

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA
MOLECULAR

LEANDRO PEREIRA REZENDE

**AS POPULAÇÕES URBANAS DE *Centris (Heterocentris) analis* Fabricius, 1804
(HYMENOPTERA, APIDAE, CENTRIDINI) SÃO GENETICAMENTE
ESTRUTURADAS?**

São Carlos – SP
2019

LEANDRO PEREIRA REZENDE

**AS POPULAÇÕES URBANAS DE *Centris (Heterocentris) analis* Fabricius, 1804
(HYMENOPTERA, APIDAE, CENTRIDINI) SÃO GENETICAMENTE
ESTRUTURADAS?**

Dissertação apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular da Universidade Federal de São Carlos como parte dos requisitos para obtenção do título de Mestre em Genética Evolutiva e Biologia Molecular, Área de Concentração: Genética Evolutiva.

Orientador: Prof. Dr. Marco Antonio Del Lama

São Carlos – SP
2019



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Leandro Pereira Rezende, realizada em 22/02/2019:

Prof. Dr. Marco Antonio Del Lama
UFSCar

Profa. Dra. Carolina de Barros Machado da Silva
UFSCar

Prof. Dr. Lago Maurício Franco
USP

A todos os maranhenses, principalmente aos pobres, negros e gays que sonham em fazer uma graduação ou pós-graduação, mas que os recursos e condições não lhes permitem realizá-la!

DEDICO

AGRADECIMENTOS

Primeiramente a Deus, por me conceder o dom da vida e benção de poder concluir meus objetivos com graça, zelo e saúde, e finalmente concluir com êxito mais uma meta acadêmica.

Ao meu orientador, Professor Dr. Marco Antonio Del Lama, em primeiro por aceitar me orientar quando eu não tinha conhecimento algum de sua linha de pesquisa e, em segundo, por me direcionar durante a pesquisa e me auxiliar na construção do conhecimento e escrita da dissertação, sempre colocando uma frase ao final do e-mail: “Vamos em frente, temos muito a fazer, e fazê-lo bem!”

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) pela concessão da bolsa na vigência do mestrado para concluir com êxito minha pesquisa, desenvolvida no Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular da Universidade Federal de São Carlos, ao qual sou muito grato por me proporcionar alcançar esta meta tão importante em minha vida.

Agradeço aos meus companheiros de Laboratório, Paulinha Gória, Diana Machado, do Laboratório de Genética Evolutiva de Himenópteros, e Cristiana Trujilu, do Laboratório de Genética de Aves, que me ajudaram bastante a aprender a rotina do laboratório, nas atividades, nos seminários, entre outros. Um agradecimento especial à minha amiga Camila Sabadini que me estendeu a mão como nenhuma outra para me guiar na pós-graduação; foi uma irmã em São Carlos. Um agradecimento especial para Fagner Miguel e Mariana Dessi que me ajudaram bastante na execução das análises, auxiliando-me na formatação de arquivos de entrada e parâmetros de programas.

Aos professores da pós-graduação que me permitiram aumentar minha bagagem de conhecimento, entender o direcionamento da minha pesquisa e ver novas possibilidades durante minha formação.

À Katia Maria Ferreira, pós-doutoranda do LGEH, que me recebeu muito bem no laboratório e me ensinou muita coisa durante minha jornada. Também à Isabel Cristina de Godói, técnica do laboratório, que foi uma pessoa incrível, principalmente na finalização do mestrado, com

muitas atividades laboratoriais sendo refeitas, e ela sempre disposta e na torcida para ocorrer tudo certo.

Ao meu irmão Gonçalo Rezende que me ajudou bastante durante minhas coletas de material biológico no Maranhão, bem como minha amiga Keliane Macedo (Imperatriz/MA), Eudilieia Vieira (Belém/PA), Camila (Marabá/PA) e Letícia Costa (Primavera do Leste/MT) que também me ofereceram estadia e me ajudaram nas expedições de coleta, e também ao Marlison que trouxe material de Cruzeiro do Sul/AC. Igualmente, aos que fizeram coleta de material que foi utilizado neste nosso trabalho. Aos meus amigos que me auxiliaram e me apoiaram, nos momentos difíceis de caminhada.

Um agradecimento especial para Cícera Matos, Beiby Rezende, Fabiana Rezende e Amanda Resende, por me fornecerem seus computadores e notebooks para eu finalizar minhas análises, quando somente o meu não foi suficiente.

À minha família, pelo apoio incondicional às minhas decisões, principalmente quando decidi largar tudo e me mudar para São Carlos/SP.

A todos os que direta ou indiretamente contribuíram para a realização e conclusão deste trabalho.

A todos, o meu muitíssimo obrigado!

“Se as abelhas desaparecerem da superfície do planeta, então ao homem restaria apenas quatro anos de vida. Com o fim das abelhas, acaba a polinização, acabam as plantas, acabam os animais, acaba o homem.”

Albert Einstein

RESUMO

A maioria das abelhas é solitária e algumas delas habitam ambientes urbanos, encontrando ali refúgio para sua sobrevivência, suprindo suas necessidades de forrageio e locais de nidificação. *Centris (Heterocentris) analis* é uma espécie com ampla distribuição, porte médio, polilética e multivoltina que nidifica em cavidades pré-existentes; as fêmeas são consideradas filopátricas e os machos o sexo dispersor. Como base nessas características, o objetivo central deste trabalho foi verificar a estrutura genética populacional e intranidal de *C. analis* e definir áreas potenciais de distribuição da espécie. Foram analisados 221 indivíduos provenientes de 21 localidades, coletados entre 2012 a 2018. Com base em registros de coleta próprios e os provenientes de consultas online, foi realizada modelagem de nicho ecológico para verificar áreas potenciais de distribuição da espécie mediante o programa Maxent implementado no ArcGis. A modelagem de nicho demonstrou áreas potenciais de distribuição desde o México na América do Norte até parte do Paraná no Brasil, incluindo países da América do Sul. Foram utilizados 16 locos microssatélites espécie-específicos, com taxa de erro de genotipagem entre 0,00 a 2,17%, ausência de desequilíbrio de ligação e de alelos nulos, para acessar a estrutura genética das populações analisadas. Os valores de F_{ST} par-a-par mostraram que Teodoro Sampaio/SP e Cruzeiro do Sul/AC foram as populações que mais se diferenciaram em relação às demais. Os índices de diversidade genética mostraram alto número de alelos (5,65) e riqueza alélica (4,58). A AMOVA para dois níveis hierárquicos demonstrou valor de $F_{ST} = 0,055$ ($p < 0,0001$) entre populações. As análises mostraram baixa estruturação populacional, sem formação de grupos genéticos geograficamente distribuídos. Os resultados do Structure modelo *admixture* mostraram que a maior probabilidade é da existência de apenas um grupo genético para fêmeas e que os machos podem ser atribuídos parcial e igualmente a três grupos genéticos em cada localidade. Os resultados do TESS não delimitaram grupos genéticos isolados. A DAPC mostrou ausência de estruturação para fêmeas e dois grupos genéticos não totalmente exclusivos nos machos. A análise de estrutura genética intranidal foi acessada com 10 locos microssatélites para 30 indivíduos de 8 ninhos. Esta análise permitiu confirmar que o sistema de acasalamento é do tipo monoginia-monandria, em que uma fêmea se acasala com um único macho durante o período de nidificação. Os resultados confirmaram parcialmente a hipótese proposta, visto que foi estimada uma baixa, porém significativa estruturação populacional, resultado esperado em razão de traços peculiares da biologia da espécie. Este trabalho contribui para o conhecimento da biologia e genética de uma espécie de abelha solitária comum em ambientes urbanos, ressaltando a relevância e a necessidade de novos estudos desta natureza em outras espécies de abelhas solitárias que ocupam as nossas cidades, com distribuição ampla ou restrita, para fins comparativos, pois a conservação destas espécies passa primeiramente pelo seu conhecimento.

Palavras-chave: Abelhas solitárias. Áreas urbanas. Estrutura genética populacional. Estrutura sociogenética intranidal. Modelagem de nicho. Variação genética.

ABSTRACT

Most bees are solitary and some inhabit urban environments, finding refuge for their survival, supply their foraging and nesting needs. *Centris (Heterocentris) analis* is a multivoltine species with broad distribution, average size and polylectic behavior that nests in preexisting cavities. Females exhibit phylopatric behavior and males are considered the dispersing sex. Based on these characteristics, the central objective of this work was to test the hypothesis that urban populations of *C. analis* are not genetically structured. Further aims were to perform niche modeling and reevaluate the intranidal sociogenetic structure. We analyzed 221 individuals from 21 localities collected between 2012 and 2018. Based on the present collection records and those obtained from online surveys, ecological niche modeling was performed to determine potential areas of species distribution using the Maxent software implemented in ArcGis. Niche modeling showed that potential distribution areas occur from Mexico in North America to part of the state of Paraná in southern Brazil, including countries in South America. Sixteen species-specific microsatellite loci, with genotyping error rate of the loci from 0.00 to 2.17%, absence of linkage disequilibrium and null alleles were used to assess the genetic structuring of the populations analyzed. The F_{ST} pairwise values showed that the Teodoro Sampaio/SP and Cruzeiro do Sul/AC populations differed most from the other populations. The genetic diversity indexes showed high number of alleles (5.65) and allelic richness (4.58). AMOVA with two hierarchical levels showed F_{ST} value = 0.055 ($p < 0.0001$) among populations. The analyses demonstrated low population structure, with no evident formation of geographically distributed genetic groups. The results of the Structure software (admixture model) show that the most probable scenario is the existence of a single genetic group for females and that males can be partially and equally assigned to three genetic groups in each locality. The TESS results did not delimit genetic groups. DAPC showed no structure for females and two not completely exclusive genetic groups in males. Intranidal genetic structure was assessed through 10 microsatellite loci in 30 individuals from eight nests. This analysis enabled confirming the monogyny-monandry mating system, in which a female mates with a single male during the nesting period. The results partially confirmed the proposed hypothesis, since low, but significant population structure was estimated, which was an expected result due to peculiar traits of the biology of the species. This work contributes knowledge on the biology and genetics of a solitary bee species that is very common in urban environments. There is a need for further studies of this nature on other species of solitary bees that occupy our cities (with broad or restricted distribution) for the purposes of comparison, since knowledge constitutes the first step toward the conservation of these species.

Keywords: Ecological niche modeling. Genic variation. Intranest sociogenetic structure. Populational genetic structure. Solitary bees. Urban areas.

LISTA DE FIGURAS

- Figura 1.** Filogenia da família Apidae por inferência bayesiana baseada na combinação de transcriptoma e elementos ultraconservados. Fonte: Adaptado de Bossert et al., 2019.....18
- Figura 2.** Indivíduos adultos de *C. analis*: Vista lateral (A) e dorsal (B) do macho de *C. analis*; Vista lateral (C) e dorsal (D) da fêmea de *C. analis*. As setas indicam a escopa presente nas pernas posteriores da fêmea. Fonte: Alonso, 2008.19
- Figura 3.** Ninho de *C. analis*. Em A, ninho aberto construído em tubo de cartolina: 1 - Célula de Cria; 2 - Célula Vestibular. Fonte: Moure-Oliveira, 2018 - Foto: José Carlos Serrano. Em B, Revestimento do ninho com óleo após o fechamento. Fonte: Hirotsu, 2013.20
- Figura 4.** Área de distribuição da espécie *C. analis* de acordo com os pontos de ocorrência.27
- Figura 5.** Procedimento de coleta de *C. analis* durante as expedições. Nesta foto foi coletado um indivíduo do gênero *Thygater* (Apidae: Eucerini). Foto: Moure-Oliveira.....31
- Figura 6.** Mapa do Brasil demonstrando as cidades e estados em que foram feitas coletas de *C. analis* entre 2012 a 2018.....32
- Figura 7.** Ninhos armadilha utilizado na coleta de *C. analis*. Em A, o bloco de madeira com perfurações usado para alocar os ninhos-armadilha. Em B, ninhos-armadilha fabricados em colorset preto. Fonte. Silva, 2010.....33
- Figura 8.** Correlação entre variáveis ambientais. a) Correlação entre temperatura mínima do mês mais frio e temperatura máxima do mês mais quente; b) correlação entre temperatura máxima do mês mais quente e precipitação do mês mais seco; c) correlação entre precipitação do mês mais seco e temperatura mínima do mês mais frio.....43
- Figura 9.** Área potencial de distribuição da espécie *C. analis*. a) Modelo 1, pontos de ocorrência do Laboratório de Genética Evolutiva de Himenópteros; b) Modelo 2, pontos de ocorrência coletados online; c) Modelo 3, pontos totais de ocorrência da espécie. Legenda: o vermelho representa maior probabilidade da ocorrência da espécie na região; à medida que se vai na direção do branco, esta probabilidade diminui.....45
- Figura 10.** Padrões de amplificação dos 16 locos microssatélites espécie-específicos de *C. analis*.46
- Figura 11.** Análise de variância molecular (AMOVA) implementada no programa GenAlEx para 21 amostras urbanas de *C. analis* para verificar a distribuição da variação genética dentro e entre populações.51
- Figura 12.** Resultado da análise do programa Structure *no admixture* para machos de *C. analis*. Em A, pode-se observar a probabilidade de $\ln(P(X/K))$ de $K=3$. Em B, observa-se o resultado do teste de Evanno com o valor de $K=3$. Em C, estruturação sugerida para machos em $K = 3$53
- Figura 13.** Resultado da análise do programa Structure *no admixture* para fêmeas de *C. analis*. Em A, pode-se observar a probabilidade de $\ln(P(X/K))$ descrita por Pritchard et al.,

(2000). Em B, observa-se o resultado do teste de Evanno com o valor de K mais provável. Em C, Estruturação sugerida em K = 2 e K=4.....54

Figura 14. Gráfico de barras de estruturação genética. Em A, resultado da análise do programa Structure modelo *admixture* para fêmeas (K=2). Em B, a análise do programa Structure modelo *admixture* para machos (K=3)55

Figura 15. Resultado da análise do programa Structure *no admixture* para machos e fêmeas de *C. analis*. Em A, pode-se observar a probabilidade de Ln (P(X/K)) descrita por Pritchard et al., (2000). Em B, observa-se o resultado do teste de Evanno com o valor de K=2. Em C, estruturação sugerida em K = 2.....56

Figura 16. Resultado da análise do programa Structure *admixture* para machos e fêmeas de *C. analis* com K=2.....56

Figura 17. Modelos testados no programa TESS para verificar a estrutura genética de *C. analis* em fêmeas. Foram testados 9 modelos: *Without admixture* (NoADM), *With mixture* BYM Constante (BYMCon), *With mixture* BYM Linear (BYMLin), *With mixture* BYM Quadrático (BYMQua), *With mixture* BYM Cúbico (BYMCub), *With mixture* CAR Constante (CARCon), *With mixture* CAR Linear (CARLin), *With mixture* CAR Quadrático (CARQua) e *With mixture* CAR Cúbico (CARCub).57

Figura 18. Gráfico de barras da estruturação genética (K=5) detectada pelo programa TESS modelo *With admixture* CAR Cúbico para as fêmeas.58

Figura 19. Modelos testados no programa TESS para verificar a estrutura genética de *C. analis* em machos. Foram testados 9 modelos: *Without admixture* (NoADM), *With mixture* BYM Constante (BYMCon), *With mixture* BYM Linear (BYMLin), *With mixture* BYM Quadrático (BYMQua), *With mixture* BYM Cúbico (BYMCub), *With mixture* CAR Constante (CARCon), *With mixture* CAR Linear (CARLin), *With mixture* CAR Quadrático (CARQua) e *With mixture* CAR Cúbico (CARCub).58

Figura 20. Gráfico de barras de estruturação genética (K=6) detectados pelo programa TESS modelo *With admixture* BYM Constante para machos.58

Figura 21. Valor de BIC versus número de clusters. Em A, fêmeas. Em B machos. Em C, machos e fêmeas.....59

Figura 22. Análise Discriminante de Componentes Principais para machos. Em A, DAPC considerando K=2. Em B, distribuição dos indivíduos em dois grupos representados em laranja e azul.....60

Figura 23. Análise Discriminante de Componentes Principais para machos e fêmeas. Em A, gráfico de dispersão dos grupos genéticos (K=3). Em B, distribuição dos indivíduos em cada grupo genético representado pelas cores amarela, azul e cinza.....61

LISTA DE TABELAS

Tabela 1. Sítios analisados listados por bioma, estado, localidade, coordenadas geográficas, indivíduos analisados por sexo e ano de coleta.	34
Tabela 2. Descrição dos primers desenvolvidos para a espécie <i>C. analis</i> mediante NGS, 2018. ID: Identificação dos primers, F: Forward; R: Reverse; TM: Temperatura de <i>Melting</i> ; TF: Tamanho do fragmento em pares de bases; N° de rep: número de repetições; T-PCR: Temperatura da Reação em Cadeia da Polimerase; N° ciclos: número de ciclos da PCR.	36
Tabela 3. Relação das amostras utilizadas para cálculo da taxa de erro de genotipagem por loco microssatélite.	38
Tabela 4. Valores de Componentes Principais (PCs) referente à análise de contribuição de variação das camadas ambientais usadas na modelagem de nicho ecológico de <i>C. analis</i>	41
Tabela 5: Descrição da taxa de erro de genotipagem por loco.	47
Tabela 6. Características dos 16 locos microssatélites amplificados para indivíduos de <i>C. analis</i> . N: número de alelos; AR: Riqueza alélica, para machos e fêmeas; Ho: heterozigosidade observada; He: heterozigosidade esperada para fêmeas; F_{ST} : Índice de Fixação (Weir e Cockerham, 1984).	48
Tabela 7. Índices de diversidade genética para as populações analisadas quanto a: Número médio de alelos (N), Riqueza Alélica (AR), Número de Alelos Privados (N_{AP}), Heterozigosidade média observada (Ho) e esperada (He) em fêmeas, Índice de Garza-Williamson (M), Coeficiente de endogamia (F_{IS}) e seu respectivo valor de P (5%), estimados com dados das fêmeas.	49
Tabela 8. Análise de variância molecular (AMOVA) implementada no programa Arlequin para verificar a distribuição da variação genética dentro e entre populações.	51
Tabela 9. Resultado da AMOVA para fêmeas, machos e machos com $K=2$ (grupos definidos pela DAPC) para dois níveis hierárquicos, dentro e entre populações.	62
Tabela 10. Definição da genética familiar para um ninho (Ninho 1) com utilização de 10 locos microssatélites.	63

SUMÁRIO

1	APRESENTAÇÃO DO PROBLEMA	14
1.1	Filogenia de Centridini	16
1.2	A espécie em estudo.....	18
1.3	Modelagem de Nicho Ecológico	21
1.4	A genética de populações e a estrutura sociogenética intranidal da espécie	22
1.5	Marcadores microssatélites	24
1.6	Justificativa e Relevância	25
2	OBJETIVOS	28
3	MATERIAL E MÉTODOS	28
3.1	Modelagem da distribuição potencial.....	28
3.2	Procedimentos de coleta	30
3.2.1	Material e locais de coleta.....	31
3.2.2	Ninhos-armadilha.....	32
3.3	Métodos	33
3.3.1	Extração de DNA.....	33
3.3.2	Prospecção dos locos microssatélites	35
3.3.3	Análise dos locos microssatélites.....	35
3.3.4	Erros de genotipagem e caracterização dos locos microssatélites	37
3.3.5	Análise de diversidade genética e estrutura populacional	38
3.3.6	Análise da estrutura sociogenética intranidal	40
4	RESULTADOS	40
4.1	Modelagem de nicho ecológico	40
4.2	Caracterização dos locos microssatélites	46
4.3	Diversidade genética populacional	48
4.4	Estrutura genética populacional.....	51
4.5	Estrutura sociogenética intranidal	63
5	DISCUSSÃO	64
5.1	Área potencial de distribuição de <i>Centris (Heterocentris) analis</i>	64
5.2	Diversidade e estrutura genética populacional.....	66
5.3	Estrutura sociogenética intranidal	70
6	CONSIDERAÇÕES FINAIS	72
	REFERÊNCIAS	74
	APÊNDICES	83

1 APRESENTAÇÃO DO PROBLEMA

O que é "fragmentação" da perspectiva de um inseto? A fragmentação descreve uma distribuição irregular de habitats adequados, às vezes considerados como "ilhas ecológicas", cercados por uma matriz de habitats inóspitos ou inadequados de permeabilidade variável (Ricketts, 2004). Fragmentação significa mais do que a mera existência de habitats isolados ou fragmentados; implica também que um habitat contínuo foi subdividido ou fragmentado por algum processo, frequentemente antropogênico, com a consequente perda de habitat interveniente. As distribuições e a dinâmica populacional de espécies de abelhas em habitats naturalmente irregulares (por exemplo, abelhas em prados alpinos) podem fornecer *insights* sobre alguns, mas não todos, os fatores associados à fragmentação do habitat (Bowers, 1985).

A fragmentação de habitats comumente ocorre durante a formação de ecossistemas urbanos caracterizados pela alta densidade da população humana, heterogeneidade espacial, diversos tipos de poluição, como visual, sonora, aquática e atmosférica, alto número de espécies exóticas, principalmente na flora, e mudanças climáticas locais (McIntyre et al., 2001). Entre as principais consequências da fragmentação, destaca-se a perda (ou limitação) do potencial de dispersão e colonização das espécie, bem como sua capacidade de forrageamento. No entanto, apesar da influência negativa da urbanização na biodiversidade, as cidades vêm demonstrando ser habitats adequados para alguns táxons devido à sua heterogeneidade espacial, clima quente e presença de espécies exóticas (Ahrné et al., 2009), abrigando muitas espécies vegetais e animais, presentes em áreas nativas periféricas. Dentre os animais que habitam populações urbanas, os insetos possuem notável destaque (Moure-Oliveira, 2013).

Nesse sentido, ecólogos de insetos polinizadores estão voltando sua atenção para as cidades, em especial por dois motivos: 1) os centros urbanos poderiam estar atuando como refúgios de algumas das ameaças enfrentadas pelos polinizadores em paisagens agrícolas, tais como o uso de agrotóxicos, pois embora a urbanização seja uma causa importante de perda de habitat, a maioria das ameaças aos polinizadores parece estar ocorrendo principalmente em terras não urbanas; 2) a maioria das pessoas vive nas cidades, e os polinizadores, especialmente abelhas e borboletas, são emblemas carismáticos da biodiversidade presentes em parques e jardins (Goddard, 2016).

Pesquisas sobre insetos polinizadores urbanos estão mudando a forma como o valor biológico e a importância ecológica das cidades são vistos. A conservação da biodiversidade deve ser reposicionada dentro dessa imagem da cidade. Em vez de tratar a conservação urbana

apenas como extensão e educação destinadas a melhorar o capital político, as agências de Manejo de Recursos Naturais (NRM) podem desenvolver programas que melhorem o capital natural, envolvendo cidadãos urbanos na melhoria da qualidade de vida das espécies ameaçadas nas cidades (Hall et al., 2016). A abordagem para melhorar o valor de habitat dentro das áreas urbanas é relativamente simples e facilmente compreendida pelos residentes urbanos. Várias análises e metanálises de insetos polinizadores urbanos mostram que a variável consistente correlacionada com a saúde dos polinizadores é o forrageamento (ou seja, a presença de flores) (Hennig e Ghazoul, 2012; Cariveau e Winfree, 2015).

A necessidade de se estudar a fauna de insetos em áreas urbanas, incluindo as abelhas, pode ser justificada por cinco motivos: (1) alta diversidade, mostrando a biodiversidade local; (2) o curto tempo de geração mostra uma resposta rápida às mudanças causadas pelas ações antropogênicas no solo e vegetação; (3) a facilidade na coleta de material; (4) representatividade do acervo de níveis tróficos animais; (5) são importantes componentes ecológicos, econômicos e sociológicos dos ambientes antropizados (McIntyre, 2000).

Vários insetos polinizadores, bem como as abelhas, contribuem para o sucesso de jardinagem urbana nos vários centros de cidades (Matteson e Langellotto, 2009). Essas espécies podem colonizar áreas urbanas com uma cobertura relativamente pequena de espaços verdes (Hernandez et al., 2009; Matteson e Langellotto, 2010) e, em termos de serviços ecossistêmicos, o sucesso e a utilidade delas como fornecedores de um "serviço" podem depender da dinâmica sócio ecológica local (Andersson et al., 2007).

A polinização como um serviço ecossistêmico em áreas urbanas pode depender do gerenciamento do espaço verde. A fim de aumentar a eficiência da polinização em jardins, é essencial projetar espaços verdes de uma forma que atraia os insetos polinizadores. A alta riqueza de espécies de plantas com flores e um período de crescimento prolongado podem ser fatores cruciais. Jardins urbanos e canteiros de flores são dois tipos de habitats que podem promover e possivelmente melhorar as populações de abelhas nos centros das cidades. No entanto, não é necessariamente o tamanho do espaço verde que determina a riqueza de espécies de polinizadores. Na cidade de Nova York, a luz solar e a abundância floral foram identificadas como os principais determinantes da diversidade urbana de abelhas (Matteson e Langellotto, 2010). O serviço de polinização fornecido por abelhas solitárias é substancial. O valor do serviço de polinização global fornecido por abelhas silvestres foi estimado em US \$ 3.251/ha, semelhante à contribuição das abelhas manejadas (abelhas exploradas

comercialmente), e sete das dez espécies de abelhas silvestres identificadas como fornecendo a maior contribuição média para a produção agrícola eram solitárias (Kleijn et al., 2015).

Nesse sentido sabendo da importância ecológica das abelhas dentro dos ecossistemas urbanos e naturais torna importante estudos que auxiliam na conservação da espécie. Estes estudos envolvem o conhecimento de sua biologia e genética. Os estudos genéticos em consolidação com ecologia fornecem base para o manejo dessas espécies e desenvolver trabalhos e programas voltados para conservação.

No Brasil, poucos são os estudos que reportam a variabilidade genética e os níveis de diferenciação entre as populações de abelhas, principalmente, de espécies solitárias, que constituem a maior parte da fauna apícola brasileira e desempenham papel tão relevante quando o das espécies sociais na funcionalidade do ecossistema. Essa limitação de conhecimento, principalmente sobre as abelhas solitárias pode ser justificada por suas populações frequentemente serem pequenas quando comparadas a outros insetos (Danks, 1971), aliada à dificuldade de se localizar seus sítios de nidificação (Jayasingh e Freeman, 1980); Ademais, a maioria dos trabalhos de nidificação em espécies solitárias é realizada com a utilização de ninhos armadilhas.

Dessa forma, é importante conhecer a biologia e genética destas espécies polinizadoras que tem os centros urbanos como refúgios dos habitats fragmentados. Este trabalho foi delineado para determinar a estrutura genética de uma espécie de abelha solitária da tribo Centridini apresentada a seguir. A introdução deste trabalho foi estruturada de modo a apresentar a filogenia da espécie, suas características biológicas, os estudos genéticos realizados até o momento com espécies relacionadas e com a própria espécie, o uso dos marcadores moleculares como ferramenta para se conhecer a estrutura genética das populações, finalizando esta seção com a justificativa e relevância da realização deste trabalho.

1.1 Filogenia de Centridini

O Reino Animal possui uma ampla variedade de organismos. Destes, o maior número de indivíduos está alocado no filo Arthropoda, que agrupa cerca de 82% de todos os animais (Hickman, Roberts e Larson 2011). Chama ainda a atenção a Classe Insecta, por apresentar elevado número de espécies descritas. Nesta classe, a ordem Hymenoptera abrange mais de 153.088 espécies existentes e descritas, além de 2.429 espécies descritas extintas (Aguilar et al., 2013). A maioria das espécies apresenta determinação do sexo por haplodiploidia arrenótoca, em que os machos são haploides e as fêmeas são diploides

(Wilgenburg et al. 2006). Com este total de espécies relatadas, os Himenópteros estão entre as quatro ordens de insetos megadiversos (Grimaldi e Engel, 2005), com estimativas de até um milhão de espécies não descritas, englobando sínfitas, vespas, formigas e abelhas (Grimaldi e Engel, 2005; Aguiar et al., 2013), atrás apenas de Coleoptera e Lepidoptera (Aguiar et al., 2013)

Os táxons desse grupo são divididos em duas subordens (Symphyta e Apocrita) com 27 superfamílias (9 superfamílias em Symphyta e 18 em Apocrita) e 132 famílias (Aguiar et al., 2013). Estes autores ainda relatam que 17 dessas famílias foram descritas desde o ano 2000 e que 16 dessas famílias recém-reconhecidas foram descritas a partir de fosséis e se encontram extintas. Novos gêneros e espécies são frequentemente encontrados e a estimativa mais recente totaliza 8.423 gêneros existentes, com um adicional de 685 gêneros extintos. Os indivíduos dessa ordem são observados em quase todos os ambientes, podendo ser capturados durante o voo, locomovendo-se ou pousados sobre plantas, em ninhos próprios ou de outros insetos, em cavidades pré-existentes, no solo e até mesmo em serrapilheira (Teixeira, 2012).

O grupo Aculeata é formado pelas superfamílias Chrysidoidea, Vespoidea e Apoidea, sendo esta última caracterizada pela constrição abdominal e ausência de pernas na fase larval (Melo, 1999 e Brothers, 1999). As abelhas constituem a superfamília Apoidea, que compreende sete famílias – Megachilidae, Apidae, Stenotritidae, Colletidae, Andrenidae, Melittidae e Halictidae. A família Apidae está subdividida em três subfamílias - Xylocopinae, Nomadinae e Apinae, esta última com 19 tribos incluindo Centridini (Michener, 2007). Segundo Martins et al., (2016) a subfamília Apinae pode ser formada por três principais clados, confirmando estudos prévios: i) Cleptoclado, formado pelos grupos que compõem os grupos "Nomadinae" e outros cleptoparasíticos, juntamente com Anthophorini; ii) Eucerini + Xylocopini, formada por Exomalopsini, Tapinotaspidini, Ancylini, Eucerini (não cleptoparasítica) e Xylocopini *lato sensu*; e iii) Apine, formada por *Epicharis*, *Centris* e as tribos corbiculadas Apini, Meliponini, Bombini e Euglossini (Figura 1).

A tribo Centridini ocorre na região Neotropical e parte da região Neártica, distribuindo-se do Sul dos Estados Unidos até a Argentina em diferentes fitofisionomias, tais como, desertos, florestas tropicais, savanas e caatinga. Esta tribo agrupa dois gêneros – *Centris* Fabricius, 1804 e *Epicharis* Klug, 1807 - 21 subgêneros e, aproximadamente, 250 espécies (Michener 2007; Moure et al., 2012). Martins e Melo (2016), no entanto, ao avaliar a filogenia molecular e a história biogeográfica de Centridini ressaltam que as evidências atuais não fornecem suporte para uma tribo monofilética composta pelos gêneros *Centris* e *Epicharis*, estes sim fortemente suportados como grupos monofiléticos, sugerindo desta forma

que Centridini seja restrito ao gênero *Centris* e seus subgêneros e o nome Epicharitini deve ser designado pela tribo constituída pelo gênero *Epicharis*.

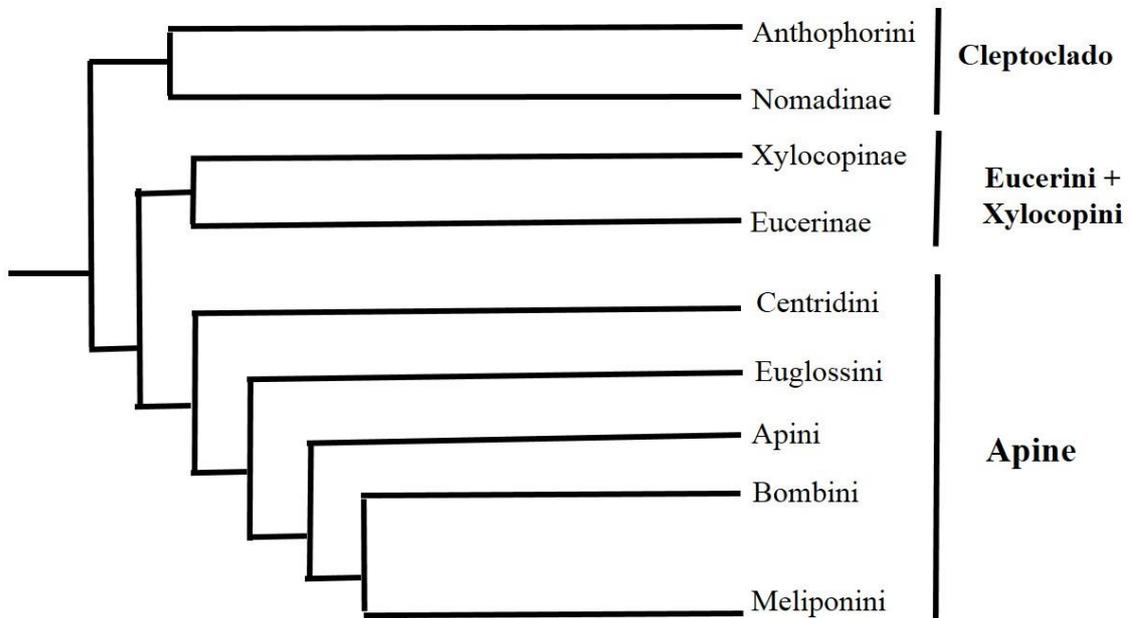


Figura 1. Filogenia da família Apidae por inferência bayesiana baseada na combinação de transcriptoma e elementos ultraconservados. Fonte: Adaptado de Bossert et al., 2019.

As abelhas da tribo Centridini são solitárias e apresentam adaptações morfológicas e comportamentais que contribuem para a construção dos ninhos e auxiliam na coleta de óleos florais para o provisionamento larval. Todas as espécies de *Epicharis* e a maioria das espécies de *Centris* constroem seus ninhos no solo, enquanto que as espécies de *Centris* dos subgêneros *Hemisiella*, *Heterocentris* e *Xanthemisia* nidificam em cavidades pré-existentes em madeiras, termiteiros e outros locais escavados por pequenos vertebrados (Silveira et al. 2002; Alves-dos-Santos et al., 2007).

1.2 A espécie de estudo

A espécie *Centris (Heterocentris) analis*, Fabricius 1804 (Figura 2) é uma abelha solitária, com distribuição desde o Texas, nos EUA (Silveira et al., 2002), até o Brasil, nos estados de Ceará, Goiás, Mato Grosso, Pará e São Paulo. Ela ocorre ainda em países da América do Sul, na Costa Rica, Guiana Francesa, Guiana, Guatemala, Honduras, Panamá, México, Venezuela e Trindade-Tobago (Moure, 1960). Assim como outras espécies do mesmo subgênero, nidificam em cavidades preexistentes, podendo também nidificar em ninhos abandonados e em ninhos-armadilha (Jesus e Garófalo, 2000; Zanella e Martins, 2003; Alves-dos-Santos et al., 2007).

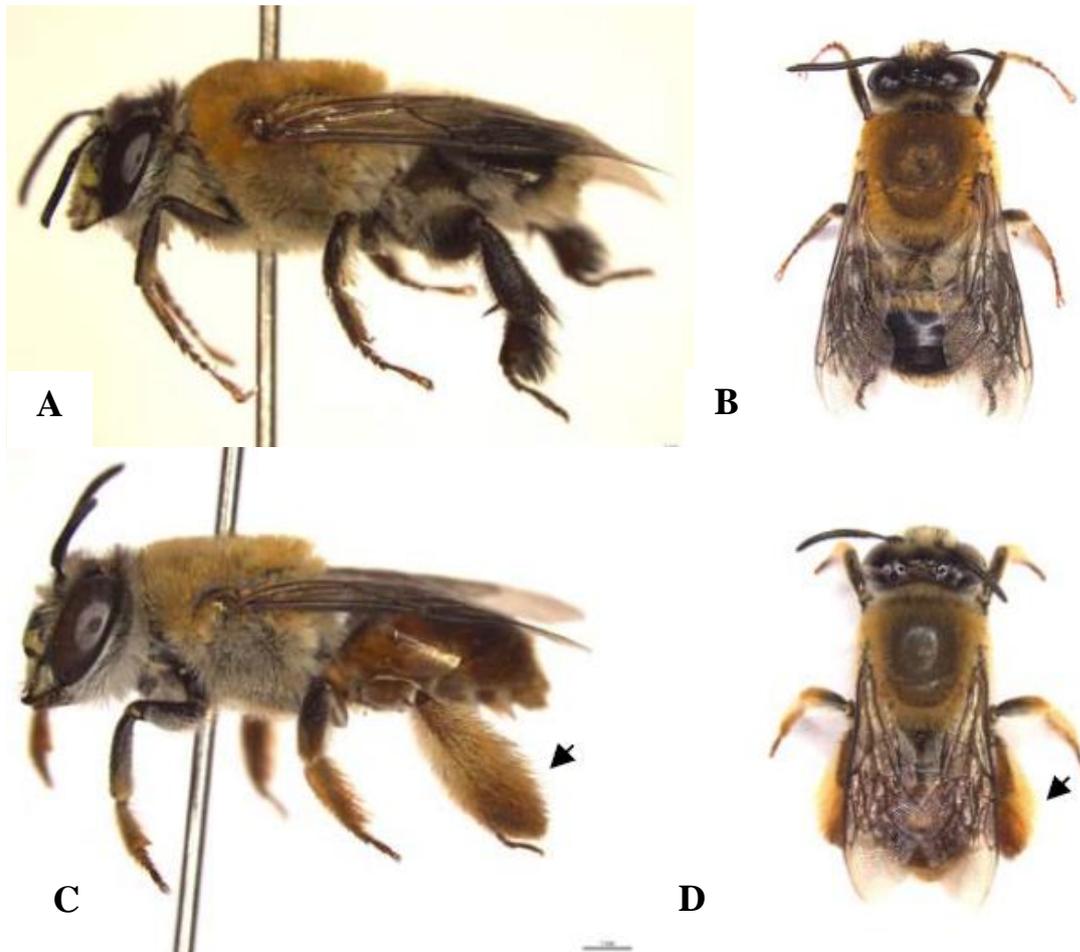


Figura 2. Indivíduos adultos de *C. analis*: Vista lateral (A) e dorsal (B) do macho de *C. analis*; Vista lateral (C) e dorsal (D) da fêmea de *C. analis*. As setas indicam a escopa presente nas pernas posteriores da fêmea. Fonte: Alonso, 2008.

Quanto ao local para nidificação, as fêmeas desta espécie procuram cavidades com tamanhos e circunferências adequados, entrando e saindo de várias delas e, após selecionar uma delas, começam a trazer material vegetal para a construção do ninho. Nesta tarefa, o número de voos pode variar de 8 a 15, com duração de 1 a 38 minutos (Hirotsu, 2013). Em relação à captação de recursos, *C. analis* utiliza subprodutos vegetais para construção de ninho e alimentação da cria; a amplitude de seu nicho ecológico demonstra sua adaptação para a coleta de pólen e óleos. Nesse sentido, a espécie é denominada polilética quanto à coleta de néctar (Rabelo et al., 2009). As células são aprovisionadas primeiramente com pólen, em seguida com néctar, depois realizada a oviposição e fechamento da célula que será a partição que servirá de fundo para a próxima célula. Terminado o ninho, é realizado o fechamento com a célula vestibular e revestimento com óleo (Figura 3) (Hirotsu, 2013).

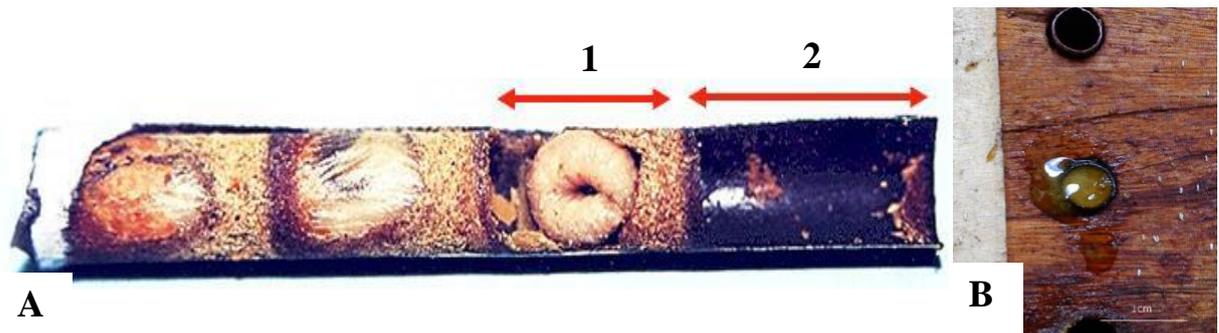


Figura 3. Ninho de *C. analis*. Em A, ninho aberto construído em tubo de cartolina: 1 - Célula de Cria; 2 - Célula Vestibular. Fonte: Moure-Oliveira, 2018 - Foto: José Carlos Serrano. Em B, Revestimento do ninho com óleo após o fechamento. Fonte: Hirotsu, 2013.

Com relação ao período de nidificação, Silva e colaboradores (2017) relatam que as frequências de nidificação de *C. analis* não diferiram entre as estações, mas a população apresentou variação temporal significativa nesse parâmetro, com a ocorrência de dois picos de atividade, sendo um deles ao final da estação chuvosa e início da estação seca (março a maio) e o outro no meio da estação chuvosa (dezembro). O período de máxima atividade reprodutiva não foi associado ao número de espécies vegetais em floração ou mesmo à quantidade de plantas de uma mesma espécie. O número de abelhas visitantes de flores não foi associado ainda à frequência de nidificação, sugerindo que fatores intrínsecos da espécie, como fatores genéticos e cronobiológicos, também podem afetar fortemente o ciclo reprodutivo.

Ainda segundo o estudo de Silva e colaboradores (2017), as fêmeas usaram 12 tipos de pólen de sete famílias botânicas (Malthighiaceae, Fabaceae, Bignoniaceae, Solanaceae, Arecaceae, Polygalaceae e Sapindaceae) para alimentar suas larvas durante o período de nidificação em dois locais de Salvador/Bahia, nos meses de maio, outubro a dezembro (2014), e entre janeiro e abril (2015). O trabalho de Lima et al., (2017) desenvolvido com ninhos de um fragmento florestal da cidade de Salvador/BA, relata que os tipos polínicos mais representativos nos ninhos foram provenientes de *Malthighiacea emarginata*, *Byrsonima sericea*, *Stigmaphyllon cavernulosum* e *Aeschynomene paucifolia*, que indicam a importância dessas espécies como fontes de alimento para abelhas imaturas.

Além disso, estas abelhas usam o óleo produzido por estas plantas para a construção e selagem do ninho (Jesus e Garófalo, 2000). Observações realizadas (Alves-dos-Santos et al., 2007; Couto e Camillo, 2014) revelaram que esta espécie utiliza o óleo coletado para o fechamento da última célula de cria. Essa substância parece conferir certa proteção aos ninhos fechados contra predadores ou parasitoides (Figura 3B) (Gazola, 1999; Pereira et al., 1999; Jesus e Garófalo, 2000).

A razão sexual desta espécie foi estudada por vários pesquisadores, com diferentes achados, haja vista que estudos realizados utilizando ninhos-armadilhas demonstraram que o tamanho deste pode influenciar a razão sexual. Neste sentido, Jesus e Garófalo (2000) encontraram uma proporção de 48% de fêmeas e 52% de machos, ou seja, uma razão de 1:1 em 89 ninhos analisados, enquanto Alonso e colaboradores (2012) verificaram uma proporção de 63% de machos e 37% de fêmeas, diferindo significativamente de 1:1, entre 440 indivíduos emergidos e/ou mortos. Porém, estes autores relataram que, à medida que o tamanho do ninho aumentava, a proporção sexual tendia a se aproximar de 1:1, fato corroborado no estudo de Moure-Oliveira e colaboradores (2017) em que uma razão sexual com viés para machos só foi estatisticamente diferente em cavidades menores. Estes achados revelam que a razão sexual da espécie pode variar, dependendo do tamanho do ninho construído pela fêmea.

Durante o processo de nidificação foi verificado por Jesus e Garófalo (2000) que houve uma assimetria na alocação sexual para *C. analis*, pois para a produção de fêmeas é necessária a construção de células de cria maiores, mais tempo para construção e mais viagens de forrageamento para alocação de recursos nas células. Segundo aqueles autores, o tamanho das fêmeas (3,54mm; $\pm 0,15$ mm) foi significativamente maior do que o de machos (3,21mm; $\pm 0,12$ mm). Esses dados são corroborados com o estudo de Moure-Oliveira e colaboradores (2017) que analisaram o tempo de construção das células de cria e observaram que o tempo médio para provisionamento da célula de cria de machos era de $0,94 \pm 0,91$ dias e de fêmeas $1,94 \pm 1,45$ dias.

A espécie em estudo, possui uma ampla distribuição conforme já mencionado, contudo, os catálogos existentes sobre sua área de distribuição total estão desatualizados e não demonstram todos os pontos de coleta já registrados. Dessa forma, a realização de uma modelagem de nicho poderia vir a demonstrar possíveis áreas de coleta para aumentar ainda mais o acervo biológico da espécie.

1.3 Modelagem de nicho ecológico

Atlas de distribuição fornece uma representação conveniente de distribuições de espécies em células de grade de área quase igual em mapas. Tais representações têm sido amplamente utilizadas para estudar distribuição de espécies, padrões de abundância, raridade, riqueza, e composição em diferentes escalas espaciais (Araújo, 2003). No entanto, existe a preocupação de que a escala da maioria dos atlas seja muito ampla para aplicações de planejamento de conservação do mundo real (Rouget, 2003).

Uma alternativa é usar técnicas de modelagem que explorem a correlação entre uma variável de resposta (por exemplo, registros de ocorrência de espécies) e conjuntos de variáveis ambientais preditoras (por exemplo, clima e cobertura da terra) para distribuições de espécies para outra resolução espacial (Barbosa et al., 2003; Pearson et al., 2004).

A modelagem ecológica de nicho representa uma aproximação do nicho ecológico de uma espécie nas dimensões ambientais examinadas (Phillips et al., 2006). Por definição, as condições ambientais nas localidades de ocorrência constituem amostras do nicho realizado. Um modelo baseado em nichos representa, assim, uma aproximação do nicho utilizado pelas espécies, na área de estudo e dimensões ambientais consideradas (Phillips et al., 2006).

A modelagem dos requisitos de habitat das espécies é uma ferramenta cada vez mais importante, tanto para investigar as necessidades das espécies quanto para planejar reservas de conservação (Jarvis e Robertson, 1999) e para a compreensão de padrões de biodiversidade (Williams e Hero, 2001). Nesse sentido, o uso dos dados existentes e a modelagem de nicho da espécie para orientar os esforços de expedições de coleta reduz o custo dos levantamentos a campo (Fielding e Bell, 1997).

Dessa forma, a elaboração de uma modelagem de nicho pode auxiliar no conhecimento das áreas potenciais de distribuição da espécie, para ampliação de coletas em áreas não amostradas para estudos futuros. O estabelecimento de associação entre variáveis ambientais e registros de ocorrência da espécie alvo identifica condições ambientais nas quais as populações têm maiores chances de se manterem (Dalapicolla, 2016). A ferramenta permite estimar a distribuição espacial do ambiente que é favorável a uma determinada espécie em uma determinada área de estudo, ou seja, verificar as potenciais áreas de ocorrência da espécie estudada (Dalapicolla, 2016).

Sabendo disso, além da área total de distribuição da espécie precisar ser revista e atualizada, os estudos genéticos da espécie não foram finalizados. Nesse sentido pode-se verificar abaixo os estudos de genética populacional realizados com abelhas solitárias, na qual, são poucos e carecendo de novos estudos para ampliar o conhecimento na área.

1.4 A genética de populações e a estrutura sociogenética intranidal da espécie

Em se tratando de genética de populações, estudos relacionados à estrutura genética das populações de abelhas vêm sendo desenvolvidos para averiguar como elas estão organizadas em sua área de distribuição, assim como determinar o grau de fluxo gênico entre estas populações. Contudo, poucos estudos sobre a temática foram realizados com abelhas solitárias e, para a espécie em estudo, verifica-se uma insuficiência de dados, visto que o

único trabalho sobre estrutura genética populacional de *C. analis* utilizando genes mitocondriais foi inconclusivo (Moure-Oliveira, 2013). Os marcadores mitocondriais utilizados neste trabalho - as subunidades I e II do complexo da citocromo c oxidase (COI e COII), citocromo b (*cytb*), a subunidade maior do RNA ribossômico (16S) e o gene do RNA transportador de Leucina (RNAt^{Leu}) não foram informativos devido a presença de picos duplos. Segundo o autor, a presença de heteroplasmia impossibilitou a utilização das sequências.

Devido à presença de heteroplasmia, uma alternativa para o estudo populacional é utilizar marcadores nucleares. Neste caso, é esperado que os resultados a serem obtidos sejam semelhantes aos obtidos por Luna-Lucena (2015). Esta autora, analisando 185 indivíduos provenientes de 16 localidades, sendo duas da Paraíba, uma de Minas Gerais, uma da Bahia e 12 de São Paulo, demonstrou ausência de estrutura genética populacional utilizando genes mitocondriais. Ao mesmo tempo, a autora apontou para os possíveis fatores responsáveis pela ausência de estruturação: i) alto fluxo gênico entre as populações; ii) número elevado de fêmeas por áreas de nidificação; iii) trata-se de uma espécie generalista quanto à dieta; iv) plasticidade quanto ao local de nidificação; v) porte corporal favorável para a dispersão. Com base nessas razões, espera-se que a espécie em estudo também apresente ausência de estrutura genética por ser uma abelha solitária da mesma tribo (Centridini), com características semelhantes.

Além dos indícios apresentados acima, Moure-Oliveira (2013) relata que a haplodiploidia e o mecanismo de determinação do sexo, compartilhado pela maioria dos himenópteros, limitam o tamanho efetivo populacional e, conseqüentemente, a variabilidade genética populacional, afetando as abelhas de comportamento solitário e eussociais. A presença da haplodiploidia é uma das prováveis causas que aumenta a taxa de fixação de alelos, dificulta a obtenção de um polimorfismo estável e aumenta a ligação gênica e o “efeito carona” devido aos menores níveis de recombinação (Crozier, 1970; Pamilo e Crozier, 1981; Owen, 1985).

Além da haplodiploidia, o sexo é também definido pelo mecanismo complementar do sexo, como a já bem estabelecida ação do gene *csd* de *Apis mellifera*, em que fêmeas podem controlar a razão sexual de sua prole (Wilgenburg et al., 2006). Este gene quando em heterozigose determina a produção de fêmeas, quando em hemizigose leva ao desenvolvimento de machos, e em homozigose serão estabelecidos machos diploides (Wilgenburg et al., 2006; Heimpel e Boer, 2008; Hasselmann et al., 2008). Esse mecanismo de formação de machos diploides pode reduzir o tamanho efetivo populacional que

influenciará na diminuição de variabilidade e nenhum valor adaptativo para as populações por formar indivíduos inviáveis (Zayed e Packer, 2005).

Algumas fêmeas de Centridini nidificam próximo ao ninho de origem, denominando-se de filopátricas, tais como *Centris flavifrons* (Martins et al., 2014) e *C. analis* (Silva, 2010; Hirotsu, 2013). Esse comportamento reduz o fluxo gênico materno, e o macho promove o fluxo gênico entre as populações durante sua dispersão. Para tanto, como afirma Luna-Lucena et al., (2017), esse comportamento de nidificação está atrelado ao fato de que várias fêmeas nidificam em um determinado local, aumentando assim a composição gênica desta população, e, além disso, como os machos são os principais dispersores, eles acabam levando genes de um local a outro, reduzindo assim as diferenças entre as populações. Neste sentido, passa a ser interessante a análise dos dados genéticos em machos e fêmeas separadamente tornando a análise mais robusta, pois machos e fêmeas podem apresentar resultados diferentes.

Além da estrutura genética populacional, estudos genéticos intranidais para *C. analis* foram realizados por Silva (2010) com 46 ninhos obtidos por ninhos-armadilha em Ribeirão Preto/SP, na qual, utilizou alozimas e não encontrou variação suficiente para diferenciação e identificação do sistema de acasalamento. Realizou também o sequenciamento de um gene mitocondrial COI para determinar as linhagens maternas. O fragmento com 465 pares de bases (pb) revelou uma única sequência para este gene. Conforme registrado pela autora em seu trabalho, a ocorrência de um único haplótipo indica que a área foi colonizada por uma linhagem materna, dada a herança uniparental deste marcador e, portanto, os dados parecem corroborar o comportamento filopátrico das fêmeas de *C. analis*. Vale ressaltar que para resultados mais consistentes, a combinação de ecologia e genética ao usar ninhos monitorados para o comportamento de nidificação, auxiliam na robustez dos dados, pois Moure-Oliveira e colaboradores (2017) afirmaram existir usurpação de ninhos entre as fêmeas de *C. analis*. Dessa forma, a detecção de usurpação de ninho utilizando um marcador com baixa variação fica prejudicada.

1.5 Marcadores microssatélites

Os microssatélites, sequências simples repetidas (SSR), repetições curtas em tandem (STR) e polimorfismos de comprimento de sequência simples (SSLP) são encontrados em procaríotos e eucariotos (Vieira et al., 2016). Correspondem a sequências de 2 a 6 pares de bases repetidos em tandem, com um alto grau de polimorfismo, distribuídos por todo o genoma e seus alelos têm “ação codominante” (Estoup et al., 1993).

O número de repetições de microssatélites é muito variável devido a ‘escorregões’ durante a replicação do DNA (Frankham et al., 2008). As mutações que evitam a correção pelo sistema de reparo de incompatibilidade de DNA formam novos alelos nos locos SSR. Por esta razão, diferentes alelos podem existir em um determinado loco SSR, o que significa que os SSRs são mais informativos do que outros marcadores moleculares, incluindo Polimorfismos de nucleotídeo único (SNPs) (Vieira et al., 2016).

Em particular, os SSRs são úteis para espécies selvagens (i) em estudos de diversidade medidos com base na distância genética; (ii) estimar o fluxo gênico e as taxas de cruzamento; e (iii) em estudos evolutivos, sobretudo para inferir relações genéticas infraespecíficas (Vieira et al., 2016). Os microssatélites têm sido marcadores moleculares de escolha para estudos populacionais por apresentar vantagens sobre os outros métodos disponíveis para avaliar o polimorfismo do DNA, uma vez que eles são altamente variáveis, estão dispersos por todo o genoma, são fáceis de serem acessados, multialélicos e codominantes (Frankham et al., 2008). O uso de marcadores moleculares não codificantes com alto grau de polimorfismo representa uma boa ferramenta para os estudos familiares e de populações (Silva, 2010).

Antes da era do sequenciamento do genoma, o desenvolvimento de marcadores SSR apresentou vários desafios, pois esse processo exigia conhecimento prévio das regiões SSR e as suas regiões flanqueadoras para o desenho de oligos iniciadores. No entanto, este cenário se alterou quando sequências genômicas inteiras podem estar prontamente disponíveis ou ser geradas a um custo reduzido. O desenvolvimento de um bom conjunto de marcadores SSR, no entanto, ainda requer o rastreamento de um grande número de marcadores em um número suficiente de indivíduos para procurar por polimorfismos, o que, em última análise, ainda envolve considerável custo e esforço (Engelbrecht et al., 2017).

Tecnologias emergentes em sequenciamento de DNA (ou seja, *Next-Generation Sequencing* - NGS) provaram ser úteis para identificar locos microssatélites a partir de grandes quantidades de dados de sequência que geram com muito menos esforço e baixo custo, desafiando abordagens tradicionais para seu desenvolvimento (Davey et al., 2011; Zapala et al., 2012).

1.6 Justificativa e Relevância

As abelhas fornecem benefícios indiretos para a sociedade, sendo consideradas importantes componentes dos ecossistemas terrestres ao realizar de forma eficiente a

polinização, com transferência dos grãos de pólen entre flores distantes e auxiliando a reprodução sexuada dos vegetais (Gaglianone et al., 2011).

As abelhas solitárias coletoras de óleo polinizam espécies botânicas que fornecem óleos florais, tais como, as famílias Calceolariaceae, Iridaceae, Krameriaceae, Orchidaceae, Malpighiaceae, Plantaginaceae, Solanaceae (Renner e Schaefer, 2010). A espécie *C. analis* tem grande importância ecológica e econômica por ser polinizadora da família Malpighiaceae. Nesse sentido, trabalhos estão sendo desenvolvidos com a utilização desta espécie na polinização de pomares de acerola (*Malpighia emarginata* DC) na tentativa de aumentar o rendimento da produção. Magalhães e Freitas (2012) demonstraram que a introdução de ninhos da espécie aumentou o rendimento da cultura, em média 3,1 kg por planta. Além disso, mostraram que a inserção de ninhos permite o estabelecimento de populações de abelhas coletoras de óleo próximas das plantas de acerola, e com isso mitigar o déficit de polinização nesta cultura com a manutenção dessas populações de abelhas.

Conhecer a biologia e genética dessas abelhas que habitam ambientes alterados, principalmente aquelas pouco conhecidas, caso de abelhas da tribo Centridini e, em particular, a espécie *C. analis*, alvo do presente estudo, permite definir estratégias de conservação para a espécie. A falta de informações sobre a estrutura genética de um agente polinizador com ampla distribuição na América do Sul requer esforços na busca de marcadores que estabeleçam parâmetros genéticos do grupo (Silva, 2010).

Poucos estudos sobre estrutura populacional com uso de marcadores moleculares em abelhas da tribo Centridini foram realizados: Ferreira et al., (2013); Moure-Oliveira (2013) Luna-Lucena (2015). Dado o reduzido número de trabalhos, fica evidente a importância da realização de novos trabalhos que aumentem nosso conhecimento acerca da estrutura genética de abelhas solitárias, sejam elas de ampla ou restrita distribuição, para que dados possam ser comparados e se compreenda como a biologia das espécies afeta o nível e a distribuição da variação genética. Além da relevância de se realizar um novo estudo na espécie, torna-se essencial a utilização de um outro tipo de marcador molecular, especificamente locos microssatélites, para contornar as dificuldades metodológicas com a utilização dos marcadores mitocondriais encontradas por Moure-Oliveira (2013).

De igual forma, no tocante à genética familiar, o trabalho realizado com ninhos-armadilha para genética intranidal necessita ser repetido com o uso de outro marcador, conforme sugerido pela autora Silva (2010), pois o gene utilizado é de herança materna e não é indicado para estudo de genética familiar por ser uniparental e todas as crias do ninho

possuirão o haplótipo materno. Sendo assim, não poderia fornecer evidências de que a apenas um macho pode ser atribuída a paternidade da prole presente no ninho.

Outro aspecto ainda questionável da biologia desta espécie é a sua distribuição. Conforme consta no Catálogo Moure (<http://moure.cria.org.br/catalogue>), a espécie ocorre no Brasil (Ceará, Goiás, Mato Grosso, Paraná, Pará, São Paulo), Costa Rica, Guiana Francesa, Guatemala, Guiana Honduras, México (Chiapas, Jalisco, Michoacán, Morelos, Oaxaca, San Luis Potosi, Veracruz, Yucatan), Panamá, Trindade e Tobago, e Venezuela (Figura 4). Porém, o Catálogo não engloba todos os pontos de ocorrência já registrados na literatura, sendo necessária sua atualização inserindo novas informações que contribuam para o melhor conhecimento desta distribuição.



Figura 4. Área de distribuição da espécie *C. analis* de acordo com os pontos de ocorrência.

Ao utilizar os pontos de ocorrência da espécie disponíveis na rede, assim como os disponíveis no Laboratório de Genética de Himenópteros, a criação de uma modelagem de

nicho pode revelar áreas potenciais ainda não amostradas, e assim, contribuir para expedições de coleta em áreas potenciais aumentando o acervo de material biológico para estudos futuros.

2 OBJETIVOS

O objetivo central deste trabalho foi testar a hipótese de que “as populações urbanas de *Centris (Heterocentris) analis* não são estruturadas geneticamente”. Esta hipótese foi formulada considerando que a espécie: i) é polilética; ii) possui alta plasticidade e, iii) alta capacidade de dispersão conferida por seu médio porte. O teste da hipótese foi realizado com a estimativa da variação genética inter e intrapopulacional utilizando locos microssatélites espécie-específicos.

Adicionalmente, como objetivos específicos:

- i) verificar a estrutura genética populacional e estimar variação genética inter e intrapopulacional;
- ii) realizar modelagem de nicho ecológico para definir áreas potenciais de distribuição da espécie;
- iii) reavaliar a estrutura sociogenética intranidal, confirmando que as fêmeas apresentam acasalamento monândrico.

3 MATERIAL E MÉTODOS

3.1 Modelagem da distribuição potencial

O modelo de distribuição foi realizado conforme tutorial de Dalapicolla (2016), com uso do software ArcGis versão 10.5 para elaboração dos mapas, criar shapefiles, máscaras, cortar e reamostrar as camadas bioclimáticas, sendo implementado no programa Maxent versão 3.4.1 para definir as áreas potenciais de distribuição da espécie.

Para estimar as áreas potenciais de ocorrência da espécie *C. analis*, as bases de dados online como catálogo Moure, site SpeciesLink (www.splink.cria.org.br), sites de museus com registro de espécie e no banco de dados GBIF (www.gbif.org) foram acessadas para obter dados de presença. Também foram realizadas coletas afim de aumentar os pontos de ocorrência. De posse dos dados, realizou-se uma triagem manual dos pontos para utilizar pontos equidistantes; em seguida, foi montada uma planilha com as coordenadas geográficas dos locais de registro, ano de coleta, local de coleta, laboratório de conservação, coletor, observação (Material Suplementar). Para as amostras disponíveis no Laboratório de Genética Evolutiva de Himenópteros (LGEH) da UFSCar, o georreferenciamento foi realizado *in situ*;

para aquelas, provenientes do banco de dados foram utilizadas as coordenadas do próprio autor e consultadas no Google no Google maps e/ou Google Earth para confirmar a localização do ponto registrado.

Foi registrado um total de 72 pontos de coleta da espécie, entre todas as fontes, incluindo as do Laboratório de Himenópteros. Desse total, foram separados 15 pontos entre os demais para fazer a validação externa do modelo a ser utilizado, sendo estes pontos não utilizados no Maxent para modelagem.

As camadas bioclimáticas utilizadas na modelagem foram provenientes do banco de dados online WorldClim (www.worldclim.org) versão 2.0. Foram utilizadas 19 variáveis (1-Temperatura anual média; 2-Intervalo Médio Diurno (Média mensal (temp máx – temp. min)); 3- Isotermalidade $(2/7)(*100)$; 4-Sazonalidade da Temperatura (desvio padrão*100); 5-Temperatura máxima do mês mais quente; 6-Temperatura mínima do mês mais frio; 7-Faixa anual de temperatura (5 - 6); 8-Temperatura média do trimestre mais chuvoso; 9-Temperatura média do Trecho Seco; 10-Temperatura média do trimestre mais quente; 11-Temperatura média do trimestre mais frio; 12-Precipitação anual; 13-Precipitação do mês mais chuvoso; 14-Precipitação do mês mais seco; 15-Sazonalidade da precipitação (Coeficiente de Variação); 16- Precipitação do trimestre mais chuvoso; 17-Precipitação do trimestre mais seco; 18-Precipitação do trimestre mais quente; 19-Precipitação do trimestre mais frio) no tamanho de 5,0 minutos, conforme recomendado por (Gianini et al., 2012).

No ArcGis, foi criado um shapefile na projeção WGS1984 dos pontos a serem usados no Maxent. Em seguida, criou-se uma máscara da área de estudo para o corte das camadas, com diluição (limpeza) e o buffer (área de amortecimento de 20 km) desta máscara. De posse da máscara, foi possível cortar todas as camadas climáticas utilizando este molde, reamostrando em seguida sua resolução e no final convertendo no formato *.asc*.

Com as camadas climáticas prontas para uso, foi realizada a eliminação de dados abióticos redundantes. Para isso, no programa RStudio versão 3.4.3 foi realizada uma PCA com uso dos pacotes “raster”, “rgdal” e “vegan” para escolher variáveis não correlacionadas. Para a eliminação de dados bióticos (pontos de ocorrência) redundantes, utilizou-se o pacote de ferramentas SDMTools implementado no ArcGis por meio de rarefação simples dos pontos de ocorrência, rarefação por heterogeneidade ambiental das camadas ambientais, considerando uma distância de 20 quilômetros (km), eliminando assim, pontos de ocorrência duplicados ou correlacionados.

A modelagem foi realizada no programa Maxent versão 3.4.1. Neste programa, foram utilizados os pontos de ocorrência e as camadas abióticas não correlacionadas. O

programa foi executado com 10.000 *background*, 30 replicações e 5.000 iterações. Ao término das corridas, o mapa com as potenciais áreas de distribuição da espécie foi plotado no ArcGis com base no valor de *threshold* mínimo, 10% e máximo, escolhendo ao final o modelo que melhor representa a área conhecida de distribuição da espécie, sendo logo em seguida validado no mesmo programa. A validação foi feita com uso dos 15 pontos inicialmente retirados do conjunto total e aqueles que foram eliminados na etapa de rarefação, sendo então, plotados no mapa para verificação se todos ficavam dentro da área potencial de distribuição.

Todos os procedimentos para criação de um modelo foram feitos três vezes, uma vez para os pontos amostrados pelo pessoal do LGEH (cidades), um segundo para aqueles pontos consultados na base de dados, e uma final para a amostragem total.

3.2 Procedimentos de coleta

Durante o pouso nas flores de *Tecoma stans* (L) Kunth para coleta de pólen e/ou óleo, os indivíduos foram coletados com auxílio de tubos Falcon 50 mL, colocando a boca do tubo frente à abertura da flor, capturando o indivíduo ao tentar sair do interior da flor (Figura 5). Foram feitas coletas também com rede entomológica enquanto as abelhas sobrevoam as plantas de *T. stans*, também conhecida como falso ipê, amarelinha ou ipezinho de jardim.

Em seguida, os animais eram sacrificados em gelo dentro de uma caixa térmica e, na sequência, acondicionados em tubos de vidro com álcool 85% para conservação do material durante o transporte e armazenamento. Todos os espécimes coletados em um mesmo sítio eram dispostos em um mesmo recipiente. No LGEH, o material foi identificado por espécie e sexo, e armazenado em álcool 85% a 4° C.



Figura 5. Procedimento de coleta de *C. analis* durante as expedições. Nesta foto foi coletado um indivíduo do gênero *Thygater* (Apidae: Eucerini). Foto: Moure-Oliveira.

3.2.1 Material e locais de coleta

Amostras de *C. analis* foram coletadas em quatro regiões brasileiras - Norte, Nordeste, Centro-Oeste e Sudeste - englobando um total de 13 estados e 21 cidades. Graças aos trabalhos anteriormente realizados no LGEH (Moure-Oliveira, 2013; Luna-Lucena, 2015; Luna-Lucena et al., 2017), foi montado um grande acervo de material biológico da espécie para ser utilizado neste trabalho. Além do mais, ampliaram-se as coletas para estados não amostrados para obtenção de dados e resultados mais completos de acordo com a distribuição da espécie no país. Portanto, os espécimes analisados neste trabalho são provenientes de diferentes períodos de coleta, incluindo aqui amostras disponíveis no laboratório desde o ano de 2012 até o ano de 2018. Ressalte-se que o material analisado de uma determinada localidade é proveniente de um único período de coleta.

Em cada cidade, selecionou-se de 3-6 sítios distantes entre si de mínimo 2 km, para a coleta de 10 a 15 indivíduos por sítio, de forma a se ter uma maior representatividade da população local. As coletas foram realizadas principalmente no período de novembro a maio, visto que a atividade de nidificação da espécie está concentrada nesse período (Gazola e Garófalo, 2009; Silva et al., 2017). Nesse período, a espécie intensifica as coletas de óleo e pólen para sua nidificação. As coletas eram feitas entre 10 e 16 horas, período em que as plantas têm maior produção de néctar e óleos florais e as abelhas são observadas com maior frequência (Camargo e Mazucato, 1984; Gaglianone, 2003; Silva et al., 2007).

Na tabela 1 e Figura 6 podem ser visualizadas as localidades nas quais foram feitas as coletas. Estão listados os estados: Acre, Amapá, Pará, Maranhão, Ceará, Paraíba, Bahia, Goiás, Espírito Santo, São Paulo, Minas Gerais, Mato Grosso e Mato Grosso do Sul, no total de 13 estados e 21 localidades, totalizando 221 espécimes, dos quais 167 são machos e 54 são fêmeas. Sempre que possível, foram analisados 8 machos e 4 fêmeas por localidade para realizar a comparação de genótipos haploides dentro da população.

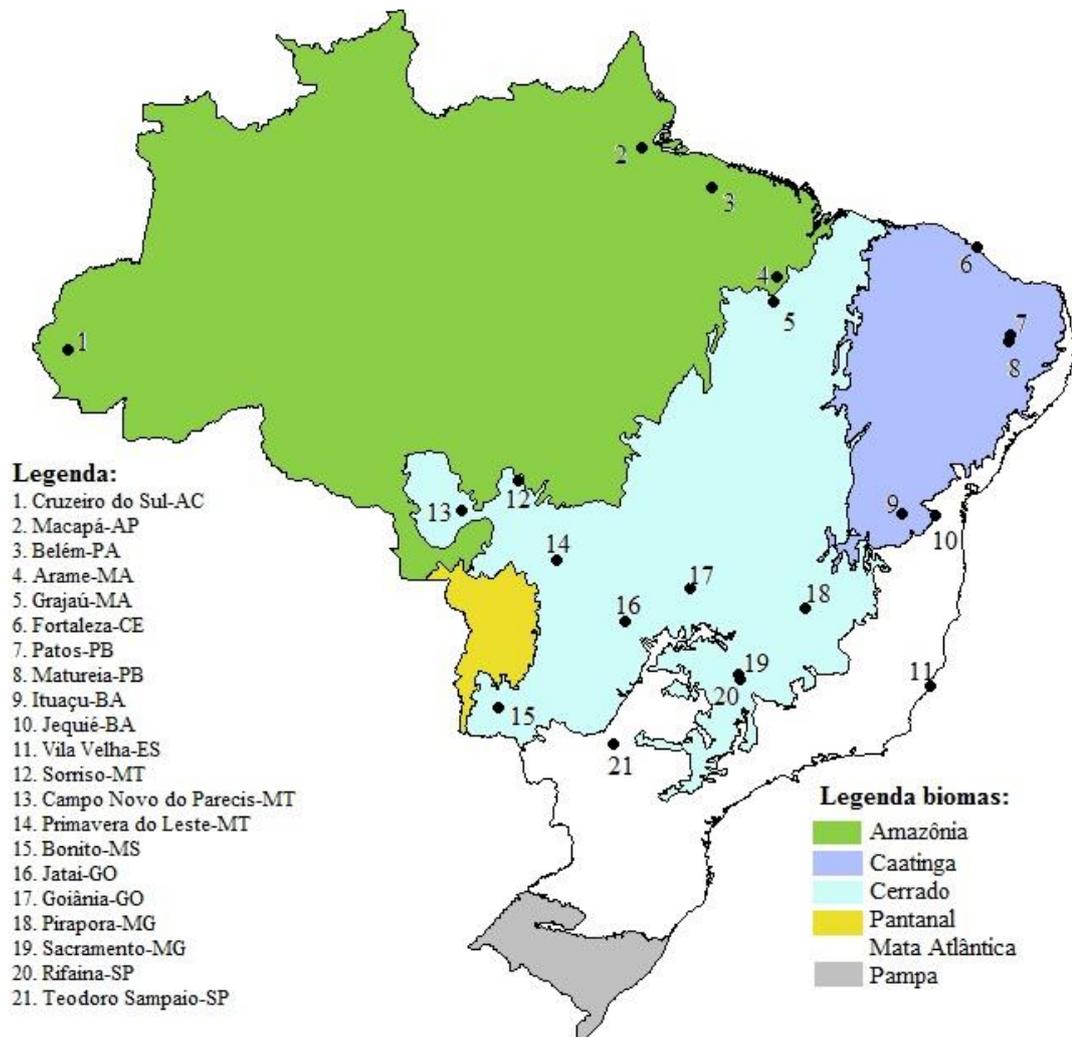


Figura 6. Mapa do Brasil demonstrando as cidades e estados em que foram feitas coletas de *C. analis* entre 2012 a 2018.

3.2.2 Ninhos-armadilha

Os ninhos-armadilha (NA) foram instalados em dois locais em que ocorria *Malphigia glabra* L. no município de São Carlos-SP para promover a nidificação de *C. analis*. Estas proles forneceram material biológico para investigar a estrutura genética dos ninhos, de acordo com a recepção da fêmea a um ou mais machos.

Os NA, confeccionados em cartolina preta na forma de tubos com uma das extremidades fechadas com o mesmo material, conforme metodologia proposta (Jesus e Garófalo, 2000; Couto e Camilo, 2007; Alonso et al., 2012), apresentavam as seguintes dimensões: 0,6 cm de diâmetro x 5,8 cm de comprimento. Estes NA eram introduzidos em orifícios feitos em placas de madeira de 05 cm de largura, 30 cm de comprimento e 12 cm de altura com 55 orifícios por placa, 05 linhas identificadas de A-E, e 11 colunas identificadas de 1 a 11 (Figura 7).

Foram coletados um total de 30 ninhos, do qual foram selecionados 8 com número de proles entre 3 e 5, totalizando 30 indivíduos analisados, 16 machos e 14 fêmeas. No estudo de genética intranidal foram utilizados 9 locos microssatélites.

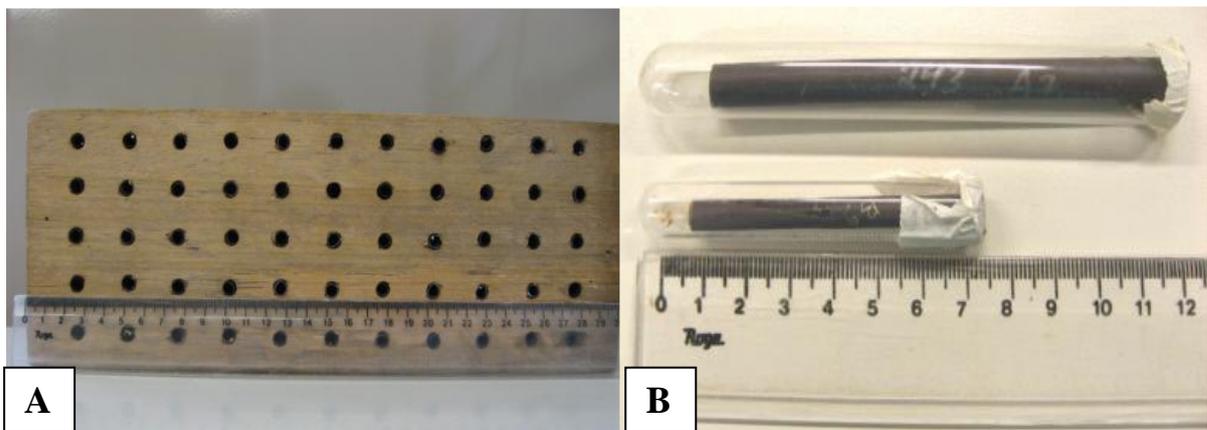


Figura 7. Ninhos armadilha utilizado na coleta de *C. analis*. Em A, o bloco de madeira com perfurações usado para alocar os ninhos-armadilha. Em B, ninhos-armadilha fabricados em colorset preto. Fonte. Silva, 2010.

3.3 Métodos

3.3.1 Extração de DNA

O DNA total foi extraído de pernas ou tórax de indivíduos adultos, machos e fêmeas de *C. analis*, utilizando o protocolo de fenol-clorofórmio (Sheppard e McPheron 1991) com algumas modificações, sendo armazenado à -20°C para análises genéticas posteriores. Foram também utilizados os protocolos do kit de Extração da Quiagen e o Kit ZR Tissue & Insect DNA MiniprepTM da Zimo Research, modificado por Kátia Maria Ferreira (Apêndice I).

Tabela 1. Sítios analisados listados por bioma, estado, localidade, coordenadas geográficas, indivíduos analisados por sexo e ano de coleta.

Bioma	Estado	Localidades	Sigla	Latitude	Longitude	Machos	Fêmeas	Total	Ano de Coleta
Amazônia	AC	Cruzeiro do Sul	CZS	-7.61394	-72.6616	8	4	12	2016
	AP	Macapá	MCP	0.034067	-51.0766	8	4	12	2018
	MA	Arame	ARA	-4.83651	-46.0232	11	1	12	2017
	PA	Belém	BLM	-1.44832	-48.4730	6	2	8	2018
Caatinga	BA	Ituaçu	ITU	-13.8131	-41.3604	8	4	12	2012
	CE	Fortaleza	FOT	-3.73699	-38.5141	5	1	6	2016
	PB	Patos	PTS	-7.02043	-37.2789	7	0	7	2013
	PB	Maturéia	MTR	-7.26214	-37.3522	6	3	9	2013
Mata Atlântica	ES	Vila Velha	VLV	-20.3463	-40.2950	8	4	12	2016
	BA	Jequié	JEQ	-13.8496	-40.0530	11	1	12	2012
Cerrado	GO	Goiânia	GYN	-16.6045	-49.3148	11	1	12	2018
	GO	Jataí	JTI	-17.8878	-51.7352	5	5	10	2018
	MA	Grajaú	GRU	-5.81504	-46.1349	8	4	12	2017
	MG	Sacramento	SCM	-19.8588	-47.4542	8	4	12	2015
	MG	Pirapora	PRP	-17.3483	-44.9430	5	2	7	2017
	MS	Bonito	BOT	-21.1271	-56.4943	6	1	7	2016
	MT	Campo Novo do Parecis	CNP	-13.6578	-57.8929	9	3	12	2016
	MT	Primavera do Leste	PVA	-15.5633	-54.2950	10	2	12	2018
	MT	Sorriso	SOR	-12.5406	-55.7221	11	0	11	2016
	SP	Teodoro Sampaio	TES	-22.5312	-52.1739	8	4	12	2015
	SP	Rifaina	RIF	-20.0799	-47.4304	8	4	12	2015
21						167	54	221	

3.3.2 Prospecção dos locos microssatélites

Devido à ausência de *primers* de microssatélites espécie-específicos para *C. analis*, foram prospectados locos de microssatélites di e tri nucleotídeos por NGS (*Next-Generation Sequencing*) a partir de DNA genômico de 04 machos no Laboratório Multiusuários Centralizado de Genômica Funcional Aplicado à Agropecuária e Agroenergia da ESALQ - USP, Piracicaba, SP. Após receber as sequências geradas por NGS (Sequenciamento de Nova Geração) em Plataforma Illumina, foi procedida uma análise preliminar com ferramentas bioinformáticas que selecionaram centenas de regiões candidatas a locos microssatélites. Estas, após triagem e busca de sequências repetitivas contendo motivos di e trinucleotídeos, serviram de base para a definição e delineamento de oligos iniciadores utilizando o software Primer 3 para 20 locos espécie-específicos dos quais 16 foram utilizados neste estudo (Tabela 2).

3.3.3 Análise dos locos microssatélites

A reação em cadeia da polimerase (PCR) foi realizada em volume final de 10 µL, contendo 1 µL de DNA, 0,1 µL de Taq polimerase, 0,8 µL M13 (12,5 µM), 0,2 µL (12,5 µM) *primer forward*, 0,8 µL (12,5 µM) *primer reverse*, 0,4 µL MgCl₂ (50 mM), 1,0 µL dNPTs (250 µM), 1,0 µL (10 X) Buffer e água esterilizada. A reação foi realizada em 40 ou 45 ciclos de acordo com o loco utilizado descritos na tabela 2, sendo um passo inicial de 94°C por 5 minutos, seguido por 30 ciclos de 94°C por 30 segundos, 62°C por 45 segundos e 72°C a 45 segundos, finalizando com 10 ciclos de 94°C por 30 segundos, 53°C a 45 segundos e 72°C a 45 segundos e um passo de extensão de 72°C a 10 minutos para anelamento da cauda M13.

A PCR foi padronizada agrupando os *primers* em multiplex de acordo com os fluoróforos utilizados (TET, FAM e HEX) de modo a realizar a genotipagem simultânea de três locos e o tamanho esperado do fragmento. Tal atividade foi realizada com a ligação de um oligonucleotídeo (fragmento de sequência universal) de 18pb capaz de hibridizar com um iniciador universal M13 marcado com um dos fluoróforos citados, conforme protocolo de Schuelke (2000) e adaptação realizada por Oblessuc e colaboradores (2009). O material amplificado foi observado sob gel de agarose 1,5%.

Tabela 2. Descrição dos primers desenvolvidos para a espécie *C. analis* mediante NGS, 2018. ID: Identificação dos primers, F: Forward; R: Reverse; TM: Temperatura de *Melting*; TF: Tamanho do fragmento em pares de bases; N° de rep: número de repetições; T-PCR: Temperatura da Reação em Cadeia da Polimerase; N° ciclos: número de ciclos da PCR.

ID	Primer 5' -3'	Descrição	TM (°C)	TF (pb)	N° de rep.	Motivo	Repetição do motivo	T-PCR	N° ciclos
Canl1	F	ATACACCCGCCACCCATCTC	68,3	211	12	CT	DI	62	30
	R	AGTCCTCCTCGTCGTCGTCT	64,8						
Canl2	F	GCATACCGCATTCCAAGTGT	64,9	285	10	TG	DI	62	30
	R	AGTGATTGCTGTGGCTCTT	63,7						
Canl4	F	GTTTCGGTTAGAGGAGTCGGC	65,5	200	10	TC	DI	62	30
	R	TAGAAAGGGACGAGAAAGGG	63,1						
Canl5	F	CTCCGCTCTTTCTTCGTCT	65,4	183	11	ACG	TRI	62	30
	R	GTGGTCATTGTCCCCTCAGT	63,6						
Canl6	F	AGAGAGCAAGTGACCGACCA	63,9	109	12	GCA	TRI	62	30
	R	GCGACTATTCTGTTTCTCCCG	66,1						
Canl7	F	AAAGGAAGGTGCCGACAAAC	65,9	195	12	CAA	TRI	62	30
	R	CCCAGCGTTCCTATTCTTGC	66,7						
Canl9	F	CGGACCCTGAGCGTAGTTCT	65,9	231	17	TC	DI	62	30
	R	ATAAGCGAAAGCGAAAGCAC	64,8						
Canl10	F	GATGAACCAGAGACGCAACG	65	242	14	TC	DI	62	35
	R	GATAATGTCGTCTGAGCCCG	64,4						
Canl11	F	AGTGCTATGCCGTCTGGA	60	121	10	CCT	TRI	58	30
	R	AGCGAGAGAACGAGGAGGAG	65,1						
Canl12	F	CCGAAGAGAACGAAGGACG	64,4	290	11	AGA	TRI	60	30
	R	GTTCCCTCTTCATCACGCAG	64,2						
Canl13	F	GCGAAATCCACCGAGAGAAT	65,8	293	12	AGA	TRI	58	30
	R	GAACAACGGTCTGGAAATCG	64,3						
Canl14	F	TCGCTCACAATAATGGCAGA	63,8	178	10	TG	DI	62	30
	R	CGCATAACACATTCACACGCT	64						
Canl16	F	GTTCCGTGTCTGGTTTCGTC	63,9	167	10	GAA	TRI	62	35
	R	CTCCTCTTCTGCCACCCTTC	65,1						
Canl18	F	CGTGCTTCTCCCTCTCCCTC	67,6	108	10	CT	DI	62	35
	R	TCGTATGTGGGTTAGAGGCG	65,2						
Canl19	F	CGTTTCTCGTAGTTTCTCGTCG	66,6	238	14	TTC	TRI	58	30
	R	CGTTTCTCTTCTCCTTCTCCG	65,7						
Canl20	F	AACTTTCCTACCTCGTCCGC	65,4	295	11	GA	DI	60	30
	R	ATCCTCTTCCCGCAGTTCAT	65,1						

O produto da PCR foi diluído em 10x, uma reação composta por 70 µL de água e 30 µL (10 de cada uma) proveniente das 3 PCR realizadas individualmente. Dessa diluição, uma alíquota de 2 µL foi misturada a 0,35 µL de Gene Tab 500 ou ROX Ladder e 7,65 µL de Tween 20 a 0,1%, conforme indicado pelos fabricantes, visando a totalidade de 10 µL no volume final. Em seguida, os fragmentos foram obtidos com a corrida em sequenciador automático MegaBACE™1000 do Laboratório de Genética Evolutiva de Himmenópteros/UFSCar. Os resultados obtidos serão analisados no software MegaBACE Genetic profiler versão 1.2, sendo possível visualização e análise dos alelos dos locos microssatélites.

3.3.4 Erros de genotipagem e caracterização dos locos microssatélites

Após a padronização de todas as PCRs para os locos microssatélites (conforme Tabela 2), foi realizado o cálculo de erro de genotipagem para cada loco conforme recomendado por Pompanon e colaboradores (2005). Os autores ressaltam que as principais causas dos erros de genotipagem podem estar associadas, a variação na sequência de DNA, mutação perto de um marcador, se a sequência flanqueadora estiver envolvida no processo de detecção de marcadores, segundo, pela baixa quantidade ou qualidade do DNA extraído, artefatos bioquímicos e erros humanos.

Segundo os mesmos autores, o cálculo do erro de genotipagem deve ser realizada com base em um conjunto de dados com replicações. Primeiro, realiza a genotipagem de todas as replicações, define-se um genótipo de referência (que mais se repete) e verifica quantas replicações se diferiram. A estimativa é realizada com base no número de erros encontrados entre o genótipo de referência e os replicados. Pode-se realizar a taxa de erro por alelo, por loco, por genótipo multilocos, e por reação. Neste estudo, foi realizada estimativa da taxa de erro por genótipo multilocos.

Para o cálculo de erro de genotipagem foram utilizadas 23 amostras escolhidas levando em consideração localidades equidistantes, conforme listadas na Tabela 3, incluindo machos e fêmeas de diferentes localidades. Foram utilizados 16 locos microssatélites, sendo que para cada um deles foram feitas 4 réplicas por amostra.

Tabela 3. Relação das amostras utilizadas para cálculo da taxa de erro de genotipagem por loco microssatélite.

Estado	Cidade	Machos	Fêmeas	Total
AC	Cruzeiro do Sul	2	1	3
BA	Vitória da Conquista	3	1	4
ES	Vila Velha	3	0	3
MA	Grajaú	2	1	3
MG	Sacramento	3	0	3
SP	Teodoro e Sampaio	1	2	3
	Rifaina	0	4	4
Total		14	9	23

Após verificar a taxa de erro de genotipagem dos locos, os mesmos foram implementados no software Micro-Checker versão 2.2.3 com um intervalo de confiança de 95% e 10.000 permutações (Oosterhout et al., 2004) para uma análise preliminar quanto à presença de alelos nulos, ausência de amplificação (*dropout*), pequenas alterações no tamanho dos alelos durante a amplificação (*stuttering*) e erros de leitura de genotipagem. Para verificar desequilíbrio de ligação, utilizou-se o programa Arlequin versão 3.5.2.2 (Excoffier e Lischer 2010) com 10.000 permutações e nível de significância de 5%.

3.3.5 Análise de diversidade genética e estrutura populacional

Para verificar o número médio de alelos e a riqueza alélica e índice de fixação (F_{ST}) por loco foi utilizado o programa FSTAT versão 2.9.4 (Goudet, 1995) com 15.000 permutações. O número de alelos privados foi verificado com o uso do software GenAlEx (Peakall e Smouse, 2012) com 999 permutações. As heterozigosidades médias observada e esperada foram estimadas com o programa Arlequin utilizando apenas os dados genéticos de fêmeas, haja vista que a diploidização dos machos envia os resultados das análises, o mesmo ocorrendo para a estimativa do coeficiente de endogamia (F_{IS}) implementado no programa FSTAT. O cálculo do valor de F_{ST} par a par entre amostras de localidades que tinham apenas fêmeas foi realizado com programa Arlequin com intervalo de confiança de 95%.

Desvios das frequências no equilíbrio de Hardy-Weinberg (EHW) foram calculadas no programa Arlequin versão 3.5.2.2 (Excoffier e Lischer 2010) com 10.000 permutações e nível de significância de 5%. O programa GenAlEx com 999 permutações e o Arlequin nas condições recém citadas foram usados para execução de uma Análise de Variância molecular (AMOVA) em dois níveis hierárquicos, dentro e entre as populações, para verificar como a diversidade estava particionada.

A verificação de estruturação genética das populações amostradas pela caracterização de grupos genéticos foi realizada por meio de inferência bayesiana com MCMC (Markov chain Monte Carlo) no programa Structure versão 2.3.4 (Pritchard, Stephens e Donnelly, 2000). As análises foram realizadas para o conjunto total de dados (não indicada devido a presença de machos diploidizados), para os genótipos dos machos e das fêmeas separadamente, implementadas nos modelos de *no admixture* e *admixture* com 20 iterações, um *burn-in* inicial de 50.000, seguidas de 1.000.000 simulações para cada valor de K, que foi estipulado entre 1 e 10 para cada análise em separado, após verificar em corridas mais curtas que o valor de K não seria superior a esse valor (10). Os modelos *no admixture* e *admixture* são dois modelos de ancestralidade aplicado pela estrutura genética. O modelo *no admixture* é mais indicado quando não há conhecimento prévio da origem das populações e considera com uma como completamente discreta, capaz de detectar estruturas sutis. O modelo *admixture* recomendado para iniciar as análises é indicado quando há mistura entre as populações de tal forma que uma grande parte dos indivíduos amostrados pode ter ancestrais recentes de várias populações (Porras-Hurtado et al., 2013).

A estrutura genética também foi verificada para o conjunto de dados de machos e fêmeas separados no programa TESS versão 2.3.1 (Durand et al., 2009). As análises foram iniciadas com a utilização da ferramenta Generate Spatial Coordinate, na qual foram introduzidos dois arquivos de entrada, sendo um deles com as coordenadas geográficas (CG) das localidades amostradas e o número de indivíduos por localidade. As CG abrangiam uma área mínima e máxima de latitude e longitude para que a ferramenta gerasse coordenadas distintas individuais dentro da mesma localidade. O outro arquivo de entrada continha as informações genéticas de todas as localidades e locos utilizados, resultando em um arquivo com os dados fornecidos combinados.

De posse do arquivo conjunto, o programa TESS foi executado nas condições de valor de K entre 2 e 10 com 10 iterações, *burn-in* inicial de 10.000, seguidas de 50.000 simulações para cada valor de K para testar os modelos disponíveis, *without admixture*, *with admixture* (BYM e CAR) tipos: constante, linear, quadrático e cúbico. Após as corridas, foram analisados os valores de DIC (*Deviance Information Criterion*) gerados em cada corrida para cada modelo, sendo plotados num gráfico após retirada de média harmônica para cada valor de K, determinando assim o melhor modelo. Após isso, com o melhor modelo (menor valor de DIC) e do seu valor de K mais provável, o programa foi executado novamente nas condições de 100 iterações, *burn-in* inicial de 20.000, seguidos de 100.000 simulações para cada valor de K. O programa TESS foi utilizado por ser possível utilizar

coordenadas geográficas por indivíduos durante as análises, permitindo dessa forma detectar estruturação genética mais sutil que possa existir no conjunto de dados analisados.

Os resultados do programa Structure foram implementados no programa online PopHelper Structure Web App v. 1.0.10 (Francis, 2016) para sumarização das corridas, avaliando o melhor valor de K pelo método de Evanno et al., (2005) que se baseia na taxa de mudança de log de probabilidade dos dados entre valores de K consecutivos (ΔK) e edição do gráfico de barras gerados pelo Structure. O programa também foi utilizado para edição do gráfico de barras com os resultados do programa TESS.

A estrutura genética populacional também foi estimada por meio de uma Análise Discriminante de Componentes Principais (DAPC) implementada no programa RStudio versão 3.4.3 com uso do pacote Adegenet, primeiramente fazendo uma Análise de Componentes Principais (PCA) para obter o valor de BIC (*Bayesian Information Criterion*). Esta análise foi realizada por estimar a estruturação genética com base em componentes principais, tais componentes podem ser sutis e indetectável em outros programas, de tal forma que a DAPC pode estimar uma estruturação mesmo que ela seja muito baixa dentro do conjunto de dados.

3.3.6 Análise da estrutura sociogenética intranidal

As análises de estrutura sociogenética intranidal foram realizadas com os indivíduos coletados em ninhos-armadilha, buscando identificar os diferentes ninhos gerados por uma mesma fêmea, de forma que os dados fossem suficientes para corroborar a paternidade e definir o tamanho da prole de uma fêmea. Para delimitar o sistema de acasalamento desta espécie utilizou-se a segregação dos genótipos segundo a genética mendeliana considerando o sistema sexual haplodiploide.

A determinação dos genótipos parentais foram determinados levando em consideração o sistema de acasalamento monândrico/monogínico em que em um ninho só pode existir 3 alelos diferentes. Aqueles alelos presentes na prole masculina são oriundos da mãe e aqueles presentes nas proles do sexo feminino são de origem materna e paterna.

4 RESULTADOS

4.1 Modelagem de nicho ecológico

A Análise de Componentes Principais (PCA) das variáveis ambientais gerou dois arquivos na pasta de saída: um deles com o valor de contribuição de cada Componente

Principal (PC) em porcentagem e o outro com a contribuição de cada variável para cada componente (CVPCA). Os valores de percentil PC obtidos pela Análise de Componentes Principais (PCA) das camadas ambientais podem ser visualizados na Tabela 4. Os valores de PC utilizados para selecionar as camadas usadas para modelagem foram aqueles que responderam por mais de 70% da variação (74,46%), ou seja, os primeiros 5 PCs são os que mais contribuíram para a variação, sendo então consideradas as camadas que mais contribuem para o modelo dentro destes PCs.

Tabela 4. Valores de Componentes Principais (PCs) referente à análise de contribuição de variação das camadas ambientais usadas na modelagem de nicho ecológico de *C. analis*.

PC	X
1	27,41
2	17,27
3	13
4	9,02
5	7,76
6	6,28
7	4,41
8	3,98
9	2,81
10	1,78
11	1,7
12	1,1
13	0,72
14	0,65
15	0,5
16	0,49
17	0,47
18	0,33
19	0,31

Uma vez definidos os valores de PCs, foi verificada a contribuição individual de cada variável resultante da PCA (Tabela Suplementar 1 – ver Apêndices) levando em conta o valor de correlação entre as camadas. Dessa forma, foram utilizadas as variáveis com

contribuição de até 0.30, 0.36, 0.39, 0.53 e 0.46 para o PC1, PC2, PC3, PC4 e PC5, respectivamente. Após verificar a contribuição de cada variável, foram eliminadas as com correção acima de 70%, conforme apresentado na Tabela Suplementar 2 (ver Apêndices).

De acordo com esses critérios, foram selecionadas 6 camadas ambientais não correlacionadas entre si (5-Temperatura máxima do mês mais quente; 6-Temperatura mínima do mês mais frio; 7-Faixa anual de temperatura (5 - 6); 14-Precipitação do mês mais seco; 15-Sazonalidade da precipitação (Coeficiente de Variação); 18-Precipitação do trimestre mais quente). Escolhidas as camadas, os pontos disponíveis foram rarefeitos de modo simples e por heterogeneidade ambiental e foram eliminados os pontos correlacionados, permanecendo para as análises 65 pontos.

Dentre as seis camadas selecionadas pela PCA, três foram escolhidas (5, 6 e 14) para extrair os valores das variáveis dos pontos de ocorrência que mostram a necessidade ou não de se coletar mais pontos. Com esses valores foram gerados gráficos biplot para cada combinação das variáveis (5 x 6; 6 x 14; 5 x 14). Os resultados das combinações podem ser vistos na Figura 8. Os resultados demonstraram que para as camadas ambientais das temperaturas mínima do mês mais frio e máxima do mês mais quente, os pontos utilizados são suficientes para realização da modelagem por se agruparem em uma área do gráfico, assim como precipitação do mês mais seco e temperatura mínima do mês mais frio, pois os pontos encontram-se espalhados por todo o gráfico. No entanto, a correlação entre temperatura máxima do mês mais quente e precipitação do mês mais seco mostra os dados agrupados tendo um ponto equidistante, demonstrando que alguns pontos precisam ser amostrados para que dados mais robustos sejam obtidos.

A área potencial de distribuição da espécie foi realizada com base nos dados de presença, ou seja, apenas com pontos de ocorrência já registrados. O modelo foi gerado sob três combinações de dados: 1) pontos de ocorrência obtidos pelo Laboratório de Genética Evolutiva de Himenópteros; 2) pontos de ocorrência registrados nos bancos de dados, e 3) conjunto completo dos dados.

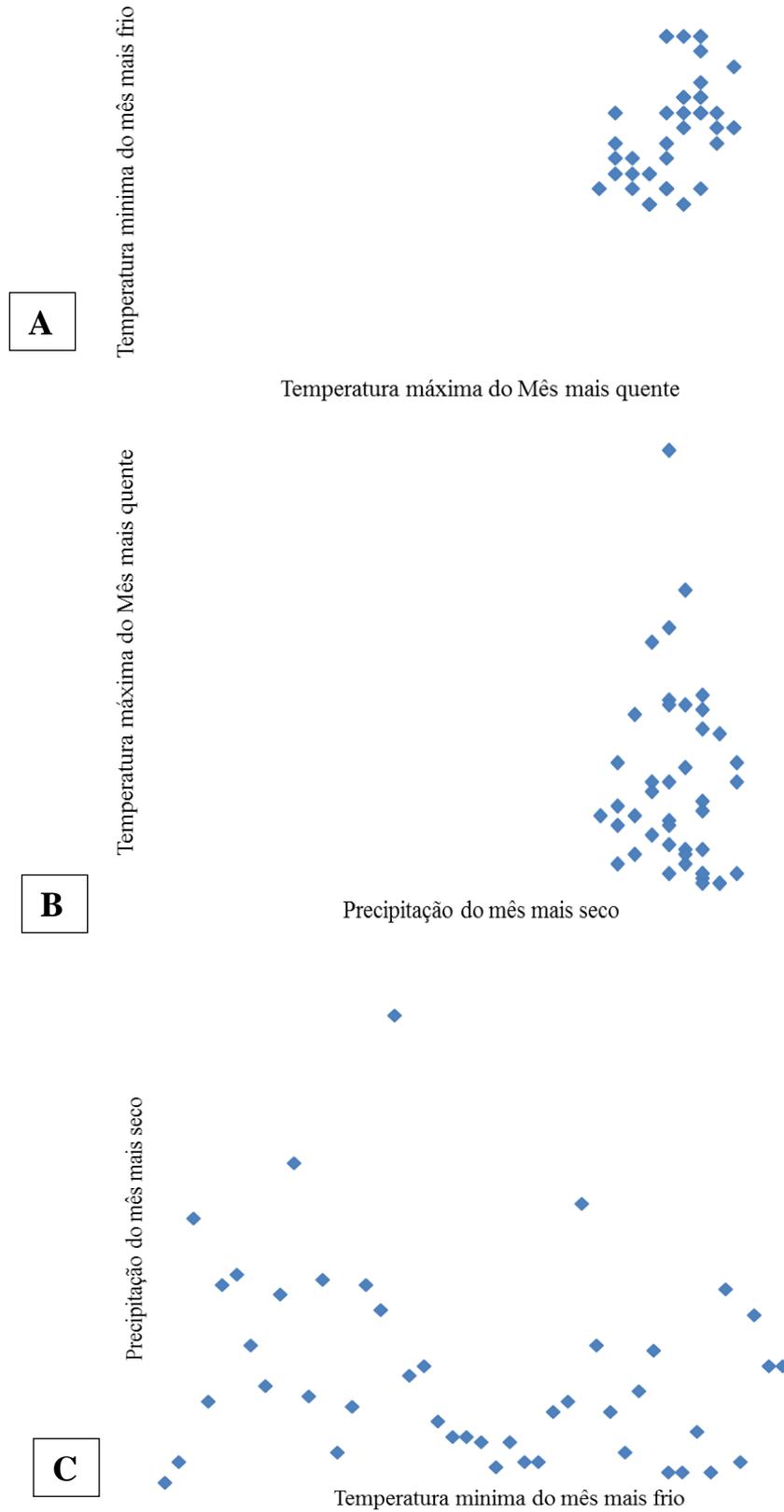


Figura 8. Correlação entre variáveis ambientais. a) Correlação entre temperatura mínima do mês mais frio e temperatura máxima do mês mais quente; b) correlação entre temperatura máxima do mês mais quente e precipitação do mês mais seco; c) correlação entre precipitação do mês mais seco e temperatura mínima do mês mais frio.

A calibração dos modelos foi realizada da forma descrita a seguir. Para o Modelo 1, utilizou-se 39 pontos para criação do modelo, $AUC = 0.9623$; *threshold* mínimo = 0,1924, erro de omissão do *threshold* = 0,0133, valor de $p = 0,00$, validação com taxa de acerto de 100,0. Para o Modelo 2, utilizou-se 33 pontos para criação do modelo, $AUC = 0.9403$; *threshold* 10% = 0,3835, erro de omissão do *threshold* = 0,015, valor de $p = 0,00$, validação com taxa de acerto de 100,0%. Finalmente, para o Modelo 3, utilizou-se 65 pontos para criação do modelo, $AUC = 0.9529$; *threshold* mínimo = 0,293, erro de omissão do *threshold* = 0,0298, valor de $p = 0,00$, validação com taxa de acerto de 98,46% e de omissão de 1,54%.

Os Modelos 1 e 2 (Figuras 9a e 9b, respectivamente) são complementares entre si, demonstrando que as áreas em que o Modelo 1 não dispunha de pontos de ocorrência, o Modelo 2 os continha, e ambos os modelos projetam a provável área de ocorrência. O Modelo 3 apresentado na Figura 9c mostra que a área de distribuição da espécie ocorre desde a América Central (México) até o sudeste do Brasil, com possibilidade de ocorrência adentrando o estado de Paraná.

Definidas as áreas potenciais de distribuição da espécie novas coletas foram realizadas em áreas que ainda não tinham pontos de coleta, tais como Altamira, Marabá e Santarém no estado do Pará, Manaus/AM, Goiânia e Jataí no Goiás. Dessa forma, com uma grande amostragem da espécie por sua área de distribuição no Brasil foram realizados estudos de diversidade genética e estrutura populacional com marcadores moleculares tipo microssatélites espécie específicos.

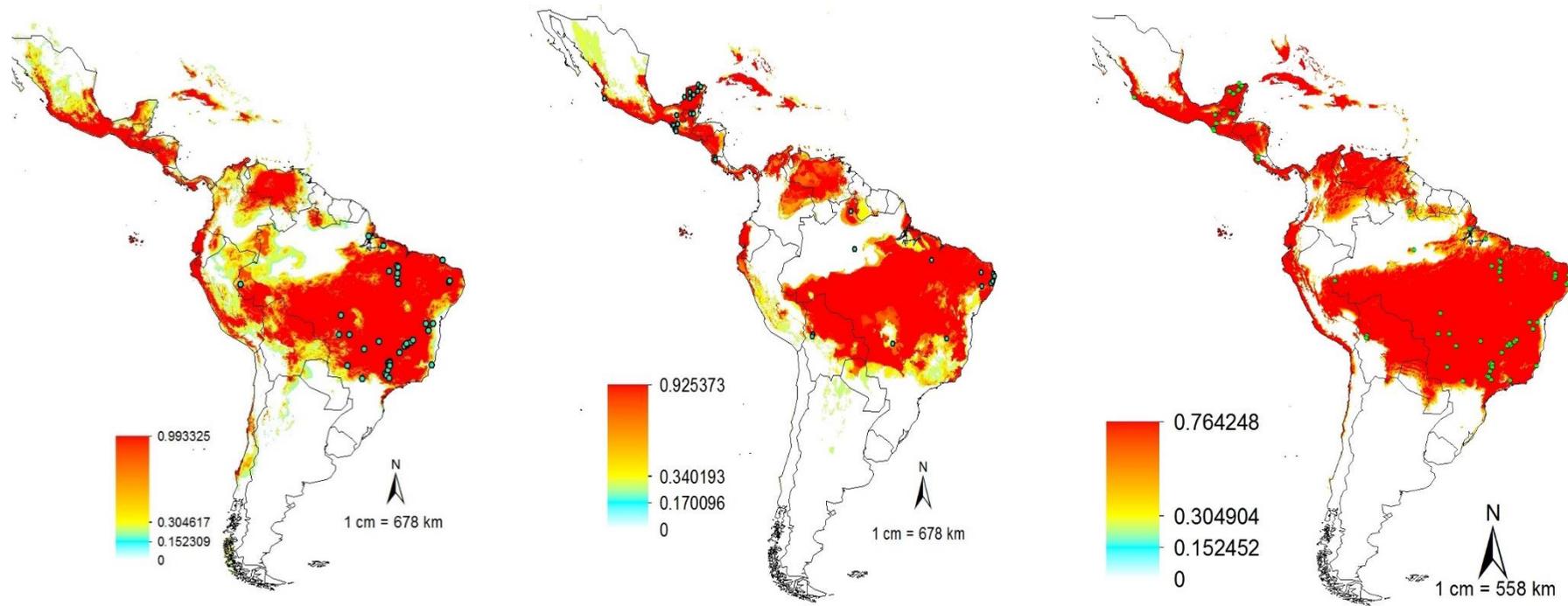


Figura 9. Área potencial de distribuição da espécie *Centris analis*. a) Modelo 1, pontos de ocorrência do Laboratório de Genética Evolutiva de Himenópteros; b) Modelo 2, pontos de ocorrência coletados online; c) Modelo 3, pontos totais de ocorrência da espécie. Legenda: o vermelho representa maior probabilidade da ocorrência da espécie na região; à medida que se vai na direção do branco, esta probabilidade diminui

4.2 Caracterização dos locos microssatélites

Foram prospectados 20 locos espécie-específicos para *C. analis*; destes, foram selecionados 16 com padrão de amplificação que permitiram a leitura inequívoca do tamanho dos alelos de cada loco. Os padrões apresentados pelos locos utilizados neste trabalho podem ser visualizados na Figura 10. A figura demonstra o formato dos picos de cada um dos locos, o que facilitava a identificação do alelo durante a conferência de genótipos.

Após a identificação do padrão dos alelos, foi realizada a estimativa da taxa de erro de genotipagem por loco (Tabela 5). Com base nesta tabela, verifica-se que a taxa de erro de genotipagem dos locos variou entre 0,00 (8 do total) a 2,17% (Canl16). Com base na baixa taxa de erro encontrada, todos os locos foram utilizados na genotipagem das amostras de todas as localidades, considerando que em nenhum deles foi atingida a taxa de 5% de erro.

A análise do Micro-Cheker não detectou erros de leitura nas análises de genotipagem e de *dropout* para nenhum dos locos utilizados. Entretanto, para 11 locos (Canl16, Canl19, Canl 10, Canl 6, Canl 13, Canl 9, Canl 7, Canl 18, Canl 5, Canl 1, Canl 2) foi detectada a presença de *stuttering* e alelos nulos devido à alta quantidade de homozigotos nas análises. Contudo, essas últimas análises foram desconsideradas, pois a “alta quantidade de homozigotos” constitui erro de interpretação do programa que considera os machos hemizigotos como indivíduos diploides homozigotos.

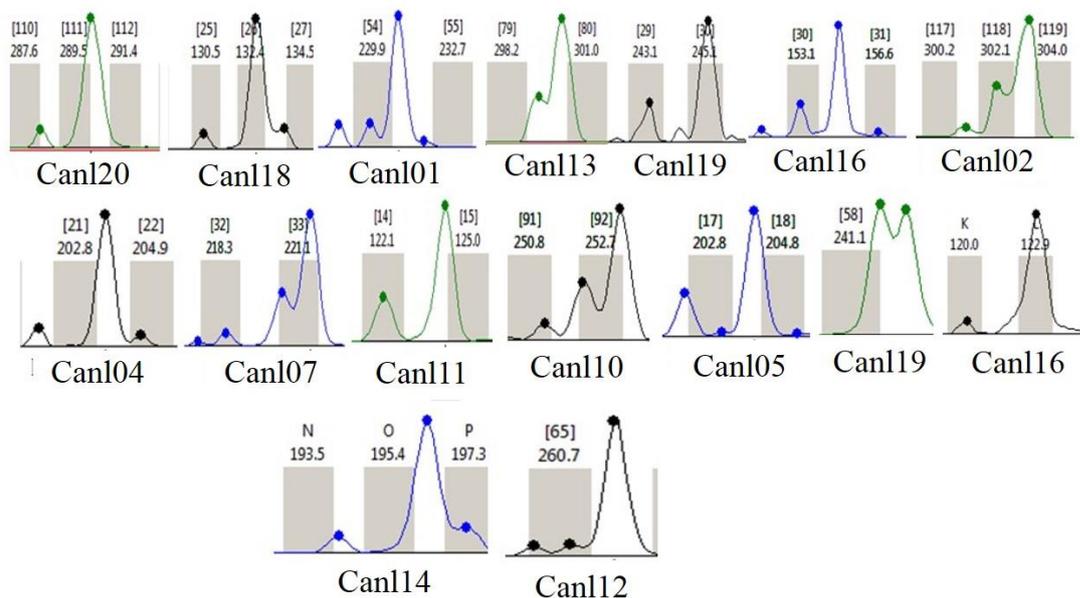


Figura 10. Padrões de amplificação dos 16 locos microssatélites espécie-específicos de *C. analis*.

Tabela 5: Descrição da taxa de erro de genotipagem por loco.

Loco	Taxa de erro por loco
Can11	0,58%
Can12	0,00%
Can14	0,00%
Can15	0,00%
Can16	1,00%
Can17	0,00%
Can19	1,33%
Can110	0,00%
Can111	0,00%
Can112	0,00%
Can113	1,25%
Can114	0,00%
Can116	2,16%
Can118	0,25%
Can119	1,17%
Can120	0,25%

Os 16 locos foram testados quanto ao desequilíbrio de ligação (DL) utilizando apenas os dados das fêmeas, pois ao considerar o conjunto total de amostras os dados enviesam. Os resultados não indicaram associação preferencial entre alelos de locos distintos para a maioria das localidades analisadas (52/4.864 testes), mas análises de pares de locos demonstraram que o DL ocorreu em uma a três localidades para o mesmo par, sendo considerados resultados não significativos. Dessa forma, ao considerar ausência de alelos nulos, ausência de desequilíbrio de ligação, todos os locos foram utilizados nas análises de estrutura populacional e diversidade genética.

Os locos utilizados foram caracterizados quanto à riqueza alélica, número de alelos, heterozigosidade observada e esperada e índice de fixação F_{ST} (Tabela 6; Tabela Suplementar 4 – ver Apêndices). Os resultados indicaram que todos os locos são polimórficos, com um total de alelos variando entre 6 (Can112) a 23 (Can113) e a riqueza alélica dos locos variou entre 3,27 (Can114) e 8,39 (Can15). Com relação à heterozigosidade observada (HO) e esperada (HE), os respectivos valores variaram entre 0,53 (Can112) a 0,89 (Can17) e entre 0,48 (Can112) a 0,88 (Can120), respectivamente, verificando-se índices similares entre os valores observados e esperados. O valor do F_{ST} por loco variou entre – 0,001 (Can111) a 0,100 (Can110) sendo, no entanto, oito locos com valores significativos.

Tabela 6. Características dos 16 locos microssatélites amplificados para indivíduos de *C. analis*. N: número de alelos; AR: Riqueza alélica, para machos e fêmeas; Ho: heterozigosidade observada; He: heterozigosidade esperada para fêmeas; F_{ST}: Índice de Fixação (Weir e Cockerham, 1984).

Locos	N	AR	Ho	He	F _{ST}
Canl1	8	4,89	0,65	0,70	-0,015
Canl2	8	4,42	0,65	0,69	-0,043
Canl4	16	5,81	0,78	0,76	0,011*
Canl5	20	8,39	0,74	0,84	0,030*
Canl6	9	5,54	0,68	0,75	0,015*
Canl7	15	6,95	0,89	0,81	0,041*
Canl9	17	7,42	0,73	0,84	-0,006
Canl10	18	7,51	0,85	0,85	0,100*
Canl11	14	5,55	0,71	0,84	-0,001
Canl12	6	4,05	0,53	0,48	-0,011
Canl13	23	7,45	0,78	0,84	0,017*
Canl14	8	3,27	0,63	0,60	-0,003
Canl16	16	5,60	0,68	0,68	0,056*
Canl18	9	5,79	0,77	0,83	0,024*
Canl19	10	5,31	0,57	0,62	0,083*
Canl20	20	6,94	0,77	0,88	-0,011
Média	13,56	5,59	0,72	0,75	0,020

F_{ST}: Valores com asterisco () foram considerados significativos a nível de 5%.

4.3 Diversidade genética populacional

Os valores de F_{ST} par-a-par entre as 21 populações (Tabela Suplementar 3 – Apêndices) foram significativos apenas para alguns pares de populações. As populações que mais se diferenciaram com relação às demais com base no F_{ST} foram Teodoro Sampaio/SP e Cruzeiro do Sul/AC.

Os resultados de diversidade genética apresentados na Tabela 7 demonstram que o número médio de alelos por localidade variou entre 4,06 a 6,81, sendo Fortaleza/CE, Bonito/MS e Patos/PB as cidades com a menor média de alelos, e Campo Novo do Parecis/MT, Rifaina/SP e Macapá/AP as com as maiores médias de alelos. As menores médias de alelos foram observadas em amostras de localidades nas quais poucos indivíduos, na quase totalidade machos foram coletados, enquanto aquelas com valores mais altos foram bem amostradas para machos e fêmeas. Esses resultados são condizentes com a riqueza alélica, a mais baixa foi encontrada em Fortaleza/CE (3,81) e a mais alta em Rifaina/SP (5,25), demonstrando que a riqueza alélica dentro de uma população está associada a amostragem.

O número de alelos privados variou entre 0 (Jataí/GO) e 8 (Rifaina/SP); porém, estes alelos mostraram-se em baixa frequência nas localidades em que foram detectados, o que justifica parcialmente a baixa diferenciação entre as localidades analisadas.

As heterozigosidades médias observada e esperada (Tabela 7) por população variaram entre 0,4062 – 1,000 e 0,6250 – 0,8750, respectivamente. Estes valores foram similares nas amostras de cada localidade analisada, exceto em Cruzeiro do Sul/AC e Matureia/PB, que apresentaram valores baixos de H_o . As populações com H_o igual a 1,000 são as que tinham apenas uma ou duas fêmeas amostradas.

O coeficiente de endogamia (F_{IS}) (Tabela 6) foi calculado com os dados das fêmeas, e variou entre -0,045 (Primavera do Leste/MT) a 0,469 (Cruzeiro do Sul/AC). O F_{IS} demonstrou valores significativos de endogamia nas populações Sacramento/MG, Vila Velha/ES, Grajaú/MA, Rifaina/SP, Matureia/PB, Cruzeiro do Sul/AC, Belém/PA e Jataí/GO. No entanto, não foi possível estimar o F_{IS} para 07 populações por não terem fêmeas amostradas ou possuir apenas uma no conjunto amostral.

O Índice de Garza-Williamson, uma medida de redução populacional com base em locos microssatélites, variou entre 0,27 a 0,40, com uma média de 0,30, demonstrando valores moderados para o índice, pois quanto mais baixo for o valor de M (Índice de Garza-Williamson), maior a redução no tamanho populacional; do contrário, quanto mais próximo de $M=1$, a população não sofreu redução de tamanho.

Tabela 7. Índices de diversidade genética para as populações analisadas quanto a: Número médio de alelos (N), Riqueza Alélica (AR), Número de Alelos Privados (N_{AP}), Heterozigosidade média observada (H_o) e esperada (H_e) em fêmeas, Índice de Garza-Williamson (M), Coeficiente de endogamia (F_{IS}) e seu respectivo valor de P (5%), estimados com dados das fêmeas.

Populações ¹	N	AR	N_{AP}	H_o	H_e	M	F_{IS}	P
TES	6,19	4,70	2	0,6667	0,7545	0,36	0,135	0,029
SCM	6,25	4,91	2	0,7188	0,8036	0,40	0,124	0,032
VLV	6,19	4,71	5	0,6250	0,7656	0,33	0,218	0,001
ITU	6,19	4,77	2	0,7031	0,7321	0,32	0,091	0,108
CNP	6,81	5,02	5	0,7500	0,8042	0,33	0,071	0,160
GRU	6,12	4,62	3	0,6250	0,7344	0,34	0,184	0,003
RIF	6,62	5,35	8	0,5365	0,8103	0,29	0,366	0,000
BOT	4,37	3,78	1	1,000	0,7500	0,33	-	-
JEQ	6,06	4,61	3	1,000	0,5000	0,33	-	-
PTS	4,19	3,87	1	-	-	0,27	-	-
SOR	5,75	4,91	3	-	-	0,32	-	-
MTR	5,50	4,61	3	0,500	0,6250	0,33	0,238	0,009
PRP	5,50	4,99	3	0,6875	0,7812	0,32	0,170	0,066
CZS	5,50	4,20	5	0,4063	0,7143	0,29	0,469	0,000
FOT	4,06	3,81	1	1,000	0,7500	0,34	-	-
ARA	6,06	4,62	3	1,000	0,8125	0,32	-	-
BLM	5,44	4,58	7	0,5625	0,7500	0,30	0,333	0,014
MCP	6,50	4,90	3	0,7188	0,7901	0,32	0,104	0,087
GYN	5,81	4,56	-	1,000	0,8750	0,36	-	-
JTI	6,12	4,98	-	0,7281	0,8002	0,35	0,102	0,049
PVA	5,75	4,42	2	0,8214	0,6979	0,32	-0,045	0,870
Média	5,65	4,58	2,95	0,7250	0,6786	0,33		

¹ Significado das siglas das populações: SCM: Sacramento/MG; VLV: Vila Velha/ES; ITU: Ituaçu/BA; CNP: Campo Novo do Parecis/MT; GRU: Grajaú/MA; RIF: Rifaina/SP; BOT: Bonito/MS; JEQ: Jequié/BA; PTS: Patos/PB; SOR: Sorriso/MT; MTR: Matureia/PB; PRP: Pirapora/BA; CZS: Cruzeiro do Sul/AC; FOT: Fortaleza/CE; ARA: Arame/MA; BLM: Belém/PA; MCP: Macapá/AP; GYN: Goiânia/GO; JTI: Jataí/GO.

4.4 Estrutura genética populacional

Uma análise de variância molecular (AMOVA) foi realizada para dois níveis hierárquicos, entre e dentro de populações, para verificar como a variação genética estava particionada. A AMOVA feita no GenAlEx (Figura 11) demonstrou que a maior parte da variação está dentro de populações (95%), permitindo afirmar que boa parte da variação presente em uma determinada localidade é compartilhada com as demais localidades analisadas, haja vista que os resultados apresentam 5% de variação entre localidades, dado o valor de F_{ST} igual a 0,055 (valor $P = 0,001$) estimado. Corroborando essa análise, os resultados do Arlequin (Tabela 8), demonstraram valores da porcentagem de variação de 5,44% entre as populações, um valor relativamente baixo, porém significativo ($F_{ST} = 0,05437$, $p < 0,0001$).

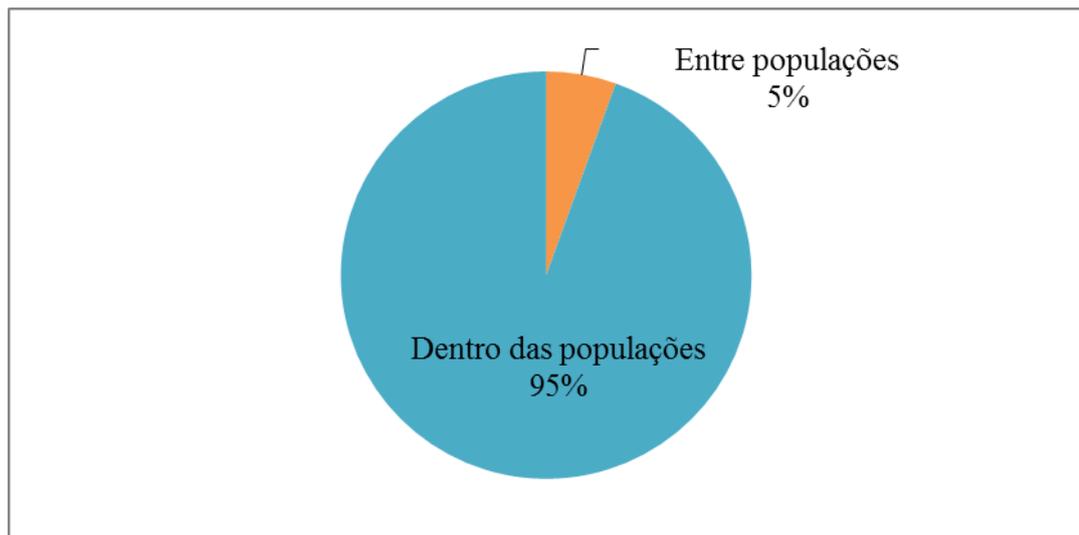


Figura 11. Análise de variância molecular (AMOVA) implementada no programa GenAlEx para 21 amostras urbanas de *C. analis* para verificar a distribuição da variação genética dentro e entre populações.

Tabela 8. Análise de variância molecular (AMOVA) implementada no programa Arlequin para verificar a distribuição da variação genética dentro e entre populações.

Fonte de variação	d.f	Componentes da variância	% da variação
Entre populações	20	0,35105 Va	5,44
Dentro das populações	421	6,10575 Vb	94,56
Total	441	6,45679	100
Índice de Fixação	F_{ST}:	0,05437	P<0,0001

A análise de grupos genéticos no Structure com o modelo *no admixture* através de inferência bayesiana indicou pelo método de Evanno et al., (2006) a existência de prováveis valores de $K=3$ para machos (Figura 12A, 12B e 12C) e $K=2$ ou $K=4$ para fêmeas (Figura 13A, 13B e 13C). O modelo *admixture* demonstrou a existência de $K=2$ para fêmeas, porém com um gráfico de barras demonstrando a possibilidade de existência de apenas um grupo genético, enquanto que para machos os resultados são semelhantes aos encontrados no modelo *no admixture* (Figura 14). Os resultados demonstram que a variação genética encontrada está concentrada principalmente dentro das populações, com baixa variação entre as mesmas, conforme resultado da AMOVA. O resultado da estruturação para machos demonstra que em todas as localidades amostradas eles compartilham três grupos genéticos caracterizados, ou seja, não há um grupo genético especificamente apresentado por machos de uma certa região geográfica.

Com relação a estruturação das fêmeas (13C e 14A), quando considerados $K=4$ (*no admixture*) e $K=2$ (*admixture*), verifica-se que a população de Matureia/PB é a que mais se diferencia das demais. Dessa forma, uma melhor amostragem na área da Caatinga nos estados de Rio Grande do Norte, Alagoas e Pernambuco talvez pudesse gerar um cenário melhor para a região Nordeste. Quando considerado $K=2$ (*no admixture*), assim como em machos, todas as localidades analisadas estão constituídas dos dois grupos, mas em proporções diferentes. As localidades de Cruzeiro do Sul/AC e Matureia/PB apresentam em quase sua totalidade o grupo representado em azul, enquanto as demais localidades possuem em maior proporção o grupo genético representado pelo laranja.

Os dados analisados em conjunto (machos e fêmeas) apresentam um provável valor de $K=2$ para o modelo *no admixture* (Figura 15A, 15B e 15C), porém, essa análise pode ter sido enviesada pela diploidização dos machos, uma vez que o programa Structure não analisa machos haploides e fêmeas diploides em um mesmo arquivo de entrada. A diploidização dos machos aumenta a proporção do grupo representado em azul se sobressaindo sobre o grupo na cor laranja. Isso pode ser verificado nas amostras com número de fêmeas igual a 4, o grupo laranja permanece mais bem representado. Os mesmos dados foram testados sob o modelo *admixture* resultando em um provável valor de $K=2$ com o gráfico de barras demonstrado na figura 16.

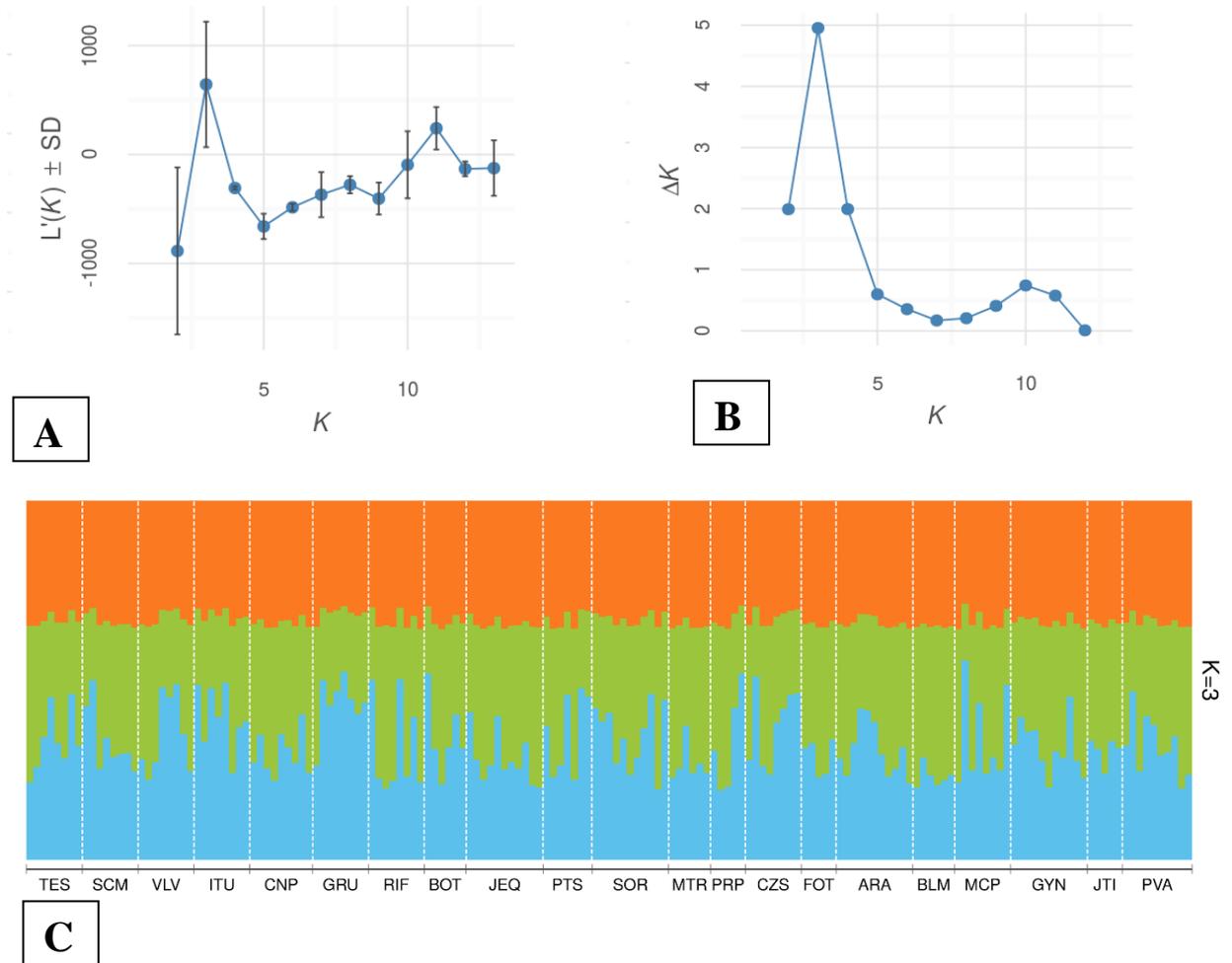


Figura 12. Resultado da análise do programa Structure *no admixture* para machos de *C. analis*. Em A, pode-se observar a probabilidade de $\ln(P(X/K))$ de $K=3$. Em B, observa-se o resultado do teste de Evanno com o valor de $K=3$. Em C, estruturação² sugerida para machos em $K = 3$.

² Cada linha vertical representa a probabilidade do genoma de um indivíduo ser atribuído a um grupo genético específico denotado por cores diferentes.

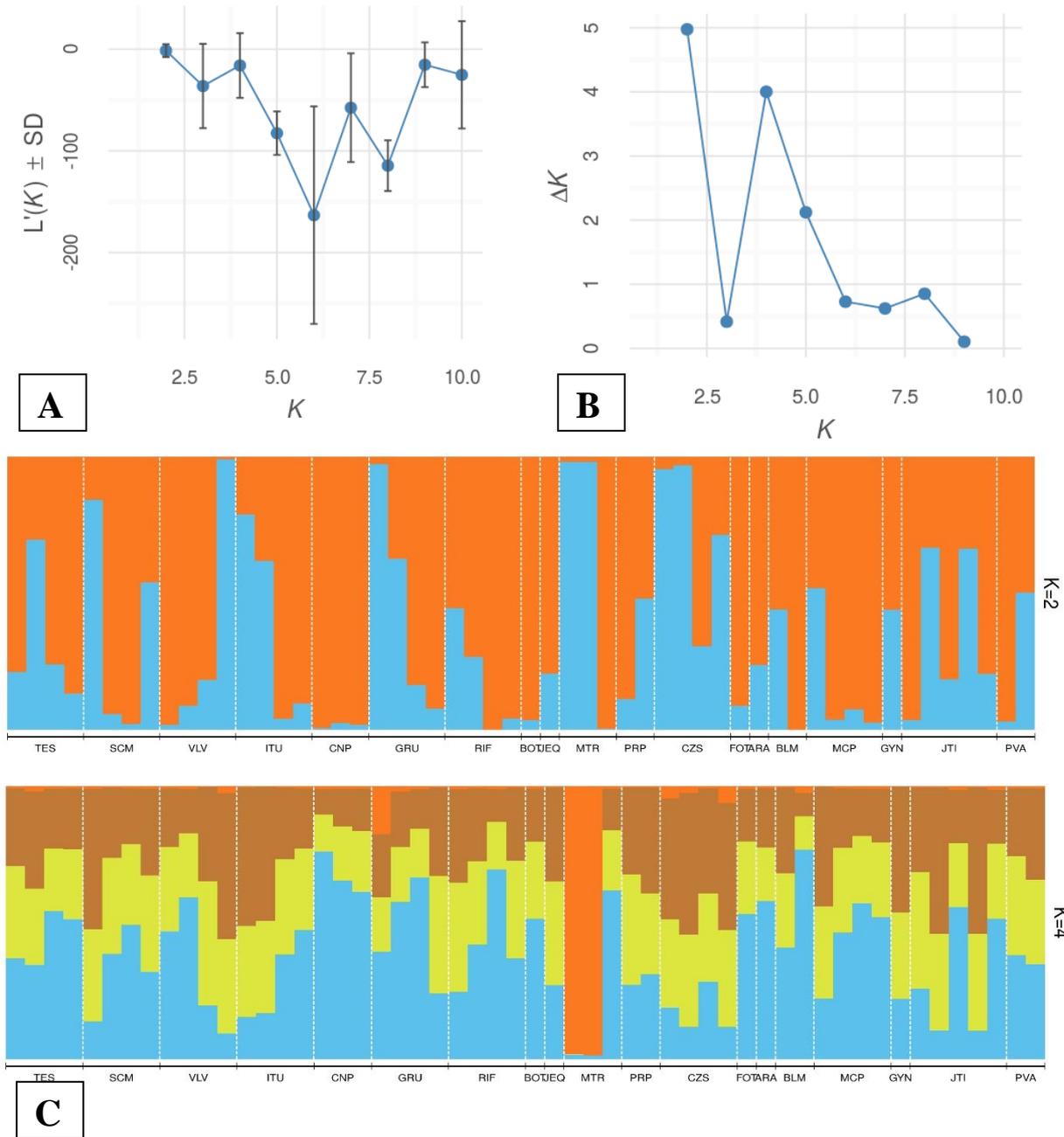


Figura 13. Resultado da análise do programa Structure *no admixture* para fêmeas de *C. analis*. Em A, pode-se observar a probabilidade de $\ln(P(X/K))$ descrita por Pritchard et al., (2000). Em B, observa-se o resultado do teste de Evanno com o valor de K mais provável. Em C, estruturação sugerida em $K = 2$ e $K=4$.

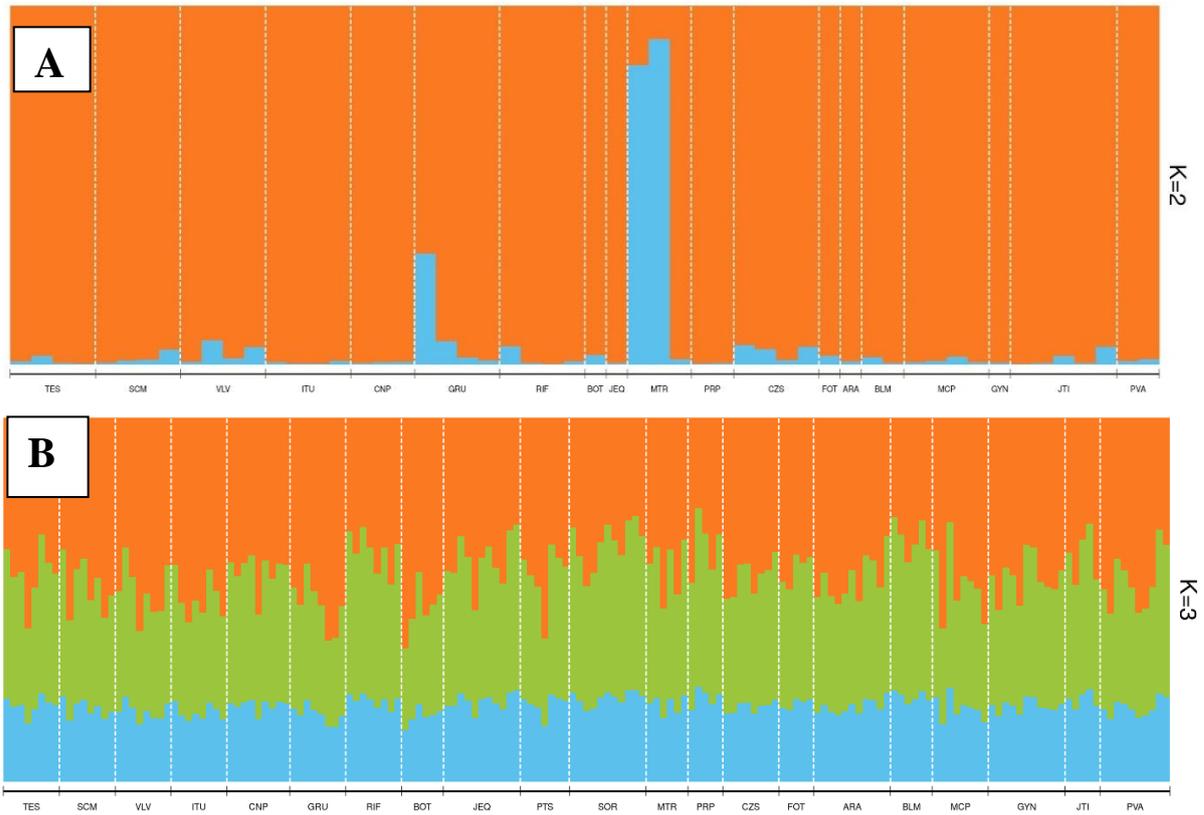


Figura 14. Gráfico de barras de estruturação genética. Em A, resultado da análise do programa Structure modelo *admixture* para fêmeas (K=2). Em B, a análise do programa Structure modelo *admixture* para machos (K=3).

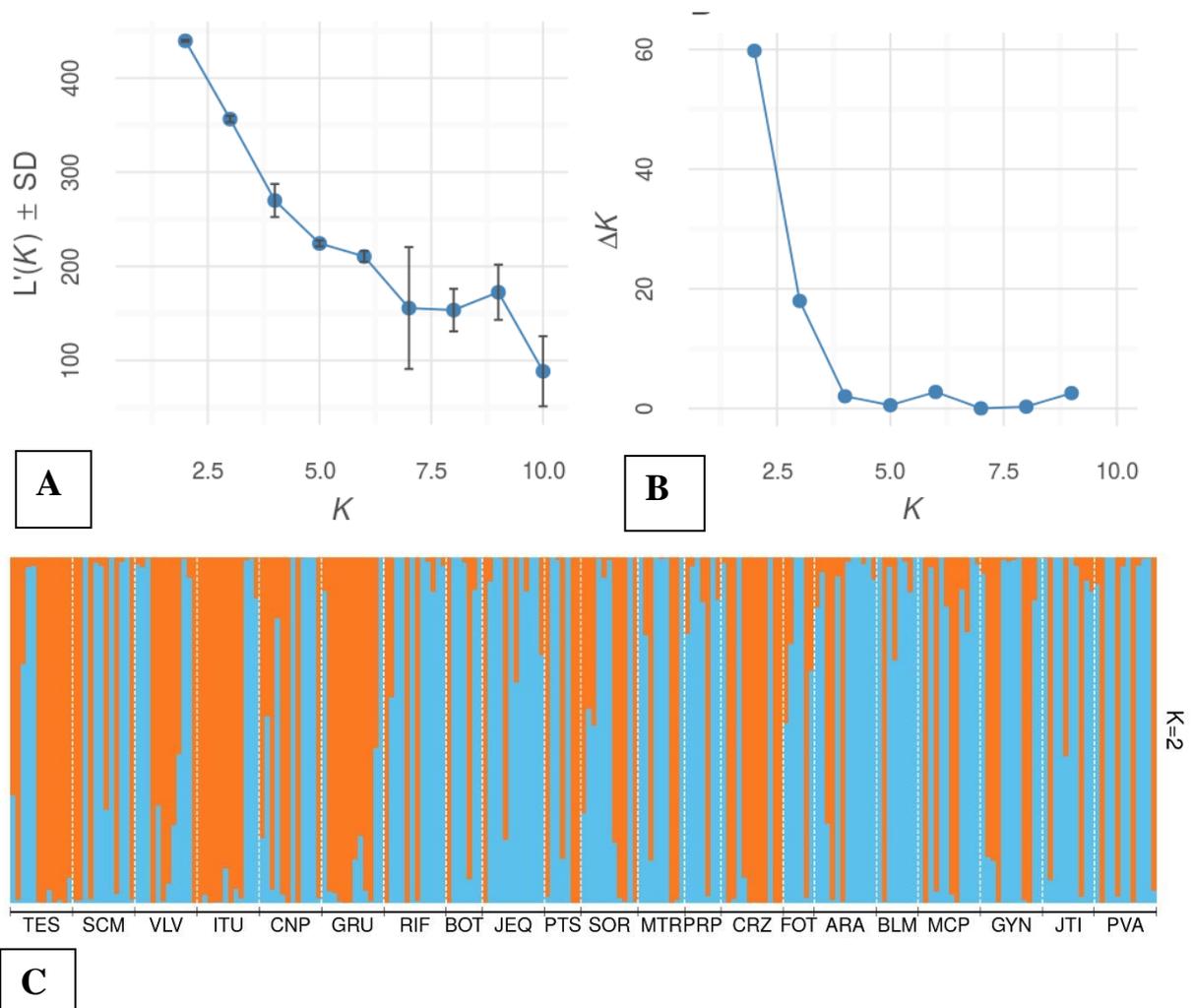


Figura 15. Resultado da análise do programa Structure *no admixture* para machos e fêmeas de *C. analis*. Em A, pode-se observar a probabilidade de $L_n(P(X/K))$ descrita por Pritchard et al., (2000). Em B, observa-se o resultado do teste de Evanno com o valor de $K=2$. Em C, estruturação sugerida em $K = 2$.

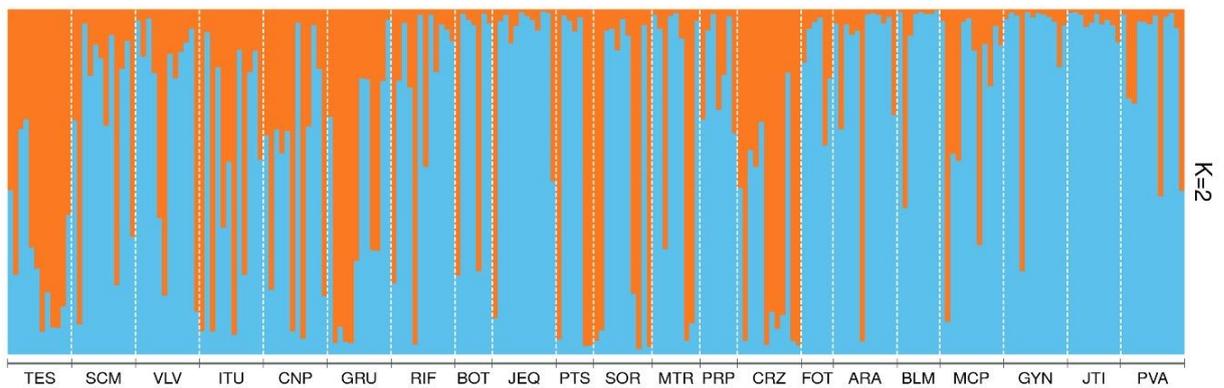


Figura 16. Resultado da análise do programa Structure *admixture* para machos e fêmeas de *C. analis* com $K=2$.

Os modelos testados pelo programa TESS para fêmeas podem ser visualizados na Figura 17. O modelo que apresenta o menor valor de DIC é o mais indicado para os dados.

Dessa forma, verifica-se que os modelos são bastante concordantes, não diferindo muito entre os valores de DIC. O modelo escolhido para representar a estruturação foi o modelo *with admixture* CAR cúbico (CARCub). O modelo sugeriu um valor de $K = 5$ (Figura 18), porém, verifica-se que o programa força uma estrutura genética pouco provável, pois ao considerar as análises do programa Structure, pode-se observar que não existe a formação de grupos genéticos bem definidos, mas sim, uma variação genética dentro das populações. Na figura que mostra os grupos formados pelo programa TESS, o principal grupo está representado em verde presente em todas as localidades. Quando se observa os demais grupos, deduz-se que são grupos genéticos não consistentes, pois são pouco representativos. Com base nesta figura, pode-se verificar que as localidades Campo Novo do Parecis/MT, Grajaú/MA, Rifaina/SP, Cruzeiro do Sul/AC, Belém/PA e Macapá/AP são diferenciadas das demais. Com base na figura 18 ($K=2$ e $K=5$) verifica-se que os resultados do TESS para fêmeas são concordantes com os resultados para o modelo *admixture* do programa Structure.

Os resultados do TESS para machos também foram testados sob nove modelos, demonstrando que o melhor modelo para representar a estruturação é o modelo *With admixture* BYM Constante (BYMCon) (Figura 19), pois este modelo apresenta o menor valor de DIC e é o que permanece estável após o valor de $K = 6$, apresentando este como valor de K mais provável (Figura 20). O gráfico de barras do programa TESS (Figura 20) demonstra, como o resultado de fêmeas já indicara, ausência de grupos genéticos isolados.

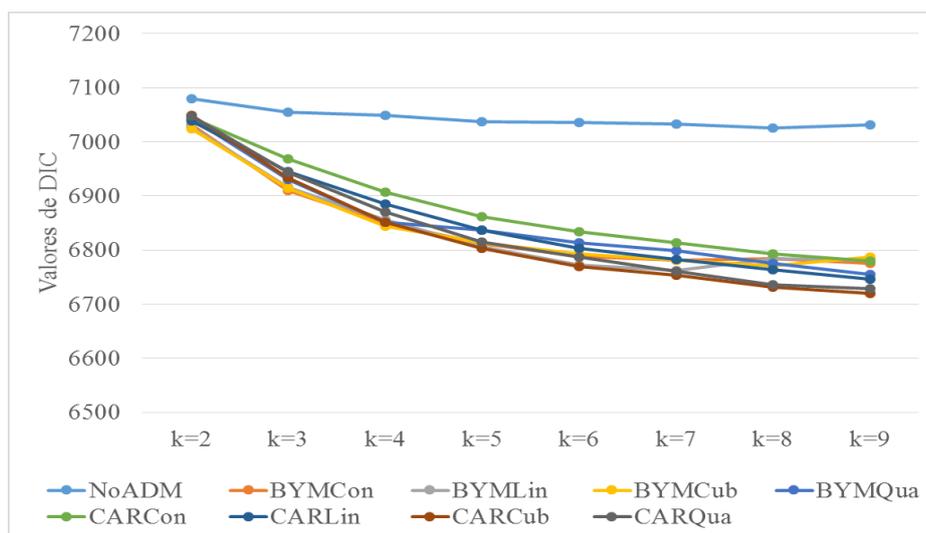


Figura 17. Modelos testados no programa TESS para verificar a estrutura genética de *C. analis* em fêmeas. Foram testados 9 modelos: *Without admixture* (NoADM), *With mixture* BYM Constante (BYMCon), *With mixture* BYM Linear (BYMLin), *With mixture* BYM Quadrático (BYMQua), *With mixture* BYM Cúbico (BYMCub), *With mixture* CAR Constante (CARCon), *With mixture* CAR Linear (CARLin), *With mixture* CAR Quadrático (CARQua) e *With mixture* CAR Cúbico (CARCub).

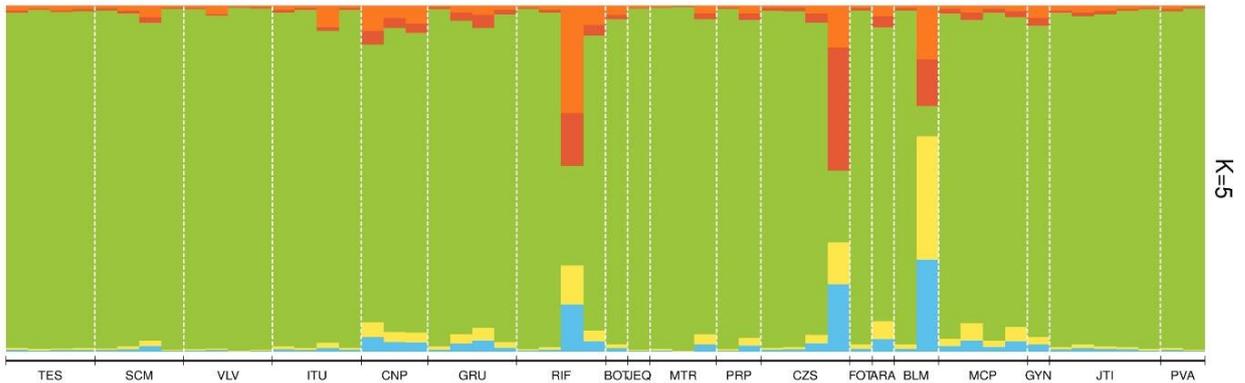


Figura 18. Gráfico de barras da estruturação genética ($K=5$) detectada pelo programa TESS modelo *With admixture* CAR Cúbico para as fêmeas.

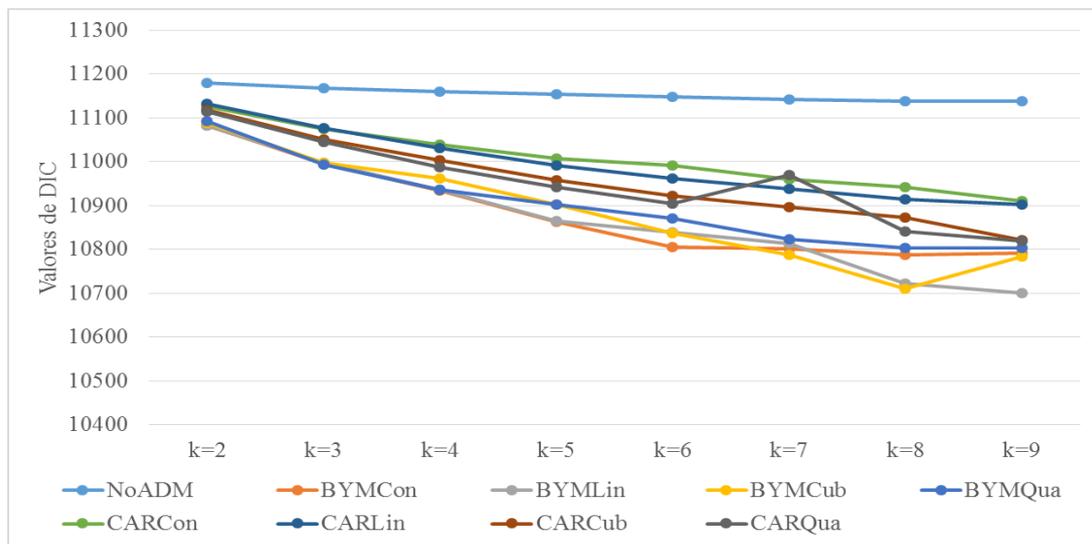


Figura 19. Modelos testados no programa TESS para verificar a estrutura genética de *C. analis* em machos. Foram testados 9 modelos: *Without admixture* (NoADM), *With mixture* BYM Constante (BYMCon), *With mixture* BYM Linear (BYMLin), *With mixture* BYM Quadrático (BYMQua), *With mixture* BYM Cúbico (BYMCub), *With mixture* CAR Constante (CARCon), *With mixture* CAR Linear (CARLin), *With mixture* CAR Quadrático (CARQua) e *With mixture* CAR Cúbico (CARCub).

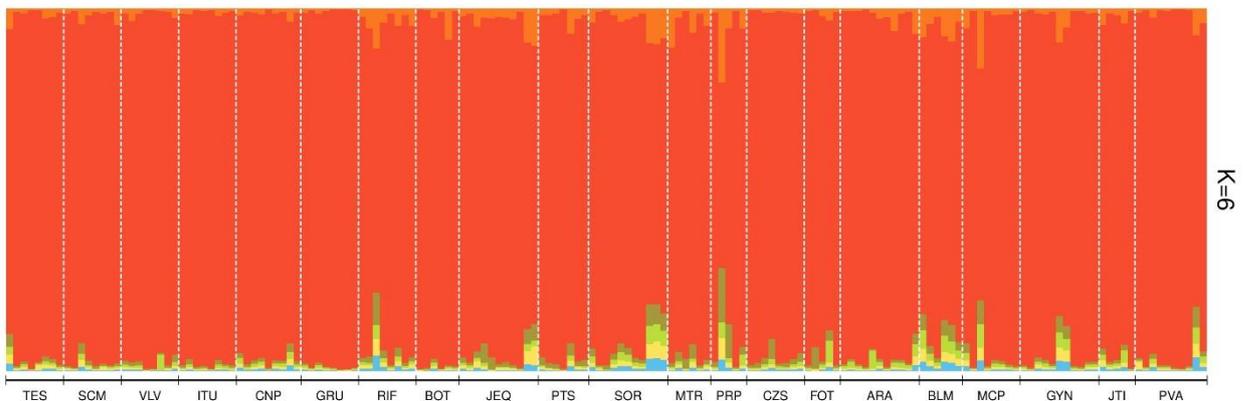


Figura 20. Gráfico de barras de estruturação genética ($K=6$) detectados pelo programa TESS modelo *With admixture* BYM Constante para machos.

A análise multivariada DAPC foi realizada utilizando 80, 95 e 120 componentes principais (PCs) para fêmeas, machos e machos e fêmeas, respectivamente, que estabeleceram dois componentes discriminantes que foram usados para finalizar a análise e determinar os valores de BIC (*Bayesian Information Criterion*) determinando o número mais provável de grupos genéticos (valor de K) de acordo com o menor valor de BIC. O valor de K mais provável para fêmeas ($K = 1$) pode ser verificado na Figura 21A, para machos ($K = 2$) na Figura 21B e para machos e fêmeas ($K = 3$) na Figura 21C.

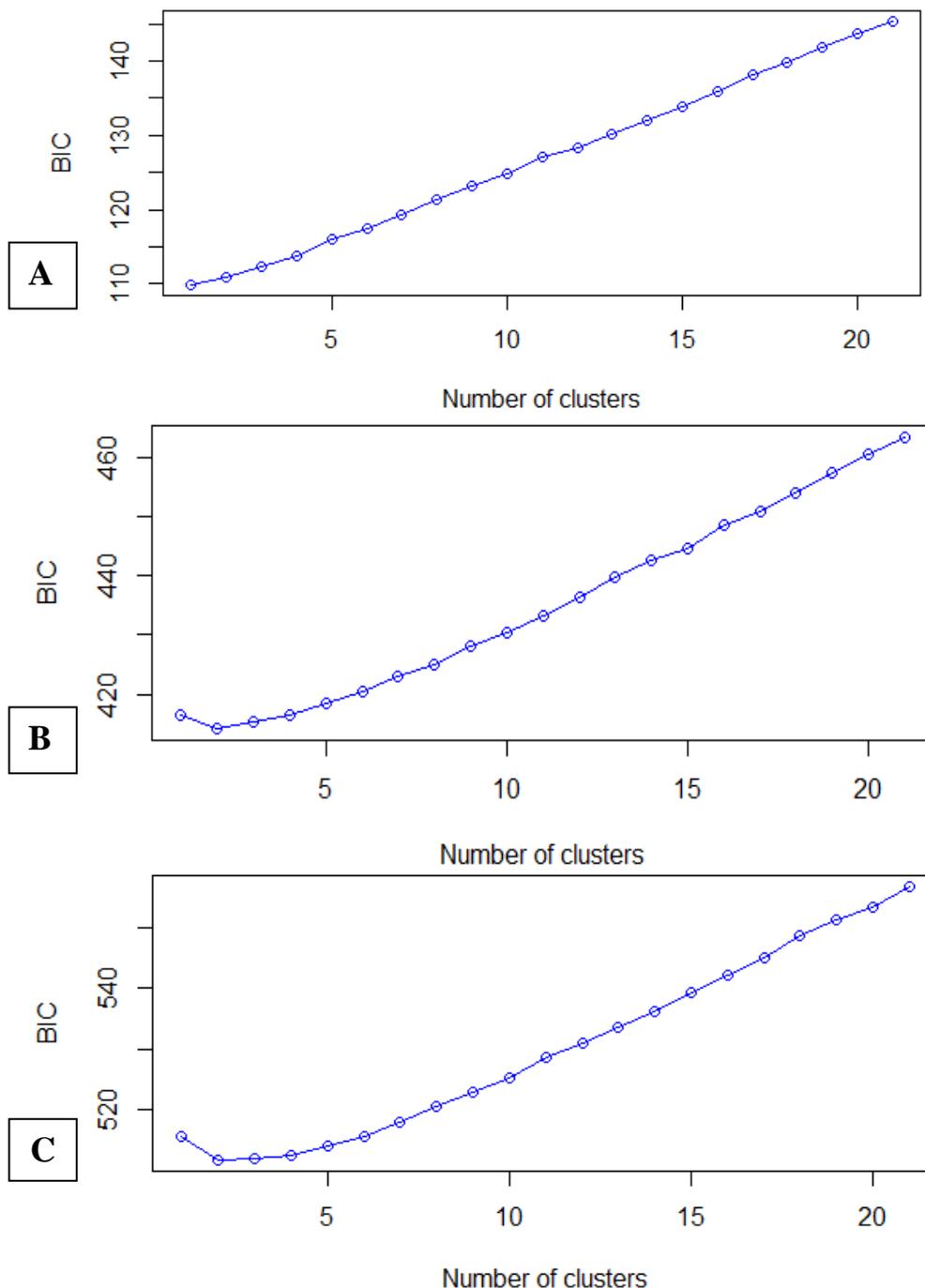


Figura 21. Valor de BIC versus número de clusters. Em A, fêmeas. Em B machos. Em C, machos e fêmeas.

Com base na figura 21A, verifica-se que para as fêmeas não há a formação de grupos genéticos, haja vista que não há a formação de uma curva com menor valor de BIC em um provável valor de K, demonstrando por esta análise que as amostras de fêmeas são geneticamente semelhantes entre as localidades. Para machos (Figura 21B), o menor valor de BIC demonstra um provável valor de $K = 2$; por outro lado, para machos e fêmeas em conjunto (Figura 21C), há a indicação de um provável valor de K entre 2 e 3. Desta forma, foi realizada uma DAPC para um valor de $K=3$ considerando ser este o valor de K mais provável de acordo com o resultado da análise de estruturação dos machos no programa Structure.

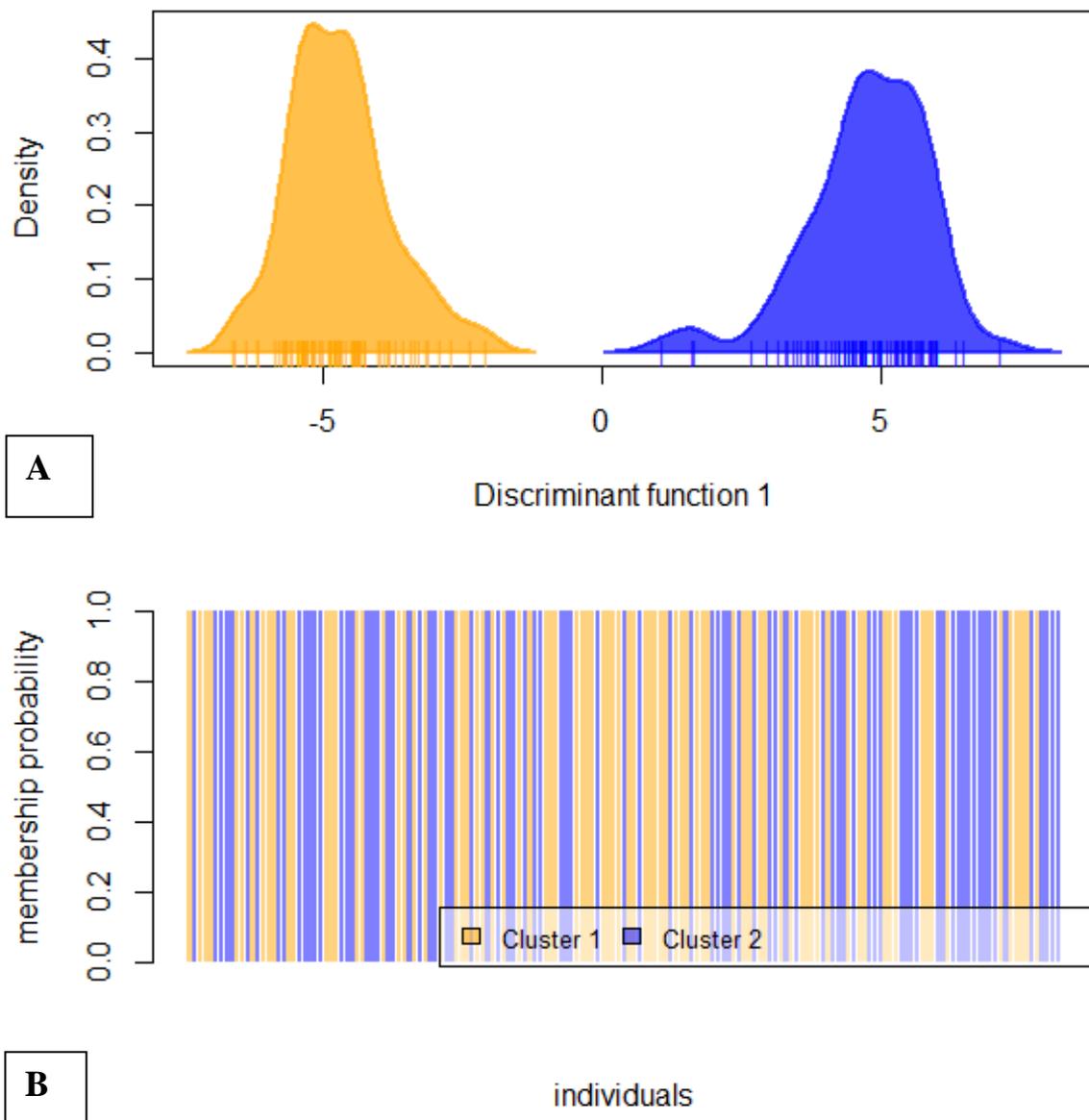


Figura 22. Análise Discriminante de Componentes Principais para machos. Em A, DAPC considerando $K=2$. Em B, distribuição dos indivíduos em dois grupos representados em laranja e azul.

Na Figura 22A é apresentado a DAPC para machos (K=2) mostrando a formação de dois grupos e na Figura 22B é apresentada a atribuição dos indivíduos a um dos grupos genéticos formados, mostrando que estes são atribuídos a um ou outro grupo genético (barras completas); porém, na maioria das populações aparecem os dois grupos genéticos definidos. A Figura 23A mostra o gráfico de dispersão dos grupos genéticos (K=3) para machos e fêmeas de acordo com a DAPC. Os grupos genéticos formados têm uma pequena não são delimitados dentro dos quadrantes, ou seja, não foram grupos genéticos distribuídos isoladamente. Pode-se verificar na Figura 23B como os indivíduos (machos e fêmeas) foram distribuídos em relação aos grupos genéticos determinados pela DAPC, em que a maioria das localidades apresentam os três grupos genéticos formados.

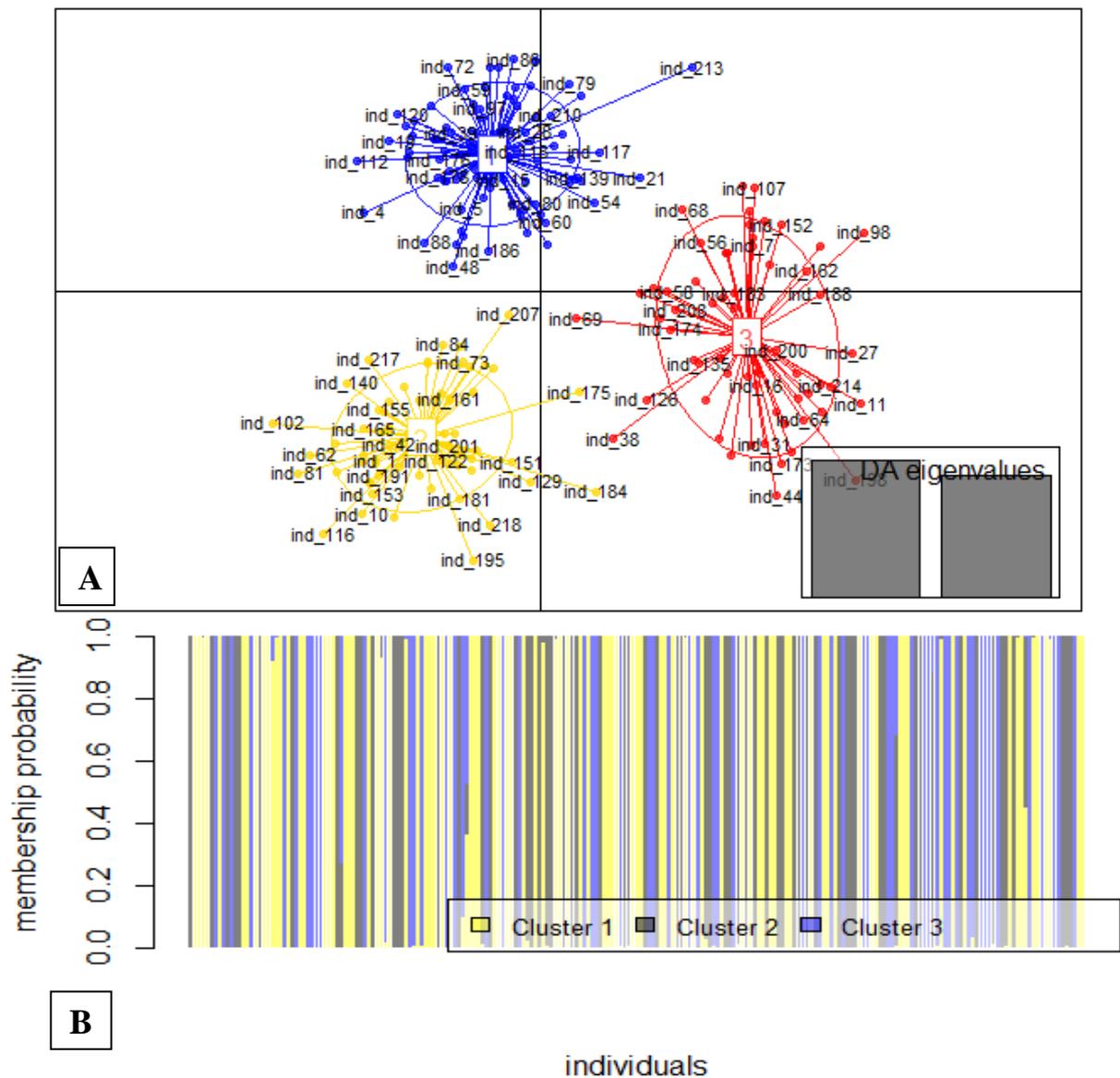


Figura 23. Análise Discriminante de Componentes Principais para machos e fêmeas. Em A, gráfico de dispersão dos grupos genéticos (K=3). Em B, distribuição dos indivíduos em cada grupo genético representado pelas cores amarela, azul e cinza.

Os resultados da estruturação definidas pelo Structure, TESS e DAPC são concordantes entre si para fêmeas e machos, exceto pela DAPC que mostrou resultado diferente ($K=2$). A variação existente, conforme a AMOVA (5,56%) não diferencia grupos genéticos geográficos, pois a maior parte da variação (94%) está distribuída dentro das populações, sendo compartilhada por todas elas.

Considerando os resultados foi realizada uma AMOVA para fêmeas e machos separadamente (Tabela 9). Os resultados indicam que o F_{ST} entre as populações de fêmeas foi condizente com o AMOVA total. Com relação aos machos, verifica-se um $F_{ST} = 0,10$ entre todas as populações amostradas, porém, quando considerados os dois grupos genéticos inferidos pela DAPC, um valor de $F_{ST} = 0,00$ foi estimado, demonstrando ausência de diferenciação entre os grupos.

Tabela 9. Resultado da AMOVA para fêmeas, machos e machos com $K=2$ (grupos definidos pela DAPC) para dois níveis hierárquicos, dentro e entre populações.

Fêmeas			
Fonte de variação	d.f	Componentes da variância	% da variação
Entre populações	13	0,355	5%
Dentro das populações	84	6,106	95%
Total	97	6,461	100%
Índice de Fixação	$F_{ST} = 0,055$	$P < 0,001$	
Machos			
Entre populações	20	385,860	10%
Dentro das populações	146	3361,066	90%
Total	166	3746,926	100%
Índice de Fixação	$F_{ST} = 0,103$	$P < 0,007$	
Machos $K=2$			
Entre populações	1	0,000	0%
Dentro das populações	165	3742,447	100%
Total	166	3742,447	100%
Índice de Fixação	$F_{ST} = -0,007$	$P < 0,666$	

De posse dos dados de estruturação genética populacional, procedeu-se a análise de dados sociogenéticos intranidais para aumentar o conhecimento de dados genéticos sobre a espécie e confirmar dados não conclusivos sobre o sistema de acasalamento da espécie, conforme já reportando com dados de alozimas e por estudos de observação de nidificação da espécie.

4.5 Estrutura sociogenética intranidal

O mecanismo de determinação sexual das abelhas permitiu confirmar o sistema de acasalamento do tipo monândrico/monogínico de *C. analis*. Após a leitura dos genótipos foi montada uma tabela com todas as proles analisadas de um mesmo ninho, e com isso inferir os genótipos dos pais da prole (Tabela 10). Os genótipos parentais foram definidos por genética mendeliana, segundo a qual em um ninho há três possíveis genótipos, um do macho genitor e dois da fêmea genitora. Todos os alelos dos filhos machos são provenientes da mãe, enquanto o genótipo das filhas tem alelos da mãe e do pai, dessa forma, possibilitando determinar os genótipos parentais.

Com a delimitação dos genótipos parentais para os 10 locos conforme a Tabela 9, os outros 7 ninhos (Tabela Suplementar 5 – ver Apêndices) também foram analisados, confirmando o sistema de acasalamento do tipo monandria para *C. analis*.

Tabela 10. Definição da genética familiar para um ninho (Ninho 1) com utilização de 10 locos microssatélites.

Locos	Sexo da prole	Genótipo da prole	Genótipo Pai	Genótipo Mãe
Loco 5	M	209		
	F	213	215	
	M	209		213
	M	215		209
Loco 9	M	242		
	F	242	246	
	M	242		246
	M	240		240
Loco 11	M	104		
	F	102	106	
	M	104		102
	M	104		104
Loco 1	M	232		
	F	232	232	
	M	232		232
	M	232		232

	M	116				
Loco 6	F	119	128		128	119
	M	119			116	
	M	119				
	M	195				
Loco 14	F	193	193		193	195
	M	193			193	
	M	193				
	M	301				
Loco 2	F	301	303		303	301
	M	301			301	301
	M	301				
	M	212				
Loco 7	F	197	215		197	215
	M	215			212	
	M	215				
	M	221				
Loco 4	F	221	227		227	221
	M	221			221	221
	M	221				
	M	0				
Canl 10	F	252	252		252	248
	M	248				
	M	248				

5 DISCUSSÃO

5.1 Área potencial de distribuição de *Centris (Heterocentris) analis*

Os resultados da modelagem de nicho demonstraram que a distribuição potencial da espécie ocorre desde a América do Norte, no México e perpassa os países da América do Sul, chegando até o estado do Paraná no Brasil. As áreas potenciais definidas são corroboradas pela literatura, exceto por Silveira e colaboradores (2002) em que afirmam que a espécie

ocorre desde o Texas, nos EUA; no entanto, apenas um ponto de ocorrência foi encontrado para a espécie no ano de 1960 e nenhum registro posterior a esta data foi reportado. Dessa forma, verifica-se que existe uma grande lacuna que vai do Texas até a área com registros de ocorrência (México).

Os resultados indicam que a espécie tem um grande potencial de distribuição na América Central e América do Sul. Isso pode ser atribuído ao porte médio da espécie que lhe confere boa capacidade de dispersão quando comparada a indivíduos de menor porte, conforme relatado por Gathmann e Tschardt (2002). Esse padrão potencial de distribuição pode ser esperado para várias espécies de *Centris* (*Hemisiella*), bem como para outras espécies de gêneros que ocorrem na América do Sul (Velez et al., 2017). Somado a isso, a grande amplitude de nicho da espécie está relacionada à sua plasticidade em habitar diferentes locais. Esta plasticidade possivelmente é explicada pelo poliletismo (Dórea et al., 2010), ou seja, seu caráter generalista na coleta de pólen e na utilização de substratos para nidificação, cavidades preexistentes encontradas nos mais diferentes sítios.

A baixa probabilidade de ocorrência da espécie nos estados do Sul do Brasil e em países da América do Sul como Argentina, Uruguai e Chile pode estar associada à diferença de clima entre essas regiões e a região intertropical próxima à linha do Equador. Como a modelagem foi feita dados de presença, nenhum ponto usado foi proveniente dessas regiões; assim, a variação climática, relevo, ou mesmo vegetação foram usadas como condições favoráveis para a espécie.

Outra área que merece atenção especial, com base nas áreas potenciais de distribuição, é a Amazônia. Esse imenso bioma brasileiro possui uma variação climática muito alta, pois é banhada por inúmeros rios e tem grandes mudanças de temperatura ao longo do dia e de umidade relativa do ar. Dessa forma, a modelagem de nicho não mostrou alta probabilidade de encontrar a espécie na Amazônia, com apenas quatro pontos amostrados (Cruzeiro do Sul/AC, Belém/PA, Macapá/AP e Manaus/AM). Considerando o tamanho da área e a variação climática, mais pontos precisam ser amostrados para que a área seja considerada como área potencial de distribuição da espécie, pois segundo Williams et al., (2007) o clima age como controle primário da distribuição de espécies e processos ecossistêmicos e, dessa forma, as mudanças climáticas podem promover a formação de novas associações de espécies e outras surpresas ecológicas.

Dada a estimativa de distribuição da espécie por Silveira et al., (2002) e Catálogo Moure, a modelagem contribuiu para verificar que a espécie está distribuída por todo o Brasil e que estudos posteriores sobre Genética, Evolução, Biogeografia e Filogeografia devem

incluir estas áreas com esforços de coleta em áreas ainda não amostradas e com potencial de distribuição, tais como, o estado do Paraná no Brasil, no Equador, Colômbia, Peru, Bolívia, Paraguai e Chile. Assim como recomenda Velez e colaboradores (2017), ao modelar o nicho de uma espécie de Centridini que, dada a completa falta de informações da distribuição, qualquer área potencial disponível é útil para propor novas áreas em que pesquisas de campo precisam ser implementadas para se obter informações mais recentes sobre a espécie.

Conforme o Catálogo Moure, a espécie ocorre até no estado brasileiro do Paraná; porém, nenhum espécime foi encontrado neste estado nas consultas online, justificando assim a baixa chance de ocorrência da espécie no referido estado por ausência de pontos na amostragem. Vale ressaltar que tentativas de coleta foram realizadas recentemente nos estados de Santa Catarina e Rio Grande do Sul, porém não foi observado ou coletado algum espécime nas flores de *Tecoma stans* ou *Malpighia emarginata*. Dessa forma, novos estudos devem prever reiterados esforços de coleta nestas áreas, pois a modelagem mostra áreas potenciais, mas não restringe a espécie a estes habitats.

5.2 Diversidade e estrutura genética populacional

Os 16 locos microssatélites espécie-específicos utilizados neste trabalho demonstraram alta variabilidade, como indicado pela riqueza alélica, alto número de alelos e alta heterozigosidade esperada para cada loco, exceto para o loco Can12 que apresentou o mais baixo número de alelos (6) e um valor de $He = 0,48$. Estes locos estão sendo utilizados pela primeira vez, podendo ser testados para uma ampla amostragem da espécie para a qual foram desenvolvidos.

A alta diversidade dos locos microssatélites deve estar relacionada ao elevado número amostral e ao tamanho da área abrangida neste trabalho. O trabalho incluiu 21 localidades de quatro regiões geopolíticas brasileiras com diferentes fitofisionomias e variação climática, um número de 221 indivíduos, machos e fêmeas, o que aumenta a probabilidade de encontrar maior variação para os locos analisados. Além disso, os locos apresentaram uma baixa taxa de erro de genotipagem, ausência de alelos nulos, bem como não demonstraram desequilíbrio de ligação ($p < 0,05$ nas análises par a par).

Os locos também poderão ser utilizados para outras espécies do gênero *Centris* da tribo Centridini por meio da transferabilidade dos locos, pois os mesmos foram testados para alguns indivíduos de *Centris (Hemisiella) tarsata* e *Centris trigonoides*, tendo resultados positivos; contudo, ainda são necessários alguns ajustes nos protocolos de PCR. É bastante

comum a transferabilidade de locos, conforme demonstrado por Dessi (2018) ao utilizar locos prospectados em *Partamona rustica* para análises em *Partamona helleri*.

As heterozigosidades observada e esperada neste estudo foram similares às encontradas para outras espécies de abelhas. Resultados semelhantes foram encontrados por Vickruck e colaboradores (2017) em *Xylocopa virginica* vivendo em ambientes com distúrbios antropogênicos e Francisco e colaboradores (2016) em *Bombus morio* de áreas de ilhas e continentes de região neotropical. Com relação ao Coeficiente de endogamia, baixos porém significativos, era valores esperados, considerando que as fêmeas são filopátricas, e assim serem mais homogêneas. Além disso, duas populações não possuíam fêmeas amostradas e outras apenas uma fêmea, não sendo possível estimar valor de F_{IS} para estas populações. O índice de Garza-William indicado para medir redução de tamanho populacional com microssatélites, demonstrou que as populações estão em redução populacional, com valores abaixo de 0,50.

Os resultados de estruturação genética confirmaram parcialmente a hipótese inicialmente formulada de que as populações urbanas de *C. analis* não estão estruturadas geneticamente ao considerar os resultados de machos e fêmeas em separado. Os resultados do Structure no modelo *no admixture* para machos mostraram haver igual atribuição a três grupos genéticos, enquanto que para fêmeas este resultado mostrou-se diferente, mesmo não formando grupos genéticos isolados. Neste sentido, ao aumentar o valor de K, o número de cores aumenta, demonstrando que cada indivíduo não é completamente atribuído a um único grupo genético, formando assim vários grupos genéticos dentro das populações. Além disso, ao analisar o modelo *admixture*, as fêmeas apresentam a formação de apenas um grupo, com exceção de alguns indivíduos que mostraram o padrão visto no modelo *no admixture*. Vale ressaltar que ao verificar a DAPC das fêmeas, uma análise mais sensível da divergência entre grupos genéticos, não há a formação de grupos genéticos. De igual forma, as análises corridas no programa TESS indicaram um valor de $K = 6$; porém, os resultados observados são indivíduos atribuídos por completo a único grupo para quase todos os indivíduos de todas as populações, exceto por algumas localidades que mostraram indivíduos atribuídos a mais de um grupo, porém não forma um segundo grupo genético.

Ao analisar os resultados dos machos, verifica-se que a formação de grupos genéticos está distribuída dentro de populações, com base na visualização do gráfico de barras do Structure e do TESS. Os resultados da estruturação genética dos machos pelo programa Structure nos modelos *no admixture* e *admixture* podem levar a conclusões equivocadas sobre a formação de grupos genéticos. Os resultados indicam que a variação genética que nos é

demonstrada através de grupos genéticos é compartilhada pelos indivíduos das diferentes populações. Esses resultados podem ser comparados com os obtidos pelo programa TESS e verificar que neste, não houve a formação de agrupamentos genéticos mesmo com um $K=6$, pois a maior parte do gráfico está representada por uma única cor. Resultados divergentes foram encontrados na DAPC, que maximiza as diferenças existentes, e demonstra haver a formação de dois grupos genéticos, separando inclusive os indivíduos (barras completas) em dois grupos; contudo, ao verificar os resultados da AMOVA para dois grupos, não houve valor significativo de F_{ST} diferenciando os dois grupos.

Resultados semelhantes para este padrão do Structure encontrado nos machos foram reportados por Galindo-Cardona e colaboradores (2013) em populações de abelhas africanizadas. Além disso, devido à ausência de trabalhos sobre estrutura genética de abelhas solitárias, após busca de relatos em outros organismos com o mesmo padrão de gráficos do Structure observado aqui, encontramos os trabalhos de Thompson e colaboradores (2015) em baleia do bico cinza, um cetáceo de oceano profundo e de Dorji e colaboradores (2018) em três raças de cavalos tradicionais do Butão. O primeiro trabalho foi realizado com uma região controle mitocondrial (530 pb) e 12 locos microssatélites para 94 indivíduos da Austrália e Nova Zelândia, em uma área de distribuição de 6.000 km, coletados em um período de 22 anos. Os resultados demonstraram alta variação genética intrapopulacional e perda de diversidade genética entre as populações em marcadores de herança materna e biparental, quando comparado com outros cetáceos. O segundo trabalho foi realizado com genotipagem de 74 cavalos para 29 locos microssatélites, resultando em uma rica diversidade genética, apesar de uma diferenciação genética muito baixa entre as raças, com um valor de F_{ST} de 0,6%.

Os resultados dos dados em conjunto, machos e fêmeas, permite inferir que os machos por serem hemizigotos podem influenciar nos resultados das análises ao serem diploidizados. Isso pode ser verificado ao analisar os resultados do Structure que tem como premissa que as populações devem estar em Equilíbrio de Hardy-Weinberg; desta forma, ao diploidizar os machos, todas as populações apresentam excesso de homozigotos. Resultados semelhantes podem ser vistos na DAPC, pois os grupos formados não divergiram o suficiente para se distribuir nos diferentes quadrantes; ademais, quando se observa a distribuição dos indivíduos, verifica-se que estes são atribuídos a um dos três grupos, e poucos deles apresentam atribuição parcial, conforme visto no gráfico de barras de machos do Structure, sugerindo que estes dados possam estar sendo influenciados pela diploidização dos machos para as análises. Os programas utilizados não foram especificamente delineados para as

peculiaridades dos organismos haplodiploides; dessa forma, ao fazer as análises com dados de machos e fêmeas, alguns resultados podem ser enviesados. Mas, esses resultados indicam a importância do estudo com machos e fêmeas em igual número de genótipos haploides analisados, e o mesmo número de indivíduos por localidade amostrada, haja vista que os resultados são mais robustos e há menos viés ao se analisar o mesmo número de genótipos haploides por local amostrado - vide nas populações com 4 fêmeas e oito machos em que os resultados apresentam mais indivíduos atribuídos a mais de um grupo, conforme as análises em separado.

As fêmeas de *C. analis* são consideradas filopátricas, espécies que nidificam próximos aos seus sítios natais, um comportamento comum em muitos himenópteros (Brockmann, 1979; Yanega, 1990; Polidori et al., 2006), tendo como consequência a redução do fluxo gênico materno. Contudo, os machos nessas espécies são considerados o sexo dispersor, caracterizando assim a aparente homogeneidade dos indivíduos a despeito dos grupos genéticos eventualmente caracterizados.

A ausência de estruturação genética geográfica para *C. analis* era esperada considerando algumas características que fundamentavam nossa hipótese do trabalho, entre as quais: i) a capacidade de dispersão, em que indivíduos maiores apresentam maior potencial para dispersão quando comparados a indivíduos de menor porte (Gathmann e Tschardt, 2002); ii) o tipo de dieta, em que espécies generalistas (poliléticas) podem se dispersar com mais facilidade quando comparadas às espécies especialistas (oligoléticas) (Rabelo et al., 2009); iii) a maior ou menor exigência de locais específicos de nidificação como elemento sucesso na colonização de uma área, entre outras (Luna-Lucena, 2015).

Segundo Michener (2007) esta espécie tem médio porte, o que lhe confere uma maior capacidade de dispersão. Essa capacidade pode ser demonstrada pela ampla área de distribuição da espécie. Conforme a modelagem realizada e a área amostral utilizada neste trabalho, observa-se que a distribuição ocorre desde a América do Norte até o sudeste do Brasil, demonstrando uma alta capacidade de dispersão associada com seu porte e tamanho corporal.

Outra questão que vale ser ressaltada é que o resultado de ausência de estruturação também era esperado quando se considera os resultados obtidos por Luna-Lucena e colaboradores (2017). Utilizando quatro genes mitocondriais para análise da estrutura genética de populações de *Centris (Hemisiella) tarsata*, estes autores estimaram um valor de F_{ST} global = -0.00483 ($P = 0,69$), demonstrando ausência de diferenciação entre e elevada diversidade dentro das populações. Ao considerar que *C. tarsata* é espécie proximamente

relacionada filogeneticamente à espécie em estudo, possui uma ampla distribuição e, particularmente, pela filopatria das fêmeas, esperaríamos encontrar aqui um valor de F_{ST} para o marcador utilizado de herança biparental ainda menor, principalmente considerando que o macho, por ser o sexo dispersor, torna mais homogênea a variação interpopulacional.

Este trabalho permite inferir também que o modelo biológico que utilizamos nos trabalhos de genética de populações (abelhas) não é usualmente considerado quando da elaboração de softwares. As abelhas são organismos haplodiploides e raramente os programas analisam machos e fêmeas em conjunto, pois todos os programas buscados analisam haploides ou diploides, mas não haplodiploides, exceto por Kingroup (Konovalov e Henshaw, 2004) e Colony (Jones e Wang, 2010) para análises de dados sociogenéticos. Dessa forma, os trabalhos de genética populacional acabam utilizando machos ou fêmeas, mas os dois em conjunto raramente são analisados. Vale ressaltar que os machos são de grande importância nos estudos, pois são considerados o sexo dispersor nas espécies solitárias com fêmeas filopátricas. Ao utilizar os machos nos estudos foi possível comparar dados de machos e fêmeas, verificar a diferença na estruturação, riqueza alélica, número de alelos, alelos privados; contudo, deve-se considerar que em algumas análises os machos não devem ser analisados em conjunto com as fêmeas, tais como nas estimativas de Equilíbrio de Hardy-Weinberg, Desequilíbrio de Ligação e Heterozigosidades observada e esperada.

5.3 Estrutura sociogenética intranidal

O sistema de acasalamento encontrado do tipo monândrico/monogínico, quando uma fêmea se acasala apenas com um macho durante o processo de nidificação, já foi caracterizada também para outras espécies de abelhas solitárias, encontrado para a referida espécie já tinha sido reportado por Silva (2010) em seus estudos genéticos com alozimas e um marcador mitocondrial e também por Moure-Oliveira (2018), mediante observação de comportamento da espécie. Este sistema de acasalamento também foi reportado para *Centris (Hemisiella) tarsata* por Silva (2010) e para *Euglossa melanotricha*, reportado por Andrade-Silva (2015) ao associar inspeção visual, análise genética intracolônia e determinação da estrutura familiar por genética mendeliana.

Os estudos de genética intranidal são relevantes para comprovar como estas abelhas realizam seu acasalamento, porém, é importante que a coleta de material (ninhas-armadilha) seja feita em conjunto com estudos de observação de comportamento das espécies durante a nidificação, ou mesmo, que estes estudos já tenham sido realizados. Conforme Moure-Oliveira e colaboradores (2017) afirmam, fêmeas usurpadoras de ninhas podem ovipositar em

ninhos em nidificação por outras fêmeas durante a viagem da fêmea fundadora do ninho na busca por material para o ninho. Dessa forma, existe a possibilidade de encontrar alelos provenientes de outro casal induzindo a interpretação errônea dos resultados. Quando existe uma alta frequência de nidificação de fêmeas usurpadoras, estas podem alterar os resultados aumentando o número para até 6 alelos dentro de um mesmo ninho, quando apenas três alelos são possíveis quando considerando o sistema monandria. Silva (2010) coloca em seu trabalho que um dos ninhos não se encaixou no padrão esperado, sendo que deste ninho emergiram 11 imagos e três deles apresentaram um padrão incompatível com a monogínia, indicando que mais de uma fêmea ovipositou no mesmo ninho. Verifica-se então que *Centris (Hemisiela) tarsata* também pode haver o comportamento de usurpação de ninho entre fêmeas.

Além disso, os estudos de biologia de nidificação associados a estudos de genética intranidal podem auxiliar na identificação da quantidade de ninhos que foram fundados pela mesma fêmea e definir assim o tamanho da prole e a produtividade média das fêmeas. Conforme registrado nas observações de Moure-Oliveira (2018), uma fêmea nidifica em cavidades uma ao lado da outra. Em nosso estudo não foi possível verificar diferentes ninhos fundados pela mesma fêmea, pois nas análises foram usados 8 ninhos escolhidos com maior número de prole. Para a detecção de mais de um ninho fundado pela mesma fêmea, todos os ninhos de um bloco de madeira devem ser analisados. Por esta razão, o número de prole de uma fêmea não foi definido neste estudo, pois não foram analisados todos os ninhos, devido utilizarmos aqueles com maior número de prole (3 a 5). Sumner e colaboradores (2007) em estudos com a vespa *Polistes canadensis* e Andrade-Silva (2015) em estudo com *Euglossa melanotricha* ressaltam a importância da combinação de estudos que associem observação com registro de imagens e análise de parentesco para determinação de um completo e robusto pedigree.

Como no nosso trabalho não foi monitorada a atividade de nidificação dos ninhos analisados, os microssatélites foram bastante eficientes em detectar o comportamento monândrico das fêmeas, podendo ser útil na detecção de fêmeas usurpadoras, o que não foi encontrado nos ninhos analisados. A importância de bons marcadores foi evidenciada em estudos com *Lasioglossum malachurum* (Soro et al., 2009) e em outros estudos com espécies de himenópteros primitivamente sociais, onde foram frequentes comutação de fêmeas entre ninhos ou usurpação de ninhos por fêmeas invasoras (Zanette e Field 2008).

6 CONSIDERAÇÕES FINAIS

A nossa hipótese de trabalho foi formulada com base em características biológicas da espécie, tais como, ampla distribuição, espécie multivoltina, polilética, porte corporal médio e biologia de nidificação. Devido a essas características, supusemos que as populações de *Centris (Heterocentris) analis* demonstrariam ausência de estruturação genética ao longo de sua distribuição. Nossos resultados apontam para uma estrutura genética baixa, dada principalmente por variação intrapopulacional.

Os achados contribuem para o conhecimento de como a genética populacional de uma espécie de abelha solitária que habita ambientes antropizados está particionada ao longo de sua distribuição, inferindo que os ambientes urbanos considerados impróprios ou hostis para muitas espécies mostram-se bastante adequados como refúgio para algumas espécies polinizadoras solitárias que apresentam adaptabilidade e tolerância às propriedades apresentadas por tais locais, indicadas pela alta abundância da espécie observada durante as coletas.

Dessa forma, verifica-se como a biologia da espécie pode influenciar a estrutura genética, pois os habitats urbanos podem fornecer uma infinidade de espécies exóticas em jardins e áreas verdes urbanas, como, por exemplo, *Tecoma stans*, que servirá de recurso para construção de ninhos. Além disso, a disponibilidade de recursos pode ocorrer durante todo o ano, garantindo assim que a espécie nidifique durante todo o ano e tenha à disposição muitas cavidades para nidificação. Dessa forma, é importante ressaltar que os projetos de ecologia urbana devem levar em consideração estas espécies que vêm utilizando as cidades como refúgio, e incluir cada vez mais áreas verdes com a disponibilidade de flores para o crescimento populacional e manutenção destas espécies.

O trabalho permitiu contribuir para os avanços no conhecimento da genética de abelhas solitárias, pois a maioria das pessoas desconhece a existências dessas espécies, acreditando que todas as abelhas são sociais. Dessa forma, é importante que não só a comunidade científica da área tenha acesso a estes trabalhos, mas que toda a comunidade possa ter conhecimento dessa fauna polinizadora e de sua importância para a manutenção de ecossistemas, haja vista que a maior parte da fauna apícola é formada por espécies solitárias.

A maioria dos estudos genéticos com abelhas utilizam fêmeas, indivíduos diploides, facilitando dessa formas as análises em todos os softwares. Ao abordar machos e fêmeas neste trabalho foi possível verificar que os machos são muito importantes no estudo de genética de populações, principalmente nas espécies solitárias filopátricas em que o macho é considerado

o sexo dispersor. Além disso, existe uma maior facilidade na coleta de machos, sendo assim, é possível coletar um número alto de machos durante poucas horas de coleta. Amostrar machos permite analisar dados em separado e verificar se existe uma diferenciação no padrão da distribuição genética. Dessa forma, esperava-se encontrar maior estruturação para fêmeas do que para machos já que estas são consideradas filopátricas e os machos dispersores. Mas, dada a baixa estruturação, percebe-se que as populações de machos e fêmeas não estão estruturadas geneticamente no espaço geográfico.

Por fim, o conhecimento aqui produzido é de grande valia para o conhecimento de estrutura genética populacional de abelhas solitárias de ampla distribuição que nidificam em cavidades pré-existentes, acessadas por marcadores microssatélites espécies-específicos. Verifica-se que o padrão da distribuição genética é diferente do encontrado para abelhas sociais e semelhante aos encontrados para abelhas solitárias com características semelhantes. Espera-se que este trabalho incentive e estimule a comunidade de pesquisadores na realização de mais trabalhos com espécies solitárias, desde espécies com distribuição reduzida a espécies de ampla distribuição, principalmente de tribo Centridini, para comparar dados e verificação se os resultados aqui encontrados se repetem ou são divergentes entre espécies próximas e as possíveis razões para tal.

REFERÊNCIAS

- AGUIAR, A. P. [et al.]. Order Hymenoptera. **Zootaxa**, v. 3703, p. 51-62, 2013.
- AHRNÉ, K.; BENGTSSON, J.; ELMQVIST, T. Bumble bees (*Bombus* spp) along a gradient of increasing urbanization. **PLoS ONE**, v. 5, p. 55-74, 2009.
- ALONSO, J. D. S. Efeito da variação no comprimento dos ninhos-armadilha na biologia de nidificação de *Centris (Heterocentris) analis* (Fabricius, 1804) (Hymenoptera, Apidae: Centridini). Dissertação (Mestrado em Ciências: Entomologia) Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo. 2008. 108 p.
- ALONSO, J. D. S.; SILVA, J. F.; GARÓFALO, C. A. The effects of cavity length on nest size, sex ratio and mortality of ratio and mortality of *Centris (Heterocentris) analis* (Hymenoptera, Apidae, Centridini). **Apidologie**, v. 43, p. 436–448, 2012.
- ALVES-DOS-SANTOS, I.; MACHADO, I. C.; GAGLIANONE, M. C. História natural das abelhas coletoras de óleo. **Oecologia Brasiliensis**, v. 11, p. 544-557, 2007.
- ANDERSON, E.; BARTHEL, S.; AHRNÉ, K. Measuring social–ecological dynamics behind the generation of ecosystem services. **Ecological Applications**, v. 17, p. 1267–1278, 2007.
- ANDRADE-SILVA, A. C. R. **Estrutura genética e sócio-etológica de *Euglossa melanotricha* Moure 1967 (Hymenoptera, Apidae, Euglossini) mediada por sinalização química e relações de parentesco intracolônial**. Tese (Doutorado em Ciências: Entomologia) Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo. 2015. 205p.
- ARAÚJO, M. B. The coincidence of people and biodiversity in Europe. **Global Ecology and Biogeography**, v. 12, p. 5–12, 2003.
- BARBOSA, A. M.; REAL, R.; OLIVERO, J.; VARGAS, J. M. Otter (*Lutra lutra*) distribution modelling at two resolution scales suited to conservation planning in the Iberian Peninsula. **Biological Conservation**, v. 114, p. 377–387, 2003.
- BOSSERT, S.; MURRAY, E. A.; ALMEIDA, E. A. B.; BRADY, S. G.; BLAIMER, B. B.; DANFORTH, B. N. Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. **Molecular Phylogenetics and Evolution**, v. 130, p. 121–131, 2019.
- BOWERS, M. A. Bumble bee (*Bombus*) colonization, extinction and reproduction in subalpine meadows in northeastern Utah, USA. **Ecology**, v. 66, p. 914-927, 1985.
- BROCKMANN, H. J. Nest-site selection in the great golden digger wasp, *Sphex ichneumoneus* L. (Sphecidae). **Ecological Entomology**, v. 4, p. 211-224, 1979.
- BROTHERS, D. J. Phylogeny and evolution of wasps, ants and bees (Hymenoptera, Chrysidoidea, Vespoidea and Apoidea). **Zoologica Scripta**, v. 28, p. 233-249, 1999.
- CARIVEAU, D.; WINFREE, R. Causes of variation in wild bee responses to anthropogenic drivers. **Current Opinion in Insect Science**, v. 10, p. 104–109, 2015.

CARMAGO, J. M. F.; MAZUCATO, M. Inventário de apifauna e flora apícola de Ribeirão Preto, SP, Brasil. **Dusenía**, v. 14, p. 55-87, 1984.

COUTO, R. M.; CAMILLO, E. Deposições de óleo por fêmeas de *Centris analis* (Fabricius) (Hymenoptera: Apidae: Centridini) parasitadas por *Physocephala* spp. (Diptera: Conopidae). **Entomo Brasilis**, v. 7, p. 81-85, 2014.

COUTO, R. M.; CAMILO, R. Influência da temperatura na mortalidade de imaturos de *Centris (Heterocentris) analis* (Hymenoptera, Apidae, Centridini). **Iheringia**, Série Zoologia, v. 97, p. 51-55, 2007.

CROZIER, R. H. On the potential for genetic variability in haplo-diploidy. **Genetica**, v. 41, p. 551-556, 1970.

DALAPICOLLA, J. **Tutorial de modelos de distribuição de espécies: guia prático usando o MaxEnt e o ArcGis 10**. Laboratório de Mastozoologia e Biogeografia, Universidade Federal do Espírito Santo, Vitória. 2016. 75p. Disponível em: <http://blog.ufes.br/lamab/tutoriais>. Acesso em: 20/05/2018.

DANKS, H. V. Populations and nesting-sites of some aculeate Hymenoptera nesting in *Rubus*. **Journal of Animal Ecology**. v. 40, p. 63–77, 1971.

DAVEY, J. W.; HOHENLOHE, P. A.; ETTER, P. D.; BOONE, J. Q.; CATCHEN, J. M.; BLAXTER, M. L. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. **Nature Reviews Genetics**, v. 12, p. 499-510, 2011.

DESSI, M. C. Genética populacional e filogeografia de *Partamona helleri* Friesi, 1900 (Hymenoptera: Apidae: Meliponini). Dissertação (Mestrado em Genética Evolutiva e Biologia Molecular) Universidade Federal de São Carlos, 2018. 97 p.

DORJI, J.; TAMANG, S.; TSHEWANG, T.; DORJI, T.; DORJI, T. Y. Genetic diversity and population structure of three traditional horse breeds of Bhutan based on 29 DNA microsatellite markers. **PLoS ONE**, v. 13, n. 6. P. 1-11, 2018.

ENGELBRECHT, J.; DUONG, T. A.; BERG, N. V. D. New microsatellite markers for population studies of *Phytophthora cinnamomi*, an important global pathogen. **Nature Scientific reports**, v. 7, p. 1-10, 2017.

ESTOUP, A.; SOLIGNAC, M.; HARRY, M.; CORNUET, J.M. Characterization of (GT) N and (CT) N microsatellites in 2 insect species – *Apis mellifera* and *Bombus terrestris*. **Nucleic Acids Research**, v. 21, p. 1427-1431, 1993.

EVANNO, G. REGNAUT, S. GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Molecular Ecology**, v. 14, p. 2611 - 2620, 2005.

EXCOFFIER, L.; LISCHER, H. E. L. Arlequin version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, p. 564-567, 2010.

FERREIRA, V. S.; AGUIAR, C. M. L.; OLIVEIRA, E. J. F.; COSTA, M. A.; SANTOS, G. M. M.; SILVA, J. G. Mitochondrial DNA variability in populations of *Centris aenea* (Hymenoptera, Apidae), a crop-pollinating bee in Brazil. **Genetics and Molecular Research**, v. 12, p. 830-837, 2013.

FIELDING, A. H., BELL, J. F. A review of methods for the assessment of prediction errors in conservation presence/ absence models. **Environmental Conservation**, v. 24, p. 38–49, 1997.

FRANCIS, R. M. POPHELPER: An R package and web app to analyse and visualise population structure. **Molecular Ecology Resources**, v. 17, p. 27 - 32, 2016.

FRANCISCO, F. O.; SANTIAGO, L. R.; MIZUSAWAL, Y. M.; OLDROYD, B. P.; ARIAS, M. C. Genetic structure of island and mainland populations of a Neotropical bumble bee species. **Journal Insect Conserv**, v. 20, p. 383–394, 2016.

FRANKHAM, R.; BALLOU, J. D.; BRISCOE, D. A. **Fundamentos de genética da conservação**. Tradução: FRANCISCO, M. R.; FARIAS, I. P.; GALETTI JUNIOR, P. M.; RODRIGUES, F. P. Ribeirão Preto, SP: SBG, 2008. 280 p.

GAGLIANONE, M. C.; AGUIAR, A. J. C.; VIVALLO, F.; ALVES-DOS-SANTOS, I. Checklist of oil bees from São Paulo State, Brazil. **Biota Neotropica**, v. 11, p. 657-666, 2011.

GAGLIANONE, M.C. Abelhas da tribo Centridini na Estação Ecológica de Jataí (Luiz Antônio, SP): composição de espécies e interações com flores de Malpighiaceae. In: MELO I, G. A. R.; ALVES-DOS-SANTOS, I. **Apoidea Neotropica: Homenagem aos 90 Anos de Jesus Santiago Moure**. Ed. UNESCO, Criciúma, 2003.

GALINDO-CARDONA, A.; ACEVEDO-GONZALEZ, J. P.; BERT RIVERA-MARCHAND, B.; GIRAY, T. Genetic structure of the gentle Africanized honey bee population (gAHB) in Puerto Rico. **BMC Genetics**, v. 14, n. 65, 2013.

GATHMANN, A.; TSCHARNTKE, T. Foraging ranges of solitary bees. **Journal of Animal Ecology**, v. 71, p. 757-764, 2002.

GAZOLA, A. L., **Parasitismo em populações de *Centris (Heterocentris) analis* (Fabricius, 1804) (Hymenoptera, Apidae, Centridini) com especial referência ao comportamento de *Leucospis cayennensis* Westwood, 1839 (Hymenoptera, Chalcidoidea, Leucospidae)**. Dissertação de Mestrado, FFCLRP-USP, Ribeirão Preto, Brazil, 1999.

GAZOLA, A. L.; GARÓFALO, C. A. Trap-nesting bees (Hymenoptera: Apoidea) in forest fragments of the state of São Paulo, Brazil. **Genetics and Molecular Research**, v. 8, p. 607-622, 2009.

GIANNINI, T. C.; SIQUEIRA, M. F.; ACOSTA, A. L.; BARRETO, F. C. C.; SARAIVA, A. M.; ALVES-DOS-SANTOS, I. Desafios atuais da modelagem preditiva de distribuição de espécies. **Rodriguésia**, v. 63, n. 5, p. 733-749, 2012.

GODDAD, M. [et al]. **Can cities save bees? How can urban habitats be made to serve pollinator conservation? How can that story be better told?**. 2016. Disponível em: <<https://www.thenatureofcities.com/2016/01/20/can-cities-save-bees-how-can-urban-habitats-be-made-to-serve-pollinator-conservation-how-can-that-story-be-better-told/>> Acesso em: 10/01/2019.

GOUDET, J. FSTAT (version 2.9.4), a program to estimate and test population genetics parameters. Available from <http://www.unil.ch/izea/software/fstat.html>. Updated from Goudet [1995] **Journal of Heredity**, v. 86, p. 485-486, 1995.

GRIMALDI, D. A.; ENGEL, M. S.; **Evolution of the insects**. Cambridge University Press. 2005. 772 p.

HALL, D. M. [et al]. The city as a refuge for insect pollinators. **Conservation Biology**, v. 31, p. 24–29, 2017.

HASSELMANN, M.; GEMPE, T.; SCHIOTT, M.; NUNES-SILVA, C. G.; OTTE, M.; BEYE, M. Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. **Nature**, v. 454, p. 519-523, 2008.

HEIMPEL, G. E.; BOER, J. G. Sex determination in Hymenoptera. **Annual Review of Entomology**, v. 53, p. 209-330, 2008.

HENNIG, E. I.; GHAZOUL, J. Pollinating animals in the urban environment. **Urban Ecosystems**, v. 15, p. 149–166, 2012.

HERNANDEZ, J. L.; FRANKIE, G. W.; THORP, R.W. Ecology of urban bees: a review of current knowledge and directions for future study. **Cities and the Environment**, v. 2, n. 1, p. 1-15, 2009.

HICKMAN, C. P.; ROBERTS, L. S.; LARSON, A. **Integrated Principles of Zoology**. Mc Graw-Hill Science, 2011. 882 p.

HIROTSU, C. M. **Seleção de cavidades para nidificação, tamanho da progênie por ninho, razão sexual e mortalidade de imaturos e adultos de *Centris (Heterocentris) analis* (Fabricius, 1804) (Hymenoptera, Apidae, Centridini)**. Dissertação de Mestrado (Ciências: Entomologia). Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 2013. 79 p.

JARVIS, A. M.; ROBERTSON, A. Predicting population sizes and priority conservation areas for 10 endemic Namibian bird species. **Biological Conservation**, v. 88, p. 121–131, 1999.

JAYASINGH, D. B.; FREEMAN, B.E. The comparative population dynamics of eight solitary bees and wasps (Aculeata, Apocrita, Hymenoptera) trap-nested in Jamaica. **Biotropica**, v. 12, p. 214-19, 1980.

JESUS, B. M. V.; GARÓFALO, C. A. Nesting behaviour of *Centris (Heterocentris) analis* (Fabricius) in southeastern Brazil (Hymenoptera, Apidae, Centridini). **Apidologie**, v. 31, p.503-515, 2000.

JONES, O.; WANG, J. COLONY: a program for parentage and sibship inference from multilocus genotype data. **Molecular Ecology Resources**, v. 10, p. 551-555, 2010.

KLEIJN, D. [et al]. Delivery of crop pollination services is an insufficient argument for wild pollinator conservation. **Nature Communications**, v. 6, p. 1-8, 2015.

KONOVALOV, D. A.; MANNING, C.; HENSHAW, M. T. KINGROUP: a program for pedigree relationship reconstruction and kip group assignments using genetic markers. **Molecular Ecology Notes**, v. 4, p. 779–782, 2004.

LIMA, R.; FERREIRA-CALIMAN, M. J.; DÓREA, M. C.; GARCIA, C. T.; SANTOS, F. A. R.; DE OLIVEIRA, F. F.; GARÓFALO, C. A. Floral resource partitioning between *Centris (Heterocentris) analis* (Fabricius, 1804) and *Centris (Heterocentris) terminata* Smith, 1874 (Hymenoptera, Apidae, Centridini), in an urban fragment of the Atlantic Forest. **Sociobiology**, v. 64, n. 3, p. 292-300, 2017.

LUNA-LUCENA, D. **A colonização de áreas urbanas por abelhas solitárias – *Centris (Hemisiella) tarsata* Smith, 1874 (Hymenoptera, Apidae, Centridini) como estudo de caso.** Dissertação de Mestrado (Genética Evolutiva e Biologia Molecular) Universidade Federal de São Carlos. São Carlos, 2015. 86 p.

LUNA-LUCENA, D.; MOURE-OLIVEIRA, D.; FERREIRA, K. M.; LUCENA, D. A. A.; GARÓFALO, C. A.; MORAES, E. M.; DEL LAMA, M. A. High mitochondrial DNA diversity and lack of population structure in a solitary cavity-nesting bee in an urban landscape. **Biological Journal of the Linnean Society**, v. 20, p. 1-12, 2017.

MAGALHÃES, C. B.; FREITAS, B. M. Introducing nests of the oil-collecting bee *Centris analis* (Hymenoptera: Apidae: Centridini) for pollination of acerola (*Malpighia emarginata*) increases yield. **Apidologie**, v. 44, p. 234–239, 2012.

MARTINS, A.C.; MELO, G. A. R. The New World oil-collecting bees *Centris* and *Epicharis* (Hymenoptera, Apidae): molecular phylogeny and biogeographic history. **Zoologia Scripta**, v. 45, p. 22–33, 2016.

MARTINS, C. F.; PEIXOTO, M. P.; AGUIAR, C. M. L. Plastic nesting behavior of *Centris (Centris) flavifrons* (Hymenoptera: Apidae: Centridini) in an urban area. **Apidologie**, v. 45, p. 156-171, 2014.

MATTESSON, K. C.; LANGELLOTTO, G.A. Determinates of inner city butterfly and bee species richness. **Urban Ecosystem**, v. 13, p. 333–347, 2010.

MATTESSON, K. C.; LANGELLOTTO, G. A. Bumble bee abundance in New York city community gardens: implications for urban agriculture. **Cities and the Environment**, v. 2, n. 5, p. 1-12, 2009.

MCINTYRE, M. E.; HOSTETLER, N. E. Effects of urban land use on pollinator (Hymenoptera: Apoidea) communities in a desert metropolis. **Basic and Applied Ecology**, v. 2, p. 209–218, 2001.

- MCINTYRE, N. E. Ecology of urban arthropods: A review and a call to action. **Annals of the Entomological Society of America**, v. 93, p. 825–835, 2000.
- MELO, G. A. R. Phylogenetic relationships and classification of the major lineages of Apoidea (Hymenoptera), with emphasis on the crabronid wasps. **Scientific Papers, Natural History Museum of the University of Kansas**, v. 14, p. 1-55, 1999.
- MICHENER, C. D. **The Bees of the World**. 2nd ed., Maryland: The Johns Hopkins University. 2007. 992p.
- MOURE, J. S. Notes on the types of the Neotropical bees described by Fabricius (Hymenoptera: Apoidea). **Studia Entomologica**, v. 3, p. 97-160, 1960.
- MOURE, J. S.; MELO, G. A. R.; VIVALLO, F. Centridini Cockerell & Cockerell, 1901. In MOURE, J. S.; URBAN, D.; MELO, G. A. R. (Orgs) **Catalogue of Bees (Hymenoptera, Apoidea) in the Neotropical Region - online version**. 2012. Available at <http://www.moure.cria.org.br/catalogue>. Acesso em: 14/03/2017.
- MOURE-OLIVEIRA, D. **Diversidade e estrutura genética de populações urbanas de abelhas Centridini (Hymenoptera: Apidae) visitantes florais de *Tecoma stans* (L) Kunth (Bignoniaceae)**. Dissertação de Mestrado (Ciências: Entomologia), Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 2013. 109 p.
- MOURE-OLIVEIRA, D. **História de vida de *Centris (Heterocentris) analis* (Fabricius, 1804) (Hymenoptera: Apidae, Centridini)**. Tese (Doutorado em Ciências: Entomologia), Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 2018. 146 p.
- MOURE-OLIVEIRA, D., ROCHA-FILHO, L.C.; FERREIRA-CALIMAN, M.J.; GARÓFALO, C. A. (2017): Nesting dynamic and sex allocation of the oil-collecting bee *Centris (Heterocentris) analis* (Fabricius, 1804) (Apidae: Centridini). **Journal of Natural History**, v. 51, p. 1151-1168, 2017.
- OBLESSUC, P. R.; CAMPOS, T. D.; KUPPER, J. M.; SFORÇA, D. A.; BARONI, R. M.; SOUZA, A. P. D.; BENCHIMOL, L. L. Adaptation of fluorescent technique for genotyping with new microsatellite markers in common bean. **Pesquisa Agropecuária brasileira**, Brasília, v. 44, p.638-644, 2009.
- OOSTERHOUT, C. V.; HUTCHINSON, W. F.; WILLS, D. P. M.; SHIPLEY, P. F. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. **Molecular Ecology Notes**, v. 4, p. 535-538, 2004.
- OWEN, R. E. The opportunity for polymorphism and genic variation in social Hymenoptera with worker-produced males. **Heredity**, v. 54, p. 25-36, 1985.
- PAMILO, P.; CROZIER, R. H. Genic variation in male haploids under deterministic selection. **Genetics**, v. 98, p. 199-214, 1981.
- PEAKALL, R.; SMOUSE, P. E. GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. **Bioinformatics**, v. 28, p. 2537-2539, 2012.

- PEARSON, R. G.; DAWSON, T. E.; LIU, C. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. **Ecography**, v. 27, p. 285–298, 2004.
- PEREIRA, M.; GARÓFALO, C. A.; CAMILLO, E.; SERRANO, J. C. Nesting biology of *Centris (Hemisiella) vittata* Lepeletier in southeastern Brazil (Hymenoptera, Apidae, Centridini). **Apidologie**, v. 30, p. 327–338, 1999.
- PHILLIPS, S. J.; ANDERSON, R. P.; SCHAPIRE, R. E. Maximum entropy modeling of species geographic distributions. **Ecological Modelling**, v. 190, p. 231–259, 2006.
- POLIDORI, C. [et al]. Philopatry, nest choice, and aggregation temporal–spatial change in the digger wasp *Cerceris arenaria* (Hymenoptera: Crabronidae). **Journal of Ethology**, v. 24, p. 155–163, 2006.
- POMPANON F. BONIN, A. BELLEMAIN, E. TABERLET, P. Genotyping errors: causes, consequences and solutions. **Nature Reviews Genetics**, v. 6, p. 847 - 59, 2005.
- PORRAS-HURTADO, L.; RUIZ, Y.; SANTOS, C.; PHILLIPS, C.; CARRACEDO, A.; LAREU, M. V. An overview of *STRUCTURE*: applications, parameter settings, and supporting software. **Frontiers in Genetic**, v. 4, p. 1–13, 2013.
- PRITCHARD, C. R.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotypic data. **Genetics**, v. 155, p. 945–959, 2000.
- RABELO, L. S.; VILHENA, A. M. G. F.; BASTOS, E. M. A. F.; AUGUSTO, S. C. Amplitude do nicho alimentar de *Centris (Heterocentris) analis* (Fabricius) (Apidae, Centridini). In: **Anais do IX Encontro Interno e XIII Seminário de Iniciação Científica** (2009). Universidade Federal de Uberlândia, Uberlândia, 2009.
- RENNER, S. S.; SCHAEFER, H. The evolution and loss of oil-offering flowers: new insights from dated phylogenies for plants and bees. **Philosophical Transactions of the Royal Society B**, v. 365, p. 423–435. 2010.
- RICKETTS, T. H. Tropical forest fragments enhance pollinator activity in nearby coffee crops. **Conservation Biology**, v. 18, p. 1262–1271, 2004.
- ROUGET, M. Measuring conservation value at fine and broad scales: implications for a diverse and fragmented region, the Agulhas Plain. **Biological Conservation**, v. 112, p. 217–232, 2003.
- SCHUELKE, M. An economic method for the fluorescent labeling of PCR fragments: A poor man's approach to genotyping for research and high-throughput diagnostics. **Nature Biotechnology**, v. 18, p. 233–234, 2000.
- SHEPPARD, W. S.; MCPHERON, B. A. Ribosomal DNA diversity in Apidae. In: SMITH, D. R. (ed). **Diversity in the genus *Apis***. Westview Press, Boulder (CO), p. 89–102, 1991
- SILVA, C. H. **Estrutura sociogenética intranidal e estrutura populacional de *Centris (Heterocentris) analis* (Fabricius, 1804) e *Centris (Hemisiella) tarsata* Smith 1874**

(**Hymenoptera: Apidae**). Dissertação (Mestrado em Ciências) Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo. 2010. 91 p.

SILVA, C. I.; HIROTSU, C. M.; PACHECO FILHO, A. J. D. S.; QUEIROZ, E. P.; GARÓFALO, C. A. Is the maximum reproductive rate of *Centris analis* (Hymenoptera, Apidae, Centridini) associated with floral resource availability?. **Arthropod-Plant Interactions**, v. 11, p. 389–402, 2017.

SILVA, C. I.; AUGUSTO, S. C.; SOFIA, S. H.; MOSCHETA, I. S. Diversidade de abelhas em *Tecoma stans* (L.) Kunth (Bignoniaceae): importância na polinização e produção de frutos. **Neotropical Entomology**, v. 36, p. 331-341, 2007.

SILVEIRA, F. A.; MELO, G. A. R.; ALMEIDA, E. A. B. **Abelhas Brasileiras, Sistemática e Identificação**. Belo Horizonte: Ministério do Meio Ambiente. Fundação Araucária, 2002. 253 p.

SORO, A.; AYASSE, M.; ZOBEL, M. U.; PAXTON, R. J. Complex sociogenetic organization and the origin of unrelated workers in a eusocial sweat bee, *Lasioglossum malachurum*. **Insectes Sociaux**, v. 56, p. 55-63, 2009.

SUMNER, S. [et al]. Radio-tagging technology reveals extreme nest drifting in a eusocial insect. **Current Biology**, v. 17, p. 140-145, 2007.

TEIXEIRA, F. M.; Técnicas de captura de Himenoptera (Insecta). **Vértices**, v. 14, p. 169-198, 2012.

THOMPSON, K. F.; PATEL, S.; BAKER, C. S.; CONSTANTINE, R.; MILLAR, C. D. Bucking the trend: genetic analysis reveals high diversity, large population size and low differentiation in a deep ocean cetacean. **Heredity**, v. 1, p. 1-9, 2015.

VELEZ, D.; VIVALLO, F.; SILVA, D. P. Nesting biology and potential distribution of na oil-colecting Centridini bee from South America. **Apidologie**, v. 48, p.181-193, 2017.

VICKRUCK, J. L.; RICHARDS, M. H. Nesting habits influence population genetic structure of a bee living in anthropogenic disturbance. **Molecular Ecology**, v. 26, p. 2674–2686, 2017.

VIEIRA, M. L. C.; SANTINI, L.; DINIZI, A. L.; MUNHOZ, C. F. Microsatellite markers: what they mean and why they are so useful. **Genetics and Molecular Biology**, v. 39, p. 312-328, 2016.

WILGENBURG, E.; DRIESSEN, G.; BEUKEBOOM, L. W. Single locus complementary sex determination in Hymenoptera: an "unintelligent" design? **Frontiers in Zoology**, v. 5, n. 3, p. 1-15, 2006.

WILLIAMS, S. E.; HERO, J. M. Multiple determinants of Australian tropical frog biodiversity. **Biological Conservation**, v. 98, p. 1–10, 2001.

WILLIAMS, J. W.; JACKSON, S. T.; KUTZBACK, J. E. Projected distributions of novel and disappearing climates by 2100 AD. **Proceedings of the National Academy of Sciences**, v. 104, p. 5738-5742, 2007.

YANEGA, D. Philopatry and nest founding in a primitively social bee, *Halictus rubicundus*. **Behavioral Ecology and Sociobiology**, v. 27, p. 37-42, 1990.

ZANELLA, F.C.V.; MARTINS, C.F. Abelhas da caatinga: Biogeografia, Ecologia e Conservação. In: LEAL, I. R.; TABARELLI, M.; SILVA, J. M. C. **Ecologia e conservação da Caatinga**. Recife: Ed. Universitária da UFPE, p. 75-134, 2003. 822 p.

ZANETTE, L. R. S.; FIELD, J. Genetic relatedness in early associations of *Polistes dominulus*: from related to unrelated helpers. **Molecular Ecology**, v. 17, p. 2590-2597, 2008.

ZALAPA, J. E. [et al.,]. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. **American Journal of Botany**, v. 99, p. 193–208, 2012.

ZAYED, A.; PACKER, L. Complementary sex determination substantially increases extinction proneness of haplodiploid population. **Proceedings of the National Academy of Sciences**, v. 102, p. 10742-10746, 2005.

APÊNDICES

APÊNDICE I

Protocolo de extração de DNA: Kit ZYMO RESEARCH modificado por Katia Maria Ferreira

Preparar solução de β -mercaptoetanol 0,5%. Adicionar 500 μ l desta solução no tampão GENOMIC LYSIS BUFFER. Utilizar material autoclavado.

- 1) Macerar em tubos de 1,5ml 3 (três) pernas em nitrogênio líquido. Adicionar 400 μ l de LYSIS SOLUTION. Misturar e deixar 10 minutos em temperatura ambiente;
- 2) Centrifugar, 10.000 rpm, por 1 minuto;
- 3) Transferir 400 μ l do sobrenadante para a coluna (tampa vermelha) + tubo coletor(1) NUMERADO e centrifugar, 7.000 rpm, 1 minuto (não esquecer de quebrar o lacre da coluna);
- 4) Adicionar 1.200 μ l de GENOMIC LYSIS BUFFER no filtrado do tubo coletor(1);
- 5) Misturar e transferir 800 μ l da solução para a outra coluna (tampa transparente) + tubo coletor(2). Centrifugar, 10.000 rpm, por 1 minuto;
- 6) Descartar o filtrado e repetir o passo 5 (reutilizar o tubo coletor);
- 7) Adicionar 200 μ l de DNA PRE-WASH BUFFER no tubo/coluna + tubo coletor(3) NUMERADO. Centrifugar, 10.000 rpm, por 1 minuto;
- 8) Adicionar 500 μ l g-DNA WASH BUFFER no tubo/coluna + tubo coletor(3). Centrifugar, 10.000 rpm, por 1 minuto;
- 9) Transferir a coluna para tubo de 1,5ml e adicionar 25 μ l de DNA ELUTION BUFFER. Centrifugar, 10.000 rpm, por 30 segundos;
- 10) Adicionar novamente 25 μ l de DNA ELUTION BUFFER. Centrifugar, 10.000 rpm, por 30 segundos.
- 11) Pronto para uso.

APÊNDICE II

Tabela Suplementar 1. Descrição das contribuições das 19 camadas ambientais referente a cada PC avaliado por Análise de Componentes Principais (PCA).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19
01	0.272446	-0.24144	0.111423	-0.08036	0.063511	-0.01301	0.001718	-0.01362	0.009539	-0.01008	0.272188	-0.01947	0.027154	-0.00238	0.328573	-0.51994	0.604739	0.147629	-0.05487
02	-0.17299	-0.24492	0.038904	0.533122	-0.1683	-0.4883	-0.22524	0.004736	0.01472	-0.37012	0.112785	-0.00582	0.332512	-0.18677	-0.04525	0.067227	0.066376	-0.01636	0.01392
03	0.238587	-0.05144	-0.33609	0.121243	0.043548	-0.50258	-0.18594	0.13116	-0.00508	0.660271	-0.23583	0.070831	-0.07576	0.010828	0.02001	-0.05868	-0.00046	0.005603	0.012835
04	-0.2446	-0.03801	0.398734	0.016841	-0.17555	0.188678	0.040485	-0.0493	0.000528	0.413359	-0.23563	0.009707	0.309498	-0.19413	-0.12327	0.000865	0.311544	-0.25671	0.413346
05	0.142533	-0.36455	0.346432	0.051962	-0.08493	-0.0414	-0.05248	-0.0586	-0.0245	0.131302	0.233616	0.048921	-0.33008	0.351298	-0.46624	0.314521	0.118645	0.263722	0.043034
06	0.302661	-0.11421	-0.05674	-0.15707	0.121945	-0.00999	0.045791	0.013364	0.011618	-0.0018	0.172652	0.022764	0.204498	-0.04528	-0.26715	0.183646	0.143422	-0.69395	-0.41066
07	-0.25436	-0.12298	0.310983	0.22077	-0.20487	-0.01771	-0.09203	-0.05576	-0.0289	0.090851	-0.04543	0.007622	-0.43382	0.228062	0.315715	-0.25532	-0.20565	-0.44267	-0.25756
08	0.212974	-0.29394	0.204396	0.054585	0.149237	0.011605	0.254858	0.589133	-0.07898	-0.25892	-0.5596	0.000411	-0.04829	0.004629	0.010057	-0.00431	-0.00726	0.008427	-0.00493
09	0.273066	-0.1645	0.027441	-0.17811	-0.09036	-0.07966	-0.14935	-0.6735	0.181337	-0.22072	-0.54125	-0.01324	-0.01939	0.000987	0.014826	0.002173	0.004184	0.008938	-0.0064
10	0.198964	-0.29864	0.330196	-0.1034	-0.03315	0.071445	0.019226	-0.04036	-0.01151	0.265162	0.167914	-0.03217	0.3022	-0.34653	0.220118	0.046258	-0.52389	0.232869	-0.23009
11	0.294862	-0.17412	-0.04785	-0.07608	0.09957	-0.07743	-0.01501	0.002677	0.000396	-0.13565	0.266775	-0.01169	-0.1004	0.039696	0.14844	-0.03527	-0.29749	-0.33092	0.73203
12	0.275117	0.210654	0.056718	0.185702	-0.14363	0.086885	-0.12891	0.025641	-0.19731	0.035694	-0.02136	-0.78937	-0.04186	0.019353	0.192683	0.279974	0.122605	-0.0184	-0.00068
13	0.28053	0.083399	-0.08835	0.244129	-0.25342	0.306718	-0.1305	-0.00292	-0.29999	-0.00442	-0.04312	0.488303	0.288796	0.366396	0.284927	0.202736	0.034275	0.012601	0.004896
14	0.148547	0.357089	0.286241	0.069136	0.076279	-0.31504	0.347634	-0.15845	-0.19644	-0.03309	0.051117	0.300967	-0.30457	-0.37597	0.198192	0.286154	0.156017	-0.01242	0.005157
15	-0.03465	-0.31509	-0.37643	0.326146	-0.16003	0.122909	0.722686	-0.2432	0.015314	0.113252	0.026625	-0.12212	-0.01634	0.009641	0.014498	0.006991	-0.00049	-0.00108	0.002039
16	0.281669	0.093661	-0.08492	0.24662	-0.24706	0.281489	-0.16228	0.005117	-0.24959	-0.0536	-0.00367	0.077545	-0.24214	-0.43686	-0.43396	-0.39825	-0.09804	-0.00321	-0.00029
17	0.159044	0.356848	0.281082	0.06844	0.050994	-0.28634	0.293847	-0.12139	-0.16016	-0.00134	-0.00428	-0.11163	0.326996	0.401377	-0.26233	-0.40034	-0.20585	0.019231	-0.00281
18	0.178413	0.165519	0.135105	0.527609	0.466427	0.256536	-0.06288	-0.05501	0.586875	0.079052	0.006629	0.050416	-0.00847	0.007594	0.003867	0.008264	-0.01802	0.001361	-0.00217
19	0.204709	0.204858	0.0347	-0.13392	-0.65723	-0.10236	0.135893	0.261543	0.598937	-0.04221	0.066528	0.045856	-0.01633	0.00895	0.014518	0.009147	-0.00099	0.002991	-0.00092

APÊNDICE III

Tabela Suplementar 2. Correlação entre as 19 variáveis ambientais

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
01	1																		
02	-0.27	1																	
03	0.6	-0.2	1																
04	-0.53	0.46	-0.91	1															
05	0.8	0.19	0.17	0.02	1														
06	0.93	-0.51	0.76	-0.78	0.53	1													
07	-0.52	0.74	-0.77	0.93	0.09	-0.8	1												
08	0.89	-0.07	0.42	-0.31	0.85	0.75	-0.27	1											
09	0.9	-0.36	0.64	-0.6	0.64	0.9	-0.6	0.67	1										
10	0.9	-0.1	0.25	-0.13	0.95	0.7	-0.15	0.89	0.76	1									
11	0.96	-0.37	0.77	-0.75	0.62	0.98	-0.72	0.8	0.91	0.75	1								
12	0.53	-0.56	0.56	-0.61	0.15	0.67	-0.68	0.35	0.58	0.32	0.62	1							
13	0.61	-0.46	0.66	-0.7	0.23	0.74	-0.7	0.43	0.65	0.38	0.71	0.91	1						
14	0.13	-0.48	0.14	-0.19	-0.08	0.25	-0.34	0.05	0.19	0.07	0.17	0.71	0.42	1					
15	0.08	0.46	0.23	-0.15	0.13	0.02	0.07	0.13	0.04	0	0.11	-0.34	-0.01	-0.66	1				
16	0.61	-0.46	0.66	-0.7	0.22	0.74	-0.71	0.42	0.65	0.37	0.71	0.93	0.99	0.44	-0.04	1			
17	0.16	-0.5	0.16	-0.22	-0.07	0.28	-0.37	0.06	0.21	0.08	0.2	0.75	0.46	0.99	-0.67	0.48	1		
18	0.33	-0.27	0.31	-0.36	0.11	0.4	-0.39	0.32	0.26	0.19	0.37	0.69	0.59	0.59	-0.24	0.6	0.6	1	
19	0.34	-0.51	0.4	-0.42	0.05	0.48	-0.52	0.16	0.47	0.21	0.41	0.75	0.68	0.57	-0.31	0.69	0.6	0.2	1

APÊNDICE IV

Tabela Suplementar 3. Valores de FST par-a-par entre as localidades amostradas de *Centris (Heterocentris) analis* analisadas. Os valores de FST com asterisco indica os valores significativos.

Localidades	TES	SCM	VLV	ITU	CNP	GRU	RIF	BOT	JEQ	PTS	SOR	MTR	PRP	CRZ	FOT
Teodoro Sampaio	0,000														
Sacramento	0,030	0,000													
Vila Velha	0,039	0,031	0,000												
Ituaçu	0,033	0,035	0,035	0,000											
Campo Novo do Parecis	0,028	0,030	0,025	0,037	0,000										
Grajaú	0,029	0,017	0,018	0,033	0,018	0,000									
Rifaina	0,035	0,015	0,033	0,043	0,037	0,027	0,000								
Bonito	0,075	0,046	0,040	0,058	0,067	0,045	0,051	0,000							
Jequié	0,075*	0,060	0,061	0,063*	0,082*	0,059*	0,062*	0,065	0,000						
Patos	0,058	0,047	0,059	0,068	0,053	0,051	0,050	0,070	0,070	0,000					
Sorriso	0,046	0,055	0,053	0,059	0,054	0,045	0,044	0,080	0,053	0,050	0,000				
Matureia	0,046	0,026	0,029	0,033	0,027	0,022	0,030	0,040	0,059	0,061	0,053	0,000			
Pirapora	0,062	0,061	0,064	0,070	0,082*	0,051	0,055	0,080	0,060	0,053	0,047	0,067	0,000		
Cruzeiro do Sul	0,061*	0,061*	0,073*	0,054*	0,061*	0,047	0,060*	0,088*	0,103*	0,064	0,083*	0,079*	0,075*	0,000	
Fortaleza	0,087*	0,062	0,079*	0,093*	0,082	0,079	0,058	0,095	0,091	0,085	0,083	0,058	0,072	0,126*	0,000
Arame	0,055	0,047	0,048	0,047	0,053	0,041	0,039	0,057	0,056	0,046	0,042	0,032	0,050	0,086*	0,063
Belém	0,071*	0,077*	0,054	0,0856*	0,063	0,075*	0,064	0,114*	0,089*	0,099	0,080	0,063	0,064	0,135*	0,076
Macapá	0,055*	0,052*	0,022	0,045	0,043	0,043	0,040	0,049	0,056*	0,074*	0,054	0,033	0,052	0,089*	0,069
Goiânia	0,072*	0,059*	0,049	0,066	0,060	0,056*	0,054	0,087	0,074*	0,073	0,071*	0,061	0,048	0,103*	0,080
Jataí	0,055*	0,030	0,032	0,038	0,040	0,033	0,033	0,047	0,040	0,054	0,033	0,025	0,023	0,066*	0,065
Primavera do Leste	0,059	0,049	0,046	0,067*	0,046	0,054	0,053	0,076	0,055	0,069	0,065	0,050	0,064	0,094*	0,088

Tabela Suplementar 3 (Continuação). Valores de FST par-a-par entre as localidades amostradas de *Centris (Heterocentris) analis* analisadas. Os valores de FST com asterisco indica os valores significativos.

Localidades	ARA	BLM	MCP	GYN	JTI	PVA
Teodoro Sampaio						
Sacramento						
Vila Velha						
Ituaçu						
Campo Novo do						
Parecis						
Grajaú						
Rifaina						
Bonito						
Jequié						
Patos						
Sorriso						
Matureia						
Pirapora						
Cruzeiro do Sul						
Fortaleza						
Arame	0,000					
Belém	0,062	0,000				
Macapá	0,038	0,061	0,000			
Goiânia	0,044	0,065	0,054*	0,000		
Jataí	0,039	0,048	0,034	0,033	0,000	
Primavera do Leste	0,040	0,061	0,050	0,036	0,039	0,000

APÊNDICE V

Tabela Suplementar 4. Descrição da diversidade genética para locos e populações. São apresentados, tamanho da amostra (N), número de alelos (Na), Riqueza alélica (Ar), Heterosigosiade observada (Ho) e esperada (He), sendo Ho e He calculados apenas com as amostras de fêmeas.

		Locos															
Pop		Canl1	Canl2	Canl 4	Canl5	Canl6	Canl7	Canl9	Canl10	Canl11	Canl12	Canl13	Canl14	Canl16	Canl18	Canl19	Canl20
TES	Na	6	4	5	10	6	8	8	5	7	5	8	3	4	6	5	9
	Ar	4,67	3,65	4,43	7,63	4,86	5,89	6,15	4,14	5,85	4,60	6,34	2,76	3,46	4,54	4,66	6,63
	Ho	0,500	0,750	0,750	0,750	0,500	1,000	1,000	0,500	0,500	0,250	0,667	0,500	0,500	1,000	0,750	0,750
	He	0,406	0,719	0,625	0,750	0,750	0,781	0,750	0,656	0,688	0,656	0,667	0,500	0,563	0,750	0,469	0,813
SCM	Na	5	5	5	10	7	7	8	7	5	4	9	4	9	4	6	8
	Ar	3,91	4,14	4,29	7,43	6,22	5,04	6,59	5,98	4,68	3,48	6,23	3,39	7,14	3,86	4,90	6,41
	Ho	0,750	0,750	0,750	0,750	0,750	0,750	0,500	1,000	0,500	0,500	1,000	0,750	0,750	0,750	0,500	0,750
	He	0,656	0,656	0,656	0,719	0,688	0,656	0,750	0,719	0,656	0,719	0,844	0,563	0,781	0,656	0,813	0,750
VLV	Na	4	4	5	10	5	8	11	7	4	4	8	5	6	7	4	8
	Ar	3,26	3,84	4,23	7,71	3,89	6,44	7,93	5,97	3,61	3,65	5,99	3,76	4,91	5,74	3,94	6,82
	Ho	0,750	0,500	0,750	0,250	0,750	0,750	1,000	0,750	0,750	0,250	1,000	0,500	0,500	0,750	0,000	0,750
	He	0,531	0,594	0,656	0,781	0,750	0,688	0,875	0,750	0,531	0,531	0,781	0,563	0,656	0,750	0,625	0,781
ITU	Na	5	5	7	9	7	8	6	9	6	5	10	3	5	5	5	7
	Ar	4,19	4,41	4,52	6,76	5,67	6,26	5,02	6,81	5,09	4,41	7,83	2,50	4,23	4,59	4,36	5,00
	Ho	1,000	0,750	1,000	0,750	0,750	0,750	0,500	1,000	1,000	0,250	0,500	0,500	1,000	0,250	0,500	0,750
	He	0,656	0,656	0,813	0,656	0,750	0,688	0,688	0,750	0,750	0,219	0,813	0,594	0,656	0,594	0,594	0,813

Pop		Canl1	Canl2	Canl 4	Canl5	Canl6	Canl7	Canl9	Canl10	Canl11	Canl12	Canl13	Canl14	Canl16	Canl18	Canl19	Canl20
CNP	Na	5	4	6	10	5	8	11	10	9	5	8	3	7	5	6	10
	Ar	4,34	3,25	4,41	7,53	4,14	6,15	8,30	7,30	6,75	4,62	6,19	2,76	5,36	4,72	4,89	7,25
	Ho	0,667	0,333	1,000	1,000	0,667	1,000	0,667	1,000	1,000	0,667	0,333	0,333	0,667	1,000	0,667	1,000
	He	0,778	0,278	0,667	0,778	0,611	0,778	0,778	0,778	0,833	0,611	0,500	0,278	0,667	0,778	0,667	0,833
GRU	Na	5	3	6	8	5	7	10	8	6	5	7	3	7	6	5	7
	Ar	4,24	2,97	4,74	6,21	4,50	5,54	7,45	6,21	4,74	4,01	5,56	2,50	5,67	5,12	4,28	5,74
	Ho	0,250	0,750	0,500	1,000	0,250	0,750	0,750	1,000	0,750	0,250	0,750	0,500	0,500	0,500	0,500	1,000
	He	0,719	0,531	0,719	0,813	0,531	0,531	0,844	0,688	0,750	0,219	0,813	0,406	0,750	0,688	0,688	0,750
RIF	Na	6	5	9	7	6	10	6	8	9	4	7	5	6	6	8	6
	Ar	5,32	4,21	6,69	5,47	5,55	7,25	4,73	6,46	6,88	3,64	6,50	4,02	5,19	5,53	6,41	5,04
	Ho	0,750	0,500	0,500	0,250	0,500	1,000	0,500	1,000	0,750	0,250	0,333	0,250	0,000	1,000	1,000	0,000
	He	0,719	0,594	0,813	0,656	0,750	0,813	0,750	0,781	0,844	0,719	0,611	0,656	0,375	0,656	0,688	0,750
BOT	Na	3	3	4	8	5	5	6	5	3	2	4	3	5	6	5	3
	Ar	2,89	3,00	3,86	7,65	4,70	4,97	5,94	4,85	2,85	1,99	3,85	2,86	4,97	5,95	4,83	2,86
	Ho	1,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	0,000	1,000	1,000	0,000	1,000	1,000	1,000
	He	0,500	0,500	0,500	0,500	0,500	0,000	0,000	0,500	0,500	0,000	0,500	0,500	0,000	0,500	0,500	0,500
JEQ	Na	6	4	7	8	7	6	8	8	6	5	6	2	4	5	7	8
	Ar	5,23	3,71	5,89	6,52	6,29	4,93	6,28	6,52	4,54	4,28	5,09	2,00	3,51	3,71	6,13	6,73
	Ho	0,000	0,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000	0,000	1,000	0,000	0,000	1,000	0,000	1,000
	He	0,000	0,000	0,500	0,000	0,000	0,500	0,500	0,500	0,500	0,000	0,500	0,000	0,000	0,500	0,000	0,500
PTS	Na	3	5	4	6	4	5	4	5	4	3	7	3	3	3	3	5
	Ar	2,98	4,97	3,98	5,94	3,98	4,97	3,99	4,95	3,98	3,00	6,92	2,92	2,98	3,00	2,99	4,97

Pop		Canl1	Canl2	Canl 4	Canl5	Canl6	Canl7	Canl9	Canl10	Canl11	Canl12	Canl13	Canl14	Canl16	Canl18	Canl19	Canl20
SOR	Na	6	5	5	9	5	8	6	6	5	3	6	4	6	6	7	5
	Ar	5,19	4,58	4,39	7,58	4,58	6,94	5,36	5,19	4,58	2,99	5,22	3,91	5,36	5,36	6,30	4,72
	Ho																
	He																
MTR	Na	4	4	6	8	5	5	6	9	5	3	8	3	5	6	6	5
	Ar	3,62	3,57	5,27	6,84	4,78	4,78	5,20	7,79	4,70	2,99	6,60	2,90	4,44	5,67	5,46	4,70
	Ho	0,667	1,000	0,333	0,333	0,333	0,333	0,667	0,667	0,000	0,333	1,000	0,333	0,333	1,000	0,333	0,333
	He	0,500	0,611	0,278	0,500	0,500	0,500	0,667	0,667	0,444	0,500	0,778	0,500	0,500	0,611	0,500	0,278
PRP	Na	6	5	8	6	4	7	6	8	6	5	6	4	4	5	4	4
	Ar	5,82	4,68	7,39	5,82	3,99	6,68	5,82	7,53	5,68	4,70	5,69	3,71	3,83	4,71	3,97	3,85
	Ho	1,000	0,000	1,000	0,500	0,500	1,000	0,500	0,500	0,500	1,000	1,000	1,000	0,500	1,000	0,500	0,500
	He	0,625	0,500	0,625	0,625	0,375	0,750	0,625	0,375	0,625	0,750	0,625	0,750	0,375	0,750	0,625	0,375
CRZ	Na	3	5	5	7	5	6	10	6	6	4	6	3	5	7	3	7
	Ar	2,76	4,40	4,28	5,98	4,25	4,91	7,82	4,76	5,23	3,60	5,02	2,50	4,47	5,75	2,52	5,43
	Ho	0,250	0,250	0,750	0,000	0,750	0,750	0,500	0,250	0,750	0,250	0,250	0,250	0,500	0,500	0,000	0,500
	He	0,469	0,594	0,563	0,625	0,719	0,750	0,813	0,531	0,531	0,469	0,656	0,594	0,625	0,656	0,625	0,781
FOT	Na	3	4	5	5	5	4	5	5	3	3	4	2	4	4	4	5
	Ar	3,00	4,00	5,00	5,00	5,00	4,00	5,00	5,00	3,00	3,00	4,00	2,00	34,00	4,00	4,00	5,00
	Ho	0,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,000	0,000	0,000	1,000	1,000	1,000	1,000
	He	0,000	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,000	0,000	0,000	0,500	0,500	0,500	0,500
ARA	Na	5	5	5	10	5	6	9	10	3	4	7	4	7	6	4	7
	Ar	4,97	4,47	4,24	8,06	4,21	5,51	7,39	7,18	2,89	3,71	5,82	3,26	5,53	4,86	3,25	5,63
	Ho	1,000	1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	0,000	1,000	1,000	1,000	1,000	0,000	1,000
	He	0,500	0,500	0,000	0,500	0,500	0,500	0,500	0,500	0,500	0,000	0,500	0,500	0,500	0,500	0,000	0,500

Pop		Canl1	Canl2	Canl 4	Canl5	Canl6	Canl7	Canl9	Canl10	Canl11	Canl12	Canl13	Canl14	Canl16	Canl18	Canl19	Canl20
BLM	Na	4	6	8	6	6	4	5	8	7	4	6	2	6	4	4	7
	Ar	3,90	5,64	7,25	5,64	5,55	3,94	4,69	7,44	6,69	3,94	5,39	2,00	5,20	3,94	3,73	6,34
	Ho	0,000	0,500	0,500	1,000	0,500	1,000	0,000	0,500	0,000	0,500	1,000	0,500	1,000	0,000	1,000	1,000
	He	0,500	0,625	0,625	0,750	0,375	0,625	0,500	0,625	0,500	0,375	0,625	0,375	0,750	0,500	0,500	0,750
	Na	6	6	8	9	5	9	8	8	4	5	8	4	5	6	5	8
MCP	Ar	5,23	5,35	5,93	6,78	4,39	6,91	6,15	6,77	43,26	4,01	5,87	3,45	4,00	4,74	4,02	6,09
	Ho	0,500	0,250	0,750	1,000	0,750	1,000	1,000	1,000	0,500	0,500	1,000	1,000	0,500	0,750	0,250	0,750
	He	0,625	0,656	0,750	0,844	0,656	0,750	0,750	0,719	0,594	0,594	0,813	0,719	0,656	0,656	0,531	0,750
	Na	4	4	4	8	7	8	8	9	4	3	6	6	5	4	6	7
GYN	Ar	3,64	3,00	3,73	6,39	6,10	6,22	6,77	7,29	3,84	2,91	5,41	4,77	4,46	3,50	5,23	6,08
	Ho	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,000	1,000	1,000	1,000	1,000	0,000	1,000
	He	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,000	0,500	0,500	0,500	0,500	0,000	0,500
	Na	5	5	8	8	4	8	7	12	5	3	7	4	6	5	5	6
JTI	Ar	4,49	4,40	6,71	6,47	3,83	6,69	6,09	9,08	4,88	2,99	6,34	3,55	5,25	4,79	4,57	5,62
	Ho	0,600	0,800	1,000	0,800	0,600	1,000	0,800	1,000	1,000	0,400	0,250	1,000	1,000	0,600	0,200	0,600
	He	0,680	0,720	0,740	0,800	0,740	0,760	0,740	0,860	0,760	0,620	0,781	0,640	0,660	0,640	0,580	0,780
	Na	7	4	5	7	5	7	5	9	7	4	7	2	6	5	4	8
PVA	Ar	5,85	3,52	4,50	5,97	4,11	5,98	4,25	7,29	5,68	3,26	6,30	2,00	5,22	4,48	3,84	6,14
	Ho	1,000	0,000	0,500	1,000	1,000	1,000	0,500	1,000	1,000	1,000	1,000	0,000	0,500	0,500	0,500	1,000
	He	0,750	0,000	0,375	0,750	0,625	0,625	0,625	0,625	0,750	0,625	0,750	0,000	0,375	0,375	0,375	0,750

APÊNDICE VI

Tabela Suplementar 5. Descrição dos genótipos para os ninhos analisados neste trabalho para 10 locos microssatélites com delimitação do genótipo dos parentais. SP: sexo da prole; GPe: Genótipo da prole; GPa: Genótipo do Pai; GM: Genótipo da mãe. As barras entre genótipos (/) significam que poder ser um ou outro genótipo. Os pontos de interrogação (?) representa que o genótipo não foi identificado. Os traços (-) representam *missing data*.

Ninho 1					Ninho 2				Ninho 3				Ninho 4				
Locos	SP	GPe	GPa	GM	SP	GPe	GPa	GM	SP	GPe	GPa	GM	SP	GPe	GPa	GM	
Loco 5	M	209			F	207	219		F	195	199		M	211			
	F	213	215		M	195			M	195			F	217	225		
	M	209		213	209	215		219	195	207		199	195	219	217	211	225
	M	215				M	195		M	219			F	217	225		
Loco 9	M	242			F	256	256		F	242	244		M	244			
	F	242	246		M	256			M	244			F	242	244		
	M	242		246	240	242		256	246	256		242	244	266	242	240	244
	M	240				M	246		M	244			F	242	244		
Loco 11	M	104			F	102	104		F	104	106		M	102			
	F	102	106		M	102			M	106			F	98	104		
	M	104		102	104	106		102	102	104		104	106	106	98/104	102	98/104
	M	104				M	104		M	106			F	98	104		
Loco 1	M	232			F	228	232		F	228	232		M	236			
	F	232	232		M	232			M	228			F	230	232		
	M	232		232	232	232		228	232	232		232	228	228	230/232	236	230/232
	M	232				M	232		M	228			F	230	232		
Loco 6	M	116			F	116	122		F	116	128		M	122			
	F	119	128		M	122			M	116			F	122	122		
	M	119		128	116	119		116	122	122		128	116	1	122	122	122
	M	119				M	122		M	116			F	122	122		

	Ninho 1				Ninho 2				Ninho 3				Ninho 4			
Locos	SP	GPe	GPa	GM												
Loco 14	M	195			F	193	193		F	195	195		M	193		
	F	193	193		M	189			M	195			F	193	193	
	M	193		193	F	193	193	189	M	193		195	F	193	195	193
	M	193			M	189			M	195			F	193	193	
Loco 2	M	301			F	301	303		F	297	301		M	301		
	F	301	303		M	303			M	297			F	297	301	
	M	301		303	F	303	303	301	M	297		301	F	297	301	297/301
	M	301			M	301			M	297			F	297	301	301
Loco 7	M	212			F	221	221		F	206	215		M	212		
	F	197	215		M	212			M	215			F	212	218	
	M	215		197	F	212	221	221	M	215		206	F	212	218	212/218
	M	215			M	221			M	209			F	212	218	212
Loco 4	M	221			F	221	223		F	221	221		M	219		
	F	221	227		M	223			M	221			F	219	219	
	M	221		227	F	221	225	221	M	221		221	F	219	221	219
	M	221			M	225			M	201			F	219	219	219
Canl 10	M	254			F	252	254		F	254	254		M	-	-	
	F	250	250	250	M	254			M	254			F	248	248	
	M	250			F	252	254	252	M	252		254	F	248	256	248
	M	250			M	254		252	M	254		252	F	248	256	248

