Ricardo Fernandes Campos Junior

Divergência populacional e expansão demográfica de
Dendrocolaptes platyrostris (Aves: Dendrocolaptidae)
no final do Quaternário

Population divergence and demographic expansion of
Dendrocolaptes platyrostris (Aves: Dendrocolaptidae)
in the late Quaternary

São Paulo

2012

Ricardo Fernandes Campos Junior

Divergência populacional e expansão demográfica de
Dendrocolaptes platyrostris (Aves: Dendrocolaptidae)
no final do Quaternário

Population divergence and demographic expansion of
Dendrocolaptes platyrostris (Aves: Dendrocolaptidae)
in the late Quaternary

Dissertação apresentada ao Instituto
de Biociências da Universidade de São
Paulo, para a obtenção de Título de
Mestre em Biologia/Genética,

Orientadora: Cristina Yumi Miyaki.

São Paulo

2012

Ficha Catalográfica

Campos-Junior, Ricardo

Divergência populacional e expansão demográfica de *Dendrocolaptes platyrostris* (Aves: Dendrocolaptidae) no final do Quaternário

45

Dissertação (Mestrado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Genética e Biologia Evolutiva.

1. Filogeografia 2. Expansão demográfica 3. *Approximate bayesian computation* I. Universidade de São Paulo. Instituto de Biociências. Departamento de Genética e Biologia Evolutiva.

Comissão Julgadora:

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Cristina Yumi Miyaki

Epígrafe

The human understanding is no dry light, but receives an infusion from the will and affections; whence proceed sciences which may be called "sciences as one would." For what a man had rather were true he more readily believes. Therefore he rejects difficult things from impatience of research; sober things, because they narrow hope; the deeper things of nature, from superstition; the light of experience, from arrogance and pride, lest his mind should seem to be occupied with things mean and transitory; things not commonly believed, out of deference to the opinion of the vulgar. Numberless, in short, are the ways, and sometimes imperceptible, in which the affections color and infect the understanding.

Francis Bacon, *Novum organon* (1620)

Agradecimentos

Primeiramente gostaria de agradecer à minha orientadora, Cris, que me recebeu no laboratório sem me conhecer, me aceitou como aluno, e esteve sempre disposta a revisar relatórios e discutir ideias.

Também gostaria de agradecer aos amigos que fiz durante o período, que além das risadas me auxiliaram no laboratório, nas discussões, a aguentar as frustrações da pós graduação, e me acompanharam nas cervejas enquanto esperávamos o BigDye sair: Adri, Flávia, Priscila, Danilo, Ana Cris, Fábio, Henrique, Tiago, André, Carol, Marcos, Fernando d'Horta e Gustavo.

Às pessoas que me ajudaram financeiramente e psicologicamente: minha mãe, Rosi, meu pai, Ricardo, minha avó Luzia e a Greici.

Aos curadores e coletores do Museu de Zoologia da USP (MZUSP) e do Laboratório de Biodiversidade e Evolução Molecular (LBEM) da UFMG que forneceram amostras de suas coleções, e às agências de fomento que financiaram esta pesquisa, FAPESP, CAPES e CNPq, sem os quais este trabalho não poderia ser realizado.

Este trabalho foi desenvolvido no âmbito do Núcleo de Pesquisa em Biodiversidade e Computação da Universidade de São Paulo (BioComp).

Índice

1.	Introdução	1
2.	Objetivos	6
3.	Material e Métodos	7
3.1	Área de estudo e amostras	7
3.2	Extração de DNA	7
3.3	Quantificação	7
3.4	Amplificação	7
3.5	Purificação	9
3.6	Sequenciamento	9
3.7	Análises	9
4.	Resultados	17
4.1	Marcadores moleculares	17
4.1.1	Intron 5 beta fibrinogênio (FIB5)	17
4.1.2	Intron 2 da mioglobina (Myo2)	17
4.1.3	Citocromo B (CitB)	19
4.2	Estrutura genética populacional	20
4.3	Expansão populacional	21
4.4	Parâmetros populacionais	21
4.5	Filogeografia estatística	22
5.	Discussão	25
5.1	Padrão de variação	25
5.2	Demografia histórica	26
5.3	Isolamento com migração e filogeografia estatística	27
5.4	Implicações biogeográficas	30
6.	Conclusões	31
7.	Resumo	32
8.	Abstract	33
9.	Referências Bibliográficas	34

1. Introdução

Trabalhos envolvendo estrutura e diversidade genética dos organismos podem agregar informação a respeito da história evolutiva e dos processos responsáveis pela especiação e diversificação de linhagens (Barton & Hewitt, 1985; Nosil et al., 2009; Sobel et al., 2009), além de servirem de base para outros estudos que ajudam a compreender estes mesmos processos em outros contextos ou em outras regiões (por exemplo, Moritz et al., 2009; Martins, 2011). A identificação e a caracterização dos potenciais eventos e processos que geraram tal estruturação e / ou diversidade são de grande interesse para a biologia da conservação (Moritz, 2002), pois podem contribuir no planejamento de estratégias conservacionistas (Moritz et al., 2000; Moritz, 2002). Isso pode ajudar a identificar espécies crípticas, zonas de sutura e populações distintas geneticamente (Moritz et al., 2009; Prado et al., 2011) e na proteção de áreas importantes no processo de divergência.

O conhecimento dos processos associados à diversificação de linhagens é especialmente importante em regiões altamente diversas, como é o caso da região neotropical. Para explicar a origem da grande diversidade desta região várias hipóteses foram propostas (revisto por Haffer, 1997), e alguns trabalhos têm buscado testá-las (Marks et al., 2002; Aleixo, 2004; Cabanne et al., 2007; Solomon et al., 2008). Enquanto na Floresta Amazônica existem evidências congruentes com diferentes hipóteses de diversificação (Nores, 1999; Marks et al., 2002; Aleixo, 2004), na Mata Atlântica a maioria dos trabalhos não refuta a teoria dos refúgios (Haffer, 1969) como uma das hipóteses para explicar sua diversidade (Cabanne et al., 2007; Carnaval et al., 2009; Thomé et al., 2010; D'Horta et al., 2011). Essa teoria propõe que durante os períodos glaciais as florestas se fragmentaram, causando diminuição do tamanho populacional das espécies florestais, seguida de expansão, quando as florestas voltaram a se expandir nos períodos inter-glaciais. Esta hipótese tem sido citada em trabalhos recentes, e a presença de sinais de expansão nas populações tem servido de suporte para não rejeitar a teoria dos refúgios para explicar a diversidade dos biomas. Entretanto, há biomas neotropicais pouco estudados, como é o caso de áreas de vegetação aberta (Chaco, Cerrado e Caatinga) e seus organismos permanecem pobremente caracterizados em termos de estrutura genética - e, talvez por isso, não foram considerados em importantes trabalhos de conservação (por exemplo, Myers et al., 2000) -.

Além disso, ainda seria importante incluir análises comparativas entre biomas (Werneck, 2011).

Embora a demonstração empírica de que diferenciação de linhagens com fluxo gênico possa ser difícil (Nosil, 2008), alguns trabalhos têm indicado que a transição entre habitats está relacionada às divergências genética e morfológica (Smith et al., 1997, 2001). Além disso, áreas de transição de habitats (ecótonos) podem também ser áreas de hibridização de espécies e de contato de linhagens (zonas híbridas). Por exemplo, na Serra do Espinhaço, que é uma área de transição entre Cerrado e Floresta Atlântica, *Callithrix penicillata* e *C. geoffroyi* produzem híbridos (Coimbra-Filho et al., 2006). Tais regiões são consideradas como “laboratórios naturais para estudos de evolução” (Hewitt, 1988), o que denota a importância de trabalhos que incorporem espécies com distribuição inter-biomas no estudo de divergência com fluxo gênico.

Um organismo que pode ser usado para explorar essa relação entre biomas é o passeriforme da família Dendrocolaptidae: *Dendrocolaptes platyrostris*. Essa ave ocorre no centro e leste do Brasil e partes adjacentes da Argentina e Paraguai (Fig. 1), sendo a subespécie *D. p. platyrostris* associada às matas úmidas do leste do Brasil, e a subespécie *D. p. intermedius* à diagonal de vegetação aberta da América do Sul (Marantz et al., 2003). Este modelo é interessante, pois as duas subespécies ocorrem em dois biomas diferentes – o que poderia sugerir diferentes regimes seletivos (Moritz et al., 2000) – e diferem entre si em plumagem (Marantz, 1997; Marantz et al., 2003) e vocalização (Willis & Oniki, 2001). Um estudo da variação de plumagem e da estrutura e diversidade genética de duas regiões mitocondriais de *D. platyrostris* revelou a associação entre a diferença de plumagem e os clados mitocondriais, ou seja, há congruência com o bioma de ocorrência (Cabanne et al., 2011). Além disso, os modelos demográficos simulados que envolviam diferenciação sem fluxo gênico foram rejeitados, o que não descarta a possibilidade de as duas formas terem evoluído na presença de fluxo gênico.

Ainda que Cabanne e colaboradores (2011) tenham testado a hipótese de expansão populacional nas populações, utilizando-se da estatística R^2 (Ramos-Onsins & Rozas, 2002), e tenham considerado a hipótese de expansão populacional em seu trabalho, nenhum dos modelos simulados no trabalho incluiu expansão populacional em seus parâmetros. Isto impediu que esta possibilidade pudesse ser contrastada com cenários de tamanho populacional estável com a metodologia de seleção de modelos utilizada no trabalho.

Figura 1. Distribuições geográficas da mata Atlântica (cinza) e de *Dendrocolaptes platyrostris* (pontilhado) e localidades das amostras obtidas (ver Tabela 1). Mapa construído utilizando o pacote sp (Pebesma & Bivand, 2005) do R (R Development Core Team, 2010) com dados de distribuição do bioma mata Atlântica do IBAMA (<http://siscom.ibama.gov.br/shapes>) e distribuição da espécie do Infonatura (<http://natureserve.org/infonatura>).

Embora os resultados de Cabanne e colaboradores (2011) sejam interessantes, há algumas limitações nas análises realizadas, sendo a principal associada ao uso de apenas um marcador (mitocondrial), que:

- 1) possui herança unicamente materna e esses resultados podem refletir apenas a história da matrilinearidade.
- 2) pode não refletir a verdadeira história da espécie devido a variações estocásticas (Hey & Machado, 2003; Hickerson et al., 2006; Knowles, 2009; Zink & Barrowclough, 2008) e não ser suficiente para avaliar se duas populações estão trocando genes (Hey & Machado, 2003). Além disso, inferências estatísticas baseadas em um único loco podem ser limitadas (Zhang & Hewitt, 2003). Adicionalmente, há evidências de que uma estratégia multilocus reduz a quantidade de variância das estimativas do parâmetro theta (Carling & Brumfield, 2007), além de diminuir os intervalos de confiança das estimativas de outros parâmetros populacionais, tais como fluxo gênico e taxa de expansão populacional (Zink & Barrowclough, 2008) – parâmetros importantes nas simulações.
- 3) pode ter sido insuficiente para a análise de distribuição posterior das estimativas de divergência, pois elas não atingiram probabilidade zero, o que limitou o conhecimento sobre o tempo de divergência entre as linhagens de *D. platyrostris* – informação importante para responder várias questões biológicas (Knowles, 2009).

Além disso, as taxas de fluxo gênico dos modelos simulados no trabalho de Cabanne e colaboradores (2011) foram escolhidas arbitrariamente, ao invés de serem estimadas diretamente dos dados. Adicionalmente, uma das estatísticas sumárias usadas para avaliar os modelos simulados foi o F_{ST} , que embora seja uma boa métrica para avaliar divergência entre populações, o F_{ST} pode ser pouco informativo para avaliar taxas de migração em modelos de divergência com fluxo gênico (Nielsen & Wakeley, 2001). Ainda, diferentes processos de migração e deriva podem gerar padrão similar nas diferenças entre frequências de alelos (Holsinger & Weir, 2009), novamente fazendo com que o F_{ST} não seja uma boa escolha para avaliar modelos no contexto desse trabalho, pois as hipóteses testadas no mesmo possuíam taxas de migração semelhantes. Alternativamente, Hickerson e colaboradores (2006) sugerem o uso do D de Tajima (1989) na análise de modelos demográficos. Nas simulações desse trabalho o D de Tajima não foi correlacionado com nenhuma outra estatística, além de fornecer informações a respeito de expansão populacional. Isso maximiza a quantidade de informações extraída dos dados, o que

melhora a avaliação dos modelos.

Ainda, Cabanne e colaboradores (2011) estimaram tamanho populacional efetivo com o método aplicado pelo programa Lamarc, o qual foi desenvolvido para populações panmíticas (Kuhner, 2006) – um cenário já negado no próprio trabalho -, o que pode ter adicionado erro na análise.

Desta forma, um trabalho que utilize uma estratégia multilocos para avaliar o padrão da diversidade genética de *D. platyrostris* pode aumentar o poder das estimativas dos parâmetros demográficos e tornar mais precisas e acuradas as avaliações dos modelos demográficos simulados, permitindo uma melhor avaliação dos principais processos que foram importantes na diferenciação das linhagens.

Assim, como a distribuição geográfica dos clados de *D. platyrostris* é contínua, não havendo barreiras geográficas que impeçam o fluxo gênico entre os mesmos, duas hipóteses, ou uma combinação delas, poderiam explicar a diferenciação destes clados: 1) a teoria dos refúgios, ou seja, a diferenciação sendo causada pela fragmentação do hábitat, seguida de expansão populacional e consequente contato secundário quando o habitat voltou a expandir, e 2) diferenciação com contato primário, sendo causada pela diminuição do fluxo gênico entre os organismos de habitats distintos.

2. Objetivos

O objetivo geral do presente trabalho é verificar se o resultado de divergência com fluxo gênico observado em *Dendrocolaptes platyrostris* com o uso de um único marcador (Cabanne et al., 2011) continua sendo obtido com a inclusão de marcadores nucleares. Para tanto: 1) será avaliado se os clados observados na análise da região controladora mitocondrial de *D. platyrostris* são também observados com marcadores nucleares; 2) espera-se obter melhores estimativas dos parâmetros populacionais de *D. platyrostris* – com menores intervalos de confiança – decorrente do uso de múltiplos marcadores, 3) serão simulados e avaliados modelos demográficos que possam representar a história das linhagens.

3. Material e Métodos

3.1 Área de estudo e amostras

Foram utilizadas 50 amostras de *Dendrocolaptes platyrostris* que representam boa parte da distribuição da espécie (Figura 1, Tabela 1). Destas, 45 foram sequenciadas para o intron 5 do beta fibrinogênio (FIB5) e 42 foram sequenciadas para o intron 2 da mioglobina (Myo2). Além disso, foram utilizadas sequências públicas do citocromo b de 32 dessas amostras (Cabanne et al., 2011). Os dados de sequência da região controladora de Cabanne e colaboradores (2011) não foram incluídos porque o citocromo B e a região controladora são fisicamente ligados, embora possuam taxas de mutação diferentes. Para utilizar estes marcadores na estimativa dos parâmetros demográficos (IMa) eles teriam que ser tratados como um único marcador com uma única taxa de mutação, ou como dois marcadores independentes com taxas de mutação diferentes. Nos dois casos estaríamos violando premissas.

3.2 Extração de DNA

As extrações de DNA foram feitas usando o protocolo padrão de digestão com proteinase K mantendo-se por uma noite a 55° C, seguido de purificação com fenol: clorofórmio: álcool isoamílico (Bruford et al., 1992).

3.3 Quantificação

A quantificação do DNA extraído foi feita utilizando uma alíquota do produto extraído e carregado em gel de agarose 1%, junto com padrões de concentrações conhecidas (“ladder” Pharmacia).

3.4 Amplificação

A amplificação por PCR dos segmentos do íntron 5 do beta fibrinogênio (FIB5) e do íntron 2 da mioglobina (Myo2) foram realizadas utilizando, respectivamente, os *primers* FIB5 e FIB6 (Marini & Hackett, 2002) e MYO2 e MYO3F (Slade et al., 1993; Heslewood et al., 1998, respectivamente). O volume total foi de 10 µl contendo 200 µM de cada dNTP, 1 µM de cada *primer*, 0,1 U de *Taq* polimerase (Pharmacia), tampão 1x (Pharmacia) e 20 –

40 ng de DNA molde. As reações foram realizadas nas seguintes condições: 1) FIB5 – desnaturação inicial a 94°C por 3 min. 30 seg.; seguido de 35 ciclos com 94°C por 35 seg., 51-56°C (10 ciclos a 56°C, 10 ciclos a 53°C e 15 ciclos a 51°C) por 40 seg. e 72°C por 1 min.; e extensão final a 72°C por 9 min.; 2) MYO2 – desnaturação inicial a 95°C por 5 min.; seguido de 10 ciclos a 95°C por 30 seg., 60°C decrescendo 1°C por ciclo por 30 seg. e 72°C 40 seg.; seguido de 30 ciclos a 95°C por 30 seg., 50°C por 30 seg. e 72°C 40 seg.; e extensão final por 9 min. A confirmação do tamanho dos fragmentos amplificados foi feita utilizando uma alíquota do produto amplificado e carregado em gel de agarose 1%, junto com padrões de tamanhos conhecidos (“ladder” Pharmacia).

Tabela 1. Localidade amostrada, número de registro na coleção (^aLaboratório de Genética e Evolução Molecular – Universidade de São Paulo, LGEMA-USP; ^bLaboratório de Biodiversidade e Evolução Molecular – Universidade Federal de Minas Gerais, LBEM-UFMG), tipo de bioma (NFA – norte da Floresta Atlântica, SFA – sul da Floresta Atlântica, CVA – corredor de vegetação aberta), marcador molecular sequenciado: ¹íntron 5 do beta fibrionogênio, ²íntron 2 da mioglobina e ³cit b (^{1,2}obtidos no presente estudo, ³obtidos por Cabanne et al., 2011).

Localidade	Registro	Bioma
1 – Buri, SP	^a 1260 ²	NFA
2 - Pinhalão, PR	^a 1283 ³	SFA
3 - Wencenslau Braz, PR	^a 1355 ^{1,2,3}	SFA
4 - Morro Grande, SP	^a 1431 ^{1,2,3} , ^a 2462 ^{1,2,3} , ^a 2876 ^{1,2} , ^a 2916 ^{1,2} , ^a 2940 ¹ , ^a 3064 ² , ^a 3095 ^{1,2}	NFA
5 - P.N. Itatiaia, RJ	^a 1724 ^{1,2,3}	NFA
6 - Morro do Diabo, SP	^a 1894 ^{1,2,3} , ^a 2297 ^{1,2,3} , ^a 2308 ^{1,2,3} , ^a 2344 ^{1,2,3} , ^a 2348 ^{1,2}	SFA
7 - Fazenda Barreiro Rico, SP	^a 2094 ^{1,2}	NFA
8 - Pelotas, RS	^a 2134 ^{1,2,3} , ^a 2136 ^{1,2,3}	SFA
9 - Rancho Queimado, SC	^a 2148 ^{1,2,3}	SFA
10 – P.N. Serra das Confusões, PI	^a 2249 ^{1,3} , ^a 2577 ^{1,2,3} , ^a 2673 ^{1,2,3} , ^a 2674 ^{1,3} , ^a 2725 ³	CVA
11 – Bonito, BA	^a 2668 ^{1,2,3}	CVA
12 - Paranapanema, SP	^a 10449 ^{1,2} , ^a 10450 ^{1,2}	NFA
13 – Ibiapaba, CE	^a 11443 ^{1,3}	CVA
14 - Três lagoas, MT	^a 11444 ³	CVA
15 - Simonésia, MG	^b B0431 ^{1,3}	NFA
16 - Araponga, MG	^b B0930 ^{1,2,3}	NFA
17 - Felixlândia, MG	^b B1625 ^{1,2,3}	CVA
18 – Brasilândia de Minas, MG	^b B1718 ^{1,2,3} , ^b B2055 ^{1,2,3} , ^b B2058 ^{1,2,3} , ^b B2064 ^{1,2,3} , ^b B2083 ^{1,2,3} , ^b B2278 ^{1,2,3}	CVA
19 – Bocaiúva, MG	^b B1790 ^{2,3} , ^b B2215 ^{1,2,3}	CVA
20 – Leme do Prado, MG	^b B2411 ^{1,2} , ^b B2934 ^{1,2} , ^b B3255 ^{1,2}	CVA
21 – Alto Rio Novo, ES	^b B2684 ^{1,2} , ^b B2762 ^{1,2}	NFA
22 - Alfredo Chaves, ES	^b B2737 ^{1,2}	NFA
23 – Sooretama, ES	^b B2779 ^{1,2} , ^b B2845 ^{1,2} , ^b B2846 ¹	NFA
24 - Chapada Diamantina, BA	^a 13404 ^{1,2}	CVA

3.5 Purificação

Para purificar os produtos amplificados seguimos o seguinte protocolo (Paithankar & Prasad, 1991): (1) acrescentar 1 µl de PEG (polietilenoglicol) 20 % para cada µl de produto de amplificação, misturar em agitador de tubos, incubar a 37°C por 15 minutos e centrifugar a 12000 rpm por 15 minutos; (2) descartar o sobrenadante e acrescentar 125 µl de etanol 80% gelado, centrifugar a 12000 rpm por 5 minutos; (3) repetir etapa 2; (4) secar em centrífuga a vácuo por 10 minutos a 45°C; (5) acrescentar 1 µl de água miliQ para cada µl de produto de amplificação inicial.

3.6 Sequenciamento

O sequenciamento foi feito utilizando o *BigDye Terminator Cycle Sequencing Kit* (Applied Biosystems) seguindo o protocolo do fabricante, no sequenciador automático ABI3130 (Applied Biosystems).

3.7 Análises

As sequências foram editadas usando o programa CodonCode Aligner 3.5.5 (CodonCode Corporation). Os sítios heterozigotos foram verificados visualmente e utilizados os códigos da IUPAC (*International Union of Pure and Applied Chemistry*) para representar as bases correspondentes. As sequências foram alinhadas utilizando o programa ClustalW (Thompson et al., 1994) implementado no MEGA5 (Tamura et al., 2011). Para confirmar a identidade das sequências foi utilizado o algoritmo BLAST no GenBank. As fases heterogaméticas foram separadas usando o algoritmo PHASE (Stephens et al., 2001) implementado no DNASP v5 (Librado & Rozas, 2009) utilizando os parâmetros *default*. As sequências com probabilidades menores que 0.6 foram excluídas das análises.

Para as análises de coalescência e as simulações das hipóteses demográficas assumimos o modelo evolutivo HKY (Hasegawa et al., 1985), já que este está disponível nos programas utilizados e é um modelo relativamente simples, o que auxilia na convergência das análises. Além disso, Strasburg e Rieseberg (2010) mostraram evidências de que a estimativa de parâmetros demográficos utilizando o modelo evolutivo HKY, mesmo quando os dados simulados foram gerados baseados em outros modelos evolutivos, resulta em menos erros nas estimativas do que quando outros modelos evolutivos são utilizados em dados gerados com o HKY. Os parâmetros da distribuição gama do modelo

evolutivo foram estimados com o programa MEGA5 (Tamura et al., 2011). As estatísticas sumárias foram estimadas usando o pacote pegas (Paradis, 2010) implementado no pacote estatístico R (R Development Core Team, 2010). O teste PHI (Bruen et al., 2006), implementado no programa SplitsTree4 (Huson & Bryant, 2006) foi utilizado para verificar a presença de sinal de recombinação nas sequências. Os marcadores que apresentaram recombinação com o teste PHI foram reanalisados com o teste de 4 gametas (Hudson & Kaplan, 1985) implementado no DNASP (Librado & Rozas, 2009) para selecionar um segmento sem sinal de recombinação. Redes de haplótipos foram construídas utilizando o algoritmo de median-joining no NETWORK 4.6 (<http://www.fluxus-engineering.com>). As redes de haplótipos resultante foram coloridas de acordo com as linhagens mitocondriais de Cabanne e colaboradores (2011).

Para avaliar a hipótese de existência da mesma estrutura genética populacional baseado nos dados mitocondriais (Cabanne et al., 2011) e nucleares (dados aqui apresentados), aplicamos a análise de variância molecular (AMOVA; Excoffier et al., 1992) em dois níveis: 1) Floresta Atlântica (FA) vs corredor de vegetação aberta (CVA) e 2) norte da Floresta Atlântica (NFA) vs sul da Floresta Atlântica (SFA), norte de CVA e sul de CVA. Essa análise foi realizada no pacote ade4 (Dray & Dufour, 2007) implementado no pacote estatístico R (R Development Core Team). Também calculamos o valor de F_{ST} (Ney, 1973) usando o pacote adegenet (Jombart, 2008), implementado no pacote estatístico R (R Development Core Team). As significâncias destas análises foram avaliadas com 1.000 permutações cada, sendo consideradas significativas aquelas acima de 0,05.

A fim de avaliar a hipótese de expansão populacional, realizamos o teste de expansão R_2 (Ramos-Onsins & Rozas, 2002) e o teste de neutralidade de D (Tajima, 1989) implementado no pacote pegas (Paradis, 2010). Também utilizamos o método *Extended Bayesian Skyline Plot* (Heled & Drummond, 2008) implementado no BEAST (Drummond & Rambaut, 2007) seguindo instruções do manual do programa. Adotamos a taxa de mutação estimada do citocromo B para Aves de 2,1% (Weir & Schluter, 2008) e essa taxa foi utilizada para estimar as taxas de cada marcador nuclear no BEAST. Todas as análises foram feitas com 100 milhões de gerações, sendo o *burn-in* de 10 milhões. A convergência foi avaliada baseando-se nas estimativas de *effective sample size* (ESS).

Com o objetivo de estimar os parâmetros demográficos (taxas de migração, tempo de divergência e tamanho populacional efetivo) utilizamos o modelo de isolamento com

migração (Hey & Nilsen, 2004) como implementado no programa IMA2 (Hey, 2010). A taxa de mutação assumida foi a mesma utilizada nas análises de *Extended Bayesian Skyline Plot* (2,1% para o citocromo B, juntamente com o intervalo de confiança a 95% de +/- 0,1; Weir & Schluter, 2008). A taxa de mutação do DNA nuclear e seu intervalo de confiança foram estimados no BEAST a partir da taxa do DNA mitocondrial conforme explicado acima, juntamente com seu intervalo de confiança. Como para cada população foi feita uma análise, para cada marcador obtivemos três médias e três intervalos de confiança. Então, utilizamos a média das médias entre as três corridas (três populações) e utilizamos os limites mais externos entre os intervalos de confiança (Tabela 2) como os limites de confiança. Os intervalos de confiança das taxas de mutação foram usados como *priors* para a taxa de mutação escalar. A fim de simplificar os modelos para que fosse possível alcançar convergência dos parâmetros, as análises foram feitas utilizando taxas de migração simétricas entre as populações, e tamanho populacional efetivo ancestral igual ao maior tamanho populacional efetivo dentre as populações descendentes. Para cada corrida foram usadas cinco cadeias de Markov linear, e as análises foram feitas com um milhão de gerações, sendo o *burn-in* de um milhão. A convergência dos valores dos parâmetros foi avaliada repetindo as corridas pelo menos mais uma vez. Além disso, verificamos a relação entre as estimativas baseando-se nas estimativas da primeira e segunda metades da corrida. Espera-se que estes valores sejam semelhantes caso a análise esteja explorando bem o espaço de parâmetros. As corridas que não convergiram foram repetidas com 10 milhões de gerações, sendo também repetida pelo menos uma vez. Entre as corridas que convergiram utilizamos aquelas que obtiveram maiores valores de *ESS*.

Tabela 2. Taxas de mutação por milhão de anos dos marcadores nucleares para o marcador como um todo (não por sítio), estimadas na corrida do BEAST com o método *Extended Bayesian Skyline Plot*. LIC= limite inferior do intervalo de confiança. UIC= limite superior do intervalo de confiança.

	CVA			NFA			SFA		
	LIC	Média	UIC	LIC	Média	UIC	LIC	Média	UIC
FIB5	3E-3	1,4E-2	3E-2	2E-3	3E-2	8E-2	6E-3	3E-2	4E-2
Myo2	9E-3	2,5E-2	4E-2	1E-7	8E-2	0,4	6E-4	1,6E-2	3,6E-2

Para avaliar a relação filogenética entre as populações mitocondriais (CVA, NFA e SFA) simulamos todas suas possíveis relações (Figs. 2 e 3). O número de amostras utilizadas nas simulações para cada marcador foi o mesmo dos dados reais do presente trabalho (Tabela 1). As hipóteses de 1 a 3 possuem uma relação hierárquica de divergência entre as populações, enquanto nas hipóteses 4 e 5 a divergência ocorre simultaneamente. A hipótese 6 assume que todas as populações constituem uma única população. A hipótese 5 é diferente da hipótese 4, pois na hipótese 5 não existe um tempo fixo para que as hipóteses coalescam. Os parâmetros populacionais (tempo de divergência, tamanho populacional efetivo ancestral e atual, e taxas de migração) utilizados nas simulações foram aqueles obtidos após as análises do IMA2 (Tabela 3). Nas simulações, para o tempo de divergência e a taxa de migração, foi assumido distribuição uniforme do valor com maior probabilidade estimado pelo IMA2 (Tabela 3) e adicionado de 10% de erro para mais e para menos (Tabela 4). O tamanho populacional efetivo de cada população foi simulado com distribuição uniforme, indo do menor ao maior valor cujas probabilidades estimadas nas corridas do IMA2 foram as mais altas (Tabela 3) e sendo o menor valor decrescido de 10%, e o maior valor acrescido de 10% (Tabela 5). As taxas de mutação utilizadas nas simulações foram as mesmas utilizadas nas análises de isolamento com migração (Tabela 2). Foi assumido que essas taxas possuem distribuição uniforme contendo os valores extremos considerando os intervalos de confiança de cada marcador. Para cada um dos cenários demográficos simulados, outro cenário igual, porém incluindo expansão populacional também foi testado (esses cenários serão denominados usando o sufixo 'E'; por ex: 1E, 2E, etc.). A taxa de expansão utilizada nas simulações foi incluída como um intervalo, de forma que cada população ancestral corresponda a 50-80% do tamanho das populações descendentes. A data de início da expansão foi simulada como um intervalo, cujos valores foram extraídos dos resultados aproximados das análises do *Bayesian Skyline Plot*, acrescidos e reduzidos de 20%. Para as hipóteses cuja divergência é mais recente do que o início da expansão estimado pelas análises do *Bayesian Skyline Plot*, utilizamos um intervalo correspondente a 10% do intervalo utilizado nos demais modelos. Como a hipótese 4.f inclui divergência muito recente (Tabela 4), simulamos a expansão como um intervalo desde logo após a separação entre as populações (2000 – 3000 anos) até o presente. Como as hipóteses 6 e 6E constituem uma população panmítica cada uma, separamos as sequências simuladas de forma aleatória em três grupos representando as três populações (CVA, SFA e NFA),

utilizando o mesmo número amostral dos dados observados para cada marcador, e então estimamos as estatísticas sumárias para os mesmos. Para simular estes cenários utilizamos o programa Bayesian Serial SimCoal (Anderson et al., 2005), uma modificação do programa SimCoal 1.0 (Excoffier et al., 2000). Para cada modelo foram feitas 10 mil simulações.

A hipótese 4 (Fig. 3) assume a existência de uma politomia. Nesse caso, como qualquer relação entre as populações (NFA, SFA e CVA) seria aceita, foram testadas todas as datas de divergência estimadas entre os possíveis três pares de populações (análises 1, 2, 3, 4, 5 e 6 da tabela 3).

Figura 2. Representação das hipóteses simuladas. SFA = sul da Floresta Atlântica, NFA = norte da Floresta Atlântica, CVA = corredor de vegetação aberta, A1 = população ancestral 1, A2 = população ancestral 2, t1 e t2 = tempo de divergência correspondente à região demarcada pelas linhas pontilhadas. Os números no topo de cada hipótese representam a numeração de cada uma delas (Tabelas 4 e 5). Os valores de parâmetros usados nas simulações (tamanhos efetivos das populações descendentes -NFA, SFA e CVA- e ancestrais, taxas de migração e tempos de divergência aproximados entre as populações descendentes e ancestrais e taxas de mutação dos marcadores) foram obtidos dos próprios dados e são apresentados nos “Resultados”.

Figura 3. Representação das hipóteses simuladas. SFA = sul da Floresta Atlântica, NFA = norte da Floresta Atlântica, CVA = corredor de vegetação aberta, A1 = população ancestral 1, t_0 = tempo de coalescência, t_1 = presente, m = taxa de migração. Os números no topo de cada hipótese representam a numeração de cada uma delas (Tabelas 4 e 5). Os valores de parâmetros usados nas simulações (tamanhos efetivos das populações descendentes -NFA, SFA e CVA- e ancestrais, taxas de migração e taxas de mutação dos marcadores) foram obtidos dos próprios dados e são apresentados nos “Resultados”.

Tabela 3. Resultado das estimativas dos parâmetros populacionais. q_1 = Tamanho populacional efetivo da população 1, q_2 = tamanho populacional efetivo da população 2, M = taxa de migração ($M=m/u$, onde u é a média geométrica entre os locos vezes o tempo de geração em anos), t = tempo de divergência. Valores acima correspondem aos valores de probabilidade mais alta, enquanto valores abaixo correspondem aos intervalos de confiança. *valores cuja distribuição posterior não chegou em zero.

Análise	Pop. 1	Pop. 2	q_1	q_2	M	t
1	CVA	NFA	53527 24594 - 95192	70309 24015 - *	1.116 0 - 3.852	10339 1697 - 197058
2	SFA	CVA	39764 14800 - 94329	110020 49750 - 233058	1.287 0.377 - 3.263	102353 31027 - *
3	SFA	NFA	20475 5890 - 63669	168009 84425 - *	2.329 0 - 7.191	23486 2842 - *
4	FA	CVA	65535 37182 - 120251	64540 26239 - 150345	1.911 0.767 - 4.485	28055 9749 - *
5	SFA	(CVA+NFA)	30316 10264 - 79012	156830 85218 - 311033	1.8840 0.6600 - 4.092	69893 27117 - *
6	NFA	(CVA+SFA)	47068 24184 - 76193	39266 23664 - *	0.012 0 - 6.324	2444 0 - 47380

Tabela 4. Taxas de migração e tempos de divergência aproximados (t_1 e t_2) entre as populações descendentes (NFA, SFA e CVA) e ancestrais (A1 e A2) obtidos após análise no IMA2 e que foram utilizados nas simulações (Fig. 2 e 3). D = população irmã da população A1. Para a hipótese 4 foram testadas seis datas de divergência diferentes (4.a, 4.b, 4.c, 4.d, 4.e, 4.f).

Hipótese	NFA<->SFA	NFA<->CVS	CVS<->SFA	A1<->D	t_1	t_2
1	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	7E-6 - 9E-6	21000 - 26000	26000 - 31000
2	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	7E-6 - 9E-6	9000 - 11000	63000 – 77000
3	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	5E-9 - 6E-8	92000 - 113000	120000 - 150000*
4.a	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	9000 - 11000	-
4.b	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	21000 - 26000	-
4.c	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	92000 - 113000	-
4.d	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	25000 - 31000	-
4.e	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	63000 - 77000	-
4.f	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	2000 - 3000	-
5	9E-6 - 11E-6	4,3E-6 – 5,3E-6	5E-6 – 6.1E-6	-	Não fixado	-

Tabela 5. Tamanhos efetivos aproximados das populações descendentes (NFA, SFA e CVA) e ancestrais (A1 e A2) obtidos após análise no IMA2 e que foram utilizados nas simulações (Fig. 2 e 3).

Hipótese	NFA	SFA	CVA	A1	A2
1, 2, 3	63000 – 185000	18000 – 44000	48000 – 121000	63000 – 185000	63000 – 185000
4	63000 – 185000	18000 – 44000	48000 – 121000	63000 – 185000	-
5	63000 – 185000	18000 – 44000	48000 – 121000	-	-
6	129000 - 350000			-	-

Para estimar a distribuição posterior dos valores de estatísticas sumárias das hipóteses aceitas em relação às estatísticas observadas utilizamos *approximate Bayesian computation* (ABC; Beaumont et al., 2002, Csilléry et al., 2010). Seguindo a sugestão de Hickerson e colaboradores (2006), a diversidade nucleotídica e o D de Tajima (1989) foram selecionados como estatísticas sumárias para a comparação das hipóteses com o observado. Essas estatísticas foram calculadas para todos os marcadores separadamente, para cada população e para a espécie como um todo. A fim de diminuir o número de dimensões de estatísticas observadas (3 marcadores * (3 populações + espécie como um todo) * 2 estatísticas = 24 dimensões) usamos análise de componentes principais (PCA) e retemos o menor número de componentes necessário para manter 99% da informação dos dados. Utilizamos o algoritmo de regressão não linear para estimar a distribuição posterior (Blum & François, 2010) como implementado no pacote abc (Csilléry et al., 2011) do R (R Development Core Team). Em simulações, foi observado que esse algoritmo é robusto

quando há muitas dimensões nos dados (Blum & François, 2010), como no caso do presente trabalho. Para selecionar o melhor modelo após a aplicação de *ABC* utilizamos o fator de Bayes. Para selecionar a tolerância de distância dos valores simulados em relação às estatísticas observadas aplicamos a mesma metodologia utilizada acima – ou seja, *ABC* seguido de fator de Bayes - em 300 simulações aleatoriamente escolhidas (50 de cada hipótese das Tabelas 4 e 5) cuja relação entre as populações e parâmetros populacionais são conhecidos e avaliamos a capacidade da metodologia de identificar a hipótese que gerou os dados, e diminuimos a tolerância gradativamente até que uma porcentagem superior a 95% de acerto fosse alcançada.

4. Resultados

4.1 Marcadores moleculares

4.1.1 Intron 5 beta fibrinogênio (FIB5)

O alinhamento das sequências resultou em 503 pb. Após a separação das fases gaméticas, 82 sequências possuíam probabilidade superior a 0,6. O teste PHI não detectou recombinação nas sequências do FIB5. A figura 4 mostra a rede de haplótipos obtida baseada nesse marcador. A tabela 6 mostra algumas estatísticas sumárias. A tabela 2 apresenta a taxa de mutação juntamente com o intervalo de confiança estimado para este marcador. A tabela 7 apresenta os parâmetros utilizados para simular este marcador.

Figura 4. Rede de haplótipos do íntron 5 do beta fibrinogênio baseada em 503 pb e 82 sequências. SFA = sul da Floresta Atlântica, NFA = norte da Floresta Atlântica, CVA = corredor de vegetação aberta.

4.1.2 Intron 2 da mioglobina (Myo2)

O alinhamento das sequências resultou em 593 pb. Como o teste PHI encontrou evidência de recombinação para este marcador foi utilizado o teste de 4 gametas que identificou as regiões onde possivelmente houve recombinação. O segmento não recombinante com o maior número de sítios polimórficos possuía 68 pb. Em seguida, foram selecionadas as sequências com probabilidade superior a 0,6 quanto às suas fases gaméticas. A figura 5 mostra a rede de haplótipos construída com este marcador. A tabela 6 mostra

algumas estatísticas sumárias. A tabela 2 apresenta a taxa de mutação juntamente com o intervalo de confiança estimado para este marcador. A tabela 7 apresenta os parâmetros utilizados para simular este marcador nas análises de filogeografia estatística.

Tabela 6. Estatísticas sumárias e suas significâncias calculadas para cada marcador (FIB5, Myo2, CitB), para toda a espécie e para cada clado da região controladora. π = diversidade nucleotídica, H = diversidade haplotípica, N = número de sequências, D = teste de neutralidade D de Tajima (1989), R2 = teste de expansão R2 (Ramos-Onsins & Rozas, 2002), NFA = norte da Floresta Atlântica, SFA = sul da Floresta Atlântica, CVA = corredor de vegetação aberta, - valor não estimado devido à insuficiência de variabilidade genética.

	π	H	Sítios polimórficos	Pares de base	N
FIB5	0,0055	0,8458	15	531	82
Myo2	0,0033	0,2167	4	68	84
CitB	0,0044	0,9032	16	962	32
NFA	π	H	D	R2	N
FIB5	0,0063	0,875	-0,049; P = 0,960	0,119; P=0,541	32
Myo2	0,0024	0,1603	-1,284; P = 0,199	0,086; P=0,215	34
CitB	0,0019	0,9333	-0,825; P = 0,408	0,158; P= 0,193	6
SFA	π	H	D	R2	N
FIB5	0,0054	0,7917	-0,01664213; P=0,987	0,136; P= 0,456	16
Myo2	0	0	-	-	18
CitB	0,0027	0,6944	0,688; P = 0,492	0,201; P= 0,789	9
CVA	π	H	D	R2	N
FIB5	0,0044	0,7219	-0,294; P = 0,768	0,092; P= 0,273	34
Myo2	0,0058	0,375	-1,079; P = 0,280	0,093; P= 0,253	32
CitB	0,0038	0,7794	0,871; P = 0,384	0,170; P= 0,780	17

Tabela 7. Valores aproximados de parâmetros das sequências usadas nas simulações (razão entre transições e transversões, taxa de mutação e parâmetros da distribuição gamma). Taxa de mutação dada por linhagem, por par de base, por milhão de anos. r = parâmetro de forma, α = parâmetro de escala.

Marcador	Razão entre transições e transversões	Taxa de mutação	Parâmetros distribuição gamma (r e α)
Cit B	4,67	9E-6 - 10E-6	-
Myo2	0,5	4E-12 - 1E-5	-
FIB5	2,17	5E-7 - 2E-5	0,4 – 10

Figura 5. Rede de haplótipos do intron 2 da mioglobina baseada em 68 pb e 84 sequências. SFA = sul da Floresta Atlântica, NFA = norte da Floresta Atlântica, CVA = corredor de vegetação aberta.

4.1.3 Citocromo B (*CitB*)

O alinhamento das 32 sequências de *CitB* resultou em 962 pb. A Figura 6 mostra a rede de haplótipos construída com este marcador. A tabela 6 mostra algumas estatísticas sumárias. A tabela 2 apresentam a taxa de mutação juntamente com o intervalo de confiança estimado para este marcador. A tabela 7 apresenta os parâmetros utilizados para simular este marcador nas análises de filogeografia estatística.

Figura 6. Rede de haplótipos do citocromo b baseada em 962 pb e 32 sequências. SFA = sul da Floresta Atlântica, NFA = norte da Floresta Atlântica, CVA = corredor de vegetação seca.

4.2 Estrutura genética populacional

As redes de haplótipos dos marcadores nucleares e do marcador mitocondrial mostraram padrões semelhantes. Entretanto, na rede de haplótipos do CitB (Fig. 6) boa parte dos haplótipos são exclusivos de cada população (75%) , enquanto que as redes de haplótipos dos marcadores nucleares (Fig. 4 e 5) apresentam a grande maioria dos haplótipos compartilhados entre as populações. Os testes aplicados para verificar a estrutura genética populacional (Tabelas 8 e 9) só a identificaram entre CVA e SFA (F_{ST} : CitB e FIB5) e no segundo nível das análises de AMOVA, isto é, entre SFA e NFA e entre norte e sul do CVA.

Tabela 8. Valores de F_{ST} entre as populações por marcador (abaixo da diagonal) e nível de significância (acima da diagonal). SFA = sul da Floresta Atlântica, NFA = norte da Floresta Atlântica, CVA = corredor de vegetação aberta. *valores significativos assumindo alfa de 5%.

FIB5	CVA	NFA	SFA		Myo2	CVA	NFA	SFA		CitB	CVA	NFA	SFA
CVA		0,14	0,012		CVA		0,30	0,14		CVA		0,072	<0.01
NFA	0,027		0,48		NFA	0,019		0,68		NFA	0,103		0,14
SFA	0,063*	0,019			SFA	0,047	0,016			SFA	0,184*	0,13	

Tabela 9. Índice *Phi* (Φ_{ST}), quantidade de variação e significância da análise de variância molecular (AMOVA). SFA = sul da Floresta Atlântica, NFA = da norte Floresta Atlântica, CVA = corredor de vegetação aberta. *valores significativos assumindo alfa de 0,05%.

CitB	Φ_{ST}	Variação	P
Dentro das amostras (cerrado, caatinga, SFA e NFA)	0,041	0,03	0,18
Entre as regiões dentro dos biomas (cerrado vs. caatinga e SFA vs. NFA)	0,024	0,0007	0,247
Entre os biomas (CVA vs. FA)	0,018	0,0005	0,68
FIB5			
Dentro das amostras (cerrado, caatinga, SFA e NFA)	0,07*	0,0333*	0,005
Entre as regiões dentro dos biomas (cerrado vs. caatinga e SFA vs. NFA)	0,065*	0,0023*	0,026
Entre os biomas (CVA vs. FA)	0,004	0,00014	0,656
Myo2			
Dentro das amostras (cerrado, caatinga, SFA e NFA)	0,012*	0,121*	0,005
Entre as regiões dentro dos biomas (cerrado vs. caatinga e SFA vs. NFA)	0,078*	0,0011*	0,043
Entre os biomas (CVA vs. FA)	0,47	0,0007	0,356

4.3 Expansão populacional

O teste de expansão populacional e o teste de neutralidade (Tabela 6) não encontraram evidência significativa de que houve expansão demográfica em nenhuma das populações. Todas as análises de mudança de tamanho populacional efetivo em função do tempo analisadas no BEAST (*Extended Bayesian Skyline Plot*) convergiram, ou seja os valores de ESS ficaram acima do mínimo desejado. O gráfico resultante da análise de mudança no tamanho efetivo com os três marcadores combinados mostra tamanho populacional estável para NFA (Figura 7 A), aumento populacional na população do SFA e CVA (Figura 7, B e C, respectivamente), embora o intervalo de confiança destes seja muito amplo e não nos permite rejeitar a hipótese de tamanho populacional constante.

4.4 Parâmetros populacionais

Todas as análises de isolamento com migração convergiram, ou seja, o resultado foi semelhante entre as duas réplicas. Os valores dos parâmetros estimados baseadas na primeira e segunda metades da análise foram semelhantes para todos os parâmetros. Apesar da convergência das análises, alguns parâmetros das mesmas não alcançaram probabilidade zero em uma das extremidades da distribuição posterior. A tabela 3 apresenta os resultados dessas análises.

Figura 7. Análise de mudança de tamanho efetivo em função do tempo (*extended Bayesian skyline plot*). A) norte da floresta Atlântica, B) sul da floresta Atlântica, C) corredor de vegetação aberta. A linha central mostra a mediana do tamanho populacional efetivo em escala logarítmica, enquanto as linhas mais externas mostram o intervalo de confiança.

4.5 Filogeografia estatística

A análise de componentes principais reduziu as 24 dimensões dos dados para 10, retendo mais de 99% da informação dos mesmos. Como a hipótese 6E possui valores muito diferentes das demais hipóteses, mesmo análises com a tolerância mais alta testada no presente trabalho rejeitavam 100% de suas réplicas. Por este motivo esta hipótese foi excluída do teste de seleção de tolerância e das análises finais. O procedimento que usamos para selecionar a tolerância usada nas análises de *approximate Bayesian computation* aumentou significativamente a acurácia do método, indo de menos que 50% de acerto com uma tolerância de 0,05 até superior a 95% com tolerância de 0,002, valor que utilizamos para avaliar os modelos em relação as estatísticas sumárias de nossos dados.

A hipótese cuja distribuição posterior das estatísticas (D e π) dos dados simulados mais se aproximou dos resultados observados foi a hipótese 4.cE (Fig. 2, Tabela 4, 5, 10 e 11), seguido pelas hipóteses 4.dE e 2E, que tiveram probabilidade posterior de aproximadamente 0,16, 0,13 e 0,09, respectivamente (Tabela 10). Ou seja, dentre as hipóteses testadas, a que sugere que as populações de *Dendrocolaptes platyrostris* se diferenciaram entre 92 e 113 mil anos atrás, com uma relação de politomia, uma taxa relativamente baixa de fluxo gênico, e expansão recente desde 20 a 30 mil anos atrás, é a que possui probabilidade posterior mais alta.

Tabela 10. Probabilidade posterior estimada para as hipóteses demográficas.

1	1E	2	2E	3	3E	4.a	4.aE	4.b	4.bE	4.c
0.004	0.046	0.0898	0.092	0.0301	0.0037	0.0118	0.0663	0.0144	0.0716	0.0317
4.cE	4.d	4.dE	4.e	4.eE	4.f	4.fE	5	5E	6	6E
0.1606	0.0901	0.1259	0.0067	0.0487	0.0429	0.0445	0.0044	0.0076	0.0071	0

Tabela 11. Comparação par a par das hipóteses demográficas utilizando fator de Bayes.

	1	1E	2	2E	3	3E	4.a	4.aE	4.b	4.bE
1	1	0.0879	0.0449	0.0439	0.1339	1.0796	0.3409	0.0609	0.2803	0.0564
1E	11.3814	1	0.5115	0.4994	1.5241	12.2877	3.8801	0.6927	3.1898	0.6422
2	22.2509	1.955	1	0.9764	2.9796	24.0227	7.5857	1.3543	6.2362	1.2555
2E	22.7889	2.0023	1.0242	1	3.0516	24.6036	7.7692	1.3871	6.387	1.2858
3	7.4678	0.6561	0.3356	0.3277	1	8.0625	2.5459	0.4545	2.093	0.4214
3E	0.9262	0.0814	0.0416	0.0406	0.124	1	0.3158	0.0564	0.2596	0.0523
4.a	2.9333	0.2577	0.1318	0.1287	0.3928	3.1668	1	0.1785	0.8221	0.1655
4.aE	16.4296	1.4436	0.7384	0.7209	2.2001	17.7379	5.6012	1	4.6047	0.927
4.b	3.568	0.3135	0.1604	0.1566	0.4778	3.8521	1.2164	0.2172	1	0.2013
4.bE	17.723	1.5572	0.7965	0.7777	2.3733	19.1343	6.0421	1.0787	4.9672	1
4.c	7.8469	0.6895	0.3527	0.3443	1.0508	8.4717	2.6752	0.4776	2.1992	0.4428
4.cE	39.7838	3.4955	1.788	1.7458	5.3274	42.9517	13.563	2.4215	11.1502	2.2448
4.d	22.3165	1.9608	1.0029	0.9793	2.9884	24.0936	7.6081	1.3583	6.2546	1.2592
4.dE	31.1752	2.7391	1.4011	1.368	4.1746	33.6577	10.6282	1.8975	8.7374	1.759
4.e	1.6476	0.1448	0.074	0.0723	0.2206	1.7788	0.5617	0.1003	0.4618	0.093
4.eE	12.0681	1.0603	0.5424	0.5296	1.616	13.0291	4.1142	0.7345	3.3823	0.6809
4.f	10.6186	0.933	0.4772	0.466	1.4219	11.4641	3.6201	0.6463	2.9761	0.5991
4.fE	11.027	0.9689	0.4956	0.4839	1.4766	11.905	3.7593	0.6712	3.0905	0.6222
5	1.0985	0.0965	0.0494	0.0482	0.1471	1.186	0.3745	0.0669	0.3079	0.062
5E	1.8714	0.1644	0.0841	0.0821	0.2506	2.0205	0.638	0.1139	0.5245	0.1056
6	1.7657	0.1551	0.0794	0.0775	0.2364	1.9063	0.602	0.1075	0.4949	0.0996

Continuação da **Tabela 11.**

	4.c	4.cE	4.d	4.dE	4.e	4.eE	4.f	4.fE	5	5E
1	0.1274	0.0251	0.0448	0.0321	0.6069	0.0829	0.0942	0.0907	0.9103	0.5343
1E	1.4504	0.2861	0.51	0.3651	6.9079	0.9431	1.0718	1.0321	10.3607	6.0816
2	2.8356	0.5593	0.9971	0.7137	13.5051	1.8438	2.0955	2.0179	20.2554	11.8897
2E	2.9042	0.5728	1.0212	0.731	13.8316	1.8884	2.1461	2.0667	20.7451	12.1772
3	0.9517	0.1877	0.3346	0.2395	4.5325	0.6188	0.7033	0.6772	6.7981	3.9904
3E	0.118	0.0233	0.0415	0.0297	0.5622	0.0768	0.0872	0.084	0.8432	0.4949
4.a	0.3738	0.0737	0.1314	0.0941	1.7803	0.2431	0.2762	0.266	2.6702	1.5674
4.aE	2.0938	0.413	0.7362	0.527	9.9719	1.3614	1.5473	1.49	14.9562	8.7791
4.b	0.4547	0.0897	0.1599	0.1145	2.1656	0.2957	0.336	0.3236	3.248	1.9066
4.bE	2.2586	0.4455	0.7942	0.5685	10.7569	1.4686	1.6691	1.6072	16.1336	9.4703
4.c	1	0.1972	0.3516	0.2517	4.7626	0.6502	0.739	0.7116	7.1432	4.193
4.cE	5.07	1	1.7827	1.2761	24.1466	3.2966	3.7466	3.6079	36.2159	21.2584
4.d	2.844	0.5609	1	0.7158	13.5449	1.8492	2.1016	2.0238	20.3151	11.9248
4.dE	3.9729	0.7836	1.397	1	18.9216	2.5833	2.9359	2.8272	28.3793	16.6584
4.e	0.21	0.0414	0.0738	0.0528	1	0.1365	0.1552	0.1494	1.4998	0.8804
4.eE	1.5379	0.3033	0.5408	0.3871	7.3247	1	1.1365	1.0944	10.9858	6.4486
4.f	1.3532	0.2669	0.4758	0.3406	6.4449	0.8799	1	0.963	9.6663	5.674
4.fE	1.4053	0.2772	0.4941	0.3537	6.6927	0.9137	1.0385	1	10.038	5.8922
5	0.14	0.0276	0.0492	0.0352	0.6667	0.091	0.1035	0.0996	1	0.587
5E	0.2385	0.047	0.0839	0.06	1.1359	0.1551	0.1762	0.1697	1.7036	1
6	0.225	0.0444	0.0791	0.0566	1.0717	0.1463	0.1663	0.1601	1.6074	0.9435

5. Discussão

5.1 Padrão da variação

A rede de haplótipos de *Dendrocolaptes platyrostris* baseada em 584 pb da região controladora do trabalho de Cabanne e colaboradores (2011) apresenta clados relativamente bem separados congruentes com três regiões geográficas (NFA, SFA e CVA). Esse resultado também é encontrado nas análises de variância molecular tanto na quantidade de variância entre as populações quanto em nível de significância das análises. No presente estudo avaliamos basicamente a mesma amostragem de indivíduos (com alguns indivíduos adicionais do CVA), mas utilizando marcadores que possuem taxa de mutação mais lenta. A rede de haplótipos do citocromo B mostra vários alelos exclusivos para cada população, porém este resultado não se reflete em significância estatística nas análises realizadas ou em padrão de estruturação na rede de haplótipos, como foi encontrado nas análises baseadas na região controladora. Embora a região controladora e o citocromo B sejam fisicamente ligados, a taxa de mutação do citocromo B é mais lenta, o que poderia explicar a ausência de estruturação encontrada na rede de haplótipos e nos testes utilizados. O único valor significativo observado nas análises do citocromo B é em relação à divergência entre o CVA e SFA. Isso sugere que o fluxo gênico entre essas populações é menor e/ou que as mesmas entraram em isolamento há mais tempo, quando comparado à relação de fluxo gênico e/ou isolamento entre a população do NFA e as outras duas populações. De fato, em nossas análises demográficas o fluxo gênico estimado entre CVA e SFA é quase duas vezes menor do que entre SFA e NFA, porém semelhante entre CVA e NFA (Tabela 3). Outro fator que pode explicar a significância dessas análises é o fato de que a única comparação entre populações que possui significância estatística é para aquelas que possuem maior número de amostras (Tabela 6), sugerindo que a ausência de significância entre as demais comparações possa ser explicada pelo menor número de amostras na população NFA.

Zink e Barrowclough (2008) chamaram a atenção para o fato de que, devido à natureza estocástica do processo coalescente, as populações menores tendem a se tornar monofiléticas mais rapidamente que populações maiores. De fato, a rede de haplótipos de Cabanne e colaboradores (2011) baseada em 585 pb da região controladora apresenta vários alelos exclusivos para a população SFA, além de esta população possuir menos alelos compartilhados com as outras populações. No entanto, apesar de o citocromo b possuir o

mesmo tamanho populacional efetivo que a região controladora, e estes marcadores estarem ligados, este padrão não se reflete na rede de haplótipos do citocromo b, o que poderia ser explicado pela diferença nas taxas de mutação destes marcadores.

O padrão de divergência observado nos marcadores nucleares também é distinto daquele observado na região controladora mitocondrial, não apresentando nenhum sinal de estrutura genética populacional nas redes de haplótipos. Considerando que o tamanho efetivo dos marcadores nucleares é 4 vezes maior que o tamanho efetivo dos marcadores mitocondriais (Avice, 2000; Moore, 1995; Ballard & Whitlock, 2004) e, portanto, a evolução da monofilia será em média 4 vezes mais lenta nestes marcadores (Palumbi et al., 2001; Ballard & Whitlock, 2004), este resultado não é inesperado, já que mesmo nos marcadores mitocondriais nenhuma população é monofilética. Entretanto, a divergência foi significativa na análise de variância molecular entre regiões dentro dos biomas tanto para o FIB5 quanto para o Myo2. Embora este resultado pareça ser inesperado, Cabanne e colaboradores (2011) já haviam encontrado maior divergência dentro dos biomas (SFA vs NFA, $Phi_{ST} = 0,752$) do que entre eles (FA vs CVA, $Phi_{ST} = 0,483$), sugerindo que, embora o padrão observado na rede de haplótipos seja distinto, ele é congruente com a relação de divergência dentro e entre os biomas encontrada por Cabanne e colaboradores (2011).

5.2 Demografia histórica

Embora em algumas comparações (F_{ST} , Tabela 8) ou níveis (Phi_{ST} , Tabela 9) existem evidências de estruturação, na maioria delas os testes indicam que as populações não são estruturadas. Isso sugere que os testes de expansão, os quais foram desenvolvidos para populações panmíticas (Tajima, 1989; Heled & Drummond, 2008), deveriam ser aplicados na espécie como um todo. No entanto, Cabanne e colaboradores (2011), utilizando dados da região controladora, evidenciaram a existência de estrutura em todos os níveis das análises de variância molecular, indicando que as populações são estruturadas. Sendo assim, apesar da violação de premissa dos métodos – possível existência de fluxo gênico com outras populações –, a estratégia de avaliar expansão populacional em análises separadas por população parece ser razoável.

Resultados de modelagem de distribuição da Floresta Atlântica no último máximo glacial sugerem que houve alterações demográficas no sul da floresta atlântica (Carnaval & Moritz, 2008). Isso foi também detectado para algumas espécies de sapos (Carnaval et al.,

2009) e aves (Cabanne et al., 2007), inclusive a espécie analisada no presente estudo, *D. platyrostris*, segundo dados da região controladora (Cabanne et al., 2011). Tais resultados contrastam com os testes de expansão para todos os marcadores utilizados no presente trabalho, cujos testes não obtiveram significância estatística (Tabela 6), portanto não refutaram a hipótese de tamanho populacional constante para todas as populações. No entanto, Ramos-Onsins e Rozas (2002) mostraram que o teste D de Tajima pode ser menos sensível que alguns outros testes para detectar expansão quando o número de amostras é pequeno, como é o caso do presente trabalho. Além disso, os gráficos do *extended Bayesian skyline plot* sugerem expansão populacional no SFA e no CVA, embora o intervalo de confiança seja amplo demais para descartar que o tamanho populacional foi constante ao longo do tempo. Assim, existem evidências que sugerem que a dinâmica demográfica das populações de *D. platyrostris* segue o modelo proposto por Carnaval e Moritz (2008), tais como: 1) o gráfico do *extended Bayesian skyline plot* sugere expansão na população do SFA (Fig. 7B); 2) detectou-se sinal de expansão demográfica na região controladora para a população do SFA (Cabanne et al., 2011); e 3) o gráfico do *extended Bayesian skyline plot* para a população do NFA (Fig. 7A) sugere estabilidade.

5.3 Isolamento com migração e filogeografia estatística

Algumas premissas da análise de isolamento com migração, tais como presença de estrutura populacional [(NFA, (SFA,CVA))] e fluxo gênico com populações não amostradas (no caso das comparações que deixam uma população fora da análise), foram violadas no presente estudo. No entanto, dados simulados (Strasburg & Rieseberg, 2010) mostraram que essas violações prejudicam principalmente as estimativas de tempo de divergência, o que poderia explicar algumas estimativas de tempo de divergência resultantes de algumas de nossas análises (por exemplo, aquelas que não alcançaram probabilidade zero), e de tamanho populacional efetivo. Porém, simulações indicam que as estimativas de taxas de migração permanecem confiáveis (Strasburg & Rieseberg, 2010). Como no presente trabalho o tamanho populacional efetivo ancestral foi assumido, e nossos intervalos de distribuição de tamanho populacional descendente nas simulações são relativamente amplos, é possível que essas violações não gerem problema para nossas análises. Além disso, os valores das estimativas de tempo de divergência e taxa de migração foram simulados como *prior* uniforme nas simulações, adicionado de 10% de erro para cima e

para baixo do valor mais provável de forma a aumentar a chance de esse intervalo de valores se aproximar do possível valor real. As análises de isolamento com migração indicam altas probabilidades para taxas de migração de muito baixa (por exemplo, análise 6, Tabela 3) a alta (por exemplo, análise 3, Tabela 3), indicando que existe ou existiu fluxo gênico durante a divergência das populações. No entanto, a análise não distingue entre fluxo gênico durante todo o processo de divergência ou fluxo gênico com contato secundário (Becquet & Przeworski, 2009), o que nos impede de contrastar a hipótese de parapatria com a hipótese de alopatria. Entretanto, Cabanne e colaboradores (2011) compararam a probabilidade posterior das hipóteses de alopatria com a hipótese com fluxo gênico durante a divergência em um modelo de isolamento com migração (Hey & Nielsen, 2004) utilizando de *Akaike information criterion* (AIC) em todas as comparações entre as populações, nos quais todas as comparações os modelos com migração obtiveram valor de AIC mais baixo, indicando que estes modelos são mais prováveis, evidenciando a presença de fluxo gênico durante a divergência destas.

Embora o valor da probabilidade posterior da análise de *ABC* pareça muito baixo (0,16), este resultado não é inesperado, já que: 1) o método de seleção da tolerância de distância das estatísticas observadas em relação àquelas estimadas para as simulações de cada modelo diminui a tolerância e aumenta a porcentagem de simulações rejeitadas; 2) o presente trabalho utilizou muitas variáveis (parâmetros) na análise da estimativa de probabilidade posterior, o que pode também ter influenciado no valor da probabilidade posterior. Essas análises sugerem que as populações se separaram simultaneamente (politomia) entre 92 e 113 mil anos, e apresentavam fluxo gênico relativamente baixo (hipótese 4.cE, Tabelas 4, 5, 10, 11). Esse resultado contrasta com os resultados simulados por Cabanne e colaboradores (2011) que haviam rejeitado a presença de politomia. Além disso, os modelos não rejeitados por Cabanne e colaboradores (2011) foram aquele que apresentam modelos de ilhas, e nenhum tempo de divergência fixado. Estes modelos possuem probabilidade posterior extremamente baixa no presente trabalho, e são muito inferiores na comparação entre os modelos (Tabela 10 e 11). Pode-se argumentar que os resultados encontrados por Cabanne e colaboradores (2011) não são totalmente comparáveis com os resultados do presente trabalho, pois: 1) utilizou-se um método de seleção de hipóteses diferente do utilizado no presente trabalho; 2) o presente trabalho não comparou diretamente as hipóteses demográficas aqui apresentadas com as hipóteses testadas por

aqueles; 3) as estatísticas sumárias utilizadas nos trabalhos, assim como o número das mesmas é diferente. Contudo, acreditamos que tudo isso não impede a comparação entre eles, pois: 1) o método de seleção de modelos utilizado no presente trabalho permite selecionar a hipótese que mais se aproxima dos dados, além de compará-las uma a uma, enquanto o método utilizado por Cabanne e colaboradores (2011) apenas permite rejeitar hipóteses; 2) as hipóteses utilizados no trabalho de Cabanne e colaboradores (2011) utilizaram taxas de migração escolhidas arbitrariamente, sendo possível, portanto, que suas hipóteses não sejam realistas. Além disso, todas as relações entre as populações testadas por Cabanne e colaboradores (2011) estão sendo testadas no presente trabalho, com a ressalva de que os parâmetros de divergência e taxa de migração foram estimadas diretamente dos dados; 3) a escolha das estatísticas sumárias utilizada nos dados do presente trabalho foram escolhidas de forma a maximizar a quantidade de informação extraída dos mesmos (Hickerson et al., 2006), e o método de seleção de modelos permite a utilização de várias dimensões nos dados (Blum & François, 2010), de forma que tanto as estatísticas sumárias, quanto o método de seleção de modelos sejam eficazes.

Embora o tempo de divergência estimado no presente estudo pareça ser muito recente, ele explicaria a ausência de estrutura observada nos dados mitocondriais do CitB, e a razão desta ser detectada por um marcador de taxa de mutação rápida, a região controladora. A retenção de polimorfismo ancestral e o *lineage sorting* incompleto seriam também resultados esperados caso a divergência entre as populações fosse tão recente (Palumbi et al., 2001; Zink & Barrowclough, 2008), o que parece ser o caso encontrado nos marcadores avaliados do presente trabalho. Além disso, o tempo de divergência encontrado no presente trabalho é consistente com o último período interglacial (Hewitt, 2000), o que poderia explicar a causa da separação entre estas populações. Adicionalmente, o modelo sugere expansão recente para as populações após o último máximo glacial. Carnaval e Moritz (2008), utilizando de modelagem de nicho ecológico com dados de paleoclima, propuseram um modelo no qual, durante o último máximo glacial, as florestas se fragmentaram devido às alterações climáticas, e depois voltaram a expandir. Neste período as espécies florestais possivelmente tiveram suas populações reduzidas, as quais voltaram a expandir quando as florestas voltaram ao seu tamanho original. Este modelo é parcialmente compatível com a hipótese proposta pelo presente trabalho (4.cE) para as populações de Floresta Atlântica. No entanto, na hipótese testada a população do norte também sofreu

expansão, mas não houve gargalo anterior a esta, e nenhuma das demais hipóteses inclui ausência de expansão na população do norte da floresta atlântica e expansão ao sul, e/ou gargalo anterior à expansão. Assim, o resultado do presente trabalho é congruente com o modelo anteriormente proposto por Cabanne e colaboradores (2011), e consistente com o encontrado para outras espécies (Carnaval et al., 2009; Cabanne et al., 2007; d'Horta et al., 2011), além de ser congruente com a mudança climática ocorrida no período em que a expansão começou (Hewitt, 2000). Embora a expansão não tenha sido detectada nos testes aplicados aos marcadores do presente trabalho, os três modelos com probabilidade posterior mais alta incluem expansão em seus parâmetros (Tabelas 10 e 11), mostrando que esse resultado é robusto não apenas pela congruência com outras espécies e eventos climáticos, mas também por ser reincidente na seleção dos modelos.

5.4 Implicações biogeográficas

O modelo de dinâmica florestal proposto por Carnaval e Moritz (2008) inclui divergência em alopatria seguida de expansão populacional. Embora o presente trabalho não tenha avaliado estatisticamente a hipótese de alopatria, Cabanne e colaboradores (2011) rejeitaram essa hipótese para todas as comparações entre as populações. No entanto, o modelo mais provável estimado pelo presente trabalho sugere que as populações tenham expandido 20 – 30 mil anos atrás, sugerindo que estas populações foram afetadas pelas alterações climáticas do final do Quaternário e, portanto, mesmo populações de espécies florestais podem ter mantido o fluxo gênico quando as florestas se fragmentaram. Assim, a ausência de fluxo gênico pode não ser uma exigência para que populações sigam o modelo de Carnaval-Moritz.

Costa (2003) propôs que as florestas de galeria de vegetações abertas podem ter servido de corredor entre as florestas Amazônica e Atlântica. Como *D. playrostris* é uma espécie associada a matas de galeria (Marrantz et al., 2003), é possível que estas matas tenham servido de refúgio à estas populações, e que através destas as populações mantiveram o fluxo gênico mesmo durante as mudanças climáticas do último máximo glacial. Este cenário poderia explicar tanto os resultados do presente trabalho, quanto a rejeição da hipótese de alopatria por Cabanne e colaboradores (2011).

6. Conclusões

Embora nem todos os possíveis cenários demográficos puderam ser avaliados diretamente, os resultados apresentados no presente trabalho, juntamente com aqueles encontrados por Cabanne e colaboradores (2011) suportam um cenário no qual as populações de *D. platyrostris* tiveram origem e tamanho afetados pelas alterações climáticas ocorridos durante o final do período Quaternário. Estes resultados são coincidentes com modelos propostos tanto para a dinâmica demográfica da Floresta Atlântica, quanto para a conexão entre diferentes biomas, e com os resultados encontrados para outras espécies de aves. Trabalhos futuros utilizando maior número de marcadores moleculares permitindo melhores estimativas dos parâmetros demográficos, amostragem mais ampla e modelagem de nicho ecológico para fornecer hipóteses espacialmente explícitas poderiam complementar nossos resultados.

7. Resumo

Dendrocolaptes platyrostris é uma espécie de ave florestal associada às matas de galeria do corredor de vegetação aberta da América do sul (*D. p. intermedius*) e à Floresta Atlântica (*D. p. platyrostris*). Em um trabalho anterior, foi observada estrutura genética populacional associada às subespécies, além de dois clados dentro da Floresta Atlântica e evidências de expansão na população do sul, o que é compatível com o modelo Carnaval-Moritz. Utilizando *approximate Bayesian computation*, o presente trabalho avaliou a diversidade genética de dois marcadores nucleares e um marcador mitocondrial dessa espécie com o objetivo de comparar os resultados obtidos anteriormente com os obtidos utilizando uma estratégia multi-locus e considerando variação coalescente. Os resultados obtidos sugerem uma relação de politomia entre as populações que se separaram durante o último período interglacial, mas expandiram após o último máximo glacial. Este resultado é consistente com o modelo de Carnaval-Moritz, o qual sugere que as populações sofreram alterações demográficas devido às alterações climáticas ocorridas nestes períodos. Trabalhos futuros incluindo outros marcadores e modelos que incluam estabilidade em algumas populações e expansão em outras são necessários para avaliar o presente resultado.

8. Abstract

Dendrocolaptes platyrostris is a forest specialist bird associated to gallery forests of the open vegetation corridor of South America (*D. p. intermedius*) and to the Atlantic forest (*D. p. platyrostris*). A previous study showed a population genetic structure associated with the subspecies, two clades within the Atlantic forest, and evidence of population expansion in the south, which is compatible with Carnaval-Moritz's model. The present study evaluated the genetic diversity of two nuclear and one mitochondrial markers of this species using approximate Bayesian computation, in order to compare the results previously obtained with those based on a multi-locus strategy and considering the coalescent variation. The results suggest a polytomic relationship among the populations that split during the last interglacial period and expanded after the last glacial maximum. This result is consistent with the model of Carnaval-Moritz, which suggests that populations have undergone demographic changes due to climatic changes that occurred in these periods. Future studies including other markers and models that include stability in some populations and expansion in others are needed to evaluate the present result.

Referências Bibliográficas

- Aleixo, A. (2004). Historical diversification of a *terra-firme* forest bird superspecies: a phylogeographic perspective on the role of different hypotheses of Amazonian diversification. **Evolution**, 58:1303-1317.
- Anderson, CNK; Ramakrishnan, U; Chan, YL; Hadly, EA. (2005). Serial SimCoal: a population genetic model for data from multiple populations and points in time. **Bioinformatics**, 21:1733-1734.
- Avice, JC. (2000). **Phylogeography: the History and Formation of Species**. Harvard University Press. Cambridge, Massachusetts.
- Ballard, JW; Whitlock, MC. (2004). The incomplete natural history of mitochondria. **Molecular Ecology**, 13:729-744.
- Barton, NH; Hewitt, GM. (1985). Analysis of hybrid zones. **Annual Reviews of Ecology and Systematics**, 16:113-148.
- Beaumont, MA; Zhang, W; Balding, DJ. (2002). Approximate Bayesian computation in population genetics. **Genetics**, 162:2025-2035.
- Becquet, C; Przeworski, M. (2009). Learning about modes of speciation by computational approaches. **Evolution**, 63:2547-2562.
- Blum, MGB; François, O. (2010). Non-linear regression models for Approximate Bayesian Computation. **Statistics and Computing**, 20:63-73.
- Bruford, MW; Hannote, O; Brookfield, JFY; Burke, T. (1992). Single-locus and multilocus DNA fingerprinting. In: Hoelzel, AR (Ed). **Molecular Genetic Analysis of Populations – A Practical Approach**. IRL Press, New York, 287-336.
- Cabanne, GS; D'Horta, FM; Meyer, D; Silva, JMC; Miyaki, CY. (2011). Evolution of *Dendrocolaptes platyrostris* (Aves: Furnariidae) between the South American open vegetation corridor and the Atlantic Forest. *Biological Journal of the Linnean Society*, 103:801-820.
- Cabanne, GS; Santos, FR; Miyaki, CY. (2007). Phylogeography of *Xiphorhynchus fuscus* (Passeriformes, Dendrocolaptidae): vicariance and recent demographic expansion in southern Atlantic forest. **Biological Journal of the Linnean Society**:91, 73-84.
- Carling, MD; Brumfield, RT. (2007). Gene sampling strategies for multi-locus population estimates of genetic diversity (Theta). **Plos One**, 2: e160.

doi:10.1371/journal.pone.0000160

Carnaval, AC; Hickerson, MJ; Haddad, CFB; Rodrigues, MT; Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. **Science**, 323: 785-789.

Carnaval, AC; Moritz, C. (2008). Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic Forest. **Journal of Biogeography**, 35:1187-1201.

Coimbra-Filho, AF; Mittermeier, RA; Rylands, AB; Mendes, SL; Kierulff, MCM; Pinto, LPS. (2006). The taxonomic status of Wied's black-tufted-ear marmoset, *Callithrix kuhlii* (Callitrichidae, Primates). **Primate Conservation**, 21:1-24.

Costa, LP. (2003). The historical bridge between the Amazon and Atlantic forest of Brazil: a study of molecular phylogeography with small mammals. **Journal of Biogeography**, 30: 71-86.

Csilléry, K; Blum, MGB; Gaggiotti, OE; François, O. (2010). Approximate Bayesian Computation (ABC) in practice. **Trends in Ecology and Evolution**, 25:410-418.

Csilléry, K; François, OE; Blum, MGB. (2011). abc: an R package for Approximate Bayesian Computaiton (ABC). **ArXiv**:1106.2793.

D'Horta, FM; Cabanne, GS; Meyer, D; Miyaki, CY. (2010). The genetic effects of the late Quaternary climatic changes over a tropical latitudinal gradient: diversification of an Atlantic Forest passerine. **Molecular Ecology**, 20:1923-1935.

Dray, S; Dufour, AB. (2007). The ade4 package: implementing the duality diagram for ecologists. **Journal of Statistical Software**, 22:1-20.

Drummond, AJ; Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. **BMC Evolutionary Biology**, 7:214.

Excoffier, L; Novembre, J; Schneider, S. (2000). SIMCOAL: a general coalescent program for simulation of molecular data in interconnected populations with arbitrary demography. **Journal of Heredity**, 91:506-509.

Excoffier, L; Smouse, PE; Quattro, JM. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, 131:479-491.

Haffer, J. (1969). Speciation in Amazonian Forest birds. **Science**, 165:131-137.

Haffer, J. (1997). Alternative models of vertebrate speciation in Amazonian: an

overview. **Biodiversity Conservation**, 6:451-476.

Hasegawa, M; Kishino, H; Yano, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. **Journal of Molecular Evolution**, 22: 160-174.

Heled, J; Drummond, AJ. (2008). Bayesian inference of population size history from multiple loci. **BMC Evolutionary Biology**, 8:289.

Heslewood, MM; Elphinstone, MS; Tidemann, SC; Baverstock, PR. (1998). Myoglobin intron variation in the Gouldian finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoresis. **Electrophoresis**, 19:142-151.

Hewitt, GM. (1988). Hybrids zones-natural laboratories for evolutionary studies. **Trends in Ecology and Evolution**, 3:158-167.

Hewitt, GM. (2000). The genetic legacy of the Quaternary ice ages. **Nature**, 405: 907-913.

Hey, J. (2010). Isolation with migration models for more than two populations. **Molecular Biology and Evolution**, 27:905-920.

Hey, J; Machado, CA. (2003). The study of structured populations – new hope for a difficult and divided science. *Nature Reviews in Genetics*, 4:535-543.

Hey, J; Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. **Genetics**, 167:747-760.

Hickerson, MJ; Dolman, G; Moritz, C. (2006). Comparative phylogeographic summary statistics for testing simultaneous vicariance. **Molecular Ecology**, 15:209-223.

Holsinger, KE. Weir, BS. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F_{st} . **Nature Reviews in Genetics**, 10:639-650.

Hudson, RR; Kaplan, N. (1985). Statistical properties of the number of recombination events in the history of a sample of DNA sequences. **Genetics**, 111:147-164.

Huson, DH; Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. **Molecular Biology and Evolution**, 19:142-151.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. **Bioinformatics**, 24:1403-1405.

Knowles, LL. (2009). Statistical phylogeography. **Annual Reviews in Ecology, Evolution and Systematics**, 40:593,612.

Kuhner, MK. (2006). Lamarc 2.0: maximum likelihood and Bayesian estimation of

population parameters. **Bioinformatics**, 22, 768-770.

Librado, P; Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, 1451-1452.

Marantz, CA. (1997). Geographic variation os plumage patterns in the woodcreeper genus *Dendrocolaptes* (Dendrocolaptidae). **Ornithological Monographs**, 48:399-429.

Marantz, CA; Aleixo, A; Bevier, LR; Patten, MA. (2003). Family Dendrocolaptidae (Woodcreepes). In: Del Hoyo, J.; Elliot, A.; Christie, D. (Eds.), **Handbook of the Birds of the World**. Vol. 8: Broadbills to Tapaculos. Lynx Edicions, Barcelona: 358-399, 422.

Marini, MA.; Hackett, SJ. (2002). A multifaceted approach to the characterization of an intergeneric hybrid manakin (Pipridae) from Brazil. **Auk**, 119:1114-1120.

Marks, BD; Hackett, SJ; Capparella, AP. (2002). Historical relationships among Neotropical lowland forest areas of endemism as determined by mitochondrial DNA sequence variation within the wedge-billed woodcreeper (Aves: Dendrocolaptidae: *Glyphorynchus spirurus*). **Molecular Phylogenetics and Evolution**, 24:153-167.

Moore, WS. (1995). Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. **Evolution**: 718-726.

Martins, FM. (2011). Historical biogeography if the Brazilian Atlantic forest and the Carnaval-Moritz model of Pleistocene refugia: what do phylogeography studies tell us?. **Journal of the Linnean Society**, 104, 499-509.

Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary process that sustain it. **Systematic Biology**, 51:238-254.

Moritz, C; Hoskin, CJ; Mackenzie, JB; Phillips, BL; Tonione, M; Silva, N; VanDerWal, J; Williams, SE; Graham, CH. (2009). Identification and dynamics of a cryptic suture zone in tropical rainforest. **Proceedings of the Royal Society**, 276:1235-1244.

Moritz, C; Patton, JL; Schneider, CJ; Smith, TB. (2000). Diversification of rainforest faunas: an integrated molecular approach. **Annual Review of Ecology, Evolution and Systematics**, 31:533-563.

Myers, N; Mittermeier, RA; Mittermeier, CG; Fonseca, GAB; Kent, J. (2000). Biodiversity hotspots for conservation priotrities. **Nature**, 403:853-858.

Ney, M. (1973). Analysis of gene diversity in subdivided population. **Proceedings of National Academy of Sciences USA**, 70:3321-3323.

Nielsen, R; Wakeley, J. (2001). Distinguishing migration from isolation: a Markov

Chain Monte Carlo approach. **Genetics**, 158:885-896.

Nores, M. (1999). An alternative hypothesis for the origin of Amazonian bird diversity. **Journal of Biogeography**, 26:475-485.

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology*, 17:2103-2106.

Nosil, P; Harmon, Lj; Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. **Trends in Ecology and Evolution**, 24:145-156.

Paithankar, KR. Prasad, KS. (1991). Precipitation of DNA by polyethylene glycol and ethanol. **Nucleic Acids Research**, 19:1346.

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. **Bioinformatics**, 26:419-420.

Pebesma, EJ; Bivand, RS. (2005). Classes and methods for spatial data in R. **R news**, 5, <http://cran.r-project.org/doc/Rnews/>

Prado, CPA; Haddad, CFB; Zamudio, KR. (2011). Cryptic lineages and Pleistocene population expansion in a Brazilian Cerrado frog. **Molecular Ecology**, doi: 10.1111/j.1365-294X.2011.05409.x

R Development Core Team (2010). R: A language and environment for statistical computing. **R Foundation for Statistical Computing**. Vienna, Austria.

Ramos-Onsins, SE; Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. **Molecular Biology and Evolution**, 19: 2092-2100.

Slade, RW; Moritz, C; Heideman, A; Hale, PT. (1993). Rapid assessment of single-copy nuclear DNA variation in diverse species. **Molecular Ecology**, 2:359-373.

Smith, TB; Schneider, CJ; Holder, K. (2001). Refugial isolation versus ecological gradients. **Genetica**, 112-113:383-398.

Smith, TB; Wayne, RK; Girman, DJ; Bruford, MW. (1997). A role for ecotones in generating rainforest biodiversity. **Science**, 276:1855-1857.

Sobel, JM; Chen, GF; Watt, LR; Schemske, DW. (2009). The biology of speciation. **Evolution**, 64: 295-315.

Solomon, SE; Bacci-Jr, M; Martins-Jr, J; Vinha, GG; Mueller, UG. (2008). Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* sp.) provide new insight into the origins of Amazonian diversity. **Plos One**, 3:e2738.

Stephens, M; Smith, NJ; Donnelly, O. (2001). A new statistical method haplotype

reconstruction from population data. **American Journal of Human Genetics**, 68:978-989.

Strasburg, JL; Rieseberg, LH. (2010). How robust are “Isolation with migration” analyses to violations of the IM model? A simulation study. **Molecular Biology and Evolution**, 27, 297-310.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, 123:585-595.

Tamura, K; Peterson, D; Peterson, N; Stecher, G; Nei, M; Kumar. S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. **Molecular Biology and Evolution**, 28:2731-2739.

Thomé, MTC; Zamudio, KR; Giovanelli, JGR; Haddad, CFB; Baldissera-Jr, FA; Alexandrino, J. (2010). Phylogeography of the endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. **Molecular Phylogenetics and Evolution**, 55:1018-1031.

Thompson, JD; Higgins, DG; Gibson, TJ. (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**, 22:4673-4680.

Weir, JT; Schluter, T. (2008). Calibrating the avian molecular clock. **Molecular Ecology**, 17:2321-2328.

Werneck, FP. (2011). The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. **Quaternary Sciences Reviews**, 30:1630-1648.

Willis, EO; Oniki, Y. (2001). On a nest of the Planalto woodcreeper, *Dendrocolaptes platyrostris*, with taxonomic and conservation notes. **Wilson Bulletin**, 113: 231-233.

Zhang, DX; Hewitt, GM. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. **Molecular Ecology**, 12: 563-584.

Zink, RM; Barrowclough, GF. (2008). Mitochondrial DNA under siege in avian phylogeography. **Molecular Ecology**, 17: 2107-2121.