



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCÊNCIAS
DEPARTAMENTO DE ZOOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL

LAYSE ALBUQUERQUE DA SILVA LUCAS

FILOGEOGRAFIA DOS CORRUPÇÕES DO COMPLEXO *Icterus icterus* (FAMÍLIA:
ICTERIDAE): INFERÊNCIAS EVOLUTIVAS SOBRE AS FLORESTAS TROPICAIS
SAZONALMENTE SECAS

Recife
2020

LAYSE ALBUQUERQUE DA SILVA LUCAS

**FILOGEOGRAFIA DOS CORRUPÇÕES DO COMPLEXO *Icterus icterus* (FAMÍLIA:
ICTERIDAE): INFERÊNCIAS EVOLUTIVAS SOBRE AS FLORESTAS TROPICAIS
SAZONALMENTE SECAS**

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Animal da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Biologia Animal.

Orientador: Prof. Dr. Luciano N. Naka

Coorientador: Dr^a. Maria W. Pil

Recife

2020

Catálogo na fonte
Elaine C Barroso
(CRB4 1728)

Lucas, Layse Albuquerque da Silva

Filogeografia dos corrupiões do complexo *Icterus icterus* (Família: Icteridae): Inferências evolutivas sobre as florestas tropicais sazonalmente secas / Layse Albuquerque da Silva Lucas – 2020.

53 f.: il., fig., tab.

Orientador: Luciano N. Naka

Coorientadora: Maria W. Pill

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-Graduação em Biologia Animal, 2020. Inclui referências.

1. Corrução 2. História evolutiva 3. Filogeografia I. Naka, Luciano N. (orient.) II. Pill, Maria W. (coorient.) III. Título

598.2

CDD (22.ed.)

UFPE/CB – 2020- 253

LAYSE ALBUQUERQUE DA SILVA LUCAS

**FILOGEOGRAFIA DOS CORRUPÇÕES DO COMPLEXO *Icterus icterus* (FAMÍLIA:
ICTERIDAE): INFERÊNCIAS EVOLUTIVAS SOBRE AS FLORESTAS TROPICAIS
SAZONALMENTE SECAS**

Dissertação apresentada ao Programa de Pós-Graduação
em Biologia Animal da Universidade Federal de
Pernambuco, como requisito parcial para a obtenção do
título de Mestre em Biologia Animal.

Aprovada em 30 de outubro de 2020

BANCA EXAMINADORA

Prof. Dr. Luciano Nicolás Naka (Orientador)
Universidade Federal de Pernambuco

Prof.^a Dr.^a. Luciana Iannuzzi (Examinador Interna)
Universidade Federal de Pernambuco

Prof. Dr. Fábio Sarubbi Raposo do Amaral (Examinador Externo)
Universidade Federal de São Paulo

Prof. Dr. Mario Eric Cohn Haft (Examinador Externo)
Instituto Nacional de Pesquisas da Amazônia

Prof. Dr. Pedro Murilo Sales Nunes (Suplente Interno)
Universidade Federal de Pernambuco

Dr. Gustavo Adolfo Bravo Mora (Suplente Externo)
Harvard University

AGRADECIMENTOS

Aos meus pais e meu irmão e a toda a minha família que me incentivaram, acreditaram e apoiaram em todos os momentos.

Ao meu orientador, Luciano Naka, o qual tenho grande admiração pelo educador/cientista que é e por ter confiado a mim uma oportunidade única e especial, que mudou minha vida acadêmica e pessoal. Serei sempre grata por ter aberto as portas científicas e por ter me mostrado o mundo.

A minha Co-orientadora, Maria Pil, por todo suporte, ajuda e incentivo. Sou grata por todas as discussões, e pelo empenho em tudo o que fez por mim.

Aos meus colegas do Ornitolab, muito obrigada pelo apoio e suporte durante todo o período de graduação e mestrado. Obrigada pelas trocas de conhecimento e por sempre me ajudarem no meu trabalho.

Aos meus amigos da Louisiana, minha gratidão por terem permanecido na minha vida e serem os melhores amigos gringos que eu poderia ter, em especial Daniel Lane, por ser um grande incentivador, amigo e recitador de *Tongue-twister*.

Ao Museu de Ciências Naturais da Universidade da Louisiana, por terem aberto as portas para mim e terem me ensinado não só no ambiente laboratorial mas também na vida científica.

Às minhas melhores amigas, Luana Guimarães e Isadora Schulze, por sempre estarem presentes na minha vida, me apoiando.

Ao meu parceiro de vida, Arthur Carvalho, que me ajudou de todas as formas possíveis para que eu pudesse finalizar essa etapa, agradeço pelo suporte, amor e cuidado. Obrigada por ter feito companhia nas noites sem dormir, enquanto eu estava trabalhando e por me incentivar e apoiar todos os meus projetos.

Aos membros da banca, que aceitaram o convite para avaliar e contribuir com meu projeto.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de iniciação científica e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Por fim, agradeço a todos que contribuíram direta ou indiretamente para a conclusão deste trabalho.

*“Quando meus olhos estão sujos da civilização, cresce
por dentro deles um desejo de árvores e pássaros”*

Manoel de Barros

RESUMO

Delimitar os limites das espécies e compreender a distribuição de unidades evolutivas independentes dentro de um complexo de espécies pode lançar luz sobre a história evolutiva de tal grupo e ajudar a desvendar a história da paisagem. Com isso em mente, estudamos a estrutura genética dos corupções Neotropicais, um complexo de espécies formado por três espécies alopátricas que habitam principalmente as terras baixas secas dos Neotrópicos. Usamos o sequenciamento de DNA de Elementos Ultraconservados (UCE) para amostrar 21 indivíduos dentro do complexo em sua gama. Com base em ~ 1.500 loci e ~ 1.000 SNPS independentes, identificamos quatro clusters genômicos morfologicamente distintos, incluindo i) *Icterus icterus*, das planícies do W colombiano / venezuelano; ii) *Icterus croconotus*, das planícies secas do oeste da América do Sul e parte da Amazônia W e C; iii) *Icterus jamacaii*, do NE do Brasil; e iv) uma população isolada de *I. c. croconotus*, das savanas Roraima-Rupununi. Nossos resultados fornecem evidências da existência de uma linhagem não descrita e indicam introgressão genética entre *I. I. jamacaii* e *I. c. croconotus* no C Brasil. Além disso, a modelagem de distribuição de espécies paleoclimáticas de *I. c. croconotus* sugere que os fragmentos disjuntos de floresta seca estavam mais conectados no passado recente, durante o Último Máximo Glacial, derivado de um processo de contração e retração das florestas durante o período Quaternário.

Palavras-chave: História Evolutiva. Genômica. UCE. Modelagem.

ABSTRACT

Delimiting species boundaries and understanding the distribution of independent evolutionary units within a species complex can shed light into the evolutionary history of such group and help unravel the history of the landscape. With that in mind, we studied the genetic structure of the Neotropical troupials, a species complex formed by three allopatric species that inhabit the dry Neotropical lowlands. We used DNA sequencing of Ultra-conserved elements (UCE) to sample 21 individuals within the complex across its range. Based on 1,500 loci, and ~1,000 independent SNPS, we identified four morphologically distinct genomic clusters, including i) *Icterus icterus*, from the W Colombian/Venezuelan lowlands; ii) *Icterus I. c. croconotus*, from western South American dry lowlands and part of W Amazonia; iii) *Icterus I. jamacaii*, from NE Brazil; and iv) an isolated population of *I. c. croconotus*, from the Roraima-Rupununi savannas. Our results provide evidence for the existence of an undescribed lineage, and indicate genetic introgression between *I. I. jamacaii* and *I. c. croconotus* in C Brazil. In addition, the paleoclimate species distribution modelling of *Icterus I. c. croconotus* suggests that the disjunct fragments of dry forest were more connected in the recent past, during the middle LGM. Furthermore, the models predicts greater habitat suitability in the Amazon region in the present-day models, indicating that the suitability in Amazonia lowlands is a recent historical event. Thus, the divergence timing could be consequences of the environmental changes due the climatic oscillations in the Pleistocene.

Keywords: Evolutionary History. Genomics. UCE. Modelling.

SUMÁRIO

1 INTRODUÇÃO GERAL.....	8
2 EVOLUTIONARY HISTORY AND SPECIES LIMITS IN THE TROUPIAL COMPLEX (<i>Icterus icterus</i>).....	12
CONCLUSÃO.....	45
REFERÊNCIAS.....	46

INTRODUÇÃO GERAL

Os limites biogeográficos de uma espécie são formados através de processos históricos e ecológicos estabelecidos através de modificações climáticas, geomorfológicas e antrópicas, responsáveis pelas mudanças na biota associadas a áreas geográficas (Morrone, 1994) refletindo na complexidade em compreender o surgimento e diversificação da biota neotropical (McCormack et al, 2009). Eventos históricos tais como o soerguimento dos Andes, o estabelecimento do istmo do Panamá, e a aridificação das terras baixas do Caribe no norte da América do Sul, dentre outros eventos, resultaram em mudanças e reconfigurações na paisagem (Smith *et al.*, 2014). Estas mudanças resultaram no isolamento geográfico de numerosas populações animais, dando lugar ao processo de diversificação biológica através da especiação alopátrica. As condições biogeográficas atuais são reflexo da combinação de diferentes padrões espaciais, temporais e genéticos (Smith *et al.*, 2014). Portanto, para entender os padrões de distribuição atual é importante entender como os eventos ecológicos e históricos se combinaram para ocasioná-los (Cox & Moore, 2000).

Um dos primeiros modelos propostos para vincular os padrões atuais de diversidade com as flutuações climáticas históricas foi a hipótese dos Refúgios Pleistocênicos (Haffer 1969), que propôs que, embora as regiões temperadas tenham sido dominadas por glaciações, os trópicos sofreram mudanças climáticas menos severas, embora importantes. Especificamente, e de acordo com essa hipótese, o clima mais seco do Pleistoceno teria fragmentado a Amazônia em núcleos de floresta úmida (refúgios), cercados por vegetação aberta, favorecendo a especiação alopátrica pelo isolamento de aves da floresta tropical. Durante os períodos interglaciais (e mais úmidos), as florestas anteriormente fragmentadas teriam se reconectado, e as populações que concluíram o processo de especiação não mais se reconheceriam como parte da mesma espécie. Um modelo semelhante foi apresentado um ano depois, usando lagartos como organismos modelo por Vanzolini & Williams (1970). Esse mecanismo de especiação foi bem aceito pela comunidade científico e dominou os estudos evolutivos por quase duas décadas, sendo até mesmo aplicado para explicar eventos de diversificação além dos Neotrópicos, incluindo roedores africanos (Nicolas et al. 2010) e raposas vermelhas na América do Norte (Aubry et al. 2009). Nos Neotrópicos, no entanto, essa hipótese perdeu fôlego no início do século XXI, por apresentar poucas evidências das mudanças climáticas sugeridas, além de aparecerem inconsistências temporais nos processos de especiação sugeridos por esta hipótese (Costa et al. 2018). Embora seja bastante claro que houve flutuações climáticas durante o Pleistoceno, há poucas evidências de que a floresta amazônica realmente tenha se contraído aos níveis sugeridos pela hipótese dos Refúgios (Smith et al. 2014). Além disso, linhas de evidência independentes indicam que grande parte da fauna de vertebrados terrestres existentes, provavelmente originaram-se antes do Pleistoceno, deixando claro que a especiação nos trópicos não pode ser atribuída apenas aos refúgios, pois outros fatores anteriores ao Quaternário também devem ter agido (Rull 2008, 2011; Silva et al. 2019).

Estudos filogeográficos recentes incorporaram ferramentas espaciais, como modelagem de nicho ecológico (ENM), para lançar luz sobre o papel da ecologia e adaptação nos processos de especiação (Carnaval e Moritz, 2008; Naka e Pil, 2020). Essas abordagens espaciais podem iluminar os estudos evolutivos, fornecendo uma avaliação objetiva do nicho ecológico de diferentes linhagens e comparando entre os cenários climáticos atuais e passados, o que fornece hipóteses testáveis sobre como as flutuações climáticas históricas que podem ter influenciado a dinâmica de distribuição de linhagens divergentes e seus genes (Savit & Bates, 2015).

As Florestas Tropicais Sazonalmente Secas (FTSS) representam um dos biomas mais ameaçados dos Neotrópicos e também um dos menos conhecidos, tanto do ponto de vista da sua biodiversidade, como dos processos evolutivos responsáveis pelos atuais padrões de distribuição das espécies adaptadas a estas florestas (Pennington *et al.*, 2009). As FTSS se caracterizam pelo predomínio de uma vegetação decídua durante a estação seca e forte sazonalidade climática (precipitações <1600 mm/ano), com um período de pelo menos 5-6 meses com menos de 100 mm de chuvas (Pennington *et al.*, 2009). Estas florestas se distribuem por quase toda a região Neotropical, desde o México até a Argentina, de forma descontínua, formando 11 núcleos de FTSS, incluindo na América do Sul a Caatinga no nordeste brasileiro, o Chaco, alguns vales Andinos no Peru, e regiões secas na Colômbia e Venezuela (Pennington *et al.*, 2000).

Nos anos 90, Prado & Gibbs (1993) propuseram a “Hipótese do Arco Pleistocênico” para explicar este intrigante padrão de distribuição disjunta. Este modelo postula que a atual distribuição disjunta das FTSS na região Neotropical seja o resultado de mudanças climáticas que fragmentaram o que teria sido uma região contínua de florestas secas. Segundo estes autores, o evento ou série de eventos vicariante(s) separando os diferentes núcleos florestais teria ocorrido no Pleistoceno, principalmente nos últimos 21.000 anos, desde o último máximo glacial (LGM) (Caetano *et al.*, 2008). As evidências de uma conexão ancestral se baseiam na distribuição disjunta de várias espécies de plantas e animais que ocorrem em alguns destes núcleos (Pennington *et al.*, 2000;

Werneck & Colli, 2006). Evidências com plantas vasculares sugerem que alguns destes núcleos tenham sido conectados no passado, porém sustentam uma separação anterior ao LGM (Caetano *et al.*, 2008). Estudos com aves ainda são insipientes e pouco tem se avançado desde Prado (1991). Entretanto, resultados preliminares sugerem falta de padrões comuns no tempo de diversificação das diferentes espécies avaliadas.

Desta forma, acredita-se que estudos evolutivos de espécies distribuídas em diferentes núcleos de FTSS possam ajudar a elucidar a história evolutiva do bioma através da interpretação das relações filogeográficas entre as populações, e ao mesmo tempo, avaliar a hipótese do Arco Pleistocênico proposta por meio de possíveis eventos vicariantes recentes (Prado & Gibbs, 1993). Embora o isolamento geográfico seja aceito como o principal motor da diversificação biológica, o padrão temporal de diversificação e a demografia histórica de múltiplas espécies podem oferecer novos e complexos cenários de evolução espaço-temporal (Rull, 2013).

Até a primeira década do Século XXI, estudos filogeográficos em aves neotropicais têm se baseado principalmente no DNA mitocondrial (mtDNA) como o principal marcador molecular. No entanto, novas tecnologias de sequenciamento de alto rendimento (também conhecidas como sequenciamento de próxima geração, NGS), permitiram a coleta de centenas ou milhares de loci, oferecendo maior eficiência na resolução de relações de demografia histórica, estrutura populacional e parentesco em um alto grau de refinamento, particularmente para organismos não modelos. Um dos marcadores genômicos mais usados atualmente envolve a captura sequencial de elementos ultraconservados (UCEs), que são trechos curtos de DNA presentes próximos de regiões altamente conservadas, encontrados em mamíferos, pássaros e peixes (Faircloth *et al.*, 2012). Devido à sua universalidade, essa região do DNA tem sido incorporada como um marcador em estudos filogenéticos. UCEs permitem acessar regiões flangeadoras do genoma com variação suficiente para estudos filogenéticos e divergências em diferentes níveis taxonômicos, de famílias a populações (Harvey *et al.*, 2016).

As espécies representam as unidades básicas dos estudos macroevolucionários e a delimitação de seus limites é um passo crucial para caracterizar a biodiversidade e compreender os padrões e processos por trás de suas origens (Cracraft, 1985; Wiens, 2007). Delimitar os limites das espécies e compreender a distribuição de unidades evolutivas independentes dentro de um complexo de espécies pode ajudar a esclarecer a história evolutiva de tal grupo e ajudar a desvendar a história da paisagem. A delimitação dos limites das espécies, no entanto, pode ser um processo ambíguo e requer o uso de diferentes abordagens. Embora o estabelecimento de limites de espécies possa ser difícil, uma série de métodos foram desenvolvidos para aumentar a capacidade de detectar linhagens recentemente separadas e identificar atributos emergentes de linhagens, como fenotípicos distintos, características comportamentais e nicho ecológico (Isler *al.* 1998, Wiens e Graham 2005, McCormack *et al.* 2010). A integração de

múltiplas abordagens para delimitação de espécies a partir de análises de dados morfológicos, vocais, ecológicos, biogeográficos e genéticos está ajudando a superar essas dificuldades. (Cadena e Cuervo 2010).

Os corrupeiros neotropicais fazem parte do complexo *Icterus icterus*, sendo composto por três espécies pertencentes à família Icteridae (Ordem Passeriformes). As três espécies eram frequentemente consideradas como parte de uma espécie politípica variável (Ridgely & Tudor, 1989). No entanto, estudos recentes admitem a separação de *Icterus icterus* em três espécies assumindo diferenças na plumagem. Análises filogenéticas indicam que as três espécies representam um clado monofilético (Omland *et al.*, 1999; Powell *et al.*, 2014) composto por: *Icterus icterus* (Linnaeus, 1766), politípico com três subespécies reconhecidas (*ridgwayi*, *icterus*, e *metae*), *I. c. croconotus* (Wagler, 1829), politípico com duas subespécies reconhecidas (*I. i. strictifrons* e *I. c. croconotus*) e *I. I. jamacaii* (Gmelin, 1788), monotípico endêmico do semiárido brasileiro.

A distribuição disjunta destas três linhagens oferece uma oportunidade única para o estudo das FTSS, pois juntas, elas ocupam praticamente todos os núcleos de florestas secas da América do Sul, representando um bom modelo para avaliar a hipótese do Arco Pleistocênico. Assim, a utilização de recursos da biologia molecular atrelados a ferramentas de modelagem de nicho e bioacústica oferecem subsídios para elucidar questões que envolvem os processos evolutivos e ecológicos em populações disjuntas ao ampliar a visualização da evolução de traços ecológicos sob um contexto filogenético (Warren *et al.*, 2010).

Neste estudo, nosso objetivo foi entender os processos históricos e evolutivos que moldaram o atual padrão de distribuição das espécies de corrupeiro do complexo *Icterus icterus*, usando dados espaciais para entender a evolução do nicho, dados moleculares para descrever a estruturação genética atual, e dados morfológicos para avaliar os padrões de variação fenotípica e avaliar se a atual taxonomia, incluindo o tratamento das subespécies, refletem a variação observada no complexo.

Evolutionary History and Species Limits In The Troupial Complex (*Icterus icterus*)

Layse Albuquerque da Silva Lucas^{1*}, Maria W. Pil¹, Leilton W. Luna², Luciano N. Naka¹

¹Laboratory of Avian Ecology & Evolution, Department of Zoology, Federal University of Pernambuco, Recife, Pernambuco, Brazil

³Laboratório de Genética e Conservação, Instituto de Estudos Costeiros, Universidade Federal do Pará, Bragança, PA, Brazil.

* Corresponding author: Layse Albuquerque da Silva Lucas (laysealbuquerque00@gmail.com)

Abstract

Delimiting species boundaries and understanding the distribution of independent evolutionary units within a species complex can shed light into the evolutionary history of such groups and help unravel the history of the landscape. With that in mind, we studied the genetic structure of the Neotropical troupials, a species complex formed by three allopatric species that inhabit the dry Neotropical lowlands. We used DNA sequencing of Ultraconserved elements (UCE) to sample 21 individuals within the complex across its range. Based on 1,500 loci, and ~1,000 independent SNPS, we identified four morphologically distinct genomic clusters, including i) *Icterus icterus*, from the W Colombian/Venezuelan lowlands; ii) *Icterus I. c. croconotus*, from western South American dry lowlands and part of W Amazonia; iii) *Icterus I. jamacaii*, from NE Brazil; and iv) an isolated population of *I. c. croconotus*, from the Roraima-Rupununi savannas. Our results provide evidence for the existence of an undescribed lineage, and indicate genetic introgression between *I. I. jamacaii* and *I. c. croconotus* in C Brazil. In addition, the paleoclimate species distribution modelling of *Icterus I. c. croconotus* suggests that the disjunct fragments of dry forest were more connected in the recent past, during the middle LGM. Furthermore, the models predict greater habitat suitability in the Amazon region in the present-day models, indicating that the suitability in Amazonia lowlands is a recent historical event. Thus, the divergence timing could be consequences of the environmental changes due the climatic oscillations in the Pleistocene.

Keywords: Evolutionary History, Genomics, UCE, modelling.

Introduction

The identification of the mechanisms shaping the diversification of lineages and species complexes is essential in understanding speciation processes (Werneck, 2012). For decades,

phylogeographic and biogeographical inferences have been widely used to test hypotheses of species diversification in the Neotropics (Ribas et al. 2012; Werneck et al. 2015). Such studies often link the evolution of biological lineages to past landscape dynamics and climate changes. One of the first models proposed to link current diversity patterns with past climatic fluctuations was the Pleistocene Refugia Hypothesis (Haffer 1969), which proposed that while temperate regions were dominated by glaciations, the tropics endured less severe, yet pervasive climatic changes. Specifically, and according to this hypothesis, the dryer climate of the Pleistocene fragmented the Amazon in patches of humid forest (refugia), surrounded by open vegetation, favoring allopatric speciation by isolating rainforest species. During interglacial (and more humid) periods, previously fragmented forests would have reconnected, and those populations that concluded the speciation process would no longer recognize one another as being part of the same species.

While Pleistocene refugia have been invoked to account for species diversification in the humid Neotropics, dryer areas were allegedly similarly susceptible to climatic fluctuations. Specifically, it has been suggested that while humid forests contracted, dry forests expanded, and that the current disjunct distribution of Seasonally Dry Tropical Forests represents a relic of a once continuous dry forest. This hypothesis, known as the “Pleistocene Arc”, states that currently fragmented dry forest populations were connected during glacial (dry) periods and disconnected during interglacial (more humid) periods (Prado & Gibbs 1993). According to these authors, such a vicariant event, or series of events, allegedly occurred during the last glacial maximum (LGM), only 21,000 years ago. An alternative scenario is one where these dry forest nuclei have remained isolated throughout the Pleistocene, and biogeographical similarities in the fauna and flora are the result of either much older connections or dispersal events (Mayle, 2004; 2006). Support for the Arc hypothesis has been controversial. Support for the Pleistocene Arc was offered by studies on dry specialist woody plants and birds, providing evidence of population size expansions during the LGM (Caetano et al., 2008; Collevatti et al., 2012), and very recent divergences in populations of the Rufous-fronted Thornbird in the Caatinga and the Chaco (Corbett et al. 2020). On the other hand, studies with different taxa such as lizard (Werneck et al. 2012), plants (Caetano et al., 2008; Collevatti et al., 2012), birds (Savit & Bates, 2015), and flies (Moraes et al., 2009) have suggested that divergences were much older than the LGM, and as early as the Middle and Late Pleistocene. Population contraction in *Tabebuia rosealba* during the LGM, also contradicts one of the tenets of the Pleistocene Arc hypothesis (de Melo et al., 2016). Despite controversies on the timing of separation, studies suggest that the entire Pleistocene may have acted as a period of range retractions and expansions of disjunct patches of dry forests (Collevatti et al., 2012; Savit & Bates, 2015).

However, for this kind of analysis to be comparable, it is important to use similar taxonomic units (Rull, 2011). These units should represent evolutionary taxa, defined as the smallest population or set of populations diagnosed as a taxon (Cracraft et al. 2020). For birds, these taxa are often represented by well-defined allopatric populations that are represented by monotypic species or subspecies within a polytypic species. Therefore, defining species limits in a clade is a vital step towards understanding the evolutionary history of a lineage (Bortolus 2008). To explain the distribution of species under evolutionary and geological events requires the understanding that the rise of lineages derived from a temporally extended process. Several evolutionary processes may render populations to become reciprocally monophyletic, reproductively isolated, ecologically divergent or morphologically distinctive. However, these sets of arrangements serve as evidence to delimit species, which can arise at different times and in different orders during the process of divergence. The delimitation of species boundaries, however, can be a difficult task, mostly due to the continuous nature of the speciation process, and the categorical nature of taxon classification (Wake, 2006). This is particularly true for recently formed species, or those that behave as independent evolutionary lineages despite ongoing gene flow. To cope with the difficulties in determining species limits, a number of methods have been developed to increase the detection of recently separated lineages and to identify emergent lineage attributes, such as distinct phenotypic and behavioral traits, and ecological niche (Isler al. 1998, Wiens and Graham 2005, McCormack et al. 2010). The integration of multiple approaches for species delimitation, such as analyses of morphological, ecological, biogeographic and genetic data is helping overcome these difficulties (Remsen 2005, Cadena and Cuervo 2010).

While the Pleistocene Refugia or the Pleistocene Arc could be invoked to account for the diversification of humid terra-firme forests or seasonally dry forests species, respectively, understanding the role of climatic fluctuations and landscape dynamics on the diversification process of lineages that are not restricted to single biomes is far more challenging. This is the case for several species that occupy not only dry forests but also neighboring biomes, such as the Cerrado, the Pampas, the Andes, or even the edges of the Amazon (Bolívar-Leguizamón et al. 2020). In fact, many species present circum-Amazonian distribution (*sensu* Remsen et al., 1991), occupying several open biomes. Such distributions are common among birds, insects and plants (Prado and Gibbs, 1993; Bates, 1997; Savit and Bates, 2015; Banda et al., 2016). The challenges related to account for the evolutionary history of these species is reflected by the lack of knowledge about the historical events and dynamics behind such distributional patterns in the Neotropics (Werneck et al. 2011). Even more challenging are those species that

not only occur in circum-Amazonian biomes, but also inhabit part of the Amazon basin. Whether these species represent long-term cases of plastic adaptation or recent range expansions events, remains to be investigated.

The use of historical climatic conditions to advance alternative evolutionary hypotheses that can be tested using molecular data represents a powerful tool for statistical phylogeography (Richards et al. 2007; Collevatti et al. 2013; Metcalf et al. 2014). Several historical demographic and phylogenetic changes, such as population expansions, bottlenecks, and time in isolation, are expected following certain changes in the landscape, such as biome expansions, retractions, or habitat fragmentation. This approach can be particularly useful for those lineages with wide distributions that include several adjacent biomes.

One such case is that of the troupials (*Icterus icterus* species complex), a Neotropical avian clade of passerines in the blackbird family (Passeriformes: Icteridae) that includes three currently recognized species level taxa, and five subspecies (Fraga, 2007). The distribution of this species complex includes some of the largest patches of seasonally dry tropical forests, including the Caatinga of NE Brazil (*I. I. jamacaii*), the Pantanal of C Brazil and the Chaco of N Argentina, extending through the dry forests of Bolivia and Perú (*I. c. croconotus* *I. i. I. i. strictifrons*), penetrating the Amazon along floodplain forests and savannas (*I. c. I. c. croconotus*), and the dry forests and *llanos* of Coastal and interior Venezuela and Colombia (including the nominate form and the subspecies *metae* and *ridgwayi*). Despite having been described as three species-level taxa (Linnaeus, 1766; Gmelin, 1788; Wagler, 1829), this complex has historically received three different taxonomic arrangements. In its seminal work of the Birds of Peru, Hellmayr (1937) recognized two species *I. c. croconotus* and *I. I. jamacaii*. Half a century later, Short (1975) lumped the whole complex into a single polytypic species. This arrangement was reversed only recently by Jaramillo and Burke (1999), who treated the complex as three recognized species with their current subspecies, which remains as the most accepted treatment to date (Fraga, 2007; BirdLife International, 2020). Differences in plumage, nesting behavior, and disjunct distribution were considered as evidence that these taxa should be treated as at least three different species in a well-supported monophyletic clade (Omeland et al., 1999)..

In this study, we take an integrative approach to understand the historical and evolutionary processes that have shaped current distribution patterns of the troupials. We studied the broad-scale phylogeographic history of the *Icterus icterus* species complex using environmental, morphological, and molecular data to describe the current phenotypic and genetic variation of this species complex across its distributional range. Specifically, we used this integrative approach to i) assess whether current taxonomy, including three species and

five subspecies, reflects the phenotypic and genotypic variation observed in the complex; ii) investigate patterns of genetic structuring within *I. I. jamacaii* and *I. croconotus*, the species for which we had adequate sampling; and iii) Evaluate levels of niche divergence among current taxonomic groups and conduct paleoclimatic reconstruction to evaluate the relative contributions of the Pleistocene climatic fluctuations on the diversification of the complex.

Material and Methods

Study group

The troupials are monomorphic, with both sexes displaying elaborate black, white and carotenoid plumage, predominantly monogamous with biparental care (Lowther, 1975). The three species of troupials represent a well-supported monophyletic clade of birds that occupy a large area of the South American lowlands, occupying at least three different biomes, including the seasonally dry forests of the Caatinga, the Chaco, and the dry valleys of Bolivia and Peru, the Pantanal and Savannas (Fig. 1). All troupials are phenotypically similar, with a very distinctive orange and black plumage pattern, with yellow eyes surrounded by naked skin (Jaramillo & Burke, 1999). Despite those superficial similarities, each species can be easily diagnosed based on plumage. A large white wing patch made up of a white shoulder and white secondaries and tertials characterizes *Icterus icterus*. However, the black back and entire black hood differentiates it from *Icterus I. c. croconotus*, which is mostly orange with a black crown. *Icterus I. jamacaii* is relatively similar to *Icterus icterus*, but has orange greater coverts rather than white and a very reduced eye patch. *Icterus icterus* (Linnaeus, 1766) is a polytypic species with three recognized subspecies (*icterus*, *ridgwayi* and *metae*), occurring from northeastern Colombia and Venezuela, including the islands of Aruba, Curaçao, and Margarita, and a feral introduced population in Puerto Rico (Voous, 1982) (Fig. 1). *Icterus I. c. croconotus* (Wagler, 1829) also represents a polytypic species, with two recognized subspecies (*I. i. strictifrons* and *I. c. croconotus*). The form *I. c. croconotus* occurs in the dry forests of southern Colombia and Ecuador, with a tongue of distribution extending through the Amazon River, and an isolated population in the Roraima-Rupununi savannas in the northernmost Brazil and adjacent Guiana (Fig. 1). Whereas the form *I. i. strictifrons*, is continuously distributed in the Gran Chaco of Bolivia, Paraguay and northeastern Argentina, including western Mato Grosso. Lastly, *I. I. jamacaii* (Gmelin, 1788), is a monotypic species restricted to northeastern and eastern Brazil (Ridgely & Tudor, 1989) (Fig. 1).

Genetic Data

We obtained genetic data from 21 specimens within the *Icterus icterus* complex, with samples of all three species: *Icterus I. jamacaii* (n=8), *Icterus I. c. croconotus* (n=12), and *Icterus icterus* (n=2) (Table S1). We extracted genomic DNA from tissues using Qiagen DNeasy Blood & Tissue Kits (Qiagen Inc., Valencia, CA, USA) and sent extractions to RAPiD Genomics (Gainesville, FL, USA) for UCE sequence capture and sequencing following an open-source protocol (<http://ultraconserved.org/>; Faircloth and Glenn, 2012). The pool of samples was enriched using a set of 5,060 loci MY_Baits_Tetrapods-UCE-5k-probes. We sequenced the pool after quantification using 150 pb paired-end Illumina HiSeq.

Data assembly and SNP calling

After sequencing, we used Illumiprocessor (Faircloth, 2016, <https://github.com/faircloth-lab/illumiprocessor>) and Trimmomatic (Bolger et al., 2012) to trim reads, discard adapter contamination and low-quality bases. We then assembled reads and performed final loci alignment using Phyluce 1.4 (Faircloth, 2016). We selected the largest contig of each locus among individuals as a reference which was used to map raw reads from individuals in BWA, using the default number of four mismatches among reads (Li, 2013). Mapped reads to references were then used for SNP calling using the GATK's UnifiedGenotyper and VariantAnnotator (McKenna et al., 2010) keeping only SNPs with a minimum phred-scaled confidence of 30. Then, a vcf file containing all SNPs was obtained. Variant per individual were then incorporated to the reference sequences with the seqcap_pop package (Harvey et al., 2016) in order to obtain full sequences for each locus.

Phylogenetic Analysis and species-tree

We evaluated the phylogenetic relationships within the *Icterus* species complex using both gene and species trees. Concatenated data were created with an incomplete matrix with ~1,500 loci of UCE, including one individual of *Agelaioides badius*, used as the outgroup. After generating completeness matrix using 70% of loci, the concatenated data matrix was analyzed with RAxML V8.0.19 (Stamatakis, 2014) using the GTR GAMMA evolutionary model looking for the best maximum likelihood tree score (ML), obtaining 1000 bootstrap searches reconciling with the best ML tree with the bootstrap replicates.

We constructed a species tree, which recovers the genealogy of taxa, individuals of a population, using SNAPP (Bryant et al, 2012) implemented in Beast2 (Bouckaert et al., 2014). Two independent MCMC chains were run for 50,000,000 iterations with default parameters

and a burnin of 500,000, sampling every 5,000, following Winger et al. (2015). Chain log files were compiled in LogCombiner v.2.4.4 (Bouckaert et al., 2014) and visualized in Tracer v1.7.1 (Rambaut et al., 2018) to check for convergence. Tree files from the independent runs were compiled with LogCombiner, a maximum clade credibility tree was generated in TreeAnnotator v.2.4.4 (Drummond & Rambaut, 2007), and the sampling of trees from the posterior distribution was visualized in DensiTree v.2.2.5 (Bouckaert, 2010).

Mitogenome assembly and annotation

The clean reads obtained from the phyluce 1.4 (Faircloth, 2016) pipeline containing nuclear and mitochondrial reads were blasted against a local database composed of mitochondrial genes of *Icterus* sp. (GenBank: AF099337.1, AF099336.1, AF089031.1, KF810934.1, AF099335.1, JQ175139.1, AF099336.1, AY211215.1, AY212245.1, KM896402.1, AF099297.1) and the complete mitogenome of *Neopsar nigerrimus* (GenBank: NC018794) (Table S1). The latter was used as the reference genome and used to map and align the hits to build a new reference using GENEIOUS PRIME 2020.2.4. Complete mitochondrial genomes were annotated manually and adjusted based on comparisons to the reference mitogenome. Sequences were aligned using MUSCLE with default parameters on GENEIOUS PRIME 2020.2.4.

Mitogenome Bayesian Phylogenetic Inference

For the mitochondrial phylogenetic analysis, we used an ingroup of 22 taxa, which includes a representative for each taxon (except for *I. i. icterus* and *I. i. metae*). For Bayesian phylogenetic inference of mitogenomes, we used all protein-coding genes, except for, ND6, ATP8 and ND4L (Table S2). The optimal partitioning scheme was selected using ModelTest (Kalyaanamoorthy et al., 2017) through the IQ-TREE web server under the Bayesian information criterion (BIC). We estimated time-calibrated trees using substitution rates on the basis of the formula provided by Nabholz et al. (2016), which accounts for rate differences associated with differences in avian body mass. The approximation of posterior tree distributions was obtained by Markov Chain Monte Carlo (MCMC) using Beast2 (Bouckaert et al., 2014). We ran 30,000,000 generations, sampling every 5000 generations. Run convergence and parameter posteriors were assessed using Tracer v1.7.1 (Rambaut et al., 2018), a maximum clade credibility tree was generated in TreeAnnotator v.2.6.3 (Drummond & Rambaut, 2007) with 20% of burnin.

Population genetics

To analyze genomic data, we used the sNMF analysis, which uses a multivariate method to infer whether there is population structure using SNPs (Frichot et al., 2014), implemented in the LEA package in R (Frichot & François, 2015). This analysis runs for K values ranging from 1 to the maximum number of samples. The analysis was performed including 100 runs for K value (ancestral populations). Minimum cross-entropy runs were used to establish the best K number, because the alpha parameter settings can play an important role in the impact of the results (Frichot et al., 2014). The number of discrete populations was also investigated using a Principal Component Discriminant Analysis (DAPC) with Adegnet v.2.0 (Jombart, 2008; Jombart & Ahmed, 2011). DAPC is a multivariate statistical approach that uses a priori defined populations to maximize the variance between populations in the sample, by partitioning them into inter-populations and within-populations components, maximizing group discrimination. A maximum of 20 clusters were investigated using the average K algorithm. The chosen number of clusters ($k = 4$) was evaluated using the Bayesian method to describe the clusters using the DAPC. We calculated the pairwise fixation indices (F_{st}) between all pairs of populations using the .vcf datafile.

Morphometrics

We analyzed 113 specimens belonging to all described taxa (except *metae*) within the *Icterus* species complex to study morphometric variation in the clade. We examined samples from the following three institutions: Louisiana State University (LSU), Museu Paraense Emílio Goeldi (MPEG), and the UFPE Ornithological Collection. We obtained standard morphological measurements (right wing, left wing, tail, tarsus, beak width and height) of all specimens. To compare species, measurement variables were subjected to an Analysis of Variance (ANOVA), followed by a posteriori Tukey tests. To assess whether phylogenetic clusters were supported by morphological characters, we used a Principal Component Analysis (PCA) and a Discriminant Function Analysis (DFA) using the PAST software.

Ecological Niche models

We generated paleoclimatic models exclusively for *I. c. croconotus*, the species that occurs both in open habitats and the Amazonian lowlands. To obtain locality points of each taxon within the species, we used online databases, which provide information on specimens deposited in zoological collections, photographic and acoustic records, complemented by sightings and birds captured in mist-nets. Specimen data were obtained from i) the Global Biodiversity Information Facility (www.gbif.org), ii) VertNet (www.vertnet.org), iii) Specieslink (<http://www.splink.cria.org.br>) and iv) the ornithological collection from the

Federal University of Pernambuco (UFPE). Localities were also obtained from digital images and audio recordings, including Wikiaves (www.wikiaves.com), Macaulay Library (www.macaulaylibrary.org), and Xenocanto, (www.xenocanto.org). Finally, we obtained data from the National Research Center for Conservation of Wild Birds (CEMAVE - ICMBio). Dubious records, repeated locations, missing or insufficient geographical coordinates were excluded from the analysis (Anderson & Gonzales, 2011; Boria et al., 2014). Finally, nearby locations (<10km) were excluded to remove spatial autocorrelation, leaving 690 unique sample localities used in the analysis. Spatial procedures were conducted using ArcGis 10.3.1 (Esri, 2011) and the SDMtoolbox 1.1c package (Brown, 2014).

Environmental Data

We obtained 19 standard bioclimatic variables to create the Environmental Niche Models of *Icterus I. c. croconotus* from WorldClim version 1.4, using a resolution of 2.5 arc minutes (5km) (Hijmans et al., 2005). We conducted a Spearman's correlation analysis among the 19 environmental variables to identify highly (>90%) correlated variables, which were excluded to avoid multicollinearity. In addition, the chosen variables were subjected to a Principal Component Analysis (PCA) to determine their contribution of the chosen variables in the multivariate space. Thus, among the variables with a correlation above 90%, it was decided to keep the ones with the greatest biological importance. The selected variables were: Average Daily Temperature Variation (Monthly average (Tmax-Tmin)) (BIO2), Seasonal Temperature (BIO4), Minimum Temperature of the Coldest Month (BIO6), Precipitation of the Driest Month (BIO14), Precipitation Seasonality (Coefficient of Variation) (BIO15), Precipitation of the Hottest Quarter (BIO18) (Table 9).

Climatic modelling

The climatic models were elaborated with the occurrence data and environmental variables using MAXENT version 3.2 (Phillips et al., 2006), which is based on the principle of maximum entropy. Processing was done in the R environment (R Core Team, 2016), using the Dismo package (Hijmans et al., 2016), which is widely used in ecological niche models because it presents more satisfactory results than other packages, not requiring absence data, even when used with a reduced number of species occurrence data (Wisz et al., 2008). For a temporal comparison, as of the climatic variables to the present-day, we generated models for the Middle-Holocene (~6,000yr) and the LGM (~21,000yr). To develop a niche model with a lower margin of error between the sources of past climatic variables, common climate models were chosen between three past scenarios: MPI-ESMP (ME), MIROC-ESM (MR) and CCSM4 (CC). Thus, we developed a consensus model between these three climate models. Overall, consensus was reached between the variables of each climatic model (ME, MR and CC). Then, a consensus model of the climatic scenarios was constructed for the LGM and the Middle-Holocene. For better precision of the models, 25 replicates were generated for each data set, whose replication methodology was Subsample, commonly used for models with

many locality occurrences. This type of replication repeatedly divides points of presence into random subsets, using 25% of them as test points. With that, a final model was generated with the average values of the probabilities of occurrence of the species obtained after the 25 repetitions projected for the variables of the present, average Holocene and LGM, resulting in three final models.

Niche background similarity test

The background similarity test asks whether ENM drawn from populations with partially or entirely non-overlapping distributions are any more different from one another than expected by chance alone under a specific null model. The test was proposed by Warren et al. (2010), who introduced a set of statistical tests that allow the quantitative comparison of MNE through the ENMtools program, which consists of a script with a graphical user interface, written with the Tk package (Warren et al, 2008, 2010).

Initially, the values of niche overlap were calculated from ENMs with the metric D of Schoener (Schoener, 1968) using ENMtools (Warren et al., 2008). To test the null hypothesis that niches are equally divergent compared to the environmental background, the background randomization procedure was used in ENMtools, which compares the observed niche overlap values with a null distribution of 100 overlap values generated by comparing the ENM of a taxon with ENM created from random points taken from the geographical range of the other taxon (Warren et al., 2008). As this process is repeated for both taxa in the comparison, two zero distributions are generated for analysis. How the background for each taxon is delimited affects the analysis, it includes regions that are more or less similar for the niche (Warren et al., 2008). Ideally, the background should include accessible habitats and should therefore reflect information on the dispersion capacity (Soberón and Peterson, 2005). Overall, 100 replicates of the overlap test were obtained. Thus, on a scale of 0 to 1, if the value of the niche overlap test is closer to 0 than the values of the background test, this indicates that the ENM would be more dissimilar than expected at random. On the other hand, when the value of the niche overlap test is closer to 1, it means that the MNG of the two species / populations are more similar than expected at random.

Results

All our results, including our Maximum Likelihood UCE concatenated tree, a time-calibrated Bayesian mitochondrial tree, our SNAPP species tree based on SNPs (Fig. 4), and our

population structure sNMF analysis (Fig. 2A) identified four main genetic clusters within the troupials, indicating that our results are consistent across markers and tree building methods. The final UCE matrix recovered a total length of 7,132,993 aligned base pairs of 1,496 loci. Among them, we identified 989 SNPs in the alignment of 21 samples of the complex taxa, using the three species of the complex. This matrix was used for the ML UCE tree, and these SNPs were used in SNAPP and all population structure analyses. Our final mitochondrial tree was based on 10 protein-coding genes. This tree presented large ESS (>200) and attained convergence for most parameters across the run for the mitochondrial Bayesian analysis. Divergence time estimates were bound by broad confidence intervals (95% highest posterior density, HPD) (Fig. 4C).

Phylogenetic Analysis

The concatenated UCE tree and the time-calibrated mitochondrial analysis yielded similar topologies, with four main genetic clusters which offer support for the current taxonomic treatment of at least three species in the *Icterus icterus* species complex (Fig. 4B, 4C). One clade is formed by *I. icterus*, which is sister to the clade including the two other clades, *I. jamaicaii* and *I. croconotus*. However, these trees also support the monophyly of a fourth lineage represented by an isolated population of *I. croconotus* from the Roraima-Rupununi savannas in N Brazil and adjacent Guiana (Fig. 4). The topology estimated by the UCE tree differs from the mitochondrial tree within the *I. croconotus* clade, in which the UCE tree shows lack of monophyly within *I. croconotus* (high bootstrap support, 97%). Once the Roraima-Rupununi clade is excluded, the complete monophyly of all current taxa are visible in the mitochondrial tree.

In our species tree (Fig. 4A), *I. jamaicaii* is sister to *I. croconotus* (including both forms, *I. c. strictifrons* and *I. c. croconotus*), whereas *I. icterus* clusters with the Roraima-Rupununi population. This topology is fully supported at all nodes. Both the SNAPP and the mitochondrial tree are congruent in placing *I. jamaicaii* as sister to *I. croconotus* (*I. c. croconotus* and *I. c. strictifrons*). Finally, the sNMF and discriminant analyses (DAPC) also supported the existence of the same four genetic clusters (k=4) mentioned above for the complex (Fig. 2).

Divergence time

According to our calibration, the Venezuelan clade, represented by two samples of *I. i. ridgwayi*, diverged from the rest of the group in the early Pleistocene (~2.0-3.0 Ma). The large clade composed by the populations from NE Brazil (*I. jamaicaii*), the western South American

dry lowlands and part of W Amazonia (*I. c. croconotus* and respective subspecies) and the Roraima-Rupununi savannas (isolated population of *I. c. croconotus*) is strongly supported ($P_p = 1.00$) and diverged the mid-Pleistocene (~ 0.9 - 1.48 Ma). Within this clade, *I. jamacaii* and *I. croconotus* are revealed as sister clades and diverged in the mid-Pleistocene (~ 0.7 - 1.08 Ma). The *I. croconotus* group, formed by two subspecies (*I. c. croconotus* and *I. i. strictifrons*) with a broad-range distribution, revealed two distinct clades, morphologically undistinguished, which diverged in the Late-Pleistocene (~ 0.5 - 0.7 Ma) (Fig. 4C, green).

Population structure

A co-ancestry analysis (sNMF) was generated separately for *I. jamacaii* and *I. c. croconotus*, to understand the ancestry coefficients within the groups. We recovered 4,749 loci and 1,081 SNPs for the 10 samples of *I. croconotus*, recovering two ancestral populations ($k=2$) (Fig. 3B upper) in the sNMF, which do not show evidence of ancestral mixing. However, we tested the $K=3$ and we recovered three ancestral populations (Fig. 3B lower). For *I. jamacaii*, we recovered 4,574 loci and 1,115 SNPs, recovering a single ancestral population ($k=1$) (Fig. 3A), indicating a lack of population structure within the Brazilian Semiarid. Our F_{st} analysis supported high population differentiation between the all pairs of populations (Table 7). We also calculated the F_{st} values within the *I. c. croconotus* cluster. The fixation indices values ($F_{st} = 0.1$) revealed a high differentiation between the two populations (*I. c. croconotus* and *I. i. strictifrons*).

Morphology

The troupials presented a relatively similar body shape, with high morphospace overlap among taxa, which was evident from the two PCA axes, which together accounted for $\sim 86\%$ of the variation (Fig. 5). In the discriminant analyses (LDA), the troupials can be divided in two groups, formed by *icterus* (including *icterus* and *ridgwayi*; we lack samples of *metae*) and the other by *I. jamacaii* and *I. c. croconotus* (Fig. 6). This grouping was mostly driven by differences in body size, which includes wing, bill, and tarsus (Fig. S1). The significant differences found in the complex (one way ANOVA, $p = 0.001$) were largely driver by *ridgwayi* (Tukey post-hoc test, Table 1-6), which occupied a distinct morphometric space in relation to all other taxa. Excluding this taxon, we found no significant differences among the remaining taxa. However, we observed significant differences between *I. jamacaii* and *I. c. croconotus* for wing variables (length of secondary and primary feathers). Despite that, this included the two

other species (*I. c. croconotus* and *I. jamacaii*) and the two forms within *I. c. croconotus*, and even the Roraima-Rupununi population (Table 1-6). Within *icterus*, we found a clear overlapping in the two forms sampled (*icterus* and *ridgwayi*).

Climatic modelling

The climatic model recovered for *I. c. croconotus* closely mirrors its current distribution (Fig. 8A) The AUC values were satisfactory (0.85), with relatively low omission rates (lack of prediction for areas with occupied locations), although over-represented in the Eastern Caatinga and the southern portion of the Atlantic forest, as well as in the Caribbean, and sub-represented in the open areas in the savannas of Roraima.

Paleoclimatic models were generated for *I. croconotus* for the Last Glacial Maximum (~21.000 years) and the Mid-Holocene (~6.000 years) (Fig. 8B, 8C). It predicts greater habitat suitability in the Amazon region in present-day models, suggesting that its current distribution may be the result of recent population expansions from W Amazonia.

The current areas of suitability are much larger than the areas found in both the Middle Holocene (~6,000 years ago) and the Last Glacial Maximum (~21,000 years ago). Compared to the present-day model, the Mid-Holocene model predicted greater habitat suitability in Central Brazil (East Mato Grosso, West Goiás) represented by the modern range of *I. c. strictifrons* (Fig. 8B). However, the LGM model predicted greater habitat suitability in Western Peru and Central Bolivia, represented by the range of *I. c. croconotus*. The LGM model also predicted greater suitability in Northern Colombia and Northwest Venezuela, which represents the current habitat of the *I. icterus* (Fig. 8C).

Niche background similarity test

We found evidence of niche divergence among all current taxonomic units, including *I. icterus*, *I. croconotus*, Roraima-Rupununi *icterus* and *I. I. jamacaii*. Background similarity tests between these lineages were significant ($P > 0.01$ Fig. 9). The test revealed that niche overlap among all lineages were significantly smaller than expected at random in the comparisons of all taxa, indicating niche divergence. However, the background test indices was higher between *I. jamacaii* and *I. croconotus* lineages ($D = 0.3$, Fig 9B). Indicating that niche differences between these lineages referred to environmental differences experienced in their respective geographical ranges.

Discussion

This study represents the first comprehensive large-scale phylogeographic study of the troupials, including samples from all species involved in the *Icterus icterus* species complex, from throughout its geographic range. Furthermore, by using genome-wide markers and the complete mitogenome, we were able to resolve important aspects of the evolutionary history of this group and make inferences on the processes behind current distribution patterns. Previous studies conducted aiming at uncovering the evolutionary history of this group were part of studies at broader phylogenetic scales, with very limited taxon sampling and few mitochondrial and nuclear loci. There are three major results of this study that are worth highlighting. First, we recovered four major lineages within the group, supporting the current treatment of three species, and unveiling what we believe represents a fourth phenotypically diagnosable undescribed species from the Roraima-Rupununi savannas. These results were consistent across markers (UCEs, mitogenome, SNPs) and tree building methods. Second, we found different levels of genetic structuring within the two species for which we had adequate sampling, *I. I. jamacaii* and *I. croconotus*. Whereas *I. I. jamacaii* showed virtually no genetic structure throughout the Caatinga, *I. croconotus* showed geographically structured populations, consistent with current subspecific classification and suggesting ecological adaptation across two different biomes. Finally, we explored the relation between the establishment of these lineages and the dynamics of the Neotropical landscape, suggesting potential and alternative scenarios that could explain current patterns of genetic structure, including evidence of introgression from the Brazilian semiarid into currently isolated Chaco populations, consistent with the Pleistocene Arc hypothesis. Therefore, we provide results that are congruent across population structure, phylogenetic relationships, and climatic history, and offer a good starting point towards the understanding of evolutionary history of this species.

Species limits and phylogenetic structure

The family Icteridae, which includes the orioles, blackbirds, and caciques, among others, has been subject to intensive systematic studies in the last two decades (Omland et al, 1999; Lovette et al, 2001; Allen and Omland, 2003; Sturge et al, 2009; Jacobsen et al, 2010; Jacobsen and Omland, 2011; Powell et al., 2014). The troupials, however, have been very poorly sampled, despite having a rather convoluted taxonomic history with splits and lumps. Although the close relationship of the three currently recognized species was never subjected to scrutiny, their monophyly as a group was confirmed only recently (Powell et al., 2014). These authors recovered a sister relationship between *I. icterus* and *I. c. croconotus*, which differed from our

results that found *I. c. croconotus* and *I. I. jamacaii* as sister species. Our results are fully supported by all analyses and were obtained from a much better sampling, a much larger number of loci, and the complete mitogenome, and we therefore believe we have fully resolved the phylogenetic relationship within the troupials, with the exception of *I. icterus*, which remains undersampled and requires a full study on its own. One of our most surprising results was the finding of a fourth lineage in this group, represented by the population of *I. croconotus* from the Roraima-Rupununi savannas, in northernmost Brazil and adjacent Guiana. The distinctiveness of this population was evident and fully supported in all phylogenetic and population structure analyses. In fact, according to our analyses, keeping the Roraima-Rupununi populations within *I. croconotus* would render this species paraphyletic. In other words, the remaining *I. croconotus* is more closely related to *I. jamacaii* (with whom it shows levels of genetic introgression), than to the Roraima-Rupununi populations. Although plumage patterns and color are highly labile within the orioles in the genus *Icterus*, they are highly conserved within clades (Omland and Lanyon, 2000). Thus, orioles have a diversity of plumage patterns and coloration that vary within similar themes (Howell & Webb, 1995). Despite this, the Roraima-Rupununi population is readily diagnosable by the extension of white on the wing and blue bare skin around the eye, differently from the black skin present in *I. croconotus*.

Population structure and genetic introgression

We investigated the intraspecific variation of the two species for which we had adequate sampling, namely the Caatinga endemic *I. I. jamacaii*, and the multi-biome inhabitant, *I. croconotus*. The two species showed very contrasting population structures. Whereas *I. I. jamacaii* presented no genetic structure throughout its distribution, suggesting widespread gene flow in the Brazilian semi-arid, *I. c. croconotus*, with its broad-scale range across different biomes, presented a very clear geographically structured pattern. Originally divided in two subspecies, *I. i. strictifrons* from the Bolivian dry forests, the Chaco and adjacent Brazilian Cerrado, and the nominate form from the flooded forests of W and C Amazonia, *I. c. croconotus* presented a very well-supported genetic divergence, consistent with these two forms, and high *Fst* (0.1), suggesting very limited gene flow. High *Fst* values within *I. croconotus* are consistent with the phylogenetic trees, including all markers and methods, and co-ancestry analysis, all of which revealed three major clusters within *I. croconotus*, including the Roraima-Rupununi savannas population and the two forms mentioned above.

The distributional range of *I. croconotus* is not common among Neotropical birds, and could reflect a disparate evolutionary history (Savit & Bates, 2015). Unfortunately, we lack adequate

phenotypic and genetic sampling along the potential contact zone of *I. c. croconotus* and *I. i. strictifrons* in Peru, but our morphological results found no significant differences between the two forms along this environmental gradient. Widely distributed species along latitudinal gradients, often show some levels of local adaptation in response to the environment, and that could be expected from the two forms (Chattopadhyay et al., 2017).

We found evidence of genetic introgression between *I. jamacaii* and the two forms of *I. croconotus*. Individuals of the nominate form from C Amazonia showed a mixed ancestry with *I. jamacaii*, and the same happened to samples of *I. c. strictifrons* from Paraguay and Bolivia. Although *I. croconotus* and *I. jamacaii* are considered allopatric, there is a known area of sympatry in the Brazilian state of Tocantins (Pacheco and Olmos, 2006). Also, both troupials have breeding feral populations around Belém, in the state of Pará (Cardoso da Silva and Oren, 1990), although no mixed pairs or intermediate specimens were reported. Whether these signs of introgression represent incomplete lineage sorting or signs of recent gene flow remains to be investigated.

Tempo and patterns of divergence

We calibrated the mitochondrial tree to investigate the tempo of diversification in this group. According to these analyses, the first split within this clade occurred between northern South American *I. icterus* and the Roraima-Rupununi savannas/*I. c. croconotus*/*I. jamacaii* ancestor, which was estimated to occur at the Pliocene/Pleistocene boundary, between 2.0 and 3.0 Mya. The population on the Roraima-Rupununi savannas apparently stopped sharing genes with the *I. c. croconotus*/*I. jamacaii* ancestor during the Mid-Pleistocene, between 0.9 and 1.48 Mya. This could explain the deep structure found in the co-ancestral analysis, not sharing signs of past introgression or secondary contact with *I. c. croconotus* taxa, revealing that the isolation in the Roraima-Rupununi Savanna is relatively old. The establishment of the current savannas landscape in the northernmost portion of the Brazilian Amazon is directly linked to tectonic events and to the past fluctuations of climate and erosion (Ruellan, 1957; Carneiro-Filho, 1991; Schaefer & Vale Jr, 1997). Throughout most of the Quaternary (~1.8-2.5 Ma) to the present-day, the formation of the current landscape of savannas has been influenced by relatively short climatic cycles (glacial and interglacial). The limits between the savannas and the humid forests likely varied as a result of the climatic fluctuations (Carneiro-Filho, 1993; Simões-Filho et al., 1997; Toledo, 2004). Therefore, it is likely that the savannas at the northern edge of the Brazilian Amazon, S Venezuela and W Guyana, can be considered as a “relict” landscape that appeared during the more recent dry periods of Pleistocene in the Amazon (Eden 1974;

Carneiro-Filho 1990). For that reason, it is possibly that the *Icterus icterus* complex ancestor, could have dispersed through the *Llanos* of Venezuela to the Savannas and the climatic changes of the glacial cycles could have separate the population in a relict in the Roraima-Rupununi Savanna, preventing gene flow.

Niche evolution and the Paeoclimatic modelling

Our *I. croconotus* paleoclimatic models predicted connected areas during the LGM in comparison to the present-day models, following a reduction and fragmentation of the suitable areas in the Mid-Holocene. Although the paleoclimatic model for the broad-range species supports the prediction of the Pleistocene Arc. Expansion of the paleoclimatic distributions during the LGM from dry forests were observed in studies with lizards (Machado et al, 2019) and mammals (Leite et al, 2016), predicted an expanded and retraction distribution during the LGM, becoming fragmented in present-day models. However, our paleoclimatic models do not corroborate with the period of divergence of the clades. Even so, Our paleoclimatic models could help to understand a possible expansion of the population of *I. c. croconotus* through the LGM to the present, enabling areas of past connection with *I. jamacaii* that could explain the introgression. Considering that the events that drive the evolutionary processes are ancient, and the further events of the expansion and contraction of habitat has occurred during the Quaternary changes, the niche divergence of the all taxa pair are congruent. However, whether niche divergence between lineages indeed corroborates ecological speciation will depend on the nature of their association with the environments available in their distributional ranges (Peterson 2011). For instance, if they selected habitats or dispersed to new environments, ecological speciation (sensu Endler 1977; Schluter 2000) would seem evident, despite its intrinsic differences to allopatric differentiation through dispersal (peripatric speciation, Mayr 1970).

Taxonomic implications and species limits

Finally, our genetic data is relevant to the question of species limits within *Icterus icterus* complex. We present data that support the treatment of not three, but four well-defined species, acceptable as such as most species concepts applied, including the Biological and the Phylogenetic Species Concepts. This means that based on the data presented, we believe that the Roraima-Rupununi population, formerly included within *I. croconotus*, represents a yet undescribed avian taxon. This species has been isolated from other troupial populations by at

least 0.9 million years, and has developed diagnostic phenotypic characters, such as the blue bare skin around the eye and the amount of white on the wing, in comparison to the Amazonian populations of *I. croconotus*. This isolated population shows a deep structure in all analyses, and was sister of the *I. jamacaii*/*I. croconotus* clade. We also presented evidence of niche divergence between the Roraima-Rupununi population and *I. c. croconotus*, suggesting a process of potential adaptation to the savannas environment.

The three currently accepted species (*I. icterus*, *I. c. croconotus*, and *I. I. jamacaii*) were initially described as species level taxa, and maintained as separate species well into the 1950s and 1960s (e.g. De Schauensee 1952, Pinto 1967). Since then, the three species were considered as one polytypic species under *Icterus icterus*, based on the argument of the existence of an intermediate population (*Icterus icterus metae*). The pattern of black and orange colors in the plumage of *metae* is basically the same as in *Icterus icterus*, but the nape is orange rather than black as in *Icterus I. c. croconotus*. However, *metae* is diagnosable by the divided white wing patch, unique to this population. Although having been described as three species-level taxa (Wagler, 1829; Linnaeus, 1766; Gmelin, 1788), this complex has historically received three different taxonomic arrangements. In its seminal work of the Birds of Peru, Hellmayr (1937) recognized two species *I. icterus* and *I. I. jamacaii*. He lumped *I. c. croconotus* into *I. jamacaii*, as a subspecies. Half a century later, Short (1975) lumped the whole complex into a single polytypic species. This arrangement was reversed only recently by Jaramillo and Burke (1999), treating the complex as three recognized species with their current subspecies, which remains as the most accepted treatment (Fraga, 2007; BirdLife International, 2020). Since then, previous studies detected a close relationship of *I. icterus* to *I. c. croconotus*, but *I. I. jamacaii* was the only oriole missing from previous phylogenies (Jacobsen et al, 2010; Omland, 1999). Powell (2014), recovered an equivalent pattern to those studies, including the missing oriole (*I. jamacaii*). A strict relationship of *I. icterus* to *I. croconotus* than to *I. jamacaii*, with a strong support only using ML method. These findings appeared contrary to the treatment of *I. c. croconotus* as a subspecies of *I. jamacaii* in classifications that recognizes two species of troupial. (Sibley and Monroe, 1990).

FIGURES

Figure 1. Geographic ranges of the troupials (*Icterus icterus* species complex) based on Jaramillo and Burke (1999), including images of the three species and sample localities used for the molecular analyses. The orange dot represents two samples from an introduced population of *I. i. ridgwayi* from Puerto Rico. The complex includes: 1. *Icterus I. jamaicaii* (A); 2. *Icterus icterus*, including the nominate form (B) *ridgwayi* (C) and *metae* (D); and 3. *Icterus I. c. croconotus*, including the nominate form (E) and *I. i. strictifrons* (F). Images by Birds of the World.

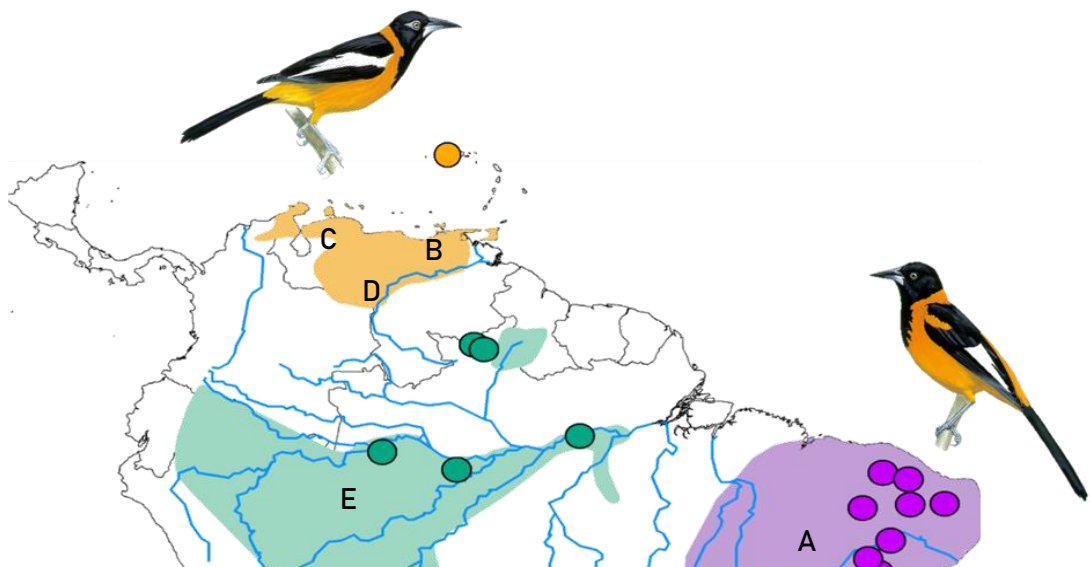


Figure 2. **A.** sNMF analysis of co-ancestry, indicating four distinct groups in the *I. icterus* species complex, and ancestral mixture between *I. jamaicae* and *I. c. croconotus*. **B.** Discriminant Analysis of Principal Components (DAPC), visualization of the number of clusters. **C.** Plot of sNMF in a map. Colors reflect the groups: orange for *icterus*; Pink for Roraima-Rupununi; Purple for *I. jamaicae*; Green for *I. c. croconotus*.

A)

C)

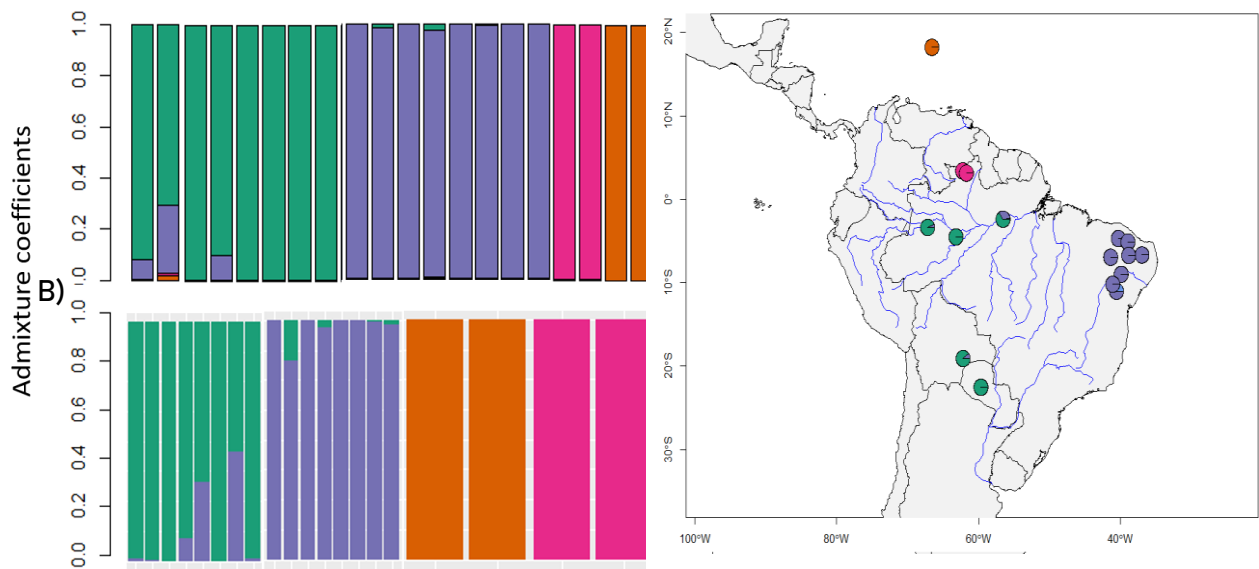
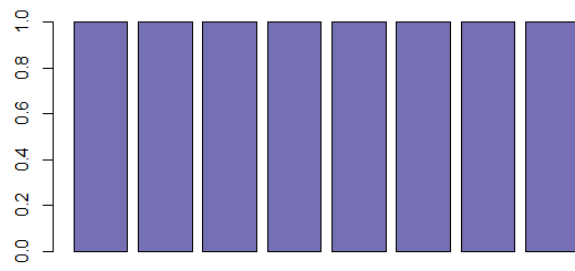
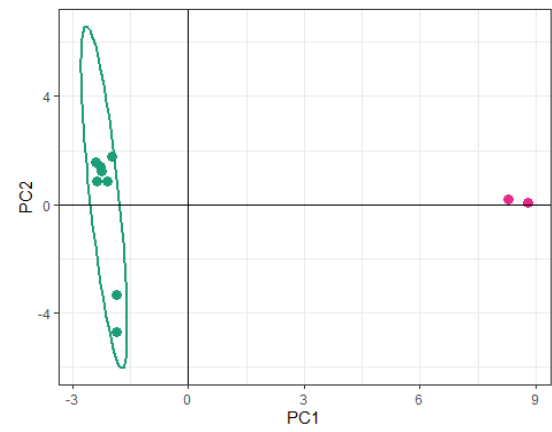


Figure 3. **A** Ancestral coefficients for *Icterus jamaicae* equal to 1 (K), indicates that the group is homogeneous,. **B.** Visualization of Ancestral coefficients for *Icterus I. c. croconotus*. Upper indicates k =2, two structured groups, indicating a scenario of differentiation between the two groups; Lower indicates k=3, revealing a third group within *I. c. croconotus*. **C** Analysis of principal components (PCA) of SNPs of *I. c. croconotus* (k=2 and k=3). Colors reflects the groups. Purple for *I. jamaicae*; Pink for Roraima-Rupununi; Dark green for *I. c. croconotus*; Light green for *I. c. strictifrons*.

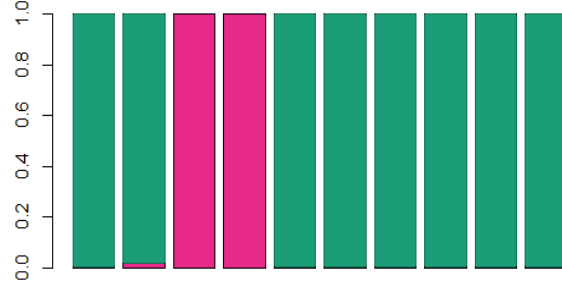
A)



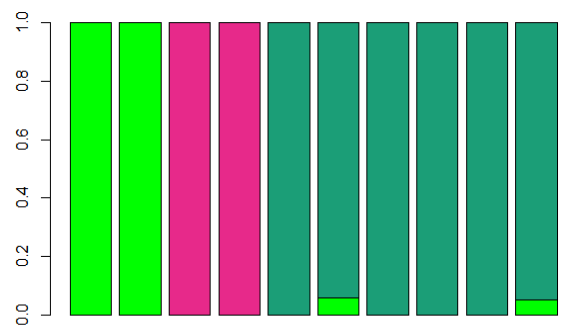
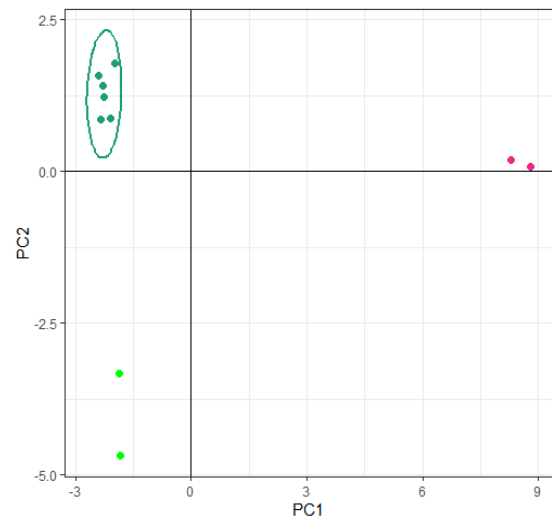
C)



B)



K=2



K=3

Figure 4. **A** Densitree representation of species trees from the posterior distribution of the SNAPP analysis. **B.** Concatenated tree of maximum likelihood RaxML based on 1,496 UCE loci, and 70 % of matrix completeness. **C.** Time-calibrated mitochondrial gene tree. Colors are consistent with those used throughout the paper: orange for *icterus*, pink for Roraima-Rupununi, purple for *I. jamacaii*, dark green for *I. c. I. c. croconotus*, and light green for *I. c. I. i. strictifrons*. Black is used for outgroups. The scale corresponds to millions of years. Nodes with >0.5 posterior support have bars indicating the 95% confidence interval on that node age.

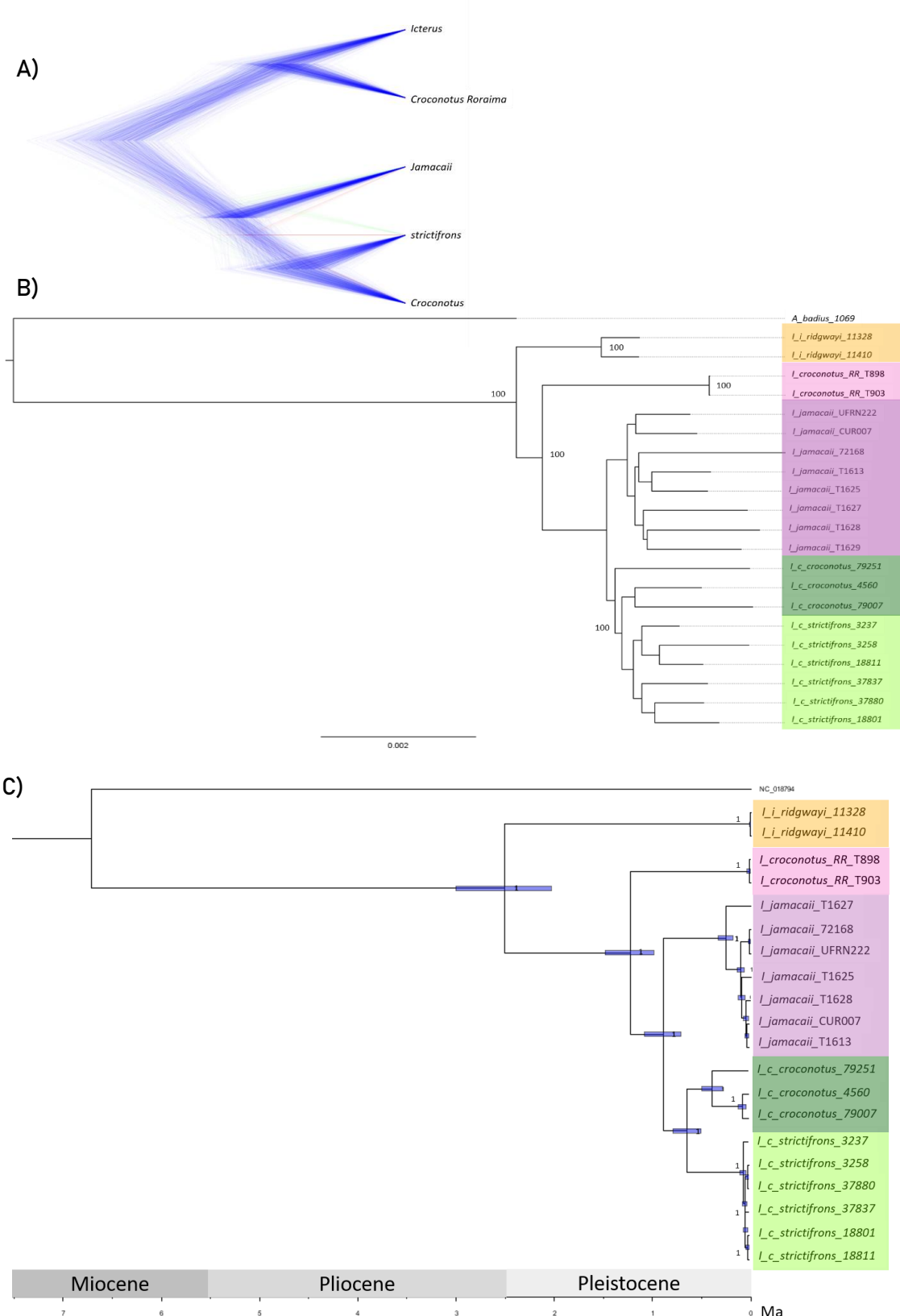


Figure 5. Principal Component Analysis (PCA) based on six morphometric variables of five taxa within the *Icterus icterus* species complex. PC loadings explained 87% of the observed morphometric variation (PC1 70% and PC2 17 %). The two yellow dots represent measurements from the only two specimens available from the Roraima-Rupununi savannas. Light green, *I. c. I. i. strictifrons*; Dark green, *I. c. I. c. croconotus*; Purple, *I. I. jamacaii*; Red, *I. i. icterus*; Orange, *I. i. Ridgwayi*.

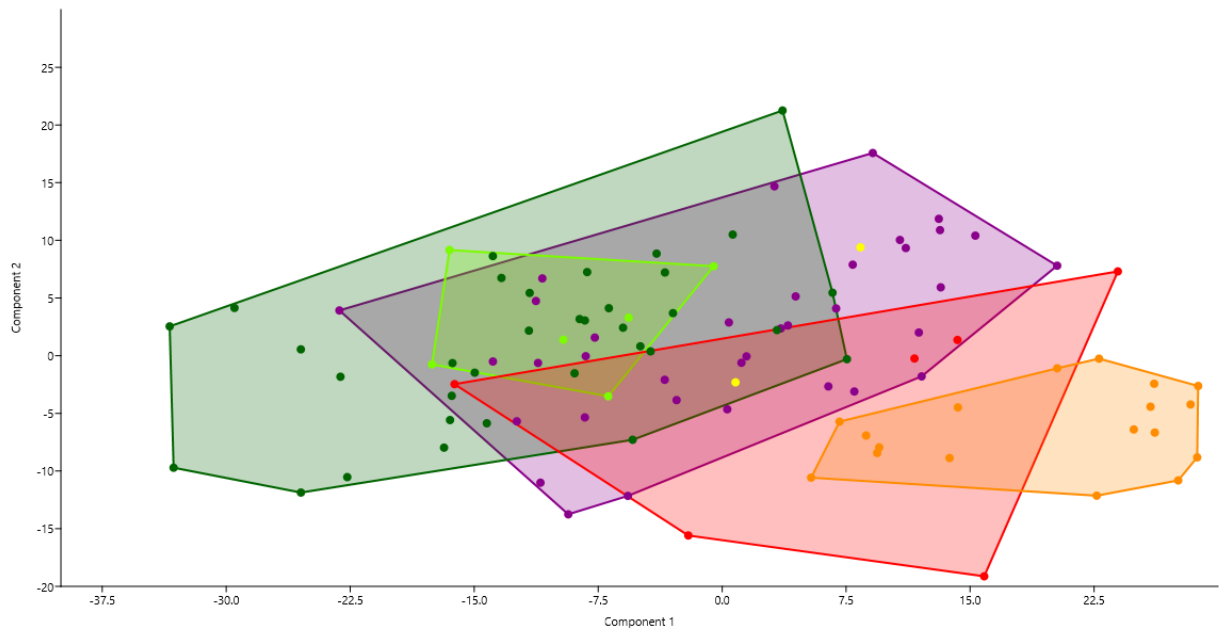


Figure 6. Discriminant Analysis (LDA) of all taxa within the *Icterus icterus* complex (except *Icterus. i. metae*). The dots in yellow represent the specimens in the savannas of Roraima. Light green, *I. c. I. i. strictifrons*; Dark green, *I. c. I. c. croconotus*; Purple, *I. I. jamacaii*; Red, *I. i. icterus*; Orange, *I. i. ridgwayi*.

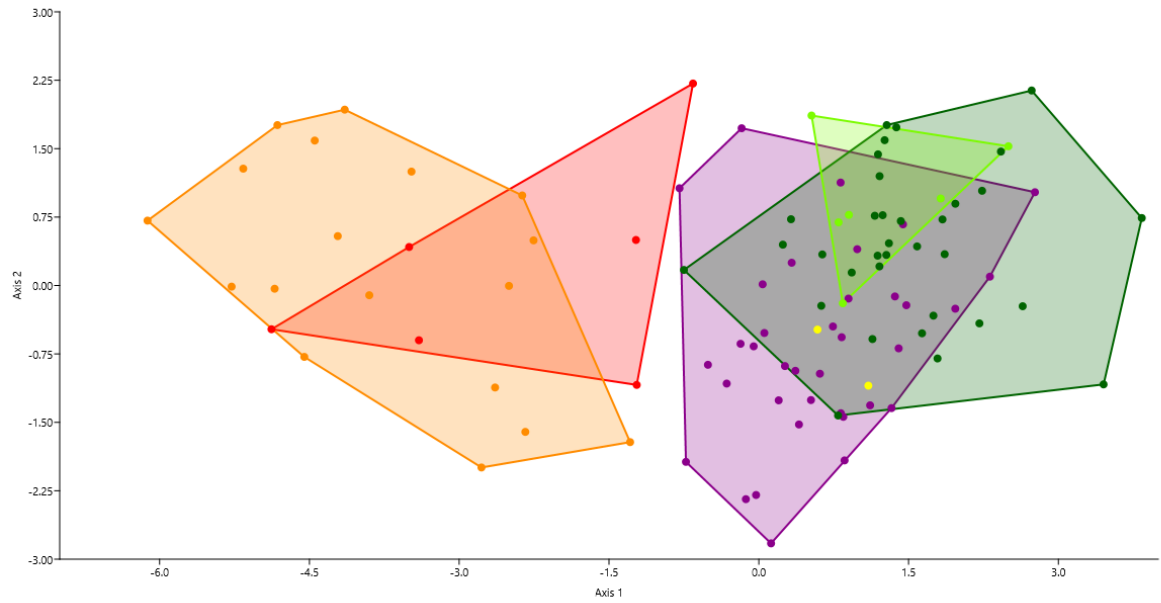


Figure 7. Paleoclimatic Models: *I. c. croconotus*; LGM (A) and Holocene (B), respectively and for the present (C).

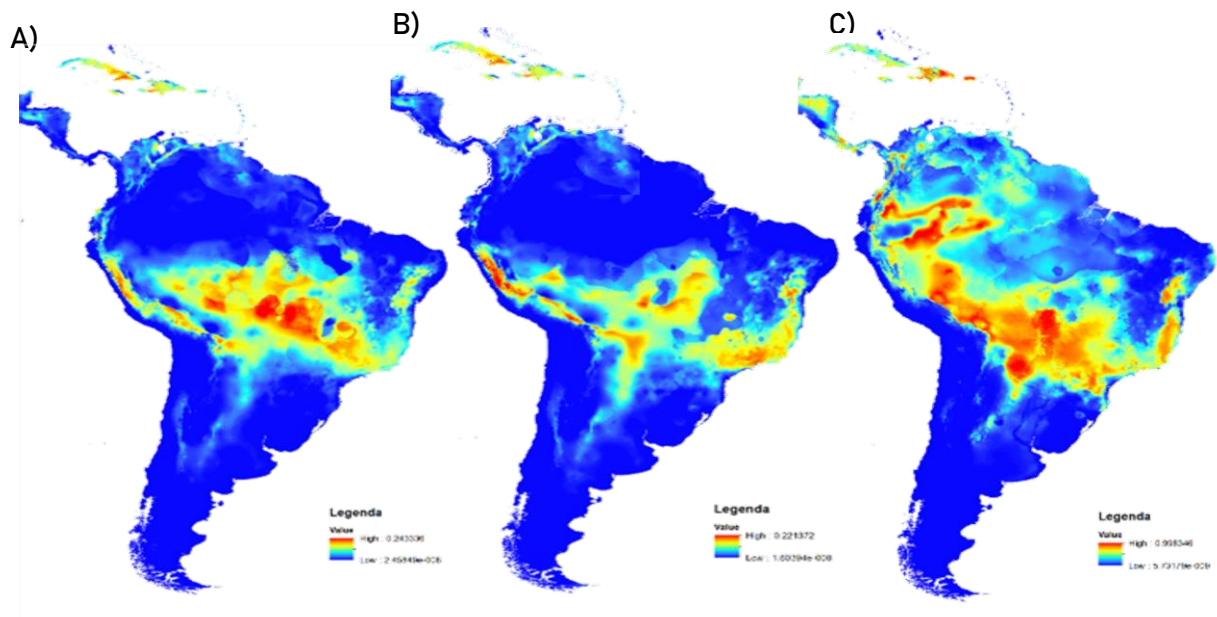
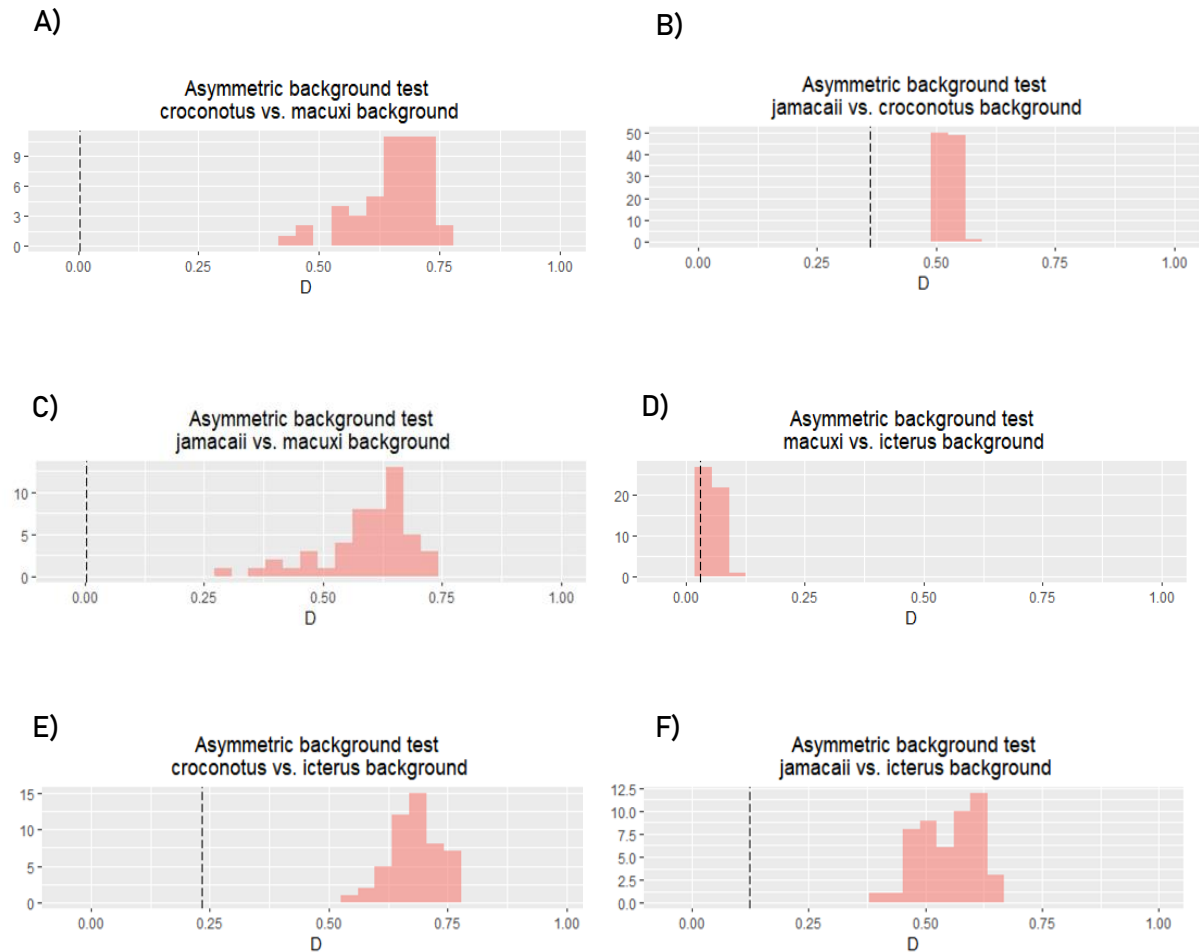


Figure 8. Background similarity tests among ecological niche models (ENMs) for lineages: A. *I. c. croconotus* lineage versus Roraima-Rupununi lineage; B. *I. jamacaii* lineage versus *I. c. croconotus* lineage. C. *I. jamacaii* lineage versus Roraima-Rupununi lineage; D. Roraima-Rupununi versus *icterus* lineage; E. *I. c. croconotus* lineage versus *icterus* lineage; F. *I. jamacaii* lineage versus *icterus* lineage. Dotted lines show the niche overlap values of the original ENMs. significance ($P < 0.01$).



TABLES

Table 1: Results of the Tukey post-hoc tests for bill height among taxa, following a significant ANOVA ($p < 0.001$). Significant pairwise differences are indicated in bold.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>ridgwayi</i>	<i>RR</i>
<i>jamacaii</i>		0.9623	0.9952	0.00251	0.0008606	0.9847
<i>croconotus</i>	1.165		1	0.0006358	0.0001197	0.9275
<i>stictifrons</i>	0.735	0.09095		0.01376	0.0313	0.9493
<i>icterus</i>	5.549	6.111	4.788		0.9127	0.5519
<i>ridgwayi</i>	5.99	6.753	4.382	1.431		0.8554
<i>RR</i>	0.9479	1.365	1.25	2.37	1.637	

Table 2: Results of the Tukey post-hoc tests for primary wing among taxa, following a significant ANOVA ($p < 0.001$). Significant pairwise differences are indicated in bold.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>ridgwayi</i>	<i>RR</i>
<i>jamacaii</i>		1.36E-05	0.1036	0.8935	6.32E-06	0.9543
<i>croconotus</i>	7.418		0.9999	0.002443	4.28E-10	0.1108
<i>stictifrons</i>	3.698	0.3238		0.06417	1.44E-06	0.2566
<i>icterus</i>	1.508	5.508	3.976		0.1941	1
<i>ridgwayi</i>	7.674	13.69	8.159	3.289		0.8055
<i>RR</i>	1.221	3.657	3.083	0.2717	1.782	

Table 3: Results of the Tukey post-hoc tests for tarsus among taxa, following a significant ANOVA ($p < 0.001$). Significant pairwise differences are indicated in bold.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>ridgwayi</i>	<i>RR</i>
<i>jamacaii</i>		0.9936	0.5915	0.9975	3.73E-08	0.5973
<i>croconotus</i>	0.7836		0.4046	0.9749	6.36E-09	0.4804
<i>stictifrons</i>	2.282	2.696		0.9489	0.1286	0.9935
<i>icterus</i>	0.6408	1.061	1.254		0.006413	0.8437
<i>ridgwayi</i>	9.306	9.859	3.564	5.099		0.9219
<i>RR</i>	2.27	2.524	0.787	1.674	1.392	

Table 4: Results of the Tukey post-hoc tests for secondary wing among taxa, following a significant ANOVA ($p < 0.001$). Significant pairwise differences are indicated in bold.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>ridgwayi</i>	<i>RR</i>
<i>jamacaii</i>		0.0004348	0.1724	0.5413	7.96E-05	0.9995
<i>croconotus</i>	6.186		1	0.00144	4.36E-10	0.4918
<i>stictifrons</i>	3.371	0.01613		0.02869	1.81E-05	0.6131
<i>icterus</i>	2.391	5.722	4.401		0.7472	0.9893
<i>ridgwayi</i>	6.806	11.82	7.321	1.931		0.6378
<i>RR</i>	0.4663	2.499	2.235	0.8764	2.182	

Table 5: Results of the Tukey post-hoc tests for bill width among taxa, following a significant ANOVA ($p < 0.001$). Significant pairwise differences are indicated in bold.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>ridgwayi</i>	<i>RR</i>
<i>jamacaii</i>		0.9997	1	0.188	3.23E-08	0.9
<i>croconotus</i>	0.4071		1	0.2709	1.77E-07	0.9338
<i>stictifrons</i>	0.06567	0.1476		0.5499	0.001566	0.9491
<i>icterus</i>	3.317	3.045	2.373		0.2321	0.9994
<i>ridgwayi</i>	9.486	8.94	5.722	3.164		0.4658
<i>RR</i>	1.482	1.335	1.252	0.4776	2.558	

Table 6: Results of the Tukey post-hoc tests for tail among taxa, following a significant ANOVA ($p < 0.001$). Significant pairwise differences are indicated in bold.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>ridgwayi</i>	<i>RR</i>
<i>jamacaii</i>		0.4794	0.9997	0.741	0.6538	0.9981
<i>croconotus</i>	2.526		0.9848	0.9985	1	0.9124
<i>stictifrons</i>	0.4188	0.9472		0.9622	0.9865	0.9944
<i>icterus</i>	1.947	0.5741	1.167		0.9992	0.8706
<i>ridgwayi</i>	2.146	0.0532	0.9227	0.5064		0.9156
<i>RR</i>	0.6038	1.433	0.7634	1.588	1.42	

Table 7: Fixation indices values between all of taxa.

	<i>jamacaii</i>	<i>croconotus</i>	<i>stictifrons</i>	<i>icterus</i>	<i>RR</i>
<i>jamacaii</i>		0.2294848	0.2381065	0.6100122	0.5920625
<i>croconotus</i>	0.2294848		0.1514265	0.6115108	0.5669399
<i>stictifrons</i>	0.2381069	0.1514265		0.6058797	0.5755839
<i>icterus</i>	0.6100122	0.6115108	0.6058797		0.6374355
<i>RR</i>	0.5920625	0.5669399	0.5755839	0.6374355	

Table 8: Niche overlap value observed for each pair of taxa.

<i>Schoener's D</i>	
<i>I. jamacaii</i> vs <i>I. croconotus</i>	0.36
<i>I. jamacaii</i> vs <i>RR</i>	0.002
<i>I. jamacaii</i> vs <i>icterus</i>	0.12
<i>I. croconotus</i> vs <i>RR</i>	0.003
<i>I. croconotus</i> vs <i>icterus</i>	0.2
<i>icterus</i> vs <i>RR</i>	0.02

Table 9. List of climatic variables available in Worldclim version 1.4 (www.worldclim.org) used in the study.

Environmental variable	Description
BIO2*	Monthly temperature range
BIO4*	Temperature Seasonality (standard deviation *100)
BIO6*	Min Temperature of Coldest Month
BIO14*	Precipitation of Driest Month
BIO15*	Precipitation Seasonality (Coefficient of Variation)
BIO18*	Precipitation of Warmest Quarter

SUPPLEMENTAR MATERIAL

Figure S1. Boxplot of all taxa within the *Icterus icterus* complex, except *Icterus. i. metae*. **A.** Bill height; **B.** Primary Wing; **C.** Tarsus; **D.** Secondary Wing; **E.** Bill width; **F.** Tail. Lightgreen, *I. c. I. i. strictifrons*; Darkgreen, *I. c. I. c. croconotus*; Purple, *I. I. jamaicai*; Red, *I. i. icterus*; Orange, *I. i. Ridgwayi*; Yellow, *I. c. croconotus* Roraima.

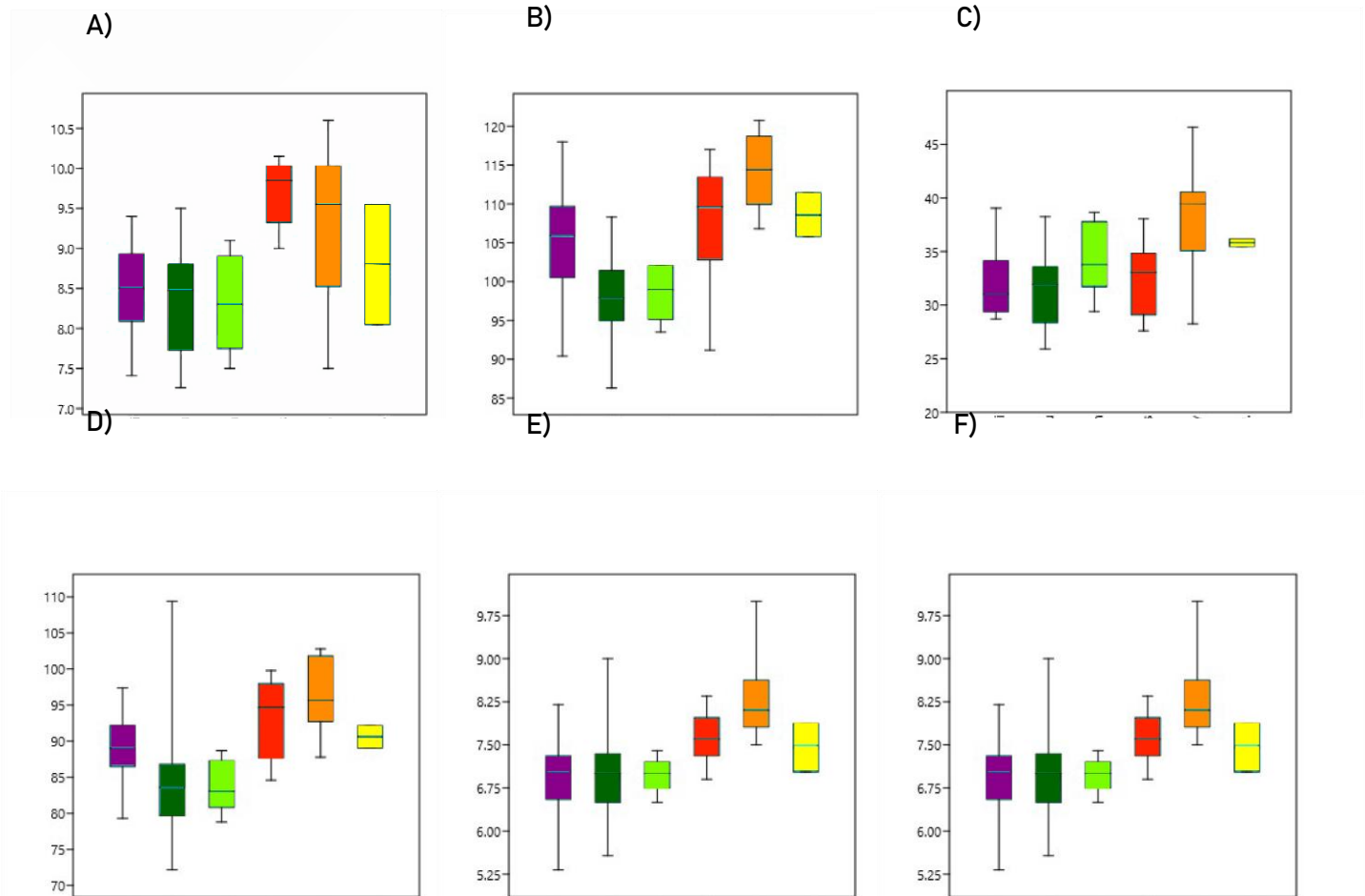


Table S1. Sample information.

Taxon	Voucher	Locality	Long	Lat
<i>I. i. ridgwayi</i>	Ictictid_LSUMNS_11410	Puerto Rico	-66.5901	18.22083
<i>I. i. ridgwayi</i>	Ictictid_LSUMNS_11328	Puerto Rico	-66.5901	18.22083
<i>I. c. strictifrons</i>	Icticrostr_KANSAS_3237_KSB359	Paraguay; Presidente Hayes, Laguna Capitan	-59.6807	-22.5335
<i>I. c. strictifrons</i>	Icticrostr_KANSAS_3258_KZ647	Paraguay; Presidente Hayes, Laguna Capitan	-59.6807	-22.5335
<i>I. c. strictifrons</i>	Icticrostr_LSUMNS_37880	Bolivia; Santa Cruz Dept	-63.1956	-17.7874
<i>I. c. strictifrons</i>	Icticrostr_LSUMNS_18648	Bolivia; Santa Cruz Dept	-63.1956	-17.7874
<i>I. c. strictifrons</i>	Icticrostr_LSUMNS_37837	Bolivia; Santa Cruz Dept	-63.1956	-17.7874
<i>I. c. strictifrons</i>	Icticrostr_LSUMNS_18811	Bolivia; Santa Cruz Dept	-63.1956	-17.7874
<i>I. c. strictifrons</i>	Icticrostr_LSUMNS_18801	Bolivia; Santa Cruz Dept	-63.1956	-17.7874
<i>I. c. croconotus</i>	Icticrocro_LSUMNS_4560	Peru; Loreto Dept	-73.5786	-4.50639
<i>I. c. croconotus</i>	Icticrocro_MPEG_79007	AM: Rio Solimões Ilha do Camaleão	-67.1958	-3.38143
<i>I. c. croconotus</i>	Icticrocro_MPEG_79251	PA; Monte Alegre Rio Amazonas Ilha Cacoal Grande	-56.5646	-2.38346
<i>Savanna-Rupununi</i>	Icticrocro_UFPE_T898	RR: Ilha do Rio Uraricoera, ca 65 km N, Boa Vista	-62.2604	3.408684
<i>Savanna-Rupununi</i>	Icticrocro_UFPE_T903	RR: Margem dir Rio Uraricoera, ca 60 km N. Boa vista, antiga Fazenda Bamerindus	-61.7271	3.132018
<i>I. jamacaii</i>	Ictjam_UFRN_222	Rio Grande do Norte	-38.853	-6.724
<i>I. jamacaii</i>	Ictjam_UFPB_CUR007	Curaçá	-39.9058	-8.99166
<i>I. jamacaii</i>	Ictjam_UFPE_T1625	BA: Boqueirão da Onça; ~80km NWN Senhor do Bonfim, "Mata Roçada"	-41.1098	-10.1794
<i>I. jamacaii</i>	Ictjam_UFPE_T1613	BA: ~20km NNW de Jacobina; "Fazenda Alto Bonito Barracão de baixo", 620m	-40.5674	-11.0231
<i>I. jamacaii</i>	Ictjam_UFPE_T1629	Ceará, Quixadá	-38.9151	-4.92912
<i>I. jamacaii</i>	Ictjam_UFPE_T1627	Rio Grande do Norte	-38.853	-6.724
<i>I. jamacaii</i>	Ictjam_UFPE_T1628	Ceará, Potengi	-40.0290	-7.08323
<i>I. jamacaii</i>	Ictjam_MPEG_72168	PI; Mun. Bocaina; Com. Balseiro	-41.3903	-6.94705
<i>Agelaiodes badius*</i>	Age_bad_LSUMNS_1069	-	-	-
<i>Neopsar nigerrimus*</i>	NC	-	-	-

(*) used as outgroup

Table S2. Substitution models of each part of mitochondrial genes.

Protein-codon gene	Evolutionary model	BEAST
ND1	HKY+F	
ND2	TN+F+I	
COX1	TN+F	
COX2	HKY+F+G4	
ATP6	TN+F	
COX3	TN+F	
ND3	TN+F	
ND4	TPM3+F+G4	GTR
ND5	TIM3+F+G4	GTR
CYTB	TN+F+I	

CONCLUSÃO

Este estudo representa o primeiro estudo filogeográfico abrangente em grande escala dos trupais, incluindo amostras de todas as espécies envolvidas no complexo de espécies *Icterus icterus*, em toda a sua distribuição geográfica. Além disso, usando marcadores de todo o genoma e o mitogenoma completo, fomos capazes de resolver aspectos importantes da história evolutiva desse grupo e fazer inferências sobre os processos por trás dos padrões de distribuição atuais. Estudos anteriores realizados com o objetivo de desvendar a história evolutiva desse grupo fizeram parte de estudos em escalas filogenéticas mais amplas, com amostragem de táxons muito limitada e poucos locos mitocondriais e nucleares. Existem três resultados principais deste estudo que merecem destaque. Primeiro, recuperamos quatro linhagens principais dentro do grupo, apoiando o tratamento atual de três espécies, e revelando o que acreditamos representar uma quarta espécie não descrita fenotipicamente diagnosticável das savanas de Roraima-Rupununi. Esses resultados foram consistentes entre marcadores (UCEs, mitogenome, SNPs) e métodos de construção de árvores. Em segundo lugar, encontramos diferentes níveis de estruturação genética nas duas espécies para as quais tivemos amostragem adequada, *I. jamacaii* e *I. croconotus*. Enquanto *I. jamacaii* não apresentou praticamente nenhuma estrutura genética em toda a Caatinga, *I. croconotus* mostrou populações geograficamente estruturadas, consistentes com a classificação subespecífica atual e sugerindo adaptação ecológica em dois biomas diferentes. Por fim, exploramos a relação entre o estabelecimento dessas linhagens e a dinâmica da paisagem neotropical, sugerindo cenários potenciais e alternativos que poderiam explicar os padrões atuais de estrutura genética, incluindo evidências de introgressão do semiárido brasileiro em populações atualmente isoladas do Chaco, consistentes com a hipótese do Arco do Pleistoceno. Portanto, nós fornecemos resultados que são congruentes em toda a estrutura populacional, relações filogenéticas e história climática, e oferecemos um bom ponto de partida para o entendimento da história evolutiva desta espécie.

REFERÊNCIAS

- ALLEN, E. S; OMLAND, K. E. Novel intron phylogeny supports plumage convergence in orioles (*Icterus*). **Auk**, 120, 961–969, 2003.
- ANDERSON, R. P. & GONZALEZ, I. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. **Ecological Modelling**, 222(15), 2796–2811, 2011.
- BANDA, K. R; DELGADO-SALINAS, A; DEXTER, K. G; LINARES-PALOMINO, R; OLIVEIRA-FILHO, A; PRADO, D; PULLAN, M; QUINTANA, C; RIINA, R; RODRIGUEZ, G. M; WEINTRITT, J; ACEVEDO-RODRIGUEZ, P; ADARVE, J; ALVAREZ, E; ARANGUREN, A. B; ARTEAGA, J. C; AYMARD, G; CASTANO, A; CEBALLOS-MAGO, N; COGOLLO, A; CUADROS, H; DELGADO, F; DEVIA, W; DUENAS, H; FAJARDO, L; FERNANDEZ, A; FERNANDEZ, M. A; FRANKLIN, J; FREID, E. H; GALETTI, L. A; GONTO, R; GONZALEZ, R. M; GRAYESON, R; HELMER, E. H; IDARRAGA, A; LOPEZ, R; MARCANO-VEGA, H; MARTINEZ, O. G; MATURO, H. M; MCDONALD, M., MCLAREN, K; MELO, O; MIJARES, F; MOGNI, V; MOLINA, D; MORENODEL, N. P; NASSAR, J. M; NEVES, D. M; OAKLEY, L. J; OATHAM, M; OLVERA-LUNA, A. R; PEZZINI, F. F; DOMINGUEZ, O. J. R; RIOS, M. E; RIVERA, O; RODRIGUEZ, N; ROJAS, A; SARKINEN, T; SANCHEZ, R; SMITH, M; VARGAS, C; VILLANUEVA, B; PENNINGTON, R.T; Plant diversity patterns in neotropical dry forests and their conservation implications. **Science**, 353 (80), 1383–1387, 2016.
- BirdLife International. **IUCN Red List for birds**. Downloaded from <http://www.birdlife.org> on 19/10/2020.
- BOLGER, A. M; LOHSE, M., & BJOERN, U. Trimmomatic: A flexible trimmer for Illumina sequence data. **Bioinformatics**, 30, 2114–2120, 2012.
- BOLÍVAR-LEGUIZAMÓN, S. D; SILVEIRA, L. F; DERRYBERRY, E. P; BRUMFIELD, R. T; & BRAVO, G. A. Phylogeography of the variable antshrike (*Thamnophilus caerulescens*), a South American passerine distributed along multiple environmental gradients. **Molecular Phylogenetics and Evolution**, 148, 106810, 2020.
- BORIA, R. A; OLSON, L. E; GOODMAN, S. M. & ANDERSON, R. P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. **Ecological Modelling**, 275, 73–77, 2014.
- BORTOLUS, A. Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. **J Hum Environ**, 37(2):114–118, 2008.
- BOUCKAERT, R; HELED, J; KÜHNERT, D; VAUGHAN, T; WU, C. H; XIE, D; SUCHARD, M. A; RAMBAUT, A. & DRUMMOND, A. J. BEAST 2: A software platform for bayesian evolutionary analysis. **PLOS Computational Biology**, 10(4), e1003537, 2014.
- BROWN, J. L. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic, and species distribution model analyses. **Methods in Ecology and Evolution**, 2014.
- BRYANT, D; BOUCKAERT, R; FELSENSTEIN, J; ROSENBERG, N. A. & ROYCHOUDHURY, A. Inferring species trees directly from biallelic genetic markers:

bypassing gene trees in a full coalescent analysis. **Molecular Biology and Evolution**, 29(8), 1917–1932, 2012.

CADENA, C. D; CUERVO, A. M. Molecules, ecology, morphology, and songs in concert: how many species is *Arremon torquatus* (Aves: Emberizidae). **Biological Journal of Linnean Society**, 99:152–176, 2010.

CAETANO, S. The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). **Molecular Ecology**, 17, 3147–3159, 2008.

CARDOSO DA SILVA, J. M. and OREN, D. C. Introduced and invading birds in Belém, Brazil. *Wilson Bull.* 102: 309-313, 1990.

CARNAVAL, A. C; MORITZ, C. Historical climate modelling predicts patterns of current biodiversity in the brazilian atlantic forest. **Journal Of Biogeography**, v, 35, n, 7, p, 1187-1202, 2008.

CARNEIRO-FILHO, A. “Roraima Savannas”: Clímax situation or botanic relic. In: Prost MT (Org) *Évolution des Littoraux de Guyane et de la Zone Caraïbe Méridionale Pendant le Quaternaire*, Cayenne (Guiane), pp 31-48, 1990.

CARNEIRO-FILHO, A. Cerrados amazônicos: Fósseis vivos? Algumas reflexões. **Revista IG**, 14, 63-68, 1993.

CARNEIRO-FILHO, A. **Contribution to the study of the forest-savanna mosaic in the area of Roraima, northern Amazon basin, Brazil**, MSc Thesis, ITC-Enschede, 108 pp, 1991.

CHATTOPADHYAY, B; GARG, K. M; GWEE, C. Y; EDWARDS, S.V; RHEINDT, F. E. Gene flow during glacial habitat shifts facilitates character displacement in a Neotropical flycatcher radiation. **BMC Evol. Biol.** 17, 210, 2017.

COLLEVATTI, R. G; TERRIBILE, L. C; DE OLIVEIRA, G; LIMA-RIBEIRO, M. S; NABOUT, J. C; RANGEL, T. F; DINIZ-FILHO, J. A. F. Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests. **J Biogeogr.** 40:345–358, 2013.

COLLEVATTI, R. G; TERRIBILE, L. C; LIMA-RIBEIRO, M. S; NABOUT, J. C; OLIVEIRA, G; RANGEL, T. F; RABELO, S. G. and DINIZ-FILHO, J. A. F. 'A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species'. **Molecular Ecology**, 21(23), pp. 5845-5863, 2012.

CORBETT, E. C; BRAVO, G. A; SCHUNCK, F; NAKA, L. N; SILVEIRA, L.F; EDWARDS, S. V. Evidence for the pleistocene arc hypothesis from genome-wide SNPs in a Neotropical dry forest specialist, the Rufous-fronted Thornbird (Furnariidae: *Phacellodomus rufifrons*). **Mol Ecol**, 2020;00:1–16, 2012.

COSTA, G. C; HAMPE, A; LEDRU, M. P; MARTINEZ, P. A; MAZZOCHINI, G. G; SHEPARD, D. B; WERNECK, F. P; MORITZ, C; CARNAVAL, A. C; Biome stability in South America over the last 30 kyr: inferences from long-term vegetation dynamics and habitat modelling. **Glob Ecol Biogeogr**, 27:285–297, 2018

- COX, C. B; MORRE, P. D. Biogeography – An ecological and evolutionary approach. **Blackwell Science**, p, 325, 2000.
- CRACRAFT, J. 'Historical biography and patterns of differentiation within the South American avifauna: Areas of Endemism'. **Ornithological Monograph**, 36, pp. 49-84, 1985.
- DE MELO, W. A; LIMA-RIBEIRO, M. S; TERRIBILE, L. C. & COLLEVATTI, R. G. Coalescent simulation and paleodistribution modeling for *Tabebuia rosealba* do not support South American dry forest Refugia hypothesis. **PLoS One**, 11(7), e0159314, 2016.
- DE SCHAUENSEE, R. M. The birds of the Republic of Colombia. Family Icteridae. Caldasia, 5:983-1001, 1952.
- EDEN, M. Paleoclimatic influences and the development of savanna in southern Venezuela. **Journal of Biogeography**, 1, 95-109, 1974.
- ENDLER, J. A. **Geographic variation, speciation, and clines**. Princeton University. Nº. 10. Press, Princeton, 1977.
- FAIRCLOTH, B. C. PHYLUCE is a software package for the analysis of conserved genomic loci. **Bioinformatics**, 32, 786–788, 2016.
- FAIRCLOTH, B. C; GLENN T. C. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. **PLoS One**, 7: e42543, 2012.
- FAIRCLOTH, B. C; MCCORMARCK, J. E; CRAWFORD, N. G; HARVEY, M. G; BRUMFIELD, R. T. & GLENN, T. C. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescale. **Systematic Biology**, 61, 717–726, 2012.
- FRICHOT, E; FRANCOIS, O. LEA: An R package for landscape and ecological association studies. **Methods Ecol. Evol.** 6, 925–929, 2015.
- FRICHOT, E; MATHIEU, F; TROUILLON, T; BOUCHARD, G; FRANCOIS, O. Fast and efficient T estimation of individual ancestry coefficients. **Genetics**, 196, 973–983, 2014.
- HAFFER, J. Speciation in Amazonian Forest Birds. **Science**, v. 165, p. 131–137, 1969.
- HARVEY, M. G; SMITH, B. T; GLENN, T. C; FAIRCLOTH, B. C; & Brumfield, R. T. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. **Systematic Biology**, 65, 910–924, 2016.
- HELLMAYR, C. E. **Catalogue of the birds of the Americas and the adjacent islands**. Vol. 12. Field Museum of Natural History, Chicago, USA, 1937.
- HIJMANS, J. R; CAMERON, S. E; PARRA, J. L; JONES, P. G. E; JAVIS, A. Very High Resolution Interpolated Climate Surfaces For Global Land Areas. **International Journal Of Climatology** v, 25, p, 1965-1978. 2005.
- HILTY, S. L. and BROWN W. L. **A guide to the birds of Colombia**. Princeton, NJ: Princeton Univ. Press, 1986
- HILTY, S. L. **Birds of Venezuela**. Second edition. Princeton, NJ: Princeton Univ. Press, 2002.

HOFFMAN, C. M; CRONIN, T. W. and OMLAND, K. E. Using spectral data to reconstruct evolutionary changes in coloration: carotenoid color evolution in New World orioles. **Evolution**, 60:1680-1691, 2006.

HOWELL, S. N. G. and WEBB, S. **The birds of Mexico and northern Central America**. Oxford Univ. Press, Oxford, U.K, 1995.

ISLER, M; ISLER, P. and WHITNEY, B. M. Use of vocalizations to establish species limits in ant- birds (Passeriformes: Thamnophilidae). **Auk**, 115:577–590, 1998.

JACOBSEN, F; FRIEDMAN, N. R; OMLAND, K. E. Congruence between nuclear and 1043 mitochondrial DNA: Combination of multiple nuclear introns resolves a well- 1044 supported phylogeny of New World orioles (Icterus). **Mol. Phylogenet. Evol**, 56, 1045 419–427, 2010.

JACOBSEN, F; OMLAND, K. E. Species tree inference in a recent radiation of 1047 orioles (Genus Icterus): Multiple markers and methods reveal cytonuclear 1048 discordance in the northern oriole group. **Mol. Phylogenet. Evol**. 61, 460–469, 2011.

JARAMILLO, A. and BURKE, P. **New World blackbirds. The Icterids**. London: A. & C. Black Publishers, 1999.

JERMIIN, L. S. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. **Nature Methods**, 14:587–589, 2017.

KALYAANAMOORTHY, S; MINH, B.Q; WONG, T. K. F; VON HAESELER, A. LEITE, Y. L; COSTA, L. P; LOSS, A. C; ROCHA, R. G; BATALLA-FILHO, H; BASTOS, A. C; QUARESMA V. S; FAGUNDES, V; PARESQUE, R; PASSAMANI, M; PARDINI, R. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. **Proc Natl Acad Sci**, 113(4):1008–1013, 2016.

LI, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. (retrieved from arXiv preprint arXiv:1303.3997), 2013.

LINDELL, C. & BOSQUE, C. Notes on the breeding and roosting biology of Troupials (*Icterus icterus*) in Venezuela. **Ornit. Neotrop**, 10:85-90, 1999..

LOVETTE, I. J; BERMINGHAM, E; RICKLEFS, R. E. Mitochondrial DNA 1080 phylogeography and the conservation of endangered Lesser Antillean Icterus 1081 orioles. **Conserv. Biol**, 13, 1088–1096, 2001.

LOWTHER, P. Geographic and ecological variation in the family Icteridae. **Wilson Bull**, 87:481–495, 1975.

MAYLE, F. E. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. **Journal of Quaternary Science**, 19(7), 713–720; 2004.

MAYR, E. **Populations, species and evolution: an abridgment of animal species and evolution**. Harvard University Press, Cambridge. ISBN 0674690133, 1970.

MCKENNA, A; HANNA, M; BANKS, E; SIVACHENKO, A; CIBULSKIS, K; Kernytzsky, A. & DEPRISTO, M. A. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Research**, 20, 1297–1303, 2010.

MORAES, E. M; YOTOKO, K. S. C; MANFRIN, M. H; SOLFERINI, V. N. & SENE, F. M. Phylogeography of the cactophilic species *Drosophila gouveai*: Demographic events and divergence timing in dry vegetation enclaves in Eastern Brazil. **Journal of Biogeography**, 36(11), 2136–2147, 2009.

MORRONE, J. J. On The Identification of Areas Of Endemism. **Systematic Biology**. v. 43, p. 438-441, 1994.

NAKA, L.N; PIL, M. W. Moving beyond the riverine barrier vicariant paradigm. **Mol Ecol**, 00:1–4, 2020.

OMLAND, K. E; LANYON, S. M. and FRITZ, S. J. A molecular phylogeny of the New World orioles (*Icterus*): the importance of dense taxon sampling. **Molecular Phylogenetics and Evolution**, 12:224-239, 1999.

OMLAND, K. E; LANYON, S. M. Reconstructing plumage evolution in orioles 1102 (*Icterus*): repeated convergence and reversal in patterns. **Evolution**, 54, 2119–1103 2133, 2000.

OMLAND, K. E; LANYON, S. M; FRITZ, S. J. A molecular phylogeny of the New World Orioles (*Icterus*): the importance of dense táxon sampling. **Molec. Phylog. Evol**, 12: 224–239, 1999.

PACHECO, J. F; and OLMOS, F. As aves do Tocantins, parte 1. **Rev. Bras. Orn**, 14:85-100, 2006.

PENNINGTON, R. T; LAYIN, M. & OLIVEIRA-FILHO, A. **Woody plant diversity, evolution and ecology in the tropics: Perspectives from seasonally dry tropical forests**, 2009.

PENNINGTON, T; PRADO, D. E. and PENDRY, C. A. 'Neotropical seasonally dry forests and Quaternary vegetation changes'. **Journal of Biogeography**, 27(2), pp. 261-273, 2000.

PHILLIPS, S. J; ANDERSON, R. P. and SCHAPIRE, R. E. 'Maximum entropy modeling of species geographic distributions'. **Ecological Modelling**, 190(3), pp. 231-259, 2006.

PINTO, O. **Do parasitismo provavel de *Icterus I. jamacaii* (Gmelin) em *Pseudoseisura cristata* (Gmelin)**. Hornero 10:447-449, 1967.

POWELL, A. F. et al. A comprehensive species-level molecular phylogeny of the New World. **New World**, v. 3, p. 1, 2013.

PRADO, D. E. & GIBBS, P. E. Patterns of species distributions in the dry seasonal forests of South America. **Annals**, 2009.

PRADO, D. E. **A critical evaluation of the floristic links between Chaco and Caatingas vegetation in South America**. PhD thesis, University of St Andrews, Scotland, 1991.

- RAMBAUT, A; DRUMMOND, A. J; XIE, D; BAELE, G. & SUCHARD, M. A. Posterior summarization in Bayesian phylogenetics using tracer 1.7. **Systematic Biology**, 67(5), 901–904, 2018.
- REMSEN, J. V; ROCHA, O; SCHMITT, C. G. & SCHMITT D.C. Zoogeography and geographic variation of *Platyrrhinus mystaceus* in Bolivia and Peru, and the circum-Amazonian distribution pattern. **Ornitol. Neotrop**, 2: 77–83, 1991.
- RIBAS, C. C; ALEIXO, A; NOGUEIRA, A. C. R; MIYAKI, C. Y; CRACRAFT, J. A. palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. **Proc R Soc B**, 799:681–689, 2012.
- RIDGELY, R. S. & TUDOR G. **The birds of South America**. Oxford, University Press, vol. 1, 516p, 1999.
- Roraima: Uma resenha do Cretáceo ao recente. In: Barbosa RI, Ferreira E, Castellón E (Eds), Homem, Ambiente e Ecologia em Roraima, Instituto Nacional de Pesquisas da Amazônia, Manaus, pp 231-265, 1997.
- RUELAN, F. **Expedições Geomorfológicas no Território do Rio Branco**. INPA, Rio de Janeiro, 170 pp, 1957.
- RULL V. Palaeoclimates and amazon biodiversity. **Journal of Biogeography**. 40: 1413–1414, 2013.
- RULL, V. Neotropical biodiversity: timing and potential drivers. **Trends Ecol Evol**, 26:508–513, 2011.
- RULL, V. Speciation timing and neotropical biodiversity: the Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. **Mol Ecol**, 17:2722–2729, 2018.
- SAVIT, A. and BATES, J. 'Right around the Amazon: the origin of the circum-Amazonian distribution in *Tangara cayana*'. **Folia Zoologica**, 64(3), pp. 273-283, 2015.
- SCHLUTER, D. **The ecology of adaptive radiation**. Oxford University Press, Oxford. ISBN 019850523X, 2000.
- SCHOENER, T. W. The Anolis Lizards of Bimini: Resource Partitioning in a Complex Fauna. **Ecology Society of America**, 49(4), 704–726, 1968.
- SHORT, L. L. **A zoogeographic analysis of the South American Chaco avifauna**. Bull. Am. Mus. Nat. Hist. 154: 163-352, 1975.
- SIBLEY, C.G; MONROE, JR, B. L. **Distribution and Taxonomy of Birds of the World**. 1165, Yale University Press, New Haven, CT, 1990.
- SILVA, S. M; PETERSON, A. T; CARNEIRO, L; BURLAMAQUI, T. C. T; RIBAS, C. C; SOUSA-NEVES, T; MIRANDA, L. S; FERNANDES, A. M; D'HORTA, F. N; ARAÚJO-SILVA, L. E; BATISTA, R; BANDEIRA, C. H. M. M; DANTAS, S. M; FERREIRA, M; MARTINS, D. M; OLIVEIRA, J; ROCHA, T. C; SARDELLI, C. H; THOM, G; RÊGO, P. S; SANTOS, M. P; SEQUEIRA, F; VALLINOTO, M; ALEIXO, A. A dynamic continental moisture gradient drove Amazonian bird diversification. **Sci Adv**, 2019.

- SIMÕES-FILHO, F; TURCQ, B; CARNEIRO-FILHO, A; SOUZA, A. G. Registros sedimentares de lagos e brejos dos Campos de Roraima: implicações 21 paleoambientais ao longo do Holoceno. In: Barbosa RI, Ferreira E, Castellón E (Eds), *Homem, Ambiente e Ecologia em Roraima*, Instituto Nacional de Pesquisas da Amazônia, Manaus, pp 295-305, 1997.
- SMITH, B. T; MCCORMACK, J. E; CUERVO, A. M; HICKERSON, M. J; ALEIXO, A; CADENA, C. D; EMÁN, J. P; C. BRUNEY, W; XIE, X; HARVEY, M. G; FAIRCLOTH, B. C; GLENN, T. C; DERRYBERRY, P; PREJEAN J; FIELDS S; BRUMFIELD, R. T. The Drivers Of Tropical Speciation. **Nature**, v. 515, n. 7527, p. 406-409, 2014.
- STURGE, R. J; JACOBSEN, F; ROSENSTEEL, B. B; NEALE, R. J; OMLAND, K. E. Colonization 1167 of South America from Caribbean islands confirmed by molecular phylogeny 1168 with increased taxon sampling. **Condor**, 111, 575–579, 2009.
- TOLEDO, M. B. **Holoceno vegetation and climate history of savanna-forest ecotones in Northeastern Amazonia**. PhD Thesis, Department of Biological Science/Florida Institute of Technology, 178 pp, 2004.
- VANZOLINI, P; WILLIAMS, E. South American Anoles: The Geographic Differentiation And Evolution Of The *Anolis Chrysolepis* Species Group (Sauria: Iguanidae). **Arquiv Zool**, São Paulo, v.19, p. 1-29, 1970.
- WARREN, D. L; GLOR, R. E. & TURELLI, M. ENMTools: A toolbox for comparative studies of environmental niche models. **Ecography**, 33(3), 607–611, 2010.
- WARREN, D. L; GLOR, R. E. & TURELLI, M. Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. **Evolution**, 62(11), 2868–2883, 2008.
- WERNECK, F. P; COLLI, G. R. The lizard assemblage from seasonally dry tropical forest enclaves in the cerrado biome, Brazil, and its association with the Pleistocenic arc. **J. biogeog**, 33, 1983-1992, 2006.
- WERNECK, F. P; GAMBLE, T; COLLI, G. R; Rodrigues, M. T. & SITES, J. W. Deep diversification and long-term persistence in the South American “Dry Diagonal”: Integrating continent-wide phylogeography and distribution modeling of geckos. **Evolution**, 66(10), 3014–3034, 2012.
- WERNECK, F. P; LEITE, R. N; GEURGAS, S. R; & RODRIGUES, M. T. Biogeographic history and cryptic diversity of saxicolous Tropicoduridae lizards endemic to the semiarid Caatinga. **BMC Evolutionary Biology**, 15, 94; 2015.
- WIENS, J. J; GRAHAM, C. H. Niche conservatism: integrating evolution, ecology, and conservation biology. **Annu Rev Ecol Evol Syst**, 36:519–539, 2005.
- WIENS, J. J; PARRA-OLEA, G; GARCIA-PARIS, M. et al Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. **Proc R Soc B Biol Sci**, 274(1612):919–928, 2007.
- WISZ, M. S; HIJMAN, R. J; PETERSON, A. T; GRAHAM C. H. E; GUIBAN, A. Predicting Species Distributions Working Group. Effects Of Sample Size On The Performance Of Species Distribution Models. **Diversity and Distributions**, v. 14, p. 763-773, 2008.

