UNIVERSIDADE DE BRASÍLIA – UnB
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL

Influência de fatores intrínsecos (migração) e das flutuações climáticas sobre a avifauna associada às áreas abertas da América do Sul: filogeografia comparada de espécies do gênero *Elaenia*

ELIANE LUIZ DE FREITAS

Brasília, Janeiro de 2022.

UNIVERSIDADE DE BRASÍLIA – UnB
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL

Influência de fatores intrínsecos (migração) e das flutuações climáticas sobre a avifauna associada às áreas abertas da América do Sul: filogeografia comparada de espécies do gênero *Elaenia*

Tese apresentada ao Programa de Pós-Graduação em Biologia Animal da Universidade de Brasília como requisito para a obtenção do título de Doutora em Biologia Animal.

Eliane Luiz de Freitas
Orientador: Dr. Renato Caparroz

Brasília, Janeiro de 2022.

“A academia não é o paraíso, mas o aprendizado, é um lugar onde o paraíso pode ser criado. A sala de aula com todas suas limitações continua sendo ambiente de possibilidades. Nesse campo de possibilidades, temos a oportunidade de trabalhar pela liberdade, exigir de nós e de nossos camaradas uma abertura da mente e do coração que nos permite encarar a realidade ao mesmo tempo em que, coletivamente, imaginemos esquemas para cruzar fronteiras, para transgredir. Isso é a educação como prática da liberdade.”

--hooks, bell. *Ensinando a transgredir: a educação como prática da liberdade. 2. ed. São Paulo:Martins Fontes, 2017. p. 273.*

AGRADECIMENTOS

Agradeço a Universidade de Brasília, minha segunda casa, e ao Programa de Pós-Graduação em Biologia Animal, ao corpo docente, técnicos, funcionários administrativos e todos aqueles que viabilizaram o desenvolvimento deste trabalho.

A CAPES e ao CNPq pelo financiamento da minha pesquisa, assim como, por minhas bolsas durante esse período. Agradeço especialmente ao programa de internacionalização CAPES-PDSE que me permitiu realizar parte da pesquisa na Cornell University - USA, firmando parceirias e contribuindo para meu crescimento pessoal e profissional.

A todas as pessoas que contribuíram com a obtenção de amostras para este estudo, aos colaboradores em expedições de campo, aos responsáveis pelas unidades de conservação e espaços particulares visitados, aos curadores das coleções ornitológicas do Museu Paraense Emílio Goeldi, da coleção da Universidade Federal do Pernambuco e da coleção da Universidade Federal do Rio Grando do Norte que forneceram amostras para ampliar nossa área amostrada e a qualidade desse trabalho.

A todos integrantes do Laboratório de Genética de Biodiversidade – LabGenBio por fornecer condições e infraestrutura para realizar esta pesquisa. Em especial, à Gislaine Fernandes, pela companhia e amizade, inclusive nos momentos mais difíceis durante o desenvolvimento deste trabalho e ao amigo e técnico Fred Horst, que proporcionou as condições necessárias para realização dos experimentos e organização do laboratório.

Ao meu orientador, Dr. Renato Caparroz, pelo suporte e orientação durante todo tempo de condução dessa pesquisa, especialmente durante a pandemia de COVID-19 (SARS-CoV-2) em que tive que pausar meu trabalho e cuidar de mim e de minha família.

A todos integrantes do Cornell Lab Of Ornithology pela experiêcia incrível de fazer pesquisa em um centro de referência na minha área. Ao Dr. Irby Lovette por ter me aceitado em seu grupo de pesquisa por 9 meses, ao Dr. Leonardo Campagna pela orientação exemplar e acolhimento, à Dra. Bronwyn Butcher pelo acompanhamento durante o trabalho no laboratório e de bioinformática. Agradeço também a todos que tornaram meu período em Ithaca inesquecível, pessoalmente e profissionalmente.

Aos coletivos, grupos, organizações que eu pude fazer parte, que completaram a minha formação como cientista acrescentando o papel da divulgação científica, da extensão, do contato

fora da academia. Entre eles o Pint of Science, projeto Escolas que Voam, Clube das Cientistas e a Rede Ornito Mulheres.

Gostaria de agradecer aos amigos, é impossível nomear todos, que acompanharam e me apoiaram em todos os momentos durante esta pesquisa, que compreendem os sacrifícios e ausências atrelados a profissão de cientista, aos amigos de jornada dentro e fora da academia, as pessoas incríveis que tive a oportunidade de conhecer neste caminho. Em especial ao Getúlio Júnior, à Juma Rezende, à Thaiz Armond e ao Victor Leandro, que me acompanharam no processo difícil de escrita da tese. Aos profissionais de educação física e à psicóloga por cuidarem da minha saúde física e mental e me permitiram chegar ao final desta longa caminhada.

Aos meus pais, Gilberto e Irany, por não medirem esforços em relação a minha educação, bem estar e saúde. Por terem sido exemplos de vida, força, coragem e família.

Gostaria de agradecer ao Getúlio, por ter sido um excelente companheiro e apoiador durante toda essa jornada.

A todos os professores que cruzaram meu caminho e dedicaram suas vidas a transformação do mundo pela educação. Aos programas sociais e de acesso à educação, que permitiram que eu não apenas ingressasse na Universidade de Brasília e na Cornell University, como garantiram a minha permanência na ciência.

SUMÁRIO

APRESENTAÇÃO

Este documento é o resultado da pesquisa que tem por objetivo geral entender os processos intrínsecos e extrínsecos que influenciam a diversificação de aves em áreas abertas na América do Sul, empregando marcadores moleculares nuclear e mitocondrial associados a análises populacionais e climáticas como ferramentas, e utilizando como modelos duas espécies do gênero *Elaenia* (Aves: Tyrannidae). Esta tese está estruturada em uma introdução geral, dois capítulos em formato de artigo científico, considerações finais e perspectivas futuras sobre o tema.

A introdução geral tem como objetivo contextualizar o cenário que este estudo foi conduzido, quais as lacunas no conhecimento sobre o impacto dos efeitos bióticos e abióticos na distribuição das linhagens de aves neotropicais, assim como proporcionar ao leitor referência para os capítulos seguintes.

O capítulo 1, entitulado "*Ecological traits drive genetic structuring in two open habitat birds from the morphologically cryptic genus Elaenia (Aves: Tyrannidae)*" foi revisado por pares e aceito com "*minor reviews*" para publicação na revista científica *Journal of Avian Biology* (https://doi.org/10.1111/jav.02931).

O capítulo 2, entitulado "*Effects of Quaternary climatic fluctuations on avifauna associated with open areas of South America: phylogeography of two species of the genus Elaenia (Aves: Tyrannidae)*" será submetido para a revista científica após a defesa desta tese.

As considerações finais abordam os achados deste estudo e as contribuições feitas ao conhecimento para as espécies foco e também para os ambientes de áreas abertas da América do Sul.

INTRODUÇÃO GERAL

A filogeografia é uma ciência relativamente nova. O termo definido por Avise e colaboradores (1987) descreve a integração dos estudos filogenéticos com estudos de genética de populações. Assim, possibilita o melhor entendimento de como os processos micro e macroevolutivos estão relacionados e podem auxiliar na compreensão de como os eventos históricos impactaram na atual distribuição geográfica das linhagens de espécies (AVISE, 2009; HICKERSON et al., 2010).

Os estudos filogeográficos, que começaram com uso de poucos marcadores devido às limitações metodológicas, atualmente contam com a disponibilidade de plataformas de sequenciamento de nova geração, capazes de diminuir o tempo de processamento e de obtenção de dados. Esses avanços permitem aumentar consideravelmente o conjunto de dados a serem analisados, quando comparado à outros métodos como sequenciamento do tipo *Sanger*, e que aliado às novas abordagens, como teoria da coalescência (HUDSON, 1991; KINGMAN, 1982, 2000), ou filogeografia estatística (KNOWLES, 2009; KNOWLES; MADDISON, 2002), permitem investigar maior número de fatores, testar hipóteses e cenários relacionados os processos de diversificação das espécies (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015).

O que determina a distribuição de uma espécie? Perguntas como essa tem sido objeto de estudo em vários campos do conhecimento. A combinação de fatores bióticos e abióticos pode atuar de diferentes formas e moldar a distribuição da biodiversidade (KUMAR; KUMAR, 2018). Fatores bióticos ou intrínsecos são características espécie-específicas, como exemplo, comportamento reprodutivo, capacidade de dispersão, migração e estratégias de forrageamento (BURNEY; BRUMFIELD, 2009; GÓMEZ-BAHAMÓN et al., 2020; SMITH et al., 2014). Os fatores abióticos ou extrínsecos geralmente estão relacionados ao espaço físico e suas características, como mudanças climáticas, variações geológicas, diferença altitudinal, entre outras (CALDERÓN et al., 2014; GUTIÉRREZ-PINTO et al., 2012; MAIRAL et al., 2017). O resultado das interações entre esses fatores deixa sinais na história evolutiva da biodiversidade e pode ser visualizando na variação genética entre populações de uma espécie através de medidas como o fluxo gênico (CURNOW; WRIGHT, 1978).

O grupo das aves é um dos mais diversos entre os vertebrados, é amplamente distribuído no planeta, o qual habita os mais diversos tipos de *habitat*, possui uma infinidade de comportamentos e os mais variados aspectos morfológicos (DEL HOYO et al., 2019). Essas

características, tornam-se sistemas adequados para testar hipóteses filogeográficas e entender os padrões de diversificação.

Tem sido crescente o uso de espécies de aves como modelos de estudos filogeográficos em diversos ambientes. Por exemplo, Matos e colaboradores (2016) buscaram entender melhor a história evolutiva de áreas florestais e abertas da Amazônia utilizando duas espécies de aves como modelos: *Tachyphonus phoenicius* e *Polytmus theresiae*. Os resultados obtidos através de marcadores moleculares, indicam que apesar das duas espécies terem histórias diferentes quanto às estruturas genéticas, ambas mostram sinais de expansão populacional recente, sugerindo que o tipo de *habitat* ocupado por elas (um tipo de vegetação de areia branca, comum na região amazônica) tenha aumentado historicamente sua área e a conexão entre fragmentos desse tipo de *habitat* tenha sido maior no passado, permitindo o encontro das populações.

Utilizando compilado de dados com registros de distribuição das espécies e subespécies de aves amazônicas, Oliveira e colegas (2017) buscaram identificar áreas de endemismo com base em teste de modelos biogeográficos. Os resultados indicaram diferenças em relação ao postulado anteriormente para a região. Alguns rios amazônicos representam limites de áreas de endemismo, funcionando como possíveis barreiras de dispersão, mas não são congruentes para todos os táxons de aves, ou seja, nem todas as espécies respondem da mesma maneira às diferenças na paisagem como se esperava, evidenciando um cenário mais complexo da distribuição das aves amazônicas.

Outro amplo estudo de filogeografia comparativa, Smith e colaboradores (2017), investigaram se os processos que possam ter conduzido às diferenças latitudinais na diversidade entre espécies podem ter efeito também dentro das espécies. Para as análises, utilizou-se um conjunto de dados genéticos de 210 espécies de aves do Novo Mundo distribuídas em um amplo gradiente latitudinal. Os resultados indicaram que as espécies de latitude inferior tinham, em média, maior diversidade filogeográfica do que as espécies de maior latitude. Os fatores associados à ecologia, história de vida e *habitat* das espécies parecem explicar pouco as variações nas estruturas filogeográficas no gradiente latitudinal. Além disso, as análises de filogeografia comparativa indicaram grandes regiões geográficas e que centenas de espécies podem mostrar conexões entre processos de nível populacional e padrões de riqueza de espécies.

Estudos com abordagens filogeográficas tem mostrado importância no entendimento de padrões de diversificação das espécies em diversos biomas. especialmente, nas áreas abertas da América do Sul. O assunto foi amplamente revisado por Werneck (2011), abordando basicamente os biomas Cerrado, Caatinga e Chaco. A revisão aborda as hipóteses

biogeográficas previamente propostas e discute as perspectivas futuras para novos estudos. Entender como as transformações geológicas e climáticas e as características espécie-específicas afetam a distribuição da biodiversidade é essencial para a conservação das espécies.

O gênero *Elaenia*

A família Tyrannidae é uma das famílias mais diversas de aves e possui representantes, dentre seus 98 gêneros, amplamente distribuídos nas Américas. Os indivíduos pertencentes a esta família, englobam uma variedade de comportamentos ecológicos, formas de reprodução, hábitos alimentares e ocupam diferentes tipos de *habitat*, dentre áreas abertas e savanas, tundra alpina e ambientes costeiros, e toda a gama de *habitats* florestais. (SICK, 1997; WINKLER; BILLERMAN; LOVETTE, 2020).

O gênero *Elaenia* é um dos mais diversos da família Tyrannidae com 21 espécies (RHEINDT; CHRISTIDIS; NORMAN, 2009; WINKLER; BILLERMAN; LOVETTE, 2020). As espécies possuem aspecto morfológico uniforme, o que dificulta a identificação apenas por avistamento. Podem apresentar comportamento migratório, parcialmente migratório e residente, com variações até mesmo dentro da própria espécie. São consideradas generalistas ecológicas, alimentam-se basicamente de insetos e frutas, e podem ocupar os mais diversos tipos de *habitats* (GUARALDO et al., 2021; HOSNER, 2020a, 2020b; HOSNER et al., 2020; HOSNER; KIRWAN, 2020; RHEINDT; CHRISTIDIS; NORMAN, 2008).

A alimentação das espécies do gênero *Elaenia* é composta em sua maioria por frutos e artrópodes (HOSNER, 2004; SICK, 1997). Marini e Cavalcanti (1998) encontraram, por meio de análises de conteúdo estomacal e revisão da literatura, que pelo menos cinco espécies do gênero são altamente frugívoras. Uma comparação da dieta em espécies de *E. cristata* e *E. chiriquensis* ao longo do ciclo anual mostrou que os indivíduos podem mudar a composição da alimentação entre períodos de muda (troca de penas)/ reprodução e o período não-reprodutivo, oscilando entre a ingestão de maior quantidade de artrópodes ou frutos (GUARALDO; KELLY; MARINI, 2016). Os autores destacam que as espécies apresentam comportamento similar de plasticidade durante a busca por alimentos durante o ciclo anual, apesar da diferença no comportamento, sendo uma espécie residente e a outra migratória.

Espécies migratórias frugívoras podem ter grande impacto como dispersor de sementes entre os locais visitados durante os deslacamentos, como a *E. albiceps chilensis* (BRAVO; CUETO; GOROSITO, 2017). Os autores encontraram alto nível de conectividade entre a área reprodutiva na Patagônia e a áreas de invernada (ou descanso reprodutivo) em Cerrado e Mata

Atlântica, sugerindo que as áreas frequentadas podem sofrer impacto significativo pela presença ou ausência da espécie devido seu papel como dispersor de sementes (BRAVO; CUETO; GOROSITO, 2017).

O comportamento migratório/ residente das espécies do gênero *Elaenia* é diverso. Análises de reconstrução do comportamento ancestral das espécies do gênero, por métodos de Parcimônia e Inferência Bayesiana, buscaram rastrear a origem da migração austral e o seu papel na diversificação do gênero. Os resultados sugerem um ancestral sedentário para o gênero e o comportamento migratório e migratório parcial surgindo várias vezes independentemente entre as espécies (RHEINDT; CHRISTIDIS; NORMAN, 2008).

Marini e Cavalcanti (1990) exploraram os padrões de distribuição e possíveis rotas migratórias das espécies *E. albiceps chilensis* e *E. chiriquensis albivertex* utilizando dados da literatura, observações de campo e espécimes de coleções científicas. As análises indicaram pelo menos duas rotas migratórias para *E. a. chilensis*: (i) sentido Norte-Sul e Sul-Norte ao longo dos Andes e (ii) da Argentina até Amazônia brasileira, com possível retorno pelo Brasil Central. Para *E. c. albivertex* não foi constatada rota migratória bem definida, mas os autores descrevem evidências de populações residentes (região Centro-Oeste e Amazônica) e migratórias (áreas ao sul de Brasília) no Brasil.

Dada a importância do conhecimento envolvido no comportamento migratório, estudos utilizando geolocalizador avaliaram o grau de conectividade entre áreas de invernada e de reprodução para a espécie *E. a. chilensis*, uma migrante de longa distância (BRAVO; CUETO; GOROSITO, 2017). Para essa espécie, foi encontrada alto grau de conectividade entre essas áreas, nas florestas da Patagônia (reprodução) e florestas de Mata Atlântica e Cerrado (áreas de invernada), portanto, mudanças em algum dos biomas que influenciem a migração da espécie podem impactar diretamente nos outros biomas envolvidos. Entretanto, esse tipo de conhecimento é escasso para outras espécies do gênero.

Espécies modelo

Dentre as espécies deste gênero, particularmente duas, ocorrem em simpatria nas áreas abertas da América do Sul, são elas: Plain-crested Elaenia (*E. cristata*) e Lesser Elaenia (*E. chiriquensis*) (HOSNER, 2020b; HOSNER et al., 2020).

A espécie *Elaenia cristata* (Pelzeln, 1868) (Figura 1a), comumente conhecida como Guaracava – de – topete – uniforme, ocupa predominantemente cerrado *sensu stricto*, geralmente ocorre em altitudes acima de 1500m , possui comportamento majoritariamente

residente, apesar de relatos de algumas populações com comportamento de deslocamentos curtos (MARINI et al., 2009; MEDEIROS; MARINI, 2007; RHEINDT; CHRISTIDIS; NORMAN, 2008; SICK, 1997).

Esta espécie possui duas subespécies reconhecidas: *E. c. cristata* (Pelzenln, 1868), com ocorrência na Venezuela, leste dos Andes, Guianas, regiões norte, central e leste do Brasil e no extremo leste da Bolívia. Há relatos de populações isoladas no sudeste da Colômbia, sudeste do Peru, nas regiões oeste e norte da Bolívia e no sudoeste da Amazônia brasileira (HOSNER, 2020b).

A subespécie *E. c. alticola* (J. T. Zimmer & Phelps, 1946), teria sua ocorrência apenas na região de teipus no sudeste da Venezuela e regiões adjacentes no norte do Brasil (HOSNER, 2020b).

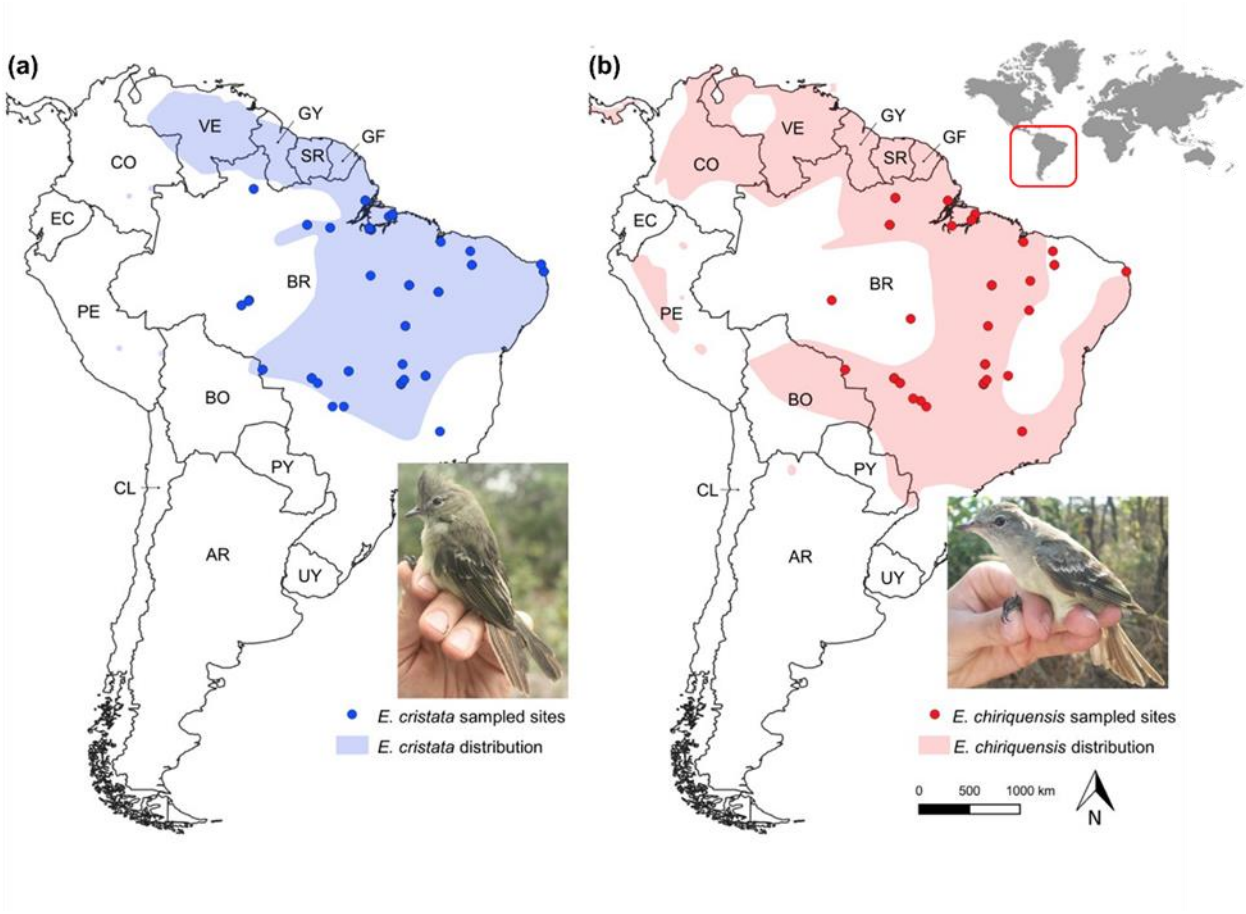


Figura 1. Áreas de distribuição geográfica de *Elaenia cristata* (a, azul) e *E. chiriquensis* (b, vermelho) baseado nos dados da International Union for Conservation of Nature (IUNC) retirado de Freitas et al (2022). Áreas amostradas são representadas por círculos coloridos para cada espécie. Fotografias: (a) *E. cristata* capturada em Parque Nacional da Chapada dos Guimarães, Mato Grosso – BR; (b) *E. chiriquensis* amostrada na Área de Proteção Gama e Cabeça de Veado, Distrito Federal – BR.

A espécie *E. chiriquensis* (Lawrence, 1865) (Figura 1b), de nome popular Chibum, possui registros de *habitat* em mata de galeria, cerrado típico e campo sujo, mas ocupa predominantemente áreas de cerrado *sensu stricto*, vegetação seca, margens de rio, áreas gramadas com arbustos dispersos, plantações, acima de 2000m (>3000m na Venezuela), possui comportamento migratório e autores relatam existência de populações residentes (DEL HOYO et al., 2019; HOSNER, 2020b; MARINI; CAVALCANTI, 1990; MEDEIROS; MARINI, 2007).

Para *E. chiriquensis* são reconhecidas atualmente também duas subespécies, a *E. c. chiriquensis* e *E. c. albivertex*. A subespécie *E. c. chiriquensis* (Lawrence, 1865) ocorre na zona tropical do sudoeste da Costa Rica e declive do Pacífico no oeste do Panamá e leste para a Zona do Canal, incluindo Ilha de Coiba e Ilhas de Pearl (HOSNER, 2020b).

A subespécie *E. c. albivertex* (Pelzeln, 1868), tem sua ocorrência nas regiões norte, central e leste da Colômbia, Venezuela, Trinidad e Guianas. No Brasil, localiza-se nas regiões norte, leste e central. Ocorre também no leste do Peru, no norte e leste da Bolívia e Paraguai e no nordeste da Argentina (HOSNER, 2020b).

Havia uma terceira subespécie em *E. chiriquensis*, a *E. chiriquensis brachyptera*. Porém, Rheindt e colaboradores (2015), utilizando metodologias de bioacústica e marcadores moleculares mitocondriais e nucleares, relataram caso de espécie críptica em *E. chiriquensis*. Apesar das semelhanças morfológicas, os resultados com base nessas análises mostraram diferenças suficientes entre essa espécie e as outras duas subespécies e propuseram a elevação da subespécie *E. chiriquensis brachyptera* ao nível de espécie (HOSNER, 2020b; RHEINDT et al., 2015).

O gênero *Elaenia* foi modelo de um amplo estudo filogenético. Rheindt e colaboradores (2008) tiveram como objetivo, além de conhecer mais sobre a filogenia do gênero, também investigar aspectos evolutivos relacionados à evolução de preferência de *habitat* e do comportamento migratório/residente das espécies. Segundo resultados do trabalho, *E. cristata* e *E. chiriquensis* constituem linhagens distintas, tendo o evento de especiação da primeira ocorrido há mais ou menos 8 milhões de anos e de *E. chiriquensis* por volta de 4,2 milhões de anos.

Estudos utilizando as espécies deste gênero têm-se focado em sua maioria sobre questões ecológicas da espécie, como comportamento reprodutivo (MARINI et al., 2009; MEDEIROS; MARINI, 2007), citogenética (KRETSCHMER et al., 2015) e, comportamento migratório/residente (GUARALDO; KELLY; MARINI, 2016).

O número de estudos sobre os efeitos que os eventos e processos históricos podem ter causado na atual distribuição das espécies tem aumentado cada vez mais na região neotropical (MARTINS; DOMINGUES, 2012). O conjunto de características, como a ampla distribuição e ocorrência em simpatria, as semelhanças morfológicas e diferentes hábitos migratórios, podem fornecer conhecimento além das espécies, mas também, sobre padrões de diversificação de outras espécies e sobre a influência dos processos ocorridos em áreas abertas. Com esse conhecimento, podemos elaborar estratégias para conservação com embasamento empiríco e levantar novas hipóteses sobre a evolução de espécies em áreas abertas, com destaque para a América do Sul.

OBJETIVO GERAL

Investigar os processos ecológicos e históricos relacionados à diversificação da avifauna associada às formações de áreas abertas da América do Sul utilizando, como modelo, duas espécies do genêro *Elaenia*: *E. cristata* e *E. chiriquensis*.

OBJETIVOS ESPECÍFICOS

(i) Avaliar o efeito da migração intratropical e o comportamento residente na estruturação genética das populações de duas espécies de aves Neotropicais: chibum (*E. chiriquensis*) e guaracava-de-topete-uniforme (*E. cristata*).

(ii) Avaliar se as flutuações climáticas ocorridas no Pleistoceno podem ter influenciado os padrões filogeográficos e a dinâmica populacional das espécies chibum (*E. chiriquensis*) e guaracava-de-topete-uniforme (*E. cristata*).

REFERÊNCIAS BIBLIOGRÁFICAS

AVISE, J. C. et al. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. **Annual Review of Ecology and Systematics**, v. 18, n. 1, p. 489–522, 1987.

AVISE, J. C. Phylogeography: retrospect and prospect. **Journal of Biogeography**, v. 36, n. 1, p. 3–15, jan. 2009.

BRAVO, S. P.; CUETO, V. R.; GOROSITO, C. A. Migratory timing, rate, routes and wintering areas of White-crested Elaenia (*Elaenia albiceps chilensis*), a key seed disperser for Patagonian forest regeneration. **PLOS ONE**, v. 12, n. 2, p. e0170188, 9 fev. 2017.

BURNEY, C. W.; BRUMFIELD, R. T. Ecology Predicts Levels of Genetic Differentiation in Neotropical Birds. **The American Naturalist**, v. 174, n. 3, p. 358–368, set. 2009.

CALDERÓN, L. et al. Phylogeography and genetic structure of two Patagonian shag species (Aves: Phalacrocoracidae). **Molecular Phylogenetics and Evolution**, v. 72, n. 1, p. 42–53, 2014.

CURNOW, R. N.; WRIGHT, S. Evolution and the Genetics of Populations, Volume 4: Variability within and among Natural Populations. **Biometrics**, v. 35, n. 1, p. 359, 1978.

DEL HOYO, J. et al. **Handbook of the Birds of the World Alive**. Barcelona: Lynx Edicions, 2019.

EDWARDS, S. V.; SHULTZ, A. J.; CAMPBELL-STATON, S. C. Next-generation sequencing and the expanding domain of phylogeography. **Folia Zoologica**, v. 64, n. 3, p. 187–206, nov. 2015.

GÓMEZ-BAHAMÓN, V. et al. Speciation Associated with Shifts in Migratory Behavior in an Avian Radiation. **Current Biology**, v. 30, n. 7, p. 1312- 1321.e6, abr. 2020.

GUARALDO, A. C. et al. Longitudinal and cyclic poleward migration of a South American intra-tropical migrant flycatcher, the Lesser Elaenia (*Elaenia chiriquensis*). **Ornitologia Neotropical**, v. 32, n. 2, p. 94–102, 2021.

GUARALDO, A. C.; KELLY, J. F.; MARINI, M. Â. Contrasting annual cycles of an intratropical migrant and a tropical resident bird. **Journal of Ornithology**, v. 157, n. 3, p. 695–705, 2016.

GUTIÉRREZ-PINTO, N. et al. Non-monophyly and deep genetic differentiation across low-elevation barriers in a Neotropical montane bird (*Basileuterus tristriatus*; Aves: Parulidae). **Molecular Phylogenetics and Evolution**, v. 64, n. 1, p. 156–165, jul. 2012.

HICKERSON, M. J. et al. Phylogeography's past, present, and future: 10 years after Avise, 2000. **Molecular Phylogenetics and Evolution**, v. 54, n. 1, p. 291–301, jan. 2010.

HOSNER, P. Rufous-crowned Elaenia (*Elaenia ruficeps*). In: DEL HOYO, J. et al. (Eds.). . **Birds of the World**. [s.l.] Cornell Lab of Ornithology, 2020a.

HOSNER, P. Plain-crested Elaenia (*Elaenia cristata*). In: DEL HOYO, J. et al. (Eds.). . **Birds of the World**. version 1. ed. Ithaca, NY, USA: Cornell Lab of Ornithology, 2020b.

HOSNER, P. et al. Lesser Elaenia (*Elaenia chiriquensis*). In: DEL HOYO, J. et al. (Eds.). . **Birds of the World**. Version1. ed. Ithaca, NY, USA: Cornell Lab of Ornithology, 2020.

HOSNER, P. A. Genus *Elaenia*. In: DEL HOYO, J.; ELLIOTT, A.; CHRISTIE, D. A. (Eds.). . **Handbook of the Birds of the World. Vol. 9. Cotingas to Pipits and Wagtails**. Eds. ed. Barcelona: Lynx Edicions, 2004. p. 267–274.

HOSNER, P.; KIRWAN, G. M. Large Elaenia (*Elaenia spectabilis*). In: DEL HOYO, J. et al. (Eds.). . **Birds of the World**. [s.l.] Cornell Lab of Ornithology, 2020.

HUDSON, R. R. Gene genealogies and the coalescent process. **Oxford Surveys in Evolutionary Biology**, v. 7, p. 1–44, 1991.

KINGMAN, J. F. C. The coalescent. **Stochastic Processes and their Applications**, v. 13, n. 3, p. 235–248, 1982.

KINGMAN, J. F. C. Origins of the Coalescent: 1974–1982. **Genetics Society of America**, v. 156, n. 1, p. 1461–1463, 2000.

KNOWLES, L. L. Statistical Phylogeography. **Annual Review of Ecology, Evolution, and Systematics**, v. 40, n. 1, p. 593–612, dez. 2009.

KNOWLES, L. L.; MADDISON, W. P. Statistical phylogeography. **Molecular Ecology**, v. 11, n. 12, p. 2623–2635, 2002.

KRETSCHMER, R. et al. Chromosome mapping of the large elaenia (*Elaenia spectabilis*): Evidence for a cytogenetic signature for passeriform birds? **Biological Journal of the Linnean Society**, v. 115, n. 2, p. 391–398, 2015.

KUMAR, R.; KUMAR, V. A review of phylogeography: biotic and abiotic factors. **Geology, Ecology, and Landscapes**, v. 2, n. 4, p. 268–274, 2 out. 2018.

MAIRAL, M. et al. Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. **Scientific Reports**, v. 7, n. 1, p. 45749, 2017.

MARINI, M. et al. Biologia reprodutiva de *Elaenia cristata* (Aves: Tyrannidae) em cerrado do Brasil central. **Neotropical Biology and Conservation**, v. 4, n. 1, p. 2–12, 2009.

MARINI, M. Â.; CAVALCANTI, R. B. Frugivory by *Elaenia* Flycatchers. **El Hornero**, v. 15, n. 1, p. 47–50, 1998.

MARINI, M.; CAVALCANTI, R. Migrações de *Elaenia albiceps chilensis* e *Elaenia chiriquensis albivertex* (Aves: Tyrannidae). **Boletim do Museu Paraense Emilio Goeldi, série Zoologia**, p. 59, 1 jan. 1990.

MARTINS, F. DE M.; DOMINGUES, M. V. Filogeografia. In: CARVALHO, C. J. B. DE; ALMEIDA, E. A. B. (Eds.). . **Biogeografia da América do Sul - Padrões & Processos**. 1ª ed. São Paulo: Roca, 2012. p. 137–150.

MATOS, M. V. et al. Comparative Phylogeography of Two Bird Species, *Tachyphonus phoenicius* (Thraupidae) and *Polytmus theresiae* (Trochilidae), Specialized in Amazonian White-sand Vegetation. **Biotropica**, v. 48, n. 1, p. 110–120, 2016.

MEDEIROS, R. DE C. S. DE; MARINI, M. Â. Biologia reprodutiva de *Elaenia chiriquensis* (Lawrence) (Aves, Tyrannidae) em Cerrado do Brasil Central. **Revista Brasileira de Zoologia**, v. 24, n. 1, p. 12–20, 2007.

OLIVEIRA, U.; VASCONCELOS, M. F.; SANTOS, A. J. Biogeography of Amazon birds: rivers limit species composition, but not areas of endemism. **Scientific Reports**, v. 7, n. 1, p. 2992, 7 dez. 2017.

RHEINDT, F. E. et al. Cryptic speciation in the Lesser Elaenia *Elaenia chiriquensis* (Aves: Passeriformes: Tyrannidae). **Zootaxa**, v. 4032, n. 3, p. 251, 16 out. 2015.

RHEINDT, F. E.; CHRISTIDIS, L.; NORMAN, J. A. Habitat shifts in the evolutionary history of a Neotropical flycatcher lineage from forest and open landscapes. **BMC Evolutionary Biology**, v. 8, n. 1, p. 193, 2008.

RHEINDT, F. E.; CHRISTIDIS, L.; NORMAN, J. A. Genetic introgression, incomplete lineage sorting and faulty taxonomy create multiple cases of polyphyly in a montane clade of tyrant-flycatchers (*Elaenia* , Tyrannidae). **Zoologica Scripta**, v. 38, n. 2, p. 143–153, mar. 2009.

SICK, H. **Ornitologia Brasileira**. Rio de Janeiro: Nova Fronteira, 1997.

SMITH, B. T. et al. The drivers of tropical speciation. **Nature**, v. 515, n. 7527, p. 406–409, 2014.

SMITH, B. T. et al. A latitudinal phylogeographic diversity gradient in birds. **PLOS Biology**, v. 15, n. 4, p. e2001073, 13 abr. 2017.

WERNECK, F. P. The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. **Quaternary Science Reviews**, v. 30, n. 13–14, p. 1630–1648, jun. 2011.

WINKLER, D. W.; BILLERMAN, S. M.; LOVETTE, I. J. Tyrant Flycatchers (Tyrannidae). In: BILLERMAN, S. M. et al. (Eds.). . **Birds of the World**. [s.l.] Cornell Lab of Ornithology, 2020.

CAPÍTULO 1

Obs.: este capítulo foi publicado recentemente na *Journal of Avian Biology* e encontra-se no apêndice desta tese.

Ecological traits drive genetic structuring in two open-habitat birds from the morphologically cryptic genus *Elaenia* (Aves: Tyrannidae)

ABSTRACT

Understanding the relative contributions of the many factors that shape population genetic structuring is a central theme in evolutionary and conservation biology. Historically, abiotic or extrinsic factors (such as geographic barriers or climatic shifts) have received greater attention than biotic or intrinsic factors (such as dispersal or migration). This focus stems in part from the logistical difficulties in taking a comparative phylogeographic approach that contrasts species that have experienced similar abiotic conditions during their evolution yet differ in the intrinsic attributes that might shape their genetic structure. To explore the effects of intratropical migration on the genetic structuring of Neotropical birds, we chose two congeneric species, the Lesser Elaenia (*Elaenia chiriquensis*) and the Plain-crested Elaenia (*E. cristata*), that are largely sympatric, and which have similar plumage, habitat preferences, and breeding phenology. Despite these many commonalities, they differ in migratory behavior: *E. chiriquensis* is an intratropical migratory species while *E. cristata* is sedentary. We used a reduced representation genomic approach to test whether migratory behavior is associated with increased gene flow and therefore lower genetic population structure. As predicted, we found notably stronger genetic structuring in the sedentary species than in the migratory one. *E. cristata* comprises genetic clusters with geographic correspondence throughout its distribution, while there are no geographic groups within Brazil for *E. chiriquensis*. This comparison adds to the growing evidence about how intrinsic traits like migration can shape the genetic structuring of birds and advances our understanding of the diversification patterns of the understudied, open habitat species from South America.

KEYWORDS

ddRAD-Seq, ecological traits, Flycatchers, genetic structure, migratory-sedentary behavior, Neotropics

INTRODUCTION

How different factors contribute to shaping the genetic structure of populations is a central question in evolutionary and conservation biology (Foll and Gaggiotti 2006). The genetic structure of populations is determined by the combined effect of evolutionary forces acting within a population (e.g., mutation, genetic drift, and selection) and by the mixing of genetic variation among populations through gene flow (Curnow and Wright 1978). The intensity of gene flow is influenced by a combination of abiotic or extrinsic factors (such as geographic barriers, geological or climatic shifts) (Mairal et al. 2017, Campillo et al. 2020) and/or biotic or intrinsic traits (such as reproductive behaviors, dispersal, migration, and adaptation to local environmental conditions) (e.g. Burney and Brumfield 2009, Calderón et al. 2014, Smith et al. 2014, Nistelberger et al. 2015).

Phylogeographic studies have traditionally focused more on extrinsic factors, as addressing the effects of biotic factors is often challenging. One method of exploring the genetic effects of intrinsic traits involves comparing species that have experienced similar abiotic conditions during their evolution while diverging in the intrinsic attributes that are hypothesized to have influenced their genetic structuring (Papadopoulou and Knowles 2016).

In birds, migration is an intrinsic trait known to influence the spatial distribution of genetic variation (e.g., Arguedas and Parker 2000, Clegg et al. 2003, Irwin et al. 2011, Contina et al. 2019). Migration can be defined operationally as the short to large-scale, cyclic, seasonal movement of a population between breeding and non-breeding areas (Newton 2003). The level of genetic structure in migratory species (as well as in sedentary species) may be influenced by natal philopatry, which describes the tendency of individuals to return to their natal breeding grounds to reproduce (Weatherhead and Forbes 1994). High natal philopatry is expected to restrict gene flow among migrant groups, promoting stronger genetic structuring, whereas low natal philopatry facilitates gene flow, resulting in lower genetic structuring. Particularly in birds, studies have shown that natal philopatry is generally higher in sedentary populations compared to those that are migratory (Weatherhead and Forbes 1994). Furthermore, even when high natal philopatry was observed in migratory species, this occurred mostly in isolated populations on islands, and was hypothesized to be the product of a local adaptation (Wright and Mauck 1998, Förschler et al. 2010).

Most research on migratory birds has focused on species that breed in northern latitudes in North America (Nearctic migrants) and Eurasia (Palearctic migrants), and which migrate south during the Northern Hemisphere winter. Less work has been done on the “Austral

migrant" species of the Southern Hemisphere that migrate north during the Southern Hemisphere winter, and even fewer studies have addressed species that perform annual latitudinal movements on a smaller scale within the tropics, a phenomenon termed "Intratropical migration" (Hayes 1995, Jahn et al. 2020). Although there are more than 200 species of austral and intratropical migrants (Chesser 1994, Stotz et al. 1996, Jahn et al. 2006, 2020) in the Neotropical region, little is currently known about these migratory processes and their evolutionary consequences (Faaborg et al. 2010, Jahn et al. 2020).

Migratory behaviors often vary even among species of the same genus, as is the case in *Elaenia* flycatchers, a genus of 21 species occurring across Central and South America and the Caribbean. In this genus, some species are austral migrants (e.g., *E. chilensis and E. parvirostris*), others are intratropical migrants (e.g., *E. chiriquensis*), and others are sedentary (e.g., *E. cristata* and *E. obscura*) (Marini et al. 2009, Guaraldo et al. 2016, Somenzari et al. 2018). The Lesser Elaenia (*E. chiriquensis*) and the Plain-crested Elaenia (*E. cristata*) are sympatric across most of their geographic distributions (Fig. 1). They have similar preferences for breeding and foraging habitat (Cerrado *sensu stricto*, highly seasonal savanna), display similar and monomorphic plumage, and are mainly frugivorous (Hosner 2020, Hosner et al. 2020). Molecular phylogenetic studies show that these species represent independent lineages that are easily distinguished genetically (Rheindt et al. 2008). These species do not hybridize nor show high levels of incomplete lineage sorting, as is the case for other species pairs in the same genus (Rheindt et al. 2008, Tang et al. 2018).

The most notable difference between these two *Elaenia* species is in their sedentary/migratory behavior. *E. cristata* lives in open savanna and can be found throughout the year across its entire distribution (Sick 1997, Hosner 2020). In contrast, *E. chiriquensis* can be found throughout the year in some smaller enclaves of savannas surrounded by forest in northern South America (hereafter only northern savannas) and increases its abundance in the center of the Cerrado (the tropical Brazilian savanna) during the breeding season between August and December (Medeiros and Marini 2007, Hosner et al. 2020). After breeding in the central Cerrado, *E. chiriquensis* moves north likely to the Amazonian region (Marini and Cavalcanti 1990). However, the exact migratory routes and the possible existence of resident populations in this species remain unclear.

Technological advances in massive parallel sequencing approaches have enabled fast and low-cost access to a high number of molecular markers, in a large number of individuals (Edwards et al. 2015, Goodwin et al. 2016), making it possible to analyze the genetic structure of populations more robustly (Lavretsky et al. 2019), including that of migratory birds (e.g.,

Kraus et al. 2011, 2013, Jonker et al. 2012, Ruegg et al. 2014, DeSaix et al. 2019, Delmore et al. 2020). Single nucleotide polymorphisms (SNP) are the main molecular marker used in studies that evaluate the genetic structure and levels of gene flow between populations of one or more species (e.g. Hohenlohe et al. 2010, Kopuchian et al. 2020), or that identify distinct groups within a migratory species (Kraus et al. 2013, Ruegg et al. 2014).

Here we used a reduced representation genomic approach (ddRAD-Seq) to study the effect of intratropical migration and sedentary behavior on the genetic structure of two Neotropical birds: the migrant (*E. chiriquensis*) and the sedentary (*E. cristata*). Our central hypothesis is that migratory behavior with lower natal philopatry will result in higher gene flow among populations, and therefore lower population structure in the migratory species compared to the sedentary one.

MATERIAL AND METHODS

Species distribution and tissue sampling

We sampled 218 specimens (*E. cristata*, n=98; *E. chiriquensis*, n=120, but see the Results section for information about misidentifications) from 2003 to 2018 across 33 sites in South America (Supplementary Table S1, Summarized in Table 1; Fig. 1). Specimens were captured in the field using mist-nest and banded to avoid duplicate sampling. Approximately 20 µl of blood was obtained from each individual using sterile needles and glass capillary tubes and stored in absolute ethanol at room temperature.

The specimens were identified in the field using slight morphological differences (Hosner 2020, Hosner et al. 2020) or through vocalization whenever possible. Briefly, the diagnostic characters for *E. cristata* are mainly an elongated crown of feathers building a conspicuous crest, without a white coronal patch, and two broad, well-marked wing bars (see Fig. 1a). While *E. chiriquensis* exhibits slightly elongated crown feathers forming a less pronounced crest with a variably sized white coronal patch occasionally hidden, and it also has two broad wing bars and whitish edges on the remiges (see Fig. 1b). Vocalizations were easy to distinguish by ear between species and when the captured individual vocalized, we also compared this vocalization with recorded vocalizations for the species available in databases such as eBird (https://ebird.org) or Xeno-canto (https://www.xeno-canto.org).

To increase the geographic coverage of our sampling, we obtained 108 *Elaenia* tissue samples from ornithological collections for a total of 326 samples (Table S1 in Supporting information).

In both species, most of the individuals (89% for *E. chiriquensis*, and 79% for *E. cristata*) were sampled during the breeding season (August to February), usually during the same expedition in each site (Fig. S1 in Supporting information).

DNA extraction and quality control

Total genomic DNA was extracted following a phenol, chloroform, isoamyl alcohol protocol (as in Friesen et al. 1997), or using the PureLink™ Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, US) (for museum tissue samples), following manufacturer instructions. Genomic DNA quality and concentrations were verified on a 1% agarose gel stained with ethidium bromide and using the Qubit™ dsDNA BR Assay Kit (ThermoFisher, Waltham, MA, US), respectively.

Species identification through mitochondrial DNA

As the two *Elaenia* species analyzed are challenging to distinguish morphologically, we confirmed the field identifications using diagnostic variation in the mitochondrial gene NADH dehydrogenase subunit-2 (ND2). Polymerase Chain Reaction (PCR) amplifications were performed with the primer pair LMET and H6313 (Sorenson et al. 1999) in a final volume of 20 µl using the Phusion High-Fidelity PCR Kit (New England BioLabs, Ipswich, Massachusetts, US), 10 mM each dNTP, 10 µM of each primer (forward and reverse), 0.4U of Taq Polymerase, and 30-80 ng of genomic DNA. The temperature cycling involved the first step at 98 °C for 30 seconds, followed by 30 cycles of 10 s at 98 °C, 30 s at 54 °C, and 30 s at 72 °C, followed by a final extension step of 72 °C for 10 min.

PCR products were then treated with Exonuclease (EXO) and Shrimp Alkaline Phosphatase (SAP) (ThermoFisher) in a final volume of 1 µl with 10U/µl of EXO and 1.0U/µl of SAP per 10 µl of PCR product, heated in a thermocycler at 37ºC for 30 minutes, then held at 90ºC for 10 minutes. PCR products were then sequenced at the Cornell University Biotechnology Resource Center Genomics Facility and MACROGEN Inc. Sequencing results were verified using Geneious Prime 2019.0.4 (https://www.geneious.com). The species were identified by comparing the ND2 sequence obtained from each sample with the database nucleotide collection (GenBank) using a Standard Nucleotide BLAST- blastn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) optimized for highly similar sequences (Megablast) (Morgulis et al. 2008).

ddRADseq dataset

We generated ddRADseq loci following the approach outlined by Peterson et al. (2012) with modifications as described by Thrasher et al. (2018). Briefly, we digested each sample with *SbfI* and *MspI* and ligated adapters that allowed multiplexing. The libraries, each containing approximately 20 samples, were size-selected and PCR-enriched, incorporating the Illumina HiSeq adapters (Illumina, San Diego, California, US). Finally, all groups of samples were combined in equimolar proportions and sequenced, single end 100bp, on two lanes of an Illumina HiSeq 2500.

After assessing read quality with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), we used FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit) to trim sequences to 98 bp to discard lower-quality base calls at the 3′ end of the sequence. Subsequently, we used FASTX-Toolkit to retain reads without a single base below a Phred quality score of 10 and with at least 95% of bases with quality above 20. We demultiplexed reads using the 'process_radtags' program from the STACKS v.2.41 bioinformatics pipeline (Catchen et al. 2011, 2013), discarding reads that did not pass the Illumina filter, had barcode contamination, lacked an *SbfI* cut site or one of the unique barcodes used for multiplexing at the 5′ end. We obtained an average of 550,757 (±263,678 reads) quality-filtered reads per individual (Supplementary Table S1).

We assembled the reads from both species into RADseq loci using the *de novo* pipeline from STACKS. We conducted a sensitivity analysis by testing different values for coverage (m= 5, 10, 20, and 30) as suggested by Rochette and Catchen (2017) yet did not find substantial differences in the number of loci recovered. Parameters were therefore set to a minimum coverage of 5 (m), up to seven differences between alleles of the same locus (M), and seven differences among aligned loci of different individuals (n). This combination of parameters produced an average coverage per locus ranging from 11.24 – 68.55x, with an overall average of 29.61x (±9.2). We exported SNPs using the program 'populations' in STACKS for all the samples combined (n=326) and again for each species separately, *E cristata* (n=156) and *E. chiriquensis* (n=150). We retained loci that were present in at least 80% of the individuals and exported both one SNP per RADSeq locus (to avoid including linked loci) and all SNP per RADSeq loci.

Principal Component Analysis

We conducted a Principal Component Analysis of the SNP dataset using the package SNPRelate (Zheng et al. 2012) in R (R Core Team 2019) to assess possible groupings among individuals. For this analysis, we exported SNPs using 'populations' from STACKS in four different ways: (i) for all the specimens (n=326 and 5,938 SNPs); (ii) removing both all the specimens misidentified in the field (see Results) and samples with more than 80% of missing data for all loci (n=306 and 6,458 SNPs); (iii) only with *E. cristata* individuals (n=156 and 6,447 SNPs); and (iv) only *E. chiriquensis* samples (n=150 and 9,671 SNPs).

Phylogenetic analyses

We built a phylogenetic tree using RAxML v.8.2.9 (Stamatakis 2014) for both species using the SNP dataset in Phylip format (variant sites only), and the following parameters: 500 replicates of rapid bootstrap analysis, and the *ASC_GTRGAMMA* model with the Lewis correction for ascertainment bias.

Population genetic structure

We performed Structure analyses using the SNP datasets obtained for each species separately and one SNP per locus to avoid the effect of linkage (2,814 SNPs for *E. cristata,* and 2,765 for *E. chiriquensis*). We conducted analyses at K values ranging from 1 to 10 for *E. cristata* and 1 to 5 for *E. chiriquensis*, with 10 replicate runs at each value. We also performed a Structure analysis for a subsample of *E. cristata* from Central Brazil (localities, 8,11, 12, 14, 17, 20, 23, 26, 27, 28, 29, 30, 31, and 33; Fig. 3c), with K values ranging from 1 to 6. Each run included 500,000 iterations of burn-in followed by 1,000,000 sampling iterations. We estimated the best value of K using the method of Evanno et al. (2005). Using the groups from the Structure results, we calculated F_{ST} among populations within *E. cristata* using the 'populations' module of STACKS (parameter 'fststats'). Due to the weak genetic structure observed in *E. chiriquensis*, we did not estimate the F_{ST} values for this species as we did not have clearly defined populations.

The level of intraspecific genetic structure was also assessed using haplotype information in fineRAD Structure v.0.3 and RADpainter (Lawson et al. 2012, Malinsky et al. 2018). Briefly, this algorithm works in four steps: (i) it calculates the co-ancestry matrix using the SNP dataset, (ii) clusters individuals based on the co-ancestry matrix, (iii) builds a dendrogram

tree; and (iv) plots results using the fineRADstructurePlot.R script in R (R Core Team 2019). Haplotype datasets were produced with 'populations' in STACKS, for each species separately, and without a filter for minor allele frequency.

RESULTS

Elaenia species identification

Of the 326 specimens analyzed in our study (including samples from both museum collections and wild birds), 273 (83.74%) had a portion of the mitochondrial ND2 gene amplified successfully. From the 186 specimens morphologically identified by our team in the field and with successfully amplified ND2 sequences, 10 (5.4%) were misidentified to species. From the 87 tissue samples requested from ornithological collections with ND2 data, 18 (20.7%) were misidentified to species. Among these 28 misidentified specimens, 12 genetically identified *E. chiriquensis* individuals were previously identified as *E. cristata*, seven specimens previously identified as *E. cristata* were genetically *E. chiriquensis*, and nine have been genetically identified as other species of *Elaenia* and even another morphologically similar tyrant flycatcher (*Sublegatus arenarum*). While high, this field identification error rate is not surprising given the high phenotypic similarity of these species and the fact that samples were collected by many different people with varied levels of experience with these taxa. The PCA results based on the SNP data from all our samples and colored based on the mtDNA genetic identification show how errors can occur when using only morphological characters to identify species in the genus *Elaenia* (Fig. 2a, and Fig. S2a). We determined that this larger dataset included representatives of six *Elaenia* species (*E. cristata*, *E. chiriquensis*, *E. albiceps*, *E. flavogaster*, *E. parvirostris*, and *E. spectabilis*), and two samples from another flycatcher species (*Sublegatus arenarum*). Because of this rate of misidentification, we only included in our final genomic analyses the *Elaenia* samples (n= 33) for which we did not confirm the field identification with mitochondrial DNA (failed to be amplified) if they grouped diagnostically in the PCA with those with confirmed identifications (Fig. 2a).

There is no overlap between *E. cristata* and *E. chiriquensis* in the PCA based on the SNP dataset derived from these two species alone (Fig. 2b). In this analysis, the first two principal components explain 51.17% of the total variation. To explore patterns of intraspecific variation we ran subsequent analyses on samples from each of these species alone. The PCA from *E. cristata* (the sedentary species; Fig. 2c) explained 9.63% of the variation and included several distinct clusters in the space of the two first principal components, whereas the

equivalent plot for *E. chiriquensis* (the migratory species; Fig. 2d) explained only 2.08% of the total genetic variation and had no apparent clustering.

Genetic clusters and moderate structure in the sedentary species

The results from the PCA were consistent with those from the remaining analyses, where *E. cristata* and *E. chiriquensis* consistently showed distinct patterns of population structure. Structure results showed that in the sedentary species (*E. cristata*), the K values with the highest likelihood were 3 and 5, with the K = 5 pattern showing genetic clusters that can be explained geographically (Fig. 3a, Table S2, and Fig. S3a). The five clusters correspond to two groups in the northern savannas, one in the southwestern Amazon Forest, one in the extreme northeast of Brazil, and a larger group including all other collection locations from the Cerrado. The Structure results obtained when analyzing individuals from this last large group alone also indicated the possibility for finer population structure, with evidence for three additional clusters (Fig. 3a, and Table S3). The central portion (site 29, Fig. 3c) contains admixed individuals with the genetic composition of all three groups, while some localities such as Nova Xavantina and Emas National Park show less evidence of admixture (sites 26 and 27; Fig. 3c). We obtained similar results in our fineRAD Structure analysis (Fig. 3d), observing five co-ancestry groups with additional genetic sub-structuring within two of them. Finally, phylogenetic analysis with RAxML showed the unrooted *E. cristata* tree with four well-supported branches (bootstrap > 85) (Fig. S6). These branches corresponded to specific geographic regions shown in the map (Fig. 3b).

The mean pairwise F_{ST} among the five genetic populations resulting from Structure for *E. cristata* ranged from 0.028 to 0.109 (Table S5). The lowest values are derived from the comparisons between the population in the central region of the distribution and other populations (Fig. S5).

Few genetic clusters and low structure in the migratory species

In contrast to the results from the sedentary species, for the migratory Lesser Elaenia (*E. chiriquensis*), PCA analysis grouped samples into one cluster (Fig. 2d). Similarly, Structure results supported K = 2 (Fig. 4a, Table S4, Fig. S4) with most of our sampling localities belonging to a single genetic population, but with a discrete separation in a group around the Amazon rainforest, formed by individuals from Óbidos, Oxiriminá and Manicoré (sites 3, 4,

and 9; Fig. 4b). The fineRAD Structure results also recovered this group, and an additional one which included individuals from Ilha do Marajó (site 7; Fig. 4b). The RAxML analysis did not show any well-supported clades (Fig. S7).

DISCUSSION

Assuming that both *Elaenia* species studied here, which have highly overlapping ranges, have experienced a similar historical geographic and climatic context (extrinsic factors which could shape their population structure), our results suggest that their intrinsic sedentary/migratory behavior has influenced their genetic structure. As we predicted, the sedentary species showed greater genetic structuring compared to the migratory one, across a wide area of South America where both species are sympatric (see Fig. 1).

Patterns of genetic structure for *E. chiriquensis*

The low genetic structure found for migratory *E. chiriquensis* had already been preliminarily observed using other marker types (Bates et al. 2003, Rheindt et al. 2015) and suggests substantial recent gene flow throughout the species' range, which is primarily in South America. As described in previous studies (Marini and Cavalcanti 1990, Medeiros and Marini 2007, De Paiva and Marini 2013), the abundance of this species increases significantly between August to December in the south-central region of the distribution in areas with Cerrado *sensu stricto* vegetation, a habitat described as preferential for nest building by the species. After breeding, studies suggest that birds fly to areas in the north of the distribution, becoming absent in some regions in the southern Cerrado during the non-breeding period (Marini and Cavalcanti 1990).

Some authors have considered *E. chiriquensis* as partially migratory (see Somenzari et al. 2018) based on the absence of records in the central Cerrado between June and August (Marini and Cavalcanti 1990), while individuals have been observed in northern Brazil throughout the year. Intraspecific differences in migratory behavior among individuals within the species can lead to temporal and spatial reproductive asynchrony promoting divergence of neutral genetic variation between sedentary and migrant populations (Burney and Brumfield 2009). For instance, migratory behavior was recently identified as a driver for the diversification of subspecies of an austral migratory bird species (*Tyrannus savanna*, Gómez-Bahamón et al. 2020).

Although a high degree of gene flow was observed in the sampled migratory populations of Lesser Elaenia, we still observed a modest level of genetic structuring among the populations from the northern (site 7, Fig. 4c) and western portion of the sampled area (sites 3, 4, and 9, Fig. 4c). The cluster analyses showed two or three populations (Structure, and FineRAD Structure, respectively) in this periphery of the Amazonian region: one population was formed by sites 3, 4, and 9, and another population by site 7 (Fig. 4). However, some admixed individuals of these populations were also collected in the central Cerrado (see site 29 in Fig. 3a) during the breeding season. Even though our data suggests a weak evidence of structuring for *E. chiriquensis* in our sampling of migratory populations (Guaraldo et al. 2021), we can't rule out the possibility of potential structuring among the resident populations in the non-sampled regions of its distribution. Future studies should include individuals from populations of *E. chiriquensis* in regions of Central America to confirm if this pattern of low genetic structure applies to the entire species. The absence of migration in some populations can lead to a decrease in gene flow and, consequently, higher population structure, as seen in the congeneric *E. cristata*.

The generally low genetic structure observed in the migratory species is consistent with the expectation of low natal philopatry as described by Weatherhead and Forbes (1994). This type of pattern occurs when individuals in a given area migrate to multiple areas in successive reproductive seasons. Occurrence data of this species throughout the year suggest that migrants come from the north of the distribution, where records of the species presence are constant throughout the year (Marini and Cavalcanti 1990), and reproduce in different regions in the south-central portion of the distribution. Furthermore, isotopic data indicate that the migrant *E. chiriquensis* exhibits a niche-following behavior (it seeks similar resources throughout the annual cycle), suggesting that despite flying north, it spends the winter in savanna areas (Guaraldo et al. 2016). Similarly, our data suggest a connection between savanna patches in northern South America and the central Cerrado region (Fig. 4). In fact, the genetic differentiation of the migrant populations that occur in these savanna patches which are interlocked in a forest environment may be associated with greater natal philopatry due to their isolation. Isolated migrant populations of passerines often show considerably higher philopatry than non-isolated populations (Weatherhead and Forbes 1994, Wright and Mauck 1998).

Alternatively, the Amazon rainforest could be a partially effective geographic barrier for this species, limiting migration and consequently decreasing gene flow among the migrant populations. The congruence of the genetic clusters formed in the northern savannas and peri-Amazonian area in both *Elaenia* species studied suggests the presence of an extrinsic barrier

locally driving the genetic structure in these two species. In fact, these two factors (philopatry and a geographic barrier) are not mutually exclusive and may be acting together to shape the genetic structure of these species, and the relative contribution of each should be better investigated.

Genetic structuring in *E. cristata*

As expected, the sedentary species showed greater genetic structure, and we identified five genetic populations (Fig. 3; K = 5; pairwise F_{ST} ranges from 0.028 to 0.109): two of them occurred in patches of savannas nested in the Amazonian region and the other three in southwestern, central, and northeastern South America. The different populations of the north have a strong geographic association with the different savanna areas, indicating that the processes that led to the disjunct formation of these areas may have also influenced these avian population dynamics.

Despite observing greater genetic structuring in central and northeastern South America for this sedentary species, the magnitude of the genetic difference among these populations is small, consistent with recent isolation and/or some ongoing gene flow. One of these populations occurs in the northeast, in the Caatinga, the largest patch of the Seasonally Dry Tropical Forest of South America with a predominance of xeric vegetation (for details, see Werneck 2011), while the other populations inhabit the Cerrado. Genetically differentiated populations occurring in the different open vegetation biomes of South America have been identified in several organisms (Wuster et al. 2005, Ramos et al. 2007), including in other birds (Rocha et al. 2020).

Most previous studies of differentiation in Neotropical birds from open areas have found evidence for Pleistocene climatic oscillations as a driver of intraspecific divergence (Lima-Rezende et al. 2019b, Rocha et al. 2020, Ritter et al. 2021). For instance, the genetic differentiation of Narrow-billed Woodcreeper (*Lepidocolaptes angustirostris*) populations seems to have occurred in allopatry in stable areas that formed during Pleistocene climatic fluctuations (Rocha et al. 2020). Similarly, studies of other bird species that occur in the Cerrado have identified a weak intraspecific genetic structure, possibly due to the increase in gene flow between populations promoted by the expansion of climatically suitable areas for these species during the Pleistocene (Lima-Rezende et al. 2019a, Rocha et al. 2020).

In this context, it is important to highlight that our study assumes that the two species studied underwent similar historical demographic processes, but we can't completely rule out

that these species, despite being ecologically similar, did not experience different geographic or climatic contexts at different times in the past. For instance, these two bird species might have occurred in distinct stable areas during the Pleistocene climatic oscillations, but the current intense gene flow promoted by migratory behavior in *E. chiriquensis* may have erased the genetic signature of historical isolation among previously isolated populations.

Elaenia sp. misidentification

About 8.5% (n = 28) of the specimens collected for this study were misidentified in the field at the time of sample collection, which is not surprising given their morphological similarity (see Fig. 1). Most of the cases involved the two focal species, with some *E. cristata* identified in the field being genetically identified as *E. chiriquensis* and vice versa. In a few cases (n = 9), other species of *Elaenia* and even another morphologically similar tyrant flycatcher (*Sublegatus arenarum*) were misidentified as the two focal species of this study. This non-trivial field identification error rate underscores the difficulty with working in this challenging group of morphologically cryptic species. The genus *Elaenia* comprises 21 species that are all quite similar morphologically, and their misidentification has been widely reported (Traylor Jr 1982, Hosner 2004, Winkler et al. 2020). For instance, Rheindt et al. (2015), in a similar study using the ND2 mtDNA gene, also found one misidentified *Elaenia* sample (*E. c. albivertex* labeled as *E. flavogaster*) out of 13 samples obtained from ornithological collections.

In general, field identification of *Elaenia* species is based either on the birds' vocalizations or on the species' geographical distribution (Sick 1997). Criteria based on geographic distribution can be challenging when species are sympatric or when there are migratory species involved. *E. cristata* is thought to be partially migratory in some regions (Hosner 2020), such as in the Mato Grosso state in Brazil. We found 12 cases of misidentified *E. cristata* labeled as *E. chiriquensis*, three of them in Mato Grosso state and five in Pará state, localities that may be in the migratory routes of other species of *Elaenia*. Therefore, misidentification between sedentary and migratory *Elaenia* species can lead to the eventual misinterpretation of migratory behavior. Owing to the difficulty of using morphology to identify *Elaenia* species, we encourage using complementary species identification methods, such as DNA barcode approaches that have been successfully applied in the identification of many Tyrannidae species (Kerr et al. 2007, Chaves et al. 2008).

CONCLUSIONS

Our study adds evidence on how migratory behavior, as an intrinsic factor, can shape the genetic structure of Neotropical bird species and improves our understanding of the diversification patterns of open habitat South American species. As expected, migratory behavior can lead to a weak genetic structure, likely the product of substantial ongoing gene flow among populations. Similar patterns may exist among the other 200 species of migratory Neotropical birds, in contrast to the high levels of geographic structuring known to exist within many sedentary Neotropical species.

DATA AVAILABILITY STATEMENT

The authors confirm that all relevant data are included in the article and/or its supplementary information files.

DNA sequences will be deposited in GenBank.

Raw sequence data will be deposited in Dryad.

REFERENCES

Arguedas, N. and Parker, P. 2000. Seasonal Migration and Genetic Population Structure in House Wrens. - Condor 102: 517–528.

Bates, J. M., Tello, J. G. and Silva, J. M. C. da 2003. Initial Assessment of Genetic Diversity in Ten Bird Species of South American Cerrado. - Stud. Neotrop. Fauna Environ. 38: 87–94.

Burney, C. W. and Brumfield, R. T. 2009. Ecology Predicts Levels of Genetic Differentiation in Neotropical Birds. - Am. Nat. 174: 358–368.

Calderón, L., Quintana, F., Cabanne, G. S., Lougheed, S. C. and Tubaro, P. L. 2014. Phylogeography and genetic structure of two Patagonian shag species (Aves: Phalacrocoracidae). - Mol. Phylogenet. Evol. 72: 42–53.

Campillo, L. C., Manthey, J. D., Thomson, R. C., Hosner, P. A. and Moyle, R. G. 2020. Genomic differentiation in an endemic Philippine genus (Aves: Sarcophanops) owing to geographical isolation on recently disassociated islands. - Biol. J. Linn. Soc. 131: 814–821.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. and Postlethwait, J. H. 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. - Genes|Genomes|Genetics 1: 171–182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. and Cresko, W. A. 2013. Stacks: an analysis tool set for population genomics. - Mol. Ecol. 22: 3124–3140.

Chaves, A. V., Clozato, C. L., Lacerda, D. R., Sari, E. H. R. and Santos, F. R. 2008. Molecular taxonomy of Brazilian tyrant-flycatchers (Passeriformes: Tyrannidae). - Mol. Ecol. Resour. 8: 1169–1177.

Chesser, R. T. 1994. Migration in South America: an overview of the austral system. - Bird Conserv. Int. 4: 91–107.

Clegg, S. M., Kelly, J. F., Kimura, M. and Smith, T. B. 2003. Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (Wilsonia pusilla). - Mol. Ecol. 12: 819–830.

Contina, A., Alcantara, J. L., Bridge, E. S., Ross, J. D., Oakley, W. F., Kelly, J. F. and Ruegg, K. C. 2019. Genetic structure of the Painted Bunting and its implications for conservation of migratory populations. - Ibis (Lond. 1859). 161: 372–386.

Curnow, R. N. and Wright, S. 1978. Evolution and the Genetics of Populations, Volume 4: Variability within and among Natural Populations. - Biometrics 35: 359.

De Paiva, L. V. and Marini, M. Â. 2013. Timing of Migration and Breeding of the Lesser Elaenia (*Elaenia chiriquensis*) in a Neotropical Savanna. - Wilson J. Ornithol. 125: 116–120.

Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J. and Liedvogel, M. 2020. The evolutionary history and genomics of European blackcap migration. - Elife 9: 1–24.

DeSaix, M. G., Bulluck, L. P., Eckert, A. J., Viverette, C. B., Boves, T. J., Reese, J. A., Tonra, C. M. and Dyer, R. J. 2019. Population assignment reveals low migratory connectivity in a weakly structured songbird. - Mol. Ecol. 28: 2122–2135.

Edwards, S. V., Shultz, A. J. and Campbell-Staton, S. C. 2015. Next-generation sequencing and the expanding domain of phylogeography. - Folia Zool. 64: 187–206.

Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. - Mol. Ecol. 14: 2611–2620.

Faaborg, J., Holmes, R. T., Anders, A. D., Bildstein, K. L., Dugger, K. M., Gauthreaux, S. A., Heglund, P., Hobson, K. A., Jahn, A. E., Johnson, D. H., Latta, S. C., Levey, D. J., Marra, P. P., Merkord, C. L., Nol, E., Rothstein, S. I., Sherry, T. W., Sillett, T. S., Thompson, F. R. and Warnock, N. 2010. Recent advances in understanding migration systems of New World land birds. - Ecol. Monogr. 80: 3–48.

Foll, M. and Gaggiotti, O. 2006. Identifying the Environmental Factors That Determine the Genetic Structure of Populations. - Genetics 174: 875–891.

Förschler, M. I., del Val, E. and Bairlein, F. 2010. Extraordinary high natal philopatry in a migratory passerine. - J. Ornithol. 151: 745–748.

Friesen, V. L., Congdon, B. C., Walsh, H. E. and Birt, T. P. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. - Mol. Ecol. 6: 1047–1058.

Gómez-Bahamón, V., Márquez, R., Jahn, A. E., Miyaki, C. Y., Tuero, D. T., Laverde-R, O., Restrepo, S. and Cadena, C. D. 2020. Speciation Associated with Shifts in Migratory Behavior in an Avian Radiation. - Curr. Biol. 30: 1312-1321.e6.

Goodwin, S., McPherson, J. D. and McCombie, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. - Nat. Rev. Genet. 17: 333–351.

Guaraldo, A. C., Kelly, J. F. and Marini, M. Â. 2016. Contrasting annual cycles of an intratropical migrant and a tropical resident bird. - J. Ornithol. 157: 695–705.

Guaraldo, A. C., Bravo, S. P., Bridge, E. and Marini, M. Â. 2021. Longitudinal and cyclic poleward migration of a South American intra-tropical migrant flycatcher, the Lesser

Elaenia (*Elaenia chiriquensis*). - Ornitol. Neotrop. 32: 94–102.

Hayes, F. E. 1995. Definitions for Migrant Birds: What Is a Neotropical Migrant? - Auk 112: 521–523.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. and Cresko, W. A. 2010. Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags (DJ Begun, Ed.). - PLoS Genet. in press.

Hosner, P. A. 2004. Genus *Elaenia*. - In: del Hoyo, J. et al. (eds), Handbook of the Birds of the World. Vol. 9. Cotingas to Pipits and Wagtails. Eds. Lynx Edicions, pp. 267–274.

Hosner, P. 2020. Plain-crested Elaenia (*Elaenia cristata*). - In: del Hoyo, J. et al. (eds), Birds of the World. version 1. Cornell Lab of Ornithology, in press.

Hosner, P., Christie, D., de Juana, E. and Kirwan, G. M. 2020. Lesser Elaenia (*Elaenia chiriquensis)*. - In: del Hoyo, J. et al. (eds), Birds of the World. Version1. Cornell Lab of Ornithology, in press.

Irwin, D. E., Irwin, J. H. and Smith, T. B. 2011. Genetic variation and seasonal migratory connectivity in Wilson's warblers (Wilsonia pusilla): species-level differences in nuclear DNA between western and eastern populations. - Mol. Ecol. 20: 3102–3115.

Jahn, A., Levey, D., Johnson, J., Mamani, A. and Davis, S. 2006. Towards a mechanistic interpretation of bird migration in South America. - El hornero 21: 99–108.

Jahn, A. E., Cueto, V. R., Fontana, C. S., Guaraldo, A. C., Levey, D. J., Marra, P. P. and Ryder, T. B. 2020. Bird migration within the Neotropics. - Auk 137: 1–23.

Jonker, R. M., Zhang, Q., Van Hooft, P., Loonen, M. J. J. E., Van der Jeugd, H. P., Crooijmans, R. P. M. A., Groenen, M. A. M., Prins, H. H. T. and Kraus, R. H. S. 2012. The Development of a Genome Wide SNP Set for the Barnacle Goose Branta leucopsis (H Ellegren, Ed.). - PLoS One 7: e38412.

Kerr, K. C. R., Stoeckle, M. Y., Dove, C. J., Weigt, L. A., Francis, C. M. and Hebert, P. D. N. 2007. Comprehensive DNA barcode coverage of North American birds. - Mol. Ecol. Notes 7: 535–543.

Kopuchian, C., Campagna, L., Lijtmaer, D. A., Cabanne, G. S., García, N. C., Lavinia, P. D., Tubaro, P. L., Lovette, I. and Di Giacomo, A. S. 2020. A test of the riverine barrier hypothesis in the largest subtropical river basin in the Neotropics. - Mol. Ecol.: mec.15384.

Kraus, R. H. S., Kerstens, H. H. D., Van Hooft, P., Crooijmans, R. P. M. A., Van Der Poel, J. J., Elmberg, J., Vignal, A., Huang, Y., Li, N., Prins, H. H. T. and Groenen, M. A. M. 2011. Genome wide SNP discovery, analysis and evaluation in mallard (Anas platyrhynchos). -

BMC Genomics 12: 150.

Kraus, R. H. S., van Hooft, P., Megens, H.-J., Tsvey, A., Fokin, S. Y., Ydenberg, R. C. and Prins, H. H. T. 2013. Global lack of flyway structure in a cosmopolitan bird revealed by a genome wide survey of single nucleotide polymorphisms. - Mol. Ecol. 22: 41–55.

Lavretsky, P., DaCosta, J. M., Sorenson, M. D., McCracken, K. G. and Peters, J. L. 2019. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. - Mol. Ecol. 28: 2594–2609.

Lawson, D. J., Hellenthal, G., Myers, S. and Falush, D. 2012. Inference of Population Structure using Dense Haplotype Data (GP Copenhaver, Ed.). - PLoS Genet. 8: e1002453.

Lima-Rezende, C. A., Rocha, A. V., Júnior, A. F. C., Martins, É. de S., Vasconcelos, V. and Caparroz, R. 2019a. Late Pleistocene climatic changes promoted demographic expansion and population reconnection of a Neotropical savanna-adapted bird, Neothraupis fasciata (Aves: Thraupidae) (Ł Kajtoch, Ed.). - PLoS One 14: e0212876.

Lima-Rezende, C. A., de Souza, R. O. and Caparroz, R. 2019b. The spatial genetic structure of the White-banded Tanager (Aves, Passeriformes) in fragmented Neotropical savannas suggests two evolutionarily significant units. - Biotropica 51: 234–244.

Mairal, M., Sanmartín, I., Herrero, A., Pokorny, L., Vargas, P., Aldasoro, J. J. and Alarcón, M. 2017. Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. - Sci. Rep. 7: 45749.

Malinsky, M., Trucchi, E., Lawson, D. J. and Falush, D. 2018. RADpainter and fineRADstructure: Population Inference from RADseq Data (N Takezaki, Ed.). - Mol. Biol. Evol. 35: 1284–1290.

Marini, M. and Cavalcanti, R. 1990. Migrações de Elaenia albiceps chilensis e Elaenia chiriquensis albivertex (Aves: Tyrannidae). - Bol. do Mus. Para. Emilio Goeldi, série Zool.: 59.

Marini, M., Sousa, N., Borges, F. and Silveira, M. 2009. Biologia reprodutiva de Elaenia cristata (Aves: Tyrannidae) em cerrado do Brasil central. - Neotrop. Biol. Conserv. 4: 2–12.

Medeiros, R. de C. S. de and Marini, M. Â. 2007. Biologia reprodutiva de Elaenia chiriquensis (Lawrence) (Aves, Tyrannidae) em Cerrado do Brasil Central. - Rev. Bras. Zool. 24: 12–20.

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R. and Schäffer, A. A. 2008. Database indexing for production MegaBLAST searches. - Bioinformatics 24:

1757–1764.

Newton, I. 2003. Speciation and biogeography of birds. - Academic Press an Imprint of Elsevier Science Ltd.

Nistelberger, H. M., Byrne, M., Coates, D. and Roberts, J. D. 2015. Genetic drift drives evolution in the bird-pollinated, terrestrial island endemic G revillea georgeana (Proteaceae). - Bot. J. Linn. Soc. 178: 155–168.

Papadopoulou, A. and Knowles, L. L. 2016. Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. - Proc. Natl. Acad. Sci. 113: 8018–8024.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. and Hoekstra, H. E. 2012. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species (L Orlando, Ed.). - PLoS One 7: e37135.

R Core Team 2019. R: A Language and Environment for Statistical Computing. in press.

Ramos, A. C. S., Lemos-Filho, J. P., Ribeiro, R. A., Santos, F. R. and Lovato, M. B. 2007. Phylogeography of the Tree Hymenaea stigonocarpa (Fabaceae: Caesalpinioideae) and the Influence of Quaternary Climate Changes in the Brazilian Cerrado. - Ann. Bot. 100: 1219–1228.

Rheindt, F. E., Christidis, L. and Norman, J. A. 2008. Habitat shifts in the evolutionary history of a Neotropical flycatcher lineage from forest and open landscapes. - BMC Evol. Biol. 8: 193.

Rheindt, F. E., Krabbe, N., Wee, A. K. S. and Christidis, L. 2015. Cryptic speciation in the Lesser Elaenia Elaenia chiriquensis (Aves: Passeriformes: Tyrannidae). - Zootaxa 4032: 251.

Ritter, C. D., Coelho, L. A., Capurucho, J. M., Borges, S. H., Cornelius, C. and Ribas, C. C. 2021. Sister species, different histories: comparative phylogeography of two bird species associated with Amazonian open vegetation. - Biol. J. Linn. Soc. 132: 161–173.

Rocha, A. V., Cabanne, G. S., Aleixo, A., Silveira, L. F., Tubaro, P. and Caparroz, R. 2020. Pleistocene climatic oscillations associated with landscape heterogeneity of the South American dry diagonal explains the phylogeographic structure of the narrow-billed woodcreeper (Lepidocolaptes angustirostris , Dendrocolaptidae). - J. Avian Biol. 51: jav.02537.

Rochette, N. C. and Catchen, J. M. 2017. Deriving genotypes from RAD-seq short-read data using Stacks. - Nat. Protoc. 12: 2640–2659.

Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., DeSante, D.

F., Moore, F. and Smith, T. B. 2014. Mapping migration in a songbird using high-resolution genetic markers. - Mol. Ecol. 23: 5726–5739.

Sick, H. 1997. Ornitologia Brasileira. - Nova Fronteira.

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

Somenzari, M., Amaral, P. P. do, Cueto, V. R., Guaraldo, A. de C., Jahn, A. E., Lima, D. M., Lima, P. C., Lugarini, C., Machado, C. G., Martinez, J., Nascimento, J. L. X. do, Pacheco, J. F., Paludo, D., Prestes, N. P., Serafini, P. P., Silveira, L. F., Sousa, A. E. B. A. de, Sousa, N. A. de, Souza, M. A. de, Telino-Júnior, W. R. and Whitney, B. M. 2018. An overview of migratory birds in Brazil. - Pap. Avulsos Zool. 58: 3.

Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T. and Mindell, D. P. 1999. Primers for a PCR-Based Approach to Mitochondrial Genome Sequencing in Birds and Other Vertebrates. - Mol. Phylogenet. Evol. 12: 105–114.

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

Stotz, D. F., Fitzpatrick, J. W., Parker III, T. A. and Moskovits, D. K. 1996. Neotropical birds: ecology and conservation. - University of Chicago Press.

Tang, Q., Edwards, S. V and Rheindt, F. E. 2018. Rapid diversification and hybridization have shaped the dynamic history of the genus Elaenia. - Mol. Phylogenet. Evol. 127: 522–533.

Thrasher, D. J., Butcher, B. G., Campagna, L., Webster, M. S. and Lovette, I. J. 2018. Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird. - Mol. Ecol. Resour. 18: 953–965.

Traylor Jr, M. A. 1982. Notes on tyrant flycatechers (Aves: Tyrannidae). - Fieldiana Zool.: 1–22.

Weatherhead, P. J. and Forbes, M. R. L. 1994. Natal philopatry in passerine birds: genetic or ecological influences? - Behav. Ecol. 5: 426–433.

Werneck, F. P. 2011. The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. - Quat. Sci. Rev. 30: 1630–1648.

Winkler, D. W., Billerman, S. M. and Lovette, I. J. 2020. Tyrant Flycatchers (Tyrannidae). - In: Billerman, S. M. et al. (eds), Birds of the World. Cornell Lab of Ornithology, in press.

Wright, N. T. W. and Mauck, R. A. 1998. Philopatry, Natal Dispersal, and Inbreeding

Avoidance in an Island Population of Savannah Sparrows. - Ecology 79: 755.

Wuster, W., Ferguson, J. E., Quijada-Mascareñas, A., Pook, C. E., Salomão, M. da G. and Thorpe, R. S. 2005. Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: Crotalus durissus). - Mol. Ecol. 14: 1095–1108.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C. and Weir, B. S. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. - Bioinformatics 28: 3326–3328.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting information section.

TABLES AND FIGURES

TABLE 1. Description of the samples of *E. cristata* and *E. chiriquensis* used in the genetic analyses after removing misidentified or samples with >80% of missing data. Locality number (Nº) corresponding to those shown in Figures 3 and 4. Sample size for each locality is also shown.

Nº	State; collection site	Latitude	Longitude	*E. cristata*	*E. chiriquensis*
1	Roraima; Viruá National Park	1.42	-60.98	2	0
2	Amapá; Cerrado Experiment Field of the Embrapa Amapá	0.39	-51.05	16	2
3	Pará; Óbidos	0.63	-55.72	0	7
4	Pará; Oriximiná	-1.76	-56.22	10	4
5	Pará; Monte Alegre	-2.02	-54.18	3	0
6	Pará; Portel	-1.85	-50.70	10	1
7	Pará; Marajó Island	-0.78	-48.61	10	12
8	Pará; Parauapebas	-6.28	-50.58	1	0
9	Amazonas; Manicoré	-8.47	-61.39	4	1
10	Rondônia; Machadinho	-8.92	-62.07	5	0
11	Maranhão; Urbano Santos	-3.28	-44.33	1	0
12	Maranhão; Chapada das Mesas National Park	-7.13	-47.15	8	11
13	Piauí; Pirarucura	-4.11	-41.71	0	1
14	Piauí; Castelo do Piauí	-5.33	-41.57	2	3
15	Piauí; Guadalupe	-6.74	-43.72	0	1
16	Piauí; Serra das Confusões	-9.37	-43.82	0	1
17	Piauí; Uruçuí	-7.73	-44.54	7	0
18	Rio Grande do Norte; Rio Fogo	-5.29	-35.40	10	0
19	Rio Grande do Norte; Parnamirim	-5.92	-35.17	6	2
20	Tocantins; Ponte Alta do Tocantins	-10.76	-47.48	4	1
21	Mato Grosso; Guarantã do Norte	-10.12	-54.36	0	1
22	Mato Grosso; Vila Bela da Santíssima Trindade	-14.63	-60.20	10	3

23	Mato Grosso; Chapada dos Guimarães National Park	-15.40	-55.83	10	10
24	Mato Grosso; Itiquira	-17.21	-54.14	0	2
25	Mato Grosso; Araguaia	-17.42	-53.45	0	1
26	Mato Grosso; Nova Xavantina	-14.78	-52.53	5	0
27	Goiás; Emas National Park	-17.92	-52.97	2	1
28	Goiás; Chapada dos Veadeiros National Park	-14.16	-47.74	2	8
29	Distrito Federal; Protection Area Gama and Cabeça de Veado	-15.92	-47.87	15	73
30	Distrito Federal; Águas Emendadas Ecological Station	-15.55	-47.59	3	1
31	Minas Gerais; Grande Sertão Veredas National Park	-15.18	-45.69	9	1
32	Minas Gerais; Mateus Leme	-20.14	-44.45	0	2
33	Minas Gerais; Serra Azul	-20.14	-44.41	1	0
	Total			**156**	**150**

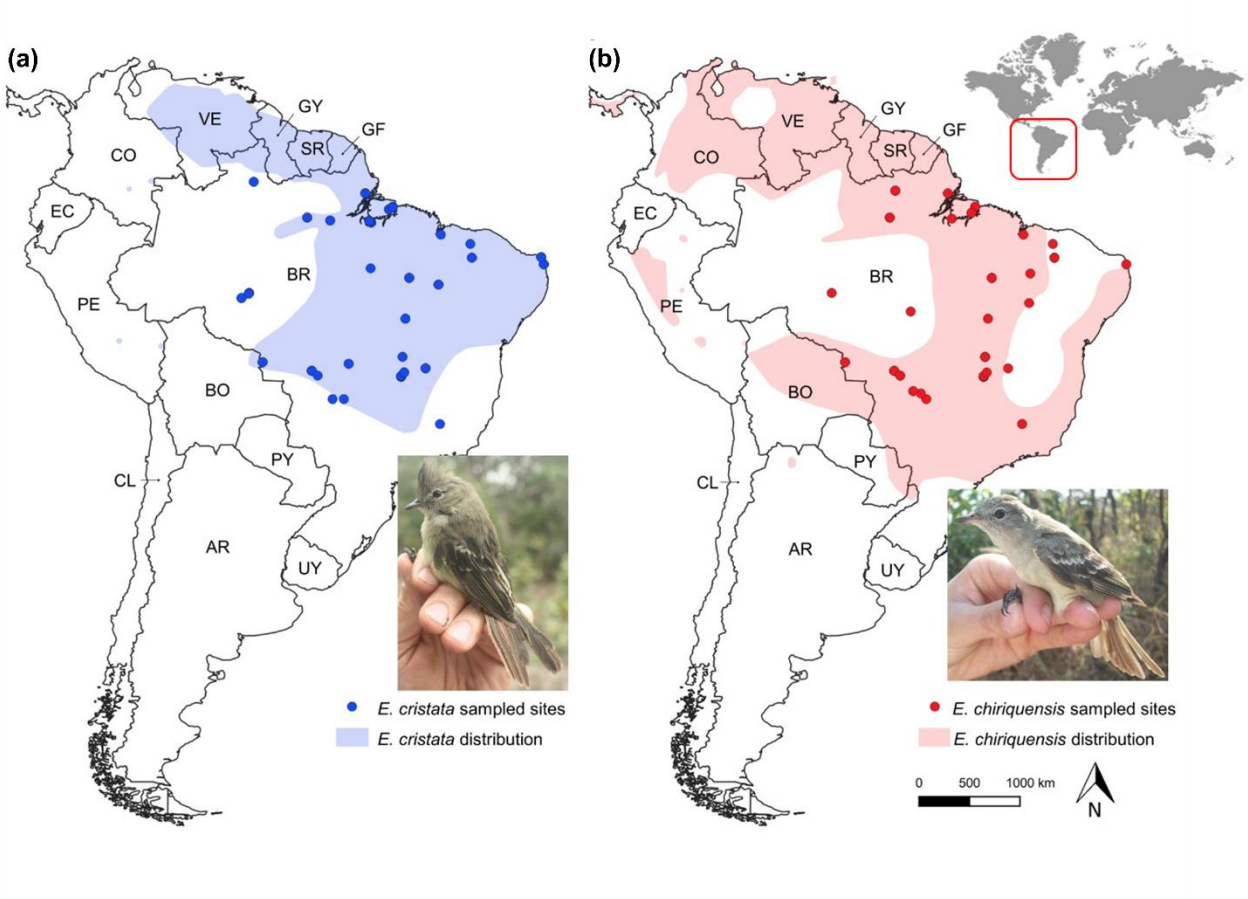


FIGURE 1. Geographical range distributions of *E. cristata* (a, blue) and *E. chiriquensis* (b, red) based on the International Union for Conservation of Nature (IUCN). Sampling sites are represented by color dots for each species (details in Table 1). Photos: (a) *E. cristata* from Mato Grosso - BR; Chapada dos Guimarães National Park (site 23; Figure 3); (b) *E. chiriquensis* from Distrito Federal - BR; Protection Area Gama and Cabeça de Veado (site 29; Figure 4).

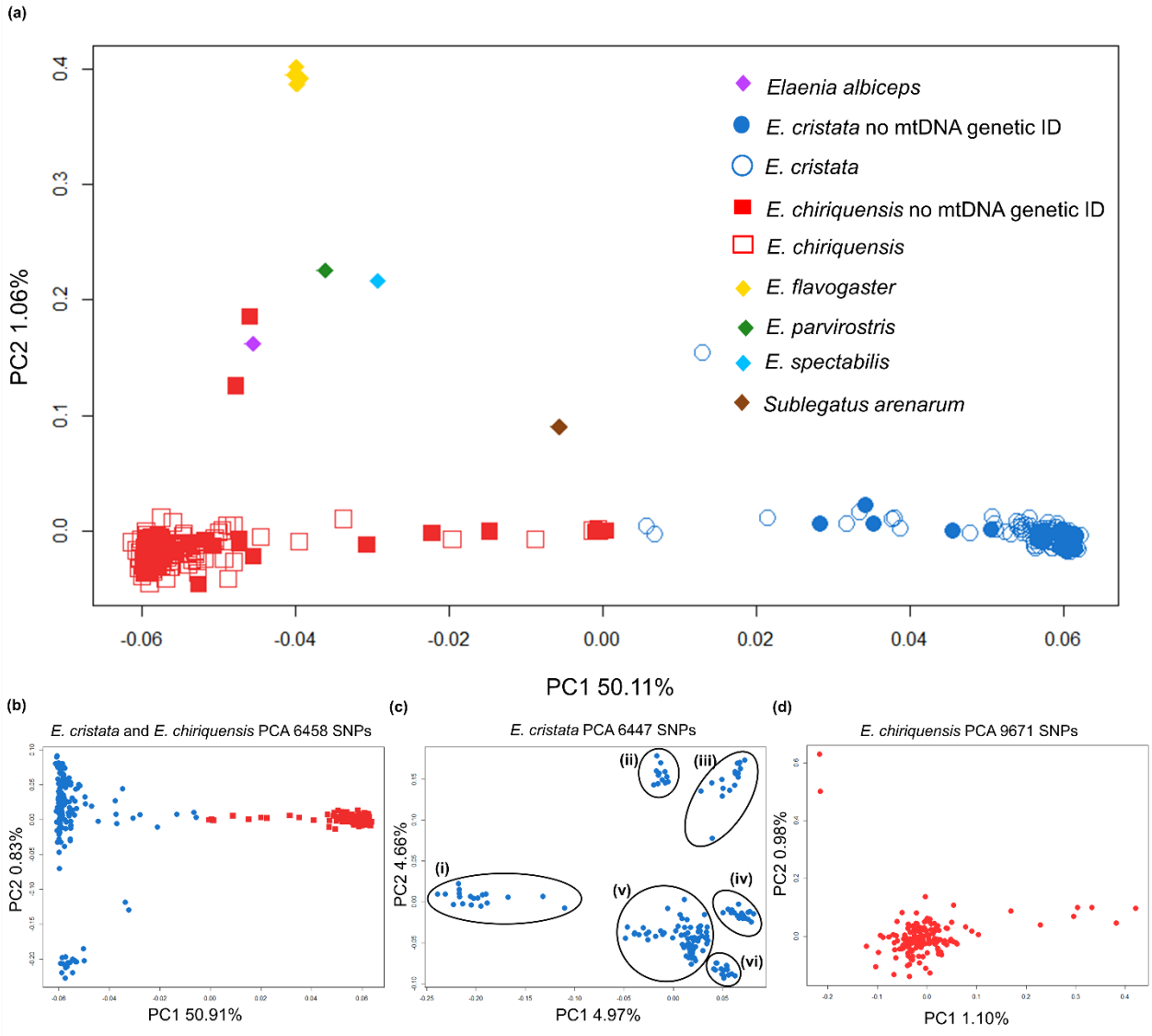


FIGURE 2. Principal component analyses (PCAs) derived from different sets of SNPs and *Elaenia*/*Sublegatus* samples (symbols). (a) 5,983 SNPs from 326 samples. Diamond symbols refer to other species detected after genetic identification. Closed circles refer *E. cristata* species with no mtDNA genetic identification while open circles refer to samples with mitochondrial genetic identification. Closed squares refer to *E. chiriquensis* without genetic identification, and open squares to those with the genetic ID (see Methods, for details). (b) 6,458 SNPs from 156 *E. cristata* and 150 *E. chiriquensis* samples after removing nine misidentified samples and samples with missing data >80%. (c) 6,447 SNPs from 156 *E. cristata*. The ellipses indicate groups with geographic correspondence, localities numbers are described in table 1: (i) sites 9, 10 and 22; (ii)1, and 4; (iii) 2, and 5; (iv) 6, and 7; (v) Brazilian central region; (vi) sites 18, and 19. (d) 9,671 SNPs from 150 *E. chiriquensis*.

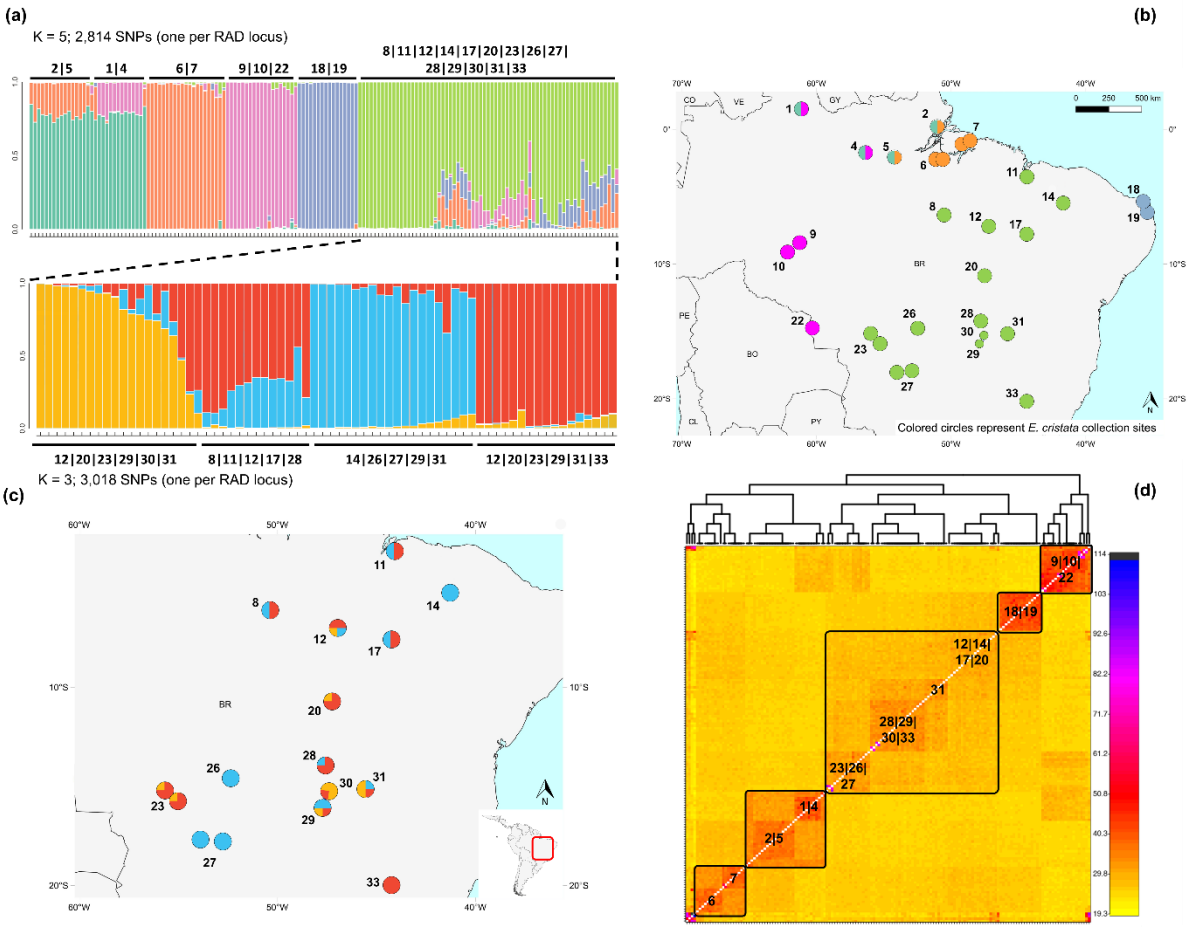


FIGURE 3. Pattern of genetic structure in the resident species (*E. cristata*). (a) Structure analysis based on 2,814 SNPs (one per RAD locus) and 156 specimens showing five populations (K = 5) across all sampled sites (top). When we looked into the Brazilian central region ("green group") in further detail, the 70 sampled specimens were grouped into three populations (K = 3) in an analysis based on 3,018 SNPs (bottom). (b) Geographical distribution of genetic clusters detected in the Structure analysis across all sampled sites; colored circles represent groups according to the Structure results; numbers correspond to sampled sites described in Table 1. (c) Geographical distribution of genetic clusters detected in the Structure analysis across sites in central Brazil. (d) The fineRADstructure plot derived from haplotype data (n=156) indicating the clusters formed according to geographic location (numbers correspond to the localities shown on the maps).

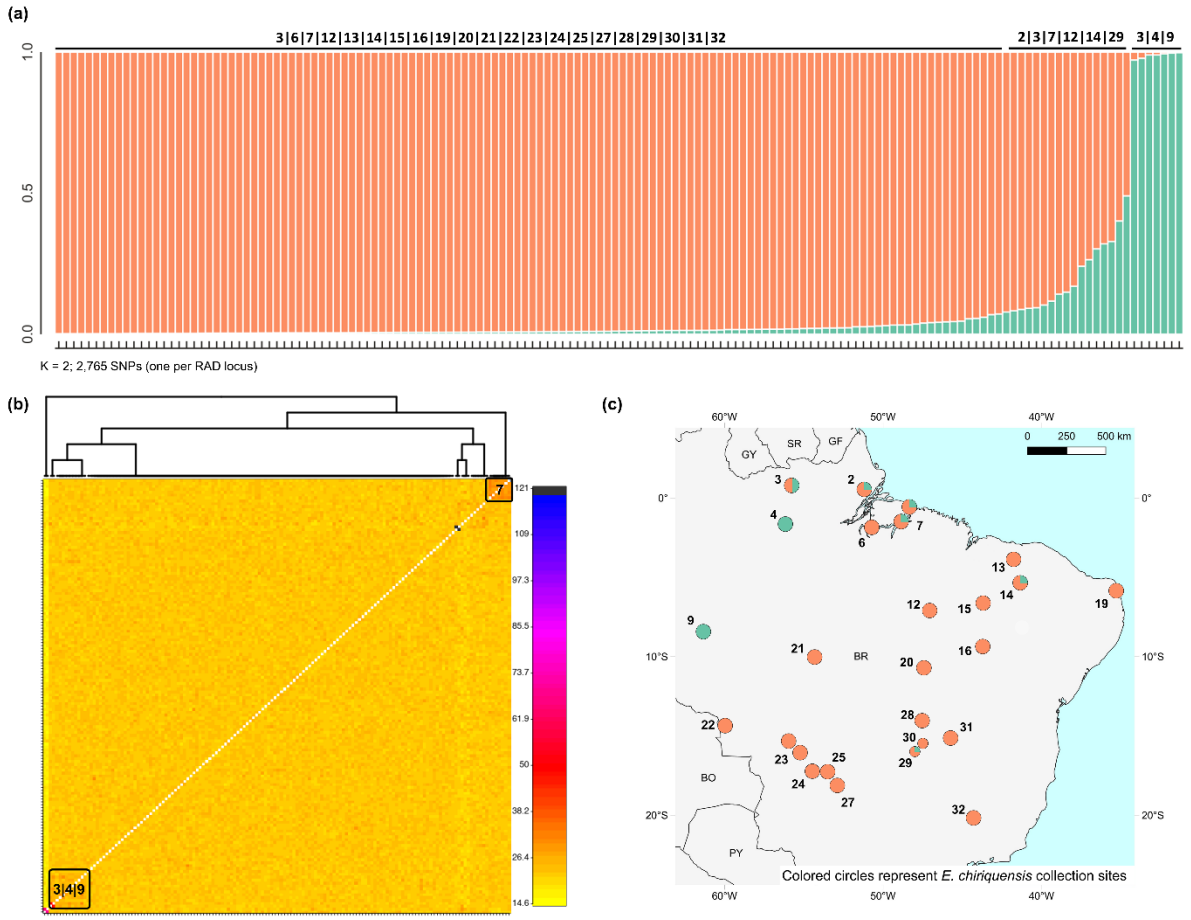


FIGURE 4. Pattern of genetic structure in the migratory species (*E. chiriquensis*). (a) The Structure analysis based on 2,765 SNPs (one per RAD locus) and 150 specimens showed 2 populations (K = 2) across all sampled sites. (b) The fineRADstructure plot derived from haplotype data (n=150) indicating the clusters formed according to geographic location (the numbers correspond to sampled sites shown on the maps). (c) Geographic distribution of genetic clusters detected in the Structure analysis across all sampled sites; colored circles represent groups according to the Structure results (not proportional to the number of sampled individuals); numbers correspond to sampled sites described in Table 1.

SUPPLEMENTAL INFORMATION

TABLE S1 go along with the Supplementary Material in separate *.csv* format.

TABLE S2. Structure results obtained for different K values based on the analyses of 2,814 SNPs for 156 *Elaenia cristata* individuals sampled across South America.

K	Nreps	deltaK	estLnProbMean	estLnProbStdev
1	10	0.0	-348556.27	2.687
2	10	0.762	-337805.330	485.760
3	10	1193.998	-327424.91	4.740
4	10	0.666	-322704.05	108.725
5	10	671.665	-317910.71	5.503
6	10	4.630	-316813.829	77.435
7	10	9.920	-316075.48	50.182
8	10	1.682	-315834.960	298.158
9	10	0.806	-316096.21	1214.345
10	10	0.0	-317336.23	3314.581

TABLE S3. Structure results obtained for different K values based on 3,018 SNPs for a subsample of 70 individuals of *Elaenia cristata* sampled across the central distribution of the species.

K	Nreps	deltaK	estLnProbMean	estLnProbStdev
1	10	0.0	-165860.639	4.336
2	10	1.723	-164227.330	172.503
3	10	114.091	-162891.279	7.413
4	10	5.750	-162401.08	112.929
5	10	0.030	-162560.29	1305.005
6	10	0.0	-162679.810	162.386

TABLE S4. Structure results obtained for different K values based on the analyses of 2,765 SNPs for 150 *Elaenia chiriquensis* individuals sampled across South America.

K	Nreps	deltaK	estLnProbMean	estLnProbStdev
1	10	0.0	-305857.43	2.217
2	10	157.195	-305481.29	10.845
3	10	0.721	-306810.029	4458.621
4	10	3.799	-304922.32	476.058
5	10	0.0	-304843.529	338.555

TABLE S5. Pairwise F_{ST} values among *E. cristata* individuals calculated using 'populations' (Stacks v2.4). The clusters were defined based on the Structure analysis (K = 5). POP1: sampling sites 9, 10 and 22; POP2: 18, 19; POP3: 6, 7; POP4: 1, 2, 4 and 5; POP5: 8, 11, 12, 14, 17, 20, 23, 26, 27, 28, 29, 30, 31, and 33 (see Fig. 1 and Table 1, for details).

	POP1	POP2	POP3	POP4	POP5
POP1		0.109	0.097	0.085	0.048
POP2			0.069	0.071	0.028
POP3				0.056	0.028
POP4					0.04

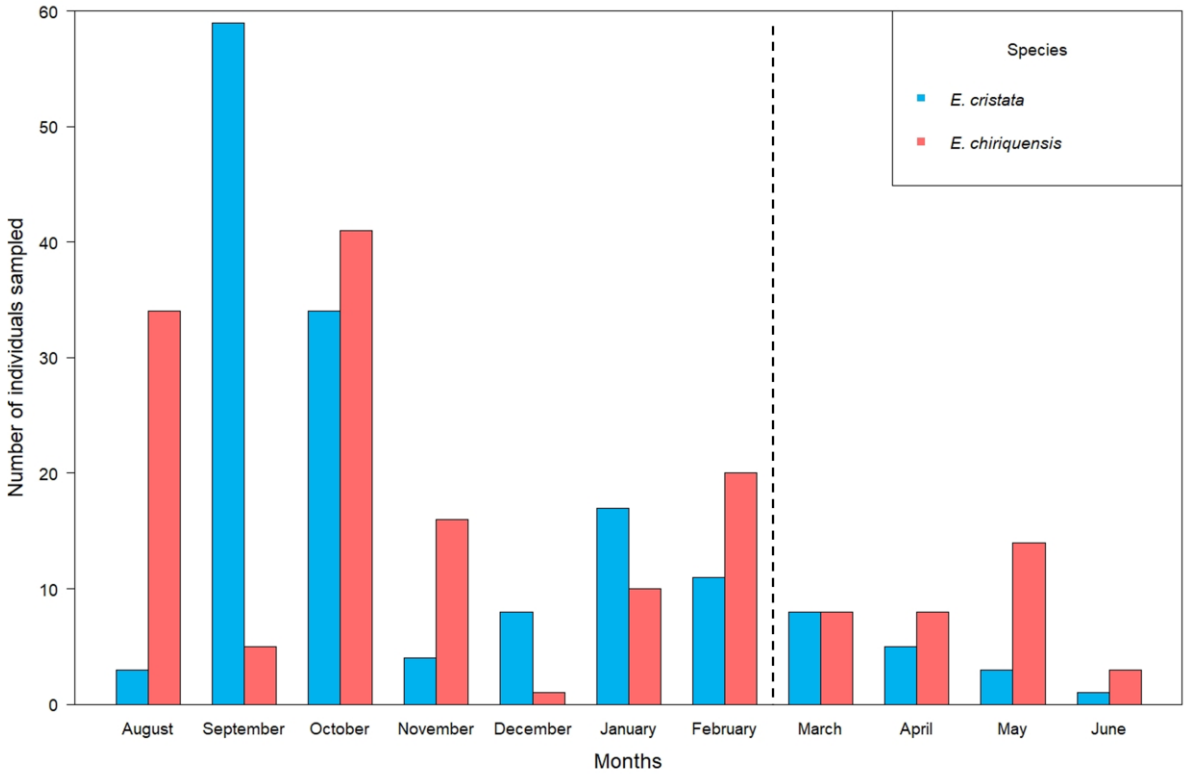


FIGURE S1. Number of individuals of *Elaenia cristata* (blue bars) and *E. chiriquensis* (red bars) analyzed in our study grouped by sampling month. The annual cycle was divided in breeding season (August-February) and non-breeding season (March-June), there were no samples obtained in July.

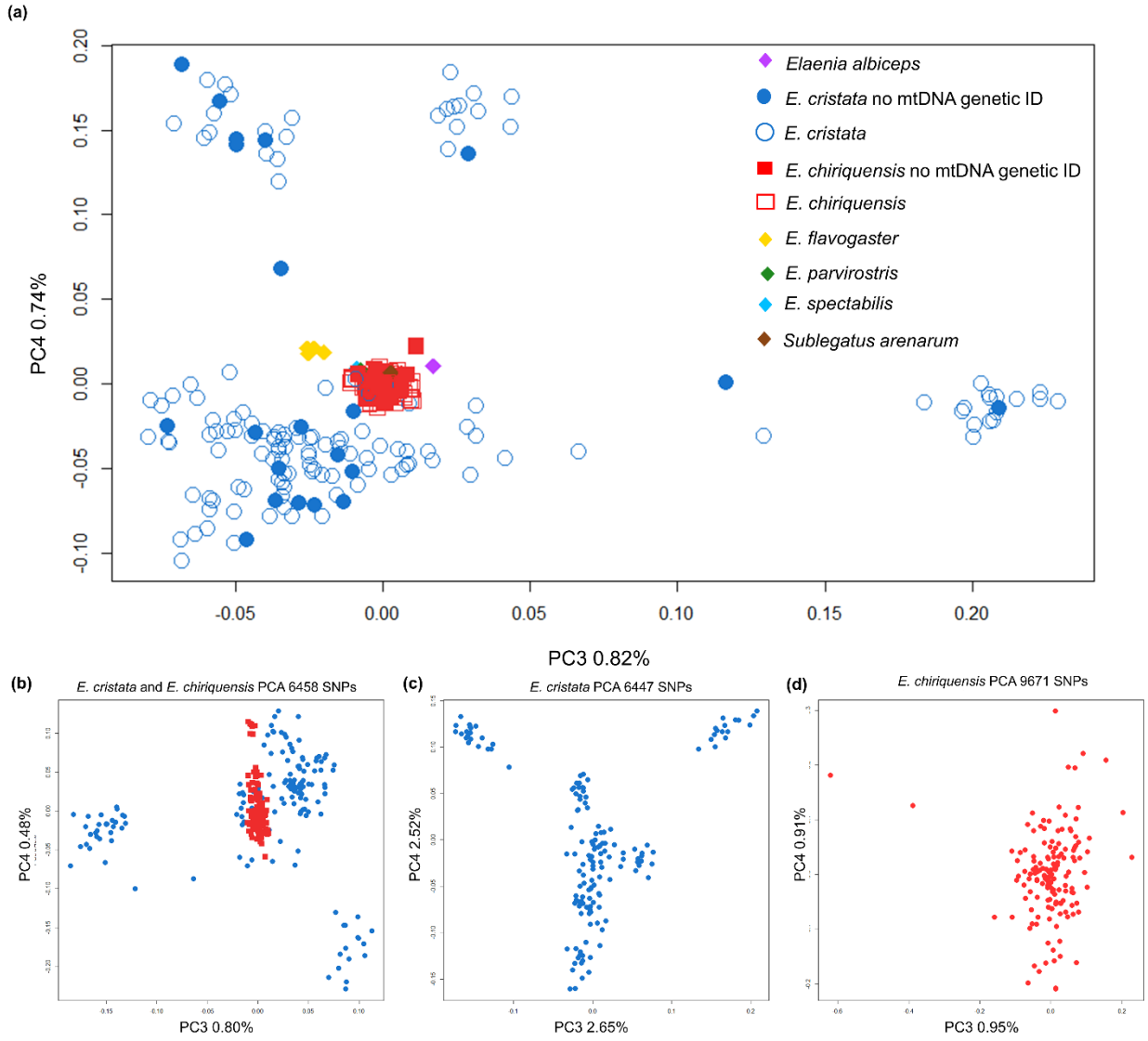
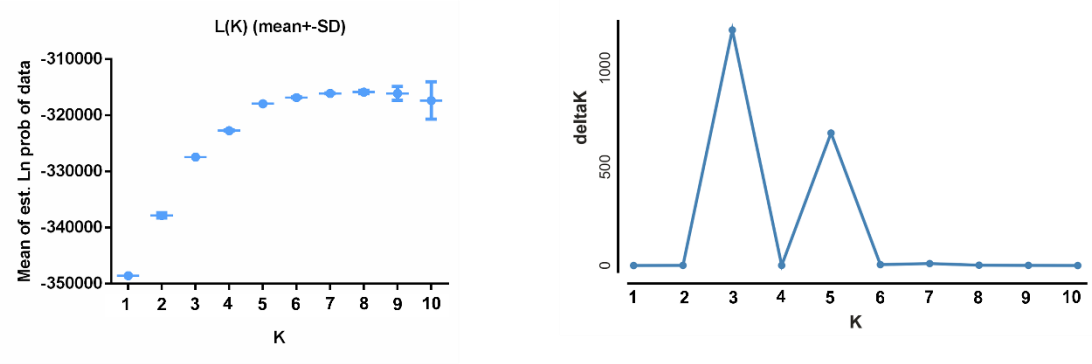


FIGURE S2. (a) Principal component analysis (PC3 and PC4) derived from 5,983 SNP obtained from 326 individuals explaining 1.56% of the total variance. Diamond symbols refer to other species detected after genetic identification. Closed circles refer *E. cristata* species with no genetic identification while open circles refer to samples with mitochondrial genetic identification. Closed squares refer to *E. chiriquensis* without genetic identification, and open squares to those with the genetic ID (see Methods, for details). (b) PC3 and PC4 using 6,458 from *E. cristata* and *E. chiriquensis* after removing the nine misidentified samples from other species. (c) PC3 and PC4 from the analysis performed with *E. cristata* samples and 6,447 SNPs. (d) PC3 and PC4 from the analysis performed with *E. chiriquensis* samples and 9,671 SNPs.

(a) *Elaenia cristata*



(b)

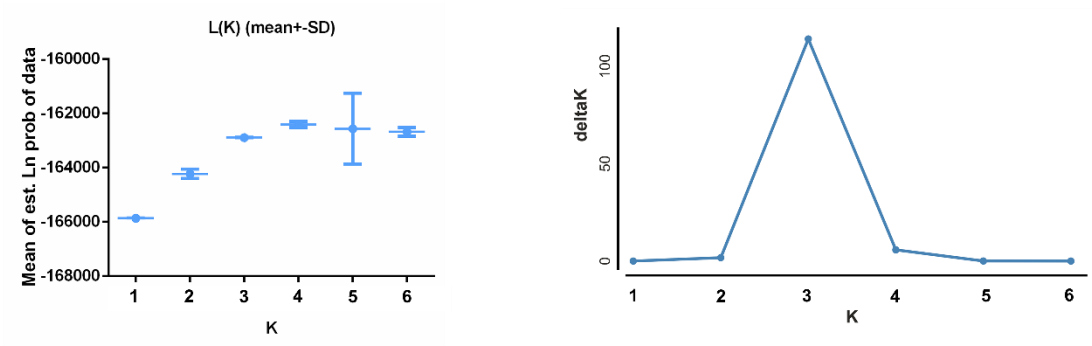


FIGURE S3. Evanno's values for the Structure analyses of the resident species *E cristata*. (a) Values refers to table S2. (b) Values refer to table S3.

(a) *Elaenia chiriquensis*

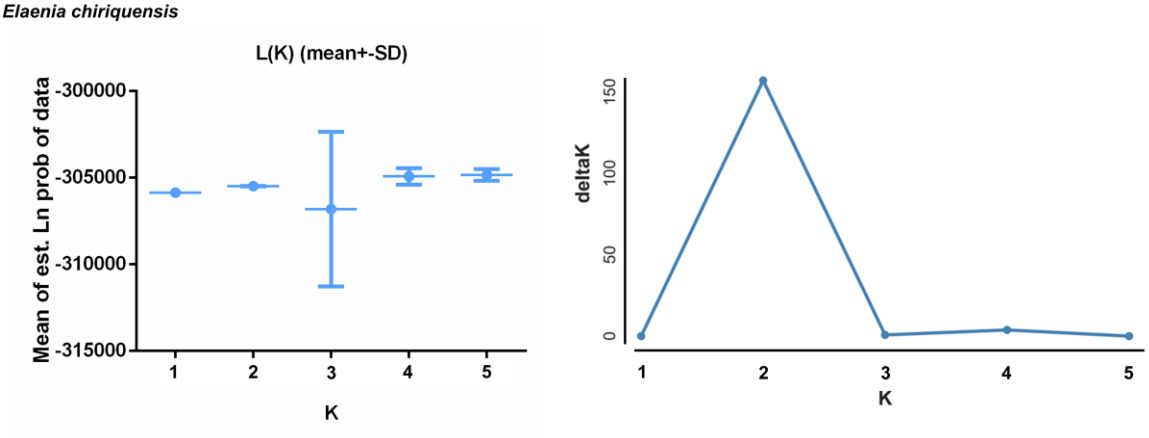


FIGURE S4. Evanno's values for the Structure analyses of the migratory species *E. chiriquensis.*

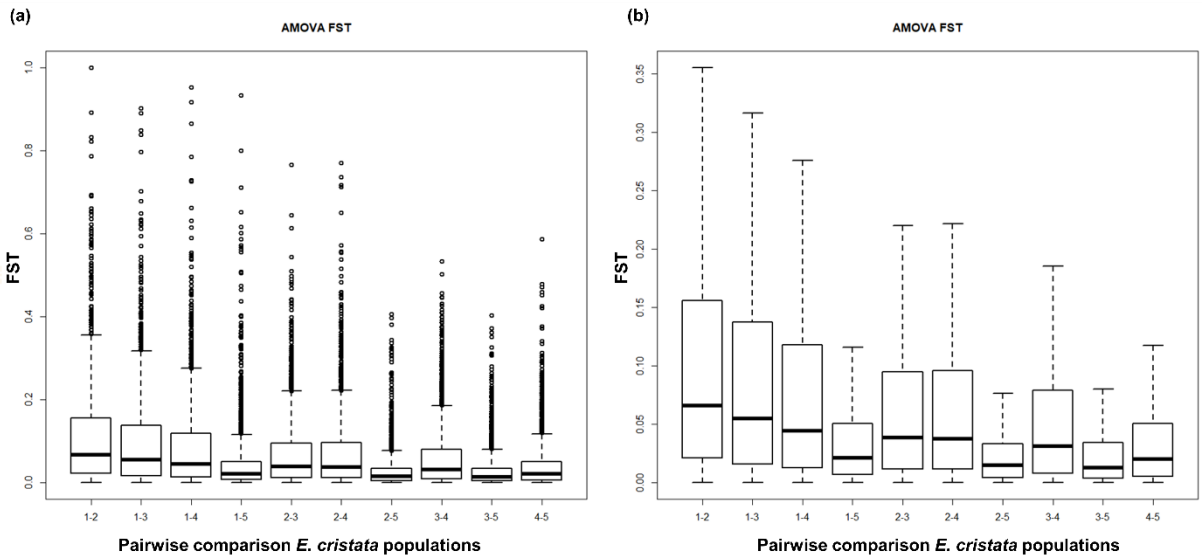


FIGURE S5. Boxplots showing the distribution of Pairwise F_{ST} values among *E. cristata* genetic populations inferred by the Structure analyses. (a) F_{ST} values including outlier values, and (b) after removing the outliers. The numbers represent the collection sites that composed each population (Table 1, Fig. 3 and 4): POP1 (9, 10 and 22); POP2 (18, 19); POP3 (6, 7); POP4 (1, 2, 4 and 5); POP5 (8, 11, 12, 14, 17, 20, 23, 26, 27, 28, 29, 30, 31, and 33).

FIGURE S6. RAxML tree derived from the SNP data set showing the relationships among *E. cristata* samples (n=156, 500 replicates, bootstrap support values >50 shown on the nodes). Specimens highlighted based on the Structure results (Fig. 3a, top). Green indicates the central area of the *E. cristata* distribution (Fig. 3b, sites 8, 11, 12, 14, 17, 20, 23, 26, 27, 28, 29, 30, 31, 33); Dark blue are samples from the Northeast (sites 18, 19); Orange samples from Marajó Island and Portel (sites 6, 7); Pink samples from South of the Amazonian Forest (sites 9, 10, 22); Light blue are sampled from North of Amazonian Forest (sites 1, 2, 4, 5); The numbers at the tree's tip are the sample collection numbers (Supplementary Table 1).

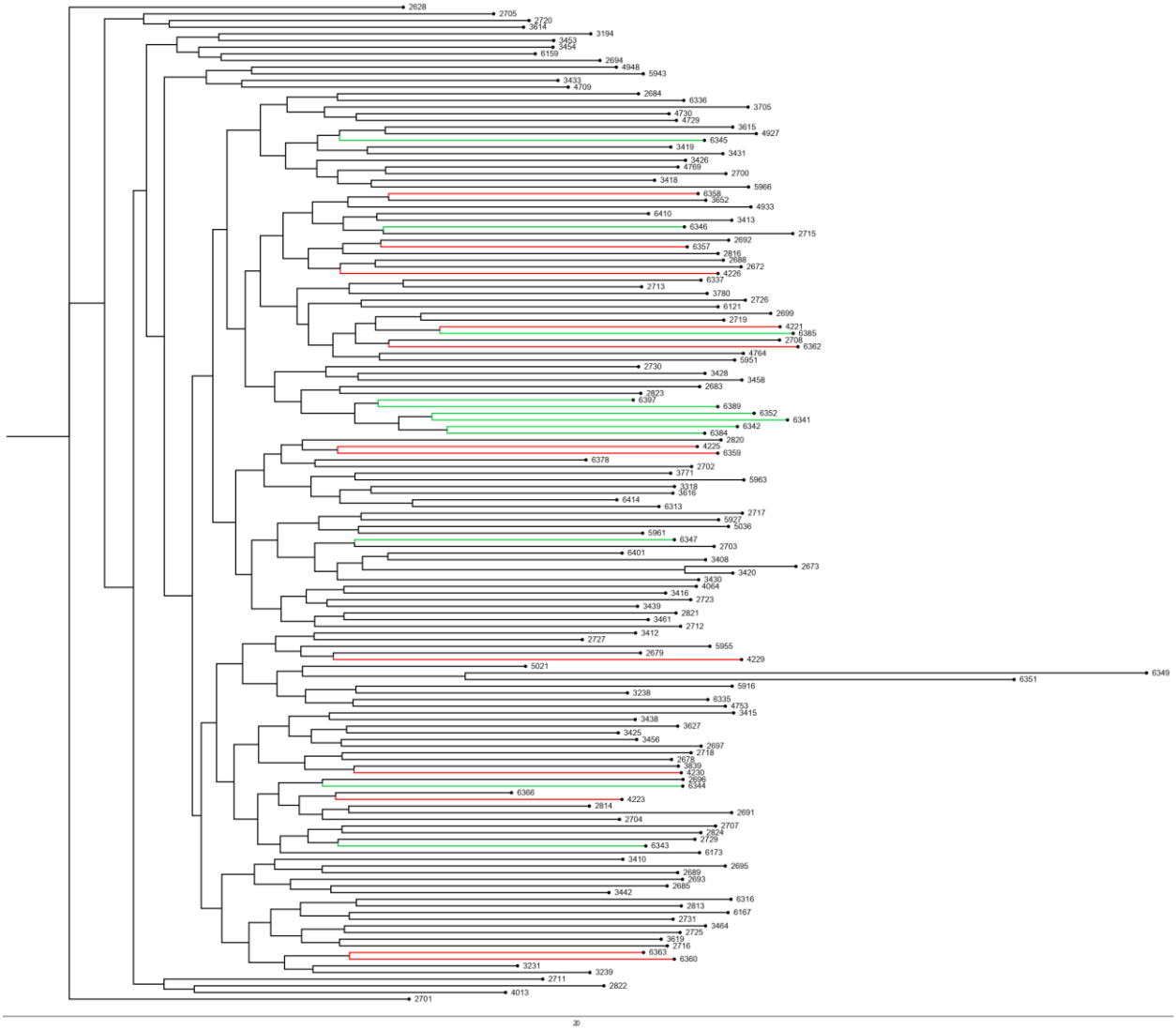


FIGURE S7. RAxML tree derived from the SNP data set showing the relationships among *E. chiriquensis* samples (n=150, 500 replicates, bootstrap support values > 50 shown on the nodes). Specimens highlighted based on the fineRADStructure results (Fig. 4c). Red indicates the Marajó Island (Fig. 4b, site 7), and samples from Óbidos, Oriximiná, and Manicoré (sites 3, 4 and 9) are in green; The numbers at the tree's tips are the sample collection numbers (Supplementary Table 1).

CAPÍTULO 2

Effects of Quaternary climatic fluctuations on avifauna associated with open areas of South America: phylogeography of two species of the genus *Elaenia* (Aves, Tyrannidae)

Eliane L. de Freitas[a], Victor Leandro-Silva[b], Leonardo Campagna[c,d], Irby Lovette[c,d], and Renato Caparroz[a]

[a]Genetics and Biodiversity Laboratory, Department of Genetics and Morphology, University of Brasília – UnB, Brasília, Brazil.

[b]Programa de Pós-graduação em Etnobiologia e Conservação da Natureza, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.

[c]Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA.

[d]Department of Ecology and Evolutionary Biology, Cornell University, 215 Tower Road, Ithaca, NY 14853, USA.

Corresponding author: elianeluizdefreitas@gmail.com

ABSTRACT

The geographical distribution patterns of biodiversity are the results of interactions between several intrinsic and extrinsic factors. These factors had affected the species in a combination of several ways. Some intrinsic elements, such as migration or dispersal capacity, can overlap with extrinsic elements, such as climatic changes and geographical barriers. Over time those interactions could shape the distribution of species. Considering the climatic fluctuations that occurred in the Pleistocene altered significantly the landscape we tested the hypothesis of those climate changes would have affected the population dynamic of the open area bird species. We used two types of molecular markers (mtDNA and SNPs) and ecological niche modeling analysis with samples of two species of the genus *Elaenia* as models: the Plain-crested Elaenia (*Elaenia cristata*) and the Lesser Elaenia (*E. chiriquensis*). Our results demonstrate that climatic fluctuations that occurred in the Pleistocene, may have affected the distribution and the genetic population structuring of *E. cristata* and *E. chiriquensis*. We found signs of expansion in populations of the resident species coincident with the expansion patterns of open areas in the climate scenarios. However, intrinsic characteristics of the species, such as migration, may restrict the identification of a general pattern for these areas due to the impact on the population structuring.

KEYWORDS

Climatic fluctuations; ecological niche modeling; Population structure

INTRODUCTION

The geographical distribution patterns of biodiversity are the results of interactions between several intrinsic and extrinsic factors. These factors had affected the species in a combination of several ways. Some intrinsic elements, such as migration (Gómez-Bahamón *et al.*, 2020) or dispersal propensity (Burney and Brumfield, 2009), can overlap with extrinsic elements, such as climatic changes (Vasconcellos *et al.*, 2019) and geographical barriers (Oliveira, Vasconcelos and Santos, 2017), over time those interactions could shape the distribution of species.

The climate change experienced by the neotropical region is specially marked by the alternation of periods of hot and humid climate or drier and colder periods, especially related to the Quaternary (ca 2.6 Myr), proved to be a determining abiotic component in the population distribution and dynamics of several species (such as dry forests Arruda *et al.*, 2013; reviewed in Baker *et al.*, 2014). In open area biomes, such as savannas, it is expected that species more adapted to the drier and colder climates would have benefited from these periods and expanded their populations. The Cerrado, a savanna biome composed of several phytophysiognomies, has undergone multiple diversification processes but has only recently earned attention in this field compared to other biomes, such as Atlantic and the Amazon Forests (Werneck, 2011; Ribeiro, Werneck and Machado, 2016).

Several tools can be applied to investigate species response and the patterns resulting from those processes, such as phylogenetic analyses, population dynamics analyses, population genetic structuring, and niche modeling (e.g. Cabanne *et al.*, 2019; Buainain *et al.*, 2020; Campillo *et al.*, 2020). The use of these tools in an integrative way makes it possible to describe the most likely scenarios for the paths resulting in the biodiversity distribution that we see nowadays.

Considering that the climatic fluctuations that occurred in the Quaternary, especially in the Pleistocene, significantly altered the landscape it is expected that the species that occupy the same type of environments have responded similarly to these events. Thus, species from open areas, such as the Cerrado, we expect to find signs in populations structure in response to these periods of environmental change. For example, signs of population expansion in drier and cooler periods and signs of population contraction in agreement with climatically warmer and moister periods. To test the hypothesis that climate change that occurred in the late Quaternary would have affected the population dynamic of the open area bird species, we used two types of molecular markers and ecological niche modeling using two species of the genus *Elaenia* as models: the Plain-crested Elaenia (*E. cristata*) and the Lesser Elaenia (*E. chiriquensis*).

These congeneric species are suitable models for this study because they are widely distributed sympatrically in open areas of South America (Hosner, 2020; Hosner *et al.*, 2020). They have similarities in terms of diet (mainly frugivorous, but also insectivorous), in terms of the type of habitat they occur (Cerrado *sensu stricto*), breeding preferences (Medeiros and Marini, 2007; Marini *et al.*, 2009), but diverges in terms of migratory behavior, being *E. cristata* considered a resident species while *E. chiriquensis* is migratory/partially migratory (Alves, 2007; Somenzari *et al.*, 2018).

METHODS

Sample collection and DNA isolation

We sampled a total of 306 individuals (*E. cristata*, n=159; *E. chiriquensis*, n=147) across the species distribution in 31 localities from 2003 to 2018 (Figure 1; Table 1; Table S1). The birds were captured in the field using mist nets and banded to avoid duplicates. From each individual, we collected approximately 20 µl of blood with a sterile needle and glass capillary. The tissue was stored in absolute ethanol until processing in the laboratory. The specimens were released after collecting tissue samples and biological information. Trapping and tissue sampling were allowed by the Research Ethics Committees of University of Brasília (UnBDOC nº 75111/ 2013) by the ICMBio/ Sistema de Autorização e Informação em Biodiversidade, and Centro Nacional de Pesquisa e Conservação de Aves Silvestres (CEMAVE) /Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio)/Sistema Nacional de Anilhamento (SNA). To increase the sampled geographic area, we requested tissue samples from ornithological collections in Brazil. For further details of tissue's origin, localities, date of collection, see Supplementary Table 1.

Total genomic DNA was extracted from each sample following the proteinase K digestion and adapted purification protocol as in Friesen et al (1997). For samples obtained from ornithological collections, we used PureLink™ Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, US) following manufacturer instructions. Total Genomic DNA quality and concentrations were verified on a 1% agarose gel stained with ethidium bromide and using the Qubit™ dsDNA BR Assay Kit (ThermoFisher, Waltham, MA, US), respectively.

Molecular markers

We used two datasets of different molecular markers: mitochondrial DNA (fragment of gene NADH dehydrogenase subunit-2 - ND2) and single nucleotide polymorphism (SNPs) from

the ddRadSeq loci from both species to perform a comparative analysis of genetic diversity and population structure.

ddRadSeq libraries and sequencing

The SNP data were generated as described by Peterson et al (2012) with modifications (see Thrasher *et al.*, 2018) and detailed in Freitas et al (2022). In a few words, each sample was digested with *SbfI* and *MspI* and ligated adapters that allowed multiplexing. The libraries, each containing approximately 20 samples, were size-selected and PCR-enriched, incorporating the Illumina HiSeq adapters (Illumina, San Diego, California, US). Finally, all groups of samples were combined in equimolar proportions and sequenced, single end 100bp, on two lanes of an Illumina HiSeq 2500.

After assessing read quality with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), we used FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit) to trim sequences to 98 bp to discard lower-quality base calls at the 3′ end of the sequence. Subsequently, we used FASTX-Toolkit to retain reads without a single base below a Phred quality score of 10 and with at least 95% of bases with quality above 20. We demultiplexed reads using the 'process_radtags' program from the STACKS v.2.41 bioinformatics pipeline (Catchen *et al.*, 2011, 2013), discarding reads that did not pass the Illumina filter, had barcode contamination, lacked and *SbfI* cut site or one of the unique barcodes used for multiplexing at the 5′ end. We obtained an average of 550,757 (±263,678 reads) quality-filtered reads per individual (Supplementary Table S1).

We used the same parameters as described in Freitas et al (2022) to export the SNPs. Briefly, we removed samples that had more than 80% of missing data after the *de novo* pipeline from STACKS. We used the 'populations' software at STACKS, and we retained loci that were present in at least 80% of the individuals, in two different datasets: (i) *E. cristata* (n=154); and (ii) *E. chiriquensis* (n=144).

Mitochondrial gene: amplification and sequencing

We amplified and sequenced a fragment of mitochondrial gene NADH dehydrogenase subunit-2 (ND2) as previously detailed described (Freitas *et al.*, 2022), from 122 individuals of *E. cristata* and 96 of *E. chiriquensis*. Briefly, we prepared a PCR reaction with primer pair LMET and H6313 (Sorenson *et al.*, 1999) at 10 µM of each, Phusion High-Fidelity PCR Kit (New

England BioLabs, Ipswich, Massachusetts, US), 10 mM each dNTP, 0.4U of Taq Polymerase, and approximately 50 ng of genomic DNA at 20 µl final volume. The reactions were performed in a thermocycler at an annealing temperature of 54ºC for 30 cycles. The PCR products were treated with Exonuclease (10U/µl) and Shrimp Alkaline Phosphatase (1.0U/µl) (ThermoFisher) per 10 µl of PCR product. The PCR products were sequenced at the Cornell University Biotechnology Resource Center - Genomics Facility and MACROGEN Inc.

The chromatograms of sequences were individually checked by eye using the software Geneious Prime v.2019.0.4 and aligned using the Clustal W multiple alignment algorithms (Thompson, Higgins and Gibson, 1994) implemented by Bioedit v. 7.0.5.3 (Hall, 1999). We produced three data sets: (i) both species, (ii) *Elaenia cristata* samples and, (iii) *Elaenia chiriquensis* samples. Sequences of Greenish Elaenia (*Myiopagis viridicata*) were used as an outgroup as in previous studies (Tang, Edwards and Rheindt, 2018), obtained from GeneBank (accession numbers: EU310967.1, EF501937.1, EF501908.1, FJ175934.1, and FJ175933.1).

Population genetic structure mtDNA ND2

Molecular diversity indexes for each species, as nucleotide diversity (Pi), the number of haplotypes (h), and haplotype diversity (H_d), were calculated using the DnaSP v 6.12.03 software (Rozas *et al.*, 2017) for both species.

We investigated the relationship among the haplotypes by performing a median-joining network (Bandelt, Forster and Röhl, 1999) for each species on PopART v1.7 (http://popart.otago.ac.nz). The data set used was the nexus file generated by DnaSP software with the haplotype information for each species: *E. cristata* (H_d=32) and *E. chiriquensis* (H_d=45).

As previous studies demonstrated a moderate genetic structure in *E. cristata* species using the SNPs dataset (Freitas *et al.*, 2022), we grouped the samples in the same seven populations indicated for further analyses with the molecular marker mtDNA ND2. The seven populations are Center-North Cerrado (POP1), South-East Cerrado (POP2), South-West Cerrado (POP3), West Amazon (POP4), North-East Amazon (POP5), North-West Amazon (POP6), and Caatinga (POP7) (Fig. 3A).

A Bayesian analysis of genetic population structuring was performed using software BAPS v.6 (Corander and Tang, 2007), with the dataset of the mtDNA from 122 individuals of *E. cristata*. We set the maximum number of populations to seven and ran five replicates of each of 100 iterations. The same analysis was performed for *E. chiriquensis* species, but the maximum number of populations was set to five.

The following analysis was performed only for *E. cristata* considering we were not able to define populations for *E. chiriquensis*. To measure the level of gene flow between the populations using the mtDNA marker, we employed the 'Population comparison' in the software Arlequin v3.5 (Excoffier and Lischer, 2010) to compute pairwise F_{ST} (nº of permutations = 1000, allowed level of missing data = 0.05). We perform Fu's Fs neutrality tests. We also performed AMOVA (Analysis of Molecular Variance) to detect the amount of genetic structuring.

Phylogenetic gene tree

A mitochondrial gene tree for each species was estimated from ND2 haplotypes using MrBayes v3.1.2. We used the HKY (Hasegawa, Kishino and Yano, 1985) model with gamma rates for the first base position, and the General Time Reversible model (Tavaré, 1986) with a proportion of invariable sites for the second and third positions, those were selected from PartitionFinder2 analyses (Lanfear *et al.*, 2016).

Estimated Effective Migration Surfaces (EEMS) analysis

We visualized the relationship between genetics and geographic space using the Estimated Effective Migration Surfaces rates (Petkova, Novembre and Stephens, 2016). This is a method to visualize spatial population structure based on geo-referenced genetic data with the estimated migration (m) rates. We used as an input file the dataset for the *E. cristata* of 4,237 SNPs and *E. chiriquensis* of 3,404 SNPs. We calculated pairwise genetic dissimilarities with the bed2diffs_v2 function (https://github.com/dipetkov/eems). We set a number of 150 demes and run independent analyses of the MCMC for 9,000,000 steps with a burn-in of 1,000,000 for each species. The results were plotted using the rEEMSplots R package (https://github.com/dipetkov/eems).

Ecological Niche Modeling

We constructed the distribution models for the *E. cristata* and *E. chiriquensis* using the ecological niche modeling technique (ENMs). The species records were mapped in a grid of cells with a resolution of 2.5 minutes of arc (~5 km × 5 km). A total of 7 142 occurrence records were obtained for *E. cristata* and 15 373 for *E. chiriquensis*, through field collection, database, and literature. Then, spatial correlations were made using Sdmtoolbox (Brown, 2014), reducing the database to 711 observations for *E. cristata* and 1374 for *E. chiriquensis* (Table S2 and Table S3). We selected one point per cell and removed repeated points from the grid, as well as those that

contained the same environmental information. We used bioclimatic data provided by WorldClim 1.4 (Hijmans *et al.*, 2005), which offers 19 environmental variables, for the present, Mid-Holocene - HM (ca 6,000 yr ago), the Last Glacial Maximum - LGM (ca 21,000 yr ago), and the Last Interglacial - LIG (ca 120,000 yr ago). To avoid correlated climatic variables, Spearman correlation tests were performed ($r < 0.7$, Figure S2).

We used six different algorithms in modeling methods to estimate the potential distribution of each species: (1) Bioclim (Nix 1986), (2) GLM (Guisan, Edwards and Hastie, 2002), (3) Domain (distance by Gower) (Carpenter, Gillison and Winter, 1993); (4) RandomForest (Liaw and Wiener, 2002); (5) Maximum entropy - Maxent (Phillips and Dudík, 2008) and (6) Support vector machines (SVM) (Tax and Duin, 2004). The SVM algorithm was executed using the ksvm function of the "kernlab" package (Karatzoglou *et al.*, 2004), and RandoForester was executed by the "randomForest" package. All other algorithms were executed in the "dismo" package (Hijmans *et al.*, 2015). Following the approach of the ensemble (Araújo and New, 2007), the binary maps were concatenated through frequency of predicted presence (Using LPT - lowest presence threshold), producing a single consensus map by the algorithm and climatic scenario (present, Middle Holocene, Last Glacial Maximum and Last Interglacial). Likewise, we obtained the final consensus map for each climatic scenario by calculating the frequencies of all algorithms. All models were developed in the software R.

To calibrate the models based on presence-background observations (SVM and Maxent), background points were selected randomly along a grid of cells from South America giving input for the evaluation of the models. For each model, 20 replicates were generated, adjusted by a partition criterion double (75% and 25%). The 75% and 25% of the points were selected at random for each of the 20 replicates in each algorithm. To evaluate the models, we used the True Skill Statistic (TSS) where the values can vary from -1 to 1; negative values or close to zero indicate that the model's predictions are no different from a randomly generated model, while models with values closer to 1 are considered very good. In general, models with TSS values above 0.5 are considered acceptable (Allouche, Tsoar and Kadmon, 2006). To generate the maps of stable areas, we overlay the binary maps of each scenario for each species.

RESULTS

Data set

We obtained the final dataset derived from the ddRadSeq approach composed of 298 samples (*E. cristata* n = 154; *E. chiriquensis* n = 144), with the mean coverage of 29.86x (+/-

9.06). We removed eight samples that obtained low quality (*E. cristata* n = 5; *E. chiriquensis* n = 3).

Concerning the mitochondrial marker, we obtained a matrix of 1035pb for both species composed of 122 sequences from *E. cristata*, and 96 from *E. chiriquensis* of the ND2 fragment. From *E. cristata* were identified 32 haplotypes (H_d:0.857; Pi: 0.0022) and 45 haplotypes from *E. chiriquensis* (H_d:0.943; Pi:0.00338).

Population genetic structuring mtDNA ND2

The haplotype network for the species *E. chiriquensis* showed no spatial population differentiation discernible at the sampled areas (Figure 2B). The resident species, Plain-crested Elaenia, showed moderate population structuring using the mitochondrial marker ND2 when compared to the results found with SNPs. There were identified two most common haplotypes with wide geographic representation. The POP7 (orange color, Fig. 4D) sampled in the Caatinga showed some differentiation compared to the other locations.

The Bayesian analysis of genetic population structuring reveals the number of clusters in the optimal partition of three (Fig. 3B). Cluster 1 comprises 52 individuals of 122 total, and it is unique in three sites sampled (Fig.3B, red). The second cluster assembles 66 samples of 122 total and shares 14 locals being exclusive in seven sites sampled (Fig. 3B, green). The third group comprises 4 individuals and is isolated in two west sampled areas (Fig. 3B, blue).

Despite the Bayesian analysis having indicated only three groups in the total sample, the comparisons of pairs of populations samples from *E. cristata* showed a significant genetic difference for 15 F_{ST} values of 21 comparisons. The populations POP4, POP6, and POP7 are significantly different from all other populations and each other (For population details see Fig. 2A; Fig. 3C; Table S4, and S5). The difference between the population of the southwest and central Cerrado was not significant. Comparisons between the population of the northeast (POP7) and the others obtained the most significant values (Tables S4, S5). Based on AMOVA results (Table S6) the highest percentage of genetic variation in *E. cristata* is within populations (79.46%) while only 20.54% is among groups. This comparison was also performed using the seven previously defined populations.

Results of effective migration rates for *E. cristata* (Fig. 4A) and *E. chiriquensis* (Fig. 4B). Values expressed in Log(m) express the levels of genetic similarity across the geographic space and range from 2 (highest) to -2 (smallest). The resulting maps predominantly exhibit the Amazon Forest area as an impacting factor in the genetic similarity between the sampled localities for both

species. The Caatinga biome also exhibits a slight differentiation, already demonstrated in previous analyses, but a more extensive sampling of this area would be necessary to confirm it.

Phylogenetic gene tree

Gene trees built using the haplotypes do not reveal a clear geographic relationship association in both species (Figure S1). For *E. cristata* species, the two most frequent haplotypes were found in more recent positions and there was a basal branch formed by three haplotypes present only in the western region of Amazonia (POP 4, Fig. 3A). The species *E. chiriquensis* does not show clear relationships between haplotypes, resulting in several polytomies in the gene tree.

Ecological Niche Modeling

The climatic models for the two species of the *Elaenia* genus exhibited AUC values > 0.9 and TSS values of 0.60 for *E. cristata* and 0.58 for *E. chiriquensis* (Table S8). For both species, the temporal scenarios showed that the LGM (~21 thousand years) was potentially the time with the greatest suitable area available for the species, while in the LIG (~120 thousand years) there was a little area with climatic potential for the species (Fig. 5). For both species, there was a decrease in the suitable potential area in the HM (~6 thousand years), where *E. chiriquensis* was more concentrated in the regions that today are humid regions of South America, while, even with smaller areas, *E. cristata* it still had a potential distribution that was concentrated from the north of the Amazon to the northeast of South America. When comparing the past and present scenarios, there was a significant expansion in the LGM, followed by a retraction in the HM and an increase in the area potentially suitable for the species in the current times, despite, preserving the distribution pattern of the HM for the present to both species (Fig. 5).

Regarding the climatic stability areas for the species, the results illustrated that for *E. cristata* the central and northeastern regions of South America as areas that over time have been in potential conditions for the presence of the species (Fig. 6A). For *E. chiriquensis*, the northern region of the Amazon biome jointly with a central portion of the Atlantic Forest were the areas that have remained stable over the years (Fig. 6B). These climatically stable areas coincide with the current distribution of the species, showing that these areas may have served as a refuge for the species during past climate changes.

DISCUSSION

The savanna environment has already been indicated as a crucial factor in the diversification of species from the genus *Elaenia* (Rheindt, Christidis and Norman, 2008). Reconstructions of ancestral habitat preferences showed *E. cristata* and *E. ruficeps* as specialists in savannas, as those species are the basal clade in the phylogeny of the genus possible being a proto-*Elaenia* first colonized this biome. These findings are supported by dating with molecular clock analysis using the mitochondrial ND2 gene (Rheindt, Christidis and Norman, 2008). Consequently, we assume that the climatic and geological processes experienced by this biome may also have an impact on the genetic distribution of these species. The Lesser Elaenia has a strong association with this biome, despite being an intratropical migratory species, it is dependent on this type of the Cerrado *sensu stricto* for breeding (Medeiros and Marini, 2007) and showed a niche-tracking behavior during migration, i.e., it pursues similar habitats when moving between areas of breeding and breeding rest. (Guaraldo, Kelly and Marini, 2016).

Based on our results for the ecological niche modeling and the estimations of suitable area available for *E. chiriquensis* during the Pleistocene climatic fluctuations, we expected to find a greater population structure than such as seen in its congeneric *E. cristata*. As there is an indication of greater loss of climatically suitable areas for these species, the populations would have been more isolated in small fragments of habitat over time, decreasing gene flow among the populations.

However, our results suggest that there is no population structuring in *E. chiriquensis* using the ND2 marker, the star-shaped network (Fig. 2), and the lack of geographic association indicate a probable high gene flow in the species, likely caused by the migratory behavior as seen previously using SNPs as molecular markers (Freitas *et al.*, 2022). Furthermore, microsatellite data demonstrated that migratory behavior increases gene flow between populations when compared to resident species (*Troglodytes aedon* and *T. musculus*, respectively), promoting a shallow division between populations for migrant species (Arguedas and Parker, 2000). Low genetic differentiation between long-distance sampled areas was found by Bates et al (2003) when using two mitochondrial markers from *E. chiriquensis* individuals sampled in the savannas of Amapá - Brazil and Santa Cruz – Bolívia.

Our results for mtDNA ND2 recovered low genetic structuring for *E. cristata* species (K=3) as found by Ritter et al (2021) in a similar study. Both studies observed no structuring of *E. cristata* exhibited in the Central Brazilian Cerrado area, but our haplotype network showed a distinct group composed of samples from the Caatinga (POP7, Fig. 3D) and another by samples from the western region of the Amazon (POP 4, Fig 3D), this difference is possibly due to the

additional samples and locations included in this study. We have shown using SNPs as molecular markers at least five populations of *E. cristata* and estimated substructuring of possibly seven populations across the sampled areas (Freitas *et al.*, 2022) (Figure 2A).

SNPs and mitochondrial DNA recovery distinguished signals across the evolutionary process, revealing results of more recent events (mtDNA) or more ancient changes (SNPs). Following our data, the resident *E. cristata*, which passed through similar climate historical processes of *E. chiriquensis*, showed moderate population structuring compared to the later. The two major groups recovered by network analysis (Fig. 3D) may reflect the results of a recent event of high gene flow among the populations of *E cristata* that occurred during the periods of expansion of the Cerrado in accordance with our ENMs results. We also found signs of expansion in POP2, POP5, and POP6 populations according to Fu's FS tests ($P < 0.02$). Those populations are currently sampled in the boundary between savannas-Amazon Forest (Fig. 3A), and this signal could be a response of the species to the expansion of climatically suitable areas for the species during the LGM (Fig. 5A). Ritter et al (2021) using population expansion analysis and time divergence estimations suggest that these episodes of higher gene flow in *E. cristata* occurred during the LGM period.

Effective migration rates for *E. cristata* and *E. chiriquensis* suggest the Amazon Forest as a barrier for both species (Fig. 4). Congruent with the results with molecular markers, visualizing areas of lower-than-average gene flow i.e., putative barriers to gene flow, isolating populations in northern and western Amazonia from the Cerrado Central portion for *E. cristata* (Fig. 4A), while genetic similarity decay in the east-west direction for *E. chiriquensis* (Fig. 4B). Both species show a discrete difference from the Caatinga samples (sites 18 and 19; Fig.4), but a larger sampling would be necessary for robust results.

To summarize, our results using mitochondrial markers and ecological niche modeling suggest that *E. cristata* e *E. chiriquensis* species likely underwent similar climatic processes during Quaternary climatic fluctuations and, despite this, the results in the structuring of populations were different for each species. Despite the climatic scenario, through ecological niche modeling and models of stable areas, indicating greater habitat fragmentation for *E. chiriquensis*, leading to possible isolation of populations and greater genetic structure, the migratory behavior may have maintained a higher gene flow in the species in its distribution. The resident species showed signs of expansion of some populations in climatic periods favorable to the growth of open areas in contrast to the loss of space of forest formations.

CONCLUSIONS

Our results demonstrate that climatic changes that occurred in the Quaternary, specifically in the Pleistocene, may have affected the distribution of bird species in open areas, such as *E. cristata* and *E. chiriquensis*. However, intrinsic characteristics of the species, such as migration, may restrict the identification of a general pattern for these areas due to the impact on the population structuring.

ACKNOWLEDGMENTS

CAPES and CNPq for the fellowship and funding this study, Cornell Lab of Ornithology – Cornell University, Ornithological Collections for sending the samples, CEMAVE.

Trapping and tissue sampling were allowed by the Research Ethics Committees of University of Brasília (UnBDOC nº 75111/ 2013) by the ICMBio/ Sistema de Autorização e Informação em Biodiversidade, and Centro Nacional de Pesquisa e Conservação de Aves Silvestres (CEMAVE) /Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio)/Sistema Nacional de Anilhamento (SNA).

REFERENCES

Allouche, O., Tsoar, A. and Kadmon, R. (2006) 'Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS)', *Journal of Applied Ecology*, 43(6), pp. 1223–1232. doi: 10.1111/j.1365-2664.2006.01214.x.

Alves, M. A. S. (2007) 'Sistemas de migrações de aves em ambientes terrestres no Brasil: Exemplos, lacunas e propostas para o avanço do conhecimento', *Revista Brasileira de Ornitologia*, 15(2), pp. 231–238.

Araújo, M. B. and New, M. (2007) 'Ensemble forecasting of species distributions', *Trends in Ecology & Evolution*, 22(1), pp. 42–47. doi: 10.1016/j.tree.2006.09.010.

Arguedas, N. and Parker, P. G. (2000) 'Seasonal Migration and Genetic Population Structure in House Wrens', *The Condor*, 102(3), pp. 517–528. doi: 10.2307/1369782.

Arruda, D. M. *et al.* (2013) 'Phytogeographical patterns of dry forests sensu stricto in northern Minas Gerais State, Brazil', *Anais da Academia Brasileira de Ciências*, 85(2), pp. 623–634. doi: 10.1590/S0001-37652013000200011.

Baker, P. A. *et al.* (2014) 'The emerging field of geogenomics: Constraining geological problems with genetic data', *Earth-Science Reviews*, 135, pp. 38–47. doi: 10.1016/j.earscirev.2014.04.001.

Bandelt, H. J., Forster, P. and Röhl, A. (1999) 'Median-joining networks for inferring intraspecific phylogenies.', *Molecular Biology and Evolution*, 16(1), pp. 37–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10331250.

Bates, J. M., Tello, J. G. and Silva, J. M. C. da (2003) 'Initial Assessment of Genetic Diversity in Ten Bird Species of South American Cerrado', *Studies on Neotropical Fauna and Environment*, 38(2), pp. 87–94. doi: 10.1076/snfe.38.2.87.15924.

Brown, J. L. (2014) 'SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses', *Methods in Ecology and Evolution*. Edited by B. Anderson, 5(7), pp. 694–700. doi: 10.1111/2041-210X.12200.

Buainain, N. *et al.* (2020) 'Paleoclimatic evolution as the main driver of current genomic diversity in the widespread and polymorphic Neotropical songbird Arremon taciturnus', *Molecular Ecology*, 29(15), pp. 2922–2939. doi: 10.1111/mec.15534.

Burney, C. W. and Brumfield, R. T. (2009) 'Ecology Predicts Levels of Genetic Differentiation in Neotropical Birds', *The American Naturalist*, 174(3), pp. 358–368. doi: 10.1086/603613.

Cabanne, G. S. *et al.* (2019) 'Phylogeographic variation within the Buff-browed Foliage-gleaner (Aves: Furnariidae: Syndactyla rufosuperciliata) supports an Andean-Atlantic forests connection via the Cerrado', *Molecular Phylogenetics and Evolution*, 133(January), pp. 198–213. doi:

10.1016/j.ympev.2019.01.011.

Campillo, L. C. *et al.* (2020) 'Genomic differentiation in an endemic Philippine genus (Aves: Sarcophanops) owing to geographical isolation on recently disassociated islands', *Biological Journal of the Linnean Society*, 131(4), pp. 814–821. doi: 10.1093/biolinnean/blaa143.

Carpenter, G., Gillison, A. N. and Winter, J. (1993) 'DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals', *Biodiversity and Conservation*, 2(6), pp. 667–680. doi: 10.1007/BF00051966.

Catchen, J. *et al.* (2013) 'Stacks: an analysis tool set for population genomics', *Molecular Ecology*, 22(11), pp. 3124–3140. doi: 10.1111/mec.12354.

Catchen, J. M. *et al.* (2011) 'Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences', *Genes|Genomes|Genetics*, 1(3), pp. 171–182. doi: 10.1534/g3.111.000240.

Corander, J. and Tang, J. (2007) 'Bayesian analysis of population structure based on linked molecular information', *Mathematical Biosciences*, 205(1), pp. 19–31. doi: 10.1016/j.mbs.2006.09.015.

Excoffier, L. and Lischer, H. E. L. (2010) 'Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows', *Molecular Ecology Resources*, 10(3), pp. 564–567. doi: 10.1111/j.1755-0998.2010.02847.x.

Freitas, E. L. de *et al.* (2022) 'Ecological traits drive genetic structuring in two open-habitat birds from the morphologically cryptic genus Elaenia (Aves: Tyrannidae)', *Journal of Avian Biology*. doi: 10.1111/jav.02931.

Friesen, V. L. *et al.* (1997) 'Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms', *Molecular Ecology*, 6(11), pp. 1047–1058. doi: 10.1046/j.1365-294X.1997.00277.x.

Gómez-Bahamón, V. *et al.* (2020) 'Speciation Associated with Shifts in Migratory Behavior in an Avian Radiation', *Current Biology*, 30(7), pp. 1312-1321.e6. doi: 10.1016/j.cub.2020.01.064.

Guaraldo, A. C., Kelly, J. F. and Marini, M. Â. (2016) 'Contrasting annual cycles of an intratropical migrant and a tropical resident bird', *Journal of Ornithology*, 157(3), pp. 695–705. doi: 10.1007/s10336-016-1327-5.

Guisan, A., Edwards, T. C. and Hastie, T. (2002) 'Generalized linear and generalized additive models in studies of species distributions: setting the scene', *Ecological Modelling*, 157(2–3), pp. 89–100. doi: 10.1016/S0304-3800(02)00204-1.

Hall, T. A. (1999) 'BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT', *Nucleic Acids Symposium Series*, 41, pp. 95–98. doi: citeulike-article-id:691774.

Hasegawa, M., Kishino, H. and Yano, T. (1985) 'Dating of the human-ape splitting by a molecular clock of mitochondrial DNA', *Journal of Molecular Evolution*, 22(2), pp. 160–174. doi: 10.1007/BF02101694.

Hijmans, R. J. *et al.* (2005) 'Very high resolution interpolated climate surfaces for global land areas', *International Journal of Climatology*, 25(15), pp. 1965–1978. doi: 10.1002/joc.1276.

Hijmans, R. J. *et al.* (2015) 'Dismo: Species Distribution Modeling'.

Hosner, P. *et al.* (2020) 'Lesser Elaenia (Elaenia chiriquensis)', in del Hoyo, J. et al. (eds) *Birds of the World*. Version1. Ithaca, NY, USA: Cornell Lab of Ornithology. doi: 10.2173/bow.lesela1.01.

Hosner, P. (2020) 'Plain-crested Elaenia (Elaenia cristata)', in del Hoyo, J. et al. (eds) *Birds of the World*. version 1. Ithaca, NY, USA: Cornell Lab of Ornithology. doi: 10.2173/bow.plcela1.01.

Karatzoglou, A. *et al.* (2004) 'kernlab - An S4 Package for Kernel Methods in R', *Journal of Statistical Software*, 11(9), pp. 389–393. doi: 10.18637/jss.v011.i09.

Lanfear, R. *et al.* (2016) 'PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses', *Molecular Biology and Evolution*. doi: 10.1093/molbev/msw260.

Liaw, A. and Wiener, M. (2002) 'Classification and Regression by randomForest', *R News*, 2(3), pp. 18–22.

Marini, M. *et al.* (2009) 'Biologia reprodutiva de Elaenia cristata (Aves: Tyrannidae) em cerrado do Brasil central', *Neotropical Biology and Conservation*, 4(1), pp. 2–12. doi: 10.4013/nbc.2009.41.01.

Medeiros, R. de C. S. de and Marini, M. Â. (2007) 'Biologia reprodutiva de Elaenia chiriquensis (Lawrence) (Aves, Tyrannidae) em Cerrado do Brasil Central', *Revista Brasileira de Zoologia*, 24(1), pp. 12–20. doi: 10.1590/S0101-81752007000100002.

Oliveira, U., Vasconcelos, M. F. and Santos, A. J. (2017) 'Biogeography of Amazon birds: rivers limit species composition, but not areas of endemism', *Scientific Reports*, 7(1), p. 2992. doi: 10.1038/s41598-017-03098-w.

Peterson, B. K. *et al.* (2012) 'Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species', *PLoS ONE*. Edited by L. Orlando, 7(5), p. e37135. doi: 10.1371/journal.pone.0037135.

Petkova, D., Novembre, J. and Stephens, M. (2016) 'Visualizing spatial population structure with estimated effective migration surfaces', *Nature Genetics*, 48(1), pp. 94–100. doi: 10.1038/ng.3464.

Phillips, S. J. and Dudík, M. (2008) 'Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation', *Ecography*, 31(2), pp. 161–175. doi: 10.1111/j.0906-

7590.2008.5203.x.

Rheindt, F. E., Christidis, L. and Norman, J. A. (2008) 'Habitat shifts in the evolutionary history of a Neotropical flycatcher lineage from forest and open landscapes', *BMC Evolutionary Biology*, 8(1), p. 193. doi: 10.1186/1471-2148-8-193.

Ribeiro, V., Werneck, F. P. and Machado, R. B. (2016) 'Distribution dynamics of South American savanna birds in response to Quaternary climate change', *Austral Ecology*, 41(7), pp. 768–777. doi: 10.1111/aec.12363.

Ritter, C. D. *et al.* (2021) 'Sister species, different histories: comparative phylogeography of two bird species associated with Amazonian open vegetation', *Biological Journal of the Linnean Society*, 132(1), pp. 161–173. doi: 10.1093/biolinnean/blaa167.

Rozas, J. *et al.* (2017) 'DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets', *Molecular Biology and Evolution*, 34(12), pp. 3299–3302. doi: 10.1093/molbev/msx248.

Somenzari, M. *et al.* (2018) 'An overview of migratory birds in Brazil', *Papéis Avulsos de Zoologia*, 58, p. 3. doi: 10.11606/1807-0205/2018.58.03.

Sorenson, M. D. *et al.* (1999) 'Primers for a PCR-Based Approach to Mitochondrial Genome Sequencing in Birds and Other Vertebrates', *Molecular Phylogenetics and Evolution*, 12(2), pp. 105–114. doi: 10.1006/mpev.1998.0602.

Tang, Q., Edwards, S. V and Rheindt, F. E. (2018) 'Rapid diversification and hybridization have shaped the dynamic history of the genus Elaenia', *Molecular Phylogenetics and Evolution*, 127(May), pp. 522–533. doi: 10.1016/j.ympev.2018.05.008.

Tavaré, S. (1986) 'Some probabilistic and statistical problems in the analysis of DNA sequences', *Lectures on mathematics in the life sciences*, 17(2), pp. 57–86.

Tax, D. M. J. and Duin, R. P. W. (2004) 'Support Vector Data Description', *Machine Learning*, 54(1), pp. 45–66. doi: 10.1023/B:MACH.0000008084.60811.49.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) 'CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice', *Nucleic acids research*, 22(22), pp. 4673–4680.

Thrasher, D. J. *et al.* (2018) 'Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird', *Molecular Ecology Resources*, 18(5), pp. 953–965. doi: 10.1111/1755-0998.12771.

Vasconcellos, M. M. *et al.* (2019) 'Isolation by instability: Historical climate change shapes population structure and genomic divergence of treefrogs in the Neotropical Cerrado savanna', *Molecular Ecology*, 28(7), pp. 1748–1764. doi: 10.1111/mec.15045.

Werneck, F. P. (2011) 'The diversification of eastern South American open vegetation biomes:

Historical biogeography and perspectives', *Quaternary Science Reviews*, 30(13–14), pp. 1630–1648. doi: 10.1016/j.quascirev.2011.03.009.

ORCID

Eliane Luiz de Freitas: https://orcid.org/0000-0002-8072-3916

Victor Leandro-Silva: https://orcid.org/0000-0001-9985-5532

Leonardo Campagna; https://orcid.org/0000-0001-8408-8553

Irby Lovette: https://orcid.org/0000-0003-0268-4875

Renato Caparroz: https://orcid.org/0000-0003-4883-4023

Data Accessibility

Sequence raw data will be deposited in Dryad.

Mitochondrial DNA sequences will be deposited in GenBank.

TABLES AND FIGURES

TABLE 1. Specification of the samples of *E. cristata* and *E. chiriquensis* collected across their distribution. Locality number (Nº) corresponds to those shown in Figures 2, 3, and 4. State, collection site refers to the site in Brazil. The sample size for each locality is also shown. The number of individuals used per analysis may vary, to further details see the Methods section.

Nº	**State; collection site**	**Latitude**	**Longitude**	***E. cristata***	***E. chiriquensis***
1	Roraima; Viruá National Park	1.42	-60.98	2	0
2	Amapá; Cerrado Experiment Field of the Embrapa Amapá	0.39	-51.05	16	2
3	Pará; Óbidos	0.63	-55.72	0	7
4	Pará; Oriximiná	-1.76	-56.22	10	4
5	Pará; Monte Alegre	-2.02	-54.18	3	0
6	Pará; Portel	-1.85	-50.70	10	1
7	Pará; Marajó Island	-0.78	-48.61	10	12
8	Pará; Parauapebas	-6.28	-50.58	1	0
9	Amazonas; Manicoré	-8.47	-61.39	4	1
10	Rondônia; Machadinho	-8.92	-62.07	5	0
11	Maranhão; Urbano Santos	-3.28	-44.33	1	0
12	Maranhão; Chapada das Mesas National Park	-7.13	-47.15	8	11
13	Piauí; Pirarucura	-4.11	-41.71	0	1
14	Piauí; Castelo do Piauí	-5.33	-41.57	2	3
15	Piauí; Guadalupe	-6.74	-43.72	0	1
16	Piauí; Serra das Confusões	-9.37	-43.82	0	1
17	Piauí; Uruçuí	-7.73	-44.54	7	0
18	Rio Grande do Norte; Rio Fogo	-5.29	-35.40	10	0
19	Rio Grande do Norte; Parnamirim	-5.92	-35.17	6	2
20	Tocantins; Ponte Alta do Tocantins	-10.76	-47.48	4	1
21	Mato Grosso; Guarantã do Norte	-10.12	-54.36	0	1
22	Mato Grosso; Vila Bela da Santíssima Trindade	-14.63	-60.20	10	3
23	Mato Grosso; Chapada dos Guimarães National Park	-15.40	-55.83	10	10

24	Mato Grosso; Itiquira	-17.21	-54.14	0	2
25	Mato Grosso; Araguaia	-17.42	-53.45	0	1
26	Mato Grosso; Nova Xavantina	-14.78	-52.53	5	0
27	Goiás; Emas National Park	-17.92	-52.97	2	1
28	Goiás; Chapada dos Veadeiros National Park	-14.16	-47.74	2	8
29	Distrito Federal; Protection Area Gama and Cabeça de Veado	-15.92	-47.87	15	73
30	Distrito Federal; Águas Emendadas Ecological Station	-15.55	-47.59	3	1
31	Minas Gerais; Grande Sertão Veredas National Park	-15.18	-45.69	9	1
32	Minas Gerais; Mateus Leme	-20.14	-44.45	0	2
33	Minas Gerais; Serra Azul	-20.14	-44.41	1	0
	Total			**156**	**150**

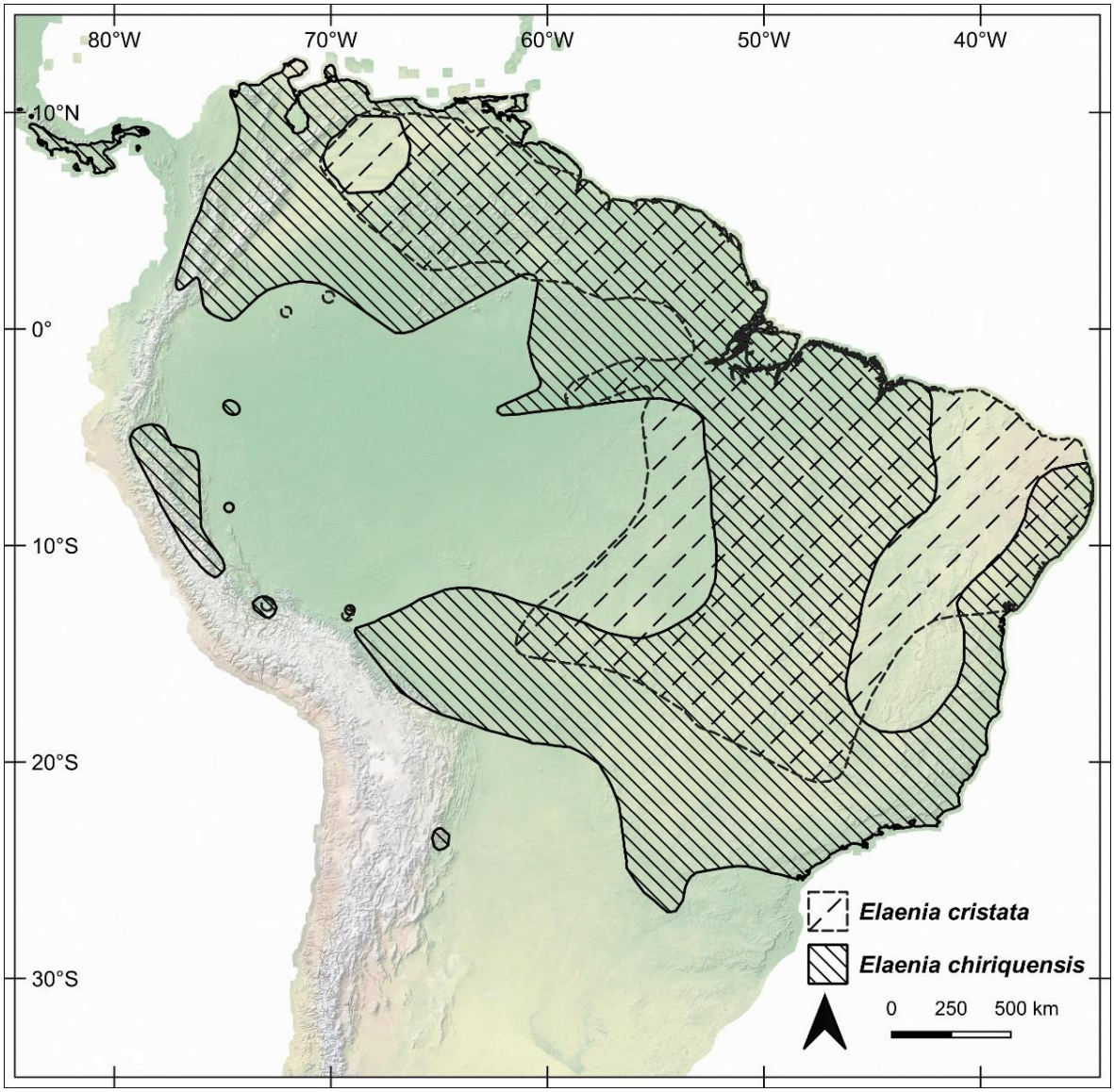


Figure 1. The distribution area of the species Plain-crested Elaenia (*Elaenia cristata*, dashed line) and the Lesser Elaenia (*Elaenia chiriquensis*, solid line) according to the International Union for Conservation of Nature – IUCN.

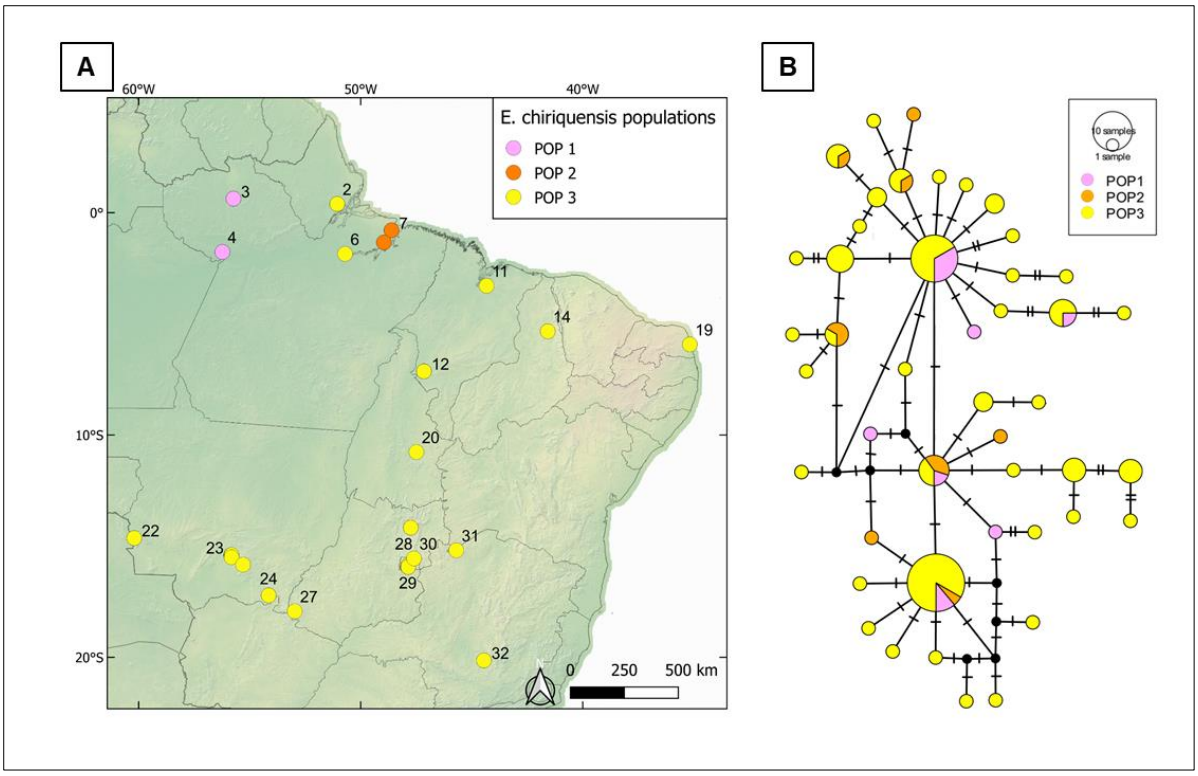


Figure 2. (A) Possible *Elaenia chiriquensis* population structuring was observed by Freitas et al. (2022) based on single nucleotide polymorphisms - SNPs. Numbers refer to the sampled locations (Table 1). (B) Haplotype network built using mitochondrial fragment ND2 and color-coded according to the three populations in A.

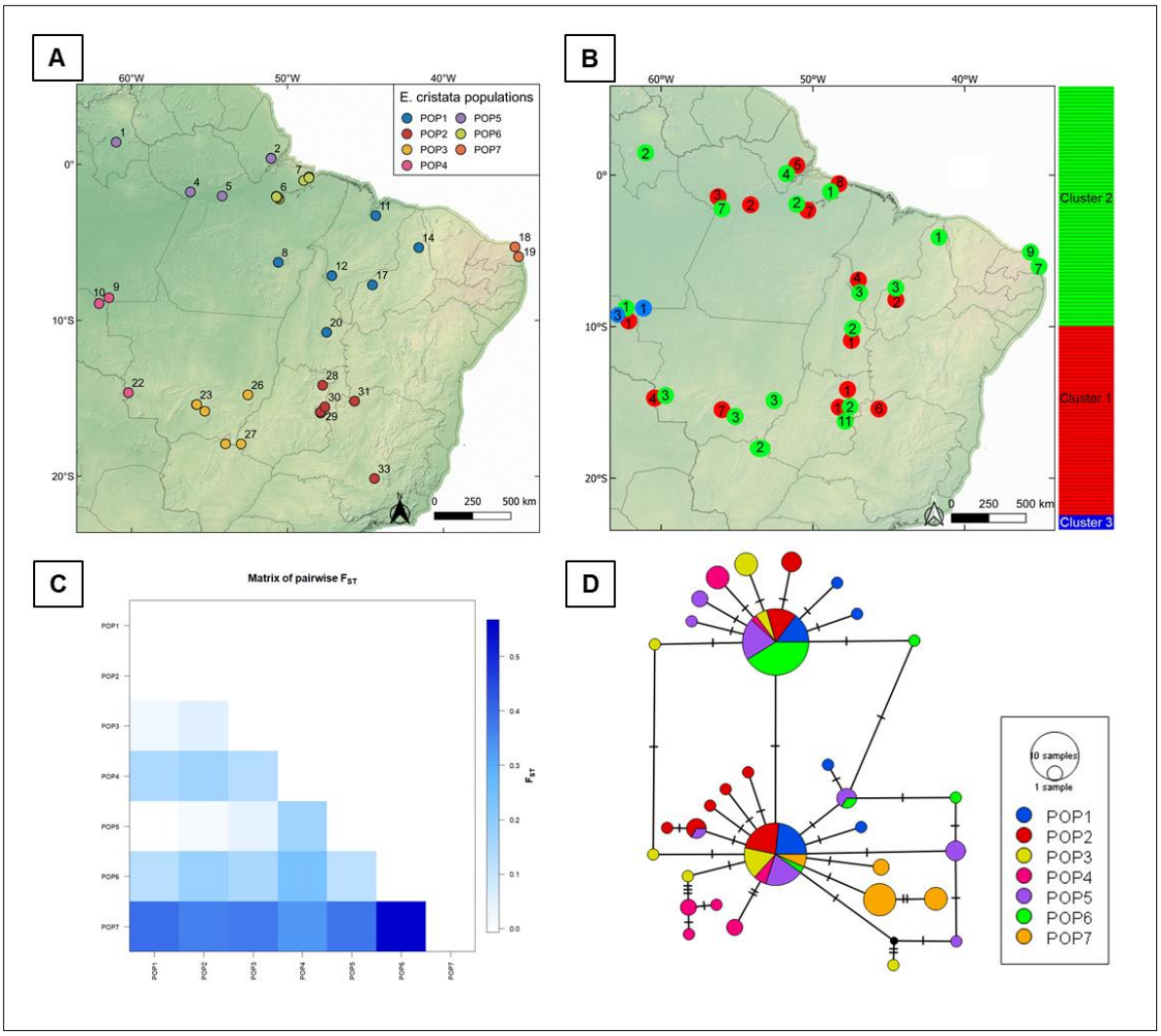


Figure 3. *Elaenia cristata* population structuring based on the mitochondrial molecular marker (mtDNA). (A) The population structure of *E. cristata* was observed by Freitas et al. (2022) based on single nucleotide polymorphisms (SNPs). Numbers refer to the sampled locations (Table 1). The results presented in panels B, C and D were obtained using 122 mtDNA samples from the species *E. cristata*. (B) Bayesian analysis of genetic population structuring. Numbers inside the circles refer to the number of samples in that cluster/locality. (C) Graphic demonstration of F_{ST} values in the pairwise comparison between the seven populations of *E. cristata* evidenced in panel A whereas using mtDNA marker. (D) Haplotype network color-coded according to the seven populations in A.

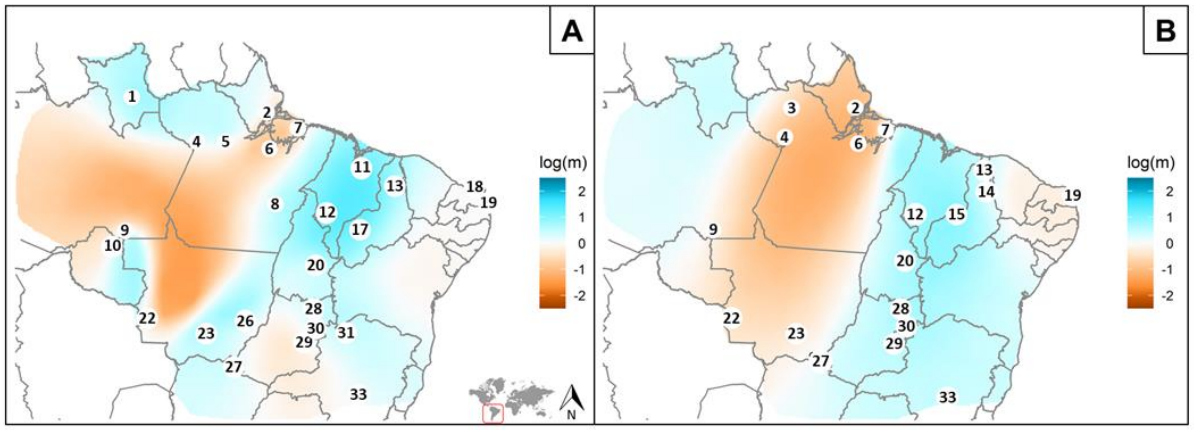


Figure 4. Effective migration rates analysis for *E. cristata* (A) and *E. chiriquensis* (B). Values expressed in log(m) range from 2 (greater genetic similarity) to -2 (less genetic similarity). Numbers indicate the locations sampled, for details see Table 1 and Table S1.

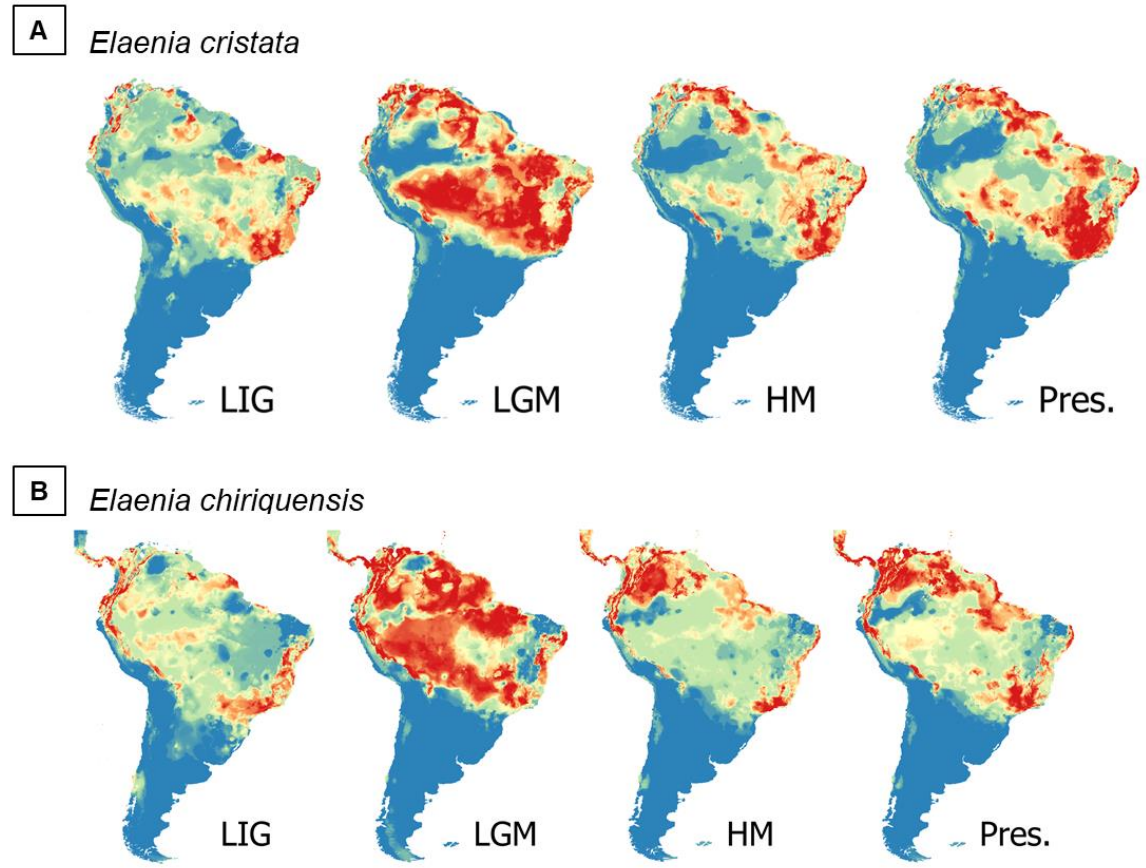


Figure 5. Graphic representation of distribution models for the *Elaenia cristata* (A) and *Elaenia chiriquensis* (B) using the ecological niche modeling technique (ENMs). The models were built using data from observations of each species and climatic variables available from WorldClim (See methods section). The periods are the Last Interglacial (ca 120,000 yr ago), Last Glacial Maximum (ca 21,000 yr ago), Mid-Holocene (ca 6,000 yr ago), and the Present. The climate suitability range is indicated by the range from red to blue colors, with red being the most suitable area.

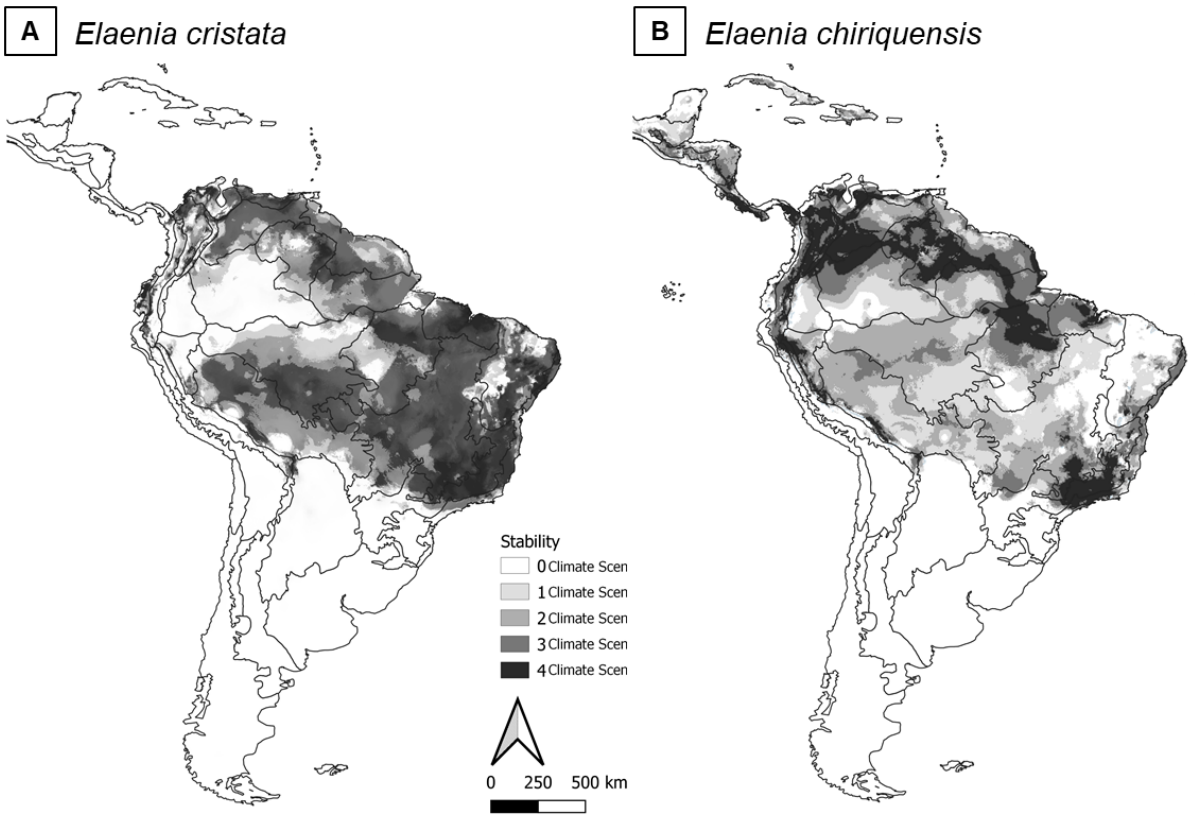


Figure 6. Stable areas of the species *Elaenia cristata* (A) and *Elaenia chiriquensis* (B) were generated from the overlapping of binary maps of each climate scenario modeled by ecological niche modeling technique (Fig. 5).

SUPPLEMENTAL INORMATION

Table S1, S2, and S3 go along with the Supplementary Material in separate *.csv* format.

Table S4. F_{ST} values from pairwise comparisons between the seven *E. cristata* populations defined. For population details see Methods and Figure 3. Analyzes conducted in software Arlequin v35 using the distance method by pairwise difference.

	POP 1	**POP 2**	**POP 3**	**POP 4**	**POP 5**	**POP 6**	**POP 7**
POP 1	0.00000						
POP 2	0.00134	0.00000					
POP 3	0.02359	0.04750	0.00000				
POP 4	**0.14489**	**0.16400**	**0.13611**	0.00000			
POP 5	-0.00721	0.01939	0.04271	**0.16564**	0.00000		
POP 6	**0.12702**	**0.17656**	**0.14682**	**0.23965**	**0.11162**	0.00000	
POP 7	**0.39216**	**0.36806**	**0.36970**	**0.33243**	**0.37977**	**0.56710**	0.00000

Table S5. F_{ST} P values from pairwise comparisons between the seven *E. cristata* populations sampled. For population details see Methods and Figure 3. Analyzes conducted in software Arlequin v35 using the distance method by pairwise difference. The number of permutations:1023. P values >0.05 are highlighted in bold.

	POP 1	**POP 2**	**POP 3**	**POP 4**	**POP 5**	**POP 6**	**POP 7**
POP 1	*						
POP 2	0.34082+-0.0163	*					
POP 3	0.18164+-0.0125	0.07520+-0.0076	*				
POP 4	**0.01270+-0.0031**	**0.00195+-0.0014**	**0.01074+-0.0033**	*			
POP 5	0.47949+-0.0157	0.16602+-0.0110	0.08496+-0.0078	**0.00195+-0.0014**	*		
POP 6	**0.02734+-0.0053**	**0.00000+-0.0000**	**0.01562+-0.0050**	**0.00000+-0.0000**	**0.02051+-0.0045**	*	
POP 7	**0.00000+-0.0000**	**0.00000+-0.0000**	**0.00000+-0.0000**	**0.00000+-0.0000**	**0.00000+-0.0000**	**0.00000+-0.0000**	*

Table S6. AMOVA results were performed at Arlequin v3.5 for the species *E. cristata* populations. Distance method of pairwise difference; 1023 permutations.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	6	25.437	0.20000 Va	20.54
Within populations	115	88.990	0.77382 Vb	79.46
Total	121	114.426	0.97382	
Fixation index (F_{ST})	0.205			

Va and F_{ST}: P (rand. value > obs. value) = 0.00000.

P (rand. value = obs. value) = 0.00000.

P-value = 0.00000 + -0.00000.

Table S7: Environmental variables were obtained by WorldClim and selected after Spearman correlation tests. ID refers to the code name used, and variable names are the description of it.

ID	Variable name
bio02	Mean Diurnal Range (Mean of monthly (max temp - min temp))
bio03	Isothermality (BIO2/BIO7) (* 100)
bio04	Temperature Seasonality (standard deviation *100)
bio05	Max Temperature of Warmest Month
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality (Coefficient of Variation)
bio18	Precipitation of Driest Quarter
bio19	Precipitation of Coldest Quarter

Table S8: True Skill Statistic (TSS) and Area Under the Curve (AUC) values used to evaluate the models of both species.

Species	AUC value	TSS value
Elaenia cristata	0.929709	0.607338
Elaenia chiriquensis	0.91388	0.581461

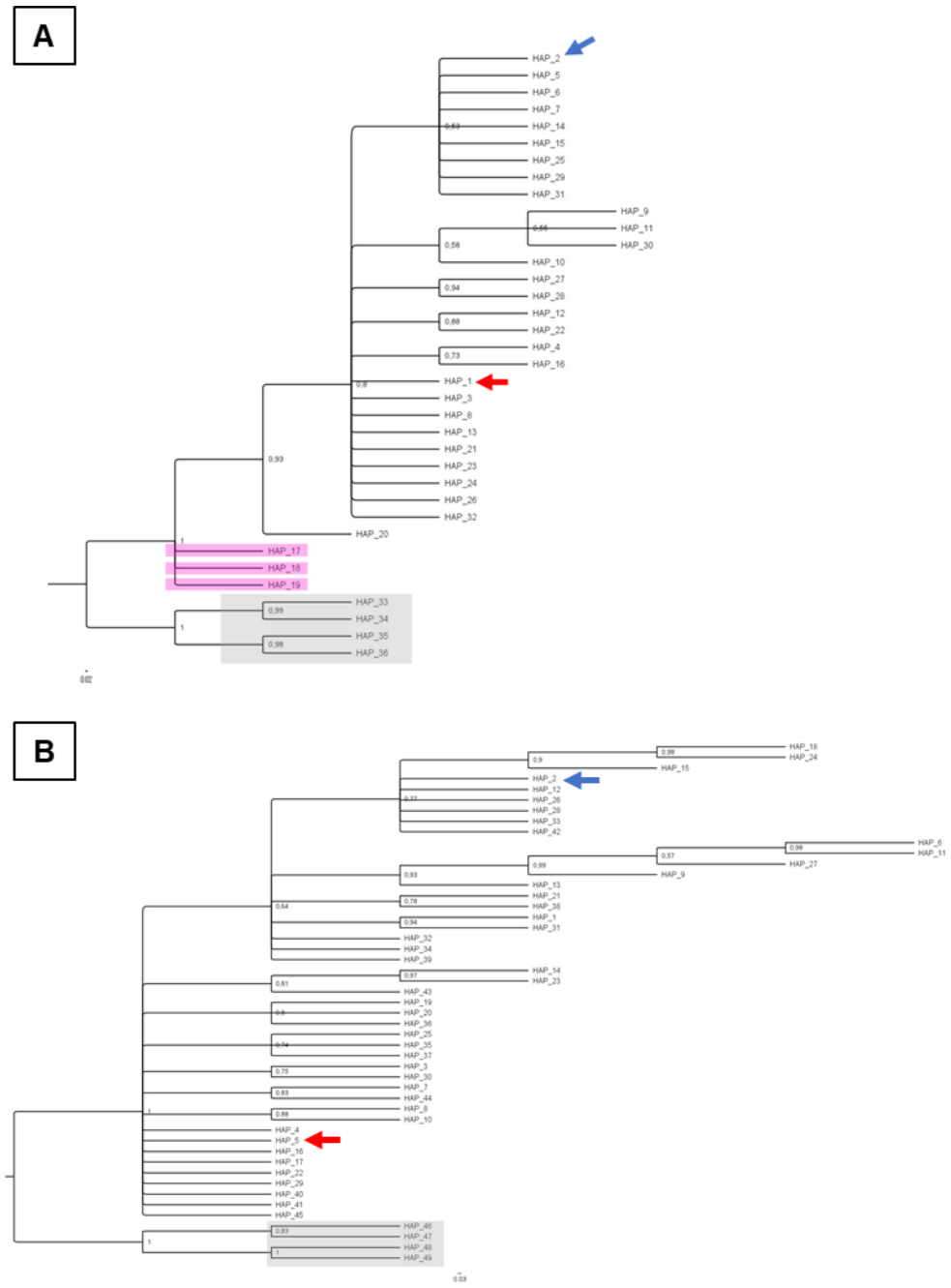


Figure S1. Bayesian gene tree estimated using *Elaenia cristata* (A) and *E. chiriquensis* (B) haplotypes. Blue and red arrows indicate the most frequent haplotypes in each sampled species. Highlighted in gray is the outgroup used (Greenish Elaenia, *Myiopagis viridicata*). Pink highlight in A indicates unique haplotypes located in POP4 (for details see methods), western Amazonia.

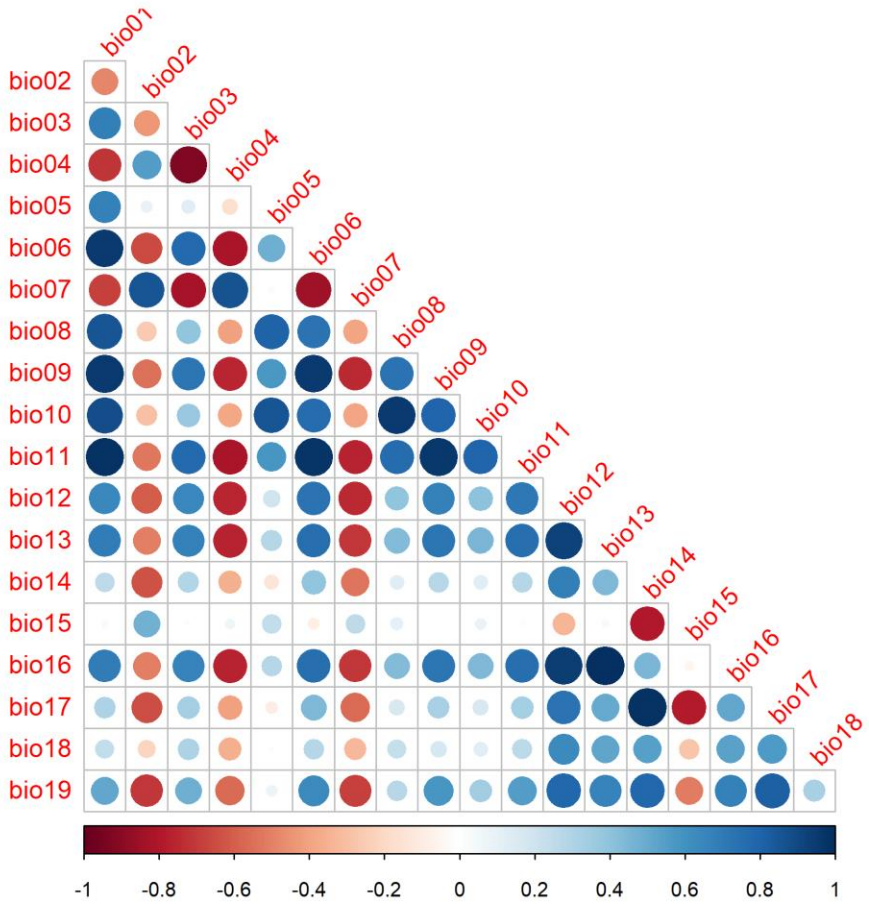


Figure S2: Graphic representation of the results of Spearman's correlation tests ($r < 0.7$) for the 19 environmental variables available in WorldClim. The selected variables are described in Table S7.

CONSIDERAÇÕES FINAIS

As diferenças encontradas na estruturação genética das populações das espécies evidência que não apenas questões extrínsecas, como relevo, clima, e mudanças climáticas, podem influenciar a distribuição das linhagens, mas também fatores intrínsecos, como a migração, também contribuem para estes processos e podem dificultar a identificação de padrões biogeográficos mais abrangentes.

Conhecer como está distribuída a biodiversidade é importante para traçar estratégias de conservação do bioma e das espécies de forma embasada cientificamente e levando em consideração a estruturação das populações de cada espécie.

Este trabalho acrescenta novos achados sobre a diversificação da avifauna em áreas abertas da América do Sul referente aos processos de diversificação da espécie migratória *E. chiriquensis* e da espécie residente *E. cristata* e pode servir como comparação para espécies de habitat e comportamentos similares.

APÊNDICE

JOURNAL OF
AVIAN BIOLOGY

Research

Ecological traits drive genetic structuring in two open-habitat birds from the morphologically cryptic genus *Elaenia* (Aves: Tyrannidae)

Eliane L. de Freitas, Leonardo Campagna, Bronwyn Butcher, Irby Lovette and Renato Caparroz

E. L. de Freitas (https://orcid.org/0000-0002-8072-3916) ✉ (elianeluizdefreitas@gmail.com) and R. Caparroz (https://orcid.org/0000-0003-4883-4023), Genetics and Biodiversity Laboratory, Dept of Genetics and Morphology, Univ. of Brasília – UnB, Brasília, Brazil. – L. Campagna, B. Butcher and I. Lovette, Fuller Evolutionary Biology Program, Cornell Lab of Ornithol., Ithaca, NY, USA and Dept of Ecology and Evolutionary Biology, Cornell Univ., Ithaca, NY, USA.

Journal of Avian Biology
2022: e02931
doi: 10.1111/jav.02931
Subject Editor: Lu Dong
Editor-in-Chief: Staffan Bensch
Accepted 17 December 2021

Understanding the relative contributions of the many factors that shape population genetic structuring is a central theme in evolutionary and conservation biology. Historically, abiotic or extrinsic factors (such as geographic barriers or climatic shifts) have received greater attention than biotic or intrinsic factors (such as dispersal or migration). This focus stems in part from the logistical difficulties in taking a comparative phylogeographic approach that contrasts species that have experienced similar abiotic conditions during their evolution yet differ in the intrinsic attributes that might shape their genetic structure. To explore the effects of intratropical migration on the genetic structuring of Neotropical birds, we chose two congeneric species, the lesser elaenia *Elaenia chiriquensis* and the plain-crested elaenia *E. cristata*, that are largely sympatric, and which have similar plumage, habitat preferences and breeding phenology. Despite these many commonalities, they differ in migratory behavior: *E. chiriquensis* is an intratropical migratory species while *E. cristata* is sedentary. We used a reduced representation genomic approach to test whether migratory behavior is associated with increased gene flow and therefore lower genetic population structure. As predicted, we found notably stronger genetic structuring in the sedentary species than in the migratory one. *E. cristata* comprises genetic clusters with geographic correspondence throughout its distribution, while there are no geographic groups within Brazil for *E. chiriquensis*. This comparison adds to the growing evidence about how intrinsic traits like migration can shape the genetic structuring of birds, and advances our understanding of the diversification patterns of the understudied, open habitat species from South America.

Keywords: ddRAD-Seq, ecological traits, flycatchers, genetic structure, migratory-sedentary behavior, Neotropics

NORDIC SOCIETY OIKOS
www.avianbiology.org

Introduction

How different factors contribute to shaping the genetic structure of populations is a central question in evolutionary and conservation biology (Foll and Gaggiotti 2006). The genetic structure of populations is determined by the combined effect of evolutionary forces acting within a population (e.g. mutation, genetic drift and selection) and by the mixing of genetic variation among populations through gene flow (Curnow and Wright 1978). The intensity of gene flow is influenced by a combination of abiotic or extrinsic factors (such as geographic barriers, geological or climatic shifts) (Mairal et al. 2017, Campillo et al. 2020) and/or biotic or intrinsic traits (such as reproductive behaviors, dispersal, migration and adaptation to local environmental conditions) (Burney and Brumfield 2009, Calderón et al. 2014, Smith et al. 2014, Nistelberger et al. 2015).

Phylogeographic studies have traditionally focused more on extrinsic factors, as addressing the effects of biotic factors is often challenging. One method of exploring the genetic effects of intrinsic traits involves comparing species that have experienced similar abiotic conditions during their evolution while diverging in the intrinsic attributes that are hypothesized to have influenced their genetic structuring (Papadopoulou and Knowles 2016).

In birds, migration is an intrinsic trait known to influence the spatial distribution of genetic variation (Arguedas and Parker 2000, Clegg et al. 2003, Irwin et al. 2011, Contina et al. 2019). Migration can be defined operationally as the short to large-scale, cyclic, seasonal movement of a population between breeding and non-breeding areas (Newton 2003). The level of genetic structure in migratory species (as well as in sedentary species) may be influenced by natal philopatry, which describes the tendency of individuals to return to their natal breeding grounds to reproduce (Weatherhead and Forbes 1994). High natal philopatry is expected to restrict gene flow among migrant groups, promoting stronger genetic structuring, whereas low natal philopatry facilitates gene flow, resulting in lower genetic structuring. Particularly in birds, studies have shown that natal philopatry is generally higher in sedentary populations compared to those that are migratory (Weatherhead and Forbes 1994). Furthermore, even when high natal philopatry was observed in migratory species, this occurred mostly in isolated populations on islands, and was hypothesized to be the product of a local adaptation (Wright and Mauck 1998, Förschler et al. 2010).

Most research on migratory birds has focused on species that breed in northern latitudes in North America (Nearctic migrants) and Eurasia (Palearctic migrants), and which migrate south during the Northern Hemisphere winter. Less work has been done on the 'Austral migrant' species of the Southern Hemisphere that migrate north during the Southern Hemisphere winter, and even fewer studies have addressed species that perform annual latitudinal movements on a smaller scale within the tropics, a phenomenon termed 'intratropical migration' (Hayes 1995, Jahn et al. 2020). Although there are more than 200 species of austral and intratropical migrants (Chesser 1994, Stotz et al. 1996, Jahn et al. 2006, 2020) in the Neotropical region, little is currently known about these migratory processes and their evolutionary consequences (Faaborg et al. 2010, Jahn et al. 2020).

Migratory behaviors often vary even among species of the same genus, as is the case in *Elaenia* flycatchers, a genus of 21 species occurring across Central and South America and the Caribbean. In this genus, some species are austral migrants (e.g. *E. chilensis* and *E. parvirostris*), others are intratropical migrants (e.g. *E. chiriquensis*) and others are sedentary (e.g. *E. cristata* and *E. obscura*) (Marini et al. 2009, Guaraldo et al. 2016, Somenzari et al. 2018). The lesser elaenia *E. chiriquensis* and the plain-crested elaenia *E. cristata* are sympatric across most of their geographic distributions (Fig. 1). They have similar preferences for breeding and foraging habitat (Cerrado sensu stricto, highly seasonal savanna), display similar and monomorphic plumage, and are mainly frugivorous (Hosner 2020, Hosner et al. 2020). Molecular phylogenetic studies show that these species represent independent lineages that are easily distinguished genetically (Rheindt et al. 2008). These species do not hybridize nor show high levels of incomplete lineage sorting, as is the case for other species pairs in the same genus (Rheindt et al. 2008, Tang et al. 2018).

The most notable difference between these two *Elaenia* species is in their sedentary/migratory behavior. *E. cristata* lives in open savanna and can be found throughout the year across its entire distribution (Sick 1997, Hosner 2020). In contrast, *E. chiriquensis* can be found throughout the year in some smaller enclaves of savannas surrounded by forest in northern South America (hereafter only northern savannas) and increases its abundance in the center of the Cerrado (the tropical Brazilian savanna) during the breeding season between August and December (Medeiros and Marini 2007, Hosner et al. 2020). After breeding in the central Cerrado, *E. chiriquensis* moves north likely to the Amazonian region (Marini and Cavalcanti 1990). However, the exact migratory routes and the possible existence of resident populations in this species remain unclear.

Technological advances in massive parallel sequencing approaches have enabled fast and low-cost access to a high number of molecular markers, in a large number of individuals (Edwards et al. 2015, Goodwin et al. 2016), making it possible to analyze the genetic structure of populations more robustly (Lavretsky et al. 2019), including that of migratory birds (Kraus et al. 2011, 2013, Jonker et al. 2012, Ruegg et al. 2014, DeSaix et al. 2019, Delmore et al. 2020). Single nucleotide polymorphisms (SNP) are the main molecular marker used in studies that evaluate the genetic structure and levels of gene flow between populations of one or more species (Hohenlohe et al. 2010, Kopuchian et al. 2020), or that identify distinct groups within a migratory species (Kraus et al. 2013, Ruegg et al. 2014).

Here we used a reduced representation genomic approach (ddRAD-Seq) to study the effect of intratropical migration and sedentary behavior on the genetic structure of two Neotropical birds: the migrant (*E. chiriquensis*) and the sedentary (*E. cristata*). Our central hypothesis is that migratory behavior with

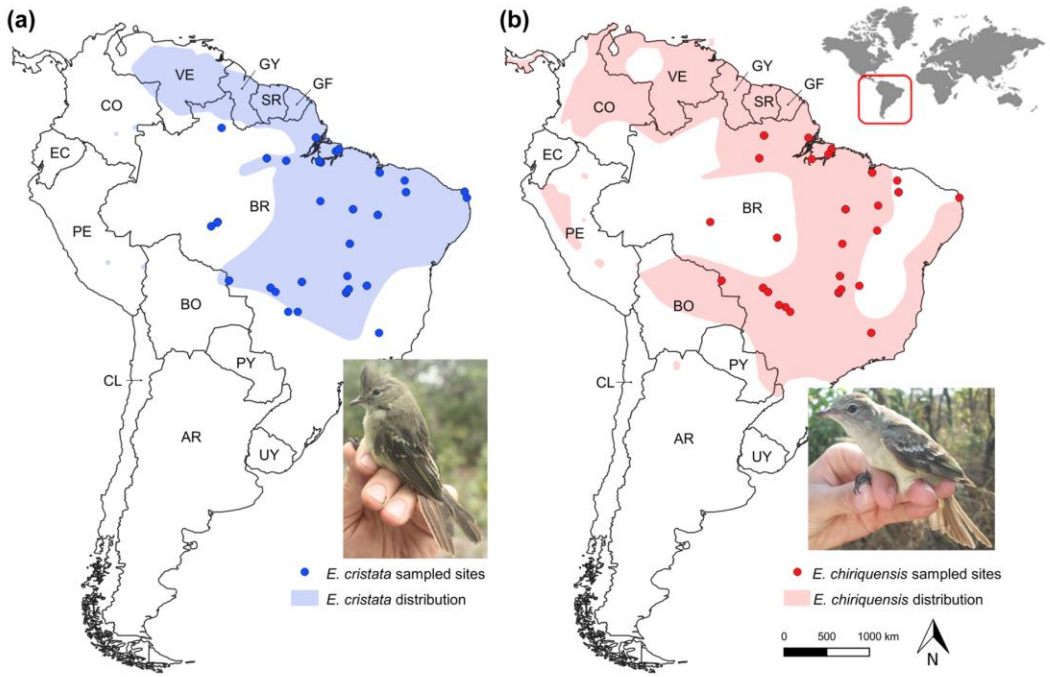


Figure 1. Geographical range distributions of *E. cristata* (a, blue) and *E. chiriquensis* (b, red) based on the International Union for Conservation of Nature (IUCN). Sampling sites are represented by color dots for each species (details in Table 1). Photos: (a) *E. cristata* from Mato Grosso – BR; Chapada dos Guimarães National Park (site 23; Fig. 3); (b) *E. chiriquensis* from Distrito Federal – BR; Protection Area Gama and Cabeça de Veado (site 29; Fig. 4).

lower natal philopatry will result in higher gene flow among populations, and therefore lower population structure in the migratory species compared to the sedentary one.

Material and methods

Species distribution and tissue sampling

We sampled 218 specimens (*E. cristata*, n = 98; *E. chiriquensis*, n = 120, but see the Results section for information about misidentifications) from 2003 to 2018 across 33 sites in South America (Supporting information, summarized in Table 1, Fig. 1). Specimens were captured in the field using mist-nest and banded to avoid duplicate sampling. Approximately 20 µl of blood was obtained from each individual using sterile needles and glass capillary tubes and stored in absolute ethanol at room temperature.

The specimens were identified in the field using slight morphological differences (Hosner 2020, Hosner et al. 2020) or through vocalization whenever possible. Briefly, the diagnostic characters for *E. cristata* are mainly an elongated crown of feathers building a conspicuous crest, without a white coronal patch, and two broad, well-marked wing bars (Fig. 1a). While *E. chiriquensis* exhibits slightly elongated crown feathers forming a less pronounced crest with a variably sized white coronal patch occasionally hidden, and it also has two broad wing bars and whitish edges on the remiges (Fig. 1b). Vocalizations were easy to distinguish by ear between species and when the captured individual vocalized, we also compared this vocalization with recorded vocalizations for the species available in databases such as eBird (<https://ebird.org>) or Xeno-canto (<www.xeno-canto.org>).

To increase the geographic coverage of our sampling, we obtained 108 *Elaenia* tissue samples from ornithological collections for a total of 326 samples (Supporting information).

In both species, most of the individuals (89% for *E. chiriquensis*, and 79% for *E. cristata*) were sampled during the breeding season (August–February), usually during the same expedition in each site (Supporting information).

DNA extraction and quality control

Total genomic DNA was extracted following a phenol, chloroform, isoamyl alcohol protocol (as in Friesen et al. 1997), or using the PureLink Genomic DNA Mini Kit

Table 1. Description of the samples of *E. cristata* and *E. chiriquensis* used in the genetic analyses after removing misidentified or samples with > 80% of missing data. Locality number (No.) corresponding to those shown in Fig. 3 and 4. Sample size for each locality is also shown.

No.	State; collection site	Latitude	Longitude	*E. cristata*	*E. chiriquensis*
1	Roraima; Viruá National Park	1.42	−60.98	2	0
2	Amapá; Cerrado Experiment Field of the Embrapa Amapá	0.39	−51.05	16	2
3	Pará; Óbidos	0.63	−55.72	0	7
4	Pará; Oriximiná	−1.76	−56.22	10	4
5	Pará; Monte Alegre	−2.02	−54.18	3	0
6	Pará; Portel	−1.85	−50.70	10	1
7	Pará; Marajó Island	−0.78	−48.61	10	12
8	Pará; Parauapebas	−6.28	−50.58	1	0
9	Amazonas; Manicoré	−8.47	−61.39	4	1
10	Rondônia; Machadinho	−8.92	−62.07	5	0
11	Maranhão; Urbano Santos	−3.28	−44.33	1	0
12	Maranhão; Chapada das Mesas National Park	−7.13	−47.15	8	11
13	Piauí; Pirarucura	−4.11	−41.71	0	1
14	Piauí; Castelo do Piauí	−5.33	−41.57	2	3
15	Piauí; Guadalupe	−6.74	−43.72	0	1
16	Piauí; Serra das Confusões	−9.37	−43.82	0	1
17	Piauí; Uruçuí	−7.73	−44.54	7	0
18	Rio Grande do Norte; Rio Fogo	−5.29	−35.40	10	0
19	Rio Grande do Norte; Parnamirim	−5.92	−35.17	6	2
20	Tocantins; Ponte Alta do Tocantins	−10.76	−47.48	4	1
21	Mato Grosso; Guarantã do Norte	−10.12	−54.36	0	1
22	Mato Grosso; Vila Bela da Santíssima Trindade	−14.63	−60.20	10	3
23	Mato Grosso; Chapada dos Guimarães National Park	−15.40	−55.83	10	10
24	Mato Grosso; Itiquira	−17.21	−54.14	0	2
25	Mato Grosso; Araguaia	−17.42	−53.45	0	1
26	Mato Grosso; Nova Xavantina	−14.78	−52.53	5	0
27	Goiás; Emas National Park	−17.92	−52.97	2	1
28	Goiás; Chapada dos Veadeiros National Park	−14.16	−47.74	2	8
29	Distrito Federal; Protection Area Gama and Cabeça de Veado	−15.92	−47.87	15	73
30	Distrito Federal; Águas Emendadas Ecological Station	−15.55	−47.59	3	1
31	Minas Gerais; Grande Sertão Veredas National Park	−15.18	−45.69	9	1
32	Minas Gerais; Mateus Leme	−20.14	−44.45	0	2
33	Minas Gerais; Serra Azul	−20.14	−44.41	1	0
Total				156	150

(Invitrogen, Carlsbad, CA, USA) (for museum tissue samples), following manufacturer instructions. Genomic DNA quality and concentrations were verified on a 1% agarose gel stained with ethidium bromide and using the Qubit™ dsDNA BR Assay Kit (ThermoFisher, Waltham, MA, USA), respectively.

Species identification through mitochondrial DNA

As the two *Elaenia* species analyzed are challenging to distinguish morphologically, we confirmed the field identifications using diagnostic variation in the mitochondrial gene NADH dehydrogenase subunit-2 (ND2). Polymerase chain reaction (PCR) amplifications were performed with the primer pair LMET and H6313 (Sorenson et al. 1999) in a final volume of 20 µl using the Phusion High-Fidelity PCR Kit (New England BioLabs, Ipswich, Massachusetts, USA), 10 mM each dNTP, 10 µM of each primer (forward and reverse), 0.4 U of Taq Polymerase and 30–80 ng of genomic DNA. The temperature cycling involved the first step at 98°C for 30 s, followed by 30 cycles of 10 s at 98°C, 30 s at 54°C and 30 s at 72°C, followed by a final extension step of 72°C for 10 min.

PCR products were then treated with exonuclease (EXO) and shrimp alkaline phosphatase (SAP) (ThermoFisher) in a final volume of 1 µl with 10 U μl^{-1} of EXO and 1.0 U μl^{-1} of SAP per 10 µl of PCR product, heated in a thermocycler at 37°C for 30 min, then held at 90°C for 10 min. PCR products were then sequenced at the Cornell University Biotechnology Resource Center Genomics Facility and MACROGEN Inc. Sequencing results were verified using Geneious Prime 2019.0.4 (<www.geneious.com>). The species were identified by comparing the ND2 sequence obtained from each sample with the database nucleotide collection (GenBank) using a Standard Nucleotide BLAST-blastn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) optimized for highly similar sequences (Megablast) (Morgulis et al. 2008).

ddRADseq dataset

We generated ddRADseq loci following the approach outlined by Peterson et al. (2012) with modifications as described by Thrasher et al. (2018). Briefly, we digested each sample with *SbfI* and *MspI* and ligated adapters that allowed multiplexing. The libraries, each containing approximately

20 samples, were size-selected and PCR-enriched, incorporating the Illumina HiSeq adapters (Illumina, San Diego, California, USA). Finally, all groups of samples were combined in equimolar proportions and sequenced, single end 100 bp, on two lanes of an Illumina HiSeq 2500.

After assessing read quality with FastQC (<www.bioinformatics.babraham.ac.uk/projects/fastqc>), we used FASTX-Toolkit (<http://hannonlab.cshl.edu/fastx_toolkit>) to trim sequences to 98 bp to discard lower-quality base calls at the 3′ end of the sequence. Subsequently, we used FASTX-Toolkit to retain reads without a single base below a Phred quality score of 10 and with at least 95% of bases with quality above 20. We demultiplexed reads using the 'process_radtags' program from the STACKS ver. 2.41 bioinformatics pipeline (Catchen et al. 2011, 2013), discarding reads that did not pass the Illumina filter, had barcode contamination, lacked an *SbfI* cut site or one of the unique barcodes used for multiplexing at the 5′ end. We obtained an average of 550 757 (± 263 678 reads) quality-filtered reads per individual (Supporting information).

We assembled the reads from both species into RADseq loci using the de novo pipeline from STACKS. We conducted a sensitivity analysis by testing different values for coverage (m = 5, 10, 20 and 30) as suggested by Rochette and Catchen (2017) yet did not find substantial differences in the number of loci recovered. Parameters were therefore set to a minimum coverage of 5 (m), up to seven differences between alleles of the same locus (M) and seven differences among aligned loci of different individuals (n). This combination of parameters produced an average coverage per locus ranging from 11.24 to 68.55x, with an overall average of 29.61x (± 9.2). We exported SNPs using the program 'populations' in STACKS for all the samples combined (n = 326) and again for each species separately, *E cristata* (n = 156) and *E. chiriquensis* (n = 150). We retained loci that were present in at least 80% of the individuals and exported both one SNP per RADSeq locus (to avoid including linked loci) and all SNP per RADSeq loci.

Principal component analysis

We conducted a principal component analysis of the SNP dataset using the package SNPRelate (Zheng et al. 2012) in R (<www.r-project.org>) to assess possible groupings among individuals. For this analysis, we exported SNPs using 'populations' from STACKS in four different ways: 1) for all the specimens (n = 326 and 5938 SNPs); 2) removing both all the specimens misidentified in the field (Results) and samples with more than 80% of missing data for all loci (n = 306 and 6458 SNPs); 3) only with *E. cristata* individuals (n = 156 and 6447 SNPs); and 4) only *E. chiriquensis* samples (n = 150 and 9671 SNPs).

Phylogenetic analyses

We built a phylogenetic tree using RAxML ver. 8.2.9 (Stamatakis 2014) for both species using the SNP dataset in Phylip format (variant sites only), and the following parameters: 500 replicates of rapid bootstrap analysis, and the *ASC_GTRGAMMA* model with the Lewis correction for ascertainment bias.

Population genetic structure

We performed Structure analyses using the SNP datasets obtained for each species separately and one SNP per locus to avoid the effect of linkage (2814 SNPs for *E. cristata*, and 2765 for *E. chiriquensis*). We conducted analyses at K values ranging from 1 to 10 for *E. cristata* and 1 to 5 for *E. chiriquensis*, with 10 replicate runs at each value. We also performed a Structure analysis for a subsample of *E. cristata* from central Brazil (localities, 8, 11, 12, 14, 17, 20, 23, 26, 27, 28, 29, 30, 31 and 33; Fig. 3c), with K values ranging from 1 to 6. Each run included 500 000 iterations of burn-in followed by 1 000 000 sampling iterations. We estimated the best value of K using the method of Evanno et al. (2005). Using the groups from the Structure results, we calculated F_{ST} among populations within *E. cristata* using the 'populations' module of STACKS (parameter 'fststats'). Due to the weak genetic structure observed in *E. chiriquensis*, we did not estimate the F_{ST} values for this species as we did not have clearly defined populations.

The level of intraspecific genetic structure was also assessed using haplotype information in fineRAD Structure ver. 0.3 and RADpainter (Lawson et al. 2012, Malinsky et al. 2018). Briefly, this algorithm works in four steps: 1) it calculates the co-ancestry matrix using the SNP dataset, 2) clusters individuals based on the co-ancestry matrix, 3) builds a dendrogram tree; and 4) plots results using the fineRADstructurePlot.R script in R (<www.r-project.org>). Haplotype datasets were produced with 'populations' in STACKS, for each species separately, and without a filter for minor allele frequency.

Results

Elaenia species identification

Of the 326 specimens analyzed in our study (including samples from both museum collections and wild birds), 273 (83.74%) had a portion of the mitochondrial ND2 gene amplified successfully. From the 186 specimens morphologically identified by our team in the field and with successfully amplified ND2 sequences, 10 (5.4%) were misidentified to species. From the 87 tissue samples requested from ornithological collections with ND2 data, 18 (20.7%) were misidentified to species. Among these 28 misidentified specimens, 12 genetically identified *E. chiriquensis* individuals were previously identified as *E. cristata*, seven specimens previously identified as *E. cristata* were genetically *E. chiriquensis* and nine have been genetically identified as other species of *Elaenia* and even another morphologically similar tyrant flycatcher *Sublegatus arenarum*. While high, this field identification error rate is not surprising given the high phenotypic similarity of these species and the fact that samples were

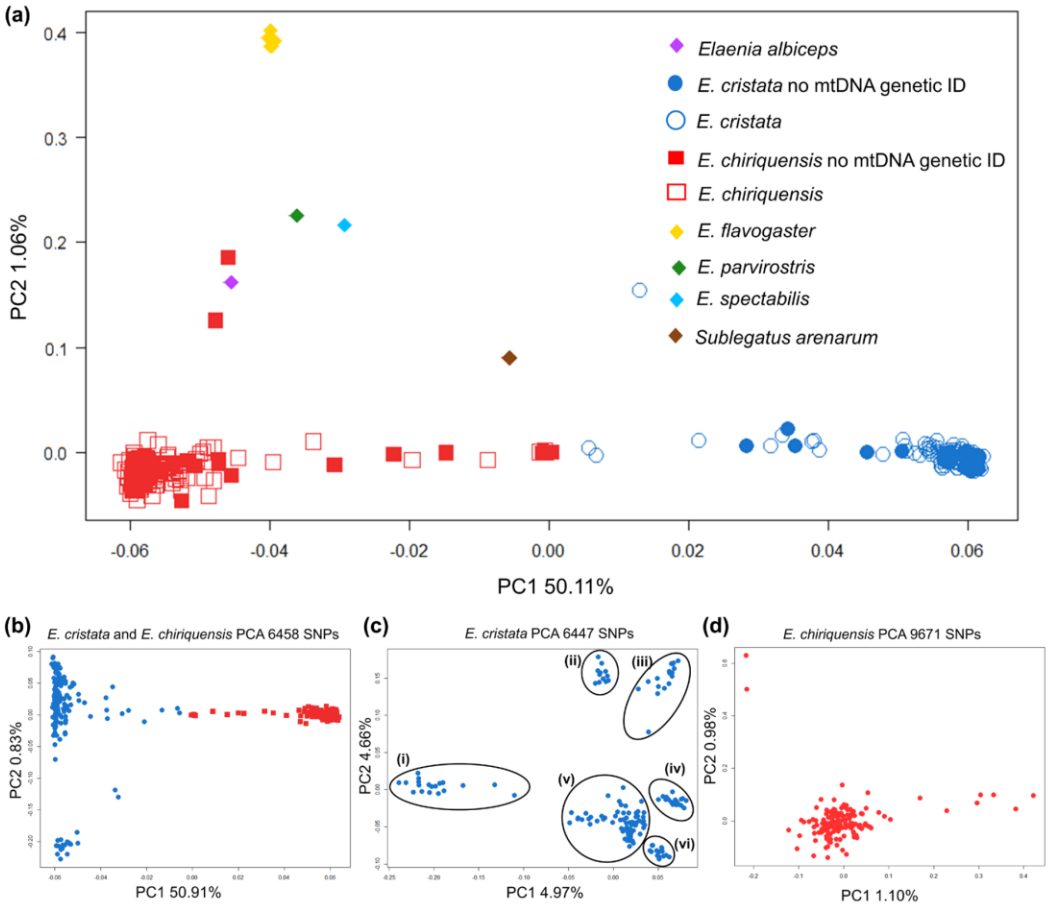


Figure 2. Principal component analyses (PCAs) derived from different sets of SNPs and *Elaenia*/*Sublegatus* samples (symbols). (a) 5983 SNPs from 326 samples. Diamond symbols refer to other species detected after genetic identification. Closed circles refer *E. cristata* species with no mtDNA genetic identification while open circles refer to samples with mitochondrial genetic identification. Closed squares refer to *E. chiriquensis* without genetic identification, and open squares to those with the genetic ID (see Methods, for details). (b) 6458 SNPs from 156 *E. cristata* and 150 *E. chiriquensis* samples after removing nine misidentified samples and samples with missing data > 80%. (c) 6447 SNPs from 156 *E. cristata*. The ellipses indicate groups with geographic correspondence, localities numbers are described in Table 1: (i) sites 9, 10 and 22; (ii) 1 and 4; (iii) 2 and 5; (iv) 6 and 7; (v) Brazilian central region; (vi) sites 18 and 19. (d) 9671 SNPs from 150 *E. chiriquensis*.

collected by many different people with varied levels of experience with these taxa. The PCA results based on the SNP data from all our samples and colored based on the mtDNA genetic identification show how errors can occur when using only morphological characters to identify species in the genus *Elaenia* (Fig. 2a, Supporting information). We determined that this larger dataset included representatives of six *Elaenia* species (*E. cristata*, *E. chiriquensis*, *E. albiceps*, *E. flavogaster*, *E. parvirostris* and *E. spectabilis*), and two samples from another flycatcher species *Sublegatus arenarum*. Because of this rate of misidentification, we only included in our final genomic analyses the *Elaenia* samples (n = 33) for which we did not confirm the field identification with mitochondrial DNA (failed to be amplified) if they grouped diagnostically in the PCA with those with confirmed identifications (Fig. 2a).

There is no overlap between *E. cristata* and *E. chiriquensis* in the PCA based on the SNP dataset derived from these two species alone (Fig. 2b). In this analysis, the first two principal components explain 51.17% of the total variation. To explore patterns of intraspecific variation we ran subsequent analyses on samples from each of these species alone. The PCA from *E. cristata* (the sedentary species; Fig. 2c) explained

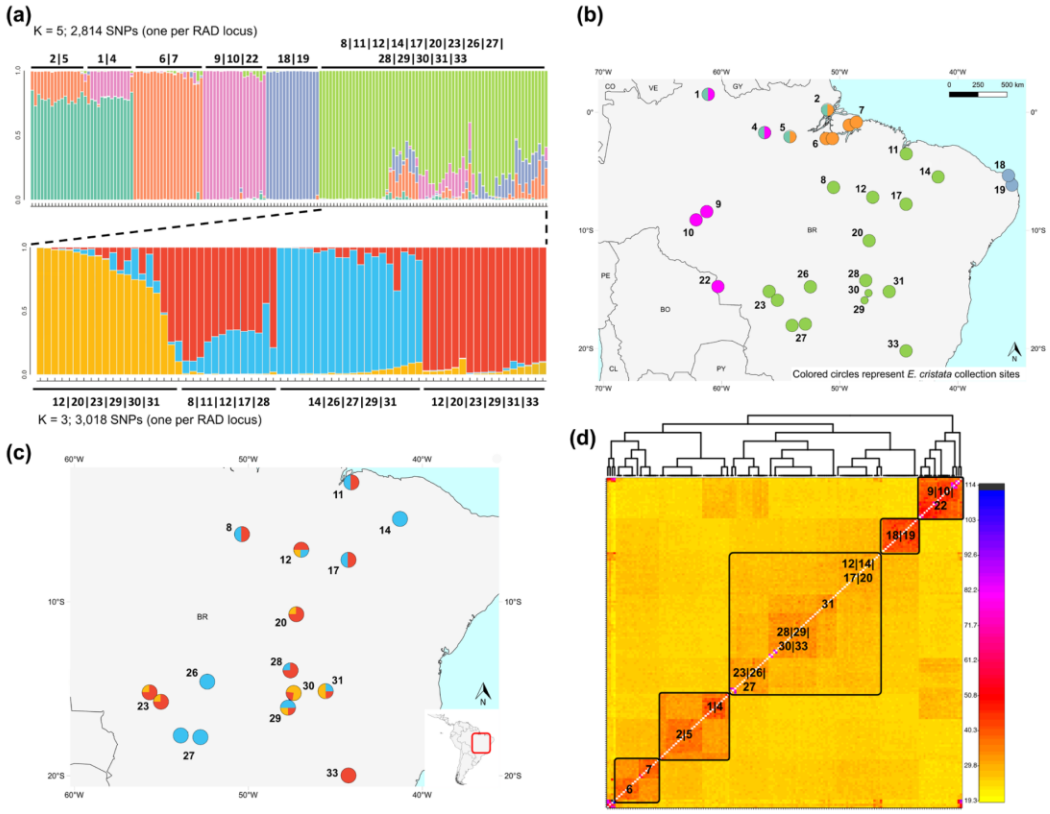


Figure 3. Pattern of genetic structure in the resident species (*E. cristata*). (a) Structure analysis based on 2814 SNPs (one per RAD locus) and 156 specimens showing five populations (K = 5) across all sampled sites (top). When we looked into the Brazilian central region ('green group') in further detail, the 70 sampled specimens were grouped into three populations (K = 3) in an analysis based on 3018 SNPs (bottom). (b) Geographical distribution of genetic clusters detected in the Structure analysis across all sampled sites; colored circles represent groups according to the Structure results; numbers correspond to sampled sites described in Table 1. (c) Geographical distribution of genetic clusters detected in the Structure analysis across sites in central Brazil. (d) The fineRADstructure plot derived from haplotype data (n = 156) indicating the clusters formed according to geographic location (numbers correspond to the localities shown on the maps).

9.63% of the variation and included several distinct clusters in the space of the two first principal components, whereas the equivalent plot for *E. chiriquensis* (the migratory species; Fig. 2d) explained only 2.08% of the total genetic variation and had no apparent clustering.

Genetic clusters and moderate structure in the sedentary species

The results from the PCA were consistent with those from the remaining analyses, where *E. cristata* and *E. chiriquensis* consistently showed distinct patterns of population structure. Structure results showed that in the sedentary species (*E. cristata*), the K values with the highest likelihood were 3 and 5, with the K = 5 pattern showing genetic clusters that can be explained geographically (Fig. 3a, Supporting information). The five clusters correspond to two groups in the northern savannas, one in the southwestern Amazon Forest, one in the extreme northeast of Brazil, and a larger group including all other collection locations from the Cerrado. The Structure results obtained when analyzing individuals from this last large group alone also indicated the possibility for finer population structure, with evidence for three additional clusters (Fig. 3a, Supporting information). The central portion (site 29, Fig. 3c) contains admixed individuals with the genetic composition of all three groups, while some localities such as Nova Xavantina and Emas National Park show less evidence of admixture (sites 26 and 27; Fig. 3c). We obtained similar results in our fineRAD Structure analysis (Fig. 3d), observing five co-ancestry groups with additional

genetic sub-structuring within two of them. Finally, phylogenetic analysis with RAxML showed the unrooted *E. cristata* tree with four well-supported branches (bootstrap > 85) (Supporting information). These branches corresponded to specific geographic regions shown in the map (Fig. 3b).

The mean pairwise F_{ST} among the five genetic populations resulting from Structure for *E. cristata* ranged from 0.028 to 0.109 (Supporting information). The lowest values are derived from the comparisons between the population in the central region of the distribution and other populations (Supporting information).

Few genetic clusters and low structure in the migratory species

In contrast to the results from the sedentary species, for the migratory lesser elaenia *E. chiriquensis*, PCA analysis grouped samples into one cluster (Fig. 2d). Similarly, Structure results supported K = 2 (Fig. 4a, Supporting information) with most of our sampling localities belonging to a single genetic population, but with a discrete separation in a group around the Amazon rainforest, formed by individuals from Óbidos, Oxiriminá and Manicoré (sites 3, 4 and 9; Fig. 4b). The fineRAD Structure results also recovered this group, and an additional one which included individuals from Ilha do Marajó (site 7; Fig. 4b). The RAxML analysis did not show any well-supported clades (Supporting information).

Discussion

Assuming that both *Elaenia* species studied here, which have highly overlapping ranges, have experienced a similar historical geographic and climatic context (extrinsic factors which could shape their population structure), our results suggest that their intrinsic sedentary/migratory behavior has influenced their genetic structure. As we predicted, the sedentary species showed greater genetic structuring compared to the migratory one, across a wide area of South America where both species are sympatric (Fig. 1).

Patterns of genetic structure for *E. chiriquensis*

The low genetic structure found for migratory *E. chiriquensis* had already been preliminarily observed using other marker types (Bates et al. 2003, Rheindt et al. 2015) and suggests substantial recent gene flow throughout the species' range, which is primarily in South America. As described in previous studies (Marini and Cavalcanti 1990, Medeiros and Marini 2007, De Paiva and Marini 2013), the abundance of this species increases significantly between August and December in the south-central region of the distribution in areas with Cerrado sensu stricto vegetation, a habitat described as preferential for nest building by the species. After breeding, studies suggest that birds fly to areas in the north of the distribution, becoming absent in some regions in the southern Cerrado during the non-breeding period (Marini and Cavalcanti 1990).

Some authors have considered *E. chiriquensis* as partially migratory (Somenzari et al. 2018) based on the absence of records in the central Cerrado between June and August (Marini and Cavalcanti 1990), while individuals have been observed in northern Brazil throughout the year. Intraspecific differences in migratory behavior among individuals within the species can lead to temporal and spatial reproductive asynchrony promoting divergence of neutral genetic variation between sedentary and migrant populations (Burney and Brumfield 2009). For instance, migratory behavior was recently identified as a driver for the diversification of subspecies of an austral migratory bird species (*Tyrannus savanna*, Gómez-Bahamón et al. 2020).

Although a high degree of gene flow was observed in the sampled migratory populations of Lesser Elaenia, we still observed a modest level of genetic structuring among the populations from the northern (site 7, Fig. 4c) and western portion of the sampled area (sites 3, 4 and 9, Fig. 4c). The cluster analyses showed two or three populations (Structure, and FineRAD Structure, respectively) in this periphery of the Amazonian region: one population was formed by sites 3, 4 and 9, and another population by site 7 (Fig. 4). However, some admixed individuals of these populations were also collected in the central Cerrado (see site 29 in Fig. 3a) during the breeding season. Even though our data suggests a weak evidence of structuring for *E. chiriquensis* in our sampling of migratory populations (Guaraldo et al. 2021), we can't rule out the possibility of potential structuring among the resident populations in the non-sampled regions of its distribution. Future studies should include individuals from populations of *E. chiriquensis* in regions of Central America to confirm if this pattern of low genetic structure applies to the entire species. The absence of migration in some populations can lead to a decrease in gene flow and, consequently, higher population structure, as seen in the congeneric *E. cristata*.

The generally low genetic structure observed in the migratory species is consistent with the expectation of low natal philopatry as described by Weatherhead and Forbes (1994). This type of pattern occurs when individuals in a given area migrate to multiple areas in successive reproductive seasons. Occurrence data of this species throughout the year suggest that migrants come from the north of the distribution, where records of the species presence are constant throughout the year (Marini and Cavalcanti 1990), and reproduce in different regions in the south-central portion of the distribution. Furthermore, isotopic data indicate that the migrant *E. chiriquensis* exhibits a niche-following behavior (it seeks similar resources throughout the annual cycle), suggesting that despite flying north, it spends the winter in savanna areas (Guaraldo et al. 2016). Similarly, our data suggest a connection between savanna patches in northern South America and the central Cerrado region (Fig. 4). In fact, the genetic differentiation of the migrant populations that occur in these savanna patches which are interlocked in a forest environment may be associated with greater natal philopatry due to their isolation. Isolated migrant populations of passerines often show

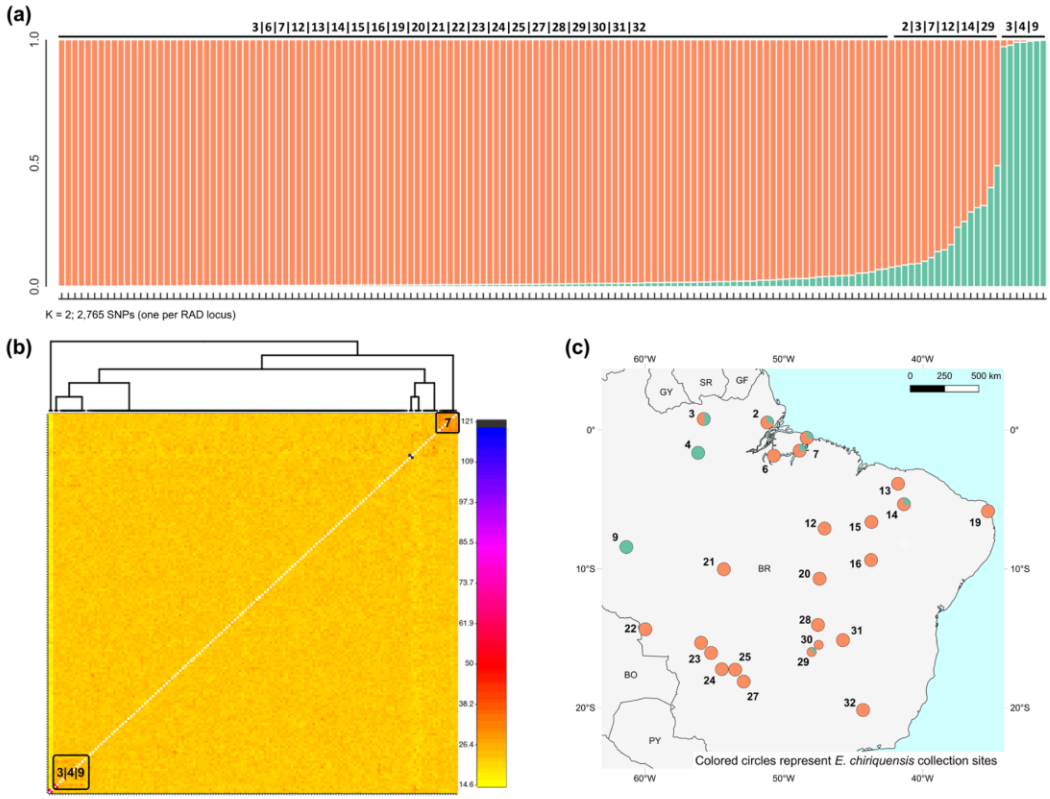


Figure 4. Pattern of genetic structure in the migratory species (*E. chiriquensis*). (a) The Structure analysis based on 2765 SNPs (one per RAD locus) and 150 specimens showed 2 populations (K=2) across all sampled sites. (b) The fineRADstructure plot derived from haplotype data (n=150) indicating the clusters formed according to geographic location (the numbers correspond to sampled sites shown on the maps). (c) Geographic distribution of genetic clusters detected in the Structure analysis across all sampled sites; colored circles represent groups according to the Structure results (not proportional to the number of sampled individuals); numbers correspond to sampled sites described in Table 1.

considerably higher philopatry than non-isolated populations (Weatherhead and Forbes 1994, Wright and Mauck 1998).

Alternatively, the Amazon rainforest could be a partially effective geographic barrier for this species, limiting migration and consequently decreasing gene flow among the migrant populations. The congruence of the genetic clusters formed in the northern savannas and peri-Amazonian area in both *Elaenia* species studied suggests the presence of an extrinsic barrier locally driving the genetic structure in these two species. In fact, these two factors (philopatry and a geographic barrier) are not mutually exclusive and may be acting together to shape the genetic structure of these species, and the relative contribution of each should be better investigated.

Genetic structuring in *E. cristata*

As expected, the sedentary species showed greater genetic structure, and we identified five genetic populations (Fig. 3; K=5; pairwise F_{ST} ranges from 0.028 to 0.109): two of them occurred in patches of savannas nested in the Amazonian region and the other three in southwestern, central and northeastern South America. The different populations of the north have a strong geographic association with the different savanna areas, indicating that the processes that led to the disjunct formation of these areas may have also influenced these avian population dynamics.

Despite observing greater genetic structuring in central and northeastern South America for this sedentary species, the magnitude of the genetic difference among these populations is small, consistent with recent isolation and/or some ongoing gene flow. One of these populations occurs in the northeast, in the Caatinga, the largest patch of the Seasonally Dry Tropical Forest of South America with a predominance of xeric vegetation (for details, Werneck 2011), while the other populations inhabit the Cerrado. Genetically differentiated populations occurring in the different open vegetation

biomes of South America have been identified in several organisms (Wuster et al. 2005, Ramos et al. 2007), including in other birds (Rocha et al. 2020).

Most previous studies of differentiation in Neotropical birds from open areas have found evidence for Pleistocene climatic oscillations as a driver of intraspecific divergence (Lima-Rezende et al. 2019b, Rocha et al. 2020, Ritter et al. 2021). For instance, the genetic differentiation of narrow-billed woodcreeper *Lepidocolaptes angustirostris* populations seems to have occurred in allopatry in stable areas that formed during Pleistocene climatic fluctuations (Rocha et al. 2020). Similarly, studies of other bird species that occur in the Cerrado have identified a weak intraspecific genetic structure, possibly due to the increase in gene flow between populations promoted by the expansion of climatically suitable areas for these species during the Pleistocene (Lima-Rezende et al. 2019a, Rocha et al. 2020).

In this context, it is important to highlight that our study assumes that the two species studied underwent similar historical demographic processes, but we can't completely rule out that these species, despite being ecologically similar, did not experience different geographic or climatic contexts at different times in the past. For instance, these two bird species might have occurred in distinct stable areas during the Pleistocene climatic oscillations, but the current intense gene flow promoted by migratory behavior in *E. chiriquensis* may have erased the genetic signature of historical isolation among previously isolated populations.

Elaenia sp. misidentification

About 8.5% (n=28) of the specimens collected for this study were misidentified in the field at the time of sample collection, which is not surprising given their morphological similarity (Fig. 1). Most of the cases involved the two focal species, with some *E. cristata* identified in the field being genetically identified as *E. chiriquensis* and vice versa. In a few cases (n=9), other species of *Elaenia* and even another morphologically similar tyrant flycatcher *Sublegatus arenarum* were misidentified as the two focal species of this study. This non-trivial field identification error rate underscores the difficulty with working in this challenging group of morphologically cryptic species. The genus *Elaenia* comprises 21 species that are all quite similar morphologically, and their misidentification has been widely reported (Traylor 1982, Hosner Hosner 2004, Winkler et al. 2020). For instance, Rheindt et al. (2015), in a similar study using the ND2 mtDNA gene, also found one misidentified *Elaenia* sample (*E. c. albivertex* labeled as *E. flavogaster*) out of 13 samples obtained from ornithological collections.

In general, field identification of *Elaenia* species is based either on the birds' vocalizations or on the species' geographical distribution (Sick 1997). Criteria based on geographic distribution can be challenging when species are sympatric or when there are migratory species involved. *Elaenia cristata* is thought to be partially migratory in some regions (Hosner 2020), such as in the Mato Grosso state in Brazil. We found 12 cases of misidentified *E. cristata* labeled as *E. chiriquensis*, three of them in Mato Grosso state and five in Pará state, localities that may be in the migratory routes of other species of *Elaenia*. Therefore, misidentification between sedentary and migratory *Elaenia* species can lead to the eventual misinterpretation of migratory behavior. Owing to the difficulty of using morphology to identify *Elaenia* species, we encourage using complementary species identification methods, such as DNA barcode approaches that have been successfully applied in the identification of many Tyrannidae species (Kerr et al. 2007, Chaves et al. 2008).

Conclusions

Our study adds evidence on how migratory behavior, as an intrinsic factor, can shape the genetic structure of Neotropical bird species and improves our understanding of the diversification patterns of open habitat South American species. As expected, migratory behavior can lead to a weak genetic structure, likely the product of substantial ongoing gene flow among populations. Similar patterns may exist among the other 200 species of migratory Neotropical birds, in contrast to the high levels of geographic structuring known to exist within many sedentary Neotropical species.

Acknowledgements – We are grateful to the directors and employees of all of our sampling locations for assistance during fieldwork and the curators of the Ornithological Collections in Brazil for providing tissues samples for this research. We also thank the Laboratório de Genética e Biodiversidade/UnB team, especially Gislaine Fernandes, for assistance during fieldwork and the Cornell Lab of Ornithology team. We thank Getúlio Oliveira Jr. for helpful comments on earlier drafts of the manuscript.

Funding – ELF was supported by scholarship CAPES (no. 1583777) and CNPq (no. 140440/2017-8), also CAPES PDSE no. 47/2017 awards 88881.190324/2018-01. This study is part of the GENPAC (564036/2010-2) and PPBio (457444/2012-6) projects supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Ministério da Ciência, Tecnologia e Inovação (MCTI)/Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Permits - Trapping and tissue sampling were allowed by the Research Ethics Committees of University of Brasília (UnBDOC no. 75111/2013) by the ICMBio/Sistema de Autorização e Informação em Biodiversidade, and Centro Nacional de Pesquisa e Conservação de Aves Silvestres (CEMAVE)/Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio)/Sistema Nacional de Anilhamento (SNA).

Conflict of interest – The authors declare that they have no conflicts of interest.

Author contributions

Eliane L. Freitas: Conceptualization (lead); Data curation (lead); Formal analysis (equal); Funding acquisition (equal); Investigation (lead); Methodology (lead); Project administration (supporting); Software (equal); Visualization (equal); Writing – original draft (lead); Writing – review and editing

(supporting). **Leonardo Campagna**: Conceptualization (supporting); Data curation (equal); Formal analysis (lead); Funding acquisition (supporting); Investigation (equal); Methodology (equal); Software (lead); Supervision (equal); Writing – original draft (supporting); Writing – review and editing (lead). **Bronwyn Butcher**: Conceptualization (supporting); Data curation (equal); Formal analysis (equal); Funding acquisition (supporting); Investigation (supporting); Methodology (equal); Resources (supporting); Software (equal); Supervision (supporting); Writing – review and editing (equal). **Irby Lovette**: Conceptualization (supporting); Data curation (supporting); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Software (equal); Supervision (lead); Writing – original draft (supporting); Writing – review and editing (lead). **Renato Caparroz**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Software (equal); Supervision (lead); Writing – original draft (supporting); Writing – review and editing (lead).

Transparent Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/jav.02931>.

Data availability statement

Data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.jdfn2z3cf> (Freitas et al. 2022).

Supporting information

The supporting information associated with this article is available from the online version.

References

Arguedas, N. and Parker, P. 2000. Seasonal migration and genetic population structure in house wrens. – Condor 102: 517–528.

Bates, J. M. et al. 2003. Initial assessment of genetic diversity in ten bird species of South American Cerrado. – Stud. Neotrop. Fauna Environ. 38: 87–94.

Burney, C. W. and Brumfield, R. T. 2009. Ecology predicts levels of genetic differentiation in neotropical birds. – Am. Nat. 174: 358–368.

Calderón, L. et al. 2014. Phylogeography and genetic structure of two Patagonian shag species (Aves: Phalacrocoracidae). – Mol. Phylogenet. Evol. 72: 42–53.

Campillo, L. C. et al. 2020. Genomic differentiation in an endemic Philippine genus (Aves: Sarcophanops) owing to geographical isolation on recently disassociated islands. – Biol. J. Linn. Soc. 131: 814–821.

Catchen, J. et al. 2013. Stacks: an analysis tool set for population genomics. – Mol. Ecol. 22: 3124–3140.

Catchen, J. M. et al. 2011. Stacks: building and genotyping loci de novo from short-read sequences. – Genes Genomes Genetics 1: 171–182.

Chaves, A. V. et al. 2008. Molecular taxonomy of Brazilian tyrant-flycatchers (Passeriformes: Tyrannidae). – Mol. Ecol. Resour. 8: 1169–1177.

Chesser, R. T. 1994. Migration in South America: an overview of the austral system. – Bird Conserv. Int. 4: 91–107.

Clegg, S. M. et al. 2003. Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler *Wilsonia pusilla*. – Mol. Ecol. 12: 819–830.

Contina, A. et al. 2019. Genetic structure of the painted bunting and its implications for conservation of migratory populations. – Ibis 161: 372–386.

Curnow, R. N. and Wright, S. 1978. Evolution and the genetics of populations, volume 4: variability within and among natural populations. – Biometrics 35: 359.

De Paiva, L. V. and Marini, M. Â. 2013. Timing of migration and breeding of the lesser elaenia *Elaenia chiriquensis* in a Neotropical savanna. – Wilson J. Ornithol. 125: 116–120.

Delmore, K. et al. 2020. The evolutionary history and genomics of European blackcap migration. – eLife 9: 1–24.

DeSaix, M. G. et al. 2019. Population assignment reveals low migratory connectivity in a weakly structured songbird. – Mol. Ecol. 28: 2122–2135.

Edwards, S. V. et al. 2015. Next-generation sequencing and the expanding domain of phylogeography. – Folia Zool. 64: 187–206.

Evanno, G. et al. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. – Mol. Ecol. 14: 2611–2620.

Faaborg, J. et al. 2010. Recent advances in understanding migration systems of New World land birds. – Ecol. Monogr. 80: 3–48.

Foll, M. and Gaggiotti, O. 2006. Identifying the environmental factors that determine the genetic structure of populations. – Genetics 174: 875–891.

Förschler, M. I. et al. 2010. Extraordinary high natal philopatry in a migratory passerine. – J. Ornithol. 151: 745–748.

Freitas, E. L. et al. 2022. Data from: Ecological traits drive genetic structuring in two open-habitat birds from the morphologically cryptic genus *Elaenia* (Aves: Tyrannidae). – Dryad Digital Repository, <https://doi.org/10.5061/dryad.jdfn2z3cf>.

Friesen, V. L. et al. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. – Mol. Ecol. 6: 1047–1058.

Gómez-Bahamón, V. et al. 2020. Speciation associated with shifts in migratory behavior in an avian radiation. – Curr. Biol. 30: 1312.e6–1321.e6.

Goodwin, S. et al. 2016. Coming of age: ten years of next-generation sequencing technologies. – Nat. Rev. Genet. 17: 333–351.

Guaraldo, A. C. et al. 2016. Contrasting annual cycles of an intra-tropical migrant and a tropical resident bird. – J. Ornithol. 157: 695–705.

Guaraldo, A. C. et al. 2021. Longitudinal and cyclic poleward migration of a South American intra-tropical migrant flycatcher, the lesser elaenia *Elaenia chiriquensis*. – Ornitol. Neotrop. 32: 94–102.

Hayes, F. E. 1995. Definitions for migrant birds: what is a neotropical migrant? – Auk 112: 521–523.

Hohenlohe, P. A. et al. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. – PLoS Genet. 6: e1000862.

Hosner, P. 2020. Plain-crested elaenia *Elaenia cristata*, ver. 1.0. – In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D. A. and de Juana, E. (eds), Birds of the world. Cornell Lab of Ornithology, Ithaca, NY, USA, <https://doi.org/10.2173/bow.plcela1.01>.

Hosner, P., Christie, D. A., de Juana, E. and Kirwan, G. M. 2020. Lesser elaenia (*Elaenia chiriquensis*), ver. 1.0. – In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D. A. and de Juana, E. (eds), Birds of the World. Cornell Lab of Ornithology, Ithaca, NY, USA, <https://doi.org/10.2173/bow.lesela1.01>.

Hosner, P. A. 2004. Genus *Elaenia*. – In: del Hoyo, J. et al. (eds), Handbook of the birds of the world. Vol. 9. Cotingas to pipits and wagtails. Lynx Edicions, pp. 267–274.

Irwin, D. E. et al. 2011. Genetic variation and seasonal migratory connectivity in Wilson's warblers *Wilsonia pusilla*: species-level differences in nuclear DNA between western and eastern populations. – Mol. Ecol. 20: 3102–3115.

Jahn, A. et al. 2006. Towards a mechanistic interpretation of bird migration in South America. – El Hornero 21: 99–108.

Jahn, A. E. et al. 2020. Bird migration within the Neotropics. – Auk 137: 1–23.

Jonker, R. M. et al. 2012. The development of a genome wide SNP set for the barnacle goose *Branta leucopsis*. – PLoS One 7: e38412.

Kerr, K. C. R. et al. 2007. Comprehensive DNA barcode coverage of North American birds. – Mol. Ecol. Notes 7: 535–543.

Kopuchian, C. et al. 2020. A test of the riverine barrier hypothesis in the largest subtropical river basin in the Neotropics. – Mol. Ecol. 29: 2137–2149.

Kraus, R. H. S. et al. 2011. Genome wide SNP discovery, analysis and evaluation in mallard *Anas platyrhynchos*. – BMC Genomics 12: 150.

Kraus, R. H. S. et al. 2013. Global lack of flyway structure in a cosmopolitan bird revealed by a genome wide survey of single nucleotide polymorphisms. – Mol. Ecol. 22: 41–55.

Lavretsky, P. et al. 2019. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. – Mol. Ecol. 28: 2594–2609.

Lawson, D. J. et al. 2012. Inference of population structure using dense haplotype data. – PLoS Genet. 8: e1002453.

Lima-Rezende, C. A. et al. 2019a. Late Pleistocene climatic changes promoted demographic expansion and population reconnection of a Neotropical savanna-adapted bird, *Neothraupis fasciata* (Aves: Thraupidae). – PLoS One 14: e0212876.

Lima-Rezende, C. A. et al. 2019b. The spatial genetic structure of the white-banded tanager (Aves, Passeriformes) in fragmented Neotropical savannas suggests two evolutionarily significant units. – Biotropica 51: 234–244.

Mairal, M. et al. 2017. Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the eastern Afromontane biodiversity hotspot. – Sci. Rep. 7: 45749.

Malinsky, M. et al. 2018. RADpainter and fineRADstructure: population inference from RADseq data. – Mol. Biol. Evol. 35: 1284–1290.

Marini, M. and Cavalcanti, R. 1990. Migrações de *Elaenia albiceps chilensis* e *Elaenia chiriquensis albivertex* (Aves: Tyrannidae). – Bol. Mus. Para. Emilio Goeldi Sér. Zool. 6: 59–67.

Marini, M. et al. 2009. Biologia reprodutiva de *Elaenia cristata* (Aves: Tyrannidae) em cerrado do Brasil central. – Neotrop. Biol. Conserv. 4: 2–12.

Medeiros, R. de C. S. and Marini, M. Â. 2007. Biologia reprodutiva de *Elaenia chiriquensis* (Lawrence) (Aves, Tyrannidae) em Cerrado do Brasil central. – Rev. Bras. Zool. 24: 12–20.

Morgulis, A. et al. 2008. Database indexing for production MegaBLAST searches. – Bioinformatics 24: 1757–1764.

Newton, I. 2003. Speciation and biogeography of birds. – Academic Press.

Nistelberger, H. M. et al. 2015. Genetic drift drives evolution in the bird-pollinated, terrestrial island endemic *Grevillea georgeana* (Proteaceae). – Bot. J. Linn. Soc. 178: 155–168.

Papadopoulou, A. and Knowles, L. L. 2016. Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. – Proc. Natl Acad. Sci. USA 113: 8018–8024.

Peterson, B. K. et al. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. – PLoS One 7: e37135.

Ramos, A. C. S. et al. 2007. Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of quaternary climate changes in the Brazilian Cerrado. – Ann. Bot. 100: 1219–1228.

Rheindt, F. E. et al. 2008. Habitat shifts in the evolutionary history of a Neotropical flycatcher lineage from forest and open landscapes. – BMC Evol. Biol. 8: 193.

Rheindt, F. E. et al. 2015. Cryptic speciation in the lesser elaenia *Elaenia chiriquensis* (Aves: Passeriformes: Tyrannidae). – Zootaxa 4032: 251.

Ritter, C. D. et al. 2021. Sister species, different histories: comparative phylogeography of two bird species associated with Amazonian open vegetation. – Biol. J. Linn. Soc. 132: 161–173.

Rocha, A. V. et al. 2020. Pleistocene climatic oscillations associated with landscape heterogeneity of the South American dry diagonal explains the phylogeographic structure of the narrow-billed woodcreeper (*Lepidocolaptes angustirostris*, Dendrocolaptidae). – J. Avian Biol. 2020: e02537.

Rochette, N. C. and Catchen, J. M. 2017. Deriving genotypes from RAD-seq short-read data using stacks. – Nat. Protoc. 12: 2640–2659.

Ruegg, K. C. et al. 2014. Mapping migration in a songbird using high-resolution genetic markers. – Mol. Ecol. 23: 5726–5739.

Sick, H. 1997. Ornitologia Brasileira. – Nova Fronteira.

Smith, B. T. et al. 2014. The drivers of tropical speciation. – Nature 515: 406–409.

Somenzari, M. et al. 2018. An overview of migratory birds in Brazil. – Pap. Avulsos Zool. 58: 3.

Sorenson, M. D. et al. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. – Mol. Phylogenet. Evol. 12: 105–114.

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

Stotz, D. F. et al. 1996. Neotropical birds: ecology and conservation. – Univ. of Chicago Press.

Tang, Q. et al. 2018. Rapid diversification and hybridization have shaped the dynamic history of the genus *Elaenia*. – Mol. Phylogenet. Evol. 127: 522–533.

Thrasher, D. J. et al. 2018. Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: a proof of concept in a highly promiscuous bird. – Mol. Ecol. Resour. 18: 953–965.

Traylor Jr., M. A. 1982. Notes on tyrant flycatechers (Aves: Tyrannidae). – Fieldiana Zool.: 1–22.
Weatherhead, P. J. and Forbes, M. R. L. 1994. Natal philopatry in passerine birds: genetic or ecological influences? – Behav. Ecol. 5: 426–433.
Werneck, F. P. 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. – Quat. Sci. Rev. 30: 1630–1648.
Winkler, D. W., Billerman, S. M. and Lovette, I. J. 2020. Tyrant flycatchers (Tyrannidae), ver. 1.0. – In: Billerman, S. M., Keeney, B. K., Rodewald, P. G. and Schulenberg, T. S. (eds), Birds of the World. Cornell Lab of Ornithology, Ithaca, NY, USA, <https://doi.org/10.2173/bow.tyrann2.01>.
Wright, N. T. W. and Mauck, R. A. 1998. Philopatry, natal dispersal and inbreeding avoidance in an island population of Savannah sparrows. – Ecology 79: 755.
Wuster, W. et al. 2005. Tracing an invasion: landbridges, refugia and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). – Mol. Ecol. 14: 1095–1108.
Zheng, X. et al. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. – Bioinformatics 28: 3326–3328.