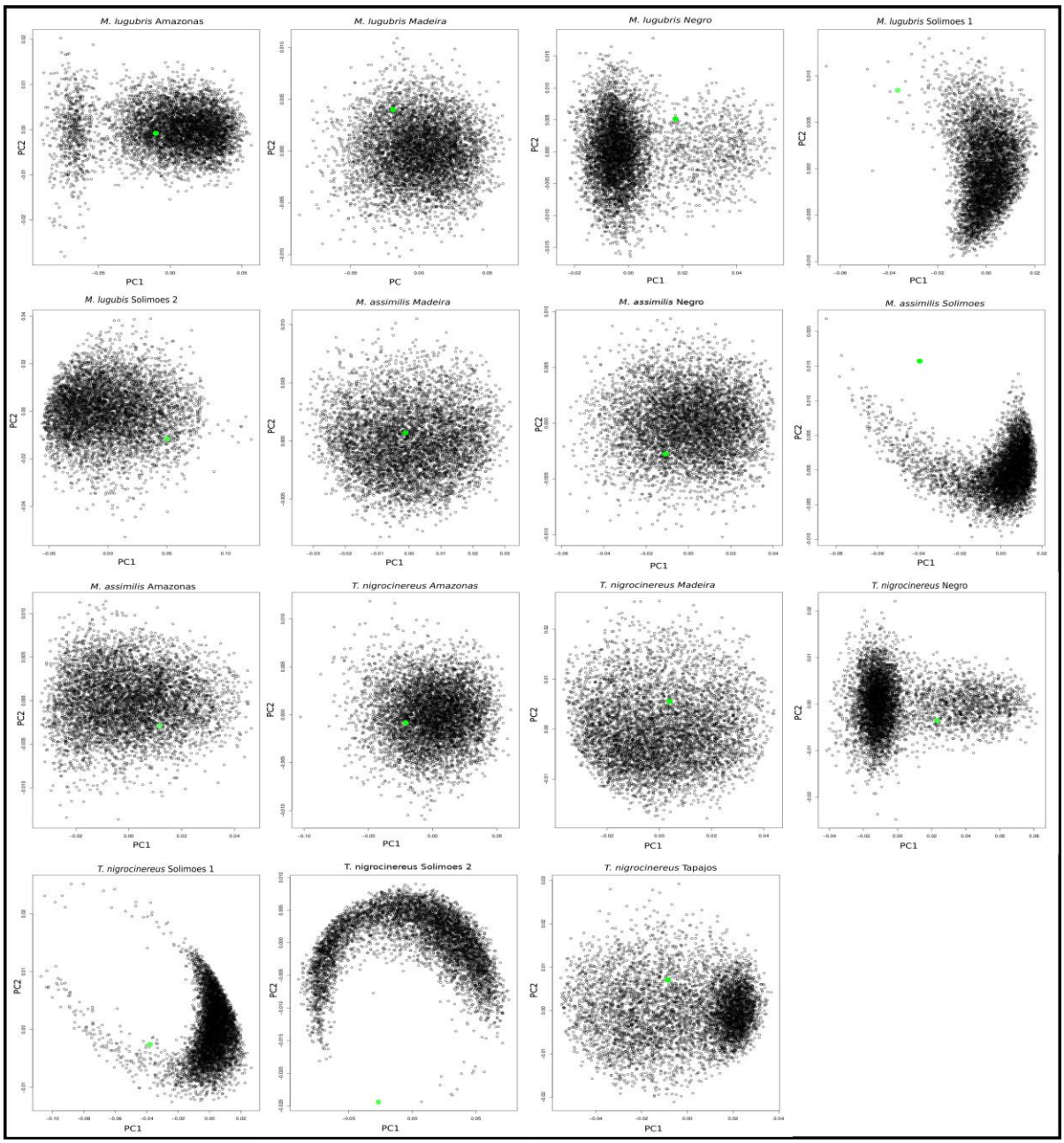

Figure S17: Principal component analyses on the retained simulated SFS of each population. Observed data is shown in green.

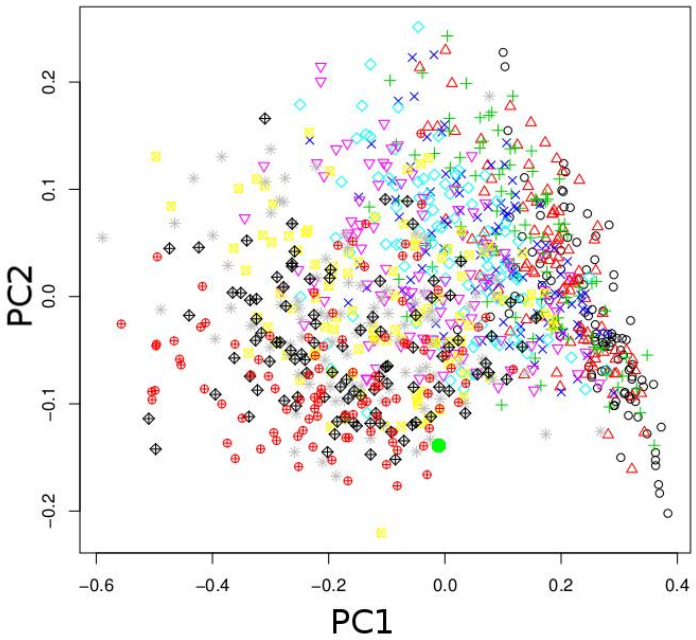

Figure S18: Principal component analysis of simulated and observed (green circles) aSFSs performed in Multi-Dice. Distinct colors and symbols represent the discrete values of ζ, the proportion of populations synchronously expanding: black circle = 0.1; red triangle = 0.2; green cross = 0.3; blue x = 0.4; blue diamond = 0.5; pink inverted triangle = 0.6; yellow crossed circle = 0.7; gray asterisk = 0.8; black crossed diamond = 0.9; red crossed circle = 1.0.

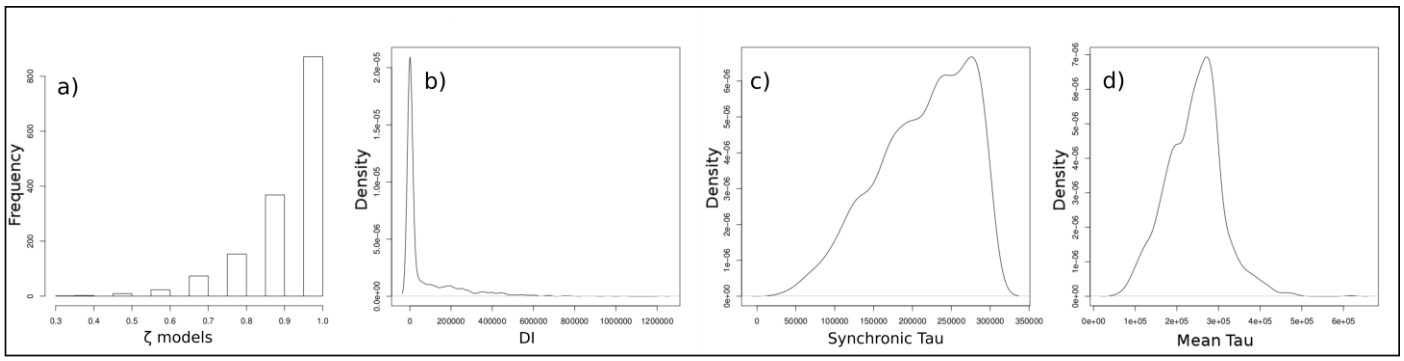

Figure S19: Plots of posterior distributions of hyperparameters obtained in Multi-dice analyses based on the aSFS. a) The proportion of taxa expanding synchronously (ζ); b) Dispersion index (variance of expansion times in relation to the mean expansion time); c) Time of synchronic expansion; d) Mean expansion time across all taxa.

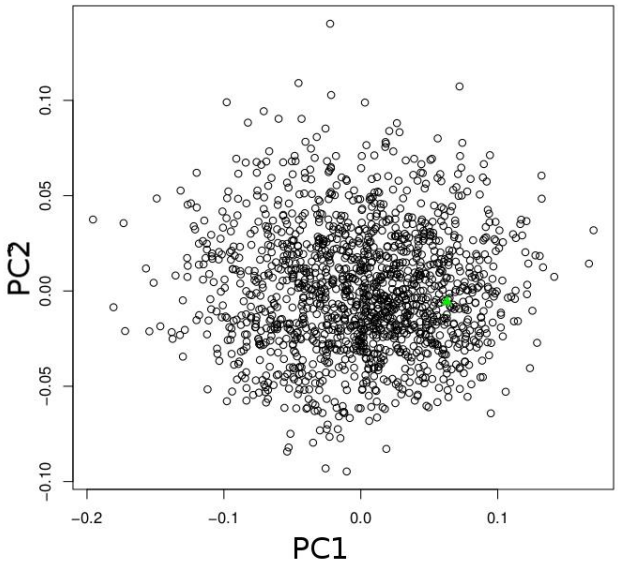

Figure S20: Principal Component Analysis of the 1500 retained simulated aSFS and position of the observed aSFS (green triangle)

Chapter 3

Population dynamics of three Amazonian floodplain forest species of birds restricted to the Solimões river basin

Abstract

Temporally dynamic environments can suffer range alterations affecting the demographic history of their entire biological community. Geographic range expansion can leave a characteristic signature in the current genetic diversity that can be later identified by population genetics methods. The presence of co-distributed populations with similar demography allows to explore more deeply how the current pattern of genetic diversity is distributed and the mechanisms behind its origin. In this study we aimed to understand if climatic oscillations during the Quaternary contracted the geographic distribution of populations restricted to the Solimões basin towards the upper and middle course of the river followed by geographic range expansion during the warmer periods of the Holocene. We selected three species, *M. lugubris, T. cryptoleucus*, and *M. assimilis* and used as genetic marckers sequences of ~2,300 Ultraconserved Elements. The results obtained here did not support a shared spatial pattern of range expansions despite the heterogeneous distribution of the genetic diversity in the three species. Our results indicated lower migration rates in the east and diversity values above the average in western Solimões for *M. lugubris*, but not for *T. cryptoleucus* and *M. assimilis*, suggesting a scenario of subtle habitat change over time and a pronounced effect of intrinsic ecological traits of each species leading to contrasting patterns in the distribution of their genetic diversity.

Introduction

Populations are subject to historical changes in their habitats. Temporally dynamic environments such as areas highly affected by climatic cycles during the Quaternary such as forested areas in the northern hemisphere may contract during glacial periods producing founder events, reducing genetic diversity and increasing genetic drift, while during interglacial periods they may expand geographically (Hewitt 1999; Lesbarrères et al. 2008). Similarly, in population dispersion to new environments, a small number of founder individuals can expand and colonize an entire continent (Handley et al. 2011). Population range expansions can produce successive founder events in the

expansion front yielding fast differentiation from the origin producing a characteristic signature that can be later identified by population genetics methods (Peter & Slatkin 2013). Tracking the patterns of past population range expansion can assist predicting how intraspecific genetic diversity will be influenced by human-mediated environmental changes, enabling more robust conservation strategies (Dawson et al. 2011; Prates et al. 2016).

In the Neotropical region various phylogeographic studies supported population demographic expansions associated to climatic oscillation during the Pleistocene in the Atlantic Forest (Carnaval et al 2009), western Amazon (Thom & Aleixo 2015), Caatinga (Gehara et al. 2017) and the Andes/Patagonia (Cosacov et al. 2010); but only a few studies aimed to understand if these demographic expansions were related to geographic range expansion (Camargo et al. 2013; Guarnizo et al. 2016; Baranzelli et al. 2017). The classical mechanism on how climatic fluctuations may act on populations is described by the refugia hypothesis (Haffer 1969). This hypothesis suggests that populations tend to be isolated in packs of optimum habitats during glacial (eg. lowland forest taxa) or interglacial (eg. montane forest taxa) periods followed by expansion and secondary contact during the opposite climatic condition. In the Amazon floodplains, historical alteration in rainfall patterns of western Amazon and global sea level changes altered the dynamics of the main rivers (Latrubesse & Franzinelli 2005; Irion et al. 2009; Soares et al. 2010; Goncalves-Jr et al. 2016), shaping the distribution of organisms that inhabit river created environments (Aleixo 2006; Choueri et al. 2017). The Amazon basin maintains the largest fluvial system in the world covering more than 300 km^2 regulated by strong seasonal flooding cycles (Wittmann et al. 2010; Junk et al. 2011). Such vast distribution holds a myriad of distinct environments reflecting in the high levels of species endemicity, with around 15% of the nonaquatic species of birds and 10% of the tree species are endemic (Remsen & Parcker 1983; Wittmann et al 2012).

Despite the high diversity of flooded environments few phylogeographic studies were conducted so far in order to understand the potential historical process behind the observed patterns

(Aleixo 2006; Cadena et al. 2011; Harvey et al. 2017; Choueri et al 2017; Thom et al. 2018; Thom et al. in prep., see Chapter 2). Early studies analyzing widely distributed and habitat generalist species (Aleixo 2006; Cadena et al. 2011) reported the absence of genetic structure and low genetic diversity over the entire Amazon basin. However, recent studies of taxa specialist of river created habitats (Choueri et al 2017; Thom et al. 2018; Thom et al. in prep., see Chapter 2), such as islands, support that the central portion of the Amazon basin - where several large tributaries meet - is a suture zone isolating recently diverged populations with marked genetic structure and distinct levels of gene flow. These studies described that for most of the populations occurring in white-water rivers (that are more dynamic and sediment rich than black-water rivers) signals of recent demographic expansions were detected. This suggests that periods with reduced sea level and low rainfall during glacial cycles could have cyclically interrupted the distribution of specific environments in the central portion of the basin and even contracted the geographic distribution of these taxa away from this region.

Thom et al. (in prep., see Chapter 2) conducted a comparative phylogeographic study of three species complexes of antbirds specialized in river edge forests, *Myrmoborus lugubris*, *Thamnophilus nigrocinereus/cryptoleucus* and *Myrmotherula assimilis*. The presence of populations restricted to the Solimões river with signals of synchronic population demographic expansion around 0.2 million years ago (Mya) was observed for the three taxa. The presence of co-distributed populations with similar demography allows to explore more deeply how the current pattern of genetic diversity is distributed and the mechanisms behind its origin; more specifically, did climatic oscillations during the Quaternary contract the geographic distribution of populations restricted to the Solimões basin towards the upper and middle course of the river?

If climatic oscillations during the Quaternary promoted the isolation of populations in the middle and upper courses of Solimões, it is expected that population demographic expansions detected by Thom et al. (in prerp., see Chapter 2) are a product of a recent range expansion from the upper to the lower river course and that populations upriver will present more genetic diversity than those in lower

portions of the distributions of these species. Additionally, if this mechanism was strong enough we expect to detect the same pattern in distinct species. Alternatively, the population demographic expansions observed by Thom et al. (in prep) could be related to a homogeneous reduction in habitat availability over the distributions of these taxa and in this case we do not expect to observe population range expansions. Here we explored the spatial distribution of the current genetic diversity of three species of antbirds with populations restricted to the Solimões river basin testing if previously estimated population demographic expansions are related to range expansions after isolation in the headwaters of this river.

Material and Methods

We selected 31, 20 and 16 tissue samples from Solimões river populations (see Chapter 2) of *Myrmoborus lugubris*, *Thamnophilus cryptoleucus*, and *Myrmotherula assimilis*, respectively (Figure 1; Table S1). Genome-wide variation was obtained through sequence capture of Ultraconserved Elements (UCEs; Faircloth et al. 2012). A minimum amount of 2µg of total DNA per individual was extracted from breast muscle using the QIAGEN DNeasy tissue and Blood kit (Valencia, CA) and QIAGEN RNAse. DNA quantification was performed in Qubit 2.0 fluorometer. Samples were extracted and sent to RAPiD Genomics (Gainsville, FL, USA) for library preparation and sequencing of UCE following the protocol by Faircloth et. al (2012) with the following modifications: 1) probe set targeting 2,312 UCEs (ultraconserved.org) and 97 additional probes targeting exons typically used in avian phylogenetic studies (Hackett et al. 2008; Kimball et al. 2009); 2) sequencing of 150 bp paired-end in Illumina Hiseq 2500. Samples were sequenced in a multiplexed batch of 96 samples. Raw reads' quality was evaluated using FastQC 0.11.4 (Andrews 2014). We trimmed off adapter sequences and filtered low-quality bases (< 15 Phred score) with Illumiprocessor. We used Trinity 2.4 (Grabherr et al. 2011) in order to assemble *de novo* reads into contigs. The contigs obtained were matched to the probes with LASTZ (Harris 2007; http://www.bx.psu.edu/~rsharris/lastz/) using

"match_contigs_to_probes.py". Contigs that did not align to probe sequences and those that matched multiple loci were removed. The resulting fasta files from the different individuals were aligned in MAFFT (Katoh et al. 2013), allowing for missing individuals and without trimming long ragged-ends. These long ragged-ends were then trimmed by applying a threshold of 50% of missing sequences among individuals with TrimAl (Capella-Gutierrez et al. 2009). The longest sequence without indels of each locus was selected as a reference for the following steps. First, reads of each individual were aligned to the reference allowing 4 mismatches per read using BWA (Li & Durbin 2009). The sam files obtained were converted to bam format with Samtools (Li et al. 2009), their reads were trimmed to match the reference using CleanSam.jar, these reads were reassigned to groups with AddOrReplaceReadGroups.jar, and duplicated reads were identified with Markduplicates.jar. We merged individual bam files into a single bam file of all samples with MergeSamFiles.jar using Picard (Broad Institute, Cambridge, MA; http://broadinstitute.github.io/picard/). Additionally, using the Genome Analyses Tool Kit (GATK; McKenna et al. 2010) we realigned all reads, identified indels (RealignerTargetCreator; IndelRealigner), and called SNPs hard-masking low-quality bases (< Q30) with UnifiedGenotyper and VariantAnotator. We obtained raw vcf files for each species that were filtered for a minimum read depth of > 8 using VCFTOOLS (Danecek et al. 2011). Loci with any site with heterozygosity higher than 0.75 were excluded. Finally, we randomly selected one SNP per locus, excluding sites with missing data, resulting in a complete SNP matrix in VCFTOOLS.

Genetic Structure

We tested the best-fit number of ancestral populations (k) for each species complex and clustered individuals to populations in sNMF (Frichot et al. 2014) by applying a sparse non-negative matrix factorization to compute least-square estimates of ancestry coefficients. We explored values of k between 1 and 8, performed 100 replicates for each value, assuming an alpha regularization parameter of 10. Additionally, we performed a Principal Component Analysis (PCA) in ADEGENET (Jombart et

al. 2011).

Isolation by distance, effective diversity and migration

We identified barriers to gene flow and tested deviations from an isolation by distance (IBD) scenario for each species with EEMS (Petkova et al. 2015). EEMS estimates the effective migration surface among drawn demes mapping genetic differentiation based on a spatially explicit approach by using a Markov Chain Monte Carlo (MCMC) to estimate demographic parameters given the observed genetic dissimilarity between individuals. Euclidian genetic dissimilarity matrices between individuals for each species complex were generated in ADEGENET (Jombart et al. 2011). Habitat polygons were produced based on the geographic distribution of each species complex and 300 demes were distributed over the habitat area. Each MCMC run was performed for $30x10^6$ generations with the first $5x10^6$ generations excluded as burn-in. Maps were generated using additional features of the EEMS R package. Additionally, to test IBD we performed a Mantel correlation test between genetic and geographic distances in ADEGENET (Jombart et al. 2011) with 999 replicates.

Geographic range expansion

We calculated a directionality index and test for population range expansion against an IBD model for each species using the Range Expansion package in R (Peter & Slatkin 2013, 2015). This method applies a founder effect algorithm assuming a stepping stone model of population expansion from a single location, testing the strength of the spatial expansion and the most likely origin of the founder effect, measuring the effective founder distance. The effective founder distance represents the size of the deme in which the effective population size is reduced by 1% during the founder event occurring in the expansion front. Thus, low values for the effective founder distance suggest a strong founder effect. To calculate the derived state of each SNP in our datasets for each species, we assumed as outgroup one diploid individual from a closely related population with no signs of shared coefficient of ancestry

in sNMF (see Chapter 2).

Results

The bioinformatics pipeline produced raw vcf files with 5,602 (16.8 mean depth; 2.8 sites per loci), 10,624 (28.4 mean depth; 4.7 sites per loci), and 10,283 (19.9 mean depth; 7.1 sites per loci) variant sites for *M. lugubris*, *T. cryptoleucus*, and *M. assimilis*, respectively. After randomly selecting one SNP per UCE loci we obtained matrices with 1,476, 1,805 and 1,776 SNPs for *M. lugubris*, *T. cryptoleucus*, and *M. assimilis*, respectively.

The number of genetic clusters obtained with sNMF was partially concordant between species. For *M. lugubris* and *T. cryptoleucus* the best value of masked cross-entropy was achieved with K=2 while for *M. assimilis* the best value suggested K=1 (Figure S1). However, the results for K=2 for *M. assimilis* recovered a similar geographic structure as the ones observed in other two species, with a gradual transition over the central portion of the Solimões basin (Figure 2). For *T. cryptoleucus* and *M. assimilis*, this population transition occurred between Tefé and Codajas (localities 9 to 11 in Figure 1) while for *M. lugubris* it occurred more to the west, at Santo Antonio do Ica (locality 6 in Figure 1). The PCA analysis was congruent with the sNMF results supporting a gradual transition of genetic diversity from the west (upriver) to east (down the river) in the Solimões basin in the first component for the three species (Figure S2). By adding the results of the second component three groups of geographically close populations are observed in *M. lugubris* and *T. cryptoleucus*, but not in *M. assimilis* (Figure S2).

The EEMS plots for the three species complexes detected in the eastern portion of the Solimões basin slightly lower effective migration values than the average, suggesting the presence of progressively more dissimilar individuals towards the east in all three species (Figure 3). The estimated effective diversity produced contrasting results between species. For *M. lugubris* there was a considerable lower diversity in eastern localities, for *T. cryptoleucus* there was a considerable lower

diversity in western localities, and for *M. assimilis* a relatively homogeneous distribution of the effective genetic diversity was observed (Figures S3, S4). The results of the Mantel test were congruent with those obtained by EEMS, significantly supporting an IBD scenario in all three species (p-value < 0.01; Mantel r statistics of 0.59, 0.45, and 0.74 for *M. lugubris*, *T. cryptoleucus*, and *M. assimilis*, respectively).

The geographic expansion model was strongly supported over an equilibrium associated to isolation by distance model in the Range Expansion analysis only for *T. cryptoleucus* (p-value < 0.001; q = 0.0004; Table 1; Figure 4) indicating expansion from western Solimões. For *M. lugubris* and *M. assimilis* a scenario of equilibrium by isolation by distance could not be rejected (p-value of 36.17 and 0.86 for *M. lugubris* and *M. assimilis*, respectively).

Discussion

Here we used genetic variation associated with thousands of Ultraconserved Elements to explore spatial patterns of the current genetic diversity in three species of antbirds that inhabit river created environments along the Solimões basin. Our previous results (Thom et al. in prep., see Chapter 2) suggested that populations of these three species experienced demographic expansions around 0.2 million years ago (Mya). Thus, we tested the hypothesis that climatic oscillations during the Quaternary could have drastically reduced the geographic distribution of these species towards the western Solimões basin during periods of lower rainfall and global sea levels, followed by range expansions when the climate changed.

The results obtained in the present study did not support a shared spatial pattern of range expansions. Only for *T. cryptoleucus* a significant range expansion from western to eastern Solimões was detected and thus, did not reject the hypothesis we tested. It was interesting that our data supported a heterogeneous distribution of effective migration signal and genetic diversity between individuals from western and eastern Solimões river in all three species (Figures 3, s3). These results supported

lower migration rates in the east (Figure 3); while diversity values above the average were observed in western Solimões for *M. lugubris*, but not for *T. cryptoleucus* and *M. assimilis* (Figure s3). These results suggested a scenario of subtle habitat change over time and a pronounced effect of intrinsic ecological traits of each species leading to contrasting patterns in the distribution of the genetic diversity, as was also observed in upland forest birds (Smith et al. 2014). It is noteworthy that in the present study these contrasting patterns arouse despite similar population genetic structure, divergence times and demography of these three species over the entire Amazon basin (Thom et al. in prep., see Chapter 2).

Choueri et al. (2017) studying landscape genetics of river island specialists, including the three species studied here, supported that the genetic diversity over Negro river archipelagos is not homogeneously distributed and that the temporal dynamics of formation and disappearance of islands produced distinct patterns of genetic structure and secondary contact matching geological data which indicate the accumulation of sediments and islands formation during the Holocene (Latrubesse & Franzinelli 2005; Latrubesse & Stevaux 2015). In contrast to our results for populations in Solimões river (Thom et al. in prep., see Chapter 2), Choueri et al. (2017) observed only a slight signal of population expansion in the Negro river, in accordance with the idea that rivers with black and clear water (sediment-poor) are less dynamic over time than white-water rivers (sediment-rich rivers).

Given the high effects of IBD (significant correlations in Mantel test), the pattern of genetic structure (K=2) is likely related to this process other than the presence of any physical barrier. A simulation study conducted on fishes reported the high effect of IBD in linear distributions including the formation of differentiated lineages (Thomaz et al. 2016). Thom et al. (in prep.; see Chapter 2) reported that the central portion of the Amazon basin, where the sedimentary basin of the Solimões river flows into hard rock bed of the Amazon river is a suture zone splitting populations and species with distinct levels of genetic distance and gene flow. Thus our data was congruent with the idea that cyclical variation in the average level of Amazonian rivers potentially interrupted the continuous

distribution of Amazonian floodplains in specific portions (e.g. in the central region of the Amazonian basin) and reduced habitat availability over the remaining distribution promoting population size changes but not range contractions and expansions (Choueri et al. 2017). Similarly our results (Thom et al., in prep.; see Chapter 2) supported that these localized interruptions are the result of much more pronounced events than the shifts that occurred along other portions of the river, given the concordant geographical patterns observed in several groups of organisms (Hubert & Renno 2006; Albernaz et al. 2011; Vinticinque et al. 2007; Cohn-Haft et al. 2007).

Additionally to the effects of historical processes, the Amazon basin is subjected to strong seasonal cycles heterogeneously affecting the distribution of specific environments, which could also influence on the observed pattern of genetic diversity (Junk et al. 2011). However, our data is limited in the sense that more sampled localities could improve our resolution in the estimation of potential range expansions. Despite recent studies reported good results using a single diploid individual per locality in a total of 10 localities with a complete matrix of 1,500 SNPs, which is a data set similar to ours, the ideal data set lies around 20 diploid individuals (1 per locality) and 7,000 SNPs (Peter and Slatkin 2013; Poter et al. 2016). Also, the methods we applied did not incorporate the effects of gene flow from other populations and this is an important issue, as Thom et al. (in prep; see Chapter 2) indicated the presence of bidirectional ongoing gene flow between Solimões and eastern Amazonian populations of these species. This process could potentially increase genetic diversity in eastern Solimões localities masking the detection of potential range expansions.

The inferred range expansion in *T. cryptoleucus* pointed to an interesting scenario since this taxon is currently in contact with *T. nigrocinereus tschudi* from the Madeira river. Thus the expansion from western Solimões suggests a diversification in allopatry followed by secondary contact. On the other hand, the Solimões population of *M. lugubris* (*M. l. berlepschi*) is apparently allopatric and more restricted to the upper and middle course of the Solimões river (Figure 2), and a potential - but not detected - geographic expansion could still be in course given the apparent availability of habitats

eastern to its current distribution.

Conclusion

The decreasing of costs for obtaining DNA sequences combined to the throughput of new sequencing methods have been increasing the number of studies of distribution patterns of genetic diversity in a landscape-scale, which represents the long-term register of the evolutionary process. The analysis of these patterns brings new insights into the historical processes and environmental characteristics behind the formation of the current biodiversity. Additionally, the use of co-distributed organisms brings the potential to track pronounced events related to the community assemblage or to control for intrinsic ecological features of the organisms such as dispersion capacity or habitat occupation in diversification studies (Burney and Brumfield 2007; Harvey et al. 2017; Oswald et. al 2017). In this sense, the Amazon floodplains are still an overlooked environment by phylogeography and landscape genetics approaches, given its huge diversity and specific habitats with exclusive evolutionary histories. The continuity and accumulation of studies regarding the diversification of this group of organisms have the potential to expose new scenarios and drivers of speciation of the Amazonian biodiversity.

Bibliography

Albernaz A.L., Pressey R.L., Costa L.R., Moreira M.P., Ramos J.F., Assunção P.A., Franciscon C.H. 2011. Tree species compositional change and conservation implications in the white-water flooded forests of the Brazilian Amazon. Journal of Biogeography. 39:869–883.

Andrews S. 2014. FastQC: A quality control tool for high-throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ Accessed March 9, 2017.

Aleixo A. 2006. Historical diversification of floodplain forest specialist species in the Amazon: a case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). Biological

Journal of the Linnean Society. 89:383–395.

Baranzelli M.C., Cosacov A., Ferreiro G., Johnson L.A., Sérsic A.N. 2017. Travelling to the south: Phylogeographic spatial diffusion model in *Monttea aphylla* (Plantaginaceae), an endemic plant of the Monte Desert. Plos One. 12.

Cadena C.D., Gutiérrez-Pinto N., Dávila N., Chesser R.T. 2011. No population genetic structure in a widespread aquatic songbird from the Neotropics. Molecular Phylogenetics and Evolution. 58:540–545.

Camargo A., Werneck F.P., Morando M., Sites J.W., Avila L.J. 2013. Quaternary range and demographic expansion of *Liolaemus darwinii* (Squamata: Liolaemidae) in the Monte Desert of Central Argentina using Bayesian phylogeography and ecological niche modeling. Molecular Ecology. 22:4038–4054.

Capella-Gutierrez S., Silla-Martinez J.M., Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25:1972–1973.

Carnaval A.C., Hickerson M.J., Haddad C.F.B., Rodrigues M.T., Moritz C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science. 323:785–789.

Choueri É.L., Gubili C., Borges S.H., Thom G., Sawakuchi A.O., Soares E.A.A., Ribas C.C. 2017. Phylogeography and population dynamics of antbirds (Thamnophilidae) from Amazonian fluvial islands. Journal of Biogeography. 44:2284–2294.

Cohn-Haft M., Naka L.N., Fernandes A.M. 2007. Padrões de distribuição da avifauna da várzea dos rios Solimões-Amazonas. Pp. 287-324 In: Albernaz A.L. (eds). Conservação da Várzea: Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ProVárzea, Manaus.

Cosacov A., Sérsic A.N., Sosa V., Johnson L.A., Cocucci A.A. 2010. Multiple periglacial refugia in the Patagonian steppe and post-glacial colonization of the Andes: the phylogeography of *Calceolaria polyrhiza*. Journal of Biogeography. 37:1463-1477.

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., Depristo M.A., Handsaker R.E., Lunter G.,

Marth G.T., Sherry S.T., Mcvean G., Durbin R. 2011. The variant call format and VCFtools. Bioinformatics. 27:2156–2158.

Dawson T.P., Jackson S.T., House J.I., Prentice I.C., Mace G.M. 2011. Beyond predictions: biodiversity conservation in a changing climate. Science. 332:53–58.

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

Frichot E., Mathieu F., Trouillon T., Bouchard G., François O. 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics. 196:973–983.

Gehara M., Garda A.A., Werneck F.P., Oliveira E.F., Fonseca E.M.D., Camurugi F., Magalhães F.D.M., Lanna F.M., Sites J.W., Marques R., Silveira-Filho R., Pedro V.A.S., Colli G.R., Costa G.C., Burbrink F.T. 2017. Estimating synchronous demographic changes across populations using hABC and its application for a herpetological community from northeastern Brazil. Molecular Ecology. 26:4756–4771.

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

Guarnizo C.E., Werneck F.P., Giugliano L.G., Santos M.G., Fenker J., Sousa L., D'Angiolella A.B., Santos A.R.D., Strüssmann C., Rodrigues M.T., Dorado-Rodrigues T.F., Gamble T., Colli G.R. 2016. Cryptic lineages and diversification of an endemic anole lizard (Squamata, Dactyloidae) of the Cerrado hotspot. Molecular Phylogenetics and Evolution. 94:279–289.

Hackett S.J., Kimball R.T., Reddy S., Bowie R.C.K., Braun E.L., Braun M.J., Chojnowski J.L., Cox W.A., Han K.-L., Harshman J., Huddleston C.J., Marks B.D., Miglia K.J., Moore W.S., Sheldon

F.H., Steadman D.W., Witt C.C., Yuri T. 2008. A Phylogenomic study of birds reveals their evolutionary history. Science. 320:1763–1768.

Haffer J. 1969. Speciation in Amazonian forest birds. Science. 165:131–137.

Handley L.-J.L., Estoup A., Evans D.M., Thomas C.E., Lombaert E., Facon B., Aebi A., Roy H.E. 2011. Ecological genetics of invasive alien species. BioControl. 56:409–428.

Harris R.S. 2007. Improved pairwise alignment of genomic DNA. Ph.D. Thesis. The Pennsylvania State University.

Harvey M., Aleixo A., Ribas C.C., Brumfield R.T. 2016. Habitat association predicts genetic diversity and population divergence in Amazonian birds. The American Naturalist 190: 631-648.

Hewitt G.M. 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society. 68:87–112.

Hubert N., Renno J.-F. 2006. Historical biogeography of South American freshwater fishes. Journal of Biogeography. 33:1414–1436.

Irion G., Müller J., Morais J.O., Keim G., Mello J.N.D., Junk W.J. 2009. The impact of Quaternary sea level changes on the evolution of the Amazonian lowland. Hydrological Processes. 23:3168–3172.

Jombart T., Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 27:3070–3071.

Junk W.J., Piedade M.T.F., Schöngart J., Cohn-Haft M., Adeney J.M., Wittmann F. 2011. A classification of major naturally-occurring Amazonian lowland wetlands. Wetlands. 31:623–640.

Júnior E.S.G., Soares E.A.A., Tatumi S.H., Yee M., Mittani J.C.R. 2016. Pleistocene-Holocene sedimentation of Solimões-Amazon fluvial system between the tributaries Negro and Madeira, central Amazon. Brazilian Journal of Geology. 46:167–180.

Katoh K., Standley D.M. 2013. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 30:772–780.

Kimball R.T., Braun E.L., Barker F.K., Bowie R.C., Braun M.J., Chojnowski J.L., Hackett S.J., Han K.-L., Harshman J., Heimer-Torres V. 2009. A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolution. 50:654–660.

Latrubesse E.M., Franzinelli E. 2002. The Holocene alluvial plain of the middle Amazon river, Brazil. Geomorphology. 44:241–257.

Latrubesse E.M., Franzinelli E. 2005. The late Quaternary evolution of the Negro river, Amazon, Brazil: Implications for island and floodplain formation in large anabranching tropical systems. Geomorphology. 70:372–397.

Latrubesse E. M., Stevaux J. C. 2015. The Anavilhanas and Mariu archipelagos: Fluvial wonders from the Negro river, Amazon basin. Pp. 157-169 In: Vieira B. C., Salgado A. A. R., Santos L. J. C. (Eds). Landscapes and Landforms of Brazil. Springer, Dordrecht.

Lesbarrères D. 2008. Post-glacial phylogeography: New insight into an old story: the post-glacial recolonization of European biota. Heredity. 102:213–213.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25:1754–1760.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079.

Mckenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., Depristo M.A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research. 20:1297–1303.

Oswald J.A., Overcast I., Mauck W.M., Andersen M.J., Smith B.T. 2017. Isolation with asymmetric gene flow during the nonsynchronous divergence of dry forest birds. Molecular Ecology. 26:1386–1400.

Peter B.M., Slatkin M. 2013. Detecting range expansions from genetic data. Evolution. 67:3274–3289.

Peter B.M., Slatkin M. 2015. The effective founder effect in a spatially expanding population. Evolution. 69:721–734.

Petkova D., Novembre J., Stephens M. 2014. Visualizing spatial population structure with estimated effective migration surfaces. Nature Genetics 48:94-100.

Potter S., Bragg J.G., Peter B.M., Bi K., Moritz C. 2016. Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. Molecular Ecology. 25:1367–1380.

Remsen J.V., Iii T.A.P. 1983. Contribution of river-created habitats to bird species richness in Amazonia. Biotropica. 15:223.

Smith B.T., Mccormack J.E., Cuervo A.M., Hickerson M.J., Aleixo A., Cadena C.D., Pérez-Emán J., Burney C.W., Xie X., Harvey M.G., Faircloth B.C., Glenn T.C., Derryberry E.P., Prejean J., Fields S., Brumfield R.T. 2014. The drivers of tropical speciation. Nature. 515:406–409.

Soares E.A., Tatumi S.H., Riccomini C. 2010. OSL age determinations of Pleistocene fluvial deposits in Central Amazonia. Anais da Academia Brasileira de Ciências. 82:691–699.

Thom G., Aleixo A. 2015. Cryptic speciation in the white-shouldered antshrike (*Thamnophilus aethiops*, Aves – Thamnophilidae): The tale of a transcontinental radiation across rivers in lowland Amazonia and the northeastern Atlantic Forest. Molecular Phylogenetics and Evolution. 82:95–110.

Thom G., Amaral F.R., Hickerson M.J., Aleixo A., Araujo-Silva L.E., Ribas C.C., Choueri E., Miyaki C.Y. 2018. Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species. Systematic Biology.

Thomaz A.T., Christie M.R., Knowles L.L. 2016. The architecture of river networks can drive the evolutionary dynamics of aquatic populations. Evolution. 70:731–739.

Venticinque E.M., Rego F.N.A.A., Brescovit A.D., Rheims C.A., Ruiz G.R.S. 2007. A araneofauna (Arachnida, Araneae) das várzeas do rio Amazonas: padrões de distribuição e estado do

conhecimento atual. Pp. 179-198 In: Albernaz A.L. (eds). Conservação da Várzea: Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ProVárzea, Manaus.

Wittmann F., Householder E., Piedade M.T.F., Assis R.L.D., Schöngart J., Parolin P., Junk W.J. 2012. Habitat specificity, endemism and the Neotropical distribution of Amazonian white-water floodplain trees. Ecography. 36:690–707.

Wittmann F., Schongart J., Junk W. J. 2010. Phytogeography, species diversity, community structure and dynamics of central Amazonian floodplains forests. Pp. 61-102 In: Junk W. J., Piedade M. T. F., Wittmann F., Schongart J., Parolin P. (eds), Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management. Springer, New York.

Table 1: Range expansion summary results for the three species studied.

Parameters	***M. lugubris***	***T. cryptoleucus***	***M. assimilis***
longitude	-62.95	-73.22144	-75.64357
latitude	-4.021451	-9.999475	-3148
r1	0.999	0.999	0.999
r10	0.996	0.991	0.997
r100	0.964	0.918	0.975
d1	27.19	11.45	40.80
r^2	0.45	0.89	0.31
p-value	36.17	2.10e-23	11.26

r1, r10, and r100 - decrease in diversity at 1, 10, and 100 kilometers of distance, respectively; d1 – effective founder distance, which represents the deme size in km for the effective population size to be reduced by 1% in the expansion front; r^2 and p-value - correlation coefficient and correlation p-value for the most likely origin, respectively.

Table S1. Individuals sampled of *Myrmoborus lugubris*, *Thamnophilus cryptoleucus* and *Myrmotherula assimilis*.

n°	Taxon	Id.	Inst.	Voucher	Locality	Latitude	Longitude
1	*M. l. berleschi*	***	LSUMZ	93417	Peru, Rio Huallaga, Isla Chaupi	5°23'56"S	75°48'04"W
1	*M. a. assimilis*	***	LSUMZ	93460	Peru, Rio Huallaga, Isla Chaupi	5°23'56"S	75°48'04"W
1	*T. cryptoleucus*	***	LSUMZ	93318	Peru, Rio Huallaga, Isla Chaupi	5°23'56"S	75°48'04"W
2	*T. cryptoleucus*	***	LSUMZ	74103	Peru, Provincia Satipo; Distrito Rio Tambo; Rio Tambo; isla entre Canuja y Atalaya	10°44'49"S	73°44'46"W
3	*T. cryptoleucus*	***	LSUMZ	7285	Peru, Isla Pasto, Rio Amazonas opposite Aysana, ca 80 km NE Iquitos	3°27'29"S	72°40'30"W
4	*M. a. assimilis*	AMA 427	MPEG	72924	AM, Estirão do Equador, Atalaia do Norte	04°29'12,1"S	71°33'02,3"W
4	*M. a. assimilis*	AMA 475	MPEG	72972	AM, Estirão do Equador, Atalaia do Norte	04°29'12,1"S	71°33'02,3"W
5	*M. l. berleschi*	ETA 001	MPEG	T22430	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. l. berleschi*	ETA 002	MPEG	T22431	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. l. berleschi*	ETA 004	MPEG	T22433	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. l. berleschi*	ETA 009	MPEG	T22438	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. l. berleschi*	ETA 052	MPEG	T22480	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. l. berleschi*	ETA 053	MPEG	T22481	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. a. assimilis*	ETA 046	MPEG	T22475	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. a. assimilis*	ETA 051	MPEG	T23495	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*M. a. assimilis*	ETA 054	MPEG	T23497	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*T. cryptoleucus*	ETA 003	MPEG	T22432	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*T. cryptoleucus*	ETA 005	MPEG	T22434	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*T. cryptoleucus*	ETA 007	MPEG	T22436	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
5	*T. cryptoleucus*	ETA 008	MPEG	T22437	AM, Tabatinga, Ilha Aramaçá, comunidade Cristo Rei	04°19'27,7"S	69°53'10,2"W
6	*M. l. berleschi*	ETA 073	MPEG	T22496	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. l. berleschi*	ETA 074	MPEG	T22497	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. l. berleschi*	ETA 093	MPEG	T22511	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. l. berleschi*	ETA 094	MPEG	T22512	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. l. berleschi*	ETA 096	MPEG	T22514	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. l. berleschi*	ETA 114	MPEG	T22527	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. l. berleschi*	ETA 115	MPEG	T22528	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. a. assimilis*	ETA 081	MPEG	T22503	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. a. assimilis*	ETA 097	MPEG	T23506	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*M. a. assimilis*	ETA 098	MPEG	T23507	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W

6	*T. cryptoleucus*	ETA 062	MPEG	T23498	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
6	*T. cryptoleucus*	ETA 063	MPEG	T22488	AM, Santo Antônio do Içа, Ilha Porto América	03°08'52,7"S	67°57'35,1"W
7	*M. a. assimilis*	JUT269	MPEG	T23446	AM, Jutaí, ESEC Jutaí/Solimões, Pati	03°13'29,9"S	67°26'10"W
8	*M. l. berleschi*	ETA 162	MPEG	T23519	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*M. l. berleschi*	ETA 163	MPEG	T23520	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*M. l. berleschi*	ETA 164	MPEG	T22566	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*M. l. berleschi*	ETA 165	MPEG	T22567	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*M. l. berleschi*	ETA 166	MPEG	T22568	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*M. l. berleschi*	ETA 194	MPEG	T23532	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*M. l. berleschi*	ETA 196	MPEG	T23534	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*T. cryptoleucus*	ETA 174	MPEG	T22574	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*T. cryptoleucus*	ETA 175	MPEG	T22575	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
8	*T. cryptoleucus*	ETA 176	MPEG	T22576	AM, Fonte Boa, Rio Solimões, Ilha Nova esperança	02°28'58,2"S	66°00'38,4"W
9	*M. l. berleschi*	ETA 233	MPEG	T22612	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
9	*M. l. berleschi*	ETA 234	MPEG	T22613	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
9	*M. l. berleschi*	ETA 235	MPEG	T22614	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
9	*M. l. berleschi*	ETA 236	MPEG	T22615	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
9	*M. l. berleschi*	ETA 237	MPEG	T22616	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
9	*M. l. berleschi*	ETA 238	MPEG	T22617	AM, Tefé, Rio Solimões, Ilha do Içé	03°15'57,2"S	64°42'25,9"W
9	*M. l. berleschi*	TVVC 25	INPA	A936	AM, Tefe, Ilha do Barbado	3°46'51,8"S	64°1'31,3"W
9	*T. cryptoleucus*	ETA 201	MPEG	T22585	AM, Tefé, Rio Solimões, Ilha 1	03°22'46,0"S	64°38'27,8"W
9	*T. cryptoleucus*	ETA 202	MPEG	T22586	AM, Tefé, Rio Solimões, Ilha 1	03°22'46,0"S	64°38'27,8"W
10	*M. l. berleschi*	ETA 246	MPEG	T22625	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*M. l. berleschi*	ETA 256	MPEG	T22631	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*M. l. berleschi*	ETA 257	MPEG	T22632	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*M. l. berleschi*	ETA 258	MPEG	T22633	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*M. a. assimilis*	ETA 212	MPEG	T22595	AM, Tefé, Rio Solimões, Ilha 1	03°22'46,0"S	64°38'27,8"W
10	*M. a. assimilis*	ETA 254	MPEG	T22630	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*M. a. assimilis*	ETA 269	MPEG	T23548	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*T. cryptoleucus*	ETA 240	MPEG	T22619	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*T. cryptoleucus*	ETA 242	MPEG	T22621	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
10	*T. cryptoleucus*	ETA 243	MPEG	T22622	AM, Coari, Rio Solimões, Ilha da Botija	04°00'49,5"S	62°57'00,5"W
11	*T. cryptoleucus*	ETA 289	MPEG	T22661	AM, Codajás, Rio Solimões, ilha do camaleão	03°50'36,4"S	62°14'29,7"W

11	*T. cryptoleucus*	ETA 290	MPEG	T22662	AM, Codajás, Rio Solimões, ilha do camaleão	03°50'36,4"S	62°14'29,7"W
12	*M. a. assimilis*	ETA 337	MPEG	T22719	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*M. a. assimilis*	ETA 338	MPEG	T22720	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*M. a. assimilis*	ETA 339	MPEG	T22721	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*M. a. assimilis*	ETA 353	MPEG	T22735	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*T. cryptoleucus*	ETA 331	MPEG	T22713	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*T. cryptoleucus*	ETA 332	MPEG	T22714	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*T. cryptoleucus*	ETA 334	MPEG	T22716	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
12	*T. cryptoleucus*	ETA 335	MPEG	T22717	AM, Iranduba, Rio Solimões, Ilha da Marchantaria	03°17'15,9"S	60°12'54,1"W
13	*T. cryptoleucus*	***	LSUMZ	25431	AM, Ilha do Careiro, ca 20 km E Manaus	3°08'24"S	59°50'24"W
out	*M. l. lugibris*	ETA 598	MPEG	T22926	PA, Almeirim, Rio Amazonas, Ilha do Camaleão	01°32'15,2"S	52°31'25,9"W
out	*M. a. assimilis*	ETA 559	MPEG	T23587	AM, Parintins, Marg. Esquerda, Rio Amazonas, Praia do Brilhante	02°34'47,5"S	56°40'45,0"W
out	*T. n. cinereoniger*	SGC 200	MPEG	T20464	AM, São Gabriel da Cachoeira, Ilha da Brigada	0°08'S	67°04'W

n° - number of localities sampled as in Figure 1; id. - Field collection number; *** - n/a; Inst. - Institution; out - outgroups

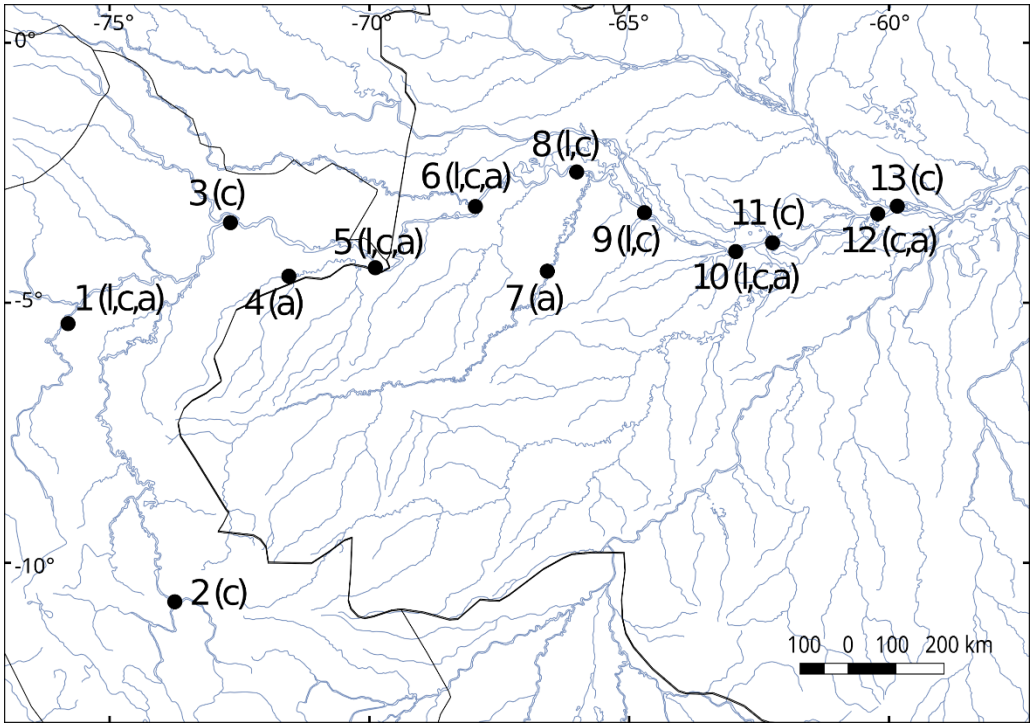

Figure 1: Geographic distribution of sampled individuals of *Myrmoborus lugubris* (l), *Thamnophilus cryptoleucus* (c) and *Myrmotherula assimilis* (a) over the Solimões basin. Numbers in the map represent localities as in Table S1.

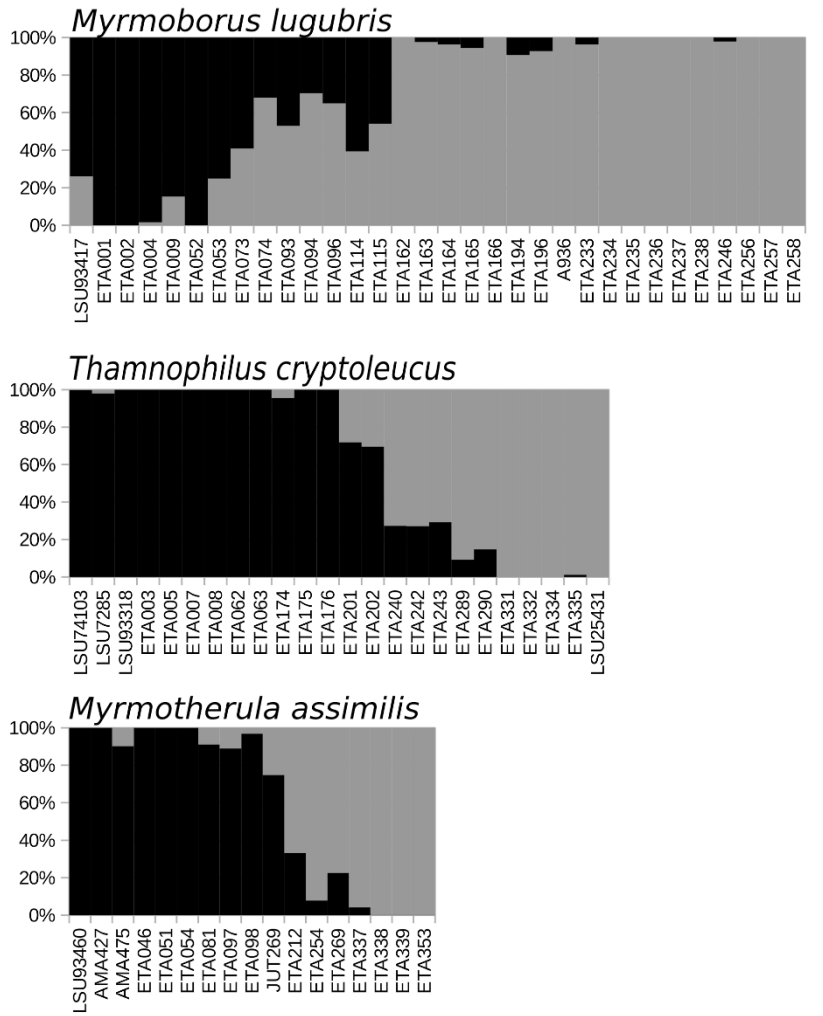

Figure 2: Population structure and individual coefficient of ancestry (bars) inferred with sNMF for the three species studied. Codes below the bars represent identification numbers of individuals as in Table S1. Note that for *M. assimilis*, K=2 was the model with the second highest cross-entropy value.

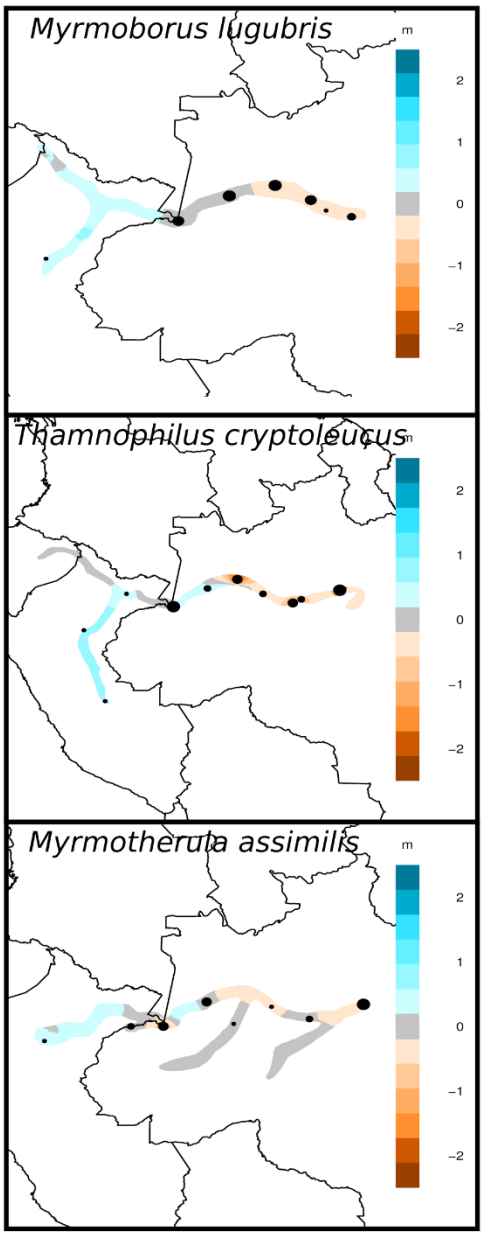

Figure 3: Estimated effective migration surface (EEMS) of three species studied. Colors represent a log10 scale from the average effective migration (gray).

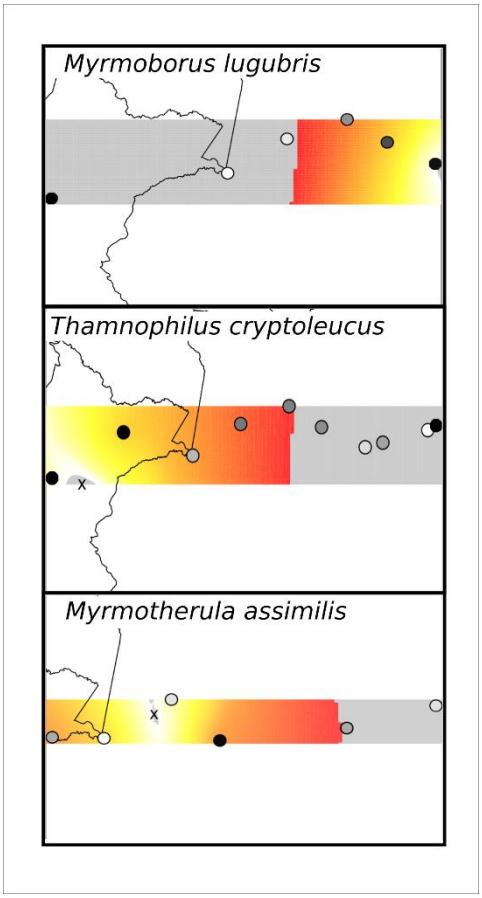

Figure 4: Origin of the potential geographic expansion for the three species studied inferred with the Range Expansion package. The only significant range expansion signal was detected for *T. cryptoleucus* (p-value<0.001). The sampled localities in each population are represented by dots and the most likely origin by an X. Colors represent how likely a given portion of the distribution can be the origin, with yellow tones being more likely than red. Gray tones are regions with negative founder effect signal and are hence unlikely to be the source of the expansion.

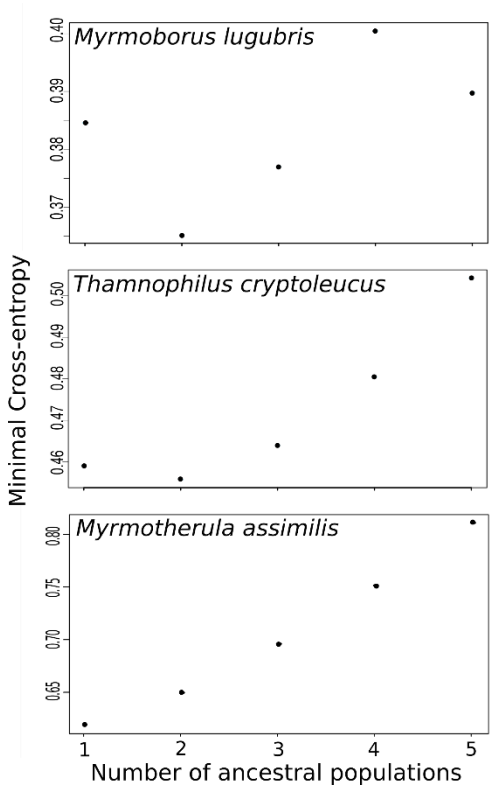

Figure S1: Minimal cross-entropy values obtained for different numbers of ancestral populations (K) in sNMF for the three species studied.

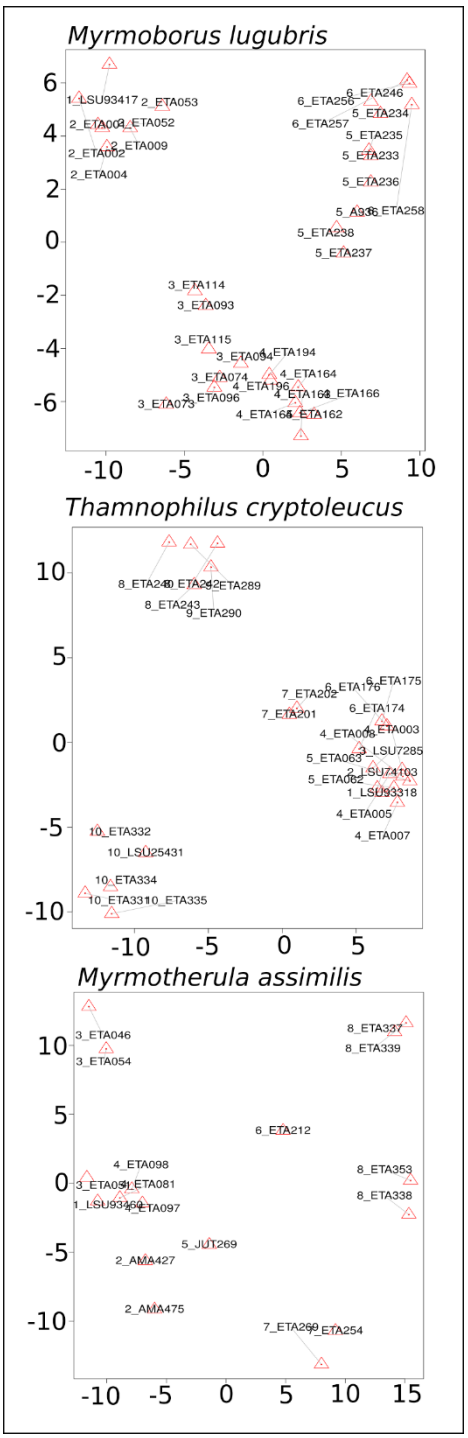

Figure S2: Principal component analysis of the three species studied. x-axis - first component; y-axis - second component. Codes represent identification numbers of individuals as in Table S1.

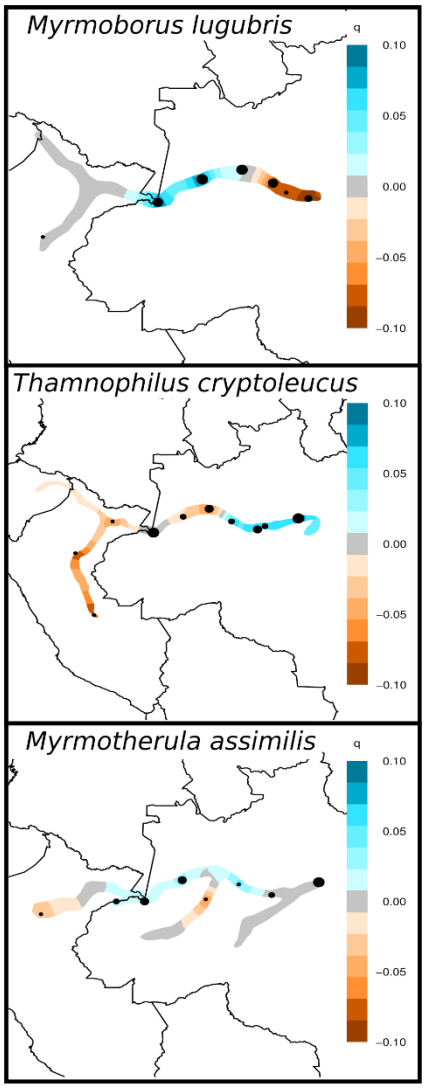

Figure S3: Estimated effective diversity surface (EEDS) of the three species studied. Colors represent a log10 scale of the average effective diversity (gray).

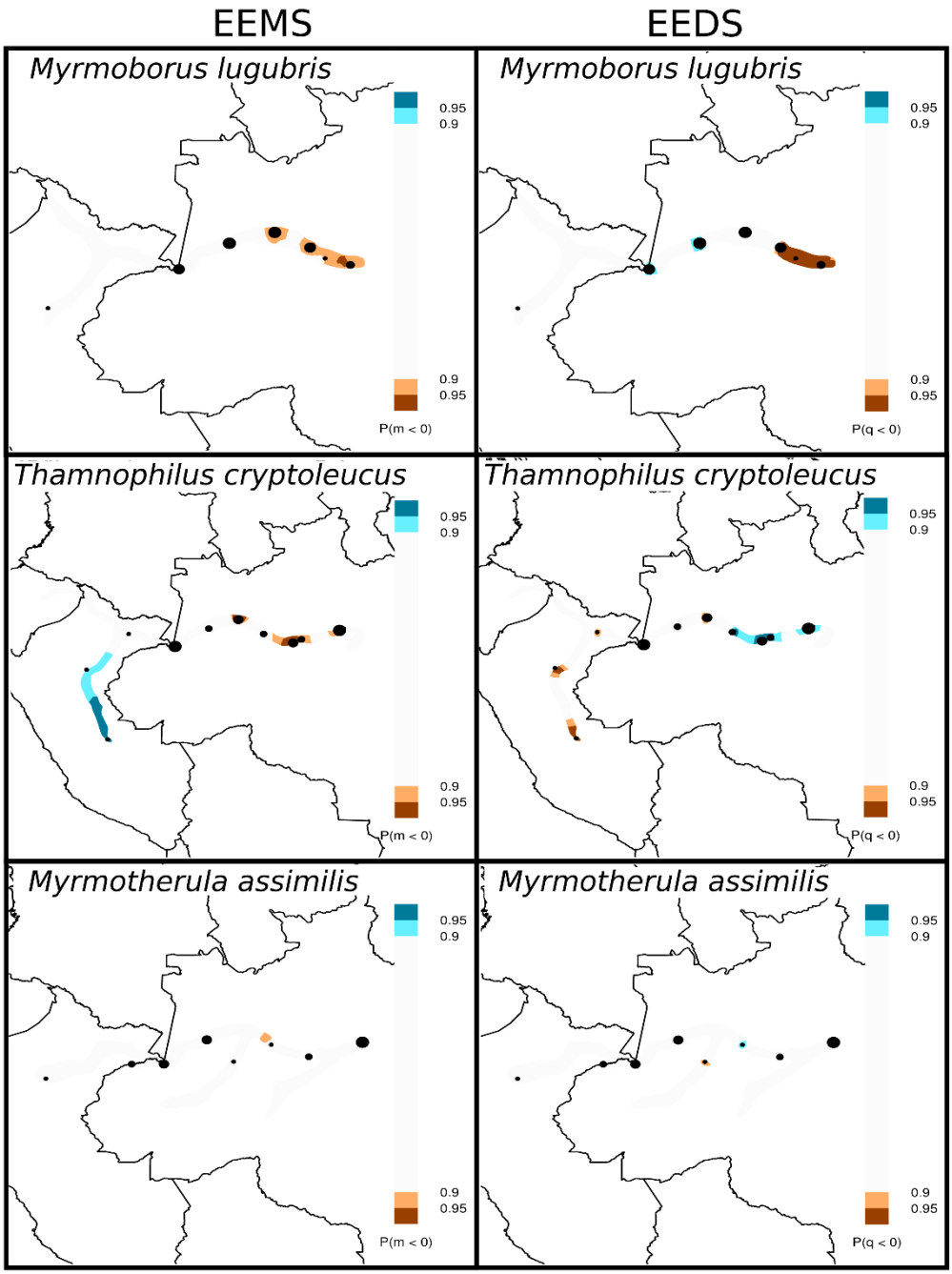

Figure S4: Confidence interval of posterior distributions obtained for effective migration (EEMS) and effective diversity (EEDS) of the three species studied.

In this Dissertation we described diversification patterns of three bird species complexs restricted to the Amazonian floodplains, *Myrmoborus lugubris, Thamnophilus nigrocinereus/T. cryptoleucus*, and *Myrmotherula assimilis*, which enabled us to inferer potential historical processes that shaped current genetic diversity and geographic distribution of this community. In Chapter 1 the pattern of genetic structure and phylogenetic relationships of *Myrmoborus lugubris* populations were studied revealing a complex scenario of diversification. In Chapter 2 the patterns of co-divergence and co-demography of three studied species were explored. Finally in Chapter 3 the dynamics of geographic range expansion of populations restricted to the Solimões river was analyzed.

Diversification of Amazonian floodplain forest organisms

The diversification patterns observed here supported that Amazonian floodplains are a highly dynamic environment over time producing complex scenarios of diversification by modulating microevolutionary processes such as gene flow and genetic drift. The effects of gene flow and incomplete lineage sorting observed here support the demand for larger datasets as those obtained with next-generation sequencing and the application of methods that can handle both processes to estimate the phylogenetic relationship of species and populations occurring in this dynamic environment. Similarly, given the discrepancy between divergence times from sister species and within species complexes, this dynamism may promote high extinction rates constantly vanishing genetic diversity.

The phylogeographic patterns observed here revealed an interesting and unexplored facet for the diversification of the large Amazonian biodiversity. On the opposite to the expectation of the rivers as barriers hypothesis - where large Amazonian tributaries are expressive barriers to gene flow for upland forest organisms - the floodplain forest species studied in this Dissertation are

compartmentalized over the main Amazonian sub-basins such as Solimões, Negro, Madeira, Tapajós and Amazonas, without signs of isolation on opposite margins of large rivers. The results obtained were in disagreement with previous phylogeographic studies that supported widely distributed populations over the entire Amazon basin, suggesting that more specialized species in terms of habitat tend to be more affected by historical events. Hence supporting that the diversification processes of the floodplain community cannot be explained by a single generalized model, even considering that some taxon present similar histories. This complex biogrographical history is also observed for upland Amazonian forest species.

The results presented in this Dissertation support that the central Amazonian basin is a suture zone splitting populations and species occurring in distinct sub-basins with variable levels of gene flow. Additionally, we suggest that climatic oscillations around 0.2 Mya produced events of isolation and secondary contact following river level fluctuations that altered their sedimentation patterns, mainly in white-water rivers. However, given potential high extinction rates in this environments it is possible that the effect of older historical processes have been erased and are no longer detectable by the methods we applied. Hence it is possible that multiple events operated in the current genetic diversity of distinct taxa that may have responded differently given intrinsic ecological attributes producing a complex scenario. In this sense, the Amazonian floodplains are still an overlooked environment by phylogeography and landscape genetics approaches, given its huge diversity and specific habitats with exclusive evolutionary histories. Thus the continuity and accumulation of studies regarding the diversification of floodplain organisms have the potential to expose new scenarios and drivers of speciation for the Amazonian biodiversity. Similarly, the association between genetic and environmental data such as niche modeling and remote sensing can yield a better understanding of what physical features are more effective shaping the observed genetic diversity.