



UNIVERSIDADE
ESTADUAL DE LONDRINA

WILSON FRANTINE DA SILVA

**FILOGEOGRAFIA, PADRÕES DEMOGRÁFICOS
HISTÓRICOS E DIVERSIDADE GENÉTICA DE ABELHAS
EUGLOSSINI, COM ÊNFASE EM ESPÉCIES
DA MATA ATLÂNTICA**

Londrina
2018



Universidade Estadual de Londrina

Instituto Agrônomo do Paraná

Embrapa Empresa Brasileira de Pesquisa Agropecuária

WILSON FRANTINE DA SILVA

**FILOGEOGRAFIA, PADRÕES DEMOGRÁFICOS
HISTÓRICOS E DIVERSIDADE GENÉTICA DE ABELHAS
EUGLOSSINI, COM ÊNFASE EM ESPÉCIES
DA MATA ATLÂNTICA**

WILSON FRANTINE DA SILVA

**FILOGEOGRAFIA, PADRÕES DEMOGRÁFICOS
HISTÓRICOS E DIVERSIDADE GENÉTICA DE ABELHAS
EUGLOSSINI, COM ÊNFASE EM ESPÉCIES
DA MATA ATLÂNTICA**

Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Estadual de Londrina como requisito parcial para a obtenção do título de Doutor.

Orientadora: Dra. Silvia Helena Sofia

Londrina
2018

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Frantine-Silva, Wilson.

Filogeografia, padrões demográficos históricos e diversidade genética de abelhas Euglossini, com ênfase em espécies da Mata Atlântica / Wilson Frantine-Silva. Londrina, 2018.
294 f. : il.

Orientador: Sílvia Helena Sofia.

Tese (Doutorado em Genética e Biologia Molecular) - Universidade Estadual de Londrina, Centro de Ciências Biológicas, , 2018. Inclui bibliografia.

1. Biogeografia - Tese. 2. Euglossini - Tese. 3. Evolução - Tese. 4. Marcadores Moleculares - Tese. I. Sofia, Sílvia Helena. II. Universidade Estadual de Londrina. Centro de Ciências Biológicas. . III. Título.

WILSON FRANTINE DA SILVA

**FILOGEOGRAFIA, PADRÕES DEMOGRÁFICOS HISTÓRICOS E
DIVERSIDADE GENÉTICA DE ABELHAS EUGLOSSINI, COM
ÊNFASE EM ESPÉCIES DA MATA ATLÂNTICA**

Tese apresentada ao Programa de Pós-Graduação em
Genética e Biologia Molecular, da Universidade
Estadual de Londrina, como requisito parcial para a
obtenção do título de Doutor.

BANCA EXAMINADORA

Orientadora: Dra. Silvia Helena Sofia
Universidade Estadual de Londrina - UEL

Dra. Cristina Yumi Miyaki
Universidade de São Paulo - USP

Dr. Gabriel Augusto Rodrigues de Melo
Universidade Federal do Paraná - UFPR

Dra. Lenice Souza-Shibatta
Universidade Estadual de Londrina - UEL

Dr. Dhiego Gomes Ferreira
Universidade Estadual do Norte do Paraná - UENP

Londrina 30 de maio de 2018.

Aos meus amados pais e irmãos, e a todos aqueles que estiveram ao meu lado durante essa longa jornada e a tornaram possível.

AGRADECIMENTOS

Antes de registrar os devidos agradecimentos, gostaria de ressaltar que o presente estudo só foi possível devido a colaboração de inúmeras pessoas e instituições que contribuíram direta e indiretamente para a execução das diversas atividades relacionadas a este estudo. Redigir, portanto, agradecimentos formais sem incluir a todos seria injustiça, mas fazê-lo também implicaria em mais páginas que eu poderia dispor. Deste modo, gostaria de deixar um agradecimento geral às muitas pessoas que cruzaram meu caminho ao longo dessa jornada e prestaram sua ajuda de forma gratuita e irrestrita, tornando o presente estudo possível. Dito isto, agradeço:

Aos membros da banca examinadora, Dra. Cristina Yumi Miyaki, Dra. Lenice Souza-Shibatta, Dr. Gabriel Augusto Rodrigues de Melo e ao Dr. Dhiego Gomes Ferreira pela prontidão ao aceitarem avaliar o presente estudo, dando suas fundamentais contribuições críticas ao trabalho.

À Universidade Estadual de Londrina e ao Programa de Pós-Graduação em Genética e Biologia Molecular, não apenas pela estrutura física, mas por todo recurso humano, em especial à Dra. Ana Lúcia Dias e Suali Trindade, fundamentais para a devida execução de todo o processo até aqui. Estendo agradecimento à Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior pelo apoio financeiro ao longo dos quatro anos de doutorado, bem como à Universidade da Califórnia Davis (EUA) e ao Departamento de Ecologia e Evolução daquela instituição pela estrutura e auxílio durante o período sanduíche.

À Dra. Silvia Helena Sofia não apenas por sua orientação científica ao longo destes quatro anos, mas também por sua amizade, a qual muito estimo. Meus sinceros agradecimentos pelas inúmeras lições, pelo exemplo de ética, dedicação, responsabilidade e profissionalismo. Agradeço pela confiança e apoio incondicional ao longo da execução deste estudo, mas também pela paciência e compreensão. Agradeço ainda pelo incentivo diário e pela incansável disposição em nos apaixonar pelas incríveis abelhas das orquídeas. Por fim, agradeço por ter visto em mim, inúmeras vezes, um potencial que nem mesmo eu havia visto. Agradeço pelas inúmeras portas que me abriu neste processo. Obrigado!

Ao Dr. Santiago Ramírez por tão gentilmente ter me recebido em seu laboratório na Universidade da Califórnia Davis (EUA), fornecendo estrutura e apoio financeiro para execução de um dos capítulos desta tese. Aproveito para estender estes agradecimentos a todo Ramírez' Lab, em especial Phillip Brand, Nicholas Saleh, Fernanda Guizar, Micah Friedman, Cheryl Dean e Julie Cridland, por todo apoio durante a estadia em Davis.

À Secretaria do Meio Ambiente do Estado de São Paulo e o complexo de Parques Estaduais da Serra do Mar, Núcleo Picinguaba, São Sebastião e ao Parque Estadual de Ilhabela. Em especial aos gestores e corpo técnico destes parques. À Nairton de Araújo por todo auxílio na expedição de campo em Camburí, Boissucanga e Juquei. À Marco Aurélio Nascimento e toda equipe do Parque Estadual de Ilhabela pelo auxílio nas coletas naquela localidade.

À secretaria do Meio Ambiente do Estado de Minas Gerais e ao corpo gestor do Parque Estadual Pico do Itacolomi. À Universidade Federal de Ouro Preto, aos alunos do Programa de Pós-Graduação em Ecologia de Biomas Tropicais (2014) e à Dr. Yasmin Antonini pelo suporte na expedição de Ouro Preto - MG.

À secretaria de Meio Ambiente do Estado da Bahia, pelo auxílio e autorização para a expedição de campo ao Parque Estadual da Serra do Conduru em Serra Grande (Itacaré – BA).

Ao Instituto Chico Mendes pela autorização e estrutura para as expedições de campo no Parque Nacional (PARNA) do Iguaçu (Foz do Iguaçu-PR), PARNA de Superagui (Guaraqueçaba-PR), PARNA Serra dos Órgãos (Terezópolis-PR), Reserva Biológica União ([Rebio] Macaé – RJ / Rio das Ostras – RJ), Rebio Sooretama (Linhares-ES), Parna Descobrimento – (Cumuruxatiba / Prado – BA), Parna Serra de Itabaiana (Areia Branca SE), Estação Ecológica de Murici (Murici-AL) e Parna Guararapes (Maranguapé-PB). Deixo aqui meus especiais agradecimentos aos gestores dos parques por todo apoio frente às limitações estruturais de cada unidade. Agradeço ainda a todos os membros do corpo técnico de cada uma das respectivas UCs, do faxineiro ao diretor, todos sem exceção contribuíram sobremaneira para tornar meus dias nesses locais os melhores possíveis.

Ao Instituto Nacional de Pesquisas Amazônicas (INPA), mas em especial ao Dr. Márcio de Oliveira, por ter tão gentilmente me auxiliado na expedição de coleta à Reserva Duke, por ter me recebido em seu laboratório e ter colaborado diretamente com a identificação do material. Agradeço também de forma especial ao Dr. Márcio pela doação de exemplares de *Euglossa crassipunctata* da bacia amazônica tão importantes para o presente estudo. O meu muito obrigado. Aproveito à oportunidade para agradecer ao amigo Me. Bruno Garcia Oliveira por todo apoio nesta expedição à Reserva Duke, por ter disponibilizado sua casa e por intermediar meu contato com o Dr. Márcio.

Ao Dr. José Eustáquio Santos-Júnior pelo acolhimento em Belo Horizonte, e pela fundamental doação do material de vários espécimes do grupo *crassipunctata*. À Dra. Isabel dos Santos-Alves, Dr. Léo Correia da Rocha Filho, Dr. Guaraci Cordeiro, Me. Enderlei Dec,

os quais tão gentilmente cederam espécimes utilizados neste estudo. Agradeço também ao Dr. Rodrigo B. Gonçalves e a Me. Priscila Soares Oliveira pelo auxílio à expedição ao Parque Estadual de São Camilo, Palotina-PR. Deixo também meus agradecimentos à Fábio Bombadil Amam pela gentileza e apoio logístico.

À Dra. Fernanda Simões de Almeida, por todo apoio e ensinamentos no Laboratório. Pelas conversas descontraídas e pelas palavras de apoio. Agradeço também os conselhos e orientações, mas principalmente pela amizade. Meu muito obrigado!

Ao Dr. Douglas Caldeira Giangarelli pela identificação de inúmeros exemplares utilizados em vários estudos que compõe esta tese, mas também pela amizade e companheirismo ao longo destes quase seis anos de convivência. Agradeço também pelos ensinamentos taxonômicos sobre as abelhas Euglossini e pela experiência de campo nas expedições à Superagui e Parna do Iguaçu. Obrigado!

Meus sinceros agradecimentos aos vários amigos que passaram pelo Laboratório de Genética e Ecologia Animal, com quem compartilhei e aprendi muito nestes últimos seis anos em Londrina. Deixo meus agradecimentos às gerações de abelhudos Alessandra Novaga Alves, Rafael Elias Silva Penha, Karen Mayumi Suzuki, Gabriele Freiria, Natália Uemura, André Gobatto, Rafael Eno e Welber Pinna pela ajuda com assuntos relativos às abelhas e pela amizade ao longo destes anos.

Agradeço aos grandes amigos de Davis - CA (EUA): Duka, Bruno, Luiz (Carioca), Luís (Gaúcho), Paulo Ricardo, Letícia, Fernanda, Fe Nakamura, Ana e Vinícius, Paulo Vitor e Thaís, Mari e Wellington, Don Pablo e Paola. Todos vocês tornaram a distância e a saudade de casa bem menores. Obrigado!

Deixo meus agradecimentos especiais:

Ao Me. Raul Henrique Cardoso Nascimento, por todo o companheirismo e amizade que completam uma década este ano, mas principalmente pelo apoio incondicional de sempre. Obrigado!

Ao Me. Luiz Eduardo de Souza Gazal, por quase uma década de amizade, mas por todo auxílio nesta fase final do Doutorado. Obrigado!

À Me. Camila Satie Savada pela grande amizade e por todo auxílio ao longo desses tantos anos no LAGEA. Agradeço pelo auxílio na revisão deste trabalho, bem como pela ajuda com execuções de experimento e por toda a ajuda no dia-a-dia do laboratório. Obrigado!

À Me. Caroline Apolinário-Silva pela grande amizade, também de longa data, mas principalmente pelo auxílio diário no laboratório, principalmente no período em que estive fora do país. Meu muito obrigado!

À mestranda Thais Kotelok Diniz, deixo meu agradecimento não apenas pela amizade, mas pelas repetidas revisões neste texto. Obrigado pelas correções e sugestões, que foram fundamentais para a apresentação do mesmo.

Agradeço ainda aos amigos de longa data do núcleo de Cornélio Procópio, Renata Alfredo, Dr. Carlos Delfino Vieira e Dr. Bruno Ambrózio Galindo. A este último, devo ainda um agradecimento em especial, pois grande parte do meu ingresso na carreira acadêmica e principalmente neste doutorado foram incentivados por ele.

Por fim, deixo meu especial agradecimento àqueles que sempre estiveram ao meu lado. Agradeço à minha família, meus amados pais, Wilson e Lidileuza e irmãos, Michel e Marlini, por todo suporte sem o qual eu jamais poderia ter realizado este trabalho. Vocês sempre foram a razão e o porque desta jornada.

“Se pude enxergar mais longe, foi porque me apoiei sobre os ombros de gigantes”.

Sir. Isaac Newton

SILVA, W. F. **Filogeografia, padrões demográficos históricos e diversidade genética de abelhas Euglossini, com ênfase em espécies da Mata Atlântica.** 2018. 294 f. Tese (Doutorado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, Londrina. 2018.

RESUMO

A tribo Euglossini compreende um grupo de abelhas neotropicais de cerca de 240 espécies as quais apresentam complexas histórias biológicas pouco exploradas do ponto de vista biogeográfico e filogeográfico. Devido uma gama de comportamentos e outras peculiaridades, estas abelhas apresentam uma intrínseca relação com seus ambientes, de modo que suas histórias biológicas são reflexo da própria história de seus respectivos habitats. Amplamente distribuídas nas florestas úmidas americanas, as abelhas das orquídeas, como são popularmente conhecidas, apresentam altas taxas de endemismo na Mata Atlântica, onde quase 50% das espécies documentadas são endêmicas a este ambiente. Dentre muitos fatores por trás dos processos geradores de diversidade, mudanças climáticas ao longo do Pleistoceno, movimentos neotectônicos e barreiras geológicas, são frequentemente apontadas como os principais agentes históricos por trás da megadiversidade deste bioma. No presente estudo, diferentes espécies de abelhas da tribo Euglossini são utilizadas como modelo para investigar os processos por trás da diversificação destas abelhas ao longo das florestas tropicais americanas, com ênfase na Mata Atlântica. Para tanto, três espécies e grupos de espécies do gênero *Euglossa* correspondendo à diferentes modelos distribucionais (*E. annectans*, *E. iopoecila*, e cinco espécies do subgênero *Glossurella*, grupo *crassipunctata*) foram selecionados para inferir sobre processos em diferentes escalas latitudinais, longitudinais e temporais. Técnicas baseadas em sequenciamento de genes mitocondriais, genotipagem de locos microsatélites e sequenciamento paralelo massivo foram utilizados em associação à modelagem de nicho ecológico e distribuição de espécies para avaliar aspectos genéticos populacionais, filogenéticos e filogeográficos, bem como sua congruência com padrões climáticos atuais e pretéritos, eventos geológicos e neotectonismos. Os resultados encontrados apontam que flutuações climáticas foram determinantes para definição da estrutura populacional das espécies de Euglossini na Mata Atlântica, porém os efeitos destes eventos dependem fortemente das idiosincrasias associadas à aspectos ecológicos de cada espécie. De acordo com os resultados, espécies associadas às Florestas Estacionais Semidecíduais tenderam a migrar parte de suas populações para o interior, se afastando da região litorânea ao longo dos picos glaciares do Pleistoceno. Por outro lado, espécies com ampla distribuição latitudinal associadas à Floresta Ombrófila Densa se encaixam nas expectativas da interpretação atual dos refúgios para a Mata Atlântica. Os modelos de nicho para as abelhas das orquídeas aqui estudadas apontam de forma recorrente a existência de áreas adequáveis na porção sul da Mata Atlântica no último máximo glacial (21 Ka a. p.). A cronologia da diversificação filogenética das espécies do grupo *crassipunctata* indica o aumento da atividade orogênica nos Andes à ~5-6 Ma a. p. como fator determinante para o surgimento de quatro das seis espécies atuais, e que as flutuações climáticas do Pleistoceno apenas influenciaram a separação entre o clado amazônico e atlântico. Os resultados do presente estudo são inéditos e agregam uma importante contribuição para o entendimento dos padrões e processos geradores de diversidade na tribo Euglossini, bem como, na Mata Atlântica.

Palavras-chave: Filogeografia. Euglossini. Evolução. Marcadores moleculares. Modelagem de nicho ecológico.

SILVA, W. F. **Phylogeography, demographic history patterns and genetic diversity of euglossine bees, with focus on species from Atlantic Forest.** 2018. 294 p. Dissertation (Doctorate Degree in Biological Sciences) - Universidade Estadual de Londrina, Londrina, 2018.

ABSTRACT

The Euglossini tribe comprises about 240 species of neotropical bees which have intriguing biological histories underexplored from biogeographic and phylogeographic perspectives. Due to a range of behaviors and other particularities, these bees have developed an intrinsic relationship with their environments, hence their biological histories reflect the history of their own habitats. Popularly known as orchid bees, these bees are widely distributed through the tropical American Forests, reaching up to 50% of endemism in the Atlantic Forest. Several causes have been associated with the diversity-generating processes, but the climatic changes along the Pleistocene, neotectonic movements and geological barriers are often referred as the main historical drivers to the megadiversity of the tropical biomes. Herein, bees of the Euglossini tribe are used as a model to investigate the processes behind the diversification of these bees along the tropical American Forests, with emphasis on the Atlantic Forest. For this purpose, three cases of species and groups of species belonging to the *Euglossa* genus with different distributional ranges (*E. annectans*, *E. iopoecila*, and five species of the subgenus *Glossurella*, group *crassipunctata*) were accessed to infer on evolutionary processes at different spatial and temporal scales. Techniques of mtDNA sequencing, microsatellites, and NGS sequencing were used in association with ecological niche modeling (ENMs) to evaluate the congruence of population structure, phylogenetic and phylogeographic patterns with current and past climate scenarios, geological events and neotectonism. Results show that climatic fluctuations were determinant to define the population structure of Euglossini species in the Atlantic Forest, but the effects of such events are strongly dependent on the respective species' idiosyncrasies associated to their ecological aspects. Species associated with the Semideciduous Seasonal Forests tended to migrate part of their populations toward the country, moving away from the coastal region along the Pleistocene glacier peaks. On the other hand, species with wide latitudinal distribution associated with the Dense Ombrophylous Forest fill the expectations of the current understanding about refugees in the Atlantic Forest. ENMs consistently point out to the existence of suitable areas for several species in the southern portion of the Atlantic Forest in the last glacial maximum (21 Ka a. p.). Chronology of the phylogenetic diversification of the species of the *crassipunctata* group indicates the increase of the orogenic activity in the Andes ~ 5 Ma a. p. as a determining factor for the emergence of four out six current species. Also, the climatic fluctuations of the Pleistocene only played a role in the speciation for Atlantic clade from the Amazonian. These results are the first of its kind for euglossine bees and aggregate an important contribution to the understanding of the processes that generate diversity in the Euglossini tribe, as well as in the Atlantic Forest.

Keywords: Phylogeography. Euglossini. Evolution. Molecular markers. Ecological niche modelling.

LISTA DE ILUSTRAÇÕES

- Figura 3.1 - Relação conceitual da Filogeografia como ponte empírica entre as ciências da Biodiversidade.** Adaptado de Avise (2009) e Riddle (2009). 18
- Figura 3.2 - Comparação da representação gráfica do genoma mitocondrial de duas abelhas corbiculadas, *Apis mellifera* (CROZIER; CROZIER, 1993) e *Euglossa dilemma* (BRAND et al., 2017).** Cores representam tipos diferentes de sequências; N/A – regiões sem informação para *E. dilemma*; A+T – origem de replicação do DNA mitocondrial (CROZIER; CROZIER, 1993). 21
- Figura 3.3 - Esquema representativo da evolução temporal do uso de marcadores moleculares na Filogeografia e Filogenia.** Fonte: (BRITO; EDWARDS, 2009). A figura retrata a unicidade original das duas áreas devido um foco comum no DNAmT na década de 1980. Durante a década de 1990, no entanto, embora o DNAmT ainda fosse usado tanto na Filogeografia quanto na filogenética, esses campos tomaram rumos distintos, muito devido aos diferentes marcadores nucleares utilizados em cada área. Após o acesso ao sequenciamento de nova geração e genomas nucleares de organismos não-modelos ao longo dos anos 2000, a Filogeografia e a Filogenética tomaram rumos convergentes com dados de SNP, embora a natureza evolutiva e as taxas mutacionais das regiões alvo ainda sejam ponto de distinção entre ambas. 27
- Figura 3.4 - Tecnologias de sequenciamento.** A) Exemplificação do esquema geral de sequenciamento automático baseado no método de Sanger; B) Exemplificação de um esquema de sequenciamento paralelo massivo (tipo *Illumina*®), também frequentemente citado como NGS (*Next Generation Sequencing*). Adaptado de Schendure et al. (2017). 30
- Figura 3.5 - Comparação entre estratégias genéricas de subamostragem e de filtragem de elementos genômicos.** (A) esquema genérico de RADSeq / GBS; (B) esquema genérico de captura de sequências; (C) distribuições de leitura típicas de bibliotecas genômicas de

sequências de vários métodos RADseq; (D) Captura de Sequência. Em (C), os sítios de restrição são representados usando setas com coloração indicando diferentes enzimas de restrição. Adaptado de Harvey et al. (2016)..... 34

Figura 3.6 - Evolução das principais tecnologias NGS e seu uso na Filogeografia. Os dados apresentados são resultado de uma busca simples no indexador *GoogleScholar* compreendendo aos termos descritos na legenda.*Estimativas não incluem abordagens relacionadas e derivações do RADseq, como ddRADseq, GBS, entre outras (veja DAVEY et al., 2011). **Dados para 2018 foram projetados a partir dos dados indexados até 28/02/2018..... 35

Figura 3.7 - Métodos para a descoberta de marcadores de alto rendimento.
 A) Um exemplo de região genômica contendo sítios de restrição (vermelho). Uma amostra de DNA de cada um dos dois indivíduos (a amostra 1 é azul escuro e a amostra 2 é azul claro) deve ser sequenciada. A amostra 2 tem uma variação no local de corte em 1.300 bases (seta cinza) e, portanto, este sítio não será cortado. B) Protocolos para o desenvolvimento de marcadores sequenciados. C) Os marcadores sequenciados estão alinhados com o genoma de referência original. RRL: qualquer um dos fragmentos (linhas grossas) ou fragmentos inteiros (linhas finas) entre 300 e 700 pb são sequenciados. Como o local em 1.300 bases não está presente na amostra 2, o fragmento longo entre 700 e 2.000 bases é filtrado por seleção de tamanho. *RAD-seq*: as regiões a jusante de todos os fragmentos acima de 300 bases são sequenciadas, mas não o fragmento entre 150 e 350 bases. As linhas finas indicam a sequência que seria coberta usando o sequenciamento bidirecional (*pair-end*). GBS: as linhas tracejadas representam regiões que seriam filtradas durante a amplificação, mas poderiam ser imputadas usando, por exemplo, o algoritmo “*multiplexed shotgun genotyping hidden Markov model*”. Os fragmentos curtos entre 150 e 350 bases será sequenciado. 40

- Figura 3.8 - Estratégia genérica para análise de dados gerados por métodos de redução de representação genômica (RRL, RADseq, GBS).** Conjunto de milhões de *reads* são demultiplexados e separados dos reads sem sinais de sítios de restrição. Em seguida, diferentes *pipelines* são assumidas dependendo da disponibilidade de genomas de referência. Deste ponto em diante, diversas *pipelines* criam catálogos de locos a partir de algumas amostras, para então verificar a presença destes locos em uma verificação cruzada contra todas as amostras do estudo. Os dados dos locos são anotados em arquivos que guardem estas informações anotando loco, variante e localização do loco, como arquivos VCF e estes são analisados por diferentes programas. 44
- Figura 3.9 - Escala de Tempo Geológico e evolução paleogeográfica, climática e alterações no nível do mar ao longo dos últimos 250Ma.** Estrelas demarcam momentos importantes de diversificação dos euglossíneos. *Q (Quaternário) e Holoceno foram graficamente aumentados para favorecer a legibilidade. **Período Terciário é uma nomenclatura em desuso de acordo com a Sociedade Americana de Estratigrafia, mas foi mantido para fins de comparação devido sua frequente presença em textos e livros-base. Adaptada de Lavina e Fauth (2016). 47
- Figura 3.10 - Transição entre os cenários de transgressões marinhas ao longo do Médio Mioceno e Mioceno Tardio.** À esquerda um cenário intermediário com leve introgressão marinha Pacífica e Atlântica, 17-15 Ma a. p. (SCOTESE, 2001). À direita um cenário de forte introgressão marinha Pacífica e Atlântica, formando o “Mar Paranaense” (ORTIZ-JAUREGUIZAR; CLADERA, 2006). 50
- Figura 3.11 - Evolução geológica da América do Sul e nível do mar nos últimos 200 Ma a. p.** *Limite inferior e **limite superior para o surgimento e diversificação dos gêneros de Euglossini (RAMIREZ et al., 2010). *** Considerando dados de (MONTES et al., 2015). Fonte: PALEOMAP Project (SCOTESE, 2001). 54
- Figura 3.12 - Distribuição dos Refúgios Florestais de Haffer (1969).** Pontilhados demarcam as áreas úmidas; setas indicam o sentido das

possíveis migrações a partir do centro-oeste brasileiro; áreas em preto demarcam elevações acima de 1000 metros; áreas hachuradas e pontilhadas demarcam zonas de transgressão marinha. Reproduzido de Haffer (1969). 56

Figura 3.13 Mapas de consenso de áreas historicamente estáveis para a Mata Atlântica sob definições estrita (*stricto sensu*) e ampla (*lato sensu*), estimando a ausência/presença de floresta sob cenários climáticos atuais de 6 Ka e 21 Ka a. p. Linhas brancas demarcam os rios São Francisco e Doce. A flecha indica o pequeno refúgio previsto nos limites dos estados do Espírito Santo e do Rio de Janeiro. Refúgios no interior da Atlântica em Floresta Estacional Semidecidual também são observáveis (meu comentário). Adaptado de Carnaval e Moritz (2008). 58

Figura 3.14 - Resumo do Modelo de Expansão da Mata Atlântica. A) Costa Brasileira e Plataforma Continental; B) Modelos de Nicho Ecológico (segundo Leite et al., 2016) para o Último Máximo Glacial (LGM) e o presente, mostrando a expansão das áreas adequáveis em direção à plataforma; C) oscilações de temperatura e nível do mar ao longo dos últimos 200 Ma, englobando o LGM, o último período interglacial (LIG) e penúltimo máximo glacial (MIS6); D) Modelos Demográficos representando os eventos esperados sob a hipótese dos Refúgios (vermelho) e a hipótese de Expansão da MA (verde); BN – *bottleneck* populacional; EX - expansão populacional. Adaptado de Leite et al. (2016). 60

Figura 3.15 - Extensão original e atual da Mata Atlântica brasileira. A) área original: FOD – Floresta Ombrófila Densa, FOA – Floresta Ombrófila Aberta, FOM – Floresta Ombrófila Mista, FED – Floresta Estacional Decidual, FES – Floresta Estacional e Semidecidual; B) Fragmentos de Mata Atlântica, ano base 2016-2017. Fonte: Fundação SOS Mata Atlântica, Relatório Anual (2018). 61

Figura 3.16 - Principais descontinuidades filogeográficas ao longo da Mata Atlântica. Descontinuidade Rio São Francisco: Cabanne et al. (2008), Carnaval e Moritz (2008), Carnaval et al. (2009), Porto et

al. (2013), López-Uribe et al. (2014), Frantine-Silva et al. (2017); Rio Doce: Cabanne et al. (2007; 2008), Carnaval e Moritz (2008), Carnaval et al. (2009), Martins (2011), Porto et al. (2013), López-Uribe (2014), Batalha-Filho e Miyaki (2016); Ribeiro, Lemos-Filho e Lovato (2011); Maldonado-Coelho (2012); Rio Paraíba do Sul: Cabanne et al. (2008), Raposo do Amaral et al. (2013); Penha et al. (2015), Frantine-Silva et al. (2015); Rio Tietê: Grazziotin et al. (2006), Cabanne et al. (2007; 2008), Martins et al. (2009), Batalha-Filho et al. (2010) e Thomé et al. (2010). *Descontinuidades baseadas em Batalha-Filho e Myiaki et al. (2011) e Da Silva, Pinto da Rocha e Souza (2016), com inclusão de alguns novos estudos. Imagem: Adaptada de SOS MATA ALTÂNTICA (2016)..... 63

Figura 3.17 - Variações do F_{ST} *coalescente* (Slatkins, 1985), G_{ST} (“análogos ao F_{ST} ”), D de Jost (2008), G'_{ST} (Hedrick, 2005) (eixo y) em relação à diferentes taxas de migração (eixo x), com taxa de mutação $\mu=10^{-3}$. Nestes gráficos, as populações tiveram $N = 100$ em um modelo de ilhas infinitas. O gráfico da esquerda e da direita mostram resultados para 2 e 100 populações por espécie, respectivamente. A linha para G_{ST} é muitas vezes encoberta pela linha de F_{STcoal} , porque eles são quase idênticos para baixas taxas de mutação (modificado de WHITLOCK, 2011). 72

Figura 3.18 - Esquema da estratégia básica de modelagem de nicho ecológico. Registros de ocorrência georreferenciados são utilizados como pontos para coleta de informações de diferentes variáveis igualmente georreferenciadas. A partir destas informações, diferentes algoritmos calculam distribuições relativas e/ou as projetam para diferentes intervalos cronológicos em um sistema retroalimentado pelas variáveis disponíveis em diferentes espaços temporais. 79

Figura 3.19 - Evolução de Apoidea (Hymenoptera: Aculeata). Topologia parcial da filogenia de Hymenoptera: Apocrita, apresentando a posição das sete famílias de abelhas (Anthophila) resultante da análise de 854 locos (203.095 Kb) com 38 pontos de calibração, modificado de Branstetter et al. (2017). Destaque para a posição de

	Formicidae como grupo irmão de Apoidea a qual sugere uma ancestralidade comum para comportamento social.....	85
Figura 3.20 -	Espécimes representativos de cada um dos cinco gêneros da tribo Euglossini: A) <i>Aglae caerulea</i> , B) <i>Eulaema nigrita</i> , C) <i>Exaerete smaragdina</i> , D) <i>Eufriesea dentilabris</i> , E) <i>Euglossa stellfeldi</i> . F) Exemplo da área aveludada da metatíbia de um macho de <i>Euglossa stellfeldi</i> com destaque para os tufo ou coxins mesotibiais (lado esquerdo), frequentemente utilizados para diagnose de espécies. Destaque para a abertura da metatíbia através da qual as abelhas masculinas depositam e armazenam fragrâncias químicas.	87
Figura 3.21 -	Evolução dos trabalhos sobre abelhas das orquídeas ao longo do tempo. A – Números de resultados correspondendo às publicações cadastradas na plataforma GoogleScholar até janeiro de 2018 contendo qualquer uma das palavras-chave: “Euglossini, <i>Orchid bees</i> , euglossine, euglossina”; B – Número de espécies válidas descritas com base em seus respectivos anos de descrição com base em Moure et al. (2012).....	88
Figura 3.22 -	Filogenia das abelhas corbiculadas baseadas em análises filogenômicas corrigindo o viés do conteúdo GC (ROMIGUIER et al. 2016). Filograma resumido das 100 melhores árvores Bayesianas retirado de Bossert et al. (2017). Nó em destaque agrupa as espécies que representam as tribos dos corbiculados com comportamentos sociais mais estruturados.	90
Figura 3.23 -	Áreas de endemismo para Abelhas das Orquídeas. As áreas de endemismo são demarcadas por diferentes cores. Reproduzido de Ramírez et al. (2010).....	100
Figura 3.24 -	Áreas de endemismo na Mata Atlântica a partir do compartilhamento de espécies de Euglossini. As cores, setas e chaves demarcam as áreas de endemismo propostas por Garraffoni et al. (2017). Adaptado de Garraffoni et al. (2017).....	101
Figura 3.25 -	Trabalhos publicados com euglossíneos e relacionados à Genética de Populações, Filogeografia e Detecção de Machos Diploides. Resultados relacionados à busca “ <i>Euglossini</i> ” OR	

“*Euglossina*” AND “*Population Genetics*” no portal *Google Scholar*. A) evolução do número de trabalhos sob os termos da pesquisa, categorizados de forma qualitativa. O número de vezes que cada espécie foi estudada e o respectivo marcador utilizado são apresentados para: B) trabalhos categorizados como Filogeografia; C) Genética de Populações; D) detecção de machos diploides. Detalhes em Apêndices. *Trabalhos envolvendo análises filogenéticas ou identificação molecular, não foram incluídos..... 107

Figura 3.26 - *Euglossa iopoecila* Dressler (1982). A- visão lateral, morfotipo azul (coleção de abelhas MZUEL-1189, Salto Morato, 14-IV-2010; *Gianganelli, DG. leg.*); B- visão lateral; morfotipo verde (WFS0012-PARNA Descobrimto; 15-III-2015, *Frantine-Silva, W. leg.*); C) visão frontal; morfotipo azul D) visão superior; morfotipo azul; E) tíbia posterior esquerda; F) tíbia média esquerda; morfotipo verde; G) morfotipo azul; Destaque para coxim anterior ~1/3 do tamanho do coxim basal em *E. iopoecila*. H) tíbia média em *E. chalybeata*, fonte: D. W. Roubik..... 113

Figura 3.27 - Distribuição e áreas de ocorrência de *Euglossa iopoecila* Dressler (1982) e seus respectivos morfotipos azul (ao sul; círculos roxos) e verde (ao norte; triângulos verdes) da distribuição. Registros aferidos na literatura, coleções e presente estudo (Apêndices)..... 114

Figura 3.28 - *Euglossa annectans* Dressler 1982 (A, B, C e D) e comparação de traços morfológicos em relação a espécie mais relacionada *Euglossa stellfeldi* (E, F, G). *E. annectans*: A) visão lateral; B) face externa da tíbia posterior, detalhe para oclusão basal mais afilada da fenda metatibial; C) face interna metatibial; D) visão lateral da mesotíbia, destaque para a proporção dos coxins em relação à *E. stellfeldi*; *E. stellfeldi*: E) face externa metatibial; F) face interna metatibial, detalhe para a concavidade na região basal; G) visão lateral da mesotíbia, detalhe para a diferença na proporção dos coxins em relação à *E. annectans*. 118

Figura 3.29 -	Distribuição de <i>Euglossa annectans</i>. Registros aferidos na literatura, coleções e presente estudo (Detalhes das origens dos registros podem ser encontrados nos Apêndices)	121
Figura 3.30 -	Comparação entre os holótipos de <i>Euglossa crassipunctata</i> (A, C e E) e <i>Euglossa sapphirina</i> (B, D e F). Setas pretas indicam a posição de possíveis caracteres diagnósticos entre as espécies (vide texto). Adaptada de Nemésio (2009), fonte original (David W. Roubik).....	124
Figura 3.31 -	Ocorrência conhecida das espécies do grupo <i>crassipunctata</i> de acordo com duas hipóteses de espécies. A- De acordo com Moure et al. (2012); B - De acordo com Nemésio e Engel (2012). Em A, <i>E. sapphirina</i> tem distribuição disjunta, ocorrendo no sul da Mata Atlântica e apenas na América Central. Destaque em B para <i>E. pepei</i> com ocorrência pontual no sul da Bahia e <i>E. clausi</i> com distribuição em toda a Mata Atlântica.	128
Figura 3.32 -	Visão frontal e características tibiais de <i>E. clausi</i> (A-D), <i>E. moratoi</i> (E-H) e <i>E. pepei</i> (I-L). Adaptado de Nemésio e Engel (2012).	130
Figura 4.1. -	Distribuição dos pontos de amostragem e áreas de distribuição conhecida para o grupo <i>crassipunctata</i> utilizados neste estudo. Áreas em diferentes cores indicam as respectivas áreas de distribuição conhecidas para cada espécie. Em azul: <i>Euglossa sapphirina</i> ; verde: <i>Euglossa crassipunctata</i> ; rosa: <i>Euglossa parvula</i> ; vermelho: <i>Euglossa moratoi</i> ; amarelo: <i>Euglossa clausi</i> . Dados de <i>Euglossa pepei</i> não são apresentados.....	161
Figura 4.2 -	Inferência filogenética e estimativa de tempo de divergência para as espécies do grupo <i>crassipunctata</i> (exceto <i>E. pepei</i>) estimada a partir de 3476 locos GBS por inferência Bayesiana (IB) e de Máxima Verossimilhança (MV). Barras azuis sobre os nós indicam o intervalo de confiança de 95% para a estimativa de cada tempo. Pontos negros sobre os nós indicam valores de suporte igual à 1 ou 100 na IB e MV, com os demais valores sobre os nós indicando o suporte para as respectivas análises. Os grupos demarcados nos táxons terminais correspondem à os grupos	

	geográficos identificados nas análises n-MDS ou Structure (veja o texto). <i>Euglossa moratoi</i> retirada de Nemésio e Engel (2012).	169
Figura 4.3 -	Resultados da análise Bayesiana (Structure) de atribuição. O resultado da variação média do <i>log</i> de probabilidade à posteriori de <i>K</i> em função do desvio padrão (ΔK) é apresentado para os testes de <i>K</i> entre 1 e 13 (à esquerda). O coeficiente de ancestralidade (<i>Q</i>) é apresentado em um gráfico de barras (à direita), onde cada indivíduo corresponde a uma barra e recebem cores de acordo com a proporção de <i>Q</i> de seu respectivo cluster.	170
Figura 4.4 -	Análise de escalonamento multidimensional (nMDS) não métrica com base em 3476 locos GBS para diferentes espécies do grupo <i>crassipunctata</i>. A - nMDS incluindo 1529 SNPs de cinco das seis espécies do grupo <i>crassipunctata</i> : <i>Euglossa clausi</i> , <i>Euglossa crassipunctata</i> , <i>Euglossa moratoi</i> , <i>Euglossa parvula</i> e <i>Euglossa sapphirina</i> ; B – nMDS incluindo dados de 687 SNPs identificados para as amostras de <i>E. clausi</i> ; as cores azul e vermelho demarcam <i>clusters</i> identificados na análise de inferência Bayesiana, Structure; grupo SMA em azul.	171
Figura 4.5. -	Consenso entre modelos de nicho ecológico com base em diferentes modelos de circulação geral (MPI-ESM-LR, CCM4-ESM, MIROC-ESM-LR) para diferentes espécies do grupo <i>crassipunctata</i>. As cores representam as áreas de distribuição para cada espécie: azul – <i>Euglossa sapphirina</i> ; verde – <i>Euglossa crassipunctata</i> ; vermelho – <i>Euglossa moratoi</i> ; roxo – <i>Euglossa parvula</i> ; amarelo <i>Euglossa clausi</i> . A Modelo de distribuição para o tempo “atual” com corte de <i>threshold</i> de adequabilidade igual a 0,5. B – modelo de paleodistribuição baseado no consenso entre os três GCM distintos.	173
Figura 5.1 -	Distribuição geográfica dos haplótipos e estruturação genética mitocondrial de <i>Euglossa annectans</i> ao longo dos 12 locais de estudo (SO, PN, BO, IC, SU, SM, SC, KB, MD, DN, DR e FZ), representados por pontos amarelos (Software QGIS 2.12). A) Mapa parcial do Brasil e distribuição das localidades de amostragem; gráficos de pizza estão posicionados próximos aos	

respectivos pontos indicando a frequência e distribuição de cada haplótipo; os haplótipos não compartilhados são mostrados em cinza claro. B) Rede de haplótipos produzida pelo método Median-Joining para os 8 haplótipos amostrados (H1-H8) obtidos para as sequências do gene *cytb*; tamanho dos círculos são proporcionais à frequência dos haplótipos; cada barra transversal sobre as linhas conectoras representa um passo mutacional; C) Análise Bayesiana de Estrutura Populacional (BAPS) de *E. annectans* a partir de 640 pb do gene *cytb*. Cada indivíduo é representado por uma barra vertical associado a um cluster genético representado por uma cor respectiva..... 196

Figura 5.2 - Estimativa de mudanças demográficas populacionais para o gene *cytb* em 126 amostras de *E. annectans* ao longo da Mata Atlântica. A - Distribuição mismatch par-a-par comparando distribuição esperada para cenário de crescimento populacional (linha sólida; Exp.) e a distribuição observada (linha pontilhada; Obs.); B - Inferência de mudanças do tamanho efetivo populacional (N_e em log) ao longo do tempo em milhões de anos. 198

Figura 5.3 - Análise Bayesiana de estruturação genética baseada em oito locos microssatélites de *E. annectans*. A) da esquerda para direita: estimativa do número de agrupamentos (K) mais provável com base na variação da média da razão do $\log Pr(X/K)$ pelo desvio padrão de $L(K)$, seguido pela média $L(K) \pm$ o desvio padrão (centro) e a variação de \log de $L(K)$ (à direita); B) gráfico de barras para o coeficiente de ancestralidade (Q) considerando $K=2$, onde cada barra representa um indivíduo; as cores vermelho e verde identificam os diferentes clusters genéticos inferidos pelo software Structure. 200

Figura 5.4 - Análise Bayesiana de inferência de fluxo gênico. Taxas relativas de migração entre as 12 localidades amostradas no Sul da Mata Atlântica baseadas na genotipagem de oito locos microssatélites de 359 macho para *E. annectans* são apresentadas sobre a distribuição geográfica das localidades. As cores vermelho (grupo “interior”) e verde (grupo “litoral”) demarcam os grupos identificados na análise

	de estrutura genética pelo software Structure. Setas brancas indicam a direção e intensidade do fluxo gênico.	202
Figura 5.5 -	Correlação entre distância genética e geográfica entre 12 localidades amostradas para <i>E. annectans</i> do sul da Mata Atlântica baseada em oito locos microssatélites. A) todas as amostras; B) grupo “interior”; C) grupo “litoral”.	203
Figura 5.6 -	Modelagem de Nicho Ecológico e Paleodistribuição de <i>Euglossa annectans</i> baseados em três diferentes Modelos de Circulação Geral (GCM: MIROC-ESM, CCSM4, MPI-ESM-RL), 71 registros de ocorrência e diferentes abordagens de seleção de variáveis. Imagens em cores correspondem a modelos baseados em apenas um GCM com a escala de variação indicando a adequabilidade ambiental; Imagens em preto e branco correspondem ao consenso de GCMs; A última linha (Consenso) resume a sobreposição das áreas presentes em todos os modelos de cada uma das estratégias de seleções de variáveis; Colunas indicam os diferentes tempos (da esquerda para à direita: Atual, 6 Ka a. p., 21 Ka a. p. e 130 Ka a. p.); Diferentes abordagens de seleção de variáveis são apresentados em linhas: PCA, Ecológico, WorldClim (19 variáveis).....	205
Figura 5.7 -	Comparação entre os consensos dos modelos de distribuição de <i>Euglossa annectans</i> baseados nas diferentes estratégias de seleção de variáveis considerando a variação climática “atual” (1950-1990) e o consenso entre os diferentes GCMs considerando a progressão moderada do acúmulo de carbono ao longo dos próximos 32 anos (IPCC rcp-4.5). Áreas de sobreposição de três (à esquerda) ou mais modelos (à direita) são destacados na coloração azul.	206
Figura 6.1 -	A) Mapa parcial do Brasil, mostrando os 14 locais de estudo (SC, PR1, PR2, SP1, SP2, SP3, SP4, SP5, RJ, ES, BA1, BA2, BA3 e AL), representados por pontos amarelos (Software QGIS 2.12). Os gráficos de pizza ligados aos pontos amarelos indicam a frequência e distribuição de cada haplótipo; os haplótipos não compartilhados são mostrados em cinza claro.	

B) Rede de relacionamento para os 76 haplótipos (Median-joining) amostrados (H1-H76) obtidos para as sequências de DNAm concatenadas. Os tamanhos do círculo são proporcionais à frequência de haplótipos (H). Os três haplogrupos inferiores (ou clusters) são representados por linhas tracejadas, pontilhadas e sólidas. Cada círculo representa um haplótipo, as cores indicam localidades e o tamanho, o número de indivíduos; cada barra transversal sobre as linhas conectoras representa um passo mutacional; círculos pretos representam os vetores médios. C) Análise Bayesiana da estrutura populacional (BAPS) de *E. iopocila* baseado em 2234 pb do DNAm. Cada indivíduo é representado por uma barra vertical fina dividida em grupo coloridos por seus grupos genéticos (azul = cluster 1; verde = cluster 2; vermelho = cluster 3). As áreas destacadas no mapa representam a distribuição geográfica dos clusters. Os códigos do sítio são os mesmos que na Tabela 5.1. (retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18:539; <https://doi.org/10.1007/s10592-016-0905-7>) 230

Figura 6.2 - Distribuições das diferenças entre pares de sequências observadas (linha pontilhada) contra um padrão esperado em um modelo de tamanho populacional constante (linha sólida) baseados em 2234 pb do DNAm de *Euglossa iopocila* amostrada ao longa da Mata Atlântica. A) cluster Sul; B) cluster Central; C) cluster Norte. As diferenças em pares de sequências (eixo X) são mostradas contra a frequência das respectivas diferenças (eixo Y). Padrões não modais correspondem ao crescimento rápido da população, enquanto os padrões multimodais correspondem a tamanhos de população constantes. (Retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18: 539 <https://doi.org/10.1007/s10592-016-0905-7>) 234

Figura 6.3 - Estimativas do *Extended Bayesian Skyline Plot* (EBSP) para os diferentes *clusters* genéticos de *Euglossa iopocila*: A) *cluster* Sul; B) *cluster* Central; C) *cluster* Norte; A linha tracejada mostra a estimativa mediana do EBSP e as linhas sólidas superior e inferior

correspondem aos limites de densidade à posteriori dentro de um intervalo de confiança de 95%. O eixo X representa o tempo em milhões de anos (Ma a. p.) e o eixo Y o logaritmo natural do tamanho efetivo da população estimado (Ne) por geração. (Retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18: 539 <https://doi.org/10.1007/s10592-016-0905-7>) 234

Figura 6.4 - Modelagem de distribuição de espécies de *Euglossa iopoecila* nos últimos 130 mil anos. A) Mapa com 45 locais de ocorrência (quadrados pretos) usados como entrada no MaxEnt. B) Modelo de nicho ecológico (ENM) de *Euglossa iopoecila* para o tempo “atual”; C) médio Holoceno (MID; 6 mil a.a.p); D) último máximo glacial (LGM; 21 Ka a. p.); E) último período interglacial (LIG; 130 mil Ka a. p.). As cores mais quentes indicam maior adequação para a ocorrência das espécies, conforme indicado na legenda. (Retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18: 539 <https://doi.org/10.1007/s10592-016-0905-7>) 237

LISTA DE TABELAS

- Tabela 4.1 - Descrição das áreas e número de machos amostrados para as seis espécies do grupo *crassipunctata*.** O código das localidades (ID), localização, bem como latitude e longitude são apresentados ao lado do número de indivíduos amostrados. A - *Euglossa crassipunctata*; B – *Euglossa sapphirina*; C – *Euglossa moratoi*; D – *Euglossa clausi*; E - *Euglossa parvula*; F – *Euglossa pepeii*. 162
- Tabela 4.2 - Índices de polimorfismos entre grupos de sequências analisando um conjunto de 3.476 GBS-locos entre cinco espécies do grupo *crassipunctata*.** Para cada espécie ou população, são apresentados: número de locos obtidos (N_L); de sítios analisados (N_S); sítios privados (P_S); sítios segregantes (S); número médio de indivíduos analisados por loco (N) e desvio padrão (d. p.); diversidade nucleotídica (π)..... 167
- Tabela 5.1 - Descrição das áreas e número de machos de *Euglossa annectans* amostrados no presente estudo.** A identificação das localidades (ID), localização e tipo florestal (TF), bem como latitude e longitude são apresentados ao lado do número de indivíduos amostrados. P.E. – Parque Estadual; PARNA - Parque Nacional; RPPN – Reserva Particular de Proteção da Natureza; APP – Área de Proteção Permanente; FOD – Floresta Ombrófila Densa; FES – Floresta Estacional Semidecidual. 189
- Tabela 5.2 - Índices de diversidade genética e testes de neutralidade para 640pb do gene citocromo *b* de *Euglossa annectans*:** (N) número de amostras, (S) número de sítios segregantes, (h) número de haplótipos, (Hd) diversidade haplotípica, (π) diversidade nucleotídica, ($s.d.$) desvio padrão, (F_s) teste de neutralidade de Fu, (D) teste D de Tajima e R_2 de Ramos-Onsin e Rozas. 195
- Tabela 5.3 - Análise de Variância Molecular (AMOVA):** considerando a existência de um único grupo de *populações a partir de 640 pb do gene *cytb* de 126 amostras de *Euglossa annectans* coletadas em 12 localidades ao longo do sul da Mata Atlântica..... 197

Tabela 5.4 -	<p>Índices de diversidade genética e estimativas de machos diploides para amostras de <i>E. annectans</i> coletadas em 12 localidades da região sul da Mata Atlântica. (N_T) Número total de machos genotipados; (N_{MD}) número de machos diploides; (f_{MD}) frequência de machos diploides; ($CI^{95\%}$) intervalo de confiança de 95% para f_{MD}; ($Phet$) Probabilidade de um macho, caso diploide, ser heterozigoto ao menos em um loco; (N_{MH}) Número de machos haploides incluídos nas análises populacionais; (\bar{N}_A) número médio de alelos; (sd) desvio padrão; (\bar{N}_E) Número médio de alelos efetivos; (N_P) número de alelos privados; (He) Heterozigosidade esperada (<i>sensu</i> Nei, 1978). 199</p>
Tabela 5.5 -	<p>Análise de Variância Molecular (AMOVA): considerando dois grupos de *populações a partir de oito locos microssatélites de 359 machos de <i>Euglossa annectans</i> coletadas em 12 localidades ao longo do sul da Mata Atlântica. 200</p>
Tabela 5.6 -	<p>Índices de estruturação genética par-a-par entre machos de <i>E. annectans</i> coletados em 12 localidades a partir de oito locos microssatélites. Valores de G_{ST} na diagonal superior e valores de $Dest$ na diagonal inferior. Todos os valores foram significativos com $\alpha=0,05$ exceto valores em negrito. 203</p>
Tabela 5.7 -	<p>Estatísticas de seleção de modelos para comparação entre três diferentes estratégias de seleção de variáveis: todas as variáveis (“<i>WorldClim</i>”), com base em aspectos ecológicos (“<i>Ecológica</i>”), e por exclusão de covariantes (“<i>PCA</i>”). N – número de amostras; \bar{AUC} – média da área sob a curva; BIC – critério de informação Bayesiano; AIC – critério de informação de Akaike; $AICc$ – com correção para amostragens reduzidas; $logL$ – log de verossimilhança; \bar{G}_{RT} – média do ganho regularizado de treino de modelo; \bar{G}_{UT} – ganho de treino de modelo não regularizado. 204</p>
Tabela 6.1 -	<p>Relação dos pontos amostrais de <i>Euglossa iopoecila</i> ao longo da Mata Atlântica. Localidades (nome da reserva, município e estado), código da localidade, coordenadas geográficas (graus decimais), e tamanho amostral (n). 225</p>

Tabela 6.2 -	Medidas de diversidade genética de DNAm_t recuperadas para matriz concatenada e para cada uma das quatro sequências de genes (16S, <i>cytb</i> e <i>COI</i>) analisadas: (<i>h</i>) número de haplótipos, (<i>H_d</i>) diversidade de haplótipos, (π) diversidade de nucleotídeos, (<i>n</i>) = número de amostras e (<i>dp</i>) desvio padrão (<i>dp</i> de $\pi < 1 \times 10^{-4}$)	231
Tabela 6.3 -	Análise da variância molecular (AMOVA) para 2234 pb de 126 indivíduos de <i>Euglossa iopoecila</i> amostrados ao longo de 14 sítios no domínio da Mata Atlântica. As divisões hierárquicas seguem os três grupos (<i>clusters</i>) identificados pela análise SAMOVA e BAPS.....	233
Tabela 6.4 -	Parâmetros demográficos e testes de neutralidade para mudanças no tamanho da população: teste de D de Tajima, <i>F_s</i> e <i>R₂</i> com base em quatro conjuntos de dados de sequências de DNAm_t de 14 localidades de <i>Euglossa iopoecila</i> na Mata Atlântica Brasileira.	235

LISTA DE ABREVIATURAS

ACMR	Ancestral Comum Mais Recente (most recent common ancestor)
AFLPP	olimorfismo de Fragmentos Restringidos Amplificados (Amplified Fragments Length Polymorphism)
AMOVA	Análise de Variância Molecular (Analysis of Molecular Variance)
arcmin	Minúto de arco
arcsec	Segundo de arco
BIB	Modelo Bayesiano de Biogeografia de Ilhas (Bayesian Island Biogeography)
DEC	Modelo de Dispersão-Extinção-Cladogênese (Dispersion-Extinction-Cladogenesis)
DIVA	Análise de Dispersão e Vicariância (Dispersion-Vicariism Analysis)
DMP	Produção de Machos Diploides (Diploid Male Production)
DNAcp	DNA cloroplastidial
ENM	Modelagem de Nicho Ecológico (Ecological Niche Modelling)
FES	Floresta Estacional Semidecidual
FOD	Floresta Ombrófila Densa
FOM	Floresta Ombrófila Mista
IB	Inferência Bayesiana (Bayesian Inference - BI)
Ka	Kilo-anum (Milhares de anos, ano x 10 ³)
LGM	Último Máximo Glacial (Last Glacial Maximum)
MCMC	Markov Chain Monte-Carlo (Cadeias de Markov e Monte-Carlo)
MV	Máxima Verossimilhança (Maximum Likelihood - ML)
NGS	Next Generation Sequencing (Sequenciamento de Nova Geração)
PAE	Análise de Endemismo via Parcimônia (Parcimony Endemism Analysis)
RAD-like	Abordagens de NGS relacionadas à RADseq (Restriction Associated DNA based approaches)
RADseq	Restricted site-Associated DNA sequencing (Sequenciamento de DNA Associado à sítios de Restrição)
RC	Região Controle da replicação mitocondrial
RFLP	Polimorfismo de Comprimento de Fragmentos Restringidos (Restricted Fragments Length Polymorphism)

RNAt	RNA transportador
ROR	Taxa de Ocorrência Relativa (Occurrence Relative Rate). Em modelagem ecológica corresponde à taxa de probabilidade de que uma determinada célula tenha registro de uma espécie
SAMOVA	Análise Espacial de Variância Molecular (Spatial Analysis of Molecular Variance)
SDM	Modelo de Distribuição de Espécies (Species Distribution Modelling)
SMM	Step-wise mutation model (modelo de mutação passo-a-passo)
UCE	Ultraconerved Regions (Regiões Ultraconservadas)

SUMÁRIO

1	APRESENTAÇÃO	3
2	OBJETIVOS	5
2.1	OBJETIVO GERAL	5
2.2	OBJETIVOS ESPECÍFICOS	5
3	FUNDAMENTAÇÃO TEÓRICA	7
3.1	UMA BREVE INTRODUÇÃO À BIOGEOGRAFIA	7
3.1.1	Principais Ideias e Métodos Biogeográficos	9
3.1.2	Filogeografia e seu impacto sobre a Biogeografia	18
3.2	FORMAÇÃO E BIOGEOGRAFIA DA AMÉRICA DO SUL: ASPECTOS GERAIS.....	46
3.2.1	O Mesozoico na América do Sul.....	48
3.2.2	A Era Cenozoica na América do Sul.....	49
3.2.3	O Quaternário e as flutuações climáticas	51
3.2.4	Reflexos biogeográficos das Glaciações do Quaternário	55
3.2.5	Padrões biogeográficos na Mata Atlântica	61
3.3	DIVERSIDADE, GENÉTICA E CONSERVAÇÃO	65
3.3.1	Genética e Conservação das Abelhas	66
3.3.2	Análises em Genética de Populações e Filogeografia.....	69
3.3.3	Modelagem de Nicho Ecológico	79
3.4	AS ABELHAS DAS ORQUÍDEAS	85
3.4.1	A origem das abelhas das orquídeas.....	88
3.4.2	Aspectos Sistemáticos	93
3.4.3	Aspectos Biogeográficos	96
3.4.4	Diversidade Genética: o estado da arte em Euglossini.....	106
3.5	O CASO DE <i>EUGLOSSA IOPOECILA</i> DRESSLER	111
3.6	O CASO DE <i>EUGLOSSA ANNECTANS</i> DRESSLER	117
3.7	O CASO DO GRUPO <i>CRASSIPUNCTATA</i>	122
	REFERÊNCIAS	132
	PARTE II : EVIDÊNCIAS EMPÍRICAS	156

4	CAPÍTULO 1: FILOGEOGRAFIA E HISTÓRIA EVOLUTIVA DO GRUPO <i>CRASSIPUNCTATA</i> (APIDAE: EUGLOSSINI) ACESSADA POR <i>GENOTYPING BY SEQUENCING</i> E MODELAGEM DE NICHOS ECOLÓGICO	158
	RESUMO	158
4.1	INTRODUÇÃO.....	158
4.2	MATERIAIS E MÉTODOS	160
4.2.1	Área de Estudo e Amostragem.....	160
4.2.2	Procedimentos Moleculares	163
4.2.3	Análise dos Dados Moleculares	163
4.2.4	Modelagem de Nicho Ecológico	165
4.3	RESULTADOS.....	167
4.3.1	Genotipagem e estatísticas populacionais	167
4.3.2	Análise Filogenética.....	168
4.3.3	Estrutura populacional.....	169
4.3.4	Modelos de Distribuição e Paleodistribuição.....	171
4.4	DISCUSSÃO.....	174
4.4.1	História evolutiva do grupo <i>crassipunctata</i>	174
4.4.2	Estruturação populacional na Mata Atlântica e flutuações climáticas	176
4.4.3	Implicações taxonômicas	178
4.4.4	Conclusões.....	178
	REFERÊNCIAS	179
5	CAPÍTULO 2: ESTRUTURA POPULACIONAL DE <i>EUGLOSSA ANNECTANS</i> DRESSLER INDICA FRAGMENTAÇÃO ENTRE LITORAL E INTERIOR AO LONGO DO PLEISTOCENO INFERIOR	187
	RESUMO	187
5.1	INTRODUÇÃO.....	187
5.2	MATERIAIS E MÉTODOS	189
5.2.1	Área de Estudo e Amostragem.....	189
5.2.2	Procedimentos Moleculares	190
5.2.3	Análise dos Dados Moleculares	190
5.2.4	Modelagem de Nicho Ecológico	193

5.3	RESULTADOS.....	195
5.3.1	DNA mitocondrial.....	195
5.3.2	DNA nuclear (Microsatélites).....	198
5.3.3	Modelagem de Nicho Ecológico e Paleodistribuição.....	204
5.4	DISCUSSÃO.....	206
5.4.1	Padrões de estruturação, demografia e instabilidade climática.....	207
5.4.2	Diversidade genética e conservação.....	211
5.5	CONCLUSÕES.....	213
	REFERÊNCIAS	213
6	CAPÍTULO 3: FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DE <i>EUGLOSSA IOPOECILA</i> DRESSLER: SINAIS DE VICARIÂNCIA ASSOCIADOS ÀS MUDANÇAS CLIMÁTICAS DO QUATERNÁRIO.....	223
	RESUMO	223
6.1	INTRODUÇÃO.....	223
6.2	MATERIAIS E MÉTODOS	225
6.2.1	Área de estudo.....	225
6.2.2	Amostragem e procedimentos moleculares.....	225
6.2.3	Análise dos dados moleculares.....	226
6.2.4	Modelagem de Nicho Ecológico	228
6.3	RESULTADOS.....	229
6.3.1	Estrutura populacional e padrões de diferenciação	232
6.3.2	Análises demográficas.....	233
6.3.3	Modelagem de nichos ecológicos e paleodistribuição	236
6.4	DISCUSSÃO.....	237
	REFERÊNCIAS	243
7	CONSIDERAÇÕES FINAIS.....	249
	APÊNDICES.....	250

1 APRESENTAÇÃO

Abelhas das orquídeas constituem, sem sombra de dúvidas, um dos grupos mais fascinantes dentre as abelhas neotropicais. A peculiaridade de sua história natural e a intrínseca relação com as orquídeas das florestas úmidas compreendem um capítulo ímpar da história natural das Américas. É justamente esta intrínseca relação com as florestas úmidas que torna os euglossíneos peça fundamental na reconstrução da complexa história biogeográfica dos Neotrópicos.

Ao longo dos últimos 40 milhões de anos, eventos geológicos e climáticos moldaram a biodiversidade em diferentes grupos americanos, e em especial os euglossíneos. Transgressões e regressões marinhas, soerguimento de cordilheiras, derramamentos basálticos e múltiplas glaciações, isolados e em conjunto, constituem provavelmente os principais eventos estocásticos e vicariantes por trás da atual distribuição das espécies e suas respectivas estruturas populacionais.

Rastrear precisamente a influência de cada um destes eventos para todas as espécies ao longo de toda a América é tarefa hercúlea. No entanto, ao longo dos textos e capítulos a seguir, tentar-se-á mensurar e inferir parte da história de uma das principais florestas tropicais das Américas, a Mata Atlântica, do ponto de vista das carismáticas abelhas das orquídeas. Ao reconstruir a história biogeográfica de algumas espécies da tribo Euglossini, espera-se incluir informações que ajudem a compreender, ratificar e explorar novos padrões e versões da história das formas biológicas ao longo do tempo e espaço geográfico, hoje conhecido como Mata Atlântica.

Sem dúvida, é uma grande responsabilidade fazer jus ao incrível legado deixado por ilustres pesquisadores que têm construído o conhecimento sobre as abelhas das orquídeas ao longo de tantos anos, bem como para o “*status quo*” da Biogeografia da América do Sul, em particular da Mata Atlântica. Em meio a tantas importantes contribuições, tenta-se aqui também dar uma contribuição para o preenchimento da lacuna sobre o conhecimento da história filogeográfica dos euglossíneos, *dando voz a estas abelhas* para contar a sua versão da história biogeográfica das florestas úmidas sul-americanas.

O presente estudo nasceu com o intuito de contribuir sobre múltiplos aspectos da biologia dessas abelhas. Na esperança de que seja útil à comunidade científica e que contribua para a conservação deste peculiar grupo de abelhas, é com muita estima que apresento este estudo após quatro anos de intenso trabalho, mais de 55.000 km percorridos e mais de 3.500

horas em expedições de campo. Ao longo deste período, tive a oportunidade de observar de perto aspectos fascinantes da biologia dos euglossíneos, os quais moldaram muitos dos conceitos e ideias aqui discutidos.

A presente tese visa oferecer um estudo comparado dos padrões filogeográficos e genético-populacionais em três diferentes espécies e grupo de espécies de euglossíneos com ênfase na Mata Atlântica. Para este fim, o estudo que segue foi estruturado em três capítulos em forma de manuscritos, incluindo um estudo já publicado antes da defesa desta tese, mas aqui incorporado sob autorização da editora detentora dos direitos de distribuição em uma versão adaptada e traduzida devido ao seu alto teor informativo em relação às conclusões gerais deste estudo.

Cada um dos capítulos foi dedicado a um caso específico, os quais representam diferentes modelos de distribuição e nicho ao longo da Mata Atlântica. Precedendo a exploração dos resultados empíricos é também oferecida uma ampla revisão de literatura sobre os temas que permeiam este estudo, visando facilitar a comparação entre os distintos métodos e resultados apresentados. Ao final, os padrões explorados ao longo dos diferentes estudos são sumarizados, buscando delinear possíveis convergências entre os diferentes modelos analisados.

Finalmente, considerando os atuais padrões internacionais de produção científica e o decorrente fato de que toda a produção empírica associada a este estudo está ou será também publicada em língua inglesa, pede-se licença à praxe contemporânea para apresentar os textos a seguir integralmente em língua vernácula como forma de atender um dos objetivos dos órgãos financiadores deste estudo, a popularização do conhecimento científico.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Com base no exposto, o presente trabalho tem como objetivo investigar os padrões de estruturação e diversidade genética de populações de Euglossini além de contribuir com a compreensão dos padrões de distribuição filogeográficos das espécies de euglossíneos.

2.2 OBJETIVOS ESPECÍFICOS

- Investigar a diversidade e estrutura genética de três espécies ou grupo de espécies de Euglossini: *Euglossa iopoecila* Dressler (1982), *Euglossa annectans* Dressler (1982) e o grupo *crassipunctata* Dressler (1982), espécies com ocorrência restrita à Mata Atlântica.
- Evidenciar os padrões de estruturação genética em três diferentes tipos de distribuição geográfica: Litorânea, interiorana e continental disjunta.
- Avaliar os padrões de variação filogeográfica por meio de marcadores moleculares mitocondriais e nucleares.
- Avaliar a influência das flutuações climáticas do Quaternário sobre a estruturação genética e filogeográfica das diferentes espécies estudadas.



Efram Goldberg

PARTE I

FUNDAMENTAÇÃO TEÓRICA

3 FUNDAMENTAÇÃO TEÓRICA

3.1 UMA BREVE INTRODUÇÃO À BIOGEOGRAFIA

O conceito por trás da ideia de Biogeografia é aparentemente simples e poderia ser definido prontamente como o estudo da distribuição geográfica dos organismos. Ou ainda nas palavras de Hampshire e Parenti (1999): “*What leaves where and why?*”; em uma tradução livre: “*o que vive onde e por quê?*”. Embora simples e precisa, esta definição pode esconder a complexidade da disciplina (POSADAS; CRISCI; KATINAS, 2006). A Biogeografia possui várias vertentes e subáreas que evoluíram ao longo de seus mais de dois séculos, passando por uma vertente descritiva e dispersionista a uma linha mais interpretativa, a qual busca explicações causais para possíveis padrões, testando hipóteses e modelos (BROWN, 2009; CARVALHO; ALMEIDA, 2016).

Estudos desta linha interpretativa, mais moderna, desenvolveram-se de forma clássica seguindo uma das duas principais tradições de pesquisa biogeográfica: Biogeografia Ecológica ou Histórica. Essas tradições têm sido reconhecidas desde 1820 pelo botânico suíço Agustín Pyramus de Candolle. Cada uma destas linhas evoluiu de forma independente, resultando em dois caminhos metodológicos e epistemológicos bastante distintos. No entanto, pode-se dizer que a principal diferença entre Biogeografia Ecológica e Histórica está na escala de tempo em que ambas trabalham. Enquanto a Biogeografia Ecológica procura explicações causais dos padrões distributivos em escalas temporais mais curtas, incluindo dezenas de centenas ou mesmo alguns milhares de anos, a Biogeografia Histórica trata de processos evolutivos ao longo de milhões de anos (BROWN, 2009).

Desta forma, as duas vertentes usam diferentes conceitos na tentativa de interpretar a distribuição de organismos no espaço geográfico. A Biogeografia Ecológica está preocupada com grupos funcionais de espécies e restrições ambientais, já a Biogeografia Histórica se concentra em grupos taxonômicos e eventos geológicos (CRISCI et al., 2006; POSADAS; CRISCI; KATINAS, 2006). Conseqüentemente, a história desta disciplina foi influenciada pelos paradigmas evolutivos decorrentes de cada uma destas interpretações.

Passados mais de 200 anos desde o início da Biogeografia, tentativas de integrar estas duas linhas, Biogeografia Histórica e Biogeografia Ecológica, ainda não foram completamente bem sucedidas (BROWN, 2009; POSADAS; CRISCI; KATINAS, 2006), embora tenham ganhado força à medida que abordagens genético-evolutivas criam uma interface entre estas duas áreas (AVISE, 2009; CRISCI et al., 2006; RISSLER, 2016).

Nas últimas duas décadas, o “tempo” como um fator adquiriu uma grande relevância em estudos biogeográficos, demonstrando uma considerável sinergia com a Biogeografia Histórica para gerar hipóteses sobre a origem, evolução e distribuição de organismos (AVISE, 2009; BROWN, 2009; CARVALHO; ALMEIDA, 2016). Por outro lado, cada vez mais fatores ecológicos mostram sua importância para determinação de grandes padrões biogeográficos, vide os reflexos da teoria de refúgios florestais de Haffer (1989) e suas derivações (CARNAVAL; MORITZ, 2008) na explicação de padrões biogeográficos em florestas tropicais (BATALHA-FILHO; MIYAKI, 2011; CARNAVAL et al., 2009; CARNAVAL; MORITZ, 2008; MARTINS, 2011; RAPOSO DO AMARAL et al., 2013).

Embora haja discordâncias históricas entre as diferentes abordagens biogeográficas, a ecologia e a própria história estão unidas indissolavelmente, e portanto, a divisão há muito estabelecida entre essas duas ideias principais é um obstáculo para o completo entendimento dos padrões biogeográficos (AVISE, 2009; BROWN, 2009; CARVALHO; ALMEIDA, 2016; CRISCI et al., 2006; POSADAS; CRISCI; KATINAS, 2006). Deste modo, qualquer aplicação dos conhecimentos biogeográficos apenas será completa quando ambas perspectivas forem consideradas (CRISCI et al., 2006).

A Biogeografia também tem grande potencial para prestar serviços à conservação da biodiversidade e tem ganhado destaque nesta área na última década (NETO; LOYOLA, 2016), uma vez que uma das principais tarefas para a conservação é a delimitação de áreas prioritárias considerando onde as espécies atualmente vivem, viveram no passado e viverão no futuro (FRANKHAM et al., 2003, 2010; VERNESI et al., 2008). Esta perspectiva é claramente biogeográfica e tão proeminente que a Biogeografia da Conservação já completa mais de uma década como Ciência distinta no amplo espectro biogeográfico (NETO; LOYOLA, 2016).

Vários métodos biogeográficos históricos e ecológicos são usados para definir estratégias de conservação em diferentes escalas e níveis hierárquicos, a exemplo da Análise de Parcimônia para detectar Endemismo (PAE), da Panbiogeografia, Filogeografia e Modelagem de Nicho Ecológico (BROWN, 2009; CARVALHO; ALMEIDA, 2016; CRISCI et al., 2006; NETO; LOYOLA, 2016). No entanto, como ressaltado por Crisci et al. (2006), a biogeografia será plenamente útil para a conservação se o quadro conceitual incluir as influências antrópicas nas distribuições dos organismos, na mudança climática, no uso da terra, introduções e decorrentes invasões biológicas.

Em suma, a distribuição das formas, pelo espaço em função do tempo têm influência de inúmeros fatores e tantas interpretações quanto as possíveis óticas relacionadas. Assim,

acumulam-se pela história tantos modos de se olhar para os eventos biogeográficos quanto o número de fatores envolvidos.

3.1.1 Principais Ideias e Métodos Biogeográficos

A intenção deste tópico é traçar um paralelo entre as principais ideias biogeográficas e suas implicações para interpretações dos padrões dos organismos estudados. Abordar como as principais vertentes biogeográficas se definem, em quais contextos foram definidas e como se interconectam. Deste modo, será apresentada uma breve definição das principais ideias biogeográficas, bem como seus pontos sinérgicos na interpretação de padrões gerais.

3.1.1.1 Biogeografia Dispersalista

Partindo da Biogeografia Histórica (forma, tempo e espaço), uma das ideias mais simples do ponto de vista biogeográfico é o Dispersalismo (BROWN, 2009). De fato, esta é uma ideia que permeou o âmbito biogeográfico desde os primórdios desta ciência, quando se considerava o Éden um centro de origem das espécies (RIDDLE, 2017). Esta corrente assume a dispersão e extinção ao longo do processo e seu principal argumento é a história de distribuição de grupos individuais (MORRONE, 2008).

Embora o Dispersalismo tenha sido uma ideia bastante vigente até a década de 1970, existem várias críticas a esta visão biogeográfica (MORRONE; CRISCI, 1995). De fato, os problemas em relação ao Dispersalismo começam a ocorrer já na determinação do centro de origem. Não há uma unicidade de métodos para encontrar o centro de origem, e algumas alternativas são propostas: a) a área em que os organismos mostram grande diversidade ou riqueza de espécies; b) área onde ocorre a forma mais primitiva; c) área ocupada pelo organismo mais primitivo; d) ou o mais derivado filogeneticamente; e) área onde ocorre o mais antigo fóssil; f) área onde está o representante de maior tamanho (MORRONE, 2008).

É interessante notar que cada um dos métodos sumarizados pode ser sistematicamente contestado. Caso seja assumido, por exemplo, que o centro de origem seja a área com maior diversidade ou riqueza, os Neotrópicos deveriam ser o centro de origem para a maior parte das espécies. Por outro lado, caso considere-se a presença do fóssil mais antigo como área de origem, dificilmente os Neotrópicos teriam dado origem a alguma espécie, já que essa ideia desconsidera as diferentes adequabilidades regionais para a formação de fósseis. Ainda, caso se considere a região do táxon mais primitivo ou derivado, excluisse a arbitrariedade da polarização de caracteres ou ainda a conservação diferencial de nichos (COX; MOORE; LADLE, 2016; CROIZAT, 1964; MORRONE; CRISCI, 1995; RIDDLE, 2017).

As críticas ainda podem ser resumidas em basicamente duas: (a) Falta de base científica: uma vez que qualquer padrão de distribuição pode ser explicado invocando um número suficiente de eventos de dispersão; hipóteses baseadas na dispersão não podem ser refutadas cientificamente mesmo dentro de um quadro hipotético-dedutivo rigoroso. Além disso, se aceitarmos a dispersão como uma possível explicação para padrões disjuntos, as explicações de vicariância nunca serão inferidas (CROIZAT, 1964; MORRONE, 2008; MORRONE; CRISCI, 1995). (b) Falta de poder preditivo: hipóteses baseadas em dispersão são cenários idiossincráticos que só podem explicar a história biogeográfica de linhagens individuais a exemplo de Brudin (1966), mas não podem fornecer uma teoria geral para explicar como organismos com diferentes ecologias e habilidades de dispersão chegaram a ocupar as mesmas regiões biogeográficas e/ou a exibir padrões de distribuição semelhantes (CROIZAT, 1964; MORRONE; CRISCI, 1995; SANMARTÍN, 2012).

Contudo, como será descrito no próximo tópico, isso não é necessariamente verdadeiro para todos os casos, e a dispersão às vezes pode gerar padrões de distribuição congruentes, semelhantes aos esperados por vicariância. Ainda assim, devido a estas e outras críticas o Dispersalismo perdeu força no fim da década de 1970 com a ascensão das ideias vicariantes de Croizat.

3.1.1.2 *Biogeografia Cladística*

Na segunda metade do século XX, duas revoluções científicas contribuíram para o aparecimento de um novo paradigma na Biogeografia Histórica. A primeira revolução foi o surgimento da Cladística (HENNIG, 1966). Tratava-se de um novo método para reconstruir as relações evolutivas entre os organismos com base em características biológicas compartilhadas e derivadas (“plesiomorfias” e “apomorfias”) utilizando o princípio da Parcimônia para decidir entre hipóteses concorrentes de relações evolutivas (COX; MOORE; LADLE, 2016; LOMOLINO; SAX; BROWN, 2004). A segunda revolução foi a redescoberta e desenvolvimento da “Teoria da Tectônica de Placas” de Wegener, quando foi possível provar a proposição de que a camada externa da Terra, a litosfera, é dividida em placas rochosas rígidas, constituídas por uma crosta continental e oceânica que se deslocam sobre a superfície da Terra, deslizando sobre o manto terrestre. A ideia de que as espécies pudessem andar passivamente nos continentes enquanto se separavam e se espalhavam pela superfície da Terra levou ao conceito de “vicariância” (COX; MOORE; LADLE, 2016).

A *Biogeografia Cladística de Vicariância* (PLATNICK; NELSON, 1978) nasceu da fusão da Cladística de Hennig (1966) e do conceito de Vicariância de Croizat (1964). No

modelo vicariante, uma área ancestral é dividida pelo surgimento de sucessivas barreiras geográficas. Uma vez que cada divisão geográfica teria sido seguida de especiação alopátrica, pode-se reconstruir a sequência de eventos de vicariância a partir da sequência de eventos cladogênicos (especiação) na filogenia das linhagens endêmicas da área (CARVALHO; ALMEIDA, 2016).

Uma análise biogeográfica cladística começa com a construção de um “cladograma de áreas”, no qual o nome dos táxons na filogenia é substituído pelas áreas onde ocorrem. Se todo táxon é endêmico em uma área e cada área abriga um táxon, a construção de cladogramas de áreas é trivial, porém, estes podem incluir táxons generalizados (presentes em mais de uma área), o que introduz ambiguidades na inferência do cladograma, já que essas áreas poderiam ocupar diferentes posições no cladograma de áreas, a depender de quais relações são permitidas entre as áreas ocupadas pelos táxons não endêmicos (COX; MOORE; LADLE, 2016; SANMARTÍN, 2012).

No entanto, assim como o Dispersalismo, a Biogeografia Cladística tornou-se ao longo do tempo uma explicação reducionista, uma vez que negava qualquer papel importante da dispersão na geração de padrões globais de biodiversidade. Segundo a ótica cladística vicariante, a dispersão tornou-se um fenômeno raro e aleatório que afetou linhagens individuais, mas não produziu padrões de distribuição congruentes (BROWN, 2009; COX; MOORE; LADLE, 2016). Deste modo, devido a suas idiosincrasias, as histórias de dispersão não poderiam ser abordadas dentro de um quadro hipotético-dedutivo como a Biogeografia Cladística. Esta visão da dispersão como um evento único em vez de um processo gerador de padrões é desafiada pelo caso de ilhas oceânicas de origem vulcânica, como o Arquipélago havaiano, que só poderia ter sido colonizado por dispersão (MACARTHUR; WILSON, 1967; SANMARTÍN, 2012).

Outra importante crítica à Biogeografia Cladística é que ela ignora os processos na inferência biogeográfica (SANMARTÍN, 2012). Os cladogramas de áreas são inferidos sem quaisquer referências aos eventos biogeográficos subjacentes. Os métodos cladísticos se concentram em encontrar padrões de relação entre as áreas de endemismo que são posteriormente interpretadas em termos de eventos, e esta sequência (descobrir um padrão, depois inferindo sua causa) é o principal fundamento da abordagem biogeográfica cladística (COX; MOORE; LADLE, 2016; MORRONE; CRISCI, 1995; POSADAS; CRISCI; KATINAS, 2006).

3.1.1.3 Biogeografia baseada em eventos

Os métodos baseados em eventos foram desenvolvidos como resposta às limitações da Biogeografia Cladística (COX; MOORE; LADLE, 2016; CRISCI; KATINAS; POSADAS, 2003; RONQUIST, 1997a; SANMARTÍN, 2012, 2016). Esses métodos usam uma abordagem determinista baseada em modelos de custo para cada evento ou processo biogeográfico (ex. dispersão ou vicariância). A lógica por trás destes modelos é que para cada evento cladogênico é atribuído um determinado custo de acordo com sua probabilidade de ocorrência. Por exemplo, pode-se atribuir valores de custo para vicariância e dispersão em uma determinada filogenia de acordo com a presença ou não de táxons em uma determinada área.

Além da dispersão e da vicariância, os modelos de custo biogeográfico incluem dois processos adicionais, a extinção (desaparecimento de uma linhagem de parte de sua distribuição ancestral) e duplicação (“especiação dentro da área”), que ocasionalmente é equiparada à especiação simpátrica ou especiação alopátrica em resposta a uma barreira de dispersão/fluxo gênico temporária que afete uma linhagem única de um organismo (BROWN, 2009; SANMARTÍN, 2012). Ao ajustar a filogenia do organismo a um cladograma de áreas refletindo as relações entre as áreas de endemismo, pode-se obter a reconstrução mais parcimoniosa dos eventos biogeográficos para explicar o padrão de distribuição observado (CRISCI; KATINAS; POSADAS, 2003; POSADAS; CRISCI; KATINAS, 2006; SANMARTÍN, 2012).

Reconstruções filogeográficas baseada em eventos não só especificam o conjunto de eventos possíveis (dispersão, extinção, duplicação e vicariância) que levaram a esses padrões de distribuições biogeográficas, mas também seu tempo relativo. Ou seja, é possível inferir não apenas os prováveis eventos que deram origem aos táxons em suas respectivas regiões, mas também a ordem destes eventos ao longo da filogenia do grupo (CRISCI et al., 2006; SANMARTÍN, 2012). Neste sentido, cladogramas de áreas podem ser inferidos pela busca do padrão de relações de área que melhor explica a distribuição da filogenia dos táxons terminais ou, em outras palavras, o padrão de relação de áreas que produz a reconstrução biogeográfica com o menor custo em termos de eventos biogeográficos inferidos (SANMARTÍN, 2012).

A principal diferença entre a reconstrução biogeográfica baseada em eventos e a abordagem biogeográfica cladística tradicional é que os processos não são inferidos *a priori* no cladograma de áreas, mas sim *a posteriori*, baseada no modelo de matriz de custo. Deste modo, um modelo biogeográfico diferente com custos diferentes pode resultar em um padrão

diferente de relações de área e um conjunto diferente de eventos biogeográficos (SANMARTÍN, 2007). Essa conexão explícita entre o modelo baseado em processo e os padrões esperados permite comparar hipóteses biogeográficas alternativas ou cenários dentro de uma estrutura de inferência estatística. Por exemplo, pode-se comparar a frequência observada de eventos biogeográficos com uma distribuição nula obtida pela aleatorização das distribuições na filogenia original para testar se os padrões observados são conservados filogeneticamente (SANMARTÍN, 2012).

Uma das principais críticas a este tipo de abordagem baseada em eventos é sobre como decidir o custo dos eventos biogeográficos (COX; MOORE; LADLE, 2016). O método mais comum é usar um critério de otimização baseado em parcimônia e escolher custos que maximizem as chances de recuperar padrões de distribuição “conservados filogeneticamente”, isto é, padrões de distribuição que não mudem de ancestrais para descendentes e são conservados ao longo da filogenia (SANMARTÍN, 2007). Contudo, isto é equivalente à explicação mais parcimoniosa para o padrão biogeográfico observado, como na Biogeografia Cladística tradicional (MORRONE, 2005; RIDDLE; HAFNER, 2006; SANMARTÍN, 2012), dado que a vicariância é um processo que produz padrões de distribuição conservados entre as linhagens (BROWN, 2009).

Sem dúvida, um dos métodos mais populares desenvolvidos para biogeografia baseados em eventos é a análise de dispersão-vicariância (DIVA; RONQUIST, 1997b). Esta análise usa uma abordagem de matriz de custo em que os eventos de extinção e dispersão são atribuídos a um custo maior em relação à vicariância e duplicação (RONQUIST, 1997b; SANMARTÍN, 2012). Porém, ao contrário dos outros métodos, o DIVA não se baseia em encontrar um cladograma de áreas, mas distribuições de áreas de mapeamento ao longo da filogenia e, a partir deste, inferir as áreas ancestrais em eventos cladogenéticos, minimizando o número de eventos de dispersão e extinção (SANMARTÍN, 2012).

Embora métodos baseados em eventos, como o DIVA, sejam amplamente utilizados (COX; MOORE; LADLE, 2016; DA SILVA; NOLL, 2015; RAMÍREZ et al., 2010; SANMARTÍN, 2007), uma forte crítica a estes métodos é a simplificação dos dados, impondo um modelo específico, ou seja, a abordagem de matriz de custo (BROOKS, 2005). Outros argumentam que a inferência biogeográfica deve se concentrar em encontrar padrões, não em busca de processos explicativos ad hoc (PARENTI, 2007). Contudo Sanmartín (2012) argumenta que a conexão explícita entre os processos e os padrões esperados se torna mais fácil para os métodos baseados em eventos avaliar hipóteses biogeográficas alternativas.

Uma segunda crítica contra esses métodos reside na inferência de processos que não estão ligados à especiação, como a extinção (SANMARTÍN, 2007). Eventos de extinção completa, em que uma linhagem desaparece de toda a distribuição ancestral, ou eventos de dispersão completos, em que uma linhagem deixa a zona de distribuição ancestral para colonizar uma área diferente sem necessariamente a ocorrência de especiação, não podem ser inferidos com métodos baseados em eventos, pois estes eventos não deixam rastros na filogenia (BROOKS, 2005; SANMARTÍN, 2007, 2012). Para inferir um evento de dispersão ou extinção (em áreas), estes eventos devem estar vinculados a um evento de especiação, ou seja, a dispersão e a extinção devem deixar pelo menos um descendente na área original para ser rastreado a partir da filogenia (SANMARTÍN, 2012). De fato, esse problema afeta todos os métodos biogeográficos que são baseados na parcimônia, incluindo a biogeografia cladística.

3.1.1.4 *Biogeografia Paramétrica ou Probabilística*

Assim como a Biogeografia baseada em eventos surgiu para suprir algumas deficiências da Cladística tradicional em considerar processos (dispersão, extinção, duplicação e vicariância), os métodos paramétricos ou probabilísticos surgiram para incorporar três fatores principais às análises biogeográficas: tempo, independência da parcimônia e suporte estatístico (SANMARTÍN, 2012, 2016).

Esses métodos são denominados “baseados em modelos” ou “paramétricos” porque se baseiam em modelos estatísticos de evolução das distribuições, cujos parâmetros (“variáveis”) são os próprios processos biogeográficos, como dispersão, expansão de distribuição ou extinção (SANMARTÍN, 2012). A evolução da distribuição (a mudança na distribuição geográfica do ancestral para os descendentes) é modelada como um processo estocástico que muda ao longo dos ramos das árvores filogenéticas de acordo com um modelo probabilístico de “Cadeias de Markov” (SANMARTÍN, 2016).

No cerne de um modelo baseado em Cadeias de Markov há uma “matriz de probabilidades de transição” que determina recursivamente a taxa de mudança de um estado para outro (HOLDER; LEWIS, 2003; HUELSENBECK et al., 2001; LEMMON; MORIARTY, 2004). Em modelos biogeográficos, os estados das Cadeias de Markov são o conjunto de áreas geográficas que formam a faixa de distribuição do grupo, e os parâmetros do modelo são os processos biogeográficos que alteram a distribuição geográfica das espécies, como a retração (extinção, em uma área) ou a expansão da distribuição (SANMARTÍN, 2012). Ao permitir que o modelo evolua ao longo dos ramos da filogenia (do ancestral para os

táxons terminais), os quais representam o tempo desde a cladogênese, pode-se estimar as taxas (probabilidade/frequência) de ocorrência dos processos biogeográficos e inferir a distribuição ancestral mais provável em cada evento cladogênico (SANMARTÍN, 2016).

Em comparação às abordagens anteriores, esses métodos oferecem várias vantagens (SANMARTÍN, 2016; SANMARTIN et al., 2010). O mais óbvio é que a frequência (taxa) de eventos pode ser estimada a partir dos dados, em vez de atribuir um custo *a priori* utilizando critérios *ad hoc*, como em métodos biogeográficos baseados em eventos (REE; SMITH, 2008). Existem também outras vantagens, como a capacidade de incorporar melhor as incertezas na distribuição dos ancestrais nas filogenias, uma vez que não assumem apenas o cenário mais parcimonioso, e sim avaliam todas as possibilidades de áreas ancestrais (SANMARTIN et al., 2010). Além disso, talvez a vantagem mais importante dos métodos paramétricos seja a sua capacidade de integrar nas estimativas de inferência biogeográfica a divergência evolutiva entre linhagens ou o tempo desde a cladogênese, que são representados pelo comprimento dos ramos na filogenia (SANMARTÍN, 2012, 2016).

Por fim, os métodos paramétricos também fornecem uma abordagem estatística mais apropriada para comparar hipóteses ou cenários biogeográficos alternativos (SANMARTÍN, 2012). Como os parâmetros de cada modelo alternativo são processos biogeográficos, pode-se identificar os processos que melhor explicam os padrões biogeográficos identificando o modelo com “melhor ajuste” através de testes estatísticos baseados em probabilidades, motivo pelo qual estes métodos também são chamados de probabilísticos (SANMARTÍN, 2016; SANMARTÍN; VAN DER MARK; RONQUIST, 2008). Isso contrasta com o uso de testes de permutação aleatória em biogeografia baseada em eventos, nos quais os padrões observados só podem ser comparados com aqueles esperados por acaso (SANMARTÍN, 2012; SANMARTIN et al., 2010).

Métodos biogeográficos paramétricos têm evoluído rapidamente desde a última década, dentre os quais duas abordagens ganham destaque: o modelo de Dispersão-Extinção-Cladogênese (DEC) (REE; SMITH, 2008) e o modelo Bayesiano de Biogeografia de Ilhas (BIB) (SANMARTÍN; VAN DER MARK; RONQUIST, 2008). O método DEC permite estimar, por máxima verossimilhança, as taxas de expansão (dispersão) e retração (extinção) em cenários de evolução de áreas ao longo de eventos cladogênicos a partir de filogenias datadas (REE; SMITH, 2008). Por outro lado, o método BIB usa inferência Bayesiana para estimar as capacidades de carga (suporte) das áreas e as taxas de dispersão entre ilhas a partir de dados filogenéticos e de distribuição de múltiplos grupos co-distribuídos (SANMARTÍN; VAN DER MARK; RONQUIST, 2008). Este método é especialmente útil para acomodar as

diferenças de idade, taxa de evolução e capacidade de dispersão entre linhagens (SANMARTIN et al., 2010).

Nesta mesma linha existem ainda extensões destes modelos como o GeoSSE (OLDBERG et al., 2011), o qual é uma extensão do DEC desenvolvida para acomodar taxas de extinção e especiação independentes em cada área. Em meio a tantas possibilidades paramétricas, existem ainda *softwares* destinados a checar a adequabilidade de cada modelo, buscando retornar o modelo que melhor se adeque ao conjunto de dados. Este é o caso do pacote BioGeoBEARS, o qual testa vários modelos diferentes implementando ainda versões de máxima verossimilhança para o DIVA (DIVALIKE) e BayArea (BAYAREALIKE) (MATZKE, 2013).

Assim como acontece com os métodos baseados em eventos, os métodos paramétricos são criticados pela dependência de um modelo biogeográfico particular, que é visto como uma limitação (SANMARTÍN, 2012). Os modelos biogeográficos, no entanto, são melhor vistos não como restrições sobre os dados, mas como hipóteses alternativas para explicar os dados (SANMARTÍN; VAN DER MARK; RONQUIST, 2008).

Finalmente, uma limitação mais atual das abordagens paramétricas é que elas pressupõem que a evolução da distribuição está desacoplada com a diversificação da linhagem (SANMARTÍN, 2012). Em métodos paramétricos, a evolução da distribuição é modelada como um processo que evolui ao longo dos ramos de uma árvore filogenética, mas o próprio modelo biogeográfico não influencia o processo estocástico de extinção-cladogênese que determina o crescimento da árvore, comprimentos de ramos e topologia (REE; SMITH, 2008; SANMARTÍN; VAN DER MARK; RONQUIST, 2008). Embora esta não seja uma limitação que exclua a possibilidade de uso dos modelos paramétricos, provavelmente a evolução destes algoritmos encaminhe-se para suprir essa carência (SANMARTÍN, 2016).

Em geral, métodos diferentes partem de pressupostos, ideais e condições diferentes (SANMARTÍN, 2012; YU et al., 2015). Deste modo, é recomendável que análises biogeográficas avaliem os padrões sobre o maior número de modelos possíveis para observar suas congruências e dissidências, e assim, alcançar os resultados mais verossímeis possíveis (YU et al., 2015).

3.1.1.5 *Integração de processos ecológicos nas análises biogeográficas*

A integração de processos ecológicos na reconstrução de cenários biogeográficos tem sido um objetivo de longo prazo na biogeografia histórica (CRISCI et al., 2006; MORRONE;

CRISCI, 1995). Porém, até recentemente, poucos esforços reais haviam sido realizados para combinar esses dois aspectos historicamente opostos (veja os tópicos 4.1 e 4.1.1) da disciplina em um quadro analítico comum. Isso se tornou possível através do desenvolvimento de técnicas de modelagem de nicho ecológico (ENM; *Ecological Niche Modelling*) (SANMARTÍN, 2012).

De acordo com o conceito de conservação de nicho (WIENS, 2004; WIENS; DONOGHUE, 2004), as linhagens tendem a conservar seu nicho ecológico ao longo do tempo, ou seja, o conjunto de condições ambientais pelas quais as linhagens podem reproduzir e manter populações viáveis. Neste contexto, a vicariância é considerada o resultado de qualquer mudança ambiental que cause uma divisão em uma área geográfica de espécies (WIENS, 2004).

Os ENMs (KOZAK; GRAHAM; WIENS, 2008; PETERSON, 1999) usam a associação entre dados de distribuição (ocorrências de espécies) e variáveis ambientais (por exemplo, temperatura, precipitação) para prever o intervalo dentro do qual uma espécie poderia ocorrer (veja a sessão de Modelagem de Nicho Ecológico). Supondo que os nichos são preservados ao longo do tempo (“conservação do nicho”) e com informações sobre os climas passados, pode-se projetar o nicho ecológico para diferentes pontos no tempo, reconstruindo padrões de distribuição de espécies passadas (PHILLIPS; ANDERSON; SCHAPIRE, 2006; TRUJILLO-ARIAS et al., 2017; VARELA; LOBO; HORTAL, 2011), ou para encontrar áreas que estavam no passado dentro da tolerância ambiental das espécies e poderiam ter atuado como corredores de dispersão em regiões que agora são inabitáveis (VARELA; LOBO; HORTAL, 2011).

Os primeiros trabalhos a combinar métodos biogeográficos paramétricos com dados paleoclimáticos e modelos de nicho ecológico de táxons existentes (SMITH; DONOGHUE, 2010) utilizaram estes métodos combinados como forma de entender como os climas passados e as conexões terrestres moldaram a distribuição biogeográfica das linhagens ao longo do tempo. Da mesma forma, outros estudos foram pioneiros na integração dos modelos ENM na análise paleobiogeográfica, através dos modelos de nicho de combinação com o registro fóssil de linhagens extintas (VARELA; LOBO; HORTAL, 2011).

A vantagem desta última abordagem em relação aos modelos ENM com base em taxas existentes é que as reconstruções ENM baseadas em fóssil não assumem que os nichos ecológicos sejam estáveis ao longo do tempo (isso pode ser verdade para escalas ecológicas, mas não sobre escalas geológicas de milhões de anos), então esses modelos podem ser usados para rastrear padrões de conservação de nicho ao longo de escalas de tempo (SANMARTÍN,

2012). Por outro lado, registros fossilíferos são geralmente escassos e esparsos ao longo do tempo, sendo dificilmente uma informação de fácil acesso.

3.1.2 Filogeografia e seu impacto sobre a Biogeografia

Inegavelmente os métodos e análises discutidos até aqui constituem a base de todo pensamento Biogeográfico, e, conseqüentemente, formaram a base teórica para o florescimento de uma Biogeografia mais refinada, focada em padrões específicos (COX; MOORE; LADLE, 2016; RIDDLE, 2017; SANMARTÍN, 2012, 2016). Nos últimos trinta anos, os avanços em genética molecular tornaram a Biogeografia Filogenética e Genética de Populações áreas cada vez mais sinérgicas, criando pontes para a compreensão e integração entre processos em escala populacional ao longo de gerações (microevolutivos) e aqueles envolvendo gêneros, famílias e ordens ao longo de milhões de anos (macroevolutivos) (AVISE, 2009; CRISCI et al., 2006). No entanto, poucos trabalhos foram tão marcantes quanto os liderados por John Avise, os quais são emblematicamente representados pela publicação de Avise et al. (1987), fundamental para criar a principal ponte entre as ciências dedicadas a processos em escala geológica e aqueles relativos a poucas gerações, nomeada por estes autores de “Filogeografia” (Figura 3.1).

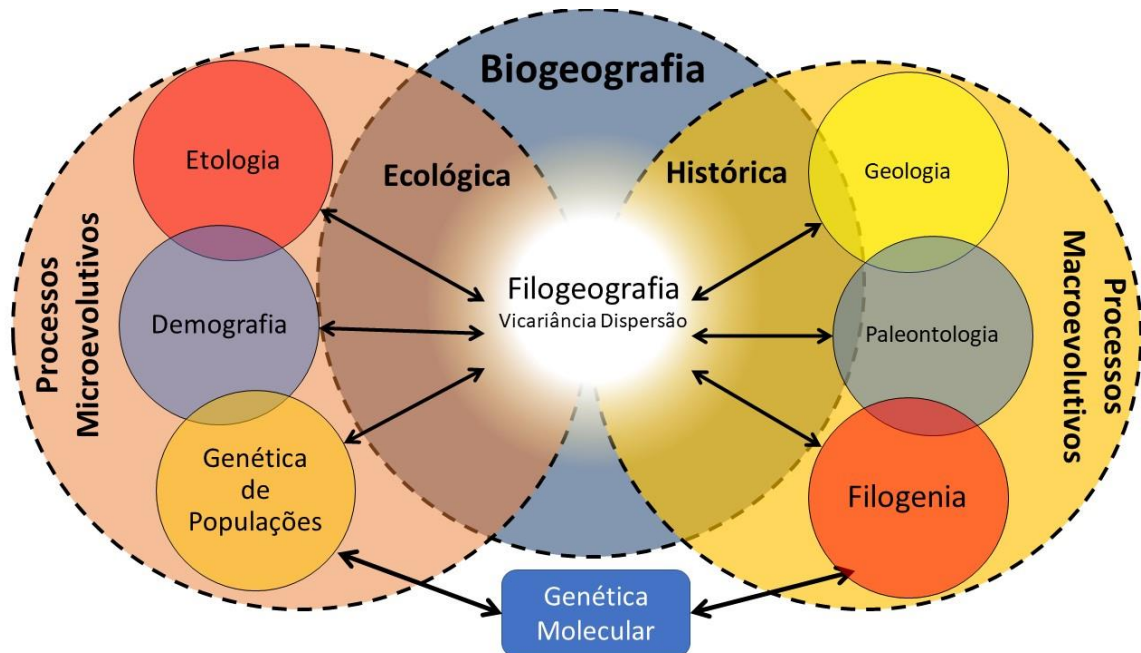


Figura 3.1 Relação conceitual da Filogeografia como ponte empírica entre as ciências da Biodiversidade. Adaptado de Avise (2009) e Riddle (2009).

Avise e colaboradores (1987) cunharam o nome Filogeografia para tratar da distribuição geográfica de linhagens genéticas ao longo do tempo e espaço geográfico, mostrando como processos intraespecíficos poderiam gerar padrões similares aos grandes

padrões biogeográficos (AVISE et al., 1987). Em outras palavras, Avise e colaboradores chamaram atenção para o fato de que processos que afetavam linhagens genéticas intraespecíficas poderiam ser extrapolados para táxons de níveis superiores (AVISE et al., 1987).

Da forma com a qual Avise a concebeu, a Filogeografia é uma disciplina que lida com os arranjos espaciais das linhagens genéticas, especialmente entre populações e espécies intimamente relacionadas (AVISE, 2009). Neste contexto, a Filogeografia visa testar hipóteses quanto a relação causal entre os fenômenos geográficos, as distribuições de espécies e os mecanismos que conduzem à especiação ou à estruturação das linhagens intraespecíficas (HICKERSON et al., 2010). Conseqüentemente, a análise de dados genéticos de múltiplos táxons co-distribuídos pode contribuir para questões mais profundas sobre os fenômenos geográficos, geológicos e/ou climáticos que geraram a distribuição observada da biodiversidade, em uma “Filogeografia Comparada” (HICKERSON et al., 2010). Assim, a abordagem proposta por Avise oferece a oportunidade de utilizar as populações como experimentos naturais independentemente submetidos aos mesmos cenários geológicos e climáticos.

Devido ao seu evidente caráter multidisciplinar, a Filogeografia teve dificuldades de se comunicar com os pesquisadores das áreas adjacentes e tomar uma clara posição entre as disciplinas mais tradicionais (Figura 3.1; AVISE, 2000). Se por um lado a Filogeografia poderia ser considerada um braço da Genética de Populações, uma vez que trata questões co-específicas (em sua maioria), a mesma se difere desta última por se concentrar nas relações genealógicas dos alelos (não apenas frequências), as quais podem ser interpretadas como microfílogias. Logo, deste ponto de vista, a Filogeografia poderia então ser considerada parte da Biologia Filogenética, porém se difere desta última por se concentrar nos processos populacionais históricos e demográficos. Exatamente devido a esse contexto histórico e causal da distribuição geográfica dos grupos, a Filogeografia poderia então ser considerada puramente um braço da Biogeografia, a qual é, nas palavras de John Avise: “*talvez seu lar mais natural...*”, ainda que se distinga desta última, por se concentrar em populações co-específicas e informações genealógicas (AVISE, 2009). Fato é que a Filogeografia é uma disciplina de fronteira, a qual se encarrega de conectar muitas áreas históricas e didaticamente desconectadas, que provavelmente encontra na Biogeografia seu centro de distribuição, mas que tende, cada vez mais, a expandir seus domínios.

O tempo trouxe tal importância à Filogeografia na Biogeografia moderna e a mesma se tornou foco de vários periódicos em sessões especialmente dedicadas a este tópico, como

no *Journal of Biogeography* o qual publicou, o icônico editorial “*What is modern Biogeography without Phylogeography?*” por Brette Riddle (2009), e consecutivos números especiais sobre Filogeografia. Nos últimos dez anos, a Filogeografia, bem como as várias outras áreas da ciência, se modernizou em métodos e ferramentas analíticas, ainda em fase de adaptação e inclusão de dados genômicos massivamente gerados pelas novas tecnologias de sequenciamento. As próximas páginas darão uma visão geral dos principais métodos e conceitos que impactaram as análises filogeográficas nos últimos 40 anos, visando introduzir o tema para a compreensão dos dados e análises empíricas da segunda parte desta tese.

3.1.2.1 Filogeografia: do DNA citoplasmático à era Genômica

A Filogeografia fundamentou seu início basicamente sobre a análise de dados de DNA de organelas, principalmente o DNA mitocondrial (DNAMt) em animais (AVISE et al., 1987). Dentre vários fatores o DNAMt ganhou força em análises filogeográficas devido a fatores intrínsecos dos genomas citoplasmáticos (ex: DNAMt e de cloroplastos - DNACp), como falta de recombinação gênica, relativa neutralidade e menor tamanho efetivo populacional (AVISE, 2009; HICKERSON et al., 2010), logo, menor tempo até a coalescência de alelos ou monofilia de áreas geográficas (EDWARDS; BEERLI, 2000; HICKERSON et al., 2010). Além disso, outras características como elevadas taxas de mutação, herança matrilinear e grande quantidade de cópias por célula, fizeram deste genoma a principal ferramenta no início da Filogeografia (AVISE, 2000, 2009; BALLARD; WHITLOCK, 2004).

Para compreender estes diferentes aspectos é interessante ter em mente um panorama geral dos genomas citoplasmáticos, utilizando como exemplo o genoma mitocondrial. Genomas de organelas em eucariotos, como o DNAMt, tem provável origem protobacteriana, endossimbioticamente assimilado por uma célula pré-eucariótica (LANG; GRAY; BURGER, 1999). Ao longo da evolução dos eucariotos, genomas citoplasmáticos tenderam a perder a maior parte do seu DNA para o núcleo. Por exemplo, o DNAMt de células animais é hoje representado por apenas 37 genes ligados ao longo de uma molécula circular com cerca de 16.000 pb de comprimento (FRIEDMAN; NUNNARI, 2014). Dois desses genes codificam RNAs ribossômicos, 22 especificam diferentes moléculas de RNA transportador e 12 ou 13 codificam subunidades de polipeptídios participantes das vias bioquímicas de fosforilação oxidativa e da cadeia transportadora de elétrons (Figura 3.2).

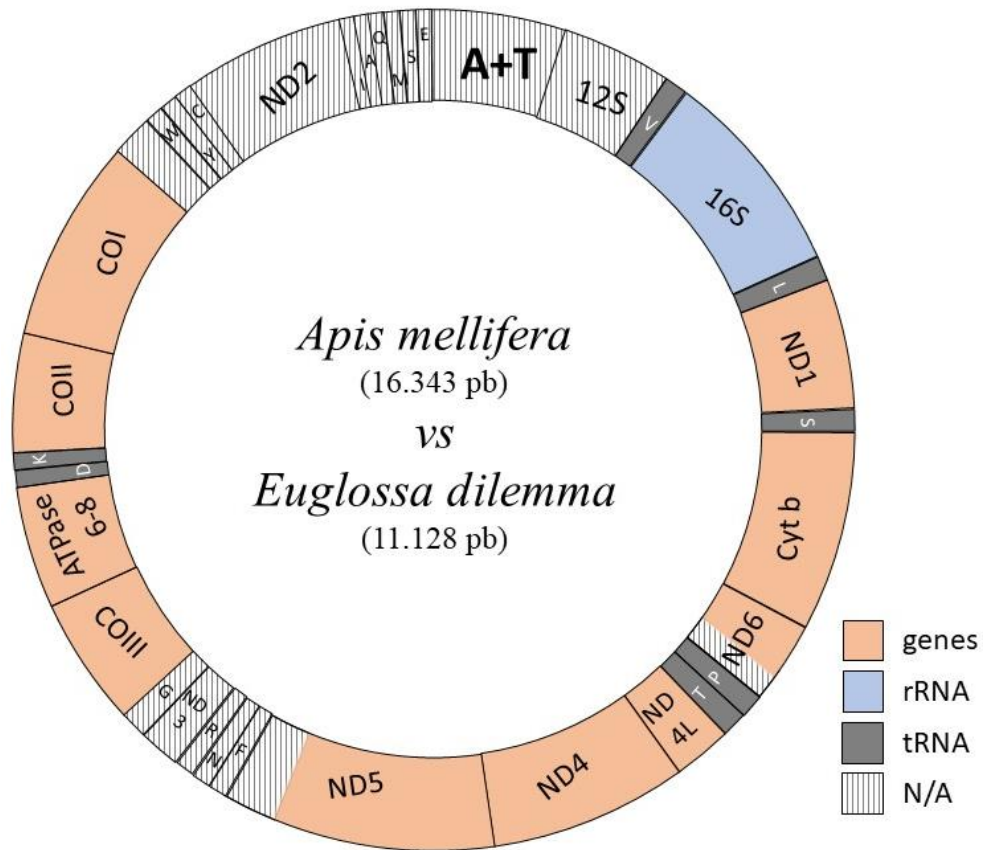


Figura 3.2 Comparação da representação gráfica do genoma mitocondrial de duas abelhas corbiculadas, *Apis mellifera* (CROZIER; CROZIER, 1993) e *Euglossa dilemma* (BRAND et al., 2017). Cores representam tipos diferentes de sequências; N/A – regiões sem informação para *E. dilemma*; A+T – origem de replicação do DNA mitocondrial (CROZIER; CROZIER, 1993).

Via de regra, essas moléculas também incluem uma região de controle de replicação (RC), a qual é utilizada como origem para a replicação do DNA. A replicação desses genomas citoplasmáticos é assíncrona à divisão celular e frequente durante a interfase, gerando centenas ou milhares de cópias de DNAm_t ou DNAc_p no citoplasma de cada célula somática ou germinativa do organismo (FRIEDMAN; NUNNARI, 2014).

Embora seja altamente variável em comparação ao genoma nuclear, evolutivamente os genomas citoplasmáticos possuem considerável estabilidade constitutiva. No caso das abelhas, o DNAm_t não apresenta grandes diferenças em relação à descrição geral da mitocôndria animal, com exceção da ordem de alguns genes de RNAt e da região de controle da replicação, que nas abelhas e outros insetos é reconhecida por ser uma região rica de repetições AT (CROZIER; CROZIER, 1993). Comparando os genomas mitocondriais da tribo Apini (*Apis mellifera*) e Euglossini (*Euglossa dilemma*) (Figura 3.2), encontra-se um padrão de distribuição dos genes altamente conservado, embora algumas regiões ainda faltem a essa comparação (áreas hachuradas).

A ausência de dados da região AT e trechos flanqueadores é um dos principais complicadores para estudos populacionais e filogeográficos com base em DNAm_t em Euglossini, visto que este trecho frequentemente é o mais variável do DNAm_t (uma vez que não é transcrito), constituindo uma das principais ferramentas em estudo de populações de vertebrados (AVISE, 2009; HICKERSON et al., 2010). No entanto, estudos do DNAm_t de abelhas frequentemente obtém dados expressivos quando analisam genes mitocondriais como citocromo *b* (LÓPEZ-URIBE et al., 2014; PENHA et al., 2015) e citocromo *c* oxidase (DICK et al., 2004; RAMÍREZ et al., 2010).

Um dos fatores que favorece o uso do DNAm_t para estudos filogeográficos é o acúmulo de mutações em escalas muito maiores do que o genoma nuclear (BALLARD; WHITLOCK, 2004). Por exemplo, em *Drosophila*, trechos mitocondriais podem acumular mutações entre 1,7-3,4 vezes mais rápido do que as regiões autossômicas mais variáveis, e quase 10 vezes mais rápido do que a média do genoma nuclear (MORIYAMA; POWELL, 1996).

Provavelmente, vários fatores contribuem para o rápido ritmo de evolução do DNAm_t, como mecanismos relativamente ineficientes de reparo do DNA; um ambiente altamente oxidativo; o relaxamento de restrições funcionais do genoma, resultante do fato de que DNAm_t não produz proteínas diretamente envolvidas em sua própria replicação, transcrição ou tradução; ou o fato de que o DNAm_t não está fortemente ligado às proteínas histonas que são por si agentes de proteção e podem restringir as taxas de mutação no DNA nuclear (AVISE, 2009; FRIEDMAN; NUNNARI, 2014). Fato resultante desta elevada taxa de mutação é a alta variação intraespecífica do DNAm_t, pré-requisito para grande parte dos estudos filogeográficos.

Outro “facilitador filogeográfico” do DNAm_t é o modo de transmissão da molécula via herança materna. Durante o processo de fertilização a maior parte do citoplasma de um zigoto vem do óvulo, portanto, cada descendente recebe o DNAm_t principalmente, se não exclusivamente, da fêmea e não do macho (BALLARD; WHITLOCK, 2004; FRIEDMAN; NUNNARI, 2014). Este fato é especialmente importante para elucidação de questões biológicas por dois motivos. O primeiro deles é que por passar integral e diretamente de um indivíduo para outro, sem recombinação ou seleção de alelos, todas as diferenças entre moléculas de gerações diferentes poderiam ser atribuídas, por exemplo, a eventos de mutação (AVISE, 2000, 2009). Por este mesmo motivo, buscar a coalescência dos alelos mitocondriais através de suas linhagens é teoricamente muito mais simples do que a partir de cópias nucleares, onde a recombinação gênica não pode ser descartada (AVISE, 2009). O segundo

motivo é que devido à ligação matrilinear, eventos que afetem diferencialmente os genomas nuclear e mitocondrial podem ser indícios de dispersão sexual diferencial. Um exemplo deste caso, são as próprias abelhas das orquídeas, as quais possuem fêmeas conhecidamente filopátricas (AUGUSTO; GARÓFALO, 2004; BOFF; SAITO; SANTOS, 2017) e que exibem nítidos reflexos deste comportamento em seus padrões filogeográficos quando dados genéticos mitocondriais e nucleares são confrontados (LÓPEZ-URIBE et al., 2014; PENHA et al., 2015). Combinados, estes dois fatores podem levar a estruturas geográficas acentuadas, as quais são por definição, reflexos históricos dos eventos que moldaram a diversificação de populações e espécies (AVISE, 2009).

Embora a maior parte dos estudos sobre Filogeografia tenham se baseado no estudo do DNAmT e na genealogia de seus genes, existem ao menos duas fontes de erros inerentes à construção de genealogias a partir deste tipo de dados: a incerteza da genealogia e o viés do genoma. Nenhuma genealogia é estimada sem erro, porém uma quantidade reduzida de informação pode levar a estimativa de genealogias não condizentes com padrões reais (BALLARD; WHITLOCK, 2004). Em contrapartida, a quantidade correta de informação tende a levar o erro associado a um mínimo cientificamente aceitável, representando fidedignamente a história evolutiva do genoma e do táxon associado (AVISE, 2009).

Outra fonte de erro inerente à estimativa de genealogias gênicas a partir do genoma mitocondrial seria o viés do genoma. O fato de os genes mitocondriais serem ligados leva qualquer evento de seleção natural ou deriva genética em genes individuais a atuar de forma conjunta no cromossomo, em um efeito carona em toda a molécula (HARTL; CLARK, 2007). Esta segregação dependente entre os genes do cromossomo mitocondrial pode levar a duas situações distintas. A primeira delas seria um possível viés filogeográfico, já que idiosincrasias evolutivas e reprodutivas poderiam afetar pontualmente genomas de organelas em relação ao genoma nuclear (BALLARD; WHITLOCK, 2004). Mais uma vez, cabe aqui o exemplo dos euglossíneos onde machos são o sexo dispersor, o que por sua vez tende a causar estruturas acentuadas no genoma mitocondrial em comparação ao genoma nuclear (LÓPEZ-URIBE et al., 2014). O mesmo pode ser observado em tartarugas e outros organismos onde fêmeas e machos exibem diferentes descolamentos reprodutivos (AVISE, 2009).

Interessantemente, a segregação dependente entre genes mitocondriais leva a uma segunda situação, distinta da anterior, onde o viés entre genomas é útil exatamente para evidenciar o efeito genético de aspectos biológicos intrínsecos dos organismos (HICKERSON et al., 2010). Novamente, traçando um paralelo com abelhas das orquídeas, onde as fêmeas

raramente são amostradas, a análise do DNA mitocondrial é, por vezes, a única forma de mensurar a dinâmica demográfica das fêmeas ao longo das gerações (LÓPEZ-URIBE et al., 2014; PENHA et al., 2015).

Obviamente há outras críticas à utilização exclusiva de DNA de organelas para detectar padrões geográficos e demográficos em populações naturais, dentre as quais, as questões de árvores de genes e de espécies, introgressão e neutralidade do DNAm são provavelmente as principais delas (BALLARD; WHITLOCK, 2004). Vale ressaltar aqui, que devido aos fatores já discutidos, o DNA de organelas tende, como um todo, a responder de forma semelhante aos eventos evolutivos e estocásticos em populações naturais. Deste modo, com algumas exceções, genealogias gênicas independentes também tendem a demonstrar histórias evolutivas semelhantes quando comparadas entre si (AVISE, 2000; AVISE et al., 1987; HICKERSON et al., 2010). Estes argumentos refutariam em partes a discrepância entre árvores de genes e espécies ao menos em nível intraespecífico (logo, árvores de genes e de linhagens), quando considerado apenas o genoma de organelas.

Ainda assim, o estudo exclusivo de um único genoma (tanto de organela, quanto nuclear) pode levar a conclusões incompletas, caso não se tenha em mente a questão do viés dos genomas. Devido a essas e outras críticas, nos últimos anos estudos filogeográficos tenderam a incluir dados de genes nucleares. Em princípio, os genes nucleares de cópia única também podem fornecer uma grande quantidade de dados para análises filogeográficas no nível intraespecífico, mas três obstáculos técnicos e biológicos impediram o seu progresso em larga escala: o ritmo relativamente lento da evolução da sequência em muitos locos nucleares; a dificuldade de isolar haplótipos nucleares de organismos diploides; e o fenômeno da recombinação gênica (AVISE, 2009).

A primeira complicação relativa às baixas taxas de mutação nuclear em comparação ao DNA de organelas podem ser facilmente contornadas buscando sequências como íntrons e microssatélites que acumulam mutações muito mais rapidamente do que a média do DNA nuclear (BALLARD; WHITLOCK, 2004). Um exemplo em Euglossini é o trabalho de López-Uribe et al. (2014), os quais empregaram o sequenciamento de locos microssatélites para avaliar a estrutura filogeográfica em espécies do gênero *Eulaema* em todo continente Sul-americano. No entanto, uma dificuldade frequente neste tipo de abordagem é como tratar indels nos modelos filogeográficos, bem como estimar seguramente a homologia dos alinhamentos em regiões altamente repetitivas.

Já no que tange o isolamento dos haplótipos em organismos diploides, duas opções são frequentemente utilizadas. A primeira delas é buscar marcadores em cromossomos

sexuais (como os cromossomos X e Y em mamíferos), os quais são naturalmente sistemas haploides. Interessantemente, este é um problema que não afetaria euglossíneos devido ao sistema haplodiploide de determinação sexual, onde machos são provenientes de óvulos não fecundados e são naturalmente haploides para todo genoma, salvo algumas exceções (GIANGARELLI et al., 2015; SOUZA et al., 2010). Nos casos onde a natureza não fornece sistemas naturalmente haploides é ainda possível isolar haplótipos nucleares a partir de métodos de clonagem de trechos de interesse ou ainda utilizar técnicas de estimativa de fase (cis/trans) para sequências com sinais de heterozigose. Neste último caso, vários estudos com aves do continente sul-americano demonstram resultados consistentes quando analisados via vários locos nucleares independentes (BATALHA-FILHO; MIYAKI, 2016; CABANNE et al., 2014; RAPOSO DO AMARAL et al., 2013).

O terceiro impedimento ao uso amplo de sequências nucleares merece mais atenção, pois de fato é o mais grave. Em qualquer população com reprodução sexuada, processos inerentes à meiose e à própria mistura de material entre indivíduos continuamente reúnem genes não ligados em novas combinações multilocus. Esses processos de recombinação também podem afetar locos individualmente sempre que ocorrem eventos de *crossing over* meióticos bem-sucedidos entre diferentes alelos. Cada alelo recombinante é então um trecho fusionado de DNA, cujos subconjuntos podem ter passados evolutivos bastante diferentes. Assim, a recombinação gênica (ou alélica) tende a complicar o que, de outro modo, seria a assinatura genealógica mais limpa do genoma nuclear (AVISE, 2009). Mesmo com todos estes complicadores, haplótipos de DNA nuclear foram empregados com sucesso em várias avaliações filogeográficas com alguns dos resultados mais informativos provenientes de sequências de íntrons em genes codificadores de proteínas (BATALHA-FILHO; MIYAKI, 2016; BRITO; EDWARDS, 2009; RAPOSO DO AMARAL et al., 2013), mas com interessante potencial para locos microssatélites (LÓPEZ-URIBE et al., 2014).

Além disso, mesmo quando os obstáculos técnicos e biológicos para obtenção de dados genealógicos explícitos de sequências nucleares se tornam insuperáveis, informações menos diretas sobre histórias populacionais ainda podem ser obtidas a partir de análises genéticas populacionais com base nas frequências observadas de alelos filogeneticamente não ordenados (AVISE, 2009). Análises multilocus baseadas em microssatélites, por exemplo, são frequentemente utilizadas para estimar parâmetros populacionais que, por sua vez, podem ser usados em análises genéticas espaciais, com vários exemplos apenas em euglossíneos (BOFF et al., 2014; LÓPEZ-URIBE et al., 2014; PENHA et al., 2015; ROCHA FILHO et al., 2013). Especialmente quando tais dados são acumulados em vários locos desvinculados, a

informação histórica pode ser facilmente comparada com os padrões genealógicos mais explícitos que normalmente são evidentes no genoma citoplasmático (LÓPEZ-URIBE et al., 2014; PENHA et al., 2015).

Provavelmente os microssatélites foram um dos marcadores mais utilizados nos últimos vinte anos de estudos filogeográficos e também um dos principais fatores de aproximação entre a Filogeografia e áreas correlatas como a Genética de Populações, Genética de Paisagens e até mesmo a Ecologia Comportamental (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). A grande capacidade de gerar matrizes altamente variáveis é sem sombra de dúvidas uma das principais vantagens destes marcadores. Microssatélites carregam ainda importantes informações sobre a relação entre os alelos (assumindo o modelo de mutação passo-a-passo – *step-wise mutation model*, SMM) e, a princípio, também possuem a capacidade de distinguir informações em diferentes escalas de tempo (como comparações entre índices de diferenciação como R_{ST} e F_{ST}) (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015).

Por outro lado, o alto conteúdo polimórfico dos microssatélites têm seu preço (MORIN; LUIKART; WAYNE, 2004). Os desafios e deficiências da aplicação destes marcadores, para fins filogeográficos foram extensivamente revisados (PERKTAŞ et al., 2015; ZINK, 2010) e incluem substanciais homoplasias, inconsistências filogenéticas, frequentes alelos nulos e dificuldade em comparação direta com marcadores baseados em sequências. Existem ainda questionamentos quanto à total neutralidade destes marcadores, frequentemente assumida como pressuposto para algumas análises, as quais tem sido questionadas por alguns autores sob sua utilidade para a Filogeografia (BRITO; EDWARDS, 2009; EDWARDS, 2009; ZINK; BARROWCLOUGH, 2008). Com certeza, microssatélites e outras sequências repetitivas tem seu espaço na Filogeografia, contudo, e a necessidade de recuperar linhagens genealógicas tem forçado uma mudança de direção neste campo.

Ao longo dos anos 2000 e início da década 2010, a Filogeografia tendeu a incluir cada vez mais locos para inferir a história demográfica dos grupos, de modo que o uso de vários locos para inferir a história da população e das espécies tornou-se a base na Filogeografia moderna (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015; MCCORMACK et al., 2013). Com a possibilidade e necessidade de se incluir cada vez mais locos nas análises filogeográficas, métodos moleculares e computacionais foram desenvolvidos para acomodar a quantidade crescente de dados. Para tanto, métodos baseados nas Teorias Coalescente, Bayesiana e em Máxima Verossimilhança (veja 3.3.2 Análises em Genética de Populações e

Filogeografia), desenvolvidos em sua maioria para análises filogenéticas, foram cada vez mais incorporados às inferências filogeográficas.

Mesmo com a crescente disponibilidade de marcadores moleculares para organismos não-modelos (EDWARDS, 2008; THOMSON; WANG; JOHNSON, 2010), o processo de geração de dados para um estudo multilocus é laborioso. No campo da biologia molecular, sempre em rápida evolução, por vezes parece pouco justificável embarcar no longo processo de seleção de locos para estudos de variabilidade (assumindo que os *primers* já existem), amplificando e sequenciando amostras de DNA individualmente para cada loco, submetendo estes dados à determinação de fase nuclear via bioinformática ou clonagem. Devido a isso, a Filogeografia e Filogenética vêm se aproximando cada vez mais das ferramentas baseadas em *next generation sequencing* (NGS) (Figura 3.3) devido ao seu potencial de condensar várias etapas da geração de dados multilocus para organismos não-modelos em um processo mais eficiente em termos de tempo e economia (MCCORMACK et al., 2013). Note que apenas nos últimos anos ferramentas baseadas nestas novas tecnologias começaram a ser desenvolvidas e muitas delas ainda se encontram em fase de desenvolvimento.

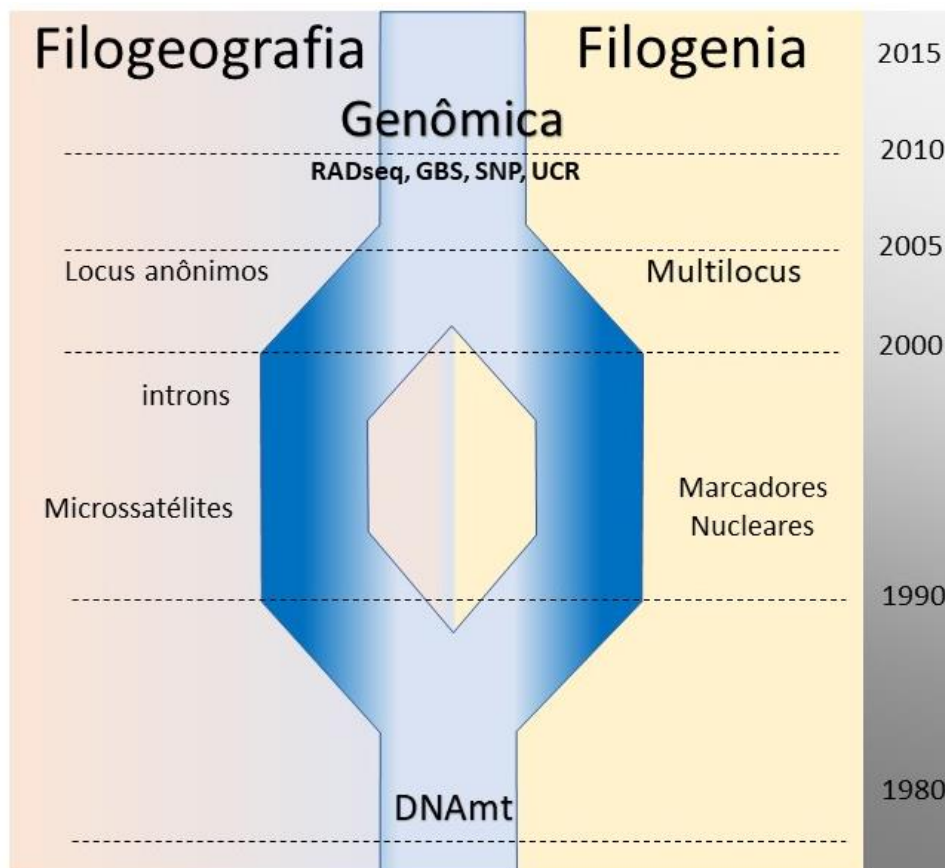


Figura 3.3 Esquema representativo da evolução temporal do uso de marcadores moleculares na Filogeografia e Filogenia. Fonte: (BRITO; EDWARDS, 2009). A figura retrata a unicidade original das duas áreas devido um foco comum no DNAmt na década de 1980. Durante a década de 1990, no entanto, embora o DNAmt ainda fosse usado tanto na Filogeografia quanto na filogenética, esses campos tomaram rumos distintos, muito devido

aos diferentes marcadores nucleares utilizados em cada área. Após o acesso ao sequenciamento de nova geração e genomas nucleares de organismos não-modelos ao longo dos anos 2000, a Filogeografia e a Filogenética tomaram rumos convergentes com dados de SNP, embora a natureza evolutiva e as taxas mutacionais das regiões alvo ainda sejam ponto de distinção entre ambas.

Desde o lançamento das primeiras plataformas de sequenciamento massivo em 2005 (Roche/454, SOLiD/Life Technologies e Illumina), iniciou-se uma revolução da era genômica em toda biologia molecular. Essas novas tecnologias de sequenciamento possibilitaram a aquisição de dados genômicos em larga escala, permitindo então a análise de centenas de milhares de locos espalhados pelo genoma (BRITO; EDWARDS, 2009). O impacto e potencialidades desses novos métodos é tão significativo que serão abordados separadamente no tópico a seguir.

3.1.2.2 *O impacto do NGS sobre a Filogeografia*

Tradicionalmente, o desenvolvimento de marcadores moleculares como microssatélites, RFLPs e AFLPs, e mesmo sequenciamento de alguns genes, é um processo recursivo e dispendioso. Sucessivas clonagens, desenvolvimento de *primers* específicos e outras etapas iniciais tornam esses processos caros e laboriosos para aplicar em diferentes espécies e populações (ARIF; KHAN, 2009). O advento do SNP-array (veja revisão em Song e Shao, 2015) possibilitou genotipagens em grande escala, mas não o processo de descoberta de SNPs, visto que a produção de uma matriz de alta qualidade requer um investimento substancial de recursos (DAVEY et al., 2011). Além disso, o caráter específico de marcadores dessa natureza é uma desvantagem importante, uma vez que a genotipagem de novas populações estará restrita aos locos presentes na pesquisa original.

Interessantemente, há algum tempo estudos com genomas-modelo têm evidenciado que, ainda que menos variáveis, os SNPs são muito mais comuns no genoma do que, por exemplo, os microssatélites (BRUMFIELD et al., 2003; EDWARDS, 2008; EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). Embora potencialmente mais informativos, o estudo de SNPs amplamente distribuídos pelo genoma foi quase irrelevante no início dos anos 2000 devido à dificuldade (se não impossibilidade) de aproveitar esses SNPs em escala proporcional à sua distribuição pelo genoma (EATON, 2014).

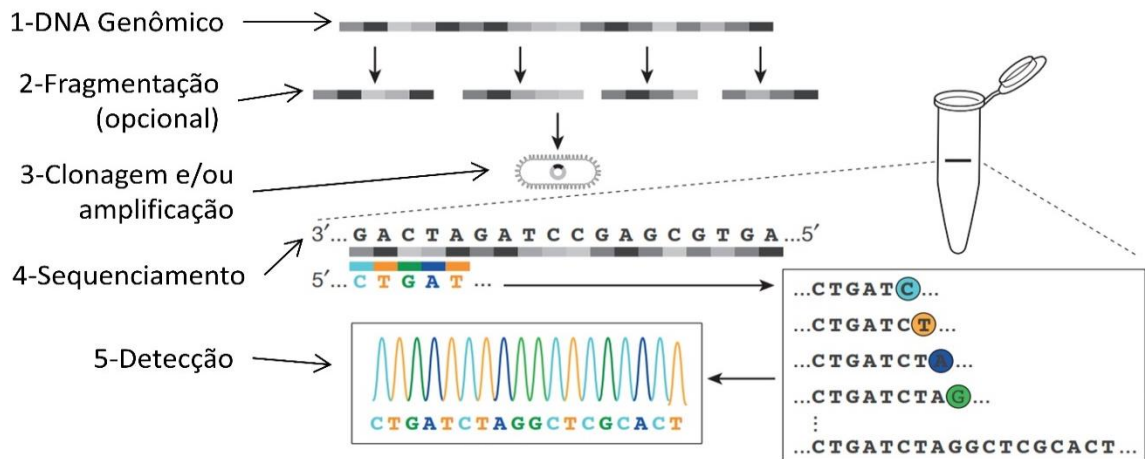
Ao longo dos anos seguintes, o desenvolvimento de plataformas capazes de sequenciar bilhões de pares de bases em apenas uma única leitura, tecnologia para o processamento destes dados e técnicas para explorar genomas de organismos não modelos foram e continuarão sendo a chave para a expansão da Filogeografia sobre as possibilidades do uso de SNPs. Devido à natureza distinta dessas novas tecnologias, bem como seu impacto e

prospecto na área da Filogeografia, cabe aqui uma rápida caracterização e contextualização dos métodos de sequenciamento, como funcionam e o que respectivamente implicam para a compreensão dos dados, traçado ainda um paralelo entre suas vantagens e desvantagens para aplicação em estudos filogeográficos.

De modo geral, existem basicamente três formas ou gerações de sequenciamento de acordo com a natureza da detecção das sequências: o método de Sanger (primeira geração), o NGS (segunda ou “próxima” geração) e o *real-time* ou *single-molecule sequencing*, considerada a terceira geração de sequenciamento (METZKER, 2010; SHENDURE et al., 2017). Embora todas as tecnologias sejam importantes por inúmeras razões distintas, será dado ênfase apenas às duas primeiras gerações de tecnologias de sequenciamento para fins de comparação dos resultados na presente tese.

A primeira geração de sequenciamento, conhecida como “método de Sanger” ou “sequenciamento de Sanger” (SANGER; NICKLEN; COULSON, 1977) utiliza, em sua forma mais moderna didesoxirribonucleotídeos (sem a hidroxila na extremidade 3’ da pentose do nucleotídeo), marcados com fluorescência específica para cada base nitrogenada em proporção de 1/10 nucleotídeos “normais” como substrato para uma reação de PCR com milhões de cópias do DNA de interesse como molde, gerando assim, milhões de cadeias aleatoriamente interrompidas que uma vez ordenadas por eletroforese de alta resolução e luminosamente estimuladas, refletem cores específicas para cada nucleotídeo em ordem crescente do menor para o maior fragmento, os quais são transcritos em sequências de DNA para cada amostra submetida ao processo (Figura 3.4-A).

A – Sequenciamento de Primeira Geração (Sanger)



B – Sequenciamento de Segunda Geração (NGS)

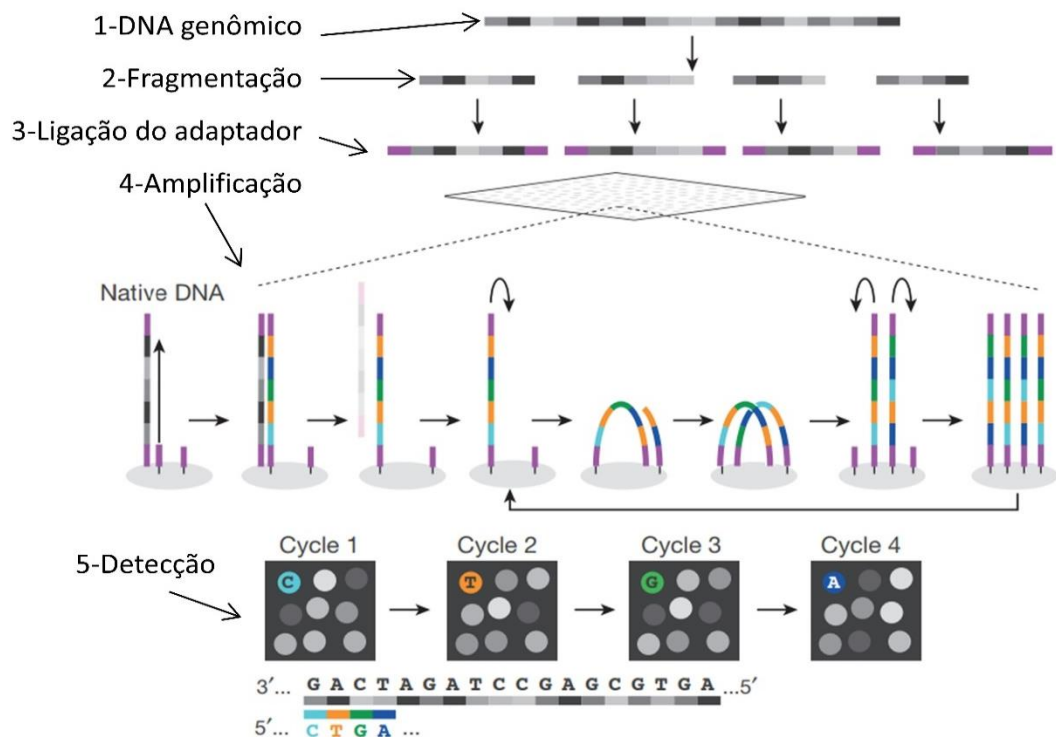


Figura 3.4 Tecnologias de sequenciamento. A) Exemplificação do esquema geral de sequenciamento automático baseado no método de Sanger; B) Exemplificação de um esquema de sequenciamento paralelo massivo (tipo *Illumina*®), também frequentemente citado como NGS (*Next Generation Sequencing*). Adaptado de Schendure et al. (2017).

Em outras palavras o método de Sanger é definido por dois pontos principais: o primeiro é a molécula didesoxirribonucleotídeo, a qual interrompe a reação em cadeia da polimerase devido à falta da extremidade 3'OH livre para adição do próximo nucleotídeo; o segundo é a ordenação dessas múltiplas cadeias interrompidas via eletroforese em capilares. Este método exige que amostras sejam analisadas individualmente, de modo que reações

independentes são necessárias para cada trecho de interesse de cada uma das amostras. Ainda que bastante laborioso e com certa limitação para aquisição de informação em larga escala, o método de Sanger foi fundamental para o desenvolvimento da Filogeografia até o estado atual, sendo a tecnologia dominante nos estudos filogeográficos por quase três décadas (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015).

Por outro lado, as tecnologias de sequenciamento de segunda geração (*Next Generation Sequencing*, *NGS*) se diferenciam do sequenciamento de primeira geração por várias características, mas o ponto chave é sem dúvida a demultiplexação. Em outras palavras, ao invés de um tubo por reação, uma biblioteca complexa com múltiplos moldes de DNA fragmentado (~300pb; Figura 3.4-B2) com adaptadores específicos já ligados (Figura 3.4-B3) é hibridizada à pequenos “*primers*” anexados à uma superfície bidimensional sólida, os quais são complementares aos adaptadores conectados nas sequências fragmentadas da biblioteca. A partir deste passo, múltiplas amplificações *in vitro* acontecem na base sólida, gerando milhões de pequenos *clusters* de DNA “clonais” adjacentes (Figura 3.4-B4). Finalmente, ao invés de medir o comprimento dos fragmentos, a detecção de nucleotídeos é baseada em ciclos de leitura de sinais bioquímicos (por exemplo, incorporação de nucleotídeos marcados com fluorescência; Figura 3.4-B5) através de imagem (Solexa® ou NovaSeq® / Illumina) ou sensores iônicos (Ion Torrent / ThermoFisher Scientific).

Há ainda diferentes estratégias pelas quais as plataformas determinam a sequência de nucleotídeos para a amostra submetida. Contudo, o padrão das plataformas *Illumina* se tornaram os mais populares e serão utilizados como exemplo neste texto (Figura 3.4-B). Note que os passos entre os métodos são diferentes na preparação das amostras. Ao contrário do método de Sanger, clonagens ou PCRs prévias não são necessárias para o isolamento de alvos (Figura 3.4-A3), pelo contrário, há apenas uma fragmentação obrigatória das amostras iniciais (Figura 3.4-B2).

Note que a fragmentação é um passo necessário à tecnologia NGS e implica algumas limitações que necessitam ser contornadas bioinformaticamente. A estratégia adotada por essas plataformas de sequenciamento massivo (como a “clusterização” e amplificação em ponte apoiada em uma base sólida) restringe o comprimento dos *reads* gerados. Trechos curtos de DNA (50, 100, 150 pb) são dificilmente alinháveis de forma ordenada para montar um genoma sem qualquer sequência de referência. Outro problema é que, devido à geração simultânea de milhões de sequências paralelas (*reads* ou *tags*), análises computacionais são necessárias para demultiplexação dos dados e agrupamento em diferentes amostras. Este último problema é facilmente contornado, adicionando uma pequena sequência (3 e 10 pb)

para identificar uma amostra ou trecho alvo, adjacente ao adaptador compatível com a máquina. Esta sequência identificadora é chamada de *barcode* e é crítica para a organização futura dos dados.

De fato, o menor tamanho de fragmentos é a maior deficiência da tecnologia NGS, a qual atualmente prometem entre 300 e 400 pb em algumas plataformas (*Illumina MySeq4000*; *Ion Torrent ThermoFisher Scientific*). Ainda assim, as tecnologias mais comuns geram *reads* entre 50 e 250 pb, os quais são bastante reduzidos quando comparados àqueles do método de Sanger que atingem facilmente 1000 pb. O reduzido tamanho dos *reads* no NGS é o foco de desenvolvimento das tecnologias de terceira geração, as quais se concentram em combinar longos fragmentos (~10 vezes maiores que os de Sanger) com a quantidade massiva de dados de NGS. O sistema PacBio (já comercialmente disponível) promete entregar leituras de até 15.000pb para DNAs íntegros, enquanto adaptações da tecnologia *Illumina* também prometem trechos com até 10.000 pb. Revisões detalhadas sobre essas novas tecnologias podem ser facilmente encontradas na literatura (METZKER, 2010; SHENDURE et al., 2017; SHENDURE; JI, 2008).

É evidente que a combinação destas tecnologias abre grandes possibilidades para estudo de organismos não-modelos, como por exemplo, a possibilidade de criar grandes moldes (*scaffolds*) de milhares de pares de base (ex: PacBio; *Illumina long-seq*) e combiná-las com sequenciamento massivo (ex: *Illumina HiSeq*) com apenas alguns milhares de dólares, algo completamente impensável no começo deste século. Por hora, as abordagens dentro da tecnologia NGS são as mais acessíveis, mas em um futuro próximo plataformas PacBio e *Illumina long-seq* guiarão uma nova mudança de rumos não apenas na Filogeografia, mas em várias áreas envolvendo a biologia molecular.

Ainda que o pequeno tamanho dos *reads* pareça ser um grande problema, diversas abordagens são possíveis via sequenciamento massivo NGS, as quais se diferenciam pela forma com a qual as bibliotecas são preparadas, isolando ou não regiões de interesse (METZKER, 2010). No que diz respeito à Filogeografia, as tecnologias de NGS têm sido exploradas basicamente por duas abordagens: as de subamostragem via enzimas de restrição e as focadas na captura de regiões específicas. As abordagens de subamostragem via enzima de restrição são representadas pelo método RADseq (embora haja algumas variações do método), enquanto os elementos ultraconservados (UCEs) e suas regiões flangeadoras são os mais comuns para as abordagens baseadas em filtragem de elementos específicos para estudos Filogeográficos (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). Existem ainda trabalhos utilizando sequenciamento baseado em transcriptoma e conjuntos de amplicons

específicos (MCCORMACK et al., 2013), porém transcriptomas são demasiadamente conservados para abordagens puramente filogeográficas. Ainda que tenham aplicações importantes para estudos avaliando seleção, o sequenciamento de amplicons é bem menos atraente devido ao trabalho envolvido e o número menor de locos que podem ser testados (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015; MCCORMACK et al., 2013).

As duas abordagens “básicas”, RADseq e UCEs prometem reorientar a Filogeografia de diferentes maneiras, devido aos tipos de dados que cada uma produz (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015; MCCORMACK et al., 2013). Os UCEs produzem dados que podem ser mais facilmente reunidos em sequências individuais com centenas, ou potencialmente, milhares de pares de bases, resultando em haplótipos ou sequências consenso dentro das quais podem haver vários SNPs, que por sua vez, em princípio, podem ser submetidos a análises filogenéticas tradicionais (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). Por outro lado, dados de RADseq geralmente produzem locos muito curtos para montagem de haplótipos e análise via métodos filogenéticos tradicionais; ao invés disso, SNPs únicos ou múltiplos são extraídos de cada loco de RADseq e depois são analisados como SNPs individuais (MCCORMACK et al., 2013).

Um outro paralelo entre essas duas abordagens diz respeito às suas aplicações práticas. Devido às restrições inerentes à preparação das respectivas bibliotecas, RADseq e UCEs têm mais e menos facilidade de aplicação para organismos não-modelos. RADseq, por exemplo, é baseado na fragmentação do DNA total via enzimas de restrição, o que é facilmente executado independentemente do organismo estudado (Figura 3.5-A). Em contrapartida, UCEs geralmente precisam de conjunto de sondas específicas para um determinado grupo de estudo, dificultando sua aplicação em organismos não-modelos, ainda que o conjunto de sondas seja menos específico do que *primers* de microssatélites, por exemplo (Figura 3.5-B).

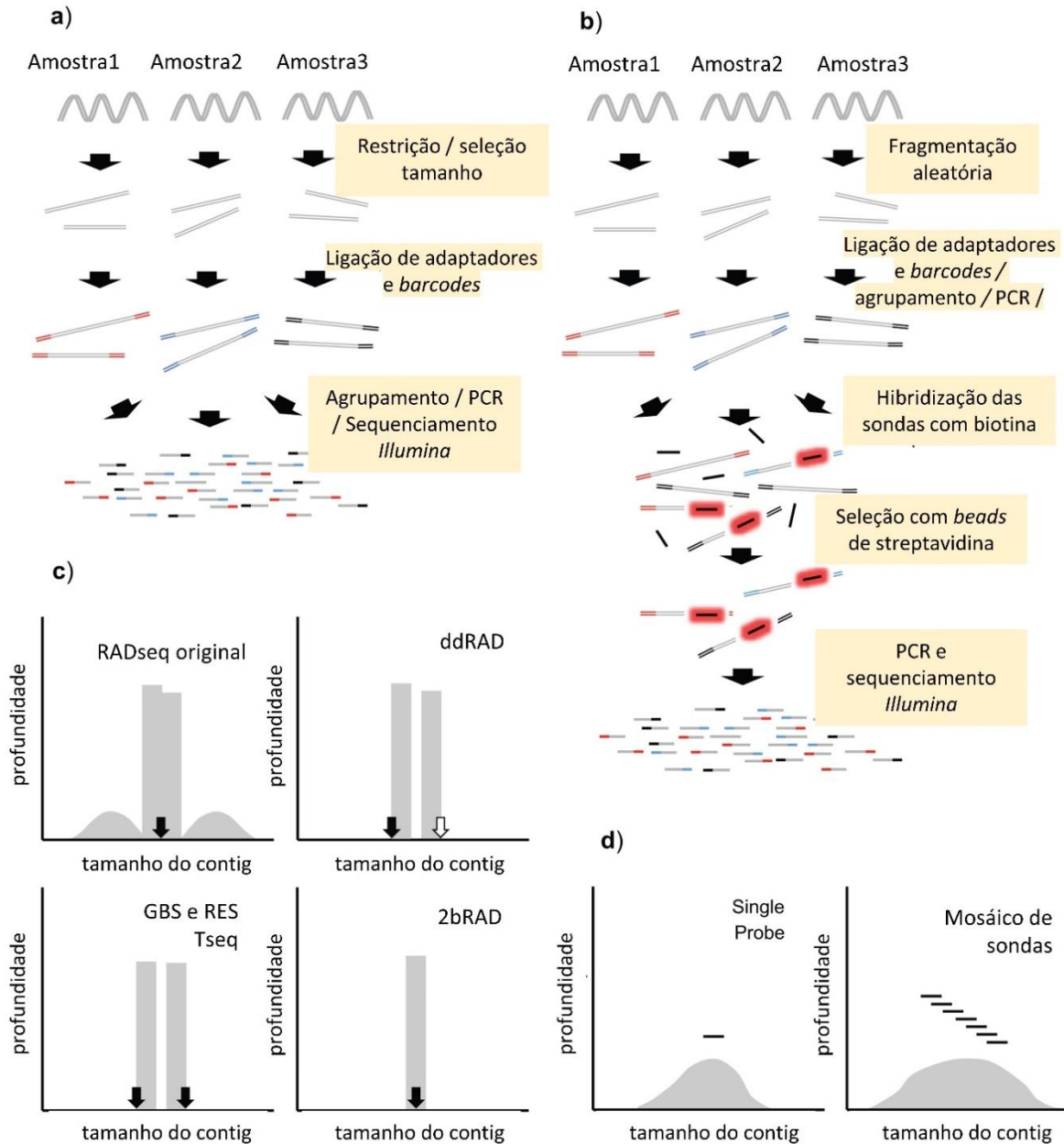


Figura 3.5 Comparação entre estratégias genéricas de subamostragem e de filtragem de elementos genômicos. (A) esquema genérico de RADSeq / GBS; (B) esquema genérico de captura de sequências; (C) distribuições de leitura típicas de bibliotecas genômicas de sequências de vários métodos RADseq; (D) Captura de Sequência. Em (C), os sítios de restrição são representados usando setas com coloração indicando diferentes enzimas de restrição. Adaptado de Harvey et al. (2016).

De muitas maneiras, as duas abordagens fornecem pontes contrastantes à Filogenética e à Filogeografia clássica, além de direcionar abordagens analíticas complementares no futuro. Por exemplo, locos produzidos por abordagens como a UCEs geralmente podem ser analisados usando métodos filogenéticos padrão para estimar árvores de genes, fornecendo uma ponte natural para a Filogeografia clássica (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). Por outro lado, embora os SNPs gerados por RADseq possam ser usados para estimar relações filogenéticas de populações ou espécies (“árvores de espécies”)

(MCCORMACK et al., 2013) e, de fato, já tenham sido submetidos a abordagens de concatenação (EMERSON et al., 2010; MERZ et al., 2013), atualmente esses marcadores são utilizados para contornar a clássica construção de árvores de genes e, em vez disso, estimam diretamente a árvore de espécies (BRYANT et al., 2012; CHIFMAN; KUBATKO, 2014). Essas duas abordagens podem, por vezes, exigir diferentes conjuntos de análises, sendo possível que, ao menos nos primeiros anos, o conjunto de ferramentas para SNPs ligados, como os obtidos a partir dos UCEs, seja ainda mais completo que o disponível para análise de SNPs não ligados (família RADseq).

A despeito deste fato, as abordagens baseadas em RADseq tem se tornado amplamente mais populares do que aquelas utilizando elementos ultraconservados e tem alcançando resultados expressivos do ponto de vista filogeográfico (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). A rápida comparação entre os últimos oito anos (Figura 3.6) mostra uma expressiva diferença entre o número de estudos utilizando RADseq em relação àqueles utilizando UCE, tanto de forma independente, quanto integrado à estudos filogeográficos.

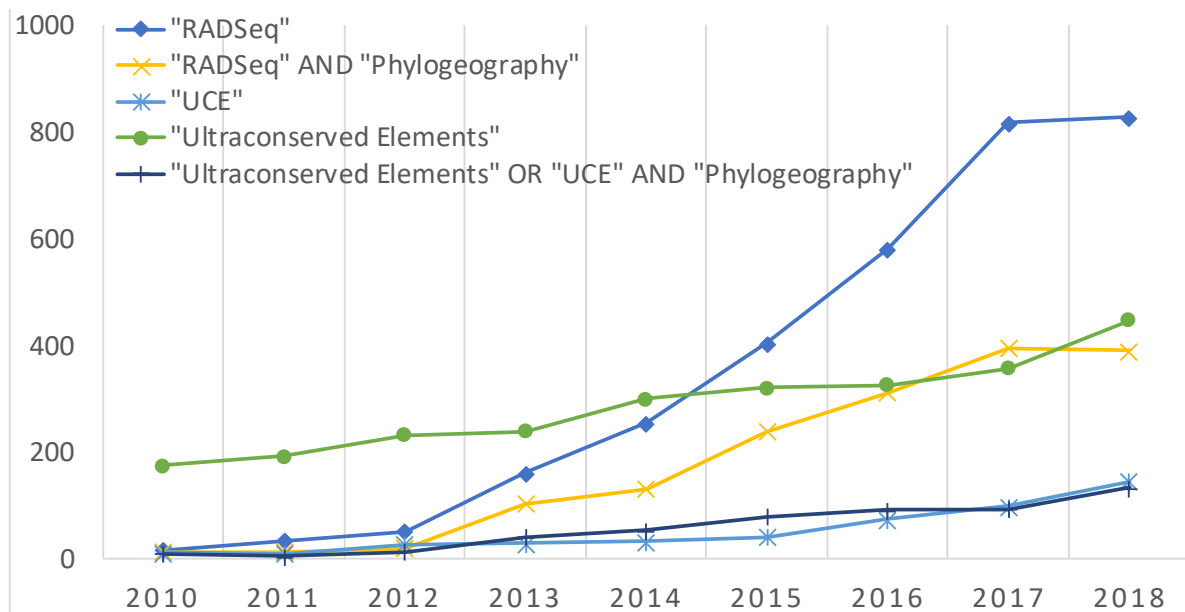


Figura 3.6 Evolução das principais tecnologias NGS e seu uso na Filogeografia. Os dados apresentados são resultado de uma busca simples no indexador *GoogleScholar* compreendendo aos termos descritos na legenda. *Estimativas não incluem abordagens relacionadas e derivações do RADseq, como ddRADseq, GBS, entre outras (veja DAVEY et al., 2011). **Dados para 2018 foram projetados a partir dos dados indexados até 28/02/2018.

Harvey et al. (2016) compararam recentemente o poder de resolução filogeográfico de conjuntos de dados variando em número e comprimento de locos analisados para estimar parâmetros demográficos (tamanho efetivo populacional, tempo de divergência, taxa de migração) de um pássaro neotropical com importantes rupturas filogeográficas. Estes autores demonstraram que o aumento do número de locos até 5000 proporcionou um aumento de

resolução da história demográfica das linhagens, mas que o aumento do número de locos além desse ponto gerou ganhos mínimos. Além disso, eles descobriram que o aumento do número de locos maiores do que 500 pb não produziu ganhos na resolução de parâmetros focais em seu estudo.

Este estudo sugere, portanto, que os números de locos revelados por métodos de subamostragem do genoma, como RADseq, provavelmente serão adequados para a resolução da história de populações em uma variedade de escalas. Estudos empíricos utilizando RADseq obtém em média de 2000-30000 SNPs, apresentando resultados bastante satisfatórios à medida em que revelam linhagens filogeográficas até então não conhecidas para organismos já estudados com outras ferramentas (EDWARDS; SHULTZ; CAMPBELL-STATON, 2015). Deste modo, é provável que os métodos de subamostragem genômica baseadas em enzimas de restrição, como RADseq, finalmente proporcionem o nível adequado de detalhes genômicos para a expansão do futuro da Filogeografia.

3.1.2.3 NGS e Subamostragem Genômica

Até aqui discutiu-se a influência e o potencial do sequenciamento paralelo massivo (ou NGS) para o desenvolvimento de estudos Filogeográficos. Devido ao caráter recente desta tecnologia e o grande número de possibilidades relacionadas, uma quantidade extremamente grande de metodologias foi desenvolvida nos últimos anos (DAVEY et al., 2011; HARVEY et al., 2016). Como visto anteriormente, essas abordagens podem ser classificadas em duas diferentes categorias: as que se concentram em subamostragem aleatória do genoma (*genome-wide methods*; ex. RADseq) e as que filtram regiões estratégicas do genoma (*genomic capture*; ex. UCEs, transcriptomas, entre outras). Uma vez que esta segunda família de métodos não é alvo do presente estudo, suas variações não serão abordadas, mas boas revisões podem ser encontradas em McComark et al. (2013), Edwards et al. (2015) e Harvey et al. (2016).

Já no que tange à primeira família de métodos, a subamostragem genômica, os métodos baseados em enzimas de restrição (daqui em diante *RAD-like*) são uma das principais ferramentas do NGS para estudos populacionais e filogeográficos (REE; HIPPEL, 2015). Os métodos *RAD-like* possuem a característica comum de tentar reduzir a proporção de genoma sequenciado, uma vez que muitas questões biológicas podem ser respondidas com a análise de apenas parte de todo o polimorfismo genômico (DAVEY et al., 2011; MCCORMACK et al., 2013; REE; HIPPEL, 2015). De maneira geral, todos os métodos envolvem os seguintes passos principais: a digestão de múltiplas amostras de DNA genômico (de indivíduos ou populações)

com uma ou mais enzimas de restrição; uma seleção ou redução dos fragmentos de restrição resultantes; e sequenciamento via NGS do conjunto final de fragmentos, que deve ter menos de 1000 pb de tamanho (devido aos limites de comprimento de leitura das plataformas NGS atuais) (DAVEY; BLAXTER, 2010; DAVEY et al., 2011; ELSHIRE et al., 2011; MCCORMACK et al., 2013). Os polimorfismos nos fragmentos sequenciados resultantes podem ser usados como marcadores genéticos.

Por outro lado, as abordagens *RAD-like* diferem umas das outras em várias etapas, embora muitas das diferenças (como a escolha da plataforma de sequenciamento) não são fundamentais para os métodos. Basicamente, esses métodos podem ser agrupados em três classes: “Sequenciamento de Representação Reduzida”, incluindo bibliotecas de representação reduzida (*Reduced Representation Libraries* - RRL) e redução de complexidade de sequências polimórficas (*Complexity Reduction of Polymorphic Sequences* - CRoPS); uma segunda classe generalizada no método “RADseq”; e por fim, uma terceira classe definida como “Genotipagem de Baixa Cobertura”, incluindo Genotipagem de *Shotgun* Multiplexada (*Multiplexed Shotgun Genotyping* - MSG) e Genotipagem por Sequenciamento (*Genotyping by Sequencing* - GBS) (DAVEY et al., 2011).

A grosso modo, a primeira classe destes marcadores tem como principal característica o sequenciamento de várias amostras de DNA em um único *pool*, no qual as variantes populacionais são definidas, mas nenhuma identificação de indivíduo é associada. Esse tipo de técnica foi empregada para a construção de mapas de SNPs em espécies modelos como, humanos, milho, soja, truta arco-íris, dentre outros (DAVEY et al., 2011; LAFRAMBOISE, 2009). A diferença básica entre as técnicas da primeira classe, RRL e CRoPS, é que a primeira é baseada apenas em restrição, funcionando como o sequenciamento de RFLP¹ e o segundo passa por uma amplificação seletiva via PCR, funcionando como o sequenciamento de AFLP¹ (DAVEY et al., 2011). CRoPs também foi o primeiro método a associar sequências *barcode* ou identificadores demultiplexação (MID, plataformas *Roche SOLiD*) para individualizar as diferentes amostras e recuperar suas respectivas sequências bioinformaticamente, o que também poderia ser feito no RRL, mas necessitaria de uma seleção de tamanho individual antes da adaptação dos *barcodes*.

¹ Para uma revisão criteriosa sobre esta classe de marcadores moleculares veja Arif e Khan (2009).

² Devido à ambiguidade da abreviação dos nomes dos gêneros da tribo *Euglossini*, usar-se-á os acrônimos *El.*, *Ef.*, *Eg.*, *Ex.*, para se referir respectivamente à *Eulaema*, *Eufriesia*, *Euglossa* e *Exaerete* onde o uso do padrão internacional não for suficientemente explícito.

³ O termo “*Protoeuglossini*” é aqui utilizado para referir-se à linhagens dos euglossíneos antes da

Embora essa primeira classe seja eficiente para mapear e analisar variações quando há genomas de referência (DAVEY et al., 2011), na ausência destes, faz-se necessário utilizar plataformas que consigam *reads* mais longos possíveis para o uso de montagem *De novo*. Em detrimento ao comprimento dos *reads* perde-se relativa informação de cobertura do genoma e profundidade dos *reads*, uma vez que a quantidade de informação paralela gerada neste tipo de análise é finita. Estas dificuldades aliadas à grande quantidade de trabalho envolvido para a preparação das amostras dificultam o emprego das técnicas do primeiro grupo em organismos sem a existência de genomas de referência.

Visando superar algumas dessas deficiências, Davey e Blaxter (2010) descreveram o Sequenciamento de DNA Associado a sítios de Restrição (*RADseq*; já anteriormente utilizados em *micro-arrays*), adicionando índices *barcodes* para cada amostra antes da seleção de tamanho, o que possibilita a execução dos demais passos após a formação do *pool* de amostras, e não mais individualmente, amostra à amostra. Um outro detalhe interessante é que o método *RADseq* sequencia todos os trechos curtos flanqueando os sítios de restrição. Para tanto as sequências restringidas são aleatoriamente fragmentadas para atingir tamanhos sequenciáveis pela plataforma (~250 pb), em seguida uma série de adaptadores são ligados com o objetivo de garantir que todos os sítios de restrição recebam ao menos um adaptador da plataforma o qual já está associado à uma sequência de identificação (*barcode*) (DAVEY; BLAXTER, 2010). Deste ponto em diante as amostras são reunidas em um único *pool*, o qual é submetido a uma PCR seletiva para enriquecimento da biblioteca apenas com os trechos que tenham recebido os adaptadores.

Esta técnica traz considerável simplificação do ponto de vista prático, porém vários cuidados devem ser tomados para garantir que não haja interações indesejadas entre adaptadores, *primers* e sequências *barcode*. Além disso, ainda que mais simples do que as técnicas de redução de complexidade do genoma tradicionais, o *RADseq* apresenta alguns passos críticos de preparação laboratorial, além de múltiplas ligações que implica em aumento de custos, bem como implicações para o sequenciamento relacionado à cobertura e número de indivíduos passíveis de leitura em um único sequenciamento (DAVEY et al., 2011). Além disso, devido à grande quantidade de informação genômica captada pelo *RADseq*, algumas espécies altamente diversas ou com grandes genomas, ou ainda as que envolvem experimentos com grande número de indivíduos, necessitam de mais de uma leitura em uma célula de fluxo para garantir uma profundidade adequada por loco (DAVEY et al., 2011; ELSHIRE et al., 2011).

Baseados nessas informações, Elshire et al. (2011) descreveram uma simplificação do método de *RADseq*, o *Genotyping by Sequencing* (GBS). No trabalho de descrição da técnica, os autores utilizaram uma enzima de restrição sensível à metilação para evitar zonas altamente repetitivas do DNA, mas que principalmente dispensa algumas etapas do *RADseq*, tornando o processo mais rápido e barato (ELSHIRE et al., 2011). Note que, ao contrário das demais técnicas, o GBS restringe e liga adaptadores nos fragmentos gerados (como no *RADseq*), porém descarta os processos de fragmentação aleatória e a ligação de um segundo adaptador, pulando diretamente para uma PCR seletiva a partir do *pool* de amostras, enriquecendo as regiões ligadas aos adaptadores (Figura 3.7).

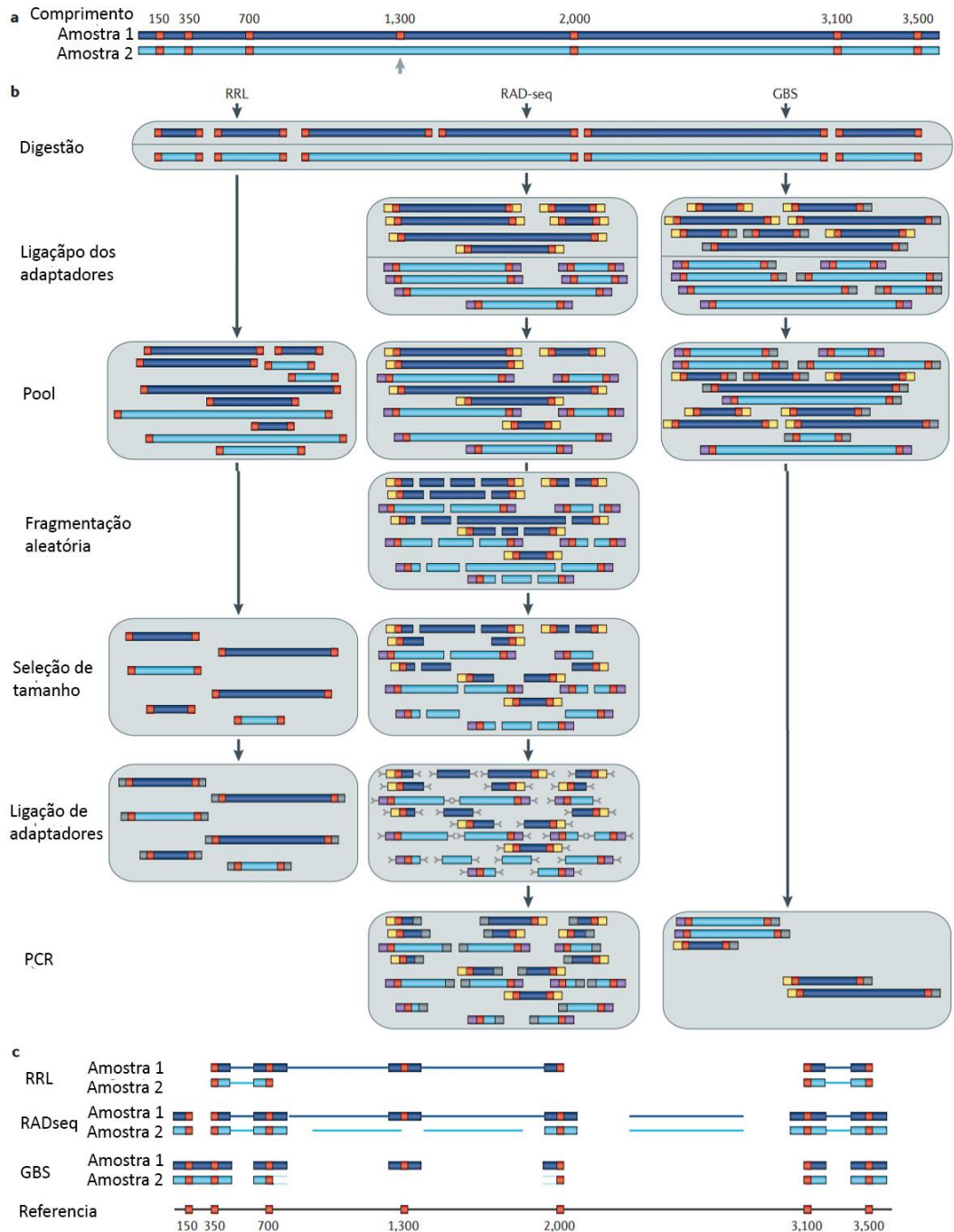


Figura 3.7 Métodos para a descoberta de marcadores de alto rendimento. A) Um exemplo de região genômica contendo sítios de restrição (vermelho). Uma amostra de DNA de cada um dos dois indivíduos (a amostra 1 é azul escuro e a amostra 2 é azul claro) deve ser sequenciada. A amostra 2 tem uma variação no local de corte em 1.300 bases (seta cinza) e, portanto, este sítio não será cortado. B) Protocolos para o desenvolvimento de marcadores sequenciados. C) Os marcadores sequenciados estão alinhados com o genoma de referência original. RRL: qualquer um dos fragmentos (linhas grossas) ou fragmentos inteiros (linhas finas) entre 300 e 700 pb são sequenciados. Como o local em 1.300 bases não está presente na amostra 2, o fragmento longo entre 700 e 2.000 bases é filtrado por seleção de tamanho. *RAD-seq*: as regiões a jusante de todos os fragmentos acima de 300 bases são sequenciadas, mas não o fragmento entre 150 e 350 bases. As linhas finas indicam a sequência que seria coberta usando o sequenciamento bidirecional (*pair-end*). GBS: as linhas tracejadas representam regiões que seriam filtradas durante a amplificação, mas poderiam ser imputadas usando, por exemplo, o algoritmo “*multiplexed shotgun genotyping hidden Markov model*”. Os fragmentos curtos entre 150 e 350 bases será sequenciado.

Ao contrário do *RADseq*, o protocolo de GBS não objetiva sequenciar todos as áreas flaqueadoras dos sítios de restrição, mas apenas uma parte deles sem a necessidade de repetir profundamente um loco. O GBS baseia-se na repetição dos locos entre os indivíduos de várias populações para criar “mapas” genéticos de cada sítio de restrição, de modo que, mesmo sem genomas de referência é possível criar combinações homólogas com segurança. Esta abordagem permite uma facilidade muito grande e redução de custos significativos para acessar informações do genoma de espécies não-modelos, ou sem mapa de referência.

Em resumo (Figura 3.7-B), todos os métodos de subamostragem genômica para o desenvolvimento de marcadores, começam com uma digestão enzimática. No RRL (Figura 3.7, painéis esquerdos), os fragmentos de todas as amostras são agrupados e selecionados para tamanhos entre 300-700 pb. Os fragmentos são ligados a adaptadores padrão de sequenciamento (quadrados cinza) e sequenciados. Neste caso simples, apenas as extremidades dos fragmentos serão sequenciadas, mas o protocolo pode ser modificado para sequenciar fragmentos inteiros. Já no *RADseq* (Figura 3.7-B; painéis médios), os fragmentos são ligados a adaptadores P1 (Figura 3.7, amarelo para a amostra 1, roxo para a amostra 2), são cortados aleatoriamente e selecionados para tamanhos entre 300-700 pb. Os adaptadores P2 com extremidades não complementares (cinza, em forma de Y) são ligados aos fragmentos com e sem adaptadores P1. Então, os fragmentos são amplificados por PCR com *primers* específicos de P1 e P2. O adaptador P2 é completado quando fragmentos contendo adaptadores P1 são amplificados via *primer* P1, e o *primer* P2 apenas se liga aos adaptadores P2 completos (quadrados cinza). Isso significa que apenas fragmentos com adaptadores P1 e P2 (os fragmentos que contêm sítios de restrição) são amplificados. Por fim, no GBS (Figura 3.7-B; painéis à direita), adaptadores de identificação “*barcode*” (amarelo) e adaptadores da plataforma de sequenciamento (cinza) são ligados a fragmentos restringidos, produzindo fragmentos com *barcode* + *adaptador*. As amostras são agrupadas e amplificadas na célula de fluxo de um sequenciador *Illumina Genome Analyzer*, onde somente pequenas amostras com a combinação *barcode* + *adaptador* serão amplificadas para sequenciamento.

Obviamente a utilização destas técnicas depende da capacidade de acomodar e explorar estes dados em estratégias de agrupamento e demultiplexação que recuperem os dados individualizados. Além disso, são necessárias estratégias de bioinformática que se adequem à disponibilidade ou não de genomas de referência. Neste sentido, duas abordagens são possíveis: alinhamento a um genoma de referência ou *assembling* ou montagem *De novo*. Estas abordagens, bem como os principais *frameworks* disponíveis para análise de dados de marcadores amplamente distribuídos pelo genoma via NGS são apresentados a seguir.

3.1.2.4 Estratégias para exploração de dados NGS

De fato, as estratégias para análise dos dados gerados via NGS são inúmeras e as próximas páginas se concentrarão apenas nas principais estratégias disponíveis para a análise de dados de marcadores amplamente distribuídos pelo genoma, a exemplo dos apresentados no tópico anterior. Como dito anteriormente, o desafio dos dados gerados via tecnologia NGS é reorganizar os dados da forma original, tanto associando às respectivas amostras de origem e regiões do genoma. Felizmente, a capacidade computacional para executar métodos para tal tarefa têm sido amplamente explorados. *Pipelines* como TASSEL-GBS, Stacks, UNEAK, IGTS e Fast-GBS são sem dúvida as mais amplamente utilizadas para análise envolvendo GBS e tem incorporado modelos cada vez mais sofisticados para a recuperação integral dos dados gerados por essas tecnologias.

Uma estratégia para análise de dados de GBS deve conter etapas para filtrar leituras de baixa qualidade, classificar leituras por grupo ou indivíduos com base nas sequências *barcode*, identificar locos e alelos via abordagem *De novo* ou alinhado a um genoma de referência, e muitas vezes, definir os genótipos para cada indivíduo incluído no estudo (TORKAMANEH; LAROCHE; BELZILE, 2016). Quando um genoma de referência está disponível, as leituras de GBS podem ser mapeadas com base em um genoma de referência e os SNPs podem ser chamados como nos projetos de *whole-genome resequencing* (MCCORMACK et al., 2013), as quais são possíveis por softwares como TASSEL-GBS v1 e v2 (BRADBURY et al., 2007; LU et al., 2012), Stacks (CATCHEN et al., 2013), IGST e Fast-GBS (TORKAMANEH; LAROCHE; BELZILE, 2016). Na ausência de um genoma de referência, é necessário identificar pares de *reads* quase idênticos (presumidas para representar alelos alternativos de um loco). As *pipelines* mais utilizadas para essa abordagem baseada em *De novo* são UNEAK e Stacks.

O uso de um genoma de referência é preferencial para este tipo de análise, uma vez que fornece importantes informações sobre a ligação física entre os locos, bem como informações sobre sua funcionalidade putativa. No entanto, nem sempre há genomas de referência disponíveis. Tomando como exemplo o caso dos euglossíneos, existe um genoma para *Euglossa dilemma* (BRAND et al., 2017), porém o uso de genomas heterólogos como referência negligência a possibilidade de alterações evolutivas ao longo do genoma, bem como alterações cromossômicas, como por exemplo, inversões, quebras e duplicações cromossômicas. O uso de genomas de referência é ainda dificultado (embora não impossibilitado) quando não há descrição física dos *scaffolds* formados, como no caso de *E.*

dilemma. Deste modo, os dados não dizem muito sobre equilíbrio de ligação a não ser entre SNPs dentro do mesmo *scaffold*, embora sejam muito úteis do ponto de vista funcional e para a descoberta de SNPs. Um último ponto que cabe consideração é sobre a distância filogenética entre os táxons, uma vez que mesmo dentro do próprio gênero, espécies com um longo tempo de divergência podem não ter um alto percentual de sobreposição de genomas.

Nestes casos, há a possibilidade de restringir a análise apenas ao conjunto de dados gerados para a espécie alvo do estudo, utilizando uma estratégia de realinhamento, denominada *De novo*. Esta metodologia busca sequências idênticas ou quase idênticas primeiro entre as sequências de um indivíduo e depois comparando-as entre diferentes indivíduos seguindo critérios específicos de demarcação de locos e variações alélicas (TORKAMANEH; BOYLE; BELZILE, 2018). Devido à natureza recursiva dos dados de NGS e GBS, abordagens dessa natureza tem o potencial de gerar centenas de milhares de locos que podem ser utilizados para comparação de SNPs entre populações (CATCHEN et al., 2013; LU et al., 2012; TORKAMANEH; LAROCHE; BELZILE, 2016). Em ambos os casos é possível obter centenas de locos e SNPs, impossíveis de se obter em outras abordagens.

Uma vez que se tem a informação sobre a disponibilidade de genomas de referência, a escolha de qual *pipeline* utilizar é crítica e bastante difícil, muitas vezes atrelada à experiência do pesquisador ou do laboratório em explorar cada uma das opções. Recentemente, Torkamaneh, Laroche e Belzile (2016) comparando o desempenho entre diferentes *pipelines* para os dois métodos disponíveis, “de referência” e “*De novo*”, concluíram que embora o uso de genoma de referência resulte em um número maior de SNPs (25.000 a 50.000) em comparação à estratégia *De novo* (13.000 a 25.000), a acurácia para ambos os métodos foi maior que 95%, mesmo entre diferentes plataformas. Interessantemente, estes mesmos autores identificaram que as *pipelines* TASSEL v2 e GBSFast são as mais robustas para análises com genomas de referência, e TASSEL v2 e Stacks são as mais robustas para análises *De novo*. Um esquema genérico de organização dessas diferentes estratégias é resumido na Figura 3.8.

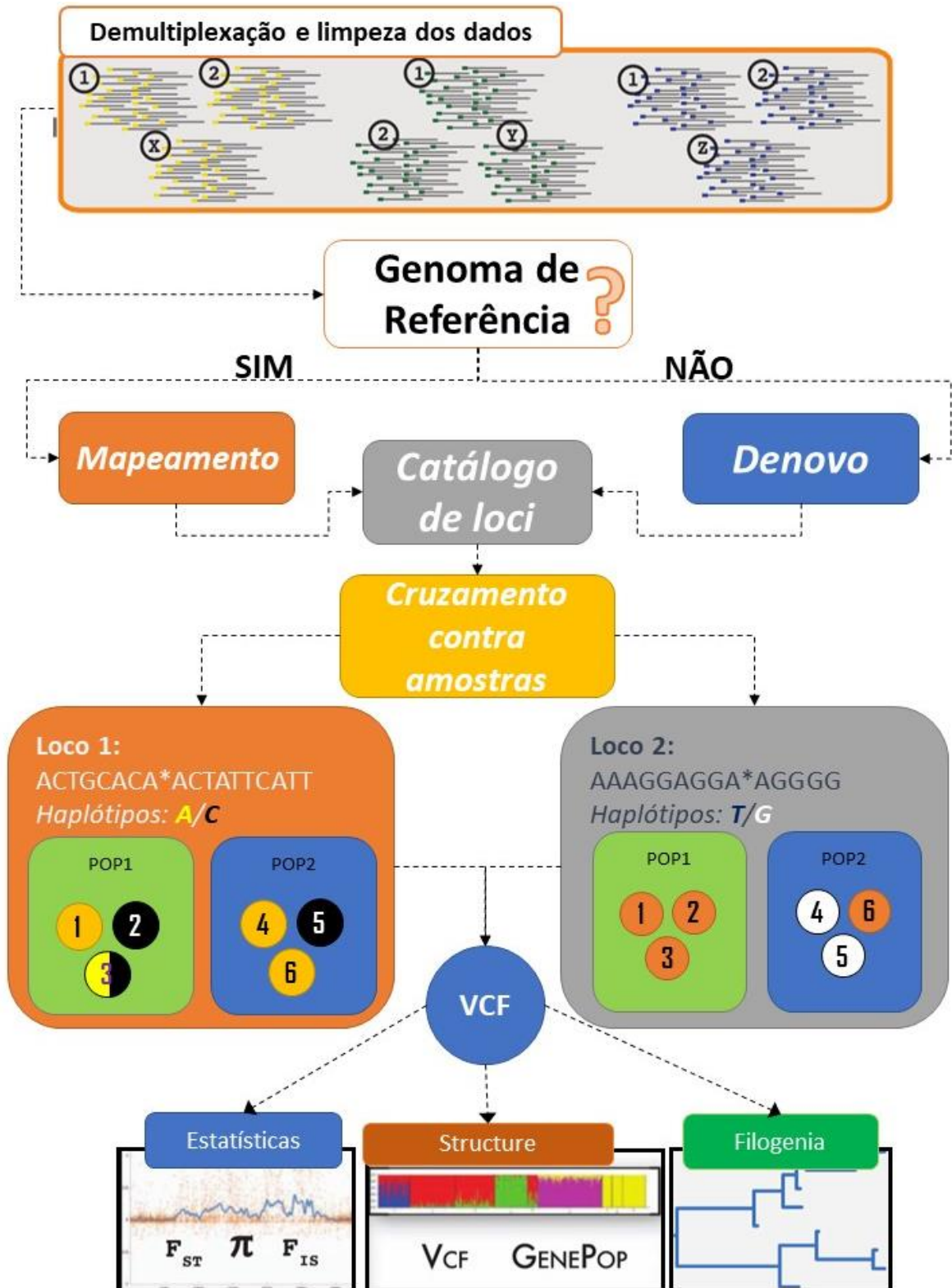


Figura 3.8 Estratégia genérica para análise de dados gerados por métodos de redução de representação genômica (RRL, *RADseq*, GBS). Conjunto de milhões de *reads* são demultiplexados e separados dos reads sem sinais de sítios de restrição. Em seguida, diferentes *pipelines* são assumidas dependendo da disponibilidade de genomas de referência. Deste ponto em diante diversas *pipelines* criam catálogos de locos a partir de algumas amostras, para então verificar a presença destes locos em uma verificação cruzada contra todas as amostras do estudo. Os dados dos locos são anotados em arquivos que guardem estas informações anotando loco, variante e localização do loco, como arquivos VCF e estes são analisados por diferentes programas.

É interessante notar que, independentemente dos passos específicos de cada uma das *pipelines* adotadas, as estratégias são parecidas devido à natureza dos dados, se diferenciando em relação a como os algoritmos buscam e determinam os locos e suas presenças nos indivíduos em diferentes populações (TORKAMANEH; BOYLE; BELZILE, 2018). Um outro aspecto bastante interessante é que mesmo que as *pipelines* encontrem quantidade de locos diferentes, a sobreposição destes locos e principalmente a acurácia são bastante semelhantes entre si, além disso a proporção de locos encontradas é tão grande em relação aos métodos tradicionais que a diferença encontrada entre estas abordagens pode não significar diferenças qualitativas nos bancos de dados (TORKAMANEH; LAROCHE; BELZILE, 2016).

Infelizmente, é impossível eleger uma única *pipeline* que seja igualmente adequada a todas as situações, e neste ponto é importante que as *pipelines* ofereçam a oportunidade de alterar os parâmetros do processo de chamada de SNPs. Contudo, nem todas as *pipelines* são igualmente “transparentes” ou, em outras palavras, oferecem a mesma oportunidade de serem alteradas. Se por um lado, UNEAK e TASSEL-GBS oferecem um desempenho muito bom (TORKAMANEH; LAROCHE; BELZILE, 2016), as mesmas dependem de ferramentas ou algoritmos construídos de forma que os usuários não podem alterar facilmente (por exemplo, para demultiplexação e chamada de SNPs). Além disso, os arquivos de dados intermediários nem sempre são facilmente acessíveis e isso torna mais difícil investigar problemas específicos (LU et al., 2012). Por outro lado, IGS e Fast-GBS tornam fácil a alteração de parâmetros com arquivos intermediários facilmente acessíveis, porém oferecem uma dificuldade metodológica extensa para usuários iniciantes ou para banco de dados muito grandes. Entre essas duas opções está a *pipeline* Stacks, a qual, aparentemente oferece um nível intermediário de “transparência” e acesso aos parâmetros de demultiplexação e chamada de SNPs.

Uma vez que o usuário obtém uma matriz de dados consistente, as análises seguem de acordo com o objetivo do projeto, possibilitando desde acessos a índices de estruturação populacional, como F_{ST} , G_{ST} , R_{ST} , ou análises filogenéticas de construção de árvores de espécies como o RAxML. Detalhes destes métodos são descritos na seção: 3.3.2 Análises em Genética de Populações e Filogeografia.

3.2 FORMAÇÃO E BIOGEOGRAFIA DA AMÉRICA DO SUL: ASPECTOS GERAIS

Considerando que a Filogeografia trata da distribuição de linhagens genéticas ao longo do tempo e espaço geográfico e que, nas palavras de Léon Croizat: “...a *Vida e a Terra evoluem juntas...*”, quaisquer considerações sobre a evolução da forma, sem considerar a evolução do próprio espaço ao longo do tempo estaria incompleta (CROIZAT, 1964). Deste modo, ainda que de forma um tanto rudimentar, as próximas páginas se ocuparão em descrever os principais eventos geológicos que formaram a paisagem atual na América do Sul, com algumas considerações para a América Central, tendo como ponto de partida os últimos 250 milhões de anos, com ênfase em acontecimentos a partir de 80 e 40 milhões de anos até o presente, quando estima-se ter ocorrido a divergência da linhagem dos euglossíneos em relação às demais abelhas corbiculadas (veja a sessão 3.4.3 Aspectos Biogeográficos).

A história geológica da América do Sul é marcada por múltiplos eventos tectônicos e climáticos, transgressões e regressões marinhas as quais moldaram a diversidade biológica nos últimos milhões de anos. Neste conjunto de eventos, a separação da Pangeia e Gondwana (220-110 Ma a. p.), o derramamento basáltico ao longo do escudo cristalino brasileiro (~140 Ma a. p.), o soerguimento dos Andes (~12 Ma a. p.), as transgressões marinhas do Mioceno (~15 Ma a. p.) e as Glaciações do Pleistoceno (2.8 Ma a. p.), têm especial importância do ponto de vista biológico e biogeográfico (LAVINA; FAUTH, 2016). Além disso, três extinções em massa (>75% das espécies) marcam os últimos 250 Ma do planeta e da América do Sul, das quais a mais recente (~66 Ma a. p.) provocou a extinção de grande parte da fauna, como dinossauros e outros animais de grande porte, bem como a flora do período Cretáceo (JABLONSKI, 1994). Após a última grande extinção em massa, outras extinções de menores proporções ocorreram afetando táxons de forma mais seletiva, como o máximo termal do Eoceno (55 Ma a. p., ~50% da microfauna marinha) e a extinção dos mamíferos gigantes já no período Quaternário na época Holoceno, 10 Ka a. p. (JABLONSKI, 1994). A sucessão de eras, períodos, épocas, bem como seus eventos biológicos associados estão resumidos na Figura 3.9.

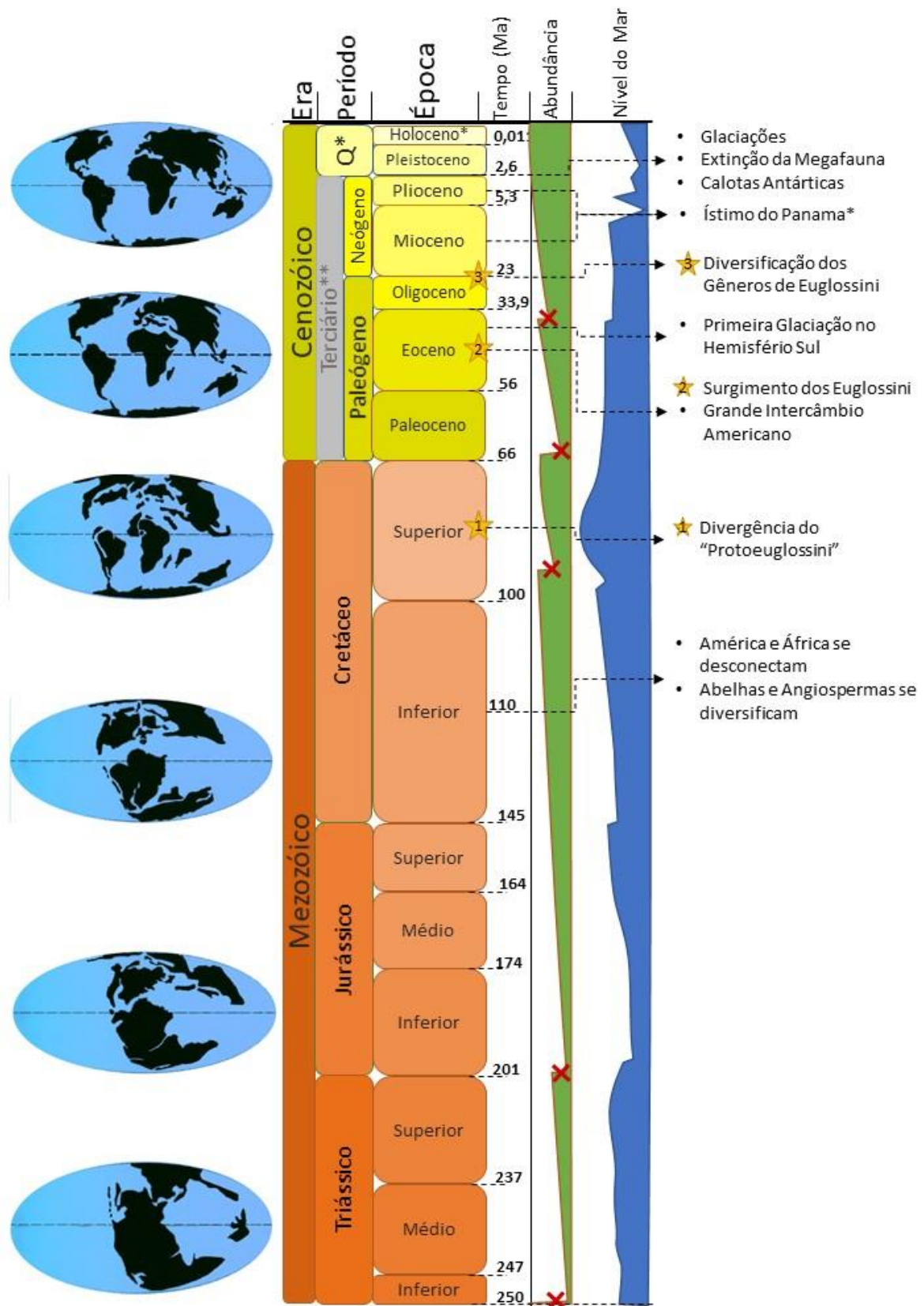


Figura 3.9 Escala de Tempo Geológico e evolução paleogeográfica, climática e alterações no nível do mar ao longo dos últimos 250Ma. Estrelas demarcam momentos importantes de diversificação dos euglossíneos. *Q (Quaternário) e Holoceno foram graficamente aumentados para favorecer a legibilidade. **Período Terciário é uma nomenclatura em desuso de acordo com a Sociedade Americana de Estratigrafia, mas foi mantido para fins de comparação devido sua frequente presença em textos e livros-base. Adaptada de Lavina e Fauth (2016).

3.2.1 O Mesozoico na América do Sul

A Era Mesozoica compreende três períodos bastante longos da história da Terra: Triássico, Jurássico e Cretáceo, compreendidos entre 250 e 66 Ma a. p., demarcados pelos dois últimos grandes eventos de extinção em massa do planeta. Ao longo desta era a América do Sul sofreu intensas modificações separando-se do continente africano e sofrendo, principalmente, modificações na paisagem devido a transgressões marinhas, eventos orogênicos e vulcanismos. Durante a Era Mesozoica, intensos movimentos tectônicos aconteceram, posicionando a América do Sul e América Central praticamente nas latitudes e longitudes atuais, além disso, intensos tectonismos são registrados para esta era ao longo de períodos de climas quentes sem registros de grandes Glaciações (LAVINA; FAUTH, 2016).

Do ponto de vista geomorfológico, a história da América do Sul como continente isolado, se inicia com a divisão do último supercontinente, a Pangeia, em Laurásia e Gondwana, separação iniciada a cerca de 220 Ma a. p. e finalizada entre 110 e 90 Ma a. p. (LAVINA; FAUTH, 2016). Este é um ponto crítico na formação da costa brasileira, com soerguimentos macrodômicos (ascensão de Magma do manto; *hot spots*) que provocaram a elevação das costas brasileira e africana, e o conseqüente rebaixamento de áreas mais ao interior do continente (ORTIZ-JAUREGUIZAR; CLADERA, 2006). A deriva entre os dois continentes causou o rebaixamento de terras na área da atual dorsal Meso-Atlântica, possibilitando a transgressão oceânica entre as bordas dos dois novos continentes, seguida por um extravasamento gradual de lava que aumentou novamente a espessura da crosta formando a dorsal Meso-Atlântica (COX, 1989).

A Era Mesozoica foi marcada por múltiplos soerguimentos macrodômicos que formaram domos de magma, que por sua vez influenciaram o soerguimento das serras do Mar e de Minas Gerais, bem como a formação do Aquífero Guarani (BAKER, 1923; COX, 1989; LAVINA; FAUTH, 2016). Estes soerguimentos estão intrinsecamente relacionados com um dos maiores eventos vulcânicos da história do planeta, que aconteceu antes da separação total entre América do Sul e África, entre 137 e 127 Ma a. p. (LAVINA; FAUTH, 2016). O vulcanismo na bacia sedimentar do Paraná é corroborado pela similaridade das rochas basálticas magmáticas ao longo dos estados de Mato Grosso do Sul, Goiás, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul, bem como no Paraguai, Bolívia e Uruguai, a qual aponta um gigantesco derramamento de magma ao longo da Bacia do Rio Paraná. Este evento foi responsável por modificar drasticamente a paisagem do Centro Oeste da então Gondwana, atual Sudeste da América do Sul (BAKER, 1923; LAVINA; FAUTH, 2016). Embora não

haja descrições diretas da relação biogeográfica entre organismos e o vulcanismo da bacia do Rio Paraná, é evidente que as condições de solo modificadas pelos derramamentos basálticos mesozoicos foram e são fundamentais para a determinação das fitofisionomias e fauna associada.

Em seguida, (~80 Ma a. p.) as altas temperaturas do planeta promovem a maior transgressão marinha da Era Cenozoica, a qual alcançou 250 m acima do nível atual (ORTIZ-JAUREGUIZAR; CLADERA, 2006). Interessantemente, análises moleculares indicam que na época, no Cretáceo Superior, a linhagem do “*Protoeuglossini*” (3.4 As Abelhas das Orquídeas) divergia dos demais corbiculados (BRANSTETTER et al., 2017), indicando uma possível relação entre estes eventos.

3.2.2 A Era Cenozoica na América do Sul

A exemplo da Era Mesozoica, a Era Cenozoica é subdividida em três períodos: Paleogeno, Neogeno e Quaternário; compreendendo o espaço de tempo entre a última grande extinção em massa do planeta há 65 Ma a. p. e o período atual (COHEN et al., 2014). Essa Era representa o intervalo de tempo mais representativo sobre a história biológica da América do Sul, moldando diretamente a diversidade atual.

Conforme é ilustrado na Figura 3.9, o período Paleogeno é dividido em três épocas: *Paleoceno*, *Eoceno*, *Oligoceno*. Foi exatamente ao longo deste primeiro período do Cenozoico que o planeta experienciou uma ampla variação térmica saindo “do ótimo” termal do Eoceno (Paleogeno médio, 55-33,7 Ma a. p.) para as grandes glaciações do Pleistoceno (Quaternário inferior).

Durante o Paleogeno dois eventos de interesse particular ao estudo de Euglossini merecem destaque: i) o Eoceno Superior (~40-38 Ma a. p.), época que inclui os primeiros indícios de soerguimento dos Andes (LAVINA; FAUTH, 2016), é estimado como o limite inferior para a coalescência dos gêneros de Euglossini (RAMÍREZ et al., 2010a); ii) no final do *Oligoceno* (33,7-23,8 Ma a. p.), última época do período Paleogeno (COHEN et al., 2014), e início do Mioceno (25-20 Ma a. p.), primeira época do período Neogeno (Figura 3.9), a conexão entre América do Norte e América do Sul via Istmo do Panamá poderia estar presente presente (BACON et al., 2015); tal fato é particularmente interessante ao entendimento da Filogeografia dos Euglossini, pois indicam a possibilidade de troca faunística entre América do Sul e Central em linha com os fósseis Dominicanos (ENGEL, 1999; POINAR, 1998). Ao fim do Oligoceno, se inicia um novo período, o Neogeno, subdividido em Mioceno e Plioceno (Figura 3.9).

O *Mioceno* é a última grande época em termos de tempo geológico se estendendo entre 23,8 e 5,3 Ma (COHEN et al., 2014), quando os Andes assumiram praticamente a forma atual e importantes transgressões marinhas foram registradas em sua porção média, cerca de 12.5 Ma a. p. (LAVINA; FAUTH, 2016). A subducção de Nazca sob a Placa Americana ao longo da porção superior do Mioceno gerou, além do soerguimento dos Andes, uma importante atividade vulcânica e um respectivo aquecimento atmosférico até o “ótimo climático do Mioceno”, que por sua vez, provocou o afinamento das calotas polares elevando os níveis dos mares em até 100 m acima do nível atual (YOU et al., 2009). Essas elevações provocaram um cenário de transgressões marinhas registradas ao menos três vezes a partir de 17 Ma a. p., representados na Figura 3.10. A porção final do Mioceno é marcada por um resfriamento gradual, com indícios de glaciação a cerca de 7 Ma provocando a regressão dos níveis oceânicos.

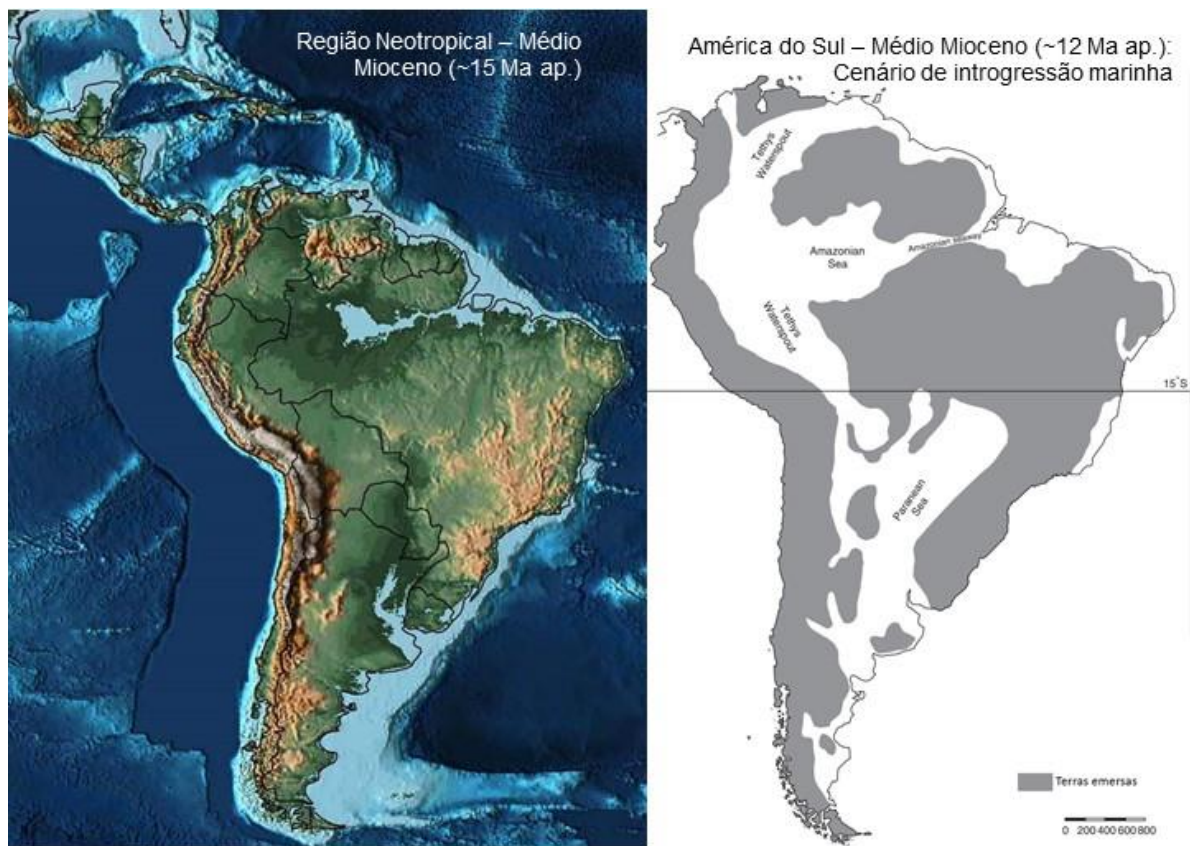


Figura 3.10 Transição entre os cenários de transgressões marinhas ao longo do Médio Mioceno e Mioceno Tardio. À esquerda um cenário intermediário com leve introgressão marinha Pacífica e Atlântica, 17-15 Ma a. p. (SCOTSE, 2001). À direita um cenário de forte introgressão marinha Pacífica e Atlântica, formando o “Mar Paranaense” (ORTIZ-JAUREGUIZAR; CLADERA, 2006).

O *Plioceno* (5,3-2,6 Ma) é a última época do período Neogeno, precedendo o período Quaternário (COHEN et al., 2014). Ao longo desta época a situação geológica da América do Sul já se encontrava bastante definida com predomínio de planícies onde haviam os grandes

mares continentais. Na porção final do Plioceno, os Andes alcançaram altitudes nas porções média e sul das cordilheiras que bloquearam os ventos para o interior do continente criando zonas áridas, influenciando a definição do deserto do Atacama (ORTIZ-JAUREGUIZAR; CLADERA, 2006). Esta época é marcada pelo fechamento definitivo do Istmo do Panamá na forma com a qual se conhece hoje. Vale ressaltar aqui que, embora dados apontem conexões Miocênicas entre América Central e América do Sul, é plausível que em vários momentos de aquecimento e elevação oceânica essas conexões intermitentes tenham se perdido (ODEA et al., 2016). Este fato é corroborado por padrões biogeográficos de peixes caribenhos e principalmente pela troca faunística entre as Américas conhecida como Grande Intercambio Americano (BACON et al., 2016). Para os Euglossini este período representa a reconexão definitiva entre áreas e a possibilidade de expansão distribucional tanto para espécies amazônicas quanto centro-americanas. Ortiz-Jaureguizar e Cladera (2006) apontam um segundo ponto interessante em relação ao fechamento do Istmo do Panamá. Em decorrência deste evento, houve a interrupção do fluxo da corrente Atlântico-Pacífico, a qual possibilitou a formação da calota polar ártica; ao mesmo tempo, a corrente polar antártica torna-se definitiva no hemisfério sul. Juntas, as circulações marinhas ártica e antártica, causariam um efeito de arrefecimento dos polos e um decorrente resfriamento global.

3.2.3 O Quaternário e as flutuações climáticas

O *Quaternário* (2,56 Ma a. p. até o presente) corresponde ao período geológico mais recente da história da Terra, comumente dividido em duas épocas: *Pleistoceno* (2,56 Ma à 11 Ka) e *Holoceno* (11 Ka até o presente) (COHEN et al., 2014). Diferentemente dos períodos anteriores, o Quaternário tem uma duração bastante curta do ponto de vista geológico e não apresenta grandes eventos orogênicos, mas com certeza apresenta uma extraordinária dinâmica de paisagens. A importância destes eventos de flutuação climática é tão significativa para os padrões filogeográficos, que será dada uma atenção especial à origem e consequência destes eventos nos próximos parágrafos.

Interessantemente, até o fim do Neogeno (período anterior), os registros geológicos e biológicos encontrados apontavam a existência de longos intervalos de tempo com relativa uniformidade climática, salvo exceções pontuais altamente perturbadoras, mas com curta duração em tempo geológico (ORTIZ-JAUREGUIZAR; CLADERA, 2006). No entanto, as mudanças climáticas ao longo do Quaternário apresentaram, além de maior amplitude, uma frequência consideravelmente maior (COMES; KADEREIT, 1998; SUGUIO; SALLUN; SOARES, 2005). Apenas nos últimos 760 Ka, ciclos glaciais aconteceram ao menos oito

vezes (FARMER; COOK, 2013). Estudos apontam que as mudanças climáticas do Quaternário foram responsáveis por múltiplas expansões e retrações de áreas de distribuição de táxons, comunidades e biomas (BEHLING, 1998; FRANTINE-SILVA et al., 2017; GRAZZIOTIN et al., 2006; MALDONADO-COELHO, 2012; ORTIZ-JAUREGUIZAR; CLADERA, 2006; RULL, 2011).

Considerando as médias de temperatura, todo o Quaternário pode ser considerado uma grande era glacial, apresentando amplitude térmica de 8°C. Este aspecto fica bastante evidente se comparado com o máximo termal do Paleoceno-Eoceno (média térmica ~8°C maior do que o presente) ou com o Cretáceo (15°C mais quente do que o presente), principalmente considerando que o Holoceno é a época mais quente de todo o Quaternário (FARMER; COOK, 2013).

O Quaternário apresenta registros de vários ciclos glaciais, inicialmente a cada 41 Ka, mas que se tornaram mais esparsos nos últimos 800 Ka, ocorrendo a cada 100 Ka (FARMER; COOK, 2013). Durante este período (~760 Ka a. p. até o presente), oito ciclos glaciares ocorreram (AUGUSTIN et al., 2004) nos quais a calota polar ártica esteve sempre presente, a exemplo de todo o Quaternário (FARMER; COOK, 2013). Atualmente o planeta se encontra em uma fase interglacial, que demarca o início da época Holoceno a 11 Ka a. p. O último período interglacial antes do Holoceno foi à cerca de 130 Ka a. p. (EHLERS; GIBBARD; HUGHES, 2018). As causas dos períodos glaciares são ainda especulativas, mas é provável que um conjunto de fatores tenha atuado para criar ciclos periódicos de resfriamento, dos quais ganham destaque os ciclos astronômicos, efeito estufa, dinâmica tectônica e correntes marinhas (FARMER; COOK, 2013).

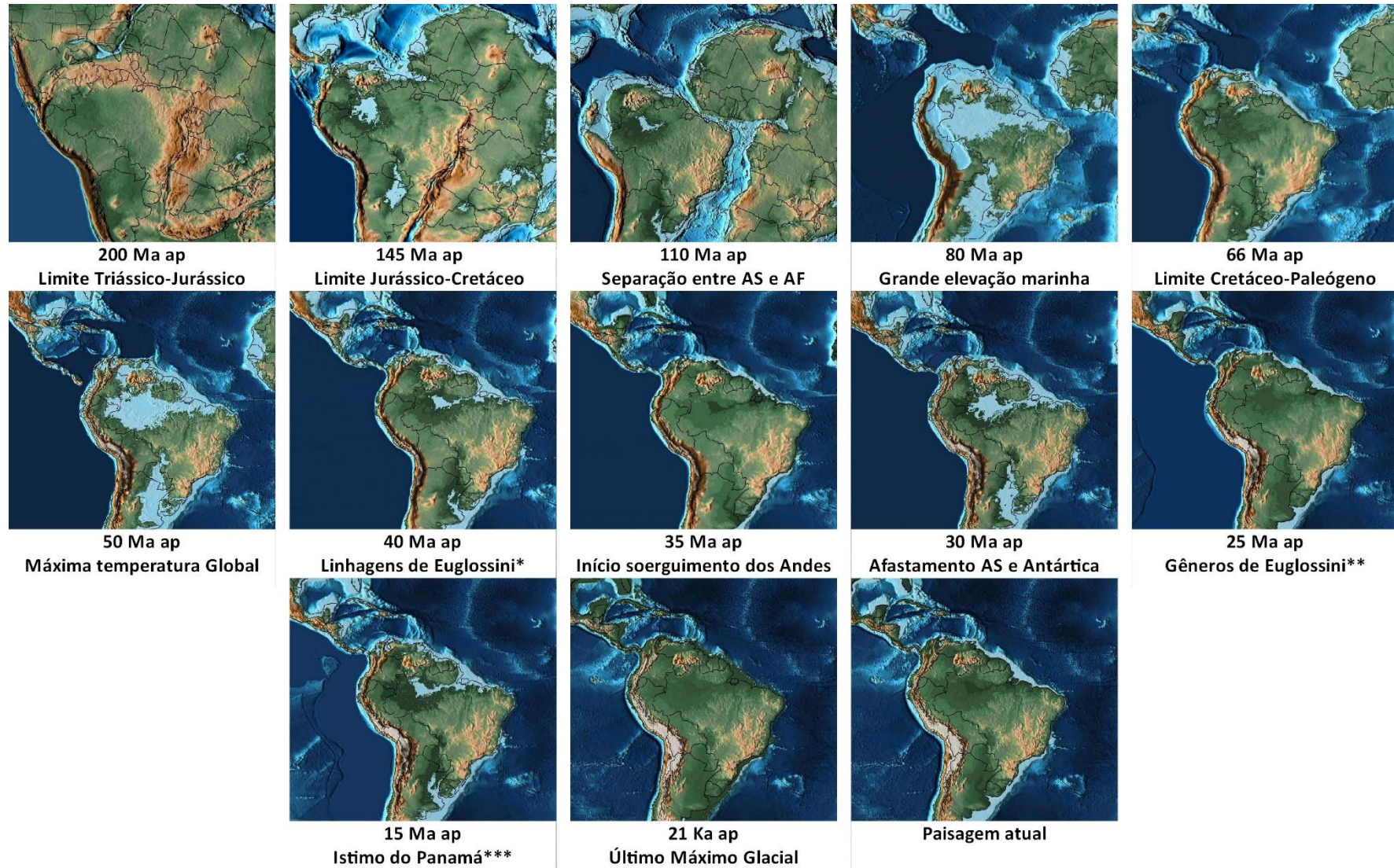
O *Pleistoceno* (2,6 Ma a 11Ka) é a primeira época do Período Quaternário (2,56 Ma até o presente), concentrando toda instabilidade climática associada a este período. Ao longo do Pleistoceno, dois principais fatores afetaram as faunas sul-americanas. O primeiro foram as glaciações que afetaram a fauna diretamente por meio das mudanças climáticas e ambientais, além das mudanças no nível do mar associadas (ORTIZ-JAUREGUIZAR; CLADERA, 2006; RUIZ PESSENDA et al., 2009; SERVANT et al., 1993). O segundo foi a chegada maciça de imigrantes norte-americanos provenientes do “Grande Intercambio Americano” (BACON et al., 2016).

O avanço e recuo cíclicos das geleiras produziram uma expansão e retração de biomas áridos (savanas, estepes) e úmidos (florestas tropicais e subtropicais) (ORTIZ-JAUREGUIZAR; CLADERA, 2006). Este padrão permitiu oportunidades alternadas para animais e plantas com diferentes afinidades a estes ambientes para explorar novos habitats, ao

mesmo tempo que favoreceu altas taxas de extinção como consequência da redução de áreas favoráveis (BARNOSKY et al., 2016). Ao mesmo tempo, a conexão definitiva entre as Américas via Istmo do Panamá possibilitou a troca de fauna entre continentes, principalmente de faunas de vertebrados (ORTIZ-JAUREGUIZAR; CLADERA, 2006), as quais também foram influenciados pelas variações do ambiente ao longo do ciclos glaciais (BACON et al., 2016).

De modo geral, as glaciações cíclicas tiveram forte influência na fauna sul-americana. À medida que as temperaturas caíam e as zonas áridas se expandiam a fauna das florestas úmidas perdiam espaço podendo se isolar das demais áreas, levando à formulação de hipóteses de “refúgios florestais” (HAFFER, 1992) para abrigar a fauna associada a florestas úmidas ao longo dos máximos glaciais. Em contrapartida à retração das áreas úmidas, novas áreas surgiram, por exemplo, a diminuição dos níveis marinhos levou à exposição da plataforma continental Atlântico-Brasileira, a qual possui apenas 200 metros de desnível ao longo de 100 km de comprimento (SOUZA, 1999). Alguns estudos apontam para a ocupação desta plataforma por mamíferos de matas úmidas ao longo do último máximo glacial (21Ka a. p.), demonstrando a ocupação destas áreas por tais biotas (LEITE et al., 2016a).

Um resumo da evolução tectônica e de paisagem da América do Sul e Central são apresentados na Figura 3.11. É possível observar a drástica mudança nos padrões espaciais ao longo dos últimos 250 Ma, o que dá uma breve noção das modificações enfrentadas pela biota ao longo deste período, com destaque para os movimentos de transgressão e regressão marinhos.



1

2 Figura 3.11 Evolução geológica da América do Sul e nível do mar nos últimos 200 Ma a. p. *Limite inferior e **limite superior para o surgimento e diversificação dos
 3 gêneros de Euglossini (RAMIREZ et al., 2010). *** Considerando dados de (MONTES et al., 2015). Fonte: PALEOMAP Project (SCOTese, 2001).

As modificações cíclicas dos eventos glaciares se estenderam até 11Ka a. p., quando são detectados os primeiros sinais de aquecimento atmosférico, dando início à última época do Período Quaternário, o *Holoceno* (época atual), que registra poucos eventos geológicos importantes, além daqueles relacionados com o acúmulo do CO₂ atmosférico e aqueles causados pelo homem. De fato, a expansão populacional da espécie humana, principalmente nos últimos 200 anos, causou uma modificação de paisagem que apenas os eventos mais drásticos da história geológica proporcionaram, comparáveis apenas aos eventos de extinção em massa da história geológica da Terra. Estes efeitos são de tal amplitude que, é possível que no futuro ratifique-se a troca do período *Holoceno* pelo nome *Antropoceno* para referir-se à era dos homens, com o marco da sexta extinção em massa.

A despeito das modificações humanas na paisagem ao longo do período recente, o texto a seguir tem como foco apenas as modificações e efeitos das glaciações do Quaternário sobre a biota e a resposta desta frente à essas alterações, com base em algumas teorias e padrões observados nas florestas úmidas Americanas, especialmente no que se refere às teorias derivadas da Hipótese dos Refúgios de Haffer (1969) e suas contraposições.

3.2.4 Reflexos biogeográficos das Glaciações do Quaternário

Como já destacado, as glaciações do Quaternário tiveram potencialmente uma grande influência sobre a diversificação da fauna sul-americana. O caráter cíclico das modificações na paisagem propiciou múltiplas oportunidades de expansão de áreas de distribuição seguidas por eventos vicariantes cladogênicos nas linhagens de espécies e populações irmãs. As principais ideias e implicações destes cenários sobre a biota da região neotropical são apresentadas abaixo com o intuito de contextualizar a situação das duas grandes florestas tropicais americanas, Amazônia e Mata Atlântica, com viés nas implicações para os Euglossini.

3.2.4.1 Teoria dos Refúgios de Haffer

Em 1969, Jurgen Haffer, geólogo e ornitólogo, apresentou uma teoria extremamente elegante para explicar eventos de especiação a partir das mudanças climáticas ao longo das glaciações. Haffer observou que as distribuições atuais de espécies de aves estreitamente relacionadas, geralmente apresentavam intervalos adjacentes, mas nunca sobrepostos, ou seja, distribuições parapátricas. Estas áreas apresentavam certa previsibilidade nos padrões de áreas de endemismo, os quais eram congruentes com os diferentes padrões pluviométricos utilizados por Haffer para reconstruir as áreas de refúgio. A partir dessas observações, Haffer

especiação; (3) apenas as espécies ecologicamente bastante restritas foram afetadas, enquanto que as populações de espécies ecologicamente mais flexíveis podem não ter sido efetivamente isoladas nos refúgios; (4) padrões paleoplinológicos sugerem uma mata contínua na região amazônica (BUSH; OLIVEIRA, 2006). Embora Haffer (2008) tenha questionado alguns destes pontos contrários à Teoria dos Refúgios, é de fato provável que os padrões Amazônicos tenham sido muito mais complexos ao longo tempo, e que as flutuações climáticas do Quaternário não tenham sido o principal fator de especiação na região (BUSH; OLIVEIRA, 2006).

De fato, alguns padrões biológicos sugerem especiações pretéritas ao Pleistoceno e nestes casos, rios, transgressões marinhas e outros eventos geológicos parecem ser uma explicação mais plausível (para detalhes veja BUSH; OLIVEIRA, 2006). Existem várias outras hipóteses para explicar a diversificação da fauna Amazônica, incluindo rios como barreiras, eventos neotectônicos, e mesmo uma hipótese híbrida entre refúgio e rios como barreiras formulada pelo próprio Haffer para explicar alguns padrões destoantes do modelo dos refúgios (HAFFER, 2008). Finalmente, é provável que muitas dessas hipóteses sejam válidas em algum nível ou para algum grupo taxonômico, dada a complexidade e o longo período em que esses táxons evoluem na região Amazônica.

3.2.4.2 *O modelo Carnaval-Moritz*

A teoria dos Refúgios de Haffer influenciou a maneira com a qual pesquisadores olharam seus dados em florestas tropicais ao longo de décadas. Muito mais comum para explicar padrões em áreas Amazônicas (BROWN; SHEPPARD; TURNER, 1974; HAFFER, 1992, 2008), por vezes este modelo foi aplicado de forma acrítica gerando explicações inconsistentes (BUSH; OLIVEIRA, 2006). Este cenário levou uma mudança de paradigma, na qual vários modelos foram explorados para explicar a ocorrência destes padrões (DA SILVA; PINTO-DA-ROCHA; SOUZA, 2016).

Em 2008, Carnaval e Moritz utilizaram uma abordagem revolucionária para testar a existência de refúgios Pleistocênicos na Mata Atlântica. Estes autores aplicaram técnicas de Modelagem de Nicho Ecológico (frequentemente referido como Modelos de Distribuição de Espécies) para estimar a provável paleodistribuição de vários táxons ao longo da Mata Atlântica. Ao sobrepor estes padrões temporais, Carnaval e Moritz identificaram regiões que apareciam como adequáveis à presença de florestas repetidamente ao longo de períodos glaciais e interglaciais, os quais foram tratados como *Refúgios Florestais* (Figura 3.13; CARNAVAL; MORITZ, 2008).

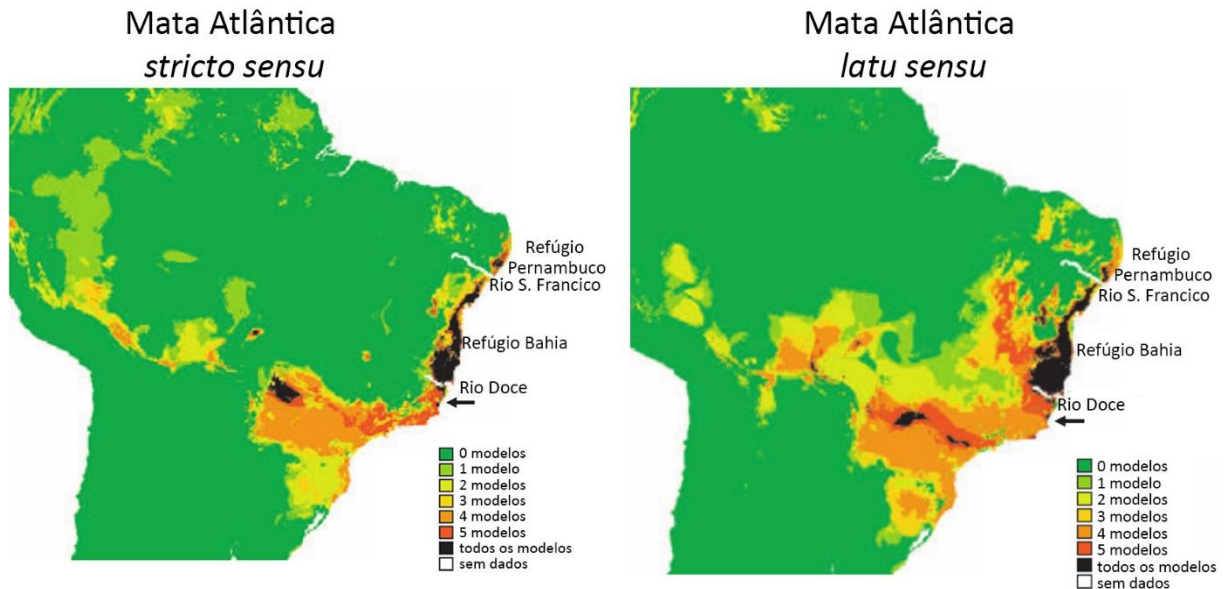


Figura 3.13 Mapas de consenso de áreas historicamente estáveis para a Mata Atlântica sob definições estrita (*stricto sensu*) e ampla (*latu sensu*), estimando a ausência/presença de floresta sob cenários climáticos atuais de 6 Ka e 21 Ka a. p. Linhas brancas demarcam os rios São Francisco e Doce. A flecha indica o pequeno refúgio previsto nos limites dos estados do Espírito Santo e do Rio de Janeiro. Refúgios no interior da Atlântica em Floresta Estacional Semidecidual também são observáveis (meu comentário). Adaptado de Carnaval e Moritz (2008).

Assim como na definição de Haffer (1969), os refúgios *sensu* (CARNAVAL; MORITZ, 2008) seriam ilhas de vegetação que permaneceram estáveis ao longo das glaciações onde as espécies adaptadas a tais condições se abrigariam em períodos secos, quando as condições climáticas seriam menos favoráveis e a vegetação xerófita expandiria seus domínios. Carnaval e Moritz, identificaram ao menos duas grandes regiões que mantiveram áreas de florestas ao longo do Último Máximo Glacial (21 Ka a. p.; *Last Glacial Maximum, LGM*), o “Refúgio de Pernambuco” e o “Refúgio da Bahia”.

Ao contrário do que foi constatado na Amazônia, na Mata Atlântica estes refúgios se mostraram bastante consistentes quando confrontados contra dados paleopalinológicos (CARNAVAL; MORITZ, 2008), além de serem especialmente congruentes com padrões filogeográficos (CARNAVAL; MORITZ, 2008; MARTINS, 2011). Interessantemente, as áreas preditas como refúgios são também altamente congruentes com a diversidade genética de vários organismos encontrados ao longo da Mata Atlântica (CARNAVAL et al., 2009), o qual corrobora a ideia de estabilidade nas áreas de refúgios. Devido ao fato das áreas preditas como refúgios por Carnaval e Moritz (2008) explicarem a estruturação filogeográfica e a diversidade genética de muitas espécies, este modelo de refúgios para a Mata Atlântica ficou conhecido como modelo CM.

De acordo com o modelo CM, enquanto houve retração de áreas de Mata Atlântica mais ao sul no LGM (21 Ka a. p.) seguidas pela expansão dessas áreas durante o Holoceno (~

6 Ka a. p.), as regiões ao norte da Mata Atlântica se mantiveram estáveis. Assim, o modelo CM prediz: 1) modelos de nicho ecológicos prevendo refúgios ao norte durante o LGM, seguido de expansão de áreas no Holoceno; 2) menor estabilidade demográfica nas regiões ao sul da floresta em relação aos refúgios; 3) sinais de expansão populacional após o LGM em áreas ao sul; 4) maior diversidade genética nos refúgios do que em áreas instáveis; (5) evidência de contato secundário com base na distribuição parapátrica das populações ou espécies relacionadas.

Embora muitos dados corroborando o modelo CM tenham sido produzidos (CARNAVAL et al., 2009; MARTINS, 2011; RAPOSO DO AMARAL et al., 2013), assim como na Amazônia, outras hipóteses tem sido levantadas para explicar a diversificação dos táxons e populações ao longo da Mata Atlântica, incluindo neotectonismos, barreiras e pressões seletivas (BATALHA-FILHO et al., 2010; CABANNE; SANTOS; MIYAKI, 2007; LEITE et al., 2016a; THOMÉ et al., 2014). Ainda assim, o modelo CM mostra-se bastante congruente quando sobreposto com a maior parte dos dados filogeográficos disponíveis na Mata Atlântica (CARNAVAL et al., 2009, 2014; CARNAVAL; MORITZ, 2008; MARTINS, 2011; PORTO; CARNAVAL; DA ROCHA, 2013).

3.2.4.3 *Teoria da Expansão da Mata Atlântica*

Embora o modelo CM tenha se tornado muito popular para explicar padrões biogeográficos ao longo da Mata Atlântica, recentemente um modelo gerado a partir de dados de pequenos marsupiais endêmicos à Mata Atlântica tem desafiado essa hipótese. Com base em simulações populacionais a partir de dados mitocondriais de cinco espécies de pequenos mamíferos, Leite et al. (2016) observaram um movimento aparentemente contrário ao descrito no modelo. Como visto anteriormente, o modelo CM prediz, entre outras situações inerentes, a ocorrência de movimentos de retração ao longo do último máximo glacial (CARNAVAL et al., 2009; CARNAVAL; MORITZ, 2008). No entanto, os resultados de Leite et al. (2016) apresentaram tendências de expansão populacional datadas exatamente para o último máximo glacial, sugerindo que suas populações teriam aumentado demograficamente ao invés de retrair.

Com base nestes resultados, Leite et al. (2016a) propuseram um avanço nos modelos paleoclimáticos até então realizados, incluindo a plataforma continental brasileira como área habitável em função da retração oceânica comumente anotada para os períodos de grandes glaciações. Ao incluir a área da plataforma continental, estes autores notaram que as áreas preditas como adequáveis para as espécies em questão se expandiam em direção à mesma, de

modo que estes autores propuseram que as florestas não teriam se retraído e sim aumentado em área, em direção à plataforma continental. A principal conclusão destes autores foi de que a hipótese dos refúgios (representada pelo modelo CM) teria pouca ou nenhuma importância na diversificação da Mata Atlântica.

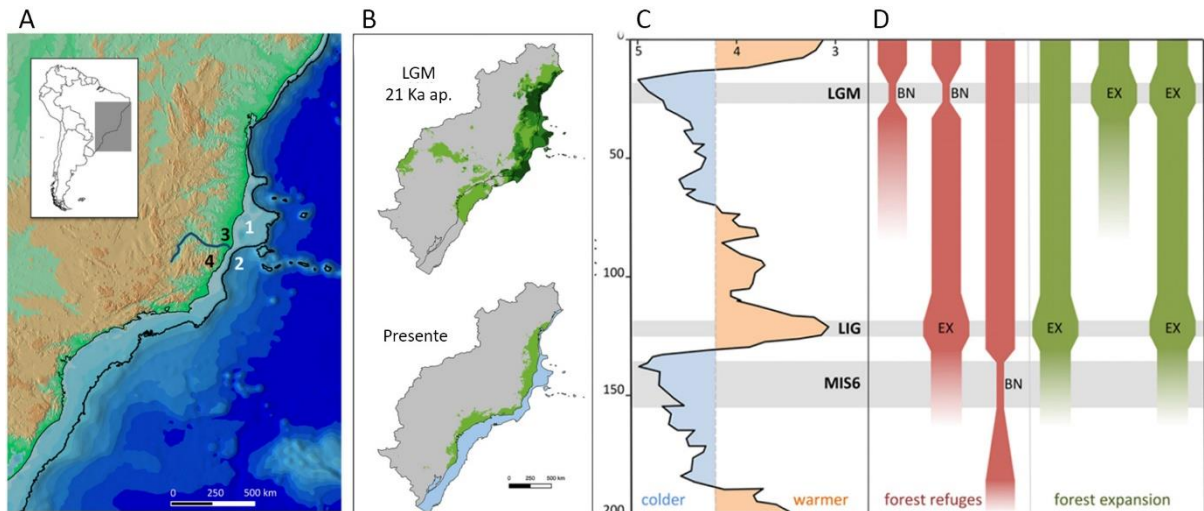


Figura 3.14 Resumo do Modelo de Expansão da Mata Atlântica. A) Costa Brasileira e Plataforma Continental; B) Modelos de Nicho Ecológico (segundo Leite et al., 2016) para o Último Máximo Glacial (LGM) e o presente, mostrando a expansão das áreas adequáveis em direção à plataforma; C) oscilações de temperatura e nível do mar ao longo dos últimos 200 Ma, englobando o LGM, o último período interglacial (LIG) e penúltimo máximo glacial (MIS6); D) Modelos Demográficos representando os eventos esperados sob a hipótese dos Refúgios (vermelho) e a hipótese de Expansão da MA (verde); BN – *bottleneck* populacional; EX - expansão populacional. Adaptado de Leite et al. (2016).

Embora o modelo de Leite et al. (2016a) tenha ineditamente incluído as variações marinhas nos modelos da Mata Atlântica, uma série de problemas foram imediatamente levantados contra esta hipótese, e principalmente quanto a ideia de que os refúgios na Mata Atlântica teriam tido pouca ou nenhuma importância na história biogeográfica deste bioma. Raposo do Amaral et al. (2016) apontam, dentre outros fatores, o uso de apenas cinco espécies e um único loco mitocondrial no estudo, como pontos frágeis para explicar toda a história da Mata Atlântica. De fato, como salientado por Raposo do Amaral et al. (2016) a hipótese de Leite e colaboradores (2016a) não explica uma série de trabalhos corroborando a hipótese dos refúgios na Mata Atlântica, bem como não abrange diferentes modelos ecológicos e nem é passível de ser refutada do ponto de vista genético, uma vez que a plataforma continental brasileira está submersa (RAPOSO DO AMARAL et al., 2016).

Ainda que as críticas de Raposo do Amaral et al. (2016) sejam bastante contundentes é inegável que a hipótese de Leite et al. (2016a) abre uma nova perspectiva do ponto de vista filogeográfico na Mata Atlântica, e deste modo, torna-se uma hipótese alternativa a padrões não congruentes com a hipótese dos refúgios que deve ser igualmente considerada em estudos futuros (LEITE et al., 2016b).

3.2.5 Padrões biogeográficos na Mata Atlântica

A Mata Atlântica compreende um complexo de diferentes fitofisionomias e áreas de transição que podem apresentar uma mistura das fitofisionomias adjacentes ou ainda uma terceira e/ou quarta matriz nas regiões de ecótono (AB'SÁBER, 2003). Dentre as diversas fitofisionomias, as principais são: 1) Floresta Ombrófila Densa (FOD), incluindo as florestas montana, submontana, planícies, e florestas costeiras de baixas altitudes; 2) Floresta Ombrófila Mista (FOM), com predomínio do Pinheiro do Paraná (*Araucaria angustifolia*) e espécies da família Lauraceae ao sul; 3) a Floresta Estacional e Semidecidual (FES), englobando a floresta subtropical ao sudeste do domínio de Mata Atlântica e a floresta decídua ao norte e nordeste; 4) ecossistemas associados que, em tese, incluem mangues, restingas, campos rupestres (campos de altitudes elevadas), campos de baixas altitudes, além de remanescentes florestais à noroeste da distribuição original da Mata Atlântica (CÂMARA, 2005).

Originalmente o Domínio de Mata Atlântica compreendia cerca de 1,3 milhões de quilômetros quadrados ao longo da porção leste do Brasil (cerca de 15% do território brasileiro) estendendo-se desde o estado do Rio Grande do Norte até o estado do Rio Grande do Sul (Figura 3.15; GALINDO-LEAL; CÂMARA, 2003; RIBEIRO et al., 2009).

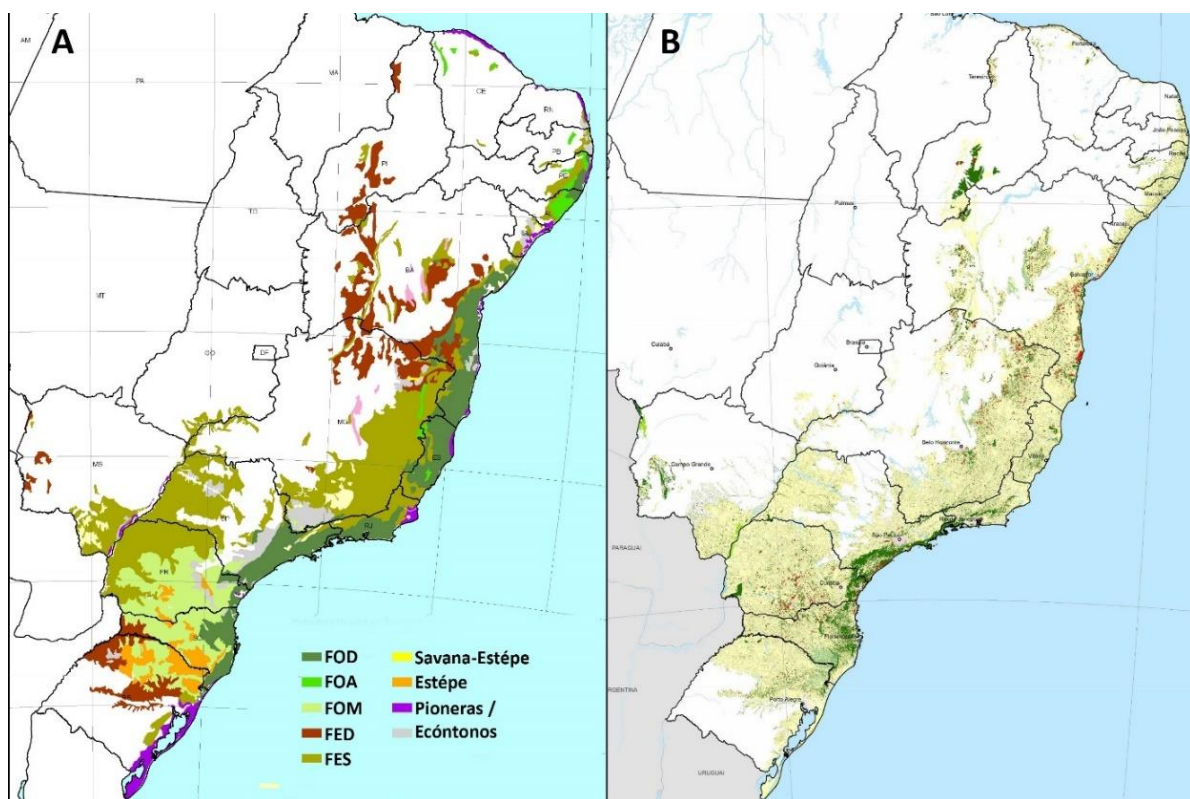


Figura 3.15 Extensão original e atual da Mata Atlântica brasileira. A) área original: FOD – Floresta Ombrófila Densa, FOA – Floresta Ombrófila Aberta, FOM – Floresta Ombrófila Mista, FED – Floresta Estacional

Decidual, FES – Floresta Estacional e Semidecidual; B) Fragmentos de Mata Atlântica, ano base 2016-2017. Fonte: Fundação SOS Mata Atlântica, Relatório Anual (2018).

Diferentemente da Floresta Amazônica, a Mata Atlântica está imbuída em um contexto muito mais heterogêneo. Há um acúmulo de fatores como a grande variação latitudinal, alcançando cerca de 27° de variação (SILVA; CASTELETTI, 2005), ~10° de variação longitudinal, com diferenças significativas entre florestas costeiras e de interior (RIZZINI, 1997), além de um importante gradiente altitudinal, variando desde o nível do mar até 2,700 m (GALINDO-LEAL; CÂMARA, 2003). Todos estes fatores, associados a variações pluviométricas e térmicas contribuem com a grande diversidade e endemismo de um dos maiores *hot-spots* de diversidade biológica do planeta (BROOKS et al., 2006; MYERS et al., 2000).

O estado extremo de degradação da Mata Atlântica vem sendo reportado há anos (GALINDO-LEAL; CÂMARA, 2003; LAURANCE, 2009; METZGER, 2009; RIBEIRO et al., 2009), e historicamente é um dos ‘biomas’ mais ameaçados do planeta (LAURANCE, 2009; RIBEIRO et al., 2009), muito devido à grande exploração antrópica e alta densidade populacional. De acordo com dados do último censo, realizado em 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2010), as maiores densidades populacionais do país (cerca de 100h/km²) e mais de 75% dos grandes centros urbanos se concentram nas regiões litorâneas e nas zonas que correspondem à área de cobertura original do Domínio de Mata Atlântica.

Além de reduzir drasticamente a cobertura original deste bioma, as atividades antrópicas ainda causam danos constantes a todas as espécies presentes, principalmente devido às atividades agrícolas e industriais (CÂMARA, 2005; TABARELLI et al., 2010). Estima-se que mais de 80% dos fragmentos são menores que 50 ha e quase metade da floresta remanescente possui menos de 100 m em suas bordas com distância média entre os fragmentos de 1440 m, em meio a matrizes urbanas ou agrícolas (RIBEIRO et al., 2009). Dados do Relatório Anual de Conservação da Mata Atlântica de 2018 (Fundação SOS Mata Atlântica) indicam que apenas 9% da floresta remanescente e 1% da floresta original são protegidos por reservas e unidades de conservação. Todos estes dados ressaltam a urgência de estudos que levantem informações sobre a diversidade deste importante bioma, fornecendo informações fundamentais para a proposição de futuras estratégias de manejo e conservação das populações envolvidas e de seus ambientes.

3.2.5.1 Padrões Biogeográficos na Mata Atlântica: o Estado da Arte

Como visto anteriormente, a Biogeografia presta papel fundamental para a conservação, uma vez que a delimitação e reconhecimento de áreas de endemismos e áreas prioritárias à conservação é um aspecto reconhecidamente biogeográfico (LAVINA; FAUTH, 2016). A inclusão de métodos que incluem aspectos ecológicos sob o ponto de vista da Biogeografia Histórica (3.1.1.5 Integração de processos ecológicos nas análises biogeográficas), tendem a tornar os aspectos biogeográficos cada vez mais um ponto central nas estratégias de conservação.

Do ponto de vista de padrões filogeográficos na Mata Atlântica, embora a grande parte dos estudos tenham sido desenvolvidos em vertebrados, principalmente pássaros e anfíbios, evidências até o momento apontam a existência de algumas discontinuidades filogeográficas recorrentes para diferentes grupos de organismos (BATALHA-FILHO; CABANNE; MIYAKI, 2012; BATALHA-FILHO; MIYAKI, 2011, 2016; DA SILVA; PINTO-DA-ROCHA; SOUZA, 2016) e são apresentados na (Figura 3.16)

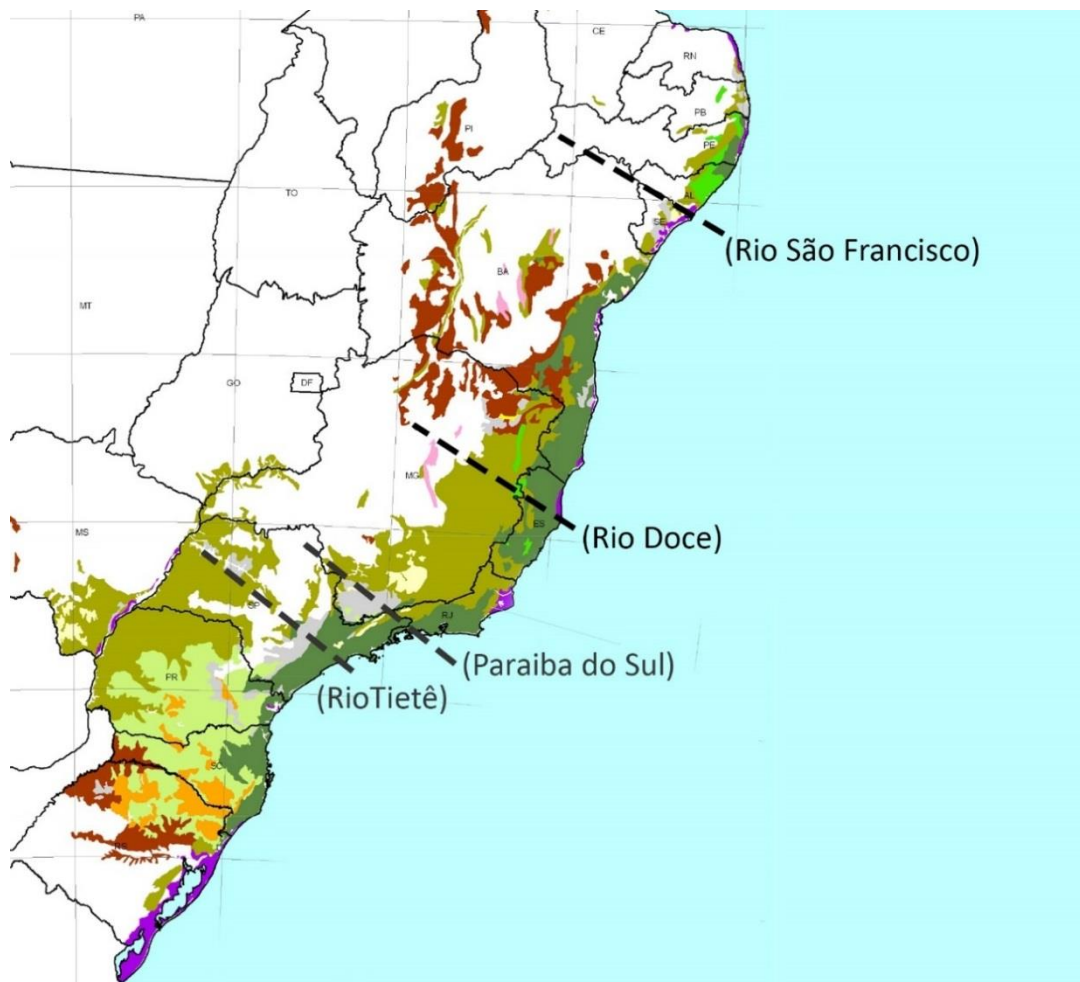


Figura 3.16 Principais descontinuidades filogeográficas ao longo da Mata Atlântica. Descontinuidade **Rio São Francisco**: Cabanne et al. (2008), Carnaval e Moritz (2008), Carnaval et al. (2009), Porto et al. (2013), López-Uribe et al. (2014), Frantine-Silva et al. (2017); **Rio Doce**: Cabanne et al. (2007; 2008), Carnaval e Moritz

(2008), Carnaval et al. (2009), Martins (2011), Porto et al. (2013), López-Uribe (2014), Batalha-Filho e Miyaki (2016); Ribeiro, Lemos-Filho e Lovato (2011); Maldonado-Coelho (2012); **Rio Paraíba do Sul**: Cabanne et al. (2008), Raposo do Amaral et al. (2013); Penha et al. (2015), Frantine-Silva et al. (2015); **Rio Tietê**: Grazziotin et al. (2006), Cabanne et al. (2007; 2008), Martins et al. (2009), Batalha-Filho et al. (2010) e Thomé et al. (2010). *Descontinuidades baseadas em Batalha-Filho e Miyaki et al. (2011) e Da Silva, Pinto da Rocha e Souza (2016), com inclusão de alguns novos estudos. Imagem: Adaptada de SOS MATA ALTÂNTICA (2016).

As explicações para a existência dessas descontinuidades filogeográficas ao longo da Mata Atlântica podem ser resumidas a duas principais linhas, a influência das mudanças climáticas do Quaternário (BATALHA-FILHO; MIYAKI, 2011) e a existência de barreiras geográficas como agentes vicariantes (BATALHA-FILHO et al., 2010; THOMÉ et al., 2010, 2014). Aparentemente, a primeira explicação baseada nas variações climáticas do Quaternário é a que encontra maior respaldo. Muitas das descontinuidades encontradas são coincidentes com a posição pretérita dos refúgios identificados pelo modelo CM (BATALHA-FILHO; MIYAKI, 2011, 2016; CARNAVAL; MORITZ, 2008; VIADANA; PAULA; CAVALCANTI, 2007). Exemplos interessantes são encontrados na distribuição geográfica de linhagens genéticas de estudos filogeográficos de pássaros (BATALHA-FILHO; MIYAKI, 2011, 2016; CABANNE et al., 2014). Entretanto, a existência, localização e o tamanho dos refúgios ecológicos durante os máximos glaciais são dependentes das tolerâncias ecológicas e ambientais de cada espécie (BATALHA-FILHO; MIYAKI, 2011; CARNAVAL; MORITZ, 2008; DA SILVA; PINTO-DA-ROCHA; SOUZA, 2016; MORITZ et al., 2000). Portanto, a interpretação dos dados filogeográficos depende da existência de informações sobre a biologia dos organismos.

No que tange a segunda linha de explicações relacionada à existência de barreiras geográficas ligados a eventos neotectônicos, aparentemente este tipo de diversificação está ligada a regiões mais ao sul da Mata Atlântica (latitudes inferiores à 19°), exatamente onde o modelo CM falha em recuperar modelos de forma mais consistentes. Exemplos com serpentes (GRAZZIOTIN et al., 2006), anfíbios (THOMÉ et al., 2010, 2014) e mesmo abelhas (BATALHA-FILHO et al., 2010), mostram recorrentes quebras filogeográficas localizadas ao longo do estado de São Paulo. Estas descontinuidades não são tão recorrentes quanto aquelas observadas na região do rio Doce e rio São Francisco (MARTINS, 2011) e aparentemente dependem das restrições ecológicas de cada grupo (PORTO; CARNAVAL; DA ROCHA, 2013).

Poucos dados estão disponíveis acerca da estruturação filogeográfica dos euglossíneos ao longo da Mata Atlântica, de modo que não é possível delinear quaisquer padrões para a tribo. Contudo, interessantes casos do ponto de vista biogeográficos são oferecidos por essas abelhas. Por exemplo, ao menos três espécies exibem morfotipos de coloração associadas a

quebras filogeográficas na região do rio Paraíba do Sul (tomado aqui apenas como ponto de referência). *Euglossa iopoecila* Dressler 1982, *Euglossa townsendi* Cockerell 1904 e *Euglossa stellfeldi* Moure 1947 são exemplos de euglossíneos que apresentam morfotipos de coloração congruentes com separação e contato secundário nesta região. Descrições de variações tegumentares são encontradas em Ferrari e Melo (2014) e principalmente Penha et al. (2015). Estes autores apresentam análises de diferentes marcadores moleculares, embasando hipóteses de que eventos relativamente recentes isolaram as abelhas das orquídeas e que estas exibem zonas de contato secundário após os isolamentos pretéritos. No entanto, existe ainda uma lacuna bastante grande no entendimento de quantas e quais seriam estas descontinuidades filogeográficas e suas respectivas zonas de contato secundário, bem como se elas afetariam a distribuição da diversidade e estruturação genética de mais de uma espécie de Euglossini ao longo de todo o domínio de Mata Atlântica.

3.3 DIVERSIDADE, GENÉTICA E CONSERVAÇÃO

Ao longo dos últimos anos, a diversidade genética tem assumido cada vez mais uma posição central nas discussões acerca da conservação da biodiversidade. A crescente ameaça antrópica e a necessidade de se conservar a biodiversidade levaram gradualmente a união de esforços de diferentes áreas para o consenso de que a conservação da biodiversidade passa pelo isolamento e manejo da origem da diversidade biológica, que por sua vez depende da compreensão do contexto genético-evolutivo do tema.

Com base no contexto genético-evolutivo, a biodiversidade pode ser encarada como um produto da evolução biológica, cuja diversidade de formas é fruto do acúmulo de variações hereditárias, inicialmente polimórficas dentro de espécies que posteriormente se fixam em unidades taxonômicas (FRANKHAM; BALOUE; BRISCOE, 2008; FUTUYMA; AGRAWAL, 2009). Segundo este ponto de vista, e levando em consideração a origem da informação que codifica a diversidade dos organismos, a biodiversidade é em grande parte fruto da diversidade genética, que por sua vez, pode ser definida como diversidade de alelos nos vários locos de uma espécie, os quais se combinam produzindo variação genotípica, e essa é herdada pela prole a partir de seus genitores ao longo das gerações, podendo expressar ou não alguma variação fenotípica (SANTOS et al., 2009; SÒLE-CAVA; CUNHA, 2012).

Ao longo do tempo, uma subárea da Genética, a Genética da Conservação, se apresentou com o papel de utilizar conhecimentos da Genética de Populações a fim de reduzir o risco de extinção das espécies ameaçadas. O principal objetivo dessa linha de pesquisa em longo prazo é preservar as espécies como unidades dinâmicas evolutivamente significativas

capazes de se adaptarem às mudanças ambientais, permitindo a elas manter o maior potencial evolutivo possível (SÒLE-CAVA; CUNHA, 2012; FRANKHAM et al., 2008). Para isso a identificação destas unidades é fundamental, e é neste ponto que a Filogeografia presta auxílio de suma importância, buscando delimitar a estruturação espacial da diversidade genética em seus mais diversos níveis taxonômicos, visando delimitar áreas prioritárias à conservação. Em um cenário cada vez mais hostil à conservação da biodiversidade é crescente a necessidade de se documentar e analisar a fonte histórica da variação genética, buscando compreender os eventos por trás da diversificação das espécies e populações, para que, com base no passado tenhamos expectativas sobre o futuro.

Considerando a notória importância da Genética para a conservação da Biodiversidade e que sua discussão é bastante extensa, aliado ao fato de que, de certo modo, este tema foge ao escopo objetivo do presente estudo, considerações mais profundas sobre este tema em particular serão omitidas a fim de priorizar temas mais particulares às abelhas e que refletirão mais enfaticamente na discussão dos resultados. Ótimas revisões em português sobre os temas relacionados à Genética e Conservação podem ser encontradas em Frankham, Balloe e Briscoe (2008) e Solé-Cava e Cunha (2012).

3.3.1 Genética e Conservação das Abelhas

Um dogma central em Genética da Conservação é que populações pequenas e isoladas possuem um maior risco de extinção por fatores extrínsecos, principalmente, devido à deriva genética e endogamia (BERTORELLE et al., 2009; FRANKHAM, 2003; FRANKHAM; BALOUE; BRISCOE, 2008). Ao longo das últimas décadas, tem sido detectado a iminente ameaça do declínio dos polinizadores, fato que tem motivado pesquisas sobre a biologia da conservação de abelhas a fim de entender as causas desse declínio e desenvolver estratégias apropriadas de conservação (POTTS et al., 2010; ZAYED, 2009). A aplicação dos conceitos da Genética da Conservação em animais diploides tem se mostrado uma importante ferramenta, tanto para responder questões sobre a viabilidade populacional quanto para orientar os programas de conservação (FRANKHAM; BALOUE; BRISCOE, 2008). No entanto, as abelhas são potencialmente mais suscetíveis às mudanças no tamanho populacional e nível de conectividade de habitats devido a várias características intrínsecas ou muitas vezes exclusivas relacionadas a fatores biológicos igualmente únicos em diversas espécies.

Devido ao seu caráter haplodiploide, é esperado que abelhas possuam cerca de 75% da diversidade genética em relação a organismos diploides (HEDRICK; PARKER, 1997), o que

naturalmente acarretaria uma “desvantagem” a esses organismos. Adicionalmente, três outros importantes fatores ameaçam a manutenção da diversidade genética nas populações de abelhas: o sistema complementar de determinação sexual, o efeito da depressão endogâmica, e a decorrente perda do potencial adaptativo-evolutivo das espécies (ZAYED, 2009).

Abelhas possuem um sistema de loco único dependente de complementariedade para determinação sexual (sl-CSD ou *single locus Complementary Sex Determination*), no qual o genótipo heterozigoto resulta em um indivíduo fêmea (SCHMIEDER; COLINET; POIRIÉ, 2012). Em outras palavras, quando uma fêmea acasala com um macho haploide que compartilha um dos alelos para o loco sl-CSD, metade da progênie diploide da fêmea será homozigota no loco de determinação do sexo e, conseqüentemente, se desenvolverá em machos diploides ao invés de fêmeas (ZAYED; PACKER, 2005). Machos diploides representam um grande custo genético nas populações de abelhas, uma vez que são inviáveis ou estéreis em sua maioria (HARPUR; SOBHANI; ZAYED, 2013). Considerando que, fêmeas ao fertilizar ovos geram outras fêmeas, a produção de machos diploides a partir de óvulos fertilizados, age aumentando a “mortalidade” ou nulidade das fêmeas na população. Adicionalmente, a produção de machos diploides também traz prejuízos caso os machos sejam viáveis e consigam acasalamentos, visto que fêmeas que se acasalarem com eles produzirão ovos inviáveis quando fertilizados ou filhas triploides estéreis (ZAYED, 2009; ZAYED; ROUBIK; PACKER, 2004).

Dada a natureza complementar do sistema de determinação sexual em abelhas, os fenômenos populacionais que causam naturalmente a redução de diversidade genética, como deriva genética, seleção natural, endogamia ou redução do tamanho populacional efetivo, atuam acentuando o desbalanço na proporção sexual ao longo das gerações (LOZIER; ZAYED, 2017). Vale ressaltar que um sistema dependente de heterozigosidade para o equilíbrio da proporção sexual é duplamente afetado pelos efeitos de redução de diversidade genética supracitados. Caso a heterozigosidade do loco CSD seja afetada de forma significativa é esperado que a probabilidade de homozigosidade seja também aumentada, gerando um desbalanço no número de machos e fêmeas, o que acarretaria na diminuição do tamanho efetivo populacional, que por fim, aumentaria a susceptibilidade da população a eventos estocásticos de redução de diversidade genética. Ou seja, populações pequenas sofreriam com a perda natural de variação genética que levaria a um aumento de endogamia, que por sua vez levaria a um potencial aumento de produção de machos de diploides, e conseqüentemente a uma diminuição do tamanho populacional, que reduziria a diversidade genética da população e assim por diante. Este quadro de retroalimentação em direção à

extinção é tratado por alguns autores como “*o vórtice de machos diploides*”(FARIA et al., 2016; ZAYED, 2009; ZAYED; PACKER, 2005).

Modelos matemáticos demonstram claramente a redução do tamanho efetivo populacional em cenários de acúmulo de produção de machos diploides (ZAYED, 2004; ZAYED; PACKER, 2005), ao passo que este fenômeno pode ser contrabalanceado quando a dispersão dos indivíduos é suficientemente grande, mesmo quando há poucos alelos no loco CSD (FARIA et al., 2016). Deste modo, a produção de machos diploides pode ser um grande problema para a conservação da diversidade genética em abelhas, mas ao mesmo tempo pode ser naturalmente minimizada quando se promove a conexão entre ambientes de forma adequada. Os efeitos e taxas de machos diploides em populações naturais de Euglossini são discutidos no tópico 3.4.4 Diversidade Genética: o estado da arte em Euglossini.

Por razões semelhantes, mas por mecanismos diferentes, o endocruzamento (cruzamento entre indivíduos geneticamente semelhantes ou aparentados) pode causar um importante efeito sobre pequenas populações de abelhas. Além de ter efeito sobre a expressão de alelos deletérios quando em homozigose ou reduzir o valor adaptativo da população (REED; FRANKHAM, 2003), a redução da diversidade genética via endogamia em abelhas pode ainda causar o efeito do “*vortex do macho diploide*”, como já descrito. Interessantemente o sistema haplodiploide oferece a oportunidade de seleção direta de genes deletérios em machos, uma vez que a hemizigose de um alelo deletério implicaria na morte do macho. Logo, a frequência deste tipo de alelo em populações haplodiploides deve se manter em níveis bastante baixos.

Em suma, as complexas histórias de vida das abelhas podem torná-las ainda mais suscetíveis à perda de habitat e outras mudanças ambientais extrínsecas. Muitas abelhas necessitam de diversos recursos além daqueles considerados básicos (como alimento e abrigo) ao longo de seu ciclo de vida. Os próprios euglossíneos requerem fontes aromáticas específicas (WEBER et al., 2016) para realizar comportamentos de corte e cópula (ELTZ; SAGER; LUNAU, 2005), em algumas espécies, isso pode ser acentuado, por exemplo, por um alto grau de filopatria (AUGUSTO; GARÓFALO, 2011; LÓPEZ-URIBE et al., 2014) que poderia causar uma dependência adicional ao local ou mesmo uma estruturação geográfica diferencial. Todos estes fatores em conjunto demonstram a urgência de se explorar a diversidade genética em abelhas a fim de identificar populações em declínio, bem como propor medidas de conservação de forma mais eficiente.

As medidas de diversidade e estruturação genética são abordadas de forma genérica na seção seguinte, a fim de apresentar um resumo visando a contextualização destes métodos,

sua natureza estatística e implicações para interpretação biológica dos resultados, com foco no comportamento destes métodos frente à diferentes situações.

3.3.2 Análises em Genética de Populações e Filogeografia

Até o presente foram discutidas as implicações da diversidade genética e sua estruturação espacial em função do tempo para a conservação das populações, bem como a natureza das ideias biogeográficas e filogeográficas para as análises destes padrões. Contudo, a evolução dos marcadores moleculares e de outros tipos de dados disponíveis às análises desses padrões cresceu em ritmo exponencial ao longo dos últimos anos. Centenas de *softwares* e dezenas de métodos estatísticos estão hoje disponíveis para comparar as frequências alélicas entre populações, estimar as relações filogenéticas entre táxons nos mais diferentes níveis e buscar congruências entre padrões espaciais e informações ecológicas. Dada a natureza interdisciplinar da Filogeografia (veja detalhes em 3.1.2 Filogeografia e seu impacto sobre a Biogeografia), há uma tendência em incorporar análises de diferentes vertentes, a fim de compreender os padrões espaciais em níveis hierárquicos distintos. Buscar-se-á ao longo das próximas páginas categorizar essas diferentes formas estatísticas de abordar os dados para fins de comparação dos resultados.

Em suma, as abordagens aplicadas em Filogeografia poderiam ser classificadas como de natureza molecular e ecológica. Nesta primeira linha, sem dúvidas a principal, podem ser englobadas as análises derivadas do uso de marcadores moleculares, principalmente os provenientes de DNA, e naturalmente os métodos estatísticos relacionados, que por sua vez poderiam ser subdivididos em duas linhas: abordagens frequentistas e probabilísticas. Dentre estas duas abordagens, há uma vasta gama de métodos disponíveis, os quais são exaltados e criticados por diferentes motivos (MEIRMANS; HEDRICK, 2011). Os principais métodos destas abordagens serão explorados aqui de forma bastante superficial, apenas com o intuito de balizar os parâmetros avaliados mais à frente. Ótimas revisões destes métodos são encontradas em Pearse e Crandall (2004), Heller e Siegismund (2009), Jost (2008, 2009), Meirmans e Hendrick (2011), Whitlock (2011) e em Sóle-Cava e Cunha (2012).

Por fim, a linha ecológica englobaria principalmente os métodos de modelagem de nicho ecológico e distribuição potencial de espécies, os quais têm servido fundamentalmente para gerar hipóteses de distribuição espaço-temporal (ELITH et al., 2011; HICKERSON et al., 2010; SMITH; DONOGHUE, 2010) com o intuito de testar a congruência dos dados genéticos sobre as paleodistribuições (CARNAVAL et al., 2009; FRANTINE-SILVA et al., 2017; MIRANDA et al., 2015). Ao final do presente texto, serão dados detalhes sobre o

processo geral da produção e fundamentação destes modelos, visando fornecer uma ideia geral da natureza deste tipo de análise.

3.3.2.1 *Abordagens frequentistas tradicionais*

Abordagens frequentistas no presente texto se referem às abordagens tradicionais em genética de populações, as quais se baseiam na diferença entre as frequências alélicas em diferentes locos para estimar a diferenciação e diversidade genética entre populações. Neste conjunto de análises, podem ser citados os estimadores de estruturação, como a bem conhecida estatística F de Wright (WRIGHT, 1950). Do ponto de vista Filogeográfico, estes estimadores são importantes, pois auxiliam, por exemplo, na identificação de padrões de isolamento por distância, bem como suas possíveis quebras. Além disso, devido à posição de transição que a Filogeografia ocupa, estes estimadores auxiliam na compreensão de processos a nível de população, os quais são fundamentais para entender a dinâmica histórica intraespecífica (AVISE, 2009). Embora o contexto matemático destes estimadores seja extremamente relevante, as particularidades de cada um dos métodos serão omitidas para enfatizar a adequabilidade de cada método e seu uso em determinadas situações.

De certo modo, há muita controvérsia sobre o uso destes estimadores, principalmente dos de estruturação (MEIRMANS; HEDRICK, 2011). Embora numerosos artigos recentes avaliem vários parâmetros genéticos populacionais e sua validade para a análise da estrutura genética entre populações, fluxo gênico e divergência genética, pouco consenso foi alcançado até o momento (HELLER; SIEGISMUND, 2009; MEIRMANS; HEDRICK, 2011; PEARSE; CRANDALL, 2004).

O F_{ST} é provavelmente a medida mais amplamente utilizada para estimar a estruturação genética entre populações, junto com o estimador G_{ST} (NEI, 1973); porém ainda existem R_{ST} (SLATKIN, 1995), G'_{ST} (HEDRICK, 2005), Φ_{ST} (EXCOFFIER; SMOUSE; QUATTRO, 1992), Φ'_{ST} (HEDRICK, 2005), D (JOST, 2008) e até G''_{ST} (MEIRMANS; HEDRICK, 2011), além de muitos outros (MEIRMANS; HEDRICK, 2011; WHITLOCK, 2011). O F_{ST} tem uma grande vantagem em termos de familiaridade, já que existe há tempo suficiente para se ter uma ideia de quais são suas propriedades. Em linhas gerais, o F_{ST} pode ser pensado como a redução na heterozigosidade devido à estrutura da população ou (para marcadores bialélicos) a variância nas frequências alélicas entre todas as populações (HARTL; CLARK, 2007), comportando-se de maneira previsível em resposta a circunstâncias específicas (KANE, 2011; WHITLOCK, 2011). O F_{ST} foi originalmente formulado para medir a distância genética usando marcadores bialélicos (WRIGHT, 1969),

mas foi generalizado para alelos múltiplos (NEI, 1973) como G_{ST} , Φ_{ST} para sequências de DNA, e R_{ST} considerando um modelo mutacional passo-a-passo. Vale ressaltar que, embora tenha origens relativamente diferentes, os termos G_{ST} , R_{ST} e F_{ST} , bem como Φ_{ST} e F_{ST} são tradicionalmente utilizados como sinônimos, ainda que não seja esta a forma mais adequada. Devido ao seu comportamento similar frente às taxas de migração e mutação, as considerações acerca destes estimadores serão generalizadas no texto como “análogos ao F_{ST} ”.

Numerosos artigos relativamente recentes (por exemplo, HEDRICK; 2005, JOST; 2008, JOST; 2009, MEIRMANS; HEDRICK 2011; WHITLOCK, 2011) apontaram uma dificuldade na interpretação dos análogos ao F_{ST} . Com duas populações e dois alelos, os análogos de F_{ST} variam de 0 a 1, com “0” representando ausência de diferença nas frequências alélicas entre duas populações e “1” indicando que as duas populações possuem diferentes alelos fixados (HEDRICK, 2005; NEI, 1973). No entanto, com mais de dois alelos os análogos ao F_{ST} não podem atingir “1”, mesmo quando não há alelos compartilhados entre as duas populações, pois sempre haverá alguma heterozigose dentro das populações (HEDRICK, 2005; JOST, 2008). Esta não é uma falha dos estimadores, mas indica que deve-se ter cuidado ao interpretar ou comparar os análogos ao F_{ST} para diferentes tipos de marcadores (KANE, 2011; WHITLOCK, 2011).

Certamente, em muitas situações isso pode ser bastante problemático. Por exemplo, para microssatélites que possuem alta heterozigidade, o G_{ST} máximo é geralmente 0,1-0,2 (WHITLOCK, 2011). Nesses casos, as diretrizes de Wright (1978) sobre os valores de F_{ST} (G_{ST} , neste caso), não são exatamente adequadas, uma vez que as indicações de baixa diferenciação genética para valores entre 0 a 0,05; 0,05-0,15 “moderada”; e 0,15-0,25 “alta”, estão fora do campo de variação dos “análogos ao F_{ST} ”. De fato, essas atribuições seriam mais plausíveis apenas para marcadores bialélicos ou locos pouco variáveis, em outras situações não é aconselhável confiar inteiramente em regras demasiadamente simplificadas (KANE, 2011; MEIRMANS; HEDRICK, 2011).

Visando incluir a variação máxima possível dos estimadores análogos ao F_{ST} no contexto da análise, Hedrick (2005) propôs medidas “padronizadas”, o G'_{ST} e Φ'_{ST} , calculadas a partir da divisão do G_{ST}/Φ_{ST} para um dado marcador pelo G_{ST} máximo teórico baseado na heterozigidade daquele marcador (HEDRICK, 2005). Além disso, em 2008, Jost introduziu outra medida de diferenciação, o D , que mede a fração da variação alélica entre as populações. Tanto o G'_{ST} como o D de Jost terão “1” como diferenciação máxima (mesmo

com alta variação dentro das populações) e “0” quando não houver diferenciação. Além disso, ambas as estatísticas têm outras propriedades intuitivamente atraentes (Figura 3.17).

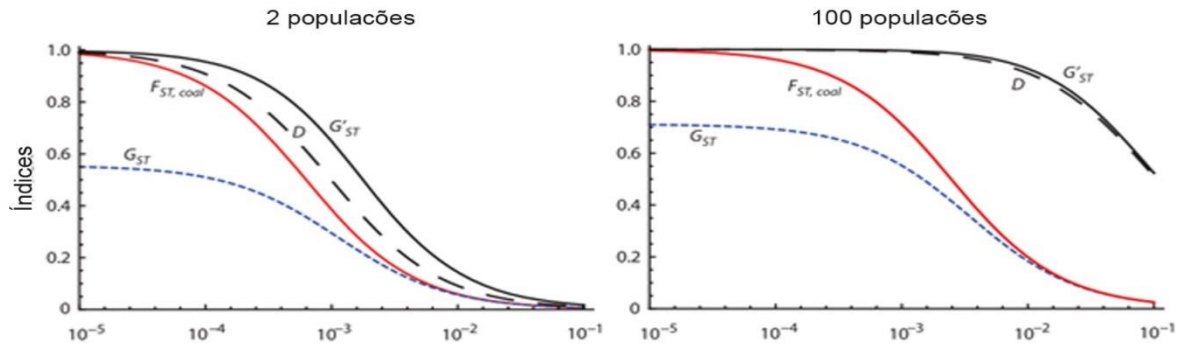


Figura 3.17 Variações do F_{ST} coalescente (Slatkins, 1985), G_{ST} (“análogos ao F_{ST} ”), D de Jost (2008), G'_{ST} (Hedrick, 2005) (eixo y) em relação à diferentes taxas de migração (eixo x), com taxa de mutação $\mu=10^{-3}$. Nestes gráficos, as populações tiveram $N = 100$ em um modelo de ilhas infinitas. O gráfico da esquerda e da direita mostram resultados para 2 e 100 populações por espécie, respectivamente. A linha para G_{ST} é muitas vezes encoberta pela linha de $F_{ST, coal}$, porque eles são quase idênticos para baixas taxas de mutação (modificado de WHITLOCK, 2011).

Ambas as medidas, no entanto, têm problemas e podem ser tão difíceis de interpretar quanto os análogos de F_{ST} . G'_{ST} e o D de Jost se comportam muito bem em alguns casos, por exemplo, quando há incidência de altas taxas de mutação e duas populações (Figura 3.17, à esquerda), acompanhando de perto a divergência “verdadeira” medida pelo F_{ST} coalescente ($F_{ST, coal}$; sensu SLATKIN, 1985) (KANE, 2011; WHITLOCK, 2011). No entanto, enquanto os análogos de F_{ST} são limitados a valores baixos quando a heterozigosidade é alta (parte direita dos gráficos), o G'_{ST} ou D são, em muitos casos, tendenciosos, raramente caindo muito abaixo de “1”, mesmo com altas taxas de migração e várias populações, 100 populações. Além disso, D é altamente afetado por taxas de mutação, por isso é difícil comparar múltiplos locos, ou mesmo comparar o mesmo loco em diferentes espécies se as taxas de mutação forem diferentes. No entanto, se a diferenciação alélica em um loco específico é o valor de interesse, o D de Jost parece o melhor estimador.

Por outro lado, os análogos de F_{ST} têm uma relação bastante direta com as taxas de migração e mutação, com padrões impulsionados pela migração quando essas são inferiores às taxas de mutação ($\mu < 10^{-3}$; detalhes em WITHLOCK, 2011). Em muitos casos, o fluxo gênico pode ser seguramente considerado alto em relação à taxa de mutação. Isso pode ser verdade no caso dos marcadores SNPs (do inglês, *Single Nucleotide Polymorphisms*), cujas taxas de mutação são tipicamente da ordem de 10^{-9} e 10^{-8} (BALLARD; WHITLOCK, 2004; HAAG-LIAUTARD et al., 2008; HAASL; PAYSEUR, 2011). Por outro lado, alguns marcadores como os microssatélites, os quais são amplamente utilizados em estudos populacionais devido à sua alta variabilidade, têm taxas de mutação que variam entre 10^{-6} e 10^{-3} (NEFF; FU; GROSS, 1999; WHITLOCK, 2011). Interessantemente, nessas altas taxas de

mutação (10^{-3} , Figura 3.17) D e G'_{ST} estão muito próximos de 1 e os análogos de F_{ST} estão próximos de zero. Este ponto levanta a óbvia questão de qual estimador seria mais adequado. A resposta, no entanto, pode ser mais complexa e muitos fatores devem ser considerados na escolha do estimador, como o tipo de marcador e sua taxa de mutação, bem como a taxa de migração estimada para o genoma/organismo em questão.

Felizmente para os euglossíneos, a partir de alguns estudos (AUGUSTO; GARÓFALO, 2011; JANZEN, 1971; LÓPEZ-URIBE et al., 2014; POKORNY et al., 2015) é possível se ter uma ideia do potencial fluxo gênico, tanto a partir de marcadores nucleares quanto mitocondriais, bem como as taxas de mutação nos marcadores frequentemente utilizados. Espera-se que em euglossíneos, machos tenham grande capacidade de dispersão enquanto fêmeas sejam filopátricas (LÓPEZ-URIBE et al., 2014), conseqüentemente espera-se altas taxas de migração para o genoma nuclear (dispersado pelos machos), com baixas taxas de migração para o genoma mitocondrial (veja 3.4.4 Diversidade Genética: o estado da arte em Euglossini). Considerando ainda que marcadores microssatélites são a primeira escolha para acessar as informações do genoma nuclear, bem como polimorfismos de sequências são a primeira opção para o genoma mitocondrial, tem-se a combinação de altas e/ou médias taxas de migração com altas taxas de mutação em análises de microssatélites, contra baixas taxas de mutação e migração para SNPs mitocondriais. Este quadro acarretaria na fatídica escolha de G'_{ST} ou D para marcadores microssatélites e análogos de F_{ST} para marcadores mitocondriais.

No entanto, vale ressaltar que o pressuposto de taxa de migração também é afetado pela conectividade de ambientes, restrições ecológicas de cada espécie e sua inerente capacidade de cruzar matrizes. Além disso, os pressupostos das simulações de Whitlock (2011) são frequentemente violados em populações naturais, tornando a escolha menos óbvia. Uma solução frequentemente recomendada para avaliar a estrutura populacional usando marcadores genéticos é calcular tanto os análogos de F_{ST} quanto D (KANE, 2011; MEIRMANS; HEDRICK, 2011; WHITLOCK, 2011). Situações onde os análogos de F_{ST} subestimam a divergência devem ter valores significativamente elevados quando analisados pela estimativa D de Jost (KANE, 2011).

Em conclusão, de acordo com Meirmans e Hedrick (2011), todas os estimadores têm suas vantagens e desvantagens, portanto, é recomendável o uso continuado dos análogos de F_{ST} em combinação com suas derivações padronizadas F'_{ST} (G'_{ST} e Φ'_{ST}) ou D de Jost. No que tange os interesses da Filogeografia, faz-se fundamental a consideração de Meirmans e Hedrick (2011), de que na maioria dos casos, o F_{ST} seria a melhor escolha entre os dois

últimos, pois é mais adequado para inferências da influência dos processos demográficos, como deriva genética e migração na estrutura da população genética.

3.3.2.2 *Abordagens probabilísticas*

Enquanto os estimadores tradicionais poderiam ser calculados usando métodos analíticos “simples”, as análises modernas em Genéticas de Populações e Filogeografia dependem fortemente da capacidade dos computadores atuais. Este fato é especialmente verdade, pois as análises baseadas apenas nas frequências dos alelos em diferentes locos, foram sistematicamente “substituídas” por estratégias matemáticas mais complexas, capazes de medir a adequabilidade de dados e hipóteses frente a um quadro probabilístico e especialmente considerando todo o conjunto de dados. Isto é especialmente verdadeiro para métodos baseados em Máxima Verossimilhança (MV) ou Inferência Bayesiana (IB) (PEARSE; CRANDALL, 2004).

Obviamente, possibilidades hipotéticas são passíveis de testes mesmo em abordagens tradicionais, porém, estes testes dependem de um quadro bastante definido de hipóteses alternativas para teste. Por exemplo, pode-se testar a significância de um índice de F_{ST} em relação a um par de populações, contudo, apenas duas hipóteses são possíveis a nulidade ou não. Os métodos mais “modernos”, por outro lado, buscam não apenas a significância dos dados, mas também seu valor mais aproximado, considerando o conjunto de dados, além da sumarização dos resultados (ALLENDORF; LUIKAR; AITKEN, 2007).

Certamente esta é a principal característica deste conjunto de análises. Vale ressaltar que devido à sumarização inerente do F_{ST} é impossível distinguir populações com níveis semelhantes de estruturação devido a diferentes fenômenos populacionais como fluxo gênico, “*bottlenecks*” ou expansão populacional recente (MEIRMANS; HEDRICK, 2011). Por outro lado, modelos baseados em Máxima Verossimilhança (MV) ou Inferência Bayesiana podem agregar variáveis suficientes para considerar resquícios evolutivos de tais fenômenos em meio às frequências alélicas (ALLENDORF; LUIKAR; AITKEN, 2007; PEARSE; CRANDALL, 2004). É importante ressaltar ainda que, embora o termo seja frequentemente utilizado para as análises tradicionais, os próprios critérios de MV poderiam ser observados como frequentistas (ALLENDORF; LUIKAR; AITKEN, 2007). Um segundo ponto que merece esclarecimentos diz respeito ao caráter “moderno” destas análises, visto que a base do modelo de MV data de 1925 (FISHER, 1925; *Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd*) e o Teorema de Bayes é ainda anterior à isso, 1763 (BAYES, 1763). Embora apenas atualmente tenha sido possível enquadrar estas análises em um panorama condizente com os

dados gerados pelas técnicas moleculares modernas, é necessário reconhecer que estes métodos foram descritos há vários anos.

Ainda que a ideia por trás destas abordagens seja aparentemente simples, suas explicações e aplicações em Genética de Populações, Filogenia e Filogeografia permanecem, até certo ponto, um tanto obscuras. A maior parte destes métodos é executada por trás da “*black box*” dos *softwares* e pouca ênfase é dada em como estes modelos trabalham o conjunto de dados. Obviamente, discorrer detalhadamente sobre estes métodos e como eles são individualmente aplicados em cada *framework* acarretaria em uma discussão demasiadamente longa, porém, cabem algumas pontuações e distinções básicas.

Os chamados métodos de *Máxima Verossimilhança* aplicam uma ideia de maximização de um determinado parâmetro em relação aos dados (ALLENDORF; LUIKAR; AITKEN, 2007). Em outras palavras, pode-se definir o conceito de MV como a probabilidade de se observar os dados segundo algum valor de parâmetro (por exemplo, $Nm = 50$), sob um determinado modelo estatístico (por exemplo, modelo de migração da ilha) (PEARSE; CRANDALL, 2004). Genericamente a MV é expressa como:

$$\mathcal{L}(\theta|x) = \ln \mathcal{L}(\theta|x)$$

ou

$$L(H|D) = \ln L(H|D)$$

onde \mathcal{L} ou L se referem ao logaritmo natural da probabilidade de um determinado modelo θ ou hipótese H , dado um conjunto de dados x ou D . Vale ressaltar que, neste caso, um modelo pode ser considerado um modelo de ilhas para migração, ou o modelo de equilíbrio de Hardy-Weinberg, ou um modelo multiparamétrico de substituição nucleotídica, ou ainda a simples relação estatística de lançamento de uma moeda. Independentemente deste pressuposto, dentro de um conjunto finito de θ ou H , sob um comportamento matemático previsível, a máxima verossimilhança buscará o maior valor dentro de um conjunto específico de valores que maximize a observação de acordo com um modelo. Em outras palavras, em MV a hipótese de um valor é variável, mas os dados continuam sempre constantes (ALLENDORF; LUIKAR; AITKEN, 2007).

Uma vantagem da análise de verossimilhança é o fato de ser baseada em modelos e permite fácil comparação de diferentes modelos (mesmo modelos complexos), melhorando assim a inferência sobre processos complexos (por exemplo, diferentes padrões de dispersão, modelos de mutação, tamanho populacional estável versus declinante) que podem explicar os dados (ALLENDORF; LUIKAR; AITKEN, 2007). Porém, talvez a principal vantagem das

abordagens de MV seja a utilização de “todos os dados” em sua forma bruta, e não uma estatística sumária (por exemplo, H_e ou F_{ST}). O contraponto é o alto custo computacional, pois quanto maior o conjunto de dados, número de modelos e complexidade destes, maior será o tempo computacional necessário para explorar todas as possibilidades possíveis, mesmo que o conjunto de possibilidades seja finito (BEERLI, 2006). Este fato, torna a MV muitas vezes inviável ou impossível de ser aplicada. Outra desvantagem é que a acurácia da MV depende do conjunto de possibilidades pré-determinadas, uma vez que o caráter exaustivo das análises é deixado de lado para torná-lo exequível, a aproximação dos resultados pelo campo amostral pode ser comprometida (BEERLI, 2006)

É neste ponto que a Inferência Bayesiana (IB) assume um papel fundamental nas análises genéticas atuais. Ao contrário da MV, que avalia a probabilidade de um determinado conjunto de dados em função de diferentes valores de um parâmetro, a IB busca a probabilidade do parâmetro em função dos dados. O teorema de Bayes pode ser definido da seguinte forma:

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)}$$

onde, a probabilidade anterior (*a priori*) da hipótese “ $P(H)$ ” é combinada com a distribuição de amostragem “ $P(D|H)$ ” e condicionada aos dados conhecidos “ $P(D)$ ” para produzir uma probabilidade posterior (*a posteriori*) da hipótese em função os dados. Aqui a hipótese engloba a mudança de valores de parâmetros em função de um modelo.

Um exemplo ilustrativo seria a comparação entre duas sequências com trinta e dois pares de bases (pb), dos quais dois são variáveis, para a qual objetiva-se estimar a distância genética entre estas sequências. Caso se aplicasse um modelo de substituição nucleotídica (*distância p*, ou qualquer outro) que acreditasse explicar a diferença entre as sequências, ter-se-ia uma função de MV, na qual um conjunto finito e pré-determinado de diferentes valores de distância seriam calculados com base no modelo de “*distância p*” gerando diferentes números de nucleotídeos diferentes. Poderia então ser dito que o valor mais verossímil de distância é aquele que resulta em dois nucleotídeos modificados.

Por outro lado, no teorema de Bayes, na mesma situação ter-se-ia a produção de diferentes “*distâncias p*” condicionadas às alterações nos valores de sítios variáveis, avaliando-se qual seria o valor de “*distância p*” mais provável. Obviamente, em um exemplo simplificado a MV mostra-se superior à inferência Bayesiana devido ao caráter exaustivo e o universo finito da análise. Contudo, em cenários mais complexos, com um maior número de parâmetros a IB é capaz de acomodar melhor as incertezas ao longo do modelo modificado.

Os métodos Bayesianos trazem algumas vantagens metodológicas. Diferente da MV, a interpretação da probabilidade Bayesiana é muito mais simples e direta, podendo ser interpretada diretamente como a probabilidade daquele parâmetro em relação aos dados (ALLENBORG; LUIKAR; AITKEN, 2007). Adicionalmente, talvez a principal vantagem da abordagem Bayesiana seja a capacidade de fatorar dados ou informações anteriores ao estimar a probabilidade posterior de que uma hipótese esteja correta (DRUMMOND; BOUCKAERT, 2015).

No entanto, a IB pode sofrer os mesmos problemas de execução que a MV em termos de tempo computacional, caso tente-se avaliar todo o espaço amostral possível em relação às modificações em modelos complexos. Neste sentido, algoritmos de escolha e sucessão, como as Cadeias de Markov e Monte-Carlo (MCMC), são extremamente úteis para associação com modelos Bayesianos. As MCMC se baseiam na variação das probabilidades *a posteriori* encontradas na inferência Bayesiana para avaliar mudanças de parâmetros no modelo (BEERLI, 2006). A grosso modo, alterações de parâmetros que resultem em maiores probabilidades *a posteriori* são assumidas como hipóteses *a priori* no próximo passo da cadeia, de modo que os parâmetros do passo anterior não mais afetam a cadeia, e esses parâmetros podem variar de forma randômica ao longo do espaço amostral (DRUMMOND; BOUCKAERT, 2015; NYLANDER et al., 2004).

O uso de MCMC tem ajudado métodos Bayesianos a abordar modelos extremamente complexos, mas a principal crítica é que essa estratégia IB+MCMC pode ser fortemente influenciada pelo início da cadeia ou pelas informações *a priori* e, portanto, menos objetivas do que as abordagens clássicas (BEERLI, 2006). Em outras palavras, duas cadeias com inícios distintos poderiam resultar em probabilidades *a posteriori* diferentes para um mesmo conjunto de dados. Uma solução frequente é executar múltiplas cadeias suficientemente longas e buscar por sua mistura e conversão (DRUMMOND; BOUCKAERT, 2015).

Análises Filogeográficas tem sido cada vez mais influenciadas por métodos que fazem uso de MV e IB (AVISE, 2009; HICKERSON et al., 2010). Métodos de inferências Bayesianas e Máxima Verossimilhança para reconstrução filogenética (DRUMMOND; RAMBAUT, 2007; REE; SMITH, 2008; STAMATAKIS, 2006), bem como para estimativa de parâmetros populacionais (BEERLI, 2006; PRITCHARD; STEPHENS; DONNELLY, 2000), tendem a assumir cada vez mais um papel protagonista nas análises Filogeográficas em detrimento das Análises de Clados Hierarquizados (NCA, *Nested Clade Analysis*).

3.3.2.3 *Análise de Clados Hierarquizados NCA*

Embora a abordagem NCA seja cada vez menos utilizada devido ao seu uso limitado da estatística, sem dúvidas, este é o primeiro passo de qualquer análise Filogeográfica, compondo um dos principais e mais utilizados métodos disponíveis para análise de dados filogeográficos (AVISE, 2009; CUNHA; SOLÉ-CAVA, 2012). De fato, a abordagem NCA consiste em uma série de abordagens que visa analisar qualitativamente os dados das genealogias gênicas, buscando um padrão adequado condizente com o espaço geográfico (AVISE, 2000).

A NCA tem alguns passos metodológicos sistêmicos, que se iniciam com a construção de uma rede de relacionamento haplotípico por inúmeros métodos (inicialmente Parcimônia – TCS, mas atualmente incluindo Median-Joining, Minimum Spawning Network, dentre outros); em seguida busca-se identificar padrões espaciais coerentes sobre essas redes; subsequentemente testa-se sistematicamente hipóteses baseadas nestas unidades de estruturação, como padrões e quebras de Isolamento por Distância, diferenças de diversidade e de estruturação genética (CUNHA; SOLÉ-CAVA, 2012).

A forma mais tradicional de se proceder uma NCA é a hierarquização de clados a partir de uma unidade mínima, no caso o haplótipo, até a maior unidade possível onde existem poucos grupos de haplótipos de igual tamanho separados por ao menos um passo mutacional (POSADA; CRANDALL; TEMPLETON, 2006). Contudo, cada vez mais métodos probabilísticos são utilizados para inferir estruturas populacionais, de modo que a hierarquização empírica dos clados tem perdido espaço neste sentido (CORANDER et al., 2008; PRITCHARD; STEPHENS; DONNELLY, 2000).

Métodos complementares baseados em análises populacionais são também empregados no contexto da NCA, buscando estimar parâmetros populacionais com base nos clados previamente determinados, como Análise de Variância Molecular (AMOVA; EXCOFFIER e SMOUSE, 1992), Análise Espacial de Variância Molecular (SAMOVA; DUPANLOOP, 2002), bem como métodos demográficos de frequência de distribuição de diferenças entre pares de sequências (distribuição *mismatch*), análises de testes de Neutralidade, como o teste *D* de Tajima e *F_s* de Fu.

Embora sejam o centro histórico da análise Filogeográfica, este quadro tende a mudar cada vez mais para os métodos probabilísticos baseados em MV e IB para a estimativa de parâmetros e modelos filogeográficos. Contudo, é provável que as abordagens de NCA continuem sendo o ponto de partida para a maior parte das inferências filogeográficas.

3.3.3 Modelagem de Nicho Ecológico

A inclusão de dados ecológicos em modelos históricos tem se tornado cada vez mais uma tendência no âmbito da Filogeografia, aumentando o papel desta disciplina como ponte entre fatores ecológicos e processos históricos (CRISCI et al., 2006). A possibilidade de desenvolver modelos matemáticos de distribuição/adequabilidade espacial e temporal trazem uma possibilidade única do ponto de vista histórico e filogeográfico, pois é possível observar variações em escala de tempo microevolutivo, extrapolando-os de forma consideravelmente precisa em escalas de tempo macroevolutivo (KOZAK; GRAHAM; WIENS, 2008; PHILLIPS; ANDERSON; SCHAPIRE, 2006), os quais podem ser testados com base em informações geológicas e genéticas (ELITH et al., 2011).

Modelos de distribuição de espécies (SDM, *species distribution modeling*) ou mais precisamente Modelos de Nicho Ecológico (ENM; *ecological niche modeling*) (PETERSON; SOBERÓN, 2012) estimam a relação entre registros de espécies em locais e as características ambientais e/ou espaciais desses locais, em tempos atuais ou pretéritos (Figura 3.18). Estes métodos são amplamente utilizados para muitos propósitos em biogeografia, biologia da conservação e ecologia (ELITH et al., 2011; KOZAK; GRAHAM; WIENS, 2008).

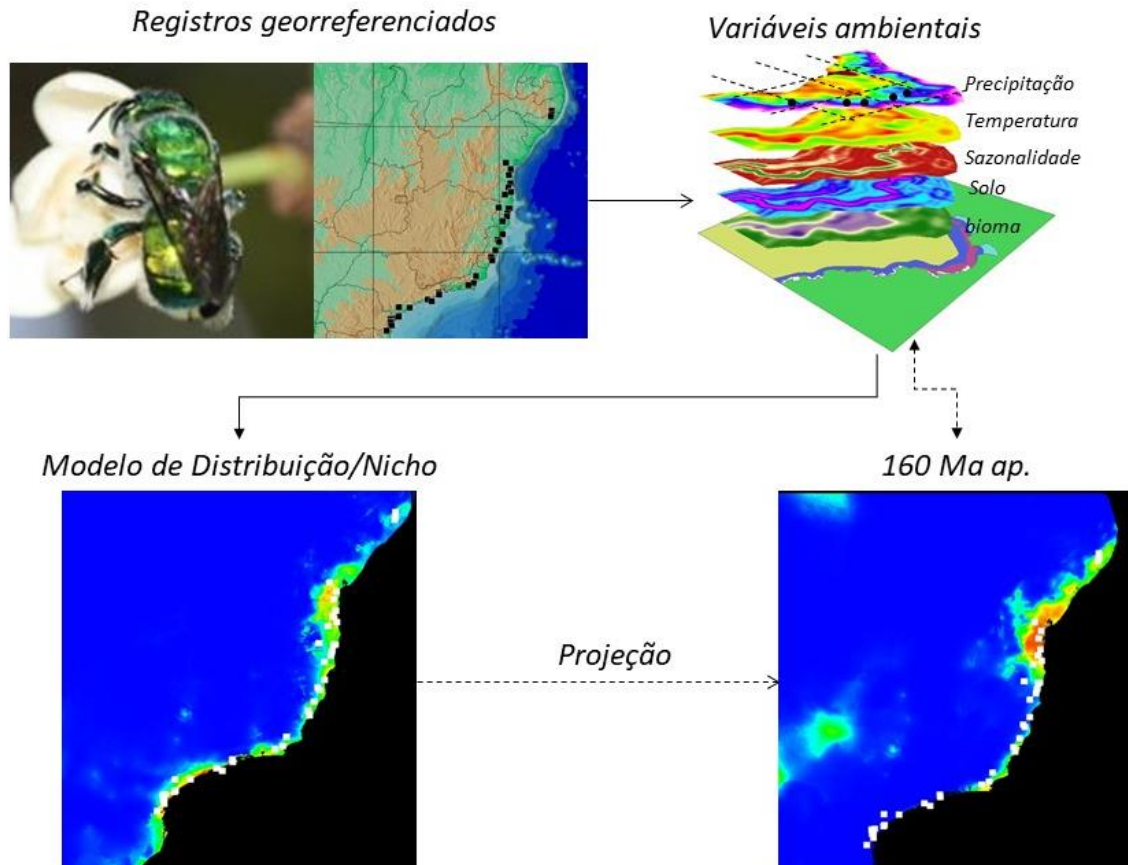


Figura 3.18 Esquema da estratégia básica de modelagem de nicho ecológico. Registros de ocorrência georreferenciados são utilizados como pontos para coleta de informações de diferentes variáveis igualmente georreferenciadas. A partir destas informações, diferentes algoritmos calculam distribuições relativas e/ou as

projetam para diferentes intervalos cronológicos em um sistema retroalimentado pelas variáveis disponíveis em diferentes espaços temporais.

Nas últimas duas décadas, houve um grande desenvolvimento em termos de modelagem de distribuição de espécies, disponibilizando vários métodos voltados a este tipo de análise: métodos de distância (*profile methods*) como BIOCLIM e distância de Mahalanobs; métodos baseados em regressão, como modelos lineares generalizados (GLM) e modelos aditivos (GAM); e os sofisticados métodos de aprendizagem de máquina (*machine learning*) como algoritmos genéticos (GARP) e método de Máxima Entropia (MAXENT) (ELITH et al., 2011; PHILLIPS; ANDERSON; SCHAPIRE, 2006; VARELA; LOBO; HORTAL, 2011). De modo geral, estes métodos podem utilizar dados de presença, ausência e até mesmo abundância, para relacionar condições ambientais georreferenciadas específicas das coordenadas onde os registros foram anotados (ELITH et al., 2011).

Dentre os métodos disponíveis, o método de Máxima Entropia, compilado no software MaxEnt (PHILLIPS et al., 2006), tem se mostrado o mais robusto frente a diferentes tipos de dados e variáveis ambientais, bem como frente à predição de adequabilidade ambiental em projeções futuras e pretéritas (DUQUE-LAZO et al., 2016; FOURCADE et al., 2014; VAN GILS et al., 2012). O MaxEnt possui ainda a vantagem de trabalhar bem com apenas dados de presença de espécies, não necessitando obrigatoriamente de dados de ausência conhecida, o que facilita o trabalho quando não se conhece toda a extensão da distribuição da espécie (PHILLIPS; ANDERSON; SCHAPIRE, 2006). Devido às suas vantagens metodológicas e facilidade, será dada ênfase a essa metodologia, em detrimento dos demais métodos de modelagem, visando esclarecer brevemente os princípios do método além do termo *machine learn*.

3.3.3.1 Princípio por trás do MaxEnt

O MaxEnt (PHILLIPS; ANDERSON; SCHAPIRE, 2006) assume um lista de presença de espécies como referência, bem como um conjunto de preditores ambientais (por exemplo, precipitação, temperatura...) em uma paisagem definida pelo usuário que é dividida em células de grade. A partir desse cenário, o programa extrai uma amostra aleatória de pontos a partir do “*background*”, que contrastam com os locais de presença. O MaxEnt introduz um conceito de “*pseudoausência*”, considerando que os pontos aleatórios correspondem à ausência verdadeira, ou seja, locais de baixa adequabilidade ao longo do espaço amostrado (ELITH et al., 2011).

Originalmente, o MaxEnt foi empregado para estimar a densidade de presenças ao longo da paisagem (PHILLIPS; ANDERSON; SCHAPIRE, 2006). Neste sentido, a estimativa de densidade supõe implicitamente que os indivíduos foram amostrados aleatoriamente na paisagem, ou seja, as amostras ocorrem em proporção à densidade populacional. Quando o tamanho total da população é conhecido, esses modelos preveem a taxa de ocorrência em uma célula, definida como o número esperado de indivíduos na célula (ELITH et al., 2011). No entanto, o tamanho da população real é frequentemente desconhecido, de modo que apenas comparações relativas entre essas taxas são significativas, resultando em uma Taxa de Ocorrência Relativa (ROR, *Relative Occurrence Rate*), a saída bruta do MaxEnt (MEROW; SMITH; SILANDER, 2013). Em outras palavras, o ROR é a probabilidade relativa de que uma célula esteja contida em um conjunto de amostras de presença (PHILLIPS et al., 2017).

O MaxEnt pode ser usado para prever a probabilidade de presença utilizando uma transformação do ROR, chamada “*output logístico*” (PHILLIPS et al., 2017), que se baseia em fortes suposições que tem sido criticadas (ELITH et al., 2011). Pode-se assumir que os dados de presença são uma amostra aleatória de indivíduos (por vezes, uma suposição questionável) e prever RORs (uma interpretação razoável da saída de MaxEnt) ou podem assumir que os dados representam uma amostra aleatória de espaço (suposição possível se o viés de amostragem não for um problema) e prever a probabilidade de presença (uma interpretação duvidosa da saída do MaxEnt) (MEROW; SMITH; SILANDER, 2013). Em casos em que é possível assumir a homogeneidade da amostragem e a interpretação probabilística do resultado do modelo, é possível simplesmente interpretar as previsões do MaxEnt como índices de adequabilidade de habitat (ELITH et al., 2011; PHILLIPS; ANDERSON; SCHAPIRE, 2006), o que pode ser útil para análises exploratórias qualitativas (KOZAK; GRAHAM; WIENS, 2008).

Embora as especificidades do algoritmo MaxEnt sejam ligeiramente mais complexas, serão feitas algumas observações gerais (mais especificamente explicadas em MEROW; SMITH; SILANDER, 2013) que se relacionam a qualquer uma das interpretações da saída do modelo, com foco especial nas RORs, uma vez que essas são as grandezas fundamentais previstas pelos modelos de entropia máxima. O modelo básico para definição das RORs [$P^*(z(x_i))$] a partir do MaxEnt pode ser definido da seguinte maneira:

$$P^*(z(x_i)) = \frac{\exp(z(x_i)\lambda)}{\sum_i \exp(z(x_i)\lambda)}$$

onde z é um vetor resumido de J variáveis de uma determinada célula x_i , e λ é um vetor do coeficiente de regressão de z (ELITH et al., 2011). Note que as RORs são relativizadas pela soma das unidades em todas as células de grade no estudo, criando um efeito de normalização (ELITH et al., 2011). A normalização garante que as taxas de ocorrência sejam, de fato, taxas de ocorrência relativas (PHILLIPS et al., 2017).

Uma vez que as previsões são feitas para cada célula na paisagem, aplica-se o princípio da *máxima entropia* (PHILLIPS, 2004) que postula que os modelos devem ser escolhidos de forma tão semelhante quanto possível às expectativas anteriores, ao mesmo tempo que são consistentes com os dados, chamada distribuição *a priori*, $Q(x_i)$. A entropia relativa (ou divergência de Kullback-Leibler) mede a semelhança da predição com a distribuição *à priori* por:

$$\sum_{i=1}^N P^*(z(x_i)) * \log \frac{P^*(z(x_i))}{Q(x_i)}$$

normalmente, a distribuição *à priori* é uma distribuição uniforme no espaço geográfico, significando que todas as células são, em tese, igualmente prováveis de conter um indivíduo (embora outras premissas sejam possíveis) (veja MEROW; SMITH; SILANDER, 2013). Esta suposição corresponde a $Q(x_i) = 1/N$, onde N é o número de células de grade envolvidas na análise (PHILLIPS, 2017). A partir destes resultados o MaxEnt introduz funções de ganho (uma função de Máxima Verossimilhança Penalizada) e funções de regularização (*thresholds*) para aumentar o contraste entre áreas adequáveis ou não (ELITH et al., 2011). Finalmente, o programa gera uma matriz de tamanho correspondente ao espaço geográfico definido pelo usuário onde cada termo da matriz contém um valor de ROR (transformado logisticamente ou não à escolha do usuário) em função das variáveis (ou preditores) de entrada.

3.3.3.2 Implicações à modelagem

As propriedades anteriormente descritas para o modelo MaxEnt possuem várias implicações a respeito de seu uso. Análises analíticas de cada uma das propriedades do modelo e suas considerações inerentes podem ser encontradas em Elith et al. (2011) e Merow et al. (2013). Contudo, alguns breves pontos serão salientados.

Há uma intrínseca dependência de amostragem não enviesada no modelo MaxEnt (como em todos os métodos ENM), logo, um conjunto abrangente de registros de presença e limpos de duplicatas e erros é crítico para obter modelos acurados (ELITH et al. 2011). Ainda

assim, existem métodos com recursos disponíveis para lidar com dados enviesados (FOURCADE et al., 2014; MEROW; SMITH; SILANDER, 2013; PHILLIPS et al., 2017). As principais alternativas para acomodar possíveis vieses de amostragem são fornecer dados de *background* com vieses semelhantes aos dos dados de presença (por exemplo, usando sítios amostrados para outras espécies no mesmo grupo biológico) ou usar uma lista de ocorrências de áreas enviesadas (veja um exemplo em ELITH et al., 2011).

Uma segunda consideração importante diz respeito à projeção das variáveis. Essas devem ser projetadas em mesma escala, qualquer região cobrindo uma faixa não trivial em latitude (digamos, mais de 200 km, especialmente longe do equador) terá células de grade de área variáveis. Por exemplo, em áreas de ampla distribuição latitudinal, as projeções geodésicas podem apresentar distorções de até 1,3 vezes arcmin entre áreas equatoriais e polares (ELITH et al., 2011). Como o MaxEnt amostra células implicitamente assumindo que as mesmas tenham áreas iguais, comparações entre células com áreas diferentes poderiam causar distorções de estimativas reais de adequabilidade (MEROW; SMITH; SILANDER, 2013). A solução é obviamente, reprojeter as variáveis para uma projeção UTM (como SIRGAS 2000, WGS84 ou SAS69), criar uma grade de viés, ou criar sua própria amostragem de *background* com pesos de amostragem apropriados (veja exemplos em ELITH, 2011).

Outro aspecto interessante é que os resultados provenientes do MaxEnt são fortemente afetados pela paisagem (região) utilizada (DUQUE-LAZO et al., 2016). Conceitualmente, a paisagem deve incluir toda a gama ambiental na qual as espécies estariam potencialmente presentes e excluir áreas que não foram pesquisadas ou que não possuem possibilidade de ocorrência (ELITH, 2011). Além disso, a região incluída como *background* deve ser cuidadosamente selecionada, visando incluir as áreas que seguramente não há registros da espécie ao longo da região selecionada para o estudo.

Por fim, a comparação de modelos entre diferentes espécies deve ser feita com cautela, principalmente quanto ao uso dos resultados logísticos, visto que probabilidade de presença é definida apenas em relação a um dado nível de esforço de amostragem, que como padrão é assumido como aquele que resulta em 50% de chance de observar as espécies em áreas adequadas (MEROW; SMITH; SILANDER, 2013). Isso apresenta alguns desafios para as comparações entre áreas adequáveis para diferentes espécies, mas isto é um resultado direto do uso de dados exclusivamente de presença e não é uma particularidade do MaxEnt.

Do ponto de vista dos euglossíneos, muitos destes problemas são contornáveis já que sistemáticos levantamentos têm sido realizados ao longo dos últimos 50 anos, e tem-se atualmente uma boa ideia da distribuição da maior parte das espécies. Deste ponto de vista, o

uso de métodos de modelagem de nicho, e especificamente o MaxEnt, possui grande potencial para diversas análises, incluindo análises complementares à Filogeografia, Biogeografia e Ecologia e Conservação.

3.4 AS ABELHAS DAS ORQUÍDEAS

Muito devido ao seu importante papel ecológico como polinizadores, mas também por sua evidente relação com a diversificação das angiospermas, as abelhas constituem um dos mais extraordinários grupos dentre os himenópteros (MICHENER, 2000; FENSTER et al., 2004; DANFORTH et al., 2006; DÖTTERL; VEREECKEN, 2010). Em números, abelhas (Apoidea: Anthophila) representam um grupo de cerca de 20 mil espécies distribuídas em sete famílias (Figura 3.19), dentre as quais Apidae se destaca como a maior família com cerca de 5,700 espécies (MICHENER, 2000; ASCHER; PICKERING, 2017).

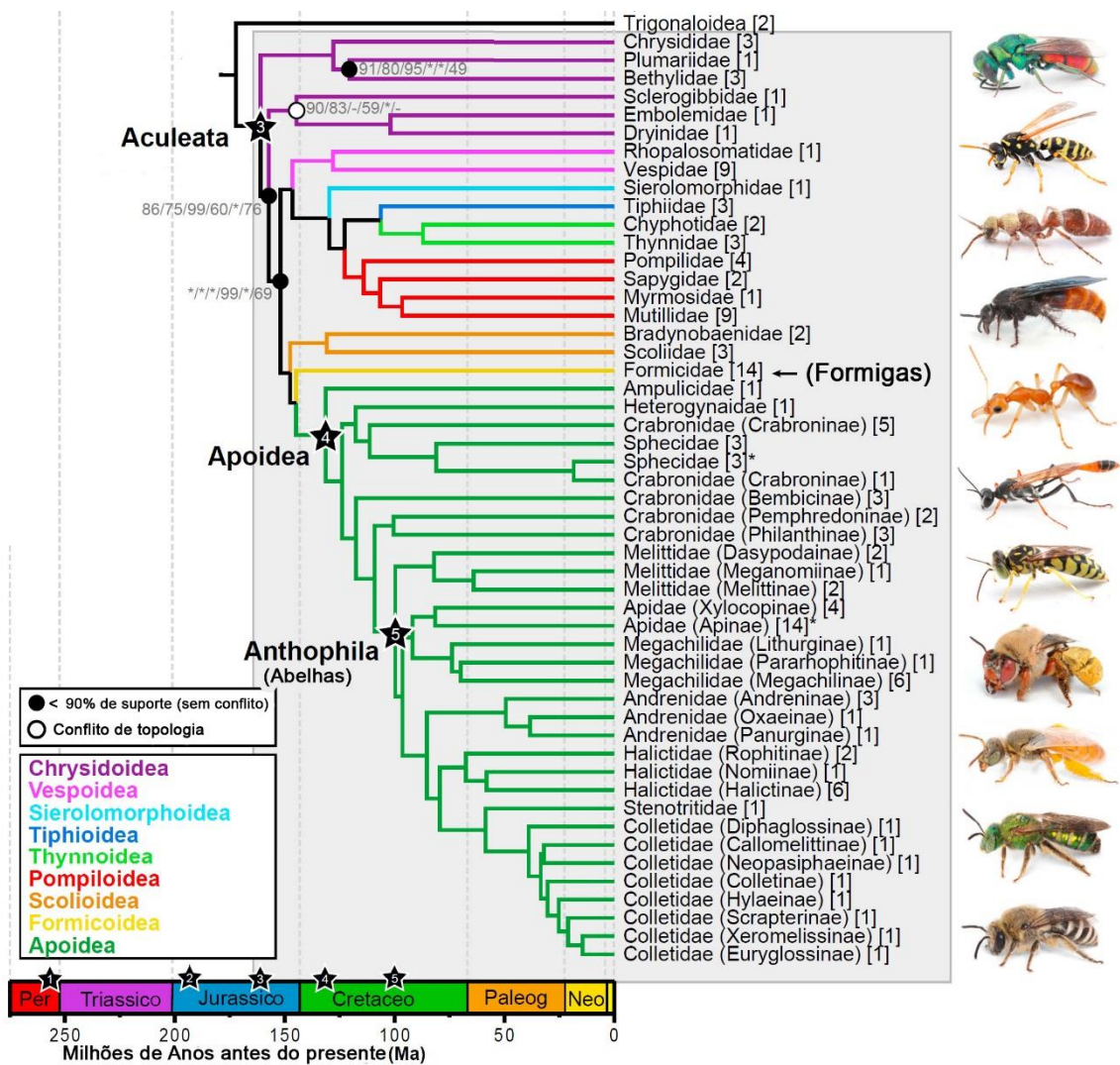


Figura 3.19 Evolução de Apoidea (Hymenoptera: Aculeata). Topologia parcial da filogenia de Hymenoptera: Apocrita, apresentando a posição das sete famílias de abelhas (Anthophila) resultante da análise de 854 locos (203.095 Kb) com 38 pontos de calibração, modificado de Branstetter et al. (2017). Destaque para a posição de Formicidae como grupo irmão de Apoidea a qual sugere uma ancestralidade comum para comportamento social.

Mais do que um grande número de espécies, as abelhas trazem uma variedade de histórias biológicas e nichos. Habitando de zonas temperadas aos trópicos, nidificando no solo, em madeira, ou a partir de resina, vivendo de forma solitária ou altamente social, em praticamente todos os lugares as abelhas apresentam-se como intrigantes modelos de evolução e adaptação (DANFORTH et al., 2006; HEDTKE; PATINY; DANFORTH, 2013; SERRANO-CINCA; FUERTES-CALLÉN; MAR-MOLINERO, 2005).

As abelhas da tribo Euglossini (*Eu* gr. verdadeiro; *glossa* gr. língua) apresentam um conjunto de características ímpares que as destacam das demais. Exemplo dessa singularidade é a alta especialização dos machos destas abelhas, os quais possuem modificações estruturais nas pernas proterópicas, médias e posteriores com estruturas e tecidos especializados para explorar e armazenar compostos aromáticos utilizados para o comportamento de corte (ROUBIK; HANSON, 2004; DRESSLER, 1982; ELTZ; ROUBIK; LUNAU, 2005; ELTZ; SAGER; LUNAU, 2005). Este tipo de comportamento é bastante singular e coloca os euglossíneos dentre as poucas abelhas neotropicais com íntima relação com angiospermas sem especificidade por pólen (RÊBELO, 2001).

A coloração metálica e iridescente do integumento de diversos membros desta tribo é outra característica que chama a atenção para o grupo (Figura 3.20, A-C), embora várias espécies possam apresentar integumento de cor escura e altamente pubescente (Figura 3.20-D). Do mesmo modo, o comprimento da língua, a qual pode por vezes ultrapassar o tamanho do próprio corpo é também uma característica relevante de Euglossini. Esta característica é tão ímpar dentre os membros do grupo que deu origem ao nome da tribo (Figura 3.20-E). Dentro da tribo, uma grande variação de tamanho dentre as diferentes espécies é também notável, variando de poucos milímetros, a exemplo de *Euglossa crassipunctata*, que tem aproximadamente 8 mm, até cerca de 20 mm, como acontece com diversas espécies do gênero *Eulaema* (ROUBIK; HANSON, 2004; DRESSLER, 1982).

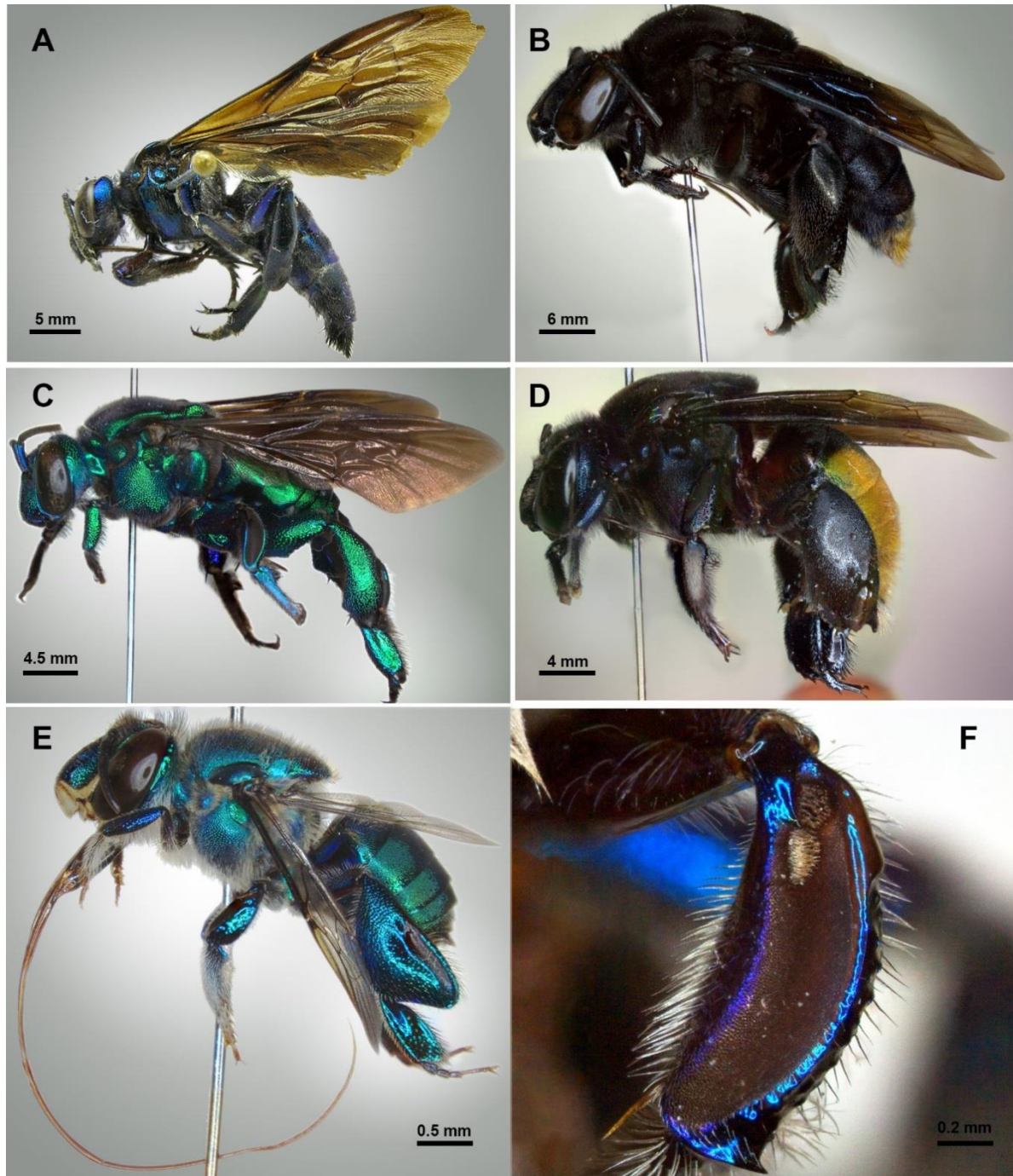


Figura 3.20 Espécimes representativos de cada um dos cinco gêneros da tribo Euglossini: A) *Aglae caerulea*, B) *Eulaema nigrita*, C) *Exaerete smaragdina*, D) *Eufriesea dentilabris*, E) *Euglossa stellfeldi*. F) Exemplo da área aveludada da metatíbia de um macho de *Euglossa stellfeldi* com destaque para os tufos ou coxins mesotibiais (lado esquerdo), frequentemente utilizados para diagnose de espécies. Destaque para a abertura da abertura da metatíbia através da qual as abelhas masculinas depositam e armazenam fragrâncias químicas.

O estudo da tribo continua em progressão desde a década de 1960 quando descobriu-se que os machos destas abelhas eram atraídos por compostos aromáticos (VOGEL, 1963, 1966; DODSON et al., 1969). Até o momento da realização desta revisão, o motor de indexação do GoogleScholar indexava incríveis 7530 resultados para produções bibliográficas

incluindo ao menos um dos termos: “Euglossini”, “Orchid bees”, “euglossine” ou “euglossina”. Estes resultados crescem exponencialmente à mediada que novas espécies foram descritas, a qual aumentou significativamente após a década de 1960 (Figura 3.21).

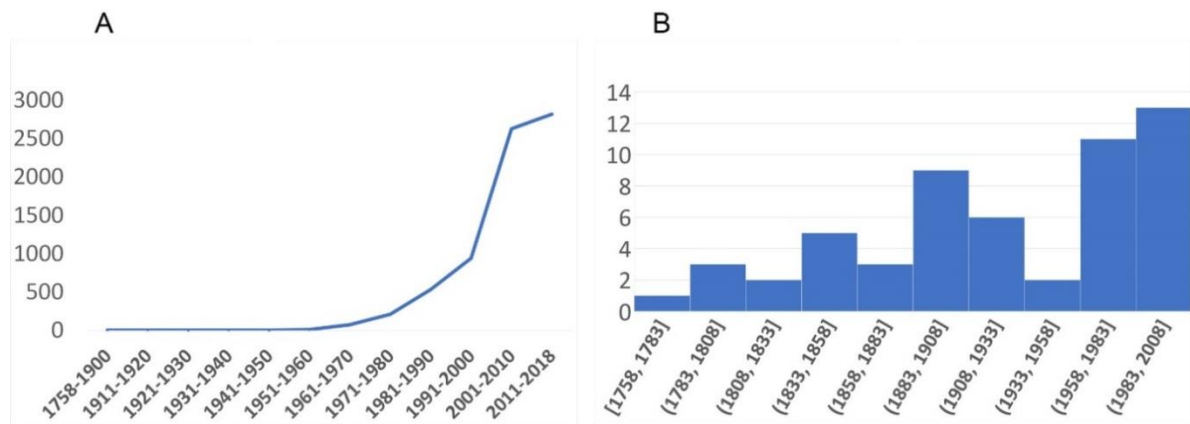


Figura 3.21 Evolução dos trabalhos sobre abelhas das orquídeas ao longo do tempo. A – Números de resultados correspondendo às publicações cadastradas na plataforma Google Scholar até janeiro de 2018 contendo qualquer uma das palavras-chave: “Euglossini, *Orchid bees*, euglossine, euglossina”; B – Número de espécies válidas descritas com base em seus respectivos anos de descrição com base em Moure et al. (2012).

Atualmente a tribo Euglossini compreende cerca de 240 espécies distribuídas em cinco gêneros (*Aglae*, *Eufriesea*, *Euglossa*, *Eulaema* e *Exaerete*²; Figura 2), com alguma variação a depender da aceitação de determinadas bibliografias sobre o *status* taxonômico de alguns grupos de espécies (MOURE; MELO; FARIA JR, 2012; NEMÉSIO, 2009).

3.4.1 A origem das abelhas das orquídeas

A determinação da origem dos euglossíneos passa fundamentalmente pela descrição precisa das relações filogenéticas do grupo frente às demais abelhas corbiculadas, tendo ainda fortes implicações sobre a compreensão da evolução do comportamento social e aspectos biogeográficos. Infelizmente, a posição filogenética dos euglossíneos em relação aos grupos mais próximos está longe de ser um consenso e vêm sofrendo várias modificações ao longo do tempo (MICHENER, 1944, 1990, 2000; CARDINAL; PACKER, 2007; KAWAKITA et al., 2008; CARDINAL; STRAKA; DANFORTH, 2010; DANFORTH et al., 2013; ROMIGUIER et al., 2016; BOSSERT et al., 2017).

Há de fato consenso sobre a monofilia da tribo Euglossini com as demais abelhas corbiculadas. A presença dessa estrutura (a corbícula) é também a principal sinapomorfia que reúne os grupos de abelhas com variável nível de socialidade, as tribos Euglossini, Apini,

² Devido à ambiguidade da abreviação dos nomes dos gêneros da tribo *Euglossini*, usar-se-á os acrônimos *El.*, *Ef.*, *Eg.*, *Ex.*, para se referir respectivamente à *Eulaema*, *Eufriesia*, *Euglossa* e *Exaerete* onde o uso do padrão internacional não for suficientemente explícito.

Meliponini e Bombini (MICHENER, 1990), monofilia que vem sendo corroborada por dados moleculares (BOSSERT et al., 2017; CARDINAL; STRAKA; DANFORTH, 2010; KAWAKITA et al., 2008; ROMIGUIER et al., 2016)

Por outro lado, os resultados em relação à filogenia interna dos corbiculados não são igualmente concisos. Entre árvores de genes, caracteres comportamentais, estruturas de ninhos e caracteres morfológicos, Cardinal e Packer (2007) identificaram 16 diferentes hipóteses de filogenia, dentre as quais a topologia proposta por Michener (1944): (Euglossini + (Bombini (Apini + Meliponini))), figura entre as mais recuperadas, sendo inclusive corroborada pelo consenso de Cardinal e Packer (2007). Adicionalmente, dados moleculares têm identificado duas outras hipóteses alternativas com alto nível de sustentação. A primeira delas propõe o agrupamento (Bombini + Meliponini) + (Euglossini + Apini) (daqui em diante EA), como a hipótese mais provável (KAWAKITA et al., 2008; CARDINAL; STRAKA; DANFORTH, 2010; DANFORTH et al., 2013). A segunda delas, a partir de dados filogenômicos corrigindo um desvio de ruído filogenético das regiões GC do genoma (ROMIGUIER et al., 2016; BOSSERT et al., 2017), corroboram a relação Bombini + Meliponini, como todas as demais hipóteses moleculares, porém coloca Apini como grupo irmão de Bombini + Meliponini e Euglossini como grupo externo (daqui em diante E+ABM): (Euglossini + (Apini + (Bombini + Meliponini))) (ROMIGUIER et al., 2016; BOSSERT et al., 2017).

Cabem aqui algumas considerações quanto às duas hipóteses moleculares. Se por um lado a hipótese EA é baseada em apenas alguns “poucos” locos (respectivamente, sete e doze locos) (KAWAKITA et al., 2008; CARDINAL; STRAKA; DANFORTH, 2010), a mesma possui uma quantidade significativamente maior de táxons para a reconstrução. Ainda assim, a hipótese E+ABM de Romiguier et al. (2016), tem a vantagem de ser apoiada em centenas de locos espalhados por todo o genoma. Há também o fato de que a hipótese E+ABM torna-se mais parcimoniosa à evolução do comportamento social partindo de Euglossini e aumentando em complexidade em direção a Apini e Meliponini.

A despeito das importantes implicações sobre a evolução do comportamento social nas abelhas trazidas por estas duas hipóteses (veja a sessão: Aspectos comportamentais), por hora, será direcionada a atenção para os aspectos geológicos decorrente desta última hipótese (Figura 3.22), com ênfase nas implicações para a diversificação da tribo Euglossini.

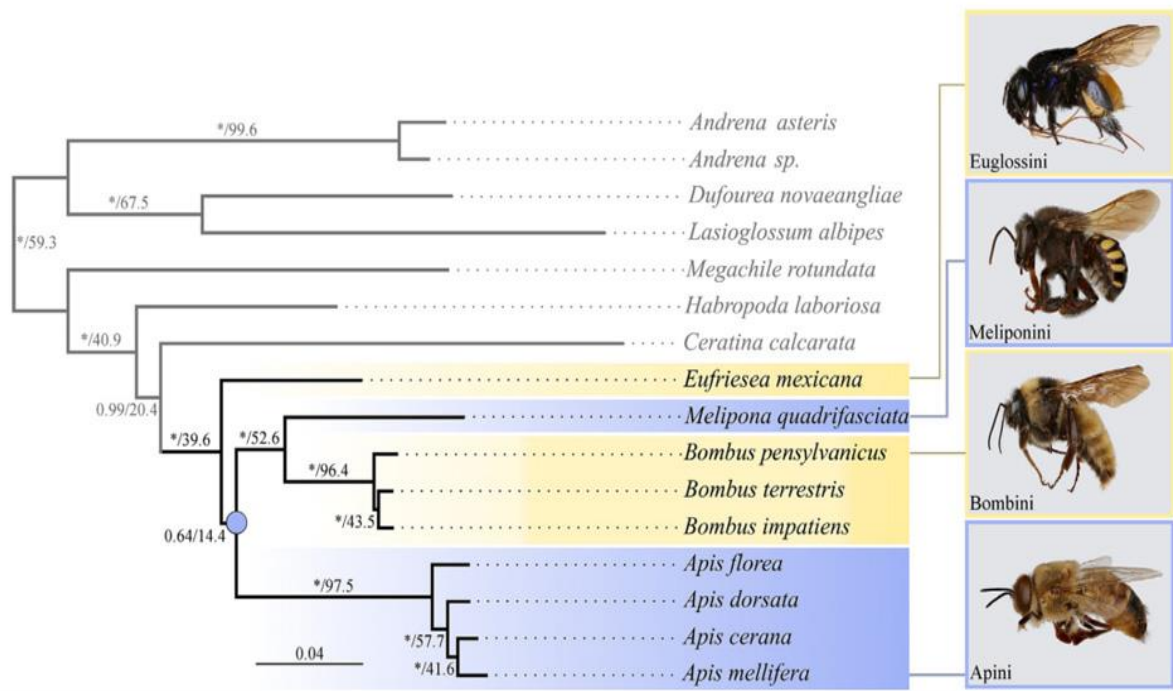


Figura 3.22 Filogenia das abelhas corbiculadas baseadas em análises filogenômicas corrigindo o viés do conteúdo GC (ROMIGUIER et al. 2016). Filograma resumido das 100 melhores árvores Bayesianas retirado de Bossert et al. (2017). Nó em destaque agrupa as espécies que representam as tribos dos corbiculados com comportamentos sociais mais estruturados.

Em Cameron (2004) encontram-se importantes considerações a respeito da possível origem dos euglossíneos do ponto de vista filogenético, porém pouca ênfase é dada sobre os aspectos geológicos e biogeográficos envolvidos neste processo. De fato, poucos trabalhos se concentraram na elucidação de quando e onde exatamente a tribo Euglossini teria surgido, e é bem provável que a depender dos táxons selecionados para tanto, múltiplas hipóteses plausíveis sejam levantadas, dada a complexidade histórica da região Neotropical.

Embora possam existir algumas diferenças quanto ao agrupamento interno de filogenias em Euglossini (ver seção Aspectos Filogenéticos), o que limita as considerações de detalhes do processo, a época estimada para o surgimento do grupo aparece com certa concordância em diferentes trabalhos. Em uma filogenia para toda a família Apidae utilizando 12 genes diferentes, Cardinal, Stracka e Danforth (2010) realizaram datações para diversificação da família utilizando o princípio do relógio molecular calibrado via registro fóssilífero. A partir dos resultados destes autores é possível observar uma data aproximada de 35 milhões de anos antes do presente para o provável surgimento e diversificação dos primeiros euglossíneos (CARDINAL; STRAKA; DANFORTH, 2010).

Interessantemente, datação semelhante é encontrada por Ramírez et al. (2010a) em um estudo independente, analisando a filogenia de Euglossini a partir de diversos táxons e genes diferentes dos utilizados por Cardinal et al. (2010) (ver detalhes na seção: Aspectos

filogenéticos). Estes dados permitem fixar com bom nível de segurança um período em torno de 35 e 40 Ma a. p. para a coalescência de todos os euglossíneos até seu ancestral comum mais recente.

Outro aspecto interessante em relação ao surgimento dos Euglossini, diz respeito à separação entre a linhagem do “*Protoeuglossini*”³ e seu grupo irmão, os demais corbiculados, a partir do ancestral comum mais recente. Os resultados de Cardinal et al. (2010) suportam a separação entre Apini e Euglossini há cerca de 75 milhões de anos antes do presente, cerca de 20 milhões de anos antes de coalescer com o ancestral dos corbiculados (~95 Ma a. p.) (CARDINAL; STRAKA; DANFORTH, 2010) e 20 milhões de anos antes de coalescer com todos os outros Apidae (~110 Ma a. p.) (BRANSTETTER et al., 2017; JOHNSON et al., 2013).

Em resumo, traçando uma linha de volta ao tempo, tem-se todos os euglossíneos coalescendo entre 35 e 40 Ma, Euglossini e Apini (ou Apini + (Meliponini + Bombini) a cerca de 75 Ma, corbiculados a 95 Ma e Apidae a cerca de 110 Ma. Considerando que a separação completa entre América do Sul e África é estimada também para cerca 110 Ma (LAVINA; FAUTH, 2016), a evolução de toda linhagem ancestral de Euglossini (“*Protoeuglossini*”) deve ter se dado já no continente sul-americano.

É importante ressaltar as condições da paisagem ao longo da evolução dos euglossíneos para entender melhor os aspectos sobre sua origem. A história geológica da América do Sul é riquíssima neste período (últimos 120 Ma), na qual intensas modificações geológicas e biológicas aconteceram (veja a seção 3.2 Formação e Biogeografia da América do Sul: Aspectos Gerais). Bastaria lembrar que neste período, houve a formação do Oceano Atlântico e Cordilheira dos Andes (LAVINA; FAUTH, 2016), porém interessante o surgimento dos euglossíneos atuais ocorre após e durante uma série de acontecimentos favoráveis à formação de florestas tropicais.

Com base nas datações atuais estimadas para o surgimento dos euglossíneos, os principais acontecimentos devem ter se dado a partir das épocas Eoceno (~50-32 Ma a. p.) e principalmente Oligoceno [~32-23 Ma a. p.] (COHEN et al., 2014). O Eoceno é principalmente marcado por uma fase de estabilidade climática, mais quente do que o presente [$>10^{\circ}\text{C}$] (PEARSON et al., 2007), que dava condições de formação vegetal em latitudes ainda menores do que as atuais, mesmo na região Patagônica (JARAMILLO;

³ O termo “*Protoeuglossini*” é aqui utilizado para referir-se à linhagens dos euglossíneos antes da diversificação dos cinco gêneros, porém após a divergência do grupo ancestral dos corbiculados.

RUEDA; MORA, 2006). Ao longo deste período a América do Sul já contava com uma posição latitudinal privilegiada, principalmente após o isolamento Antártico no fim do Eoceno e o congelamento daquele continente (CARTER et al., 2017) acompanhado de um rápido período glacial que pode ter forçado a biota tropical em direção ao norte (LAVINA; FAUTH, 2016).

Rebêlo (2001) faz importantes contribuições sobre os aspectos biogeográficos do surgimento das abelhas das orquídeas. Em seu livro, Rebêlo ressalta o isolamento da América do Sul ao longo de praticamente todo o Paleogeno e Neogeno superior (~65-10 Ma a. p.) como fator propício à diversificação dos euglossíneos neste continente, o qual voltou a se conectar com a América Central apenas no final do Neogeno (Mioceno, ~3.5 Ma a. p.). Contudo, alguns aspectos devem ser ressaltados à luz das descobertas mais recentes.

Existem, por exemplo, indícios de que haviam ligações fluviais entre a região sul do Panamá e o norte da Colômbia já em meados do Mioceno [15-10 Ma a. p.] (MONTES et al., 2015). Deste modo, é possível que a colonização da América Central tenha ocorrido muito antes dos atuais 3 ou 5 Ma antes estimados para o fechamento do Istmo do Panamá. Vale ressaltar que uma importante regressão marinha é esperada durante a transição entre Eoceno e Mioceno, devido ao congelamento Antártico (MÉRICO; TYRRELL; WILSON, 2008). Em outras palavras, há 40 Ma a. p. a Bacia Amazônica, em sua porção mais leste-central, apresentaria condições favoráveis para a diversificação dos euglossíneos, bem como, potenciais conexões poderiam ter surgido, principalmente nos últimos 20 Ma a. p., propiciando a colonização da América Central, a qual teria sido forçada pelo gradual resfriamento ao longo do Mioceno.

Interessantemente, o registro fóssil disponível corrobora a ocorrência dos euglossíneos em latitudes mais elevadas entre 15 e 20 Ma, registros de depósitos Dominicanos de âmbar do Mioceno (15-20 Ma a. p.) apontam indubitavelmente a existência da linhagem do gênero *Euglossa* viventes àquela época nas vulcânicas ilhas protoantíflicas (ENGEL, 1999), e mesmo antes disso, existem registros fósseis que apontam a existência de euglossíneos na mesma região mesmo antes disso, entre 20 e 30 Ma a. p. (POINAR, 1998).

Embora haja controvérsias se o fóssil supracitado seria um ancestral do gênero *Euglossa* (*Paleoeuglossa mellisiflora*⁴) ou *Eufriesea* (ENGEL, 1999), seguramente este registro indica a ocorrência de euglossíneos colonizando e se diversificando na região

⁴ A descrição morfológica do fóssil *Paleoeuglossa mellisiflora* (Poinar, 1998) é bastante sucinta, e deste modo, segue-se de acordo com Engel (1999), o qual transfere a espécie e cria a combinação *Eufriesea mellisiflora* (Poinar, 1998).

caribenha, bem antes do fechamento definitivo do Istmo do Panamá. Este fato ressalta ainda a possibilidade de diversificação dos euglossíneos em outras áreas que não a bacia Amazônica.

Finalmente, é importante considerar que as informações sobre a espécie fóssil *Eufriesea mellisiflora* (Poinar, 1998) são geralmente desconsideradas nas datações moleculares devido à sua posição dúbia, o que poderia indicar uma origem pretérita para o grupo. Contudo, os melhores indícios disponíveis indicam que os primeiros euglossíneos surgiram por volta de 40 Ma a. p. na região da Bacia Amazônica. Vale ainda ressaltar que embora a Bacia Amazônica seja considerada aqui o berço dos primeiros euglossíneos, é plausível que a diversificação de muitas espécies do grupo tenha sido mais complexa e que ondas de colonização em diferentes sentidos devem estar presentes na história biogeográfica do grupo.

3.4.2 Aspectos Sistemáticos

Como dito anteriormente, a descoberta dos compostos aromáticos como atrativos para os machos de euglossíneos contribuiu substancialmente para o aumento no número de trabalhos descrevendo novas espécies (NEMÉSIO, 2009; ROUBIK; HANSON, 2004). No entanto, o conhecimento sobre a sistemática e filogenia da tribo Euglossini continua ainda em construção e apenas alguns padrões que se apresentam de forma mais clara serão descritos.

Sistematicamente o grupo vem sofrendo uma série de modificações ao longo do tempo. Durante anos, a tribo foi subdividida em seis gêneros, quatro de vida livre (*Eufriesea*, *Euplusia*, *Euglossa* e *Eulaema*) mais dois cleptoparasitas (*Aglae* e *Exaerete*), até Kimsey (1979, 1982) reorganizar a tribo em apenas cinco gêneros atualmente aceitos (Figura 3.22). Kimsey (1979) trata *Euplusia* Moure (1944) como sinônimo júnior de *Eufriesea* Cockerell (1908). De acordo com esta autora, a terminália das espécies em *Euplusia* não apresentaria diferenças suficientes para sustentar a separação de *Eufriesea*. Ainda embora o próprio padre Moure tenha mantido sua posição em relação ao gênero *Euplusia* e tenha continuado a descrição de novas espécies, mesmo após as publicações de Kimsey (MOURE, 1999; MOURE; LUZ, 2001), assumir-se-á aqui a posição mais aceita na literatura, considerando válidos os cinco gêneros (como na Figura 3.20).

Aglae Lepelier & Serville 1825 é um gênero bastante peculiar em Euglossini pois apresenta apenas uma única espécie (*Aglae caerulea*) de grande tamanho (20-25 mm; Figura 3.20-A), cleptorasita exclusivo de *Eulaema* (DRESSLER, 1982a). A distribuição desta abelha restringe-se à florestas úmidas da bacia Amazônica e áreas adjacentes, além da região de Chocó (MOURE; MELO; FARIA JR, 2012; RAMIREZ; DRESSLER; OSPINA, 2002)

Exaerete Hoffmannsegg 1817 compõe com *Aglae* o grupo dos cleptoparasitas em Euglossini, utilizando tanto ninhos de *Eulaema* quanto *Eufriesea* (ROUBIK; HANSON, 2004). *Exaerete* apresenta oito espécies descritas (MOURE; MELO; FARIA JR, 2012) que são reconhecidos por exibir a metatíbia duas ou três vezes mais longa do que larga (Figura 3.20-B) e asas escurecidas. O tamanho intraespecífico de *Exaerete* pode variar consideravelmente, talvez refletindo a falta de provimento em determinadas ou mesmo à depender da espécie parasitada (ROUBIK; HANSON, 2004). Chaves de identificação para *Exaerete* são encontradas em Roubik e Hanson (2004) e Nemésio (2009).

Eulaema Lepeletier 1841 compreende atualmente a 21 espécies com os maiores tamanhos dentre os euglossíneos (entre 18 e 31 mm) (MELO et al., 2008), com coloração escura, pouco metálica e alta pubescência no tegumento. O gênero é amplamente distribuído, com espécies como *E. nigrita* ocorrendo em quase toda região Neotropical. Taxonomicamente, o gênero é dividido em dois subgêneros *Apeulaema* e *Eulaema* (*sensu stricto*) (MOURE, 1950; MELO, 2014), os quais foram redefinidos por Melo (2014) para corresponder à hipótese de filogenia molecular de Ramírez et al. (2010a), que por sua vez aponta uma parafilía entre os subgêneros segundo a classificação de Moure (1950).

As observações de Melo (2014) definem *El. (Apeulema)* basicamente como abelhas do gênero *Eulaema* que possuem uma conspícua depressão basal e circular no basitarso posterior com denso tufo de cerdas finas, além de longas cerdas na margem posterior do basitarso posterior. Por outro lado, *El. (Eulaema)* não apresenta a estrutura circular com tufos, mas sim um basitarso ainda mais côncavo na face externa com cerdas esparsas; variação na extensão apical da fenda metatibial; franjas da abertura não se projetam sobre a margem inferior da tíbia e a ponta das franjas não ultrapassa o ápice da projeção tibial externa. Existem ainda subdivisões de grupos de espécies, a exemplo dos gêneros mais diversos. *El. (Apeulaema)* possui dois grupos (“speciosa” e “cingulata”) enquanto *El. (Eulaema)* possui quatro grupos (bombiformis, meriana, polyzona e seabrai). Chaves de identificação para *Eulaema* (*sensu amplo*) são encontradas em Roubik e Hanson (2004), Melo (2014) e Nemésio (2009).

Eufriesea Cockerell 1908 é o segundo gênero mais rico em espécies na tribo Euglossini com 67 espécies ocorrendo desde o México à Argentina (MOURE; MELO; FARIA JR, 2012), tornando-se recentemente o segundo gênero com registro para os Estados Unidos (*Ef. coerulescens*) em áreas de parques de Eddy County NM e Culberson County TX (GRISWOLD; HERNDON; GONZALEZ, 2015). Este gênero possui uma marcante sazonalidade ocorrendo em poucos meses do ano, geralmente entre outubro e fevereiro (DRESSLER, 1982a; ROUBIK; HANSON, 2004), o que dificulta a descrição detalhada de

seus aspectos biológicos. Aparentemente, estas abelhas podem não ser tão sensíveis quanto os demais euglossíneos às essências frequentemente utilizadas em inventários e podem responder a diferentes compostos ainda não explorados (DEC; DA SILVA MOUGA; ALVES-DOS-SANTOS, 2017).

De acordo com Roubik e Hanson (2004), *Eufriesea* se define por possuir face com coloração metálica, sem machas circulares no labro e raramente com apenas uma linha medial no clípeo. *Eufriesea* é dividido em doze grupos de espécies (KIMSEY, 1982), embora recentes análises moleculares tenham indicado que vários destes grupos não são monofiléticos (RAMÍREZ et al., 2010) e alguns destes tenham sofrido revisão nos últimos anos (GONZALEZ; GRISWOLD; SIMÕES, 2017). Chaves de identificação para *Eufriesea* são encontrados em Kimsey (1982), Roubik e Hanson (2004) e Nemésio (2009).

De todos os gêneros, *Euglossa* Latreille 1802 é o gênero com o maior número de espécies, aproximadamente 130 espécies descritas (MOURE; MELO; FARIA JR, 2012) e não surpreendentemente exibe o maior número de subgêneros e grupos de espécies. Ao total são seis subgêneros (*Dasystilbe*, *Glossura*, *Glossurella*, *Glossuropoda*, *Euglossella*, *Euglossa*) e 12 grupos de espécies (DRESSLER, 1982a, 1982b, 1982c; ROUBIK; HANSON, 2004), dos quais vários têm se mostrado não monofiléticos frente às análises moleculares (RAMÍREZ et al., 2010).

De modo geral, membros do gênero *Euglossa* são facilmente reconhecidos por apresentar a fenda metatibial curta, sem alcançar a margem inferior da metatíbia e curvada na porção inferior (Figura 3.20-E); mesotíbia com tufos/coxins presentes (1 a 3) na porção apical da área aveludada (Figura 3.20-F); labro usualmente esbranquiçado com duas manchas ovais escuras.

Espécies do gênero *Euglossa* são as menores abelhas da tribo com cerca de 8 mm de comprimento como espécies do grupo *crassipunctata* (DRESSLER, 1982c), mas podendo chegar a abelhas de médio porte, como *Eg. imperialis* e *Eg. intersecta* que podem ter cerca de 15 mm (MELO et al., 2008; ROUBIK; HANSON, 2004). Todas as espécies do gênero são de vida livre e apresentam um amplo espectro de coloração integumentar metálica entre verde, vermelho, dourado, tons cúpreos e azuis. Chaves de identificação para diversas espécies de *Euglossa* são encontradas em Rebêlo e Moure (1995), Roubik e Hanson (2004) e Nemésio (2009).

3.4.3 Aspectos Biogeográficos

As abelhas das orquídeas representam uma fonte riquíssima de informações biogeográficas. Contudo, a exploração de seus padrões tem sido pouco estudada ao longo dos últimos anos. É possível que em um futuro breve, com a crescente disponibilização de dados geográficos, de ocorrência de espécies e filogenias moleculares; o estabelecimento de padrões cada vez mais criteriosos para a realização de inventários e a facilidade de se analisar dados georreferenciados, padrões mais detalhados sobre a tribo e áreas de endemismo sejam propostos e testados. As próximas páginas buscam revisar o conhecimento sobre os padrões biogeográficos e filogeográficos atuais sobre os euglossíneos, visando fornecer uma base comparativa entre os resultados encontrados no presente estudo e os padrões observados para euglossíneos.

3.4.3.1 *Euglossini como modelo*

Dentre uma infinidade de táxons presentes ao longo das florestas tropicais, as abelhas das orquídeas se destacam como modelos para o estudo filogeográfico devido a uma série de modificações, adaptações e peculiaridades. Euglossíneos são altamente diversificados, exclusivamente neotropicais e com ampla distribuição (MOURE; MELO; FARIA JR, 2012; NEMÉSIO, 2009; ROUBIK; HANSON, 2004), mas ao mesmo tempo possuem altas taxas de endemismo (NEMÉSIO, 2009; RAMÍREZ et al., 2010) produzindo uma série de experimentos naturais onde padrões filogeográficos podem ser testados.

Talvez, um dos principais aspectos que sustentem o uso dos euglossíneos como modelos biogeográficos seja sua intrínseca relação com o seu ambiente de origem (DRESSLER, 1982a; GIANGARELLI et al., 2009; GIANGARELLI; AGUIAR; SOFIA, 2015; PARRA-H; NATES-PARRA, 2012; SILVEIRA et al., 2015). Um exemplo interessante é demonstrado por Silveira e colaboradores (2015) que, ao compararem áreas abertas e de floresta, observaram que mais de 70% da distribuição das espécies poderia ser explicada pela categoria ambiental estabelecida, savana ou floresta.

A relação de cleptoparasitismo entre congêneres no grupo é um outro exemplo interessante. Espécies dos gêneros *Aglae* e *Exaerete* dependem não apenas da presença das espécies parasitadas, mas por conseguinte, de que seus “hospedeiros” possuam recursos suficientes para construção e provimento dos ninhos (DRESSLER, 1982a; ROUBIK; HANSON, 2004). Esta relação já foi apontada como um fator biogeográfico sobre a distribuição de *Exaerete* e *Eulaema* ao longo das florestas sul-americanas (NEMÉSIO; SILVEIRA, 2006b).

Interessantemente, espécies dos gêneros *Euglossa*, *Eulaema* e *Eufriesea* dependem não apenas da exploração de vários recursos comuns a todas as abelhas, mas também de fontes aromáticas para fins de reprodução (ELTZ et al., 2007; ELTZ; SAGER; LUNAU, 2005), bem como *Aglae* e *Exaerete* (ROUBIK; HANSON, 2004). A busca constante destes compostos aromáticos, coloca os euglossíneos sob constante pressão sexual (WEBER et al., 2016), relacionada à capacidade dos machos em explorarem o ambiente e encontrarem recursos para atração das fêmeas (ELTZ et al., 2007; ELTZ; ROUBIK; WHITTEN, 2003) e alterações no ambiente que tornem tais recursos escassos podem levar a uma estruturação geográfica ou mesmo até à especiação (BRAND et al., 2015). Finalmente, a relação euglossíneos/meio passa pelas diferentes restrições ecológicas apresentadas por cada espécie ou grupo de espécies (LÓPEZ-URIBE et al., 2014; MILET-PINHEIRO; SCHLINDWEIN, 2005; NEMÉSIO; SILVEIRA, 2006a, 2006b), as quais potencialmente podem levar à diferentes estruturações filogeográficas dentro de um mesmo bioma (LÓPEZ-URIBE et al., 2014). Essa relação indissolúvel com o meio por tantas vias diferentes, torna a tribo Euglossini um modelo ímpar para estudos biogeográficos, uma vez que trazem atrelado à sua história biogeográfica também, a história do domínio biogeográfico em si.

Um outro aspecto singular aos euglossíneos, do ponto de vista filogeográfico, é a distinção entre os padrões de dispersão de machos e fêmeas de euglossíneos. Como descrito anteriormente, machos destas abelhas tem ampla capacidade de dispersão, voando quilômetros diariamente (WIKELSKI et al., 2010) ou mantendo fluxo gênico em longas distâncias (DICK et al., 2004; POKORNY et al., 2015). Por outro lado, as fêmeas aparentam ter comportamento filopátrico⁵, reativando seus ninhos ou nidificando nas proximidades (AUGUSTO; GARÓFALO, 2011). Este padrão proporciona a rara oportunidade de se comparar dois modelos de dispersão dentro da mesma unidade taxonômica, a qual obrigatoriamente compartilha uma mesma história evolutiva.

Ainda do ponto de vista Filogeográfico, euglossíneos apresentam uma outra facilidade bastante conveniente para análises moleculares. Como discutido anteriormente (Filogeografia e seu impacto sobre a Biogeografia), estudos filogeográficos tradicionalmente se baseiam em genealogia gênicas necessariamente haplotípica, uma vez que heterozigosidade podem ser encaradas como ambiguidades nas reconstruções genealógicas (AVISE, 2009). Deste ponto de vista, euglossíneos apresentam uma vantagem adicional, pois devido à metodologia de

⁵ Filopatria é usada neste contexto para se referir ao fato de preferencialmente as fêmeas escolherem reativar o próprio ninho ou nidificar nas proximidades, ainda que fêmeas tenham uma ampla capacidade de voo (DRESSLER, 1982d).

iscas odoríferas aplicadas ao longo das amostragens, apenas machos haploides são amostrados, como todos os machos de Hymenoptera. Este fato exclui a necessidade de determinação de fase haplotípica para os locos nucleares, facilitando abordagens genealógicas.

A representatividade dos euglossíneos também é um ponto interessante do ponto de vista filogeográfico. Estas abelhas são bastante abundantes em florestas de baixa e média altitude, compondo cerca de até 25% das comunidades locais (ROUBIK; HANSON, 2004), apresentando ainda uma correlação positiva entre a riqueza de espécies e a pluviosidade (DRESSLER, 1982d), facilitando, por exemplo, o estudo da variação climática como um fator biogeográfico. Devido sua peculiar atração por compostos aromáticos, euglossíneos têm sido estudados com certa padronização ao longo dos últimos anos, de modo que, a maior parte dos dados levantados são passíveis de comparação em algum nível (GARRAFFONI; MOURA; LOURENÇO, 2017; SYDNEY; GONÇALVES; FARIA, 2010). Adicionalmente, poucas subáreas dos grandes biomas ainda não possuem ao menos um levantamento de espécies de euglossíneos (DE OLIVEIRA, 2007; DRESSLER, 1982a; RAMALHO; GAGLIANONE; OLIVEIRA, 2009; ROUBIK; HANSON, 2004; STORCK-TONON; MORATO; OLIVEIRA, 2009), principalmente na Mata Atlântica (MOURE; MELO; FARIA JR, 2012; NEMÉSIO, 2009).

Frente ao exposto, fica evidente que as abelhas das orquídeas representam um modelo ímpar para o estudo filogeográfico ao longo de suas respectivas distribuições. Poucos ou talvez nenhum grupo monofilético apresente tantas particularidades propícias ao estudo da variação de formas através do tempo e espaço, retendo uma riquíssima história biogeográfica indissolivelmente ligada à história de seu próprio ambiente.

3.4.3.2 *Padrões Biogeográficos*

Ainda que os euglossíneos representem um modelo tão atraente a estudos filogeográficos e biogeográficos, pouco tem sido explorado deste grupo quando comparado a outros táxons tradicionalmente mais estudados na Mata Atlântica e Amazônia. Diversos estudos têm coberto padrões de outros organismos, como aves (BATALHA-FILHO; CABANNE; MIYAKI, 2012; BATALHA-FILHO; MIYAKI, 2016; CABANNE et al., 2014; CABANNE; SANTOS; MIYAKI, 2007; RAPOSO DO AMARAL et al., 2013), serpentes (GRAZZIOTIN et al., 2006; WÜSTER et al., 2005), anfíbios (THOMÉ et al., 2010, 2014), e invertebrados, como opiliões (DA SILVA; PINTO-DA-ROCHA; SOUZA, 2016). Apenas

recentemente alguns estudos começaram a explorar os padrões biogeográficos e filogeográficos a partir dos euglossíneos como modelo.

Como descrito anteriormente, abelhas das orquídeas são amplamente distribuídas ao longo da região Neotropical, tanto latitudinalmente (Argentina aos Estados Unidos) quanto longitudinalmente (Brasil à Colômbia), onde evoluem há milhões de anos, atingindo uma elevada diversidade de espécies (DRESSLER, 1982d; MOURE; MELO; FARIA JR, 2012; NEMÉSIO, 2009; RÊBELO, 2001; ROUBIK; HANSON, 2004). As tentativas de sumarizar os padrões biogeográficos para os euglossíneos ao longo dessa ampla distribuição encontraram uma estruturação de fauna dividida em três grandes áreas de endemismo: América Central (AC), Amazônia (AM) e Mata Atlântica (MA) (DRESSLER, 1982a; RAMÍREZ et al., 2010; ROUBIK; HANSON, 2004; SYDNEY; GONÇALVES; FARIA, 2010), o que é congruente com a grande maioria da biota Neotropical (MORRONE, 2014).

Quando comparadas entre si, as regiões América Central e Amazônica tendem a formar um grupo mais coeso quando comparados à Mata Atlântica a qual se mantém isolada das demais regiões, até certo ponto (RAMÍREZ et al., 2010; SYDNEY; GONÇALVES; FARIA, 2010). Dois pontos são de especial interesse quanto em relação a este aspecto. O primeiro deles diz respeito ao já esperado maior compartilhamento entre Amazônia e América Central, o qual é frequentemente reportado para a maior parte das espécies e grupo de espécies da região Neotropical (MORRONE, 2014). O segundo seria o isolamento da Mata Atlântica em relação às demais áreas que sofre grande influência da diagonal seca da América do Sul. Ainda assim, Sydney, Gonçalves e Faria (2010) encontraram importantes similaridades entre a fauna de um domínio denominado “Paranaense” e a Amazônia, indicando uma possível conectividade entre essas áreas. Mais uma vez, esses corredores pretéritos já foram demonstrados em outros estudos (COSTA, 2003; LEDO; COLLI, 2017) e de fato constituem um dos principais pontos de intercâmbio entre Mata Atlântica e Amazônia (LEDO; COLLI, 2017).

Embora padrões de distribuição sejam úteis para o estabelecimento de padrões filogeográficos, os mesmos frequentemente falham em considerar as relações filogenéticas entre os organismos, caso não assumam filogenias acuradas (veja a seção 3.1.1 Principais Ideias e Métodos Biogeográficos). Ramírez et al. (2010) utilizaram uma ampla filogenia molecular para explorar padrões filogeográficos de forma mais refinada. Com base nos resultados, estes autores dividiram a área de distribuição dos euglossíneos em sete áreas de endemismo: América Central e Antilhas (Ca), Zona Endêmica do Panamá (Pn), Choco (Ch),

Andes (An), Corredor Paraguai (Pr), Mata Atlântica (Af) e Bacia Amazônica (Am) (Figura 3.23).

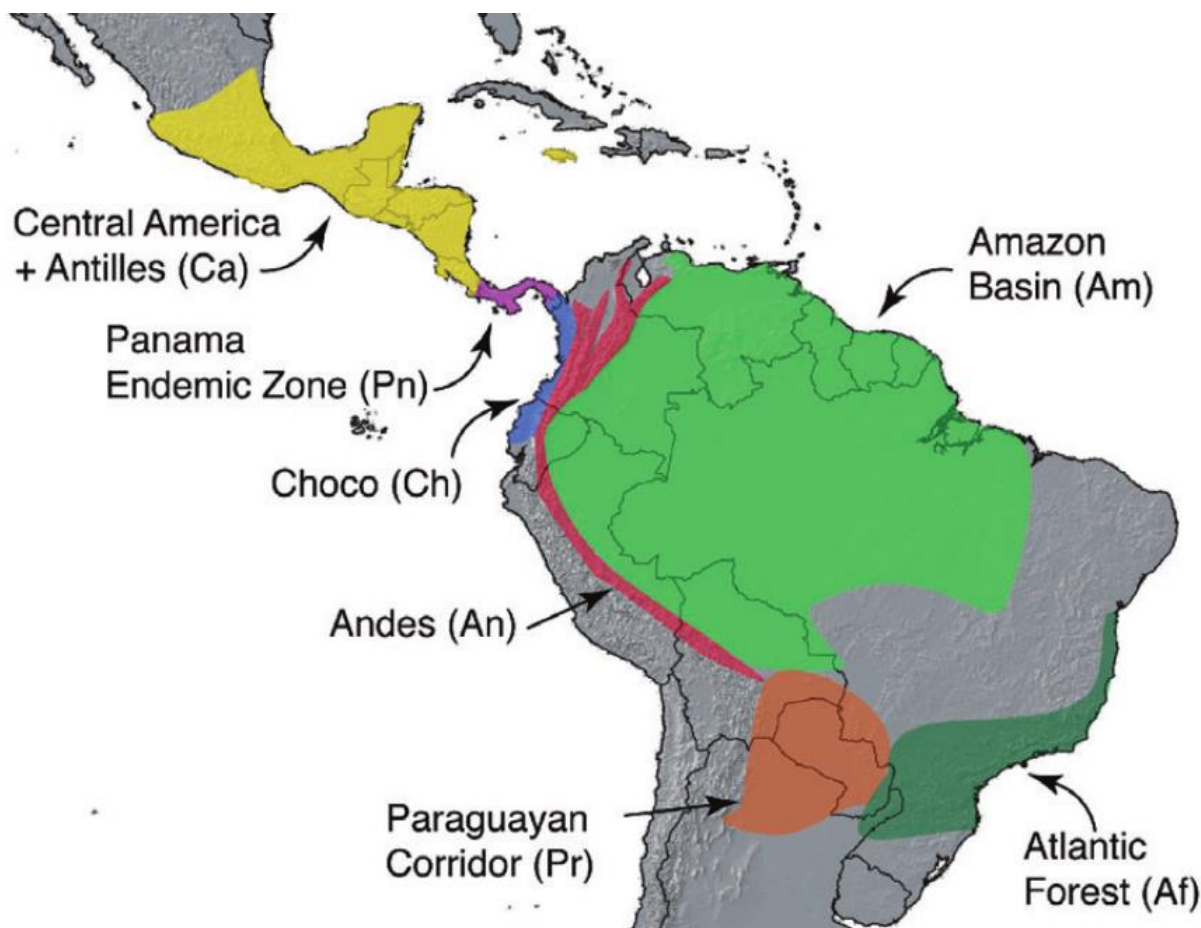


Figura 3.23 Áreas de endemismo para Abelhas das Orquídeas. As áreas de endemismo são demarcadas por diferentes cores. Reproduzido de Ramírez et al. (2010)

Embora a reconstrução das zonas de endemismo de Ramírez et al. (2010) tenha sido eficiente em recuperar as grandes áreas de endemismo, regiões mais complexas como a Mata Atlântica e Amazônia não foram suficientemente exploradas, de modo que as subestruturas destas áreas permaneceram generalizadas. Ainda assim, importantes padrões foram observados a partir dos dados destes autores. Ao combinar a cronologia filogenética e a reconstrução de áreas ancestrais, os autores notaram pontos de diversificação das espécies e distribuição em suas respectivas áreas em torno de $2,41 \pm 1,77$ Ma a. p., o que colocaria esses eventos de diversificação no final do Mioceno e início do Plioceno (detalhes na seção 3.2.2 A Era Cenozoica na América do Sul). Estes resultados trazem duas importantes implicações. A primeira delas diz respeito à súbita diversificação dos euglossíneos após o início do Quaternário (Plioceno), salientando a importância das flutuações climáticas como um fator de estruturação filogeográfica para os euglossíneos. A segunda tem relação óbvia com o fechamento do Istmo do Panamá, uma vez que a datação de Ramírez et al. (2010) para a diversificação dos euglossíneos sobrepõe a época do Grande Intercâmbio Americano, que por

sua vez, é o principal ponto de sustentação para o fechamento completo do Istmo do Panamá há ~3 Ma a. p. Interessantemente, os mesmos autores datam uma grande diversificação dentro dos gêneros de Euglossini para cerca de 15-20 Ma a. p., quando dados apontam a possíveis paleoconexões entre América Central e América do Sul via Istmo do Panamá (veja discussão nas pg. 69-50).

Outro aspecto bastante relevante proveniente da datação de Ramírez et al. (2010) está relacionado com o soergimento dos Andes, uma vez que, grande parte dos eventos orogênicos de soergimento relacionados às cordilheiras são datados para o início e médio Mioceno (20 a 12 Ma) em diante (LAVINA; FAUTH, 2016). Os dados de Ramírez et al. (2010) são contundentes em apontar o soergimento dos Andes e as conexões e desconexões entre as Américas como fatores fundamentais para a diversificação de Euglossini.

No tocante à Mata Atlântica, recentemente Garraffoni e colaboradores (2017) utilizaram o método PAE e análise de endemicidade NDM para descrever as relações de áreas de endemismo do domínio de Mata Atlântica com base em dados de assembleias de euglossíneos. Estes autores incluíram dados de 1630 localidades e 66 espécies, resultando em cinco grandes áreas de endemismo na Mata Atlântica (Figura 3.24).

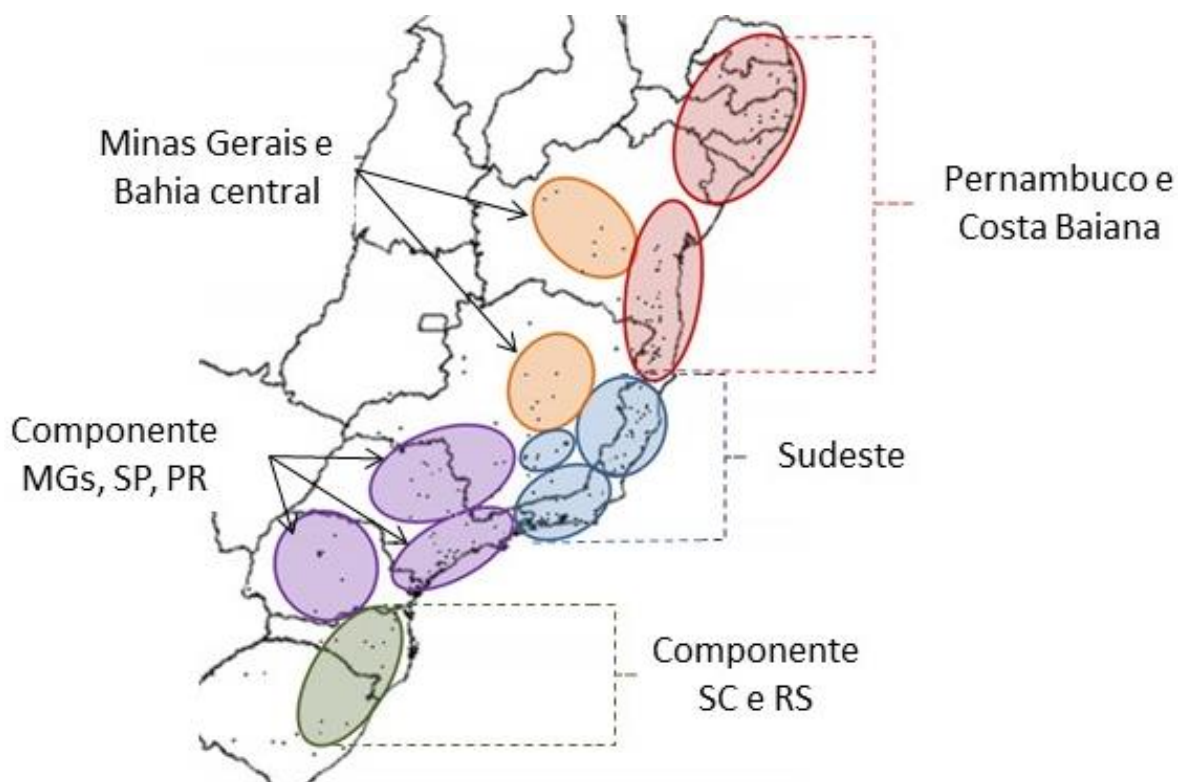


Figura 3.24 Áreas de endemismo na Mata Atlântica a partir do compartilhamento de espécies de Euglossini. As cores, setas e chaves demarcam as áreas de endemismo propostas por Garraffoni et al. (2017). Adaptado de Garraffoni et al. (2017).

Embora este estudo tenha sido pioneiro em delimitar áreas de endemismo para euglossíneos da Mata Atlântica, dois pontos importantes merecem ressalvas. O primeiro deles

diz respeito à natureza das análises. É importante ter em mente que as análises de Endemismo via Parcimônia e as análises NDM, possuem limitações e podem apresentar diferentes resultados frente a pequenas modificações nos bancos de dados (MORRONE, 2008), tamanhos de quadrícula e níveis de informação, como discutido pelos próprios autores (GARRAFFONI; MOURA; LOURENÇO, 2017). O segundo ponto diz respeito à classificação taxonômica utilizada por Garraffoni, Moura e Lourenço (2017), os quais embora tenham dito se basear em Moure et al. (2012) como referência final para as decisões taxonômicas, incluíram espécies sob sinonímia a exemplo de *Eg. solangeae*, sinônimo júnior de *Eg. stellfeldi* (FARIA JR; MELO, 2007; MOURE; MELO; FARIA JR, 2012) e outras espécies de taxonomia sobreposta, como *Eg. sapphirina*, *Eg. clausi* e *Eg. crassipunctata* (NEMÉSIO; ENGEL, 2012). Embora estes fatores não invalidem os padrões gerais apresentados por Garraffoni, Moura e Lourenço (2017), é possível que análises futuras apresentem resultados divergentes quanto aos limites das zonas de endemismo para euglossíneos. Os próprios autores salientam a importância de aplicações de outros métodos para corroborar ou descartar estas zonas de endemismo ao longo da Mata Atlântica.

3.4.3.3 Padrões Filogeográficos: existe algum?

Se os estudos de padrões biogeográficos são escassos para Euglossini, os de cunho Filogeográfico são ainda mais. Considerando o nível hierárquico em que a Filogeografia trabalha, como populações e espécies irmãs, pode-se verificar que existem poucos estudos voltados especificamente a este tipo de abordagem, como os trabalhos de Dick et al. (2004), López-Uribe et al. (2014) e Frantine-Silva et al. (2017). Há ainda o estudo de Ferrari e Melo (2014) que não necessariamente foi desenhado para testar hipóteses filogeográficas, mas foi incluído na presente discussão por apresentar um desenho experimental de certo modo adequável à questão. Obviamente, estudos populacionais considerando a distribuição geográfica da diversidade genética, são também parte da Filogeografia, como por exemplo os estudos de Penha et al. (2015) e Rocha-Filho et al. (2014). Contudo, os dois últimos não foram desenhados para testar hipóteses filogeográficas, diferentemente dos estudos de Dick et al. (2014), López-Uribe et al. (2014) e Frantine-Silva et al. (2017).

Em um grupo tão diverso e historicamente rico, como no caso da tribo Euglossini, delinear padrões a partir de tão poucos dados é tarefa bastante difícil, talvez até mesmo vã. Contudo, com base nos trabalhos apresentados é possível encontrar as seguintes características: a) Euglossíneos possuem grande capacidade dispersiva; b) A tolerância ecológica é determinante para a presença de estruturas; c) A filopatria parece exercer forte

influência nos padrões filogeográficos; d) Flutuações climáticas parecem ter grande potencial em definir estruturas geográficas.

No que tange a capacidade dispersiva do grupo, os dados moleculares e filogeográficos parecem apenas confirmar as observações etológicas de grande capacidade de voo tanto em fêmeas (DRESSLER, 1982a; ROUBIK; HANSON, 2004) quanto em machos (POKORNY et al., 2015; WIKELSKI et al., 2010). A novidade dos dados moleculares em relação às espécies é que, embora machos e fêmeas possam voar por quilômetros, aspectos ecológicos e comportamentais parecem restringir a distribuição das espécies de Euglossini. Em um estudo analisando três espécies de *Eulaema* (abelhas de grande porte e amplamente distribuídas; seção 3.4.2), López-Urbe et al. (2014) constataram que as diferenças ecológicas das ecorregiões promoviam a restrição de fluxo gênico entre as populações ao longo da América Central, Amazônia e Mata Atlântica, principalmente em épocas de instabilidade climática. López-Urbe e colaboradores, também observaram que os marcadores nucleares falharam em recuperar estruturas geográficas tão conspícuas quanto às apresentadas nas análises mitocondriais, concluindo que o comportamento filopátrico das fêmeas é ponto determinante nesta questão.

De fato, os dados de López-Urbe et al. (2014) encontram suporte nas análises realizadas por Penha et al. (2015) e Frantine-Silva et al. (2017). Note que neste ponto, embora Penha et al. (2015) não tenha tido como objetivo primeiro a análise filogeográfica, seus resultados em muito auxiliam na questão. Estes autores analisaram populações de *Eg. iopoecila* (à época com questões taxonômicas envolvidas; veja seção 3.5 O caso de *Euglossa iopoecila* Dressler) ao sul da Mata Atlântica, observando estruturações muito mais acentuadas para o DNAm (herança matrilinear) em relação aos microssatélites, passíveis de dispersão pelos machos. Alguns anos depois, Frantine-Silva et al. (2017), também utilizando *Eg. iopoecila* como modelo, observaram que as flutuações climáticas do Quaternário potencialmente foram as principais responsáveis pela determinação da estrutura genética mitocondrial da espécie, corroborando uma dispersão gradual do DNAm (detalhes no CAPÍTULO 3: FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DE *EUGLOSSA IOPOECILA* DRESSLER: SINAIS DE VICARIÂNCIA ASSOCIADOS ÀS MUDANÇAS CLIMÁTICAS DO QUATERNÁRIO).

Embora a regra da filopatria seja um paradigma emergente para explicar os padrões filogeográficos em diferentes espécies de euglossíneos, as conclusões de Dick et al. (2004) são um contraponto aparente. Com base na análise de 550 pb do gene mitocondrial *Citocromo c Oxidase I (COI)* de várias espécies de euglossíneos, principalmente ao longo da região

Central e região Andina (algumas amostras na Bolívia e Guiana Francesa), os autores concluem que a falta de estruturação geográfica deve ser fruto de múltiplas especiações ao longo do Quaternário, expansão populacional e/ou fluxo gênico; o que merece algumas ressalvas. Embora tenha-se dito anteriormente que o estudo de Dick et al. (2004) contempla um desenho experimental focado em objetivos filogeográficos, devido ao próprio título do estudo, o caráter supragenérico das análises destoa da abordagem filogeográfica tradicional; o que se torna uma questão secundária frente ao caráter multidisciplinar da Filogeografia e outros pontos mais críticos do estudo.

Dick et al. (2004) estendem algumas conclusões a partir dos dados explorados naquele estudo, os quais merecem algumas considerações. Primeiramente, o estudo de um único gene mitocondrial para explicar a história filogeográfica de uma, ou ainda várias unidades taxonômicas, como realizado por Dick et al. (2004), é bastante arriscado devido ao paradoxo “árvore de genes e de espécies” (BRITO; EDWARDS, 2009), também brevemente discutido na seção: “Filogeografia: do DNA citoplasmático à era Genômica”. Aliado a isso, existe o fato de que os tamanhos amostrais naquele estudo são inferiores a dez indivíduos por unidade taxonômica, o que pode levar a um viés de tamanho amostral (ex. um haplótipo comum amostrado em duas áreas resultaria fatidicamente no compartilhamento de alelos). Adicionalmente, a taxa de mutação da região utilizada (*COI*) frequentemente não se adequa ao estudo de padrões populacionais, visto que esta região é conhecida por acumular diferenças intraespecíficas com baixa diferenciação intraespecífica (HEBERT; RATNASINGHAM; DE WAARD, 2003). Taxas de mutação reduzidas implicariam em ausência de estruturação (PALSBØLL; BÉRUBÉ; ALLENDORF, 2007), a qual no estudo em questão foi interpretada como presença de fluxo gênico. É importante ainda ter em mente que o *COI* é um gene de aproximadamente 1200 pb (CROZIER; CROZIER, 1993) com uma região aparentemente mais variável (porção 5’), a mesma utilizada para a identificação molecular de espécies (HEBERT et al., 2003; HEBERT; RATNASINGHAM; DE WAARD, 2003), e uma segunda porção nucleotidicamente menos diversa (porção 3’; veja FRANTINE-SILVA et al., 2017), coincidentemente a utilizada por Dick et al. (2004). Outro ponto diz respeito ao fato de que parte do conceito de espécies utilizadas por Dick et al. (2004) era falha, principalmente para as espécies do grupo analisado (veja FARIA; MELO, 2012). Deste modo, é difícil assumir a premissa de que os dados de Dick et al. (2004) representem de fato fluxo gênico através dos Andes ou em distâncias geográficas tão grandes quanto 3000 Km apenas com o desenho experimental utilizado naquele trabalho, mesmo que isso seja possível para algumas espécies.

No que tange a Mata Atlântica, existem três estudos que permeiam a questão filogeográfica mais diretamente, Ferrari e Melo (2014), Penha et al. (2015) e Frantine-Silva et al. (2017). À exemplo de Penha et al. (2015), Ferrari e Melo (2014) não tinham foco filogeográfico específico como objetivo primário, mas seus dados são extremamente convenientes à abordagem. Ferrari e Melo (2014) analisaram populações de três espécies do gênero *Euglossa* com ampla distribuição ao longo da costa brasileira, *Eg. townsendi*, *Eg. stellfeldi* e *Eg. iopoecila*, visando verificar a congruência da hipótese de espécies com morfotipos de coloração diferentes. Para as três espécies há uma distinção de coloração tegumentar entre regiões norte (verde) e sul (azul) de suas distribuições (ponto médio, região de Ubatuba, SP - Macaé, RJ), o que por si já indicaria a existência de uma certa estrutura filogeográfica. Contudo, Ferrari e Melo (2014) evidenciaram pelo estudo de dois genes mitocondriais (*cytb* e *COI*) que não havia correspondência direta entre os morfotipos de coloração e linhagens genéticas que sustentassem o status de espécie para qualquer uma das três espécies em questão.

Embora não reste dúvida quanto a unicidade de *E. iopoecila* como espécie (veja o item 3.5 O caso de *Euglossa iopoecila* Dressler), duas razões principais podem ter impedido Ferrari e Melo (2014) de recuperar estruturas filogeográficas mais detalhadas para estas espécies. A primeira delas é que o estudo não foi desenhado para tal propósito, portanto, faltam análises direcionadas a este fim. A segunda diz respeito à análise conjunta das três espécies aliadas a outros congêneres e *outgroups* na mesma árvore Bayesiana, o que tende a minimizar as diferenças intraespecíficas frente às diferenças interespecíficas. Este detalhe pode ser observado na análise de Frantine-Silva et al. (2017), os quais encontraram linhagens mitocondriais congruentes com os morfotipos de coloração para *Eg. iopoecila*, além de uma estruturação adicional entre as abelhas de tegumento verde, todas classificadas apenas como diferenças populacionais em consonância aos padrões de distância genética encontrados por Ferrari e Melo (2014).

Embora seja difícil estabelecer padrões frente a tão poucos estudos, quebras de padrões geográficos como as estudadas por Ferrari e Melo (2014), Penha et al. (2015) e Frantine-Silva et al. (2017) apontam para eventos geológicos e/ou climáticos comuns aos euglossíneos. Adicionalmente, quase todos os estudos até o presente ponderam a influência das flutuações climáticas do Quaternário sobre a estruturação populacional das espécies estudadas (FRANTINE-SILVA et al., 2017; LÓPEZ-URIBE et al., 2014; PENHA et al., 2015) e possíveis eventos de especiação (DICK et al., 2004; RAMÍREZ et al., 2010). Neste contexto, o presente estudo visa contribuir com informações sobre a congruência de padrões

distribucionais ao longo do último ciclo glacial e dados genéticos, com o intuito de testar a influência do paleoclima Pleisto-Holocênico sobre a estruturação genética dos euglossíneos com foco na Mata Atlântica.

3.4.4 Diversidade Genética: o estado da arte em Euglossini

Ao contrário dos estudos filogeográficos, o campo da Genética de Populações e o estudo da diversidade genética nos Euglossini tem mostrado consideráveis avanços nas últimas décadas. Parte provável deste acúmulo de estudos relacionados ao tema é com certeza decorrente de vários fatores, como por exemplo, a facilidade em amostrar os indivíduos e de gerar dados moleculares, mas principalmente devido ao importante papel ecológico destas abelhas na polinização de espécies nativas e a crescente preocupação mundial com o declínio das populações de abelhas (LOZIER; ZAYED, 2017; POTTS et al., 2010; ZAYED, 2009).

Como visto anteriormente, abelhas são especialmente frágeis a degradação ambiental e estão cada vez mais susceptíveis a ameaças antrópicas, o que é agravado de certo modo pelo sistema haplodiploide de determinação sexual dos himenópteros, o qual torna estes insetos mais propensos à perda de diversidade genética e aos efeitos estocásticos (ZAYED, 2009). Apesar disto, ainda são escassos os estudos genéticos envolvendo populações de abelhas, em particular os euglossíneos, sendo necessários maiores esforços no sentido de preencher tal lacuna.

Embora ainda esteja longe de alcançar níveis equivalentes a outros grupos, há um crescente acúmulo de dados entre diversas espécies de euglossíneos (Figura 3.25-A). Uma breve busca no portal “*Google Scholar*” com a expressão: “*Euglossini*” OR “*Euglossina*” AND “*population genetics*”; retornou 278 resultados, dos quais apenas 25 representaram de fato estudos relacionados à Genética de Populações ou Filogeografia de abelhas das orquídeas. Dentre estes 25 trabalhos, três são relacionados à Filogeografia (podendo-se adicionar Ferrari e Melo, 2014; Figura 3.25-B), 15 são relacionados ao estudo de Genética de Populações e Diversidade Genética (Figura 3.25-C), três ao reconhecimento de machos diploides (podendo-se adicionar BOFF et al. 2014; HATANAKA et al, 2001; ROUBIK et al. 1996; SORO et al. 2017; Figura 3.25-D), e quatro estudos foram dedicados à descrição de primers (incluindo Penha et al. 2015) (Veja Apêndices). É possível ainda observar com base nestes breves resultados, a escassez de informações a respeito de diferentes espécies em Euglossini, visto que pouco mais de trinta espécies foram acessadas em estudos de diferentes naturezas.

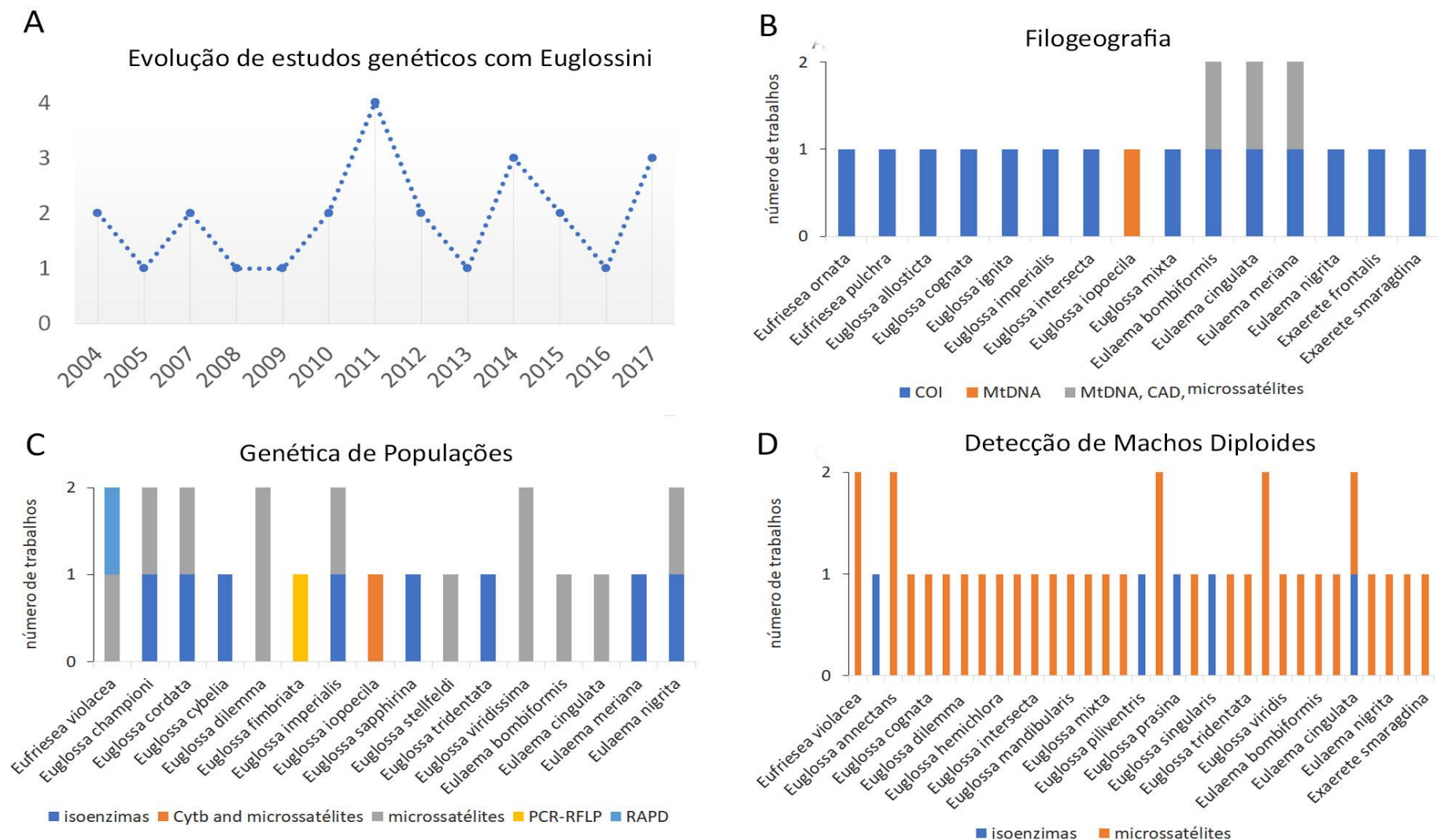


Figura 3.25 Trabalhos publicados com euglossíneos e relacionados à Genética de Populações, Filogeografia e Detecção de Machos Diploides. Resultados relacionados à busca “*Euglossini*” OR “*Euglossina*” AND “*Population Genetics*” no portal *Google Scholar*. A) evolução do número de trabalhos sob os termos da pesquisa, categorizados de forma qualitativa. O número de vezes que cada espécie foi estudada e o respectivo marcador utilizado são apresentados para: B) trabalhos categorizados como Filogeografia; C) Genética de Populações; D) detecção de machos diploides. Detalhes em Apêndices. *Trabalhos envolvendo análises filogenéticas ou identificação molecular, não foram incluídos.

De modo geral, os estudos genéticos-populacionais em euglossíneos podem ser agrupados em três principais categorias: i) produção de ferramentas moleculares; ii) análise de estrutura genética em paisagens fragmentadas; iii) detecção de machos diploides. Estudos da primeira categoria se distinguem por apresentar a produção de ferramentas, especialmente a confecção de *primers* para análise de microssatélites, o principal marcador molecular utilizado para estudos genéticos em euglossíneos (Figura 3.25-C). Embora marcadores microssatélites tenham um alto potencial polimórfico devido ao seu caráter multialélico, conjuntos de locos devem ser descritos e primers específicos devem ser construídos antes da execução do estudo, o que torna muitas vezes estes estudos caros e laboriosos. Existem atualmente cinco conjuntos de locos microssatélites descritos para espécies de Euglossini, incluindo: *Eg. annectans* (PAXTON et al., 2009), *Eg. iopocila* (PENHA et al., 2015), *Eg. cordata* e *El. nigrita* (SOUZA et al., 2007), e *El. meriana* (LÓPEZ-URIBE et al., 2011). Estes locos apresentaram transferabilidade com polimorfismo para várias espécies do grupo e são frequentemente utilizados com tal propósito.

Na segunda categoria, encontram-se os estudos relacionados à Genética de Populações e Genética da Conservação em suas versões clássicas. De forma geral, estes estudos objetivam identificar padrões de estruturação genética em paisagens fragmentadas ou entre ilhas e continente (BOFF et al., 2014; FREIRIA et al., 2012; ROCHA FILHO et al., 2013; SUNI, 2017; SUNI; BROSI, 2012). Os resultados destes estudos podem ser condensados em algumas linhas gerais: i) Microssatélites tendem a revelar baixa estruturação genética, possivelmente devido ao alto potencial de dispersão dos machos e frequentemente conseguem manter algum fluxo gênico mesmo entre paisagens fragmentadas por matrizes desfavoráveis (FREIRIA et al., 2012; PENHA et al., 2015; SORO et al., 2017; SUNI; BROSI, 2012; ZIMMERMANN et al., 2011); ii) a estruturação genética do DNA nuclear em euglossíneos, quando presente, está em baixos níveis e responde mais à padrões de isolamento por distância do que à possíveis barreiras (BOFF et al., 2014; PENHA et al., 2015; ROCHA FILHO et al., 2013; SUNI; BROSI, 2012); iii) ilhas podem diminuir a diversidade genética, mas apenas apresentam estruturação quando as populações estão distantes 10 km ou mais (BOFF et al., 2014; ROCHA FILHO et al., 2013; SUNI; BROSI, 2012). Embora generalizações sempre incorram em algum erro, aparentemente poucas barreiras são capazes de limitar a alta capacidade de dispersão dos euglossíneos, e via de regra, a restrição ecológica/capacidade de dispersão de cada espécie parece ser o fator limitante para a manutenção do fluxo gênico destas abelhas.

Por fim, a terceira linha de estudos relacionados à diversidade genética de euglossíneos se concentra na detecção de machos diploides. Também com principal foco na conservação das populações, estes estudos buscam identificar a ocorrência de machos diploides em populações naturais e aquelas sob influência antrópica, bem como buscam padrões de produção de machos diploides (*diploid male production*, DMP) entre diferentes espécies. Como discutido no tópico Diversidade, Genética e Conservação, a baixa diversidade genética no loco de determinação sexual dos himenópteros pode acarretar em um aumento na produção de machos diploides, os quais são inaptos à reprodução, e conseqüentemente levaria a população a um declínio populacional (ZAYED, 2009; ZAYED; ROUBIK; PACKER, 2004).

Estudos nessa linha trazem alguma controvérsia em relação aos resultados encontrados desde o início das análises e serão mais detalhadamente abordados. As primeiras análises nessa linha foram baseadas no polimorfismo de sistemas isoenzimáticos, as quais indicavam DMP na ordem de 15% à 50% em algumas populações da América Central (ROUBIK; WEIGT; BONILLA, 1996), 0-30% em populações Colombianas (LÓPEZ-URIBE; ALMANZA; ORDOÑEZ, 2007), outros estudos indicavam proporções baixíssimas nas populações brasileiras, inferiores à 0.5% (TAKAHASHI et al., 2001). A mesma variabilidade de DMP foi observada com dados de microssatélites, que segundo alguns autores produz taxas inferiores à 1% (SOUZA et al., 2010), mas de acordo com outros autores pode produzir taxas próximas à 11% em algumas espécies, como *E. annectans* (GIANGARELLI et al., 2015).

A discrepância destes resultados chama a atenção e merecem algumas considerações, uma vez que devem ter origens distintas. No que tange às análises isoenzimáticas, não é possível excluir artefatos da técnica, como desvios eletroforéticos ou mesmo a expressão de formas alternativas do sistema enzimático em questão devido a diferentes condições ambientais. Por outro lado, as análises de microssatélites são reconhecidamente mais estáveis e sensíveis (ARIF; KHAN, 2009) e, neste caso, a diferença entre os resultados de Souza et al. (2010) e Giangarelli et al. (2015) pode estar associada, entre outros fatores, às diferentes estratégias de amostragem adotadas nestes dois trabalhos. Embora Souza et al. (2010) tenham analisado 1010 indivíduos, estes autores priorizaram o número de espécies (27 espécies), porém algumas delas apresentaram uma representação reduzida de indivíduos genotipados, apenas um ou dois. Por outro lado, Giangarelli et al. (2015) analisaram 1457 indivíduos de seis espécies, abrangendo uma maior distribuição das espécies em diferentes subpopulações com 45 ± 2.47 (exceto *Eg. truncata*, $\bar{x} = 10 \pm 2,88$; cálculos não apresentados naquele estudo).

Embora os resultados encontrados por Giangarelli et al. (2015) sejam ainda baixos em relação às isoenzimas (maior DMP por espécie = 11,3%), estes são significativamente maiores do que o número encontrado por Souza et al. (2010), DMP < 1% (limite superior do intervalo de confiança). É necessário ponderar que outros dois estudos com estratégias similares à utilizada em Giangarelli et al. (2015) reportaram baixos índices de DMP em *Euglossa cordata* (BOFF et al., 2014) e *Euglossa viridissima* (SORO et al., 2017), corroborando o padrão observado por Souza et al. (2010). Portanto, a detecção de machos diploides parece depender não apenas da estratégia utilizada, mas também da espécie em questão. Os próprios dados apresentados por Giangarelli et al. (2015) indicam essa tendência, uma vez que apenas quatro das seis espécies analisadas naquele estudo apresentaram machos diploides e, dentre estas, a DMP variou entre 0,1% em *El. cingulata* e 11,3% em *Eg. annectans*. Além disso, a DMP pode variar fortemente entre populações, saindo de entre 0%-5% em *El. cingulata* e 8%-11% em *Eg. annectans* (GIANGARELLI et al., 2015). Vale ressaltar que os dados de Giangarelli et al. (2015) estão em consonância com dados isoenzimáticos de López-Uribe et al. (2007), que observaram uma grande variação (0 e 30%) de DMP em diferentes espécies na Colômbia.

Deste modo, a detecção de machos diploides em euglossíneos apresenta dois fortes padrões. O primeiro diz respeito à variabilidade de DMP de acordo com a população e espécies analisadas, bem como a ferramenta empregada para a detecção dos machos diploides. O segundo está relacionado com a forte variação de DMP de acordo com a estratégia utilizada para a detecção destes. Levando em consideração que a probabilidade de se encontrar um macho diploide em uma espécie é diretamente proporcional ao número de indivíduos analisados, a estratégia utilizada por Boff et al. (2014), Giangarelli et al. (2015) e Soro (2017), parece mais factível e, portanto, seus resultados devem representar níveis mais próximos a realidade em relação à DMP em euglossíneos.

A generalização dos padrões genéticos em Euglossini é especialmente perigosa, dada a alta variabilidade de parâmetros encontrados. Ainda assim, até o presente, os estudos genéticos-populacionais em euglossíneos mostram que estas abelhas apresentam uma grande capacidade de dispersão e seus parâmetros genéticos devem estar ligados a especificidades ecológicas nem sempre conhecidas. Deste modo, o isolamento por distância e restrições ecológicas específicas parecem ser os principais fatores de estruturação genética até o momento.

3.5 O CASO DE *EUGLOSSA IOPOECILA* DRESSLER

Euglossa iopoecila Dressler, 1982 é uma abelha de médio porte em relação às outras abelhas do gênero *Euglossa* (~17 mm), caracterizada por apresentar glossa bastante longa, ultrapassando o tamanho do corpo (1,5 – 2mm), além de exibir um tegumento azul ou verde metálico-iridescente e relativamente pubescente (DRESSLER, 1982b). *Euglossa iopoecila* apresenta distribuição fortemente associada à floresta ombrófila densa ao longo da costa brasileira entre os estados de Santa Catarina e Paraíba (DURAN CORDEIRO et al., 2013; ROCHA-FILHO; GARÓFALO, 2013; SYDNEY; GONÇALVES; FARIA, 2010).

Essa espécie tem sido foco de um debate taxonômico desde o início da década passada devido à sua grande similaridade com *E. chalybeata* e também por apresentar uma variação de coloração associada à distribuição geográfica (FARIA JR; MELO, 2007; NEMÉSIO, 2009; ROUBIK, 2004). Ainda que de forma não intencional, *E. iopoecila* acabou se tornando um modelo de estudos dentre os euglossíneos em decorrência dos múltiplos estudos relacionados à sua taxonomia e aspectos genéticos. Poucas espécies foram tantas vezes sistematicamente acessadas em estudos detalhados sobre os limites de espécie e estrutura filogeográfica (DRESSLER, 1982b; FARIA JR; MELO, 2007; FERRARI; MELO, 2014; FRANTINE-SILVA et al., 2017; NEMÉSIO, 2009; PENHA et al., 2015). Sem dúvida, a história biológica de *E. iopoecila* constituiu um dos mais belos e complexos exemplos filogeográficos da tribo Euglossini.

Quando Dressler (1982b) revisou o grupo *cybelia*, *analys* e *piliventris*, o mesmo descreveu “*E. chalybeata iopoecila*” como uma subespécie de *E. chalybeata* devido à sua coloração azul-violeta, ambas pertencentes ao subgênero *Glossura*, grupo *piliventris*. Naquela oportunidade, Dressler havia analisado apenas amostras do Paraná e São Paulo (holótipo: Alexandra, PR), definindo sua distribuição apenas para estes estados. Nas palavras de Dressler (1982b) sobre *E. iopoecila*:

[...] Esta bela abelha distribuiu-se de São Paulo ao Paraná, conforme indicado pelo material em coleções brasileiras. Eu não encontrei uma clara diferença estrutural entre esta forma e *E. chalybeata* da região Amazônica, e assim tratá-la-ei como uma subespécie de *E. chalybeata*. Esta população é distinguida pela coloração escura, azul-violeta, a que o epíteto se refere. [...]

Anos depois, ao avaliar várias espécies fortemente relacionadas dos subgêneros *Glossura* e *Glossuropoda*, Roubik (2004) elevou *E. iopoecila* ao status de espécie baseado em modificações morfológicas consistentes entre as formas da região Amazônica e Atlântica, definindo estas duas últimas regiões como áreas limites de ocorrência das respectivas formas (ROUBIK, 2004). Este autor aponta a diferença de coloração, mas principalmente a proporção do coxim anterior 1/3 do tamanho do coxim basal, como diagnose de *E. iopoecila* em relação a *E. chalybeata* (Figura 3.26). Adicionalmente a distribuição de *E. iopoecila* é reinterpretada por Roubik (2004) para a porção norte da Mata Atlântica, estados do Rio de Janeiro, Bahia e Paraíba, as quais abrigam formas verdes de *E. iopoecila*.

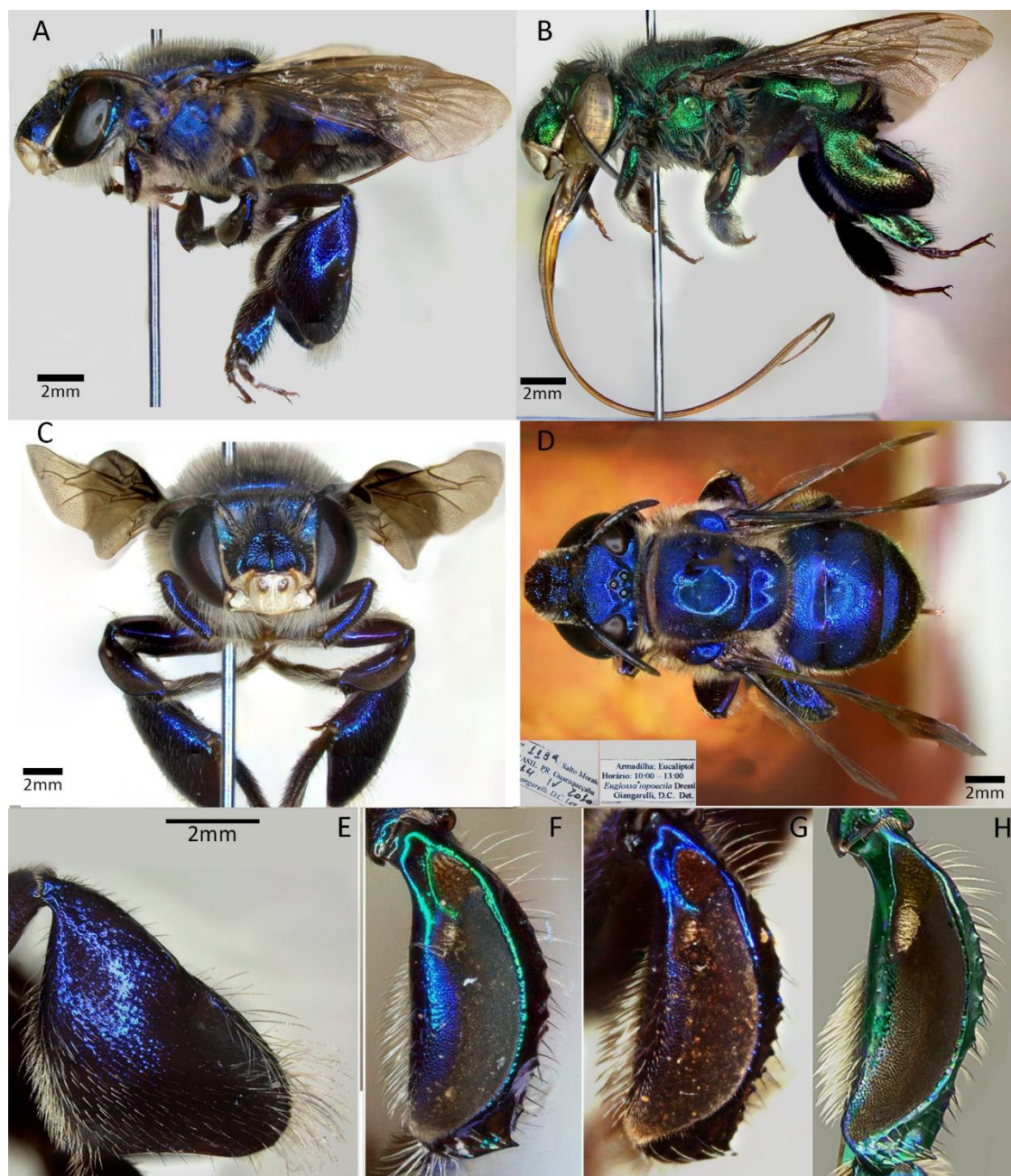


Figura 3.26 *Euglossa iopoecila* Dressler (1982). A- visão lateral, morfotipo azul (coleção de abelhas MZUEL-1189, Salto Morato, 14-IV-2010; *Giangarelli, DG. leg.*); B- visão lateral; morfotipo verde (WFS0012-PARNA Descobrimto; 15-III-2015, *Frantine-Silva, W. leg.*); C) visão frontal; morfotipo azul D) visão superior; morfotipo azul; E) tibia posterior esquerda; F) tibia média esquerda; morfotipo verde; G) morfotipo azul; Destaque para coxim anterior $\sim 1/3$ do tamanho do coxim basal em *E. iopoecila*. H) tibia média em *E. chalybeata*, fonte: D. W. Roubik.

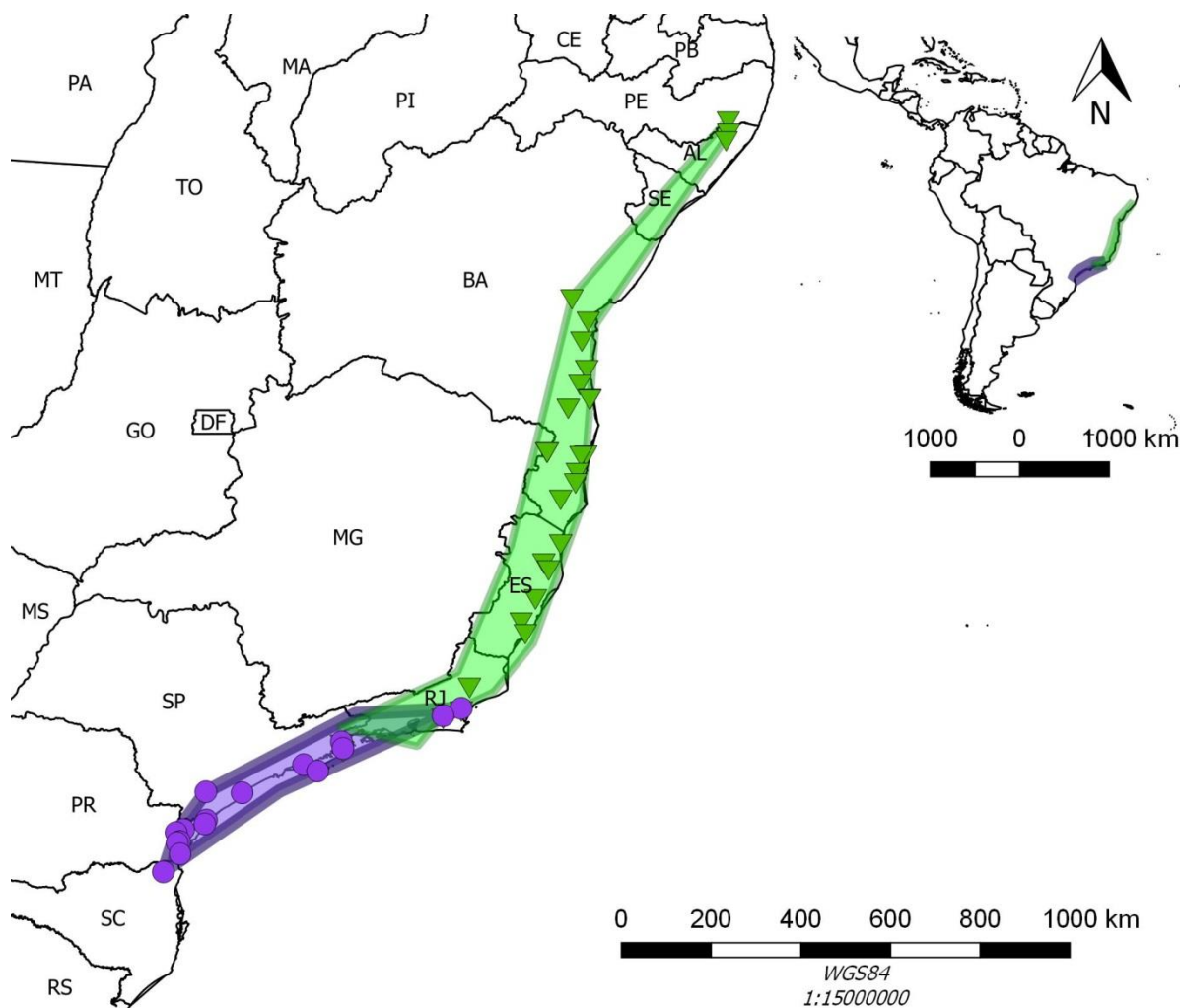


Figura 3.27 Distribuição e áreas de ocorrência de *Euglossa iopoecila* Dressler (1982) e seus respectivos morfotipos azul (ao sul; círculos roxos) e verde (ao norte; triângulos verdes) da distribuição. Registros aferidos na literatura, coleções e presente estudo (Apêndices).

Posteriormente Faria Jr. e Melo (2007) revisando espécies de *Glossura* da Mata Atlântica, incluem novos caracteres que distinguem *E. iopoecila* de *E. chalybeata*. Estes autores ressaltam i) a presença de pontuações esparsas sobre a margem posterior do escutelo (próxima à depressão escutelar); ii) margem posterior de T4 e T5 homogêneas, sem pontuações como em *E. chalybeata*; iii) cerdas plumosas mais densas sobre o mesoepisterno em *E. iopoecila*. Por fim, Faria Jr. e Melo (2007) comentam que a variação da coloração tegumentar em *E. iopoecila* (mais azuis entre SC e SP; mais verdes entre RJ e PB; com intermediários entre SP e RJ) compõe um gradiente latitudinal e foram interpretadas apenas como variações, exatamente por serem encontrados fenótipos intermediários na zona de contato entre os morfotipos azul e verde.

Em 2009, Nemésio realiza uma revisão de euglossíneos da Mata Atlântica e propõe a distinção entre os morfotipos verdes e azuis, atribuindo o nome *Euglossa roubiki* (Nemésio, 2009) para a forma verde e mantendo o epíteto *iopoecila* para as formas azuis. Embora não

tenha levantado outras características estruturais contundentes, a exemplo de Roubik (2004) e Faria Jr. e Melo (2007), Nemésio (2009) argumenta que as diferenças de coloração são caracteres morfológicos como quaisquer outros e, portanto, devem ser utilizados como tal. Como discutido a seguir (item 3.7 O caso do grupo *crassipunctata*), há uma diferença na visão da origem da variação de cor nos euglossíneos, mais especificamente no caso das espécies da Mata atlântica. Contudo, este tema não é foco do presente estudo e gera uma discussão circular, visto que ainda não há dados contundentes determinando o mecanismo molecular e/ou fisiológico de variação de coloração dos euglossíneos.

Felizmente, para o caso de *E. iopoecila* existem estudos moleculares que dão base à decisão taxonômica. Ferrari e Melo (2014) analisaram não apenas o caso de *E. iopoecila*, mas de outras duas espécies, como já mencionado anteriormente, não encontrando diferenças nas distâncias genéticas ou agrupamentos Bayesianos que sustentassem o status de espécies para os morfotipos azuis e verdes. Quase que concomitantemente, Penha et al. (2015 *in press*, mas 2014 *online*) realizaram uma análise do DNA mitocondrial e nuclear de subpopulações da região sul da distribuição de *E. iopoecila*, incluindo amostras do Paraná (Salto Morato - PR) até o Rio de Janeiro (REBIO União, Rio das Ostras - RJ), à época ainda consideradas *E. iopoecila*, com possíveis híbridos no Rio de Janeiro. A rede de haplótipos apresentada por Penha et al. (2015) apresenta claramente uma alta relação entre os haplótipos encontrados no Rio de Janeiro e aqueles encontrados em São Paulo, corroborando os resultados de Ferrari e Melo (2014).

Ainda mais contundente, os resultados das análises do DNA nuclear apresentados por Penha et al. (2015), indicam a presença de fluxo gênico entre as populações estudadas. Os índices de estruturação genética são todos compatíveis com diferenças esperadas entre populações e não entre espécies diferentes (BOFF et al., 2014; FREIRIA et al., 2012; ROCHA FILHO et al., 2013; ZIMMERMANN et al., 2011). Interessantemente, estes autores encontraram níveis de estruturação menores entre os morfotipos verdes (REBIO União - RJ) e os azuis (Ubatuba - SP), do que aqueles encontrados entre subpopulações ao sul do Paraná e o norte de São Paulo. Deste modo, fica claro que não é possível assumir o isolamento reprodutivo entre os morfotipos azul e verde e que as diferenças genéticas entre eles são compatíveis com o esperado entre populações da mesma espécie. Assim, os dados genéticos encontrados por Penha et al. (2015) e Ferrari e Melo (2014) apoiam indubitavelmente a sinonímia de *Euglossa roubiki* (Nemésio, 2009) sob *Euglossa iopoecila* Dressler, 1982.

Assumindo então a sinonímia de *E. roubiki* sob *E. iopoecila*, Frantine-Silva et al. (2017) analisaram a estrutura filogeográfica desta espécie ao longo de toda sua distribuição,

avaliando a influência da instabilidade climática do Quaternário, anteriormente sugerida por Penha et al. (2015) como possível fonte de divergência entre as subpopulações do sul e aquelas mais ao norte da distribuição da espécie. Embora os resultados apresentados em Frantine-Silva et al. (2017) sejam produto da presente tese e estejam detalhadamente descritos no Capítulo 3, ao menos um ponto proveniente destes resultados merece atenção no presente contexto, pois inclui informações pertinentes à questão do status taxonômico de *E. iopoecila*, a existência de múltiplas linhagens mitocondriais dentre as subpopulações de *E. iopoecila*. Frantine-Silva e colaboradores (2017) encontraram três linhagens ao longo da distribuição de *E. iopoecila*, das quais duas sobrepõem a distribuição dos morfotipos azul e verde, e a terceira encontra-se integralmente dentro da área de distribuição do morfotipo verde. Contudo, os níveis de diferenciação entre a “linhagem azul” e “verde” não parecem ser significativamente maiores do que aqueles encontrados entre as duas linhagens pertencentes ao morfotipo verde. Um outro ponto que merece destaque é que, assim como em Penha et al. (2015), o número de passos mutacionais nas redes de haplótipos são consideravelmente pequenos (3 passos entre *cluster* sul e central) mesmo analisando mais de 2000 pb. Note também que o *cluster* norte (Alagoas) se conecta antes ao *cluster* sul do que ao *cluster* central, o que também reflete um aspecto interessante de proximidade entre as três estruturas. Deste modo, os dados genéticos encontrados nos estudos anteriormente citados não suportam a hipótese de duas espécies, pelo contrário, apoiam a idéia de unicidade taxonômica entre *E. iopoecila* e *E. roubik* já que a diferença entre as linhagens genéticas é de poucos paços mutacionais.

Ainda assim, recentemente Ferrari, Nemésio e Silveira (2017) em uma análise filogenética entre espécies do gênero *Euglossa*, recuperaram *E. roubiki* e *E. iopoecila* como espécies distintas, porém vários aspectos deste estudo merecem esclarecimentos a este respeito. Primeiramente, os autores basearam suas análises na combinação de dados morfológicos e dados moleculares minados de Ramirez et al. (2010), à exceção do gene *COI* para *E. cyanochlora* obtida de Nemésio et al. (2012). Mais precisamente os autores utilizaram três genes nucleares (Arginina Quinase, AK; Fator de alongamento 1 α , Fl α ; e RNA polimerase II) e um gene mitocondrial (*COI*) todos gerados por Ramírez et al. (2010), mais dados morfológicos. No entanto, Ramírez et al. (2010) nunca geraram dados moleculares para *E. iopoecila* ou *E. roubiki*, contendo apenas dados de *E. chalybeata* como espécie mais relacionada às duas anteriores. Em outras palavras, os resultados apresentados em Ferrari, Nemésio e Silveira (2017) para *E. iopoecila* e seu sinônimo júnior, *E. roubiki*, não apresentam quaisquer bases moleculares.

Vale ressaltar ainda que devido à natureza dos dados de Ferrari, Nemésio e Silveira (2017), as árvores apresentadas são consensuais e possuem caráter apenas topológico, ou seja, não guardam informação de distância entre os táxons terminais. Logo, a magnitude de diferenciação entre estes táxons é graficamente relativizada, fazendo com que diferenças grandes ou pequenas sejam apresentadas em escalas similares. Por fim, os autores afirmam que os dados de Penha et al. (2015) sustentam suas conclusões. No entanto, como discutido anteriormente, essa interpretação dos dados de Penha et al. (2015) merece algumas ressalvas, visto que há uma clara hibridização entre os morfotipos verdes e azuis (“*E. roubiki* e *E. iopoecila*” *sensu* Nemésio, 2009). De fato, deve ser ponderado que o objetivo de Ferrari, Nemésio e Silveira (2017) não era testar os limites de espécie em *E. iopoecila*. Contudo, a interpretação das formas verde e azul como unidades taxonômicas distintas não pode ser sustentada enquanto há evidências de fluxo gênico nuclear ocorrendo naturalmente entre as duas formas e/ou há ausência de divergência genética suficiente para inferir completa separação entre as formas.

Do ponto de vista filogeográfico *E. iopoecila* torna-se extremamente interessante devido diversos fatores, dentre eles: a) sua ampla distribuição latitudinal; b) a existência de morfotipos diferentes distribuídos de forma coesa geograficamente; c) indícios de zonas híbridas na região de contato entre os morfotipos; d) seu forte endemismo à região de Mata Atlântica. Frente ao exposto, o estudo da estrutura filogeográfica de *E. iopoecila* apresenta grande potencial para elucidar diversas questões relacionadas à história filogeográfica dos euglossíneos.

3.6 O CASO DE *EUGLOSSA ANNECTANS* DRESSLER

No mesmo artigo em que Dressler descreve *E. iopoecila*, ele também descreve duas outras espécies de grupos distintos, uma das quais é *Euglossa annectans*. Esta espécie não possuía uma associação claramente definida com as demais espécies do subgênero *Glossura*, pelo contrário, possuía características de transição entre os subgêneros *Glossura* e *Euglossa*, o que motivou Dressler a escolher o epíteto “*annectans*”. *Euglossa annectans* pode ser resumidamente descrita como uma abelha de médio porte (dentre os seus congêneres, ~12 mm), de cor verde com traços dourados no abdômen e clipeo azul-esverdeado nos machos; tórax com cerdas plumosas, densas e longas; glossa longa, cerca de 1mm menor do que o corpo; coxim dividido, sendo o anterior duas vezes mais longo do que o basal (Figura 3.28).

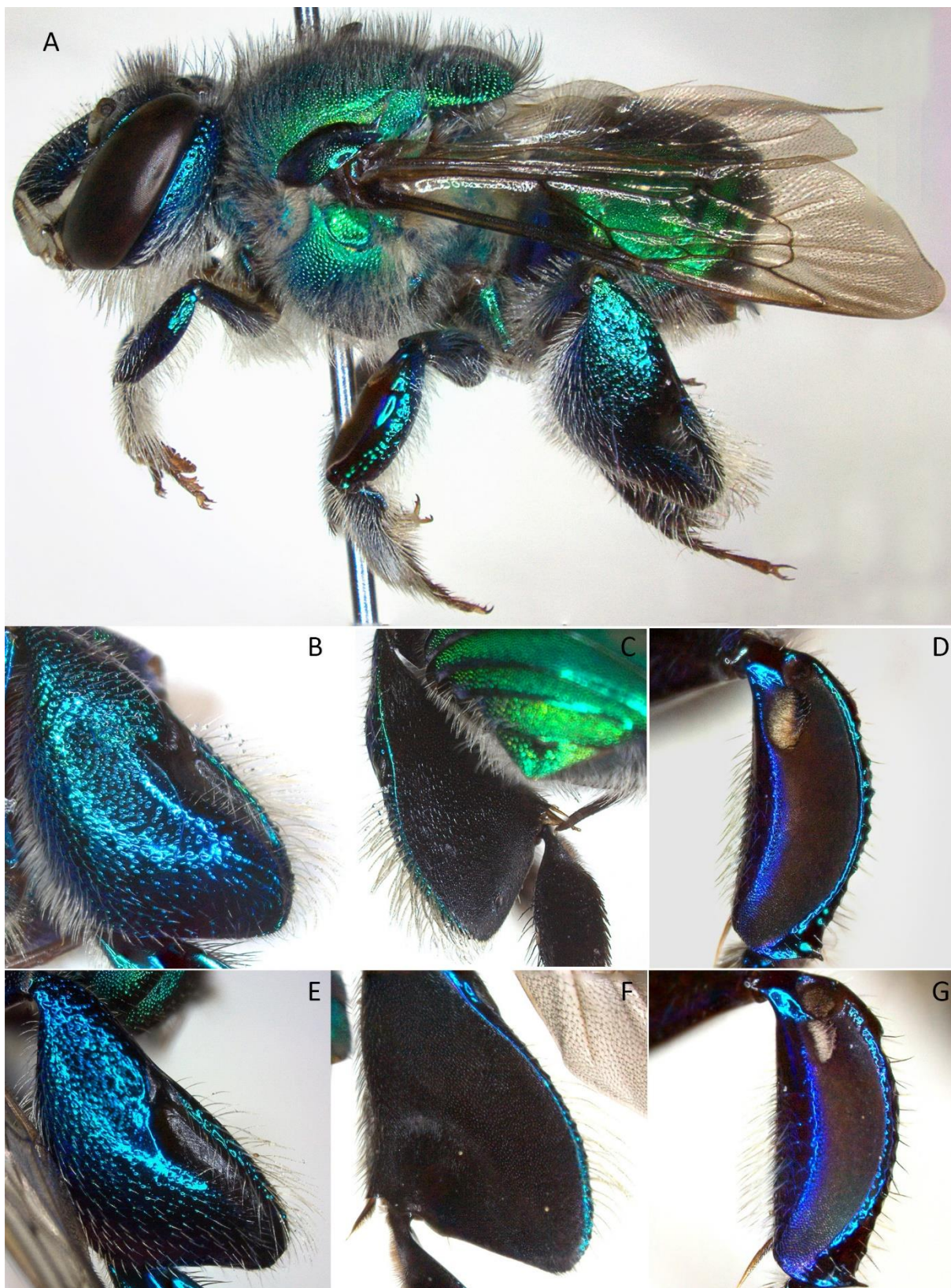


Figura 3.28 *Euglossa annectans* Dressler 1982 (A, B, C e D) e comparação de traços morfológicos em relação a espécie mais relacionada *Euglossa stellfeldi* (E, F, G). *E. annectans*: A) visão lateral; B) face externa da tíbia posterior, detalhe para oclusão basal mais afilada da fenda metatibial; C) face interna metatibial; D) visão lateral da mesotíbia, destaque para a proporção dos coxins em relação à *E. stellfeldi*; *E. stellfeldi*: E) face externa metatibial; F) face interna metatibial, detalhe para a concavidade na região basal; G) visão lateral da mesotíbia, detalhe para a diferença na proporção dos coxins em relação à *E. annectans*.

No momento de sua descrição, Dressler chama a atenção para o fato de *E. annectans* ser altamente relacionada a *Euglossa stellfeldi* Moure (1941), as quais juntas comporiam um grupo bastante distinto dentro do subgênero *Glossura*, o grupo “*annectans*”, com apenas as duas espécies. Dressler salienta as seguintes diferenças entre *E. annectans* e *E. stellfeldi*, em uma tradução livre:

[...] *Euglossa annectans* difere de *E. stellfeldi* por seu maior tamanho, pelo clípeo azul nos machos, pela margem posterior do escutelo retilínea ao invés de superficialmente côncava, e pelo coxim posterior muito maior na tíbia média [...]

Embora as expectativas de Dressler fossem que o grupo *annectans* aumentasse em número de espécie com o decorrer dos anos, os *checklists* subsequentes mantiveram a posição isolada dessas duas espécies em relação às demais espécies do subgênero (RAMIREZ; DRESSLER; OSPINA, 2002; ROUBIK; HANSON, 2004).

A partir de 2004, o status taxonômico de *E. annectans* passou por certa polêmica e merece alguns comentários para reunir toda a discussão a este respeito em um único texto. Fragmentos dessa discussão podem ser encontrados em Faria Jr. e Melo (2007) e Nemésio (2009). Obviamente há um interesse particular de estabelecer uma indubitável unicidade taxonômica em *E. annectans* devido à imprescindível necessidade de certeza taxonômica antes de se prosseguir com estudos de natureza genético-populacional e filogeográfica.

Como visto anteriormente, os caracteres que distinguem as formas *E. stellfeldi* e *E. annectans* são bastante contundentes, porém uma série de decisões taxonômicas se seguiram e podem deixar confusa a literatura acerca de *E. annectans*. Em 2004, um trabalho realizado no Parque Estadual do Rio Preto, MG, sinonimiza *E. annectans* sob *E. stellfeldi* (NEMÉSIO; FARIA JR., 2004). No entanto, dois anos depois, o primeiro autor reconsidera a validade de *E. annectans* baseado na análise de um novo material que tivera sido coletado a 50 km da localidade tipo de *E. stellfeldi* (NEMÉSIO; SILVEIRA, 2006a). Até então, *E. annectans* e *E. stellfeldi* eram consideradas integrantes do subgênero *Glossura*, ainda que tivessem um caráter bastante distinto em relação às outras espécies do subgênero.

No entanto, no ano seguinte, Faria Jr. e Melo (2007) analisaram as espécies do subgênero *Glossura* da Mata Atlântica. Ao abordar novamente a questão de *E. annectans* e *E. stellfeldi*, estes autores ratificaram a distinção entre as espécies e realocaram *E. stellfeldi* em *Glossurella* devido às diferenças morfológicas encontradas. Faria Jr. e Melo (2007) apontam,

dentre outros caracteres, as seguintes diferenças entre *E. annectans* e *E. stellfeldi*: i) a ausência de oclusão da fenda metatibial alcançando a margem inferior da metatíbia em *E. stellfeldi* (Figura 3.28-E), em contraste a *E. annectans* (Figura 3.28-B) que apresenta um afunilamento bem demarcado na porção distal da fenda metatibial antes de alcançar a margem inferior da metatíbia; e ii) a concavidade na face interna da metatíbia de *E. stellfeldi* (Figura 3.28-F), em relação à superfície plana da face interna da metatíbia de *E. annectans* (Figura 3.28-C). Estes caracteres ao lado da diferença na proporção dos coxins de ambas as espécies tornam bastante clara a diferenciação entre as duas formas.

Até este ponto, não restava qualquer dúvida sobre a validade de ambas as espécies. Contudo, Nemésio (2009) faz uma longa discussão apontando uma possível confusão entre os holótipos utilizados para a descrição de *E. annectans* (veja Nemésio, 2009: *The Euglossa stellfeldi problem*). Neste trabalho, este autor justifica a sinonímia de *E. annectans* sob *E. stellfeldi*, reservando a forma correspondente a *E. stellfeldi* os nomes *Euglossa carinilabris* Dressler (morfortipo verde) e *Euglossa solangeae* Nemésio (morfortipo azul)⁶. Subsequentemente Ferrari e Melo (2014) demonstram a falta de diferenciação genética entre espécies com diferentes morfotipos de coloração, como *E. iopoecila* (anteriormente comentado), *E. stellfeldi* e *E. townsendi*. Os mesmos autores salientam que Nemésio (2009) teria aplicado erroneamente os nomes *E. carinilabris* e *E. solangeae* à espécie que na verdade seria *E. stellfeldi*, e aplicando o último nome à espécie que seria *E. annectans*.

Embora Ferrari e Melo (2014) não tenham dado detalhes do porquê Nemésio (2009) teria aplicado os nomes erroneamente, a falta de diferenciação genética dos morfotipos de *E. stellfeldi* depõe claramente a favor da decisão daqueles autores. A decisão de assumir *E. annectans* e *E. stellfeldi* como espécies válidas é corroborada por Nemésio recentemente em Ferrari, Nemésio e Silveira (2017). Abordando as relações filogenéticas entre espécies do gênero *Euglossa*, estes autores aceitam mais uma vez *E. annectans* como espécie válida, confirmando a posição distinta de *E. annectans* e *E. stellfeldi*, as quais são posicionadas como *Incertae sedis* dentro do gênero *Euglossa*.

Deste modo, com base nos resultados de Ferrari e Melo (2014), Ferrari, Nemésio e Silveira (2017) e a natureza cronológica das descrições de *Euglossa stellfeldi* Moure, 1947 e *Euglossa annectans* Dressler, 1982, parece bastante plausível apenas assumir *E. stellfeldi* e *E. annectans* (*sensu* Faria Jr. e Melo, 2007). Por fim, seguem as definições assumidas no

⁶ À exemplo de *E. iopoecila*, *E. stellfeldi* também apresenta dois morfotipos de coloração com distribuição semelhante a *E. iopoecila*. Como Ferrari e Melo (2014) demonstraram que estas formas não correspondem a espécies distintas, assume-se o nome *E. stellfeldi* como válido.

presente estudo em relação à interpretação de *E. annectans* e sua distribuição (Figura 3.29) ao longo da literatura: i) Os registros de *E. stellfeldi* em Nemésio e Faria Jr (2004), foram considerados como *E. annectans*; ii) O material tratado como *E. stellfeldi* em Nemésio (2009) foi também considerado *E. annectans*; iii) *E. carinilabris* e *E. solangeae* são tratados como sinônimos jr. de *E. stellfeldi*.

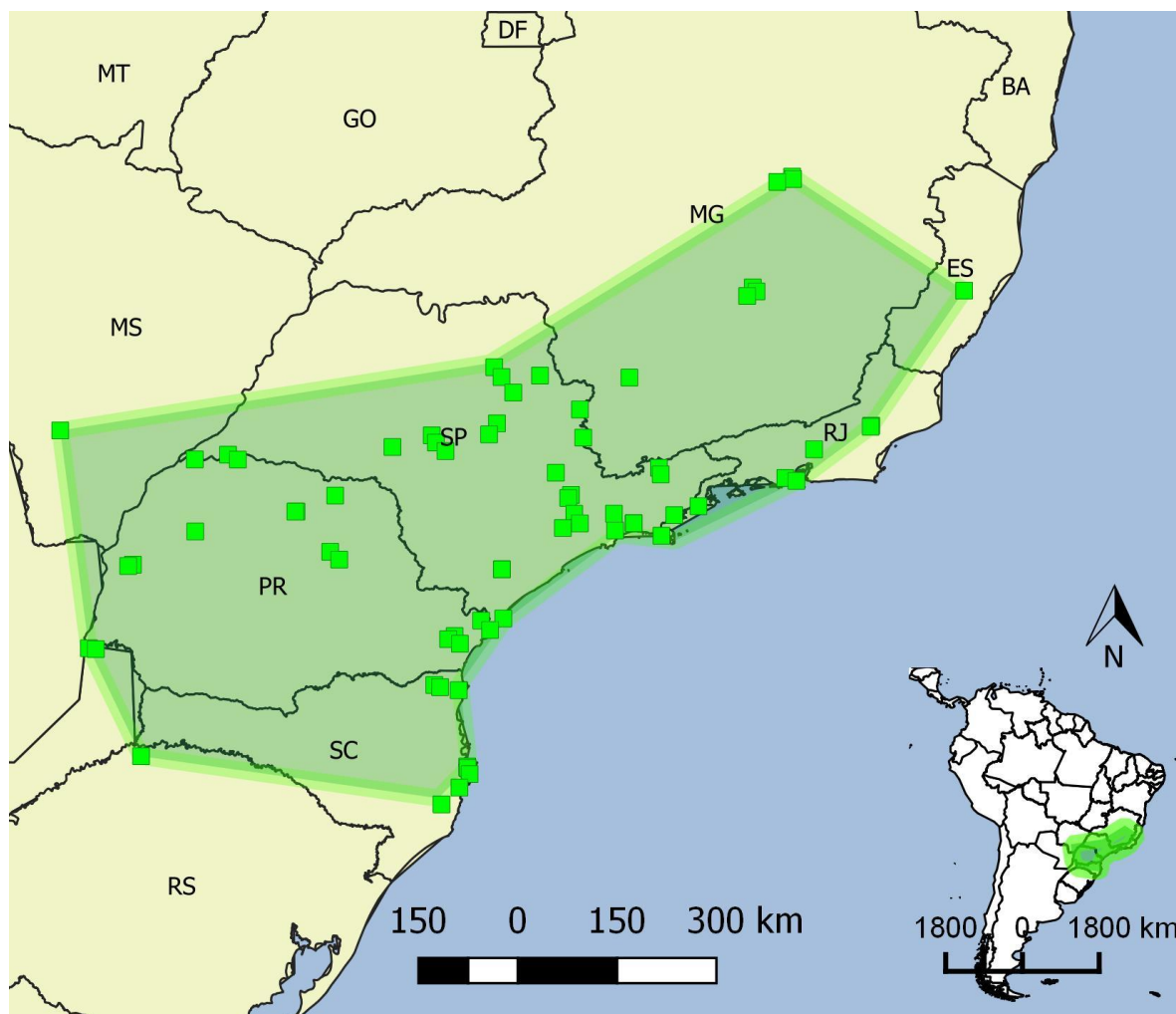


Figura 3.29 Distribuição de *Euglossa annectans*. Registros aferidos na literatura, coleções e presente estudo (Detalhes das origens dos registros podem ser encontrados nos Apêndices)

Frente ao exposto, pode-se assumir que *E. annectans* possui uma distribuição bastante associada com a Florestas Estacionais Semidecidual, alcançando longitudes tão altas quanto 55°O (Dourados-MS), porém com distribuição restrita a latitudes entre ~19°S e 28°S, tornando-a um modelo de distribuição bastante interessante e alternativo àqueles apresentados por espécies costeiras, como *E. iopoecila*.

E. annectans, assim como outros euglossíneos, possui um grande potencial de polinização. Apenas na Ilha de Santa Catarina, um estudo confirmou a presença de 74 tipos de polens pertencentes a 24 famílias, das quais Melastomataceae, Bromeliaceae, Ochnaceae, Fabaceae e Myrtaceae foram as mais comuns (CORTOPASSI-LAURINO; ZILLIKENS;

STEINER, 2009). Outro aspecto biológico interessante de *E. annectans* diz respeito à sua nidificação. Estudos apontam que *E. annectans* pode construir ninhos com participação de mais de uma fêmea (GARÓFALO et al., 1998) e que estes agregados podem chegar a números bastante grandes (CORTOPASSI-LAURINO et al., 2008). Estes mesmos autores reportaram a atividade em um ninho durante cinco anos, o qual chegou a ter ao menos 27 abelhas e 259 células, demonstrando que *E. annectans* pode compor grandes agregados de indivíduos ou manter ninhos quase solitários.

Embora *E. annectans* apresente esta plasticidade em termos de nidificação, a mesma é aparentemente uma espécie susceptível à produção de machos diploides. Como discutido anteriormente (3.4.4 Diversidade Genética: o estado da arte em Euglossini), himenópteros possuem um sistema de determinação sexual haplodiploide controlado por uma interação alélica heterozigótica, de modo que hemizigotos ou homozigotos dão origem à machos. Em outras palavras, níveis de diversidade genética baixos ou alto índice de endogamia poderiam levar ao aumento na produção de machos diploides, podendo ainda interferir na sobrevivência da espécie (ZAYED, 2009; ZAYED; ROUBIK; PACKER, 2004). Embora não se tenha estimativas da diversidade genética de um loco de determinação sexual tipo CSD em *E. annectans*, resultados de estimativas de machos diploides apontam que *E. annectans* pode atingir proporções entre 10% e 27% de machos diploides (GIANGARELLI et al., 2015).

Neste contexto, o estudo da diversidade genética faz-se urgente para esta espécie, visando investigar seus níveis de diversidade genética. Do mesmo modo, o estudo de sua história biogeográfica tem grande potencial para levantar hipóteses de padrões biogeográficos em euglossíneos de distribuição interiorana ao longo da Mata Atlântica, principalmente na Floresta Estacional Semidecidual.

3.7 O CASO DO GRUPO *CRASSIPUNCTATA*

Euglossa crassipunctata Moure 1968 e seus congêneres mais próximos (*E. sapphirina* Moure, 1968; *E. parvula* Dressler, 1982; *E. moratoi*, *E. pepei* e *E. clausi*, Nemésio e Engel, 2012) constituem um grupo de “espécies crípticas” (NEMÉSIO; ENGLE 2012), de difícil identificação (NEMÉSIO, 2009) e com tempo de divergência relativamente alto em relação ao ancestral comum mais recente (ACMR) dos demais membros do gênero *Euglossa* (~15-20 milhões de anos antes do presente, ap.), porém com diversificação muito mais recente para o ACMR do grupo *crassipunctata*, estimada entre 10 e 7 Ma a. p. (RAMÍREZ et al. 2010a).

Atualmente, todos os membros do grupo *crassipunctata* encontram-se alocados no subgênero *Glossurella*, movidos de *Glossura* juntamente com outras espécies do grupo

bursigera, desde que Dressler (1982) revisou a subdivisão de gêneros da tribo Euglossini (tratada neste estudo *sensu* Michener, 2000). Na primeira subdivisão, Dressler (1978) enquadrara *E. sapphirina* e *E. crassipunctata* no subgênero *Glossura*, o qual agrupa abelhas pequenas de tegumento brilhante e glossa longa ou muito longa; com coxim mesotibial anterior completo, ablongado ou ovalado; tíbia posterior triangular, usualmente aguda; dois ou três dentes mandibulares; tufos variados nos esternitos metassomais; ângulo pronatal obtuso; e tufos escutelares moderadamente longos (espécie tipo: *E. piliventris*).

Alguns anos depois, em Dressler (1982c), o mesmo autor descreve o subgênero *Glossurella* para realocar espécies relacionadas à *Euglossa bursigera*, caracterizadas por apresentar na tíbia média um coxim anterior completo combinado com coxim posterior ausente, presente e dividido em dois lobos, ou ainda, outros dois coxins completos; dois ou três dentes mandibulares; tufos no segundo esternito semicirculares presentes ou rudimentares. Dressler (1982) ainda acrescenta uma nota biológica à descrição do novo subgênero, indicando que a nidificação das espécies em *Glossurella* seria mais aérea do que nas espécies de *Glossura*. Entretanto, o mesmo autor salienta que a área de ocupação das espécies era pouco conhecida e que as espécies sul-americanas ainda não apresentavam fêmeas descritas para a maior parte das espécies, o que torna difícil a adição deste caráter como uma apomorfia em *Glossurella*.

No que tange especificamente ao grupo *crassipunctata*, a similaridade e monofilia do grupo podem ser observados já na descrição de suas primeiras espécies. De acordo com Nemésio (2009), quando Moure (1968) descreve *E. crassipunctata*, concomitantemente descreve *E. sapphirina*, ambas como espécies altamente similares e simpátricas na América Central (Canal da Zona, Gamboa, Panamá), porém com marcante diferença de coloração tegumentar, respectivamente verde (Figura 3.30-A, C, E) e azul (Figura 3.30-B, C, F).

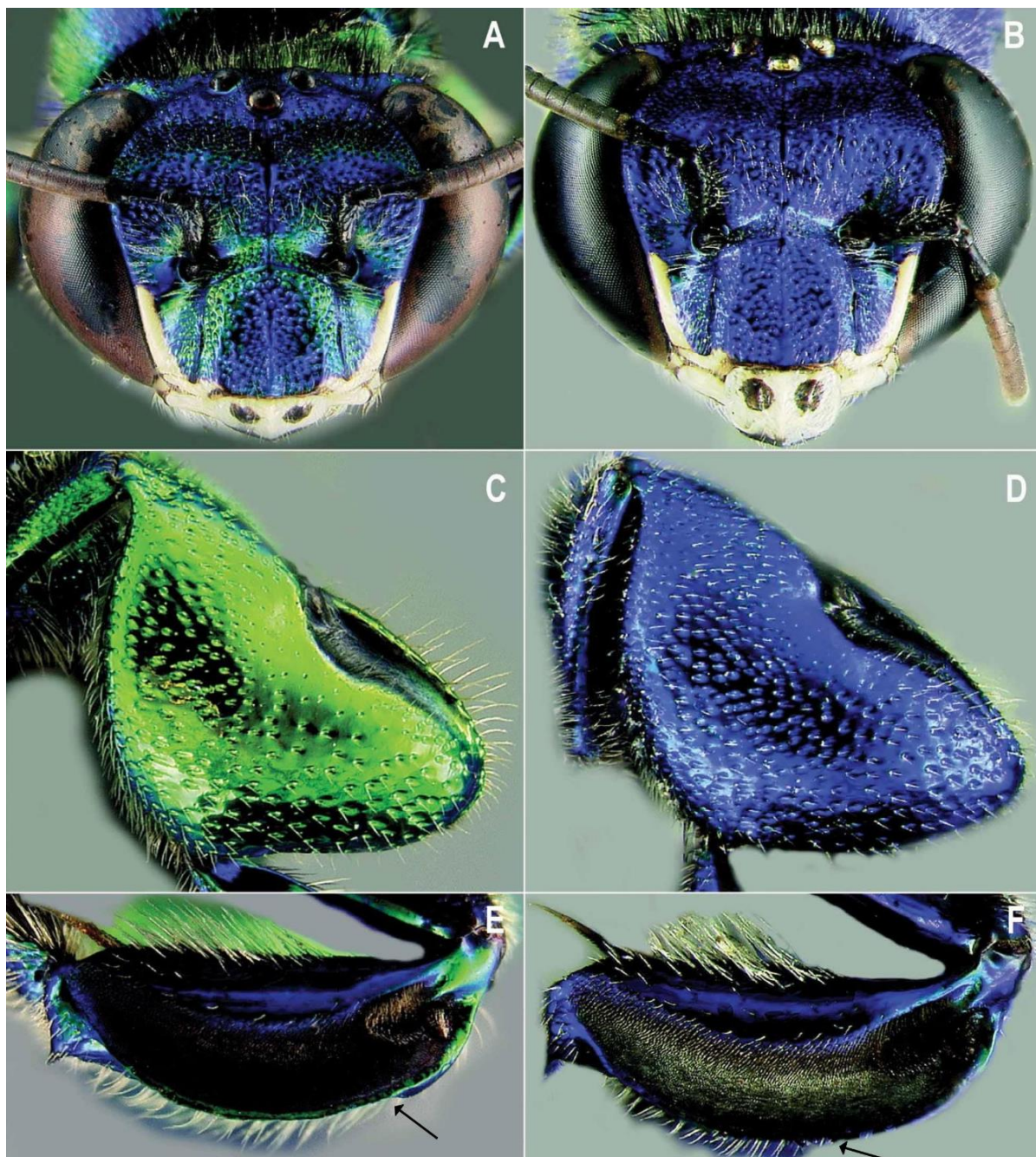


Figura 3.30 Comparação entre os holótipos de *Euglossa crassipunctata* (A, C e E) e *Euglossa sapphirina* (B, D e F). Setas pretas indicam a posição de possíveis caracteres diagnósticos entre as espécies (vide texto). Adaptada de Nemésio (2009), fonte original (David W. Roubik).

Esta alta similaridade também é observada por Dressler (1982) no momento da descrição do subgênero *Glossurella*, quando o autor formaliza a existência do grupo *crassipunctata* incluindo uma nova espécie, *E. parvula* Dressler (1982). Nas palavras do próprio autor: “[...] existem vários grupos de espécies mais ou menos reconhecíveis dentro de *Glossurella*. O mais distinto dentre estes é o grupo *crassipunctata* com *E. crassipunctata*, *E. parvula* e *E. sapphirina* [...]”.

O mesmo autor ainda define o grupo como abelhas de pequeno tamanho (~10 mm), clípio azul escuro, episterno com pontuações largas e um escutelo ainda mais pontuado. Desde a publicação de Dressler (1982) não houveram alterações na constituição do grupo *crassipunctata* e estudos de várias naturezas ampliaram a área de ocorrência destas espécies tanto para a América Central (ACKERMAN; APR; ACKERMAN, 1983; DRESSLER, 1978, 1982b; HINOJOSA-DÍAZ; ENGEL, 2012; JANZEN et al., 1982), quanto na Bacia Amazônica (OLIVEIRA; CAMPOS, 1996; STORCK-TONON; MORATO; OLIVEIRA, 2009) e Mata Atlântica (DURAN CORDEIRO et al., 2013; NEMÉSIO, 2009; NEMÉSIO; ENGEL, 2012; NEMÉSIO; SILVEIRA, 2006a).

Embora haja uma clara confluência entre as espécies do grupo *crassipunctata*, o status taxonômico de algumas delas vem sendo ponto de discussão ao longo dos últimos anos (NEMÉSIO, 2009). De fato, historicamente a descrição de algumas destas espécies dentre os euglossíneos é pouco contundente em relação as apomorfias que as distinguem. Este fato é observável no caso de *E. crassipunctata* e *E. sapphirina* ao analisar-se a chave de identificação para as espécies do grupo *Glossurella* de Dressler (1982). Nesta chave, o autor salienta a diferença de coloração como principal apomorfia para distinção entre *E. sapphirina* e *E. crassipunctata*. Embora posteriormente Roubik e Hanson (2004) também chamem a atenção para a pontuação mais densa no escutelo de *E. crassipunctata* em relação à *E. sapphirina*, permanece ainda um grande compartilhamento de caracteres (Figura 3.30), que por sua vez, ressalta a íntima relação entre as espécies do grupo.

Interessantemente, quando se analisa as características verificáveis entre os holótipos da Figura 3.30 (A, C e E: *E. crassipunctata*; B, D e F: *E. sapphirina*) pode-se observar outras diferenças morfológicas, que não apenas a coloração tegumentar. O clípeo levemente esverdeado em *E. crassipunctata* (Figura 3.30-A) é acompanhado de uma maior concentração de cerdas metatibiais combinados com um maior alargamento da porção terminal da fenda metatibial (Figura 3.30-C, D). Adicionalmente, as cerdas na face posterior da tibia média são mais densas em *E. crassipunctata* e tem início mais próximo à articulação femorotibial, alinhado com a linha basal do coxim anterior (Figura 3.30-E, seta preta). Já em *E. sapphirina* as cerdas mesotibiais são mais esparsas (Figura 3.30-F, seta preta), tornando-se mais densas na porção médio-basal. No entanto, é importante ressaltar que estes caracteres não puderam ser amplamente observados em diferentes parátipos da série tipo e, portanto, não é possível determinar a amplitude de variação intraespecífica dos mesmos.

No que tange à mais evidente apomorfia do grupo, o uso da coloração como caráter apomórfico e diagnóstico dentre as espécies do grupo é bastante controverso na literatura.

Embora este caráter já tenha sido utilizado como diagnose e apomorfia em algumas descrições, como no próprio caso de *E. sapphirina* e *E. crassipunctata*, ainda não há consenso sobre os fatores determinantes desta variação, se de influência ambiental ou genética, ou ainda um artefato do processo de secagem dos exemplares (veja NEMÉSIO, 2009; *red and blue phenomenon*).

Ambas ideias têm argumentos a seu favor, seja a congruência dos morfotipos e suas distribuições latitudinais ou a existência de linhagens genéticas relativamente congruentes com os morfotipos de cor (PENHA et al. 2015; FRANTINE-SILVA et al. 2017). Um terceiro argumento seria a existência de zonas híbridas com morfotipos intermediários que, por sua vez, poderia sugerir tanto uma zona ambiental intermediária quanto um sistema de interação alélica ou gênica, congruente com uma “hipótese genética”.

De modo geral, é bastante plausível que o fenótipo de coloração seja naturalmente resultado de uma interação ambiente e genótipo, alinhados ou não com isolamento reprodutivo. Esta linha de raciocínio ganha força ao considerar-se a transitoriedade da associação destas variações de coloração com a congruência de hipóteses de espécie válidas. Em alguns casos a diferença de coloração parece um caráter confiável, mesmo para espécies simpátricas, como mais uma vez no exemplo de *E. crassipunctata* e *E. sapphirina* na América Central. Por outro lado, várias outras hipóteses de espécies baseadas em diferenças de coloração não apresentam suporte quando analisadas a partir de dados genéticos (FARIA JR; MELO, 2007; FERRARI; MELO, 2014; FRANTINE-SILVA et al., 2017; PENHA et al., 2015). Exemplos recentes são descritos para *Euglossa townsendi* Cockerell, 1904, *Euglossa iopoecila* Dressler, 1982 e *Euglossa stellfeldi* Moure, 1947, onde as diferenças entre linhagens mitocondriais ou nucleares não são compatíveis com os níveis esperados entre espécies diferentes (FERRARI; MELO, 2014; FRANTINE-SILVA et al., 2017; PENHA et al., 2015). Vale ressaltar ainda que os estudos até o presente exploraram diferenças mitocondriais e apenas Penha et al. (2015) possui uma análise de DNA nuclear em uma área parcial de distribuição de *E. iopoecila*. Deste modo, esta questão estará em aberto até que estudos mais aprofundados elucidem os mecanismos moleculares de determinação de coloração tegumentar nos euglossíneos, para que então possa-se compreender os limites da influência genética e ambiental nestes padrões. Não obstante, o estudo sobre a biogeografia histórica dos euglossíneos tende também a contribuir com a elucidação desta questão em particular, visto que a mesma pode oferecer informações sobre eventos vicariantes e sua congruência com a hipótese de espécies.

É no âmbito biogeográfico que a discussão entorno do status taxonômico do grupo *crassipunctata* ganha ênfase. Até 2012 o grupo *crassipunctata* compreendia apenas três espécies, como dito anteriormente, *E. crassipunctata* e *E. sapphirina*, além de *E. parvula*. De acordo com os dados compilados a partir do Catálogo de Abelhas Moure (MOURE et al., 2012), há a hipótese de que *E. crassipunctata* seja uma espécie amplamente distribuída desde a América Central, cruzando a barreira dos Andes colombianos até a bacia Amazônica, isolada da porção centro-norte da Mata Atlântica pela diagonal árida (Caatinga, Cerrado e Chaco), ocorrendo de Pernambuco até o Rio de Janeiro. Ao mesmo tempo, *E. sapphirina* apresentar-se-ia como uma espécie de distribuição disjunta, restrita a América Central e, portanto, sem registros na Bacia Amazônica, mas voltando a ocorrer na porção sul da Mata Atlântica, entre a região de Ubatuba-SP e Salto Morato-PR (Figura 3.31-A).

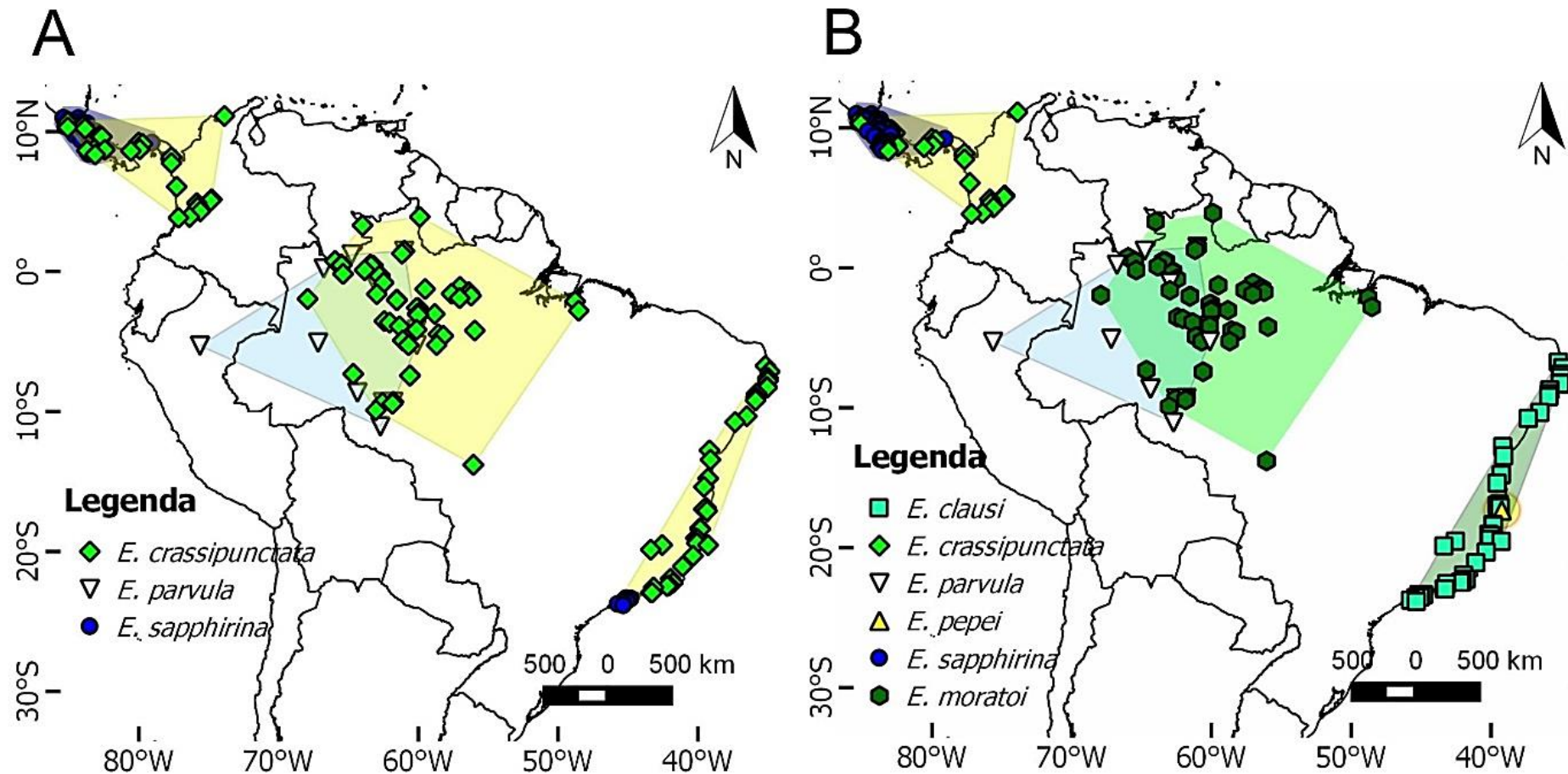


Figura 3.31 Ocorrência conhecida das espécies do grupo *crassipunctata* de acordo com duas hipóteses de espécies. A- De acordo com Moure et al. (2012); B - De acordo com Nemésio e Engel (2012). Em A, *E. sapphirina* tem distribuição disjunta, ocorrendo no sul da Mata Atlântica e apenas na América Central. Destaque em B para *E. pepei* com ocorrência pontual no sul da Bahia e *E. clausi* com distribuição em toda a Mata Atlântica.

Embora amplas distribuições, como a aceita por Moure et al. (2012) para e *E. crassipunctata*, sejam relativamente observadas para alguns grupos, a exemplo de espécies de pimenta da família Piperaceae (QUIJANO-ABRIL; CALLEJAS-POSADA; MIRANDA-ESQUIVEL, 2006), morcegos vespertilionídeos (LARSEN et al., 2012), aves Dendrocolaptidae (ROCHA et al., 2015) e mesmo euglossíneos (ver *E. townsendi*; FERRARI; MELO, 2014), a ocorrência de espécies com distribuição disjunta, como a resultante para *E. sapphirina*, é sempre desafiadora do ponto de vista biogeográfico.

Especificamente, para o grupo *crassipunctata*, Nemésio (2009) já salientara a baixa parcimônia da distribuição destas duas espécies, *E. crassipunctata* e *E. sapphirina*, como um ponto de destaque em relação ao *status* taxonômico deste grupo. Alguns anos depois, Nemésio e Engel (2012) sugeriram a divisão do grupo sustentando esta hipótese em diferenças na terminália das espécies, tamanho corporal e a própria distribuição geográfica⁷ (Figura 3.32).

⁷ Embora Nemésio e Engel (2012) descrevem uma série de modificações na angulação do gonóstilo e de processos relacionado ao oitavo esternito para todas as espécies descritas, apenas as modificações externas e independentes de dissecação foram abordadas.

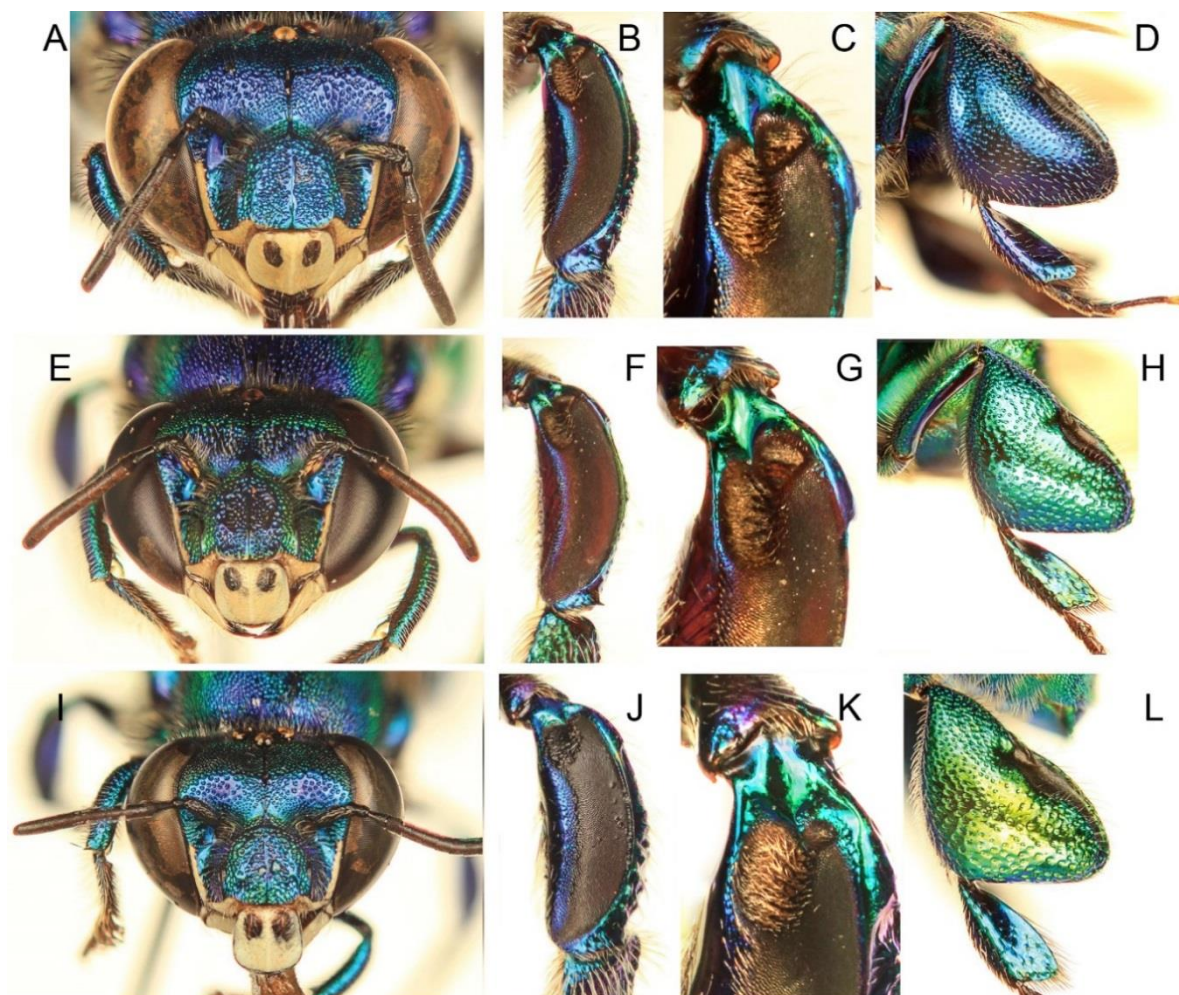


Figura 3.32 Visão frontal e características tibiais de *E. clausi* (A-D), *E. moratoi* (E-H) e *E. pepei* (I-L). Adaptado de Nemésio e Engel (2012).

De acordo com a hipótese de Nemésio e Engel (2012), haveria uma divisão entre os clados amazônicos e da Mata Atlântica (Figura 3.32, A-D; I-L) em três novas espécies: *E. clausi* e *E. pepei* para a Mata Atlântica, e *E. moratoi* para a Bacia Amazônica (Figura 3.32, E-H). Estes autores reafirmam a hipótese de Nemésio (2009) de que *E. sapphirina* e *E. crassipunctata* seriam uma única espécie na Mata Atlântica, *E. clausi*, exibindo dois morfotipos de coloração, mas sim compartilhariam uma única coloração tegumentar azul-esverdeada e o maior tamanho corporal do grupo *crassipunctata* (> 10mm). Já a espécie da bacia amazônica *E. moratoi*, antes identificada como *E. crassipunctata*, constituiria uma nova espécie por apresentar o menor tamanho do grupo (20% menor), coloração tegumentar plenamente verde e o maior coxim posterior do grupo (Figura 3.32, E-H). Por fim, *E. pepei* seria uma espécie, endêmica à região de Porto Seguro, com exemplares coletados apenas no Parque Nacional do Pau Brasil. Interessantemente, embora *E. pepei* tenha sido descrita a partir de uma série tipo bastante reduzida (cinco indivíduos) de uma única localidade, a mesma seria a espécie mais distinta das descritas naquele estudo, com coxim anterior muito

semelhante ao de *E. parvula*, porém com o coxim posterior presente e o menor dentre as espécies do grupo (Figura 3.32, I-J). Segundo Nemésio e Engel (2012), *E. pepei* ainda se distingue de *E. parvula* pelo maior tamanho do coxim anterior da mesotíbia além da fenda metatibial menor.

Em suma, à exceção de *E. parvula*, todas as espécies do grupo *crassipunctata* parecem apresentar questões taxonômica relacionadas e, ainda que o presente estudo não se preste primariamente à proposição de novas hipóteses de espécies, análises genéticas podem auxiliar na elucidação e delimitação entre os limites populacionais e de espécie, suportando ou não as hipóteses atualmente propostas.

Adicionalmente, a despeito do status taxonômico das espécies do grupo, fica também evidente o caráter monofilético de todas estas espécies e/ou populações deste grupo e, ao mesmo tempo, a complexidade de sua história biogeográfica e evolutiva. Assim, explorar as relações filogenéticas e geográficas dos clados que constituem o grupo *crassipunctata* mostra-se bastante relevante para o entendimento da dinâmica biogeográfica das abelhas das orquídeas, sua história evolutiva e, conseqüentemente, das florestas úmidas neotropicais.

REFERÊNCIAS

- AB'SÁBER, A. **Os domínios de natureza no Brasil: potencialidades paisagísticas**. 4. ed. São Paulo: Ateliê Editorial, 2003.
- ACKERMAN, J. D.; APR, N.; ACKERMAN, D. Diversity and Seasonality of Male Euglossine Bees (Hymenoptera: Apidae) in Central Panamá DIVERSITY AND SEASONALITY OF MALE EUGLOSSINE. v. 64, n. 2, p. 274–283, 1983.
- ALLENDORF, F. W.; LUIKAR, G.; AITKEN, S. N. **Conservation and the genetics of population**. 1. ed. Victoria: Blackwell Publishing, 2007.
- ARIAS, M. C. et al. An oligonucleotide primer set for PCR amplification of the complete honey bee mitochondrial genome. **Apidologie**, v. 39, n. 4, p. 475–480, 2008.
- ARIF, I. A.; KHAN, H. A. Molecular markers for biodiversity analysis of wildlife animals : a brief review. **Animal Biodiversity and Conservation**, v. 32, n. 1, p. 9–17, 2009.
- ASCHER, J. S.; PICKERING, J. **Discover life: bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila)**.
- AUGUSTIN, L. et al. Eight glacial cycles from an Antarctic ice core. **Nature**, v. 429, n. 6992, p. 623–628, 10 jun. 2004.
- AUGUSTO, S. C.; GARÓFALO, C. A. Nesting biology and social structure of *Euglossa (Euglossa) townsendi* Cockerell (Hymenoptera, Apidae, Euglossini). **Insectes Sociaux**, v. 51, n. 4, p. 400–409, 2004.
- AUGUSTO, S. C.; GARÓFALO, C. A. Task allocation and interactions among females in *Euglossa carolina* nests (Hymenoptera, Apidae, Euglossini)*. **Apidologie**, v. 42, n. 2, p. 162–173, 2011.
- AVISE, J. C. et al. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. **Annual review of ecology and systematics**, v. 18, n. 1, p. 489–522, 1987.
- AVISE, J. C. **Phylogeography: the history and formation of species**. [s.l.] Harvard university press, 2000.
- AVISE, J. C. Phylogeography: Retrospect and prospect. **Journal of Biogeography**, v. 36, n. 1, p. 3–15, jan. 2009.
- BACON, C. D. et al. Biological evidence supports an early and complex emergence of the Isthmus of Panama. **Proceedings of the National Academy of Sciences**, v. 112, n. 19, p. 6110–6115, 12 maio 2015.
- BACON, C. D. et al. Quaternary glaciation and the Great American Biotic Interchange. **Geology**, v. 44, n. 5, p. 375–378, maio 2016.
- BAKER, C. L. The Lava Field of the Paraná Basin, South America. **The Journal of Geology**,

v. 31, n. 1, p. 66–79, 1923.

BALLARD, J. W. O.; WHITLOCK, M. C. The incomplete natural history of mitochondria. **Molecular Ecology**, v. 13, n. 4, p. 729–744, abr. 2004.

BARNOSKY, A. D. et al. Variable impact of late-Quaternary megafaunal extinction in causing ecological state shifts in North and South America. **Proceedings of the National Academy of Sciences**, v. 113, n. 4, p. 856–861, 26 jan. 2016.

BATALHA-FILHO, H. et al. Geographic distribution and spatial differentiation in the color pattern of abdominal stripes of the Neotropical stingless bee *Melipona quadrifasciata* (Hymenoptera: Apidae). **Zoologia (Curitiba)**, v. 26, n. 2, p. 213–219, jun. 2009.

BATALHA-FILHO, H. et al. Phylogeography and historical demography of the neotropical stingless bee *Melipona quadrifasciata* (Hymenoptera, Apidae): incongruence between morphology and mitochondrial DNA. **Apidologie**, v. 41, n. 5, p. 534–547, 4 set. 2010.

BATALHA-FILHO, H.; CABANNE, G. S.; MIYAKI, C. Y. Phylogeography of an Atlantic forest passerine reveals demographic stability through the last glacial maximum. **Molecular Phylogenetics and Evolution**, v. 65, n. 3, p. 892–902, 2012.

BATALHA-FILHO, H.; MIYAKI, C. Y. Filogeografia da Mata Atlântica. **Revista da Biologia**, v. Especial, n. 1, p. 31–34, 2011.

BATALHA-FILHO, H.; MIYAKI, C. Y. Late Pleistocene divergence and postglacial expansion in the Brazilian Atlantic Forest: Multilocus phylogeography of *Rhopias gularis* (Aves: Passeriformes). **Journal of Zoological Systematics and Evolutionary Research**, v. 54, n. 2, p. 137–147, maio 2016.

BAYES, T. An Essay towards Solving a Problem in the Doctrine of Chances. **Philosophical Transactions of the Royal Society of London**, v. 53, n. 0, p. 370–418, 1 jan. 1763.

BEERLI, P. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. **Bioinformatics**, v. 22, n. 3, p. 341–345, 1 fev. 2006.

BEHLING, H. Late Quaternary vegetational and climatic changes in Brazil. **Review of Palaeobotany and Palynology**, v. 99, n. 2, p. 143–156, 1998.

BERTORELLE, G. et al. **Population Genetics for Animal Conservation**. [s.l: s.n.]. v. 144

BOFF, S. et al. Island isolation reduces genetic diversity and connectivity but does not significantly elevate diploid male production in a neotropical orchid bee. **Conservation Genetics**, v. 15, n. 5, p. 1123–1135, 16 abr. 2014.

BOFF, S.; SAITO, C. A.; SANTOS, I. A. Multiple aggressions among nestmates lead to weak dominance hampering primitively eusocial behaviour in an orchid bee. **Sociobiology**, v. 64, n. 2, p. 202–211, 2017.

BOSSERT, S. et al. The impact of GC bias on phylogenetic accuracy using targeted enrichment phylogenomic data. **Molecular Phylogenetics and Evolution**, v. 111, p. 149–157, 2017.

BRADBURY, P. J. et al. TASSELL Software for association mapping of complex traits in diverse samples. **Bioinformatics**, v. 23, p. 2633–2635, 2007.

BRAND, P. et al. Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: Euglossini). **BMC Evolutionary Biology**, v. 15, n. 1, p. 1–16, 2015.

BRAND, P. et al. The Nuclear and Mitochondrial Genomes of the Facultatively Eusocial Orchid Bee *Euglossa dilemma*. **G3; Genes|Genomes|Genetics**, v. 7, n. 9, p. 2891–2898, set. 2017.

BRANSTETTER, M. G. et al. Phylogenomic Insights into the Evolution of Stinging Wasps and the Origins of Ants and Bees. **Current Biology**, v. 27, n. 7, p. 1019–1025, 2017.

BRITO, P. H.; EDWARDS, S. V. Multilocus phylogeography and phylogenetics using sequence-based markers. **Genetica**, v. 135, n. 3, p. 439–455, 2009.

BROOKS, D. R. Historical biogeography in the age of complexity: expansion and integration. Biogeografía histórica en la era de la complejidad: expansión e integración. **Revista Mexicana De Biodiversidad**, v. 76, p. 79–94, 2005.

BROOKS, T. M. et al. Global biodiversity conservation priorities. **Science**, v. 313, n. 5783, p. 58–61, 7 jul. 2006.

BROWN, D. Biogeography. Sinauer Associates. 2009.

BROWN, K. S.; SHEPPARD, P. M.; TURNER, J. R. G. Quaternary Refugia in Tropical America: Evidence from Race Formation in *Heliconius* Butterflies. **Proceedings of the Royal Society B: Biological Sciences**, v. 187, n. 1088, p. 369–378, 1974.

BRUMFIELD, R. T. et al. The utility of single nucleotide polymorphisms in inferences of population history. **Trends in Ecology & Evolution**, v. 18, n. 5, p. 249–256, maio 2003.

BRYANT, D. et al. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. **Molecular Biology and Evolution**, v. 29, n. 8, p. 1917–1932, ago. 2012.

BUSH, M. B.; OLIVEIRA, P. E. DE. The rise and fall of the Refugial Hypothesis of Amazonian speciation: a paleoecological perspective. **Biota Neotropica**, v. 6, n. 1, 2006.

CABANNE, G. S. et al. Phenotypic evolution of an Atlantic Forest passerine (*Xiphorhynchus fuscus*): Biogeographic and systematic implications. **Biological Journal of the Linnean Society**, v. 113, n. 4, p. 1047–1066, dez. 2014.

CABANNE, G. S.; SANTOS, F. R.; MIYAKI, C. Y. Phylogeography of *Xiphorhynchus fuscus* (Passeriformes, Dendrocolaptidae): Vicariance and recent demographic expansion in southern Atlantic forest. **Biological Journal of the Linnean Society**, v. 91, n. 1, p. 73–84, 2007.

CÂMARA, I. G. **Breve história da conservação da Mata Atlântica**. 1. ed. Belo Horizonte: Fundação SOS Mata Atlântica, 2005.

CARDINAL, S.; PACKER, L. Phylogenetic analysis of the corbiculate Apinae based on morphology of the sting apparatus (Hymenoptera: Apidae). **Cladistics**, v. 23, n. 2, p. 99–118, 2007.

CARDINAL, S.; STRAKA, J.; DANFORTH, B. N. Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. **Proceedings of the National Academy of Sciences**, v. 107, n. 37, p. 16207–16211, 2010.

CARDOSO, D. C. et al. Phylogeography of the sand dune ant *Mycetophylax simplex* along the Brazilian Atlantic Forest coast: Remarkably low mtDNA diversity and shallow population structure. **BMC Evolutionary Biology**, v. 15, n. 1, p. 106, 10 dez. 2015.

CARDOSO DA SILVA, J. M.; CARDOSO DE SOUSA, M.; CASTELLETTI, C. H. M. Areas of endemism for passerine birds in the Atlantic forest, South America. **Global Ecology and Biogeography**, v. 13, n. 1, p. 85–92, 2004.

CARNAVAL, A. C. et al. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. **Science**, v. 323, n. 5915, p. 785–789, 6 fev. 2009.

CARNAVAL, A. C. et al. Prediction of phylogeographic endemism in an environmentally complex biome. **Proceedings of the Royal Society B: Biological Sciences**, v. 281, n. 1792, p. 20141461–20141461, 13 ago. 2014.

CARNAVAL, A. C.; MORITZ, C. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. **Journal of Biogeography**, v. 35, n. 7, p. 1187–1201, jul. 2008.

CARTER, A. et al. Widespread Antarctic glaciation during the Late Eocene. **Earth and Planetary Science Letters**, v. 458, p. 49–57, 2017.

CARVALHO, C. J. B.; ALMEIDA, E. A. **Biogeografia da América do Sul: Análise de tempo, espaço e forma**. 2. ed. Rio de Janeiro: [s.n.].

CATCHEN, J. et al. Stacks: an analysis tool set for population genomics. **Molecular Ecology**, v. 22, n. 11, p. 3124–3140, jun. 2013.

CHIFMAN, J.; KUBATKO, L. Quartet Inference from SNP Data Under the Coalescent Model. **Bioinformatics**, v. 30, n. 23, p. 3317–3324, 1 dez. 2014.

COHEN, K. M. et al. The ICS International Chronostratigraphic Chart. **Episodes**, v. 36, n. 3, p. 199–204, 2014.

COMES, H. P.; KADEREIT, J. W. The effect of quaternary climatic changes on plant distribution and evolution. **Trends in Plant Science**, v. 3, n. 11, p. 432–438, 1998.

CORANDER, J. et al. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. **BMC bioinformatics**, v. 9, n. 1, p. 539, 2008.

CORANDER, J.; TANG, J. Bayesian analysis of population structure based on linked molecular information. **Mathematical Biosciences**, v. 205, n. 1, p. 19–31, 2007.

CORTOPASSI-LAURINO, M. et al. **Longevidade e tamanho de ninho gregário de Euglossa annectans em área urbana**. Anals do VIII Encontro sobre Abelhas. **Anais...Ribeirão Preto: 2008Disponível em: <<https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/160198/1/OPB1853.pdf>>. Acesso em: 18 mar. 2018**

CORTOPASSI-LAURINO, M.; ZILLIKENS, A.; STEINER, J. Pollen sources of the orchid bee *Euglossa annectans* Dressler 1982 (Hymenoptera: Apidae, Euglossini) analyzed from larval provisions. **Genetics and Molecular Research**, v. 8, n. 2, p. 546–556, 2009.

COSTA, L. P. The historical bridge between the Amazon and the forest of Brazil: a study of molecular phylogeography with small mammals. **Journal of Biogeography**, v. 30, n. 1, p. 71–86, 2003.

COX, C. B.; MOORE, P. D.; LADLE, R. J. **Biogeography: An Ecological and Evolutionary Approach**. 9. ed. Chichester: John Wiley & Sons, 2016.

COX, K. G. The role of mantle plumes in the development of continental drainage patterns. **Nature**, v. 342, n. 6252, p. 873–877, 28 dez. 1989.

CRISCI, J.; KATINAS, L.; POSADAS, P. **Historical biogeography: an introduction**. 1. ed. Cambridge: Harvard University Press, 2003.

CRISCI, J. V. et al. Bridging historical and ecological approaches in biogeography. **Australian Systematic Botany**, v. 19, n. 1, p. 1–10, 2006.

CROIZAT, L. Space, time, form: The biological synthesis. p. 881, 1964.

CROZIER, R. H.; CROZIER, Y. C. The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization. **Genetics**, v. 133, n. 1, p. 97–117, 1993.

CROZIER, Y. C.; KOULIANOS, S.; CROZIER, R. H. An improved test for Africanized honeybee mitochondrial DNA. **Experientia**, v. 47, n. 9, p. 968–969, 1991.

CUNHA, H.; SOLÉ-CAVA, A. Análise Filogenética. In: MATIOLI, S. R.; FERNANDES, F. M. C. (Eds.). **Biologia Molecular e Evolução**. 2. ed. Ribeirão Preto: Holos, 2012. p. 197–216.

DA SILVA, M. B.; PINTO-DA-ROCHA, R.; SOUZA, A. M. História Biogeográfica da Mata Atlântica: opções como modelo para sua inferência. In: CARVALHO, C. J. B. DE; ALMEIDA, E. A. (Eds.). **Biogeografia da América do Sul, Análise de tempo, espaço e forma**. 2. ed. Rio de Janeiro: Roca, 2016. p. 228–244.

DA SILVA, M.; NOLL, F. B. Biogeography of the social wasp genus *Brachygastra* (Hymenoptera: Vespidae: Polistinae). **Journal of Biogeography**, v. 42, n. 5, p. 833–842, 2015.

DANFORTH, B. N. et al. The history of early bee diversification based on five genes plus morphology. **Proceedings of the National Academy of Sciences**, v. 103, n. 41, p. 15118–15123, 10 out. 2006.

DANFORTH, B. N. et al. The Impact of Molecular Data on Our Understanding of Bee Phylogeny and Evolution. **Annual Review of Entomology**, v. 58, n. 1, p. 57–78, 2013.

DAVEY, J. L.; BLAXTER, M. W. RADseq: Next-generation population genetics. **Briefings in Functional Genomics**, v. 9, n. 5–6, p. 416–423, 2010.

DAVEY, J. W. et al. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. **Nature Reviews Genetics**, v. 12, n. 7, p. 499–510, 2011.

DE OLIVEIRA, M. L. Catálogo comentado das espécies de abelhas do gênero *Eulaema* Lepeletier, 1841 (Hymenoptera: Apidae). **Lundiana**, v. 8, n. 2, p. 113–136, 2007.

DEC, E.; DA SILVA MOUGA, D. M. D.; ALVES-DOS-SANTOS, I. New Geographical Records of Two Species of *Eufriesea* Cockerell, 1908 (Apidae: Euglossini) in Santa Catarina, Brazil, and the Use of a New Fragrance to Attract Orchid Bees. **Journal of the Kansas Entomological Society**, v. 90, n. 1, p. 24–30, 2017.

DELLICOUR, S.; MARDULYN, P. spads 1.0: A toolbox to perform spatial analyses on DNA sequence data sets. **Molecular Ecology Resources**, v. 14, n. 3, p. 647–651, 2014.

DICK, C. W. et al. Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA phylogeography. **Molecular Ecology**, v. 13, n. 12, p. 3775–3785, dez. 2004.

DODSON, C. H. et al. Biologically active compounds in orchid fragrances. **Science (New York, N.Y.)**, v. 164, n. 3885, p. 1243–1249, 1969.

DÖTTERL, S.; VEREECKEN, N. J. The chemical ecology and evolution of bee–flower interactions: a review and perspectives The present review is one in the special series of reviews on animal–plant interactions. **Canadian Journal of Zoology**, v. 88, n. 7, p. 668–697, 2010.

DRESSLER, R. L. An infrageneric classification of *Euglossa*, with notes on some features of special taxonomic importance (Hymenoptera; Apidae). **Revista de Biología Tropical**, v. 26, n. 1, p. 187–198, 1978.

DRESSLER, R. L. Biology of Orchid Bees. **Annual Review of Ecology and Systematics**, v. 13, n. 1982, p. 373–394, 1982a.

DRESSLER, R. L. New species of *Euglossa*.II. (Hymenoptera: Apidae). **Revista de Biología Tropical**, v. 30, n. 2, p. 121–129, 1982b.

DRESSLER, R. L. New species of *Euglossa* IV. The cordata and purpurea species groups (Hymenoptera: Apidae). **Revista de Biología Tropical**, v. 30, n. 2, p. 141–150, 1982c.

DRESSLER, R. L. Biology of the orchid bees (Euglossini). **Annual Review of Ecology and Systematics**, v. 13, n. 1982, p. 373–394, 1982d.

DRUMMOND, A. J.; BOUCKAERT, R. R. Bayesian evolutionary analysis with BEAST. **Bayesian Evolutionary Analysis with BEAST**, p. 1–249, 2015.

DRUMMOND, A. J.; RAMBAUT, A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC Evolutionary Biology**, v. 7, n. 1, p. 214, jan. 2007.

DUPANLOUP, I.; SCHNEIDER, S.; EXCOFFIER, L. A simulated annealing approach to define the genetic structure of populations. **Molecular Ecology**, v. 11, n. 12, p. 2571–2581, 2002.

DUQUE-LAZO, J. et al. Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. **Ecological Modelling**, v. 320, n. April, p. 62–70, 2016.

DURAN CORDEIRO, G. et al. Euglossine bees (Apidae) in Atlantic forest areas of São Paulo State, southeastern Brazil. **Apidologie**, v. 44, n. 3, p. 254–267, 16 nov. 2013.

EATON, D. A. R. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. **Bioinformatics**, v. 30, n. 13, p. 1844–1849, 1 jul. 2014.

EDGAR, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**, v. 5, p. 113, 19 ago. 2004.

EDWARDS, S.; BEERLI, P. Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. **Evolution**, v. 54, n. 6, p. 1839–1854, 2000.

EDWARDS, S. V. A smörgåsbord of markers for avian ecology and evolution. **Molecular Ecology**, v. 17, n. 4, p. 945–946, 23 jan. 2008.

EDWARDS, S. V. Is a new and general theory of molecular systematics emerging? **Evolution**, v. 63, n. 1, p. 1–19, 2009.

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

EHLERS, J.; GIBBARD, P. L.; HUGHES, P. D. Quaternary Glaciations and Chronology. In: **Past Glacial Environments**. [s.l.] Elsevier, 2018. p. 77–101.

ELITH, J. et al. A statistical explanation of MaxEnt for ecologists. **Diversity and Distributions**, v. 17, n. 1, p. 43–57, 2011.

ELSHIRE, R. J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PLoS ONE**, v. 6, n. 5, p. e19379, 4 maio 2011.

ELTZ, T. et al. Enfleurage, lipid recycling and the origin of perfume collection in orchid bees. **Proceedings of the Royal Society B: Biological Sciences**, v. 274, n. 1627, p. 2843–2848, 2007.

ELTZ, T.; ROUBIK, D. W.; LUNAU, K. Experience-dependent choices ensure species-specific fragrance accumulation in male orchid bees. **Behavioral Ecology and Sociobiology**, v. 59, n. 1, p. 149–156, 2005.

ELTZ, T.; ROUBIK, D. W.; WHITTEN, M. W. Fragrances, male display and mating behaviour of *Euglossa hemichlora*: A flight cage experiment. **Physiological Entomology**, v. 28, n. 4, p. 251–260, 2003.

ELTZ, T.; SAGER, A.; LUNAU, K. Juggling with volatiles: Exposure of perfumes by displaying male orchid bees. **Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology**, v. 191, n. 7, p. 575–581, 2005.

EMERSON, K. J. et al. Resolving postglacial phylogeography using high-throughput sequencing. **Proceedings of the National Academy of Sciences**, v. 107, n. 37, p. 16196–16200, 14 set. 2010.

ENGEL, M. S. The First Fossil *Euglossa* and Phylogeny of the Orchid Bees (Hymenoptera: Apidae; Euglossini). **American Museum Novitates**, v. 3272, p. 1–14, 1999.

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. **Molecular Ecology**, v. 14, n. 8, p. 2611–2620, jul. 2005.

EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin: an integrated software package for population genetics data analysis. **Evolutionary Bioinformatics**, v. 1, p. 47–50, 2005.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. **Genetics**, v. 131, n. 2, p. 479–491, 1992.

FARIA, L. R. R. et al. Diploid male dynamics under different numbers of sexual alleles and male dispersal abilities. **Theory in Biosciences**, v. 135, n. 3, p. 111–119, 2016.

FARIA JR, L. R. R.; MELO, G. A. R. Species of *Euglossa* (Glossura) in the Brazilian Atlantic forest, with taxonomic notes on *Euglossa stellfeldi* Moure (Hymenoptera, Apidae, Euglossina). **Revista Brasileira de Entomologia**, v. 51, n. 3, p. 275–284, 2007.

FARMER, G. T.; COOK, J. Pleistocene Glaciations. In: **Climate Change Science: A Modern Synthesis**. Dordrecht: Springer Netherlands, 2013. p. 407–427.

FENSTER, C. B. et al. Pollination Syndromes and Floral Specialization. **Annual Review of Ecology, Evolution, and Systematics**, v. 35, n. 1, p. 375–403, 2004.

FERRARI, B. R.; MELO, G. A. R. Deceiving colors: Recognition of color morphs as separate species in orchid bees is not supported by molecular evidence. **Apidologie**, v. 45, n. 5, p. 641–652, 19 mar. 2014.

FOURCADE, Y. et al. Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. **PLoS ONE**, v. 9, n. 5, p. e97122, 12 maio 2014.

FRANKHAM, R. Genetics and conservation biology. **Comptes Rendus Biologies**, v. 326, n. 1, p. 22–29, ago. 2003.

FRANKHAM, R. Where are we in conservation genetics and where do we need to go? **Conservation Genetics**, v. 11, n. 2, p. 661–663, 2010.

FRANKHAM, R.; BALOUE, J. D.; BRISCOE, D. A. **Introdução à Genética da Conservação**. 1. ed. New York: Cambridge University Press, 2008.

FRANTINE-SILVA, W. et al. Phylogeography and historical demography of the orchid bee *Euglossa iopocila*: signs of vicariant events associated to Quaternary climatic changes. **Conservation Genetics**, v. 18, n. 3, p. 539–552, 10 jun. 2017.

FREIRIA, G. A. et al. Population structure and genetic diversity of the orchid bee *Eufriesea violacea* (Hymenoptera, Apidae, Euglossini) from Atlantic Forest remnants in southern and southeastern Brazil. **Apidologie**, v. 43, n. 4, p. 392–402, 10 nov. 2012.

FRIEDMAN, J. R.; NUNNARI, J. **Mitochondrial form and function** *Nature*, 16 jan. 2014. Disponível em: <<https://www.nature.com/articles/nature12985.pdf>>. Acesso em: 20 fev. 2018

FU, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. **Genetics**, v. 147, n. 2, p. 915–925, 1997.

FUTUYMA, D. J.; AGRAWAL, A. A. Evolutionary history and species interactions. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 43, p. 18043–18044, 27 out. 2009.

GALINDO-LEAL, C.; CÂMARA, I. DE G. **The Atlantic Forest of South America: Biodiversity Status, Threats, and Outlook**. 1. ed. Washington, DC: Inland Press, 2003.

GARÓFALO, C. A. et al. Nest structure and communal nesting in *Euglossa* (*Glossura*) *annectans* Dressler (Hymenoptera, Apidae, Euglossini). **Revista Brasileira de Zoologia**, v. 15, n. 3, p. 589–596, 1998.

GARRAFFONI, A. R. S.; MOURA, F. R.; LOURENÇO, A. P. Areas of endemism in the Atlantic Forest: quantitative biogeography insights from orchid bees (Apidae: Euglossini). **Apidologie**, v. 48, n. 4, p. 513–522, 15 jul. 2017.

GENT, P. R. et al. The community climate system model version 4. **Journal of Climate**, v. 24, n. 19, p. 4973–4991, 2011.

GIANGARELLI, D. C. et al. *Eufriesea violacea* (Blanchard) (Hymenoptera: Apidae): an orchid Bee apparently sensitive to size reduction in forest patches. **Neotropical Entomology**, v. 38, n. 5, p. 610–615, 2009.

GIANGARELLI, D. C. et al. Orchid bees: a new assessment on the rarity of diploid males in populations of this group of Neotropical pollinators. **Apidologie**, v. 46, n. 5, p. 606–617, 2015.

GIANGARELLI, D. C.; AGUIAR, W. M. DE; SOFIA, S. H. Orchid bee (Hymenoptera: Apidae: Euglossini) assemblages from three different threatened phytophysiognomies of the subtropical Brazilian Atlantic Forest. **Apidologie**, v. 46, n. 1, p. 71–83, 18 jan. 2015.

GIORGETTA, M. A. et al. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. **Journal of Advances in Modeling Earth Systems**, v. 5, n. 3, p. 572–597, 2013.

GONZALEZ, V. H.; GRISWOLD, T.; SIMÕES, M. On the identity of the adventive species of *Eufriesea Cockerell* in the USA: systematics and potential distribution of the *coerulescens* species group (Hymenoptera, Apidae). **Journal of Hymenoptera Research**, v. 55, p. 55–105, 2017.

GRAZZIOTIN, F. G. et al. Phylogeography of the *Bothrops jararaca* complex (Serpentes: Viperidae): Past fragmentation and island colonization in the Brazilian Atlantic Forest. **Molecular Ecology**, v. 15, n. 13, p. 3969–3982, 2006.

GRISWOLD, T.; HERNDON, J. D.; GONZALEZ, V. H. First record of the orchid bee genus *Eufriesea Cockerell* (Hymenoptera: Apidae: Euglossini) in the United States. **Zootaxa**, v. 3957, n. 3, p. 342–346, 2015.

HAAG-LIAUTARD, C. et al. Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. **PLoS Biology**, v. 6, n. 8, p. 1706–1714, 19 ago. 2008.

HAASL, R. J.; PAYSEUR, B. A. Multi-locus inference of population structure: A comparison between single nucleotide polymorphisms and microsatellites. **Heredity**, v. 106, n. 1, p. 158–171, 2011.

HAFFER, J. Ciclos de tempo e indicadores de tempos na história da Amazônia. **Estudos avançados**, v. 6, n. 15, p. 7–39, 1992.

HAFFER, J. Hypotheses to explain the origin of species in Amazonia. **Brazilian Journal of Biology**, v. 68, n. 4 suppl, p. 917–947, nov. 2008.

HARPUR, B. A.; SOBHANI, M.; ZAYED, A. A review of the consequences of complementary sex determination and diploid male production on mating failures in the Hymenoptera. **Entomologia Experimentalis et Applicata**, v. 146, n. 1, p. 156–164, jan. 2013.

HARTL, D. L.; CLARK, A. G. **Principles of Population Genetics**. 4. ed. Sunderland, MA: Sinauer, 2007.

HARVEY, M. G. et al. Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. **Systematic Biology**, v. 65, n. 5, p. 910–924, set. 2016.

HEBERT, P. D. N. et al. Biological identifications through DNA barcodes. **Proceedings of the Royal Society B: Biological Sciences**, v. 270, n. 1512, p. 313–321, 7 fev. 2003.

HEBERT, P. D. N. et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. **Proceedings of the National Academy of Sciences**, v. 101, n. 41, p. 14812–14817, 2004.

HEBERT, P. D. N.; RATNASINGHAM, S.; DE WAARD, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. **Proceedings of the Royal Society B: Biological Sciences**, v. 270, n. Suppl_1, p. S96–S99, 7 ago. 2003.

HEDRICK, P. W. A standardized genetic differentiation measure. **Evolution**, v. 59, n. 8, p. 1633, 2005.

- HEDRICK AND, P. W.; PARKER, J. D. Evolutionary genetics and genetic variation of haplodiploids and x-linked genes. **Annual Review of Ecology and Systematics**, v. 28, n. 1, p. 55–83, nov. 1997.
- HEDTKE, S. M.; PATINY, S.; DANFORTH, B. N. The bee tree of life: A supermatrix approach to apoid phylogeny and biogeography. **BMC Evolutionary Biology**, v. 13, n. 1, 2013.
- HELLER, R.; SIEGISMUND, H. R. Relationship between three measures of genetic differentiation G_{ST} , D_{ST} and $G'ST$: How wrong have we been? **Molecular Ecology**, v. 18, n. 10, p. 2080–2083, 2009.
- HENNIG, W. Phylogenetic Systematics. **Annual review of entomology**, v. 10, n. 1, p. 97–116, 1966.
- HICKERSON, M. J. et al. Phylogeography's past, present, and future: 10 years after Avise, 2000. **Molecular Phylogenetics and Evolution**, v. 54, n. 1, p. 291–301, jan. 2010.
- HIJMANS, R. J. et al. Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology**, v. 25, n. 15, p. 1965–1978, 2005.
- HINOJOSA-DÍAZ, I. A.; ENGEL, M. S. A Checklist of the Orchid Bees of Nicaragua (Hymenoptera: Apidae: Euglossini). **Journal of the Kansas Entomological Society**, v. 85, n. 2, p. 135–144, 2012.
- HO, S. Y. W.; SHAPIRO, B. Skyline-plot methods for estimating demographic history from nucleotide sequences. **Molecular ecology resources**, v. 11, n. 3, p. 423–434, 2011.
- HOLDER, M.; LEWIS, P. O. Phylogeny estimation: Traditional and Bayesian approaches. **Nature Reviews Genetics**, v. 4, n. 4, p. 275–284, abr. 2003.
- HUELSENBECK, J. P. et al. Bayesian inference of phylogeny and its impact on evolutionary biology. **Science**, v. 294, n. 5550, p. 2310–2314, 14 dez. 2001.
- IBGE. **Atlas do Censo Demográfico 2010**. 1. ed. Rio de Janeiro: IBGE press, 2010.
- JABLONSKI, D. Extinctions in the Fossil Record. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 344, n. 1307, p. 11–17, 29 abr. 1994.
- JANZEN, D. H. et al. Seasonal and site variation in Costa Rican Euglossine bees at chemical baits in lowland deciduous and evergreen forests. **Ecology**, v. 63, n. 1, p. 66–74, 1982.
- JARAMILLO, C.; RUEDA, M. J.; MORA, G. Cenozoic plant diversity in the neotropics. **Science**, v. 311, n. 5769, p. 1893–1896, 2006.
- JOHNSON, B. R. et al. Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. **Current Biology**, v. 23, n. 20, p. 2058–2062, 2013.
- JOST, L. G_{ST} and its relatives do not measure differentiation. **Molecular Ecology**, v. 17, n. 18, p. 4015–4026, 2008.

JOST, L. D vs. GST: Response to Heller and Siegismund (2009) and Ryman and Leimar (2009) *Molecular Ecology*, 2009. Disponível em: <<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2009.04186.x>>. Acesso em: 26 mar. 2018

KANE, N. Should I use FST , G ' ST or D? Disponível em: <<http://www.molularecologist.com/2011/03/should-i-use-fst-gst-or-d-2/>>.

KAWAKITA, A. et al. Phylogenetic analysis of the corbiculate bee tribes based on 12 nuclear protein-coding genes (Hymenoptera: Apoidea: Apidae). *Apidologie*, v. 39, n. 1, p. 163–175, 2008.

KIMSEY, L. S. Synonymy of the genus *Euplusia* under *Eufriesia*, Hymenoptera, Apidae, Euglossini. *Pan-Pacific Entomologist*, v. 55, n. 2, p. 126, 1979.

KIMSEY, L. S. *Systematics of bees of the genus Eufriesea*. [s.l.] Univ of California Press, 1982. v. 95

KOPELMAN, N. M. et al. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, v. 15, n. 5, p. 1179–1191, set. 2015.

KOZAK, K. H.; GRAHAM, C. H.; WIENS, J. J. Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, v. 23, n. 3, p. 141–148, mar. 2008.

LACEY KNOWLES, L.; ALVARADO-SERRANO, D. F. Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: Insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Molecular Ecology*, v. 19, n. 17, p. 3727–3745, 2010.

LAFRAMBOISE, T. Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances. *Nucleic Acids Research*, v. 37, n. 13, p. 4181–4193, 1 jul. 2009.

LANG, B. F.; GRAY, M. W.; BURGER, G. Mitochondrial genome evolution and the origin of eukaryotes. *Annual review of genetics*, v. 33, n. 1, p. 351–397, 1999.

LARSEN, R. J. et al. Genetic Diversity of Neotropical Myotis (Chiroptera: Vespertilionidae) with an Emphasis on South American Species. *PLoS ONE*, v. 7, n. 10, p. e46578, 3 out. 2012.

LAURANCE, W. F. Conserving the hottest of the hotspots. *Biological Conservation*, v. 142, n. 6, p. 1137, jun. 2009.

LAVINA, E. L.; FAUTH, G. Evolução Geológica da América do Sul nos Últimos 250 milhões de Anos. In: CARVALHO, C. J. B. DE; ALMEIDA, E. A. B. (Eds.). *Biogeografia da América do Sul, Análise de tempo, espaço e forma*. 2. ed. Rio de Janeiro: Roca, 2016. p. 181–205.

LEDO, R. M. D.; COLLI, G. R. The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of Biogeography*, v. 44, n. 11, p. 2551–2563, nov. 2017.

LEIGH, J. W.; BRYANT, D. POPART: Full-feature software for haplotype network construction. **Methods in Ecology and Evolution**, v. 6, n. 9, p. 1110–1116, set. 2015.

LEITE, Y. L. R. et al. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. **Proceedings of the National Academy of Sciences**, v. 113, n. 4, p. 1008–1013, 2016a.

LEITE, Y. L. R. et al. Reply to Raposo do Amaral et al.: The “Atlantis Forest hypothesis” adds a new dimension to Atlantic Forest biogeography. **Proceedings of the National Academy of Sciences**, v. 113, n. 15, p. E2099–E2100, 2016b.

LEMMON, A. R.; MORIARTY, E. C. The importance of proper model assumption in bayesian phylogenetics. **Systematic Biology**, v. 53, n. 2, p. 265–277, abr. 2004.

LIBRADO, P.; ROZAS, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, n. 11, p. 1451–1452, 1 jun. 2009.

LOMOLINO, M. V.; SAX, D. F.; BROWN, J. H. **Foundations of biogeography: classic papers with commentaries**. [s.l.] University of Chicago Press, 2004.

LÓPEZ-URIBE, M. M. et al. Isolation and cross-species characterization of polymorphic microsatellites for the orchid bee *Eulaema meriana* (Hymenoptera: Apidae: Euglossini). **Conservation Genetics Resources**, v. 3, n. 1, p. 21–23, 2011.

LÓPEZ-URIBE, M. M. et al. Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees. **Molecular Ecology**, v. 23, n. 7, p. 1874–1890, abr. 2014.

LÓPEZ-URIBE, M. M.; ALMANZA, M. T.; ORDOÑEZ, M. Diploid male frequencies in Colombian populations of euglossine bees. **Biotropica**, v. 39, n. 5, p. 660–662, 2007.

LOZIER, J. D.; ZAYED, A. Bee conservation in the age of genomics. **Conservation Genetics**, v. 18, n. 3, p. 713–729, 17 jun. 2017.

LU, F. et al. TASSEL 3.0 Universal Network Enabled Analysis Kit (UNEAK) pipeline documentation. **White Paper**, p. 1–12, 2012.

MACARTHUR, R. H.; WILSON, E. O. **The Theory of Island Biogeography**. [s.l.: s.n.]. v. 1

MALDONADO-COELHO, M. Climatic oscillations shape the phylogeographical structure of Atlantic Forest fire-eye antbirds (Aves: Thamnophilidae). **Biological Journal of the Linnean Society**, v. 105, n. 4, p. 900–924, 2012.

MARTINS, F. DE M. Historical biogeography of the Brazilian Atlantic forest and the Carnaval-Moritz model of Pleistocene refugia: What do phylogeographical studies tell us? **Biological Journal of the Linnean Society**, v. 104, n. 3, p. 499–509, 2011.

MATZKE, N. J. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. **R package, version 0.2**, v. 1, p. 2013, 2013.

- MAYLE, F. E. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. **Journal of Quaternary Science**, v. 19, n. 7, p. 713–720, 2004.
- MCCORMACK, J. E. et al. Applications of next-generation sequencing to phylogeography and phylogenetics. **Molecular Phylogenetics and Evolution**, v. 66, n. 2, p. 526–538, 2013.
- MEIRMANS, P. G.; HEDRICK, P. W. Assessing population structure: FST and related measures. **Molecular Ecology Resources**, v. 11, n. 1, p. 5–18, 1 jan. 2011.
- MELO, G. A. R. et al. Small orchid bees are not safe: parasitism of two species of Euglossa (Hymenoptera: Apidae: Euglossina) by conopid flies (Diptera: Conopidae). **Revista Brasileira de Zoologia**, v. 25, n. 3, p. 573–575, 2008.
- MELO, G. A. R. Notes on the systematics of the orchid-bee genus Eulaema (Hymenoptera, Apidae). **Revista Brasileira de Entomologia**, v. 58, n. 3, p. 235–240, 2014.
- MERICO, A.; TYRRELL, T.; WILSON, P. A. Eocene/Oligocene ocean de-acidification linked to Antarctic glaciation by sea-level fall. **Nature**, v. 452, n. 7190, p. 979–982, 2008.
- MEROW, C.; SMITH, M. J.; SILANDER, J. A. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. **Ecography**, v. 36, n. 10, p. 1058–1069, 2013.
- MERZ, C. et al. Replicate Phylogenies and Post-Glacial Range Expansion of the Pitcher-Plant Mosquito, *Wyeomyia smithii*, in North America. **PLoS ONE**, v. 8, n. 9, p. e72262, 6 set. 2013.
- METZGER, J. P. Conservation issues in the Brazilian Atlantic forest. **Biological Conservation**, v. 142, n. 6, p. 1138–1140, 2009.
- METZKER, M. L. Sequencing technologies the next generation. **Nature Reviews Genetics**, v. 11, n. 1, p. 31–46, jan. 2010.
- MICHENER, C. D. Comparative External Morphology, Phylogeny, and a Classification of the Bees (Hymenoptera). **Bulletin of the American Museum of Natural History**, v. 82, p. 1–326, 1944.
- MILET-PINHEIRO, P.; SCHLINDWEIN, C. Do euglossine males (Apidae, Euglossini) leave tropical rainforest to collect fragrances in sugarcane monocultures? **Revista Brasileira de Zoologia**, v. 22, n. 4, p. 853–858, 2005.
- MIRANDA, E. A. et al. Natural history and biogeography of *Partamona rustica*, an endemic bee in dry forests of Brazil. **Insectes Sociaux**, v. 62, n. 3, p. 255–263, 2015.
- MONTES, C. et al. Middle Miocene closure of the Central American Seaway. **Science**, v. 348, n. 6231, p. 226–229, 2015.
- MORIN, P. A.; LUIKART, G.; WAYNE, R. K. SNPs in ecology, evolution and conservation. **Trends in Ecology and Evolution**, v. 19, n. 4, p. 208–216, 2004.

- MORITZ, C. et al. Diversification of Rainforest Faunas: An Integrated Molecular Approach. **Annual Review of Ecology and Systematics**, v. 31, n. 1, p. 533–563, 2000.
- MORIYAMA, E. N.; POWELL, J. R. Intraspecific nuclear DNA variation in *Drosophila*. **Molecular Biology and Evolution**, v. 13, n. 1, p. 261–277, 1996.
- MORRONE, J. J. Cladistic biogeography: identity and place. **Journal of Biogeography**, v. 32, n. 7, p. 1281–1284, 2005.
- MORRONE, J. J. **Evolutionary biogeography: an integrative approach with case studies**. [s.l.] Columbia University Press, 2008.
- MORRONE, J. J. Biogeographical regionalisation of the Neotropical region. **Zootaxa**, v. 3782, n. 1, p. 1, 25 mar. 2014.
- MORRONE, J. J.; CRISCI, J. V. Historical biogeography: Introduction to Methods. **Annual Review of Ecology and Systematics**. Palo Alto, CA, v. 26, n. 1995, p. 373–401, 1995.
- MOURE, J.; MELO, G. A. R.; FARIA JR, L. R. R. **Catalogue of bees (Hymenoptera, Apoidea) in the Neotropical Region-online version**. Disponível em: <<http://www.moure.cria.org.br/catalogue>>. Acesso em: 14 abr. 2016.
- MOURE, J. S. Contribuição para o conhecimento do gênero *Eulaema* Lepeletier (Hymen.-Apoidea). **Dusenía**, v. 1, n. 2, p. 181–231, 1950.
- MOURE, J. S. Novas espécies e notas sobre Euglossinae do Brasil e Venezuela (Hymenoptera, Apidae). **Revista Brasileira de Zoologia**, v. 16, p. 91–104, 1999.
- MOURE, J. S.; LUZ, E. Uma nova espécie de *Euplusia* da Bahia, Brasil. **Revista Brasileira de Zoologia**, v. 1, n. 3, p. 841–844, 2001.
- MYERS, N. et al. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, n. 2, p. 853–858, 2000.
- NEFF, B. D.; FU, P.; GROSS, M. R. Microsatellite evolution in sunfish (Centrarchidae). **Canadian Journal of Fisheries and Aquatic Sciences**, v. 56, n. 7, p. 1198–1205, 1999.
- NEI, M. Analysis of Gene Diversity in Subdivided Populations. **Proc. Nat. Acad. Sci. USA**, v. 70, n. 12, p. 3321–3323, dez. 1973.
- NEMÉSIO, A. Orchid bees (Hymenoptera: Apidae) of the Brazilian Atlantic forest. **Zootaxa**, v. 2041, n. 2041, p. 1–242, 2009.
- NEMÉSIO, A. et al. Effects of climate change and habitat loss on a forest-dependent bee species in a tropical fragmented landscape. **Insect Conservation and Diversity**, v. 9, n. 2, p. 149–160, 2016.
- NEMÉSIO, A.; ENGEL, M. S. Three new cryptic species of *Euglossa* from Brazil (Hymenoptera, apidae). **ZooKeys**, v. 222, p. 47–68, 2012.

- NEMÉSIO, A.; FARIA JR., L. R. R. First assessment of the orchid bee fauna (Hymenoptera: Apidae) at Parque Estadual do Rio Preto, a cerrado area in southeastern Brazil. **Lundiana**, v. 5IS-2, n. 2, p. 113–117, 2004.
- NEMÉSIO, A.; SILVEIRA, F. A. Edge effects on the orchid-bee fauna (Hymenoptera: Apidae) at a large remnant of Atlantic Rain Forest in southeastern Brazil. **Neotropical Entomology**, v. 35, n. 3, p. 313–323, 2006a.
- NEMÉSIO, A.; SILVEIRA, F. A. Deriving ecological relationships from geographical correlations between host and parasitic species: an example with orchid bees. **Journal of Biogeography**, v. 33, n. 1, p. 91–97, jan. 2006b.
- NETO, P. L.; LOYOLA, R. Biogeografia da Conservação. In: CARVALHO, C. J. B.; ALMEIDA, E. A. B. (Eds.). **Biogeografia da América do Sul, Análise de tempo, espaço e forma** 2. ed. Rio de Janeiro: Roca, 2016. p. 169–190.
- NYLANDER, J. A. A. et al. Bayesian phylogenetic analysis of combined data. **Systematic Biology**, v. 53, n. 1, p. 47–67, 1 fev. 2004.
- ODEA, A. et al. Formation of the Isthmus of Panama. **Science Advances**, v. 2, n. 8, p. e1600883–e1600883, 17 ago. 2016.
- OLDBERG, E. M. M. A. E. G. et al. Phylogenetic Inference of Reciprocal Effects between Geographic Range Evolution and Diversification. **Systematic biology**, v. 60, n. 4, p. 451–65, 2011.
- OLIVEIRA, M. L.; CAMPOS, L. A. O. Preferência por estratos florestais e por substâncias odoríferas em abelhas Euglossinae (Hymenoptera, Apidae). **Revista Brasileira de Zoologia**, v. 13, n. 4, p. 1075–1085, 1996.
- ORTIZ-JAUREGUIZAR, E.; CLADERA, G. A. Paleoenvironmental evolution of southern South America during the Cenozoic. **Journal of Arid Environments**, v. 66, n. 3 SPEC. ISS., p. 498–532, 2006.
- OSSA, P. G.; PÉREZ, F.; ARMESTO, J. J. Phylogeography of two closely related species of *nolana* from the coastal atacama desert of chile: Post-glacial population expansions in response to climate fluctuations. **Journal of Biogeography**, v. 40, n. 11, p. 2191–2203, 2013.
- PALSBØLL, P. J.; BÉRUBÉ, M.; ALLENDORF, F. W. Identification of management units using population genetic data. **Trends in Ecology and Evolution**, v. 22, n. 1, p. 11–16, jan. 2007.
- PARENTI, L. R. Common Cause and Historical Biogeography. In: **Biogeography in a changing world**. [s.l: s.n.]. p. 61–82.
- PARRA-H, A.; NATES-PARRA, G. The Ecological Basis for Biogeographic Classification: An Example in Orchid Bees (Apidae: Euglossini). **Neotropical Entomology**, v. 41, n. 6, p. 442–449, 2012.
- PAXTON, R. J. et al. Microsatellite loci for *Euglossa annectans* (Hymenoptera: Apidae) and their variability in other orchid bees. **Molecular Ecology Resources**, v. 9, n. 4, p. 1221–1223, 2009.

- PEAKALL, R.; SMOUSE, P. E. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. **Molecular Ecology Notes**, v. 6, n. 1, p. 288–295, 2006.
- PEARSE, D. E.; CRANDALL, K. A. Beyond FST: Analysis of population genetic data for conservation. **Conservation Genetics**, v. 5, n. 5, p. 585–602, 2004.
- PEARSON, P. N. et al. Stable warm tropical climate through the Eocene Epoch. **Geology**, v. 35, n. 3, p. 211–214, 2007.
- PENHA, R. E. S. et al. Mitochondrial DNA of *Euglossa iopoecila* (Apidae, Euglossini) reveals two distinct lineages for this orchid bee species endemic to the Atlantic Forest. **Apidologie**, v. 46, n. 3, p. 346–358, 28 maio 2015.
- PENNINGTON, R. T.; PRADO, D. E.; PENDRY, C. Neotropical Seasonally Dry Forests and Quaternary Vegetation Changes. **Journal of Biogeography**, v. 27, n. 2, p. 261–273, 2000.
- PERKTAŞ, U. et al. Climate-driven range shifts and demographic events over the history of Kruper's Nuthatch *Sitta krueperi*. **Bird Study**, v. 62, n. 1, p. 14–28, 2 jan. 2015.
- PETERSON, A. T. Conservatism of Ecological Niches in Evolutionary Time. **Science**, v. 285, n. 5431, p. 1265–1267, 20 ago. 1999.
- PETERSON, A. T.; SOBERÓN, J. Species distribution modeling and ecological niche modeling: Getting the Concepts Right. **Natureza a Conservacao**, v. 10, n. 2, p. 102–107, 2012.
- PHILLIPS, S.; ANDERSON, R. P.; SCHAPIRE, R. Maximum entropy modeling of species geographic distributions. **Ecological Modelling**, v. 190, n. 3–4, p. 231–259, 2006.
- PHILLIPS, S. J. et al. Opening the black box: an open-source release of Maxent. **Ecography**, v. 40, n. 7, p. 887–893, 2017.
- PHILLIPS, S. J.; DUDÍK, M.; SCHAPIRE, R. E. A maximum entropy approach to species distribution modeling. **21st International Conference on Machine Learning, Banff, Canada**, p. 655–662, 2004.
- PLATNICK, N. I.; NELSON, G. A method of analysis for historical biogeography. **Systematic Zoology**, v. 27, n. 1, p. 1–16, 1978.
- POINAR, G. *Paleoeglossa melissiflora* gen. n., sp. n. (Euglossinae: Apidae), Fossil Orchid bees in Dominican Amber. **Journal of the Kansas Entomological Society**, v. 71, n. April, p. 29–34, 1998.
- POKORNY, T. et al. Dispersal ability of male orchid bees and direct evidence for long-range flights. **Apidologie**, v. 46, n. 2, p. 224–237, 23 mar. 2015.
- PORFIRIO, L. L. et al. Improving the Use of Species Distribution Models in Conservation Planning and Management under Climate Change. **PLoS ONE**, v. 9, n. 11, p. e113749, 2014.
- PORTO, T. J.; CARNAVAL, A. C.; DA ROCHA, P. L. B. Evaluating forest refugial models using species distribution models, model filling and inclusion: A case study with 14 Brazilian species. **Diversity and Distributions**, v. 19, n. 3, p. 330–340, 2013.

POSADA, D. jModelTest: Phylogenetic model averaging. **Molecular Biology and Evolution**, v. 25, n. 7, p. 1253–1256, 2008.

POSADA, D.; BUCKLEY, T. R. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. **Systematic Biology**, v. 53, n. 5, p. 793–808, 2004.

POSADA, D.; CRANDALL, K. A.; TEMPLETON, A. R. Nested clade analysis statistics. **Molecular Ecology Notes**, v. 6, n. 3, p. 590–593, set. 2006.

POSADAS, P.; CRISCI, J. V; KATINAS, L. Historical biogeography: A review of its basic concepts and critical issues. **Journal of Arid Environments**, v. 66, n. 3, p. 389–403, 2006.

POTTS, S. G. et al. Global pollinator declines: Trends, impacts and drivers. **Trends in Ecology and Evolution**, v. 25, n. 6, p. 345–353, 2010.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945–959, 2000.

QUIJANO-ABRIL, M. A.; CALLEJAS-POSADA, R.; MIRANDA-ESQUIVEL, D. R. Areas of endemism and distribution patterns for Neotropical Piper species (Piperaceae). **Journal of Biogeography**, v. 33, n. 7, p. 1266–1278, jul. 2006.

RAMALHO, A. V.; GAGLIANONE, M. C.; OLIVEIRA, M. L. DE. Comunidades de abelhas Euglossina (Hymenoptera, Apidae) em fragmentos de Mata Atlântica no Sudeste do Brasil. **Revista Brasileira de Entomologia**, v. 53, n. 1, p. 95–101, 2009.

RAMÍREZ-SORIANO, A. et al. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. **Genetics**, v. 179, n. 1, p. 555–567, 2008.

RAMIREZ, S.; DRESSLER, R. L.; OSPINA, M. Abejas euglossinas (Hymenoptera: Apidae) de la Región Neotropical: Listado de especies con notas sobre su biología. **Biota Colombiana**, v. 3, n. 1, p. 7–118, 2002.

RAMÍREZ, S. R. et al. Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). **Biological Journal of the Linnean Society**, v. 100, n. 3, p. 552–572, 2010.

RAMOS-ONSINS, S. E.; ROZAS, J. Erratum: Statistical properties of new neutrality tests against population growth (*Molecular Biology and Evolution* (2002) 19 (2092–2100)). **Molecular Biology and Evolution**, v. 23, n. 8, p. 1642, 2006.

RAPOSO DO AMARAL, F. et al. Multilocus tests of Pleistocene refugia and ancient divergence in a pair of Atlantic Forest antbirds (Myrmeciza). **Molecular Ecology**, v. 22, n. 15, p. 3996–4013, ago. 2013.

RAPOSO DO AMARAL, F. et al. The “Atlantis Forest hypothesis” does not explain Atlantic Forest phylogeography. **Proceedings of the National Academy of Sciences**, v. 113, n. 15, p. E2097–E2098, 2016.

RAYMOND, M.; ROUSSET, F. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. **Journal of Heredity**, v. 86, n. 3, p. 248–249, 1995.

RÊBELO, J. M. M. **História Natural das Euglossíneas. As abelhas das orquídeas**. 1. ed. São Luís: Lithograf, 2001.

REBÊLO, J. M. M.; MOURE, J. S. As espécies de *Euglossa* Latreille do nordeste de São Paulo (Apidae, Euglossinae). **Revista Brasileira de Zoologia**, v. 12, n. 3, p. 445–466, 1995.

REE, R. H.; HIPPI, A. L. Inferring phylogenetic history from restriction site associated DNA (RADseq). In: HÖRANDL, E.; APPELHANS, M. S. (Eds.). **Next-Generation Sequencing in Plant Systematics**. 1. ed. Königstein: Koeltz Scientific Books, 2015. p. 1–24.

REE, R. H.; SMITH, S. A. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. **Systematic Biology**, v. 57, n. 1, p. 4–14, 2008.

REED, D. H.; FRANKHAM, R. Correlation between Fitness and Genetic Diversity/Correlación entre Adaptabilidad y Diversidad Genética. **Conservation Biology**, v. 17, n. 1, p. 230–237, 2003.

RIBEIRO, M. C. et al. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. **Biological Conservation**, v. 142, n. 6, p. 1141–1153, jun. 2009.

RIBEIRO, R. A. et al. Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): Insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. **Heredity**, v. 106, n. 1, p. 46–57, 2011.

RIDDLE, B.; HAFNER, D. Phylogeography in historical biogeography: investigating the biogeographic histories of populations, species, and young biotas. **Biogeography in a changing world**, p. 161–176, 2006.

RIDDLE, B. R. What is modern biogeography without phylogeography? **Journal of Biogeography**, v. 36, n. 1, p. 1–2, 2009.

RIDDLE, B. R. Biogeography: History. **The International Encyclopedia of Geography**, 2017.

RISSLER, L. J. Union of phylogeography and landscape genetics. **Proceedings of the National Academy of Sciences**, v. 113, n. 29, p. 8079–8086, 2016.

RIZZINI, C. **Tratado de Fitogeografia do Brasil**. 2. ed. Rio de Janeiro: Editora Âmbito Social, 1997.

ROCHA-FILHO, L. C. DA; GARÓFALO, C. A. Community Ecology of Euglossine Bees in the Coastal Atlantic Forest of São Paulo State, Brazil. **Journal of Insect Science**, v. 13, n. 23, p. 1–19, 2013.

ROCHA, T. C. et al. Molecular phylogeny and diversification of a widespread Neotropical rainforest bird group: The Buff-throated Woodcreeper complex, *Xiphorhynchus guttatus/susurrans* (Aves: Dendrocolaptidae). **Molecular Phylogenetics and Evolution**, v. 85, p. 131–140, abr. 2015.

ROCHA FILHO, L. C. et al. Genetic differentiation of the Euglossini (Hymenoptera, Apidae) populations on a mainland coastal plain and an island in southeastern Brazil. **Genetica**, v. 141, n. 1–3, p. 65–74, mar. 2013.

ROMIGUIER, J. et al. Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. **Molecular Biology and Evolution**, v. 33, n. 3, p. 670–678, 2016.

RONQUIST, F. Phylogenetic approaches in coevolution and biogeography. **Zoologica Scripta**, v. 26, n. 4, p. 313–322, 1997a.

RONQUIST, F. Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. **Systematic Biology**, v. 46, n. 1, p. 195–203, 1997b.

ROUBIK, D. W. Sibling Species of Glossura and Glossuropoda in the Amazon Region (Hymenoptera: Apidae: Euglossini). **Journal of the Kansas Entomological Society**, v. 77, n. 3, p. 235–253, 2004.

ROUBIK, D. W.; HANSON, P. **Abejas de orquídeas de la América tropical. Biología y guía de campo (Orchid bees of Tropical América) Biology and field guide**. 1. ed. Heredia, Costa Rica: INBio Press, 2004.

ROUBIK, D. W.; WEIGT, L. A.; BONILLA, M. A. Population Genetics, Diploid Males, and Limits To Social Evolution of Euglossine Bees. **Evolution**, v. 50, n. 2, p. 931–935, 1 abr. 1996.

RUIZ PESSEDA, L. C. et al. The evolution of a tropical rainforest/grassland mosaic in southeastern Brazil since 28,000 14C yr BP based on carbon isotopes and pollen records. **Quaternary Research**, v. 71, n. 3, p. 437–452, 2009.

RULL, V. Neotropical biodiversity: Timing and potential drivers. **Trends in Ecology and Evolution**, v. 26, n. 10, p. 508–513, 2011.

SAKAGAMI, S. F. Classification of the Apidae (Hymenoptera). **University of Kansas Science Bulletin**, v. 54, n. 4, p. 75–164, 1990.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the national academy of sciences**, v. 74, n. 12, p. 5463–5467, 1977.

SANMARTIN, I. et al. Bayesian island biogeography in a continental setting: the Rand Flora case. **Biology Letters**, v. 6, n. 5, p. 703–707, 2010.

SANMARTÍN, I. Event-based biogeography: integrating patterns, processes, and time. In: EBACH, M. C.; TANGNEY, R. S. (Eds.). **Biogeography in a changing world**. 1. ed. Boca Raton: CRC Press-Taylor and Francis Group, 2007. p. 135–159.

SANMARTÍN, I. Historical Biogeography: Evolution in Time and Space. **Evolution: Education and Outreach**, v. 5, n. 4, p. 555–568, 2012.

SANMARTÍN, I. Modelos Probabilísticos e Biogeografia. In: CARVALHO, C. J. B.; ALMEIDA, E. A. B. (Eds.). . **Biogeografia da América do Sul, Análise de tempo, espaço e forma**. 2. ed. Rio de Janeiro: Roca, 2016. p. 119–128.

SANMARTÍN, I.; VAN DER MARK, P.; RONQUIST, F. Inferring dispersal: A Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. **Journal of Biogeography**, v. 35, n. 3, p. 428–449, 2008.

SANTOS, F. R. et al. Diversidade Genética. In: DRUMMOND, G. M. et al. (Eds.). . **Biota Minas: diagnostico sobre o conhecimento da biodiversidade no estado de Minas Gerais**. 1. ed. Belo Horizonte: Fundação Biodivesitas, 2009. p. 389–410.

SCHMIEDER, S.; COLINET, D.; POIRIÉ, M. Tracing back the nascence of a new sex-determination pathway to the ancestor of bees and ants. **Nature Communications**, v. 3, n. 1, p. 895, 12 jan. 2012.

SCOTESE, C. R. **Atlas of Earth History, Vol. 1. Paleogeography, PALEOMAP Project, Arlington, Texas**. 1. ed. Arlington: [s.n.].

SERRANO-CINCA, C.; FUERTES-CALLÉN, Y.; MAR-MOLINERO, C. **Measuring DEA efficiency in Internet companies**. 1. ed. Baltimore: Johns Hopkins University Press, 2005. v. 38

SERVANT, M. et al. Tropical Forest Changes During the Late Quaternary in African and South-American Lowlands. **Global and Planetary Change**, v. 7, n. 1–3, p. 25–40, 1993.

SHENDURE, J. et al. DNA sequencing at 40: Past, present and future. **Nature**, v. 550, n. 7676, 2017.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nature Biotechnology**, v. 26, n. 10, p. 1135–1145, 1 out. 2008.

SILVA, J. M. C.; CASTELETI, C. H. M. Estado sa biodiversidade da Mata Atlântica brasileira. In: GALINDO-LEAL, C.; CÂMARA, I. G. (Eds.). . **Mata Atlântica, biodiversidade, ameaças e perspectivas**. Belo Horizonte: Fundação SOS Mata Atlântica, 2005. p. 472.

SILVEIRA, G. C. et al. The orchid bee fauna in the Brazilian savanna: do forest formations contribute to higher species diversity? **Apidologie**, v. 46, n. 2, p. 197–208, 2015.

SLATKIN, M. A measure of population subdivision based on microsatellite allele frequencies. **Genetics**, v. 139, n. 1, p. 457–462, 1995.

SMITH, S. A.; DONOGHUE, M. J. Combining Historical Biogeography with Niche Modeling in the Caprifolium Clade of Lonicera (Caprifoliaceae, Dipsacales). **Systematic Biology**, v. 59, n. 3, p. 322–341, maio 2010.

SÒLE-CAVA, A. M. S.; CUNHA, H. Genética e a Conservação da Natureza. In: MATIOLI, S. R.; FERNANDES, F. M. C. (Eds.). . **Biologia Molecular e Evolução**. 2. ed. Ribeirão Preto: Sociedade Brasileira de Genética, Holos, 2012. p. 217–238.

- SONG, J.; SHAO, H. SNP Array in Hematopoietic Neoplasms: A Review. **Microarrays**, v. 5, n. 1, p. 1, 2015.
- SORO, A. et al. The population genetics of two orchid bees suggests high dispersal, low diploid male production and only an effect of island isolation in lowering genetic diversity. **Conservation Genetics**, v. 18, n. 3, p. 607–619, 23 jun. 2017.
- SOUZA, R. O. et al. Microsatellite loci for euglossine bees (Hymenoptera: Apidae). **Molecular Ecology Notes**, v. 7, n. 6, p. 1352–1356, nov. 2007.
- SOUZA, R. O. et al. Conservation genetics of neotropical pollinators revisited: Microsatellite analysis suggests that diploid males are rare in orchid bees. **Evolution**, v. 64, n. 11, p. 3318–3326, 2010.
- SOUZA, J. M. DE. Mar territorial, zona econômica exclusiva ou plataforma continental? **Revista Brasileira de Geofísica**, v. 17, n. 1, p. 79–82, mar. 1999.
- STAMATAKIS, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics**, v. 22, n. 21, p. 2688–2690, 1 dez. 2006.
- STORCK-TONON, D.; MORATO, E. F.; OLIVEIRA, M. L. Fauna de Euglossina (Hymenoptera : Apidae) da Sul-ocidental, Amazônia. **Acta Amazonica**, v. 39, n. 3, p. 693–706, 2009.
- SUGUIO, K.; SALLUN, E. A. M.; SOARES, A. E. A. PERÍODO QUATERNÁRIO : “QUO VADIS”. **Revista Brasileira de Geociências**, v. 35, n. 3, p. 427–432, 2005.
- SUNDAR, D.-R. **binom: Binomial Confidence Intervals For Several Parameterizations**, 2014.
- SUNI, S. S. Dispersal of the orchid bee *Euglossa imperialis* over degraded habitat and intact forest. **Conservation Genetics**, v. 18, n. 3, p. 621–630, 5 jun. 2017.
- SUNI, S. S.; BROSI, B. J. Population genetics of orchid bees in a fragmented tropical landscape. **Conservation Genetics**, v. 13, n. 2, p. 323–332, 15 nov. 2012.
- SYDNEY, N. V.; GONÇALVES, R. B.; FARIA, L. R. R. Padrões espaciais na distribuição de abelhas Euglossina (Hymenoptera, Apidae) da região neotropical. **Papéis Avulsos de Zoologia**, v. 50, n. 43, p. 667–679, 2010.
- TABARELLI, M. et al. Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified landscapes. **Biological Conservation**, v. 143, n. 10, p. 2328–2340, out. 2010.
- TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, n. 3, p. 585–595, 1989.
- TAKAHASHI, N. C. et al. A Reanalysis of Diploid Male Frequencies in Euglossine Bees (Hymenoptera: Apidae). **Heredity**, v. 55, n. 9, p. 1897–1899, 2001.

TAMURA, K. et al. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.

THOMÉ, M. T. C. et al. Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. **Molecular Phylogenetics and Evolution**, v. 55, n. 3, p. 1018–1031, jun. 2010.

THOMÉ, M. T. C. et al. Barriers, rather than refugia, underlie the origin of diversity in toads endemic to the Brazilian Atlantic Forest. **Molecular Ecology**, v. 23, n. 24, p. 6152–6164, dez. 2014.

THOMSON, R. C.; WANG, I. J.; JOHNSON, J. R. Genome-enabled development of DNA markers for ecology, evolution and conservation. **Molecular Ecology**, v. 19, n. 11, p. 2184–2195, 23 abr. 2010.

TORKAMANEH, D.; BOYLE, B.; BELZILE, F. Efficient genome-wide genotyping strategies and data integration in crop plants. In: **Theoretical and Applied Genetics**. [s.l.] Springer Berlin Heidelberg, 2018. v. 131p. 499–511.

TORKAMANEH, D.; LAROCHE, J.; BELZILE, F. Genome-Wide SNP Calling from Genotyping by Sequencing (GBS) Data: A Comparison of Seven Pipelines and Two Sequencing Technologies. **PLoS ONE**, v. 11, n. 8, p. e0161333, 22 ago. 2016.

TRUJILLO-ARIAS, N. et al. The niche and phylogeography of a passerine reveal the history of biological diversification between the Andean and the Atlantic forests. **Molecular Phylogenetics and Evolution**, v. 112, n. April, p. 107–121, 2017.

VAN GILS, H. et al. Fine resolution distribution modelling of endemics in Majella National Park, Central Italy. **Plant Biosystems**, v. 146, n. sup1, p. 276–287, jan. 2012.

VAN OOSTERHOUT, C. et al. micro-checker: software for identifying and correcting genotyping errors in microsatellite data. **Molecular Ecology Notes**, v. 4, n. 3, p. 535–538, set. 2004.

VARELA, S.; LOBO, J. M.; HORTAL, J. Using species distribution models in paleobiogeography: A matter of data, predictors and concepts. **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 310, n. 3–4, p. 451–463, 2011.

VERNESI, C. et al. Where's the conservation in conservation genetics? **Conservation Biology**, v. 22, n. 3, p. 802–804, jun. 2008.

VIADANA, A. G.; PAULA, A.; CAVALCANTI, B. a Teoria Dos Refúgios Florestais Aplicada Ao Estado De São Paulo. **Revista da Casa da Geografia de Sobral**, v. 8, n. 9, p. 61–80, 2007.

WATANABE, S. et al. MIROC-ESM: model description and basic results of CMIP5-20c3m experiments. **Geoscientific Model Development Discussions**, v. 4, n. 2, p. 1063–1128, 17 maio 2011.

WEBER, M. G. et al. Macroevolution of perfume signalling in orchid bees. **Ecology Letters**, v. 19, n. 11, p. 1314–1323, 2016.

WHITLOCK, M. C. G'ST and D do not replace FST. **Molecular Ecology**, v. 20, n. 6, p. 1083–1091, mar. 2011.

WIENS, J. J. Speciation and Ecology Revisited: Phylogenetic Niche Conservatism and the Origin of Species. **Evolution**, v. 58, n. 1, p. 193, 2004.

WIENS, J. J.; DONOGHUE, M. J. Historical biogeography , ecology and species richness. **Elsevier**, v. 19, n. 12, 2004.

WIKELSKI, M. et al. Large-range movements of neotropical orchid bees observed via radio telemetry. **PLoS ONE**, v. 5, n. 5, p. e10738, 2010.

WILSON, G. A.; RANNALA, B. Bayesian inference of recent migration rates using multilocus genotypes. **Genetics**, v. 163, n. 3, p. 1177–1191, 2003.

WRIGHT, S. Genetical structure of populations. **Nature**, v. 166, n. 4215, p. 247–249, 12 ago. 1950.

WÜSTER, W. et al. Tracing an invasion: Landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: Crotalus durissus). **Molecular Ecology**, v. 14, n. 4, p. 1095–1108, 2005.

YOU, Y. et al. Simulation of the Middle Miocene Climate Optimum. **Geophysical Research Letters**, v. 36, n. 4, p. L04702, 19 fev. 2009.

YU, Y. et al. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. 2015.

ZAYED, A. Effective population size in Hymenoptera with complementary sex determination. **Heredity**, v. 93, n. 6, p. 627–630, 8 dez. 2004.

ZAYED, A. Bee genetics and conservation. **Apidologie**, v. 40, n. 3, p. 237–262, 7 jul. 2009.

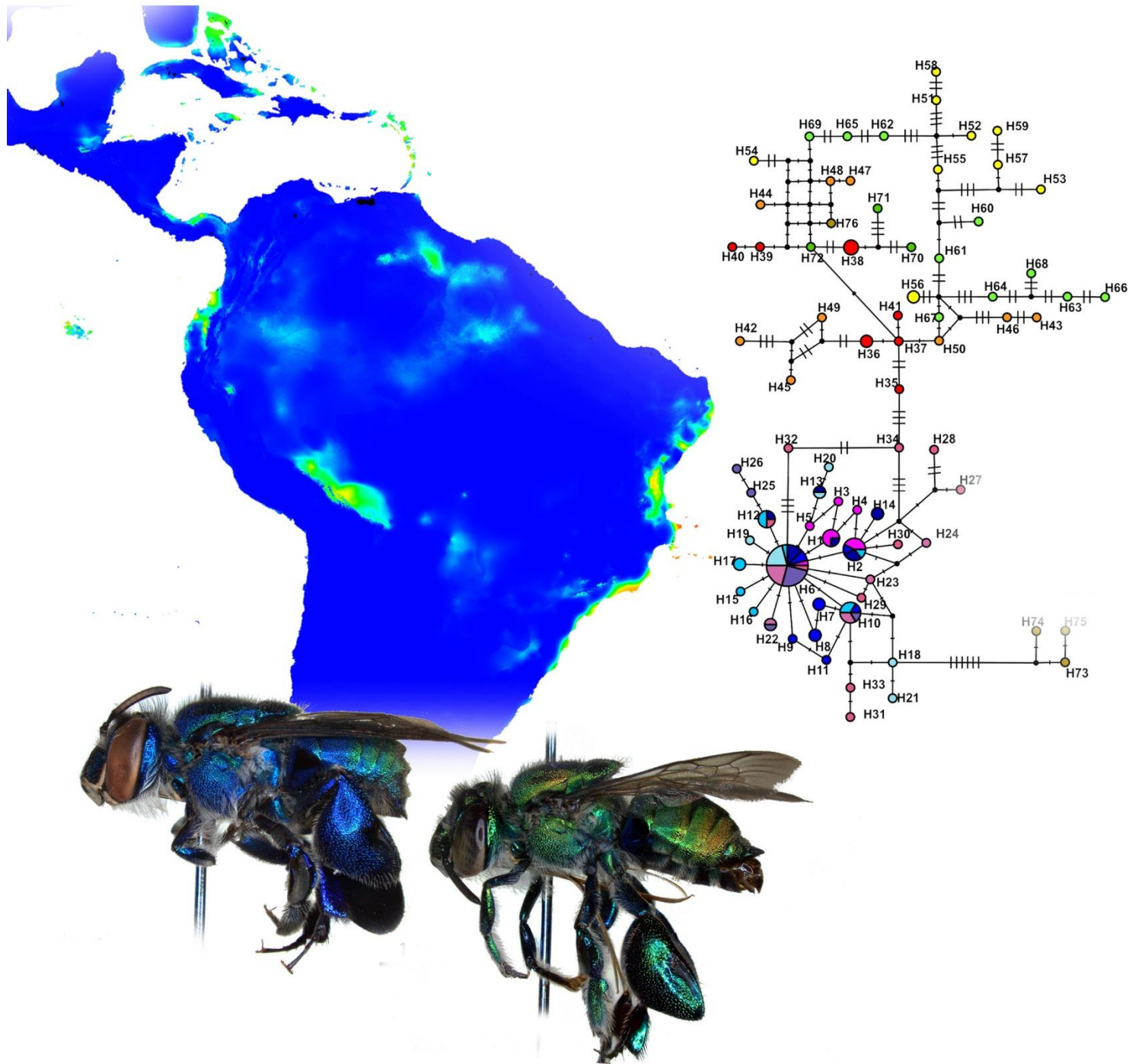
ZAYED, A.; PACKER, L. Complementary sex determination substantially increases extinction proneness of haplodiploid populations. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 30, p. 10742–6, 26 jul. 2005.

ZAYED, A.; ROUBIK, D. W.; PACKER, L. Use of diploid male frequency data as an indicator of pollinator decline. **Proceedings of the Royal Society B: Biological Sciences**, v. 271, n. Suppl_3, p. S9–S12, 2004.

ZIMMERMANN, Y. et al. Population genetic structure of orchid bees (Euglossini) in anthropogenically altered landscapes. **Conservation Genetics**, v. 12, n. 5, p. 1183–1194, 15 out. 2011.

ZINK, R. M. Drawbacks with the use of microsatellites in phylogeography: The song sparrow *Melospiza melodia* as a case study. **Journal of Avian Biology**, v. 41, n. 1, p. 1–7, jan. 2010.

ZINK, R. M.; BARROWCLOUGH, G. F. Mitochondrial DNA under siege in avian phylogeography. **Molecular Ecology**, v. 17, n. 9, p. 2107–2121, maio 2008.



PARTE II

EVIDÊNCIAS EMPÍRICAS

CAPÍTULO 1

FILOGEOGRAFIA E HISTÓRIA EVOLUTIVA DO GRUPO *CRASSIPUNCTATA* (APIDAE:
EUGLOSSINI) ACESSADA POR *GENOTYPING BY SEQUENCING* E MODELAGEM DE
NICO ECOLÓGICO

4 CAPÍTULO 1: FILOGEOGRAFIA E HISTÓRIA EVOLUTIVA DO GRUPO *CRASSIPUNCTATA* (APIDAE: EUGLOSSINI) ACESSADA POR *GENOTYPING BY SEQUENCING* E MODELAGEM DE NICHO ECOLÓGICO

RESUMO

No presente estudo as relações filogenéticas e filogeográficas de cinco das seis espécies do grupo *crassipunctata* foram investigadas utilizando 3248 locos *Genotyping by sequencing* (GBS), correspondendo a 312.840 pb. Adicionalmente, a influência dos eventos climáticos do Quaternário foi avaliada pela relação dos putativos padrões de distribuição, obtidos por modelos de nicho ecológico, com os padrões filogenéticos e de estrutura populacional da espécie atlântica *Euglossa clausi*. Os resultados apontam uma nova hipótese filogenética para grupo, com pontos de diversificação na transição Mioceno-Plioceno e Plioceno-Pleistoceno. Um corredor central entre a Amazônia (AM) e Mata Atlântica (MA) apresenta-se como principal rota de colonização do grupo para este ambiente. O padrão de estrutura genética encontrada para a espécie atlântica indica *clusters* concordantes com as expectativas dos modelos climáticos atuais. Em consonância com outros estudos, os resultados aqui apresentados demonstram que as flutuações climáticas do Quaternário foram determinantes para a colonização da MA e definição da estrutura genética das espécies endêmicas a este bioma, mas tiveram uma importância reduzida na diversificação das demais espécies do grupo.

Palavras-chave: Euglossini, GBS, Filogeografia, RAD-tags, Filogenia

4.1 INTRODUÇÃO

Espécies distribuídas ao longo dos neotrópicos têm experimentado uma história geológica bastante dinâmica, especialmente nos últimos 25 Ma a. p., com a formação progressiva dos Andes entre 23 e 4 Ma a. p., passando pela formação da Bacia Amazônica há ~7 Ma a. p. (Hoorn et al., 2010), até o fechamento definitivo do Istmo do Panamá, há aproximadamente 2,8 Ma a. p. (O’Dea et al., 2016). Um crescente número de estudos tem apontado a importância de eventos climáticos ao longo do Quaternário, mais especificamente no Pleistoceno, sobre a diversificação e estruturação genética de espécies neotropicais (Muellner et al., 2005; Carnaval et al., 2009; Maldonado-Coelho, 2012; Batalha-Filho & Miyaki, 2016; Trujillo-Arias et al., 2017).

Frequentemente ideias como a Teoria dos Refúgios (Haffer, 1969) são evocadas para explicar eventos de especiação na região Neotropical, uma vez que envolvem sucessivos cenários de vicariância causados por ciclos de expansão e retração das florestas úmidas (Haffer, 2008) decorrentes dos ciclos glaciais do Pleistoceno (Ehlers et al., 2018). Embora ao longo dos anos a ideia da Teoria dos Refúgios (*sensu* Haffer, 1969) como principal motor de diversificação nas florestas sul-americanas tenha caído em desuso, devido a uma série de questões relacionadas ao modelo (Bush & Oliveira, 2006), uma nova interpretação de como estas variações poderiam ter influenciado a biota ao longo dos últimos 2,58 Ma (Carnaval & Moritz, 2008) tem ganhado suporte de estudos filogeográficos recentes, principalmente no bioma de Mata Atlântica (MA) (Maldonado-Coelho, 2012; Batalha-Filho & Miyaki, 2016; Frantine-Silva et al., 2017). Contudo, a maior parte destes estudos têm dado ênfase ao entendimento de processos em nível interno do bioma, com uma menor atenção aos efeitos das glaciações entre os diferentes biomas (Trujillo-Arias et al., 2017).

Abelhas da tribo Euglossini são modelos bastante interessantes para estudos de processos evolutivos na região Neotropical por diversos motivos. Abelhas das orquídeas, como são popularmente conhecidas, constituem um grupo de abelhas “exclusivamente” neotropicais, incluindo uma grande diversidade de espécies (Roubik & Hanson, 2004; Nemésio, 2009; Moure et al., 2012), origem amazônica há cerca de 40 Ma a. p. (Roubik & Hanson, 2004; Ramírez et al., 2010), uma gama de comportamentos peculiares (Dressler, 1982a; Roubik & Hanson, 2004), e uma história biogeográfica praticamente inexplorada (Ramírez et al., 2010). Estas abelhas possuem ainda várias espécies com grande associação com florestas úmidas (Dressler, 1982a), onde alguns grupos se diversificaram em um gradiente de tolerância fisiológica com especificidades de nicho (López-Uribe et al., 2014).

A tribo Euglossini compreende quase 250 espécies distribuídas em cinco gêneros (Cameron, 2004), dos quais *Euglossa* é o mais diverso com cerca de 120 espécies (Moure et al., 2012). Estas espécies estão distribuídas em seis subgêneros: *Dasystilbe*, *Euglossa*, *Euglossella*, *Glossura*, *Glossurella* e *Glossuropoda*; e doze grupos de espécies (Dressler, 1982a; Cameron, 2004). Dentre essa matriz de espécies, poucas são tão amplamente distribuídas geograficamente quanto aquelas do grupo *crassipunctata*, *Euglossa* (*Glossurella*) (Dressler, 1982b), as quais se distribuem desde a América Central (Nicarágua, Costa Rica e Panamá) até o sudeste da MA (litoral de São Paulo, Brasil). Atualmente, este grupo compreende seis espécies: *Euglossa clausi* Nemésio & Engel 2012, *Euglossa crassipunctata* Moure 1968, *Euglossa moratoi* Nemésio & Engel 2012, *Euglossa parvula* Dressler 1982, *Euglossa pepei* Nemésio & Engel 2012 e *Euglossa sapphirina* Moure 1968; reconhecidas

como um grupo distinto desde sua proposição (Dressler, 1982b), sendo confirmado como monofilético em análises moleculares recentes (Ramírez et al., 2010). Embora a atual hipótese de separação do grupo acomode de forma mais parcimoniosa a ampla distribuição das espécies, o processo pelo qual essas espécies atingiram seus respectivos *status* ainda se encontra inexplorado.

Neste contexto, o presente estudo emprega uma abordagem integrada utilizando milhares de marcadores GBS (*Genotyping by Sequencing*) (Elshire et al., 2011) associados à modelagem de nicho ecológico e paleodistribuição (Phillips et al., 2006), visando inferir sobre os processos que influenciaram a diversificação das abelhas das orquídeas. Para tanto, análises filogenéticas envolvendo o grupo monofilético *crassipunctata* foram avaliadas em um contexto cronológico para identificar os possíveis eventos relacionados à diversificação deste grupo. Adicionalmente, a estrutura populacional da espécie atlântica *E. clausi* foi acessada buscando analisar a congruência dos padrões genéticos espaciais com as predições dos modelos paleoclimáticos. Por fim, as implicações taxonômicas da filogenia proposta são brevemente discutidas para propor uma nova subdivisão do grupo *crassipunctata*.

4.2 MATERIAIS E MÉTODOS

4.2.1 Área de Estudo e Amostragem

Considerando-se a distribuição continental do grupo *crassipunctata* (Nemésio & Engel, 2012), um levantamento dos registros de ocorrência para as diferentes espécies do grupo foi realizado com base em dados da literatura, registros em coleções e museus, além do banco de dados *online* (veja a seção de modelagem de nicho ecológico). Os dados levantados para as análises de nicho ecológico encontram-se resumidos na Figura 4.1 com a área de distribuição de cada espécie representada em forma de polígonos gerados pela intersecção dos pontos mais externos dos registros levantados para cada espécie. Ao longo do presente estudo considerou-se a classificação proposta por Nemésio e Engel (2012) para o reconhecimento dos registros ainda inalterados em várias fontes de dados.

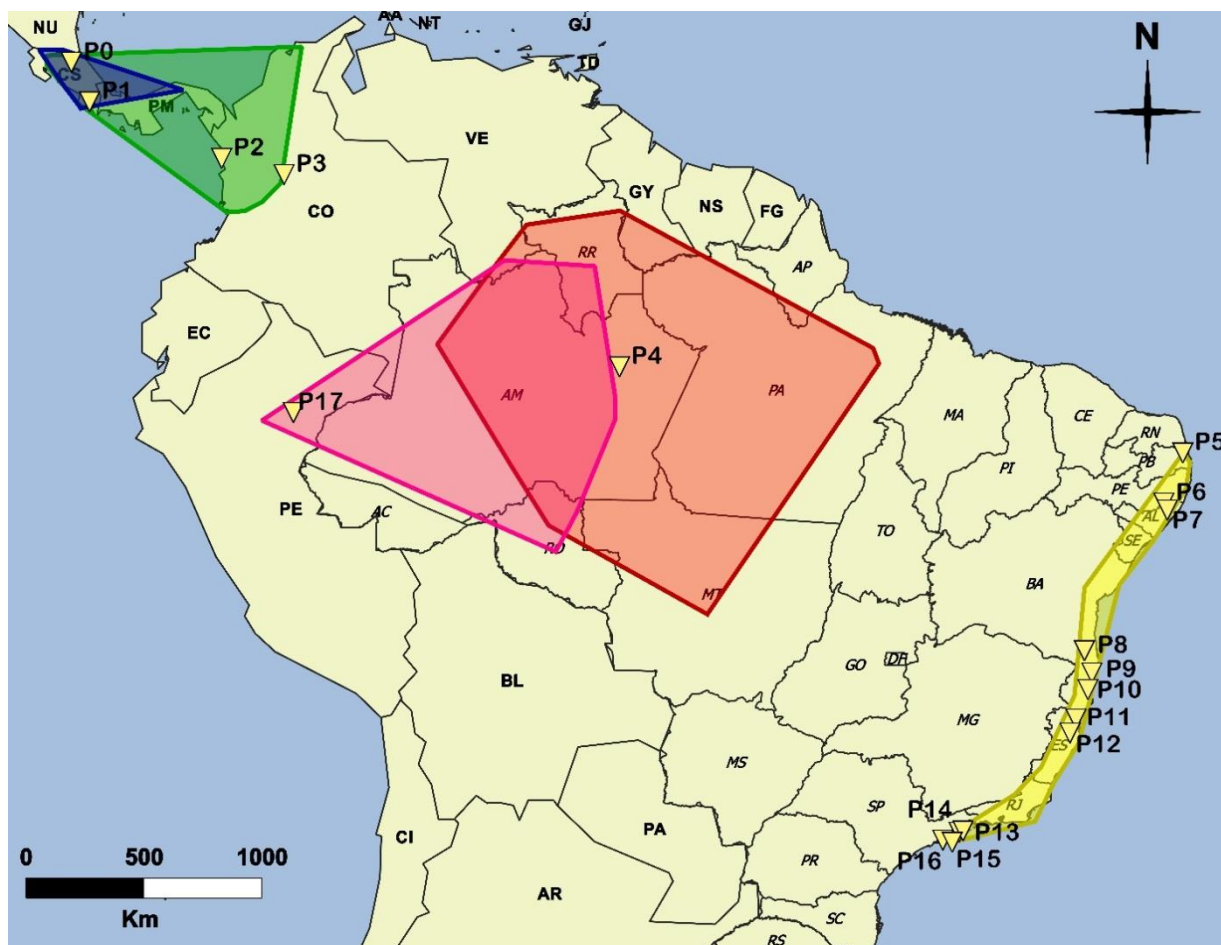


Figura 4.1. Distribuição dos pontos de amostragem e áreas de distribuição conhecida para o grupo *crassipunctata* utilizados neste estudo. Áreas em diferentes cores indicam as respectivas áreas de distribuição conhecidas para cada espécie. Em azul: *Euglossa sapphirina*; verde: *Euglossa crassipunctata*; rosa: *Euglossa parvula*; vermelho: *Euglossa moratoi*; amarelo: *Euglossa clausi*. Dados de *Euglossa pepei* não são apresentados.

As amostras utilizadas nas análises moleculares provenientes de 17 localidades, foram coletadas ao longo de suas respectivas distribuições desde a América Central e América do Sul (Figura 4.1; Tabela 4.1), seguindo o protocolo descrito por Giangarelli et al. (2009) e Ramírez et al. (2010), entre os anos de 2002 e 2006 (*E. parvula*, *E. crassipunctata*, *E. sapphirina* e *E. moratoi*) e 2014 e 2016 (*Euglossa clausi* e *Euglossa pepei*), utilizando armadilhas (Ramalho et al., 2009) e coletas ativas com iscas odoríferas (Nemésio & Santos Junior, 2014; Frantine-Silva et al., 2017).

Tabela 4.1 Descrição das áreas e número de machos amostrados para as seis espécies do grupo *crassipunctata*. O código das localidades (ID), localização, bem como latitude e longitude são apresentados ao lado do número de indivíduos amostrados. A - *Euglossa crassipunctata*; B - *Euglossa sapphirina*; C - *Euglossa moratoi*; D - *Euglossa clausi*; E - *Euglossa parvula*; F - *Euglossa pepeii*.

ID	Localidade	Latitude	Longitude	A	B	C	D	E	F	N
P0	Estação Ecológica La Selva - CR	10,430442	- 84,008531	2	2					4
P1	Estação Ecológica La Gamba - CR	8,7023190	- 83,206836	2	5					8
P2	Bahia Solano, Choco - CO	6,2151030	- 77,406233	7	6					12
P3	La Dorada, Caldas - CO	5,4990800	- 74,662347	2						2
P4	Manaus, Amazonas - BR	- 2,9504760	- 59,926394			5				5
P5	Mamanguape, Paraíba - BR	- 6,7305010	- 35,179149					2		2
P6	São José da Laje, Alagoas - BR	- 8,9337490	- 36,012453					5		5
P7	Murici, Alagoas - BR	- 9,2443980	- 35,860826					1 1		11
P8	Camacan, Bahia - BR	- 15,423622	- 39,501705					8		8
P9	Porto Seguro, Bahia - BR	- 16,395150	- 39,174996					3		3
P10	Prado, Bahia - BR	- 17,128357	- 39,349844					6	1	7
P11	Conceição da Barra, Espírito Santo - BR	- 18,424333	- 39,872963					3		3
P12	Linhares, Espírito Santo - BR	- 19,035305	- 40,136453					5		5
P13	Picinguaba, São Paulo - BR	- 23,339233	- 44,834434					1 2		12
P14	Ubatuba, São Paulo - BR	- 23,398530	- 45,005258					1 6		16
P15	Ilhabela, São Paulo - BR	- 23,832819	- 45,309999					8		8
P16	São Sebastião, São Paulo - BR	- 23,744016	- 45,722783					4		4
P17	Lagunas, Alto Amazonas - PE	-4,956756	- 74,254935						1	1
Total										115

Grande parte dessas amostras foram cedidas por vários pesquisadores, incluindo Dr. José Eustáquio Santos-Junior, Dr. Léo Correia Rocha-Filho, Dr. Márcio de Oliveira e Dr. Santiago Ramírez, de modo que seus respectivos *vouchers* encontram-se depositados em três diferentes instituições: Coleção de Abelhas do Museu de Zoologia da Universidade Estadual de Londrina, Coleção Entomológica da Universidade Federal de Minas Gerais, Coleção de Abelhas do Laboratório do Dr. Santiago Ramírez na Universidade da Califórnia.

4.2.2 Procedimentos Moleculares

Amostras de DNA foram extraídas das pernas protorácicas, musculatura torácica ou genitália por meio do protocolo de Chelex100, segundo Frantine-Silva et al. (2015) ou com auxílio de *kit* de extração DNAeasy Blood and Tissue (Qiagen, CA), seguindo as recomendações do fabricante. As bibliotecas GBS foram construídas a partir do protocolo original (Elshire et al., 2011), utilizando a enzima de restrição NSII (NEB, Beverly, MA, USA), a qual possui corte coesivo em sítios de seis pares de bases (3'T[^]ACGTA5'). Assim como o sequenciamento RAD, a técnica de GBS é também baseada em representação reduzida do genômica via enzima de restrição (McCormack et al., 2013), porém se difere do *RAD-seq* pela ausência de seleção de fragmentos na ligação dos adaptadores por tamanho, facilitando o manuseio e a construção da biblioteca (Elshire et al., 2011). As amostras indicadas na Tabela 4.1 foram individualmente identificadas por índices “*barcodes*” e sequenciadas em duas corridas *single-end* de 100pb em sequenciador NGS Illumina HiSeq 2000 (San Diego, CA, USA) no Laboratório de Sequenciamento Genômico Vincent J. Coates na Universidade da Califórnia, Berkeley.

4.2.3 Análise dos Dados Moleculares

4.2.3.1 Busca de marcadores e qualidade dos dados

Sequências de 100pb foram demultiplexadas por meio *software* Stacks 1.47 (Catchen et al., 2013), descartando sequências com ausência dos índices *barcodes*, sítio de restrição ou com distância maior do que uma base entre sítio de restrição e índice *barcodes*. Após a associação entre sequências e amostras, os índices *barcodes* e sítios de restrição foram removidos, resultando em sequências de 90 pb utilizando *scripts* Perl personalizados, descartando sequências com escores de qualidade inferiores a 80% (Rochette & Catchen, 2017).

Embora o genoma de *Euglossa dilemma* tenha sido recentemente descrito (Brand et al., 2017), o alinhamento entre as sequências aqui geradas via protocolo GBS para as espécies do grupo *crassipunctata* apresentaram valores de cobertura muito reduzidos (< 5%), de modo que a abordagem *De Novo* foi empregada para a obtenção de um maior número de SNPs. As *pipelines* contidas no programa Stacks 1.47 foram utilizadas manualmente devido à sua versatilidade e boa recuperação de SNPs frente às outras *pipelines* disponíveis (Torkamaneh et al., 2016). Primeiro, a função *ustacks* foi executada com número mínimo de “*reads*” idênticos para delimitar um marcador inicial (-m) igual a três e máximo de diferença para

reagrupar possíveis marcadores sobrepostos (-M) igual a dois. Em seguida, um catálogo de locos foi criado por meio da função *cstacks* com até sete bases desalinhadas para confirmar a correspondência entre locos de diferentes amostras (-n). Por fim, a função *populations* do programa Stacks foi executada para gerar arquivos de saída para os programas filogenéticos e populacionais, filtrando amostras com ao menos 500.000 *reads*, 20% ou mais de dados perdidos e uma frequência mínima de alelo de 1% ou maior do que a frequência mínima de um alelo comum. Adicionalmente foram incluídos apenas os locos ou marcadores que estivessem presentes em 90% das espécies e/ou populações analisadas.

4.2.3.2 Inferência filogenética

Supermatrizes concatenadas foram construídas a partir de sequências GBS (90 pb), considerando o consenso entre populações. Adicionalmente, amostras de diferentes localidades para as espécies amazônicas foram reunidas em uma única população (*E. parvula*, *E. crasipunctata*, *E. sapphirina* e *E. moratoi*). Métodos de Máxima Verossimilhança executado pelo programa RAxML (Stamatakis, 2006) e de inferência Bayesiana pelo programa MrBayes v.3.2.2 (Ronquist et al., 2012) foram utilizados para reconstruir a história filogenética das espécies do grupo *crasipunctata*. O programa MrBayes foi utilizado com os seguintes parâmetros, modelo GTR + Γ , cadeias de Markov e Monte Carlo (MCMC) com comprimento de 10.000.000 de gerações com quatro cadeias independentes, amostrando parâmetros a cada 1000 gerações. No programa PhyML foi utilizado o modelo GTR, com árvore inicial aleatória e 1.000 *bootstraps*. Modelos de substituição GTR foram utilizados nas supermatrizes concatenadas devido à complexa história evolutiva relacionada à origem das centenas de locos GBS (Tariel et al., 2016). A convergência das cadeias da MCMC foi avaliada pelo programa Tracer 1.16 do pacote BEAST (Drummond & Rambaut, 2007), bem como a árvore de máxima credibilidade foi gerada pelo programa LogCombiner, do mesmo pacote. Árvores filogenéticas e valores de suporte foram visualizados utilizando o programa FigTree v1.4.0 (Rambaut, 2014).

Tempos de divergência foram estimados pelo *software* BEAST v 2.3 (Drummond & Rambaut, 2007) utilizando um relógio molecular relaxado, com distribuição logNormal, combinado à um Modelo Yule de coalescência e modelo de substituição nucleotídica GTR+ Γ , a exemplo das análises filogenéticas. Quatro corridas independentes foram realizadas com 10 milhões de gerações, amostrando parâmetros a cada 1.000 gerações, partindo de árvores iniciais que foram fornecidas pelo usuário, baseadas nos resultados das análises de Máxima Verossimilhança e Bayesiana. Com base caráter monofilético das

espécies do grupo *crassipunctata*, as árvores foram calibradas para a datação da coalescência das espécies do grupo, centralizada em 8,1 Ma a. p. (s. d. 2,5), a partir de estimativas de Ramírez et al. (2010), estimadas aproximadamente entre 7 e 9,2 Ma a.p.

4.2.3.3 Análises discriminatórias e Parâmetros populacionais

Para avaliar a capacidade dos dados em diferenciar espécies e populações, análises não paramétricas de escalonamento multidimensional (*non-parametric multidimensional scaling*, nMDS) foram conduzidas no *software* TASSEL v. 5.0 (Lu et al., 2012). Adicionalmente, o *software* Structure v. 3.3.4 (Pritchard et al., 2000), foi utilizado para avaliar a estrutura populacional da espécie da MA, utilizando o modelo *admixture*, com *K* variando entre 1 e 13 com 20 interações de 100.000 passos para as cadeias Markov e Monte Carlo (MCMC), com descarte inicial das 10.000 primeiras interações. Os resultados desta análise foram visualizados pelo pacote CLUMPAK (Kopelman et al., 2015) observando o ΔK para determinar o *K* mais provável. Índices de diversidade nucleotídica e proporção de locos polimórficos foram estimados no *software* Stacks v. 1.47 (Catchen et al., 2013).

4.2.4 Modelagem de Nicho Ecológico

Considerando os efeitos dos ciclos glaciares ao longo do Quaternário sobre a diversificação de espécies das florestas tropicais (López-Uribe et al., 2014; Trujillo-Arias et al., 2017), bem como na estruturação genética de populações (Carnaval et al., 2009; Batalha-Filho & Miyaki, 2016; Frantine-Silva et al., 2017), modelos de nicho ecológico foram construídos, objetivando avaliar a congruência entre os padrões filogenéticos e áreas de adequabilidade para as respectivas espécies do grupo *crassipunctata*. Para tanto, o *software* MaxEnt v 3.3.3 k (Phillips et al., 2006) foi utilizado para construir modelos putativos de distribuição com base no algoritmo de Máxima Entropia (Phillips et al., 2004). Variáveis com maior importância para os quatro principais componentes (~90% da variação) foram selecionadas via PCA. Para evitar o efeito de colinearidade entre variáveis, os resultados da PCA foram também plotados em um mapa de fatores, onde variáveis sobrepostas foram consideradas colineares e, subsequentemente removidas. Um conjunto de 13 variáveis foi retido compreendendo uma variação média de 91,2% (d. p. 4,3), as quais são: BIO 11, BIO 6, BIO 10, BIO 5, BIO 1, BIO 9, BIO 7, BIO 8, BIO 16, BIO 12, BIO 17, BIO15 e BIO 13.

Os modelos foram baseados em diferentes números de registros de ocorrência, totalizando 260 registros: *E. clausi* (51), *E. crassipunctata* (30), *E. moratoi* (96), *E. parvula* (14), *E. pepei* (apenas 1 registro) e *E. sapphirina* (59) (veja Apêndice C). As ocorrências

georreferenciadas foram obtidas na literatura disponível, na coleção de Abelhas do Museu de Zoologia da Universidade Estadual de Londrina e nos bancos de Dados *SpeciesLink* (<http://www.splink.org>) e *Catalogue of Life* quando havia referência do material analisado. Os modelos gerados a partir destas ocorrências e das diferentes estratégias para seleção de variáveis, foram avaliadas segundo os valores de área sob a curva (AUC), que corresponde à relação entre especificidade e sensibilidade do modelo em relação à possível ocorrência da espécie (Merow et al., 2013).

Os modelos gerados foram então projetados para o último máximo glacial (*Last Glacial Maximum*, LGM; 21 Ka a. p.) considerando três diferentes Modelos de Circulação Geral (*Circulation General Models*, GCM), incluindo o *Community Climate System Model* versão 4.0, CCSM4 (GENT et al., 2011); *Max-Planck Institute - Earth System Model*, MPI-ESM (GIORGETTA et al., 2013); e *Model for Interdisciplinary Research on Climate - Earth System Model*, MIROC-ESM (WATANABE et al., 2011). Estes três modelos apresentam diferentes especificidades para interpolação e projeção dos dados climáticos em relação ao tempo considerando a exposição da Plataforma Continental no último máximo glacial. O pacote “*Raster*” do *software* R foi utilizado para equacionar *datum* (WGS84) e escala das 19 variáveis climáticas para todos os GCMs em resolução final de 30 segundos de arco (arcsec, $\sim 1 \text{ Km}^2$). Considerando a distribuição continental do grupo *crassipunctata*, os modelos de nicho ecológico (ENM) foram gerados para o espaço geográfico entre as latitudes 23 e -67, e longitudes -107 e -34, compreendendo toda a região Neotropical. Os modelos de adequabilidade ambiental foram gerados no *software* MaxEnt v 3.3.3 retendo 25% das amostras para validação cruzada e os demais 75% como dados de treinamento e teste de modelo ao longo de 10 replicações independentes. Limiares de igual sensibilidade e especificidade foram empregados para gerar as projeções logísticas e demais parâmetros foram utilizados em valores padrão do *software*. A partir das projeções logísticas, mapas binários de presença foram produzidos utilizando o *software* QGIS 2.14 (Quantum GIS Team, 2016) a partir de um limiar de corte de 0,5 para a taxa de adequabilidade, que segundo com López-Uribe et al. (2014) é adequado pois, além de ser um valor padrão para a determinação da presença de espécies, este também evita a consideração de extremos das curvas de adequabilidade onde os valores entre diferentes espécies tendem a convergir. Finalmente, mapas de ocorrência combinados por multiplicação do canal alpha com fator 0,33 foram usados para gerar os mapas de consenso entre os diferentes GCMs.

4.3 RESULTADOS

4.3.1 Genotipagem e estatísticas populacionais

Foram identificados ao longo de todas as populações analisadas um total de 119.257 locos, incluindo 31.120 SNPs. Após a aplicação dos parâmetros de filtragem para dados perdidos < 20%, frequência mínima de alelo alternativo 1% e presença do SNP em ao menos 80% das populações/espécies, 3.476 locos foram identificados com cobertura média de 10,7 *reads* por loco (d. p. 5,7), resultando em um conjunto de 1.529 SNPs retidos entre espécies e 687 SNPs entre as populações de *E. clausi* na MA (Tabela 4.2).

Tabela 4.2 Índices de polimorfismos entre grupos de sequências analisando um conjunto de 3.476 GBS-locos entre cinco espécies do grupo *crassipunctata*. Para cada espécie ou população, são apresentados: número de locos obtidos (N_L); de sítios analisados (N_S); sítios privados (P_S); sítios segregantes (S); número médio de indivíduos analisados por loco (N) e desvio padrão (d. p.); diversidade nucleotídica (π)

ID	N_L	N_S	P_S	S	N	d. p.	$\pi \cdot 10^{-3}$	d. p. 10^{-3}
<i>E. parvula</i>	2.789	250.870	863	1.069	1	0	0,0	0,0
<i>E. sapphirina</i>	2.819	253.758	680	1.322	1,7235	1,37	0,3	0,2
<i>E. crassipunctata</i>	3.476	312.840	780	1.664	3,5615	2,0622	0,9	0,4
<i>E. moratoi</i>	2.676	240.901	893	1.476	1,3708	0,3873	0,2	0,1
<i>E. clausi</i>	3.476	312.840	1023	1.874	35,735	2,7401	3,5	0,6
P5	599	53.910	18	46	1,3781	0,2351	0,9	0,5
P6	610	54.900	12	81	3,4374	1,0429	0,9	0,3
P7	523	52.070	43	141	5,0199	2,9314	0,9	0,4
P8	597	53.730	27	98	3,3944	1,5314	1,5	0,4
P9	625	56.250	20	86	1,5759	0,3661	1,2	0,8
P10	637	57.330	35	168	3,6803	1,4137	1,3	0,6
P11	641	57.690	19	96	2,034	0,580	1,1	0,5
P12	613	55.170	16	99	3,2069	0,9023	1,0	0,4
P13	625	56.250	17	67	8,7657	8,4909	1,0	0,6
P14	645	58.050	9	53	7,7393	5,9354	0,9	0,4
P15	640	57.600	8	93	2,3593	0,9548	0,9	0,5
P16	643	57.870	19	76	4,4885	2,9303	0,9	0,7

Uma baixa diferença entre o número de locos obtidos dentre espécies foi observada, variando entre 2.789 em *E. parvula* a 3.476 em *E. crassipunctata* e *E. clausi*, os quais refletem diretamente o número de indivíduos analisados por táxon. O mesmo pode ser observado para as demais estatísticas (N_S , P_S e S) que apresentaram números relativamente homogêneo entre as espécies analisadas, com aumento expressivo para os dados de *E. clausi*, a qual apresenta também o maior número de indivíduos analisados. Por outro lado, a comparação entre as amostras de *E. clausi* revela um baixo número de sítios privados em relação às suas demais populações, variando de oito em P15 à 35 em P10. Adicionalmente, os

valores de π mostram-se pouco variáveis ao longo das amostras de *E. clausi*, embora sejam ligeiramente maiores nas populações da porção central e norte da distribuição desta espécie (P5-P12).

4.3.2 Análise Filogenética

A abordagem aqui empregada utilizou 3476 locos GBS, os quais produziram 1529 sítios polimórficos filogeneticamente informativos e uma supermatriz de 312.840 pb. As topologias resultantes das análises de Máxima Verossimilhança e inferência Bayesiana se apresentaram bastante similares e foram resumidas em uma única árvore (Figura 4.2). Os resultados apontam *E. sapphirina* e *E. parvula* como grupo irmão de *E. crassipunctata*, *E. moratoi* e *E. clausi*, todos com 100% de suporte em ambas análises. A análise filogenética também revelou um agrupamento peculiar entre as amostras de *E. clausi*, apresentando um grupo monofilético para as amostras do sul da MA (SMA) e do norte (NMA), separando as amostras da porção central (CMA) em dois grupos (Figura 4.2), embora para alguns destes grupos o suporte seja baixo (< 70%) ou relativamente baixo, entre 70% e 80%.

Um importante ponto de diversificação para a maior parte das espécies do grupo é encontrado no final do Mioceno, há cerca de 5,9 Ma a. p. (95% i.c. 4,9 – 6,1), demarcando a separação de duas linhagens diferentes (*E. sapphirina* + *E. parvula*; e (*E. crassipunctata* + [*E. moratoi* + *E. clausi*])). Já separação da espécie Amazônica e Atlântica foi estimada para 2,9 Ma a. p. (95% i.c. 3,4 e 2,2), na transição entre Neogeno (Plioceno) e Quaternário (Pleistoceno), com o ancestral comum mais recente de *E. clausi* coalescendo em cerca de 0,9 Ma a. p. (95% i.c.: 1,1 e 0,6 Ma a. p.) e sua subsequente diversificação para os grupos ao longo da MA nos últimos 500 mil anos (Figura 4.2).

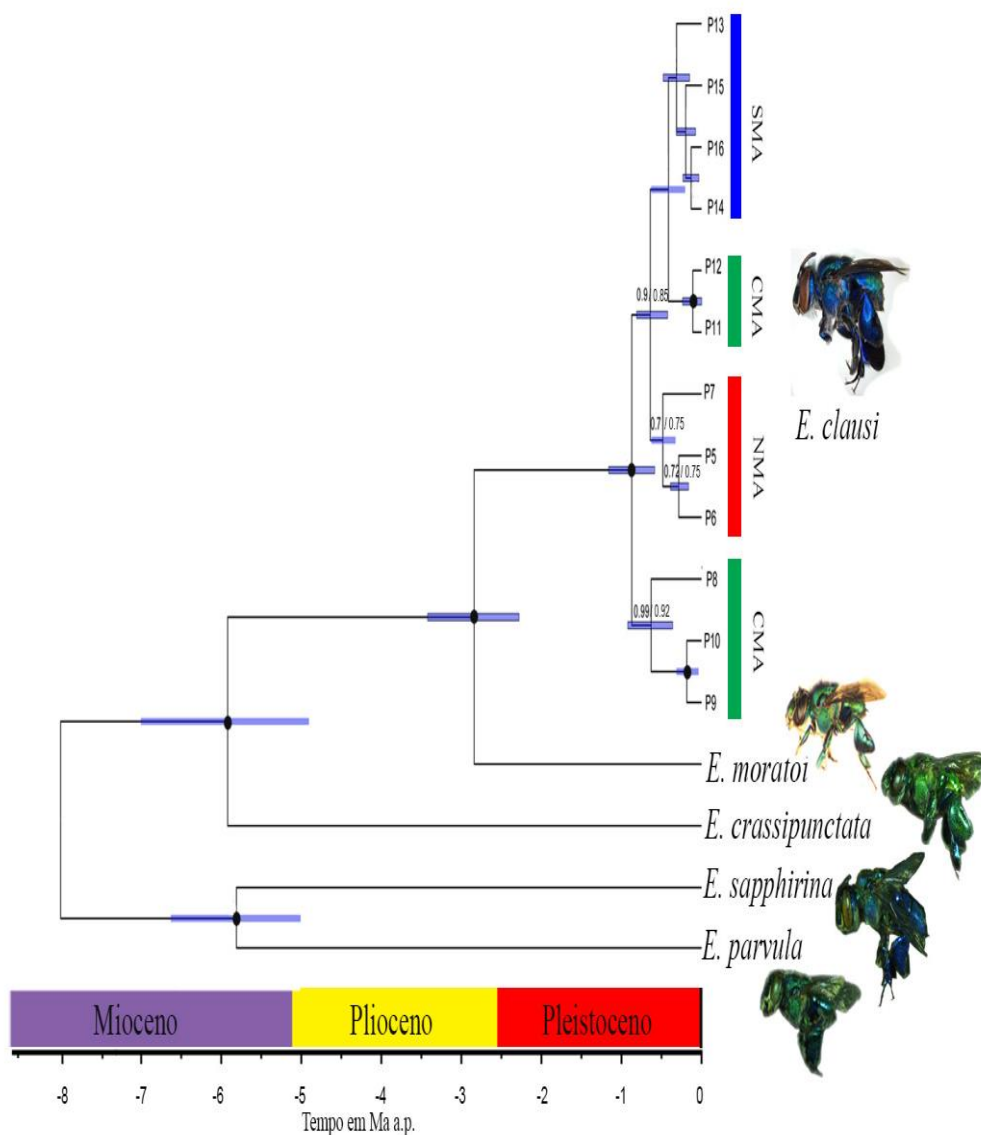


Figura 4.2 Inferência filogenética e estimativa de tempo de divergência para as espécies do grupo *crassipunctata* (exceto *E. pepei*) estimada a partir de 3476 locos GBS por inferência Bayesiana (IB) e de Máxima Verossimilhança (MV). Barras azuis sobre os nós indicam o intervalo de confiança de 95% para a estimativa de cada tempo. Pontos negros sobre os nós indicam valores de suporte igual à 1 ou 100 na IB e MV, com os demais valores sobre os nós indicando o suporte para as respectivas análises. Os grupos demarcados nos táxons correspondem à os grupos geográficos identificados nas análises n-MDS ou Structure (veja o texto). *Euglossa moratoi* retirada de Nemésio e Engel (2012).

4.3.3 Estrutura populacional

Com o intuito de compreender melhor os eventos de diversificação e seus efeitos sobre a estrutura populacional da espécie atlântica do grupo *crassipunctata*, análises de agrupamento via inferência Bayesiana (Structure) foram realizadas para as amostras de *E. clausi*. Os resultados da análise de agrupamento em até 13 *clusters* (Figura 4.3) revelou um número mais provável de *clusters* (K) igual a dois, segundo a estatística ΔK , como indicado

por Evanno et al. (2005). Segundo os resultados dessa análise, as amostras dos pontos P5 à P12 apresentam uma clara diferença na proporção de coancestralidade em relação às amostras de P13 à P16.

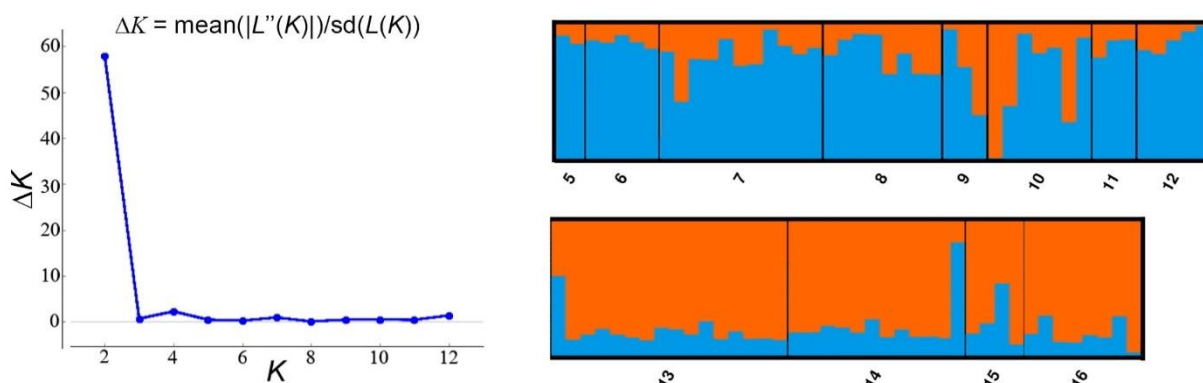


Figura 4.3 Resultados da análise Bayesiana (Structure) de atribuição. O resultado da variação média do \log de probabilidade à posteriori de K em função do desvio padrão (ΔK) é apresentado para os testes de K entre 1 e 13 (à esquerda). O coeficiente de ancestralidade (Q) é apresentado em um gráfico de barras (à direita), onde cada indivíduo corresponde a uma barra e recebem cores de acordo com a proporção de Q de seu respectivo cluster.

Adicionalmente, uma análise de escalonamento multidimensional não métrica (*non-metric Multidimensional Scalling*, n-MDS) foi realizada para investigar a variação dos dados SNPs tanto entre as espécies do grupo, como entre as amostras de *E. clausi*. Os resultados provenientes dessa análise indicaram a completa separação das espécies (Figura 4.4-A), em concordância com as análises filogenéticas, bem como uma separação mais refinada dos grupos dentre as amostras de *E. clausi*. Três grupos não sobrepostos foram observados entre as amostras de *E. clausi*, correspondendo ao agrupamento geográfico das amostras. Um grupo ao sul (SMA) em concordância com a análise de inferência Bayesiana, e dois outros grupos, um subdividindo a porção central (CMA; P10, P11 e P12) e agregando P8 e P9 às amostras da porção norte (NMA; P5, P6 e P7) (Figura 4.4-B).

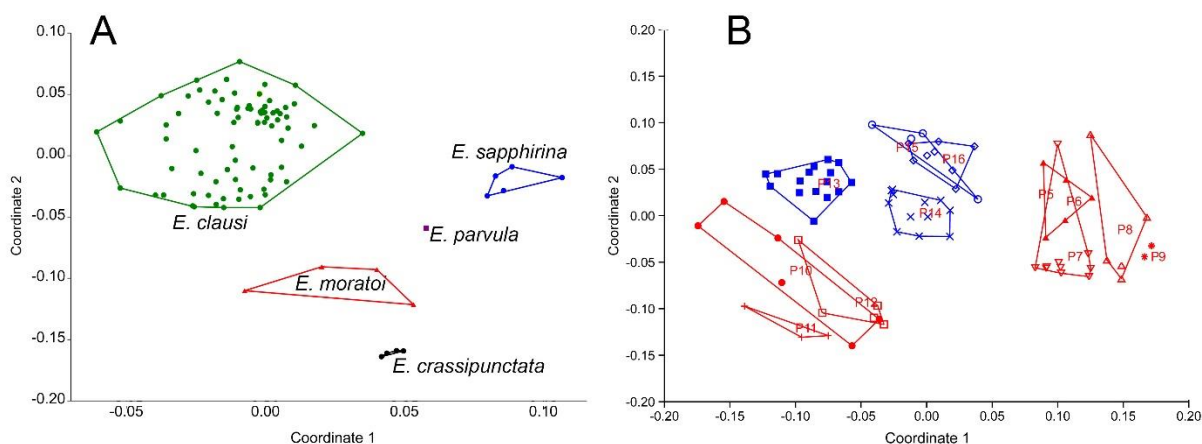


Figura 4.4 Análise de escalonamento multidimensional (nMDS) não métrica com base em 3476 locos GBS para diferentes espécies do grupo *crassipunctata*. A - nMDS incluindo 1529 SNPs de cinco das seis espécies do grupo *crassipunctata*: *Euglossa clausi*, *Euglossa crassipunctata*, *Euglossa moratoi*, *Euglossa parvula* e *Euglossa sapphirina*; B – nMDS incluindo dados de 687 SNPs identificados para as amostras de *E. clausi*; as cores azul e vermelho demarcam *clusters* identificados na análise de inferência Bayesiana, Structure; grupo SMA em azul.

4.3.4 Modelos de Distribuição e Paleodistribuição

Modelos de nicho ecológico e putativa distribuição foram construídos com o objetivo de observar a congruência dos eventos climáticos com possíveis eventos vicariantes envolvidos na especiação ou estruturação genética das espécies do grupo *crassipunctata*. Os resultados dos modelos de nicho para o tempo atual e para o LGM apresentaram elevados valores de AUC (entre 0,921 e 0,997), indicando uma alta acuidade em relação a um modelo nulo. As projeções para o período atual apresentaram um interessante padrão de isolamento de nicho em relação às diferentes espécies do grupo. *Euglossa clausi*, *E. crassipunctata* e *E. moratoi* não apresentaram áreas de sobreposição em relação à projeção dos modelos de nicho para o período atual. O mesmo foi observado entre *E. parvula* e *E. sapphirina*. A maior amplitude de nicho foi observada para *E. parvula*, enquanto *E. sapphirina* apresentou a projeção de distribuição mais restrita (a 4.5-A). *Euglossa pepeï* apresentou apenas dois pontos de ocorrência registrados no levantamento para o presente estudo, impossibilitando a construção de modelos baseadas nestes dados. As projeções para o LGM apresentaram um cenário bastante instável, resultando na redução de áreas de consenso entre os modelos na porção amazônica, enquanto na porção atlântica, *E. clausi* apresentou um padrão esperado de redução e fragmentação de áreas adequáveis. As duas espécies da América Central obtiveram expansões expressivas de áreas adequáveis em direção à região Amazônica enquanto as espécies amazônicas, *E. parvula* e *E. moratoi*, fizeram o movimento inverso, retraindo e fragmentado suas putativas distribuições em relação ao período atual (a 4.5-B). A região central da bacia amazônica e a região Andina do litoral do Chocó, apresentaram-se como

únicas regiões de estabilidade climática entre o período atual e o LGM para as espécies *E. parvula* e *E. moratoi*, bem como *E. crassipunctata*.

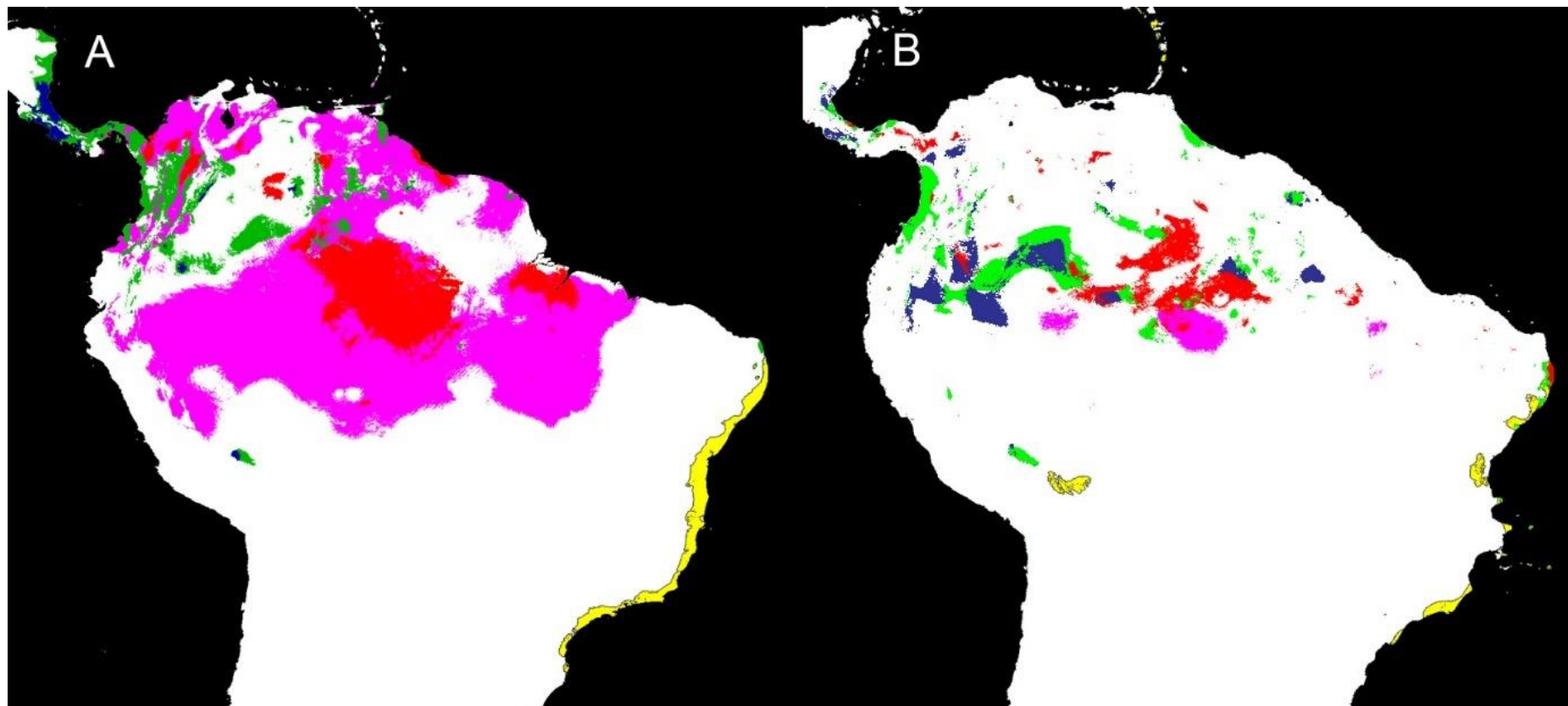


Figura 4.5. Consenso entre modelos de nicho ecológico com base em diferentes modelos de circulação geral (MPI-ESM-LR, CCM4-ESM, MIROC-ESM-LR) para diferentes espécies do grupo *crassipunctata*. As cores representam as áreas de distribuição para cada espécie: azul – *Euglossa sapphirina*; verde – *Euglossa crassipunctata*; vermelho – *Euglossa moratoi*; roxo – *Euglossa parvula*; amarelo *Euglossa clausi*. A – Modelo de distribuição para o tempo “atual” com corte de *threshold* de adequabilidade igual a 0,5. B – Modelo de paleodistribuição baseado no consenso entre os três GCM distintos.

4.4 DISCUSSÃO

Euglossini se apresenta como um dos mais interessantes grupos de abelhas neotropicais devido à sua grande diversidade de espécies (Roubik & Hanson, 2004; Nemésio, 2009; Moure et al., 2012), nichos e comportamentos (Dressler, 1982a; Roubik & Hanson, 2004), além de sua ampla história biogeográfica ainda pouco explorada (Ramírez et al., 2010). O presente estudo apresenta uma abordagem integrada entre dados moleculares e de modelagem de nicho ecológico, visando inferir sobre os processos que influenciaram a diversificação das abelhas das orquídeas em escala continental por meio de um grupo de espécies do gênero *Euglossa* subgênero *Glossurella* amplamente distribuído ao longo das florestas úmidas neotropicais. Os resultados aqui apresentados apontam a influência de três principais processos na diversificação desse grupo: i) a complexa história de soerguimento do Istmo do Panamá; ii) a acentuada elevação dos Andes no Plioceno tardio, por volta dos 5 Ma a. p. (Gregory-Wodzicki, 2000; Hoorn et al., 2010); iii) o início dos ciclos glaciares do Quaternário, possibilitando a existência de corredores por matas de galeria entre a Amazônia e a porção central da Mata Atlântica (Costa, 2003; Batalha-Filho et al., 2013; Ledo & Colli, 2017).

4.4.1 História evolutiva do grupo *crassipunctata*

Em relação às abelhas da tribo Euglossini parece haver um consenso em relação à sua origem na Bacia Amazônica, de onde teriam se dispersado para toda a distribuição atual (Roubik & Hanson, 2004; Ramírez et al., 2010). Embora possivelmente este seja o padrão geral para a tribo, os resultados aqui encontrados suportam uma hipótese alternativa e mais complexa para o grupo *crassipunctata*. As análises filogenéticas indicam que a separação entre as linhagens que deram origem às espécies da América Central, *E. crassipunctata* e *E. sapphirina* (~8 Ma a. p.), ocorreu anteriormente ao surgimento independente das espécies Amazônicas (~5 Ma a. p.). Em suporte à essa hipótese estão os resultados da filogenia de Ramírez et al. (2010), realizada de forma independente e baseada em dados de genes mitocondriais e nucleares, os quais também suportam o agrupamento *E. crassipunctata* + (*E. sapphirina* + *E. parvula*). Interessantemente, ambas análises (Ramírez et al., 2010; presente estudo) apontam a separação entre o ancestral de *E. sapphirina*+*E. parvula* e o ancestral de *E. crassipunctata*+*E. moratoi*+*E. clausi*. Deste modo, seria plausível que a dispersão seguida de vicariância entre a bacia amazônica e América Central poderia explicar primariamente a separação entre essas linhagens. Pontes entre a América do Sul e América Central são

relatadas anteriormente ao *landmarck* do fechamento do Istmo do Panamá (Bermingham & Moritz, 1998), consensualmente datado em 2,8 Ma a. p. (O’Dea et al., 2016). Bermingham e Moritz (1998) baseados em dados estimados para peixes de água doce, apontam para ao menos três ondas de colonização a partir América do Sul para a América Central inferior (Panamá e Costa Rica), com a primeira delas finalizada há cerca de 7 Ma a. p. Estes dados sugerem uma conexão ou aproximação mais acentuada entre os continentes anterior ao surgimento do Istmo da Panamá (Bacon et al., 2015; Montes et al., 2015), ainda que a data mais amplamente aceita para o fechamento completo da passagem seja 2,8 Ma a. p. (O’Dea et al., 2016). Vale ressaltar que, abelhas Euglossini possuem uma notável capacidade de voo (Wikelski et al., 2010; Pokorny et al., 2015), sendo frequentemente reportada uma baixa ou uma ausência de estruturação genética em sistemas de ilhas e continentes (Boff et al., 2014; Penha et al., 2015). Portanto, embora trocas faunísticas anteriores ao fechamento completo do istmo do Panamá com níveis reduzidos de conectividade constituam uma hipótese plausível para explicar a divergência entre estas duas espécies, estudos especificamente voltados a este sistema seriam necessários para confirmar tal hipótese.

Um segundo importante evento de diversificação para as abelhas do grupo *crassipunctata* foi identificado para a transição entre o Mioceno e Plioceno. Linhagens independentes há cerca de 6 Ma a. p. se diversificarem, dando origem ao ancestral de *E. moratoi*+*E. clausi* e a linhagem de *E. crassipunctata*. Quase sincronicamente, houve a separação de *E. parvula* e *E. sapphirina*, indicando um possível evento em comum entre a separação das linhagens atualmente ocupando América Central e Bacia Amazônica (mais Mata Atlântica). A transição entre Mioceno e Plioceno é marcada pela acentuada elevação dos Andes exatamente entre 5 e 4,5 Ma a. p. (Hoorn et al., 2010), o qual está cronologicamente alinhado com a divergência das formas amazônicas e centro-americanas. A presença da Cordilheira dos Andes como barreira vicariante é amplamente suportada por diversos estudos de diferentes naturezas relacionando o soergimento da cordilheira com a separação de diversos grupos (Gregory-Wodzicki, 2000; Miller et al., 2008; Hoorn et al., 2010; Luebert & Weigend, 2014; Domínguez et al., 2016), mas também encontra suporte na distribuição atual conhecida das espécies e na projeção das áreas de adequabilidade ambiental encontradas nas análises de nicho ecológico. Interessantemente, as mesmas análises também apontam para uma especificidade de nicho envolvendo as putativas áreas de distribuição das espécies do grupo *crassipunctata* formadas logo após este segundo evento de diversificação. De acordo com as projeções para o período atual, apenas linhagens independentes compartilham áreas de distribuição (ex. *E. parvula* e *E. moratoi*, e *E. crassipunctata* e *E. sapphirina*), enquanto as

espécies derivadas diretamente da mesma linhagem não apresentam sobreposição de áreas adequáveis. Este padrão encontra-se em linha com achados de López-Uribe et al. (2014) que apontam a tolerância fisiológica como um dos principais fatores para diversificação e isolamento de espécies de Euglossini na América do Sul.

Um último evento de diversificação foi observado para o grupo *crassipunctata*, envolvendo *E. moratoi* e *E. clausi*, foi observado no início do Pleistoceno (2,58 Ma a. p.; Cohen et al., 2014) com tempo de divergência estimado em 2,5 Ma a. p. (95% i.c. 2,8 e 1,9). A estimativa de divergência entre a espécie amazônica e atlântica sobrepõe o início das glaciações do Quaternário (Ehlers et al., 2018), as quais são amplamente relacionadas a grandes eventos de diversificação biológica (Rull, 2008; Ruiz Pessenda et al., 2009; Werneck, 2011; Bacon et al., 2016). Ao menos três conexões entre a Amazônia e MA são frequentemente reportadas: um corredor norte, cruzando os estados do Maranhão, Ceará e Rio Grande do Norte e Paraíba, formando brejos de altitude; um corredor central via matas de galeria através dos estados de Tocantins e Bahia (Costa, 2003; Batalha-Filho et al., 2013; Ledo & Colli, 2017); e o corredor Paraguai na porção sul conectando a Amazônia andina e Mata Atlântica (Costa, 2003; Batalha-Filho et al., 2013; Ledo & Colli, 2017; Trujillo-Arias et al., 2017). Os padrões filogenéticos observados para as populações de *E. clausi* apontam a diversificação inicial das amostras localizadas na porção média da MA, indicando uma possível rota de colonização pelo corredor central. Contudo, é importante ressaltar que os modelos de nicho não recuperaram áreas adequáveis para *E. moratoi* no LGM sobre a região do corredor central, ainda que diversos estudos documentem essa área como adequável a diversos organismos de MA (Costa, 2003; Batalha-Filho et al., 2013; Ledo & Colli, 2017). Por outro lado, a cronologia estabelecida para a formação deste corredor central (Batalha-Filho et al., 2013), mais recente, é perfeitamente alinhada com a datação da diversificação de *E. clausi*. Deste modo, diferentes conformações entre MA e Amazônia poderiam ser geradas ao longo das glaciações do Pleistoceno em respostas às suas diferenças em intensidade e duração (Augustin et al., 2004; Ehlers et al., 2018). Deste modo, uma complexa associação de eventos climáticos e geológicos está associada à diversificação do grupo *crassipunctata*, o qual apresenta um padrão a ser corroborado por futuros estudos com outras espécies de Euglossini.

4.4.2 Estruturação populacional na Mata Atlântica e flutuações climáticas

Diversos estudos têm apresentado a influência das variações climáticas do Quaternário sobre a estrutura genética de diferentes espécies ao longo da MA (Carnaval et al., 2009;

López-Uribe et al., 2014; Batalha-Filho & Miyaki, 2016; Frantine-Silva et al., 2017; Trujillo-Arias et al., 2017). Os resultados aqui obtidos para *E. clausi* apontam para esta mesma direção. A estruturação genética encontrada pela análise Bayesiana evidencia uma quebra filogeográfica na proporção do coeficiente de ancestralidade entre as amostras da porção norte e central da MA (NMA e CMA) em relação àquelas localizadas na porção sul (SMA). Este padrão de estruturação é frequentemente encontrado para espécies sob a influência da hipótese de refúgios, *sensu* Carnaval e Moritz (2008), segundo a qual, espécies com ampla distribuição latitudinal (ex.: *E. clausi*) encontraram refúgios na porção central e norte da MA ao longo dos períodos de resfriamento nos máximos glaciares. De fato, estudos filogeográficos tem suportado tanto esta hipótese (Carnaval et al., 2009; Martins, 2011; Maldonado-Coelho, 2012; Batalha-Filho & Miyaki, 2016), quanto cenários com áreas estáveis ao sul da MA (Thomé et al., 2010; Batalha-Filho et al., 2012; Álvarez-Presas et al., 2014). Interessantemente, os modelos de nicho ecológico para *E. clausi* indicam a presença de pequenos refúgios nas porções ao norte da MA, e adicionalmente uma área de adequabilidade que se expandiu em direção à Plataforma Continental na porção Sul, também observado recentemente em Batalha-Filho e Miyaki (2016). A permanência em refúgios implica em algumas expectativas, como o acúmulo de maior diversidade genética em relação a populações em zonas de instabilidade. Estudos que corroboram o modelo de refúgios *sensu* Carnaval e Moritz (2008), frequentemente reportam uma maior diversidade genética em populações que permaneceram nas regiões central e norte da MA (Carnaval et al., 2009; Batalha-Filho & Miyaki, 2016; Frantine-Silva et al., 2017). No entanto, embora as estimativas de diversidade genética levantadas para amostras de *E. clausi* apontem maiores estimativas de diversidade nucleotídica para a porção central da MA, essas são apenas ligeiramente maiores ($0,9 - 1,5 \times 10^{-3}$), corroborando a ideia de áreas estáveis também na porção sul da MA. A existência de refúgios nas três porções da MA, ainda que prevista em estudos anteriores, é reportada pela primeira vez em um único organismo, indicando uma peculiaridade de *E. clausi*. Adicionalmente, a análise de escalonamento multidimensional não métrica indicou um agrupamento ainda mais refinado, com três ou quatro grupos não sobrepostos, formados pelas amostras de SMA, CMA e NMA. Apenas as amostras CMA apresentaram um agrupamento não coerente com a disposição geográfica dos refúgios no LGM, aproximando as amostras da porção NMA com as de CMA.

4.4.3 Implicações taxonômicas

Até recentemente o grupo *crassipunctata* apresentava apenas três espécies reconhecidas, *E. crassipunctata*, *E. parvula* e *E. sapphirina*, as quais apresentavam um padrão de distribuição disjunto entre América Central e Mata Atlântica (*E. sapphirina*) ou eram reconhecidas como amplamente distribuídas ao longo de todo o continente (*E. crassipunctata*). Nemésio e Engel (2012) apresentaram a descrição de três novas espécies, *E. moratoi* na Bacia Amazônica, *E. clausi* e *E. pepeï* para a Mata Atlântica, uma hipótese mais coerente com a distribuição geográfica do grupo e corroborada pelos resultados moleculares aqui reportados. Estes autores propuseram ainda a separação natural do grupo em dois subgrupos: *crassipunctata* (*sensu strictissimo*) incluindo *E. crassipunctata*, *E. sapphirina*, *E. moratoi* e *E. clausi*; e o subgrupo *parvula*, incluindo *E. parvula* e *E. pepeï*. Embora esta separação apresente uma forte congruência do ponto de vista morfológico, os resultados das análises filogenéticas aqui apresentados, em consonância com os resultados encontrados por Ramirez et al. (2010), apontam uma subdivisão também em dois grupos, mas com a mudança de *E. sapphirina* para o subgrupo *parvula*. Considerando a descrição anterior de *E. sapphirina* em relação a *E. parvula*, propõe-se então a alteração do nome do subgrupo para *sapphirina*. Interessantemente, os resultados moleculares aqui reportados, alinhados aos de Ramírez et al. (2010), apontam para uma conservação de caracteres morfológicos ou convergência entre as formas *E. sapphirina* e *E. crassipunctata*.

4.4.4 Conclusões

Os resultados aqui apresentados são os primeiros a reunir diferentes metodologias para explorar a história evolutiva dos euglossíneos em escala intercontinental. De acordo com estes dados, as variações climáticas do Quaternário parecem ter tido uma importância reduzida na diversificação da maior parte das espécies do grupo *crassipunctata*, mas foram determinantes na colonização e estruturação genética de *E. clausi*. O pico de soergimento dos Andes na transição Mioceno-Plioceno aparece como um dos principais agentes vicariantes entre as espécies da América Central e Amazônia, cronológica e geograficamente alinhado com a diversificação das espécies atuais. Uma provável conexão entre a região Amazônica e Atlântica, via um corredor para porção central da MA, ao longo do Pleistoceno médio aparece com potencial caminho para a colonização da MA para as espécies do grupo *crassipunctata*. O padrão de estruturação genética e a estimativa de relação filogenética entre as amostras de *E. clausi* corroboram as predições dos modelos climáticos que apontam a instabilidade climática como um fator determinante da estruturação genética de espécies da MA, mas ao

mesmo tempo apontam uma idiosincrasia no processo de isolamento glacial para esta espécie demonstrando putativos refúgios nas três porções da MA. Estes dados, em conjunto com outros padrões observados para abelhas das orquídeas corroboram a importância da instabilidade climática para determinação da estrutura genética e diversificação do grupo. As análises filogenéticas possuem ainda importantes implicações para classificação atual do grupo *crassipunctata*, propondo uma hipótese filogenética distinta, em concordância com outros dados moleculares, que apontam uma diversificação mais complexa envolvendo a conservação de caracteres morfológicos em algumas linhagens em relação às demais. O presente estudo contribuí com uma nova hipótese de diversificação para espécies do gênero *Euglossa* que deve ser considerado em estudos futuros que buscam a identificação de padrões e processos para as espécies de Euglossini na MA.

REFERÊNCIAS

- Álvarez-Presas, M., Sánchez-Gracia, A., Carbayo, F., Rozas, J., & Riutort, M. (2014) Insights into the origin and distribution of biodiversity in the Brazilian Atlantic forest hot spot: a statistical phylogeographic study using a low-dispersal organism. *Heredity*, **112**, 656–665.
- Augustin, L., Barbante, C., Barnes, P.R.F., et al. (2004) Eight glacial cycles from an Antarctic ice core. *Nature*, **429**, 623–628.
- Bacon, C.D., Molnar, P., Antonelli, A., Crawford, A.J., Montes, C., & Vallejo-Pareja, M.C. (2016) Quaternary glaciation and the Great American Biotic Interchange. *Geology*, **44**, 375–378.
- Bacon, C.D., Silvestro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P., & Antonelli, A. (2015) Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences*, **112**, 6110–6115.
- Batalha-Filho, H., Cabanne, G.S., & Miyaki, C.Y. (2012) Phylogeography of an Atlantic forest passerine reveals demographic stability through the last glacial maximum. *Molecular Phylogenetics and Evolution*, **65**, 892–902.
- Batalha-Filho, H., Fjeldså, J., Fabre, P.-H., & Miyaki, C.Y. (2013) Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of Ornithology*, **154**, 41–50.
- Batalha-Filho, H. & Miyaki, C.Y. (2016) Late Pleistocene divergence and postglacial expansion in the Brazilian Atlantic Forest: Multilocus phylogeography of *Rhopias gularis* (Aves: Passeriformes). *Journal of Zoological Systematics and Evolutionary*

- Research*, **54**, 137–147.
- Bermingham, E. & Moritz, C. (1998) Comparative mtDNA phylogeography of neotropical freshwater fishes: Testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*, **7**, 499–517.
- Boff, S., Soro, A., Paxton, R.J., & Alves-dos-Santos, I. (2014) Island isolation reduces genetic diversity and connectivity but does not significantly elevate diploid male production in a neotropical orchid bee. *Conservation Genetics*, **15**, 1123–1135.
- Brand, P., Saleh, N., Pan, H., Li, C., Kapheim, K.M., & Ramírez, S.R. (2017) The Nuclear and Mitochondrial Genomes of the Facultatively Eusocial Orchid Bee *Euglossa* dilemma. *G3; Genes/Genomes/Genetics*, **7**, 2891–2898.
- Bush, M.B. & Oliveira, P.E. de (2006) The rise and fall of the Refugial Hypothesis of Amazonian speciation: a paleoecological perspective. *Biota Neotropica*, **6**, .
- Cameron, S.A. (2004) Phylogeny and Biology of Neotropical Orchid Bees (Euglossini). *Annual Review of Entomology*, **49**, 377–404.
- Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., & Moritz, C. (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.
- Carnaval, A.C. & Moritz, C. (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., & Cresko, W.A. (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.
- Cohen, K.M., Finney, S.C., Gibbard, P.L., & Fan, J.-X. (2014) The ICS International Chronostratigraphic Chart. *Episodes*, **36**, 199–204.
- Costa, L.P. (2003) The historical bridge between the Amazon and the forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography*, **30**, 71–86.
- Domínguez, M.C., Agrain, F.A., Flores, G.E., & Roig-Juñent, S.A. (2016) Vicariance events shaping Southern South American insect distributions. *Zoologica Scripta*, **45**, 504–511.
- Dressler, R.L. (1982a) Biology of Orchid Bees. *Annual Review of Ecology and Systematics*, **13**, 373–394.
- Dressler, R.L. (1982b) New species of *Euglossa*. III. the *bursigera* species group (Hymenoptera: Apidae). *Rev. Biol. Trop.*, **30**, 131–140.
- Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Ehlers, J., Gibbard, P.L., & Hughes, P.D. (2018) Quaternary Glaciations and Chronology.

- Past Glacial Environments* pp. 77–101. Elsevier,
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., & Mitchell, S.E. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.
- Evanno, G., Regnaut, S., & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Ferrari, R.R., Nemésio, A., & Silveira, F.A. (2017) Phylogeny of the orchid-bee genus *Euglossa* Latreille (Hymenoptera: Apidae), with emphasis on the subgenera E. (*Glossura*) Cockerell and E. (*Glossuropoda*) Moure. *Zoologischer Anzeiger*, **267**, 82–100.
- Frantine-Silva, W., Giangarelli, D.C., Penha, R.E.S., Suzuki, K.M., Dec, E., Gaglianone, M.C., Alves-dos-Santos, I., & Sofia, S.H. (2017) Phylogeography and historical demography of the orchid bee *Euglossa iopoeila*: signs of vicariant events associated to Quaternary climatic changes. *Conservation Genetics*, **18**, 539–552.
- Frantine-Silva, W., Sofia, S.H., Orsi, M.L., & Almeida, F.S. (2015) DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring. *Molecular Ecology Resources*, **15**, 1226–1237.
- Giangarelli, D.C., Freiria, G.A., Colatreli, O.P., Suzuki, K.M., & Sofia, S.H. (2009) *Eufriesea violacea* (Blanchard) (Hymenoptera: Apidae): an orchid Bee apparently sensitive to size reduction in forest patches. *Neotropical Entomology*, **38**, 610–615.
- Gregory-Wodzicki, K.M. (2000) Uplift history of the Central and Northern Andes: A review. *Geological Society of America Bulletin*, **112**, 1091–1105.
- Haffer, J. (1969) Speciation in Amazonian Forest Birds. *Science*, **165**, 131–137.
- Haffer, J. (2008) Hypotheses to explain the origin of species in Amazonia. *Brazilian Journal of Biology*, **68**, 917–947.
- Hoorn, C., Wesselingh, F.P., Ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T., & Antonelli, A. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.
- Jombart, T. & Ahmed, I. (2011) adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, **27**, 3070–3071.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., & Mayrose, I. (2015)

- CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, **15**, 1179–1191.
- Ledo, R.M.D. & Colli, G.R. (2017) The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of Biogeography*, **44**, 2551–2563.
- Longo, G. & Bernardi, G. (2015) The evolutionary history of the embiotocidi surfperch radiation based on genome wide RAD sequence data.pdf. *Molecular phylogenetics and evolution*, **88**, 55–63.
- López-Urbe, M.M., Zamudio, K.R., Cardoso, C.F., & Danforth, B.N. (2014) Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees. *Molecular Ecology*, **23**, 1874–1890.
- Lu, F., Glaubitz, J., Harriman, J., Casstevens, T., & Elshire, R. (2012) TASSEL 3.0 Universal Network Enabled Analysis Kit (UNEAK) pipeline documentation. *White Paper*, 1–12.
- Luebert, F. & Weigend, M. (2014) Phylogenetic insights into Andean plant diversification. *Frontiers in Ecology and Evolution*, **2**, 1–17.
- Maldonado-Coelho, M. (2012) Climatic oscillations shape the phylogeographical structure of Atlantic Forest fire-eye antbirds (Aves: Thamnophilidae). *Biological Journal of the Linnean Society*, **105**, 900–924.
- Martins, F. de M. (2011) Historical biogeography of the Brazilian Atlantic forest and the Carnaval-Moritz model of Pleistocene refugia: What do phylogeographical studies tell us? *Biological Journal of the Linnean Society*, **104**, 499–509.
- McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., & Brumfield, R.T. (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**, 526–538.
- Merow, C., Smith, M.J., & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, **36**, 1058–1069.
- Miller, M.J., Bermingham, E., Klicka, J., Escalante, P., do Amaral, F.S.R., Weir, J.T., & Winker, K. (2008) Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1133–1142.
- Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J.C., Valencia, V., Ayala, C., Pérez-Angel, L.C., Rodríguez-Parra, L.A., Ramirez, V., & Niño, H. (2015) Middle Miocene closure of the Central American Seaway. *Science*, **348**, 226–229.
- Moure, J., Melo, G.A.R., & Faria Jr, L.R.R. (2012) Available at:

<http://www.moure.cria.org.br/catalogue>.

- Muellner, A.N., Tremetsberger, K., Stuessy, T., & Baeza, C.M. (2005) Pleistocene refugia and recolonization routes in the southern Andes: Insights from *Hypochaeris palustris* (Asteraceae, Lactuceae). *Molecular Ecology*, **14**, 203–212.
- Nemésio, A. (2009) Orchid bees (Hymenoptera: Apidae) of the Brazilian Atlantic forest. *Zootaxa*, **2041**, 1–242.
- Nemésio, A. & Engel, M.S. (2012) Three new cryptic species of *Euglossa* from Brazil (Hymenoptera, apidae). *ZooKeys*, **222**, 47–68.
- Nemésio, A. & Santos Junior, J. (2014) Is the “Centro de Endemismo Pernambuco” a biodiversity hotspot for orchid bees? *Brazilian Journal of Biology*, **74**, S078–S092.
- ODEa, A., Lessios, H.A., Coates, A.G., et al. (2016) Formation of the Isthmus of Panama. *Science Advances*, **2**, e1600883–e1600883.
- Penha, R.E.S., Gaglianone, M.C., Almeida, F.S., Boff, S. V., & Sofia, S.H. (2015) Mitochondrial DNA of *Euglossa iopoecila* (Apidae, Euglossini) reveals two distinct lineages for this orchid bee species endemic to the Atlantic Forest. *Apidologie*, **46**, 346–358.
- Phillips, S., Anderson, R.P., & Schapire, R. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Phillips, S.J., Dudík, M., & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *21st International Conference on Machine Learning, Banff, Canada*, 655–662.
- Pokorny, T., Loose, D., Dyker, G., Quezada-Euán, J.J.G., & Eltz, T. (2015) Dispersal ability of male orchid bees and direct evidence for long-range flights. *Apidologie*, **46**, 224–237.
- Pritchard, J.K., Stephens, M., & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Ramalho, A.V., Gaglianone, M.C., & Oliveira, M.L. de (2009) Comunidades de abelhas Euglossina (Hymenoptera, Apidae) em fragmentos de Mata Atlântica no Sudeste do Brasil. *Revista Brasileira de Entomologia*, **53**, 95–101.
- Rambaut, A. (2014) FigTree 1.4. 2 software. Institute of Evolutionary Biology, Univ. Edinburgh. .
- Ramírez, S.R., Roubik, D.W., Skov, C.E., & Pierce, N.E. (2010) Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). *Biological Journal of the Linnean Society*, **100**, 552–572.
- Rochette, N.C. & Catchen, J.M. (2017) Deriving genotypes from RAD-seq short-read data

- using Stacks. *Nature Protocols*, **12**, 2640–2659.
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., & Huelsenbeck, J.P. (2012) MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.
- Roubik, D.W. & Hanson, P. (2004) *Abejas de orquídeas de la América tropical. Biología y guía de campo (Orchid bees of Tropical América) Biology and field guide*. INBio Press, Heredia, Costa Rica.
- Ruiz Pessenda, L.C., De Oliveira, P.E., Mofatto, M., de Medeiros, V.B., Francischetti Garcia, R.J., Aravena, R., Bendassoli, J.A., Zuniga Leite, A., Saad, A.R., & Lincoln Etchebehere, M. (2009) The evolution of a tropical rainforest/grassland mosaic in southeastern Brazil since 28,000 14C yr BP based on carbon isotopes and pollen records. *Quaternary Research*, **71**, 437–452.
- Rull, V. (2008) Speciation timing and neotropical biodiversity: The Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, **17**, 2722–2729.
- Stamatakis, A. (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.
- Tariel, J., Longo, G.C., & Bernardi, G. (2016) Tempo and mode of speciation in *Holacanthus* angelfishes based on RADseq markers. *Molecular Phylogenetics and Evolution*, **98**, 84–88.
- Thomé, M.T.C., Zamudio, K.R., Giovanelli, J.G.R., Haddad, C.F.B., Baldissera, F.A., & Alexandrino, J. (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution*, **55**, 1018–1031.
- Torkamaneh, D., Laroche, J., & Belzile, F. (2016) Genome-Wide SNP Calling from Genotyping by Sequencing (GBS) Data: A Comparison of Seven Pipelines and Two Sequencing Technologies. *PLoS ONE*, **11**, e0161333.
- Trujillo-Arias, N., Dantas, G.P.M., Arbeláez-Cortés, E., Naoki, K., Gómez, M.I., Santos, F.R., Miyaki, C.Y., Aleixo, A., Tubaro, P.L., & Cabanne, G.S. (2017) The niche and phylogeography of a passerine reveal the history of biological diversification between the Andean and the Atlantic forests. *Molecular Phylogenetics and Evolution*, **112**, 107–121.
- Werneck, F.P. (2011) The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. *Quaternary Science Reviews*, **30**, 1630–1648.

Wikelski, M., Moxley, J., Eaton-Mordas, A., López-Urbe, M.M., Holland, R., Moskowicz, D., Roubik, D.W., & Kays, R. (2010) Large-range movements of neotropical orchid bees observed via radio telemetry. *PLoS ONE*, **5**, e10738.

CAPÍTULO 2

ESTRUTURA POPULACIONAL DE *EUGLOSSA ANNECTANS* DRESSLER INDICA
FRAGMENTAÇÃO ENTRE LITORAL E INTERIOR AO LONGO DO PLEISTOCENO INFERIOR

5 CAPÍTULO 2: ESTRUTURA POPULACIONAL DE *EUGLOSSA ANNECTANS* DRESSLER INDICA FRAGMENTAÇÃO ENTRE LITORAL E INTERIOR AO LONGO DO PLEISTOCENO INFERIOR

RESUMO

Padrões biogeográficos na porção sul da Mata Atlântica (MA) são pouco explorados e os principais modelos de predição de adequabilidade ambiental falham em recuperar áreas de refúgio ao longo do último máximo glacial. Considerando a distribuição de *Euglossa annectans* Dressler 1982 associada à porção sul da MA em seus limites longitudinais, o presente estudo buscou avaliar a estrutura populacional e filogeográfica desta espécie visando testar a influência das flutuações climáticas do Quaternário sobre este tipo de distribuição. Para tanto, 129 machos tiveram 640 pb do gene mitocondrial *cytb* sequenciados e 359 machos foram genotipados para oito locos microssatélites. Adicionalmente, modelos de nicho ecológico foram construídos considerando três modelos de circulação geral e três estratégias de seleção de variáveis buscando avaliar a existência de áreas estáveis ao longo tempo. Os resultados indicam uma peculiar disjunção entre áreas de interior e litorâneas, coincidentes com os modelos climáticos gerados para o fim do Quaternário, que indicam um afastamento destes dois grupos intermediados por um processo de fragmentação em direção ao centro-oeste do continente. Em concordância com estudos anteriores, os padrões apresentados por *E. annectans* indicam uma grande instabilidade na porção sul da MA, diferente da situação de refúgios estáveis na porção ao norte da MA. A produção de machos diploides acima da média na espécie é confirmada, a qual, em conjunto com modelos de nicho ecológico projetados para o futuro, são considerados para proposição de medidas para a conservação desta espécie.

Palavras-chave: Euglossini, Refúgios, Quaternário, Machos Diploides, Estruturação Genética

5.1 INTRODUÇÃO

O domínio de Mata Atlântica (MA) constitui um dos mais complexos conjuntos de fitofisionomias do mundo, em grande parte determinado por sua ampla variação latitudinal, longitudinal e altitudinal (Galindo-Leal & Câmara, 2003). Dentre as fitofisionomias que constituem a MA destacam-se a Floresta Ombrófila Densa (FOD), Floresta Ombrófila Mista (FOM) e a Floresta Estacional Semidecidual (FES) (Ab'Sáber, 2003; Ribeiro et al., 2009). Ao lado de outras fitofisionomias de menor abrangência, estas três formações abrigam grandes níveis de endemismo para diferentes organismos (Cardoso da Silva et al., 2004; Quijano-Abril

et al., 2006; Tabarelli et al., 2010), incluindo as abelhas Euglossini (Ramírez et al., 2010; Garraffoni et al., 2017) representadas por cerca de 50 espécies ao longo da MA, das quais cerca de 25 são endêmicas (Nemésio, 2009; Moure et al., 2012).

Hipóteses baseadas nas variações climáticas do Quaternário tem se tornado uma das principais explicações para as altas taxas de diversificação na MA (Carnaval & Moritz, 2008; Carnaval et al., 2009, 2014; Porto et al., 2013). Em especial, o modelo de Carnaval e Moritz (2008; daqui em diante “Modelo CM”), o qual se baseia em modelagem de nicho ecológico a partir de amostragens aleatórias de pontos da MA, apresenta predições frequentemente corroboradas por estudos genético-populacionais e filogeográficos (Carnaval et al., 2009; Ossa et al., 2013; Batalha-Filho & Miyaki, 2016), onde refúgios ao norte da distribuição da MA teriam constituído zonas de estabilidade climática ao longo do último máximo glacial (LGM, do inglês *Last Glacial Maximum*). No entanto, as projeções do modelo CM não recuperam áreas estáveis para regiões tão ao sul (Carnaval & Moritz, 2008; Thomé et al., 2010), quanto as ocupadas por *Euglossa annectans* Dressler 1982. Por outro lado, outros estudos indicam situações mais complexas na MA (Thomé et al., 2010, 2014; Porto et al., 2013), levando à proposição de outras hipóteses como, por exemplo, a existência de barreiras geográficas (Thomé et al., 2014), neotectonismos (Batalha-Filho et al., 2010), e ainda a expansão da MA para a Plataforma Continental devido à regressão marinha ao longo dos períodos glaciais (Leite et al., 2016a).

Abelhas das orquídeas têm se mostrado modelos especialmente úteis para o estudo das variações climáticas do Quaternário (López-Uribe et al. 2014; Frantine-Silva et al. 2017) devido à sua intrínseca relação com seus respectivos habitats (Parra-H & Nates-Parra, 2012; Giangarelli et al., 2015a; Silveira et al., 2015). Dentre as espécies endêmicas da MA, *Euglossa annectans* apresenta uma distribuição mais ao sul e sudoeste, fortemente associada com a FES (Faria Jr & Melo, 2007; Giangarelli et al., 2015a), porém com diversos pontos de ocorrência em áreas próximas ao litoral na FOD (Faria Jr & Melo, 2007; Nemésio, 2009). *Euglossa annectans* apresenta ainda uma aparente sensibilidade à fragmentação ambiental (Giangarelli et al., 2015a) e uma produção de machos diploides expressivamente maior do que a média das espécies de Euglossini (Souza et al., 2010; Giangarelli et al., 2015b).

Neste contexto, o presente objetivou avaliar a influência das variações climáticas do Quaternário sobre a estruturação genética de *E. annectans* ao longo de sua distribuição. Para tanto, modelos de nicho ecológico gerados para essa espécie são confrontados com padrões de estruturação genética provenientes de marcadores moleculares mitocondriais e nucleares, comparando a congruência cronológica das variações climáticas com os padrões

demográficos de *E. annectans* nos últimos 130 Ka. Adicionalmente, dados de frequência de machos diploides e projeções futuras de adequabilidade ambiental são explorados à fim de fornecer informações para medidas de manejo desta espécie.

5.2 MATERIAIS E MÉTODOS

5.2.1 Área de Estudo e Amostragem

Doze remanescentes de FES e FOD da porção sul da Mata Atlântica foram amostrados para coleta de machos de *E. annectans* abrangendo uma área aproximada de 280 mil km² concentrando a maior parte da distribuição conhecida da espécie (Tabela 5.1).

Tabela 5.1 Descrição das áreas e número de machos de *Euglossa annectans* amostrados no presente estudo. A identificação das localidades (ID), localização e tipo florestal (TF), bem como latitude e longitude são apresentados ao lado do número de indivíduos amostrados. P.E. – Parque Estadual; PARNA - Parque Nacional; RPPN – Reserva Particular de Proteção da Natureza; APP – Área de Proteção Permanente; FOD – Floresta Ombrófila Densa; FES – Floresta Estacional Semidecidual.

ID	UF	Município	Localidade	TF	latitude	longitude	N
SO	RJ	Terezópolis	PARNA da Serra dos Órgãos	FOD	-22,493	-43,0733	40
BO	SP	São Sebastião	P.E. da Serra do Mar, Boraceia	FOD	-23,631	-45,8697	9
PN	SP	Mogi das Cruzes	Parque das Neblinas	FOD	-23,748	-46,1633	22
IC	SP	Cananéia	P.E. da Ilha do Cardoso	FOD	-25,305	-48,8967	11
SU	PR	Superagui	PARNA do Superagui	FOD	-25,331	-48,1558	60
SM	PR	Salto Morato	RPPN de Salto Morato	FOD	-25,182	-48,2978	46
SC	SC	Joinville	Área adjacente a Joinville	FOD	-26,207	-49,0431	35
KB	PR	Telemaco Borba	RPPN Fazenda Monte Alegre	FES	-24,212	-50,5572	26
MD	SP	Teodoro Sampaio	P.E. do Morro do Diabo	FES	-22,621	-52,1717	40
DN	PR	Diamante do Norte	Área de Proteção Permanente	FES	-22,618	-52,8564	35
FZ	PR	Foz do Iguaçu	PARNA do Iguaçu	FES	-25,695	-54,4367	60
DR	MS	Dourados	APP Serra da Picadinha	FES	-22,142	-54,7317	14

As amostragens foram conduzidas seguindo o protocolo descrito por Giangarelli et al. (2009) utilizando iscas odoríferas (eucaliptol, eugenol, salicilato de metila, benzoato de benzila, vanilina, cinamato de metila e acetato de benzila), além de armadilhas (Ramalho et al. 2009) com as mesmas essências entre os anos de 2009 e 2016. Amostras do Parque das Neblinas foram gentilmente doadas pelo Dr. Guaraci Ribeiro, bem como as amostras de Joinville foram igualmente cedidas pelo Me. Enderley Dec e Dra. Isabel Alves-dos-Santos.

5.2.2 Procedimentos Moleculares

A extração do DNA total foi realizada segundo os protocolos descrito por Freiria et al. (2012) e Penha et al. (2015) a partir de amostras da musculatura torácica ou da perna posterior direita. Indivíduos representativos de amostragens pertencentes ao mesmo evento de coleta foram depositados na coleção de abelhas do Museu Zoológico da Universidade Estadual de Londrina (MZUEL).

As amplificações e sequenciamento de 640 pb do gene mitocondrial *citocromo b* (*cytb*) foram realizadas de acordo com Frantine-Silva et al. (2017) utilizando uma combinação de *primers* desenhados a partir das sequências de *Euglossa iopoecila* (*Eiop-F1*: 3'-CGA GGT CGC AAA CAT CT-5') com o *primer* AMB 16 (Crozier et al., 1991; Arias et al., 2008), com 48°C como temperatura de anelamento. Após amplificação e verificação de positivo em gel de agarose 1%, 8 µL dos produtos de PCR foram enzimaticamente purificados para remoção de resíduos de primers e nucleotídeos com o *kit* Illustra™ Exo Pro-Star 1-Step (GE Healthcare, Buckinghamshire, Reino Unido) seguindo o protocolo do fabricante. As PCRs purificadas foram então submetidas a reações de sequenciamento utilizando o *kit* BigDye Terminator v3.1 (Applied Biosystems, Foster City, CA, EUA). Finalmente, os fragmentos de PCR foram sequenciados em sequenciador automático ABI-PRISM 3500-XL (Applied Biosystems, Foster City, CA, EUA).

O protocolo descrito por Penha et al (2015) foi utilizado para amplificação de oito locos microssatélites descritos para *Euglossa annectans*: Ann2, Ann3, Ann4, Ann6, Ann8, Ann24, Ann37 e Ann41 (Paxton et al., 2009), seguindo as mesmas temperaturas de anelamento originalmente descritas para os respectivos locos. Após a amplificação os fragmentos gerados foram lidos em sequenciador automático ABI3500-XL (Applied Biosystems, Foster City, CA, EUA).

5.2.3 Análise dos Dados Moleculares

5.2.3.1 DNA mitocondrial

As sequências provenientes do gene *cytb* foram alinhadas via algoritmo MUSCLE (Edgar, 2004) integrado no *software* MEGA v.6.05 (Tamura et al., 2013), o qual também foi empregado para posterior checagem e edição manual das sequências. A fim de evitar a análise de possíveis picos duplos, o pacote SangerSeq do *software* R (R Core Team, 2017) foi empregado para detectar picos com relação de intensidade maior do que 1/3 do pico principal.

Apenas um loco foi identificado com este padrão para algumas amostras e por estar próximo do início do sequenciamento com baixa qualidade, a posição foi excluída das análises.

As redes de relacionamento de haplótipos foram inferidas por meio do *software* PopArt v1.7 (Leigh & Bryant, 2015). O *software* DnaSP 5.0 (Librado & Rozas, 2009) foi utilizado para categorização das sequências em haplótipos, além de calcular a diversidade nucleotídica e haplotípica, frequência de diferenças entre pares de sequências (distribuição de *mismatch*), detecção de variações demográficas com base em modelos de tamanho efetivo populacional constante via testes *D* de Tajima (Tajima, 1989) e *F_s* de Fu (Fu, 1997) sem grupo externo, bem como estatística *R₂* (Ramos-Onsins & Rozas, 2006) recomendada para pequenos tamanhos de população (Ramírez-Soriano et al., 2008). Por fim, o módulo de simulações coalescentes do DnaSP 5.0 foi empregado para testar a significância estatística dos testes de neutralidade com base no valor *theta* sem recombinação, com um intervalo de confiança de 95% e 10.000 replicações.

Com o intuito de testar a sincronia da variação do tamanho efetivo populacional em relação às mudanças climáticas, foi empregado o método *Extended Bayesian Skyline Plot* (EBSP) (Drummond & Bouckaert, 2015) disponível no pacote de *software* BEAST 2.3.2 (Drummond & Rambaut, 2007). Este método é especialmente indicado frente a árvores parcialmente resolvidas ou pouco variáveis (Ho & Shapiro, 2011), como o esperado para análise intraespecífica de genes codificantes. Para tanto, quatro corridas independentes foram realizadas, cada uma das quais com 100 milhões de passos para as Cadeias de Markov e Monte Carlo, utilizando uma árvore inicial de UPGMA e modelo de substituição HKY+I+G, como indicado pelo *software* jModelTest 1.1 (Posada, 2008) com base no Critério de Informação de Akaike com correção para tamanho amostral reduzido (AICc). Árvores e parâmetros foram amostrados a cada 10.000 cadeias. As corridas foram combinadas usando o *software* LogCombiner (pacote BEAST) para analisar a convergência das cadeias com *burning* de 25%. A datação das variações do tamanho efetivo populacional foi baseada em um relógio molecular estrito com distribuição *log-normal* e taxa de substituição de 1,3% (d.p. 0,2%) por milhão de ano (Ma⁻¹), indicada para o DNAmT em abelhas (Dick et al., 2004; Batalha-Filho et al., 2010). A convergência dos parâmetros ao longo das cadeias MCMC e o desempenho EBSP (valores ESS > 200) foi observada pelo *software* Tracer v1.6 do pacote BEAST (disponível em: <http://beast.bio.ed.ac.uk/tracer>).

A estrutura populacional e espacial dos dados provenientes do DNA mitocondrial foi acessada pelo *software* BAPS v6.2 (Análise Bayesiana da Estrutura da População) (Corander et al., 2008). BAPS foi utilizado com a opção de agrupamento espacial, onde as coordenadas

geográficas foram combinadas com dados genéticos para definir o número K mais provável para as populações, variando de 1 a 15. Com base nos resultados das análises nucleares uma série de AMOVAs hierárquicas e comparações Φ_{CT} (10000 permutações) foram realizadas no *software* Arlequin v3.11 (Excoffier et al., 2005). Os cenários testados incluíram um único grupo e dois grupos: grupo “*litoral*” e grupo “*interior*”. Padrões de isolamento por distância foram testados através do teste de Mantel (1000 permutações) no *software* GenAlex (Peakall & Smouse, 2006). A distância geográfica em quilômetros foi plotada contra a estimativas de Φ_{ST} par-a-par.

5.2.3.2 *Microssatélites*

Os dados provenientes das leituras dos fragmentos via sequenciador automático foram analisados pelo *software* GeneMarker v. 2.2.0. O *software* Micro-Checker (Van Oosterhout et al., 2004) foi utilizado para checar possíveis falhas de genotipagens como *dropout* ou erros de leitura. A partir destes dados, medidas descritivas e de diversidade genética foram realizadas pelo *software* GenAlEx (Peakall & Smouse, 2006), tais como: média do número de alelos para todos os locos em cada população (\bar{N}_A), média do número de alelos efetivos (\bar{N}_E), número de alelos privados (N_P), além da média da heterozigosidade esperada (\bar{H}_e).

As estatísticas relacionadas à produção de machos diploides (DMP) foram acessadas segundo o protocolo utilizado por Boff et al. (2014) e Giangarelli et al., (2015), avaliando-se o número absoluto (N_{MD}) e a frequência relativa de machos diploides em relação ao total (f_{MD}). A probabilidade de o conjunto de dados detectar um indivíduo heterozigoto, caso o mesmo seja diploide (P_{het}) foi calculada de acordo com Souza et al. (2010). Embora alelos nulos putativos possam causar dificuldades na leitura de microssatélites podendo levar à uma superestimativa do P_{het} (Souza et al., 2010), os dados apresentados por Giangarelli et al. (2015) demonstram que a presença de alelos nulos é frequentemente muito baixa e causa pouca diferença prática entre o P_{het} com ou sem a presença de alelos nulos. Deste modo, no presente estudo serão apresentados apenas os dados de P_{het} . Um intervalo de confiança de 95% em relação à frequência de machos diploides foi calculado assumindo uma distribuição binomial (Souza et al., 2010; Boff et al., 2014; Giangarelli et al., 2015b) pelo pacote “binom” do R (Sundar, 2014).

Diferentes abordagens foram utilizadas para acessar a estruturação genética dentre as diversas amostras. O *software* Structure (Pritchard et al., 2000) foi utilizado para estimar o número de grupos (K) mais prováveis via análise Bayesiana. Estimativas do número de K -grupos mais prováveis foram obtidas por simulações de K -grupos variando entre 1 e 15,

incluindo 20 interações por hipótese de K ao longo de cadeias de MCMC com 1.000.000 de passos e descarte de 100.000. O aplicativo *online* CLUMPAK (<http://clumpak.tau.ac.il/>) (Kopelman et al., 2015), foi utilizado para sumarizar os resultados do Structure via método de ΔK de Evanno (Evanno et al., 2005), bem como o $LogK$ e Δln . Complementarmente, uma Análise de Variância Molecular (AMOVA) foi executada frente a dois cenários: o primeiro considerando apenas um grupo, indicado pelo DNA mitocondrial; e um segundo cenário considerando dois grupos, como indicado pelos resultados do software *Structure*. Valores de índices de estruturação genética global e par-a-par foram computados pelo software Arlequin 3.5 através da análise de AMOVA (Φ_{ST}), além de outras estatísticas análogas ao F_{ST} , como G_{ST} de Ney, G'_{ST} de Hendricks (2005) e $Dest$ de Jost (2008). Os valores linearizados de G'_{ST} foram utilizados para investigar os padrões de isolamento por distância entre as amostras estudadas através do teste de Mantel com 1000 permutações.

Por fim, com intuito de inferir sobre taxas de migração e fluxo gênico recente entre as localidades estudadas, o *software* BayesAss 3.0 (Wilson & Rannala, 2003) foi utilizado para estimar a proporção relativa de migrantes entre localidades, bem como a direção do fluxo gênico através de uma abordagem Bayesiana. Quatro corridas independentes foram realizadas com 1.000.000 de gerações de MCMC, amostrando parâmetros a cada 1000 gerações com *burning* de 100.000 gerações, com demais parâmetros em *default*.

5.2.4 Modelagem de Nicho Ecológico

Visando analisar a congruência entre os padrões de estruturação genética e as variações climáticas do Quaternário em relação às populações de *E. annectans*, modelos de adequabilidade ambiental e de paleodistribuição foram construídos por meio do algoritmo Máxima Entropia (Phillips et al., 2004) implementado no software MaxEnt v 3.3.3 k (Phillips et al., 2006). Para tanto, três diferentes abordagens foram utilizadas quanto à escolha das variáveis, como sugerido por Porfirio et al. (2014): a) utilizando todas as 19 variáveis bioclimáticas disponíveis (Hijmans et al., 2005), denominada estratégia “*WorldClim*”; b) selecionando variáveis com maior potencial de explicação dos dados via análise de componentes principais, denominada estratégia “*PCA*”; c) utilizando variáveis ambientais ecologicamente significativas, BIO 4 (sazonalidade de temperatura), BIO 5 (temperatura do quartil anual mais úmido), BIO 9 (temperatura do quartil mais seco), BIO 16 (precipitação do quartil mais úmido), BIO 17 (precipitação do quartil mais seco), as mesmas utilizadas por Nemésio et al. (2016). Após seleção via PCA um conjunto de 11 variáveis foi selecionado compreendendo a variação de 95% do conjunto de dados (em ordem de contribuição: BIO 10,

BIO 5, BIO 14, BIO 17, BIO 1, BIO 9, BIO 19, BIO 3, BIO 12, BIO 15 e BIO 4), excluindo-se as variáveis BIO 8 e BIO 11 devido à alta correlação com as variáveis BIO 10 e BIO 9 ($r^2 > 0,8$; $p < 0,001$), respectivamente.

Todos os modelos foram baseados em 71 ocorrências georreferenciadas de *E. annectans* (Ver Apêndices) obtidas na literatura disponível, na coleção de Abelhas do Museu de Zoologia da Universidade Estadual de Londrina e no banco de dados *SpeciesLink* quando havia referência do material analisado. Neste sentido, amostras identificadas como *Euglossa stellfeldi* em Nemésio (2009) foram tratadas como *E. annectans*. Os modelos gerados a partir destas ocorrências e das diferentes estratégias para seleção de variáveis, foram avaliadas segundo os valores de área sob a curva (AUC), que corresponde à relação entre especificidade e sensibilidade do modelo em relação à possível ocorrência da espécie (Merow et al., 2013), além dos testes do critério de informação de Akaike (AIC) com correção para tamanhos populacionais reduzidos (AICc) e critério de informação Bayesiano (BIC), além do log de probabilidade dos modelos (*LogL*) por meio do software ENMtools .

Os modelos gerados foram então projetados para quatro cenários diferentes: condições climáticas “atuais”, médio Holoceno (6 Ka a. p.); último máximo glacial (21 Ka a. p.) e último período interglacial (130 Ka a. p.). Adicionalmente, os modelos foram projetados para um cenário futuro, a fim de identificar padrões úteis a medidas de conservação. O cenário futuro corresponde à projeção para 2050 considerando um padrão moderado de acúmulo de gases-estufa com pico em 2045, seguido de estabilização a partir de então (RCP 4.5) segundo o Painel Intergovernamental sobre Mudanças Climáticas (IPCC).

Para todos os cenários, três Modelos de Circulação Geral (*Circulation General Models*, GCM) foram utilizados, incluindo o *Community Climate System Model* versão 4.0, CCSM4 (Gent et al., 2011); o *Max-Planck Institute - Earth System Model*, MPI-ESM (Giorgetta et al., 2013) e o *Model for Interdisciplinary Research on Climate - Earth System Model*, MIROC-ESM (Watanabe et al., 2011). Estes três modelos apresentam especificidades diferentes, nas diferentes projeções ao longo do tempo e considerando a exposição da plataforma continental no último máximo glacial (CCSM4 e MPI-ESM-LR) ou não (MIROC-ESM). As 19 variáveis climáticas de cada um dos GCMs foram baixadas em resolução de 2,5 minuto de arco (arcmin), e reprojatadas para o mesmo *datum* e resolução das variáveis “atuais” (WGS84, 30 segundos de arco – arcsec, ~1 Km) por meio do pacote “raster” do *software R*. Considerando a distribuição de *E. annectans*, os modelos de nicho ecológico (ENM) foram gerados para o espaço geográfico entre as latitudes 5 e -35, e longitudes -73 e -34.

O *software* MaxEnt v 3.3.3 foi utilizado para gerar os modelos de adequabilidade ambiental utilizando os parâmetros *default*, retendo 25% das amostras para validação cruzada e os demais 75% como dados de treinamento do modelo, ao longo de 10 replicações. Para produzir os mapas binários, utilizou-se os limiares de igual sensibilidade e especificidade, os quais reduzem a diferença entre a taxa de omissão e a área prevista (Elith et al., 2011; Merow et al., 2013). Os modelos resultantes do MaxEnt foram visualizados no *software* QGIS 2.14 (Quantum GIS Team, 2016), onde as áreas com valores de adequabilidade maiores que 0,5 foram mantidas para a construção de mapas de ocorrência, as quais foram combinadas por multiplicação do canal *alpha* com fator 0,33 para gerar os mapas de consenso entre os diferentes GCMs. Por fim, um consenso múltiplo foi gerado para cada uma das projeções pretéritas a partir do consenso dos GCMs.

5.3 RESULTADOS

5.3.1 DNA mitocondrial

Sequências do gene *citocromo b* foram analisadas a partir de amostras de 112 machos de *E. annectans*, resultando em 640 pb, com 19,2% de conteúdo G+C, sete sítios polimórficos parcimoniosamente informativos, os quais combinaram-se em oito diferentes haplótipos (*h*) distribuídos entre as 12 localidades amostradas (Tabela 5.2).

Tabela 5.2 Índices de diversidade genética e testes de neutralidade para 640pb do gene citocromo *b* de *Euglossa annectans*: (*N*) número de amostras, (*S*) número de sítios segregantes, (*h*) número de haplótipos, (*Hd*) diversidade haplotípica, (π) diversidade nucleotídica, (*s.d.*) desvio padrão, (*F_s*) teste de neutralidade de Fu, (*D*) teste *D* de Tajima e *R₂* de Ramos-Onsin e Rozas.

Localidade	<i>N</i>	<i>S</i>	<i>h</i>	<i>Hd</i> - <i>sd.</i>		π^*	<i>F_s</i>	<i>D</i>	<i>R₂</i>
SO	9	3	4	0,690	0,147	0,156%	-1,039	-0,359	0,167
PN	10	0	1	0,000	0,000	0,000%	0,000	0,000	0,000
BO	10	2	3	0,378	0,154	0,063%	-1,164**	-1,401**	0,200
IC	10	2	3	0,664	0,117	0,118%	-0,046	0,222	0,204
SU	7	4	4	0,810	0,129	0,179%	0,263	-1,434	0,262
SM	10	4	3	0,689	0,104	0,253%	1,485	0,566	0,203
SC	10	0	1	0,000	0,000	0,000%	0,000	0,000	0,000
KB	10	2	3	0,600	0,131	0,104%	-0,272	-0,184	0,190
MD	9	1	1	0,000	0,000	0,035%	-0,263	-1,088	0,314
DN	9	1	4	0,750	0,112	0,087%	1,015	1,401	0,278
FZ	8	2	5	0,857	0,108	0,134%	-0,071	0,414	0,214
DR	10	1	2	0,356	0,159	0,056%	0,417	0,015	0,178
Todas as amostras	112	7	8	0,520	0,051	0,110%	-3,382	-1,079	0,050**

*Os dados de desvio padrão para π foram omitidos por serem menores que $x10^{-3}$; ** $p \leq 0,05$

Os índices descritivos de diversidade genética para os dados de DNAmT indicaram uma importante variação entre as localidades amostradas. A fixação haplotípica observada em PN, SC e MD ocasionaram a ocorrência dos menores valores de diversidade haplotípica (Hd) e nucleotídica (π) para essas localidades, enquanto os maiores valores de Hd e π foram respectivamente observados em FZ ($Hd = 0.857$) e SM ($\pi = 0.253\%$). Os padrões de compartilhamento de haplótipos, por outro lado, revelam uma homogeneidade na distribuição dos haplótipos ao longo da distribuição de *E. annectans* amostrada. O haplótipo H3 mostrou-se o mais frequente dentre todos os haplótipos, sendo encontrado em todos os pontos amostrados, bem como o haplótipo H2 em oito das doze localidades estudadas com importante representatividade. Haplótipos privados foram observados apenas em SO (H4, H1) e FZ (H8) (Figura 5.1-A).

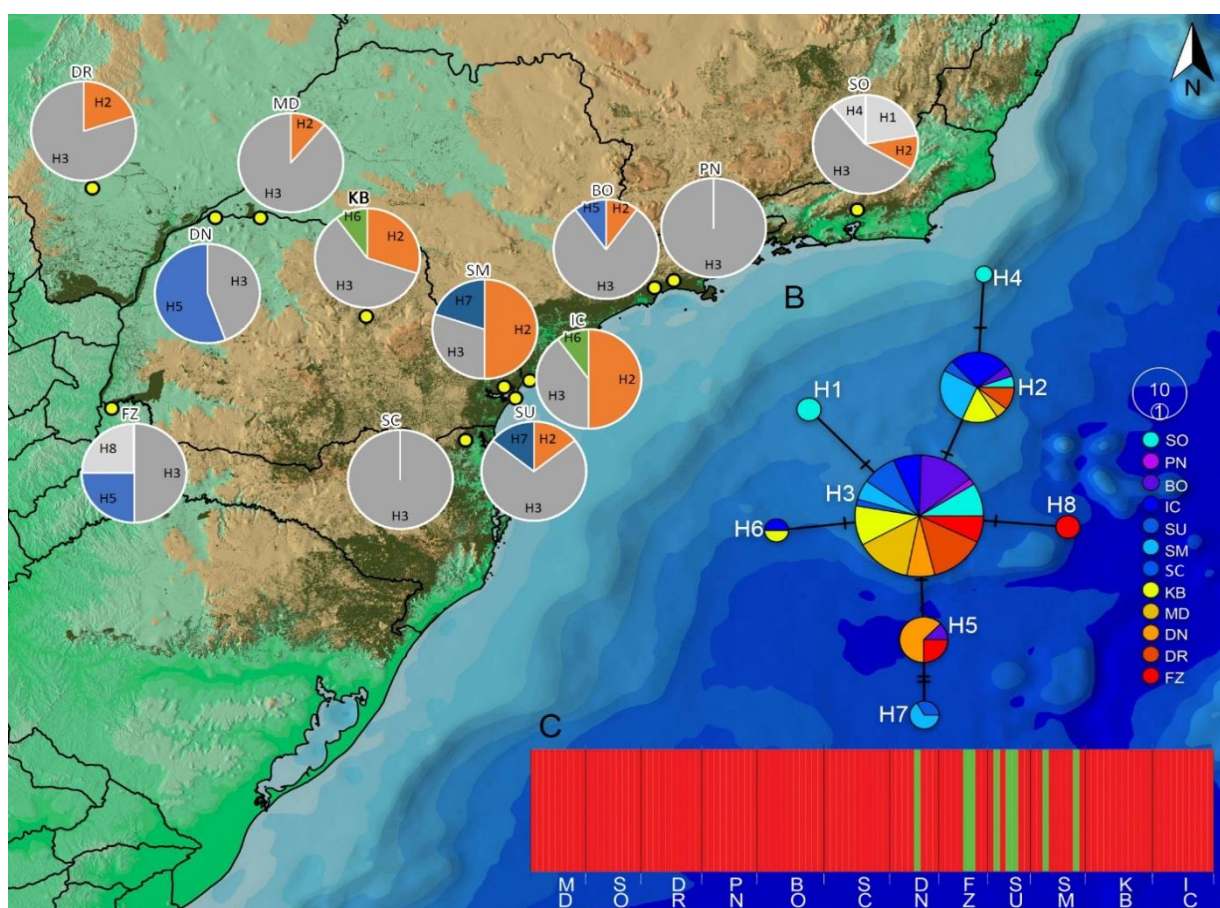


Figura 5.1 Distribuição geográfica dos haplótipos e estruturação genética mitocondrial de *Euglossa annectans* ao longo dos 12 locais de estudo (SO, PN, BO, IC, SU, SM, SC, KB, MD, DN, DR e FZ), representados por pontos amarelos (Software QGIS 2.12). A) Mapa parcial do Brasil e distribuição das localidades de amostragem; gráficos de pizza estão posicionados próximos aos respectivos pontos indicando a frequência e distribuição de cada haplótipo; os haplótipos não compartilhados são mostrados em cinza claro. B) Rede de haplótipos produzida pelo método *Median-Joining* para os 8 haplótipos amostrados (H1-H8) obtidos para as sequências do gene *cytb*; tamanho dos círculos são proporcionais à frequência dos haplótipos; cada barra transversal sobre as linhas conectoras representa um passo mutacional; C) Análise Bayesiana de Estrutura Populacional (*BAPS*) de *E. annectans* a partir de 640 pb do gene *cytb*. Cada indivíduo é representado por uma barra vertical associado a um cluster genético representado por uma cor respectiva.

A relação mutacional entre os haplótipos, reconstruída pelo método *Median-Joining*, apresentou uma configuração simples de estrela com um haplótipo central mais frequente que está relacionado com os demais haplótipos por apenas um ou dois passos mutacionais (Figura 5.1-B). Esta estrutura está tipicamente relacionada com cenários de expansão populacional pós gargalos genéticos ou efeito do fundador (Ramírez-Soriano et al., 2008). Esta relação simples entre os haplótipos é corroborada pela análise Bayesiana de Estrutura Populacional (BAPS) a qual encontrou dois *clusters* genéticos para as amostras estudadas, porém sem estruturação geográfica definida (Figura 5.1-C). O mesmo pode ser observado pela análise de variância molecular (AMOVA) a qual indicou valores de estruturação e proporção de variação genética não significativos, mesmo quando incluído dados espaciais (SAMOVA) em uma análise hierárquica de definição de grupos de populações (dados não apresentados) corroborando a existência de um único grupo de amostras (Tabela 5.3).

Tabela 5.3 Análise de Variância Molecular (AMOVA) considerando a existência de um único grupo de *populações a partir de 640 pb do gene *cytb* de 126 amostras de *Euglossa annectans* coletadas em 12 localidades ao longo do sul da Mata Atlântica.

Fonte da variação	Graus de liberdade	Soma dos quadrados	Componente de variação	% de variação	Φ_{ST} -global
Entre populações	11	8,393	0,04887 Va	13,72	0,137**
Dentro de populações	100	30,731	0,30731 Vb	86,28	
Total	111	39,124	0,35618		

*População neste contexto refere-se às amostras coletadas em uma mesma localidade; Todos os valores de p foram significativos com $\alpha = 0,001$ testado por 1000 permutações aleatórias.

Os testes demográficos para expansão ou retração populacional não apresentaram fortes evidências para estes cenários quando as amostras foram analisadas de acordo com suas respectivas localidades (veja Apêndice K). Embora a distribuição das frequências das diferenças entre pares de sequências (distribuição *Mismatch*) tenha evidenciado padrões alinhados com o esperado para populações sob efeito de expansão populacional, tanto para as populações individualizadas (veja Apêndices) quanto para todas as amostras analisadas como um único grupo (Figura 5.2), os testes de neutralidade F_s de Fu e D de Tajima não apresentaram sinais significativos de *bottleneck* ou expansão populacionais, salvo para a localidade BO (Tabela 5.2). De igual modo, o teste de R_2 apresentou sinais significativos de redução populacional recente quando todas as amostras foram analisadas como uma única população.

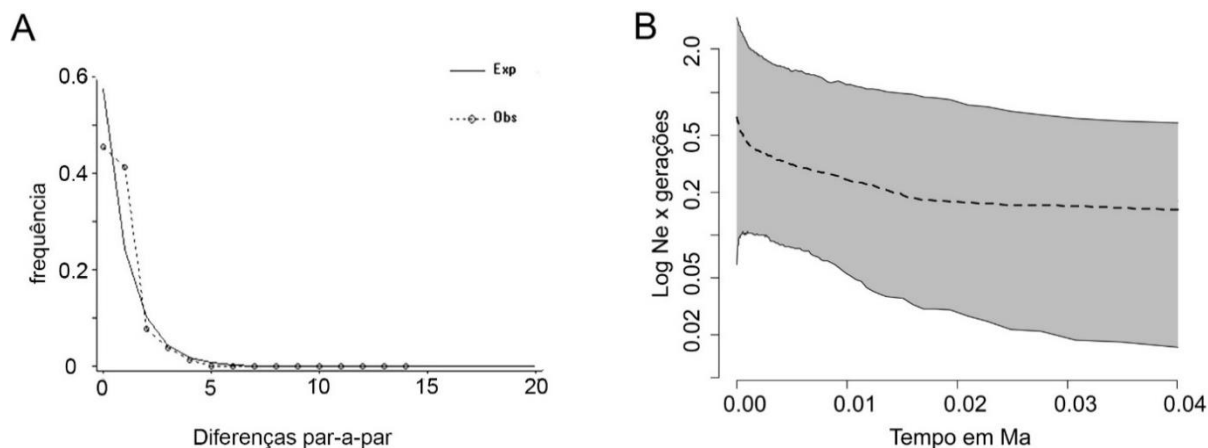


Figura 5.2 Estimativa de mudanças demográficas populacionais para o gene *cytb* em 126 amostras de *E. annectans* ao longo da Mata Atlântica. A - Distribuição *mismatch* par-a-par comparando distribuição esperada para cenário de crescimento populacional (linha sólida; Exp.) e a distribuição observada (linha pontilhada; Obs.); B - Inferência de mudanças do tamanho efetivo populacional (N_e em *log*) ao longo do tempo em milhões de anos.

A estimativa Bayesiana de mudança do tamanho efetivo populacional ao longo do tempo (*Extended Bayesian Skyline by Plot*) demonstra um sensível aumento de cerca de quatro vezes no tamanho efetivo populacional para o gene *cytb* de *E. annectans* iniciando-se por volta de 15 Ka a. p., o que indica uma possível expansão populacional pós-glacial.

5.3.2 DNA nuclear (Microssatélites)

No total, foram genotipados 398 machos de *E. annectans* ao longo das 12 localidades amostradas, dos quais 39 apresentaram um ou mais locos em heterozigose. Com exceção das amostras provenientes de BO, em todas as demais localidades foram detectados entre um e seis machos diploides, representando frequências entre 0,057 e 0,182 em suas respectivas populações. A probabilidade de um macho, caso diploide, ser detectado por heterozigosidade em ao menos um dos locos (*Phet*) esteve sempre muito próximo de um variando entre 0,999551 (PN) e 0,999997 (DN), indicando que o conjunto de dados é suficientemente discriminatório para detectar a presença de machos diploides na população (Tabela 5.4).

Tabela 5.4 Índices de diversidade genética e estimativas de machos diploides para amostras de *E. annectans* coletadas em 12 localidades da região sul da Mata Atlântica. (N_T) Número total de machos genotipados; (N_{MD}) número de machos diploides; (f_{MD}) frequência de machos diploides; ($CI^{95\%}$) intervalo de confiança de 95% para f_{MD} ; (*Phet*) Probabilidade de um macho, caso diploide, ser heterozigoto ao menos em um loco; (N_{MH}) Número de machos haploides incluídos nas análises populacionais; (\bar{N}_A) número médio de alelos; (*sd*) desvio padrão; (\bar{N}_E) Número médio de alelos efetivos; (N_P) número de alelos privados; (*He*) Heterozigosidade esperada (*sensu* Nei, 1978).

	N_T	N_{MD}	f_{MD}	$CI^{95\%}$	<i>Phet</i>	N_{MH}	\bar{N}_A	<i>sd</i>	\bar{N}_E	<i>sd</i>	N_P	<i>He</i>	<i>sd</i>
SO	40	6	0,150	0,04-0,27	0,999949	34	8,5	1,32	5,2	0,71	2	0,78	0,03
BO	9	--	--	--	0,999978	9	4,1	0,44	3,0	0,45	1	0,64	0,06
PN	22	4	0,182	0,05-0,40	0,999853	18	5,5	0,93	3,9	0,55	0	0,73	0,05
IC	11	2	0,182	0,03-0,60	0,999990	9	5,3	0,53	4,3	0,57	0	0,78	0,03
SU	46	6	0,125	0,06-0,30	0,999936	40	7,1	0,74	3,7	0,37	2	0,72	0,03
SM	60	4	0,067	0,02-0,16	0,999990	56	8,0	0,76	3,9	0,62	1	0,70	0,04
SC	35	1	0,029	0,00-0,15	0,999992	34	7,1	0,97	4,7	0,82	5	0,74	0,05
KB	26	3	0,115	0,02-0,30	0,999730	23	5,8	0,56	3,5	0,54	0	0,67	0,05
MD	40	6	0,150	0,06-0,30	0,999995	34	7,9	1,27	5,1	0,91	1	0,77	0,04
DN	35	2	0,057	0,00-0,19	0,999997	33	7,4	1,19	4,7	0,85	3	0,75	0,05
FZ	60	4	0,067	0,02-0,16	0,999551	56	9,9	1,63	5,3	1,01	1	0,76	0,04
DR	14	1	0,077	0,01-0,33	0,999961	13	4,6	0,46	3,2	0,45	0	0,65	0,08
Total	398	39	0,098	0,07-0,13	0,999990	359	13,75	5,62	4,2	0,20	--	0,73	0,01

As análises de diversidade genética incluíram os 359 machos haploides genotipados, os quais apresentaram ao longo dos oito locos analisados uma média de 12,87 alelos por loco variando de oito (Ann6 e Ann41) a 20 alelos (Ann24). Considerando cada população individualmente, foram observadas médias de número de alelos por loco variando de 4,13 ($dp = 0,44$) em BO a 9,88 ($dp = 1,63$) em FZ. A média do número de alelos efetivos teve uma variação similar, levemente inferior, com o menor número observado em DR (3,2) e maior número em FZ (5,3). Dentre os 110 alelos identificados para os oito locos em toda a população analisada, 16 alelos foram observados como privados em oito das 12 localidades analisadas, dos quais apenas seis tiveram frequência maior que 5% em três localidades (BO, SC e DN). Por fim, a heterozigosidade esperada apresentou baixa variação entre as localidades, variando de 0,64 a 0,78 em BO e SO e IC respectivamente.

O resultado da análise Bayesiana de estruturação genética indicou que o número de grupos (K) mais provável para as amostras analisadas é $K=2$, tanto pelo $\ln Pr(X/K)$ quanto para “delta K ” (ΔK). De acordo com este resultado, as localidades amostradas podem ser agrupadas em dois conjuntos, o primeiro formado por amostras localizadas na zona litorânea

de distribuição de *E. annectans* (SO, BO, PN, IC, SU, SM, SC) e um segundo grupo formado por amostras com distribuição interioranas (KB, MD, DN, FZ e DR) (Figura 5.3).

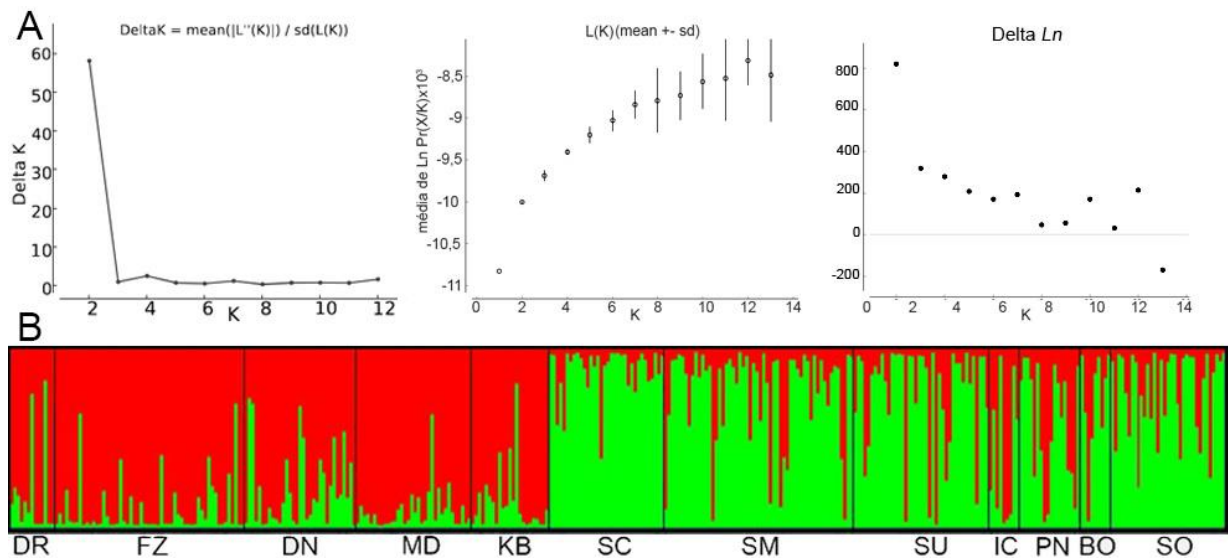


Figura 5.3 Análise Bayesiana de estruturação genética baseada em oito locos microssatélites de *E. annectans*. A) da esquerda para direita: estimativa do número de agrupamentos (K) mais provável com base na variação da média da razão do $\log \text{Pr}(X/K)$ pelo desvio padrão de $(L(K))$, seguido pela média $L(K) \pm$ o desvio padrão (centro) e a variação de \log de $L(K)$ (à direita); B) gráfico de barras para o coeficiente de ancestralidade (Q) considerando $K=2$, onde cada barra representa um indivíduo; as cores vermelho e verde identificam os diferentes *clusters* genéticos inferidos pelo software *Structure*.

A análise de variância molecular (AMOVA) corrobora a estruturação entre grupos apontada pela inferência Bayesiana, indicando uma estruturação baixa, porém significativa entre os grupos “interior” e “litoral” ($\Phi_{CT} = 0,0476$; $p\text{-value} = 0,002$). Esta mesma análise também indicou um maior percentual de variação na diferença entre as amostras dentro da mesma localidade (87,79%), refletindo o alto conteúdo informativo do conjunto de dados para distinção individual. Contudo, uma importante parcela da variação foi encontrada na comparação entre as localidades do mesmo grupo (7,45%; $\Phi_{SC} = 0,078$; $p\text{-value} < 0,05$), bem como na diferenciação entre os grupos (4,76%; Tabela 5.5).

Tabela 5.5 Análise de Variância Molecular (AMOVA) considerando dois grupos de *populações a partir de oito locos microssatélites de 359 machos de *Euglossa annectans* coletadas em 12 localidades ao longo do sul da Mata Atlântica.

Fonte de variação	Graus de liberdade	Soma dos quadrados	Componentes de variância	% de variação	Índice de estruturação
Entre grupos	1	74,854	0,15225 Va	4,76	Φ_{CT} : 0,048
Entre "populações" dentro dos grupos	10	162,988	0,23828 Vb	7,45	Φ_{SC} : 0,078
Dentro das "populações"	706	1982,242	2,80771 Vc	87,79	Φ_{ST} : 0,122
Total	717	2220,084	3,19823		

*População neste contexto refere-se às amostras coletadas em uma mesma localidade; Todos os valores de p foram significativos com $\alpha = 0,05$ testado por 1000 permutações aleatórias.

A análise Bayesiana de inferência de fluxo gênico indicou a existência de duas principais rotas migratórias entre as localidades amostradas, ambas concentradas dentro dos respectivos grupos definidos pela análise de estruturação populacional com baixas taxas de migração entre os dois grupos. Dentre as amostras do grupo “*interior*”, taxas de migração entre 20,2% e 28,2% foram identificadas na direção FZ -> (DR, DN, MD e KB), com importantes taxas de migração de KB à FZ (26,2%). Já dentre as amostras do grupo “*litoral*”, taxas de fluxo gênico mais variáveis foram identificadas entre as localidades, variando de 3% (SM-IC) à 29% (SM-SU) (Figura 5.4) .

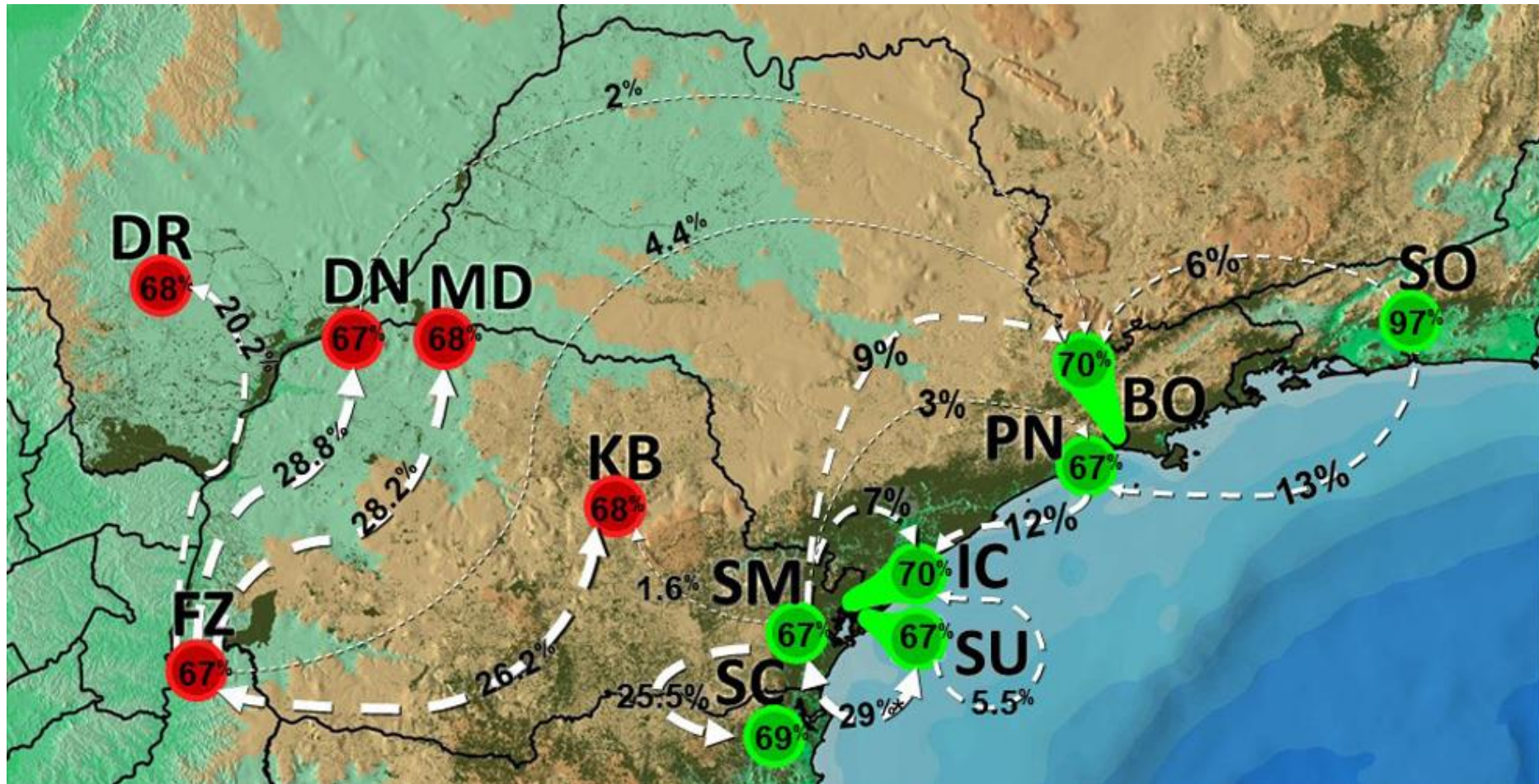


Figura 5.4 Análise Bayesiana de inferência de fluxo gênico. Taxas relativas de migração entre as 12 localidades amostradas no Sul da Mata Atlântica baseadas na genotipagem de oito locos microssatélites de 359 machos para *E. annectans* são apresentadas sobre a distribuição geográfica das localidades. As cores vermelho (grupo “interior”) e verde (grupo “litoral”) demarcam os grupos identificados na análise de estrutura genética pelo software *Structure*. Setas brancas indicam a direção e intensidade do fluxo gênico.

Indicativos de isolamento por distância (IBD) foram observados na correlação entre o G'_{ST} e distância geográfica, tanto nas comparações entre todas as amostras, quanto na comparação dentre as amostras do grupo “litoral” e grupo “interior” (Figura 5.5). Embora as correlações tenham se mostrado não significativas, em qualquer um dos casos nota-se uma tendência no aumento da diferenciação genética em relação ao aumento da distância geográfica.

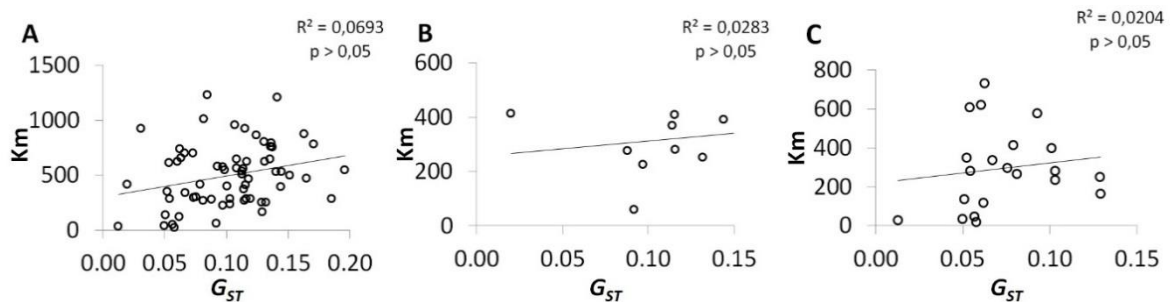


Figura 5.5 Correlação entre distância genética e geográfica entre 12 localidades amostradas para *E. annectans* do sul da Mata Atlântica baseada em oito locos microssatélites. A) todas as amostras; B) grupo “interior”; C) grupo “litoral”.

No caso específico do grupo “interior” é possível observar a presença de um ponto com alta distância geográfica e baixa distância genética influenciando fortemente a ausência significativa de IBD, o qual diz respeito à relação dos pontos KB e FZ com cerca de 400 Km de distância, porém baixos valores para os índices de estruturação genética ($G_{ST} = 0,002$; $Dest = 0,007$). Para o grupo “litoral” os índices de estruturação variaram entre $G_{ST} = 0,001$ (SU – SM) e $G_{ST} = 0,054$ (IC – SC), bem como $Dest = 0,006$ (SU – SM, não significativo) e $Dest = 0,357$ (IC – SC) (Tabela 5.6).

Tabela 5.6 Índices de estruturação genética par-a-par entre machos de *E. annectans* coletados em 12 localidades a partir de oito locos microssatélites. Valores de G_{ST} na diagonal superior e valores de $Dest$ na diagonal inferior. Todos os valores foram significativos com $\alpha=0,05$ exceto valores em negrito.

	SO	BO	PN	IC	SU	SM	SC	KB	MD	DN	FZ	DR
SO	--	0,018	0,017	0,033	0,024	0,022	0,025	0,063	0,049	0,035	0,068	0,033
BO	0,127	--	0,008	0,039	0,022	0,011	0,024	0,066	0,054	0,020	0,066	0,035
PN	0,109	0,041	--	0,033	0,044	0,033	0,043	0,049	0,046	0,024	0,054	0,030
IC	0,216	0,212	0,168	--	0,012	0,010	0,054	0,016	0,042	0,034	0,014	0,048
SU	0,164	0,126	0,245	0,064	--	0,001	0,025	0,057	0,071	0,050	0,064	0,083
SM	0,157	0,065	0,191	0,052	0,006	--	0,020	0,056	0,054	0,045	0,057	0,063
SC	0,209	0,174	0,290	0,357	0,168	0,144	--	0,096	0,068	0,053	0,099	0,064
KB	0,315	0,284	0,199	0,060	0,239	0,245	0,495	--	0,062	0,053	0,002	0,065
MD	0,315	0,297	0,238	0,208	0,388	0,304	0,443	0,247	--	0,040	0,054	0,033
DN	0,256	0,121	0,135	0,192	0,304	0,286	0,389	0,235	0,225	--	0,053	0,039
FZ	0,345	0,290	0,222	0,053	0,273	0,253	0,522	0,007	0,216	0,233	--	0,063

DR 0,195 0,179 **0,007** 0,222 0,429 0,339 0,388 0,242 0,151 0,202 0,239 --

5.3.3 Modelagem de Nicho Ecológico e Paleodistribuição

Devido à discussão em torno da seleção de variáveis e para construção de modelos de nicho e distribuição (Merow et al., 2013; Porfirio et al., 2014), três estratégias diferentes foram utilizadas para construir os modelos de nicho, selecionando variáveis via PCA, segundo um critério ecológico, e utilizando as 19 variáveis disponíveis no banco *Worldclim*, como sugerido por Porfirio et al. (2014). A avaliação destes modelos pode ser feita de diversas maneiras, porém a proporção de área sob a curva (AUC) é um dos principais parâmetros disponíveis (Merow et al., 2013; Phillips et al., 2017), uma vez que essa estatística compara a sensibilidade e especificidade do modelo contra uma distribuição aleatória, em outras palavras, quanto maior a AUC melhor o desempenho do modelo em prever a distribuição da espécie em questão (Elith et al., 2011). Todos os modelos gerados para *Euglossa annectans* apresentaram-se altamente precisos com valores de área sob a curva (AUC), variando entre 0,976 (estratégia PCA) e 0,986 (estratégia “*WorldClim*”) (Tabela 5.7). Apesar dos valores de AUC e de ganho de modelo (ganho de treino G_{UT} ; ganho de treino regularizado G_{RT}), o modelo baseado em PCA obteve os melhores escores para AICc e BIC, bem como para $logL$ (Tabela 5.7).

Tabela 5.7 Estatísticas de seleção de modelos para comparação entre três diferentes estratégias de seleção de variáveis: todas as variáveis (“*WorldClim*”), com base em aspectos ecológicos (“Ecológica”), e por exclusão de covariantes (“PCA”). N – número de amostras; \bar{AUC} – média da área sob a curva; BIC – critério de informação Bayesiano; AIC – critério de informação de Akaike; $AICc$ – com correção para amostragens reduzidas; $logL$ – log de verossimilhança; \bar{G}_{RT} – média do ganho regularizado de treino de modelo; \bar{G}_{UT} – ganho de treino de modelo não regularizado.

Estratégia	N	\bar{AUC}	$s.d.$	BIC	$AICc$	AIC	$logL$	\bar{G}_{RT}	$s.d.$	\bar{G}_{UT}
PCA	71	0,976	0,005	1663,4	1644,1	1593,2	-765,61	2,3379	0,049	2,8109
Ecológica	71	0,981	0,002	2068,6	2041,3	2014,6	-983,31	2,7596	0,033	3,105
WorldClim	71	0,986	0,005	2044,6	2017,9	1988,4	-969,18	2,9724	0,052	3,3134

A comparação das diferentes estratégias de seleção de variáveis (Figura 5.6; “PCA”: A-B; “Ecológico”: E-H; “*WorldClim*”: I-L) não apresentou diferenças expressivas do ponto de vista qualitativo, ainda que as proporções de áreas preditas sejam as mesmas, a avaliação do padrão geral permanece inalterado entre as estratégias. Embora o modelo “PCA” tenha apresentado um melhor desempenho do ponto de vista estatístico, as demais estratégias foram também incluídas na análise de consensos entre os diferentes GCMs e estratégias a fim de representar a variação entre as diferentes projeções. O consenso entre as estratégias (Figura 5.6: M-P) ressalta uma expressiva redução de áreas de sobreposição entre os paleomodelos, a qual representa o cerne de áreas adequáveis a *E. annectans* nas respectivas projeções. Em

relação à distribuição “atual” de *E. annectans* (Figura 5.6-A, E, I e M), os modelos indicam áreas de adequabilidade concentradas na porção sul da Mata Atlântica sobrepondo praticamente toda a área de distribuição conhecida entre as latitudes -29 e -19 e longitude -40 e -55. Os modelos abrangem tanto áreas litorâneas quanto de interior, com estrangulamento da área de distribuição entre as porções sudoeste do estado de Minas Gerais e noroeste de São Paulo, no limite entre Mata Atlântica e a porção de Cerrado da Diagonal de Formações Abertas da América do Sul (Figura 5.6-M).

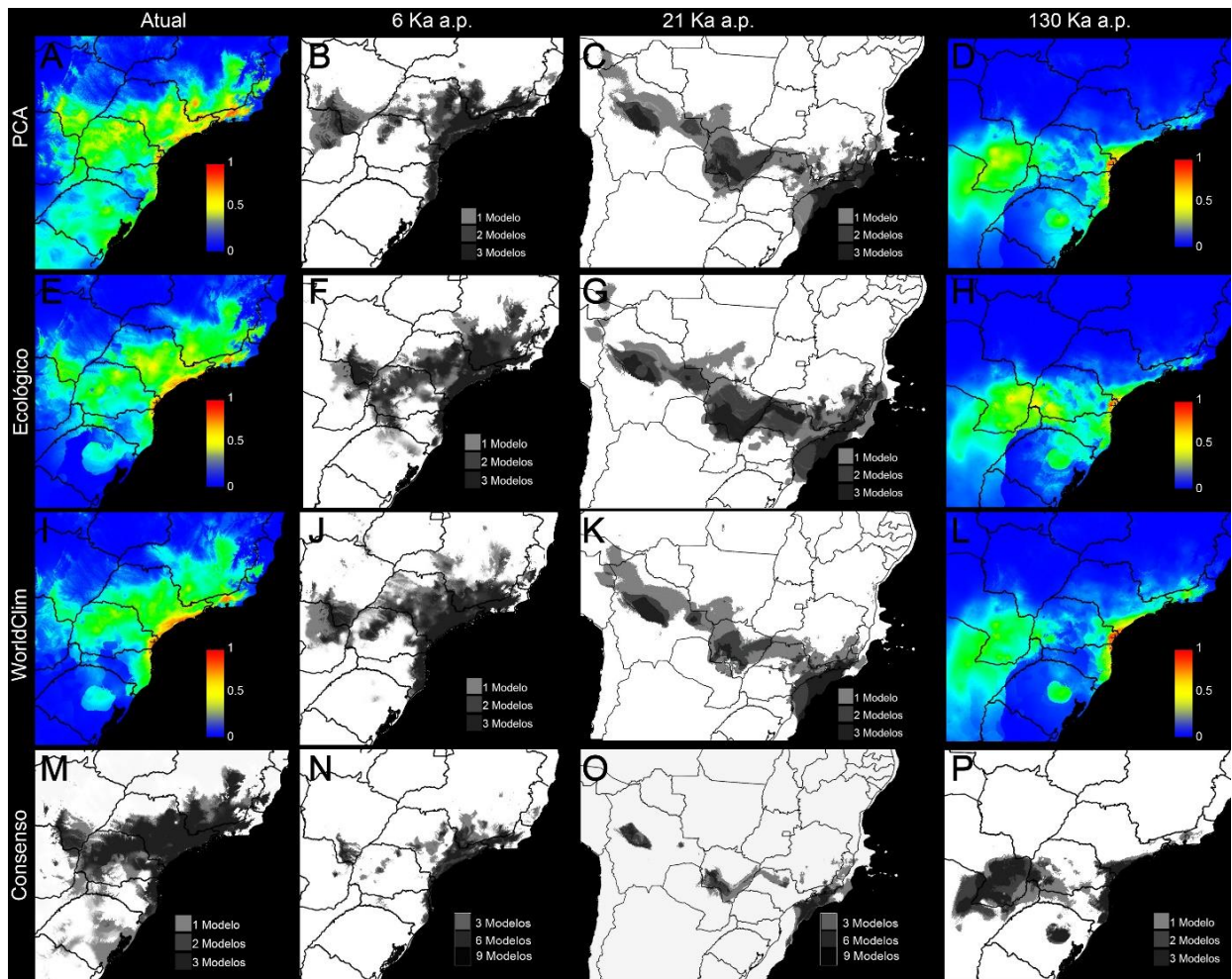


Figura 5.6 Modelagem de Nicho Ecológico e Paleodistribuição de *Euglossa annectans* baseados em três diferentes Modelos de Circulação Geral (GCM: MIROC-ESM, CCSM4, MPI-ESM-RL), 71 registros de ocorrência e diferentes abordagens de seleção de variáveis. Imagens em cores correspondem a modelos baseados em apenas um GCM com a escala de variação indicando a adequabilidade ambiental; Imagens em preto e branco correspondem ao consenso de GCMs; A última linha (Consenso) resume a sobreposição das áreas presentes em todos os modelos de cada uma das estratégias de seleções de variáveis; Colunas indicam os diferentes tempos (da esquerda para à direita: Atual, 6 Ka a. p., 21 Ka a. p. e 130 Ka a. p.); Diferentes abordagens de seleção de variáveis são apresentadas em linhas: PCA, Ecológico, WorldClim (19 variáveis).

As projeções para o médio-Holoceno (6 Ka a. p.; Figura 5.6-B, F, J e N) apresentaram um cenário bastante semelhante ao atual, com maior fragmentação entre as áreas de consenso e uma notável redução de áreas em relação ao modelo “atual” (Figura 5.6-N). Na mesma linha, os modelos projetados para o LGM (21 Ka a. p.; Figura 5.6-C, G, K e O) apresentaram

fragmentação e isolamento ainda maiores nas áreas de adequabilidade e, adicionalmente, uma substancial mudança na distribuição de *E. annectans* em direção noroeste, mais claramente demonstrado na Figura 5.6-O. Os modelos gerados para o LGM também apresentaram uma importante expansão de áreas adequáveis a *E. annectans* em direção à Plataforma Continental, criando zonas de instabilidade entre a porção litorânea e de interior da paleodistribuição desta espécie. Por fim, os modelos projetados para o último período interglacial (130 Ka a. p.; Figura 5.6-D, H, L e P) apontam para uma separação ainda maior entre a zona litorânea e interiorana de distribuição de *E. annectans*, com apenas duas grandes áreas se repetindo entre os modelos gerados pelas diferentes estratégias (Figura 5.6-P). A comparação da estabilidade climática ao longo do tempo, aponta para áreas estáveis apenas na região litorânea, com uma importante variação na porção interior da distribuição de *E. annectans*.

Para fins de conservação, projeções considerando a evolução média do acúmulo de carbono para os próximos trinta e dois anos (IPCC rpc-4.5) também foram realizadas com base nos mesmos três modelos de circulação geral disponíveis para tempos pretéritos. A comparação entre o consenso dos modelos construídos com base na variação “atual” (1950-1990) e aqueles construídos com base no painel rpc4.5 para 2050 mostra uma considerável redução de áreas adequáveis à ocupação de *E. annectans*, principalmente em áreas onde estão os maiores fragmentos atuais Figura 5.7.

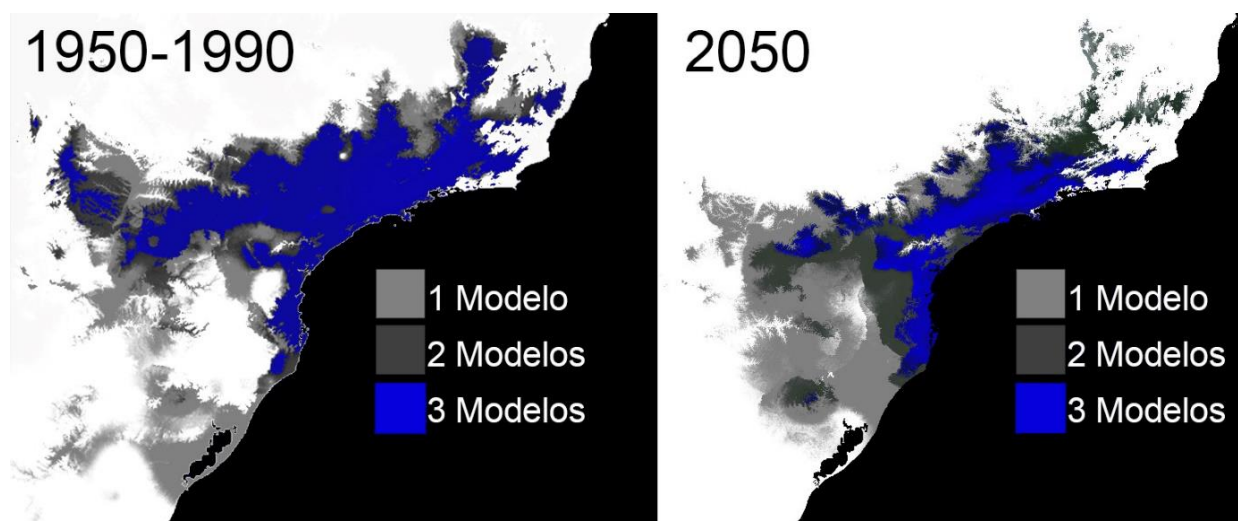


Figura 5.7 Comparação entre os consensos dos modelos de distribuição de *Euglossa annectans* baseados nas diferentes estratégias de seleção de variáveis considerando a variação climática “atual” (1950-1990) e o consenso entre os diferentes GCMs considerando a progressão moderada do acúmulo de carbono ao longo dos próximos 32 anos (IPCC rpc-4.5). Áreas de sobreposição de três (à esquerda) ou mais modelos (à direita) são destacados na coloração azul.

5.4 DISCUSSÃO

O presente estudo é o primeiro a investigar a influência das flutuações climáticas do Quaternário sobre a estrutura e diversidade genética de uma espécie de abelha com

distribuição nos limites longitudinais da MA em uma abordagem que combina modelos de distribuição espacial e diferentes marcadores moleculares. Os resultados corroboram a hipótese de que as flutuações climáticas ao longo do Pleistoceno tiveram um papel central na história demográfica e estruturação genética de *E. annectans*, revelando um padrão de estruturação pela primeira vez reportado para o sul da MA. A hipótese de que as flutuações climáticas tiveram forte influência na estruturação genética desta espécie é suportada por ao menos três resultados aqui encontrados: i) modelos paleoclimáticos indicando redução de áreas adequáveis no LGM, e consequente disjunção entre populações congruentes com a estruturação genética observada; ii) padrões demográficos cronologicamente alinhados com as flutuações climáticas; iii) ausência de barreiras geográficas significativas que não as influenciadas pelas variações climáticas.

5.4.1 Padrões de estruturação, demografia e instabilidade climática

De modo geral, as projeções paleoclimáticas em todos os cenários indicam uma progressiva redução de áreas adequáveis a *E. annectans* em direção ao LGM, levando consequentemente à disjunção das populações nas áreas costeira e de interior. Padrões similares de disjunção em modelos paleoclimáticos foram também observados em outros organismos endêmicos à MA, como anfíbios (Thomé et al., 2010), aves (Batalha-Filho & Miyaki, 2016) e opiliões (Porto et al., 2013). Este padrão é suportado pelos resultados observados a partir dos marcadores microssatélites, como indicado pelas análises de estruturação genética *Structure* e AMOVA, corroborando a existência de um grupo de interior e um segundo litorâneo. Contudo, nenhuma estruturação espacial foi detectada para os dados do gene mitocondrial *cytb*, a exemplo do observado para outros organismos desta região da MA (Batalha-Filho et al., 2012; Thomé et al., 2014). Interessantemente, a tribo Euglossini é apontada como um grupo de abelhas que possui um viés sexual de dispersão (López-Uribe et al., 2014), onde machos são capazes de voar dezenas ou centenas de quilômetros (Pokorny et al., 2015) promovendo fluxo gênico (Zimmermann et al., 2011; Suni, 2017) e fêmeas, a despeito de sua grande capacidade de voo (Janzen, 1971), são tidas como filopátricas (López-Uribe et al., 2014), uma vez que frequentemente emergem e permanecem próximas ao ninho de origem (Augusto & Garófalo, 2011), como no caso de *E. annectans* (Garófalo et al., 1998; Boff et al., 2017). Assim, seria esperado que o DNAm mantivesse um forte padrão de IBD, como o observado por Frantine-Silva et al. (2017) e Penha et al. (2015) para *E. iopoecila*. No entanto, como observado anteriormente, não há uma clara estruturação geográfica em relação ao DNAm no caso de *E. annectans*, porém a variação e a redução de áreas adequáveis ao

longo do tempo potencialmente explicam a ausência dessa estruturação espacial, bem como a diferença observada entre os genomas. Devido ao fato do tamanho efetivo populacional do DNAm_t ser cerca de três vezes menor que o DNA nuclear em populações haplodiploides (Hedrick and & Parker, 1997; Zayed, 2004), efeitos estocásticos de perda de diversidade genética afetariam mais intensamente o DNAm_t do que o DNAnu (Luikart et al., 1998), podendo levar à perda do sinal filogeográfico mesmo sem a extinção local da população (Edwards & Beerli, 2000). Esta hipótese é corroborada pelos sinais de *bottleneck* encontrados para o DNAm_t na comparação da distribuição *mismatch* e no teste de mudança do tamanho populacional R_2 .

Hipóteses alternativas para explicar a diferença entre os resultados observados para o DNAm_t e DNAnu poderiam incluir a utilização de um único gene ou o reduzido tamanho amostral nas análises do DNAm_t. No entanto, o *cytb* constitui um dos genes mais informativos do DNAm_t de abelhas (Frantine-Silva et al., 2017), apresentando um padrão evolutivo intrinsecamente relacionado a todo DNAm_t devido à ligação física entre os genes (Hurst & Jiggins, 2005; López-Uribe et al., 2014), com resultados similares em tamanhos amostrais semelhantes aos aqui analisados (Frantine-Silva et al., 2017) ou mesmo maiores (Penha et al., 2015), como os reportados por estes autores em análises independentes com *E. iopoecila*.

A ausência de estruturação genética e geográfica em casos intraespecíficos é frequentemente relatada em outros organismos com distribuição semelhante a *E. annectans* na MA (Batalha-Filho et al., 2012; Thomé et al., 2014; Batalha-Filho & Miyaki, 2016). Um exemplo interessante foi reportado para *Basileuturus leucoblepharus*, uma ave (Passeriformes) também endêmica à porção sul da MA (Batalha-Filho et al., 2012) para o qual é refutada a hipótese de um *bottleneck* no LGM, porém com influência de glaciações anteriores à 130 Ka. Os dados de *E. annectans* adicionam uma nova perspectiva dos eventos demográficos ocorridos nos táxons endêmicos à esta porção da MA, indicando que tanto a demografia quanto a estruturação genética de algumas espécies podem ter sido afetadas pela última glaciação. Em consonância com estes resultados, Batalha-Filho e Miyaki (2016) apontam a influência do LGM na divergência e demografia entre duas linhagens do pássaro *Rhopias gularis* em uma distribuição semelhante a *E. annectans*.

Interessantemente, o mesmo padrão de expansão demográfica após o LGM corroborado pelos dados de *E. annectans* é observado para o *cluster* sul de *Euglossa iopoecila* (Frantine-Silva et al., 2017). Os recorrentes sinais de expansão populacional em espécies pertencentes a diferentes táxons endêmicos à porção sul da MA indicam que ao menos para

parte da fauna houve uma forte influência das flutuações climáticas do Quaternário, em especial do último máximo glacial, sobre a estruturação genético-populacional e filogeográfica das espécies desta porção da MA. Contudo, é importante ainda ressaltar que, embora as duas últimas glaciações pareçam semelhantes em intensidade e duração (Augustin et al., 2004; Joos & Prentice, 2004) eventos anteriores ao LGM podem ter influenciado a estruturação genética observada (Fitzpatrick et al., 2009), como sugerido pelos modelos do último período interglacial (LIG), apresentando dois grandes grupos alinhados com a estruturação genética atual de *E. annectans*.

Os resultados dos modelos paleoclimáticos do LGM apontam ainda um aspecto adicional em relação à biogeografia dos Euglossini na MA, suportando a existência de conexão entre a MA e a Floresta Tropical Andina. Embora os dados moleculares de *E. annectans* não permitam testar diretamente esta hipótese, os modelos de nicho corroboram a região central do cerrado como uma possível rota de fluxo gênico entre a MA e sul da Amazônia e oeste Andino, como já reportado para aves (Trujillo-Arias et al., 2017), dentre outros organismos (Ledo & Colli, 2017). Essa ideia é também suportada pela presença de áreas estáveis na porção central do cerrado no LGM para outras espécies de Euglossini, como *Eulaema cingulata* (López-Uribe et al., 2014). Futuros estudos mais detalhados deverão responder especificamente sobre a importância desta conexão para euglossíneos de mata atlântica.

A instabilidade climática encontrada para os modelos de *E. annectans* em consonância com a variação demográfica observada para o DNAm_t apontam a existência de um padrão apenas parcialmente alinhado ao modelo de refúgios de Carnaval e Moritz (2008) “*strictu sensu*” (daqui em diante “modelo CM”). Segundo o modelo CM, populações de espécies endêmicas deveriam ter permanecido em refúgios ao norte da MA ao longo do LGM, onde maiores níveis de diversidade genética seriam encontrados, aliado a um sinal de expansão populacional pós-LGM nas populações ao sul da MA. Em consonância ao modelo CM pode-se citar a instabilidade das áreas de adequabilidade na porção sul da distribuição de *E. annectans* e os sinais de expansão populacional pós-LGM. Em contrapartida, *E. annectans* não apresentou um padrão de estruturação latitudinal esperado sob o modelo CM, e sim um padrão longitudinal, de acordo com as previsões dos modelos de nicho específicos. Além disso, embora haja uma nítida migração de áreas adequáveis em direção ao norte, poucas áreas foram identificadas como temporalmente estáveis para caracterizar a existência de refúgios *sensu* modelo CM (Carnaval & Moritz, 2008; Carnaval et al., 2009). Com exceção de pequenas áreas na porção interior e uma possível área na porção litorânea, o cerne da

distribuição desta espécie migrou substancialmente para norte, para o interior e para a plataforma continental com praticamente completa ausência de sobreposição em relação à distribuição atual. Interessantemente, essa expansão em direção à plataforma continental é esperada em uma segunda hipótese aparentemente antagônica ao modelo CM. Baseada em modelagem de nichos e dados genéticos de pequenos mamíferos da porção centro-norte da MA a Hipótese da Mata Atlântica (Leite et al., 2016a) afirma que a exposição da plataforma continental devido à regressão marinha disponibilizou áreas para as quais a MA expandiu seus domínios, e portanto, sinais de expansão populacional deveriam ser identificados no LGM ao invés de *bottlenecks*. Embora a extensão das conclusões desta última hipótese seja discutível (Leite et al., 2016b; Raposo do Amaral et al., 2016), mesmo estudos anteriores apontam a expansão de áreas adequáveis a diferentes espécies em direção à plataforma continental na porção sul e central da MA (Thomé et al., 2010; Batalha-Filho & Miyaki, 2016), corroborando a utilização destas áreas ao longo do LGM, a exemplo de *E. annectans*. Ainda assim, *E. annectans* e outros organismos da MA apresentam sinais de expansão populacional pós-LGM (Batalha-Filho & Miyaki, 2016). Os resultados aqui encontrados para *E. annectans* indicam uma situação mais complexa, mesclando pontos destas duas hipóteses aparentemente antagônicas, uma vez que exibe uma possível expansão em direção à plataforma continental em resposta a uma mudança de adequabilidade no LGM (Leite et al., 2016a), porém com sinais de variação demográfica que se encaixam às expectativas do modelo CM (Carnaval & Moritz, 2008), como também descrito em Batalha-Filho e Miyaki (2016).

É importante ainda considerar a possibilidade da existência de barreiras, uma vez que vários estudos apontam vales em bacias hidrográficas ou eventos neotectônicos como possíveis barreiras para serpentes (Grazziotin et al., 2006), aves (Cabanne et al., 2007), anfíbios (Carnaval et al., 2009; Thomé et al., 2014) planárias (Álvarez-Presas et al., 2014) ou mesmo abelhas (Batalha-Filho et al., 2010). Contudo, este não parece ser o caso de *E. annectans* devido à alta congruência dos padrões espaciais com as predições dos modelos de nicho para os marcadores microssatélites, ou mesmo a ausência de quebra filogeográfica para o DNAm, mais plausivelmente explicado pela perda de sinal filogeográfico discutido anteriormente.

Deste modo, os dados apresentados neste estudo indicam um cenário climático bastante dinâmico ao longo dos últimos 130 Ka, no qual diferentes organismos devem ter respondido de maneiras idiossincráticas às mudanças ambientais. De forma conjunta, os dados observados para *E. annectans* e o *cluster* sul de *E. iopoecila* (Frantine-Silva et al., 2017) indicam uma baixa adequabilidade ambiental para as abelhas das orquídeas nas latitudes onde

estas espécies são observadas atualmente durante o LGM. Contudo, é importante explorar novos modelos com distribuições semelhantes à de *E. annectans* buscando avaliar se este padrão irá se confirmar como uma regra para os Euglossini da porção sul da MA.

5.4.2 Diversidade genética e conservação

No que tange a diversidade genética é possível observar mais uma vez dois padrões distintos a depender da origem do marcador analisado. Por um lado, há uma clara homogeneidade entre os índices descritivos de diversidade genética para os marcadores microsatélites (N_A , N_E e H_e) indicando níveis de diversidade relativamente altos mesmo em populações com tamanho amostral bastante reduzido (BO, IC e DR) em relação a outros euglossíneos presentes na MA (Freiria et al., 2012; Boff et al., 2014; Penha et al., 2015). Por outro lado, os dados para o gene *cytb* demonstram uma grande divergência entre índices equivalentes ao longo das localidades amostradas como H_d e π , com haplótipos fixados em algumas populações e índices de H_d próximos à 1 em outras, ainda que especialmente o H_d possa apresentar algum viés em relação ao tamanho amostral (Excoffier et al., 1992). Parte desta diferença pode ser explicada pelo desvio sexual de dispersão nas abelhas das orquídeas anteriormente discutido. Este é um aspecto conhecido e recorrente em Euglossini, como observado para *E. iopocila* (Penha et al., 2015) e *Eulaema spp.* (López-Uribe et al., 2014). Especificamente para *E. annectans* esta diferença é claramente observada pela estimativa da análise Bayesiana de fluxo gênico, a qual demonstra um intenso fluxo gênico espacialmente estruturado, também em consonância com a separação histórica dos *clusters*. Adicionalmente, a diferença na diversidade genética entre os genomas também se deve em parte aos respectivos tamanhos efetivos populacionais, também anteriormente discutido. Interessantemente, a homogeneidade dos parâmetros genéticos em relação ao respectivo marcador também depõe contra o esperado pelo modelo CM “*strictu sensu*”, ainda que do ponto de vista nuclear, a dispersão parece ser o principal fator homogeneizador.

E. annectans apresenta ainda uma outra peculiaridade relevante à sua conservação, a qual diz respeito à expressiva frequência de machos diploides (f_{MP}) nessa espécie. Na maior parte das abelhas o sexo é determinado por um loco único dependente de complementariedade (loco *CSD*) (Hasselmann et al., 2008), onde genótipos heterozigotos expressam fêmeas e hemizigotos ou homozigotos dão origem a machos (Harpur et al., 2013). Estudos salientam que em casos de baixa diversidade genética haveria um aumento na produção de machos diploides devido à baixa diversidade do loco *CSD* (Zayed & Packer, 2005; Zayed, 2009; Harpur et al., 2013; Lozier & Zayed, 2017). De modo geral, Euglossini apresentam médias de

f_{MP} entre 0,01 e 0,02 (Souza et al., 2010; Giangarelli et al., 2015b), enquanto *E. annectans* apresentou uma média geral de 0,098 e até 0,182 em algumas localidades (PN e IC), os quais corroboram os valores observados por Giangarelli et al. (2015), além de dados independentemente observados pelo Professor Robert Paxton (comunicação pessoal). Zayed e Packer (2005) apontam a possibilidade de que o aumento da produção de machos diploides poderia levar a população a um quadro de retroalimentação de aumento de probabilidade de extinção denominado “vortex do macho diploide”, uma vez que a baixa diversidade do loco *CSD* aumentaria a f_{MD} que por sua vez diminuiria o número de machos viáveis, consequentemente o Ne da população e a diversidade do loco *CSD*. Contudo, a despeito da maior f_{MD} observada para *E. annectans*, os valores de diversidade genética encontrados demonstram que as populações mantêm níveis de diversidade genética compatíveis com espécies com baixa f_{MD} (Boff et al., 2014; Penha et al., 2015; Soro et al., 2017; Suni, 2017). Este fato indica algum tipo de mecanismo para contornar este “problema” possivelmente com a capacidade de seleção de parceiros ou mesmo relacionado a uma possível viabilidade dos machos diploides, embora estudos futuros devam explorar especificamente esta questão. Adicionalmente, a interpretação da proporção de machos diploides deve ser feita com cautela uma vez que valores que representativamente parecem muito distintos quando analisados em formas de frequência podem significar números absolutos não tão distintos. A comparação entre localidades com tamanhos amostrais pequenos (IC: $N=11$; DR: $N=14$) apresenta valores em f_{MP} que poderiam ser interpretados como bastante diferentes, respectivamente 0,182 e 0,077, porém representam a diferença de um único macho diploide detectado. Vale ainda ressaltar que a diferença de f_{MP} variou entre 0,029 e 0,182 o que parece uma mudança significativa, mas a variação absoluta foi de 1 a 6 machos diploides detectados. Ainda assim, *E. annectans* apresentou 39 machos diploides em 398 machos genotipados, o que é expressivamente maior do que a taxa estimada por Souza et al. (2010) menor que 0,001 em 1034 indivíduos e deve ser levado em consideração para ações de conservação para essa espécie.

Ainda no que tange à conservação, projeções futuras baseadas nos modelos atuais de distribuição para *E. annectans* e um cenário conservador de mudança climática até 2050, revelaram uma redução substancial de áreas climaticamente adequáveis à essa espécie na periferia da distribuição. Áreas como o Parque Nacional do Iguazu e a região de Belo Horizonte saem da predição de consenso de áreas adequáveis, restando praticamente a região da Serra do Mar como consenso nos três modelos gerados. Embora, como exposto até aqui, *E. annectans* tenha passado por uma história de redução e expansão populacional recente, vale

ressaltar que o mesmo nunca ocorreu em um ambiente tão fragmentado e com matrizes tão desafiadoras quanto ao atual. Ademais, dados genéticos e de modelagem indicam que a fragmentação pretérita sempre deixou redutos interioranos de diversidade genética para esta espécie, o que foi aproveitado pela grande capacidade dispersiva dos euglossíneos para manter algum nível de fluxo gênico, como também demonstrado aqui pela estimativa de fluxo gênico Bayesiana. Considerando ainda a tendência de *E. annectans* em produzir machos diploides em frequências expressivamente maiores do que as habitualmente encontradas em outros Euglossini, estes dados em conjunto ascendem um alerta do ponto de vista conservacionista para esta espécie. Neste sentido, manter pontos de conexão entre áreas verdes no interior e o corredor de biodiversidade da Serra do Mar, parece ser fundamental para a conservação de *E. annectans*.

5.5 CONCLUSÕES

Os resultados aqui apresentados demonstram um padrão diferenciado de estruturação genética até então inédito em relação para organismos de MA. As flutuações climáticas do Pleistoceno tardio tiveram importância fundamental na definição desta estrutura, causando um quadro dinâmico de adequabilidade ambiental. Os resultados de *E. annectans* também apontam para a união de ideias de duas hipóteses aparentemente contrárias em relação à dinâmica de refúgios ao longo do Quaternário, demonstrando que as áreas de ocupação podem ter alterado significativamente considerando a regressão marinha, mas as populações podem ter sofrido gargalos genéticos e redução de áreas adequáveis, uma vez que o avanço de vegetação para novas áreas na plataforma continental deve ter sido contrabalanceado pelo avanço de outras áreas menos adequáveis nas maiores latitudes da MA. Por fim, as variações climáticas aparentemente continuarão tendo um importante papel para distribuição de *E. annectans* e medidas que priorizem o aumento e conexão de áreas verdes no interior, e deste para o litoral, são de fundamental importância para manter populações viáveis dessa espécie no futuro.

REFERÊNCIAS

- Ab'Sáber, A. (2003) *Os domínios de natureza no Brasil: potencialidades paisagísticas*. Ateliê Editorial, São Paulo.
- Álvarez-Presas, M., Sánchez-Gracia, A., Carbayo, F., Rozas, J., & Riutort, M. (2014) Insights into the origin and distribution of biodiversity in the Brazilian Atlantic forest hot spot: a statistical phylogeographic study using a low-dispersal organism. *Heredity*, **112**, 656–

665.

- Arias, M.C., Silvestre, D., Oliveira Francisco, F. de, Weinlich, R., & Sheppard, W.S. (2008) An oligonucleotide primer set for PCR amplification of the complete honey bee mitochondrial genome. *Apidologie*, **39**, 475–480.
- Augustin, L., Barbante, C., Barnes, P.R.F., et al. (2004) Eight glacial cycles from an Antarctic ice core. *Nature*, **429**, 623–628.
- Augusto, S.C. & Garófalo, C.A. (2011) Task allocation and interactions among females in *Euglossa carolina* nests (Hymenoptera, Apidae, Euglossini)*. *Apidologie*, **42**, 162–173.
- Batalha-Filho, H., Cabanne, G.S., & Miyaki, C.Y. (2012) Phylogeography of an Atlantic forest passerine reveals demographic stability through the last glacial maximum. *Molecular Phylogenetics and Evolution*, **65**, 892–902.
- Batalha-Filho, H. & Miyaki, C.Y. (2016) Late Pleistocene divergence and postglacial expansion in the Brazilian Atlantic Forest: Multilocus phylogeography of *Rhopias gularis* (Aves: Passeriformes). *Journal of Zoological Systematics and Evolutionary Research*, **54**, 137–147.
- Batalha-Filho, H., Waldschmidt, A.M., Campos, L.A.O., Tavares, M.G., & Fernandes-Salomão, T.M. (2010) Phylogeography and historical demography of the neotropical stingless bee *Melipona quadrifasciata* (Hymenoptera, Apidae): incongruence between morphology and mitochondrial DNA. *Apidologie*, **41**, 534–547.
- Boff, S., Saito, C.A., & Santos, I.A. (2017) Multiple aggressions among nestmates lead to weak dominance hampering primitively eusocial behaviour in an orchid bee. *Sociobiology*, **64**, 202–211.
- Boff, S., Soro, A., Paxton, R.J., & Alves-dos-Santos, I. (2014) Island isolation reduces genetic diversity and connectivity but does not significantly elevate diploid male production in a neotropical orchid bee. *Conservation Genetics*, **15**, 1123–1135.
- Cabanne, G.S., Santos, F.R., & Miyaki, C.Y. (2007) Phylogeography of *Xiphorhynchus fuscus* (Passeriformes, Dendrocolaptidae): Vicariance and recent demographic expansion in southern Atlantic forest. *Biological Journal of the Linnean Society*, **91**, 73–84.
- Cardoso da Silva, J.M., Cardoso de Sousa, M., & Castelletti, C.H.M. (2004) Areas of endemism for passerine birds in the Atlantic forest, South America. *Global Ecology and Biogeography*, **13**, 85–92.
- Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., & Moritz, C. (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.

- Carnaval, A.C. & Moritz, C. (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.
- Carnaval, A.C., Waltari, E., Rodrigues, M.T., Rosauer, D., VanDerWal, J., Damasceno, R., Prates, I., Strangas, M., Spanos, Z., Rivera, D., Pie, M.R., Firkowski, C.R., Bornschein, M.R., Ribeiro, L.F., & Moritz, C. (2014) Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20141461–20141461.
- Corander, J., Marttinen, P., Sirén, J., & Tang, J. (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC bioinformatics*, **9**, 539.
- Crozier, Y.C., Koulianos, S., & Crozier, R.H. (1991) An improved test for Africanized honeybee mitochondrial DNA. *Experientia*, **47**, 968–969.
- Dick, C.W., Roubik, D.W., Gruber, K.F., & Bermingham, E. (2004) Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA phylogeography. *Molecular Ecology*, **13**, 3775–3785.
- Drummond, A.J. & Bouckaert, R.R. (2015) Bayesian evolutionary analysis with BEAST. *Bayesian Evolutionary Analysis with BEAST*, 1–249.
- Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Edgar, R.C. (2004) MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.
- Edwards, S. & Beerli, P. (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.
- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.
- Evanno, G., Regnaut, S., & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier, L., Laval, G., & Schneider, S. (2005) Arlequin: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**, 47–50.
- Excoffier, L., Smouse, P.E., & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Faria Jr, L.R.R. & Melo, G.A.R. (2007) Species of Euglossa (Glossura) in the Brazilian

- Atlantic forest, with taxonomic notes on *Euglossa stellfeldi* Moure (Hymenoptera, Apidae, Euglossina). *Revista Brasileira de Entomologia*, **51**, 275–284.
- Fitzpatrick, S.W., Brasileiro, C.A., Haddad, C.F.B., & Zamudio, K.R. (2009) Geographical variation in genetic structure of an Atlantic Coastal Forest frog reveals regional differences in habitat stability. *Molecular Ecology*, **18**, 2877–2896.
- Frantine-Silva, W., Giangarelli, D.C., Penha, R.E.S., Suzuki, K.M., Dec, E., Gaglianone, M.C., Alves-dos-Santos, I., & Sofia, S.H. (2017) Phylogeography and historical demography of the orchid bee *Euglossa iopoecila*: signs of vicariant events associated to Quaternary climatic changes. *Conservation Genetics*, **18**, 539–552.
- Freiria, G.A., Ruim, J.B., De Souza, R.F., & Sofia, S.H. (2012) Population structure and genetic diversity of the orchid bee *Eufriesea violacea* (Hymenoptera, Apidae, Euglossini) from Atlantic Forest remnants in southern and southeastern Brazil. *Apidologie*, **43**, 392–402.
- Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Galindo-Leal, C. & Câmara, I. de G. (2003) *The Atlantic Forest of South America: Biodiversity Status, Threats, and Outlook*. Inland Press, Washington, DC.
- Garófalo, C.A., Camillo, E., Augusto, S.C., Jesus, B.M.V. de, & Serrano, J.C. (1998) Nest structure and communal nesting in *Euglossa* (Glossura) *annectans* Dressler (Hymenoptera, Apidae, Euglossini). *Revista Brasileira de Zoologia*, **15**, 589–596.
- Garraffoni, A.R.S., Moura, F.R., & Lourenço, A.P. (2017) Areas of endemism in the Atlantic Forest: quantitative biogeography insights from orchid bees (Apidae: Euglossini). *Apidologie*, **48**, 513–522.
- Gent, P.R., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z.L., & Zhang, M. (2011) The community climate system model version 4. *Journal of Climate*, **24**, 4973–4991.
- Giangarelli, D.C., Aguiar, W.M. de, & Sofia, S.H. (2015a) Orchid bee (Hymenoptera: Apidae: Euglossini) assemblages from three different threatened phytophysiognomies of the subtropical Brazilian Atlantic Forest. *Apidologie*, **46**, 71–83.
- Giangarelli, D.C., Freiria, G.A., Ferreira, D.G., Aguiar, W.M., Penha, R.E.S., Alves, A.N., Gaglianone, M.C., & Sofia, S.H. (2015b) Orchid bees: a new assessment on the rarity of diploid males in populations of this group of Neotropical pollinators. *Apidologie*, **46**, 606–617.

- Giorgetta, M.A., Jungclaus, J., Reick, C.H., et al. (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, **5**, 572–597.
- Grazziotin, F.G., Monzel, M., Echeverrigaray, S., & Bonatto, S.L. (2006) Phylogeography of the *Bothrops jararaca* complex (Serpentes: Viperidae): Past fragmentation and island colonization in the Brazilian Atlantic Forest. *Molecular Ecology*, **15**, 3969–3982.
- Harpur, B.A., Sobhani, M., & Zayed, A. (2013) A review of the consequences of complementary sex determination and diploid male production on mating failures in the Hymenoptera. *Entomologia Experimentalis et Applicata*, **146**, 156–164.
- Hasselmann, M., Gempe, T., Schiøtt, M., Nunes-Silva, C.G., Otte, M., & Beye, M. (2008) Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. *Nature*, **454**, 519–522.
- Hedrick and, P.W. & Parker, J.D. (1997) Evolutionary genetics and genetic variation of haplodiploids and x-linked genes. *Annual Review of Ecology and Systematics*, **28**, 55–83.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Ho, S.Y.W. & Shapiro, B. (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular ecology resources*, **11**, 423–434.
- Hurst, G.D. & Jiggins, F.M. (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1525–1534.
- Janzen, D.H. (1971) Euglossine bees as long-distance pollinators of tropical plants. *Science*, **171**, 203–205.
- Joos, F. & Prentice, I.C. (2004) A Paleo-Perspective on Changes in Atmospheric CO₂ and Climate. *Global Carbon Cycle: Integrating Humans, Climate, and the Natural World* pp. 165–186.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., & Mayrose, I. (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, **15**, 1179–1191.
- Ledo, R.M.D. & Colli, G.R. (2017) The historical connections between the Amazon and the Atlantic Forest revisited. *Journal of Biogeography*, **44**, 2551–2563.
- Leigh, J.W. & Bryant, D. (2015) POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, **6**, 1110–1116.

- Leite, Y.L.R., Costa, L.P., Loss, A.C., Rocha, R.G., Batalha-Filho, H., Bastos, A.C., Quaresma, V.S., Fagundes, V., Paresque, R., Passamani, M., & Pardini, R. (2016a) Neotropical forest expansion during the last glacial period challenges refuge hypothesis. *Proceedings of the National Academy of Sciences*, **113**, 1008–1013.
- Leite, Y.L.R., Costa, L.P., Loss, A.C., Rocha, R.G., Batalha-Filho, H., Bastos, A.C., Quaresma, V.S., Fagundes, V., Paresque, R., Passamani, M., & Pardini, R. (2016b) Reply to Raposo do Amaral et al.: The “Atlantis Forest hypothesis” adds a new dimension to Atlantic Forest biogeography. *Proceedings of the National Academy of Sciences*, **113**, E2099–E2100.
- Librado, P. & Rozas, J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- López-Uribe, M.M., Zamudio, K.R., Cardoso, C.F., & Danforth, B.N. (2014) Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees. *Molecular Ecology*, **23**, 1874–1890.
- Lozier, J.D. & Zayed, A. (2017) Bee conservation in the age of genomics. *Conservation Genetics*, **18**, 713–729.
- Luikart, G., Sherwin, W.B., Steele, B.M., & Allendorf, F.W. (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**, 963–974.
- Merow, C., Smith, M.J., & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species’ distributions: What it does, and why inputs and settings matter. *Ecography*, **36**, 1058–1069.
- Moure, J., Melo, G.A.R., & Faria Jr, L.R.R. (2012) Available at: <http://www.moure.cria.org.br/catalogue>.
- Nemésio, A. (2009) Orchid bees (Hymenoptera: Apidae) of the Brazilian Atlantic forest. *Zootaxa*, **2041**, 1–242.
- Nemésio, A., Silva, D.P., Nabout, J.C., & Varela, S. (2016) Effects of climate change and habitat loss on a forest-dependent bee species in a tropical fragmented landscape. *Insect Conservation and Diversity*, **9**, 149–160.
- Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., & Shipley, P. (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
- Ossa, P.G., Pérez, F., & Armesto, J.J. (2013) Phylogeography of two closely related species of *Nolana* from the coastal atacama desert of Chile: Post-glacial population expansions in

- response to climate fluctuations. *Journal of Biogeography*, **40**, 2191–2203.
- Parra-H, A. & Nates-Parra, G. (2012) The Ecological Basis for Biogeographic Classification: An Example in Orchid Bees (Apidae: Euglossini). *Neotropical Entomology*, **41**, 442–449.
- Paxton, R.J., Zobel, M.U., Steiner, J., & Zillikens, A. (2009) Microsatellite loci for *Euglossa annectans* (Hymenoptera: Apidae) and their variability in other orchid bees. *Molecular Ecology Resources*, **9**, 1221–1223.
- Peakall, R. & Smouse, P.E. (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.
- Penha, R.E.S., Gaglianone, M.C., Almeida, F.S., Boff, S. V., & Sofia, S.H. (2015) Mitochondrial DNA of *Euglossa iopoecila* (Apidae, Euglossini) reveals two distinct lineages for this orchid bee species endemic to the Atlantic Forest. *Apidologie*, **46**, 346–358.
- Phillips, S., Anderson, R.P., & Schapire, R. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E., & Blair, M.E. (2017) Opening the black box: an open-source release of Maxent. *Ecography*, **40**, 887–893.
- Phillips, S.J., Dudík, M., & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *21st International Conference on Machine Learning, Banff, Canada*, 655–662.
- Pokorny, T., Loose, D., Dyker, G., Quezada-Euán, J.J.G., & Eltz, T. (2015) Dispersal ability of male orchid bees and direct evidence for long-range flights. *Apidologie*, **46**, 224–237.
- Porfirio, L.L., Harris, R.M.B., Lefroy, E.C., Hugh, S., Gould, S.F., Lee, G., Bindoff, N.L., & Mackey, B. (2014) Improving the Use of Species Distribution Models in Conservation Planning and Management under Climate Change. *PLoS ONE*, **9**, e113749.
- Porto, T.J., Carnaval, A.C., & da Rocha, P.L.B. (2013) Evaluating forest refugial models using species distribution models, model filling and inclusion: A case study with 14 Brazilian species. *Diversity and Distributions*, **19**, 330–340.
- Posada, D. (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.
- Pritchard, J.K., Stephens, M., & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Quijano-Abril, M.A., Callejas-Posada, R., & Miranda-Esquivel, D.R. (2006) Areas of endemism and distribution patterns for Neotropical Piper species (Piperaceae). *Journal*

- of Biogeography*, **33**, 1266–1278.
- Ramírez-Soriano, A., Ramos-Onsins, S.E., Rozas, J., Calafell, F., & Navarro, A. (2008) Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics*, **179**, 555–567.
- Ramírez, S.R., Roubik, D.W., Skov, C.E., & Pierce, N.E. (2010) Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). *Biological Journal of the Linnean Society*, **100**, 552–572.
- Ramos-Onsins, S.E. & Rozas, J. (2006) Erratum: Statistical properties of new neutrality tests against population growth (Molecular Biology and Evolution (2002) 19 (2092-2100)). *Molecular Biology and Evolution*, **23**, 1642.
- Raposo do Amaral, F., Edwards, S. V., Pie, M.R., Jennings, W.B., Svensson-Coelho, M., d’Horta, F.M., Schmitt, C.J., & Maldonado-Coelho, M. (2016) The “Atlantis Forest hypothesis” does not explain Atlantic Forest phylogeography. *Proceedings of the National Academy of Sciences*, **113**, E2097–E2098.
- Ribeiro, M.C., Metzger, J.P., Martensen, A.C., Ponzoni, F.J., & Hirota, M.M. (2009) The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation*, **142**, 1141–1153.
- Silveira, G.C., Freitas, R.F., Tosta, T.H.A., Rabelo, L.S., Gaglianone, M.C., & Augusto, S.C. (2015) The orchid bee fauna in the Brazilian savanna: do forest formations contribute to higher species diversity? *Apidologie*, **46**, 197–208.
- Soro, A., Quezada-Euan, J.J.G., Theodorou, P., Moritz, R.F.A., & Paxton, R.J. (2017) The population genetics of two orchid bees suggests high dispersal, low diploid male production and only an effect of island isolation in lowering genetic diversity. *Conservation Genetics*, **18**, 607–619.
- Souza, R.O., Del Lama, M.A., Cervini, M., Mortari, N., Eltz, T., Zimmermann, Y., Bach, C., Brosi, B.J., Suni, S., Quezada-Euán, J.J.G., & Paxton, R.J. (2010) Conservation genetics of neotropical pollinators revisited: Microsatellite analysis suggests that diploid males are rare in orchid bees. *Evolution*, **64**, 3318–3326.
- Sundar, D.-R. (2014) binom: Binomial Confidence Intervals For Several Parameterizations. R package version 1.1-1.
- Suni, S.S. (2017) Dispersal of the orchid bee *Euglossa imperialis* over degraded habitat and intact forest. *Conservation Genetics*, **18**, 621–630.
- Tabarelli, M., Aguiar, A.V., Ribeiro, M.C., Metzger, J.P., & Peres, C.A. (2010) Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified

- landscapes. *Biological Conservation*, **143**, 2328–2340.
- Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., & Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.
- Thomé, M.T.C., Zamudio, K.R., Giovanelli, J.G.R., Haddad, C.F.B., Baldissera, F.A., & Alexandrino, J. (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution*, **55**, 1018–1031.
- Thomé, M.T.C., Zamudio, K.R., Haddad, C.F.B., & Alexandrino, J. (2014) Barriers, rather than refugia, underlie the origin of diversity in toads endemic to the Brazilian Atlantic Forest. *Molecular Ecology*, **23**, 6152–6164.
- Trujillo-Arias, N., Dantas, G.P.M., Arbeláez-Cortés, E., Naoki, K., Gómez, M.I., Santos, F.R., Miyaki, C.Y., Aleixo, A., Tubaro, P.L., & Cabanne, G.S. (2017) The niche and phylogeography of a passerine reveal the history of biological diversification between the Andean and the Atlantic forests. *Molecular Phylogenetics and Evolution*, **112**, 107–121.
- Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., & Kawamiya, M. (2011) MIROC-ESM: model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development Discussions*, **4**, 1063–1128.
- Wilson, G.A. & Rannala, B. (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177–1191.
- Zayed, A. (2004) Effective population size in Hymenoptera with complementary sex determination. *Heredity*, **93**, 627–630.
- Zayed, A. (2009) Bee genetics and conservation. *Apidologie*, **40**, 237–262.
- Zayed, A. & Packer, L. (2005) Complementary sex determination substantially increases extinction proneness of haplodiploid populations. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 10742–6.
- Zimmermann, Y., Schorkopf, D.L.P., Moritz, R.F.A., Pemberton, R.W., Quezada-Euan, J.J.G., & Eltz, T. (2011) Population genetic structure of orchid bees (Euglossini) in anthropogenically altered landscapes. *Conservation Genetics*, **12**, 1183–1194.

CAPÍTULO 3

FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DE *EUGLOSSA IOPOECILA* DRESSLER:
SINAIS DE VICARIÂNCIA ASSOCIADOS ÀS MUDANÇAS CLIMÁTICAS DO
QUATERNÁRIO

6 CAPÍTULO 3: FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DE *EUGLOSSA IOPOECILA* DRESSLER: SINAIS DE VICARIÂNCIA ASSOCIADOS ÀS MUDANÇAS CLIMÁTICAS DO QUATERNÁRIO⁸

Resumo

O objetivo deste estudo foi investigar a influência das flutuações climáticas do Pleistoceno sobre a estrutura filogeográfica e a história demográfica de *Euglossa iopoecila* Dressler (1982), uma espécie endêmica da Mata Atlântica, fortemente associada a regiões litorâneas, que apresenta padrões distintos de coloração tegumentar ao longo de sua distribuição geográfica. Amostras de populações de *E. iopoecila* foram coletadas em 14 remanescentes de Mata Atlântica entre os estados de Pernambuco e Santa Catarina, as quais tiveram 2234 pb de quatro genes mitocondriais sequenciados: *16S* (357 pb), *cytb* (651 pb) e *COI* (1206 pb). As redes de haplótipos, SAMOVA e BAPS indicaram três linhagens (clado sul, central e norte), com duas importantes quebras filogeográficas. Áreas preditas de estabilidade climática ao longo do Quaternário (modelo de extinção e estabilidade) da MA foram coincidentes com a região de maior diversidade genética. As análises demográficas sugerem que apenas o clado sul sofreu expansão populacional recente, a qual provavelmente começou após o Último Máximo Glacial (LGM). Os dados sugerem que a diferenciação observada nas três linhagens mitocondriais de *E. iopoecila* é o resultado de eventos vicariantes e dispersões passadas, resultantes de múltiplos eventos de extinção/recolonização ao longo das flutuações climáticas do Pleistoceno. Em termos de conservação, vale ressaltar: (1) a região do clado central como a área de maior diversidade genética do DNA mitocondrial nas populações de *E. iopoecila*; (2) o comportamento filopátrico das fêmeas que tende a restringir o fluxo gênico de DNAm em particular, com implicações diretas para a conservação da diversidade genética em populações de Euglossini.

6.1 INTRODUÇÃO

A tribo Euglossini compreende cerca de 250 espécies de abelhas (Moure et al. 2012), muitas das quais são intimamente relacionadas às florestas úmidas da região Neotropical (Roubik e Hanson, 2004), atuando como importantes polinizadores de dezenas de famílias de

⁸ Este trabalho é uma versão adaptada de Frantine-Silva et al. (2017) *Conserv. Genet.* 18: 539 (<https://doi.org/10.1007/s10592-016-0905-7>) autorizada pela Editora Springer Nature (integra em anexo). Todos os direitos autorais estão sob concessão da editora (Springer Nature ©Copyrights 2016). Todas as referências a estes resultados devem ser feitas ao artigo original.

Angiospermas (Dressler 1982; Ramirez et al. 2002; Roubik e Hanson 2004). Cerca de um quinto destas espécies são encontradas na Mata Atlântica (MA), das quais metade é endêmica desta região (Nemésio, 2009), caracterizando uma das principais áreas de endemismo de abelhas das orquídeas (Ramírez et al. 2010). Dentre as espécies endêmicas da MA (Nemésio, 2009; Moure et al., 2012), *Euglossa iopoecila* Dressler (1982) é uma abelha que ocorre quase exclusivamente em áreas costeiras, em formações de Floresta Ombrófila Densa (Galindo-Leal e Câmara 2003), variando do sul ao nordeste do Brasil (Faria Jr e Melo 2007; Nemésio 2009; Moure et al. 2012).

Ao longo de sua distribuição geográfica, as populações de *E. iopoecila* mostram algumas peculiaridades, como variações na coloração tegumentar, que podem ir de azul-violeta nas regiões mais ao sul (Santa Catarina e Paraná) a verde nas regiões mais ao norte (Rio de Janeiro, Espírito Santo, Bahia, Alagoas e Paraíba) com a ocorrência de fenótipos intermediários em zonas de contato (Ferrari e Melo 2014). Embora estas variações fenotípicas de coloração tegumentar não suportem a hipótese de espécies distintas ao longo da costa brasileira (Faria Jr e Melo 2007; Ferrari e Melo 2014), a repetição destes padrões de coloração em diferentes espécies de Euglossini ainda deixa dúvidas sobre a natureza de tal variação (Faria Jr e Melo 2007; Nemésio 2009).

Recentemente, vários estudos têm sugerido a influência da variação climática do Quaternário sobre a estrutura genética de diferentes organismos ao longo do domínio da Mata Atlântica (Carnaval et al. 2009; Ribeiro et al. 2011; Batalha-Filho e Miyaki 2016), inclusive para abelhas das orquídeas (López-Uribe et al. 2014; Penha et al. 2015). Duas hipóteses têm sido levantadas para explicar os padrões encontrados nestes estudos, retrações florestais com refúgios glaciais (Carnaval e Moritz 2008) e a expansão da MA para a plataforma continental sul-americana (Leite et al. 2016). No que diz respeito às abelhas das orquídeas, aparentemente há um componente extra, além da instabilidade climática, o comportamento filopátrico das fêmeas (Augusto e Garófalo 2004; López-Uribe et al. 2014), que potencialmente desempenha um importante papel na diversificação das linhagens de DNAm (López-Uribe et al. 2014; Penha et al. 2015).

Neste contexto, o presente estudo combina modelagem de nicho ecológico (ENM), análises de demografia histórica e estruturação genética para investigar a influência da instabilidade climática do Pleistoceno na estrutura genética de *E. iopoecila* ao longo de sua distribuição no domínio da MA. Para tanto, três questões principais foram levantadas: (a) A estrutura genética de *E. iopoecila* está relacionada à hipótese de refúgio MA ou expansão da MA? (b) Essa putativa estrutura é congruente com padrões apresentados por outros

organismos na MA? (c) Existe alguma relação entre os padrões de cor e estrutura genética em *E. iopocila* na MA?

6.2 MATERIAIS E MÉTODOS

6.2.1 Área de estudo

Os machos de *E. iopocila* foram amostrados em 14 remanescentes de Floresta Ombrófila Densa no domínio de MA ao longo da costa brasileira entre os estados de Santa Catarina e Alagoas, em uma distância de cerca de 2500 km (Tabela 6.1), correspondendo à área principal.

Tabela 6.1 Relação dos pontos amostrais de *Euglossa iopocila* ao longo da Mata Atlântica. Localidades (nome da reserva, município e estado), código da localidade, coordenadas geográficas (graus decimais), e tamanho amostral (*n*).

Localidade	Código	Latitude	Longitude	n
Vila Glória, São Francisco do Sul, SC	SC	-26,2439	-48,6383	10
Ilha do Superagui, Guaraqueçaba, PR	PR1	-25,3864	-48,2325	10
Reserva Natural Salto Morato, Guaraqueçaba, PR	PR2	-25,2900	-48,3172	10
Parque Estadual Ilha do Cardoso, Cananéia, SP	SP1	-25,1467	-47,9464	10
Continente, Cananéia, SP	SP2	-25,0172	-47,9194	10
Parque Estadual de Ilhabela, Ilhabela, SP	SP3	-25,8442	-45,3564	10
Parque Estadual da Serra do Mar, São Sebastião, SP	SP4	-23,7697	-45,6067	10
Parque Estadual da Serra do Mar (Picinguaba), Ubatuba, SP	SP5	-25,3642	-44,8236	10
Reserva Biológica União, Rio das Ostras, RJ	RJ	-22,4275	-42,0408	10
Reserva Biológica Sooretama, Linhares, ES	ES	-19,0625	-40,1186	9
Parque Nacional do Descobrimento, Itamaraju, BA	BA1	-17,1489	-39,3725	10
Parque Estadual da Serra do Conduru, Ilhéus, BA	BA2	-14,7275	-39,1133	10
Reserva Ecológica Michelin, Igrapiúna, Bahia	BA3	-13,7753	-39,1536	3
Estação Ecológica de Murici, Murici, AL	AL	-9,1900	-35,8814	4
Total				126

6.2.2 Amostragem e procedimentos moleculares

As abelhas foram coletadas entre 2009 e 2015. Nas amostragens de campo, iscas odoríferas de diferentes fragrâncias foram utilizadas para atrair os machos de Euglossini, os quais foram coletados ativamente com redes entomológicas ou passivamente por armadilhas de isca, de acordo com os procedimentos descritos em Giangarelli et al. (2015) e Ramalho et al. (2009), respectivamente.

O DNA total foi extraído de acordo com o protocolo em Freiria et al. (2012) e Penha et al. (2015), a partir de amostras da perna metatorácica direita ou do músculo torácico. Os vouchers dos espécimes utilizados foram depositados na coleção de abelhas do Museu Zoológico da Universidade Estadual de Londrina (MZUEL).

Os procedimentos laboratoriais para amplificação e sequenciamento de DNA foram realizados de acordo com Penha et al. (2015), com pequenas modificações. Resumidamente, as reações de PCR foram realizadas em um volume final de 15 µL, contendo 15 ng de DNA, 1 x tampão, 3 mM de MgCl₂, 0,5 µM de cada *primer*, 200 µM de dNTPs e 1U de Taq DNA polimerase (Invitrogen, Grand Island, NY, EUA). Para as análises, quatro segmentos de três genes mitocondriais foram amplificados: (a) *cytb*-755 pb com os *primers* CB-J-10612 e CB-N-11367 (SIMON et al., 1994); (b) *16S*-357 pb com *primers* confeccionados para o presente estudo - 16S-F 5'CGA GGT CGC AAA CAT CT e 16S-R 5'TAG CTG CAG TAT ATT GAC T; (c) *COI*-3'-610 pb com os *primers* CO1F e CO1R (Dick et al. 2004); e (d) *COI*-5'-*barcoding* (daqui em diante *COI*-5') - 657 pb com os *primers* LepF1 e LepR1 (Hebert et al. 2004). As amplificações foram realizadas em um termociclador (PCT-100 Peltier Thermal Cycler, J Research Inc., Quebec, Canadá), como segue: desnaturaç o inicial a 94°C/5 min, seguida de 35 ciclos (desnaturaç o a 94°C/60s, pareamento a 54°C/60s, e extens o a 72°C/80s) e uma extens o final a 72°C/10min. No entanto, para o gene *cytb*, a temperatura de pareamento (48°C/80s) e tempo de extens o final (72°C/100s) foram alterados.

Os produtos de PCR foram posteriormente purificados usando o Illustra™ Exo Pro-Star 1-Step (GE Healthcare, Buckinghamshire, Reino Unido) seguindo o protocolo do fabricante e submetidos a uma reaç o de sequenciamento utilizando o kit BigDye Terminator v3.1 (Applied Biosystems, Foster City, CA, EUA). Finalmente, os produtos de PCR foram sequenciados em sequenciador autom tico ABI-PRISM 3500-XL (Applied Biosystems, Foster City, CA, EUA). Todas as seq ncias utilizadas neste estudo foram depositadas no GenBank sob os seguintes n meros de acesso: *16S* (KX155649-KX155774); *cytb* (KX249836-KX249961); *COI*-5' (KX249962-KX250087); *COI*-3' (KX250088-KX250213).

6.2.3 An lise dos dados moleculares

O alinhamento de seq ncias, verificaç o de mutaç o, ediç o manual e concatenaç o foram realizados no MEGA v.6.05 (Tamura et al. 2013). Uma vez que os genes mitocondriais n o sofrem eventos de recombinaç o, todos os segmentos (*16S*, *cytb*, *COI*-5' e *COI*-3') foram analisados como uma  nica seq ncia (L pez-Urbe et al. 2014). Para obter uma vis o geral das relaç es de hapl tipos, as redes com algoritmos de uni o mediana foram constru das usando o PopArt v1.7 (Leigh e Bryant 2015). Os hapl tipos foram determinados pelo software DnaSP 5.0 (Librado e Rozas 2009), que tamb m foi usado para calcular a diversidade nucleot dica e haplot pica e executar testes de distribuiç o de *mismatch* de an lise demogr fica com base no modelo de tamanho populacional constante Tajima D (Tajima

1989) e F_s de Fu (Fu 1997) teste de neutralidade sem grupo externo. O DnaSP 5.0 também foi utilizado para gerar as estatísticas R_2 (Ramos-Onsins e Rozas 2006), recomendado para pequenos tamanhos de população (Ramírez-Soriano et al. 2008). O módulo de simulações coalescentes implementado no DnaSP 5.0 foi utilizado para testar a significância estatística dos testes de neutralidade com base no valor θ sem recombinação, com um intervalo de confiança de 95% e 10.000 replicações.

Além dos testes de neutralidade, também foi inferido a variação do tamanho efetivo populacional ao longo do tempo através do método *Extended Bayesian Skyline Plot* (EBSP) (Drummond e Bouckaert 2015) disponível no pacote de software BEAST 1.6.2 (Drummond e Rambaut 2007). Esta é uma abordagem não paramétrica mais adequada às árvores que não são totalmente resolvidas ou com dados pouco variáveis (Ho e Shapiro 2011). Foram realizadas três corridas independentes com 100 milhões de passos nas Cadeias de Markov e Monte Carlo para cada *cluster* identificados pelas análises de estrutura genética, utilizando os seguintes parâmetros: árvore inicial UPGMA, modelo de substituição GTR+I+G, como indicado por jModelTest 1.1 (Posada 2008) com base na Critério de Informação de Akaike corrigido para um pequeno tamanho amostral (AICc). Os parâmetros e as árvores foram amostrados a cada 10.000 cadeias com um descarte de 25%, e as três corridas independentes foram combinadas usando LogCombiner (pacote BEAST) para analisar a convergência do EBSP. Para estimar a datação das variações do tamanho efetivo populacional, foi utilizado um relógio molecular relaxado com distribuição *log-normal* não correlacionado e uma taxa de substituição de 1,3% (d.p. 0,2%) por milhão de anos (Ma^{-1}), comumente aplicadas ao DNAm de abelhas (Dick et al. 2004; Batalha-Filho et al. 2010). O desempenho EBSP (valores ESS > 200) e a convergência de parâmetros entre corridas foi verificado usando o Tracer v1.6 do pacote BEAST (disponível em: <http://beast.bio.ed.ac.uk/tracer>).

O software jModelTest v1.1 foi utilizado para selecionar os modelos de substituição de base em cada região gênica e na matriz concatenada com base no AICc, como recomendado (Posada e Buckley 2004). As diferenças genéticas par-a-par e as estatísticas Φ (1000 permutações) foram calculadas pelo Arlequin v3.11 (Excoffier et al., 2005) utilizando o modelo Kimura-2-Parameters, o melhor modelo disponível de acordo com os resultados do jModelTest.

Para testar limites e estruturas populacionais, utilizamos o programa BAPS v6.2 (Análise Bayesiana da Estrutura da População) (Corander et al. 2008), o qual atribui indivíduos à *K-clusters*, com base nos dados genéticos, sem informação prévia da população. Para esse fim, utilizamos agrupamentos com a opção de locos-ligados (*linked loci*) no modelo

de ligação independente (Corander e Tang 2007) com 10 corridas para cada K -grupo, até 50 indivíduos de referência amostrados por população e 20 interações por indivíduo. BAPS também foi executado com a opção de agrupamento espacial, onde as coordenadas geográficas foram combinadas com dados genéticos para definir o número K mais provável para as populações. Em todas as análises, o número K variou de 1 a 20. A Análise Espacial da Variância Molecular (SAMOVA) (Dupanloup et al. 2002) foi executada no software SPADS v1.0 (Dellicour e Mardulyn 2014) para identificar agrupamentos genéticos com e sem informações geográficas. Com base nos resultados da análise BAPS e SAMOVA, realizamos uma série de AMOVAs hierárquicas e comparações Φ_{ST} emparelhadas (10000 permutações) no Arlequin v3.11 (Excoffier et al. 2005). Os cenários testados incluíram três *clusters* (*cluster* 1, sítios de amostragem do Sul, *cluster* 2, Central e Norte, *cluster* 3, sítio de amostragem Nordeste; ver em resultados).

Para testar a possível correlação entre distâncias geográficas e genéticas, realizamos um teste de Mantel (1000 permutações) em Genepop v.1.2. (Raymond e Rousset 1995). Os logaritmos naturais (\ln) de distâncias geográficas foram plotados contra estimativas de F_{ST} par-a-par linearizadas ($F_{ST}/(1 - F_{ST})$).

Diferenças significativas entre a diversidade nucleotídica dos clados Sul e Central (ver seção de estrutura populacional) foi testada via teste de Mann-Whitney executado no programa Graphpad InStat v. 3:06. As amostras foram divididas nesses dois grupos com base nas diferenças indicadas pela análise genética. As amostras do estado de Alagoas (AL) não foram incluídas nesta análise devido ao pequeno tamanho amostral. O teste de Mann-Whitney foi executado em dados gerados separadamente para cada segmento de DNAm (*16S*, *cytb*, *COI-5* 'e *COI-3*'), bem como para a matriz concatenada. Nestas análises, p-valores inferiores a 0,05 foram considerados estatisticamente significativos.

6.2.4 Modelagem de Nicho Ecológico

Uma vez que os processos evolutivos são intrinsecamente influenciados pela variação ambiental em relação ao tempo e espaço, o algoritmo máximo de entropia foi implementado no programa MaxEnt v 3.3.3 k (Phillips et al. 2006) para modelar os nichos ecológicos atuais e passados para *E. iopoecila*. Para tanto, 19 variáveis bioclimáticas (Hijmans et al. 2005) e 45 ocorrências georreferenciadas de *E. iopoecila* (veja Apêndices) foram utilizadas, projetando os resultados para três cenários climáticos diferentes: médio Holoceno (11 Ka a. p.); último máximo glacial (21 Ka a. p.) e último período interglacial (130 Ka a. p.). As ocorrências foram exclusivamente extraídas de uma ampla revisão bibliográfica para evitar possíveis erros

de registro em bancos de dados públicos, devido às questões históricas de taxonomia relativas a esta espécie (para detalhes veja: o caso de *E. iopoecila*).

As 19 variáveis climáticas do modelo de circulação global MIROC-ESM para os três cenários pretéritos foram baixadas em resolução de 2,5 minuto de arco (arcmin), então o software QGIS 2.14 (Quantum GIS Development Team 2009) foi utilizado para reprojetar essas camadas ao mesmo *datum* utilizados para os dados “atuais” (WGS84, 30 segundos de arco - arcsec). Como *E. iopoecila* é uma espécie endêmica da costa brasileira, os modelos de nicho ecológico (ENM) foram gerados limitando-se às fronteiras da América do Sul. Assim, foi considerado o espaço entre as latitudes 13.3702 e -56.5692, e longitudes -81.8505 e -33.9972.

As camadas resultantes do MaxEnt foram visualizadas no QGIS 2.14. Para avaliar a precisão dos modelos, foi verificado se os ENMs eram melhores do que um modelo aleatório usando um teste binomial baseado em omissão e áreas previstas (Merow et al. 2013). Além disso, realizamos 10 replicações retendo 25% das amostras como dados de teste e os demais 75% como dados de treinamento do modelo para calcular a área sob a curva (AUC) tanto para treinamento como para dados de teste. AUC foi utilizada para medir a sensibilidade e especificidade dos modelos.

Como em López-Urbe et al. (2014), mantivemos todas as variáveis bioclimáticas independentemente da correlação, uma vez que excluir variáveis correlacionadas muitas vezes faz pouca diferença no ENM final (Lacey Knowles e Alvarado-Serrano 2010). O mesmo efeito foi observado quando se excluiu as variáveis correlacionadas para a distribuição de *E. iopoecila*, uma vez que os modelos com menos variáveis não diferiram significativamente em termos de interpretação dos ENM. Para produzir os mapas binários, utilizou-se os limiares de igual sensibilidade e especificidade, os quais reduzem a diferença entre a taxa de omissão e a área prevista (Elith et al. 2011; Merow et al. 2013).

6.3 RESULTADOS

No total, 126 machos de *E. iopoecila* foram analisados e um conjunto de dados compreendendo 2234 (pares de bases) pb: 357 pb no segmento *16S*, 651 pb em *cytb*, 650 pb em *COI-5'* e 576 pb em *COI-3'*. As sequências analisadas resultaram em 44 sítios polimórficos, seis dos quais eram *singletons* e 38 parcimoniosamente informativos. A composição de frequência nucleotídica média foi: T = 45,9%, C = 10,1%, A = 33,7%, G = 10,3%.

A matriz concatenada resultou em 76 haplótipos (h), sete dos quais foram compartilhados por pelo menos duas amostras (Figura 6.1-A). Entre os quatro segmentos sequenciados, o número de haplótipos encontrados variou de 10 a 28 (Tabela 6.2). As estimativas de diversidade haplotípica (H_d) e a diversidade de nucleotídeos ($\pi\%$) variaram de 0,995 a 1,0 e 0,239 a 0,336%, respectivamente (Tabela 6.2). Para todos os segmentos e seqüências concatenadas, os maiores valores de H_d e π foram encontrados entre as amostras do *cluster* central, com diferenças estatísticas significativas entre os grupos do sul e central para os valores π (Tabela 6.2).

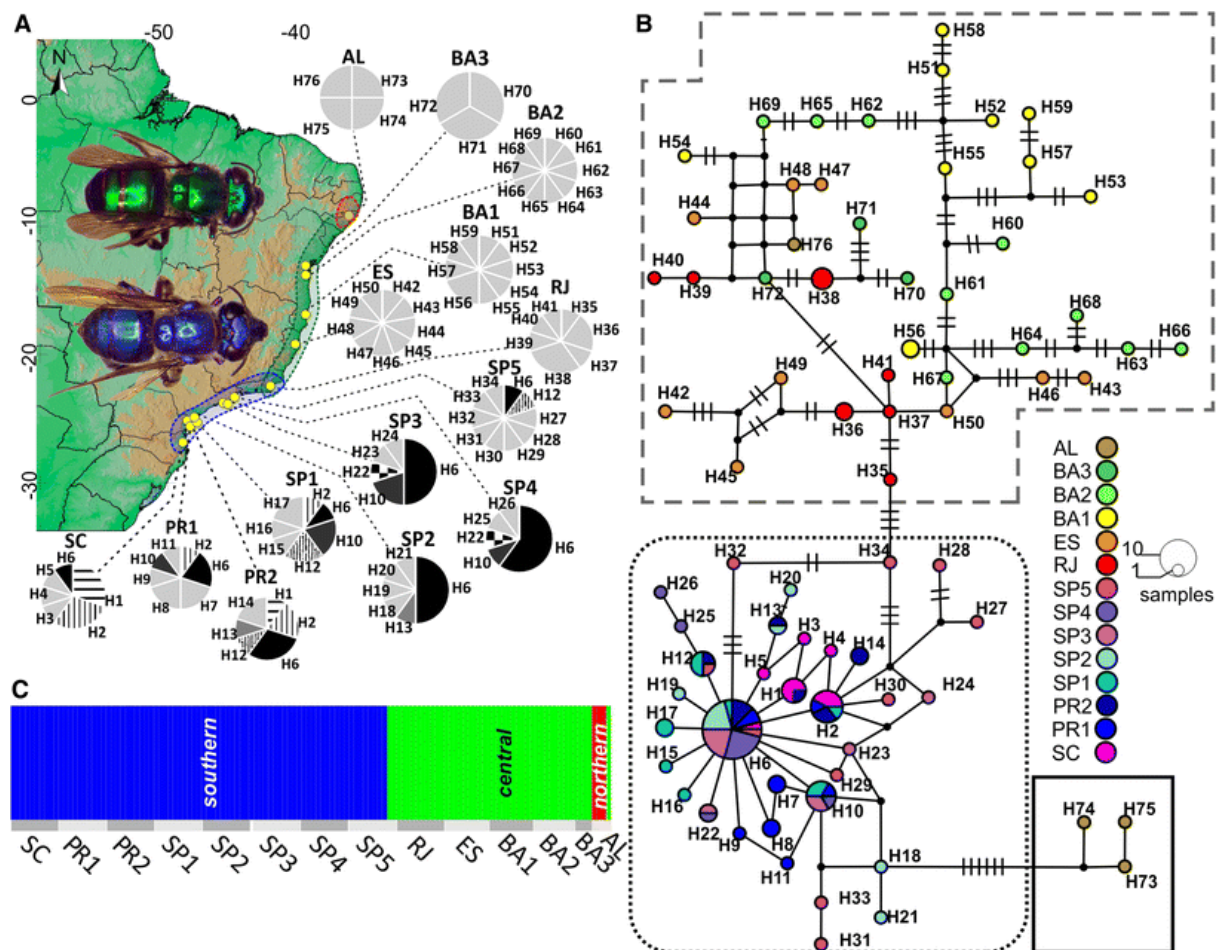


Figura 6.1 A) Mapa parcial do Brasil, mostrando os 14 locais de estudo (SC, PR1, PR2, SP1, SP2, SP3, SP4, SP5, RJ, ES, BA1, BA2, BA3 e AL), representados por pontos amarelos (Software QGIS 2.12). Os gráficos de pizza ligados aos pontos amarelos indicam a frequência e distribuição de cada haplótipo; os haplótipos não compartilhados são mostrados em cinza claro. B) Rede de relacionamento para os 76 haplótipos (*Median-joining*) amostrados (H1-H76) obtidos para as seqüências de DNAm_t concatenadas. Os tamanhos do círculo são proporcionais à frequência de haplótipos (H). Os três haplogrupos inferiores (ou *clusters*) são representados por linhas tracejadas, pontilhadas e sólidas. Cada círculo representa um haplótipo, as cores indicam localidades e o tamanho, o número de indivíduos; cada barra transversal sobre as linhas conectoras representa um passo mutacional; círculos pretos representam os vetores médios. C) Análise Bayesiana da estrutura populacional (BAPS) de *E. iopoeicola* baseado em 2234 pb do DNAm_t. Cada indivíduo é representado por uma barra vertical fina dividida em grupo coloridos por seus grupos genéticos (azul = *cluster* 1; verde = *cluster* 2; vermelho = *cluster* 3). As áreas destacadas no mapa representam a distribuição geográfica dos clusters. Os códigos do sítio são os mesmos que na Tabela 5.1. (retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18:539; <https://doi.org/10.1007/s10592-016-0905-7>)

Tabela 6.2 Medidas de diversidade genética de DNAm_t recuperadas para matriz concatenada e para cada uma das quatro sequências de genes (*16S*, *cytb* e *COI*) analisadas: (*h*) número de haplótipos, (*Hd*) diversidade de haplótipos, (π) diversidade de nucleotídeos, (*n*) = número de amostras e (*dp*) desvio padrão (dp de $\pi < 1 \times 10^{-4}$)

L	N	<i>16S</i> (357 bp)			<i>COI-5'</i> (650 bp)			<i>COI-3'</i> (576 bp)			<i>cytb</i> (651 bp)			Concatenado		
		<i>h</i>	<i>Hd</i> ± d.p.	π	<i>h</i>	<i>Hd</i> ± d.p.	π	<i>h</i>	<i>Hd</i> ± d.p.	π	<i>h</i>	<i>Hd</i> ± d.p.	π	<i>h</i>	<i>Hd</i> ± s.d.	π
SC	10	2	0,533 ± 0,07	0,149	1	0,000 ± 0,00	0,000	2	0,356 ± 0,02	0,062	2	0,556 ± 0,00	0,085	6	0,867 ± 0,00	0,065
PR1	10	2	0,200 ± 0,02	0,056	1	0,000 ± 0,00	0,000	2	0,533 ± 0,00	0,093	3	0,711 ± 0,00	0,137	7	0,933 ± 0,00	0,073
PR2	10	2	0,533 ± 0,01	0,149	2	0,200 ± 0,02	0,031	2	0,200 ± 0,02	0,035	4	0,644 ± 0,02	0,116	6	0,889 ± 0,00	0,076
SP1	10	2	0,200 ± 0,02	0,056	3	0,511 ± 0,02	0,085	2	0,356 ± 0,02	0,062	3	0,511 ± 0,02	0,085	7	0,933 ± 0,00	0,075
SP2	10	3	0,511 ± 0,03	0,156	1	0,000 ± 0,00	0,000	3	0,622 ± 0,01	0,247	2	0,356 ± 0,02	0,055	6	0,778 ± 0,01	0,104
SP3	10	2	0,200 ± 0,02	0,112	1	0,000 ± 0,00	0,000	3	0,622 ± 0,01	0,123	2	0,200 ± 0,02	0,031	5	0,756 ± 0,01	0,059
SP4	10	2	0,200 ± 0,02	0,056	2	0,356 ± 0,02	0,109	2	0,200 ± 0,02	0,035	2	0,200 ± 0,02	0,031	5	0,667 ± 0,02	0,059
SP5	10	5	0,756 ± 0,01	0,498	3	0,622 ± 0,01	0,109	4	0,644 ± 0,02	0,239	2	0,356 ± 0,02	0,109	10	1,000 ± 0,00	0,205
RJ	10	3	0,511 ± 0,03	0,230	2	0,200 ± 0,02	0,031	6	0,889 ± 0,00	0,320	3	0,689 ± 0,01	0,212	8	0,956 ± 0,00	0,190
ES	9	4	0,778 ± 0,01	0,669	3	0,750 ± 0,00	0,154	3	0,750 ± 0,00	0,260	5	0,806 ± 0,01	0,256	9	1,000 ± 0,00	0,293
BA1	10	4	0,644 ± 0,02	0,305	3	0,689 ± 0,01	0,126	6	0,889 ± 0,00	0,285	4	0,822 ± 0,00	0,608	9	0,978 ± 0,00	0,336
BA2	10	4	0,778 ± 0,01	0,442	3	0,733 ± 0,00	0,144	6	0,889 ± 0,00	0,285	4	0,800 ± 0,00	0,273	10	1,000 ± 0,00	0,266
BA3	3	2	0,667 ± 0,09	0,373	2	0,667 ± 0,09	0,103	3	1,000 ± 0,07	0,231	3	1,000 ± 0,07	0,307	3	1,000 ± 0,07	0,239
AL	4	2	0,500 ± 0,07	0,280	3	0,833 ± 0,04	0,154	2	0,500 ± 0,07	0,174	3	0,833 ± 0,04	0,486	4	1,000 ± 0,03	0,276
Sul	80	6	0,410 ± 0,00	0,168**	5	0,232 ± 0,00	0,044*	8	0,461 ± 0,00	0,118*	10	0,503 ± 0,00	0,097**	34	0,896 ± 0,00	0,098*
Cen	42	10	0,732 ± 0,00	0,455**	4	0,621 ± 0,00	0,113*	9	0,868 ± 0,00	0,283*	16	0,928 ± 0,00	0,567**	38	0,995 ± 0,00	0,344*
Nor	4	2	0,500 ± 0,07	0,280	3	0,833 ± 0,04	0,154	2	0,500 ± 0,07	0,174	3	0,833 ± 0,04	0,486	4	1,000 ± 0,03	0,276
Total	126	14	0,720 ± 2,33	0,469	10	0,641 ± 0,03	0,150	14	0,758 ± 0,02	0,321	28	0,793 ± 0,01	0,489	76	0,958 ± 0,02	0,344

Sul, Cen e Nor = clusters identificados pela análise BAPS e SAMOVA. Valores de Mann-Whitney: *16S*: $U' = 36,0$; *cytb*: $U' = 40,0$; *COI-5'*: $U' = 34,5$; *COI-3'*: $U' = 38,0$; dados concatenados: $U' = 39,0$. * $p < 0,05$; ** $p < 0,001$

6.3.1 Estrutura populacional e padrões de diferenciação

A rede de relacionamento de haplótipos para as sequências concatenadas revelou divisões importantes entre três grupos de amostras (Figura 6.1-B): (a) um grupo que inclui um conjunto de amostras da distribuição sul (SC para SP5), incluindo abelhas com coloração tegumentar violeta; (b) um segundo grupo, compreendendo amostras da distribuição centro-norte (de RJ a BA3) com coloração verde; (c) e um terceiro conjunto de haplótipos também de abelhas verdes coletadas em AL (Figura 6.1-B), extremo norte da distribuição de *E. iopoecila*. Estes três haplogrupos (SC para SP5; RJ para BA3; e AL) são a seguir denominados clusters ou clados sul, centro e norte, de acordo com sua distribuição geográfica e os grupos identificados pela análise BAPS (Figura 6.1-C).

Duas diferenças marcantes podem ser observadas na rede de haplótipos para os clados sul e central. O primeiro haplogrupo possui uma configuração de estrela, em que o haplótipo mais frequente (H6) é centralizado com vários haplótipos que derivam dele por um passo mutacional (Figura 6.1-B), sugerindo que H6 é o haplótipo ancestral desse haplogrupo. Por outro lado, as ligações entre os haplótipos no *cluster* central foram mais complexas. Vários haplótipos estão separados por dois ou mais passos mutacionais, ligados em múltiplos pontos por vários vetores médios, um padrão que sugere haplótipos “antigos” derivando em um cenário de estabilidade populacional ao longo do tempo (Figura 6.1-B).

Do mesmo modo, o compartilhamento de haplótipos entre os sítios amostrais revelou padrões diferentes para os *clusters* do sul e central. Por exemplo, haplótipos no grupo sul (H1, H2, H6, H10, H12 e H22) tendem a ser comuns a dois ou mais locais, enquanto nenhum haplótipo compartilhado foi observado no *cluster* central (Figura 6.1-A). Vale ressaltar que no *cluster* sul, o haplótipo H6 é comum a todos os sítios, em contraste ao padrão encontrado para o *cluster* central.

SAMOVA também revelou um padrão de estrutura genética com três grupos distintos ($\Phi_{CT} = 0,62606$, $p < 0,001$), semelhante ao padrão encontrado para redes de haplótipos. Este foi também o caso da AMOVA hierárquica (melhor AMOVA, Tabela 6.3), na qual o valor Φ_{CT} mais alto foi obtido quando as amostras foram agrupadas como indicado pela SAMOVA e o próximo valor mais alto foi obtido quando a amostra AL (*cluster* Norte) foi incluído no *cluster* Central. A AMOVA também revelou uma importante proporção da variação global (30,14%) dentro das amostras (ex. PR1, PR2...) e uma variação de 7,25% entre os sítios dentro dos *clusters* (Tabela 6.3), o que salienta a grande diferenciação entre *clusters* no contexto geral. No entanto, apesar deste baixo valor (7,25%) para a variação entre as amostras

(ou sítios) dentro dos *clusters*, note que valores de Φ_{ST} significativos ($\Phi_{ST} = 0,69596$, $p < 0,001$) foram observados entre populações, por vezes ainda maior do que o valor entre *clusters* (Φ_{CT} ; Tabela 6.3).

Além disso, as comparações F_{ST} em pares revelaram valores significativos para a maioria dos pares analisados, ou seja, dos 91 pares comparados, apenas 17 apresentaram valores F_{ST} não significativos (veja em Apêndices). Visivelmente, desses 17 pares, 14 pertenciam ao *cluster* sul e isso se aplica a todas as amostras desse *cluster*. Também foi encontrada uma correlação positiva significativa entre a distância geográfica (km) e a estrutura genética ao longo da distribuição de *E. iopoecila* ($r > 0,69$, $p < 0,05$), indicando um efeito de isolamento por distância entre as populações (veja em Apêndices).

Tabela 6.3 Análise da variância molecular (AMOVA) para 2234 pb de 126 indivíduos de *Euglossa iopoecila* amostrados ao longo de 14 sítios no domínio da Mata Atlântica. As divisões hierárquicas seguem os três grupos (*clusters*) identificados pela análise SAMOVA e BAPS.

Fonte de variação	g.l.	Soma dos quadrados	Componentes de variação	% de variação	Estatísticas Φ
Entre <i>clusters</i>	2	226,792	3,56421 Va	62,61	$\Phi_{CT} = 0,62606^*$
Entre sítios dentro de <i>cluster</i>	11	61,272	0,41274 Vb	7,25	$\Phi_{ST} = 0,69856^*$
Dentro de sítios	112	192,206	1,71612 Vc	30,14	$\Phi_{SC} = 0,19388^*$
Total	125	480,27	5,69307		

* $p < 0,001$

A análise do programa BAPS também corroborou a distribuição em três *clusters* na análise de agrupamento de locos-ligados. Semelhante à rede de haplótipos e SAMOVA, o BAPS agrupou as amostras nos mesmos três grupos: *cluster* Sul (SC, PR1, PR2, SP1, SP2, SP3, SP4 e SP5), *cluster* Central (RJ, ES, BA1, BA2, BA3) e *cluster* do Norte (AL somente) (Fig. 6.1-C). Assim, os limites dos principais *clusters* estão entre os sítios SP5 e RJ, acrescentando peso à ideia de descontinuidade filogeográfica entre *clusters* do Sul e Central.

6.3.2 Análises demográficas

As análises demográficas revelaram sinais de crescimento populacional recente em amostras do *cluster* Sul da MA. Enquanto as análises *mismatch* para *cluster* Sul indicaram um padrão unimodal, as sequências dos *clusters* Central e Norte apresentaram padrões multimodais, sugerindo expansão e estabilidade populacionais, respectivamente (Figura 6.2). O mesmo padrão foi observado quando cada amostra foi analisada individualmente (dados

não apresentados) reforçando a ideia de crescimento populacional recente em amostras do *cluster* Sul.

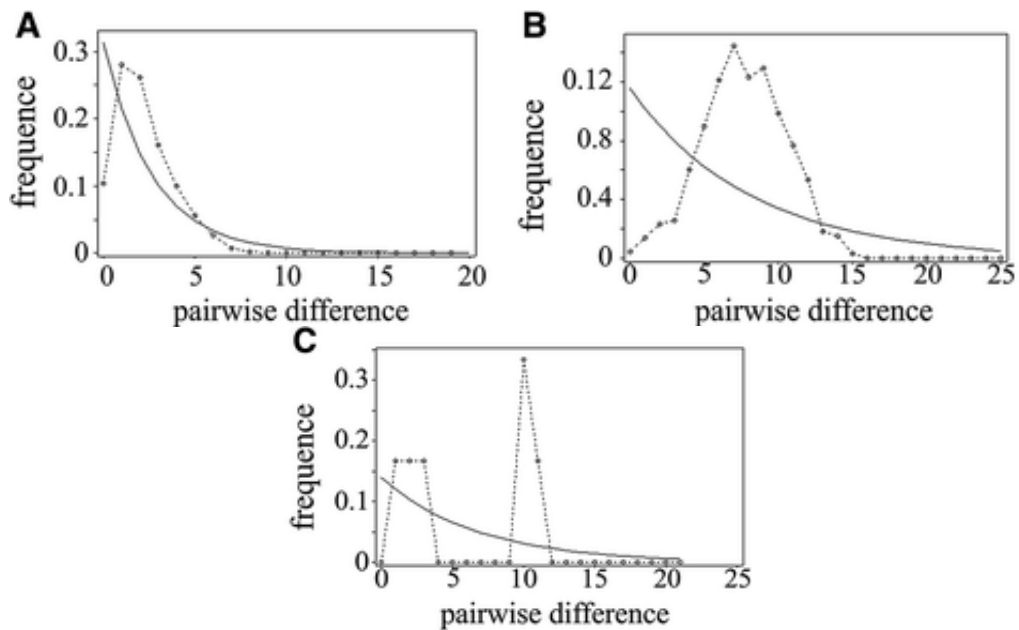


Figura 6.2 Distribuições das diferenças entre pares de sequências observadas (linha pontilhada) contra um padrão esperado em um modelo de tamanho populacional constante (linha sólida) baseados em 2234 pb do DNAmT de *Euglossa iopoeicila* amostrada ao longa da Mata Atlântica. A) *cluster* Sul; B) *cluster* Central; C) *cluster* Norte. As diferenças em pares de sequências (eixo X) são mostradas contra a frequência das respectivas diferenças (eixo Y). Padrões não modais correspondem ao crescimento rápido da população, enquanto os padrões multimodais correspondem a tamanhos de população constantes. (Retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18: 539 <https://doi.org/10.1007/s10592-016-0905-7>)

Esses resultados foram confirmados estaticamente pelos testes de neutralidade, D de Tajima, F_s de Fu e R_2 , apresentando sinais significativos de expansão populacional tanto para o *cluster* Sul como um todo, quanto para amostras testadas individualmente (Tabela 6.4). Do mesmo modo, o EBSP mostrou uma expansão demográfica para o *cluster* Sul na região sul da MA, apontando um aumento do tamanho efetivo populacional em torno de 20 Ka a. p., após o pico do último máximo glacial (Figura 6.3).

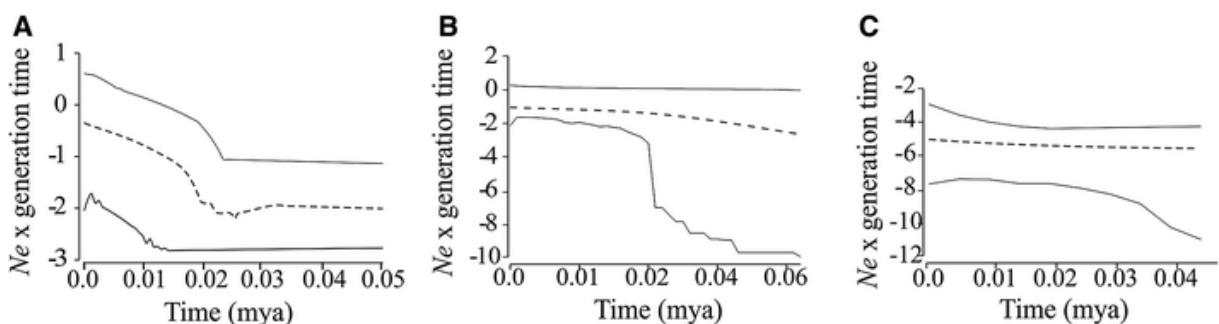


Figura 6.3 Estimativas do *Extended Bayesian Skyline Plot* (EBSP) para os diferentes *clusters* genéticos de *Euglossa iopoeicila*: A) *cluster* Sul; B) *cluster* Central; C) *cluster* Norte; A linha tracejada mostra a estimativa mediana do EBSP e as linhas sólidas superior e inferior correspondem aos limites de densidade à posteriori dentro de um intervalo de confiança de 95%. O eixo X representa o tempo em milhões de anos (Ma a. p.) e o eixo Y o logaritmo natural do tamanho efetivo da população estimado (N_e) por geração. (Retirado de Frantine-Silva et al. *Conserv Genet* (2017) 18: 539 <https://doi.org/10.1007/s10592-016-0905-7>)

Tabela 6.4 Parâmetros demográficos e testes de neutralidade para mudanças no tamanho da população: teste de D de Tajima, Fs e R₂ com base em quatro conjuntos de dados de sequências de DNAm_t de 14 localidades de *Euglossa iopoeila* na Mata Atlântica Brasileira.

Sítios	n	16S			<i>cytb</i>			<i>COI-5'</i>			<i>COI-3'</i>			DNAm _t		
		D	Fs	R ₂	D	Fs	R ₂	D	Fs	R ₂	D	Fs	R ₂	D	Fs	R ₂
SC	10	1,3	1,03	0,27	1,46	1,1	0,28	0,00	0,00	0,00	0,01	0,42	0,18	1,30	-2,74*	0,24
PR1	10	-1,11	-0,34	0,30	0,83	0,25	0,22	0,00	0,00	0,00	1,30	1,03	0,27	0,57	-3,99*	0,19
PR2	10	1,30	1,03	0,27	-1,03	-1,47	0,15*	-1,11	-0,34	0,30	-1,11	-0,34	0,30	-0,84	-2,29*	0,13
SP1	10	-1,11	-0,34	0,30	-0,69	-0,59	0,17	-0,69	-0,59	0,17	0,01	0,42	0,18	-0,89	-3,89*	0,12*
SP2	10	-0,69	-0,59	0,17	0,01	0,42	0,18	0,00	0,00	0,00	0,02	1,19	0,18	-0,24	-1,43	0,16
SP3	10	-1,4	0,59	0,30	-1,11	-0,34	0,30	0,00	0,00	0,00	0,02	-0,16	0,18	-1,04	-1,59	0,15
SP4	10	-1,11	-0,34	0,30	-1,11	-0,34	0,30	0,02	1,52	0,18	-1,11	-0,34	0,3	-1,04	-1,59	0,11*
SP5	10	0,02	-0,87	0,16	0,02	1,52	0,18	0,02	-0,16	0,18	-0,10	-0,21	0,17	-0,02	-6,29*	0,16
RJ	10	0,53	0,11	0,21	1,08	1,13	0,23	-1,11	-0,34	0,30	0,17	-2,05	0,17	0,41	-2,45	0,09
ES	09	1,26	0,78	0,24	-0,40	-1,26	0,17	1,23	0,35	0,25	1,37	1,15	0,25	0,91	-3,98	0,20
BA1	10	0,10	-0,70	0,18	1,72	2,29	0,25	0,53	0,11	0,21	0,63	-2,36	0,21	1,16	-2,36	0,22
BA2	10	1,76	0,08	0,26	0,99	0,33	0,22	1,03	0,35	0,23	0,63	-2,36	0,21	1,32	-5,23	0,23
BA3	03	0,00	1,06	0,47	0,00	-0,69	0,27	0,00	0,20	0,47	0,00	-1,22	0,24	0,00	0,46	0,12
AL	04	-0,71	1,10	0,43	-0,31	0,81	0,28	-0,71	-0,89	0,25	-0,71	1,10	0,43*	-0,58	-0,25	0,30
Cluster1	80	-1,18	-2,18	0,05	-1,82	-7,19*	0,03*	-1,33	-3,25*	0,04	-0,72	-3,96	0,07	-1,73	-27,15*	0,04*
Cluster2	42	0,44	-3,25	0,13	0,42	-4,43*	0,13	0,12	-0,05	0,13	1,01	-2,28	0,16	0,61	-24,95*	0,13
Cluster3	04	-0,71	1,10	0,43	-0,31	0,81	0,28	-0,71	-0,89	0,25	-0,71	1,10	0,43	-0,58	-0,25	0,30

Clusters 1, 2 e 3(Sul, Central e Norte, respectivamente) = grupos identificados pela análise do BAPS e SAMOVA. * $p < 0,05$

6.3.3 Modelagem de nichos ecológicos e paleodistribuição

Os modelos de nicho ecológico (ENMs) foram bastante precisos, conforme indicado pela alta proporção de áreas sob a curva (AUC) em todos os modelos (treinamento $AUC > 0,993 \pm 0,003$, teste $AUC > 0,994 \pm 0,001$). Como dito anteriormente, os modelos são apresentados para quatro cenários diferentes: “atual” (média 1950-2000), médio Holoceno (MID; 6 Ka a. p.), último máximo glacial (LGM, 21 Ka a. p.) e último período interglacial (LIG; 130 Ka a. p.; Figura 6.4).

A distribuição prevista do ENM para o tempo “atual” considerando os dados de ocorrência de *E. iopoecila* (Figura 6.4-A) quase se sobrepõe completamente à distribuição conhecida desta espécie. Apenas algumas áreas mais ao sul do limite de distribuição conhecida para *E. iopoecila* foram apontados como adequáveis de acordo com o modelo. O ENM “atual” ainda revela duas outras características interessantes: (a) uma diminuição da adequabilidade ambiental para *E. iopoecila* de acordo com as duas principais discontinuidades filogeográficas observadas entre os *clusters* Sul e Central, bem como os *clusters* Central e Norte; (b) a população do *cluster* Norte é disjunta, com baixa adequabilidade em regiões de conexão com o *cluster* Central.

Nosso ENM para o Holoceno médio (MID, 6 Ka a. p.; Figura 6.4-C) mostra um cenário mais conectado do que o atual, com mais áreas de conexão entre os três *clusters*, os quais foram drasticamente reduzidos no ENM do último máximo glacial (LGM, 21 Ka a. p.; Figura 6.4-D). Durante o LGM, a distribuição de *E. iopoecila* parece restrita à três regiões, enquanto observa-se maior probabilidade de ocorrência nas áreas preditas como refúgios da Bahia e Pernambuco (*sensu* CARNAVAL; MORITZ, 2008). Finalmente, o paleomodelo para o último máximo interglacial (LIG, 130 Ka a. p., Figura 6.4-E) mostra um ENM semelhante ao modelo “atual”, apenas as áreas centrais da MA aparecem menos adequadas do que no ENM atual, com os *clusters* central e norte localizados mais ao norte e maior concentração de áreas adequadas no refúgio de Pernambuco.

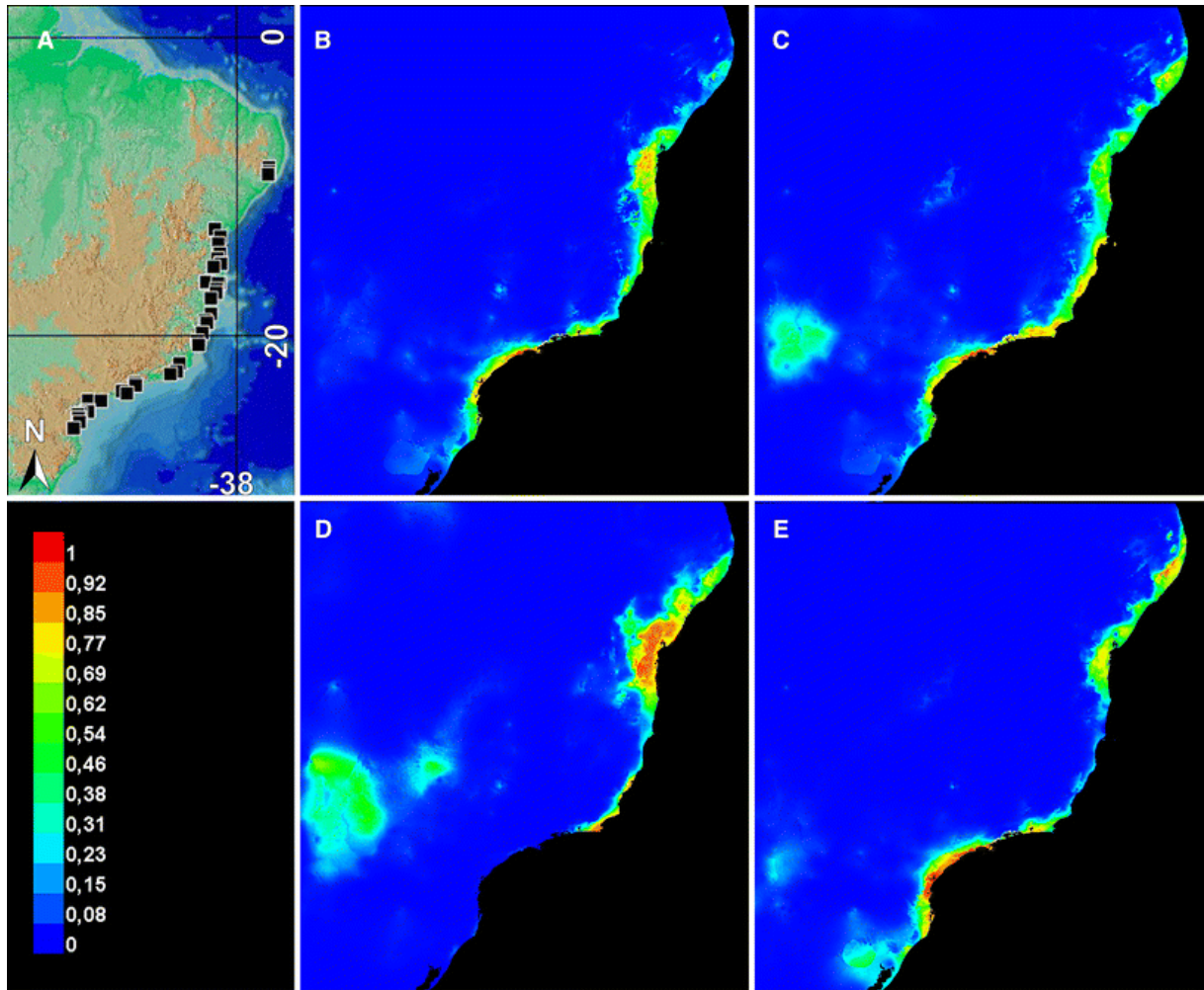


Figura 6.4 Modelagem de distribuição de espécies de *Euglossa iopoecila* nos últimos 130 mil anos. A) Mapa com 45 locais de ocorrência (quadrados pretos) usados como entrada no MaxEnt. B) Modelo de nicho ecológico (ENM) de *Euglossa iopoecila* para o tempo “atual”; C) médio Holoceno (MID; 6 mil a.a.p); D) último máximo glacial (LGM; 21 Ka a. p.); E) último período interglacial (LIG; 130 mil Ka a. p.). As cores mais quentes indicam maior adequação para a ocorrência das espécies, conforme indicado na legenda. (Retirado de Frantini-Silva et al. *Conserv Genet* (2017) 18: 539 <https://doi.org/10.1007/s10592-016-0905-7>)

6.4 DISCUSSÃO

Os resultados até aqui apresentados para os dados de *E. iopoecila* sustentam a hipótese de que as mudanças climáticas do Pleistoceno desempenharam um importante papel na formação da atual estrutura genética desta abelha das orquídeas. Tanto as redes de haplótipos, quanto as análises de SAMOVA e BAPS foram muito congruentes ao apontar três *clusters* em toda a costa leste da Mata Atlântica brasileira, os quais têm alta concordância geográfica com uma série de estudos que relatam padrões similares de estrutura filogeográfica ao longo da MA. Por exemplo, vários grupos de animais, incluindo invertebrados (Batalha-Filho et al. 2010; Cardoso et al., 2015), vertebrados (Amaral et al., 2013; Cabanne et al., 2014) e plantas (Ribeiro et al., 2011), mostram distribuição de agrupamento genético similar, nas regiões sul ou central da MA. Além disso, os dois principais *clusters* (Sul e Central) são consistentes com

a distribuição geográfica dos fenótipos de cor de *E. iopoecila*, azul no Sul e verde nas regiões Centro-Norte da MA, indicando uma possível relação entre a variação genética e fenotípica.

A ideia de mudanças climáticas afetando a diversidade dos organismos em florestas tropicais não é nova (hipótese de Refúgios Florestais – HRF; HAFFER, 1969) e, há anos, tem sido vista como um fator determinante na diversificação atual da floresta tropical (Carnaval e Moritz 2008; Batalha-Filho e Miyaki 2011). Entre várias outras teorias propostas para explicar a diversificação da MA, o modelo de estabilidade-extinção (daqui em diante modelo CM), proposto por Carnaval e Moritz (2008) e Carnaval et al. (2009), atraiu muito apoio. Segundo o modelo CM, enquanto as áreas localizadas na parte mais austral da MA experimentaram contração no LGM (21 Ka a. p.), seguido de expansão durante o Holoceno (~ 6 Ka a. p.), as áreas do norte da MA constituíram refúgios estáveis, grandes ilhas de floresta úmidas ao longo do tempo (Carnaval e Moritz 2008). Deste modo, o modelo CM prediz (a) ENMs específicos de espécies que predizem refúgios ao norte da Mata Atlântica no LGM, seguido de expansão distribucional no Holoceno; (b) flutuações demográficas ao longo do tempo em que as porções ao sul tendem a ser menos estáveis que os refúgios ao norte; (c) sinais de expansão populacional após o LGM nas populações do sul; (d) maior diversidade genética em refúgios LGM; (e) evidências de contatos secundários com base na distribuição parapátrica.

Embora diversos estudos tenham gerado dados genéticos que apoiam o modelo de CM (Carnaval et al. 2009; Batalha-Filho et al. 2010; Raposo Do Amaral et al. 2013), dados divergentes também foram relatados (Cabanne et al. 2007; Thomé et al. 2010). Assim, explicações que apresentam alternativas ao modelo CM foram propostas de forma concomitante (Thomé et al. 2010; Cardoso et al. 2015; Leite et al. 2016). Em particular, um modelo proposto por Leite et al. (2016) que prevê a expansão da MA no pico do LGM ao contrário do modelo CM, foi apontada como alternativa ao modelo CM, e conseqüentemente, à hipótese dos refúgios florestais (HRF, sensu Haffer, 1969).

Com base nos ENM de pequenos marsupiais, alterações do nível do mar no Quaternário e as simulações de modelos coalescentes para o DNAmT, a hipótese da expansão da Mata Atlântica (AFH) sugere que o domínio da Mata Atlântica experimentou um período de expansão no LGM, em vez de retração ou fragmentação (para detalhes, veja Leite et al., 2016). Assim, as espécies e populações influenciadas pela AFH devem apresentar distribuição expandida em ENMs no LGM, bem como picos de crescimento populacional em torno de 21 Ka a. p.

Baseando-se nessas duas hipóteses, foram construídos modelos de nicho ecológicos que indicassem a possível distribuição geográfica de *E. iopoecila*, bem como a flutuação do tamanho efetivo de suas populações ao longo dos últimos 130 mil anos, utilizando o ENM e o EBSP, respectivamente. Os resultados obtidos a partir destas duas análises satisfazem várias previsões do modelo CM. A modelagem da distribuição de espécies mostra que as populações do sul da distribuição de *E. iopoecila* passaram por vários processos de extinção e recolonização. Ao mesmo tempo, esta análise mostra áreas de estabilidade ao longo dos últimos 130 mil de anos, congruentes com os refúgios da Bahia e Pernambuco previstos por Carnaval e Moritz (2008). Além disso, o ENM para o Holoceno médio (6 Ka a. p.) sugere uma expansão da distribuição de *E. iopoecila* ainda maior que a do ENM para o tempo atual. Finalmente, o modelo LIG (130 kpb) é bastante semelhante ao ENM atual, o que pode indicar um processo relativamente cíclico.

Do mesmo modo, as análises demográficas dos clados Sul e Central-Norte suportam integralmente o modelo CM. Enquanto as redes de haplótipos e testes de neutralidade apontam que o *cluster* Sul passou por uma recente expansão populacional, os resultados do EBSP não apenas confirmam, mas também datam o início da expansão após o LGM (21 Ka a. p.). Congruentemente, as mesmas análises mostraram evidências de estabilidade nas populações dos *clusters* Central e Norte. Ainda em suporte ao modelo CM, os índices de diversidade genética (π), não só são maiores nas zonas de refúgio, mas também são estatisticamente significativos em relação aos índices encontrados nas populações do sul, mesmo quando os genes mitocondriais foram analisados individualmente, o que indica que algum evento interferiu na diversidade genética destas últimas populações. Por fim, o modelo CM ainda encontra suporte por evidências genéticas e morfológicas de uma zona híbrida entre os grupos sul (SP5) e central (RJ), sugerindo contato secundário com base na distribuição parapátrica.

Embora essa hipótese não possa ser diretamente testada com estes dados, o modelo de distribuição para o LGM sugere duas áreas estáveis nas quais populações ancestrais de *E. iopoecila* no *cluster* Sul poderiam ter permanecido abrigadas. Áreas no interior do continente aparecem com considerável adequabilidade entre 21 e 6 Ka a. p. nos modelos de *E. iopoecila*. Contudo, uma segunda com relativa adequabilidade, mais próxima da porção central da MA, parece ser um candidato à área ancestral igualmente bom. Neste tocante, vale ressaltar que até agora, *E. iopoecila* foi registrada apenas em regiões costeiras (floresta ombrófila densa). Assim, é mais plausível que as populações ancestrais de *E. iopoecila* do Sul tenham permanecido nas áreas mais ao norte, ao invés das regiões mais interioranas da floresta. A

hipótese de recente extinção e recolonização local para as populações do Sul é suportada tanto pelos ENM quanto por sinais de crescimento populacional encontrados no presente estudo, bem como a baixa estrutura de DNAnu encontrada por Penha et al. (2015). Outro ponto importante que se enquadra na recente colonização do Sul para *E. iopoecila* na MA é a falta de estruturas internas nesta região, bem como entre ilhas e continentes para as populações do sul. Caso as populações tivessem permanecido no Sul em todo o LGM, esperar-se-ia que as populações insulares tivessem gradualmente acumulado diferenças nas frequências alélicas, mesmo que apenas por deriva. No entanto, com poucas exceções, comparações entre populações continentais e insulares de abelhas das orquídeas no sul da MA revelam apenas níveis baixos de estrutura genética ($F_{ST} \leq 0,05$) no DNAnu (Rocha Filho et al. 2013; Boff et al. 2014; Penha et al. 2015) e principalmente no DNAm (Penha et al. 2015).

Com base em todos esses dados, acreditamos que a diferenciação observada nas três linhagens genéticas de *E. iopoecila* é o resultado de desconexões passadas e múltiplos eventos de extinção/recolonização ao longo de diferentes regiões da MA devido a flutuações climáticas do Pleistoceno (*sensu* Carnaval e Moritz, 2008). Recentemente, López-Urbe et al. (2014) apontaram as instabilidades climáticas do Quaternário como um importante mecanismo que impulsionou a diversificação de linhagens mitocondriais de abelhas das orquídeas. Os resultados aqui apresentados apontam para a mesma conclusão, dada a importante quebra filogeográfica observada para *E. iopoecila*.

Contudo, deve-se ter em mente que as análises neste estudo são baseadas exclusivamente em dados de DNAm, e devem ser avaliados com cautela. Por outro lado, dados de DNAm fazem sentido, tanto espacial quanto taxonomicamente na maioria dos casos estudados e, portanto, os padrões de diferenciação de DNAm bem suportados geograficamente devem ser suficientes para fornecer uma forte inferência da história evolutiva dos táxons (Zink e Barrowclough 2008; Avise 2009). Além disso, como observado por Avise (2009), o DNAm sozinho pode fornecer *insights* biológicos fundamentais, pois reflete a demografia matrilinear, bem como a estrutura geográfica e suas influências sobre a aspectos reprodutivos. Isto é particularmente verdadeiro para as abelhas das orquídeas, nas quais as fêmeas parecem exibir um forte comportamento filopátrico (Augusto e Garófalo 2004), enquanto os machos viajam em grandes distâncias (Pokorny et al. 2015) e podem se acasalar praticamente ao acaso com fêmeas de diferentes locais. Vários estudos realizados no DNAnu de diferentes espécies de abelhas das orquídeas corroboraram esta afirmação, relatando níveis relativamente baixos de estrutura genética (Freiria et al. 2012; Rocha Filho et

al. 2013; López-Uribe et al. 2014; Penha et al. 2015), enquanto os estudos de DNAm revelaram estrutura genética profunda (López-Uribe et al. 2014; Penha et al. 2015).

O presente estudo contribui com dois aspectos relevantes para essa questão específica. Primeiro, os dados de *E. iopoecila* indicam uma correlação forte e significativa entre as distâncias geográficas e genéticas (F_{ST} par-a-par), o que é consistente com o comportamento filopátrico das fêmeas (Augusto e Garófalo 2004). Em segundo lugar, os resultados dos modelos de nicho ecológico para o Holoceno médio mostram uma ampla conectividade para populações de *E. iopoecila* ao longo da costa da MA em um passado recente (6 Ka a. p.). Este cenário facilitaria a dispersão dos machos e, portanto, níveis mais altos de fluxo gênico de DNAnu, resultando na atual falta de estrutura genética neste genoma entre os diferentes *clusters*, como descrito por Penha et al. (2015). Por outro lado, embora esses autores não tenham encontrado níveis elevados de estrutura de DNAnu ($F_{ST} \leq 0,05$) em *E. iopoecila*, deve-se enfatizar que as amostras analisadas pertenciam, em sua maioria, a um único *cluster*, com apenas uma localidade aqui identificada como pertencente ao *cluster* Central, exatamente na zona de contato entre os *clusters* adjacentes. Além disso, Penha et al. (2015) observaram uma correlação significativa entre distâncias genéticas e geográficas para ambos os tipos de marcadores moleculares (DNAm e DNAnu) pesquisados em seu estudo. Assim, a falta de estrutura de DNAnu em *E. iopoecila* em toda a sua distribuição não pode ser assumida sem a confirmação por estudos mais detalhados nesta área.

Outro ponto interessante é que a estrutura aqui encontrada para o DNAm é altamente congruente com a variação dos fenótipos de coloração tegumentar apresentadas por *E. iopoecila* em toda a sua distribuição geográfica, azul na região do *cluster* Sul e verde nas regiões dos *clusters* Central e Norte. Como já foi documentado por Faria e Melo (2007), outras espécies de *Euglossa* além de *E. iopoecila*, como *E. stellfeldi* e *E. townsendi*, apresentam variação latitudinal em sua coloração de tegumento ao longo da região costeira da MA. Tal como acontece em *E. iopoecila*, os indivíduos azuis dessas duas espécies ocorrem na mesma região (sul MA), onde as populações de *E. iopoecila* apresentam tegumento violeta azulado (Ferrari e Melo 2014). Sugeriu-se que essa tendência a ter um intenso integumento azul em direção às regiões mais a sul da MA poderia estar associada à pressão seletiva favorecendo fenótipos de cor azul em abelhas do gênero *Euglossa* que ocorrem nesta região (Faria Jr e Melo, 2007; Ferrari e Melo, 2014). Embora esta hipótese seja plausível e encontre apoio na literatura recente (Cabanne et al. 2014), nossos dados genéticos e resultados ENM também apoiam fortemente a ideia de que os eventos vicariantes poderiam ter desempenhado um papel importante nas diferenças fenotípicas agora exibidas por *E. iopoecila* em sua

distribuição no domínio de Mata Atlântica. De acordo com estes achados, é perfeitamente possível que os eventos vicariantes que ocorreram entre 130 mil e 21 Ka a. p. afetaram a seleção divergente em populações desconectadas de *E. iopoecila*.

Embora seja ainda tentador atribuir os atuais padrões de cores de *E. iopoecila* (e seus congêneres do Sul) a um processo de especiação incompleto, particularmente devido à forte quebra filogeográfica detectada entre os clados Sul e Central-Norte, essa ideia deve ser considerada com cautela. O presente experimento não foi concebido para testar especificamente as relações moleculares entre as cores tegumentares nestes euglossíneos, ou mesmo determinar os mecanismos moleculares pelos quais essas diferenças são expressas. Mais do que isso, os resultados aqui apresentados são provenientes apenas da análise de locos mitocondriais, e mesmo que haja uma nítida congruência entre os padrões espaciais morfológico e moleculares, é improvável que a expressão da coloração tegumentar esteja ligada ao DNAm.

Ainda que eventos vicariantes tenham sido amplamente reconhecidos como importantes fatores geradores de diversidade genética e fenotípica no domínio de MA (Carnaval et al. 2009; Batalha-Filho et al. 2009; Cabanne et al. 2014), achados divergentes destes primeiros têm sido relatados em outros trabalhos (Cardoso et al. 2015). Vale ainda ressaltar que, de acordo com Cabanne et al. (2014), a heterogeneidade da MA também pode ter desempenhado um papel significativo na evolução das espécies ao longo deste bioma, talvez por meio de seleção divergente entre regiões.

Assim, em termos gerais, os resultados aqui apresentados sugerem que a estrutura genética observada para *E. iopoecila* tenha sido moldada por eventos vicariantes, o que levou à divergência genética encontrada nos três diferentes grupos dessas espécies de abelhas das orquídeas. No entanto, não se pode descartar a possibilidade de que outros fatores, como pressões seletivas diferenciais, também tenham agido em conjunto com estes eventos para moldar a diversidade genética e fenotípica apresentada pelas populações de *E. iopoecila* na MA. Além de elucidar a estrutura genética, a filogeografia e a demografia histórica de *E. iopoecila* em toda a MA, os resultados aqui apresentados representam uma nova e importante contribuição à compreensão dos padrões filogeográficos de outras espécies de Euglossini, também endêmicas deste bioma.

Finalmente, para fins de conservação, considerando os resultados do presente estudo e aqueles relatados por Penha et al. (2015), ressalta-se o efeito significativo do comportamento filopátrico das fêmeas sobre a distribuição da diversidade genética nas populações de *E. iopoecila*. Apesar do importante papel dos machos na manutenção do fluxo gênico nuclear em

longas distâncias (López-Uribe et al. 2014; Penha et al., 2015), o mesmo não acontece com o genoma mitocondrial, provavelmente devido à filopatria das fêmeas. Neste contexto, é importante ressaltar a necessidade de maximizar a retenção de diversidade genética nas populações (Frankham 2010). Deste modo, visando otimizar as medidas de conservação para as populações de *E. iopoecila* ao longo do ameaçado domínio de Mata Atlântica, propõe-se que quatro principais fatores sejam levados em consideração. Primeiro, os maiores níveis de diversidade genética encontrados nas populações desta espécie no Corredor de Biodiversidade Central da MA (*cluster* Central). Em segundo lugar, a estabilidade climática histórica no "refúgio da Bahia" (*sensu* Carnaval e Moritz, 2008) sobre o *cluster* central. Em terceiro, o padrão de isolamento por distância como importante fator na estruturação da diversidade genética de *E. iopoecila*. Finalmente, o pobre estado de conservação da Mata Atlântica brasileira. Com base em todos esses fatores, recomenda-se como orientação prioritária que os esforços de conservação de *E. iopoecila* devem ser focados em medidas para melhorar a conectividade em áreas geográficas mais amplas e, em particular, aumentar as conexões entre os três corredores da Biodiversidade da Mata Atlântica (CBSM, Corredor Central e Nordeste).

REFERÊNCIAS

- Augusto SC, Garófalo CA (2004) Nesting biology and social structure of *Euglossa* (*Euglossa*) *townsendi* Cockerell (Hymenoptera, Apidae, Euglossini). *Insectes Soc* 51:400–409. doi: 10.1007/s00040-004-0760-2
- Awise JC (2009) Phylogeography: Retrospect and prospect. *J Biogeogr* 36:3–15. doi: 10.1111/j.1365-2699.2008.02032.x
- Batalha-Filho H, Melo GAR, Waldschmidt AM, et al (2009) Geographic distribution and spatial differentiation in the color pattern of abdominal stripes of the Neotropical stingless bee *Melipona quadrifasciata* (Hymenoptera: Apidae). *Zool* 26:213–219. doi: 10.1590/S1984-46702009000200003
- Batalha-Filho H, Miyaki CY (2011) Filogeografia da Mata Atlântica. *Rev da Biol Especial*:31–34
- Batalha-Filho H, Miyaki CY (2016) Late Pleistocene divergence and postglacial expansion in the Brazilian Atlantic Forest: Multilocus phylogeography of *Rhopias gularis* (Aves: Passeriformes). *J Zool Syst Evol Res* 54:137–147. doi: 10.1111/jzs.12118
- Batalha-Filho H, Waldschmidt AM, Campos LAO, et al (2010) Phylogeography and historical demography of the neotropical stingless bee *Melipona quadrifasciata* (Hymenoptera, Apidae): incongruence between morphology and mitochondrial DNA.

- Apidologie 41:534–547. doi: 10.1051/apido/2010001
- Boff S, Soro A, Paxton RJ, Alves-dos-Santos I (2014) Island isolation reduces genetic diversity and connectivity but does not significantly elevate diploid male production in a neotropical orchid bee. *Conserv Genet* 15:1123–1135. doi: 10.1007/s10592-014-0605-0
- Cabanne GS, Santos FR, Miyaki CY (2007) Phylogeography of *Xiphorhynchus fuscus* (Passeriformes, Dendrocolaptidae): Vicariance and recent demographic expansion in southern Atlantic forest. *Biol J Linn Soc* 91:73–84. doi: 10.1111/j.1095-8312.2007.00775.x
- Cabanne GS, Trujillo-Arias N, Calderón L, et al (2014) Phenotypic evolution of an Atlantic Forest passerine (*Xiphorhynchus fuscus*): Biogeographic and systematic implications. *Biol J Linn Soc* 113:1047–1066. doi: 10.1111/bij.12362
- Cardoso DC, Cristiano MP, Tavares MG, et al (2015) Phylogeography of the sand dune ant *Mycetophylax simplex* along the Brazilian Atlantic Forest coast: Remarkably low mtDNA diversity and shallow population structure. *BMC Evol Biol* 15:106. doi: 10.1186/s12862-015-0383-4
- Carnaval AC, Hickerson MJ, Haddad CFB, et al (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* (80-) 323:785–789. doi: 10.1126/science.1166955
- Carnaval AC, Moritz C (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *J Biogeogr* 35:1187–1201. doi: 10.1111/j.1365-2699.2007.01870.x
- Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539. doi: 10.1186/1471-2105-9-539
- Corander J, Tang J (2007) Bayesian analysis of population structure based on linked molecular information. *Math Biosci* 205:19–31. doi: 10.1016/j.mbs.2006.09.015
- Dellicour S, Mardulyn P (2014) spads 1.0: A toolbox to perform spatial analyses on DNA sequence data sets. *Mol Ecol Resour* 14:647–651. doi: 10.1111/1755-0998.12200
- Dick CW, Roubik DW, Gruber KF, Bermingham E (2004) Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA phylogeography. *Mol Ecol* 13:3775–3785. doi: 10.1111/j.1365-294X.2004.02374.x
- Dressler RL (1982) Biology of Orchid Bees. *Annu Rev Ecol Syst* 13:373–394
- Drummond AJ, Bouckaert RR (2015) Bayesian evolutionary analysis with BEAST. Bayesian

- Evol Anal with BEAST 1–249. doi: 10.1017/CBO9781139095112
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214. doi: 10.1186/1471-2148-7-214
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* 11:2571–2581. doi: 10.1046/j.1365-294X.2002.01650.x
- Elith J, Phillips SJ, Hastie T, et al (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57. doi: 10.1111/j.1472-4642.2010.00725.x
- Excoffier L, Laval G, Schneider S (2005) Arlequin: an integrated software package for population genetics data analysis. *Evol Bioinforma* 1:47–50. doi: 10.1111/j.1755-0998.2010.02847.x
- Faria Jr LRR, Melo GAR (2007) Species of *Euglossa* (Glossura) in the Brazilian Atlantic forest, with taxonomic notes on *Euglossa stellfeldi* Moure (Hymenoptera, Apidae, Euglossina). *Rev Bras Entomol* 51:275–284. doi: 10.1590/S0085-56262007000300004
- Ferrari BR, Melo GAR (2014) Deceiving colors: Recognition of color morphs as separate species in orchid bees is not supported by molecular evidence. *Apidologie* 45:641–652. doi: 10.1007/s13592-014-0280-7
- Frankham R (2010) Where are we in conservation genetics and where do we need to go? *Conserv Genet* 11:661–663. doi: 10.1007/s10592-009-0010-2
- Freiria GA, Ruim JB, De Souza RF, Sofia SH (2012) Population structure and genetic diversity of the orchid bee *Eufriesea violacea* (Hymenoptera, Apidae, Euglossini) from Atlantic Forest remnants in southern and southeastern Brazil. *Apidologie* 43:392–402. doi: 10.1007/s13592-011-0104-y
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925. doi: genetics.org//147/2/915
- Galindo-Leal C, Câmara I de G (2003) *The Atlantic Forest of South America: Biodiversity Status, Threats, and Outlook*, 1st edn. Inland Press, Washington, DC
- Giangarelli DC, Aguiar WM de, Sofia SH (2015) Orchid bee (Hymenoptera: Apidae: Euglossini) assemblages from three different threatened phytophysiognomies of the subtropical Brazilian Atlantic Forest. *Apidologie* 46:71–83. doi: 10.1007/s13592-014-0303-4
- Hebert PDN, Penton EH, Burns JM, et al (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad*

- Sci 101:14812–14817. doi: 10.1073/pnas.0406166101
- Hijmans RJ, Cameron SE, Parra JL, et al (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978. doi: 10.1002/joc.1276
- Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Mol Ecol Resour* 11:423–434
- Lacey Knowles L, Alvarado-Serrano DF (2010) Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: Insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Mol Ecol* 19:3727–3745. doi: 10.1111/j.1365-294X.2010.04702.x
- Leigh JW, Bryant D (2015) POPART: Full-feature software for haplotype network construction. *Methods Ecol Evol* 6:1110–1116. doi: 10.1111/2041-210X.12410
- Leite YLR, Costa LP, Loss AC, et al (2016) Neotropical forest expansion during the last glacial period challenges refuge hypothesis. *Proc Natl Acad Sci* 113:1008–1013. doi: 10.1073/pnas.1513062113
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. doi: 10.1093/bioinformatics/btp187
- López-Uribe MM, Zamudio KR, Cardoso CF, Danforth BN (2014) Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees. *Mol Ecol* 23:1874–1890. doi: 10.1111/mec.12689
- Merow C, Smith MJ, Silander JA (2013) A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography (Cop)* 36:1058–1069. doi: 10.1111/j.1600-0587.2013.07872.x
- Moure J, Melo GAR, Faria Jr LRR (2012) Catalogue of bees (Hymenoptera, Apoidea) in the Neotropical Region-online version. In: Euglossini Latreilli 1802. <http://www.moure.cria.org.br/catalogue>. Accessed 14 Apr 2016
- Nemésio A (2009) Orchid bees (Hymenoptera: Apidae) of the Brazilian Atlantic forest. *Zootaxa* 2041:1–242
- Penha RES, Gaglianone MC, Almeida FS, et al (2015) Mitochondrial DNA of *Euglossa iopoecila* (Apidae, Euglossini) reveals two distinct lineages for this orchid bee species endemic to the Atlantic Forest. *Apidologie* 46:346–358. doi: 10.1007/s13592-014-0329-7
- Phillips S, Anderson RP, Schapire R (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259. doi: 10.1016/j.ecolmodel.2005.03.026

- Pokorny T, Loose D, Dyker G, et al (2015) Dispersal ability of male orchid bees and direct evidence for long-range flights. *Apidologie* 46:224–237. doi: 10.1007/s13592-014-0317-y
- Posada D (2008) jModelTest: Phylogenetic model averaging. *Mol Biol Evol* 25:1253–1256. doi: 10.1093/molbev/msn083
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst Biol* 53:793–808. doi: 10.1080/10635150490522304
- Ramalho AV, Gaglianone MC, Oliveira ML de (2009) Comunidades de abelhas Euglossina (Hymenoptera, Apidae) em fragmentos de Mata Atlântica no Sudeste do Brasil. *Rev Bras Entomol* 53:95–101. doi: 10.1590/S0085-56262009000100022
- Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, et al (2008) Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics* 179:555–567. doi: 10.1534/genetics.107.083006
- Ramirez S, Dressler RL, Ospina M (2002) Abejas euglossinas (Hymenoptera: Apidae) de la Región Neotropical: Listado de especies con notas sobre su biología. *Biota Colomb* 3:7–118
- Ramírez SR, Roubik DW, Skov CE, Pierce NE (2010) Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). *Biol J Linn Soc* 100:552–572. doi: 10.1111/j.1095-8312.2010.01440.x
- Ramos-Onsins SE, Rozas J (2006) Erratum: Statistical properties of new neutrality tests against population growth (*Molecular Biology and Evolution* (2002) 19 (2092-2100)). *Mol Biol Evol* 23:1642. doi: 10.1093/molbev/msl052
- Raposo Do Amaral F, Albers PK, Edwards S V., Miyaki CY (2013) Multilocus tests of Pleistocene refugia and ancient divergence in a pair of Atlantic Forest antbirds (*Myrmeciza*). *Mol Ecol* 22:3996–4013. doi: 10.1111/mec.12361
- Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *J Hered* 86:248–249. doi: 10.1093/oxfordjournals.jhered.a111573
- Ribeiro RA, Lemos-Filho JP, Ramos ACS, Lovato MB (2011) Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): Insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. *Heredity (Edinb)* 106:46–57. doi: 10.1038/hdy.2010.64
- Rocha Filho LC, de Campos Muradas Cerântola N, Garófalo CA, et al (2013) Genetic

- differentiation of the Euglossini (Hymenoptera, Apidae) populations on a mainland coastal plain and an island in southeastern Brazil. *Genetica* 141:65–74. doi: 10.1007/s10709-013-9706-9
- Roubik DW, Hanson P (2004) Abejas de orquídeas de la América tropical. *Biología y guía de campo (Orchid bees of Tropical América) Biology and field guide*, 1st edn. INBio Press, Heredia, Costa Rica
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595. doi: PMC1203831
- Tamura K, Stecher G, Peterson D, et al (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. doi: 10.1093/molbev/mst197
- Thomé MTC, Zamudio KR, Giovanelli JGR, et al (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Mol Phylogenet Evol* 55:1018–1031. doi: 10.1016/j.ympev.2010.02.003
- Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Mol Ecol* 17:2107–2121. doi: 10.1111/j.1365-294X.2008.03737.x

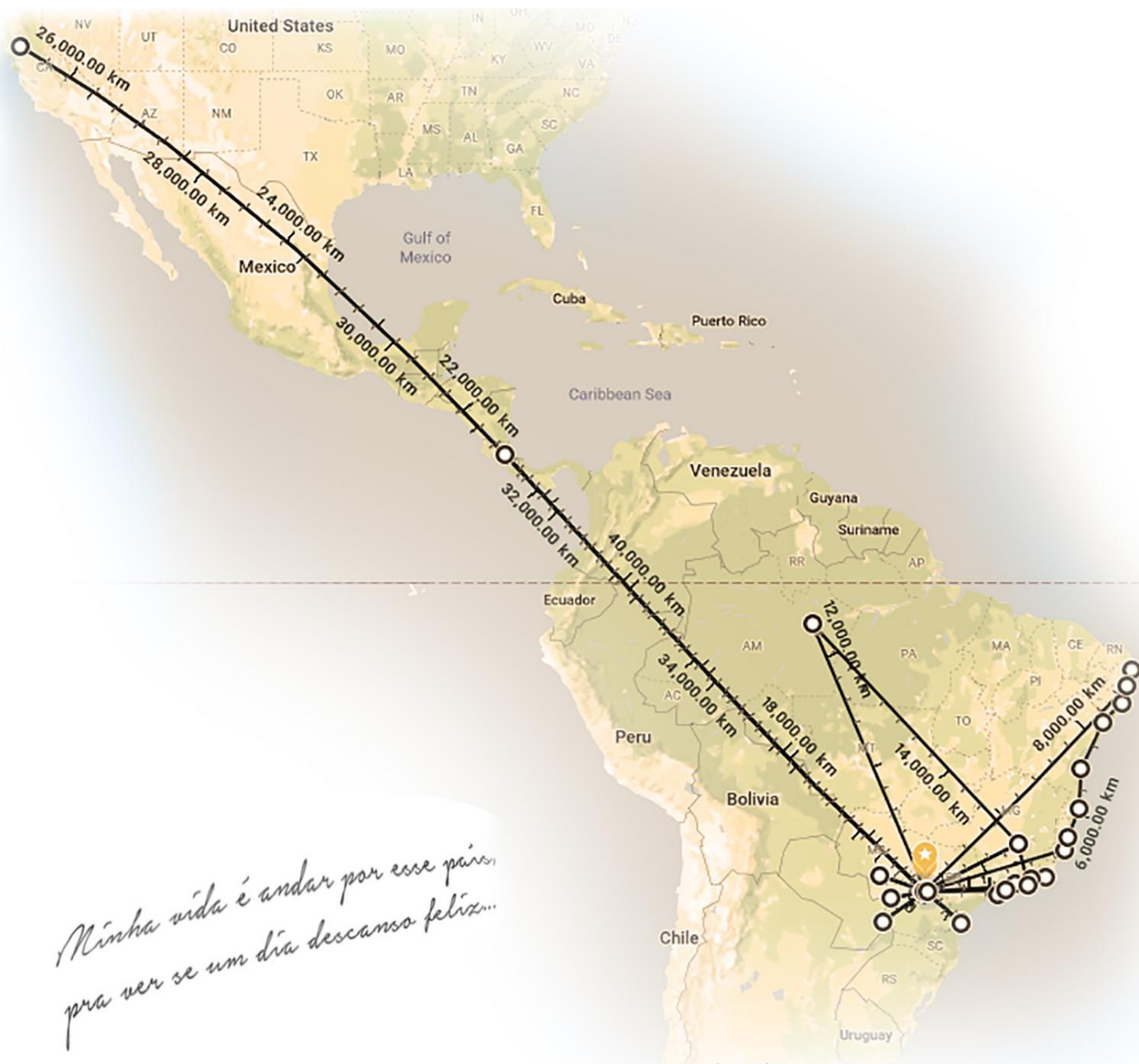
7 CONSIDERAÇÕES FINAIS

Os resultados obtidos no presente estudo apontam para um papel central das variações climáticas do Pleistoceno sobre os padrões de estruturação genética das espécies de abelhas da tribo Euglossini na Mata Atlântica. Embora as espécies da América Central e Amazônia pareçam ter reagido de maneira distinta a esses ciclos glaciais, ao menos uma espécie do grupo *crassipunctata* aponta para o possível efeito das glaciações sobre a diversificação de clados atlânticos em relação aos amazônicos.

Padrões de distribuição latitudinal e longitudinal parecem ter sido afetados de maneira distinta por esses eventos climáticos. Espécies como *Euglossa iopocila* e *Euglossa clausi*, associadas a Floresta Ombrófila Densa demonstram características mais alinhadas às expectativas dos modelos climáticos descritos para a Mata Atlântica, como o modelo de Carnaval e Moritz (2008). Por outro lado, espécies com distribuição longitudinal, e portanto, mais fortemente associadas à Floresta Estacional Semidecidual, como no caso de *Euglossa annectans*, parecem tender à refúgios em direção ao interior e não estritamente à região norte. Adicionalmente, modelos de distribuição paleoclimáticos, demonstram áreas estáveis para euglossíneos na porção sul da Mata Atlântica, corroborando a ideia de processos idiossincráticos na diversificação de organismos e grupos na MA.

Do ponto de vista evolutivo, as análises do grupo *crassipunctata* apontam para uma complexa combinação de eventos geológicos e climáticos moldando a diversidade biológica de abelhas das orquídeas. Com base nos dados obtidos para este grupo de espécie, a elevação dos Andes aparece como principal fator na diversificação das espécies amazônicas em relação as da América Central, enquanto as variações climáticas do Pleistoceno começam a influenciar mais incisivamente na diversificação dos clados atlânticos em relação aos amazônicos. A combinação das relações filogenéticas entre as populações de *E. clausi* aliado às zonas de estabilidade climáticas e dados da literatura, indicam que um corredor mais recente formado entre a porção média da Mata Atlântica e Amazônia ao longo do Pleistoceno foi o caminho mais provável de colonização para o ancestral de *E. clausi* atingir o domínio de Mata Atlântica.

Deste modo, a combinação destas diferentes análises indica que a história evolutiva das abelhas das orquídeas nas Américas, mais especificamente na Mata Atlântica, é fruto da combinação destes fatores geológicos e climáticos, e portanto, os modelos observados no presente estudo devem ser confrontados com informações provenientes de outras espécies de Euglossini para confirmar estes padrões para as abelhas das orquídeas.



APÊNDICES

LISTA DE APÊNDICES

- 1
- 2 Apêndice A – Licença de tradução e publicação do Capítulo 3: *Phylogeography and historical*
 3 *demography of the orchid bee Euglossa iopoecila: signs of vicariant events associated to*
 4 *Quaternary Climate Changes*.
- 5 Apêndice B - Fundamentação Teórica: Trabalhos relacionados à genética de populações
 6 realizados com abelhas das orquídeas, marcadores empregados, área do conhecimento e
 7 principais conclusões.
- 8 Apêndice C - Amplitude de variação para as 19 variáveis climáticas *WorldClim* em escala de
 9 2.5 arcmin observadas pelo pacto “dismo” do *software R*.
- 10 Apêndice D - Registros de espécies utilizados na produção dos modelos de nicho ecológico
 11 para as espécies do grupo *crassipunctata*. Os táxons seguem classificação sugerida por
 12 Nemésio e Engel (2012) de acordo com a região geográfica do registro original. Referências
 13 incluem literatura e plataforma SpeciesLink e Catalogue of Life.
- 14 Apêndice E - Registros considerados de *Euglossa iopoecila* Dressler (1982) no capítulo 3.
- 15 Apêndice I - Percentual de contribuição de 19 variáveis climáticas em duas dimensões
 16 exploradas via análise de componentes principais baseada na ocorrência de espécies do grupo
 17 *crassipunctata*.
- 18 Apêndice J - Mapa de Fatores para 19 variáveis climáticas exploradas via análise de
 19 componentes principais baseada na ocorrência de *Euglossa annectans*
- 20 Apêndice K - Registros de espécies utilizados na produção dos modelos de nicho ecológico
 21 para *Euglossa annectans*. Referências incluem literatura e plataforma SpeciesLink e
 22 Catalogue of Life.
- 23 Apêndice L - Localidades (nome da reserva, município e estado), código da localidade,
 24 coordenadas genética e tamanho de amostral (*n*) de *Euglossa iopoecila* colizada ao longo da
 25 Mata Atlântica Brasileira.
- 26 Apêndice M - Lista de ocorrências utilizadas para a construção dos modelos de nicho
 27 ecológico para *Euglossa iopoecila*.
- 28 Apêndice N – Distribuição das diferenças entre pares de sequências de 640 pb do gene *cytb*
 29 em 120 amostras de *Euglossa annectans*. *Haplótipos fixados, gráfico ilustrativo.
- 30

- 1 Apêndice A – Licença de tradução e publicação do Capítulo 3: *Phylogeography and historical*
- 2 *demography of the orchid bee Euglossa iopocila: signs of vicariant events associated to*
- 3 *Quaternary Climate Changes.*

4/18/2018

RightsLink Printable License

**SPRINGER NATURE LICENSE
TERMS AND CONDITIONS**

Apr 18, 2018

This Agreement between Ms. Wilson Frantine ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4332131469012
License date	Apr 18, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Conservation Genetics
Licensed Content Title	Phylogeography and historical demography of the orchid bee <i>Euglossa iopocila</i> : signs of vicariant events associated to Quaternary climatic changes
Licensed Content Author	Wilson Frantine-Silva, Douglas C. Giangarelli, Rafael E. S. Penha et al
Licensed Content Date	Jan 1, 2016
Licensed Content Volume	18
Licensed Content Issue	3
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	yes, without original language
Number of languages	1
Circulation/distribution	<501
Author of this Springer Nature content	yes
Title	Abelhas das Orquídeas: Filogeografia e Diversidade Genética
Instructor name	Dr. Silvia Helena Sofia
Institution name	Universidade Estadual de Londrina
Expected presentation date	May 2018
Order reference number	1
Specific Languages	Portuguese
Requestor Location	Ms. Wilson Frantine-Silva R. Antonio Amadeu 161 Jd. Monte Belo Londrina, Paraná 86041650 Brazil Attn: Ms. Wilson Frantine-Silva
Billing Type	Invoice
Billing Address	Ms. Wilson Frantine-Silva R. Antonio Amadeu 161

<https://s100.copyright.com/CustomAdmin/PLF.jsp?ref=03cfc9ea-0532-4f9b-9430-8643e3b1f1e2>

1/3

- 1 Apêndice B - Fundamentação Teórica: Trabalhos relacionados à genética de populações realizados com abelhas das orquídeas, marcadores
2 empregados, área do conhecimento e principais conclusões.

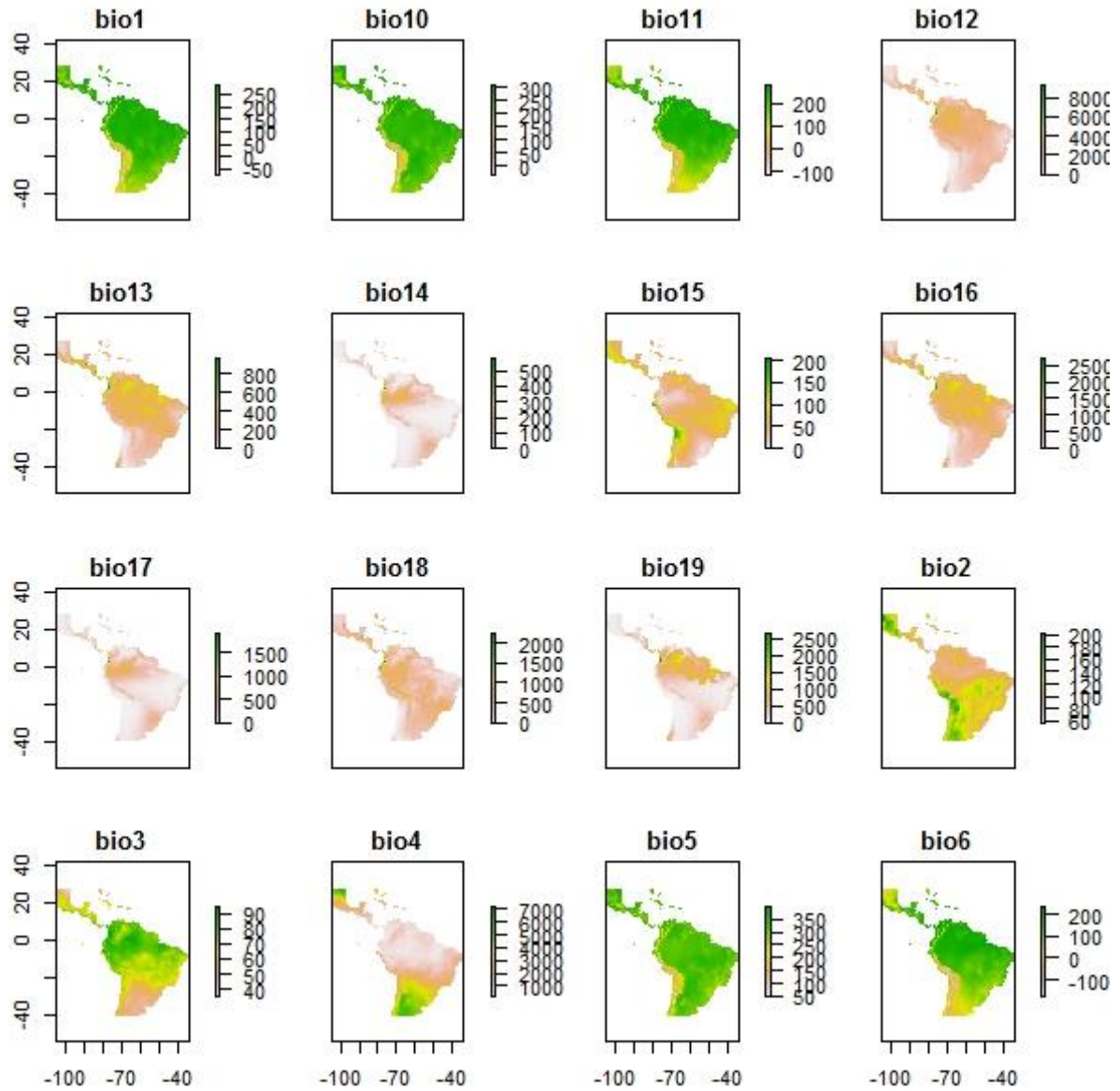
Título	Autores	Ano	Jornal	Localidade	Marcadores	Área relacionada	Principais conclusões
Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA Phylogeography	Dick et al.	2004	Molecular Ecology	América Central e Andes	<i>COI</i>	Filogeografia	Ampla fluxo gênico na região dos Andes
Population genetics, diploid males, and limits to social evolution of euglossine bees	Roubik et al.	2004	Evolution	Costa Rica	isoenzimas	Genética de Populações	Alta diversidade genética pode ser fator inibidor da eussocialidade
Genetic structure analysis of <i>Eufriesea violacea</i> (Hymenoptera, Apidae) populations from southern Brazilian Atlantic rainforest remnants	Sofia et al.	2005	Genetics and Molecular Biology	Brasil; PR	RAPD	Genética de Populações	Ausência de estruturação entre os pontos
Microsatellite loci for euglossine bees (Hymenoptera: Apidae)	Souza et al.	2007	Molecular Ecology Resources	Brasil	microsatélites	primer description	
Diploid Male Frequencies in Colombian Populations of Euglossine Bees	Lopez-Uribe et al.	2007	Biotropica	Colômbia	isoenzimas	Ocorrência de Machos Diploides	Autores sugerem susceptibilidade diferencial para cada espécie à produção de machos diploides
An Olfactory Shift Is Associated with Male Perfume Differentiation and Species Divergence in Orchid Bees	Eltz et al.	2008	Current Biology	Península de Yucatan	microsatélites	Genética de Populações	Microsatélites apresentam diferenças nas frequências alélicas entre machos com dois e três dentes
Microsatellite loci for <i>Euglossa annectans</i> (Hymenoptera: Apidae) and their variability in other orchid bees	Paxton et al.	2009	Molecular Ecology Resources	Brasil, SC	microsatélites	primer description	

Titulo	Autores	Ano	Jornal	Localidade	Marcadores	Área relacionada	Principais conclusões
Mitochondrial DNA Diversity of Orchid Bee <i>Euglossa fimbriata</i> (Hymenoptera: Apidae) Populations Assessed by PCR-RFLP	Suzuki et al.	2010	Biochemical Genetics	Brasil; PR	PCR-RFLP	Genética de Populações	
Conservation genetics of neotropical pollinators revisited: microsatellite analysis suggests that diploid males are rare in orchid bees	Souza et al.	2010	Evolution	Brasil and Colombia	microsatélites	Ocorrência de Machos Diploides	Machos diploides são extremamente raros
Genetic differentiation of urban populations of <i>Euglossa cordata</i> from the state of São Paulo, Brazil	Cerântula et al.	2011	Apidologie	Brasil; SP	microsatélites	Genética de Populações	Ausência de estruturação, machos diploides raros, alto Ne, dados corroboram alta dispersão
Characterization of the orchid bee <i>Euglossa viridissima</i> (Apidae: Euglossini) and a novel cryptic sibling species, by morphological, chemical, and genetic characters	Eltz et al.	2011	Zoological Journal of Linnean Society	Central America	<i>COI</i> e microsatélites	taxonomy	Sem diferenciação entre ilhas e continente
Population genetic structure of orchid bees (Euglossini) in anthropogenically altered landscapes	Zimmerman et al.	2011	Conservation Genetics	Yucatan peninsule and Florida	microsatélites	Genética de Populações	Estrutura genética entre grupos de populações distantes, mas não entre populações do mesmo grupo
solution and cross-species characterization of polymorphic microsatellites for the orchid bee <i>Eulaema meriana</i>	Lopez-Uribe et al.	2011	Conservation Genetics Resources	Costa Rica, La Selva	microsatellites	Descrição de Primers	4 a 9 alelos por loco; oito e nove locos funcionaram em sete outras especies de <i>Eulaema</i>
Population genetics of orchid bees in a fragmented tropical landscape	Suni; Brosi	2012	Conservation Genetics	Costa Rica	microsatélites	Genética de Populações	
Population structure and genetic diversity of the orchid bee <i>Eufriesea violacea</i> (Hymenoptera, Apidae, Euglossini) from Atlantic Forest remnants in southern and southeastern Brazil	Freiria et al.	2012	Apidologie	Brasil; SP, PR, SC, RS	microsatélites	Genética de Populações	Restrição de fluxo gênico entre os sítios

Titulo	Autores	Ano	Jornal	Localidade	Marcadores	Área relacionada	Principais conclusões
Genetic differentiation of the Euglossini (Hymenoptera, Apidae) populations on a mainland coastal plain and an island in southeastern Brazil	Rocha-Filho et al.	2013	Genetica	Brasil; SP, Picinguaba e Ilhabela	microsatélites	Genética de Populações	Sem estruturação entre ilhas e continente
Island isolation reduces genetic diversity and connectivity but does not significantly elevate diploid male production in a neotropical orchid bee	Boff et al.	2014	Conservation Genetics	Brasil; SP	microsatélites	Genética de Populações	Redução de diversidade nas ilhas; baixo índice de machos diploides
Spatio-temporal Genetic Structure of a Tropical Bee Species Suggests High Dispersal Over a Fragmented Landscape	Suni et al.	2014	Biotropica	Costa Rica; Southern	microsatélites	Genética de Populações	Alto nível de dispersão deve explicar os resultados
Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees	Lopez-Uribe et al.	2014	Molecular Ecology	All Americas	MtDNA, CAD, microsatélites	Filogeografia	Modelagem de Nicho e resultados genéticos alinhados; Mudanças climáticas do Quaternário moldaram a distribuição
Mitochondrial DNA of <i>Euglossa iopoeila</i> (Apidae, Euglossini) reveals two distinct lineages for this orchid bee species endemic to the Atlantic Forest	Penha et al.	2015	Apidologie	Brasil; RJ, SP e PR	<i>cytb</i> e microsatélites	Genética de Populações	Pequena, porém significativa estruturação para os microsatélites, alta estruturação para o DNAm; hibridização entre morfotipos de cor
Orchid bees: a new assessment on the rarity of diploid males in populations of this group of Neotropical pollinators	Sofia et al.	2015	Apidologie	Brasil, AF	microsatélites	Ocorrência de Machos Diploides	Machos diploides são mais ou menos frequentes a depender de cada espécie
Functional connectivity and genetic diversity of <i>Eulaema atleticana</i> (Apidae, Euglossina) in the Brazilian Atlantic Forest Corridor: assessment of gene flow	Rosa et al.	2016	Biotropica	Brasil; BA, Atlantic Forest Corridor	<i>COI</i> , <i>cytb</i> , microsatélites		Fluxo gênico mantém a variabilidade

Título	Autores	Ano	Jornal	Localidade	Marcadores	Área relacionada	Principais conclusões
The population genetics of two orchid bees suggests high dispersal, low diploid male production and only an effect of island isolation in lowering genetic diversity	Soro et al.	2017	Conservation Genetics	Yucatan peninsule	microsatélites	Genética de Populações	Redução de diversidade na ilha; baixo índice de machos diploides; machos diploides não são boa medida de para declínio das abelhas
Dispersal of the orchid bee <i>Euglossa imperialis</i> over degraded habitat and intact forest	Suni	2017	Conservation Genetics	Costa Rica	microsatélites	Genética de Populações	baixa diversidade genética; Isolamento por distância; inclusão de irmão completos nas análises traz pouca alteração aos resultados
Phylogeography and historical demography of the orchid bee <i>Euglossa iopocila</i> : signs of vicariant events associated to Quaternary climatic changes	Frantine-Silva et al.	2017	Conservation Genetics	Brasilian Coast	<i>COI</i> , 16S, <i>cytb</i>	Filogeografia	Três estruturas mitocondriais; confirma unicidade taxonômica; instabilidade climática do Quaternário influencia a estruturação; Corredor Central da Bahia é região mais diversa

- 1 Apêndice C - Amplitude de variação para as 19 variáveis climáticas *WorldClim* em escala de
- 2 2.5 arcmin observadas pelo pacto “dismo” do *software* R.



3

1 Apêndice D - Registros de espécies utilizados na produção dos modelos de nicho ecológico para as espécies do grupo *crassipunctata*. Os táxons
 2 seguem classificação sugerida por Nemésio e Engel (2012) de acordo com a região geográfica do registro original. Referências incluem literatura
 3 e plataforma SpeciesLink e Catalogue of Life.

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	PB	REBIO Guaribas	-6,739066	-35,171611	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	MT	Mata do Buraquinho	-7,137364	-34,856676	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PB	João Pessoa, Mata do Buraquinho	-7,146366	-34,860276	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Goiana, Mata de Bujary	-7,598506	-35,000261	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Goiana, U. Sta Teresa, CAIG, M. Bujary e IPA	-7,611111	-34,969167	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Igarassu, Mata do Engenho Piedade, São José	-7,835556	-34,999444	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Cabo do Santo Agostinho, Res. Gurjaú, Compesa, Mata do Cuxio	-8,233056	-35,057778	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Cabo do Santo Agostinho, Compesa, Mata São Braz	-8,289133	-35,038325	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Cabo do Santo Agostinho, Compesa, Xangô	-8,289158	-35,03837	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Cabo do Santo Agostinho, Compesa, Mata Pau Santo	-8,289167	-35,038333	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Jaqueira, Usina Frei Caneca- Mata Quengo	-8,726931	-35,7933	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	PE	Jaqueira, Usina Frei Caneca- Mata Fervedouro	-8,726944	-35,793333	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	ES	Colônia Leopoldina, Usina Serra Grande	-8,995555556	-35,84055556	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	AL	Penedo, Estação Ecológica de Murici	-9,211694444	-35,87	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	AL	Penedo, Estação Ecológica de Murici	-10,337778	-36,500278	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	SE	Areia Branca, PARNA Serra de Itabaiana	-10,76160400	-37,3389060	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	BA	Conceição do Almeida	-12,77940000	-39,1700000	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	BA	Valença, Cajalbas	-13,44875800	-39,0978170	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	BA	Itabuna	-14,78647000	-39,272770	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	BA	Itabuna	-14,78648000	-39,272780	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	BA	Camacan, RPPN Serra Bonita	-15,39172000	-39,565109	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	BA	Porto Seguro, PARNA Monte Pascoal	-16,88025000	-39,41338889	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	BA	Itamaraju	-16,98939200	-39,570526	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	BA	Itamaraju, PARNA Descobrimento	-17,11238100	-39,347471	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	ES	Conceição da Barra, FLONA do Rio Preto	-18,34833333	-39,85191667	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	ES	Conceição da Barra, Floresta Nacional do Rio Preto	-18,422864	-39,830142	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	ES	Linhares, REBIO Sooretama	-19,04513889	-40,14738889	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	ES	Linhares	-19,130898	-40,010228	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	ES	Linhares	-19,416	-40,066	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	ES	Linhares, Sooretama	-19,41666	-40,06667	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	MG	Marileia, Parque Estadual do Rio Doce	-19,544362	-42,554889	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	MG	São Gonçalo do Rio Abaixo, ESEC Peti	-19,877131	-43,3716	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	ES	Vitoria, Morro do Moscoso	-20,312503	-40,342422	Nemésio 2009

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	ES	Presidente Kennedy	-21,074322	-41,088836	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	RJ	Santa Maria Madalena, Parque Estadual do Desengano	-21,958452	-41,957625	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	RJ	Macaé, Corrego do Ouro	-22,25	-41,78333	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	RJ	Casimiro de Abreu, REBIO Uniao	-22,429482	-42,04976	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	RJ	Casimiro de Abreu	-22,480833	-42,203889	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	RJ	Duke de Caixias, Pau Grande	-22,578951	-43,169325	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	RJ	Jacarepagua, Baia de Guanabara	-22,9428	-43,3581	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	RJ	Rio de Janeiro, PARNA Floresta da Tijuca	-22,955452	-43,280381	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	SP	São Luiz do Paraitinga, PESM, Núcleo Santa Virgínia	-23,335249	-45,144563	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	SP	Ubatuba, Praia da Fazenda, PE Serra do Mar	-23,36436111	-44,83247222	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	SP	Ubatuba, Picinguaba	-23,364401	-44,832501	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	SP	Ubatuba, Picinguaba	-23,365424	-44,822559	Este estudo
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	SP	Ubatuba, Picinguaba	-23,4339	-45,0711	Catalogue of Life
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	SP	São Sebastião, Ilha Bela	-23,503901	-45,222801	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	SP	Ubatuba, Ilha Anchieta	-23,5403	-45,071098	SpeciesLink
<i>Euglossa clausi</i>	<i>Euglossa crassipunctata</i>	Brasil	SP	Caraguatatuba, PESM, Caraguatatuba	-23,585378	-45,390629	Nemésio 2009
<i>Euglossa clausi</i>	<i>Euglossa clausi</i>	Brasil	BA	Barra da Úna	-23,750008	-45,75272	Este estudo

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa clausi</i>	<i>Euglossa sapphirina</i>	Brasil	SP	São Sebastião, Parque Estadual de Ilhabela	-23,848231	-45,34903	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Distrito turístico, cultural e histórico de Santa Marta	11,121404	-73,885137	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Bijagua, 20 km S Upala	10,7333	-85,1	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Finca Montezuma, 3 km SE Rio Naranjo	10,6641	-85,0972	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Chilamate	10,45	-84,0667	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Sarapiquí, Zona Prot La Selva, Est Biol La Selva. Finca La Selva, 3 Km S de Sarapiquí	10,439193	-84,010919	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		La Selva, CEN 50	10,429414	-84,007241	Este estudo
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		La Selva Reserve Station	10,4166	-84	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Cañas, 14 km S, EJM	10,3	-85,0956	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Guapiles, 5 mi W	10,2166	-83,8568	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Cahuita, 5 mi S	9,67722	-82,8167	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		R.V.S. Gandoca Manzanillo. Centro Operativo Manzanillo.	9,632582	-82,659053	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		Achiote Road	9,222	-80,02	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		Pipeline Rd, km 6.2, Parque Nac. Soberania	9,116667	-79,75	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		Cerro Campana, nr Capira	8,733333	-79,95	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		La Gamba, Costa Rica	8,706042	-83,206505	Este estudo
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		Boquete, 8 km S	8,69859	-82,45046	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Osa Peninsula, Campanario Biological Reserve	8,643	-83,725	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		El Cope, 7.2 km NE	8,616667	-80,583333	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Costa Rica		Banco, Tiskita	8,355833	-83,135	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		Yavisa	8,1719	-77,6862	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Panama		Cana Biological Station	7,75	-77,683333	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Bahia Solano, Choco - Colombia	6,064043	-77,30424	Este estudo
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Surinami		Bigi Poika, West Suriname Road, 108 km WSW Zanderij Airport	5,226944	-55,881667	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Vereda El Triunfo, Finca El Triunfo	5,164944	-74,810583	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Vereda El Triunfo, Finca El Triunfo	5,154083	-74,803111	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Vereda El Triunfo, Finca El Triunfo	5,153444	-74,8025	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Vereda El Triunfo, Finca El Triunfo	5,152583	-74,801944	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Vereda El Triunfo, Finca El Triunfo	5,149972	-74,805389	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Vereda El Triunfo, Finca El Triunfo	5,067278	-74,861889	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Surinami		Brownsberg Nature Park, Witi Creek Trail	4,914133	-55,160933	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		La virginia, Risaralda - Colombia	4,894451	-75,866434	Este estudo
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Guyana		Iwokrama Forest, Pakatau hills	4,748333	-59,026667	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Quimbaya ocase, Quindio - Colombia	4,645392	-75,762729	Este estudo
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Caserío El Laurel	4,58192	-75,79794	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Guyana		Iwokrama Forest, 26 km SW Kurupukari, Iwokrama Mt. base camp	4,338056	-58,810556	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		La Dorada, Quindio - Colombia	4,308391	-75,60692	Este estudo
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		Sin datos	3,902949	-76,390497	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Colômbia		La Bocana	3,84193	-77,17787	Catalogue of Life
<i>Euglossa crassipunctata</i>	<i>Euglossa crassipunctata</i>	Peru		Maynas, Allpahuaya-Mishana Inia Trial	-3	-73,66667	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Normandia, Aldeia Indígena São Marcos, Guaribas	3,911474	-59,903189	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Alto Alegre	3,2997222	-64,0083333	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Caracaraí, Parque Nacional do Viruá, LO-4	1,4688	-61,0069	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Caracaraí, Parque Nacional do Viruá, LO-3, 2500m	1,4687	-61,0068	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Caracaraí, Parque Nacional do Viruá, LO-3, 500m	1,4686	-61,0067	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Caracaraí, Parque Nacional do Viruá, LO-1, Ponto 500m	1,280566	-61,183397	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra da Neblina	0,7469444	-65,9708333	SpeciesLink

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra da Neblina	0,746944	-65,970833	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra do Imeri	0,482778	-65,4975	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra do Imeri	0,4827778	-65,4975	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, Campina	0,477778	-63,47125	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, Campina	0,4777778	-63,47125	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, Campo	0,472167	-63,426722	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, Campo	0,4721667	-63,4267222	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra do Imeri	0,458889	-65,495278	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra do Imeri	0,4588889	-65,4952778	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, próximo à Serrinha	0,4478333	-63,2544722	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, próximo à Serrinha	0,447833	-63,254472	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, Serrinha	0,41925	-63,383528	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Aracá, Serrinha	0,41925	-63,3835278	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Barcelos, Igarapé Coruja	0,1046111	-63,8504722	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Igarapé Coruja	0,104611	-63,850472	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra do Imeri	-0,172222	-65,394722	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Pico da Neblina, Serra do Imeri	-0,1722222	-65,3947222	SpeciesLink

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Demeni, Jalauaca, Lago	-0,2666667	-62,7463889	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Demeni, Alubiá, Igapó	-0,3186111	-62,7605556	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Demeni, Pirico	-0,3211111	-62,7922222	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Demeni, Pirico	-0,3211111	-62,7922222	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Jufari, Comunidade Caju	-0,802861	-62,4855	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Jufari, Comunidade Caju	-0,8028611	-62,4855	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	PA	Rio Trombetas, Juquiri	-1,095207	-57,024988	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Presidente Figueiredo, Reserva Biológica do Uatumã, LO-1, 500m	-1,250964	-59,521059	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	PA	Porto Trombetas, Platô Aviso	-1,462873	-56,399527	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	PA	Oriximiná, Porto Trombetas, Monte Branco - P7	-1,506422	-56,369512	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Margem Esquerda do Rio Nhamundá, Igarapé Areias, Igapó	-1,5863889	-57,6255556	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Margem Esquerda do Rio Nhamundá, Igarapé Areias, Igapó	-1,586389	-57,625556	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Reserva Extrativista Unini, Rio Unini, Lago Galomanha, Terra Firme	-1,6166667	-62,9833333	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa moratoi</i>	Brasil	PA	Comunidade Casinha	-1,725166667	-56,23333333	Este estudo
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Nhamundá, Cuipiranga, Mata	-1,899444	-57,049722	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Nhamundá, Cuipiranga, Mata	-1,8994444	-57,0497222	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Estação Ecológica Nacional Juami-Japurá, Médio Rio Juami	-1,9556667	-67,9299444	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Estação Ecológica Nacional Juami-Japurá, Médio Rio Juami	-1,955667	-67,929944	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Terra Firme	-2,010611	-61,539222	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Terra Firme	-2,0106111	-61,5392222	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Igapó	-2,012472	-61,53675	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Igapó	-2,0124722	-61,53675	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Terra Firme	-2,041944	-61,5525	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Terra Firme	-2,0419444	-61,5525	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Campinarana	-2,046583	-61,551111	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Campinarana	-2,0465833	-61,5511111	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Campinarana	-2,048611	-61,551667	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Carabinani, Campinarana	-2,0486111	-61,5516667	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	MT	Moju	-2,125	-48,7825	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Reserva da Campina, BR 174, Km 44, Floresta	-2,5833333	-60,0333333	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, ZF-2, Km 14, Torre	-2,5891667	-60,1152778	SpeciesLink

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	ZF-2, Km 14, Torre 20m	-2,589167	-60,115278	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	MT	Tailândia	-2,803583	-48,5282	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Reserva Florestal Adolpho Ducke, AM 010, Km 26, LO-4, 500m	-2,963764	-59,931592	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Reserva Florestal Adolpho Ducke, AM 010, Km 26, LO-5, 1500m	-2,970042	-59,931207	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Reserva Florestal Adolpho Ducke, AM 010, Km 26, LO-6, 1500m	-2,97632	-59,930822	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Reserva Florestal Adolpho Ducke, AM 010, Km 26, LO-7, 1500m	-2,982598	-59,930437	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Reserva Florestal Adolpho Ducke, AM 010, Km 26, LO-8, 500m	-2,988876	-59,930052	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Reserva Florestal Adolpho Ducke, AM 010, Km 26, LO-8, 1500m	-2,995154	-59,929667	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Rio Negro, Tarumã	-3,013746	-60,067848	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Puraquequara	-3,019542	-59,849692	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Itacoatiara, AM 010, Km 215, Fazenda Aruanã	-3,0333333	-58,8333333	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus, Campus do Instituto Nacional de Pesquisas da Amazônia, Aleixo	-3,096302	-59,989632	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Manaus	-3,10194	-60,025	SpeciesLink

Espécie	Original	País	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Jaú, Roçado	-3,543343	-62,444987	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Parque Nacional do Jaú, Roçado	-3,651056	-62,060625	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Beruri, Rio Purus	-3,910682	-61,352274	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Campo do Tupana, cerca de 120 Km ao sul de Manaus, Campina	-4,15803	-60,12887	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	MT	Itaituba, Station Maria	-4,2167	-55,9667	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Borba, Rio Abacaxis, Paxiúba, 30 m	-4,48	-58,5733333	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Abacaxis, Paxiúba	-4,48	-58,573333	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Abacaxis, Pacamiri, Campina	-4,596944	-58,220556	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Rio Abacaxis, Pacamiri, Campina	-4,5969444	-58,2205556	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Campo da Catuquira, cerca de 240 Km ao sul de Manaus, Campina	-4,912583	-61,108517	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Campo da Catuquira, cerca de 240 Km ao sul de Manaus, Campina	-4,9125833	-61,1085167	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Campo da Catuquira, cerca de 240 Km ao sul de Manaus, Mata	-4,919733	-61,110317	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Campo da Catuquira, cerca de 240 Km ao sul de Manaus, Mata	-4,9197333	-61,1103167	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Borba, Rio Abacaxis, Cachoeira Travessão	-5,2525	-58,6977778	SpeciesLink

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RR	Rio Madeira, Lago Xadá, Comunidade Bela Vista, Ponto 1, Margem Esquerda	-5,260833	-60,708889	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Novo Aripuanã, Rio Madeira, Lago Xadá, Comunidade Bela Vista, Ponto 1, Margem Esquerda	-5,2608333	-60,7088889	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	BR 230, Km 12, Ramal Palheiral, Km 9	-7,319444	-64,668611	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Lábrea, BR 230, Km 12, Ramal Palheiral, Km 9	-7,3194444	-64,6686111	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Apuí, BR 174, Km 72, PDBFF, Reserva 2108	-7,425312	-60,593782	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Apuí, BR 174, Km 70, PDBFF, Pasto	-7,438968	-60,615411	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RO	Machandino D'Oeste, Fazenda Lami	-9,315783	-61,79475	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RO	Machandino D'Oeste, Fazenda Lami	-9,3157833	-61,79475	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RO	Cujubim, Linha 2	-9,38675	-62,5809	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Machandino D'Oeste, Igarapé Mucura	-9,449133	-61,857417	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	RO	Machandino D'Oeste	-9,4491333	-61,8574167	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	AM	Ariquemes, 62 km SE	-9,9137	-63,0439	Catalogue of Life
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	MT	Nova Mutum, Fazenda Buriti, Estrada da Roda d'água	-13,8166667	-56,0666667	SpeciesLink
<i>Euglossa moratoi</i>	<i>Euglossa crassipunctata</i>	Brasil	MT	Nova Mutum, Fazenda Buriti	-13,8166668	-56,0666668	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	RR	Caracaraí, Parque Nacional do Viruá, LO-2,1500m	1,4775	-61,0157	SpeciesLink

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Barcelos, Serra do Tapirapecó	1,2075	-64,7775	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	São Gabriel da Cachoeira, Estrada para Cucuí, BR 307, Km 64, Torre do LBA	0,2119444	-66,7747222	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Reserva Extrativista Unini, Rio Unini, Três Bocas, Igapó	-1,0666667	-62,9666667	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Campo do Tupana, cerca de 120 Km ao sul de Manaus, Mata	-4,15679	-60,13203	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Carauari, Margem Esquerda do Rio Juruá, Comunidade Nova Esperança, Trilha do Roque	-5,0908333	-67,1825	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Novo Aripuanã, Floresta Úmida	-5,2647222	-60,1188889	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Peru		Lagunas, Loreto	-5,301829	-75,636713	Este estudo
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Canutama, Ramal do Mucuí, Floresta	-8,6346111	-64,3751944	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	AM	Canutama, Ramal do Mucuí, Campina	-8,6601389	-64,366	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	RO	none, Machadinho, Fazenda Lami	-9,31578	-61,79475	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	RO	Cujubim, Linha 14	-9,3359	-62,5518333	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	RO	Mirante da Serra, Linha 68	-11,0746833	-62,72395	SpeciesLink
<i>Euglossa parvula</i>	<i>Euglossa parvula</i>	Brasil	MT	Nova Mutum, Fazenda Buriti, Estrada da Roda D'agua	-13,81667	-56,06667	SpeciesLink
<i>Euglossa pepeï</i>	<i>Euglossa pepeï</i>	Brasil	BA	Prado	-17,33989	-39,22829	SpeciesLink
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Estación Pitilla, 9 Km S Santa Cecilia.	10,992609	-85,429477	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Nicaragua		San Carlos, 60 km SE, Refugio Bartola	10,973333	-84,338333	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Fca. Dago Mora	10,794556	-84,899778	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		El Pilon, Cerro La Carmela entre Rio Celeste y cabeceras del Rio Chimurria. Bijagua.	10,720833	-84,995833	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Cerro La Carmela	10,710164	-84,997578	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Estación Pilon	10,704605	-84,99231	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Valle Río Roble, Pie de Terraza 1 (Send. a Montezuma).	10,698407	-85,012032	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Punto 2: Falda N Cerro Montezuma.	10,698113	-85,018467	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		ca. 7 km NE de Boca Tapada, Laguna del Lagarto Lodge.	10,686111	-84,180556	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Sarapiquí, La Virgen, Finca del IDA.	10,643616	-84,107585	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Pococí, Colorado, Sector Cerro Cocorí, 30 Km N de Cariari	10,594274	-83,716512	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Guatuso de Alajuela, Cabanga. Finca de José Martínez.	10,59	-84,852806	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Puerto Viejo, 3 km N	10,473333	-84,015	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		La Selva Biological Station	10,433333	-84,016667	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		A.C.C.V.C, Sarapiquí, Zona Prot La Selva, Estación Biol La Selva	10,431957	-84,008187	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica	Heredia	Estacion Biologica La Selva	10,4166667	-84	SpeciesLink
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Puerto Viejo, W, Selva Verde Lodge	10,335	-84,165	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		El Ceibo	10,329083	-84,080102	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Guacimo, 7 mi N	10,3186	-83,6833	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Pococí, Finca INBio.	10,19434	-83,86086	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Fca. Sandoval-RECOPE, Transecto Bosque A.	10,001689	-83,106659	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Fca. Sandoval-RECOPE, Sitio Manglar.	10,000416	-83,097735	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Fca. Sandoval-RECOPE, Transecto Bosque C.	10,000041	-83,113086	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Fca. Sandoval-RECOPE, Polígono (Sitio D, herbazal ó área abierta).	9,999566	-83,109394	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Fca. Sandoval-RECOPE, Transecto Bosque B.	9,997867	-83,107185	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		INBioparque, Santa Rosa, Santo Domingo.	9,97362	-84,093477	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		200 mts. Norte Restaurant El Cruce.	9,946497	-83,573117	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Veragua Rainforest.	9,92573	-83,191405	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Veragua Rainforest, Río Victoria.	9,923732	-83,18895	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Finca Josephina, 2 km S Pital	9,7985	-84,4667	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Finca Quijote, Propietario Sr. Philip E. Hacock	9,796708	-83,639719	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		C. Turrubares [Cerro Bares]	9,794664	-84,473244	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		R.B. Carara, 2 Km N. de Bijagual	9,774293	-84,527464	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Garabito, Reserva Biol Carara, Est Quebrada Bonita	9,774233	-84,608124	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Estación Quebrada Segundo	9,762583	-83,788328	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		A.C.L.A.P, Paraíso, Pque Nal Tapantí, Send Oropendula	9,755344	-83,784695	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Bribri, borde de carretera.	9,615164	-82,861492	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Amubri (Talamanca)	9,51666	-82,96666	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Cerro Nara, NE of Quepos	9,4999	-84,0166	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Cerro Nara, falda NE Finca Oldemar Valverde	9,499849	-84,011765	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Puntarenas, Aguirre. Palma Tica, Finca Naranjo	9,457241	-84,11517	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Arboleda, Costa Rica	9,397162	-84,153996	Este estudo
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		P. N. Manuel Antonio, borde este	9,393803	-84,129222	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Carti-El Llano Rd, km 7.5, N of Chepo	9,216667	-79,083333	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Ojochal.	9,075283	-83,657497	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Casa Coca, camino a Valle del Silencio.	9,04862	-82,990224	Catalogue of Life

Espécie	Original	Pais	UF	Localidade	Latitude	Longitude	Referência
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		P.N. La Amistad. Estación Altamira. Sendero a Casa Coca.	9,042898	-82,991308	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Los Gigantes del Bosque, Altamira.	9,034719	-83,01388	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Estación Altamira	9,032987	-83,010881	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Ciudad Pto Cortés, Río Balsar.	9,006418	-83,530972	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Palmar Sur, 4 millers E.	8,95078	-83,470703	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Las Cruces Biological Station	8,785667	-82,959667	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Estacion Esquinas, Peninsula de Osa	8,759387	-83,283128	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Golfito	8,6392	-83,1624	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Sendero hacia Llorona	8,5932	-83,721174	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		A.C.O, Golfito, Pque Nal Corcovado, Est Sirena, Playa Sirena	8,480171	-83,591289	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Sirena, Corcovado Nat. Pk., Osa Peninsula	8,479267	-83,588565	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Costa Rica		Corcovado National Park, Sirena Station, lower Ollas Trail	8,413333	-83,589444	Catalogue of Life
<i>Euglossa sapphirina</i>	<i>Euglossa sapphirina</i>	Guyana Francesa		Kourou, 11 km SW	5,0797	-52,72058	Catalogue of Life

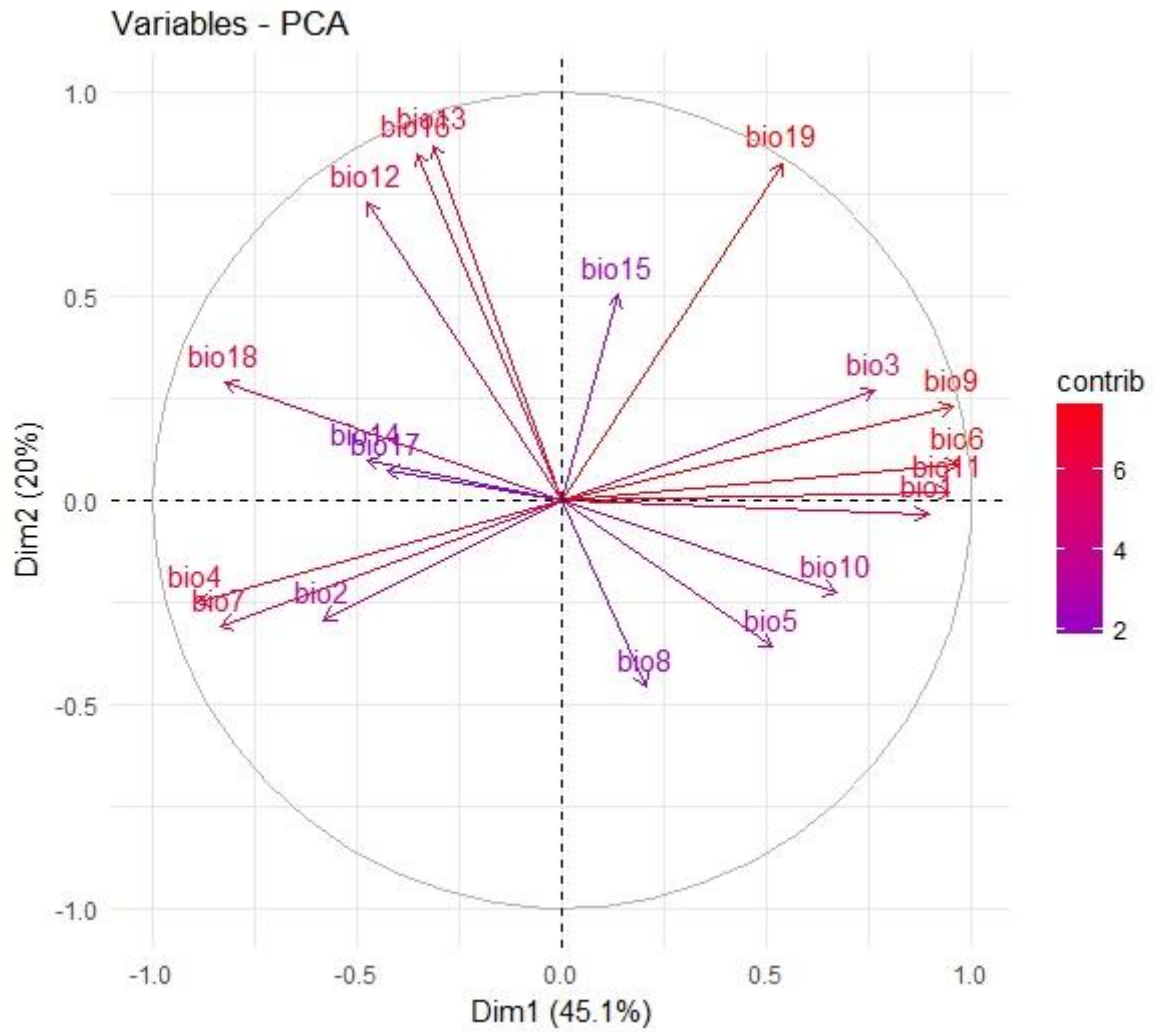
1
2 Apêndice E - Registros considerados de *Euglossa iopoecila* Dressler (1982) no capítulo 3.

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Synonyms
1	<i>Euglossa iopoecila</i>	-8.723137	-35.85241	PE	Jaqueira	RPPN Frei Caneca	Nemésio A (2009) Zootaxa 2041:242pp	como <i>Eg. roubiki</i>
2	<i>Euglossa iopoecila</i>	-9.00266	-35.84526	AL	Ibatengura	Usina Serra Grande	Nemesio and Santos-Junior(2014) Bras Jour Biol 74:S78-S92	como <i>Eg. roubiki</i>
3	<i>Euglossa iopoecila</i>	-9.179589	-35.90246	AL	Murici	ESEC Murici	Nemesio and Santos-Junior(2014) Bras Jour Biol 74:S78-S92	como <i>Eg. roubiki</i>
4	<i>Euglossa iopoecila</i>	-12.85261	-39.47884	BA	Santa Teresinha	Serra Jibóia	Ferrari and Melo (2014) Apidologie 45:641–652	como <i>Eg. iopoecila</i>
5	<i>Euglossa iopoecila</i>	-13.37272	-39.09984	BA	Valença		Peruquetti et al (1999) Rev Bras Zool 16:101-118	como <i>Eg. chalybeata</i>
6	<i>Euglossa iopoecila</i>	-13.83383	-39.249	BA	Igrapiúna	Reserva Ecologica Michelin	Nemesio (2014) Bras Jour Biol 74:16-22	como <i>Eg. roubiki</i>
7	<i>Euglossa iopoecila</i>	-14.49622	-39.13697	BA	Uruçaca	Parque Estadual Serra do Conduru	Nemesio (2011) Zootaxa 2821:47-54	como <i>Eg. roubiki</i>
8	<i>Euglossa iopoecila</i>	-14.49652	-39.13283	BA	Uruçaca	Parque Estadual Serra do Conduru	Nemesio (2011) Zootaxa 2821:47-54	como <i>Eg. roubiki</i>
9	<i>Euglossa iopoecila</i>	-14.84203	-39.29246	BA	Itabuna		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284	como <i>Eg. iopoecila</i>
10	<i>Euglossa iopoecila</i>	-15.16866	-39.0647	BA	Una	REBIO Una site 1	Nemesio (2013) Bras Jour Biol 73:347-352	como <i>Eg. roubiki</i>
11	<i>Euglossa iopoecila</i>	-15.174	-39.06128	BA	Una	REBIO Una site 2	Nemesio (2013) Bras Jour Biol 73:347-352	como <i>Eg. roubiki</i>
12	<i>Euglossa iopoecila</i>	-15.18336	-39.067	BA	Una	REBIO Una site 3	Nemesio (2013) Bras Jour Biol 73:347-352	como <i>Eg. roubiki</i>
13	<i>Euglossa iopoecila</i>	-15.39302	-39.56395	BA	Camacan	Proximidades do PARNA Serra das Lontras	Ferrari and Melo (2014) Apidologie 45:641–652	como <i>Eg. iopoecila</i>
14	<i>Euglossa iopoecila</i>	-16.40609	-40.04732	MG	Santa Maria do Salto	RPPN Duas Barras, Site 2	Nemesio (2012) Spixiana 35:109-116	como <i>Eg. roubiki</i>

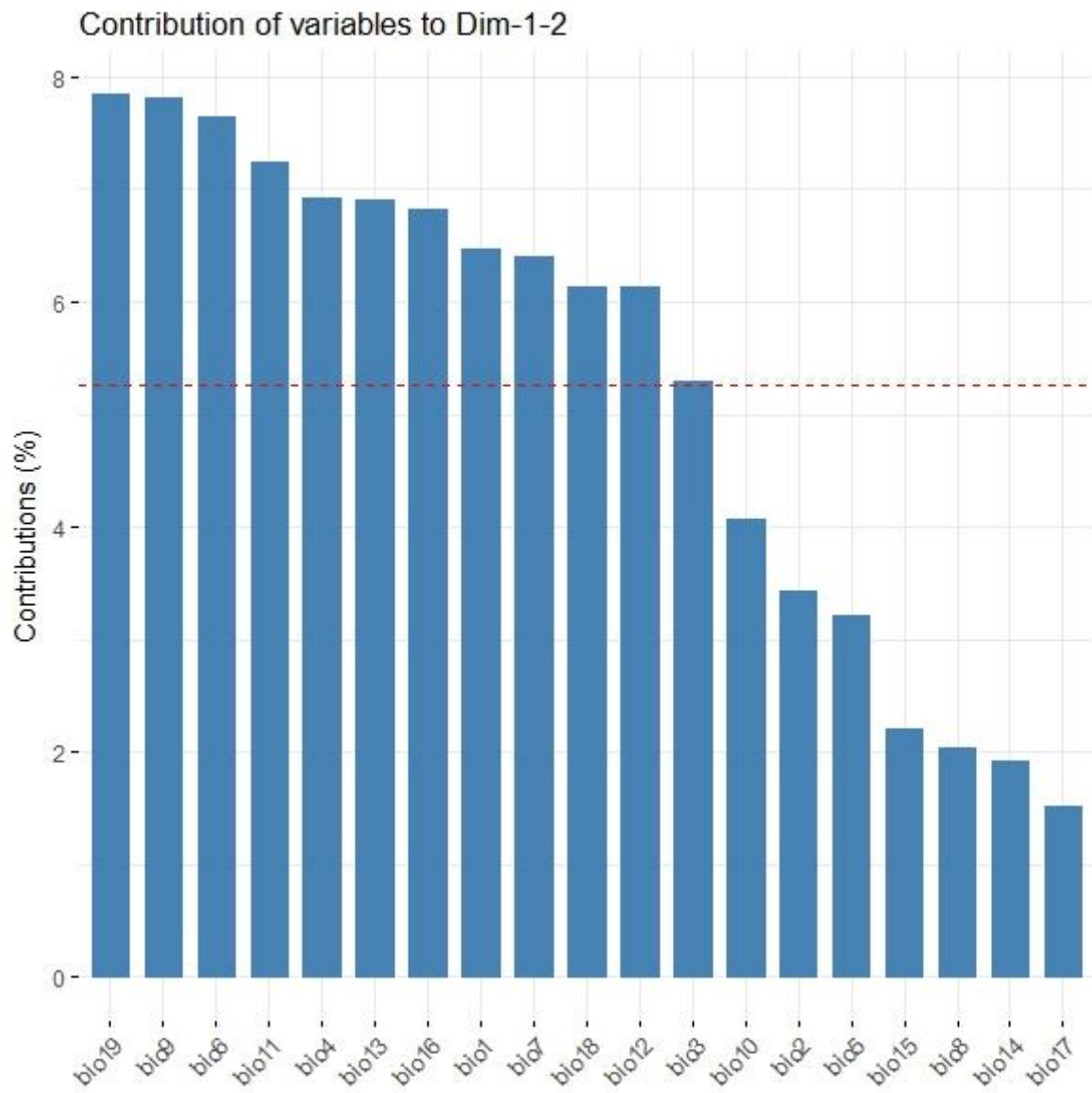
ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Synonyms
15	<i>Euglossa iopoecila</i>	-16.41171	-40.06213	MG	Santa Maria do Salto	RPPN Duas Barras, Site 3	Nemesio (2012) Spixiana 35:109-116	como <i>Eg. roubiki</i>
16	<i>Euglossa iopoecila</i>	-16.41454	-40.05384	MG	Santa Maria do Salto	RPPN Duas Barras, Site 1	Nemesio (2012) Spixiana 35:109-116	como <i>Eg. roubiki</i>
17	<i>Euglossa iopoecila</i>	-16.48427	-39.15655	BA	Porto Seguro	PARNA Pau Brasil	Nemesio (2013) Bras Jour Biol 73:375-381	como <i>Eg. roubiki</i>
18	<i>Euglossa iopoecila</i>	-16.49417	-39.2679	BA	Porto Seguro	PARNA Pau Brasil	Nemesio (2013) Bras Jour Biol 73:375-381	como <i>Eg. roubiki</i>
19	<i>Euglossa iopoecila</i>	-16.88505	-39.35394	BA	Porto Seguro	PARNA Descobrimento	Nemesio (2013) Bras Jour Biol 73:437-446	como <i>Eg. roubiki</i> and <i>Eg. iopoecila</i> respectively
20	<i>Euglossa iopoecila</i>	-17.13083	-39.39401	BA	Prado	PARNA Descobrimento	Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284; this study	como <i>Eg. iopoecila</i>
21	<i>Euglossa iopoecila</i>	-17.51194	-39.74034	BA	Teixeira de Freitas		Nemésio A (2009) Zootaxa 2041:242pp	como <i>Eg. roubiki</i>
22	<i>Euglossa iopoecila</i>	-18.5473	-39.73903	ES	Conceição da Barra		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–285	como <i>Eg. iopoecila</i>
23	<i>Euglossa iopoecila</i>	-18.99192	-40.129	ES	Linhares	REBIO Sooreta	Ferrari and Melo (2014) Apidologie 45:641–652	como <i>Eg. iopoecila</i>
24	<i>Euglossa iopoecila</i>	-19.15199	-40.019	ES	Linhares	Reserva Vale	Ferrari and Melo (2014) Apidologie 45:641–652	como <i>Eg. iopoecila</i>
25	<i>Euglossa iopoecila</i>	-19.82402	-40.323	ES	Aracruz		Peruquetti et al (1999) Rev Bras Zool 16:101-118	como <i>Eg. chalybeata</i>
26	<i>Euglossa iopoecila</i>	-20.36151	-40.65451	ES	Domigos Martins	PNM Domingos Martins	Nemésio A (2009) Zootaxa 2041:242pp	como <i>Eg. roubiki</i>
27	<i>Euglossa iopoecila</i>	-20.63304	-40.55975	ES	Guarapari		Peruquetti et al (1999) Rev Bras Zool 16:101-118	como <i>Eg. chalybeata</i>
28	<i>Euglossa iopoecila</i>	-21.88162	-41.85283	RJ	Campos dos Goytacazes	PE Desengano	Tonhasca et al (2012) Biotropica 34:416-422	como <i>Eg. chalybeata</i>
29	<i>Euglossa iopoecila</i>	-22.42701	-42.04299	RJ	Rio das Ostras	REBIO União	Ferrari and Melo (2014) Apidologie 45:641–652	como <i>Eg. iopoecila</i>
30	<i>Euglossa iopoecila</i>	-22.59994	-42.46656	RJ	Embaú		Ramalho et al. (2009) Rev Bras Entomol 53:95-101	como <i>Eg. iopoecila</i>

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Synonyms
31	<i>Euglossa iopoecila</i>	-23.20805	-44.822	RJ	Paraty	Serra da Bocaina	Ferrari and Melo (2014) Apidologie 45:641–652	como <i>Eg. iopoecila</i>
32	<i>Euglossa iopoecila</i>	-23.35955	-44.78769	SP	Ubatuba	PESM - Picinguaba	Rocha-Filho and Garofalo (2013) J Inse Scien 13(23)1-4;	como <i>Eg. chalybeata iopoecila</i> , <i>Eg. Iopoecila</i>
33	<i>Euglossa iopoecila</i>	-23.73821	-45.70028	SP	São Sebastião	PESS-Juqueí	Penha et al (2015) Apidologie 46:346–358	como <i>Eg. iopoecila</i>
34	<i>Euglossa iopoecila</i>	-23.88617	-45.37501	SP	Ilhabela	PESS-Juqueí	Penha et al (2015) Apidologie 46:346–358	como <i>Eg. iopoecila</i>
35	<i>Euglossa iopoecila</i>	-24.36682	-47.96675	SP	Sete Barras	Fazenda Morro do Capim	Mattozo et al (2011) Pap Avulsos Zool 51:505–515	como <i>Eg. iopoecila</i>
36	<i>Euglossa iopoecila</i>	-24.39143	-47.12484	SP	Iguape	Estação Ecologica Juréia-Itatins	Peruquetti et al (1999) Rev Bras Zool 16:101-118	como <i>Eg. chalybeata iopoecila</i>
37	<i>Euglossa iopoecila</i>	-25.023	-47.95338	SP	Cananéia	PESM-cananeia	Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284; Penha et al (2015) Apidologie 46:346–358	como <i>Eg. iopoecila</i>
38	<i>Euglossa iopoecila</i>	-25.11281	-47.99344	SP	Ilha do Cardoso	PEIC	Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284; Penha et al (2015) Apidologie 46:346–359	como <i>Eg. iopoecila</i>
39	<i>Euglossa iopoecila</i>	-25.24447	-48.4864	PR	Guaraqueçaba		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284	como <i>Eg. iopoecila</i>
40	<i>Euglossa iopoecila</i>	-25.31342	-48.65716	PR	Antonina	RN Rio Cachoeira	Mattozo et al (2011) Pap Avulsos Zool 51:505–515	como <i>Eg. iopoecila</i>
41	<i>Euglossa iopoecila</i>	-25.52196	-48.57336	PR	Paranaguá	CBSM - Paranaguá	Faria Jr and Melo(2007). Rev Bras Entomol 51:275–285	como <i>Eg. iopoecila</i>
42	<i>Euglossa iopoecila</i>	-25.54803	-48.63331	PR	Alexandra		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–286	como <i>Eg. iopoecila</i>
43	<i>Euglossa iopoecila</i>	-25.81008	-48.56436	PR	Matinhos	CBSM - Matinhos	Faria Jr and Melo(2007). Rev Bras Entomol 51:275–286	como <i>Eg. iopoecila</i>
44	<i>Euglossa iopoecila</i>	-26.22634	-48.95324	SC	Joinville		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–287	como <i>Eg. iopoecila</i>

Mapa de Fatores para 19 variáveis climáticas exploradas via análise de componentes principais baseada na ocorrência de espécies do grupo *crassipunctata*



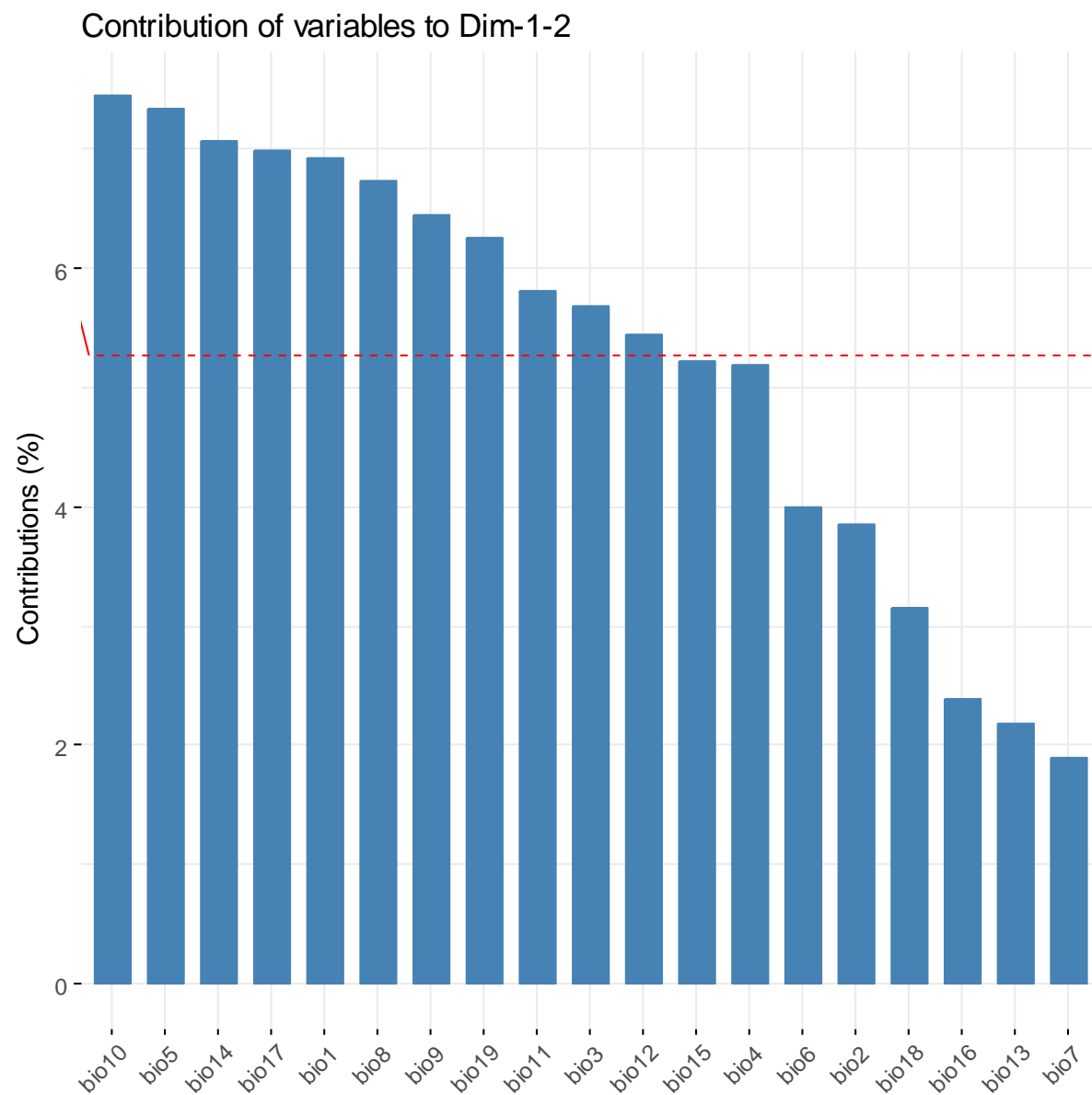
Percentual de contribuição de 19 variáveis climáticas em duas dimensões exploradas via análise de componentes principais baseada na ocorrência de espécies do grupo *crassipunctata*



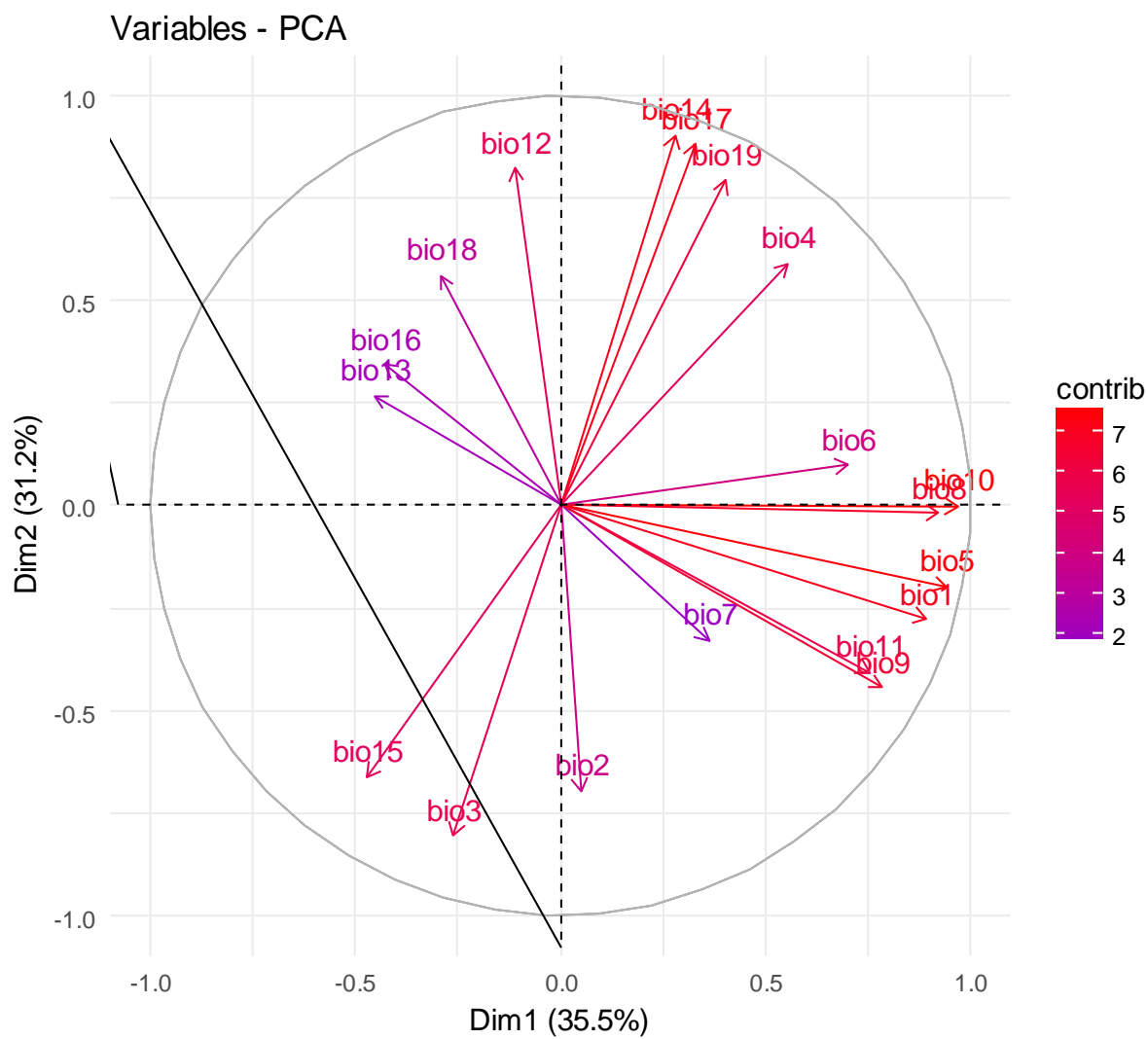
Localidades amostradas para *E. annectans* ao longo da Mata Atlântica

Código	Localidade	Latitude	Longitude
SO	Parque Nacional da Serra dos Órgãos, Teresópolis - RJ	-22.4925	-43.0594
PN	Parque Nacional do Pico da Neblina, São Paulo - SP	-23.5611	-46.1488
BO	Boracéia - SP	-23.7342	-46.1855
IC	Parque Estadual da Ilha do Cardoso - SP	-25.1118	-47.9514
SU	Parque Nacional Ilha de Superagui - PR	-25.4312	-48.2464
SM	Reserva Natual de Salto Morato - PR	-25.1636	-48.3002
SC	Município de Joinvile - SC	-26.2067	-49.0452
KB	Reserva Fazenda Monte Alegre (Klabin) - PR	-24.1685	-50.6953
MD	Parque Estadual do Morro do Diabo, Teodoro Sampaio - SP	-22.5	-52.3421
DN	Diamante do Norte - PR	-22.6418	-52.8845
FZ	Parque Nacional de do Iguaçu, Foz do Iguaçu - PR	-25.6244	-54.4616
DR	Mata da Picadinha, Dourados - MS	-22.1547	-55.0002

Apêndice F - Percentual de contribuição de 19 variáveis climáticas em duas dimensões exploradas via análise de componentes principais baseada na ocorrência de espécies do grupo *crassipunctata*



Apêndice G - Mapa de Fatores para 19 variáveis climáticas exploradas via análise de componentes principais baseada na ocorrência de *Euglossa annectans*



Apêndice H - Registros de espécies utilizados na produção dos modelos de nicho ecológico para *Euglossa annectans*. Referências incluem literatura e plataforma SpeciesLink e Catalogue of Life.

Espécies	Determin. Anterior	País	UF	Localização	Local específico	Longitude	Latitude	N	Referências
<i>Euglossa annectans</i>	<i>Euglossa stellfeldi</i>	Brasil	MG	São Gonçalo do Rio Preto	PE Rio Preto	-43.34503	-18.11779	2	Nemésio & Faria Jr 2009
<i>Euglossa annectans</i>	<i>Euglossa stellfeldi</i>	Brasil	MG	São Gonçalo do Rio Preto	PE Rio Preto	-43.33049	-18.15654	14	Antonine et al. 2016
<i>Euglossa annectans</i>	<i>Euglossa stellfeldi</i>	Brasil	MG	Diamantina		-43.58233	-18.20261	1	Viotti et al. 2013
<i>Euglossa annectans</i>		Brasil	MG	Belo Horizonte	Belo Horizonte	-43.97222	-19.88236	2	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	ES	Santa Teresa		-40.60688	-19.93202	3; 1	Dressler, 1982; SpeciesLink
<i>Euglossa annectans</i>		Brasil	MG	Belo Horizonte	Zona Metalúrgica - Parque Mangabeiras	-43.91343	-19.94432	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	MG	Ibirité	COPASA/Taboões PROBIO	-44.06115	-20.01664	1	SpeciesLink; Ferrari et al. 2017
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Sertãozinho	Est. Exp. de Zootecnia	-48.08935	-21.15159	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Cajuru	Fazenda Santa Carlota	-47.36064	-21.28397	1	Rebelo e Garófalo, 1997
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Dumont	Mata Boa Vista	-47.97291	-21.30782	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	MG	Alfenas		-45.93615	-21.31556	16	Ranieiro 2013
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Luiz Antônio	Estação Ecológica de Jataí	-47.78543	-21.55428	1	SpeciesLink; Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Roque da Fartura	Mirante	-46.72479	-21.82373	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Carlos		-48.04552	-22.04437	17	Souza et al. 2010
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RJ	Trajano de Morais	Mata da Cabecinha	-42.08441	-22.08083	9	Aguiar et al. 2014
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RJ	Trajano de Morais		-42.09445	-22.09739	16	Aguiar et al. 2015

Espécies	Determin. Anterior	País	UF	Localização	Local específico	Longitude	Latitude	N	Referências
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	MS	Dourados	Serra da Picadinha	-55.00047	-22.15702	14	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Brotas		-48.17103	-22.21969	1	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Bauru	Ecological Station of Bauru	-49.08345	-22.23523	29	Knoll et al. 2012
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	MG	Jacutinga	Pico da Forquilha	-46.67097	-22.27118	378	Cardoso Jr 2010
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Bauru	Permanent Preservation Area Vargem Limpa	-49.01681	-22.35017	27	Knoll et al. 2012
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Alvilândia	Estação Ecológica Caetetus	-49.70926	-22.42052	202	Knoll et al. 2012
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RJ	Terezópolis	Serra dos órgãos	-42.99526	-22.45772	40	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Agudos	Permanent Preservation Durapex area	-48.86485	-22.48451	58	Knoll et al. 2012
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Teodoro Sampaio	Parque Est. Morro do Diabo	-52.32778	-22.5425	5	SpeciesLink; Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Diamante do Norte	APP	-52.85639	-22.61833	35	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Teodoro Sampaio	Morro do Diabo	-52.17167	-22.62083	40	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Pindamonhangaba	Faz São Sebastião do Ribeirão Grande	-45.46679	-22.75003	1	Uehara-Prado et al. 2006
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Campinas	Reserva Santa Genebra	-47.10895	-22.82906	3	Genini 2011; Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Pindamonhangaba	Tequetera	-45.4376	-22.85636	1	Costa & Francoy 2017
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RJ	Rio de Janeiro		-43.45452	-22.91313	27	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RJ	Rio de Janeiro	Floresta da Tijuca	-43.27721	-22.95831	6	SpeciesLink; Faria Jr & Melo 2007

Espécies	Determin. Anterior	País	UF	Localização	Local específico	Longitude	Latitude	N	Referências
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Jundiaí	Floresta do Japí	-46.86652	-23.18277	2	Garofallo et al. 1998; SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Cornélio Procópio	Mata São Francisco	-50.62151	-23.19254	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Jundiaí		-46.9105	-23.23097	1	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Ubatuba	Picinguaba	-44.83707	-23.36245	1	Costa & Francoy 2017
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Londrina	Apucarantina-Godoy	-51.23861	-23.4475	3	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Londrina	Mata dos Godoy	-51.25591	-23.4481	36	Ferronato et al. 2017
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Mogi das Cruzes	Parque das Neblinas	-46.18255	-23.48151	50	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Paulo	Universidade de São Paulo (USP)	-46.81522	-23.48214	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Sebastião	Ilha Bela	-45.2228	-23.5039	2	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Sebastião	Boraceia	-45.86972	-23.63083	9	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Salesópolis	Estação Ecológica da Boracéia	-45.86655	-23.63333	19	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Paulo	Morumbi Park	-46.72504	-23.63549	53	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Cotia	Reserva Florestal de Morro Grande	-46.99636	-23.7083	5	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Mogi das Cruzes	Parque da Neblina	-46.16333	-23.74806	22	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Tapejara		-52.84941	-23.7645	1	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	São Sebastião	Parque Estadual de Ilhabela	-45.4264	-23.8311	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Telemaco borba	Faz Morro Alegre-Klabin	-50.70081	-24.08805	47; 5	Giangarelli et al. 2015; SpeciesLink

Espécies	Determin. Anterior	País	UF	Localização	Local específico	Longitude	Latitude	N	Referências
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Telemaco Borba	Fazenda Monte Alegre - Klabin	-50.55722	-24.21167	26	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Palotina	Campus UFPR	-53.84294	-24.29315	2	Margato-Aloisio 2016
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Palotina	PE São Camilo	-53.91749	-24.313	1	Sidney & Gonçalves
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RJ	Sossego do Itambé		-47.96938	-24.36542	1	Tonhasca et al. 2002
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Sete Barras	Fazenda Morro do Capim	-47.96662	-24.36669	5	Boff et al. 2015
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SP	Ilha do Cardoso		-47.94452	-25.14939	11	Giangarelli et al. 2015b; Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Guaraqueçaba	Reserva Natural Salto Morato	-48.29778	-25.18167	8; 46; 119	SpeciesLink; Present study; Giangarelli 2015
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Guaraqueçaba	Parque Nacional do Superagui	-48.15583	-25.33139	12; 60	SpeciesLink, Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Antonina		-48.71915	-25.42791	22	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Morretes		-48.82274	-25.47752	7	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Alexandra		-48.63513	-25.55042	17	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Argentina	Misiones	Porto iguazú		-54.54378	-25.62275	7	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	PR	Foz do Iguaçu	Parque Nacional do Iguaçu	-54.43292	-25.63589	17; 60	SpeciesLink; This study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	Joinville	Joinville	-49.04306	-26.20667	35	Present study
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	Joinville		-48.95108	-26.24157	3	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	São Francisco do Sul		-48.65515	-26.2903	2	Faria Jr & Melo 2007
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	RS	Tenente Portela		-53.7124	-27.34041	4	Faria Jr & Melo 2007

Espécies	Determin. Anterior	País	UF	Localização	Local específico	Longitude	Latitude	N	Referências
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	Florianópolis	Santo Antônio de Lisboa	-48.51962	-27.50318	20 (larvae)	Cortopassi-Laurino et al. 2010
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	Florianópolis	Unidade de Conservação Ambiental Desterro	-48.51244	-27.53063	23 (larvae); 21; 30	Cortopassi-Laurino et al. 2009; Paxton et al. 2009; Kamke et al. 2011
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	Florianópolis	Dunas, próximo à Lagoa da Conceição	-48.48126	-27.62352	1	SpeciesLink
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	Morro da Pipa		-48.643	-27.83916	11	Kamke et al. 2011b
<i>Euglossa annectans</i>	<i>Euglossa annectans</i>	Brasil	SC	São Martinho		-48.92885	-28.10701	2	Faria Jr & Melo 2007

Apêndice I - Localidades (nome da reserva, município e estado), código da localidade, coordenadas genética e tamanho de amostral (n) de *Euglossa iopocila* colizada ao longo da Mata Atlântica Brasileira.

Localidade	ID	Latitude	Longitude	n
-------------------	-----------	-----------------	------------------	-----------------------

Vila Glória, São Francisco do Sul, SC	SC	26°14'38.00"S	48°38'18.00"W	10
Superagui Island, Guaraqueçaba, PR	PR1	25°23'11.00"S	48°13'57.00"W	10
Reserva Natural Salto Morato, Guaraqueçaba, PR	PR2	25°17'24.02"S	48°19'02.01"W	10
Parque Estadual Ilha do Cardoso, Cardoso Island, Cananéia, SP	SP1	25°08'48.00"S	47°56'47.00"W	10
Mainland, Cananéia, SP	SP2	25°01'02.65"S	47°55'10.50"W	10
Parque Estadual de Ilhabela, Ilhabela, SP	SP3	25°50'39.00"S	45°21'23.00"W	10
Parque Estadual da Serra do Mar, São Sebastião, SP	SP4	23°46'11.00"S	45°36'24.00"W	10
Parque Estadual da Serra do Mar (Picinguaba), Ubatuba, SP	SP5	25°21'51.00"S	44°49'25.00"W	10
Reserva Biológica União, Casimiro de Abreu/Rio das Ostras/ Macaé, RJ	RJ	22°25'39.02"S	42°02'27.70"W	10
Reserva Biológica Sooretama, Linhares, ES	ES	19°03'45.72"S	40°07'07.36"W	9
Parque Nacional do Descobrimento, Itamaraju, BA	BA1	17°08'56.18"S	39°22'21.03"W	10
Parque Estadual da Serra do Conduru, Ilhéus, BA	BA2	14°43'39.22"S	39°06'48.86"W	10
Reserva Ecológica Michelin, Igrapiúna, Bahia	BA3	13°46'31.56"S	39°09'13.93"W	3
Estação Ecológica de Murici, Murici, AL	AL	09°11'24.63"S	35°52'53.68"W	4
Total				126

Apêndice J - Lista de ocorrências utilizadas para a construção dos modelos de nicho ecológico para *Euglossa iopoecila*.

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Sinonimia
1	<i>Euglossa iopoecila</i>	-8.723137	-35.85241	PE	Jaqueira	RPPN Frei Caneca	Nemésio A (2009) Zootaxa 2041:242pp	as <i>E. roubiki</i>
2	<i>Euglossa iopoecila</i>	-9.00266	-35.84526	AL	Ibatengura	Usina Serra Grande	Nemesio and Santos-Junior(2014) Bras Jour Biol 74:S78-S92	as <i>E. roubiki</i>
3	<i>Euglossa iopoecila</i>	-9.179589	-35.90246	AL	Murici	ESEC Murici	Nemesio and Santos-Junior(2014) Bras Jour Biol 74:S78-S92	as <i>E. roubiki</i>
4	<i>Euglossa iopoecila</i>	-12.85261	-39.47884	BA	Santa Teresinha	Serra Jibóia	Ferrari and Melo (2014) Apidologie 45:641–652	as <i>E. iopoecila</i>
5	<i>Euglossa iopoecila</i>	-13.37272	-39.09984	BA	Valença		Peruquetti et al (1999) Rev Bras Zool 16:101-118	as <i>E. chalybeata</i>
6	<i>Euglossa iopoecila</i>	-13.83383	-39.249	BA	Igrapiúna	Reserva Ecologica Michelin	Nemesio (2014) Bras Jour Biol 74:16-22	as <i>E. roubiki</i>
7	<i>Euglossa iopoecila</i>	-14.49622	-39.13697	BA	Uruçaca	Parque Estadual Serra do Conduru	Nemesio (2011) Zootaxa 2821:47-54	as <i>E. roubiki</i>
8	<i>Euglossa iopoecila</i>	-14.49652	-39.13283	BA	Uruçaca	Parque Estadual Serra do Conduru	Nemesio (2011) Zootaxa 2821:47-54	as <i>E. roubiki</i>
9	<i>Euglossa iopoecila</i>	-14.84203	-39.29246	BA	Itabuna		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284	as <i>E. iopoecila</i>
10	<i>Euglossa iopoecila</i>	-15.16866	-39.0647	BA	Una	REBIO Una site 1	Nemesio (2013) Bras Jour Biol 73:347-352	as <i>E. roubiki</i>

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Sinonimia
11	<i>Euglossa iopoecila</i>	-15.174	-39.06128	BA	Una	REBIO Una site 2	Nemesio (2013) Bras Jour Biol 73:347-352	as <i>E. roubiki</i>
12	<i>Euglossa iopoecila</i>	-15.18336	-39.067	BA	Una	REBIO Una site 3	Nemesio (2013) Bras Jour Biol 73:347-352	as <i>E. roubiki</i>
13	<i>Euglossa iopoecila</i>	-15.39302	-39.56395	BA	Camacan	Proximidades do PARNA Serra das Lontras	Ferrari and Melo (2014) Apidologie 45:641–652	as <i>E. iopoecila</i>
14	<i>Euglossa iopoecila</i>	-16.40609	-40.04732	MG	Santa Maria do Salto	RPPN Duas Barras, Site 2	Nemesio (2012) Spixiana 35:109-116	as <i>E. roubiki</i>
15	<i>Euglossa iopoecila</i>	-16.41171	-40.06213	MG	Santa Maria do Salto	RPPN Duas Barras, Site 3	Nemesio (2012) Spixiana 35:109-116	as <i>E. roubiki</i>
16	<i>Euglossa iopoecila</i>	-16.41454	-40.05384	MG	Santa Maria do Salto	RPPN Duas Barras, Site 1	Nemesio (2012) Spixiana 35:109-116	as <i>E. roubiki</i>
17	<i>Euglossa iopoecila</i>	-16.48427	-39.15655	BA	Porto Seguro	PARNA Pau Brasil	Nemesio (2013) Bras Jour Biol 73:375-381	as <i>E. roubiki</i>
18	<i>Euglossa iopoecila</i>	-16.49417	-39.2679	BA	Porto Seguro	PARNA Pau Brasil	Nemesio (2013) Bras Jour Biol 73:375-381	as <i>E. roubiki</i>
19	<i>Euglossa iopoecila</i>	-16.88505	-39.35394	BA	Porto Seguro	PARNA Descobrimento	Nemseio (2013) Bras Jour Biol 73:437-446	as <i>E. roubiki</i> and <i>E. iopoecila</i> respectively
20	<i>Euglossa iopoecila</i>	-17.13083	-39.39401	BA	Prado	PARNA Descobrimento	Faria Jr and Melo(2007). Rev Bras Entomol 51:275–284; this study	as <i>E. iopoecila</i>

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Sinonimia
21	<i>Euglossa iopoecila</i>	-17.51194	-39.74034	BA	Teixeira de Freitas		Nemésio A (2009) Zootaxa 2041:242pp	as <i>E. roubiki</i>
22	<i>Euglossa iopoecila</i>	-18.5473	-39.73903	ES	Conceição da Barra		Faria Jr and Melo(2007). Rev Bras Entomol 51:275–285	as <i>E. iopoecila</i>
23	<i>Euglossa iopoecila</i>	-18.99192	-40.129	ES	Linhares	REBIO Sooreta	Ferrari and Melo (2014) Apidologie 45:641–652	as <i>E. iopoecila</i>
24	<i>Euglossa iopoecila</i>	-19.15199	-40.019	ES	Linhares	Reserva Vale	Ferrari and Melo (2014) Apidologie 45:641–652	as <i>E. iopoecila</i>
25	<i>Euglossa iopoecila</i>	-19.82402	-40.323	ES	Aracruz		Peruquetti et al (1999) Rev Bras Zool 16:101-118	as <i>E. chalybeata</i>
26	<i>Euglossa iopoecila</i>	-20.36151	-40.65451	ES	Domigos Martins	PNM Domingos Martins	Nemésio A (2009) Zootaxa 2041:242pp	as <i>E. roubiki</i>
27	<i>Euglossa iopoecila</i>	-20.63304	-40.55975	ES	Guarapari		Peruquetti et al (1999) Rev Bras Zool 16:101-118	as <i>E. chalybeata</i>
28	<i>Euglossa iopoecila</i>	-21.88162	-41.85283	RJ	Campos dos Goytacazes	PE Desengano	Tonhasca et al (2012) Biotropica 34:416-422	as <i>E. chalybeata</i>
29	<i>Euglossa iopoecila</i>	-22.42701	-42.04299	RJ	Rio das Ostras	REBIO União	Ferrari and Melo (2014) Apidologie 45:641–652	as <i>E. iopoecila</i>
30	<i>Euglossa iopoecila</i>	-22.59994	-42.46656	RJ	Embaú		Ramalho et al. (2009) Rev Bras Entomol 53:95-101	as <i>E. iopoecila</i>
31	<i>Euglossa iopoecila</i>	-23.20805	-44.822	RJ	Paraty	Serra da Bocaina	Ferrari and Melo (2014) Apidologie 45:641–652	as <i>E. iopoecila</i>

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Sinonimia
32	<i>Euglossa iopoecila</i>	-23.35955	-44.78769	SP	Ubatuba	PESM - Picinguaba	Rocha-Filho and Garofalo (2013) J Inse Scien 13(23)1-4;	as <i>E. chalybeata iopoecila</i> , <i>E. iopoecila</i>
33	<i>Euglossa iopoecila</i>	-23.73821	-45.70028	SP	São Sebastião	PESS-Juqueí	Penha et al (2015) Apidologie 46:346–358	as <i>E. iopoecila</i>
34	<i>Euglossa iopoecila</i>	-23.88617	-45.37501	SP	Ilhabela	PESS-Juqueí	Penha et al (2015) Apidologie 46:346–358	as <i>E. iopoecila</i>
35	<i>Euglossa iopoecila</i>	-24.36682	-47.96675	SP	Sete Barras	Fazenda Morro do Capim	Mattozo et al (2011) Pap Avulsos Zool 51:505–515	as <i>E. iopoecila</i>
36	<i>Euglossa iopoecila</i>	-24.39143	-47.12484	SP	Iguape	Estação Ecologica Juréia-Itatins	Peruquetti et al (1999) Rev Bras Zool 16:101-118	as <i>E. chalybeata iopoecila</i>
37	<i>Euglossa iopoecila</i>	-25.023	-47.95338	SP	Cananéia	PESM-cananeia	Faria Jr and Melo (2007). Rev Bras Entomol 51:275–284; Penha et al (2015) Apidologie 46:346–358	as <i>E. iopoecila</i>
38	<i>Euglossa iopoecila</i>	-25.11281	-47.99344	SP	Ilha do Cardoso	PEIC	Faria Jr and Melo (2007). Rev Bras Entomol 51:275–284; Penha et al (2015) Apidologie 46:346–359	as <i>E. iopoecila</i>
39	<i>Euglossa iopoecila</i>	-25.24447	-48.4864	PR	Guaraqueçaba		Faria Jr and Melo (2007). Rev Bras Entomol 51:275–284	as <i>E. iopoecila</i>
40	<i>Euglossa iopoecila</i>	-25.31342	-48.65716	PR	Antonina	RN Rio Cachoeira	Mattozo et al (2011) Pap Avulsos Zool 51:505–515	as <i>E. iopoecila</i>

ID	Espécie	Latitude	Longitude	UF	Cidade	Localidade	Fonte	Sinonimia
41	<i>Euglossa iopoecila</i>	-25.52196	-48.57336	PR	Paranaguá	CBSM - Paranaguá	Faria Jr and Melo (2007). Rev Bras Entomol 51:275–285	as <i>E. iopoecila</i>
42	<i>Euglossa iopoecila</i>	-25.54803	-48.63331	PR	Alexandra		Faria Jr and Melo (2007). Rev Bras Entomol 51:275–286	as <i>E. iopoecila</i>
43	<i>Euglossa iopoecila</i>	-25.81008	-48.56436	PR	Matinhos	CBSM - Matinhos	Faria Jr and Melo (2007). Rev Bras Entomol 51:275–286	as <i>E. iopoecila</i>
44	<i>Euglossa iopoecila</i>	-26.22634	-48.95324	SC	Joinville		Faria Jr and Melo (2007). Rev Bras Entomol 51:275–287	as <i>E. iopoecila</i>

Apêndice K – Distribuição das diferenças entre pares de sequências de 640 pb do gene *cytb* em 120 amostras de *Euglossa annectans*.
 *Haplótipos fixados, gráfico ilustrativo.

