Universidade de São Paulo
Instituto de Biociências
Departamento de Genética e Biologia Evolutiva

Fernando Nodari

Filogenia do gênero *Paroaria* (Aves: Passeriformes: Oscines) e filogeografia de *Paroaria dominicana*.

São Paulo

2008

Fernando Nodari

Filogenia do gênero *Paroaria* (Aves: Passeriformes: Oscines) e filogeografia de *Paroaria dominicana*.

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

Orientadora: Cristina Yumi Miyaki

São Paulo

2008

Nodari, Fernando

Filogenia do gênero *Paroaria* (Aves: Passeriformes: Oscines) e filogeografia de *Paroaria dominicana*.

71 páginas.

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

1. *Paroaria* 2. Filogenia 3. Filogeografia 4. Caatinga 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).

Orientador(a)

Dedico à família,
aos amigos e
às *Paroaria*.

“No we’re never gonna survive
unless we are a little crazy”

Seal – Crazy

AGRADECIMENTOS

Não tenho palavras para expressar todo meu agradecimento à Dra. Cristina Y. Miyaki (Cris). Ela foi uma excelente orientadora; justa, comprometida, atenciosa, presente e rigorosa. Obrigado por me mostrar o que é ser um orientador.

Agradeço à FAPESP, à CNPq, à CAPES e ao meu pai pelo apoio financeiro.

Agradeço ao Instituto de Biociências da Universidade de São Paulo e ao Departamento de Genética e Biologia Evolutiva pela infra-estrutura.

Agradeço a todos do LGEMA (Érika, Tânia, Cibele, Fabio, Priscila, Adriana, Flavia, Gustavo, Fernando, Erwin, Camila, Rodrigo, José, Polyana, Beatriz e Patrícia) pela ajuda e pelos ensinamentos. Agradeço, em especial, ao Gustavo Cabanne pela imensa ajuda na coleta e ao Fabio Raposo pela imensa ajuda laboratorial e computacional.

Agradeço a Caio Graco e Fernando Mendonça D’Horta pela coleta de material indispensável a esse trabalho.

Agradeço a Luis Fabio Silveira e Diogo Meyer pela ajuda com a discussão dos dados.

Agradeço a Julio Rozas, Haipeng Li e Kai Zeng pela ajuda com a interpretação dos testes de neutralidade e a Kai Zeng pelo envio dos programas *ReadFasta* e *ReadFastaDHEW*.

Agradeço aos respectivos curadores e museus pela disponibilização dos dados de coleta das espécies de *Paroaria*: Paul Sweet (AMNH), Thomas Trombone (AMNH), Nate Rice (ANSP), Sievert Rohwer (UWBM), Ada Echevarria (Fundación Miguel Lillo), David Willard (FMNH), Steven Cardiff (LSU), Carlos Darrieu (Museo de La Plata), Elizabeth Ribas Bilo (UNICAMP), Marcos Raposo (MNRJ), Luis Fabio Silveira (MZUSP) e Alexandre Aleixo (MPEG).

Agradeços aos “caatingólogos” Ivan Prates e Helder Araújo pela ajuda com discussões e troca de artigos.

Agradeço a Amaro Luiz Alves pelas fotos disponibilizados e pelas dicas para coleta de *P. baeri*.

Agradeço à minha família (pai, mãe, irmão e cunhada) por todo o apoio e incentivo dado ao longo de todo meu mestrado.

Agradeço aos grandes e inseparáveis amigos de infância Rô, Flavinho, Marião, Carla e Jalbut.

Agradeço à Rena, afinal, até um pé na bunda te empurra pra frente!

Agradeço aos amigos da Biologia da UnB, UFRJ, UNESP/BTU, UNESP/RP, UNESP/RC, USP/SP, USP/RP e Unicamp... Sem vocês minha vida seria um tédio! Não vou citar nomes para não cometer injustiças. Além disso, quem é meu amigo sabe que é meu amigo. Porém, desses

amigos, um, com certeza, merece um agradecimento em especial: Flatú! Sem palavras para te agradecer mermão!

E foi desse grupo de amigos que surgiu a pessoa mais especial da minha vida. Tudo aquilo que eu sonhei sobre a pessoa que eu queria ao meu lado para o resto da vida, ela tem. Obrigado Robertinha, por ser a realização dos meus sonhos.

ÍNDICE

Resumo/Abstract	1
Capítulo 1 – Introdução geral	3
1.1 – Bibliografia	9
Capítulo 2 – Filogenia do gênero <i>Paroaria</i> (Aves: Passeriformes: Oscines)	13
Resumo/Abstract	13
2.1 – Introdução	15
2.2 – Materiais e Métodos	19
2.3 – Resultados	26
2.4 – Discussão	32
2.5 – Referências Bibliográficas	37
Capítulo 3 – Filogeografia de <i>Paroaria dominicana</i>	43
Resumo/Abstract	43
3.1 – Introdução	44
3.2 – Materiais e Métodos	47
3.3 – Resultados	53
3.4 – Discussão	61
3.5 – Referências Bibliográficas	65
Capítulo 4 – Considerações finais	70

RESUMO

A presente Dissertação apresenta dados sobre dois estudos complementares: a filogenia do gênero *Paroaria* (Aves: Passeriformes: Oscines; Capítulo 2) e a filogeografia de *Paroaria dominicana* (Capítulo 3). O estudo filogenético do gênero *Paroaria* utilizou como fonte de dados seqüências do ND2 (1041 pb) obtidas a partir de indivíduos das cinco espécies desse gênero (exceto *P. baeri* com 925 pb) e de seis espécies de diferentes tribos da subfamília Emberizinae como grupos externos. As reconstruções filogenéticas obtidas indicam que: 1) táxons da atual tribo Emberizini não formam um grupo monofilético; 2) *Gubernatrix cristata* pertence à tribo Thraupini; 3) dentre as espécies amostradas, *G. cristata* é a mais próxima ao gênero *Paroaria*; 4) o gênero *Paroaria* é monofilético; 5) o gênero *Paroaria* está dividido em dois clados principais: (*P. coronata*, *P. dominicana*) e (*P. baeri* (*P. gularis*, *P. capitata*)). Para auxiliar a interpretação dos resultados, as distribuições dos registros das espécies do gênero *Paroaria* foram obtidas por meio da compilação de informações de localidades de coleta de onze instituições. O clado (*P. coronata*, *P. dominicana*) sugere que as florestas abertas dos arredores do Chaco e da Caatinga estão relacionadas. Já o clado (*P. baeri* (*P. gularis*, *P. capitata*)) sugere relação entre as matas ciliares do Cerrado, da Amazônia, do Pantanal e dos arredores do Chaco. Estimativas de datas de divergências sugerem que as diversificações mais antigas dentro do gênero *Paroaria* ocorreram no período Plioceno, enquanto apenas a mais recente ocorreu no Pleistoceno inferior. Eventos paleogeográficos e paleoclimáticos poderiam estar associados às diversificações. Já o estudo filogeográfico da espécie *P. dominicana* utilizou como fonte de dados seqüências do ND3 (351 pb) e do primeiro domínio da região controladora (391 pb) obtidas a partir de 51 indivíduos dessa espécie. A espécie irmã *P. coronata* foi selecionada como grupo externo. Porém, devido a problemas laboratoriais ela foi substituída por *P. capitata* em algumas matrizes. Testes de neutralidade foram aplicados aos dados e sugerem que: 1) há fuga do modelo de populações em equilíbrio neutro causada por seleção positiva (fase pré ou pós-fixação) ou seleção negativa atuando concomitantemente ou não a uma expansão populacional recente; e 2) a não utilização do grupo externo mais adequado causou saturação nos dados, prejudicando a realização de inferências mais precisas sobre as forças causadoras do padrão de polimorfismos encontrado. Com isso, análises de demografia histórica e de tamanho populacional efetivo foram comprometidas. As análises de estrutura populacional indicam que: 1) a espécie *P. dominicana* não apresenta estruturação populacional para o DNAm; 2) essa espécie deve ser tratada como uma única Unidade de Manejo; e 3) o rio São Francisco e os tipos de Caatinga não constituem barreiras geográfica e ecológica para essa espécie, respectivamente.

ABSTRACT

This Master's Thesis comprises two complementary studies: phylogeny of the genus *Paroaria* (Aves: Passeriformes; Oscines; Chapter 2) and phylogeography of *Paroaria dominicana* (Chapter 3). The phylogeny of the genus *Paroaria* is based on complete ND2 (1041 bp) sequences from all five *Paroaria* species (except *P. baeri* with 925 bp) and from six species from different tribes of the subfamily Emberizini as outgroups. The results indicate that: 1) the current tribe Emberizini is not monophyletic; 2) *Gubernatrix cristata* should be placed within the tribe Thraupini; 3) among the outgroups sampled, *G. cristata* is the closest to genus *Paroaria*; 4) genus *Paroaria* is monophyletic; and 5) genus *Paroaria* is comprised of two major clades: (*P. dominicana*, *P. coronata*) and (*P. baeri* (*P. gularis*, *P. capitata*)). To help interpreting the results, the distribution of sampling localities of the genus *Paroaria* was compiled from eleven institutions. The clade (*P. dominicana*, *P. coronata*) indicates an association between the open forests of Caatinga and around the Chaco. The clade (*P. baeri* (*P. gularis*, *P. capitata*)) suggests that the riparian forests of Cerrado, Amazon, Pantanal and around the Chaco are related. Date estimates of diversification indicate that the oldest splits within the genus occurred during the Pliocene, and only the youngest diversification occurred during the early Pleistocene. Paleogeographic and paleoclimate events could be related to these diversifications. The phylogeography of the species *P. dominicana* is based on sequences from of ND3 (351 bp) and first domain of the control region (391 bp) from 51 individuals. The sister species *P. coronata* was used as outgroup. However, due to technical problems it was replaced by *P. capitata* in some data matrices. Neutrality tests results suggest that: 1) there is departure from the neutral equilibrium population model due to negative or positive selection (pre or post fixation phases) acting simultaneously or not with a recent population expansion; and 2) the use of a non-sister outgroup caused saturation of the data, jeopardizing more precise inferences on the forces shaping the polymorphism pattern found. Therefore, historical demography and effective population size analyses could not be performed. The population structure analyses suggest that: 1) the species *P. dominicana* is not geographically structured based on mtDNA data; 2) this species can be considered as a single Management Unit; and 3) the São Francisco river and the types of Caatinga are not geographic and ecological barriers for this species, respectively.

CAPÍTULO 1

Introdução Geral

A Sistemática é uma ciência que investiga e expõe as relações entre os seres vivos (Hennig, 1979). Ela existe desde que o Homem começou a tentar organizar seus conhecimentos sobre os seres vivos (Amorim, 2002). Até os anos 60, a Sistemática não tinha preocupações evolutivas e havia diferentes classificações dos mesmos objetos de estudo, pois estas se baseavam em diferentes caracteres (morfologia, ecologia, fisiologia etc.). Nos anos 60 ocorreram duas grandes revoluções na Sistemática: uma delas foi conceitual, quando Willi Hennig desenvolveu a base teórica e metodológica da Sistemática Filogenética (Hennig, 1979); a outra foi tecnológica, que permitiu o acesso a dados moleculares (citogenética, hibridação de DNA, sequenciamento de DNA etc.) e deu início à Sistemática Molecular (Miyamoto & Cracraft, 1991).

Uma vez que a Biologia Comparativa é a ciência que reúne informações de diversos organismos em busca de generalizações (padrões); que a Biologia Evolutiva é a ciência que descreve a variabilidade e seus padrões com base na interação de fatores endógenos e exógenos (Diniz Filho, 2000) e que a Sistemática Filogenética é o estudo da história evolutiva dos organismos por meio de comparações entre seus caracteres (Graur & Li, 1999); podemos dizer então que a Sistemática Filogenética é uma ciência situada entre os campos da Biologia Evolutiva e da Biologia Comparativa (Amorim, 2002; Diniz Filho, 2000; Graur & Li, 1999). Pois, devido à natureza do processo evolutivo, as diferenças entre espécies não são independentes; estão ligadas em algum ponto do tempo passado por um ancestral comum (Diniz Filho, 2000). A Sistemática Filogenética tem como objetivos descrever a diversidade biológica, encontrar que tipo de ordem há nela, apresentar um sistema geral de referência sobre ela e compreender os processos que são responsáveis por sua geração (Amorim, 2002).

Filogenias são hipóteses sobre as relações entre os organismos e, a partir delas, é possível formular hipóteses sobre a história das unidades biogeográficas às quais esses organismos estão associados. Com isso, a precisão e o refinamento das reconstruções biogeográficas são diretamente dependentes do rigor das reconstruções filogenéticas dos grupos estudados (Amorim, 2002) e a compreensão de ambas as reconstruções é crítica para tentar compreender os processos que causaram a diversificação biótica (Bates *et al.*, 1998).

Sob uma perspectiva evolutiva, uma hipótese biogeográfica representa um conjunto de hipóteses sobre a distribuição dos táxons e suas relações históricas (Morrone, 2006). Para a região

neotropical há várias hipóteses propostas, porém, apenas quatro serão discutidas devido à sua relevância ao presente estudo. Cabrera e Willink (1973), a partir de uma revisão de dados fito e zoogeográficos dividiram os neotrópicos em cinco domínios (Caribenho, Patago-andeano, Guiano, Chaquenho e Amazônico). Apenas dois domínios, onde os organismos estudados no presente estudo ocorrem, serão destacados: 1) Chaquenho, formado pelas províncias da Caatinga, do Chaco, do Chaco Serrano, dos Pampas, do Monte e da Prepunha; e 2) Amazônico, formado pelas províncias Amazônica, das Yungas, do Cerrado, Paranaense, Atlântica, Pacífica, Venezuelana, do Páramo e da Sabana (Figura 1.1).

Figura 1.1: Províncias biogeográficas de parte da América do Sul (figura modificada de Cabrera e Willink [1973]).

Morrone (2006), a partir de estudos da entomofauna, divide a região neotropical em cinco domínios (o autor utiliza o termo sub-regiões): Amazônico, Chaquenho, Paranaense, Caribenho e Áreas de Transição (Figura 1.2). Ao classificarem as províncias da Caatinga e do Chaco dentro do mesmo domínio (Chaquenho), Cabrera e Willink (1973) e Morrone (2006) propõem uma relação histórica entre elas, que também foi sugerida por outros autores baseados em estudos fitogeográficos (Andrade-Lima, 1981), de lagartos (Vanzolini, 1974) e de aves (Haffer, 1985).

Domínio	Províncias
Pacífico Norte Americano	Califórnia (1), Baixa Califórnia (2), Sonora (3), Platô Mexicano (4), Tamaulipas (5), Serra Madre Ocidental (6), Serra Madre Oriental (7), Cinturão Vulcânico Transmexicano (8), Bacia do Balsas (9) e Serra Madre do Sul (10)
Caribenho	Costa Pacífica Mexicana (11), Golfo do México (12), Península de Yucatan (13), Chiapas (14), Leste da América Central (15), Oeste do Istmo do Panamá (16), Bahamas (17), Cuba (18), Ilhas Cayman (19), Jamaica (20), Ilha de São Domingos (21), Porto Rico (22), Pequenas Antilhas (23), Chocó (24), Maracaibo (25), Costa da Venezuela (26), Trinidad & Tobago (27), Madalena (28), Lhanos Vanezuelanos (29), Cauca (30), Ilhas Galápagos (31), Oeste do Equador (32), Equador Árido (33) e Tumbes-Piura (34)
Amazônico	Napo (35), Imeri (36), Guiana (37), Guiana Úmida (38), Roraima (39), Amapá (40), Várzea (41), Ucayali (42), Madeira (43), Tapajós-Xingu (44), Pará (45), Pantanal (46) e Yungas (47)
Chaquenho	Caatinga (48), Cerrado (49), Chaco (50) e Pampas (51)
Paranaense	Mata Atlântica (53), Floresta do Paraná (54), Floresta de <i>Araucaria angustifolia</i> (55) e Páramo dos Andes do Norte (56)
Áreas de Transição	Deserto Costeiro do Peru (57), Puna (58), Atacama (59), Prepunha (60) e Montes (52)
Chile Central	Coquimbo (61) e Santiago (62)
Subantártico	Ilhas de Juan Fernandez (63), Maule (64), Floresta Valdivia (65), Floresta Magelânica (66), Páramo Magelânico (67) e Ilhas Malvinas (68)
Patagônico	Patagônia Central (69) e Patagônia Sub-andina (70)

Figura 1.2: Domínios e províncias biogeográficas da América do Sul. Números se referem às províncias listadas na tabela (Morrone, 2006). Algumas províncias da tabela (1 a 23 e 31) não são apresentadas na figura.

Porém, estudos fitogeográficos baseados em mapas pontuais de distribuição de táxons lenhosos rejeitaram as conexões entre Caatinga e Chaco (Prado, 1991; citado por Prado, 2000) e entre Caatinga e Cerrado (Prado & Gibbs, 1993) e uma outra proposta para a biogeografia da América do Sul foi sugerida (Prado, 2000). Nela, cria-se um novo domínio, Florestas Sazonais Tropicais, formado pelas províncias da Caatinga, Paranaense e da Floresta de Pedemonte Subandino (antigo distrito da província Yungas, ecótono entre Chaco e Yungas; Figura 1.3) e as demais províncias são mantidas nos domínios Chaquenho e Amazônico segundo Cabrera e Willink (1973). Áreas menores e mais isoladas de Florestas Sazonais Tropicais são encontradas nos vales secos dos Andes (Bolívia, Equador, Peru e Colômbia), nas costas do Equador e do norte do Peru, no centro-oeste do Cerrado (vegetação denominada “Mato Grosso de Goiás”) e espalhadas ao longo do Cerrado em regiões de solos mais férteis (Ratter *et al.*, 1997). Na hipótese de Prado (2000), o Pantanal não é atribuído a nenhum domínio, pois é considerado uma área de transição entre as províncias do Chaco, do Cerrado, Paranaense e Amazônica *sensu* Cabrera e Willink (1973), constituindo um mosaico dessas quatro formações (Prado *et al.*, 1992). Vale ressaltar que a ausência de relação entre Caatinga e Cerrado também foi encontrada em estudos da avifauna do Cerrado (Silva, 1995a). Assumindo que o Chaco e seus arredores realmente pertençam a domínios diferentes, pode-se supor que a presença da Caatinga e do Chaco no mesmo domínio, como sugerida por Cabrera e Willink (1973) e Morrone (2006), pode estar relacionada ao fato de os

arredores do Chaco serem considerados parte do Chaco. A distinção dessas duas províncias foi obtida pela análise de mapas pontuais de distribuição (Prado, 2000), mostrando a importância do conhecimento das distribuições reais dos táxons para estudos biogeográficos.

Figura 1.3: Mapa do domínio das Florestas Sazonais Tropicais. Hachurado oblíquo: província da Caatinga *sensu lato* (Prado, 1993; citado por Prado, 2000); vertical: província Paranaense *sensu* Prado (2000); horizontal: província da Floresta de Pedemonte Subandino *sensu* Prado (2000). ●: áreas com vegetações semelhantes às Florestas Sazonais Tropicais, mas que não podem ser seguramente atribuídas a nenhuma das três províncias (Prado, 2000).

Porzecanski e Cracraft (2005) realizaram uma análise cladística de distribuições e endemismos (CADE) da avifauna da América do Sul e encontraram um clado bem apoiado formado por Caatinga, Cerrado, Chaco e Mata Atlântica. No entanto, as relações entre essas áreas de endemismo não foram claras e só foram resolvidas ao excluir as aves do Cerrado dependentes de florestas (por exemplo, terras alagadas como o Pantanal, matas de galeria e florestas secas). Tal resultado era esperado, pois há espécies de aves da Mata Atlântica e da Floresta Amazônica que entram no Cerrado através das matas de galerias (Silva, 1995b), o que causa ruído nos dados. Com essa segunda análise, as áreas apresentaram a seguinte relação: (Mata Atlântica (Caatinga (Chaco, Cerrado))). Isso indica uma relação mais próxima entre Chaco e Cerrado do que entre qualquer uma dessas duas áreas e a Caatinga, corroborando estudos anteriores com aves do Cerrado (Silva, 1995a). Em ambas as análises de Porzecanski e Cracraft (2005), a área de endemismo dos Pampas formou um clado à parte com as áreas da Patagônia, do Sul dos Andes e do Pacífico Subtropical. No entanto, quando apenas os Emberizinae ou Oscines foram analisados, os Pampas e o Chaco agruparam com um apoio alto e as relações entre as outras áreas apresentaram apoio baixo. Essa incerteza quanto às relações dos Pampas ocorre devido ao fato de essa área ser de transição entre as províncias do Monte, do Chaco, Paranaense e Atlântica, formando um mosaico dessas quatro

formações (Porzecanski & Cracraft, 2005). Com isso, a avifauna de diferentes habitats nos Pampas pode ter relações biogeográficas distintas (Arballo & Cravino, 1999). Vale ressaltar que a área de endemismo definida como Chaco por Porzecanski e Cracraft (2005) coincide com a província do Chaco definida por Prado (2000).

Como discutido anteriormente, reconstruções filogenéticas e biogeográficas são críticas para tentar compreender os processos que causaram a diversificação biótica (Bates *et al.*, 1998). Estudos filogeográficos também auxiliam a compreensão desses processos, porém, em escala micro-evolutiva. A filogeografia é o estudo dos princípios e dos processos que governam a distribuição geográfica de linhagens intra-específicas e entre espécies filogeneticamente próximas (Avise, 2000). Apesar de haver estudos filogeográficos anteriores, o termo filogeografia só foi cunhado em 1987 por Avise *et al.* para resumir observações de que ramos de árvores intra-específicas frequentemente apresentam um padrão geográfico.

Estudos filogeográficos fornecem dados sobre processos de diferenciação e especiação, estrutura genética, época de colonização ou de separação, taxa de fluxo gênico, demografia histórica e tamanho populacional efetivo (Avise, 2000; Barrowclough *et al.*, 1999; Dietzen *et al.*, 2006; Garcia-Moreno *et al.*, 2004; Gay *et al.*, 2004; Milot *et al.*, 2000). Para realizar as análises de época de colonização ou de separação, taxa de fluxo gênico, demografia histórica e tamanho populacional efetivo, é necessário assumir que as mutações presentes no conjunto de dados analisados são seletivamente neutras ou quase neutras, ou seja, que nenhuma dessas mutações seja vantajosa ou deletéria a ponto de a ação da seleção natural ser significativa (Gillespie, 2004; Li, 1997). A teoria da coalescência (Kingman, 1982) permitiu que testes estatísticos fossem desenvolvidos para tentar verificar a neutralidade das mutações presentes em um conjunto de dados (Gillespie, 2004; Li, 1997); são os chamados testes de neutralidade (Fu, 1997; McDonal & Kreitman, 1991; Tajima, 1989; Zeng *et al.*, 2007b). Segundo essa teoria, utilizando uma amostra de uma população, é possível inferir o passado dela e identificar os eventos ocorridos desde a sua origem no ancestral comum mais recente da amostra (Fu & Li, 1999). Há duas principais categorias de metodologias relacionadas aos testes de neutralidade: 1) uma que analisa somente os padrões de polimorfismo, ou seja, que investiga a diversidade intra-específica; e 2) outra que considera dados de padrões de divergência e de polimorfismo (mutações inter e intra-específicas, respectivamente). Essas metodologias são brevemente descritas a seguir.

As análises isoladas dos padrões de polimorfismo podem ser sub-divididas em duas principais categorias: 1) testes de frequência dos sítios; e 2) testes de frequência de haplótipos. Os testes de frequência dos sítios se baseiam na frequência de mutações nos sítios polimórficos. Seus cálculos comparam duas diferentes estimativas do parâmetro central da teoria da evolução neutra (θ) obtidas utilizando classes diferentes de mutações (raras, abundantes e intermediárias; ver a

seguir); sendo $\theta = xN\mu$, onde x é igual a 2 para genomas haplóides e 4 para diplóides, N é o tamanho efetivo populacional e μ é a taxa de mutação por sequência por geração. Sob neutralidade, a diferença entre as duas estimativas de θ deve ser zero (Li, 1997). A frequência de uma mutação é determinada a partir do alinhamento de sequências homólogas de DNA. Porém, não se pode diferenciar as mutações de frequência baixa (mutações raras) das de alta (mutações abundantes) na ausência da sequência de um grupo externo. Por exemplo, na Figura 1.4, a mutação no sítio polimórfico 1 pode ser rara ou abundante caso o nucleotídeo ancestral seja A ou G, respectivamente. Caso a sequência 6 seja de um grupo externo, pode-se inferir que o nucleotídeo ancestral do sítio 1 é G e que, portanto, a mutação para o nucleotídeo A é rara. Seguindo o exemplo, as mutações nos sítios 2 e 4 possuem frequência intermediária (mutações intermediárias) e no sítio 3, alta (Fu, 1995).

Sequência					
1	--	A	--	G -- C -- T --	
2	--	G	--	G -- C -- T --	
3	--	G	--	C -- C -- A --	
4	--	G	--	C -- C -- A --	
5	--	G	--	C -- T -- A --	
6	--	G	--	C -- T -- T --	
<hr/>					
Sítio		1	2	3	4
polimórfico					

Figura 1.4: Alinhamento de seis sequências hipotéticas de DNA. Traços indicam sítios não polimórficos (modificado de Fu; 1995).

Resumidamente, diferentes tipos de seleção influenciam de maneira diferente nas proporções dessas três classes de mutações (raras, abundantes e intermediárias). A seleção positiva faz com que uma mutação vantajosa aumente rapidamente sua frequência numa população até atingir a fixação, ou seja, antes da fixação (fase pré-fixação) ela aumenta a proporção das mutações abundantes em relação às intermediárias e raras. Após a fixação (fase pós-fixação), novas mutações podem surgir, aumentando a proporção das mutações raras em relação às intermediárias e abundantes. A seleção negativa elimina mutações deletérias e, devido a mutações recorrentes, há um aumento na proporção das mutações raras em relação às mutações intermediárias. Já a seleção balanceadora mantém, num mesmo sítio, diferentes mutações em proporção equilibrada na população, ou seja, aumenta a proporção de mutações intermediárias em relação às mutações raras e abundantes (Fu, 1996, 1997; Li, 1997).

Como o teste de frequência dos sítios D_T (Tajima, 1989) não utiliza dados do grupo externo, não distingue as mutações raras das abundantes. É realizada a comparação dos valores de θ estimados a partir de mutações com frequência intermediária com aqueles estimados a partir de

mutações com frequência não intermediária (raras e/ou abundantes). Com isso, o teste detecta desvios de neutralidade da seguinte maneira: D_T é negativo caso haja excesso de mutações raras/abundantes e/ou déficit de mutações intermediárias (seleção positiva ou negativa) e é positivo caso haja déficit de mutações raras/abundantes e/ou excesso de mutações intermediárias (seleção balanceadora).

Os testes H (Fay & Wu, 2000) e E (Zeng *et al.*, 2006) consideram dados do grupo externo e, portanto, distinguem as mutações raras das abundantes. Com isso, o teste H compara valores de θ estimados a partir de mutações abundantes e mutações intermediárias, de modo que o valor de H é negativo caso haja excesso de mutações abundantes (fase pré-fixação da seleção positiva) e positivo caso haja excesso de mutações intermediárias (seleção balanceadora). Zeng *et al.* (2006) aumentaram o poder desse teste ao normalizá-lo. O teste E compara valores de θ estimados a partir de mutações abundantes e mutações raras, de modo que o valor de E é negativo caso haja excesso de mutações raras (seleção negativa e fase pós-fixação da seleção positiva) e positivo caso haja excesso de mutações abundantes (fase pré-fixação da seleção positiva).

Os testes de frequência de haplótipos consideram haplótipos como alelos. Com isso, se um gene apresenta uma proporção alta de haplótipos de frequência alta, ele é considerado homozigoto. Caso apresente uma proporção alta de haplótipos de frequência intermediária ou baixa, é considerado heterozigoto. Como exemplo dessa categoria de testes, pode-se citar o teste EW (Watterson, 1978), no qual a homozigosidade (proporção de haplótipos de frequência alta) esperada é calculada a partir do número total de haplótipos e, sob neutralidade, deve ser semelhante à homozigosidade observada. Dessa forma, a ação da seleção balanceadora, por exemplo, causaria um aumento da heterozigosidade (proporção de haplótipos de frequência intermediária) e a homozigosidade observada se torna menor do que a esperada.

Os testes F_s (Fu, 1997) e S (Strobeck, 1987) são intermediários às duas categorias acima (testes de frequência dos sítios e de frequência de haplótipos), pois se baseiam na comparação entre um valor de θ estimado a partir do número de haplótipos (θ_h) e um a partir da média de diferenças nucleotídicas entre duas seqüências (θ_π). Ambos não utilizam grupo externo, portanto, mutações raras não podem ser distinguidas das abundantes. Com isso, excesso de mutações raras/abundantes e déficit de mutações intermediárias (seleção positiva e negativa) fazem com que θ_π seja menor do que θ_h . Enquanto o déficit de mutações raras/abundantes e o excesso de mutações intermediárias (seleção balanceadora) fazem com que θ_π seja maior do que θ_h . Porém, F_s é sensível apenas ao primeiro cenário e S apenas ao segundo, ou seja, ambos são unilaterais. A insensibilidade de um teste a um desses cenários, aumenta o poder de detecção de desvios de neutralidade causados pelo outro cenário (Fu, 1996, 1997). O teste S foi modificado por Fu (1996) para aumentar seu poder. Vale ressaltar que o valor de F_s é negativo quando se analisa casos do primeiro cenário e que o teste

S é um cálculo de probabilidade e, com isso, seu valor varia entre 0 (dados não são neutros) e 1 (dados são neutros).

Dentre os testes que consideram dados de padrões de divergência e de polimorfismos (mutações inter e intra-específicas, respectivamente), quatro podem ser destacados: HKA (Hudson *et al.*, 1987), MK (McDonald & Kreitman, 1991), D_{FL} e F (Fu & Li, 1993). O teste HKA assume que mudanças populacionais afetam igualmente todo o genoma enquanto a seleção natural afeta apenas alguns locos, com isso, em condições de neutralidade, a razão entre as variabilidades inter e intra-específicas deve ser constante em vários locos independentes. Assim, o teste HKA deve ser utilizado com dados de locos independentes. O teste MK assume que, sob neutralidade, a razão entre divergências não-sinônimas e divergências sinônimas nos genes codificadores para proteína deve ser semelhante à razão entre polimorfismos não-sinônimos e polimorfismos sinônimos. Ou seja, o teste MK só é aplicável a dados de genes codificantes. Além disso, o teste MK é pouco conclusivo quanto às possíveis causas de desvio de neutralidade (Hughes, 2007). O teste D_{FL} se baseia na diferença entre um valor de θ estimado a partir do número total de divergências (θ_d) e um a partir do número total de polimorfismos (θ_p). O teste F se baseia na diferença entre um valor de θ estimado a partir do número total de divergências (θ_d) e um a partir da média de diferenças nucleotídicas entre duas seqüências do grupo interno (θ_π). Sendo θ_d uma estimativa das mutações antigas, θ_p e θ_π das mutações recentes, pode-se interpretar os efeitos dos diferentes tipos de seleção da seguinte maneira: 1) seleção positiva faz com que uma mutação vantajosa aumente rapidamente sua frequência numa população até atingir a fixação, aumentando a proporção de mutações recentes em relação às mutações antigas, ou seja, θ_p e θ_π são maiores que θ_d ; 2) seleção negativa elimina mutações deletérias e, devido a mutações recorrentes, há um aumento na proporção de mutações recentes em relação às mutações antigas, ou seja, θ_p e θ_π são maiores que θ_d ; e 3) seleção balanceadora mantém, num mesmo sítio, diferentes tipos de mutações em proporção equilibrada na população, diminuindo a proporção de mutações recentes em relação às mutações antigas, ou seja, θ_p e θ_π são menores que θ_d . Os valores de D_{FL} e F são negativos quando a proporção de mutações recentes (θ_p e θ_π) é maior e positivos quando é menor (Fu & Li, 1993).

Uma das maiores dificuldades na interpretação e no desenvolvimento dos testes de neutralidade é separar forças seletivas de eventos demográficos (efeito gargalo, migração, expansão e estruturação populacional), pois esses também causam mudanças nas proporções das três classes de mutações (raras, abundantes e intermediárias). O efeito gargalo pode causar diminuição na proporção de mutações raras em relação às demais, pois, como uma população atual é fruto de uma sub-amostra da população ancestral, a probabilidade de as mutações raras permanecerem na população atual é baixa. As altas taxas de migração podem causar aumento na proporção de mutações raras em relação às demais, pois os migrantes podem introduzir novos haplótipos na

população. A expansão demográfica recente também pode causar aumento na proporção de mutações raras em relação às demais, pois o aumento do número de indivíduos resulta em grande número de seqüências de DNA nas quais podem ocorrer mutações independentes. A estruturação populacional pode resultar no aumento da proporção de mutações intermediárias em relação às demais, pois quando as sub-populações de mesmo tamanho são uniformemente amostradas, as mutações diagnósticas de cada uma estarão igualmente representadas (Fu, 1996, 1997; Fu & Li, 1993; Nei & Kumar, 2000).

Portanto, os testes de neutralidade não testam apenas a neutralidade dos dados, mas também se os dados obedecem ao modelo de populações em equilíbrio neutro (mutações seletivamente neutras, tamanho populacional constante e ausência de estruturação populacional; Nielsen, 2001). Com isso, valores significativos dos índices D_T , H-normalizado, E, F_s , S modificado, D_{FL} e F podem ser causados por efeitos seletivos e/ou demográficos. Na tentativa de obter testes sensíveis a apenas um efeito (por exemplo, seleção positiva) foram desenvolvidos novos testes a partir da combinação de dois ou mais índices (Zeng *et al.*, 2006; Zeng *et al.*, 2007b). O teste DH (Zeng *et al.*, 2006) combina os testes D_T e H-normalizado e, com isso, a sensibilidade do D_T à expansão populacional é anulada pela insensibilidade do H-normalizado ao mesmo efeito, ou seja, o teste DH se torna mais específico a forças seletivas. Simulações com esse teste indicam que ele é sensível às seleções positiva (fases pré e pós-fixação) e balanceadora (Zeng *et al.*, 2006; Zeng *et al.*, 2007a; Zeng *et al.*, 2007b). O teste DHEW (Zeng *et al.*, 2007b) combina os testes DH e EW associando a alta sensibilidade do EW a desvios do modelo neutro com a alta especificidade do DH. Com isso, o teste DHEW se torna mais sensível que o DH às seleções positiva (fases pré e pós-fixação) e balanceadora. Esses testes compostos não apresentam um valor de índice, mas somente um valor de significância (P). Vale ressaltar que todos os testes mencionados assumem o modelo de sítios infinitos (em uma seqüência suficientemente longa, as mutações devem ocorrer em diversos sítios sem reincidência; Kimura, 1969). A partir de tabelas apresentadas por Zeng *et al.* (2006, 2007b) e dos trabalhos de Fu (1996, 1997) e Zeng *et al.* (2007a), as sensibilidades de alguns dos testes discutidos foram resumidas na Tabela 1.1. Como ainda não foram desenvolvidos testes de neutralidade para detectar especificamente cada tipo de desvio do modelo de populações em equilíbrio neutro, é necessária a aplicação de vários testes para identificar os possíveis fatores atuando em um conjunto de dados (Fu, 1997; Hughes, 2007).

Tabela 1.1: Poder e sinal (negativo [<0] ou positivo [>0]) dos testes D_T (Tajima, 1989), H -normalizado, E (Zeng *et al.*, 2006), F_s (Fu, 1997), S modificado (Fu, 1996), D_{FL} , F (Fu & Li, 1993), DH (Zeng *et al.*, 2006) e $DHEW$ (Zeng *et al.*, 2007b) para detectar diferentes efeitos. ^a: recombinação ausente ou muito baixa; ^b: todas sub-populações amostradas uniformemente. ¹, ², ³, ⁴ e ⁵: poder dos testes analisados segundo as simulações de Fu (1996), Fu (1997), Zeng *et al.* (2006), Zeng *et al.* (2007a) e Zeng *et al.* (2007b), respectivamente. +++: poder do teste frequentemente acima de 50%; ++: poder do teste em geral entre 20% e 50%, +: poder do teste entre 10% e 20%, na maioria dos casos; -: poder do teste raramente acima de 10%; *: poder não analisado até o momento. Os valores do poder dos testes são comparáveis somente entre os analisados em um mesmo estudo.

Efeito \ Teste	D_T ^{3,4,5}	H ^{3,4,5} normalizado	E ³	F_s ²	S ¹ modificado	D_{FL} ^{1,2}	F ^{1,2}	DH ^{3,4,5}	$DHEW$ ⁵
Seleção positiva (pré-fixação) ^a	++ (<0)	+++ (<0)	* (>0)	+++	-	++ (<0)	++ (<0)	+++	+++
Seleção positiva (pós-fixação) ^a	++ (<0)	-	+++ (<0)	+++	-	++ (<0)	++ (<0)	+	+
Seleção balanceadora	++ (>0)	++ (>0)	*	-	+++	+++ (>0)	+++ (>0)	-	-
Seleção negativa	++ (<0)	-	++ (<0)	++	-	+++ (<0)	+++ (<0)	-	-
Expansão populacional	++ (<0)	-	+++ (<0)	+++	-	++ (<0)	++ (<0)	-	-
Efeito gargalo	+ (>0)	++ (>0)	* (>0)	-	++	++ (>0)	++ (>0)	-	-
Estruturação populacional ^b	++ (>0)	-	*	-	++	++ (>0)	++ (>0)	-	-

Atualmente, um dos marcadores moleculares mais utilizados nos dois tipos de estudos discutidos acima (filogenia e filogeografia) é o DNA mitocondrial (DNAMt) (Burns & Naoki, 2004; Dietzen *et al.*, 2006; Kvist *et al.*, 2001; Omland *et al.*, 1999; Yuri & Mindell, 2002; Zink *et al.*, 2005). Isso se deve ao fato de o DNAMt ser facilmente isolado e amostrado, ter uma estrutura relativamente simples (sem íntrons e transposons), apresentar uma transmissão genética direta (maternalmente, sem recombinação ou outros rearranjos) e ser informativo para análises macro e micro-evolutivas (Avise *et al.*, 1987). Além disso, sua estrutura é relativamente bem conhecida, sabe-se que o DNAMt animal apresenta 37 genes, sendo 13 regiões codificadoras de proteínas, 22 RNAt e 2 RNAr, e mais uma região não codificadora. Nas aves ele apresenta os mesmos genes que nos outros vertebrados, porém, possui ordem gênica única (Quinn, 1997).

Estudos filogenéticos e filogeográficos auxiliam no entendimento da evolução das diferenças moleculares, morfológicas, comportamentais e fisiológicas dos organismos (Mindell *et al.*, 1997), numa boa análise comparativa desses caracteres (Harvey & Pagel, 1991), nas inferências sobre os padrões de caracteres num contexto histórico (Miyamoto & Cracraft, 1991) e nas reconstruções biogeográficas (Amorim, 2002). A principal diferença entre esses dois tipos de estudos é que filogenias permitem realizar interpretações de processos macro-evolutivos, enquanto análises filogeográficas, de processos micro-evolutivos (Avise, 2000). Com isso, suas análises se tornam distintas e são tratadas, na presente Dissertação, em capítulos separados: Filogenia do

gênero *Paroaria* (Aves: Passeriformes: Oscines; Capítulo 2) e Filogeografia de *Paroaria dominicana* (Capítulo 3).

1.1. REFERÊNCIAS BIBLIOGRÁFICAS

- Amorim, D. S. 2002. *Fundamentos de Sistemática Filogenética*. Ed. Holos, Riberão Preto.
- Andrade-Lima, D. 1981. The Caatinga dominium. *Rev. Bras. Bot.* **4**: 149-153.
- Arballo, E.; Cravino, J. L. 1999. *Handbook of the Birds of Uruguay*. Vol. 1. Ed. Hemisferio Sur, Montevideo.
- Avise, J. C. 2000. *Phylogeography: the History and Formation of Species*. Harvard University Press, Londres.
- Avise, J. C.; Arnold, J.; Ball, R. M.; Bermingham, E.; Lamb, T.; Neigel, J. E.; Reeb, C. A.; Saunders, N. C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* **18**: 489-522.
- Barrowclough, G. F.; Gutiérrez, R. J.; Groth, J. G. 1999. Phylogeography of spotted owl (*Strix occidentalis*) populations based on mitochondrial DNA sequences: gene flow, genetic structure, and a novel biogeographic pattern. *Evolution* **53**: 919-931.
- Bates, J. M.; Hackett, S. J.; Cracraft, J. 1998. Area-relationships in the Neotropical lowlands: an hypothesis based on raw distribution of Passerine birds. *J. Biogeog.* **25**: 783-793.
- Burns, K. J.; Naoki, K. 2004. Molecular phylogenetics and biogeography of Neotropical tanagers in the genus *Tangara*. *Mol. Phylogenet. Evol.* **32**: 838-854.
- Cabrera, A. L.; Willink, A. 1973. *Biogeografía de América Latina*. Ed. Secretaria General de la Organización de los Estados Americanos, Washington.
- Dietzen, C.; Voigt, C.; Wink, M.; Manfred, G.; Leitner, S. 2006. Phylogeography of island canary (*Serinus canaria*) populations. *J. Ornithol.* **147**: 485-494.
- Diniz Filho, J. A. F. 2000. *Métodos Filogenéticos Comparativos*. Ed. Holos, Riberão Preto.
- Fay, J. C.; Wu, C.-I. 2000. Hitchhiking under positive Darwinian selection. *Genetics* **155**: 1405-1413.
- Fu, Y.-X. 1995. Statistical properties of segregating sites. *Theor. Popul. Biol.* **48**: 172-197.
- Fu, Y.-X. 1996. New statistical tests of neutrality for DNA samples from a population. *Genetics* **143**: 557-570.
- Fu, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915-925.
- Fu, Y.-X.; Li, W.-H. 1993. Statistical tests of neutrality of mutations. *Genetics* **133**: 693-709.

- Fu, Y.-X.; Li, W.-H. 1999. Coalescing into the 21st century: an overview and prospects of coalescent theory. *Theor. Popul. Biol.* **56**: 1-10.
- Garcia-Moreno, J.; Navarro-Siguenza, A. G.; Peterson, A. T.; Sánchez-González, L. A. 2004. Genetic variation coincides with geographic structure in the common bush-tanager (*Chlorospingus ophthalmicus*) complex from Mexico. *Mol. Phylogenet. Evol.* **33**: 186-196.
- Gay, L.; Du Rau, P. D.; Mondain-Monval, J. Y.; Crochet, P. A. 2004. Phylogeography of a game species: the red-crested pochard (*Netta rufina*) and consequences for its management. *Mol. Ecol.* **13**: 1035-1045.
- Gillespie, J. H. 2004. *Population Genetics: a Concise Guide*. The Johns Hopkins University Press, Baltimore e Londres.
- Graur, D.; Li, W. H. 1999. *Fundamentals of Molecular Evolution*. Sinauer Associates Press, Sunderland.
- Haffer, J. 1985. Avian zoogeography of the Neotropical lowland. *Ornith. Monog.* **36**: 113-146.
- Harvey, P. H.; Pagel, M. D. 1991. *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.
- Hennig, W. 1979. *Phylogenetic Systematics*. University of Illinois Press, Urbana.
- Hudson, R.; Kreitman, M.; Aguade, M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153-159.
- Hughes, A. L. 2007. Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level. *Heredity* **99**: 1-10.
- Kimura, M. 1969. The rate of molecular evolution considered from the standpoint of population genetics. *Proc. Natl. Acad. Sci. USA* **63**: 1181-1188.
- Kingman, J. F. C. 1982. The coalescent. *Stoch. Process. Appl.* **13**: 235-248.
- Kvist, L.; Martens, J.; Ahola, A.; Orell, M. 2001. Phylogeography of a Palearctic sedentary passerine, the willow tit (*Parus montanus*). *J. Evol. Biol.* **14**: 930-941.
- Li, H.-W. 1997. *Molecular Evolution*. Ed. Sinauer Associates, Sunderland.
- McDonald, J. H.; Kreitman, M. 1991. Adaptive evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652-654.
- Milot, E.; Gibbs, H. L.; Hobson, K. A. 2000. Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). *Mol. Evol.* **9**: 667-681.
- Mindell, D. P.; Sorenson, M. D.; Huddleston, C. J.; Miranda-Jr., H. C.; Knight, A.; Sawchuk, S. J.; Yuri, T. 1997. Phylogenetic relationships among and within selected avian orders based on mitochondrial DNA. In: Mindell, D. P. *Avian Molecular Evolution and Systematics*. Academic Press, San Diego. Pp. 214-247.

- Miyamoto, M. M.; Cracraft, J. 1991. Phylogenetic inference, DNA sequence analysis, and the future of molecular systematics. *In*: Miyamoto, M. M.; Cracraft, J. *Phylogenetic Analysis of DNA Sequences*. Oxford University Press, Oxford. Pp. 3-17.
- Morrone, J. J. 2006. Biogeographic areas and transition zones of Latin America and the Caribbean Islands based on panbiogeographic and cladistic analyses of entomofauna. *Annu. Rev. Entomol.* **51**: 467-494.
- Nei, M.; Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, Nova York.
- Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. *Heredity* **86**: 641-647.
- Omland, K. E.; Lanyon, S. M.; Fritz, S. J. 1999. A molecular phylogeny of the New World orioles (*Icterus*): the importance of dense taxon sampling. *Mol. Phylogenet. Evol.* **12**: 224-239.
- Porzecanski, A. L.; Cracraft, J. 2005. Cladistic analysis of distributions and endemism (CADE): using raw distributions of birds to unravel the biogeography of the South American aridlands. *J. Biogeogr.* **32**: 261-275.
- Prado, D. E. 2000. Seasonally dry forests of tropical South America: from forgotten ecosystems to a new phytogeographic unit. *Edinb. J. Bot.* **57**: 437-461.
- Prado, D. E.; Gibbs, P. E. 1993. Patterns of species distributions in the dry seasonal forests of South America. *Ann. Missouri Bot. Gard.* **80**: 902-927.
- Prado, D. E.; Gibbs, P. E.; Pott, A.; Pott, V. J. 1992. Chaco-Pantanal transition in Southern Mato Grosso (Brazil). *In*: Furley, P. A.; Proctor, J.; Ratter, J. A. (eds) *Nature and Dynamics of Forest-Savanna Boundaries*. Chapman & Hall Press, London. Pp. 451-470.
- Quinn, T. W. 1997. Molecular evolution of the mitochondrial genome. *In*: Mindell, D. P. *Avian Molecular Evolution and Systematics*. Academic Press, San Diego. Pp. 3-28.
- Ratter, J. A.; Ribeiro, J. F.; Bridgewater, S. 1997. The Brazilian Cerrado vegetation and threats to its biodiversity. *Ann. Bot.* **80**: 223-230.
- Silva, J. M. C. 1995a. Avian inventory of the Cerrado region, South America: implications for biological conservation. *Bird Conserv. Internatl.* **5**: 291-304.
- Silva, J. M. C. 1995b. Birds of the Cerrado region, South America. *Steenstrupia* **21**: 69-92.
- Strobeck, C. 1987. Average number of nucleotide differences in a sample from a single subpopulation: a test for population subdivision. *Genetics* **117**: 149-153.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.
- Vanzolini, P. E. 1974. Ecological and geographical distribution of lizards in Pernambuco, Northeastern Brazil (Sauria). *Pap. Avulsos Zool.* **28**: 61-70.

- Watterson, G. A. 1978. The homozygosity test of neutrality. *Genetics* **88**: 405-417.
- Yuri, T.; Mindell, D. P. 2002. Molecular phylogenetic analysis of Fringillidae, "New World nine-primaried oscines" (Aves: Passeriformes). *Mol. Phylogenet. Evol.* **23**: 229-243.
- Zeng, K.; Fu, Y.-X.; Shi, S.; Wu, C.-I. 2006. Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* **174**: 1431-1439.
- Zeng, K.; Mano, S.; Shi, S.; Wu, C.-I. 2007a. Comparisons of site- and haplotype-frequency methods for detecting positive selection. *Mol. Biol. Evol.* **24**: 1562-1574.
- Zeng, K.; Shi, S.; Wu, C.-I. 2007b. Compound tests for the detection of hitchhiking under positive selection. *Mol. Biol. Evol.* **24**: 1898-1908.
- Zink, R. M.; Rising, J. D.; Mockford, S.; Horn, A. G.; Wright, J. M.; Leonard, M.; Westberg, M. C. 2005. Mitochondrial DNA variation, species limits, and rapid evolution of plumage coloration and size in the savannah sparrow. *Condor* **107**: 21-28.

CAPÍTULO 2

Filogenia do gênero *Paroaria* (Aves: Passeriformes: Oscines)

RESUMO

A biogeografia da América do Sul é muito debatida, o que se reflete nas diferentes hipóteses biogeográficas propostas. O gênero *Paroaria* é exclusivo de áreas abertas da América do Sul *cis-andina* e é um modelo interessante para testar algumas das diferentes relações biogeográficas propostas por essas hipóteses, bem como para examinar algumas hipóteses de diversificação. Para isso, o ND2 (1041 pb) foi completamente seqüenciado para representantes das cinco espécies do gênero (exceto *P. baeri*) e para seis espécies de diferentes tribos da subfamília Emberizinae como grupos externos (*Basileuterus leucoblepharus* [tribo Parulini, utilizado para enraizar a topologia], *Arremon flavirostris* [tribo Emberizini], *Gubernatrix cristata* [tribo Emberizini], *Haplospiza unicolor* [tribo Thraupini], *Saltator similis* [tribo Cardinalini] e *Tachyphonus coronatus* [tribo Thraupini]). Reconstruções filogenéticas baseadas em máxima parcimônia, máxima verossimilhança, distância e análise bayesiana resultaram em árvores semelhantes. Os resultados indicam que: 1) táxons da atual tribo Emberizini não formam um grupo monofilético; 2) *G. cristata* pertence à tribo Thraupini; 3) dentre as espécies amostradas, *G. cristata* é a mais próxima ao gênero *Paroaria*; 4) o gênero *Paroaria* é monofilético; e 5) o gênero *Paroaria* está dividido em dois clados principais: (*P. coronata*, *P. dominicana*) e (*P. baeri* (*P. gularis*, *P. capitata*)). Para auxiliar a interpretação dos resultados, as distribuições dos registros das espécies do gênero *Paroaria* foram obtidas por meio da compilação de informações de localidades de coleta das seguintes instituições: American Museum of Natural History, Academy of Natural Sciences of Philadelphia, Burke Museum of Natural History and Culture, Colección Ornitológica Fundación Miguel Lillo, Field Museum of Natural History, Louisiana State University - Museum of Natural Science, Museo de La Plata, Museu de História Natural da UNICAMP, Museu Nacional do Rio de Janeiro (MNRJ), Museu de Zoologia da Universidade de São Paulo e Museu Paraense Emílio Goeldi. O clado (*P. coronata*, *P. dominicana*) sugere que as florestas abertas dos arredores do Chaco e da Caatinga estão relacionadas. Já o clado (*P. baeri* (*P. gularis*, *P. capitata*)) sugere relação entre as matas ciliares do Cerrado, da Amazônia, do Pantanal e dos arredores do Chaco. Datações das divergências sugerem que as diversificações mais antigas dentro do gênero *Paroaria* ocorreram no período

Plioceno, enquanto apenas a mais recente ocorreu no Pleistoceno inferior. Eventos paleogeográficos e paleoclimáticos poderiam estar associados às diversificações.

ABSTRACT

South American biogeography has always been debated and there are many biogeographical hypotheses proposed. The genus *Paroaria* occurs exclusively in open areas of the cis-Andean South America, and it is an interesting model to test some of the different biogeographic hypotheses, as well as to verify some diversification hypotheses. Here the complete ND2 (1041 bp) was sequenced for all five *Paroaria* species (except *P. baeri*) and for six species from different tribes of the subfamily Emberizinae as outgroups (*Basileuterus leucoblepharus* [tribe Parulini, used root the topology], *Arremon flavirostris* [tribe Emberizini], *Gubernatrix cristata* [tribe Emberizini], *Haplospiza unicolor* [tribe Thraupini], *Saltator similis* [tribe Cardinalini] and *Tachyphonus coronatus* [tribe Thraupini]). Similar trees were obtained based on maximum parsimony, maximum likelihood, distance and Bayesian analysis. The results indicate that: 1) the present tribe Emberizini is not a monophyletic group; 2) *G. cristata* should be placed within the tribe Thraupini; 3) among the outgroups sampled, *G. cristata* is the closest to genus *Paroaria*; 4) genus *Paroaria* is monophyletic; and 5) genus *Paroaria* is comprised of two major clades: (*P. dominicana*, *P. coronata*) and (*P. baeri* (*P. gularis*, *P. capitata*)). To help interpreting the results, the distribution of sampling localities of the genus *Paroaria* was compiled from the following institutions: American Museum of Natural History, Academy of National Sciences of Philadelphia, Burke Museum of Natural History and Culture, Colección Ornitológica Fundación Miguel Lillo, Field Museum of Natural History, Louisiana State University - Museum of Natural Science, Museo de La Plata, Museu de História Natural da UNICAMP, Museu Nacional do Rio de Janeiro (MNRJ), Museu de Zoologia da Universidade de São Paulo and Museu Paraense Emílio Goeldi. The clade (*P. dominicana*, *P. coronata*) indicates that the open forests of Caatinga and around the Chaco and are related. The clade (*P. baeri* (*P. gularis*, *P. capitata*)) suggests that the riparian forests of Cerrado, Amazon, Pantanal and around the Chaco are related. Date estimates of diversification indicate that the oldest splits within the genus occurred during the Pliocene, and only the youngest diversification occurred during the early Pleistocene. Paleogeographic and paleoclimate events could be linked to these diversifications.

2.1. INTRODUÇÃO

O gênero *Paroaria* é tradicionalmente classificado na tribo Emberizini (ordem Passeriformes, subordem Oscines, família Fringillidae, subfamília Emberizinae; Sibley & Monroe, 1990), mas já foi alocado na tribo Cardinalini (Sick, 2001). Porém, estudos baseados em análises filogenéticas utilizando seqüências de genes do DNA mitocondrial (DNAmt) questionam essas classificações e sugerem que o gênero *Paroaria* pertence à tribo Thraupini (Burns & Naoki, 2004; Klicka *et al.*, 2007; Yuri & Mindell, 2002). Esse gênero é tradicionalmente dividido em duas super-espécies (Figura 2.1): 1) *P. dominicana* (galo-da-campina) e *P. coronata* (cardeal-do-sul); e 2) *P. gularis* (cardeal-da-amazônia; com três subespécies: *P. g. nigrogenis* de Trinidad ao centro da Colômbia, *P. g. cervicalis* no Mato Grosso, leste e nordeste da Bolívia e *P. g. gularis* nas demais regiões atribuídas à espécie), *P. capitata* (cardeal-do-pantanal; com duas subespécies: *P. c. fuscipes* do sudoeste da Bolívia e *P. c. capitata* nas demais regiões atribuídas à espécie) e *P. baeri* (cardeal-de-goias; com duas subespécies: *P. b. baeri* a oeste de Goiás e nordeste do Mato Grosso e *P. b. xinguensis* no alto do rio Xingu e norte do Mato Grosso; Peters, 1970). Essa divisão é baseada em caracteres morfológicos (por exemplo, presença de uma “gravata” preta nas espécies do segundo grupo) e ecológicos (as espécies do primeiro grupo estão associadas a áreas abertas e secas e as do segundo grupo a matas ciliares; Sick, 2001). Aparentemente, as distribuições de *P. coronata* e *P. capitata* se sobrepõem, porém, elas ocupam habitats distintos, sendo *P. coronata* associada a áreas abertas e secas e *P. capitata*, a áreas alagadas e a rios (Sick, 2001).

Figura 2.1: Fotos e distribuições geográficas de *P. dominicana* (a), *P. coronata* (b), *P. gularis* (c), *P. capitata* (d) e *P. baeri* (e) (Alves, 2005; Ridgely & Tudor, 1989). Fotos gentilmente cedidas por Amaro Luiz Alves (a, b, d, e) e por Arthur Grosset (c).

As distribuições geográficas das espécies de *Paroaria* tornam esse gênero um modelo interessante para serem testadas em relação às hipóteses biogeográficas discutidas no Capítulo 1. Assumindo que uma hipótese biogeográfica representa um conjunto de hipóteses sobre a distribuição dos táxons e suas relações históricas (Morrone, 2006), cada uma das hipóteses biogeográficas discutidas anteriormente sugere relações distintas entre as espécies de *Paroaria*. Essas relações e aquelas sugeridas por caracteres morfológicos e ecológicos estão resumidas na Figura 2.2. Como *P. coronata* e *P. capitata* ocorrem em mais de um bioma, elas surgem mais de uma vez em algumas hipóteses biogeográficas.

Figura 2.2: a) Relações entre as espécies do gênero *Paroaria* sugeridas por caracteres morfológicos e ecológicos (Sick, 2001). Relações entre áreas baseadas em algumas hipóteses biogeográficas e, entre parênteses, as espécies do gênero *Paroaria* que ocorrem nas respectivas áreas – hipótese biogeográfica de: b) Cabrera e Willink (1973), c) Morrone (2006), d) Prado (2000) e e) Porzecanski e Cracraft (2005). As espécies do gênero *Paroaria* foram designadas aos biomas baseado nos mapas zonais (Figura 2.1) e pontuais (ver resultados; Figura 2.4). Em negrito estão as espécies cujo bioma correspondente foi identificado somente pelos mapas pontuais. Sublinhadas estão as espécies que não ocorreriam no bioma indicado baseado nos mapas pontuais.

As hipóteses biogeográficas auxiliam na compreensão dos processos de diversificação biótica e, para explicar esses processos, várias hipóteses de diversificação foram levantadas. Dentre elas, estão: Hipótese dos Refúgios (Haffer, 1997) e Hipótese Paleogeográfica (Emsley, 1965). Segundo a hipótese dos refúgios, devido aos ciclos climáticos do Cenozóico, as florestas úmidas se expandiriam em períodos quentes e úmidos e retraíam em refúgios durante os períodos mais frios e secos; com isso, se populações associadas a essas vegetações ficassem isoladas em diferentes refúgios por tempo suficiente, poderiam originar espécies distintas. Já as florestas mais secas e áreas abertas em geral, expandiriam em períodos mais frios e secos e retraíam em períodos quentes e úmidos (Haffer, 1997). A hipótese paleogeográfica prediz que o soerguimento e a depressão de certas áreas e mudanças do nível do mar durante o Cenozóico isolariam populações de uma mesma

espécie e, dado o devido período de tempo, poderia resultar em diversificação de espécies (Emsley, 1965). Segundo Haffer (1997), essas duas hipóteses não são excludentes e os ciclos climáticos e as mudanças geológicas em conjunto poderiam compor uma explicação da diversificação biótica neotropical. A análise filogenética das espécies do gênero *Paroaria* pode adicionar mais dados para auxiliar a compreender tal diversificação.

Atualmente, o DNAm^t é um dos marcadores moleculares mais utilizados para estudos filogenéticos (Burns & Naoki, 2004; Klicka *et al.*, 2000; 2007; Omland *et al.*, 1999; Yuri & Mindell, 2002). Além dos pontos positivos do DNAm^t citados anteriormente (Capítulo 1), os dados moleculares apresentam outras vantagens: a) evoluem a taxas relativamente regulares; b) são relativamente abundantes; c) a detecção de homologia é relativamente fácil; d) em geral, seqüências de DNA não sofrem tanta influência de fatores ambientais como caracteres morfológicos; e) a descrição dos caracteres do DNA é direta; f) em alguns casos podem ser usados para avaliar relações filogenéticas entre organismos filogeneticamente distantes (Graur & Li, 1999). Porém, os dados moleculares também apresentam desvantagens como: a) as topologias obtidas podem ser árvores dos marcadores utilizados e não dos organismos; b) algumas técnicas apresentam uma relação custo/benefício alta; c) há, ainda, um grande debate sobre a natureza neutra das variações moleculares e sobre sua taxa de evolução (Hillis *et al.*, 1996).

Uma outra vantagem dos dados moleculares é a possibilidade de estimar as datas de divergência entre os clados de uma filogenia, auxiliando as análises de biogeografia histórica (Crisci *et al.*, 2003). Isso pode ser feito, por exemplo, assumindo que existe um relógio molecular, que presume que as mutações em uma determinada seqüência de DNA se acumulam a taxas constantes ao longo do tempo. Essas taxas devem ser calibradas a partir de datas de registros fósseis ou de eventos geológicos. Porém, para aves, os pontos de calibração são raros, pois há poucos fósseis e há poucos eventos geológicos precisos o suficiente e associados a algum evento de diversificação (Garcia-Moreno, 2004; Lovette, 2004). Com isso, muitos estudos se baseiam em taxas universais, como a de 2% de divergência por milhão de ano (2%/M.a.) estimado inicialmente para o DNAm^t de mamíferos (Brown *et al.*, 1979) e posteriormente para o DNAm^t de aves (Shields & Wilson, 1987). Para casos onde não é possível fazer o cálculo de uma taxa de divergência, o mais indicado é utilizar taxas produzidas para grupos mais próximos filogeneticamente ao grupo estudado (Garcia-Moreno, 2004; Lovette, 2004). Vale ressaltar que não há aqui intenção em mostrar superioridade dos dados moleculares sobre outros tipos de dados, deseja-se apenas mostrar algumas de suas vantagens e desvantagens.

Os objetivos do presente trabalho são: a) propor uma hipótese filogenética e datas de divergência para o gênero *Paroaria*; b) testar, para o gênero *Paroaria*, as relações sugeridas por caracteres morfológicos e ecológicos e pelas interpretações das hipóteses biogeográficas propostas

para a América do Sul; c) propor hipóteses para a diversificação das espécies do gênero *Paroaria*; e d) verificar a aplicabilidade das hipóteses de diversificação paleogeográfica e em refúgios para esse gênero.

2.2. MATERIAIS E MÉTODOS

2.2.1. Distribuição Geográfica

Como os tradicionais mapas zonais de distribuição apresentam baixa definição (Figura 2.1), um mapa pontual da distribuição das espécies de *Paroaria* foi obtido pela compilação de registros de coleta das seguintes instituições: American Museum of Natural History (AMNH), Academy of Natural Sciences of Philadelphia (ANSP), Burke Museum of Natural History and Culture (UWBM), Colección Ornitológica Fundación Miguel Lillo, Field Museum of Natural History (FMNH), Louisiana State University - Museum of Natural Science (LSUMNS), Museo de La Plata, Museu de História Natural da UNICAMP, Museu Nacional do Rio de Janeiro (MNRJ), Museu de Zoologia da Universidade de São Paulo (MZUSP) e Museu Paraense Emílio Goeldi (MPEG). As amostras sem informações precisas do local de coleta foram excluídas. As coordenadas geográficas das amostras foram obtidas utilizando o Google Earth 4.2 (<http://earth.google.com>) e/ou os *Ornithological Gazetteer of Argentina* (Paynter, 1995), *Bolivia* (Paynter, 1992), *Brazil* (Paynter & Traylor, 1991), *Colombia* (Paynter, 1997), *Ecuador* (Paynter, 1993), *Guianas* (Stephens & Traylor, 1985), *Paraguay* (Paynter, 1989), *Peru* (Stephens & Traylor, 1983), *Uruguay* (Paynter, 1994) e *Venezuela* (Paynter, 1982). Essas coordenadas foram plotadas em um mapa da América do Sul utilizando o programa DIVA-GIS 5.4 (<http://www.diva-gis.org>).

2.2.2. Obtenção das Amostras

Dos nove táxons descritos para o gênero *Paroaria*, cinco foram obtidos para o presente estudo (Tabela 2.1). Eles são provenientes de coletas realizadas por colaboradores e de aves capturadas em redes de neblina de 7 ou 10 metros ou com espingarda de ar comprimido (CBC-Monte Negro Super calibre 4,5 mm) munida com chumbinhos Gamo-Hunter. Para aumentar o sucesso de captura foram realizados *playbacks* para atraí-las e as redes foram colocadas próximas às fontes de alimento. De cada ave coletada em campo foi extraído um pedaço de músculo de 0,125 cm³ ou 0,1 mL de sangue com seringa de insulina da maior veia da região cervical direita ou cortou-se a unha de um dos dedos até que começasse a sangrar. As amostras de *P. baeri* foram obtidas a

partir de peles de aves taxidermizadas do Museu de Zoologia da USP. Dessas peles, foram retirados pequenos pedaços dos coxins de uma das patas. Todas as amostras estão armazenadas em etanol absoluto e mantidas a -20 °C no Laboratório de Genética e Evolução Molecular de Aves do Instituto de Biociências da Universidade de São Paulo. Os espécimes coletados foram depositados no Museu de Zoologia da Universidade de São Paulo. Foram analisados pelo menos dois indivíduos de cada espécie, com exceção de *Gubernatrix cristata* e *Haplospiza unicolor*. Os indivíduos da mesma espécie são, de preferência, provenientes de localidades diferentes.

Tabela 2.1: Espécie, número de registro (¹ – Laboratório de Genética e Evolução Molecular de Aves do Instituto de Biociências da Universidade de São Paulo; ² – Museu de Zoologia da Universidade de São Paulo; ³ – Academy of Natural Sciences of Philadelphia), local de coleta e coletor das amostras.

Espécie	Registro	Localidade	Coletor
<i>Saltator similis</i>	LGEMA1298 ¹	Buri/SP; S 23°43', W 48°34'	R. Gabam-Lima e A. Mendez
	LGEMA1343 ¹	Wenceslau Braz/PR	R. Gabam-Lima e A. Mendez
<i>Haplospiza unicolor</i>	LGEMA1275	Pinhalao/PR; S 23°58', W 50°3'	R. Gabam-Lima
<i>Basileuterus leucoblepharus</i>	LGEMA1452	Morro Grande/SP; S 23°41.7', W 45°57.12'	A. Martensen
	LGEMA2196 ¹	Rancho Queimado/SC	G Cabanne e R Pessoa
<i>Tachyphonus coronatus</i>	LGEMA2157 ¹	Rancho Queimado/SC	G Cabanne e R Pessoa
	LGEMA2162 ¹	Rancho Queimado/SC	G Cabanne e R Pessoa
<i>Arremon flavirostris</i>	LGEMA1307 ¹	Pinhalao/PR; S 23°58', W 50°3'	R. Gabam-Lima
	LGEMA1686 ¹	Pontal de Paranapanema/SP	A. Uezo
<i>Gubernatrix cristata</i>	LGEMA2546 ¹	Cativeiro	L. F. Silveira
<i>Paroaria baeri baeri</i>	35325 ²	São Domingos/MT	E. Dente
	35330 ²	Porto Pimentel/MT	E. Dente
<i>Paroaria capitata capitata</i>	LGEMA2257 ¹	Corumbá/MS; S 19°36', W 26°59'	C. Graco
	LGEMA2258 ¹	Corumbá/MS; S 19°36', W 26°59'	C. Graco
<i>Paroaria gularis gularis</i>	LGEMA619 ¹	Cativeiro	L. A. Shinaoba S.O.B.
	LGEMA620 ¹	Cativeiro	L. A. Shinaoba S.O.B.
	ANSP3279 ³	Equador, Zancudo Cocha; S 0°32', W 75°30'	M. A. Carriker
<i>Paroaria coronata</i>	ANSP10244 ³	Uruguai, Rota 24 km 75	L. Joseph
	LGEMA2144 ¹	Apreensão no Município de Rio Grande/RS	G Cabanne e R Pessoa
	LGEMA2183 ¹	Estrada Arroio Grande/RS	G Cabanne e R Pessoa
<i>Paroaria dominicana</i>	LGEMA617 ¹	Cativeiro	L. A. Shinaoba S.O.B.
	LGEMA618 ¹	Cativeiro	L. A. Shinaoba S.O.B.
	LGEMA2238 ¹	PN Serra das Confusões/PI	L. F. Silveira
	LGEMA2583 ¹	Tanquinho/BA; S 12°29'44.7", W 41°10'34.4"	F. Nodari e G. Cabanne
	LGEMA2595 ¹	Beira Rio/BA; S 12°16'14.8", W 42°42'52.0"	F. Nodari e G. Cabanne
	LGEMA2603 ¹	Parnaguá/PI; S 10°18'10.4", W 44°39'37.3"	F. Nodari e G. Cabanne
	LGEMA2610 ¹	Simplicio Mendes/PI; S 07°49'51.4", W 41°53'15.0"	F. Nodari e G. Cabanne
	LGEMA2618 ¹	Trindade/PE; S 07°47'50.5", W 40°17'40.4"	F. Nodari e G. Cabanne
	LGEMA2626 ¹	Itiúba/BA; S 10°40'19.5", W 39°46'40.1"	F. Nodari e G. Cabanne

2.2.3. Grupo Externo

Em análises filogenéticas, grupos externos proximalmente relacionados ao grupo interno devem ser utilizados para evitar a saturação dos dados (Smith, 1994). Devido à indefinição do posicionamento do gênero *Paroaria* na tribo Cardinalini ou na Emberizini (Sick, 2001) e aos recentes indícios de esse gênero pertencer à tribo Thraupini (Burns & Naoki, 2004; Klicka *et al.*, 2007; Yuri & Mindell, 2002), foram adicionados como grupos externos na presente análise as seguintes espécies: *Basileuterus leucoblepharus* (tribo Parulini, utilizado para enraizar a topologia), *Arremon flavirostris* (tribo Emberizini), *Gubernatrix cristata* (tribo Emberizini), *Haplospiza*

unicolor (tribo Thraupini), *Saltator similis* (tribo Cardinalini) e *Tachyphonus coronatus* (tribo Thraupini).

2.2.4. Seleção do Marcador

Devido às características do DNAm citadas, selecionou-se o gene da sub-unidade 2 da NADH desidrogenase (ND2). Esse gene codifica uma proteína e se mostrou útil para estudos filogenéticos em vários níveis, inclusive ao nível intra-genérico (Burns & Naoki, 2004; Ericson *et al.*, 2002; Klicka *et al.*, 2000; Omland *et al.*, 1999; Ribas *et al.*, 2005). Durante a etapa de seleção de marcadores também foi sequenciado o gene citocromo b. Porém, em um indivíduo de *P. dominicana*, um de *P. coronata* e um de *P. gularis* foram encontradas seqüências ambíguas, o que indica a presença de uma outra cópia do gene. Com isso, não foi possível utilizar esse gene.

2.2.5. Seleção dos *Primers*

Para amplificar o citocromo b, foi utilizado o par de *primers* L14841 (modificado de Kocher *et al.*, 1989) e H16065 (Lougheed *et al.*, 2000). Para amplificar todo o ND2 (1041 pb) de amostras frescas foram utilizados os *primers* LMet (Ribas *et al.*, 2005) e H6313 (Sorenson *et al.*, 1999; Figura 2.3). Para amostras de indivíduos taxidermizados foram utilizados os seguintes pares de *primers*: LMet (Ribas *et al.*, 2005) e H5550 (5'-GGGTGGGTGAGTTGGGTGATGT-3'), L5454 (5'-TTGAAGCCGCCACTAAATA-3') e H5730 (5'-AGGGTGGGGTTTAGTGAGGGTGAGG-3'), L5602 (Ribas *et al.*, 2005) e H5766 (Sorenson *et al.*, 1999), ND2L5730 (5'-CCTCACCCCTCACTAAACCC-3') e H5951 (5'-GATAGTTTTAGTACTTTGG-3'), L5838 (5'-TCTCCTCTATCTCCACCT-3') e H6078 (5'-GCTGTTGGGGCTATGTCCT-3'), L6057 (5'-TTCCTCCCCTAACAGGATT-3') e H6313 (Sorenson *et al.*, 1999; Figura 2.3). Os *primers* H5550, L5454, H5730, ND2L5730, H5951, L5838, H6078 e L6057 foram desenvolvidos para o presente estudo baseados nas seqüências alinhadas de *Paroaria dominicana*, *P. coronata*, *P. capitata* e *P. gularis*. Amostras de indivíduos taxidermizados tiveram fragmentos menores amplificados porque, geralmente, seu DNA está degradado em pequenos pedaços de até 500 pb (Cooper & Poinar, 2000).

Figura 2.3: Região mitocondrial do gene ND2 com a localização dos *primers* utilizados: LMet (Ribas *et al.*, 2005), L5454 (presente estudo), L5602 (Ribas *et al.*, 2005), ND2L5730 (presente estudo), L5838 (presente estudo), L6057 (presente estudo), H6313 (Sorenson *et al.*, 1999), H6078 (presente estudo), H5951 (presente estudo), H5766 (Sorenson *et al.*, 1999), H5730 (presente estudo) e H5550 (presente estudo). L e H se referem às cadeias leve e pesada, respectivamente, e os números se referem à posição do último nucleotídeo na sequência da galinha (Desjardins & Morais, 1990).

2.2.6. Extração do DNA

Para a extração de DNA total das amostras frescas de aves, uma pequena quantidade de sangue ou músculo foi colocada em tubos de 1,5 mL com 300 µL de TNE (50 mM de Tris, 100 mM de NaCl e 5 mM de EDTA, pH 7,5), 30 µL de Tris 1 M (pH 7,5), 8 µL de SDS 25% e 20 µL de Proteinase K (25 ng/mL). Essa mistura foi incubada em estufa a 37 °C por 24 horas. Após a incubação, foi misturado um volume de fenol:clorofórmio:álcool isoamílico (25:24:1) e centrifugado por 10 minutos a 12000 rpm. A fase superior foi transferida para outro tubo, onde foram adicionadas duas partes de etanol 100%. Essa solução foi misturada até a precipitação do DNA. Esse DNA foi isolado, lavado e dissolvido seguindo os passos: centrifugação por 45 minutos a 12000 rpm, descarte do etanol, adição de etanol 70%, centrifugação por 10 minutos a 12000 rpm, descarte do etanol, adição de 100 µL de TE (10 mM de Tris e 1 mM de EDTA, pH 8) e armazenamento a 4 °C. A qualidade e a quantidade do DNA foram verificadas por eletroforese em gel de agarose 1%. Após a quantificação, esse DNA foi diluído para aproximadamente 20 ng/µL para posterior amplificação.

Para a extração de DNA das amostras de aves taxidermizadas, foi utilizado o *QIAamp DNA Mini Kit* (QIAGEN) seguindo o protocolo para tecidos indicado pelo fabricante com as seguintes modificações: adição de 25 µL de DTT (1,4-Ditiotreitol; 1 g/mL; BioAgency) a cada amostra para melhorar a digestão da amostra; e, ao invés de uma diluição final com 200 µL do tampão AE (QIAGEN), foram realizadas duas diluições finais distintas com 100 µL do tampão AE (QIAGEN) para obter uma maior concentração de DNA. Para evitar a contaminação das amostras taxidermizadas, todo esse procedimento foi realizado em um laboratório separado (Laboratório de

Ictiogenética do Instituto de Biociências/USP) utilizando-se materiais (pipetas, ponteiras, tubos e soluções) exclusivos para esse tipo de procedimento, como sugerido por Cooper e Poinar (2000). Como sugestão desses mesmos autores, a presença de contaminação nessas extrações foi testada por meio de uma extração controle negativa realizada simultaneamente.

2.2.7. Amplificação do DNA

A amplificação dos segmentos-alvo de amostras frescas de aves foi realizada em termociclador (TC2400, Perkin-Elmer) em volume total de 10 µL, contendo 1 µL de dNTPs (2 mM de cada), 1 µL de tampão (10x), 1 µL de *primer* L (10 µM), 1 µL de *primer* H (10 µM), 2 µL de DNA (~20 ng/µL), 0,1 µL de *Taq* polimerase (5 U/µ) e 3,9 µL de água milli-Q. Foram realizadas reações convencionais (mesma temperatura de hibridação em todos os ciclos; Saiki *et al.*, 1988) e reações *touchdown* (temperatura de hibridação diminui em alguns ciclos; Hecker & Roux, 1996). As reações convencionais foram inicialmente desnaturadas a 95 °C por 3 minutos, e os 40 ciclos consistiram de: 30 segundos a 95 °C (desnaturação), 30 segundos a 55 °C ou 58 °C (hibridação) e 30 segundos a 72 °C (extensão); e seguido de uma extensão final de 9 minutos a 72 °C. As reações *touchdown* variaram a temperatura de hibridação de 65 °C a 60 °C diminuindo 0,5 °C por ciclo. Devido a dificuldades no sequenciamento, algumas amostras foram reamplificadas com um dos *primers* externos (LMet [Ribas *et al.*, 2005] ou H6313 [Sorenson *et al.*, 1999]) e um dos internos (L5602 [Ribas *et al.*, 2005], ND2L5730 [presente estudo] ou H5766 [Sorenson *et al.*, 1999]). Nas reamplificações, os volumes e os ciclos foram mantidos sob as mesmas condições da amplificação com exceção da temperatura de hibridação que foi de 70 °C a 60 °C, diminuindo 1 °C por ciclo.

A amplificação do DNA de amostras de aves taxidermizadas foi realizada em termociclador (TC2400, Perkin-Elmer) em volume total de 25 µL, contendo 12,5 µL de *HotStarTaq Master Mix* (QIAGEN), 1 µL de *primer* L (10 µM), 1 µL de *primer* H (10 µM), 5 µL de DNA e 5,5 µL de água milli-Q. A reação foi inicialmente desnaturada a 95 °C por 15 minutos para ativar a *HotStarTaq* DNA Polimerase (QIAGEN), e os 50 ciclos consistiram de: 30 segundos a 95 °C (desnaturação), 30 segundos a 40 – 48 °C (hibridação) e 30 segundos a 72 °C (extensão); e seguido de uma extensão final de 10 minutos a 72 °C. A temperatura de hibridação dos *primers* variou de acordo com o conjunto de *primers* utilizado. Para o par de *primers* ND2L5730/H5951 foi necessário realizar uma reamplificação com os mesmos *primers* para se obter quantidade suficiente de produto amplificado. Os volumes e os ciclos utilizados para essas reamplificações foram os mesmos dos utilizados para a amplificação de amostras frescas de aves, porém a temperatura de hibridação foi de 55 °C e o tempo de extensão foi de 15 segundos. São necessários maiores cuidados para minimizar o risco de contaminação nessas amplificações, pois a temperatura de hibridação é baixa e a *HotStarTaq* DNA

Polimerase (QIAGEN) apresenta alta sensibilidade. Com isso, toda a preparação foi realizada em outro laboratório (Laboratório de Ictiogenética do Instituto de Biociências/USP) e foram utilizados materiais (pipetas, ponteiros, tubos e soluções) exclusivos para esse procedimento (exceto para a reamplificação), como sugerido por Cooper e Poinar (2000). Vale ressaltar que foram realizadas PCRs (Polimerase Chain Reaction) com os controles negativos da extração com todos os conjuntos de *primers* para testar a presença de contaminação (Cooper & Poinar, 2000).

O tamanho e a concentração dos produtos da amplificação foram verificados por eletroforese em gel de agarose 1%, utilizando 1/10 do volume total da amplificação de amostras frescas de aves, 1/5 do volume total da amplificação de amostras de aves taxidermizadas e padrões de concentrações e de tamanho conhecidos para comparação. No caso de contaminação no controle negativo, todas as reações foram descartadas.

O produto da amplificação foi purificado por digestão com exonuclease I (10 un/μL) e fosfatase alcalina de camarão (1 un/μL; Amersham/Pharmacia), adicionando-se 0,5 μL de cada enzima para cada 10 μL de produto da amplificação, incubando 1 hora a 37 °C seguido de 10 minutos a 80 °C. Alternativamente, a purificação foi realizada com a adição de um volume Polietilenoglicol (PEG; 1 g/mL) ao produto da amplificação, misturado, incubado a 37 °C por 15 minutos, centrifugado a 12000 rpm por 15 minutos, o sobrenadante foi descartado, adicionados 125 μL de etanol 80% gelado, centrifugado a 12000 rpm por 2 minutos, o sobrenadante foi descartado, os últimos três passos foram repetidos e foram adicionados 10 μL de água milli-Q.

2.2.8. Sequenciamento dos Produtos da Amplificação

Os produtos de amplificação foram utilizados como moldes para as reações de sequenciamento direto utilizando 200 a 400 ng desses produtos, 1,5 μL de *Big Dye Terminator Cycle Sequencing Kit* (Applied Biosystems) e 1 μL de *primer* (10 μM). O volume de produto de amplificação utilizado variou conforme seu tamanho e sua concentração. As reações foram colocadas no termociclador nas seguintes condições: 96 °C por 1 minuto, 30 ciclos de: 96 °C por 10 segundos, 50 °C por 10 segundos e 60 °C por 4 minutos. Os *primers* utilizados para o sequenciamento do citocromo b foram L14841 (modificado de Kocher *et al.*, 1989), L15298 (Cheng *et al.*, 1994), H15764 (Miyaki *et al.*, 1998), H16065 (Lougheed *et al.*, 2000). Para o sequenciamento do ND2 (Figura 2.3) das amostras frescas de aves foram LMet, L5602 (Ribas *et al.*, 2005), ND2L5730 (presente estudo), H5766 e H6313 (Sorenson *et al.*, 1999). Para as amostras de aves taxidermizadas; LMet (Ribas *et al.*, 2005), L5602 (Ribas *et al.*, 2005), H5766 (Sorenson *et al.*, 1999), H6313 (Sorenson *et al.*, 1999), H5550, ND2L5730, L5838, H5951, L6057 e H6078 (presente estudo).

Para a precipitação do produto dessa reação foram adicionados quatro volumes de isopropanol 75%. Essa solução foi centrifugada por 45 minutos a 12000 rpm, descartado o álcool, foram adicionados 200 µL de etanol 70% seguido de nova centrifugação por 20 minutos a 12000 rpm. Essas amostras foram aplicadas em gel de poliacrilamida 5% em seqüenciador automático ABI 377/36 (Applied Biosystems).

2.2.9. Análise dos Dados

Utilizando o programa CodonCode Aligner v.1.6.3 (CodonCode Corporation) as seqüências da cadeia pesada obtidas foram revertidas e complementadas para que ficassem idênticas às da cadeia leve. Utilizando-se o mesmo programa, essas seqüências foram alinhadas, com correção manual e as seqüências de DNA foram traduzidas em seqüências de aminoácidos para verificar a presença de códons de parada inesperados. O modelo de evolução molecular mais adequado para os dados foi selecionado utilizando-se o Critério de Informação de Akaike (AIC; Akaike, 1974) no programa Modeltest 3.7 (Posada & Crandall, 1998). Para avaliar se havia saturação das matrizes de seqüências, foram obtidos gráficos de saturação (Griffiths, 1997) utilizando-se o programa DAMBE (Xia & Xie, 2001).

A partir da matriz de seqüências alinhadas foi obtida a composição nucleotídica, testou-se a homogeneidade da composição de bases entre os táxons e obteve-se árvores filogenéticas baseadas nos métodos de máxima parcimônia e de distância (*Neighbor Joining*) pelo uso do programa PAUP* (versão 4.0 b10; Swofford, 2002), máxima verossimilhança utilizando o *PHYML online web server* (Guindon *et al.*, 2005) e por análise bayesiana no programa MrBayes 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). Análises de *bootstrap* (Felsenstein, 1985) com 1000 réplicas para as análises de máxima parcimônia e de distância e com 500 réplicas para as análises de máxima verossimilhança foram realizadas para estimar o apoio de cada ramo (Page & Holmes, 1998).

A hipótese de existência de relógio molecular foi verificada pelo teste de razão de verossimilhança (*likelihood ratio test*, LRT), que se baseia na comparação por meio do teste de qui-quadrado (X^2) das verossimilhanças ($-\ln L$) da topologia obtida e da topologia obrigada a obedecer ao relógio molecular, ou seja, com taxas de mutação constantes (topologia linearizada). O número de graus de liberdade é definido pelo número de táxons menos dois. Caso as topologias apresentem probabilidades significativamente diferentes, a hipótese do relógio molecular não é aceita (Felsenstein, 1988). Caso a hipótese seja aceita, os tempos de divergência dos nós da árvore podem ser calculados pela fórmula $t = D/2r$ (Rutschmann, 2006), onde t é o tempo de divergência, D é a divergência entre duas seqüências e r é a taxa de divergência. Devido à ausência de registro fóssil

ou de eventos geológicos para calibrar a datação molecular das diversificações do gênero *Paroaria*, não foi possível estimar a taxa de divergência dos dados. Nesse caso, o mais indicado é utilizar taxas produzidas para grupos mais próximos filogeneticamente ao grupo estudado (Garcia-Moreno, 2004; Lovette, 2004). Porém, não há estimativas de taxas de evolução para o ND2 em Passeriformes. Packert *et al.* (2007) utilizaram eventos geológicos para estimar taxas de 1,5% e 1,9% de divergência por milhão de anos (M.a.) para o citocromo b na família Paridae (subordem Oscines). Esse grupo é irmão da família Fringillidae, a qual o gênero *Paroaria* pertence (Sibley & Monroe, 1990). Porém, o ND2 é mais variável que o citocromo b (22,3%; Klicka *et al.*, 2007; 34,4%; Klicka & Spellman, 2007; 12,3%; Pérez-Emán, 2005; 18,5%; Ribas *et al.*, 2005) e o uso da taxa de divergência estimada para o citocromo b não pode ser utilizada diretamente em dados do ND2. Assim, as taxas obtidas por Packert *et al.* (2007) foram corrigidas para podermos utilizá-las com os dados do presente trabalho. A partir de seqüências do *GenBank* de ND2 e de citocromo b de *Paroaria capitata* (EF529863.1 e EF529977.1, respectivamente; Klicka *et al.*, 2007) e de *P. dominicana* (EF529880.1 e EF529994.1, respectivamente; Klicka *et al.*, 2007) foi estimado que o ND2 é 11% mais variável que o citocromo b no gênero *Paroaria*. Com isso, as taxas de Packert *et al.* (2007) foram aumentadas em 11% e obteve-se as taxas de 1,665%/M.a. e 2,109%/M.a. que foram utilizadas no presente estudo. Os desvios padrões das datas de divergência foram estimados a partir de tamanhos de ramos obtidos por *bootstrap* (100 réplicas da matriz original). As réplicas foram criadas utilizando-se o programa SEQBOOT presente no pacote PHYLIP 3.6 (Felsenstein, 2004) e, utilizando-se o programa r8s 1.71 (Sanderson, 2003), os desvios padrões de cada nó foram calculados baseados nas 100 réplicas fixando-se como idades máximas aquelas estimadas para o nó mais externo.

2.3. RESULTADOS

O mapa pontual da distribuição das espécies de *Paroaria* totalizou 790 registros, sendo 17 de *P. baeri*, 96 de *P. capitata*, 138 de *P. dominicana*, 235 de *P. coronata* e 304 de *P. gularis* (Figura 2.4). Nesse mapa pode-se observar que *P. dominicana* e *P. baeri* foram coletadas apenas na Caatinga e no centro-oeste do Cerrado, respectivamente. *P. gularis* foi coletada quase que exclusivamente no domínio Amazônico *sensu* Cabrera e Willink (1973) e Morrone (2006), porém há uma ausência de indivíduos coletados numa região que vai do noroeste do Mato Grosso à Colômbia. As coletas de *P. capitata* são mais abundantes no Pantanal e mais escassas no oeste da província Paranaense *sensu* Prado (2000) e na província da Floresta de Pedemonte Subandino *sensu* Prado (2000) ou nas bordas leste e oeste do Chaco *sensu* Cabrera e Willink (1973). As coletas de *P.*

coronata são mais escassas no Pantanal e mais abundantes nos Pampas, no oeste da província Paranaense *sensu* Prado (2000) e na província da Floresta de Pedemonte Subandino *sensu* Prado (2000) ou nas bordas leste e oeste do Chaco *sensu* Cabrera e Willink (1973) e Morrone (2006). Há uma ausência de indivíduos de *P. capitata* e *P. coronata* coletados numa faixa ao centro-norte da Argentina correspondente ao Chaco *sensu* Prado (2000). Apesar de mapas pontuais de distribuição apresentarem um viés devido a possíveis concentrações de coletas ou ausência delas em determinadas regiões, é possível que esse problema tenha sido minimizado pelo grande número de indivíduos analisados e de instituições consultadas.

Figura 2.4: Mapa pontual de distribuição de *P. dominicana* (vermelho), *P. coronata* (azul), *P. capitata* (laranja), *P. baeri* (preto) e *P. gularis* (verde) a partir de informações de coleta das instituições AMNH, ANSP, UWB, Colección Ornitológica Fundación Miguel Lillo, FMNH, LSUMNS, Museo de La Plata, Museu de História Natural da UNICAMP, MNRJ, MZUSP e MPEG.

O ND2 foi completamente seqüenciado para todas as amostras exceto as de *P. baeri* (taxidermizadas), para as quais não houve sucesso na amplificação com o par de *primers* L5454/H5730. Com isso, foram obtidas duas matrizes de seqüências: A) com 1041 caracteres para 29 táxons, sendo que o trecho correspondente ao par de *primers* L5454/H5730 (116 caracteres) das amostras de *P. baeri* foi preenchido com “N” (base indeterminada); B) com 925 caracteres para 29 táxons, com o trecho correspondente ao mesmo par de *primers* de todas as amostras excluído. Não foram detectados códons de parada inesperados nem há lacunas no alinhamento em nenhuma das matrizes. Gráficos de saturação foram obtidos com distâncias genéticas corrigidas segundo o modelo de evolução de Tamura e Nei (1993), pois esse é, dentre os modelos presentes no programa

DAMBE (Xia & Xie, 2001), o que mais se aproxima daqueles selecionados pelo programa Modeltest 3.7 (Posada & Crandall, 1998). A partir da matriz A obteve-se um gráfico de saturação (Figura 2.5a) onde se verifica sinais de saturação das transições, pois as curvas de transição se estabilizam com o aumento da distância genética. O gráfico de saturação obtido a partir da matriz B apresenta o mesmo padrão do gráfico anterior e não é apresentado.

Em análises posteriores, somente *Gubernatrix cristata* foi utilizada como grupo externo. Com tal estratégia esperava-se diminuir a saturação observada e aumentar o apoio dos nós das árvores obtidas (Smith, 1994). Portanto, a partir de ambas as matrizes de seqüências acima descritas, obteve-se as duas matrizes de seqüências a seguir: C) com 1041 caracteres para 20 táxons; D) com 925 caracteres para 20 táxons. A partir da matriz C obteve-se um gráfico de saturação (Figura 2.5b) onde se confirma que a utilização de apenas *G. cristata* como grupo externo reduziu a saturação do conjunto de dados, pois há apenas um sinal fraco de saturação das transições na terceira posição do códon. O gráfico de saturação obtido a partir da matriz D apresenta o mesmo padrão do gráfico anterior e não é apresentado.

Figura 2.5: Gráficos de saturação referente a toda a seqüência do ND2 (Total) e somente para a 3ª posição do códon (3ª pos.) para a matriz A (a) e para a matriz C (b).

Para a matriz A, a composição nucleotídica encontrada foi: A = 28,6%, C = 36%, G = 10,8%, T = 24,6%. Essa composição de bases é homogênea entre os táxons ($X^2 = 23,6260$ [df = 84], $P = 1,00$). Dos 1041 sítios, 626 não apresentam variação, 49 não são informativos para parcimônia e 366 são informativos para parcimônia. Para esse conjunto de dados o modelo evolutivo selecionado foi o GTR + I + G, ou seja, modelo do Tempo Geral Reversível (Tavaré, 1986; citado por Posada & Crandall, 1998) incluindo-se a proporção dos sítios invariáveis ($I = 0,5358$) e a variação de taxa entre os sítios ($\alpha = 2,1613$). Para a matriz B, a composição nucleotídica encontrada foi: A = 28,8%, C = 35,9%, G = 10,8%, T = 24,5%. Essa composição de bases é homogênea entre os táxons ($X^2 = 20,3318$ [df = 84], $P = 1,00$). Dos 925 sítios, 549 não apresentam variação, 46 não são informativos

para parcimônia e 330 são informativos para parcimônia. O modelo evolutivo selecionado também foi o GTR + I + G com proporção dos sítios invariáveis ($I = 0,5265$) e variação de taxa entre os sítios ($\alpha = 2,2690$). Dos 116 sítios excluídos nessa matriz, 77 não apresentam variação, três não são informativos para parcimônia e 36 são informativos para parcimônia. Foram obtidas filogenias semelhantes baseadas em ambos os conjuntos de dados utilizando-se diferentes métodos de análises (máxima parcimônia, máxima verossimilhança, distância e método bayesiano) e somente as topologias obtidas por máxima verossimilhança são apresentadas (Figura 2.6a e 2.6b). Nessas árvores verifica-se que: 1) há indícios de que a tribo Emberizini não seja monofilética devido a *G. cristata* e *Arremon flavirostris* não formarem um clado monofilético; 2) *G. cristata* seria, dentre as espécies incluídas na análise, a mais próxima filogeneticamente do gênero *Paroaria*; 3) há evidências de que o gênero *Paroaria* seja monofilético; e 4) parece haver dois grupos dentro do gênero *Paroaria*, sendo um formado por *P. dominicana* e *P. coronata* e outro por *P. gularis*, *P. capitata* e *P. baeri*.

Para a matriz C, a composição nucleotídica encontrada foi: A = 28,2%, C = 36,2%, G = 11,1%, T = 24,5%. Essa composição de bases é homogênea entre os táxons ($X^2 = 11,5910$ [df = 57], $P = 1,00$). Dos 1041 sítios, 804 não apresentam variação, 67 não são informativos para parcimônia e 170 são informativos para parcimônia. Para esse conjunto de dados o modelo evolutivo selecionado foi o TIM + G, ou seja, modelo de Transição (Posada & Crandall, 1998) incluindo-se a variação de taxa entre os sítios ($\alpha = 0,1646$). Para a matriz D, a composição nucleotídica encontrada foi: A = 28,4%, C = 36%, G = 11,1%, T = 24,5%. Essa composição de bases é homogênea entre os táxons ($X^2 = 8,7536$ [df = 57], $P = 1,00$). Dos 925 sítios, 706 não apresentam variação, 64 não são informativos para parcimônia e 155 são informativos para parcimônia. O modelo evolutivo selecionado foi o TIM + G incluindo-se a variação de taxa entre os sítios ($\alpha = 0,1836$). Dos 116 sítios excluídos nessa matriz, 98 não apresentam variação, três não são informativos para parcimônia e 15 são informativos para parcimônia. Foram obtidas filogenias semelhantes baseadas em ambos os conjuntos de dados utilizando-se diferentes métodos de análises (máxima parcimônia, máxima verossimilhança, distância e método bayesiano) e somente as topologias obtidas por máxima verossimilhança são apresentadas (Figura 2.6c e 2.6d). Nas árvores obtidas verifica-se que: 1) há evidências de que o gênero *Paroaria* seja monofilético; 2) parece haver dois grupos dentro do gênero *Paroaria*, sendo um formado por *P. dominicana* e *P. coronata* e outro por *P. gularis*, *P. capitata* e *P. baeri*; e 3) a utilização de apenas *G. cristata* como grupo externo causou um pequeno aumento nos valores de apoio dos nós das árvores, confirmando o esperado de que a diminuição da saturação dos dados causaria um aumento nos valores de apoio da filogenia. A composição nucleotídica encontrada em todas as matrizes de seqüências é semelhante à encontrada em estudos com DNAm em Oscines (Burns & Naoki, 2004; Yuri & Mindell, 2002).

a)

b)

c)

d)

Figura 2.6: Árvores obtidas pelo método de máxima verossimilhança baseadas em 1041 caracteres (a e c) e 925 caracteres (b e d) do ND2. Os números nos nós correspondem respectivamente aos valores de *bootstrap* para análises de máxima verossimilhança, máxima parcimônia e distância e de probabilidades posteriores segundo a análise bayesiana. Valores abaixo de 70% não são apresentados. Valores sublinhados são iguais para todas as análises.

Para realizar a datação, utilizou-se como base a matriz C por apresentar o menor sinal de saturação (Figura 2.5) e o modelo evolutivo mais simples, uma vez que isso implica em uma menor variância (Posada & Crandall, 1998). A partir dela, apenas um indivíduo de cada espécie foi mantido em uma nova matriz, para a qual se selecionou o mesmo modelo evolutivo da matriz C. Através do teste de razão de verossimilhanças (Felsenstein, 1988), verificou-se que não há diferença significativa entre a topologia estimada ($-\ln L = 2662,44$) e aquela que segue o relógio molecular ($-\ln L = 2665,52$; $P > 0,1$). As datas de divergência de cada clado foram estimadas utilizando as taxas corrigidas de 1,665%/M.a. e 2,109%/M.a. e estão apresentadas na Figura 2.7 e na Tabela 2.2. Os desvios padrões de cada estimativa são apresentados na Tabela 2.2. A partir desses dados verifica-se que possivelmente a maioria das divergências dentro desse gênero ocorreu no Plioceno.

Figura 2.7: Cronograma obtido por máxima verossimilhança baseado em 1041 caracteres do ND2. Escala de tempo em milhões de anos (TS-Creator [<http://www.chronos.org>]). O número de cada nó é apresentado abaixo do mesmo. As barras cinza nos nós são proporcionais ao intervalo entre as idades de cada nó. Os números acima das barras cinza indicam as idades máxima e mínima de cada nó em milhões de anos.

Tabela 2.2: Datas das divergências (em milhões de anos, M.a.) estimadas com as taxas de 1,665%/M.a. e 2,109%/M.a. e seus desvios padrões (entre parênteses) para cada nó. Números dos nós se referem à Figura 2.7.

Nó	1,665%/M.a.	2,109%/M.a.
1	10,3	8,1
2	4,68 (4,02-5,34)	3,68 (3,16-4,20)
3	3,40 (2,85-3,95)	2,67 (2,24-3,10)
4	2,26 (1,86-2,66)	1,78 (1,46-2,10)
5	1,56 (1,21-1,91)	1,23 (0,96-1,50)

2.4. DISCUSSÃO

As relações filogenéticas aqui encontradas são mais uma evidência apoiando a necessidade de revisão da composição das tribos da subfamília Emberizinae (Figuras 2.6a e 2.6b; Burns, 1997; Burns & Naoki, 2004; Klicka *et al.*, 2007; Yuri & Mindell, 2002). Segundo a filogenia molecular mais recente dessa subfamília (Klicka *et al.*, 2007), os representantes dos gêneros *Saltator* (tradicionalmente classificado como Cardinalini), *Haplospiza*, *Tachyphonus* (tradicionalmente classificados como Thraupini) e *Paroaria* (tradicionalmente classificado como Emberizini) seriam atribuídos à tribo Thraupini e o gênero *Arremon* seria mantido na tribo Emberizini. No presente trabalho, o posicionamento do gênero *Gubernatrix* (tradicionalmente classificado como Emberizini) junto aos gêneros *Paroaria*, *Saltator*, *Haplospiza* e *Tachyphonus* sugere que ele também pertença à tribo Thraupini. Porém, a baixa amostragem de táxons da subfamília Emberizinae no presente trabalho impede a realização de outras inferências a respeito das relações entre suas tribos. Dentre as espécies do grupo externo aqui utilizadas, *G. cristata* aparenta ser a mais próxima filogeneticamente do gênero *Paroaria* e a utilização de apenas ela como grupo externo reduziu a saturação dos dados (Figura 2.5) e aumentou os valores de apoio dos nós das árvores filogenéticas (Figura 2.6), confirmando a importância da utilização de grupos externos filogeneticamente próximos ao grupo interno (Smith, 1994).

A evidência de monofilia do gênero *Paroaria* encontrada no presente trabalho (Figura 2.6) é esperada uma vez que suas cinco espécies apresentam comportamento e aparência semelhantes, sendo o grupo de fácil diagnose (Sick, 2001). A filogenia obtida (Figura 2.6) é congruente com as relações sugeridas por caracteres morfológicos e ecológicos (Sick, 2001), pois o gênero *Paroaria* parece estar dividido em dois grupos, sendo um formado por *P. coronata* e *P. dominicana* e outro por *P. baeri*, *P. gularis* e *P. capitata*. O apoio encontrado para o primeiro grupo não é alto em nenhuma das análises, porém, o nó que agrupa o segundo conjunto de espécies tem alto apoio (Figura 2.6). Para aumentar o apoio do primeiro grupo seria interessante aumentar a amostragem de

P. coronata e/ou utilizar mais um marcador molecular (Omeland *et al.*, 1999). A monofilia desse gênero, sua divisão em dois grupos e o apoio baixo para o clado (*P. coronata*, *P. dominicana*) também foram encontrados por Porzecanski (com. pess.) em um estudo molecular (citocromo b) e morfológico de todas as espécies do gênero.

Para analisar as implicações biogeográficas dessa divisão do gênero *Paroaria*, é interessante analisar o mapa pontual de distribuição obtido para suas espécies (Figura 2.4). Nesse mapa verifica-se que os tradicionais mapas zonais de distribuição para o gênero (Figura 2.1) podem levar a interpretações errôneas. Segundo o mapa zonal de distribuição, *P. coronata* e *P. capitata* estariam associadas às províncias do Chaco, Paranaense, da Floresta de Pedemonte Subandino *sensu* Prado (2000), dos Pampas e do Pantanal. Porém, segundo o mapa pontual de distribuição, essas espécies não estariam associadas ao Chaco *sensu* Prado (2000), mas aos arredores do Chaco *sensu* Prado (2000). Com isso, as hipóteses biogeográficas que não distinguem os arredores do Chaco (Cabrera & Willink, 1973; Morrone, 2006) foram aqui interpretadas com cautela, pois toda possível relação levantada por essas hipóteses envolvendo essa província pode ser explicada tanto pelo Chaco *sensu* Prado (2000) como pelos seus arredores. Portanto, os mapas pontuais de distribuição de *P. coronata* e *P. capitata* (Figura 2.4) apóiam a diferenciação entre os arredores do Chaco e o Chaco como proposto por Prado (2000). Além disso, segundo as distribuições geográficas pontuais e a hipótese de Prado (2000), *P. coronata* e *P. capitata* passariam a ocorrer somente nos arredores do Chaco e não mais também no Chaco (Figura 2.2d). Aparentemente, as distribuições de *P. coronata* e *P. capitata* se sobrepõem, porém, essas espécies ocupam habitats distintos, sendo *P. coronata* associada a áreas mais secas e *P. capitata*, a áreas alagadas e a rios (Sick, 2001).

O mapa pontual de distribuição do gênero *Paroaria* (Figura 2.4) permite supor alguns padrões de estrutura populacional. *P. gularis* poderia ter dois grupos, sendo um a sudoeste e outro a nordeste da região com ausência de coletas do noroeste do Mato Grosso à Colômbia. *P. capitata* e *P. coronata* apresentariam dois grupos, sendo um associado ao oeste da província Paranaense *sensu* Prado (2000) e ao Pantanal e outro à província da Floresta de Pedemonte Subandino *sensu* Prado (2000). Ao norte do Chaco *sensu* Prado (2000), haveria migração entre essas duas populações. Vale ressaltar que as distribuições sugeridas no presente estudo para os grupos de *P. gularis* e *P. capitata* não coincidem com as distribuições de suas subespécies (Peters, 1970).

Na filogenia obtida (Figura 2.6) observam-se dois clados dentro do gênero *Paroaria*: um formado por aves associadas às florestas abertas do nordeste e sul da América do Sul *cis-andina* (*P. coronata*, *P. dominicana*) e outro associado a matas ciliares das bacias hidrográficas do centro, norte e oeste da América do Sul *cis-andina* (*P. baeri* (*P. gularis*, *P. capitata*)). O primeiro clado não é congruente com nenhuma das hipóteses biogeográficas discutidas, pois todas sugerem que uma ou mais espécies adicionais também pertenceriam a este clado. As relações filogenéticas do clado (*P.*

baeri (*P. gularis*, *P. capitata*)) não devem ser interpretadas como um indício de relação entre os biomas Cerrado, Amazônia, Pantanal e arredores do Chaco *sensu* Prado (2000), pois essas espécies habitam vegetações específicas desses biomas (matas ciliares). Com isso, as relações desse clado devem ser vistas como um indício de relação entre as matas ciliares do Cerrado, da Amazônia, do Pantanal e dos arredores do Chaco *sensu* Prado (2000) e não entre os biomas propriamente ditos, uma vez que diferentes formações dentro de um mesmo bioma podem ter histórias evolutivas distintas (Silva, 1995a). Isso também é evidenciado pelas diferentes interpretações biogeográficas da América do Sul (Cabrera & Willink, 1973; Morrone, 2006; Porzecanski & Cracraft, 2005; Prado, 2000). Como nenhuma das hipóteses biogeográficas discutidas considera as matas ciliares, as relações filogenéticas do clado formado por *P. baeri*, *P. gularis* e *P. capitata* não puderam ser utilizadas para testá-las.

Para auxiliar a levantar hipóteses sobre os eventos geológicos e/ou climáticos possivelmente relacionados às diversificações observadas no gênero *Paroaria*, realizou-se a datação das divergências (Figura 2.7; Tabela 2.2). Apesar das limitações e dos problemas associados a essa metodologia, ela é uma ferramenta útil por possibilitar uma estimativa ao menos grosseira da idade de linhagens para as quais o registro fóssil é raro, como as aves (Garcia-Moreno, 2004; Lovette, 2004). As estimativas foram obtidas com base no relógio molecular utilizando taxas de divergência corrigidas de 1,665% e 2,109%/M.a..

O nó que separa *G. cristata* e o gênero *Paroaria* foi datado entre 10,3 e 8,1 M.a. (Mioceno; Figura 2.7; Tabela 2.2, nó 1). Esse período coincide com algumas divergências inter-genéricas dentro de uma tribo proximamente relacionada, a tribo Thraupini (Burns, 1997; Fjeldså & Rahbek, 2006). Se *G. cristata* (gênero monotípico) for realmente o grupo irmão do gênero *Paroaria*, como essa espécie e o clado (*P. coronata*, *P. dominicana*) ocorrem em florestas mais abertas e secas, é mais parcimonioso que o ancestral do gênero *Paroaria* ocorresse nessas florestas. Com isso, supõe-se que sua distribuição geográfica acompanhasse a distribuição desse tipo de florestas. As vegetações mais abertas e secas parecem sofrer expansão durante períodos mais secos e frios, característicos de períodos glaciais, e retração durante períodos mais úmidos e quentes, característicos de períodos inter-glaciais (Potts *et al.*, 1992). Registros palinológicos indicam que, durante os períodos glaciais, as Florestas Sazonais Tropicais expandiram sua distribuição de modo a contornar completamente o Cerrado, chegando a invadir certas regiões da Floresta Amazônica (Pennington *et al.*, 2005). Entre 7,0 e 4,6 M.a. é registrada uma glaciação (Clapperton, 1993). Com isso, supõe-se que nesse período o ancestral de *Paroaria* ocorreria ao longo das Florestas Sazonais Tropicais e das matas de galeria do Cerrado. A separação entre os clados (*P. coronata*, *P. dominicana*) e (*P. baeri* (*P. gularis*, *P. capitata*)) foi datada entre 4,7 e 3,7 M.a. (início do Plioceno; Figura 2.7; Tabela 2.2, nó 2). Entre 4,6 e 3,5 M.a. é registrada uma inter-glaciação. Durante esse

período, a distribuição do ancestral do gênero teria se restringido com a retração das florestas abertas e secas. Poderia se supor que uma população do ancestral de *Paroaria* tenha se mantido entre o Cerrado e a Amazônia, associando-se às matas de galeria do centro-oeste do Cerrado. Essa população teria originado o ancestral do clado (*P. baeri* (*P. gularis*, *P. capitata*)). A outra população teria acompanhado a retração das Florestas Sazonais Tropicais, ocupando regiões das atuais províncias da Caatinga e Paranaense. Essa população teria originado o ancestral do clado (*P. coronata*, *P. dominicana*).

A separação entre as espécies *P. coronata* e *P. dominicana* foi datada entre 3,4 e 2,7 M.a. (Plioceno intermediário; Figura 2.7; Tabela 2.2, nó 3). Isso coincide com o auge do soerguimento dos Andes centrais (3,5 e 2,3 M.a.; entre as latitudes 16 °S e 28 °S; Gregory-Wodzicki, 2000), conseqüente depressão da região dos Chacos e o soerguimento do Planalto Brasileiro (Petri & Fúlfaro, 1988). Esses soerguimentos formaram uma barreira para a chegada de nuvens à região dos Chacos, resultando nas condições climáticas xéricas da região. Com isso, as atuais formações vegetais começaram a se estabelecer no sul da América do Sul (Ortiz-Jaureguizar & Cladera, 2006). As alterações geológicas poderiam ter separado os grupos de organismos que ocupavam os arredores do Chaco daqueles que ocupavam o sul e leste do Planalto Brasileiro (Silva, 1995a), dando origem, respectivamente, às espécies *P. coronata* e *P. dominicana*. Nesse mesmo período (entre 3,5 e 3,2 M.a.) é registrada uma glaciação (Clapperton, 1993). Como as glaciações provocariam a expansão de habitats favoráveis às espécies do gênero *Paroaria*, essa glaciação não deve estar relacionada com a divergência entre *P. coronata* e *P. dominicana*.

A separação entre a espécie *P. baeri* e o ancestral das espécies *P. gularis* e *P. capitata* foi datada entre 2,3 e 1,8 M.a. (Plioceno superior; Figura 2.7; Tabela 2.2, nó 4). Essa data coincide com o final do auge do soerguimento do nordeste dos Andes (3,0 e 2,0 M.a.; Gregory-Wodzicki, 2000; Hoorn *et al.*, 1995), conseqüente depressão do Pantanal e da bacia do Acre e soerguimento do Planalto Central (Petri & Fúlfaro, 1988). Com isso, o Pantanal e a bacia amazônica começam a apresentar sistemas fluviais e regimes pluviais semelhantes aos atuais (Campbell Jr. *et al.*, 2006; Lundberg *et al.*, 1998). As alterações geológicas podem ter levado à separação de grupos de organismos que ocupavam o Planalto Central daqueles que ocupavam o Pantanal e a bacia do Acre (Silva, 1995a), dando origem, respectivamente, à espécie *P. baeri* e ao ancestral das espécies *P. gularis* e *P. capitata*.

A separação entre as espécies *P. gularis* e *P. capitata* foi datada entre 1,6 e 1,2 M.a. (Pleistoceno inferior; Figura 2.7; Tabela 2.2, nó 5), que coincide com um período de inter-glaciação (entre 2,2 e 1,2 M.a. [Clapperton, 1993]). Como nesses períodos mais quentes e úmidos se prevê a expansão de florestas úmidas (Potts *et al.*, 1992), uma possível expansão da Floresta Amazônica nas regiões de Rondônia, da Bolívia e do sudoeste do Mato Grosso, poderia ter separado grupos de

indivíduos que ocupavam a bacia do Acre daqueles que ocupavam o Pantanal, dando origem, respectivamente, às espécies *P. gularis* e *P. capitata*. Com isso, podemos levantar a hipótese de que estudos filogeográficos devem identificar haplótipos ancestrais de *P. gularis* mais a sudoeste da Amazônia e de *P. capitata* no Pantanal.

A hipótese levantada aqui de que as diversificações do gênero *Paroaria* ocorreram principalmente durante o Plioceno, não é exclusiva para esse gênero e já foi sugerida para as diversificações de alguns gêneros de Angiospermas associados às Florestas Sazonais Tropicais (Pennington *et al.*, 2005). Além disso, as diversificações internas da subfamília Emberizinae (Barker *et al.*, 2004) e das tribos Cardinalini (Klicka *et al.*, 2007) e Thraupini (Burns, 1997; Fjeldså & Rahbek, 2006) também são datadas para o mesmo período. As diversificações pliocênicas desses grupos são explicadas com base no soerguimento dos Andes e suas consequências regionais (Barker *et al.*, 2004; Burns, 1997; Fjeldså & Rahbek, 2006). Essa explicação também é utilizada para as diversificações do gênero *Paroaria* ocorridas no Plioceno (entre 3,5 e 1,8 M.a.). Como o soerguimento dos Andes teve grandes consequências climáticas e geológicas na América do Sul (Gregory-Wodzicki, 2000; Hoorn *et al.*, 1995; Ortiz-Jaureguizar & Cladera, 2006; Petri & Fúlfaro, 1988), é plausível aceitar que ele tenha gerado o isolamento de grupos de organismos e a subsequente especiação. Porém, essas modificações geológicas não seriam os únicos possíveis eventos vicariantes influenciando a diversificação do gênero *Paroaria*. As diversificações mais antiga (entre 4,7 e 3,7 M.a.; Figura 2.7; Tabela 2.2, nó 2) e mais recente (entre 1,6 e 1,2 M.a.; Figura 2.7; Tabela 2.2, nó 5) desse gênero poderiam estar associadas aos ciclos climáticos. Portanto, as hipóteses dos Refúgios (Haffer, 1997) e Paleogeográfica (Emsley, 1965) são aplicáveis às diversificações do gênero *Paroaria*, porém, em momentos distintos.

Weir (2006) sugere que as principais causas da diversificação das aves de terras baixas da América do Sul são as modificações geológicas. Porém, esse autor considera que os ciclos climáticos só seriam capazes de causar vicariância durante o Pleistoceno. A diversificação mais antiga do gênero *Paroaria* vai de encontro com essa hipótese, pois poderia estar relacionada aos ciclos climáticos do Plioceno. Tais ciclos climáticos parecem ter ocorrido desde o Terciário e estão intimamente relacionados à distribuição dos vertebrados na América do Sul, portanto, não se pode descartar uma associação entre os ciclos e eventos vicariantes desde o Terciário (Haffer, 1997).

Vale ressaltar que o presente estudo não amostrou todas as subespécies do gênero *Paroaria*. Portanto, não podemos concluir que as espécies desse gênero sejam reciprocamente monofiléticas. A utilização de táxons abaixo do nível específico poderia resultar em novas árvores que indicam a poli ou parafilia de uma ou mais espécies do gênero (Omeland *et al.*, 1999). Com isso, as interpretações das relações filogenéticas inter-específicas poderiam ser alteradas e, consequentemente, as interpretações biogeográficas também poderiam ser modificadas. Para testar a

monofilia das espécies do gênero *Paroaria* e as interpretações biogeográficas realizadas no presente trabalho seria importante realizar uma nova análise filogenética desse gênero amostrando todas as suas subespécies.

2.5. REFERÊNCIAS BIBLIOGRÁFICAS

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automatic Control* **19**: 716-723.
- Alves, A. L. 2005. *Aves Brasileiras: uma Visão Fotográfica*. Ed. Pallotti, Brasília.
- Arballo, E.; Cravino, J. L. 1999. *Handbook of the Birds of Uruguay*. Vol. 1. Ed. Hemisferio Sur, Montevideo.
- Barker, F. K.; Cibois, A.; Schikler, P.; Feinstein, J.; Cracraft, J. 2004. Phylogeny and diversification of the largest avian radiation. *Proc. Natl. Acad. Sci. USA* **101**: 11040-11045.
- Brown, W. Jr.; George, M. Jr.; Wilson, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **76**: 1967-71.
- Burns, K. J. 1997. Molecular systematics of tanagers (Thraupinae): evolution and biogeography of a diverse radiation of neotropical birds. *Mol. Phylogenet. Evol.* **8**: 334-348.
- Burns, K. J.; Naoki, K. 2004. Molecular phylogenetics and biogeography of Neotropical tanagers in the genus *Tangara*. *Mol. Phylogenet. Evol.* **32**: 838-854.
- Cabrera, A. L.; Willink, A. 1973. *Biogeografía de América Latina*. Ed. Secretaria General de la Organización de los Estados Americanos, Washington.
- Campbell Jr., K. E.; Frailey, C. D.; Romero-Pittman, L. 2006. The Pan-Amazonian Ucayali Peneplain, late Neogene sedimentation in Amazonia, and the birth of the modern Amazon River system. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **239**: 166-219.
- Cheng, S.; Higuchi, R.; Stoneking, M. 1994. Complete mitochondrial genome amplification. *Nat. Genet.* **7**: 350-351.
- Clapperton, C. 1993. *Quaternary Geology and Geomorphology of South America*. Elsevier Press, Amsterdam.
- Cooper, R. A.; Poinar, H. N. 2000. Ancient DNA: do it right or not at all. *Science* **289**: 1139.
- Crisci, J. V.; Katinas, L.; Posadas, P. 2003. *Historical Biogeography: an Introduction*. Harvard University Press, Cambridge.
- Desjardins, P.; Morais, R. 1990. Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *J. Mol. Biol.* **212**: 599-634.

- Emsley, M. G. 1965. Speciation in *Heliconius* (Lep., Nymphalidae): Morphology and geographic distribution. *Zoologica* **50**: 191-254.
- Ericson, P. G. P.; Christidis, L.; Irestedt, M.; Norman, J. A. 2002. Systematics affinities of the lyrebirds (Passeriformes: *Menura*), with a novel classification of the major groups of passerine birds. *Mol. Phylogenet. Evol.* **25**: 53-62.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **79**: 783-791.
- Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.* **22**: 521-565.
- Felsenstein, J. 2004. PHYLIP (Phylogeny Inference Package) version 3.6. *Distributed by the author*. Department of Genome Sciences, University of Washington, Seattle.
- Fjeldså, J.; Rahbek, C. 2006. Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America. *Integr. Comp. Biol.* **46**: 72-81.
- Garcia-Moreno, J. 2004. Is there a universal mtDNA clock for birds? *J. Avian Biol.* **35**: 465-468.
- Graur, D.; Li, W. H. 1999. *Fundamentals of Molecular Evolution*. Sinauer Associates Press, Sunderland.
- Gregory-Wodzicki, K. M. 2000. Uplift history of the Central and Northern Andes: a review. *GSA Bulletin* **112**: 1091-1105.
- Griffiths, C. S. 1997. Correlation of functional domains and rates of nucleotide substitution in cytochrome b. *Mol. Phylogenet. Evol.* **7**: 353-365.
- Guindon, S.; Lethiec, F.; Duroux, P.; Gascuel, O. 2005. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* **33** (Web Server issue): W557-W559.
- Haffer, J. 1997. Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiv. Conserv.* **6**: 451-476.
- Hecker, K. H.; Roux, K. H. 1996. High and low annealing temperatures increase both specificity and yield in “Touchdown” and “StepDown” PCR. *BioTechniques* **20**: 478-485.
- Hillis, D. M.; Moritz, C.; Mable, B. K. 1996. *Molecular Systematics*. 2^a ed. Sinauer Associates Press, Sunderland.
- Hoorn, C.; Guerrero, J.; Sarmiento, G. A.; Lorente, M. A. 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology* **23**: 237-240.
- Huelsenbeck, J. P.; Ronquist, F. 2001. MRBAYES. Bayesian inference of phylogeny. *Bioinformatics* **17**: 754-755.
- Klicka, J.; Burns, K.; Spellman, G. M. 2007. Defining a monophyletic Cardinalini: a molecular perspective. *Mol. Phylogenet. Evol.* **45**: 1014-1032.

- Klicka, J.; Johnson, K. P.; Lanyon, S. M. 2000. New world nine-primaried oscine relationship: constructing a mitochondrial DNA framework. *Auk* **117**: 321-336.
- Klicka, J.; Spellman, G. M. 2007. A molecular evaluation of the North American “grassland” sparrow clade. *Auk* **124**: 537-551.
- Kocher, T. D.; Thomas, W. K.; Meyer, A.; Edwards, S. V.; Pääbo, S.; Villablanca, F. X.; Wilson, A. C. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86**: 6196-6200.
- Lougheed, S. C.; Freeland, J. R.; Handford, P.; Boag, P. T. 2000. A molecular phylogeny of warbling-fishes (Poospiza): paraphyly in Neotropical Emberizid genus. *Mol. Phylogenet. Evol.* **17**: 367-378.
- Lovette, I. J. 2004. Mitochondrial dating and mixed support for the “2% rule” in birds. *Auk* **121**: 1-6.
- Lundberg, J. G.; Marshall, L. G.; Guerrero, J.; Horton, B.; Malabarba, M. C.; Wesselingh, F. 1998. The stage for Neotropical fish diversification: a history of tropical south American rivers. in Malabarba, L. R.; Reis, R. E.; Vari, R. P.; Lucena, S. M.; Lucena, C. A. S. (eds.) *Phylogeny and classification of Neotropical fishes*. Edipucrs, Porto Alegre. Pp. 13-48
- Miyaki, C. Y.; Mantioli, S. R.; Burke, T.; Wajntal, A. 1998. Parrot evolution and paleogeographical events: mitochondrial DNA evidence. *Mol. Biol. Evol.* **15**: 544-551.
- Morrone, J. J. 2006. Biogeographic areas and transition zones of Latin America and the Caribbean Islands based on panbiogeographic and cladistic analyses of entomofauna. *Annu. Rev. Entomol.* **51**: 467-494.
- Omland, K. E.; Lanyon, S. M.; Fritz, S. J. 1999. A molecular phylogeny of the New World orioles (*Icterus*): the importance of dense taxon sampling. *Mol. Phylogenet. Evol.* **12**: 224-239.
- Ortiz-Jaureguizar, E.; Cladera, G.A. 2006. Paleoenvironmental evolution of southern South America during the Cenozoic. *J. Arid Environ.* **66**: 498-532.
- Päckert, M.; Martens, J.; Tietze, D. T.; Dietzen, C.; Wink, M.; Kvist, L. 2007. Calibration of a molecular clock in tits (Paridae) – Do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? *Mol. Phylogenet. Evol.* **44**: 1-14.
- Page, R. D. M.; Holmes, E. C. 1998. *Molecular Evolution: a Phylogenetic Approach*. Blackwell Science Press, Oxford.
- Paynter, R. A. Jr. 1982. *Ornithological Gazetteer of Venezuela*. Museum of Comparative Zoology Press, Cambridge.
- Paynter, R. A. Jr. 1989. *Ornithological Gazetteer of Paraguay*. 2^a ed. Museum of Comparative Zoology Press, Cambridge.

- Paynter, R. A. Jr. 1992. *Ornithological Gazetteer of Bolivia*. 2^a ed. Museum of Comparative Zoology Press, Cambridge.
- Paynter, R. A. Jr. 1993. *Ornithological Gazetteer of Ecuador*. 2^a ed. Museum of Comparative Zoology Press, Cambridge.
- Paynter, R. A. Jr. 1994. *Ornithological Gazetteer of Uruguay*. 2^a ed. Museum of Comparative Zoology Press, Cambridge.
- Paynter, R. A. Jr. 1995. *Ornithological Gazetteer of Argentina*. 2^a ed. Museum of Comparative Zoology Press, Cambridge.
- Paynter, R. A. Jr. 1997. *Ornithological Gazetteer of Colombia*. 2^a ed. Museum of Comparative Zoology Press, Cambridge.
- Paynter, R. A. Jr.; Traylor, M. A. Jr. 1991. *Ornithological Gazetteer of Brazil*. 2 vols. Museum of Comparative Zoology Press, Cambridge.
- Pennington, R. T.; Lavin, M.; Prado, D. E.; Pendry, C. A.; Pell, S. K. 2005. Climate change and speciation in neotropical seasonally dry forest plants. *In*: Malhi, Y.; Phillips, O. (eds.) *Tropical Forests & Global Atmospheric Change*. Oxford University Press, Oxford. Pp. 199-214.
- Pérez-Emám, J. L. 2005. Molecular phylogenetics and biogeography of the Neotropical redstarts (*Myioborus*; Aves, Parulinae). *Mol. Phylogenet. Evol.* **37**: 511-528.
- Peters, J. L. 1970. *Check-list of the Birds of the World*. Vol. XIII. The Hefferman Press, Worcester.
- Petri, S.; Fúlfaro, V. J. 1988. *Geologia do Brasil*. Ed. da Universidade de São Paulo, São Paulo.
- Porzecanski, A. L.; Cracraft, J. 2005. Cladistic analysis of distributions and endemism (CADE): using raw distributions of birds to unravel the biogeography of the South American aridlands. *J. Biogeogr.* **32**: 261-275.
- Posada, D.; Crandall, K. A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.
- Potts, R.; Behrensmeyer, A. K.; Taggart, R. E.; Spaulding, W. G.; Harris, J. A.; Valkenburgh, B. V.; Martin, L. D.; Damuth, J. D.; Foley, R. 1992. Late Cenozoic Terrestrial Ecosystems. *In*: Behrensmeyer, A. K.; Damuth, J. D.; DiMichele, W. A.; Potts, R.; Sues, H.-D.; Wing, S. L. (eds.) *Terrestrial Ecosystems Through Time: Evolutionary Paleocology of Terrestrial Plants and Animals*. University of Chicago Press, Chicago. Pp. 419-542.
- Prado, D. E. 2000. Seasonally dry forests of tropical South America: from forgotten ecosystems to a new phytogeographic unit. *Edinb. J. Bot.* **57**: 437-461.
- Prado, D. E.; Gibbs, P. E.; Pott, A.; Pott, V. J. 1992. Chaco-Pantanal transition in Southern Mato Grosso (Brazil). *In*: Furley, P. A.; Proctor, J.; Ratter, J. A. (eds) *Nature and Dynamics of Forest–Savanna Boundaries*. Chapman & Hall Press, London. Pp. 451–470.

- Ribas, C. C.; Gaban-Lima, R.; Miyaki, C. Y.; Cracraft, J. 2005. Historical biogeography and diversification within the Neotropical parrot genus *Pionopsitta* (Aves: Psittacidae). *J. Biogeog.* **32**: 1409-1427.
- Ridgely, R. S.; Tudor, G. 1989. *The Birds of South America: The Oscine Passerines*. Vol. 1. University of Texas Press, Austin.
- Ronquist, F.; Huelsenbeck, J. P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574.
- Rutschman, L. 2006. Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. *Diversity Distrib.* **12**: 35-48.
- Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Ehrlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.
- Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**: 301-302.
- Shields, G. F.; Wilson, A. C. 1987. Calibration of mitochondrial DNA evolution in geese. *J. Mol. Evol.* **24**: 212-217.
- Sibley, C. G.; Monroe, B. L. Jr. 1990. *Distribution and Taxonomy of Birds of the World*. Yale University Press, New Haven.
- Sick, H. 2001. *Ornitologia Brasileira*. 3^a ed. Ed. Nova Fronteira, Rio de Janeiro.
- Silva, J. M. C. 1995a. Avian inventory of the Cerrado region, South America: implications for biological conservation. *Bird Conserv. Internatl.* **5**: 291-304.
- Smith, A. B. 1994. Rooting molecular trees: problems and strategies. *Biol. J. Linnean Soc.* **51**: 279-292.
- Sorenson, M. D.; Ast, J. A.; Dimcheff, D. E.; Yuri, T.; 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.
- Stephens, L.; Traylor, M. A. Jr. 1983. *Ornithological Gazetteer of Peru*. Museum of Comparative Zoology Press, Cambridge.
- Stephens, L.; Traylor, M. A. Jr. 1985. *Ornithological Gazetteer of Guianas*. Museum of Comparative Zoology Press, Cambridge.
- Swofford, D. L. 2002. PAUP*: *Phylogenetic Analysis Using Parsimony (* and other methods)*, version 4. Sinauer Associates, Sunderland.
- Tamura, K.; Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512-526.

- Weir, J. T. 2006. Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. *Evolution* **60**: 842-855.
- Xia, X.; Xie, Z. 2001. DAMBE: Data analysis in molecular biology and evolution. *J. Hered.* **92**: 371-373.
- Yuri, T.; Mindell, D. P. 2002. Molecular phylogenetic analysis of Fringillidae, “New World nine-primaried oscines” (Aves: Passeriformes). *Mol. Phylogenet. Evol.* **23**: 229-243.

CAPÍTULO 3

Filogeografia de *Paroaria dominicana*

RESUMO

A Caatinga é um bioma brasileiro considerado como importante centro de endemismo. Uma espécie endêmica é a *Paroaria dominicana* que ocorre ao longo de toda a Caatinga. No presente trabalho, realizamos um estudo filogeográfico dessa espécie visando compreender como a sua diversidade genética se distribui nesse bioma. Para isso 391 pb do primeiro domínio da região controladora e os 351 pb do ND3 foram seqüenciados para 51 indivíduos amostrados de forma a representar sua distribuição. A espécie irmã *P. coronata* foi selecionada como grupo externo mais apropriado. Porém, devido a problemas laboratoriais ela foi substituída por *P. capitata* em algumas matrizes de dados. Testes de neutralidade foram aplicados aos dados e sugerem que: 1) há fuga do modelo de populações em equilíbrio neutro causada por seleção positiva (fases pré ou pós-fixação) ou por seleção negativa atuando concomitantemente ou não a uma expansão populacional recente; e 2) a não utilização do grupo externo mais adequado causou saturação nos dados, prejudicando a realização de inferências mais precisas sobre as forças causadoras do padrão de polimorfismos encontrado. Com isso, análises de demografia histórica e de tamanho populacional efetivo foram comprometidas. Já a análise da estrutura populacional não é afetada pela não neutralidade dos dados. Análises de máxima parcimônia, máxima verossimilhança, rede de haplótipos, AMOVA, SAMOVA e teste de Mantel indicam que: 1) a espécie *P. dominicana* não apresenta estrutura populacional para o DNAm; 2) essa espécie deve ser tratada como uma única Unidade de Manejo; e 3) o rio São Francisco e os tipos de Caatinga não constituem barreiras geográfica e ecológica para essa espécie, respectivamente.

ABSTRACT

Caatinga is a Brazilian biome considered as an important endemism centre. One of its endemic species is *Paroaria dominicana* that occurs all over the Caatinga. In the present work we performed a phylogeographic study of the species to characterize how its genetic diversity is distributed in the biome. We sequenced 391 bp from the first domain of the control region and 351 bp from the ND3 for 51 individuals. The sister species *P. coronata* was used as outgroup. However, due to technical problems it was replaced by *P. capitata* in some data matrices. Neutrality tests

results suggest that: 1) there is departure from the neutral equilibrium population model due to negative or positive selection (pre or post fixation phases) acting simultaneously or not with a recent population expansion; and 2) the use of a non-sister outgroup caused saturation of the data, jeopardizing more precise inferences on the forces shaping the polymorphism pattern found. Therefore, historical demography and effective population size analyses could not be performed. However, this departure from neutrality does not affect population structure analyses. Maximum parsimony, maximum likelihood, haplotype network, AMOVA, SAMOVA and Mantel test analyses suggest that: 1) the species *P. dominicana* is not geographically structured based on mtDNA data; 2) this species can be considered as a single Management Unit; and 3) the São Francisco river and the types of Caatinga are not geographic and ecological barriers for this species, respectively.

3.1. INTRODUÇÃO

A filogeografia é o estudo dos princípios e dos processos que governam a distribuição geográfica de linhagens intra-específicas e entre espécies filogeneticamente próximas (Avice, 2000). Portanto, está diretamente ligada à biologia da conservação, pois os esforços conservacionistas são mais efetivos quando se busca a manutenção das linhagens evolutivas e dos habitats que as mantém (Crozier, 1992; Moritz, 1994; Moritz, 2002; Moritz & Faith, 1998). Porém, esses esforços enfrentam três desafios: 1) o conhecimento dos padrões e dos processos associados à diversidade biológica é incompleto; 2) as alterações naturais e antrópicas são inevitáveis; e 3) os conflitos entre as necessidades da conservação da diversidade biológica e do progresso da sociedade humana. Para vencer esses desafios é necessário que os biólogos da conservação criem estratégias que priorizem a diversidade biológica levando-se em conta os esforços conservacionistas e os recursos necessários à sociedade humana (Bowen, 1999). Uma das formas para incorporar a diversidade genética nos programas de conservação foi proposta por Moritz (1994), quando definiu como Unidades de Manejo (UM) populações com significativa divergência na frequência alélica em locos nucleares ou mitocondriais. Como esse conceito reflete a diversidade genética, ele não se limita a um nível taxonômico, pois uma UM pode ser tanto uma espécie como uma população de uma espécie, isso dependerá do estudo (Moritz, 1994; Moritz & Faith, 1998).

A Caatinga é um bioma brasileiro que tem sido negligenciado nos esforços conservacionistas, apesar de ser considerada um importante centro de endemismo para plantas (Prado, 2003) e para aves sul-americanas (Cracraft, 1985; Haffer, 1985). Apenas 1% está sob alguma forma de proteção integral (Major *et al.*, 2004), o que resulta em grande perturbação por

fogo, exploração de madeira e criação de gado (Moffat, 2002). Estimativas da Empresa Brasileira de Pesquisa Agropecuária, EMBRAPA (1994), indicam que de 1979 a 1994, 40.000 km² da Caatinga estavam em processo de desertificação. Castelletti *et al.* (2003), por meio de um estudo das atividades agrícolas e do efeito das estradas na Caatinga, estimaram que cerca de 45% da região foi alterada pelo homem, valor superado apenas pela Floresta Atlântica e pelo Cerrado. Essa falta de proteção é agravada pelo baixo nível de conhecimento da biologia da Caatinga quando comparada aos outros biomas brasileiros (Silva *et al.*, 2003).

A Caatinga, também conhecida como sertão, ocupa uma área de cerca de 800.000 km², o que corresponde a 70% da área do Nordeste e a 11% do território brasileiro (EMBRAPA, 1994). Ela é caracterizada por um solo pedregoso, raso e que não armazena água, e por um quadro climático que exibe alguns dos valores meteorológicos mais extremos do país: a mais forte insolação, a mais baixa nebulosidade, as mais altas médias térmicas, as mais baixas porcentagens de umidade relativa, as mais escassas e irregulares precipitações pluviais (limitadas a dois ou três meses do ano) e as mais elevadas taxas de evaporação (Assis *et al.*, 1994; EMBRAPA, 1994; Reis, 1976). O maior e mais importante rio da região é o São Francisco, que atravessa a Caatinga nascendo no interior do estado de Minas Gerais, passando pelos estados da Bahia e de Pernambuco e desembocando no Oceano Atlântico entre os estados de Alagoas e Sergipe (Figura 3.1; MMA, 2006).

A flora da Caatinga é caracterizada como florestas arbóreas ou arbustivas, com árvores e arbustos baixos que apresentam espinhos, microfilia e algumas características xerofíticas (Assis *et al.*, 1994; Prado, 2003). Segundo Andrade-Lima (1981), a Caatinga pode ser classificada basicamente em florestas arbóreas (hipoxerofítica) e florestas arbustivas (hiperxerofítica; Figura 3.1).

Figura 3.1: Tipos de vegetação presentes na Caatinga (Major *et al.*, 2004).

Uma estratégia interessante para contribuir com os esforços conservacionistas da Caatinga pode ser avaliar os padrões filogeográficos de alguma espécie que habite esse bioma. O modelo selecionado no presente estudo foi *Paroaria dominicana*, uma ave endêmica do bioma com distribuição contínua sobre ele. Essa espécie apresenta relativa grande capacidade de voo e é alvo de intensa captura pelo comércio ilegal de aves silvestres, uma vez que é uma ave bonita com grande qualidade canora (Sick, 2001).

Para que populações com distribuições contínuas apresentem estruturação é necessário que seus indivíduos possuam movimentação reduzida, que pode ser causada por baixa mobilidade ou presença de barreiras geográficas ou ecológicas (Avice, 1994; Wakeley, 2005). Com isso, podemos levantar cinco possíveis modelos de estrutura populacional para *P. dominicana*: 1) ausência de estruturação populacional; 2) isolamento por distância, ou seja, há correlação entre as distâncias geográfica e genética; 3) populações separadas pelo rio São Francisco (modelo Rio São Francisco); 4) duas populações sendo uma associada à Caatinga hipoxerofítica e outra à Caatinga hiperxerofítica (modelo Caatingas); 5) três populações sendo uma associada à Caatinga hiperxerofítica, outra à Caatinga hipoxerofítica ocidental e outra à Caatinga hipoxerofítica oriental (modelo Regiões).

O objetivo do presente trabalho é avaliar os padrões filogeográficos de *P. dominicana*, caracterizando seus padrões de polimorfismo, bem como suas possíveis causas, e testando os modelos de estrutura populacional propostos de modo a sugerir as UMs para essa espécie.

3.2. MATERIAIS E MÉTODOS

3.2.1. Obtenção das Amostras

As amostras obtidas para o trabalho estão apresentadas na Tabela 3.1. Elas são provenientes de coletas realizadas por colaboradores e de aves capturadas em redes de neblina de 7 ou 10 metros ou com espingarda de ar comprimido (CBC-Monte Negro Super calibre 4,5 mm) munida com chumbinhos Gamo-Hunter. Para aumentar o sucesso de captura foram realizados *playbacks* para atraí-las e as redes foram colocadas próximas às fontes de alimento. De cada ave capturada em campo foi extraído um pedaço de músculo de 0,125 cm³ ou 0,1 mL de sangue com seringa de insulina da maior veia da região cervical direita ou cortou-se a unha de um dos dedos até que começasse a sangrar. Todas as amostras foram armazenadas em etanol absoluto e mantidas a -20 °C no Laboratório de Genética e Evolução Molecular de Aves do Instituto de Biociências da Universidade de São Paulo. Os espécimes coletados foram depositados no Museu de Zoologia da Universidade de São Paulo. As localidades das amostras de *P. dominicana* foram plotadas em um mapa utilizando o programa DIVA-GIS 5.4 (<http://www.diva-gis.org>; Figura 3.2).

3.2.2. Grupo Externo

Em análises filogenéticas e filogeográficas, grupos externos proximamente relacionados ao grupo interno devem ser utilizados para evitar a saturação dos dados (Smith, 1994; Fu & Li, 1993). Portanto, baseado na proposta de que *Paroaria dominicana* forme uma super-espécie com *P. coronata* (Sick, 2001) e nas relações filogenéticas obtidas no Capítulo 2, *P. coronata* foi selecionada como grupo externo.

3.2.3. Seleção do Marcador

Devido às características do DNAm citadas no Capítulo 1, selecionou-se um gene codificador de proteína (sub-unidade 3 da NADH desidrogenase; ND3) e a região controladora do DNAm. Ambos têm se mostrado informativos em estudos populacionais (Kvist *et al.*, 1998; Kvist *et al.*, 2001; Meyer, 1994; Zink *et al.*, 2002; Zink *et al.*, 2005).

Tabela 3.1: Espécie, número de registro (LGEMA: Laboratório de Genética e Evolução Molecular de Aves do Instituto de Biociências da Universidade de São Paulo; ANSP: Academy of Natural Sciences of Philadelphia), local de coleta e coletor das amostras do gênero *Paroaria*. Entre parênteses se encontra o número total de amostras por espécie.

Espécie	Localidade	Registro	Coletor
<i>P. dominicana</i> (51)	PN Serra das Confusões/PI	LGEMA2238	L. F. Silva
	Tanquinho/BA (cativeiro)	LGEMA2578, LGEMA2579, LGEMA2587	F. Nodari e G. Cabanne
	Morro do Chapéu/BA (cativeiro em Bonito/BA)	LGEMA2580, LGEMA2581	
	Bonito/BA (cativeiro)	LGEMA2582	
	Tanquinho/BA	LGEMA2583, LGEMA2584 LGEMA2585, LGEMA2586	
	Beira Rio/BA	LGEMA2588, LGEMA2589, LGEMA2590 LGEMA2591, LGEMA2592, LGEMA2593 LGEMA2594, LGEMA2595	
	Parnaguá/PI	LGEMA2596, LGEMA2597, LGEMA2598 LGEMA2599, LGEMA2600, LGEMA2601 LGEMA2602, LGEMA2603	
	Simplicio Mendes/PI	LGEMA2604, LGEMA2605, LGEMA2606 LGEMA2607, LGEMA2608, LGEMA2609 LGEMA2610	
	Vila Nova do Piauí/PI	LGEMA2611	
	Trindade/PE	LGEMA2612, LGEMA2613, LGEMA2614 LGEMA2615, LGEMA2616, LGEMA2617 LGEMA2618	
	Juremal/BA	LGEMA2619	
	Itiúba/BA	LGEMA2620, LGEMA2621, LGEMA2626	
	Itiúba/BA (cativeiro)	LGEMA2622, LGEMA2623 LGEMA2624, LGEMA2625	
	RPPN Serra das Almas/CE	LGEMA10528	F. M. Horta
<i>P. coronata</i> (1)	Uruguai, Rota 24 km 75	ANSP10244	L. Joseph
<i>P. capitata</i> (1)	Corumbá/MS	LGEMA2257	C. Graco

Figura 3.2: Mapa do nordeste do Brasil com destaque ao rio São Francisco indicando as localidades das coletas de *P. dominicana*. 0: PN Serra das Confusões/PI; 1: Tanquinho/BA; 2: Beira Rio/BA; 3: Parnaguá/PI; 4: Simplicio Mendes/PI; 5: Vila Nova do Piauí/PI; 6: Trindade/PE; 7: Juremal/BA; 8: Itiúba/BA; 9: RPPN Serra das Almas/CE.

3.2.4. Seleção dos *Primers*

Para amplificar o ND3 (351 pb) foram utilizados os *primers* A6PWL (Eberhard & Bermingham, 2004) e H11151 (Chesser, 1999; aqui denominado ND3H) que produzem um fragmento de aproximadamente 1600 pb incluindo os genes ATPase 6 (ATP6), Citocromo Oxidase III (COIII) e ND3 (Figura 3.3a). Para amplificar a região controladora (~1300 pb), primeiramente, foram utilizados os *primers* LGluTc (5'-AATTCCTACTCGGCCTCTCTCCG-3'; Cabanne, com. pess.) e H1251VcPHE (Sorenson *et al.*, 1999) para algumas amostras de *P. dominicana*. Devido ao baixo sucesso desse par de *primers*, desenvolveu-se o *primer* HPhePd (5'-GGTTGATGTTTTATCGATATGGTTTG-3') baseado em seqüências obtidas com o *primer* H1251VcPHE (Sorenson *et al.*, 1999). Com isso, para amplificar a região controladora, foram utilizados os *primers* LGluTc (Cabanne, com. pess.) e HPhePd (presente estudo; Figura 3.3b).

Figura 3.3: a) Região mitocondrial dos genes ATP6, COIII e ND3 com a localização dos *primers* utilizados: A6PWL (Eberhard & Bermingham, 2004), ND3L (L10755 de Chesser, 1999) e ND3H (H11151 de Chesser, 1999). b) Região controladora mitocondrial com a localização dos *primers* utilizados: LGluTc (Cabanne, com. pess.), H1251VcPHE (Sorenson *et al.*, 1999), HPhePd (presente estudo) e H739Vc (Sorenson *et al.*, 1999). L e H se referem às cadeias leve e pesada, respectivamente, e os números se referem à posição da base 3' no genoma mitocondrial de galinha (Desjardins & Morais, 1990). Esquemas fora de escala.

3.2.5. Extração do DNA

Para a extração de DNA total das amostras, uma pequena quantidade de sangue ou músculo foi colocada em tubos de 1,5 mL com 300 µL de TNE (50 mM de Tris, 100 mM de NaCl e 5 mM

de EDTA, pH 7,5), 30 µL de Tris 1 M (pH 7,5), 8 µL de SDS 25% e 20 µL de Proteinase K (25 ng/mL). Essa mistura foi incubada em estufa a 37 °C por 24 horas. Após a incubação, foi misturado um volume de fenol:clorofórmio:álcool isoamílico (25:24:1) e centrifugado por 10 minutos a 12000 rpm. A fase superior foi transferida para outro tubo, onde foram adicionadas duas partes de etanol 100%. Essa solução foi misturada até a precipitação do DNA. Esse DNA foi isolado, lavado e dissolvido seguindo os passos: centrifugação por 45 minutos a 12000 rpm, descarte do etanol, adição de etanol 70%, centrifugação por 10 minutos a 12000 rpm, descarte do etanol, adição de 100 µL de TE (10 mM de Tris e 1 mM de EDTA, pH 7,5), armazenamento a 4 °C. A qualidade e a quantidade do DNA foram verificadas por eletroforese em gel de agarose 1%. Após a quantificação, esse DNA foi diluído para aproximadamente 20 ng/µL para posterior amplificação.

3.2.6. Amplificação do DNA

A amplificação dos segmentos-alvo foi realizada em termociclador (TC2400, Perkin-Elmer) em volume total de 10 µL, contendo 1 µL de dNTPs (2 mM de cada), 1 µL de tampão (10x), 1 µL de *primer* L (10 µM), 1 µL de *primer* H (10 µM), 2 µL de DNA (~20 ng/µL), 0,1 µL de *Taq* polimerase (5 U/µ) e 3,9 µL de água milli-Q. No termociclador, foram realizadas reações convencionais (mesma temperatura de hibridação em todos os ciclos; Saiki *et al.*, 1988) e reações *touchdown* (temperatura de hibridação diminui em alguns ciclos; Hecker & Roux, 1996). As reações convencionais foram inicialmente desnaturadas a 95 °C por 3 minutos, e os 40 ciclos consistiram de: 30 segundos a 95 °C (desnaturação), 30 segundos a 58 °C (hibridação) e 60 segundos a 72 °C (extensão); e seguido de uma extensão final de 9 minutos a 72 °C. As reações *touchdown* variaram a temperatura de hibridação de 60 °C a 50 °C diminuindo 1 °C por ciclo nas amplificações da região controladora e de 60 °C a 55 °C diminuindo 0,5 °C por ciclo nas amplificações do ND3.

O tamanho e a concentração dos produtos da amplificação foram verificados por eletroforese em gel de agarose 1%, utilizando 1/10 do volume total da amplificação e padrões de concentrações e de tamanho conhecidos para comparação. No caso de contaminação no controle negativo, todas as reações foram descartadas. Todo o produto da amplificação foi purificado por digestão com exonuclease I (10un/µL) e fosfatase alcalina de camarão (1un/µL; Amersham/Pharmacia), adicionando-se 0,5 µL de cada enzima para cada 10 µL de produto da amplificação, incubando 1 hora a 37 °C seguido de 10 minutos a 80 °C. Alternativamente, a purificação foi realizada com a adição de um volume de Polietilenoglicol (PEG; 1 g/mL) ao produto da amplificação, misturado, incubado a 37 °C por 15 minutos, centrifugado a 12000 rpm por 15 minutos, o sobrenadante foi descartado, adicionados 125 µL de etanol 80% gelado, centrifugado a 12000 rpm por 2 minutos, o

sobrenadante foi descartado, os últimos três passos foram repetidos e foram adicionados 10 µL de água milli-Q.

3.2.7. Sequenciamento dos Produtos da Amplificação

Os produtos de amplificação foram utilizados como moldes para as reações de sequenciamento direto utilizando 200 a 400 ng desses produtos, 1,5 µL de *Big Dye Terminator Cycle Sequencing Kit* (Applied Biosystems) e 1 µL de *primer* (10 µM). O volume de produto de amplificação utilizado variou conforme seu tamanho e sua concentração. As reações foram colocadas no termociclador nas seguintes condições: 96 °C por 1 minuto, 30 ciclos de: 96 °C por 10 segundos, 50 °C por 10 segundos e 60 °C por 4 minutos. Os *primers* utilizados (Figura 3.3) para o sequenciamento do ND3 foram ND3L e ND3H (respectivamente, L10755 e H11151 de Chesser, 1999) e para o sequenciamento da região controladora foram H739Vc (Sorenson *et al.*, 1999) e HPhePd (presente estudo). As seqüências obtidas com os *primers* H739Vc (Sorenson *et al.*, 1999) e HPhePd (presente estudo) são trechos do primeiro e do terceiro domínios da região controladora, respectivamente. Baseado em seqüências de 19 indivíduos de *P. dominicana*, observou-se que em 391pb do primeiro domínio há cinco sítios variáveis, enquanto que em 397pb do terceiro domínio há dois sítios variáveis. Devido à sua maior variabilidade, o primeiro domínio da região controladora foi selecionado para o presente estudo. Como não foi possível sequenciar a fita pesada do primeiro domínio da região controladora devido à presença de uma repetição de aproximadamente 12 bases C, foram seqüenciadas as cadeias leves de produtos de amplificações independentes para confirmar a seqüência.

Para a precipitação do produto dessa reação foram adicionados quatro volumes de isopropanol 75%. Essa solução foi centrifugada por 45 minutos a 12000 rpm, descartado o álcool, foram adicionados 200 µL de etanol 70% seguido de nova centrifugação por 20 minutos a 12000 rpm. Essas amostras foram carregadas em gel de poliacrilamida 5% em seqüenciador automático ABI 377/36 (Applied Biosystems).

3.2.8. Análise dos Resultados

Utilizando o programa CodonCode Aligner v.1.6.3 (CodonCode Corporation) as seqüências da cadeia pesada obtidas para o ND3 foram revertidas e complementadas para que ficassem idênticas às da cadeia leve. As seqüências obtidas para o primeiro domínio da região controladora já correspondiam à cadeia leve. Utilizando-se o mesmo programa, as seqüências foram alinhadas com correção manual e as seqüências de DNA do ND3 foram traduzidas em seqüências de aminoácidos

para verificar a presença de códons de parada inesperados. Os modelos de evolução mais adequados para as matrizes de dados foram selecionados utilizando-se o Critério de Informação de Akaike (AIC; Akaike, 1974) no programa Modeltest 3.7 (Posada & Crandall, 1998).

A partir da matriz de seqüências alinhadas foram obtidos o número de haplótipos (h), a diversidade nucleotídica (π) e os índices de neutralidade D_T (Tajima, 1989), F_s (Fu, 1997), S modificado (Fu, 1996), D_{FL} e F (Fu & Li, 1993) utilizando o programa DnaSP 4.10.9 (Rozas *et al.*, 2003); e H -normalizado, DH , E (Zeng *et al.*, 2006) e DHEW (Zeng *et al.*, 2007b), com os programas *ReadFasta* e *ReadFastaDHEW*, gentilmente cedidos por Kai Zeng. Para avaliar se havia saturação das matrizes de seqüências, foram obtidos gráficos de saturação (Griffiths, 1997) utilizando-se o programa DAMBE (Xia & Xie, 2001). Os padrões de distribuição de diferenças nucleotídicas par-a-par (*mismatch distribution*; Rogers & Harpending, 1992) foram obtidos utilizando o programa DnaSP 4.10.9 (Rozas *et al.*, 2003). Para estimar os padrões esperados segundo os modelos de tamanho populacional constante e de expansão populacional, os valores de Tau (τ), Theta inicial (θ_0) e final (θ_1) foram calculados no programa Arlequin v. 3.11 (Excoffier *et al.*, 2005) e, posteriormente, analisados no programa DnaSP 4.10.9 (Rozas *et al.*, 2003).

Testou-se a homogeneidade da composição de bases entre os táxons e obteve-se a composição nucleotídica e as árvores pelos métodos de máxima parcimônia utilizando o programa PAUP* v. 4.0 b10 (Swofford, 2002) e máxima verossimilhança no *PHYML online web server* (Guindon *et al.*, 2005). Análises de *bootstrap* (Felsenstein, 1985) com 1000 réplicas para as análises de máxima parcimônia e com 500 réplicas para as análises de máxima verossimilhança foram realizadas para estimar o apoio de cada clado (Page & Holmes, 1998). Redes de haplótipos foram obtidas utilizando-se o método de *median-joining* com posterior análise de máxima parcimônia utilizando o programa Network 4.2.0.1 (<http://www.fluxus-technology.com>; Polzin & Daneschmand, 2003).

As distâncias geográficas foram calculadas com base nas coordenadas geográficas das amostras utilizando-se o programa DIVA-GIS 5.4 (<http://www.diva-gis.org>) e as distâncias genéticas foram calculadas utilizando-se o programa MEGA v. 4 (Tamura *et al.*, 2007). A correlação entre essas distâncias foi testada pelo teste de Mantel (Mantel, 1967) utilizando-se o programa XLSTAT v. 2008.2.01 (Addinsoft).

A análise de variação molecular (AMOVA; Excoffier *et al.*, 1992) e a análise espacial de variação molecular (SAMOVA; Dupanloup *et al.*, 2002) foram realizadas utilizando-se, respectivamente, os programas Arlequin v. 3.11 (Excoffier *et al.*, 2005) e SAMOVA v. 1.0 (Dupanloup *et al.*, 2002). A SAMOVA utiliza informações genéticas e geográficas das localidades de coleta e identifica o número de grupos de populações que gera a maior proporção de variação genética entre grupos. Ou seja, nessa análise, a estruturação populacional não é definida *a priori*,

porém, é necessário informar *a priori* a quantidade de possíveis grupos. Caso a quantidade de grupos seja desconhecida, é sugerido que ela seja modificada (aumentada) até que o valor de Φ_{CT} diminua em relação ao Φ_{CT} obtido com o número de grupos imediatamente menor. O número mínimo inicial de grupos a ser testado é dois (Dupanloup *et al.*, 2002).

3.3. RESULTADOS

Devido a problemas na amplificação da região controladora de *Paroaria coronata*, *P. capitata* também foi seqüenciada e utilizada como grupo externo. O ND3 de todos os indivíduos foi completamente seqüenciado e 391 pb do primeiro domínio da região controladora (RC) foram seqüenciados de todas as amostras exceto de *P. coronata*. A partir desse conjunto de dados foram obtidas quatro matrizes de seqüências: A) seqüências do ND3 (351 caracteres) para todos indivíduos de *P. dominicana* e de *P. coronata*; B) seqüências do ND3 (351 caracteres) para todos indivíduos de *P. dominicana* e de *P. capitata*; C) seqüências do primeiro domínio da RC (391 caracteres) para todos indivíduos de *P. dominicana* e de *P. capitata*; D) seqüências do ND3 e do primeiro domínio da RC concatenadas, totalizando 742 caracteres para todos indivíduos de *P. dominicana* e de *P. capitata*. Alguns métodos de análise empregados não utilizam grupo externo, com isso, quando necessário, foram utilizadas as matrizes citadas excluindo-se as seqüências de *P. capitata* ou *P. coronata*. Não foram detectados códons de parada inesperados nas seqüências do ND3 e não há lacunas no alinhamento em nenhuma das matrizes.

Para a matriz A, a composição nucleotídica encontrada foi: A = 28,5%, C = 32,8%, G = 12,0%, T = 26,7%. Essa composição de bases é homogênea entre os táxons ($X^2 = 1,09$ [df = 153], $P = 1,00$). Dos 351 sítios, 321 não apresentam variação, 24 não são informativos para parcimônia e seis são informativos para parcimônia. Para esse conjunto de dados o modelo evolutivo selecionado foi o TrN + I, ou seja, o modelo de Tamura e Nei (1993) com proporção dos sítios invariáveis ($I = 0,7888$). Para a matriz B, a composição nucleotídica encontrada foi: A = 28,5%, C = 32,8%, G = 12,0%, T = 26,7%. Essa composição de bases é homogênea entre os táxons ($X^2 = 2,00$ [df = 153], $P = 1,00$). Dos 351 sítios, 312 não apresentam variação, 33 não são informativos para parcimônia e seis são informativos para a parcimônia. O modelo evolutivo selecionado foi o TrN + I com proporção dos sítios invariáveis ($I = 0,7895$). Excluindo-se a seqüência de *P. capitata*, selecionou-se o mesmo modelo evolutivo, porém, com outro valor de proporção de sítios invariáveis ($I = 0,8780$). Para a matriz C, a composição nucleotídica encontrada foi: A = 30,0%, C = 34,8%, G = 14,4%, T = 20,8%. Essa composição de bases é homogênea entre os táxons ($X^2 = 1,06$ [df = 153], $P = 1,00$). Dos 391 sítios, 370 não apresentam variação, 16 não são informativos para parcimônia e

cinco são informativos para a parcimônia. O modelo evolutivo selecionado foi o TrN + I com proporção dos sítios invariáveis ($I = 0,9135$). Excluindo-se a sequência de *P. capitata*, selecionou-se o modelo evolutivo HKY + I, ou seja, o modelo de Hasegawa *et al.* (1985) com proporção dos sítios invariáveis ($I = 0,9045$). Para a matriz D, a composição nucleotídica encontrada foi: A = 26,3%, C = 33,9%, G = 13,2%, T = 23,6%. Essa composição de bases é homogênea entre os táxons ($\chi^2 = 0,69$ [df = 153], $P = 1,00$). Dos 742 sítios, 682 não apresentam variação, 49 não são informativos para parcimônia e 11 são informativos para a parcimônia. O modelo evolutivo selecionado foi o TrN + I com proporção dos sítios invariáveis ($I = 0,8662$). Excluindo-se a sequência de *P. capitata*, selecionou-se o modelo evolutivo HKY + I com proporção dos sítios invariáveis ($I = 0,9431$). A composição nucleotídica encontrada em todas as matrizes de sequências é semelhante à encontrada em outras espécies de Oscines (Marshall & Baker, 1997; Pérez-Emán, 2005).

O número de haplótipos (h), os valores de diversidade nucleotídica (π) e dos índices de neutralidade calculados e suas significâncias estão resumidos na Tabela 3.2. Verificou-se que a variabilidade genética é baixa em todos os conjuntos de dados. Para a matriz baseada em sequências do ND3 apenas de indivíduos de *P. dominicana*, o teste F_s apresentou valor marginalmente significativo (-4,504). Para a matriz A, os testes DH e DHEW apresentaram valores marginalmente significativos. Para a matriz B o teste DHEW apresentou valor significativo e os testes DH e H-normalizado (-1,257), marginalmente significativos. Para a matriz C, os testes D_T (-1,814), F_s (-5,378), H-normalizado (-5,526), E (3,283), DH e DHEW apresentaram valores significativos. Para a matriz D os testes F_s (-9,681), H-normalizado (-3,688), E (2,065), DH e DHEW apresentaram valores significativos e o teste D_T (-1,644), marginalmente significativo. O teste S modificado não apresentou valor significativo para nenhuma das matrizes e seus valores foram próximos a 1 para todas as matrizes. Os testes D_{FL} e F também não apresentaram valores significativos para nenhuma das matrizes e seus valores foram negativos para todas as matrizes.

Tabela 3.2: Número de haplótipos (h) e valores de diversidade nucleotídica (π) e dos índices de neutralidade D_T (Tajima, 1989), F_s (Fu, 1997), S modificado (Fu, 1996), D_{FL} , F (Fu & Li, 1993), H-normalizado, E, DH (Zeng *et al.*, 2006) e DHEW (Zeng *et al.*, 2007b) para sete matrizes de seqüências (ND3: seqüências de ND3 de todos indivíduos de *P. dominicana*; ND3-cap: seqüências de ND3 de todos indivíduos de *P. dominicana* e de *P. capitata*; ND3-cor: seqüências de ND3 de todos indivíduos de *P. dominicana* e de *P. coronata*; RC: seqüências do primeiro domínio da RC de todos indivíduos de *P. dominicana*; RC-cap: seqüências da RC de todos indivíduos de *P. dominicana* e de *P. capitata*; Concatenado: seqüências concatenadas das matrizes RC e ND3; Concatenado-cap: seqüências concatenadas das matrizes RC-cap e ND3-cap). Nível de significância dos índices: ***: $P < 0,01$; **: $0,01 < P < 0,05$; *: $0,05 < P < 0,10$; sem asterisco: $P > 0,10$; –: índice não se aplica. Entre parênteses estão os tamanhos das amostras.

Índice \ Matriz	ND3 (51)	ND3-cap (52)	ND3-cor (52)	RC (51)	RC-cap (52)	Concatenado (51)	Concatenado-cap (52)
h	11	-	-	8	-	17	-
π	0,0041	-	-	0,0015	-	0,0027	-
D_T	-1,200	-	-	-1,814**	-	-1,644*	-
F_s	-4,504**	-	-	-5,378***	-	-9,681***	-
S modificado	0,997	-	-	0,999	-	1,000	-
D_{FL}	-	-0,701	-1,165	-	-0,364	-	-0,691
F	-	-0,913	-1,347	-	-0,948	-	-1,145
H normalizado	-	-1,257*	-0,218	-	-5,526***	-	-3,688***
E	-	0,364	-0,640	-	3,283***	-	2,065**
DH	-	*	*	-	***	-	**
DHEW	-	**	*	-	***	-	***

A partir das matrizes A e B obteve-se gráficos de saturação (Figura 3.4). Há um sinal de saturação para a matriz B, pois a curva começa a se estabilizar com o aumento da distância genética. As distâncias genéticas foram corrigidas segundo o modelo de Tamura e Nei (1993). O objetivo do cálculo dos gráficos de saturação no presente estudo foi testar o efeito da utilização de diferentes grupos externos no conjunto de dados. Portanto, não foram obtidos gráficos de saturação a partir das matrizes C e D.

Figura 3.4: Gráficos de saturação para as matrizes A (a) e B (b).

Os padrões de distribuição de diferenças nucleotídicas par-a-par (*mismatch distribution*; Figura 3.5) foram obtidos a partir das matrizes A (ND3), C (RC) e D (conjunto de dados concatenados) excluindo-se as seqüências do grupo externo. Todas as matrizes apresentaram padrões semelhantes aos esperados segundo o modelo de expansão populacional recente.

Figura 3.5: Padrões de distribuição de diferenças nucleotídicas par-a-par (*mismatch distribution*) baseados no ND3 (a), no primeiro domínio da RC (b) e no conjunto de dados concatenados (c).

Foram obtidas filogenias semelhantes baseadas nos quatro conjuntos de dados utilizando-se os métodos de máxima parcimônia e máxima verossimilhança (Figura 3.6). As topologias revelam ausência de estrutura populacional e os nós não apresentam bons valores de apoio.

Figura 3.6: Árvores obtidas pelo método de máxima verossimilhança. a) e b) baseadas em 351 caracteres do ND3 utilizando-se, respectivamente, *P. coronata* (L1935) e *P. capitata* (L2297) como grupos externos. c) baseada em 391 caracteres do primeiro domínio da RC. d) 742 caracteres concatenados do ND3 e do primeiro domínio da RC. Os números nos nós indicam os valores de *bootstrap* para análises de máxima verossimilhança e máxima parcimônia, respectivamente. Valores abaixo de 70% não estão apresentados.

As redes de haplótipos foram estimadas sem os grupos externos para todas as matrizes (Figura 3.7). Nessas redes, observa-se que o haplótipo central foi amostrado em todas as localidades, exceto Juremal/BA nas matrizes baseadas no ND3 e no conjunto de dados concatenados e RPPN Serra das Almas/CE nas matrizes baseadas no primeiro domínio da RC e no conjunto de dados concatenados. Além disso, o haplótipo central apresenta alta frequência e é cercado por alguns haplótipos raros, indicando ausência de estrutura populacional, assim como as filogenias.

Figura 3.7: Redes de haplótipos baseadas no ND3 (a), no primeiro domínio da RC (b) e no conjunto de dados concatenados (c). O tamanho dos círculos é proporcional à frequência do haplótipo. Losangos pretos indicam haplótipos intermediários não amostrados. As cores indicam as localidades dos haplótipos: PN Serra das Confusões/PI (marrom), cativoiro em Tanquinho/BA (azul-claro), Morro do Chapéu/BA (roxo), Bonito/BA (verde-claro), Tanquinho/BA (cinza), Beira-Rio/BA (azul), Parnaguá/PI (verde), Simpício Mendes/PI (amarelo), Vila Nova do Piauí/PI (verde-oliva), Trindade/PE (vermelho), Juremal/BA (branco), Itiúba/BA (laranja) e RPPN Serra das Almas/CE (rosa).

As relações entre as distâncias geográficas e genéticas estão apresentadas na Figura 3.8. As distâncias genéticas para as matrizes do primeiro domínio da RC e para o conjunto de dados concatenados foram corrigidas segundo o modelo de evolução de Tajima e Nei (1982), pois este é, dentre os modelos presentes no programa MEGA v. 4, o mais próximo daquele selecionado pelo Modeltest 3.7. Para a matriz do ND3, as distâncias genéticas foram corrigidas segundo o modelo de Tamura & Nei (1993). O teste de Mantel indica ausência de correlação entre as distâncias geográficas e genéticas (ND3: $r = -0,002$ e $P = 0,955$; RC: $r = 0,027$ e $P = 0,378$; conjunto de dados concatenados: $r = 0,003$ e $P = 0,924$), ou seja, não há indícios de isolamento por distância para *P. dominicana* baseado em dados do DNAmT.

Figura 3.8: Relação entre as distâncias geográficas e genéticas baseadas no ND3 (a), no primeiro domínio da RC (b) e no conjunto de dados concatenados (c).

Para os cálculos da AMOVA realizaram-se 1000 permutações e utilizou-se o modelo evolutivo de Tamura e Nei (1993) para a matriz do ND3 e de Tajima e Nei (1982) para as matrizes do primeiro domínio da RC e do conjunto de dados concatenados. Estes são, dentre os modelos

evolutivos presentes no programa Arlequin v. 3.11, os mais próximos daqueles selecionados pelo Modeltest 3.7. Para não comprometer os cálculos da variação intra-localidade, as localidades com apenas um indivíduo amostrado (PN Serra das Confusões/PI, RPPN Serra das Almas/CE, Vila Nova do Piauí/PI e Juremal/BA) foram retiradas dessas análises. Os três modelos de estrutura populacional propostos foram testados: modelo do Rio São Francisco (dois grupos de localidades Tanquinho/BA + Beira Rio/BA + Itiúba/BA X Parnaguá/PI + Simplicio Mendes/PI + Trindade/PE), modelo das Caatingas (dois grupos de localidades Beira Rio/BA + Trindade/PE X Tanquinho/BA + Itiúba/BA + Parnaguá/PI + Simplicio Mendes/PI) e modelo das Regiões (três grupos de localidades Beira Rio/BA + Trindade/PE X Tanquinho/BA + Itiúba/BA X Parnaguá/PI + Simplicio Mendes/PI). Os resultados do AMOVA para as três matrizes de dados estão apresentados na Tabela 3.3. Verificou-se que a maior parte da variação observada é atribuída a divergências intra-localidades (maiores porcentagens de variação e valores significativos de Φ_{ST} para todos os modelos e todas as matrizes). Tais resultados indicam que a espécie *P. dominicana* não apresenta estrutura populacional segundo nenhum dos modelos de estrutura propostos (porcentagens de variação atribuídas a divergências entre grupos são baixas e nenhum dos valores de Φ_{CT} é significativo).

Tabela 3.3: Resultados da AMOVA para três modelos de estrutura populacional para as três matrizes de dados (ND3, RC, Conc.: conjunto de dados concatenados). IL: porcentagem da variação atribuída a divergências intra-localidades; IG: porcentagem da variação atribuída a divergências intra-grupos; EG: porcentagem da variação atribuída a divergências entre grupos. Os cálculos de variação Φ_{ST} , Φ_{SC} e Φ_{CT} são referentes às divergências intra-localidades, intra-grupos e entre grupos, respectivamente. *: $0,05 < P < 0,10$; **: $P < 0,05$; sem asterisco: $P > 0,10$.

Modelo	Matriz	IL	IG	EG	Φ_{ST}	Φ_{SC}	Φ_{CT}
Rio São Francisco	ND3	75,3%	22,5%	2,2%	0,2468**	0,2303**	0,0214
	RC	90,3%	11,9%	-2,2%	0,0970**	0,1160**	-0,0216
	Conc.	79,3%	19,7%	1,0%	0,2066**	0,1985**	0,0101
Caatingas	ND3	70,1%	12,4%	17,5%	0,2986**	0,1500*	0,1748
	RC	88,0%	8,1%	3,9%	0,1198**	0,0843*	0,0387
	Conc.	74,8%	11,3%	13,9%	0,2521**	0,1310*	0,1395
Regiões	ND3	75,0%	18,3%	6,7%	0,2506**	0,1966*	0,0672
	RC	90,0%	12,4%	-2,4%	0,1002**	0,1215*	-0,0243
	Conc.	79,0%	16,7%	4,3%	0,2103**	0,1750*	0,0428

Para os cálculos da SAMOVA, as localidades com apenas um indivíduo amostrado (PN Serra das Confusões/PI, RPPN Serra das Almas/CE, Vila Nova do Piauí/PI e Juremal/BA) também foram retiradas para não comprometer os cálculos da variação intra-populacional. Os resultados da SAMOVA estão apresentados na Tabela 3.4. Verificou-se que a hipótese de dois grupos (número

mínimo inicial) seria a melhor configuração para esses dados. Os valores de variação entre grupos foram mais elevados do que os encontrados utilizando os modelos propostos anteriormente (Tabela 3.3). Apesar dos valores altos da porcentagem da variação entre grupos, os valores de Φ_{CT} não foram significativos e a maior parte da variação observada também é atribuída a divergências intra-localidades (maiores porcentagens de variação e valores significativos de Φ_{ST} para todas as estruturas populacionais propostas e para todas as matrizes). Além disso, as estruturas populacionais sugeridas pela SAMOVA não parecem ser congruentes com aspectos da geografia ou fitogeografia da Caatinga. O cálculo da variação intra-grupo foi comprometido por haver grupos formados por apenas uma localidade (BR, Tr e SM). Com isso, a porcentagem da variação intra-grupo se tornou negativa.

Tabela 3.4: Resultados da SAMOVA para dois e três grupos (K) para as três matrizes de dados (ND3, RC, Conc.: conjunto de dados concatenados). BR: Beira Rio/BA; SM: Simplicio Mendes/PI; Tr: Trindade/PE; Outras: todas as outras localidades. IL: porcentagem da variação atribuída a divergências intra-localidades; IG: porcentagem da variação atribuída a divergências intra-grupos; EG: porcentagem da variação atribuída a divergências entre grupos. Os cálculos de variação Φ_{ST} , Φ_{SC} e Φ_{CT} são referentes às divergências intra-localidades, intra-grupos e entre grupos, respectivamente. *: $0,05 < P < 0,10$; **: $P < 0,05$; sem asterisco: $P > 0,10$.

K	Matriz	Estrutura	IL	IG	EG	Φ_{ST}	Φ_{SC}	Φ_{CT}
2	ND3	BR X Outras	53,6%	-1,2%	47,6%	0,4642**	-0,0235**	0,4766
	RC	BR X Outras	68,2%	-6,5%	38,3%	0,3175**	-0,1060*	0,3830
	Conc.	BR X Outras	57,3%	-2,6%	45,3%	0,4268**	-0,0473**	0,4527
3	ND3	BR X Tr X Outras	65,4%	-4,4%	39,0%	0,3458**	-0,0725**	0,3900*
	RC	BR X SM X Outras	79,4%	-11,0%	31,6%	0,2063**	-0,1607	0,3162*
	Conc.	BR X Tr X Outras	69,5%	-5,1%	35,6%	0,3048**	-0,0800**	0,3563

3.4. DISCUSSÃO

Para analisar o padrão de polimorfismos encontrado pelos testes de neutralidade foram utilizadas a avaliação do poder desses testes resumida na Tabela 1.1, a teoria que embasa cada teste e a filogenia proposta no Capítulo 2. Ao comparar os resultados obtidos com as matrizes A e B (Tabela 3.2) verifica-se que a utilização de *Paroaria capitata* (mais distante) como grupo externo ao invés de *P. coronata* (grupo irmão de *P. dominicana*) resultou em uma mudança na frequência das mutações, causando um aumento das mutações abundantes. O que pode ser observado pela

mudança no sinal do índice E (Zeng *et al.*, 2006) e pela variação do índice H-normalizado (Zeng *et al.*, 2006). Isso pode ocorrer devido a uma mutação rara se tornar abundante em razão do uso de um grupo externo inadequado que causaria a troca do nucleotídeo ancestral e uma mutação abundante é “criada” ao introduzir uma mutação inter-específica diferente. Ou seja, o modelo dos sítios infinitos (assumido para todos os testes) foi violado devido à saturação causada pela não utilização do grupo externo filogeneticamente mais próximo. Essa hipótese é confirmada pelos gráficos de saturação obtidos para as matrizes A e B (Figura 3.4). Com isso, a análise dos testes realizada com as matrizes ND3-*cap*, RC-*cap* e Concatenado-*cap*, indicando ação de seleção positiva (testes H-normalizado, E, DH [Zeng *et al.*, 2006] e DHEW [Zeng *et al.*, 2007b]), pode ser duvidosa e não pode ser considerada.

Os testes Fs (autor [Fu, 1997] recomenda utilizar $\alpha = 0,01$ para o teste Fs), DH (Zeng *et al.*, 2006) e DHEW (Zeng *et al.*, 2007b) aplicados às matrizes ND3 e ND3-*cor* (Tabela 3.2) apresentam valores marginalmente significativos. Ou seja, há indícios de desvios do modelo de populações em equilíbrio neutro devido a um excesso de mutações abundantes, o que indica uma possível ação da seleção positiva, porém, esses desvios não seriam significativos. Analisando-se a matriz RC, verificam-se desvios significativos do modelo de populações em equilíbrio neutro segundo os testes D_T (Tajima, 1989) e Fs. Como ambos os testes não utilizam grupo externo, não se pode determinar se o desvio é devido a um excesso de mutações raras ou abundantes. Portanto, não é possível inferir precisamente qual força atuou na geração dos polimorfismos encontrados nos dados, pois eles podem ser explicados por efeitos de seleção positiva (fases pré e pós-fixação) ou negativa atuando concomitantemente ou não a uma expansão populacional recente. Agrupando-se os dados do ND3 e da RC em uma única matriz (matriz Concatenado) verificam-se, basicamente, os mesmos resultados encontrados para a matriz RC, pois há desvios significativos do modelo de populações em equilíbrio neutro segundo o teste Fs e marginalmente significativo segundo o teste D_T .

Como o ND3 e a região controladora fazem parte da molécula de DNAm_t, que é uma unidade genealógica não recombinante (Avise, 1994), considera-se que ambos marcadores estão sujeitos aos mesmos efeitos (seletivos e/ou demográficos). Por exemplo, ambos teriam sua variabilidade reduzida por um evento de garrafa genética devido a uma mutação vantajosa em qualquer região do DNAm_t. Portanto, como os dados sugerem que a região controladora não pode ser considerada neutra, o ND3 também não poderia, mesmo apresentando desvios apenas marginalmente significativos. Além disso, esses desvios indicam que as mesmas forças parecem atuar nos dois marcadores. Com isso, nenhum dos marcadores poderia ser utilizado para estimar tamanho populacional efetivo e demografia histórica (Gillespie, 2004; Li, 1997). Assim, as curvas observadas de *mismatch distribution* (Figura 3.5) podem estar associadas tanto a efeitos seletivos como demográficos (Zink, 2005). Ou seja, análises dessas curvas são pouco informativas quanto

aos efeitos geradores dos polimorfismos encontrados, pois são sensíveis também a efeitos seletivos, assim como alguns testes de neutralidade. Isso pode ser observado ao comparar os resultados dos testes de neutralidade (Tabela 3.2) com as curvas da *mismatch distribution* (Figura 3.5). A matriz RC é a que apresenta os valores mais significativos para os testes de neutralidade e a *mismatch distribution* mais semelhante à esperada segundo o modelo de expansão populacional recente. Enquanto a matriz ND3 apresenta valores não significativo para o teste D_T e marginalmente significativo para o teste F_s e uma *mismatch distribution* menos semelhante à esperada segundo o mesmo modelo.

No presente trabalho não foi possível identificar uma única força determinante dos padrões de polimorfismos encontrados, pois eles poderiam ser explicados por efeitos de seleção positiva (fases pré e pós-fixação) ou seleção negativa atuando concomitantemente ou não a uma expansão populacional recente. Porém, esse problema poderia ser resolvido pela utilização de *P. coronata* como grupo externo nas matrizes C e D, pois é a espécie filogeneticamente mais próxima à *P. dominicana* (Capítulo 2) e, com isso, o problema da saturação dos dados poderia ser resolvido. Tal resultado ressalta a importância da seleção adequada do grupo externo para estudos filogeográficos (Fu & Li, 1993).

Apesar da importância do pressuposto da neutralidade dos dados (Gillespie, 2004; Li, 1997) nos estudos filogeográficos, os esforços para testá-lo são esporádicos ou superficiais. Um resultado comum nos estudos com aves Oscines é a ausência de estrutura populacional, baixa variabilidade genética e os valores significativos e negativos de D_T (Zink, 2005). Isso vem sendo interpretado como um sinal de expansão populacional recente sem considerar os possíveis efeitos da seleção natural, mesmo que esse padrão de polimorfismos também possa ser causado por efeitos seletivos, como seleção negativa ou positiva. Com isso, o conhecimento atual sobre a demografia histórica de alguns táxons pode estar comprometido. Porém, a análise da estrutura populacional não é afetada pelos efeitos da seleção natural (Zink, 2005).

A ausência de estrutura populacional encontrada no presente trabalho poderia ser interpretada como uma expansão populacional recente, como sugerido por Zink *et al.* (2001). Os resultados dos testes de neutralidade (Tabela 3.2) indicam que a expansão populacional recente é um dos possíveis efeitos causadores dos polimorfismos encontrados. Uma hipótese para explicar essa expansão de *P. dominicana* seria a recente colonização da Caatinga seguida de expansão populacional. Porém, as datações moleculares obtidas sugerem que *P. dominicana* e sua espécie irmã (*P. coronata*) se diversificaram entre 3,4 e 2,7 M.a. (Capítulo 2), invalidando a hipótese de colonização recente da Caatinga. Outra hipótese seria que a expansão populacional de *P. dominicana* estaria associada à expansão desse bioma, uma vez que essa é uma espécie de ave associada à Caatinga. Wang *et al.* (2004), baseados em estudos de clima e de deposição geológica

em uma caverna no centro da Caatinga, concluíram que nos últimos 210 mil anos o nordeste do Brasil passou por muitos períodos de alternância pluvial que não coincidem com os longos ciclos glaciais e inter-glaciais. Durante os períodos de maior pluviosidade, a Caatinga se retrairia, enquanto as florestas úmidas se expandiriam. O último período de aumento da pluviosidade no nordeste do Brasil é datado há 15 mil anos atrás, aproximadamente. Desde então, essa região encontra-se num período de menor pluviosidade, com conseqüente expansão da Caatinga (Wang *et al.*, 2004). Portanto, como *P. dominicana* é uma espécie de ave associada à Caatinga, é plausível aceitar que essa espécie tenha sofrido uma expansão populacional recente. Seria interessante realizar uma datação molecular dessa possível expansão para verificar a congruência das datas. Porém, isso não é possível no presente estudo porque não há taxas evolutivas calculadas para o ND3 e para a região controladora para grupos próximos à *P. dominicana*. Além disso, para realizar essas datações é importante que os resultados dos testes de neutralidade descartem efeitos seletivos como possíveis causadores dos polimorfismos encontrados.

Essa ausência de estrutura populacional encontrada para o DNAmT da espécie *P. dominicana* poderia indicar que as fêmeas dispersam, pois ao analisar o DNAmT, observa-se apenas a variabilidade associada à herança materna (Avise, 1995). Com isso, uma possível estruturação dos machos não seria detectada. Infelizmente, devido à ausência de dados com um marcador nuclear não é possível levantar nenhuma hipótese em relação aos machos e nem testar a hipótese de dispersão diferencial dos sexos. Dados da literatura indicam que os machos seriam sedentários apenas na época reprodutiva (Sick, 2001), porém, não se sabe se eles formam os mesmos territórios em períodos reprodutivos subseqüentes, ressaltando a importância dos estudos de campo.

A ausência de estrutura populacional também poderia ser causada por influência humana, uma vez que *P. dominicana* sofre intensa captura pela população local e pelo tráfico ilegal de aves (Rocha *et al.*, 2006; Sick, 2001). Indivíduos poderiam ter sido soltos em locais distintos de onde foram capturados, causando uma migração artificial que, em larga escala, impossibilitaria a diferenciação populacional dessa espécie devido ao alto fluxo gênico. Porém, como os machos de *P. dominicana* são territorialistas durante a época reprodutiva, a captura ocorre principalmente nesse período pelo método de gaiola de campo, no qual um macho é mantido na gaiola enquanto canta quase que ininterruptamente de modo que outros machos são atraídos à armadilha ao tentar defender seus territórios. Com isso, a captura de machos dessa espécie pelo tráfico é maior do que a de fêmeas (Rocha *et al.*, 2006), de modo que a possível migração artificial teria maior proporção de machos. Portanto, o tráfico não deve ter influência sobre a ausência de estrutura encontrada neste estudo, pois essa está baseada em dados do DNAmT, que possui herança materna (Avise, 1995). Há, ainda, a possibilidade de os machos coletados no presente estudo serem indivíduos recentemente libertados do tráfico. O que causaria os mesmos efeitos da migração artificial discutida

anteriormente. Porém, o presente estudo é baseado em coletas indiscriminadas de fêmeas, machos e jovens. Portanto, é pouco provável que a amostragem do presente estudo seja dominada por indivíduos machos. De qualquer modo, será realizada a sexagem molecular para confirmar a presença dos dois sexos em proporções semelhantes. Além disso, parece ser pouco provável que todos os machos coletados sejam provenientes de solturas do tráfico. Com isso, caso tenham sido coletados machos libertados pelo tráfico no presente estudo, sua proporção deve ser tão baixa, que não teria influência sobre a ausência de estrutura encontrada neste estudo.

Outra possível interpretação para a ausência de estrutura populacional encontrada é que o rio São Francisco e os tipos de Caatinga não constituem, respectivamente, barreiras geográfica e ecológica para essa espécie. Ou seja, os indivíduos dessa espécie se dispersariam livremente por toda a Caatinga, não sendo possível diferenciar geneticamente indivíduos provenientes de locais distintos desse bioma. Pode-se dizer que essa observação seja a esperada devido à relativa grande capacidade de vôo e à distribuição contínua de *P. dominicana* (Awise, 1994; Sick, 2001; Wakeley, 2005). Com isso, segundo a definição de Unidades de Manejo (UM) proposta por Moritz (1994), essa espécie poderia ser tratada como uma UM. No entanto, é importante ter dados nucleares para confirmar esse resultado.

Devido ao presente estudo ser pioneiro na análise da estrutura populacional de uma espécie de ave associada à Caatinga, comparações com outros estudos são limitadas. Braggio e Bonvicino (2004) realizaram um estudo filogenético do gênero *Thrichomys* (pequenos roedores de áreas abertas) baseado em seqüências do citocromo b. Mesmo que esse estudo filogenético não teve como objetivo analisar estrutura populacional, a espécie *T. aperoides* (que apresenta distribuição geográfica semelhante à de *P. dominicana*) foi amostrada ao longo de toda a Caatinga. As análises das relações filogenéticas obtidas entre os indivíduos dessa espécie mostram apoio alto para o agrupamento de indivíduos de localidades distantes, o que indica ausência de estrutura populacional para *T. aperoides*. A comparação entre esse resultado e os obtidos no presente estudo parece indicar uma tendência à ausência de estrutura populacional em animais associados à Caatinga. Porém, a correlação entre esses dois estudos é limitada e deve ser interpretada com cautela. Estudos populacionais com outros táxons associados à Caatinga são necessários para compreender a dinâmica desse bioma e auxiliar na sua conservação.

3.5. REFERÊNCIAS BIBLIOGRÁFICAS

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automatic Control* **19**: 716-723.

- Andrade-Lima, D. 1981. The Caatinga dominium. *Rev.Bras. Bot.* **4**: 149-153.
- Assis, C.; Toledo, C. B.; Romaniuc-Neto, S.; Cordeiro, I. 1994. *Caatinga (Nossas Plantas)*. Ed. FTD, São Paulo.
- Avise, J. C. 1994. *Molecular Markers, Natural History and Evolution*. Ed. Chapman & Hall, Nova York.
- Avise, J. C. 1995. Mitochondrial DNA polymorphism and a connection between genetics and demography of relevance to conservation. *Conserv. Biol.* **9**: 686-690.
- Avise, J. C. 2000. *Phylogeography: the History and Formation of Species*. Ed. Harvard University Press, Londres.
- Bowen, B. W. 1999. Preserving genes, species, or ecosystems? Healing the fractured foundations of conservation policy. *Mol. Ecol.* **8**: S5-S10.
- Braggio, E.; Bonvicino, C. R. 2004. Molecular divergence in the genus *Thrichomys* (Rodentia, Echimyidae). *J. Mammal.* **85**: 316-320.
- Castelletti, C. H. M.; Santos, A. M. M.; Tabarelli, M.; Silva, J. M. C. 2003. Quanto ainda resta da Caatinga? Uma estimativa preliminar. In: Leal, I. R.; Tabarelli, M.; Silva, J. M. C. (ed.) *Ecologia e Conservação da Caatinga*. Ed. Universitária da UFPE, Recife. Pp. 719-734.
- Chesser, R. T. 1999. Molecular phylogenetics of the rhinocryptid genus *Pteroptochos*. *Condor* **101**: 439-445.
- Cracraft, J. 1985. Historical biogeography and patterns of differentiation within the South American avifauna: areas of endemism. *Ornithol. Monog.* **36**: 49-84.
- Crozier, R. H. 1992. Genetic diversity and the agony of choice. *Biol. Conserv.* **61**: 11-15.
- Desjardins, P.; Morais, R. 1990. Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *J. Mol. Biol.* **212**: 599-634.
- Dupanloup, I.; Schneider, S.; Excoffier, L. 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* **11**: 2571-2581.
- Eberhard, J. D.; Bermingham, E. 2004. Phylogeny and biogeography of the *Amazona ochrocephala* (Aves: Psittacidae) complex. *Auk* **121**: 318-332.
- EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. 1994. *Atlas do Meio Ambiente do Brasil*. Ed. Terra Viva, Brasília.
- Excoffier, L.; Laval, G.; Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**: 47-50.
- Excoffier, L.; Smouse, P.; Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.

- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **79**: 783-791.
- Fu, Y.-X. 1996. New statistical tests of neutrality for DNA samples from a population. *Genetics* **143**: 557-570.
- Fu, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915-925.
- Fu, Y.-X.; Li, W.-H. 1993. Statistical tests of neutrality of mutations. *Genetics* **133**: 693-709.
- Gillespie, J. H. 2004. *Population Genetics: a Concise Guide*. The Johns Hopkins University Press, Baltimore e Londres.
- Griffiths, C. S. 1997. Correlation of functional domains and rates of nucleotide substitution in cytochrome b. *Mol. Phylogenet. Evol.* **7**: 353-365.
- Guindon, S.; Lethiec, F.; Duroux, P.; Gascuel, O. 2005. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* **33** (Web Server issue): W557-W559.
- Haffer, J. 1985. Avian zoogeography of the Neotropical lowland. *Ornithol. Monog.* **36**: 113-146.
- Hasegawa, M.; Kishino, K.; Yano, T. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**: 160-174.
- Hecker, K. H.; Roux, K. H. 1996. High and low annealing temperatures increase both specificity and yield in “Touchdown” and “StepDown” PCR. *BioTechniques* **20**: 478-485.
- Kvist, L.; Martens, J.; Ahola, A.; Orell, M. 2001. Phylogeography of a Palearctic sedentary passerine, the willow tit (*Parus montanus*). *J. Evol. Biol.* **14**: 930-941.
- Kvist, L.; Ruokonen, M.; Thessing, A.; Lumme, J.; Orell, M. 1998. Mitochondrial control region polymorphism reveal high amount of gene flow in Fennoscandian willow tit (*Parus montanus borealis*). *Hereditas* **128**: 133-143.
- Li, H.-W. 1997. *Molecular Evolution*. Ed. Sinauer Associates, Sunderland.
- Major, I.; Sales Jr.; L. G.; Castro, R. 2004. *Aves da Caatinga – Birds of the Caatinga*. Ed. Demócrito Rocha, Fortaleza.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**: 209-220.
- Marshall, H. D.; Baker, A. J. 1997. Structural conservation and variation in the mitochondrial control region of fringillinae finches (*Fringilla* spp.) and the greenfinch (*Carduelis chloris*). *Mol. Biol. Evol.* **14**: 173-184.
- Meyer, A. 1994. Shortcomings of the cytochrome b gene as a molecular marker. *Trends Ecol. Evol.* **9**: 278-280.

- MMA – Ministério do Meio Ambiente. 2006. *Plano de Ações Estratégicas e Integradas para o Desenvolvimento do Turismo Sustentável na Bacia do Rio São Francisco*. Ministério do Meio Ambiente, Secretaria Executiva. Brasília.
- Moffat, A. S. 2002. South American landscapes: ancient and modern. *Science* **296**: 1959-1960.
- Moritz, C. 1994. Defining “Evolutionary Significant Units” for conservation. *Trends Ecol. Evol.* **9**: 373-375.
- Moritz, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* **51**: 238-254.
- Moritz, C.; Faith, D. 1998. Comparative phylogeography and the identification of genetically divergent areas for conservation. *Mol. Ecol.* **7**: 419-429.
- Page, R. D. M.; Holmes, E. C. 1998. *Molecular Evolution: A Phylogenetic Approach*. Ed. Blackwell Science, Oxford.
- Pérez-Emán, J. L. 2005. Molecular phylogenetics and biogeography of the Neotropical redstarts (*Myioborus*; Aves, Parulinae). *Mol. Phylogenet. Evol.* **37**: 511-528.
- Polzin, T.; Daneschmand, S. V. 2003. On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* **31**: 12-20.
- Posada, D.; Crandall, K. A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.
- Prado, D. E. 2003. As Caatingas da América do Sul. In: Leal, I. R.; Tabarelli, M.; Silva, J. M. C. (ed.) *Ecologia e Conservação da Caatinga*. Ed. Universitária da UFPE, Recife. Pp. 3-74.
- Reis, A. C. S. 1976. Clima da Caatinga. *An. Acad. Brasil. Ciênc.* **48**: 325-335.
- Rocha, M. S. P.; Cavalcanti, P. C. M.; Sousa, R. L.; Alves, R. R. N. 2006. Aspectos da comercialização ilegal de aves nas feiras livres de Campina Grande, Paraíba, Brasil. *Rev. Biol. Ciênc. Terra* **6**: 204-221.
- Rogers, A. R.; Harpending, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**: 552-569.
- Rozas, J.; Sánchez-DelBarrio, J. C.; Messeguer, X.; Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496-2497.
- Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Ehrlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.
- Sick, H. 2001. *Ornitologia Brasileira*. 3ª edição. Ed. Nova Fronteira, Rio de Janeiro.
- Silva, J. M. C.; Souza, M. A.; Bieber, A. G. D.; Carlos, C. J. 2003. Aves da Caatinga: status, uso do habitat e sensibilidade. In: Leal, I. R.; Tabarelli, M.; Silva, J. M. C. (ed.) *Ecologia e Conservação da Caatinga*. Ed. Universitária da UFPE, Recife. Pp. 237-274.

- Smith, A. B. 1994. Rooting molecular trees: problems and strategies. *Biol. J. Linnean Soc.* **51**: 279-292.
- Sorenson, M. D.; Ast, J. A.; Dimcheff, D. E.; Yuri, T.; 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.
- Swofford, D. L. 2002. PAUP*: *Phylogenetic Analysis Using Parsimony (* and other methods)*, version 4. Sinauer Associates, Sunderland.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.
- Tajima, F.; Nei, M. 1982. Biases of the estimates of DNA divergence obtained by the restriction enzyme technique. *J. Mol. Evol.* **18**: 823-833.
- Tamura, K.; Dudley, J.; Nei, M.; Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**: 1596-1599.
- Tamura, K.; Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512-526.
- Wakeley, J. 2005. *Coalescent Theory: an Introduction*. Roberts & Company Publishers. Greenwood Village.
- Wang, X.; Auler, A. S.; Lawrence, E. R.; Cheng, H.; Cristalli, P. S.; Smart, P. L.; Richards, D. A.; Shen, C.-C. 2004. Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. *Nature* **432**: 740-743.
- Xia, X.; Xie, Z. 2001. DAMBE: Data analysis in molecular biology and evolution. *J. Hered.* **92**: 371-373.
- Zeng, K.; Fu, Y.-X.; Shi, S.; Wu, C.-I. 2006. Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* **174**: 1431-1439.
- Zeng, K.; Shi, S.; Wu, C.-I. 2007b. Compound tests for the detection of hitchhiking under positive selection. *Mol. Biol. Evol.* **24**: 1898-1908.
- Zink, R. M. 2005. Natural selection on mitochondrial DNA in *Parus* and its relevance for phylogeographic studies. *Proc. R. Soc. B* **272**: 71-78.
- Zink, R. M.; Drovetski, S. V.; Rohwer, S. 2002. Phylogeographic patterns in the great spotted woodpecker (*Dendrocopos major*) across Eurasia. *J. Avian Biol.* **33**: 175-178.
- Zink, R. M.; Kessen, A. E.; Line, T. V.; Blackwell-Rago, R. C. 2001. Comparative phylogeography of some aridland bird species. *Condor* **103**: 1-10.
- Zink, R. M.; Rising, J. D.; Mockford, S.; Horn, A. G.; Wright, J. M.; Leonard, M.; Westberg, M. C. 2005. Mitochondrial DNA variation, species limits, and rapid evolution of plumage coloration and size in the savannah sparrow. *Condor* **107**: 21-28.

CAPÍTULO 4

Considerações Finais

Pelo presente estudo percebe-se a grande influência que uma filogenia pode ter em um estudo filogeográfico, pois é somente com um estudo filogenético prévio que se pode inferir se a espécie alvo de um estudo filogeográfico é monofilética e, além disso, pode-se selecionar o grupo externo mais adequado para a análise filogeográfica. No Capítulo 3, verifica-se que a falta do grupo externo mais proximamente relacionado ao grupo interno pode prejudicar a análise chegando a conclusões errôneas sobre as forças que determinam os padrões de polimorfismos observados em um estudo filogeográfico, comprometendo, assim, análises de demografia histórica, de tamanho populacional efetivo, de taxa de fluxo gênico e de épocas de colonização ou de separação.

Os resultados da análise filogenética do gênero *Paroaria* apresentada no Capítulo 2 permitiram concluir que:

- Foi encontrada mais uma evidência de que a atual composição das tribos da subfamília Emberizinae necessita de revisão.
- *Gubernatrix cristata* parece pertencer à tribo Thraupini.
- *G. cristata* parece ser, dentre as espécies analisadas, a mais próxima filogeneticamente do gênero *Paroaria*.
- Há evidências de que o gênero *Paroaria* seja monofilético.
- A tradicional divisão do gênero *Paroaria* em dois grupos: 1) *P. coronata* e *P. dominicana* e 2) *P. baeri*, *P. gularis* e *P. capitata* baseada em caracteres morfológicos e ecológicos (Sick, 2001), foi corroborada.
- *P. coronata* e *P. capitata* parecem não ocorrer no Chaco *sensu* Prado (2000), mas nos seus arredores (províncias Paranaense e da Floresta de Pedemonte Subandino *sensu* Prado; 2000).
- O clado (*P. coronata*, *P. dominicana*) sugere que as florestas abertas da Caatinga e dos arredores do Chaco estão relacionadas.
- O clado (*P. baeri* (*P. gularis*, *P. capitata*)) sugere que as matas ciliares do Cerrado, da Amazônia, do Pantanal e dos arredores do Chaco *sensu* Prado (2000) estão relacionadas.
- As diversificações mais antiga e mais recente parecem ser congruentes com a hipótese dos Refúgios (Haffer, 1997); enquanto as outras diversificações, com a hipótese Paleogeográfica (Emsley, 1965).

- Baseado na distribuição obtida no mapa pontual supõe-se que a espécie *P. gularis* apresente o seguinte padrão filogeográfico: uma população no nordeste e uma no sudoeste amazônico, sendo a última ancestral.
- Baseado na distribuição obtida no mapa pontual supõe-se que a espécie *P. capitata* apresente o seguinte padrão filogeográfico: uma população associada à província da Floresta de Pedemonte Subandino *sensu* Prado (2000) e outra ao oeste da província Paranaense *sensu* Prado (2000) e ao Pantanal, de modo que a última seria a população ancestral.
- Baseado na distribuição obtida no mapa pontual supõe-se que a espécie *P. coronata* apresente o seguinte padrão filogeográfico: uma população associada à província da Floresta de Pedemonte Subandino *sensu* Prado (2000) e outra ao oeste da província Paranaense *sensu* Prado (2000), ao Pantanal e aos Pampas.
- Estudos filogeográficos das espécies e uma filogenia com maior amostragem de marcadores e de todas as subespécies do gênero *Paroaria* seriam importantes para testar as hipóteses levantadas neste trabalho, bem como para auxiliar na compreensão da biogeografia histórica da América do Sul.

Os resultados do estudo filogeográfico de *Paroaria dominicana* apresentados no Capítulo 3 indicam que:

- O padrão de polimorfismos encontrado pode ser causado por seleção positiva (fases pré ou pós-fixação) ou seleção negativa atuando concomitantemente ou não a uma expansão populacional recente.
- O rio São Francisco e os tipos de Caatinga não constituem, respectivamente, barreiras geográfica e ecológica.
- A espécie *P. dominicana* não apresenta estrutura populacional.
- *P. dominicana* poderia ser tratada como uma Unidade de Manejo segundo a definição de Moritz (1994).