



REDE NORDESTE DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

Aline Barbosa Negreiros

ANÁLISE GENÉTICA E GENÔMICA DA ABELHA-SEM-FERRÃO *Melipona rufiventris*: SEQUENCIAMENTO, DESENVOLVIMENTO DE FERRAMENTAS MOLECULARES E INFERÊNCIAS POPULACIONAIS

Teresina-PI

2019

Aline Barbosa Negreiros

ANÁLISE GENÉTICA E GENÔMICA DA ABELHA-SEM-FERRÃO *Melipona rufiventris*: SEQUENCIAMENTO, DESENVOLVIMENTO DE FERRAMENTAS MOLECULARES E INFERÊNCIAS POPULACIONAIS

Tese apresentada ao Programa de Pós-Graduação da Rede Nordeste de Biotecnologia (RENORBIO), Universidade Federal do Piauí (UFPI), como requisito parcial para a obtenção do grau de Doutor em Biotecnologia.

Área de Concentração: Biotecnologia em Agropecuária

Orientador: Dr. Fábio Mendonça Diniz

Teresina-PI

2019

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Comunitária Jornalista Carlos Castello Branco
Serviço de Processamento Técnico

N385a Negreiros, Aline Barbosa.

Análise genética e genômica da abelha-sem-ferrão *Melipona rufiventris*: sequenciamento, desenvolvimento de ferramentas moleculares e inferências populacionais / Aline Barbosa Negreiros. – 2019.

126 f. : il.

Tese (Doutorado em Desenvolvimento e Meio Ambiente) - Universidade Federal do Piauí, Teresina-PI, 2019.
“Orientação: Prof. Dr. Fábio Mendonça Diniz”.

1. Abelhas Nativas. 2. Caracterização Populacional. 3. Marcadores Heteroespecíficos. 4. Microssatélites. I. Título.

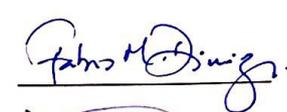
CDD 638.1

FOLHA DE APROVAÇÃO – DEFESA DE TESE

ALUNA: Aline Barbosa Negreiros

TÍTULO DO PROJETO: "ANÁLISE GENÉTICA E GENÔMICA DA ABELHA-SEM-FERRÃO
Melipona rufiventris: SEQUENCIAMENTO, DESENVOLVIMENTO DE FERRAMENTAS
MOLECULARES E INFERÊNCIAS POPULACIONAIS"

PROFESSOR ORIENTADOR: Prof. Dr. Fábio Mendonça Diniz

BANCA EXAMINADORA:	CONCEITO	ASSINATURA
Prof. Dr. Fábio Mendonça Diniz - EMBRAPA (Presidente)	<u>APROVADO</u>	
Prof. Dr. Francisco das Chagas Alves Lima - UESPI (Examinador)	<u>Aprovada</u>	
Prof. Dr. Bruno de Almeida Souza - EMBRAPA (Examinador)	<u>APROVADA</u>	
Prof. Dr. Paulo Sarmanho da Costa Lima – EMBRAPA (Examinador)	<u>Aprovada</u>	
Prof. Dr. Vladimir Costa Silva - UFPI (Examinador)	<u>Aprovado</u>	

DATA DA AVALIAÇÃO: 20 de março de 2019.

HORÁRIO: 14h00

LOCAL: Núcleo Integrado de Morfologia e Pesquisas com Células-tronco -
NUPCelt/UFPI

Dedico a toda minha família em especial aos meus Pais (Arnaldo Negreiros e Maria do Santo), meu irmão Francisco Gomes (in Memoriam) e ao meu esposo Rogério Azevedo.

AGRADECIMENTOS

- A **Deus** pela minha vida e por sempre guiar meus caminhos e nunca me abandonar.
- À Universidade Federal do Piauí (**UFPI**), o Programa de Pós-Graduação da Rede Nordeste de Biotecnologia (**RENORBIO**), em nome das suas secretárias **Deusilene e Eliane**, e a todo seu corpo docente pela oportunidade de realização do curso;
- À Empresa Brasileira de Pesquisa Agropecuária (**EMBRAPA Meio-Norte**) e ao Laboratório de Biotecnologia e Biologia Molecular, onde pelo financiamento da pesquisa macroprograma 2 (02.11.01.029.00.00), tive oportunidade de desenvolver esse trabalho;
- À **FUNCAP** (Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico) pela concessão da bolsa de estudos nos primeiros seis meses de doutorado;
- À **CAPES** (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela concessão da bolsa de estudos no restante do curso de doutorado;
- À professora do RENORBIO (ponto focal UFC) **Vânia Maria Maciel Melo**, minha primeira orientadora (primeiros seis meses), por ter me recebido de uma forma muito acolhedora em FORTALEZA-CE para o início do DOUTORADO e por ter me compreendido e me auxiliado no processo de transferência;
- Ao pesquisador **Fábio Mendonça Diniz** (Embrapa Meio-Norte) por ter aceitado minha orientação e por todos ensinamentos e atenção direcionada a minha pessoa nesses três anos e meio de orientação;
- À pesquisadora **Fábia de Mello Pereira** (Embrapa Meio-Norte) por ter me “adotado” como estagiária na EMBRAPA, durante o período de análises laboratoriais e em todo o desenvolvimento do projeto com a abelha-sem-ferrão;
- Aos Pesquisadores da Embrapa Meio-Norte, **Bruno de Almeida Souza** e **Maria Teresa do R. Lopes**, pelo apoio e direcionamentos que me deram durante o decorrer da pesquisa;
- Ao professor **Rodrigo Maggioni** do Instituto de Ciências do Mar (LABOMAR-UFC) por ter aceitado a parceria do sequenciamento e ter nos disponibilizado todo apoio “humano” e laboratorial do seu departamento para nos auxiliar nas análises laboratoriais.
- Ao laboratório NPDM – Núcleo de Pesquisa e Desenvolvimento de Medicamentos da Universidade Federal do Ceará-UFC, em nome do professor **Diego Veras Wilke**, por disponibilizar toda estrutura laboratorial para a realização do sequenciamento do genoma da abelha *Melipona rufiventris*, durante duas semanas no município de Fortaleza-CE.

- As técnicas de laboratório da UFC (**Graça, Renata, Andreia e Giovanna**) pelo auxílio nas análises do sequenciamento;
- Aos professores, **Tânia Maria Fernandes Salomão** e **Helder Canto Resende**, da UFV- Universidade Federal de Viçosa por ter nos enviado amostras de abelhas *Melipona rufiventris*;
- **Aos meliponicultores** (Claudionor (BA), Claudio (GO), Gesimar (SP) e outros inúmeros produtores do Estado do PIAUÍ pelo fornecimento das amostras de abelhas *Melipona rufiventris* e ao bolsista de pós-doc da Embrapa Meio-Norte, **Rafael**, por contribuir na coleta de campo das amostras do Piauí;
- Ao coordenador do laboratório de Biologia Molecular da Embrapa Meio-Norte **Dr. Paulo Sarmanho** e ao laboratorista **Leonardo Furtado**, pela paciência e apoio nas atividades realizadas no laboratório;
- Ao meu colega **Geice Ribeiro** (hoje doutor em Ciência Animal-UFPI) pelo apoio desde o primeiro dia que cheguei na EMBRAPA, e por não medir esforços para me ajudar na execução (Bancada e análises) de todas as práticas realizadas no laboratório durante esse período, um verdadeiro co-orientador (sem as tuas orientações nada disso seria possível, muito GRATA);
- Aos colegas do Laboratório de Biotecnologia e Biologia Molecular da Embrapa Meio-Norte que juntos compartilharam experiências fundamentais no dia-a-dia de pesquisa em especial a **Kaline, Jessica, Vanessa, Willame, Isis, Sarah, Gisele, Remédios, Rayane, Jarbson, Mauricio, Alberto e Sárvia** pela ajuda de bancada;
- Aos meus inúmeros **colegas de almoço** (Sr. Francisco, Sr.Ozires, Sr.Moreira, Sr.Afonso, Dona Neidinha, Ananda, Nadja, Leovânio, Gabriela, dentre outros que apareciam algumas vezes nos almoços de “Quentinha da EMBRAPA”) que me atualizavam das notícias e me faziam dar as melhores gargalhadas do dia.
- As minhas colegas de moradia (**Jessica, Cibele, Mariane, Alcilane e Viviane**) que muito me ouviam e aguentavam meus lamentos do dia-a-dia, vocês contribuíram muito para as forças não faltarem no decorrer do trajeto;
- Aos meus queridos pais (**Arnaldo Negreiros de Sousa e Maria do Santo Barbosa de Sousa**) pelo amor e incentivo de sempre. Isso tudo é pra vocês e por vocês. Minha essência, minha base, obrigada por tudo.

- Ao meu anjo príncipe, meu único irmão **Francisco Gomes de Sousa Neto** (*in memoriam*), que desde o início do Doutorado me deixou. Mas sei que sua ausência foi apenas física, pois em todas as etapas dessa fase da minha vida sentia sua mão me empurrando e dizendo: “Vai mana, deixa de ser fraca, você consegue”. Obrigada meu anjo lindo no céu”.
- Aos meus lindos e amados sobrinhos, **Rihan Lucas e Marianna**, pelo carinho dado nos poucos momentos que era possível a minha presença.
- Ao meu querido e amado esposo, **Rogério Sousa Azevedo**, pelo incentivo, amor e carinho de sempre, e por compreender e aceitar todos os momentos de ausência. Seu apoio foi essencial;
- Ao meu cachorro, **Bolotinha**, que chegou na minha vida na parte de escrita da tese, meu companheiro inseparável debaixo da mesa, era um dos que me inspirava;
- **À minha família** de uma forma geral e **a todos meus amigos** pelo incentivo em realizar mais esse desafio;
- **A todos** que de alguma forma contribuíram no desenvolvimento da pesquisa e no meu crescimento pessoal.

Meus sinceros agradecimentos

RESUMO

NEGREIROS, A. B. Análise genética e genômica da abelha-sem-ferrão *Melipona rufiventris*: sequenciamento, desenvolvimento de ferramentas moleculares e inferências populacionais. 2019. 123f. Tese (Doutorado) – RENORBIO. Universidade Federal do Piauí, Teresina-PI, 2019.

A abelha brasileira *Melipona rufiventris*, vulgarmente conhecida como Uruçu Amarela, apesar do potencial ecológico e econômico, assim como as demais espécies desse grupo, tem sofrido bastante com a devastação ambiental provocada pelo homem, principalmente por conta da constante exploração agropecuária das regiões onde a espécie é endêmica, o que a colocou dentro da categoria de espécies ameaçadas de extinção pelo ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade). Porém, deve-se haver cautela quanto a esse enquadramento, principalmente pelo fato de se tratar de uma espécie com taxonomia indefinida. Assim, vê-se a necessidade do uso de marcadores moleculares, para não só solucionar essa questão, como também para identificar o perfil genético das populações presentes no Brasil de forma a auxiliar nas estratégias de conservação. Baseado nisso, o objetivo principal desse trabalho foi realizar um estudo genômico de frequência e abundância dos microssatélites nas abelhas-sem-ferrão *Melipona rufiventris*, *M. subnitida* e *M. fasciculata* utilizando a tecnologia do Sequenciamento de Nova Geração (NGS) e identificar e desenvolver marcadores microssatélites específicos e heteroespecíficos para *M. rufiventris* com validação feita por meio de inferência populacional de abelhas presentes em Minas Gerais, Goiás e Piauí. Para isso, a partir de uma parcela do genoma, uma plataforma de base de dados de sequências foi gerada, a partir do qual foram selecionadas e identificadas regiões microssatélites com potencial para estudos genéticos. No geral, obteve-se um total de 3.495 (137.313 *contigs*), 11.869 (47.081 *contigs*) e 93.588 SSRs (141.412 *contigs*), presentes em *M. rufiventris*, *M. fasciculata* e *M. subnitida*, respectivamente. Os tipos de microssatélites mais frequentes foram di e trinucleotídeos, comuns em regiões neutras do DNA. Trinta e sete marcadores heteroespecíficos (19 de *M. subnitida* e 18 de *M. fasciculata*) foram utilizados em *M. rufiventris*, sendo que maior parte da transferência foi relacionada à espécie *M. subnitida*, comprovada sua eficácia por meio de pequeno estudo populacional. De 25 marcadores desenvolvidos especificamente para *M. rufiventris*, 16 foram polimórficos, sendo capaz de demonstrar, como esperado, alta diferenciação genética ($F_{ST}=0,252$, $R_{ST}=0,317$ e $Dest= 0,284$) entre as populações de Goiás, Minas Gerais e Piauí. Por meio desse trabalho de Tese, obtivemos um conjunto de novos marcadores eficazes para inferências populacionais em abelhas *M. rufiventris*, corroborando inclusive com as suspeitas de que as populações, presentes nos biomas Cerrado, Caatinga e Mata Atlântica possam ser subespécies diferentes.

Palavras-chave: Abelhas Nativas, Caracterização Populacional, Marcadores Heteroespecíficos, Microssatélites, Sequenciamento Nova Geração.

ABSTRACT

NEGREIROS, A.B. Genetic and genomic analysis of the stingless bee *Melipona rufiventris*: sequencing, development of molecular tools and population inferences. 2019. 123f. Thesis (Doctor) – RENORBIO. Universidade Federal do Piauí, Teresina, 2019.

The Brazilian bee *Melipona rufiventris*, commonly known as “Uruçu Amarela”, despite the ecological and economic potential, as well as the other species of this group, has suffered a lot from the environmental devastation provoked by man, mainly due to constant agricultural speculation in the regions where species is endemic, placing it within the category of endangered species by ICMBio (Chico Mendes Institute for Biodiversity Conservation). However, caution should be exercised in this framework, mainly because it is a species with indefinite taxonomy. Therefore, it is necessary to use molecular markers to solve this problem, as well as to identify the genetic profile of the populations present in Brazil, in order to assist conservation strategies. Based on this, the main objective of this work was to carry out a genomic study of frequency and abundance of the microsatellites in the stingless bees *Melipona rufiventris*, *M. subnitida* and *M. fasciculata* using the New Generation Sequencing (NGS) technology and to identify and develop specific and heterospecific microsatellite markers for *M. rufiventris* with validation made through population inference of bees present in Minas Gerais, Goiás and Piauí. To this end, from a portion of the genome, a sequence database platform was generated, from which microsatellite regions with potential for genetic studies were selected and identified. In general, a total of 3,495 (137,313 contigs), 11,869 (47,081 contigs) and 93,588 SSRs (141,412 contigs), present in *M. rufiventris*, *M. fasciculata* and *M. subnitida*, respectively, were obtained. The most frequent types of microsatellites were di and trinucleotides, common in neutral regions of DNA. Thirty-seven heterospecific markers (19 of *M. subnitida* and 18 of *M. fasciculata*) were used in *M. rufiventris*, and most of the transfer was related to *M. subnitida* species, proved its efficacy by means of a small population study. Of the 25 markers developed specifically for *M. rufiventris*, 16 were polymorphic, being able to demonstrate, as expected, high genetic differentiation ($F_{ST} = 0.252$, $R_{ST} = 0,317$ and $Dest = 0.284$) among populations of Goiás, Minas Gerais and Piauí. Through this thesis, we obtained a set of new effective markers for population inferences in *M. rufiventris* bees, corroborating also the suspicions that the populations present in the Cerrado, Caatinga and Atlantic Forest biomes may be different subspecies.

Key-words: Native Bees, Population Characterization, Heterospecific Markers, Microsatellites, New Generation Sequencing.

LISTA DE FIGURAS

INTRODUÇÃO GERAL

Figura 1. Abelha *Melipona rufiventris* nos ninhos em troncos de árvores..... 21

CAPÍTULO 2

Figura 1. Purificação do DNA genômico, modificado de Sambrook et al. (1989). Modificações estão nos quadrados em amarelo. PCI – Fenol : Clorofórmio : Álcool isoamílico; CI – Clorofórmio : Álcool isoamílico; IE – Isopropanol : Etanol absoluto; TE – Tris e EDTA. 55

Figura 2. Box plot (valores máximos, mínimos e médios) das concentrações de DNA (ng / µl) obtidas de chelex (CHLX), HotShot (HS) fenol-clorofórmio (PCI-A e PCI-B) e Kit Comercial (KIT COM) protocolos de extração em diferentes tecidos (A, abdome; C, cabeça; P, perna; T, Tórax) de *Melipona subnitida*. Para cada gráfico de caixa, a linha dentro da caixa representa a mediana; as linhas superior e inferior representam 75 e 25 percentis dos dados, respectivamente; os traços superiores e inferiores representam 95 e 5 por cento, respectivamente. 59

Figura 3. Géis de Agarose com os diferentes tipos de tecido de *Melipona subnitida* (A, abdome; C, cabeça; P, perna; T, Tórax) e protocolos (chelex (CHLX), HotShot (HS) fenol-clorofórmio (PCI-A e PCI-B) e Kit Comercial (Kit com)). Integridade do DNA em gel de eletroforese (A-B-C); digestão utilizando enzimas de restrição, DraI e RsaI, (D-E-F); e amplificação por PCR utilizando marcadores ISSR (Inter Simple Sequence Repeat) (G-H-I). 60

CAPÍTULO 3

Figura 1. Fluxograma com as etapas para o preparo da biblioteca genômica..... 66

Figura 2. As etapas para identificação de elementos repetitivos em *tandem* no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de nova geração (NGS)..... 70

Figura 3. As métricas de controle de qualidade de sequências de DNA gerado pelo FastQC para *Melipona fasciculata* e *M. subnitida* com a distribuição da pontuação de qualidade de Phred sobre todas as leituras em cada base..... 71

Figura 4. Distribuição do número de *contigs* de diferentes comprimentos em *Melipona fasciculata* e *M. subnitida*..... 72

CAPÍTULO 4

- Figura 1. Os dez *motifs* de microssatélites tetranucleotídeos mais frequentes e abundantes encontrados no genoma sequenciado das abelhas-sem-ferrão *Melipona rufiventris* (A), *Melipona subnitida* (B) e *Melipona fasciculata* (C)..... 88
- Figura 2. Correlação entre Bases não-consenso e comprimento dos microssatélites puros interrompidos nas espécies de abelhas-sem-ferrão *Melipona rufiventris* (A), *M. fasciculata* (B) e *M. subnitida* (C). 91

CAPÍTULO 5

- Figura 1. Mapa do Brasil mostrando a região onde os ninhos de *Melipona rufiventris* (*Hymenoptera: Apidae*) foram amostrados no semiárido..... 98
- Figura 2. Gráfico de dispersão da análise de coordenadas principais (PCoA) usando locos microssatélites de *M. rufiventris*; (B) Determinação do melhor número de *clusters* a partir da análise STRUCTURE; (C) Gráfico de barras da estrutura da população Inferida usando o programa baseado em modelo de mistura de agrupamento bayesiano STRUCTURE (K = 2). Indivíduos são representados por cada barra. 105

CAPÍTULO 6

- Figura 1. Locais de amostragem de espécimes do "grupo *rufiventris*" nos biomas continentais brasileiros. Mata Atlântica (□), Caatinga (■) e Cerrado (▲). Este mapa do Brasil utilizado para a elaboração da Figura 1, e o shapefile para gerá-lo, foi extraído do banco de dados de domínio público do Instituto Brasileiro de Geografia e Estatística - IBGE (<http://mapas.ibge.gov.br/bases-e-referenciais.html>), e foi modificado usando o Inkscape 0.91 (<https://inkscape.org/en/>). 114
- Figura 2. (A) Determinação do melhor número de *clusters* genéticos da análise STRUCTURE para *loci* de microssatélites em colônias do grupo *Melipona rufiventris*; K é o número de *clusters*. (B) Gráfico de barras da estrutura populacional inferida usando o programa baseado em modelo de mistura de agrupamento bayesiano STRUCTURE (K = 2). (C) Gráfico de dispersão da análise de coordenadas principais (PCoA) usando *loci* SSR. ♦ Minas Gerais (Cerrado); ▲ Goiás (Mata Atlântica); ■ Piauí (Caatinga). 122

LISTA DE TABELAS

CAPÍTULO 2

- Tabela 1. Concentrações de DNA [média em ng / $\mu\text{L} \pm$ desvio padrão (coeficiente de variação em %)] de abelhas-sem-ferrão *Melipona subnitida* usando diferentes métodos de extração de DNA e fontes de tecidos. A seta vermelha indica o desempenho dos diferentes protocolos em uma parte anatômica da abelha analisada. A seta azul compara o desempenho de um determinado protocolo referente as diferentes partes anatômicas da abelha.....57
- Tabela 2. Proporção A260 / 280 [média \pm desvio padrão (coeficiente de variação em %)] de abelhas-sem-ferrão *Melipona subnitida* usando diferentes métodos de extração de DNA e fontes de tecidos. A seta vermelha indica o desempenho dos diferentes protocolos em uma parte anatômica da abelha analisada. A seta azul compara o desempenho de um determinado protocolo referente entre as diferentes partes anatômicas da abelha.....58

CAPÍTULO 3

- Tabela 1. Qualidade de Phred relacionada com precisão do sequenciamento.....68
- Tabela 2. Dados básicos gerados a partir da análise bioinformática.....72

CAPÍTULO 4

- Tabela 1. Cálculo da contagem total (N), comprimento total (bp), densidade (SSR em bp/Mb), abundância (N° de SSR/Mb) e frequência (%) dos microssatélites do genoma das abelhas *Melipona rufiventris* (*Mruf*), *M. fasciculata* (*Mfsc*) e *M. subnitida* (*Msub*). Mono – Mononucleotídeos, Di – Dinucleotídeos, Tetra – Tetranucleotídeos, Penta – Pentanucleotídeos e Hexa – Hexanucleotídeos85
- Tabela 2. Descrição dos quatro primeiros *motifs* com maior número (No) e abundância (N° SSR/Mb) dos microssatélites dinucleotídeos (Di), trinucleotídeos (Tri), tetranucleotídeos (Tetra), pentanucleotídeos (Penta) e hexanucleotídeos (Hexa) das espécies de abelhas-sem-ferrão *Melipona rufiventris*, *M. fasciculata* e *M. subnitida*87
- Tabela 3. Descrição do número (No), comprimento total, frequência, densidade e abundância dos tipos de microssatélites: composto, complexo, composto interrompido, puro e puro interrompido nas espécies de abelhas-sem-ferrão *M. rufiventris*, *M. fasciculata* e *M. subnitida*.
.....90

CAPÍTULO 5

Tabela 1. Marcadores, sequência de <i>primers</i> e parâmetros experimentais utilizados para a amplificação de <i>loci</i> de microssatélites em <i>Melipona rufiventris</i> (<i>Hymenoptera: Apidae</i>) e seu número de acesso ao GenBank.	100
Tabela 2. Variabilidade de nove <i>loci</i> de microssatélites e estimativa da diversidade genética em <i>Melipona rufiventris</i> (<i>Hymenoptera: Apidae</i>).	102
Tabela 3. Variabilidade de nove <i>loci</i> microssatélites e estimativas de diversidade genética em <i>Melipona rufiventris</i> (<i>Hymenoptera: Apidae</i>) para as três populações analisadas no Estado do Piauí.	104

CAPÍTULO 6

Tabela 1. Localização e número de colônias do grupo <i>rufiventris</i> amostradas nos biomas continentais brasileiros.	114
Tabela 2. Características e amplificação de 16 marcadores microssatélites polimórficos desenvolvidos para <i>Melipona rufiventris</i> (<i>Hymenoptera: Apidae</i>) (N = 50).	117
Tabela 3. Variabilidade de 16 <i>loci</i> de microssatélites e estimativas de diversidade genética no grupo " <i>Melipona rufiventris</i> " (<i>Hymenoptera: Apidae</i>).	121

LISTA DE ABREVIATURAS E SÍMBOLOS

ICMBio	Instituto Chico Mendes de Conservação da Biodiversidade
AFLP	Polimorfismos de comprimento de fragmento amplificado
EM	Em perigo/ <i>endangered</i>
ISSR	Entre sequências simples repetidas
AMOVA	Análise de variância molecular
DNA	Ácido desoxiribonucléico
HWE	Equilíbrio de Hardy-Weinberg
NGS	Sequenciamento de nova geração
PCI	Conteúdo de informação polimórfica
PCR	Reação em cadeia da polimerase
AFLPs	Polimorfismo de comprimento de fragmentos amplificados
RAPD	DNA polimórfico aleatoriamente amplificado
dNTPs	Desoxinucleotídeos
ddNTPs	Didesoxinucleotídeos
SNPs	Polimorfismo de nucleotídeo único
SSR	Sequência simples repetida
RFLP	Polimorfismo do comprimento de fragmentos de restrição
F_{ST}	Estimativa diferenciação genética entre subpopulações
SCAR	Sequence Characterized of Amplified Region
MCMC	Cadeia Markov de Monte Carlo

Sumário

INTRODUÇÃO GERAL	20
Referências	23
CAPÍTULO 1	25
1.1 Resumo	26
1.2 Introdução	27
1.3 DNA como fonte de marcadores moleculares	28
1.4 Técnicas moleculares para genética populacional de abelhas-sem-ferrão (Técnicas baseadas em PCR)	30
1.4.1 RAPD (Polimorfismo de DNA amplificado ao acaso).....	30
1.4.2 Análise entre sequências simples repetidas (ISSR).....	31
1.4.3 Análise do polimorfismo de comprimento de fragmento amplificado (AFLP)	32
1.4.4 Microssatélites	33
1.5 Sequenciamento de DNA	35
1.5.1 Técnica de Sequenciamento Sanger	35
1.5.2 Técnica de Sequenciamento de Nova Geração.....	36
1.6. Aplicações em abelhas-sem-ferrão	37
1.6.1 RAPD	37
1.6.2 ISSR.....	38
1.6.3 AFLP	39
1.6.4 Microssatélites	40
1.6.5 SNP.....	42
1.7 Considerações finais	42
1.8 Referências	43
CAPÍTULO 2	50
Avaliação da extração de DNA de diferentes fontes de tecidos para análise de marcadores moleculares baseados em PCR em abelhas-sem-ferrão	50
2.1 Resumo	51

2.2	Introdução	52
2.3	Material e Métodos	53
2.3.1	Material biológico	53
2.3.2	Protocolos de extração de DNA	53
2.3.3	Protocolo de extração 1 - Kit Comercial	53
2.3.4	Protocolo de extração 2 - Chelex 100.....	53
2.3.5	Protocolo de extração 3 - HotSHOT	54
2.3.6	Protocolo de extração 4 e 5 - extração utilizando fenol-clorofórmio-álcool isoamílico (A e B)	54
2.3.7	Avaliação quantitativa e qualitativa do DNA extraído.....	55
2.3.8	Digestão de restrição, validação de PCR e eletroforese	55
2.3.9	Análise estatística	56
2.4	Resultados e Discussão.....	56
2.4.1	Rendimento do DNA e taxa de absorvância.....	56
2.4.2	Análise de eletroforese de DNA extraído.....	60
2.5	Conclusão	60
2.6	Referências	61
	CAPÍTULO 3	63
	Metodologia para identificação de elementos repetitivos em tandem no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de nova geração em baixa cobertura.....	63
3.1	Introdução.....	64
3.2	Metodologia.....	65
3.2.1	Extração de DNA.....	65
3.2.2	Preparo da amostra para o sequenciamento.....	65
3.2.3	Preparo da biblioteca genômica.....	65
3.2.3.1	Tagmentação do DNA genômico	65
3.2.3.2	Amplificação via PCR e adição de adaptadores	66
3.2.3.3	Purificação das bibliotecas (PCR <i>Clean-up</i>)	66

3.2.3.4 Normalização da biblioteca e agrupamento das amostras	67
3.2.3.5 Agrupamento de bibliotecas	67
3.2.4 Sequenciamento do DNA (Illumina Nextera XT_V3)	67
3.2.5 Análise bioinformática	67
3.3 Resultados.....	69
ANEXO	73
3.4 Referências	77
CAPÍTULO 4	80
Mineração do genoma de espécies de abelhas do gênero <i>Melipona</i> para a busca de sequências simples repetidas: abundância, composição e variação do comprimento	80
4.1 Resumo	81
4.2 Introdução	82
4.3 Material e Métodos	83
4.4 Resultados e Discussões	84
4.5 Conclusões.....	92
4.6 Referências	92
CAPÍTULO 5	94
Estudo da diversidade genética de <i>Melipona rufiventris</i> (Hymenoptera: Apidae) no semiárido nordestino usando microssatélites heteroespecífico	94
5.1 Resumo	95
5.2 Introdução.....	96
5.3 Material e Métodos.....	98
5.3.1 Amostras de abelha e isolamento de DNA genômico	98
5.3.2 Teste e genotipagem de marcadores microssatélites heteroespecíficos	99
5.4 Análise estatística	100
5.5 Resultado e Discussão	101
5.6 Conclusão	106
5.7 Referências	106

CAPÍTULO 6	110
Desenvolvimento de marcadores microssatélite por meio do sequenciamento Illumina MiSeq para uso em genética de populações de abelha-sem-ferrão urucu-amarela (<i>Melipona rufiventris</i> , Hymenoptera, Meliponini).....	110
6.1 Resumo	111
6.2 Introdução.....	112
6.3 Material e Métodos.....	113
6.3.1 Amostragem e Extração de DNA	113
6.3.2 Preparação de biblioteca e sequenciamento de alto rendimento	115
6.3.3 Mineração SSR e desenho de primer.....	115
6.3.4 Amplificação por PCR e validação de SSRs selecionados.....	115
6.3.5 Análise estatística	115
6.4 Resultados e Discussão.....	118
6.4.1 Montagem de sequência e mineração SSR.....	118
6.4.2 Validação SSR.....	118
6.4.3 Divergência genética entre populações no grupo 'rufiventris'.....	119
6.5 Conclusão	122
6.6 Referências	122
7 CONSIDERAÇÕES FINAIS	126

INTRODUÇÃO GERAL

Considerados como visitantes florais obrigatórios, o grupo das abelhas-sem-ferrão (Ordem Hymenoptera, Família Apidae, Tribo Meliponini) possui importante papel ecológico de polinização das angiospermas, sendo essencial para a manutenção da flora nativa e, indiretamente, da fauna que dela se alimenta (BOLLIVIÁN, 2008). Além disso, são polinizadores importantes das culturas comerciais e, portanto, de grande importância biológica e econômica (JAFFÉ et al., 2015).

Abelhas-sem-ferrão estão entre os invertebrados sociais que fazem parte da biodiversidade local de muitos ecossistemas tropicais e subtropicais. No Brasil e em outras partes da América do Sul, a sua criação racional, atividade denominada meliponicultura, apresenta grande potencial em função da extensão territorial do país, que proporciona uma diversidade de ecossistemas com características adequadas à atividade, e pela diversidade de espécies desse grupo no território nacional, principalmente do gênero *Melipona*.

Apesar da meliponicultura ser uma atividade tradicional, com técnicas de manejo já consagradas popularmente, alterações antrópicas nos ambientes naturais têm colocado espécies deste grupo sobre risco de desaparecimento (QUEZADA-EUÁN; MAY-ITZÁ; GONZÁLEZ-ACERETO, 2001). Mesmo assim, o crescente interesse pelos serviços de polinização e de geração de produtos como mel e própolis, disponibilizados pelas abelhas tem contribuído para a ampliação de esforços para o desenvolvimento de estratégias de manejo com fins conservacionistas.

Dentro dos Meliponini, tem-se a espécie *Melipona rufiventris*, que ocorre com predominância no bioma Cerrado, constantemente ameaçada e alvo de fortes pressões antrópicas, devido ao grande poder agricultável das suas terras. A ação de meleiros, queimadas, uso inadequado de pesticidas e mudança do uso da terra vêm provocando grandes alterações na mata nativa, alterando com isso todo o ecossistema local, podendo comprometer desfavoravelmente esses importantes polinizadores (CARRECK, 2016).

Como consequência, a espécie *M. rufiventris* foi incluída pela primeira vez no livro vermelho da fauna brasileira ameaçada de extinção, coordenado pelo Instituto Chico Mendes de Conservação da Biodiversidade – ICMBio, o qual avalia o risco de extinção de todos os animais vertebrados e invertebrados com ocorrência no país, no qual a abelha-sem-ferrão *Melipona (Michmelia) rufiventris* Lepelletier, 1836 está inserida na categoria “em perigo – *endangered* (EN)” (ICMBIO, 2016).

Como é característico do gênero, a espécie *M. rufiventris* constrói seus ninhos dentro de cavidades e nos troncos das árvores (**Figura 1**) com baixa capacidade de dispersão, o que dificulta uma possível “fuga” quando o seu ninho está ameaçado por fatores externos. Geralmente, as rainhas de *M. rufiventris* se acasalam com um único macho durante o voo nupcial e, portanto, a variabilidade genética entre seus descendentes é naturalmente baixa (LOPES et al., 2009).

Figura 1. Abelha *Melipona rufiventris* nos ninhos em troncos de árvores.



Fotos: Fabia Mello e Ricardo Camargo

Desta forma, levando em consideração todas essas informações, a análise genômica e genética dessa espécie de abelha compreendem importantes informações que podem servir de pré-requisito para o estabelecimento de programas de manejo e conservação eficientes. Um melhor aproveitamento da espécie depende de um programa amplo de melhoramento, a começar por atividades de pré-melhoramento, como: caracterização morfológica, comportamental e molecular das populações. Existe, portanto, um interesse crescente no

desenvolvimento de marcadores genéticos que possibilitem estimar a diversidade genética nessas abelhas e assim contribuir nos esforços de conservação.

A fim de se avaliar as informações de variabilidade genética existente nas populações naturais ou introduzidas de abelhas, de todos marcadores moleculares os microssatélites de DNA, ou SSR (Sequência simples repetida) tem sido os eleitos por muitos pesquisadores.

Microssatélites são repetições curtas de um a seis pares de bases de DNA, repetidas em tandem, distribuídos com abundância no genoma dos eucariotos (DINIZ et al., 2007). São marcadores codominantes, de natureza multialélica, e altamente polimórficos, com poder de transferibilidade entre espécies e facilidade de detecção pela PCR, além de uma cobertura extensiva do genoma e necessidade de concentrações pequenas de DNA para às reações de PCR, sendo que constitui uma excelente ferramenta para análise populacional (LI et al., 2002). As desvantagens correspondem ao alto investimento inicial e ao tempo necessário para o desenvolvimento destes marcadores no caso de espécies nunca antes estudadas. Até o presente momento apenas Lopes et al. (2009) desenvolveram ferramentas moleculares específicas para a espécie *M. rufiventris*, no qual foram identificados e caracterizados oito *loci* de microssatélites a partir de técnicas tradicionais que utilizam, por base a clonagem molecular e sequenciamento Sanger.

No entanto, hoje já existem estratégias que ajudam a superar essas limitações, como o uso do sequenciamento de Nova Geração (*Next-Generation Sequencing-NGS*) para a identificação e isolamento de microssatélites e uso de marcadores heteroespecíficos disponíveis nas plataformas internacionais de bancos de dados de sequências, para minimizar a longo e a curto prazo o custo no desenvolvimento de marcadores. Este último vem sendo bastante útil pelo fato de as regiões adjacentes às repetições serem muito conservadas entre espécies distintas, próximas evolutivamente, o que garante a obtenção de *primers* heterólogos, bem como volume maior de marcadores disponíveis para as abelhas-sem-ferrão, de forma a auxiliar com maior precisão os estudos voltados à caracterização de populações, dando subsídios para a conservação da espécie (FRANCINI et al., 2009; FRANCINI et al., 2010).

Considerando a relevância desse grupo de abelhas para o ecossistema e para a economia de subsistência do Brasil, por meio do presente trabalho de tese, teve-se como objetivo principal realizar estudo genômico de frequência e abundância dos microssatélites na abelha-sem-ferrão *M. rufiventris* utilizando a tecnologia do Sequenciamento de Nova Geração (NGS), identificar e desenvolver marcadores moleculares específicos de microssatélites e por fim realizar inferências populacionais.

O trabalho desempenhado neste estudo está organizado em forma de capítulos, conforme segue: O capítulo I apresenta uma revisão baseada no levantamento dos principais trabalhos realizados com marcadores moleculares, que enfocam a diversidade e conservação da espécie. O capítulo II apresenta uma comparação entre cinco métodos de extração de DNA a partir de diferentes amostras teciduais de abelhas-sem-ferrão. O capítulo III apresenta a metodologia para identificação de elementos repetitivos em tandem no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de Nova Geração em baixa cobertura. O capítulo IV reporta os resultados da mineração do genoma de três espécies de abelha-sem-ferrão (*Melipona fasciculata*, *Melipona subnitida* e *Melipona rufiventris*) por sequências simples repetidas avaliando a abundância, composição e variação do comprimento. O capítulo V analisa a diversidade genética e estrutura populacional de *M. rufiventris* (Hymenoptera, Apidae) com marcadores microssatélites heterólogos. E por último, o capítulo VI trata do desenvolvimento de novos marcadores microssatélites em abelha *M. rufiventris* usando o sequenciamento NGS (Nova Geração – *Next Generation Sequencing*) com estudo populacional para validação.

Referências

- BOLLIVIÁN, J. M. P. P. **Abelhas Nativas sem ferrão**. São Leopoldo: Oikos, 128p, 2008.
- CARRECK, N. **Decline of Bees and Other Pollinators**. Biological and Environmental Hazards, Risks, and Disasters. cap.5: 2016.
- DINIZ, F. M. et al. Application of a double-enrichment procedure for microsatellite isolation and the use of tailed primers for high throughput genotyping. **Genetics and Molecular Biology**, v. 30, p. 380-384, 2007.
- FRANCINI, I. B. et al. Microsatellite loci for an endemic stingless bee *Melipona seminigra merrillae* (Apidae, meliponini) from Amazon. **Conservation Genetics Resources**, v. 1, n. 1, p. 487–490, 2009.
- FRANCINI, I. B. et al. Isolation and characterization of microsatellite loci in the stingless bee *Melipona interrupta manaosensis* (Apidae: Meliponini). **Conservation Genetics Resources**, v. 2, n. 1, p. 27–30, 2010.
- ICMBIO. Livro Vermelho da Fauna Brasileira Ameaçada de Extinção. **Livro Vermelho**, p. 1–76, 2016.
- JAFFÉ, R. et al. Bees for development: Brazilian survey reveals how to optimize stingless beekeeping. **PLoS ONE**, v. 10, n. 3, p. 1–21, 2015.
- LI, Y. et al. Microsatellites: genomic distribution, putative functions, and mutational

mechanism: a review. **Molecular Ecology**, v. 11, p. 253–256, 2002.

LOPES, D. M. et al. Microsatellite loci for the stingless bee *Melipona rufiventris* (Hymenoptera: Apidae). **Molecular Ecology Resources**, v. 9, n. 3, p. 923–925, 2009.

QUEZADA-EUÁN, J. J. G.; MAY-ITZÁ, W. DE J.; GONZÁLEZ-ACERETO, J. A. Meliponiculture in Mexico: Problems and perspective for development. **Bee World**, v. 82, n. 4, p. 160–167, 2001.

CAPÍTULO 1

Aplicações de marcadores moleculares em estudos populacionais em abelhas-sem-ferrão

Artigo de Revisão “*Molecular marker applications to population studies in stingless bees*”, Aline B. Negreiros e Fábio M. Diniz a ser submetido à revista **Arquivos do Instituto Biológico (ISSN 1808-1657)**.

1.1 Resumo

Pertencentes à tribo Meliponini (Hymenoptera, Apidae), as abelhas-sem-ferrão constituem um dos grupos mais abundantes de abelhas no mundo com os mais diferentes padrões comportamentais e morfológicos. Tem grande importância para a biodiversidade mundial sendo responsável pela polinização de uma boa parte da vegetação voltada ao consumo humano. Além disso, tem grande importância econômica na complementação de renda de pequenos criadores, devido à possibilidade de comercialização, principalmente de mel e própolis. No entanto, devido à constante devastação ambiental, muitas espécies estão ameaçadas em função da redução dos seus estoques populacionais. Por isso, muitos esforços vêm sendo empregados no sentido de ampliar o conhecimento das mais diferentes espécies, com destaque para os estudos genéticos usando os marcadores moleculares. Dessa forma, pretendeu-se com esse estudo de revisão incrementar uma discussão referente aos principais meios empregados nas pesquisas usando essas ferramentas moleculares em abelhas-sem-ferrão no intuito de esclarecer a atual situação daquelas relacionadas a esse tema. Diante disso, dentro do que foi levantado até o momento, ficou claro que os microssoatélites compreendem os marcadores com maior respaldo nos estudos genéticos populacionais com esse grupo de inseto, devido ao volume de informações que podem ser gerados e pela possibilidade de transferência de marcadores de uma espécie para outra, influenciado pelo baixo custo envolvido. Além do mais, verificou-se uma tendência de ampliação do número desses marcadores, principalmente devido a avanços crescentes na tecnologia de sequenciamento. Porém, fica evidente a necessidade da utilização de outros marcadores moleculares, principalmente por possibilitarem a adição de mais informações geradas de outro ponto de vista. Dentre esses, chama-se a atenção para os SNPs, pouco empregado em abelhas-sem-ferrão, que tem potencial não só para os estudos conservacionistas como referente àqueles voltados ao melhoramento genético, de forma a garantir a manutenção desse importante recurso natural.

Palavras – chave: Estudos genéticos, marcadores moleculares, redução populacional, Tribo Meliponini.

1.2 Introdução

O Brasil tem uma fauna diversificada de abelhas nativas eussociais que são caracterizadas principalmente pela atrofia do ferrão, sendo denominadas, também, de abelhas-sem-ferrão, o que as posicionam taxonomicamente na ordem Hymenoptera, família Apidae, subfamília Apinae e tribo Meliponini (COSTA et al., 2018). Com aproximadamente 192 espécies encontradas no país, essas abelhas estão entre as mais abundantes, sendo que representam um grande grupo monofilético, com presença em vários países das regiões Tropicais e Subtropicais do mundo, principalmente na faixa Pantropical (LEONHARDT; RASMUSSEN; SCHMITT, 2013).

Essas abelhas apresentam um comportamento de nidificação em ocos de árvores, no chão e nas mais diferentes cavidades, características essas que variam de espécie para espécie. Outra característica desse grupo é a formação de ninhos perenes, com geralmente uma única fêmea fértil e ativa, com vários zangões e numerosas operárias que coletam o alimento e cuidam da manutenção da colônia (MICHENER, 2000).

No geral, a hipótese mais aceita para o mecanismo de acasalamento é que as rainhas virgens desse grupo são fecundadas ao ar livre por um único macho, o que garante naturalmente uma baixa variabilidade genética populacional quando comparadas às abelhas do gênero *Apis* (MICHENER, 2007).

Contudo, além da produção de mel, esse grupo de abelhas se destaca pela polinização, oferecendo benefícios ecológicos cruciais, não só para vegetação selvagem, como para àquela voltada ao consumo humano. Onde grande parte do alimento humano consumido é oriundo direta ou indiretamente deste serviço ecossistêmico (WRATTEN et al., 2012).

A elevação anual das temperaturas causadas pelo aumento da emissão de gases de efeito estufa, assim como o desmatamento, uso indiscriminado de agrotóxicos e a ação predatória de meleiros, dentre outros, têm provocado alterações ambientais, o que afeta as espécies de diferentes formas, não só pela mortandade como pelo isolamento das populações (ZAYED, 2009). Diante disso, prevê-se uma redução do tamanho populacional e, por sua vez, perda da variabilidade genética. Neste contexto, estudos voltados à caracterização genética das populações podem ser essenciais para auxiliar no direcionamento de políticas de conservação deste importante recurso natural. Portanto, os marcadores moleculares, podem ser ferramentas que podem contribuir efetivamente para esse fim, de forma a permitir o monitoramento e melhorias na criação racional das espécies evitando a perda recorrente de abelhas.

A fim de ampliar a discussão, esta revisão teve por objetivo fazer uma abordagem dos estudos que envolvem as principais tecnologias de biologia molecular utilizadas até o

momento com abelhas-sem-ferrão, por meio de pesquisas indexadas nos principais portais de pesquisa (Web of Science e Google acadêmico).

1.3 DNA como fonte de marcadores moleculares

Uma das grandes vantagens dos marcadores moleculares baseados no DNA em estudos genéticos é a sua abundância, e em alguns casos à sua insensibilidade aos fatores do ambiente e o fato de o DNA estar presente em todas as células dos organismos.

O genoma nuclear e mitocondrial podem ser explorados como importantes fontes para gerar marcadores moleculares, com a possibilidade de serem testados, avaliados e aplicados em estudos biotecnológicos voltados a diversas áreas, com destaque para a agropecuária e medicina. São caracterizados basicamente como todo e qualquer fenótipo molecular oriundo de um segmento do DNA que pode ou não ser expresso em proteínas, avaliados em função do polimorfismo ou variação (deleção, inserção e substituição de base) apresentado entre diferentes indivíduos ou populações (FERREIRA; GRATTAPAGLIA, 1998; SONG et al., 2010; JIANG, 2017).

Graças ao avanço das tecnologias e o desenvolvimento de técnicas modernas em análises moleculares, como a tecnologia do DNA recombinante, a reação em cadeia da polimerase (PCR) e o sequenciamento, surgiu uma grande variedade desses marcadores, o que possibilitou identificar alterações genéticas diretamente no DNA (FALEIRO, 2007).

São muitos os marcadores moleculares existentes, podendo ser organizados basicamente em três classes: (1) marcadores baseados em hibridação de sondas, como o utilizado no polimorfismo do comprimento de fragmentos de restrição (RFLP); (2) Marcadores baseados em PCR, como é o caso do polimorfismo de DNA amplificado ao acaso (RAPD), polimorfismo de comprimento de fragmentos amplificados (AFLPs) e microssatélites ou repetição de sequência simples (SSR); e (3) marcadores baseados em sequência, por exemplo, polimorfismo de nucleotídeos único (SNP) (JIANG, 2017). Contudo, quando se trata de parâmetros genético-populacionais, os microssatélites se destacam dentre os demais devido serem uma classe de marcadores mais informativa (CAIXETA et al. 2014).

Esses diferentes tipos de marcadores moleculares existentes apresentam uma enorme amostragem do genoma, onde se destacam na caracterização da estrutura populacional e avaliação da diversidade genética, em aplicações de âmbito filogenético e evolutivo, assim como no auxílio em programas de melhoramento e conservação genética.

A fim de compreender de forma prática a aplicação dos marcadores para o melhoramento e conservação, os genótipos são avaliados por meio dos marcadores, de forma a identificar os perfis comuns, agrupando com base nas semelhanças e distinções moleculares observadas dentro e/ou entre as populações em estudo, bem como o grau de variação presente nestes. Geralmente isso é realizado por meio de softwares estatísticos, específicos para o tipo de marcador analisado. Eles podem ser baseados no padrão de fragmentos amplificados ou restringidos, que podem ser dominantes, baseados na ausência (recessivo) ou presença (dominante) de fragmentos (Ex.: RAPD – Polimorfismo de DNA amplificado ao acaso, ISSR – Entre sequências simples repetidas e AFLP – Polimorfismo no comprimento dos fragmentos amplificados), e codominantes, que tem como base a frequência dos tamanhos gerados (Ex.: Microsatélite e RFLP). A outra forma seria por meio do sequenciamento, utilizando como base de dados, as diferenças observadas nas sequências amplificadas e as frequências dos haplótipos (as diferentes sequências identificadas no *pool* de genes amplificados e sequenciados) (Ex.: SNPs) (GUIMARÃES et al., 2006).

No entanto, algumas diretrizes são necessárias para guiar os pesquisadores nos processos de escolha de amostras que norteiam métodos de análises genéticas com marcadores moleculares. Tais métodos envolvem o domínio de técnicas laboratoriais como, por exemplo, reações de PCR, eletroforese e dependendo do marcador, o sequenciamento de DNA. Dentre os principais fatores limitantes que se destacam para a aplicação dessa ferramenta em programas de melhoramento e conservação ainda têm sido o custo e em alguns casos, o tempo e a laboriosidade de certos procedimentos exigindo mão-de-obra especializada.

Dessa forma, o desenvolvimento de mais marcadores que abordem diferentes regiões do genoma que possibilitem o monitoramento da dinamicidade genética dos diferentes recursos naturais, aplicáveis, por exemplo, à agropecuária, torna-se crucial para o enriquecimento de ferramentas que sejam úteis para solucionar questões como, a caracterização de populações de organismos de interesse econômico, estimativa do grau de parentesco, grau de variabilidade genética, além da inferência de similaridade entre populações e espécies (MARQUES, 2002). À medida que se amplia a cobertura de avaliação do genoma do organismo abordado, mais se garante confiabilidade das informações geradas, o que possibilita maior ganho para programas de conservação, principalmente de espécies que estão ameaçadas pelas atividades antrópicas desordenadas, como é o caso das abelhas nativas do Brasil.

1.4 Técnicas moleculares para genética populacional de abelhas-sem-ferrão (Técnicas baseadas em PCR)

Desenvolvida em meados da década de 80, o desenvolvimento da técnica de reação em cadeia da polimerase (PCR), que amplifica segmentos específicos de DNA *in vitro*, permite que as análises sejam realizadas a partir de quantidades muito pequenas de material biológico, o que oferece uma variedade de ferramentas úteis para a detecção da variação genética sem a necessidade de clonagem molecular (SMITH, 2005). Essa técnica teve respaldo em diversas áreas da biologia quase que imediatamente, principalmente pela simplicidade da técnica (BARDAKCI, 2001).

Além da presença dos componentes necessários para a formação das novas fitas de DNA ($MgCl_2$, dNTP, tampão e *primers*), a técnica se caracteriza pelo uso, na reação, de uma única enzima, a DNA polimerase (ex.: Taq DNA polimerase, extraído da bactéria termofila *Thermus aquaticus*), e ciclos de temperatura, responsáveis pelo rompimento das pontes de hidrogênio, hibridização dos pequenos oligonucleotídeos (*primers*) responsáveis pelo início da reação, bem como pela ativação e ação da polimerase.

A técnica revolucionou as pesquisas nas áreas das ciências biológicas permitindo maior acesso ao genoma de forma a possibilitar, não só a manipulação como o desenvolvimento de uma imensa quantidade de marcadores moleculares úteis para os estudos genéticos.

Abaixo, no presente estudo, descrevem-se alguns dos principais marcadores usados em estudos populacionais, com foco em abelhas que utilizam como base a PCR.

1.4.1 RAPD (Polimorfismo de DNA amplificado ao acaso)

Dentre os marcadores que utilizam como estratégia a reação de PCR, tem-se o polimorfismo de DNA amplificado ao acaso (RAPD) que se baseia na repetição cíclica da extensão enzimática a partir de iniciadores de pequena extensão (pequenas sequências ou oligonucleotídeos complementares ao DNA) que se anelam nos dois extremos opostos de uma porção da fita de DNA que serve como molde, de forma a amplificar a região flanqueada no final do processo (ARAÚJO et al., 2003). Caracteriza-se pela produção de múltiplos fragmentos de DNA que servem como referência para a distinção entre indivíduos ou populações, gerados pelo polimorfismo de DNA ocasionados pelas mutações ocorridas nas regiões de DNA complementares ao *primer* único (oligonucleotídeo) que os flanqueia (WILLIAMS et al., 1990; WELSH ; MCCLELLAND,1990).

Para o processo de amplificação essa técnica que origina os marcadores RAPD usam os *primers* arbitrários e baixas temperaturas de anelamento. Os produtos amplificados são separados por eletroforese em gel de agarose e detectados por coloração direta com brometo de etídio ou similares (ex.: GELRED e SYBR). Não requer sondas específicas e não usa radioisótopos (BINNECK; NEDEL; DELLAGOSTIN, 2002). Além disso, não requer conhecimento prévio das sequências dos marcadores e pode produzir abundantes fragmentos polimórficos, o que torna assim um método simples, rápido, eficiente e barato (YANG; QUIROS, 1993). Soma-se a todas essas vantagens, o fato que os marcadores RAPD asseguram a avaliação de um número elevado de genótipos em um curto prazo de tempo, além de sofrerem pouca influência ambiental (COSTA et al., 2003).

Portanto, o RAPD tornou-se uma ferramenta poderosa e precisa para analisar a relação genética e a diversidade de muitos organismos, principalmente por serem eficazes na identificação de variações no DNA genômico a nível inter e intraespécies, além de poder identificar linhagens distintas e determinar o impacto genético da introdução de espécies exóticas em populações naturais (FREITAS et al., 2007).

1.4.2 Análise entre sequências simples repetidas (ISSR)

Dentre os outros tipos de marcadores que utilizam a PCR como estratégia de acesso, encontram-se aqueles que se localizam entre sequências simples repetidas (ISSR) que se baseiam na amplificação de fragmentos flaqueados por regiões microssatélites (regiões repetidas). O primeiro trabalho utilizando essa técnica foi realizado com plantas na década de 90, sendo estendido para animais, em meados dos anos 2000, a fim de revelar a diversidade e diferenciação entre as populações (ZIETKIEWICZ; RAFALSKI; LABUDA, 1994; CEKSTERYTE et al., 2012; SILVA et al., 2014).

Por se basearem nos microssatélites que estão amplamente distribuídos pelo genoma dos organismos eucariotos, os ISSR, geram vários fragmentos de DNA que podem servir como *barcoding* para caracterização das populações, o que garante boa cobertura do genoma nuclear, sendo que não há necessidade de conhecimento prévio das sequências que serão amplificadas (SILVA et al., 2014; MIRANDA et al., 2012; ARAUJO et al., 2016; NAYAK; PRAKASH, 2017).

Geralmente os fragmentos são amplificados por *primers* semi-arbitrários de 16 a 25 pares de bases desenvolvidos para serem complementares a regiões microssatélites ancoradas ou não na extremidade 5' ou 3' com um a três pares de bases diferentes. Além da simplicidade e baixo custo, o atrativo da técnica é a capacidade de maior reprodutibilidade das

reações quando comparado aos marcadores RAPD (GUPTA et al., 1994; ZIETKIEWICZ; RAFALSKI; LABUDA, 1994).

Assim como o marcador RAPD, não há necessidade de conhecimento prévio das sequências que serão amplificadas, sendo útil para a caracterização de espécies selvagens ou domésticas (SILVA et al., 2014; MIRANDA et al., 2012; ARAUJO et al., 2016; NAYAK; PRAKASH, 2017) .

Pelas suas características, esse tipo de marcador compreende mais uma ferramenta útil para determinar o grau de relação e diversidade genética entre as populações, principalmente no que se refere às abelhas.

1.4.3 Análise do polimorfismo de comprimento de fragmento amplificado (AFLP)

Outro marcador bastante útil em estudos genéticos populacionais são os AFLPs, polimorfismo do comprimento de fragmentos amplificados, que é uma técnica de impressão digital de DNA que facilita a varredura de múltiplos *loci* em todo o genoma para identificar mutações, inserções, deleções e rearranjos pontuais derivadas do polimorfismo. A análise AFLP desenvolvida por Vos et al. (1995) e patenteada pela empresa Keygene NV vem demonstrando ser uma ferramenta extremamente eficaz para distinção entre genótipos intimamente relacionados.

Em se tratando de rapidez, riqueza de polimorfismo, boa reprodutibilidade e resultados confiáveis, a AFLP tem as mesmas vantagens do polimorfismo do comprimento do fragmento de restrição (RFLP) e do RAPD. Além disso, a tecnologia AFLP não requer nenhum conhecimento prévio da sequência do genoma e muitos marcadores podem ser analisados em tempo razoável, quando comparado a outros métodos (LI et al., 2017). Ela pode ser aplicada a qualquer amostra de DNA, incluindo humano, animal, vegetal e microrganismos, gerando um potencial de se tornar um sistema universal de impressão de DNA (*fingerprint* DNA). Dentre os fatores que afetam a reprodutibilidade do AFLP tem-se a qualidade do DNA, sendo necessário um DNA de alta pureza para assegurar a completa digestão pelas endonucleases de restrição (BLEARS et al., 1998).

Esta técnica envolve três etapas principais: (i) restrição do DNA e ligação de adaptadores de oligonucleotídeo, (ii) amplificação seletiva de conjuntos de fragmentos de restrição e (iii) análise de gel dos fragmentos amplificados (VOS et al., 1995). O procedimento se inicia com o DNA genômico sendo digerido por duas enzimas de restrições e a ligação de adaptadores (oligonucleotídeos) de cadeia dupla, aos extremos 5' ou 3' dos fragmentos gerados durante a digestão. Esses fragmentos são amplificados por PCR

utilizando iniciadores (*primers*) complementares ao adaptador combinado com o sítio de restrição, mais um a três nucleotídeos ancorados. Sobre diferentes combinações de iniciadores, somente aqueles fragmentos com nucleotídeos complementares que se estendem entre adaptadores de um mesmo fragmento serão amplificados sob rigorosas condições. O polimorfismo é revelado pela análise de fragmentos amplificados em um gel de poliacrilamida desnaturante com as amostras amplificadas comparadas com padrões de peso molecular ou *ladders* (BLEARS et al., 1998).

1.4.4 Microssatélites

Os microssatélites ou SSR (*Simple Sequence Repeats*) são sequências de um a seis pares de bases repetidas “em tandem”, ou seja, adjacentes umas às outras, distribuídas no genoma (QUELLER; STRASSMANN; HUGHES, 1993). É considerado um dos marcadores mais polimórficos (SANTURE et al., 2010) e devido a sua característica de codominância, neutralidade, herança mendeliana, abundância e distribuição ao longo do genoma são os mais apropriados para estudos genéticos populacionais (HAASL; PAYSEUR, 2011). Além do que, vêm sendo apontados como a classe de marcadores mais informativa em estimativa de parâmetros genético-populacionais (GUICHOUX; LEPOITTEVIN; REVARDEL, 2011).

Essas características dos microssatélites, juntamente com a alta reprodutibilidade e simplicidade da PCR, tornam os microssatélites os marcadores mais utilizados em ecologia, biologia da conservação e estudos filogenéticos de vários seres (SELKOE; TOONEN, 2006). A principal limitação desta classe de marcadores reside na necessidade de isolar e desenvolver *primers* específicos para cada espécie, o que pode ser um processo laborioso, longo e caro. No entanto, esta tarefa foi facilitada pela disponibilidade de milhares de sequências em bancos de dados de DNA públicos e privados. O conhecimento dessas sequências e seu uso em conjunto com novas técnicas de bioinformática permitiram o desenvolvimento de *primers* para o SSR sem a necessidade de etapas iniciais complexas da estratégia tradicional, tornando o processo muito menos laborioso (CAIXETA et al. 2014).

O desenvolvimento de marcadores de microssatélites pode ser baseado em informações de sequência de DNA depositadas em bancos de dados, assim como na triagem de bibliotecas de DNA genômico construídas especificamente para a descoberta de sequências repetidas no genoma (RITSCHHEL et al., 2004). Uma vez desenvolvidos, os marcadores de microssatélites podem ser utilizados para uma variedade de fins em diferentes espécies (KIM et al., 2015).

Nesse sentido, a possibilidade de transferência de marcadores de *locus* de microssatélites entre espécies relacionadas, se torna uma estratégia útil, principalmente pelo baixo custo, sendo uma alternativa interessante em programas de melhoramento e conservação, para acompanhar o desdobramento das sequências repetidas que estão ligadas direta ou indiretamente a caracteres específicos, uma vez que os sítios que flanqueiam essas regiões são conservadas, principalmente entre espécies próximas evolutivamente (ZANE; BARGELLONI; PATARNELLO, 2002).

1.4.5 Polimorfismo de Nucleotídeos Único (SNP)

O polimorfismo de nucleotídeos único, frequentemente chamados SNPs, é o tipo mais comum de variação genética entre os indivíduos. É considerado um método rápido, simples, barato, e confiável de genotipagem, sendo caracterizado por apresentar diferenças de um único par de bases na sequência do DNA entre indivíduos de uma mesma espécie, sendo considerado marcador bialélico. As modificações ocorrem uma vez em cada 300 nucleotídeos em média, o que significa que há cerca de 10 milhões de SNPs no caso do genoma humano (BROOKES, 1999).

Os SNPs são formados por mutações e herdados como variantes alélicas, podendo ou não gerar diferenças fenotípicas, sendo uma ferramenta valiosa para estudos de mapeamento de genes (HAYASHI et al., 2004). Eles são usados para estudar a base genética de doenças comuns e traços complexos, como a altura ou a função metabólica, para estudar as origens e a ascendência das populações de diferentes espécies. Além disso, são usados em estudos de biodiversidade e tarefas comerciais como rastreabilidade, teste de paternidade e seleção de genótipos adequados (PARISET et al., 2006).

Mais comumente, essas variações são encontradas no DNA entre os genes. Eles podem atuar como marcadores biológicos, ajudando os cientistas a localizar genes associados às diversas características, inclusive, aquelas de cunho quantitativo. Quando os SNPs ocorrem dentro de um gene ou em uma região reguladora perto de um gene, eles podem desempenhar um papel mais direto, afetando a função do gene. Muitas destas variações polimórficas são neutras e não interferem na regulação e função dos genes, no entanto, alguns destes SNPs conferem modificações regulatórias que levam às diferenças na expressão gênica (WANG et al., 2005).

Dentre os inúmeros métodos de identificação de SNP que vêm sendo desenvolvidos, os atualmente disponíveis são: hibridação, anelamento ou não de *primers* complementares ao sítio SNP, digestão por enzimas de restrição, e sequenciamento. Além do uso de microchips,

atualmente, por meio do sequenciamento NGS, é possível obter uma ampla cobertura genômica de SNPs, principalmente com o uso de técnicas como o RESTseq, que compreende sequenciamento de uma biblioteca genômica composta de fragmentos restringidos e amplificados, o que vem a engrandecer ainda mais esse marcador, dando maior eficácia nos trabalhos de melhoramento e conservação (SCHINCARIOL, 2011).

1.5 Sequenciamento de DNA

O sequenciamento de DNA é um método laboratorial usado para identificar alterações na sequência de nucleotídeos que compõe a molécula de DNA. As técnicas de sequenciamento estão abrindo grandes oportunidades nas ciências biológicas, principalmente pelo fato de possibilitar de forma mais direta, acesso às informações herdadas nas moléculas de DNA. Basicamente até o momento são disponibilizadas duas técnicas principais de sequenciamento: a convencional (técnica de Sanger) e a técnica de sequenciamento de Nova Geração (MOREY et al., 2013).

1.5.1 Técnica de Sequenciamento Sanger

Até alguns anos atrás, os únicos métodos utilizados para o sequenciamento era a técnica de “didesoxi” enzimática de Sanger descrita pela primeira vez em 1977 (SANGER; NICKLEN; COULSON, 1977) e o método de degradação química de Maxam e Gilbert (MAXAM; GILBERT, 1977) descrito no mesmo ano, que foi usado em casos de sequências que não podiam facilmente ser resolvido com a técnica de Sanger.

A técnica de Sanger dominou por 30 anos o cenário dos projetos de sequenciamento, tendo sido a base utilizada para o desenvolvimento do Projeto Genoma Humano. Esse método geralmente é chamado de sequenciamento de "primeira geração" porque foi o primeiro a ser desenvolvido. No entanto, mesmo com grandes avanços em automação e otimização dessas técnicas durante décadas, esse método continua relativamente caro e demorado.

Esse método é baseado na incorporação de desoxinucleotídeos (dNTPs) e de didesoxinucleotídeos (ddNTPs) a uma cadeia de DNA em crescimento, tendo como molde o DNA de interesse (CARRARO; KITAJIMA, 2002). Os ddNTPs diferem dos dNTPs pela ausência da hidroxila usada para alongamento da cadeia, ou seja, quando um ddNTP é aleatoriamente adicionado a uma cadeia, sua replicação cessa. Como esses ddNTPs são marcados com fósforo radioativo (^{32}P) ou fluorescência, podem ser detectados e a sequência dos nucleotídeos identificada.

Anteriormente o sequenciamento Sanger era feito manualmente, no qual a sequência de bases do DNA era lida em um gel de poli(acrilamida) altamente concentrado, após a

exposição das bases identificadas em um filme de raio-X. As reações eram feitas separadamente para cada um dos quatro ddNTPs com a leitura realizada do fragmento maior para o menor.

Atualmente o mais utilizado é o sequenciamento automático, que utiliza sequenciadores com eletroforese vertical em placa (ABI377 Applied Biosystem), equipamento descontinuado, ou eletroforese em capilar (ABI3100 Applied Biosystem, e modelos mais atuais como ABI 3500, e outros), nos quais os próprios ddNTPs possuem marcação com fluorescência, ao contrário do método tradicional, em que estes apresentam marcação radioativa.

Dentre as limitações da técnica de Sanger tem-se a dificuldade da separação dos fragmentos de DNA marcados fluorescentemente, o número relativamente baixo de amostras que podem ser analisadas em paralelo e a dificuldade de automação total da amostra.

1.5.2 Técnica de Sequenciamento de Nova Geração

Com as limitações existentes no método de Sanger, iniciaram-se esforços para desenvolver técnicas sem géis que permitissem a determinação de sequências em números muito maior. Com a utilização de um microsistema com “chips”, hoje se pode distinguir pelo menos quatro gerações com características distintas de sequenciamento (ANSORGE, 2009).

Desde o ano de 2005 começaram a ser comercializadas as novas tecnologias de sequenciamento e estão em constante crescimento, tendo como denominação tecnologias de sequenciamento de Nova Geração (Next Generation Sequencing - NGS). Essas tecnologias promovem o sequenciamento de DNA em plataformas capazes de gerar informação sobre milhões de pares de bases em uma única corrida, permitindo assim, o sequenciamento de um genoma inteiro por um custo acessível, fazendo com que questões anteriormente intangíveis, fossem solucionadas (CARVALHO; SILVA, 2010).

As etapas envolvidas no NGS, inclui a preparação da biblioteca genômica, o sequenciamento através da síntese de DNA e a análise de dados por meio de softwares especializados. Procedimentos estes descritos e detalhados por autores como Metzker (2010), Head et al. (2014) e Pop & Salzberg (2008).

Basicamente a técnica se processa da seguinte forma: fragmentação do genoma; ligação dos adaptadores às extremidades dos fragmentos; construção da biblioteca genômica; sequenciamento dos *reads*; montagem dos *contigs* e seleção dos marcadores e/ou montagem do genoma. De acordo com Datto & Lundblad (2016) existem duas características que são exclusivas das abordagens do NGS, primeiro é a determinação dos dados de sequência de

fragmentos de DNA simples amplificados, evitando a necessidade de clonagem de fragmentos de DNA; e segundo que eles geram dados de sequência em tempo real à medida que cada novo nucleotídeo é incorporado em uma cadeia de DNA nascente.

Existem inúmeras vantagens nessa técnica como alta velocidade, alto rendimento e alta precisão (MA; GONG; JIANG, 2017). Para isso diversas empresas disponibilizam o serviço, na fabricação e execução de plataformas modernas responsáveis pela eficiência dos procedimentos. As que mais se destacam mundialmente são: a plataforma 454 FLX da Roche, Illumina da Solexa (HiSeq, MiSeq), Life/APG, HeliScope da Helicos, SOLiD da Applied Biosystems, PacBio da Pacific Biosciences, e Ion Torrent da Life Technology (LUO et al., 2012; GLENN, 2011; MARDIS, 2008; NIEDRINGHAUS et al., 2012; VAN DIJK et al., 2014).

Existe uma mudança conceitual significativa com essas novas abordagens tecnológicas do NGS, pois se um único *amplicon* de uma única amostra é amplificado, então uma sequência única é obtida. Se 100% da amostra for amplificada, 100% pode ser sequenciado a cada vez (HEATHER; CHAIN, 2016).

Um fator limitante dessa nova tecnologia, a curto prazo, continua sendo o alto custo, embora comparado com o sequenciamento do Sanger, o custo por base é menor em várias ordens de grandeza, o que acaba sendo compensatório, quando se estima a longo prazo. Outro gargalo a ser considerado, talvez o maior deles, é a gestão e interpretação da vasta quantidade de dados gerados por esses métodos, além da necessidade de computadores com alta capacidade de processamento (PETTERSSON; LUNDEBERG; AHMADIAN, 2009).

1.6. Aplicações em abelhas-sem-ferrão

1.6.1 RAPD

Ao utilizar o marcador RAPD em abelhas-sem-ferrão, identificam-se em plataformas de pesquisas como *Web of Science*, alguns estudos, principalmente voltados à caracterização e diferenciação populacional. Tem sido relatado como uma ferramenta eficiente para diferenciar populações oriundas de locais geograficamente distintos, tendo sido útil para verificar a existência de populações de espécies que surgiram por meio da seleção natural sob diferentes condições ambientais ou como resultado de deriva genética, aceleradas por intervenção humana (FUCHS et al., 1998).

Alguns estudos demonstraram, por exemplo, a capacidade desses marcadores de distinguir subespécies de abelhas mandaçaia (*Melipona quadrifasciata*) oriundas de diversas regiões do Brasil (WALDSCHMIDT; BARROS; CAMPOS, 2000; WALDSCHMIDT et al.,

2002; TAVARES et al., 2013a). Com o mesmo intuito, Tavares et al. (2007) confirmaram, por meio dos marcadores RAPD que as populações de abelhas Uruçu-amarela presentes no Cerrado e na Mata Atlântica seriam duas espécies diferentes, sendo elas a *M. rufiventris* e *M. mondury*, respectivamente, o que sugere que cada uma deva ser tratada diferentemente para fins de conservação.

No estudo realizado por Oliveira et al. (2004) em populações de *Tetragonisca angustula*, de 25 localidades oriundas de três diferentes países da América Latina, detectou-se dois grupos genéticos distintos, o que sugere que esta espécie está passando por diferenciação genética, impulsionado por fatores ecológicos, evolutivos ou mesmo antrópicos.

Em outro estudo com *T. angustula* oriunda da região Sudeste do Brasil, utilizando 12 iniciadores de marcadores RAPD, verificou-se um alto polimorfismo genético, sendo detectado 171 fragmentos, dos quais 150 foram polimórficos (87,72%) (BAITALA et al., 2006).

No entanto, além desses estudos com abelhas, outras espécies de Hymenoptera, usaram essa técnica, como foi o caso de Landry et al. (1993) onde usaram marcadores RAPD para caracterizar biótipos de *Anaphes sordidatus* e uma nova espécie deste gênero, *Anaphes* sp.nov, além de três espécies de *Trichogramma*. No trabalho de Lu & Rank (1996) os marcadores RAPD demonstraram que são uma ferramenta conveniente para estimar a variação genética populacional em *Megachile rotundata* e que, com um tamanho de amostra adequado, a técnica é aplicável à avaliação da divergência em populações diplóides.

1.6.2 ISSR

Com o início de sua utilização em animais em meados dos anos 2000, viu-se a possibilidade de se utilizar os marcadores ISSR em insetos, incluindo as espécies de abelhas que são pouco exploradas nos estudos genéticos, como é o caso da tribo Meliponini.

A tribo Meliponini possui uma imensa diversidade de espécies, o que impossibilita a disponibilidade de marcadores específicos para muitas abelhas. Diante disso, marcadores como o ISSR tornam-se eficazes para estudos com esse grupo de insetos.

Num levantamento feito recentemente por Silva et al. (2019) na busca de estudos realizados com abelhas-sem-ferrão em todo o mundo indexados no portal da *Web of science*, até início de 2018, existiam apenas seis trabalhos que englobavam os marcadores ISSR, sendo que todos foram desenvolvidos por autores brasileiros com o propósito principal de caracterização populacional. O Brasil é conhecido pela grande diversidade de espécies de

abelhas que englobam esse grupo, o que reflete como o país que realiza mais estudos com esse tipo de marcador.

Um dos estudos que podem ser ressaltados, é o realizado por Nascimento et al. (2010) com populações de *Melipona quadrifasciata* presentes em algumas localidades de Minas Gerais. Nesse estudo, foram obtidos no geral 119 fragmentos amplificados com 68% polimórficos. Foi possível verificar uma alta estruturação populacional ($F_{ST}= 0,65$), provavelmente devido ao isolamento geográfico corroborado por outras pesquisas usando outros marcadores. Além desse, outros estudos também tem contribuído para caracterizar essa espécie, como foi o caso do trabalho realizado por Araújo et al. (2016), no qual comprovou a alta distinção ($F_{ST}=0,52$) das colônias presente no estado de Sergipe (Semiárido) e da foz do Rio São Francisco, evidenciando uma possível barreira geográfica entre as duas localidades.

Outro estudo foi o realizado por Silva et al. (2014), no qual utilizando esses marcadores, foi possível evidenciar uma grande distinção genética ($F_{ST}=0,34$) entre colônias de *Melipona subnitida* presentes no semiárido do Rio grande do Norte e litoral piauiense.

Apesar de não ter sido observado indexação nos registros levantados durante o período pesquisado por Silva et al. (2019), há de se destacar também, um único estudo na Índia, realizado por Nayak e Prakash (2017) que usando ISSR, conseguiram detectar uma moderada estruturação genética entre as diferentes espécies de *Tetragonula* presentes em habitats urbanos e naturais.

1.6.3 AFLP

No que diz respeito à aplicação do marcador AFLP em abelhas-sem-ferrão, vale ressaltar o estudo realizado por Makert; Paxton; Hartfelder (2006) que descreveram um protocolo otimizado (protocolo original VOS et al., 1995) para a abelha-sem-ferrão *M. quadrifasciata* oriunda de três colônias mantidas no apiário do Departamento de Genética no campus da Universidade de São Paulo em Ribeirão Preto. Este protocolo difere do original pelo fato de terem usado o processo de digestão com enzimas de restrição em dois passos, com *EcoRI* e *TruII*. Nessa análise obtiveram 719 marcadores e a análise da variabilidade genética revelou que, em média, 32% dos marcadores eram polimórficos dentro de uma colônia, onde esse grande número de marcadores deixa claro o potencial deste protocolo AFLP otimizado em estudos genéticos avançados sobre abelhas-sem-ferrão, especialmente para análises de alta cobertura genômica.

Outro estudo utilizando o AFLP foi realizado por Thummajitsakul et al. (2008) com a espécie de abelha-sem-ferrão *Trigona pagdeni*. Os autores coletaram as abelhas das regiões

Norte, Nordeste, Central e peninsular da Tailândia, sendo que foi identificado, altos níveis de variação genética entre os indivíduos em todas as populações, com a heterozigose média esperada maior no Nordeste.

Ainda dentro do mesmo grupo de pesquisa, com um trabalho mais detalhado, os autores Thummajitsakul et al. (2010) apresentaram um estudo que teve como objetivo a identificação de uma região do DNA que fosse útil para distinguir abelhas do gênero *Trigona* na Tailândia. Para esse feito, um fragmento de AFLP encontrado especificamente em *T. pagdeni* foi clonado e sequenciado, no qual foi possível distinguir espécies semelhantes de *T. pagdeni* e *T. fuscobalteata*. Essa abordagem derivada é também chamada de caracterização de sequência de região amplificada (SCAR – *Sequence Characterized of Amplified Region*).

Realizando a mesma estratégia, Theeraapisakkun et al. (2010) desenvolveram um marcador para identificar a origem das espécies de abelhas-sem-ferrão (*Trigona collina*). Inicialmente, foi realizada uma análise do polimorfismo de comprimento de fragmento amplificado de 11 espécies usando 64 combinações de *primers*. Onde, verificou-se que o marcador desenvolvido é conveniente e rentável para a diferenciação entre as espécies de abelhas-sem-ferrão mostrando que esse trabalho ajudará a evitar erros de identificação taxonômica.

A principal limitação dos marcadores AFLPs refere-se ao alto custo e laboriosidade da técnica, porém são úteis para estudos populacionais e para a construção de mapas genéticos, uma vez que abrangem uma ampla cobertura do genoma.

1.6.4 Microssatélites

Nos últimos anos, os microssatélites atraíram a atenção de pesquisadores em todo mundo, pelo grande volume de informações fornecido. Embora a ênfase inicial fosse em humanos, a análise de microssatélites se tornou uma ferramenta poderosa para estudos em muitos seres vivos nas mais diversas áreas das ciências biológicas (OLIVEIRA et al., 2006). No mundo até o presente momento já foram desenvolvidos 18 conjuntos de marcadores microssatélite específicos para as abelhas-sem-ferrão, incluído espécies de vários gêneros, inclusive as do gênero *Melipona*, com maior número de espécies no Brasil.

Contudo, os primeiros a serem descritos foram os desenvolvidos por Peters et al. (1998), onde identificaram 33 *loci* obtidos de *Melipona bicolor*. No ano seguinte Paxton, Weibschuh; Quezada-Euán (1999) descreveram mais cinco *loci* para a espécie *Scaptotrigona postica*. A partir daí marcadores para outras espécies também foram desenvolvidos: *Trigona carbonaria* (GREEN; FRANCK; OLDROYD, 2001), *Nannotrigona testaceicornis*

(OLIVEIRA et al., 2009), *Tetragonisca angustula* (BRITO et al., 2009), *Melipona seminigra merrillae* (FRANCINI et al., 2009), *Melipona rufiventris* (LOPES et al., 2009), *Melipona interrupta manaosensis* (FRANCINI et al., 2010), *Melipona mondury* (LOPES et al., 2010), *Plebeia remota* (FRANCISCO et al., 2011), *Scaptotrigona xanthotricha* (DUARTE et al., 2012), *Scaura latitarsis* (FRANSCISCO et al., 2013), , *Melipona quadrifasciata* (TAVARES et al., 2013a), *Melipona scutellaris* (TAVARES et al., 2013b) e *Melipona capixaba* (NOGUEIRA et al., 2014). Além desses vale destacar os estudos realizados mais recentemente com *Melipona subnitida* (SOUZA et al., 2014), *Trigona spinipes* (SANTIAGO et al., 2015; JAFFÉ et al., 2016) e *Melipona fasciculata* (SILVA et al., 2018), onde desenvolveram marcadores específicos utilizando a técnica de sequenciamento de Nova Geração (SNG), que é bastante promissora, uma vez que é possível obter um grande volume de marcadores, em função da grande cobertura genômica proporcionada por esse tipo de sequenciamento.

A eficácia desses marcadores pode ser observado em trabalhos recentes como o desenvolvidos por Francisco et al. (2017), no qual utilizando marcadores específicos para avaliar as populações de *Tetragonisca angustula*, conseguiram verificar que o padrão de dispersão característico da espécie pode ter sido um dos principais fatores responsáveis pela diferenciação genética detectada entre os estoques presentes nas mais diferentes regiões do Brasil.

Com esses marcadores disponíveis nos bancos de dados internacionais de sequências como o Genbank, uma diversidade de trabalhos tem implementado ampliações de *loci* gerados em espécies próximas evolutivamente, por meio de *primers* heterólogos de forma a reduzir os custos com o desenvolvimento. Um desses estudos foi o realizado por Viana et al. (2013) para identificar a variabilidade genética interpopulacional de *M. scutellaris* em diferentes partes da região do Recôncavo baiano, Brasil, utilizando dez pares de *primers* microssatélites desenvolvidos para diferentes espécies de Meliponini (*Melipona bicolor*, *Melipona mondury*, *Melipona rufiventris* e *Trigona carbonaria*). Os resultados indicaram pouca estrutura genética, revelando os efeitos das atividades humanas sobre as populações de *M. scutellaris*.

Cabe destacar também o trabalho desenvolvido por Hurtado-Burilo et al. (2016) com as espécies *Scaptotrigona mexicana* e *S. pectoralis*, no qual utilizaram marcadores específicos e heteroespecíficos de microssatélites. Neste estudo, os autores conseguiram detectar uma significativa diferenciação populacional em ambas as espécies, de colônias presentes na costa do Pacífico e do Atlântico. No estudo foi proposto mudanças nas estratégias de conservação e

manejo dos genótipos presentes nessas duas regiões, uma vez que os resultados foram confirmados pelos dados morfométricos.

1.6.5 SNP

Demonstrando a utilidade dos SNPs em insetos, destaca-se aqui o estudo, no qual combinado com microssatélites, foi possível construir um mapa genético, onde foi identificado um novo *locus* de resistência do pulgão (*Aphis craccivora*) em feijão-caupi (KUSI et al., 2018).

Entretanto, no trabalho desenvolvido por Stolle & Moritz (2013) foi apresentado uma nova abordagem denominada sequenciamento de fragmento de restrição (RESTseq) para melhorar os métodos existentes, sendo bem promissor para estudos com abelhas, com a utilização de maior cobertura genômica. Nesse estudo, os autores utilizaram duas fêmeas diplóides de uma única colônia de abelhas (*Apis mellifera*) da Alemanha e quatro machos de abelha-sem-ferrão de *Nannotrigona perilampoides* do México para a validação do RESTseq.

Outro estudo recente com SNP, tendo como foco de estudo *A. mellifera*, objetivou validar um teste híbrido de DNA em comparação com os dados da sequência do genoma inteiro (WGS) e comparar proporções de ancestralidade deduzidas de dois tipos diferentes de marcadores (microssatélites e SNPs) para dar recomendações para futuras estratégias de manejo com foco na conservação da espécie. Os autores recomendaram o SNP como uma importante ferramenta a ser aplicada para esse fim, de forma a auxiliar ferramentas como os microssatélites, já consagrados em estudos com abelhas (PAREJO et al. 2018).

Entretanto, apesar da escassez de utilização dos marcadores SNPs em pesquisas com abelha-sem-ferrão, é conhecido sua utilidade em diversos estudos com outros animais, inclusive, insetos, com destaque para aqueles voltados à caracterização e mapeamento genético. Isso o qualifica com potencial para ser empregado nas mais diferentes pesquisas com essas abelhas, principalmente para solucionar questões complexas relacionadas ao padrão comportamental e reprodutivo, que é bastante complexo nesse grupo.

1.7 Considerações finais

As informações aqui descritas mostram uma visão geral sobre os mais relevantes estudos genéticos que utilizaram ferramentas moleculares envolvidas na conservação de diferentes abelhas, principalmente nas do gênero *Melipona*, a fim de auxiliar nos trabalhos voltados à genética populacional da espécie, e com isso auxiliar e elaborar estratégias para a conservação e também para programas de melhoramento, devido a sua grande importância

econômica. Sendo destaque, a utilização de marcadores moleculares microsatélites específicos e heteroespecíficos que têm-se mostrado essenciais na tomada de decisão quanto à conservação e manejo dos meliponíneos.

1.8 Referências

- ANSORGE, W. J. Next-generation DNA sequencing techniques. **New Biotechnology**, v. 25, n. 4, p. 195–203, 2009.
- ARAUJO et al. Risk of local extinction and genetic diversity of *Melipona quadrifasciata* (Apidae: Meliponini) in a possible Northeastern limit of its distribution in Brazil. **Sociobiology**, v.63, n.2, p.804-812, 2016.
- ARAÚJO, E. . et al. Uso de rapd para análise de diversidade genética em arroz. **Agronomia**, v. 3, p. 33–37, 2003.
- BAITALA, T. V. et al. RAPD Polymorphism in *Tetragonisca angustula* (Hymenoptera; Meliponinae, Trigonini) Populations. **Sociobiology**, n. 3, v. 48, 2006.
- BARDAKCI, F. Random Amplified Polymorphic DNA (RAPD) Markers. **Turkish Journal of Biology** n.25, p.185-196, 2001.
- BINNECK, E.; NEDEL, J. L.; DELLAGOSTIN, O. A. Análise de RAPD na identificação de cultivares: uma metodologia útil? **Revista Brasileira de Sementes**, v. 24, n. 1, p. 183–196, 2002.
- BLEARS, M. J. et al. Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. **Journal of Industrial Microbiology & Biotechnology**, v. 21, p. 99–114, 1998.
- BRITO, R. M. et al. Characterization of microsatellite loci of *Tetragonisca angustula* (Hymenoptera, Apidae, Meliponini). **Conservation Genetics Resources**, v. 1, n. 1, p. 183–187, 2009.
- BROOKES, A. J. The essence of SNPs. **Gene**, v. 234, n. 2, p. 177–186, 1999.
- CAETANO, A. R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, Viçosa , v. 38, n. spe, p. 64-71, 2009 .
- CAIXETA, E.T. et. al. Molecular Markers. **Biotechnology Applied to Plant Breeding**, p.19-45, 2014.
- CARRARO, D. M.; KITAJIMA, J. P. Sequenciamento e Bioinformática de Genomas Bacterianos. **Biociência & Desenvolvimento**, v. 28, p. 16–20, 2002.
- CARVALHO, M. C. D. C. G. DE; SILVA, D. C. G. DA. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural**, v. 40, n. 3, p. 735–744, 2010.
- CEKSTERYTE et al. Genetic characterization of Lithuanian honey bee lines based on ISSR polymorphism. **Apidologie**, v.43, p.652-662, 2012.
- COSTA, A. C. V. DA et al. Sensory and volatile profiles of monofloral honeys produced by native stingless bees of the Brazilian semiarid region. **Food Research International**, v. 105,

n. August 2017, p. 110–120, 2018.

COSTA, M. R. et al. Similaridade genética de cultivares de mandioca (*Manihot esculenta*) por meio de marcadores RAPD. **Ciência e Agrotecnologia, Lavras**, v. 27, n. 1, p. 158–164, 2003.

DATTO, M.; LUNDBLAD, R. L. **DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing)**. [s.l.] Elsevier Ltd., 2016. v. 1

DUARTE, O. M. P. et al. Isolation and characterization of microsatellites from *Scaptotrigona xanthotricha* (Apidae, Meliponini): A stingless bee in the Brazilian Atlantic rainforest. **Apidologie**, v. 43, n. 4, p. 432–435, 2012.

FALEIRO, G. **Marcadores genético-moleculares: aplicados a programas de conservação e uso de recursos genéticos**, Planaltina, DF: Embrapa cerrados, 102p., 2007.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética** Embrapa-Cenargen, 1998.

FRANCINI, I. B. et al. Isolation and characterization of microsatellite loci in the stingless bee *Melipona interrupta manausensis* (Apidae: Meliponini). **Conservation Genetics Resources**, v. 2, n. 1, p. 27–30, 2010.

FRANCINI, I. B. et al. Microsatellite loci for an endemic stingless bee *Melipona seminigra merrillae* (Apidae, meliponini) from Amazon. **Conservation Genetics Resources**, v. 1, n. 1, p. 487–490, 2009.

FRANCISCO, F. O. et al. Isolation and characterization of 15 microsatellite loci in the stingless bee *Plebeia remota* (Apidae: Meliponini). **Conservation Genetics Resources**, v. 3, n. 3, p. 417–419, 2011.

FRANCISCO, F.O. et al. Population structuring of the ubiquitous stingless bee *Tetragonisca angustula* in southern Brazil as revealed by microsatellite and mitochondrial markers. **Insect science**, v.24, p.877-890, 2017.

FRANCISCO, F. O. et al. Characterization of microsatellite loci for the stingless bee *Scaura latitarsis* (Hymenoptera, Apidae, Meliponini). **Journal of Hymenoptera Research**, v. 30, p. 7–11, 2013.

FREITAS, K. C. et al. Reação de amplificação aleatória de dna polimórfico a partir de amostras de Robalo Peva *Centropomus Parallelus*. **Estudos de Biologia**, v. 29, n. 67, p. 151–156, 2007.

FUCHS, H. et al. Application of molecular genetic markers for the differentiation of bream (*Abramis brama* L.) populations from the rivers Main and Danube. **Journal of Applied Ichthyology**, v. 14, n. 1–2, p. 49–55, 1998.

GLENN, T. C. Field guide to next-generation DNA sequencers. **Molecular Ecology Resources**, v. 11, n. 5, p. 759–769, 2011.

GREEN, C.; FRANCK, P.; OLDROYD, B. Characterization of microsatellite loci for *Trigona carbonaria*, a stingless bee endemic to Australia. **Molecular Ecology Notes**, v. 1, n. 2001, p. 89–92, 2001.

GUICHOUX, E. et al. Current trends in microsatellite genotyping. **Molecular Ecology Resources**, v.11, p. 591–611, 2011.

GUIMARÃES, C. et al. Marcadores moleculares e suas aplicações no melhoramento genético. **Conservation Genetics**, v. 7, n. 2, p. 167–184, 2006.

- GUPTA, M. et al. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. **Theoretical and Applied Genetics**, v.89, p. 7– 8, 1994.
- HAASL, R. J.; PAYSEUR, B. A. Multi-locus inference of population structure: a comparison between single nucleotide polymorphisms and microsatellites. **Heredity**, v. 106, n. 1, p. 158–171, 2011.
- HAYASHI, K. et al. Development of PCR-based SNP markers for rice blast resistance genes at the Piz locus. **Theoretical and Applied Genetics**, v. 108, n. 7, p. 1212–1220, 2004.
- HEAD, S. R. et al. Library construction for next-generation sequencing: Overviews and challenges. **BioTechniques**, v. 56, n. 2, p. 61–77, 2014.
- HEATHER, J. M.; CHAIN, B. The sequence of sequencers: The history of sequencing DNA. **Genomics**, v. 107, n. 1, p. 1–8, 2016.
- HURTADO-BURILLO, M. et al. A geometric morphometric and microsatellite analyses of *Scaptotrigona mexicana* and *S. pectoralis* (Apidae: Meliponini) sheds light on the biodiversidade of Mesoamericana stingless bees, **Journal of Insect Conservation**, p1-11, 2016.
- JAFFÉ, R. et al. Landscape genetics of a tropical rescue pollinator. **Conservation Genetics**, v. 17, n. 2, p. 267–278, 2016.
- JIANG, G. **Molecular Markers**. Encyclopedia of Applied Plant Sciences, 2 ed, v. 2 Petersburg, USA: Virginia State University, 2017.
- KIM, H. K. et al. Microsatellite markers developed by next-generation sequencing differentiate inbred lines of *Apis mellifera*. **Journal of Asia-Pacific Entomology**, v. 18, n. 4, p. 801–805, 2015.
- KUSI, F. et al. A novel aphid resistance locus in cowpea identified by combining SSR and SNP markers. **Plant Breeding**, v. 137, n. June 2017, p. 203–209, 2018.
- LANDRY, B. S.; DEXTRAZE, L.; BOIVIN, G. Random amplified polymorphic DNA markers for DNA fingerprinting and genetic variability assessment of minute parasitic wasp species (Hymenoptera: Mymaridae and Trichogrammatidae) used in biological control programs of phytophagous insects. **Genome**, v. 36, n. 3, p. 580–587, 1993.
- LEONHARDT, S. D.; RASMUSSEN, C.; SCHMITT, T. Genes versus environment: geography and phylogenetic relationships shape the chemical profiles of stingless bees on a global scale. **Proceedings of the Royal Society B: Biological Sciences**, v. 280, n. 1762, p. 20130680–20130680, 2013.
- LI, S. W. et al. Genetic diversity and spread of *Lissorhoptrus oryzophilus* (Coleoptera: Curculionidae) in China, based on amplified fragment length polymorphism. **Journal of Asia-Pacific Entomology**, v. 20, n. 3, p. 935–941, 2017.
- LOPES, D. M. et al. Comparative study on the use of specific and heterologous microsatellite primers in the stingless bees (Hymenoptera, Apidae). **Genetics and Molecular Biology**, v. 33, n. 2, p. 4, 2010.
- LOPES, D. M. et al. Microsatellite loci for the stingless bee *Melipona rufiventris* (Hymenoptera: Apidae). **Molecular Ecology Resources**, v. 9, n. 3, p. 923–925, 2009.
- LU, R. & RANK, G.H. Use of RAPD analyses to estimate population genetic parameters in the alfalfa leaf-cutting bee, *Megachile rotundata*. **Genome**, v.39, p.655-663, 1996.

- LUO, C. et al. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. **PLoS ONE**, v. 7, n. 2, 2012.
- MA, R.; GONG, J.; JIANG, X. Novel applications of next-generation sequencing in breast cancer research. **Genes & Diseases**, v. 4, n. 3, p. 149–153, 2017.
- MAKERT, G. R.; PAXTON, R. J.; HARTFELDER, K. An optimized method for the generation of AFLP markers in a stingless bee (*Melipona quadrifasciata*) reveals a high degree of intracolony genetic polymorphism. **Apidologie**, v. 37, n. 6, p. 687–698, 2006.
- MARDIS, E. R. Next-Generation DNA Sequencing Methods. **Annual Review of Genomics and Human Genetics**, v. 9, n. 1, p. 387–402, 2008.
- MARQUES, D. K. S. **Aplicação da biologia molecular em programas de conservação de recursos pesqueiros**. Corumbá: Embrapa Pantanal, (Documentos), v. 36, p. 22, 2002.
- MAXAM AM AND GILBERT W. **A new method for sequencing DNA**, n.74, p.560–564, 1977.
- METZKER, M. L. Sequencing technologies the next generation. **Nature Reviews Genetics**, v. 11, n. 1, p. 31–46, 2010.
- MICHENER C.D. **The bees of the world (Internet)**. 2a. The Johns Hopkins University Press, editor. v. 85, Florida Entomologist. 2000. 963 p. Available from: http://books.google.com/books?id=bu_1gmY13FIC
- MICHENER, C.D. **The bees of the World**. Baltimore: The Johns Hopkins University Press. 953 p, 2007.
- MIRANDA, E.A et al. Genetic diversity of *Melipona mandacaia* SMITH 1863 (Hymenoptera, Apidae), na Endemic bee species from Brazilian caatinga, using ISSR. **Hindawi publishing corporation**, p.1-7, 2012.
- MOREY, M. et al. A glimpse into past, present, and future DNA sequencing. **Molecular Genetics and Metabolism**, v. 110, n. 1–2, p. 3–24, 2013.
- NASCIMENTO, M.A. et al. Variation and genetic structure of *Melipona quadrifasciata* Lepeletier (Hymenoptera, Apidae) populations based o ISSR pattern. **Genetics and molecular biology**, v.33, n.2, p.394-397, 2010.
- NAYAK, P.P.; PRAKASH, J. Molecular characterization of the indigenous stingless bees (*Tetragonula* spp. Complex) using ISSR marker from southern peninsular India. **Neotropical Entomology**, p.1-12, 2017, DOI: 10.1007/s13744-017-0538-7.
- NIEDRINGHAUS, T. P. et al. **NIH Public Access**. v. 83, n. 12, p. 4327–4341, 2012.
- NOGUEIRA, J. et al. Conservation study of an endangered stingless bee (*Melipona capixaba*-Hymenoptera: Apidae) with restricted distribution in Brazil. **Journal of Insect Conservation**, v. 18, n. 3, p. 317–326, 2014.
- NOGUEIRA-NETO, P. **Vida e Criação de Abelhas Indígenas Sem Ferrão**. São Paulo: Urna Edição Nogueirapis, v. 34, 1997.
- OLIVEIRA, E. J. et al. Origin, evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, v.29, n.2, p. 294–307, 2006.
- OLIVEIRA, E. J. F. et al. Isolation and characterization of microsatellite markers from the stingless bee *Nannotrigona testaceicornis*. **Conservation Genetics Resources**, p. 1–3, 2009.
- OLIVEIRA, R. C. et al. Genetic divergence in *Tetragonisca angustula* Latreille, 1811 (Hymenoptera, Meliponinae, Trigonini) based on RAPD markers. **Genetics and Molecular**

Biology, v. 27, n. 2, p. 181–186, 2004.

PAREJO, M. et al. Empirical comparison of microsatellite and SNP markers to estimate introgression in *Apis mellifera mellifera*. **Journal of Apicultural Research**, v. 57, n. 4, p. 504–506, 2018.

PARISET, L. et al. Assessment of population structure by single nucleotide polymorphisms (SNPs) in goat breeds. **Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences**, v. 833, n. 1, p. 117–120, 2006.

PAXTON, R. J.; WEIBSCHUH, N.; QUEZADA-EUÁN, J. J. G. Characterization of dinucleotide microsatellites loci from stingless bees. **Molecular Ecology**, v. 8, p. 690–691, 1999.

PETERS, J.M. et al. “Microsatellites loci stingless bees.” **Molecular Ecology**, v.7, p.783–792, 1998.

PETTERSSON, E.; LUNDEBERG, J.; AHMADIAN, A. Generations of sequencing technologies. **Genomics**, v. 93, n. 2, p. 105–111, 2009.

POP, M.; SALZBERG, S. L. Bioinformatics challenges of new sequencing technology. **Trends in Genetics**, v. 24, n. 3, p. 142–149, 2008.

QUELLER, D. C.; STRASSMANN, J. E.; HUGHES, C. R. Microsatellites and kinship. **Trends in Ecology and Evolution**, v. 8, n. 8, p. 285–288, 1993.

RITSCHHEL, P. S. et al. Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). **BMC Plant Biology**, v. 4, p. 9, 2004.

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

SANTIAGO, L. R. et al. Isolation and characterization of ten microsatellite loci in stingless bee *Trigona spinipes* (Apidae: Meliponini). **Entomological Science**, v. 18, n. 1, p. 138–140, 2015.

SANTURE, A. W. et al. On the use of large marker panels to estimate inbreeding and relatedness: Empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. **Molecular Ecology**, v. 19, n. 7, p. 1439–1451, 2010.

SCHINCARIOL, L. P. **IDENTIFICAÇÃO DE MARCADORES SNP PELAS TÉCNICAS DE PCR-RFLP E Tetra-primer ARMS-PCR E SUAS ASSOCIAÇÕES COM TÉCNICAS DE PCR-RFLP E Tetra-primer ARMS-PCR E.** 2011.

SELKOE, K. A.; TOONEN, R. J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. **Ecology Letters**, v. 9, n. 5, p. 615–629, 2006.

SILVA et al. Pesquisa com abelhas-sem-ferrão (Hymenoptera: Meliponini) e aplicabilidade dos marcadores moleculares: Uma revisão sistemática da literatura. **PUBVET**, v.13, n.1, p.1-19, 2019.

SILVA, G. R. et al. New molecular evidence for fragmentation between two distant populations of the threatened stingless bee *Melipona subnitida* Ducke (Hymenoptera, Apidae, Meliponini). **Journal of Hymenoptera Research**, v.38, p.1-9, 2014.

SILVA, G.R. et al. Genome-wide discovery and characterization of microsatellite markers from *Melipona fasciculata* (Hymenoptera: Apidae), cross-amplification and a snapshot assessment of the genetic diversity in two stingless bee populations. **European Journal of Entomology**, v.115, p. 614–619, 2018.

- SMITH, P. J. Random Amplified Polymorphic DNA (RAPD). **Stock Identification Methods**, p. 371–387, 2005.
- SONG, Q. et al. Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. **Crop Science**, v. 50, n. 5, p. 1950–1960, 2010.
- SOUZA, I. G. B. et al. Isolation and characterization of 23 microsatellite loci in the stingless bee *Melipona subnitida* using next-generation sequencing. **Conservation Genetics Resources**, v. 7, n. 1, p. 239–241, 2014.
- STOLLE, E.; MORITZ, R. F. A. RESTseq - Efficient Benchtop Population Genomics with RESTriction Fragment SEQuencing. **PLoS ONE**, v. 8, n. 5, p. 4–8, 2013.
- TAVARES, M. G. et al. Genetic divergence between *Melipona quadrifasciata* Lepelletier (Hymenoptera, Apidae) populations. **Genetics and Molecular Biology**, v. 36, n. 1, p. 111–117, 2013a.
- TAVARES, M. G. et al. Genetic divergence between populations of the stingless bee urucu amarela (*Melipona rufiventris* group, Hymenoptera, Meliponini): is there a new *Melipona* species in the Brazilian state of Minas Gerais? **Genetics and Molecular Biology**, v. 30, n. 3, p. 667–675, 2007.
- TAVARES, M. G. et al. Genetic variability and population structure in *Melipona scutellaris* (Hymenoptera: Apidae) from Bahia, Brazil, based on molecular markers. **Apidologie**, v. 44, p. 720–728, 2013b.
- THEERAAPISAKKUN, M.; KLINBUNGA, S.; SITTIPRANEED, S. Development of a species-diagnostic marker and its application for population genetics studies of the stingless bee *Trigona collina* in Thailand. **Genetics and Molecular Research**, v.9, n.2, p.919-930, 2010.
- THUMMAJITSAKUL, S.; KLINBUNGA, S.; SITTIPRANEED, S. Development of a species-diagnostic marker for identification of the stingless bee *Trigona pagdeni* in Thailand. **Biochemical Genetics**, v. 48, n. 3–4, p. 181–192, 2010.
- THUMMAJITSAKUL, S. et al. Genetic diversity and population structure of *Trigona pagdeni* Schwarz in Thailand. **Apidologie**, v. 39, p.446–455, 2008.
- VAN DIJK, E. L. et al. Ten years of next-generation sequencing technology. **Trends in Genetics**, v. 30, n. 9, p. 418–426, 2014.
- VIANA, J. L. et al. Genetic variability in *Melipona scutellaris* from Reconcavo, Bahia, Brazil. **Genetics and Molecular Research**, v. 12, n. 3, p. 3444–3454, 2013.
- VOS, P. et al. A new technique for DNA fingerprinting. **Nucleic Acids Research**, v. 44, n. 21, p. 388–396, 1995.
- WALDSCHMIDT, A. M. et al. Genetic analysis of *Melipona quadrifasciata* Lep. (Hymenoptera: Apidae, Meliponinae) with RAPD markers. **Brazilian Journal of Biology**, v. 62, n. 4b, p. 923–928, 2002.
- WALDSCHMIDT, A. M.; BARROS, E. G. DE; CAMPOS, L. A. DE O. A molecular marker distinguishes the subspecies *Melipona quadrifasciata quadrifasciata* and *Melipona quadrifasciata anthidioides* (Hymenoptera: Apidae, Meliponinae). **Genetics and Molecular Biology**, v. 23, n. 3, p. 609–611, 2000.
- WANG, J. et al. High-throughput SNP genotyping by single-tube PCR with Tm-shift primers. **BioTechniques**, v. 39, n. 6, p. 885–892, 2005.

- WELSH, J.; MCCLELLAND, M. Fingerprint genomes using PCR with arbitrary primers. **Nucleic Acids Research**, v. 18, n. 24, p. 6531–6535, 1990.
- WILLIAMS, J. G. K. et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research**, v. 18, n. 22, p. 6531–6535, 1990.
- WRATTEN, S. D. et al. Pollinator habitat enhancement: Benefits to other ecosystem services. **Agriculture, Ecosystems and Environment**, v. 159, p. 112–122, 2012.
- YANG, X.; QUIROS, C. Identification and classification of celery cultivars with RAPD markers. **Theoretical and Applied Genetics**, v. 86, n. 2–3, p. 205–212, 1993.
- ZANE, L.; BARGELLONI, L.; PATARNELLO, T. Strategies for microsatellite isolation: A review. **Molecular Ecology**, v. 11, n. 1, p. 1–16, 2002.
- ZAYED, A. Bee genetics and conservation. **Apidologie**, v. 40, p. 237-262, 2009.
- ZIETKIEWICZ, E.; RALFSKI, A.; LABUDA, D. Genome Fingerprinting by simple sequence (SSR) - Anchored Polymerase Chain Reaction Amplification. **Genomics**, v.20, p.176-183, 1994.

CAPÍTULO 2

Avaliação da extração de DNA de diferentes fontes de tecidos para análise de marcadores moleculares baseados em PCR em abelhas-sem-ferrão

Artigo: *“Evaluation of DNA extraction from different tissues sources for PCR-based molecular marker analysis in stingless bees”* (a ser submetido).

2.1 Resumo

No Brasil muitos estudos genéticos têm sido direcionados às abelhas nativas, auxiliando na elaboração de estratégias de manejo e conservação destas espécies. Para a implementação desses estudos é necessário que o material genético seja extraído com quantidade e qualidade suficientes para que os procedimentos posteriores não sejam comprometidos. Com isso, objetivou-se testar a eficiência do processo de extração de DNA de cinco protocolos, no intuito de obter DNA genômico de boa qualidade para os estudos genéticos com abelhas-sem-ferrão. Para isso, o protocolo PCI modificado (A e B) foi comparado com protocolo Chelex 10%, Hotshot (extração alcalina) e um kit comercial, avaliando o desempenho de extração com diferentes partes anatômicas de abelhas [cabeça (1/2), tórax (1/4), abdômen (1/4) e pernas (2)]. A espécie de abelha-sem-ferrão utilizada como modelo biológico foi a *Melipona subnitida*. Para checar a quantidade e a qualidade, o material extraído foi visualizado em gel de agarose (1%) e submetido à espectrofotometria de DNA (NanoDrop 2000). Posteriormente o material extraído foi submetido à reação de digestão empregando enzimas de restrição (*DraI* e *RsaI*) e amplificação por PCR adotando marcadores ISSR (*Inter Simple Sequence Repeat* - Entre Simples Sequências Repetidas) e microssatélites. Dentre os procedimentos testados, os protocolos PCI modificado e o Kit comercial proporcionam DNA genômico, íntegro e de alto peso molecular em todas as partes da abelha, ao contrário do Chelex e Hotshot, o qual apresentou um perfil de degradação em todas as amostras analisadas. Porém, dos três procedimentos, o PCI foi o único que apresentou a razão 260/280 acima de 1,8 para todas as amostras, indicando maior grau de pureza. Adotando o teste de Tukey (programa Past v.3) para a comparação de médias, não houve diferença significativa entre as amostras extraídas das diferentes partes da abelha dentro de cada protocolo, com exceção do Kit comercial, no qual apresentou maior concentração de DNA nas partes referentes à cabeça $23,7 \pm 7,3$ e abdômen $42,0 \pm 17,7$. Com exceção do Chelex e Hotshot todos os protocolos demonstraram eficiência na amplificação de DNA e quanto à digestão enzimática. O protocolo PCI demonstrou ser o mais adequado para a extração de DNA das diferentes secções de abelhas-sem-ferrão, com eficiência tanto para o processo de digestão quanto para a amplificação de DNA via PCR.

Palavras-chave: DNA genômico, Protocolos, *Melipona subnitida*, Partes anatômicas.

2.2 Introdução

As abelhas-sem-ferrão (Apidae: Meliponini) são insetos sociais de grande diversidade, com ampla distribuição geográfica pela faixa Pantropical do mundo. Esses insetos himenópteros apresentam grande importância econômica e ecológica como produtores de mel, geleia real, pólen, própolis e cera (CHUTONG et al., 2016) e como polinizadores chave de plantas nativas e cultivadas (KLEIN et al., 2007). Estima-se que um terço da dieta humana depende direta ou indiretamente da polinização por abelhas (KREMEN & WILLIAMS THORP, 2002).

O extrativismo descontrolado, o desmatamento e o uso de agroquímicos em lavouras comerciais colocaram muitas espécies de abelhas-sem-ferrão em risco de extinção (MOURA, 2003; FREITAS & PINHEIRO, 2010). Portanto, existe uma necessidade urgente de estudos com as espécies desse grupo de abelhas, principalmente aquelas ameaçadas de declínio populacional.

O conhecimento da diversidade, variabilidade genética e estrutura populacional das abelhas têm sido cada vez mais eficazes para ajudar em estratégias que as mantenham em nosso meio ambiente (ZAYED, 2009). Para estes estudos a biologia molecular tem fornecido ferramentas importantes que permitem acesso direto ao DNA. Portanto, um passo crítico para o sucesso de métodos moleculares baseados em DNA é o isolamento e purificação, de forma a obter uma quantidade adequada de ácido nucléico, livre de inibidores que possam prejudicar as reações de PCR que são a base dos procedimentos moleculares.

Vários estudos genéticos com abelhas utilizando técnicas de PCR (Reação em Cadeia da Polimerase) utilizam diferentes métodos de extração de DNA (TAVARES et al., 2007; TAVARES et al., 2013; CARVALHO-ZILSE & KERR, 2006). A maioria dos procedimentos laboratoriais inclui o uso de solventes orgânicos para a purificação de DNA, o que o tornam danosos quanto à toxicidade. No entanto, existem vários kits comercialmente disponíveis e protocolos laboratoriais padrão para extração e purificação de DNA genômico que evitam esses reagentes (CHEN et al., 2008). Embora muitas técnicas tenham sido descritas na literatura, até o momento, poucos estudos compararam a eficácia relativa de vários protocolos relatados de extração de DNA de abelha (CAWTHORN et al., 2011; SCALDAFERRI et al., 2013, VIANA et al., 2015 YALÇINKAYA et al., 2017).

A escolha da técnica de extração de DNA depende não apenas da fonte e finalidade do DNA alvo a ser utilizado em aplicações moleculares (DEMEKE, JENKINS, 2010; RANTAKOKKO-JALAVA; JALAVA, 2002), mas também de certos critérios, tais como rendimento, custo, reprodutibilidade, tempo e segurança.

Neste estudo, cinco protocolos de extração de DNA de diferentes fontes de tecidos, foram investigados para identificar o mais apropriado para abelhas-sem-ferrão, de forma a verificar o mais útil para extração e purificação de DNA em análises de estudos genéticos maximizando a quantidade e qualidade de DNA.

2.3 Material e Métodos

2.3.1 Material biológico

Abelhas operárias adultas foram coletadas aleatoriamente em diferentes colônias na região Nordeste do Brasil, no Estado do Piauí (6 °11'56.2 "S; 42°11'43.8"W), Brasil. Cinco protocolos de extração foram testados para o DNA genômico de *Melipona*. A espécie de abelha-sem-ferrão *Melipona subnitida* foi usada como organismo modelo. Para cada protocolo, segmentos corporais (½ cabeça, ¼ tórax, ¼ abdômen e duas pernas) de amostras frescas foram separados e processados em triplicata para extração de DNA.

2.3.2 Protocolos de extração de DNA

Para cada protocolo de extração, o DNA total foi extraído individualmente das abelhas. O DNA foi ressuspensão em 50 a 100 µL de TE (10 mM Tris, 0,1 mM EDTA.Tris HCl EDTA), exceto para o DNA obtido usando Chelex 100, que foi ressuspensão em 100 µL de ddH₂O. As soluções de DNA foram armazenadas a -20 ° C até nova análise.

2.3.3 Protocolo de extração 1 - Kit Comercial

O DNA foi extraído com o Kit *Extractme Genomic DNA* (EM13-050, DNA-Gdansk), conforme o protocolo do fabricante para o tecido animal. No kit foram usados colunas com membrana de sílica para a captura e posterior purificação de DNA genômico por meio de lavagens utilizando tampões específicos. No final da extração, o *pellet* de DNA foi ressuspensão em 100 µl de tampão de eluição de DNA e deixado durante a noite a 4 °C. As amostras foram armazenadas a -20 ° C. Um microlitro de amostra foi utilizado em PCRs subsequentes.

2.3.4 Protocolo de extração 2 - Chelex 100

Cada amostra de tecido (quantidade de material comparada à metade de um grão de arroz, exceção da perna que foi inteira) foi submetida a 200 µL de Chelex® (Sigma Aldrich) 10% e após a maceração no tubo, a solução foi incubada com 5 µL de proteinase K a 56°C por 30 min, seguida por uma rápida desnaturação de proteínas por meio de aquecimento a 99,9°C (WALSH et al., 1991), programados no termociclador VERITI 384-Well (Applied

Biosystems). A solução foi armazenada a -20°C até ser requerido para os posteriores procedimentos.

2.3.5 Protocolo de extração 3 - HotSHOT

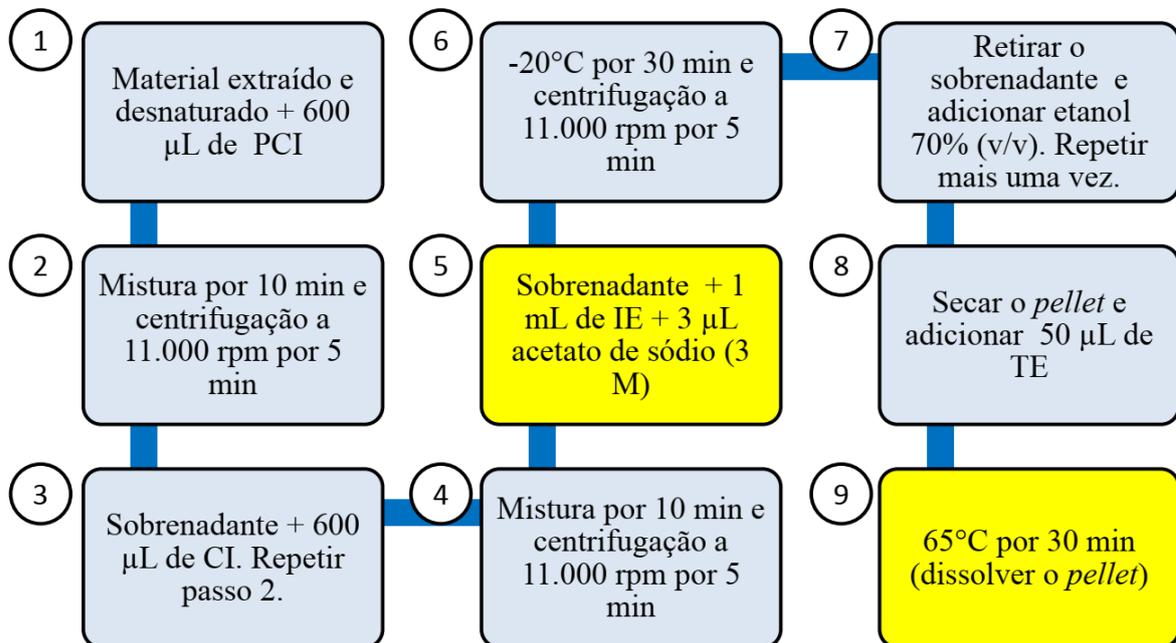
O DNA genômico foi extraído com o protocolo HotSHOT ou extração alcalina, conforme descrito por Truett et al. (2000), diferindo no volume de reagentes utilizados. Com amostras de tecidos biológico (similar ao Chelex), aqueceu-se 50 μL do tampão de lise alcalina (25mM NaOH, 0,2mM Na_2EDTA , pH= 12) a 95°C durante 60 min no termociclador VERITI 384-Well (Applied Biosystems); as amostras foram submetidas a choque térmico e então 50 μL de tampão de neutralização foram adicionados à mistura. Por fim, para separar dos resíduos de tecidos, o extraído foi submetido à centrifugação de 15.000 x g e o sobrenadante transferido para um novo tubo.

2.3.6 Protocolo de extração 4 e 5 - extração utilizando fenol-clorofórmio-álcool isoamílico (A e B)

Para ambos os procedimentos, o material biológico (cabeça, tórax, abdômen e pernas) foi incubado a 65°C por 2 horas com agitação suave a cada 30 minutos em tampão de extração contendo proteinase K para digestão. Em seguida, as amostras foram purificadas usando o protocolo modificado de SAMBROOK et al. (1989) utilizando os solventes orgânicos fenol : clorofórmio : álcool isoamílico (24:1:1) (PCI) (ver Figura 1). Ambos os protocolos usaram Tris-HCl 30 mM; 10Mm EDTA; SDS a 1%, pH 8,0 como tampão de lise celular e uma modificação do protocolo convencional, que foi a substituição do álcool absoluto (para precipitação do DNA) pela combinação etanol : isopraponol (1:1). Essa modificação foi realizada, a fim de garantir maior integridade e rentabilidade de DNA genômico no processo final.

O que diferenciou um protocolo do outro foi a adição de NaCl (273 mM) utilizada no tampão de lise do PCI-A, a fim de possibilitar maior atividade da enzima proteinase K, o que garante maior rendimento de DNA na etapa final.

Figura 1. Purificação do DNA genômico, modificado de Sambrook et al. (1989). Modificações estão nos quadrados em amarelo. PCI – Fenol : Clorofórmio : Álcool isoamílico; CI – Clorofórmio : Álcool isoamílico; IE – Isopropanol : Etanol absoluto; TE – Tris e EDTA.



Fonte: Elaboração Própria.

2.3.7 Avaliação quantitativa e qualitativa do DNA extraído

O DNA extraído de todos os diferentes protocolos foi quantificado usando um espectrofotômetro NanoDrop™ 2000 (Thermo Fisher Scientific Inc., Delaware, EUA). A concentração de DNA foi estimada em função da absorbância de 260 nm. A avaliação da pureza dos ácidos nucleicos foi medida por meio da relação A260 / A280. A relação A260 / A280 indica a presença de contaminantes orgânicos, como proteína ou fenol (proporção ideal está entre 1,8 e 2,2).

2.3.8 Digestão de restrição, validação de PCR e eletroforese

O DNA genômico (~ 200 ng / µL) foi digerido separadamente com *RsaI* (0,4 U / µL) e *DraI* (0,4 U / µL) (*New England Biolabs - NEB, Beverly, MA*) na presença de albumina bovina sérica (BSA; µg / µL) a 37°C por 2 horas e 70°C por 15 min, programado num termociclador *Techné TC-412* versão 34.11 e analisados em géis de agarose a 1,5 %.

As extrações de DNA foram analisadas com base em reações de PCR usando *primer* ISSR (Entre simples sequências repetidas – *Inter Simple Sequence Repeat*) UBC-845 desenvolvido pela *University of British Columbia*. As reações de PCR foram realizadas em volumes de 10 µL contendo aproximadamente 10-50 ng de DNA, 1 × tampão PCR (40 mM Tris-HCl; 100 mM KCl), 0,5 µM de iniciador, 2,5 mM MgCl₂, 0,2 mM de cada dNTP, 0,5 U

de DNA polimerase Taq da Invitrogen. Todas as amplificações foram realizadas em um termociclador VERITI 96-*well* (Applied Biosystem). Foram utilizadas as seguintes condições de PCR: uma desnaturação inicial a 94 ° C durante 1,5 min, seguida de 45 ciclos de 94°C / 40 seg, 36°C / 1 min. e 72°C / 2 min e uma extensão final de 72°C / 5 min. Os marcadores ISSR foram rastreados por meio do gel de agarose a 1,5%.

2.3.9 Análise estatística

As médias e desvios padrão dos rendimentos de DNA e a razão A260 / A280 foram obtidos para todos os cinco protocolos de extração de diferentes fontes de tecidos. A ANOVA foi aplicada a essas medidas usando o *Kruskal-Wallis rank sum test*. As comparações pareadas foram realizadas usando o teste de comparação múltipla de Kruskal-Wallis de Dunn (1964), a um nível de significância de 0,05, ajustados com o método Holm usando a plataforma estatística R. A estatística básica (média, desvio padrão e coeficiente de variação) foi calculada usando o pacote PAST v.3 (HAMMER et al., 2001). Diagramas de caixa foram utilizados para representar graficamente os parâmetros estatísticos para a comparação da reprodutibilidade e eficácia da extração de DNA.

2.4 Resultados e Discussão

2.4.1 Rendimento do DNA e taxa de absorvância

Diferentes tecidos de abelha (cabeça, tórax, abdômen e pernas) foram usados como materiais para a extração de DNA para determinar quais produziram maior concentração e pureza. Concentrações de DNA e razões de absorvância de acordo com a fonte do tecido para todos os cinco protocolos de extração de DNA estão listadas na Tabela 1 e 2, respectivamente.

As concentrações de DNA obtidas pelo uso da resina Chelex nas extrações de DNA foram altas. Os resultados indicaram que qualquer material biológico oriundo de abelhas-sem-ferrão poderia ser usado para extrair DNA por meio desse procedimento. As concentrações variaram de $327,67 \pm 73,74$ ng / μ L (pernas) a $599,47 \pm 15,11$ ng / μ L (cabeça), no entanto, os valores não foram significativamente diferentes (pvalor > 0,05).

Os protocolos HotShot e PCI-A produziram as segundas maiores concentrações para cabeça, tórax e abdômen, com média de 331,51 ng / μ L (HotShot) e 304,43 ng / μ L (PCI-A). Estatisticamente, o kit comercial produziu as menores concentrações de DNA, variando de 4,83 a 42,03 ng / μ L

Tabela 1. Concentrações de DNA [média em ng / $\mu\text{L} \pm$ desvio padrão (coeficiente de variação em %)] de abelhas-sem-ferrão *Melipona subnitida* usando diferentes métodos de extração de DNA e fontes de tecidos. A seta vermelha indica o desempenho dos diferentes protocolos em uma parte anatômica da abelha analisada. A seta azul compara o desempenho de um determinado protocolo referente as diferentes partes anatômicas da abelha.




Amostra de tecido	Protocolo de extração de DNA				
	Chelex	Hotshot	PCI-A	PCI-B	Kit
Cabeça	599,47 \pm 15,11 (2,52) ^{aA}	326,40 \pm 57,76 (17,69) ^{aAB}	254,73 \pm 5,78 (2,27) ^{abAB}	130,57 \pm 81,59 (62,448) ^{aAB}	23,7 \pm 7,39 (31,18) ^{abB}
Tórax	418,20 \pm 206,02 (49,26) ^{aAB}	373,53 \pm 85,50 (22,88) ^{aA}	167,57 \pm 147,34 (87,93) ^{abAB}	136,93 \pm 88,51 (64,64) ^{aAB}	9,47 \pm 2,14 (22,56) ^{abB}
Abdômen	524,00 \pm 112,26 (21,42) ^{aA}	294,60 \pm 86,79 (29,46) ^{aAB}	491,00 \pm 272,48 (55,49) ^{bAB}	206,77 \pm 163,83 (79,23) ^{aAB}	42,03 \pm 17,71 (42,13) ^{bB}
Pernas	327,67 \pm 73,74 (22,50) ^{aA}	188,67 \pm 65,32 (34,62) ^{aAB}	44,63 \pm 11,94 (26,74) ^{aAB}	28,93 \pm 20,51 (70,89) ^{aAB}	4,83 \pm 1,36 (28,17) ^{aB}

Valores em colunas com letras diferentes (a, b, c) e em linhas (A, B, C) são significativamente diferentes no nível de 5%.

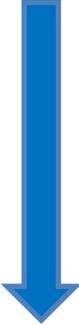
As razões de absorbância média variaram muito, dependendo do tipo de amostra de tecido e do protocolo usado para extração de DNA (Tabela 2). A relação média A260 / 280 dos extratos de DNA obtidos pelos protocolos fenol / clorofórmio (PCI-A e PCI-B) situou-se no intervalo 1,8-2,0, indicando a menor contaminação proteica entre os cinco métodos (GLASEL, 1995; VESTY et al. 2017). O tipo de tecido utilizado não teve efeito significativo na razão de absorbância para o protocolo PCI-B (pvalor > 0,05).

As amostras extraídas pelos protocolos Chelex e HotShot apresentaram menores valores da relação A260 / A280 quando comparado aos protocolos fenol / clorofórmio, o que pode indicar a presença de resquícios de proteína ou outros contaminantes que absorvem fortemente comprimento de ondas próximo de 280 nanômetro (nm). Esses métodos não adotam etapas de purificação, sendo indicados quando se utiliza pequena quantidade de material biológico, sendo geralmente aplicados em práticas forenses, pois a quantidade mínima de resíduos não compromete a PCR (GALLAGHER, 1989).

Os mesmos índices baixos da razão de 260/280 nm foram encontrados para a extração de DNA realizada pelo kit comercial, que pode ter sido causada por uma concentração muito baixa de ácido nucléico (MATLOCK, 2015). Porém, não se pode descartar a presença dos contaminantes (polissacarídeos e resíduos de proteínas), uma vez que visivelmente observou-se uma provável interferência na reação de amplificação, pois comparado aos outros

protocolos, o rendimento de fragmentos gerados pelo mesmo *primer* de ISSR foi muito inferior. Provavelmente, há a necessidade de um ajuste nas etapas desse protocolo, uma vez que os kits comerciais primam pela sua eficiência.

Tabela 2. Proporção A260 / 280 [média \pm desvio padrão (coeficiente de variação em %)] de abelhas-sem-ferrão *Melipona subnitida* usando diferentes métodos de extração de DNA e fontes de tecidos. A seta vermelha indica o desempenho dos diferentes protocolos em uma parte anatômica da abelha analisada. A seta azul compara o desempenho de um determinado protocolo referente entre as diferentes partes anatômicas da abelha.

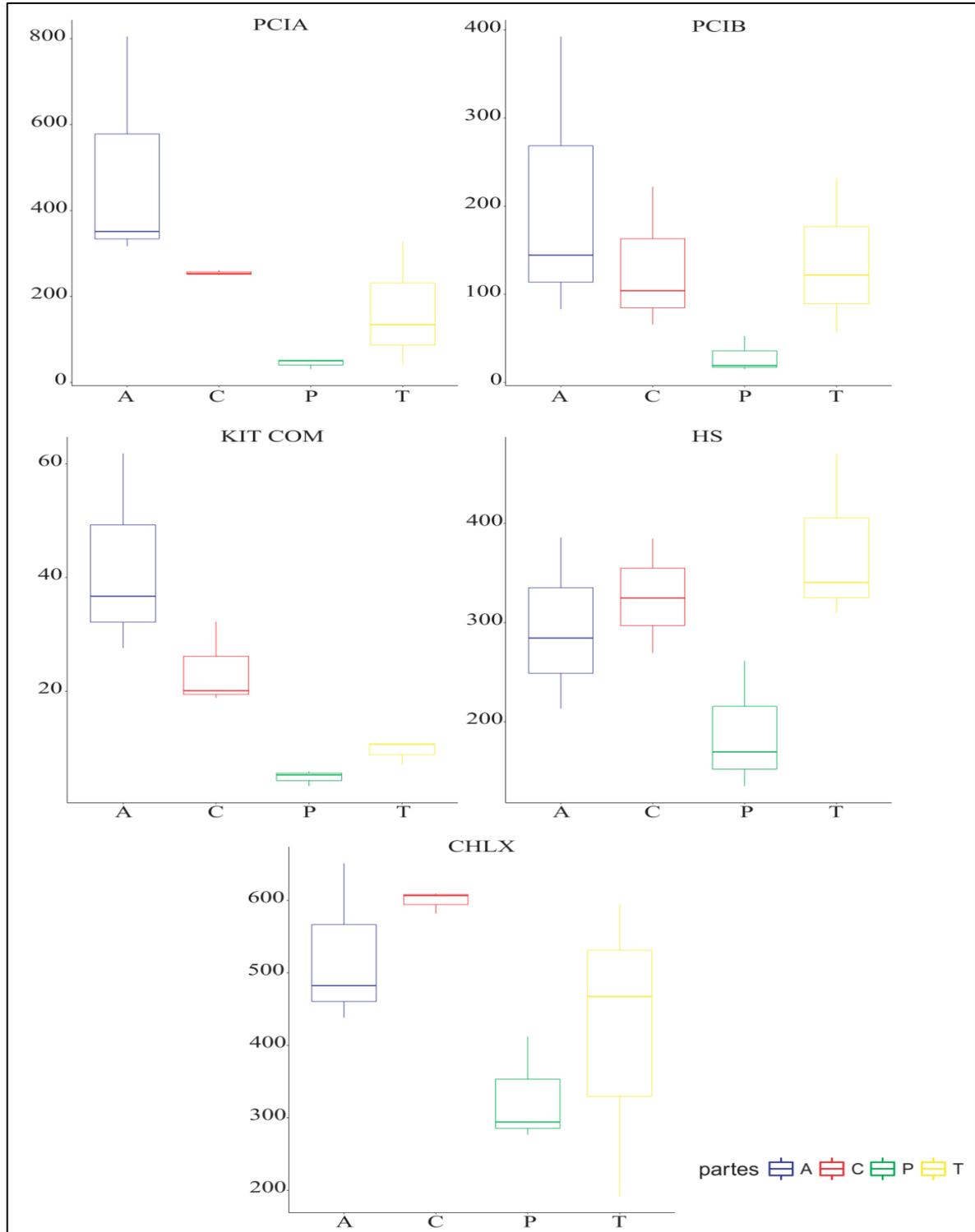



Amostra de tecido	Protocolo de extração de DNA				
	Chelex	Hotshot	PCI-A	PCI-B	Kit
Cabeça	1,29 \pm 0,05 (3,59) ^{aA}	1,55 \pm 0,04 (2,33) ^{aAB}	1,94 \pm 0,04 (1,86) ^{aB}	1,84 \pm 0,09 (4,89) ^{aAB}	1,34 \pm 0,23 (17,00) ^{abAB}
Tórax	0,98 \pm 0,06 (6,12) ^{bA}	1,45 \pm 0,15 (10,62) ^{aA}	1,91 \pm 0,14 (7,12) ^{bA}	1,83 \pm 0,14 (7,83) ^{aA}	1,06 \pm 0,32 (30,39) ^{abA}
Abdômen	1,15 \pm 0,14 (11,76) ^{abA}	1,34 \pm 0,01 (0,86) ^{aAC}	1,95 \pm 0,13 (6,84) ^{abBC}	1,93 \pm 0,09 (4,40) ^{aC}	1,69 \pm 0,14 (8,05) ^{aABC}
Perna	1,11 \pm 0,06 (5,41) ^{abAB}	1,49 \pm 0,04 (2,70) ^{aAB}	1,87 \pm 0,13 (6,72) ^{abA}	1,8 \pm 0,25 (14,80) ^{aAB}	0,89 \pm 0,11 (12,07) ^{bB}

Valores em colunas com letras diferentes (a, b, c) e em linhas (A, B, C) são significativamente diferentes no nível de 5%.

As quantidades de DNA obtidas pelos cinco protocolos em cada fonte de tecido, também são mostradas para comparação na Figura 2. O protocolo PCI-A mostra a menor dispersão para a cabeça e patas, considerando todos os protocolos e fontes de tecido, o que é uma evidência adicional para se considerar entre os rendimentos de DNA e, conseqüentemente, uma maior eficiência de extração. A maior dispersão foi obtida da extração usando Chelex, para o tórax, e PCI-A e PCI-B para o abdômen.

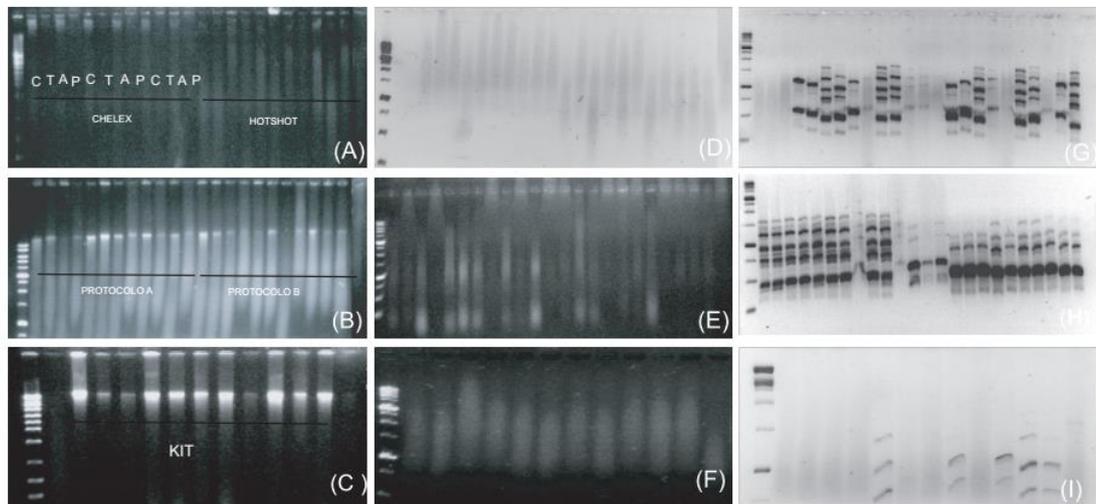
Figura 2. Box plot (valores máximos, mínimos e médios) das concentrações de DNA (ng / μ l) obtidas de chelex (CHLX), HotShot (HS) fenol-clorofórmio (PCI-A e PCI-B) e Kit Comercial (KIT COM) protocolos de extração em diferentes tecidos (A, abdome; C, cabeça; P, perna; T, Tórax) de *Melipona subnitida*. Para cada gráfico de caixa, a linha dentro da caixa representa a mediana; as linhas superior e inferior representam 75 e 25 percentis dos dados, respectivamente; os traços superiores e inferiores representam 95 e 5 por cento, respectivamente.



2.4.2 Análise de eletroforese de DNA extraído

Exemplos do perfil geral da extração de DNA, digestão por restrição e amplificação por PCR apresentada usando os cinco métodos visualizados em gel de agarose, ilustrados na Figura 2. A integridade do DNA obtido por meio dos protocolos de extração indica que PCI-A, PCI-B e kit apresentaram qualidade superior (Figura 3A, 3B e 3C). O DNA genômico extraído foi completamente digerido com as enzimas de restrição e não houve evidência da presença de impurezas que impedissem a ação das enzimas em nenhuma extração, mesmo naquelas que não foram submetidas a purificação (Chelex e HotShot), portanto, adequada para manipulações enzimáticas. Com exceção do DNA obtido pelo kit, que apresentou problemas na reação, a amplificação por PCR utilizando os outros protocolos, também indica que o DNA era de boa qualidade, livre de compostos interferentes, e que seria adequado para outras análises de DNA, como *Southern blotting* e sequenciamento.

Figura 3. Géis de Agarose com os diferentes tipos de tecido de *Melipona subnitida* (A, abdome; C, cabeça; P, perna; T, Tórax) e protocolos (chelex (CHLX), HotShot (HS) fenol-clorofórmio (PCI-A e PCI-B) e Kit Comercial (Kit com)). Integridade do DNA em gel de eletroforese (A-B-C); digestão utilizando enzimas de restrição, DraI e RsaI, (D-E-F); e amplificação por PCR utilizando marcadores ISSR (Inter Simple Sequence Repeat) (G-H-I).



2.5 Conclusão

Para a extração das diferentes partes anatômicas das abelhas-sem-ferrão, o protocolo PCI-A demonstrou ser o mais apropriado para os diversos procedimentos moleculares que tem como base o uso das enzimas de restrição e PCR. Porém, para técnicas menos exigentes que englobam os marcadores moleculares, com base nos resultados, não se pode descartar o uso dos procedimentos que utilizam o Chelex e HotShot, pois além de terem demonstrado a mesma eficiência do PCI-A, consomem menos tempo e baixo custo.

2.6 Referências

- CARVALHO-ZILSE, G.A.; KERR, W.E. Utilização de marcadores microssatélites para estudos populacionais em *Melipona scutellaris* (Apidae, Meliponini). **Magistra**, v.18, p.213-220, 2006.
- CAWTHORN, D.-M., STEINMAN, H. A., & WITTHUHN, R. C. Comparative study of different methods for the extraction of DNA from fish species commercially available in South Africa. **Food Control**, v.22 n.2, p. 231–244, 2011.
- CHEN, M. et al. Methodological comparison of DNA extraction from *Holcocerrus hippophaecolus* (Lepdoptera: Cossidae) for AFLP analysis. **Forestry Studies in China**, Beijing, v. 10, n. 3, p. 189-192, 2008.
- CHUTTONG, B. et al. Effects of long term storage on stingless bee (Hymenoptera: Apidae: Meliponini) honey. **Journal of Apicultural Research**, v.54, n.5, p.441-451, 2016.
- DUNN, O.J. Multiple comparisons using rank sums. **Technometrics**, v.6, p. 241-252, 1964.
- DEMEKE, T.; JENKINS, R. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. **Analytical and Bioanalytical Chemistry**, Heidelberg, n. 396, p. 1977-1990, 2010.
- FREITAS, B. M.; PINHEIRO, J. N. Efeitos sub-letais dos pesticidas agrícolas e seus impactos no manejo de polinizadores dos agroecossistemas brasileiros. **Oecologia Australis**, v. 14, p. 282-298, 2010.
- GALLAGHER, S.R. 1989. **Quantitation of DNA and RNA with absorption and fluorescence spectroscopy**, p. A.3.9-A.3.15. In F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl, Current Protocols in Molecular Biology. John Wiley & Sons, New York.
- GLASEL, J.A. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. **BioTechniques**, v.18, p.62-63, 1995.
- HAMMER Ø, HARPER DAT, RYAN PD. PAST: Paleontological statistics software package for education and data analysis. **Palaeontologia Electronica**, v.4, p. 1–9, 2001.
- KLEIN, A.M. et al. Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society of London B: **Biological Sciences**, v. 274, n. 1608, p. 303-313, 2007.
- KREMEN, C.; WILLIAMS, N. M.; THORP R.W. Crop pollination from native bees at risk from agricultural intensification, v.99, n. 26, **PNAS**, p. 6812–16816, 2002.
- MATLOCK, B. Assessment of Nucleic Acid Purity. **Thermo Fisher Scientific**, p.1-3, 2015.
- MOURA, D. C. **Riqueza e Abundância de Abelhas em Diferentes Estágios de Degradação da Caatinga como Indicadores Ambientais no Entorno da Usina Hidrelétrica de Xingó**, dissertação de Mestrado apresentada ao Programa em Gestão e Políticas Ambientais/UFPE, Recife, 2003. 92p.

RANTAKOKKO-JAVALA, K., JAVALA, J. Optimal DNA isolation method for detection of bacteria in clinical specimens by broad-range PCR. **Journal of Clinical Microbiology**. v.40, p.4211- 4217, 2002.

SAMBROOK; FRITSCH; MANIATIS. Molecular Cloning: A Laboratory Manual 2 ed, vol. 3, p. E3 - E4; **Cold Spring Harbor Laboratory Press** 1989.

SCALDAFERRI, M.M.; FREITAS, J.S. SANTOS, E.S.L.; VIEIRA, J.G.P. et al. Comparison of protocols for genomic DNA extraction from ‘velame pimenta’ (*Croton linearifolius*), a species native to the Caatinga, Brazil. **African Journal of Biotechnology**, v.12, p. 4761-4766, 2013.

SUDHAGAR S. et al. Qualitative and quantitative differences in DNA extracted from different body parts of *Apis* spp. (Hymenoptera:Apidae) and its validation using microsatellite markers. **Pest Management in Horticultural Ecosystems**, v.20, n.1, p.55-58, 2014.

TAVARES, M.G. et al. Genetic divergence between *Melipona quadrifasciata* Lepelletier (Hymenoptera, apidae) populations. **Genetics and Molecular Biology**, v.36, p.111-117, 2013.

TAVARES M.G. et al. Genetic divergence between populations of the stingless bee uruqu amarela (*Melipona rufiventris* group, Hymenoptera, Meliponini): Is there a new *Melipona* species in the Brazilian state of Minas Gerais? **Genetics and Molecular Biology**, v.30, p.667–675, 2007.

TRUETT, G. E. et al. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and Tris (HotSHOT). **BioTechniques**, v.29, p.52-54, 2000.

VESTY A. et al. Evaluating the impact of DNA extraction method on the representation of human oral bacterial and fungal communities. **PLoS ONE** v.12, n.1. 2017.

VIANA, J.P.G. et al. Comparison of eight methods of genomic DNA extraction from babassu. **Genetics and Molecular Research**, v. 14, n. 4, p. 18003-18008, 2015.

WALSH, P. S., METZGER, D. A., HIGUCHI, R. Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. **BioTechniques**, v.10, p.506-513, 1991.

YALÇINKAYA, B. et al. Comparison of DNA extraction methods for meat analysis. **Food Chemistry**, v. 221, p. 1253–1257, 2017.

ZAYED, A. Bee genetics and conservation. **Apidologie**, v. 40, p. 237-262, 2009.

CAPÍTULO 3

Metodologia para identificação de elementos repetitivos em tandem no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de nova geração em baixa cobertura

Artigo: Metodologia para identificação de elementos repetitivos em tandem no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de nova geração em baixa cobertura, submetido a série documentos da EMBRAPA MEIO-NORTE, TERESINA-PI

3.1 Introdução

Sequências simples repetitivas em tandem, ou microssatélites de DNA, são compostas de dois a seis nucleotídeos [e.g. (CA)_n, (TGC)_n, (GATA)_n, etc.], distribuídas aleatoriamente pelo genoma dos eucariotos. Estas regiões repetitivas são hipervariáveis, e desta forma, altamente polimórficas, sendo por isso frequentemente utilizadas como marcadores moleculares em estudos genéticos. Estas características atribuem aos microssatélites um papel fundamental como ferramenta para análises de variabilidade e diversidade genéticas, sendo este conhecimento base para elaboração de programas de conservação e melhoramento de espécies de interesse comercial (AVISE, 1994; DINIZ et al., 2007; SOUZA et al., 2008; ZHAO et al., 2010).

Dentre as inúmeras metodologias possíveis para o desenvolvimento de marcadores microssatélites o sequenciamento de nova geração (NGS) desponta como um método bastante promissor por abranger maior cobertura do genoma e assim facilitar o processo de varredura de genomas desconhecidos, o que garante a identificação de maior número de sequências repetitivas em tandem que possam ser utilizadas como marcadores moleculares. A plataforma de sequenciamento da Illumina é amplamente considerada como a tecnologia mais precisa, mesmo que não tenha sido ainda possível produzir “reads” mais longos quando comparados à outras plataformas de nova geração. Dentre as várias opções de tecnologia NGS disponíveis, plataformas mais compactas, do tipo “de bancada”, são consideradas mais adequadas para mineração de sequências simples repetitivas em tandem, tendo em vista a relação custo/benefício, destacando-se a plataforma Illumina MiSeq.

Neste contexto, uma metodologia foi desenvolvida para identificação de elementos repetitivos em tandem no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de nova geração em baixa cobertura, a partir do protocolo modificado de sequenciamento recomendado pelo fabricante Illumina, tendo por objetivo o desenvolvimento de marcadores moleculares, utilizando como modelo as abelhas Jandaíra (*Melipona subnitida* Ducke, 1910) e Tiúba (*Melipona fasciculata* Smith, 1854).

3.2 Metodologia

3.2.1 Extração de DNA

O DNA genômico é extraído do tórax, aproximadamente 20 mg, de uma abelha adulta. O kit de extração Extractme® Genomic DNA (BLIRT DNA-Gdansk) pode ser utilizado, e as recomendações do fabricante seguidas. A integridade do DNA das amostras pode ser avaliada em gel de agarose a 1,0 % com TBE (Tris-Borato-EDTA) corado com GelRed™, e o DNA quantificado no espectrofotômetro NanoDrop 1000® (Thermo Fisher Scientific, Wilmington, EUA) na absorbância 260/280.

3.2.2 Preparo da amostra para o sequenciamento

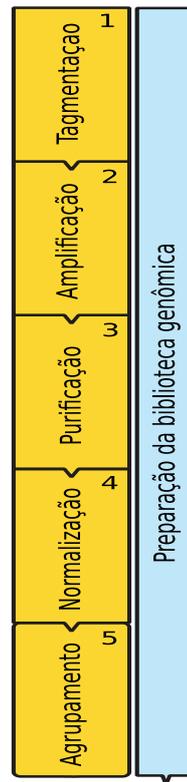
A etapa inicial acontece com uma nova quantificação do material extraído no fluorímetro Qubit 2.0 (Thermo Fisher Scientific, Waltham, EUA) e posterior diluição para 0,2 ng/μL. Para o preparo da biblioteca utiliza-se o kit Nextera® XT DNA Sample Preparation (Illumina, San Diego, USA), do tipo paired-end, que proporciona o sequenciamento nas duas extremidades do read com tamanho entre 200 e 500 nucleotídeos, e Nextera® XT Index (Illumina, San Diego, USA).

3.2.3 Preparo da biblioteca genômica

3.2.3.1 Tagmentação do DNA genômico

O preparo da biblioteca (Figura 1) se inicia com a uma reação de *tagmentação* do DNA genômico, etapa onde a fragmentação (*tagmentação*) randômica do DNA a ser sequenciado acontece com enzimas transposases, a marcação com adaptadores no final das sequências (e.g. Kit Nextera XT transposoma), a colocação de um par de índices, e posteriormente, usado para amplificação pela reação em cadeia da polimerase (PCR). O DNA *tagmentado* é purificado com o kit AMPure *beads* (em 0.5× do volume da biblioteca) para retirada das enzimas da *tagmentação*.

Figura 1. Fluxograma com as etapas para o preparo da biblioteca genômica.



Fonte: Elaboração própria.

3.2.3.2 Amplificação via PCR e adição de adaptadores

Nessa etapa, o DNA tagmentado é amplificado via PCR [i.e. 72°C por 3 min, 95°C por 30 sec, e 12 ciclos de (95°C por 10 sec, 55°C por 30 sec, 72°C por 30 sec), e 72°C por 10 min], resultando em milhares de *clusters*, cada um representando um fragmento a ser sequenciado (Fedurco et al., 2006; Turcatti et al., 2008). O passo da PCR também adiciona adaptadores que apresentam uma sequência específica na sua região terminal. O processo de agrupamento ou clusterização dependerá desses adaptadores (P5 e P7), que adicionalmente possuem uma sequência variável de nucleotídeos (índices i5 e i7), necessários à separação e identificação futura dos fragmentos no sequenciamento da amostra, que são misturadas em uma mesma reação. Ao finalizar a PCR, tem-se a etapa de purificação.

3.2.3.3 Purificação das bibliotecas (PCR Clean-up)

Essa etapa usa AMPure XP *beads* (Beckman Coulter, Inc., USA), que promove uma seleção por tamanho, removendo fragmentos pequenos da biblioteca. As microesferas magnéticas (*beads*) se ligam ao DNA, separando-o do restante da solução.

3.2.3.4 Normalização da biblioteca e agrupamento das amostras

Com a biblioteca finalizada se inicia o procedimento de normalização da biblioteca para e o agrupamento equimolar de todas as amostras do estudo, e assim seguir para sequenciamento no MiSeq Illumina. Esse processo normaliza a quantidade de cada biblioteca para garantir uma representação da biblioteca mais igual na biblioteca agrupada (*pooled library*). A concentração das bibliotecas é normalizada pelo tamanho médio dos fragmentos obtidos pela eletroforese e pela taxa da diluição, chegando-se a uma concentração final.

3.2.3.5 Agrupamento de bibliotecas

O agrupamento de bibliotecas combina volumes iguais de bibliotecas normalizadas em um único tubo. Após a reunião, diluir e desnaturar (calor) o conjunto de bibliotecas antes de carregar as bibliotecas para a execução de sequenciamento.

3.2.4 Sequenciamento do DNA (Illumina Nextera XT_V3)

O sequenciamento é realizado na plataforma MiSeq (Illumina, San Diego, USA). Para que isso aconteça, é necessário um kit de reagentes descartável (i.e. MiSeq Reagent v2, 500 cycle) que inclui a lâmina de fluxo, o frasco PR2 e o cartucho de reagentes. Depois de carregar o cartucho e colocar os reagentes no equipamento se inicia o procedimento de obtenção de sequências fazendo a identificação das amostras no sistema para emitir as sequências desejadas (*reads*). Para obter uma descrição dos kits de reagentes disponíveis, acesse o site da Illumina support.illumina.com/sequencing/sequencing_kits/miseq_reagent_kit.html.

3.2.5 Análise bioinformática

A análise bioinformática envolve as etapas de controle da qualidade dos dados obtidos pelo sequenciamento por meio de software especializado, a montagem das sequências (*reads*), e a identificação das regiões de repetição em *tandem* que poderão ser usadas como marcadores moleculares, posteriormente.

Inicialmente, é necessário avaliar a qualidade das sequências brutas ou leituras (*reads*) por meio de programas específicos [e.g. FastQC (ANDREWS, 2010); Rqc (SOUZA & CARVALHO, 2014)]; essa etapa é também denominada de ´trimagem´. Com o controle de qualidade das sequências é possível detectar problemas ou erros sistemáticos, caso existam, que potencialmente afetem a reação de sequenciamento, e que desta forma, podem introduzir

vieses nas etapas de “chamada das bases” (*base calling*) pelo sequenciador. Estes programas fornecem informações sobre estatísticas básicas das sequências; qualidade; conteúdo de GC; distribuição do tamanho das sequências; níveis de duplicação das sequências; e sequências sobre-representadas, podendo indicar se a biblioteca genômica é enviesada ou contaminada.

A qualidade das sequências pode ser avaliada por métricas como a qualidade Phred (EWING ET AL. 1998; EWING & GREEN, 1998) que calcula índices de qualidade relacionados à probabilidade de erros, de que uma determinada base seja nomeada incorretamente pelo sequenciador de nova geração, aumentando assim a acurácia do processo de montagem. Os valores de qualidade Phred (Tabela 1) são calculados de acordo com a seguinte fórmula: $Q = -10 \log_{10}(P)$, onde P é a probabilidade de erro.

Tabela 1. Qualidade de Phred relacionada com precisão do sequenciamento.

Valor de qualidade PHRED	Probabilidade de erro (“base calling error”)	Precisão “do base calling”
10	1 em 10	90%
20	1 em 100	99%
30	1 em 1.000	99.9%
40	1 em 10.000	99.99%
50	1 em 100.000	99.999%

Fonte: Elaboração própria.

A montagem do genoma é também uma etapa obrigatória após o sequenciamento de DNA na tecnologia NGS, onde leituras (*reads*) longas e curtas são reunidas, ou emparelhadas pelas extremidades, para constituir sequências contíguas (*contigs*), ou seja, uma cadeia de superposições de sequências de DNA que reconstruem a sequência original de DNA, que por sua vez são reunidas em trechos ainda mais longos do genoma (*scaffolds*), até a montagem de um genoma completo. Desta forma, as ferramentas de software de montagem de *contigs* ou do genoma completo detectam sobreposições entre os *reads*, montam os *contigs* com base em sobreposições, em seguida, combinam *contigs* em *scaffolds* para obter uma sequência de esboço do genoma completo. Quando o objetivo é exclusivamente a identificação de elementos repetitivos em *tandem* no genoma, faz-se apenas necessário o montagem dos *reads* em *contigs*. Com estas sequências de DNA mais longas que os dados das sequências brutas (*reads*), os *contigs*, é possível identificar estes elementos repetitivos por meio de uma mineração dos *contigs*, exclusivamente, por meio de softwares desenhados para este objetivo. No caso do sequenciamento do DNA na plataforma MiSeq da Illumina, e em outras

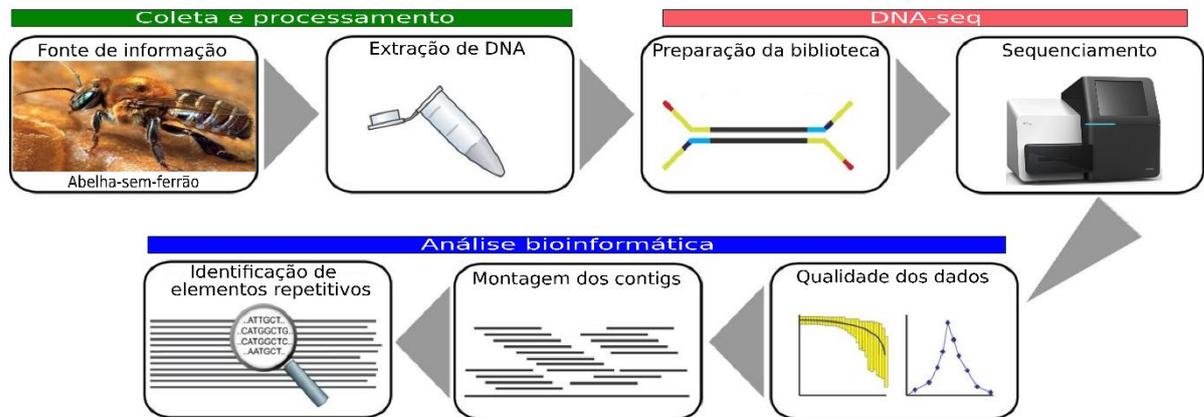
plataformas, os *contigs* criados a partir dos dados das sequências são gerados no formato FASTA por meio do programa CLC Genomics Workbench 7.0.4 (Qiagen) ou similares [e.g. Trinity (GRABHERR et al., 2011); ABySS v1.5.2 (SIMPSON et al. 2009); PEAR (ZHANG et al. 2014), etc.].

Tendo sido gerados *contigs* a partir das sequências brutas oriundas do sequenciamento de nova geração, procede-se à identificação de elementos repetitivos em *tandem* no genoma investigado por meio de programa como: MSATCOMMANDER (FAIRCLOTH, 2008), MSDB (*Microsatellite Search Building Database*) (DU et al., 2013), MISA (THIEL et al., 2003); SSRHunter (LI & WAN, 2005); Phobos (MAYER, 2006); SSRIT (*Simple Sequence Repeat Identification Tool program*) (TEMNYKH et al. 2001). Estes elementos repetitivos em *tandem* poderão ser posteriormente utilizados como marcadores moleculares, desde que tenham primers de PCR desenhados para cada *locus*, utilizando o software PRIMER3 (ROZEN & SKALETSKY, 2000) ou algum similar (e.g. TROLL, CASTELO et al. 2002), e validados através da genotipagem de indivíduos oriundos da população de interesse.

3.3 Resultados

Uma metodologia para a identificação de elementos repetitivos, descrita em detalhes (Figura 2), é aqui apresentada para abelhas-sem-ferrão. Os resultados demonstram a eficiência da estratégia metodológica adotada na identificação destes elementos presentes no genoma de eucariotos, não somente pelo volume de dados gerados, mas também no número de elementos repetitivos encontrados nestas espécies de abelha, Jandaíra e Tiúba (Tabela 2). A identificação destas sequências de gDNA poderão ser utilizados no desenvolvimento de marcadores moleculares para as espécies utilizadas como modelo, ou até mesmo para espécies no mesmo gênero por meio da amplificação interespecífica via uma reação de PCR.

Figura 2. As etapas para identificação de elementos repetitivos em *tandem* no genoma de abelhas-sem-ferrão pelo uso do sequenciamento de nova geração (NGS).



Fonte: Elaboração própria.

Um total de 2.669.884 e 2.533.648 *reads* ou sequências foram geradas a partir da amostras de *Melipona fasciculata* e *M. Subnitida*, respectivamente. As sequências brutas foram montadas (*assembled*) em *contigs*. Esse processo elimina as sequências repetitivas e gera *reads* mais longos, o que pode aumentar a probabilidade de detectar sequências repetitivas, e *primers* adequados para a reação de PCR, dentro de uma sequência (PERRY, 2011). Os *reads* de alta qualidade filtrados após análise no FastQC (Figura 3), foram montados em 141.412 (*Melipona subnitida*) e 47.081 (*M. fasciculata*) *contigs* com um comprimento médio de 477 e 453 bases, respectivamente. Adicionalmente, *contigs* mais longos aumentam a probabilidade de detecção de loci com um número maior de repetições, e possivelmente mais polimórficas.

Figura 3. As métricas de controle de qualidade de seqüências de DNA gerado pelo FastQC para *Melipona fasciculata* e *M. subnitida* com a distribuição da pontuação de qualidade de Phred sobre todas as leituras em cada base.



Fonte: Elaboração própria.

O número de *reads* e também de *contigs* pode variar significativamente entre espécies do mesmo gênero (Tabela 2; Figura 4) ou de gêneros distantes. Usando a tecnologia Illumina de sequenciamento de duas espécies de insetos, porém com uma abordagem metodológica diferente da empregada neste trabalho, Andersen e Mills (2014) obtiveram 65.535 *contigs*

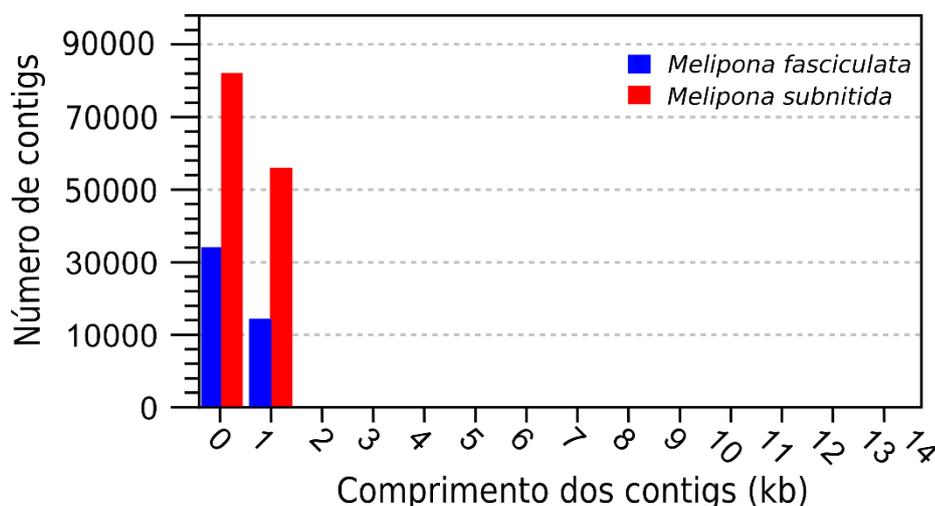
para a espécie *Trioxys pallidus* (uma vespa) e 474.388 contigs para a espécie *Chromaphis juglandicola* (um pulgão). Portanto, o sequenciamento em baixa cobertura é suficientemente eficaz para desenvolver um conjunto mais abrangente de marcadores moleculares, tendo em vista o maior número de elementos repetitivos detectados no genoma (FARRER et al., 2009).

Tabela 2. Dados básicos gerados a partir da análise bioinformática.

Espécie	Tamanho do genoma (Mbp)	Número de contigs	Tamanho mínimo do read	Tamanho máximo do read	Tamanho médio do read	Programa para identificação de repetições no genoma	Número de contigs com elementos repetitivos	Referências
<i>Melipona subnitida</i>	264,06 ^{*1}	141.412	200	11.035	477	Msatcommander	5.574	Souza et al., 2015
<i>Melipona fasciculata</i>	801,96 ^{*2}	47.081	200	16.428	453	Msatcommander	9.954	Silva et al. 2018

Fonte: Elaboração própria.

Figura 4. Distribuição do número de contigs de diferentes comprimentos em *Melipona fasciculata* e *M. subnitida*.



Fonte: Elaboração própria.

Desta forma, o sequenciamento de nova geração em baixa cobertura é bem mais econômico, rápido e eficiente para a descoberta de elementos repetitivos do que as abordagens tradicionais baseadas em clonagem do DNA e enriquecimento por regiões repetitivas (YU et al., 2011; JI et al., 2012). Ele também tem o benefício adicional de limitar erros de sequenciamento devido aos filtros de controle de qualidade. Com a metodologia descrita neste estudo é possível a identificação de elementos repetitivos em *tandem* utilizando a técnica de sequenciamento de nova geração (NGS). A utilização da metodologia descrita

acima vem permitindo a detecção de elementos repetitivos em *tandem* no genoma de diversos organismos, e desta forma, desenvolver inúmeros marcadores microssatélites a partir do sequenciamento de espécies de interesse comercial (SOUZA et al., 2014; SANTOS et al. 2018; SILVA et al., 2018; PEREIRA et al. 2018).

ANEXO

1. Passo-a-passo na bancada

1.1 Tagmentação do DNA genômico

Consumíveis:

ATM (*Amplicon Tagment Mix*); TD (*Tagment DNA buffer*); NT (*Neutralize Tagment Buffer*); Input DNA (0,2–0,4 ng/μL); Placa de 96 poços; Adesivo para vedar as placas.

Preparação:

- 1) Remova o ATM, TD, e o DNA estocado no freezer (-25 a -15°C) e descongele no gelo;
- 2) Visualmente verifique NT para se certificar de que não exista precipitados. Se houver precipitados, agite até que todas as partículas sejam ressuspensas;
- 3) Após o descongelamento, misture os reagentes gentilmente invertendo o tubo 3-5 vezes, seguindo um breve spin na microcentrífuga.

OBS: *Certifique-se que a reação é montada na ordem descrita para o desempenho ideal do kit. Não é necessário manter a reação em gelo.*

Procedimento:

- 1) Nomeie uma nova placa de 96 poços com NTA (*Nextera XT Tagment Amplicon Plate*);
- 2) Adicione 10 μL do tampão TD em cada poço a ser utilizado neste ensaio. Troque as ponteiras entre cada amostra;

OBS: *Calcule o volume total de TD para todas as reações e divida o volume igualmente entre os poços de PCR. Use uma pipeta multicanal para colocar o reagente dentro de cada poço da placa NTA.*

- 3) Adicione 5 μL do *input* de DNA a 0,2-0,4 ng/μL (1 ng total) em cada poço da placa NTA;
- 4) Use uma pipeta multicanal e misture gentilmente por 5 vezes (invertendo o tubo). Mude as ponteiras entre as amostras;
- 5) Adicione 5 μL de ATM em cada poço contendo o *input* de DNA e o buffer TD. Mude as ponteiras entre as amostras;

OBS: *Calcule o volume total de ATM para todas as reações e divida o volume igualmente entre os poços de PCR. Use uma pipeta multicanal para colocar o reagente dentro de cada poço da placa NTA.*

- 6) Use uma pipeta multicanal e misture gentilmente por 5 vezes (invertendo o tubo). Mude as ponteiras entre as amostras;
- 7) Vede a placa NTA com o selo adesivo;
- 8) Centrifugar a $280 \times g$ a 20°C por 1 minuto;
- 9) Coloque a placa no termociclador e executar o programa TAGMENT recomendado pela Illumina;

OBS: *Certifique-se que a tampa do termociclador está aquecida durante a incubação.*

Programa TAGMENT para a PCR:

55°C por 3,5 minutos; Mantenha a 10°C

- 10) Quando as amostras alcançarem 10°C inicie imediatamente a neutralização dos transposomas ativos;
- 11) Retire a placa do termociclador e coloque na bancada.

OBS: Calcule o volume total de NT para todas as reações e divida o volume igualmente entre os poços de PCR. Use a pipeta multicanal para colocar o reagente dentro de cada poço da placa NTA.

- 12) Remova cuidadosamente o adesivo da placa e adicione 5 µL do buffer NT em cada poço da placa NTA. Mude as ponteiros entre as amostras;
- 13) Use uma pipeta multicanal e misture gentilmente por 5 vezes (invertendo o tubo). Mude as ponteiros entre as amostras;
- 14) Vede a placa NTA com o selo adesivo;
- 15) Centrifugue a 280 × g a 20°C por 1 minuto;
- 16) Coloque a placa NTA a temperatura ambiente por 5 minutos.

1.2 Amplificação via PCR

Consumíveis:

NPM (*Nextera PCR Master Mix*); Index 1 primers (N7XX); Index 2 primers (N5XX); *Truseq Index Plate Fixture*; Adesivo para vedar a placa

Preparação (24 bibliotecas):

- 1) Remova o *mix* NPM e os index primers (i5 e i7) estocado no freezer (-25 a -15°C) e descongele em temperatura ambiente (aproximadamente 20 minutos para descongelar o NPM e o index primers).
- 2) Após todos os reagentes estarem descongelados, gentilmente inverta cada tubo 3-5 vezes, seguindo um breve *spin* na microcentrífuga;
- 3) Para 24 amostras organize o index primers na placa *TruSeq* usando a seguinte organização:
A – Organize o index 1 (i7) primers (tampa laranja) na horizontal.
B – Organize o index 2 (i5) primers (tampa branca) na vertical.
C – Para evitar contaminação cruzada, descarte a tampa original e aplique novas tampas em cada index primers.

Procedimento:

- 1) Coloque a placa NTA em *TruSeq Index Plate Fixture*;
- 2) Adicione 15 µL de NPM em cada poço da placa NTA contendo o index primers. Mude as ponteiros entre as amostras;
- 3) Use uma pipeta multicanal e adicione 5 µL do *index 2 primers* (tampa branca). Mude as ponteiros entre as amostras para evitar contaminação;
- 4) Use uma pipeta multicanal e adicione 5 µL do *index 1 primers* (tampa laranja). Mude as ponteiros entre as amostras para evitar contaminação;
- 5) Para evitar contaminação descarte sempre as tampas branca e laranja dos “*index primers*” e coloque uma nova tampa;
- 6) Use uma pipeta multicanal e misture gentilmente 5 vezes (invertendo o tubo). Mude as ponteiros entre as amostras para evitar contaminação;
- 7) Vede a placa NTA com o adesivo;
- 8) Centrifugue a 280 x g a 20°C por 1 minuto;
- 9) Execute a PCR usando o programa do termociclador INDEX recomendado pela Illumina;

OBS: Esteja certo que a tampa do termociclador está aquecida durante a incubação.

Programa INDEX:

72°C por 3 minutos; 95°C por 30 segundos; 12 ciclos de: (95°C por 10 segundos, 55°C por 30 segundos, 72°C por 30 segundos); 72°C por 5 minutos; Manter a 10°C

PONTO DE PARADA

Se não houver planos de proceder imediatamente com a *PCR Clean-Up*, existem duas opções de estocagem: a placa NTA poderá permanecer no termociclador *overnight* ou poderá ser estocada a 2-8°C por dois dias.

1.3 PCR Clean-up

Consumíveis:

RSB (*Resuspension buffer*); 80% etanol (EtOH) fresco; AMPure XP *Beads*; Placa de PCR de 96 poços (*skirted*)

Preparação:

- 1) Traga a AMPure XP *beads* a temperatura ambiente;
- 2) Prepare etanol 80% fresco a partir do etanol absoluto.

OBS: Sempre prepare etanol 80% fresco para a etapa de lavagem. O etanol pode absorver água do ar ambiente impactando seus resultados.

Procedimento:

- 1) Centrifugue a placa NTA a 280 x g a 20°C por 1 minuto para coletar a condensação;
- 2) Nomeie uma nova placa CAA (*Clean Amplified NTA Plate*);
- 3) Usando uma pipeta multicanal definida para 50 µL, transfira o produto de PCR da placa NTA para a placa CAA. Mude as ponteiros entre as amostras;
- 4) Realize um vortex do AMPure XP *beads* por 30 segundos para ter certeza que as esferas estão uniformemente dispersas;
- 5) Usando uma pipeta multicanal, adicione 30 µL de AMPure XP *beads* para cada poço da placa CAA.
- 6) Gentilmente misture invertendo o tubo por 10 vezes;

OBS: Alternativamente a solução pode ser misturada por agitação da placa CAA em um agitador de microplacas a 1800 rpm por 2 minutos.

- 7) Incube à temperatura ambiente sem misturar por 5 minutos;
- 8) Coloque a placa CAA no suporte magnético por 2 minutos ou até o sobrenadante ficar claro;
- 9) Com a placa CAA no suporte magnético, use uma pipeta multicanal, remova e descarte o sobrenadante cuidadosamente. Mude as ponteiros entre as amostras;
- 10) Com a placa CAA no suporte magnético, lave as *beads* com o etanol 80% fresco seguindo as etapas:
 - A. Use a pipeta multicanal e adicione 200 µL de etanol 80% preparado em cada poço. Não ressuspender as *beads* ainda;
 - B. Incube a placa no suporte magnético por 30 segundos;
 - C. Cuidadosamente remova e descarte o sobrenadante.
- 11) Com a placa CAA no suporte magnético, execute a segunda lavagem de etanol como segue:
 - A. Use uma pipeta multicanal e adicione 200 µL de etanol 80% preparado em cada poço. Não ressuspender as *beads* ainda.
 - B. Incube a placa no suporte magnético por 30 segundos.
 - C. Cuidadosamente remova e descarte o sobrenadante.
 - D. Use a pipeta multicanal P20 com uma fina ponteira e remova o excesso de etanol.
- 12) Com a placa CAA ainda no suporte magnético, permita que as *beads* sequem por 15 minutos;
- 13) Remova a placa CAA do suporte magnético. Use uma pipeta multicanal e adicione 52,5 µL de RSB em cada poço da placa;
- 14) Gentilmente misture invertendo o tubo por 10 vezes, mudando as ponteiros.

OBS: Alternativamente a solução pode ser misturada por agitação da placa CAA em um agitador de microplacas a 1800 rpm por 2 minutos.

- 15) Incube à temperatura ambiente por 2 minutos;
- 16) Coloque a placa CAA no suporte magnético por 2 minutos ou até o sobrenadante está claro.
- 17) Nomeie uma nova placa CAN (*Clean Amplified NTA Plate*);
- 18) Use uma pipeta multicanal e transfira cuidadosamente 50 µL do sobrenadante da placa CAA para a placa CAN. Mude as ponteiros para evitar contaminação cruzada entre as amostras.

PONTO DE PARADA

Se não houver planos de proceder imediatamente com a normalização da biblioteca, vede a placa CAN com o adesivo e estoque a -25 a -15°C por uma semana.

1.4 Normalização das bibliotecas**Consumíveis:**

LNA1 (*Library Normalization Additives 1*); LNB1 (*Library Normalization Beads 1*); LNW1 (*Library Normalization Wash 1*); LNS1 (*Library Normalization Storage Buffer 1*); 0.1 N NaOH (preparado a menos de 7 dias; 3 mL/96 amostras); placa de 96 poços (*skirted*); adesivo para vedar a placa; tubos cónicos de 15 mL

Preparação:

Os reagentes LNA1, LNB1, LNW1 e LNS1 são preparados de acordo com as recomendações do kit.

Procedimento:

- 1) Transferir 20 μ L de sobrenadante da placa de PCR para uma nova placa (*midi*);
- 2) Adicione 44 μ L de LNA1 por amostra a um novo tubo cónico de 15 mL. Calcular cerca de 5% de amostra extra por conta de perda da amostra devido à pipetagem. Por exemplo: para 96 amostras, adicione 4,4 mL de LNA1 ao tubo ($100 \text{ amostras} \times 44 \mu\text{L} = 4,4 \text{ mL}$);
- 3) Ressuspender completamente o LNB1. Pipetar para misturar;
- 4) Transferir 8 μ L de LNB1 por amostra (incluindo o extra de 5%) para o tubo cónico de 15 ml contendo LNA1. Inverta para misturar. Por exemplo: para 96 amostras, transfira 800 μ L de LNB1 para o tubo de LNA1 ($100 \text{ amostras} \times 8 \mu\text{L} = 800 \mu\text{L}$);
- 5) Despeje a mistura de esferas em um poço;
- 6) Adicione 45 μ L de LNA1 e LNB1 combinados a cada poço contendo as bibliotecas;
- 7) Agite a 1800 rpm por 30 minutos;
- 8) Coloque em um suporte magnético e espere até que o líquido esteja limpo (~ 2 minutos);
- 9) Remova e descarte todo o sobrenadante de cada poço;
- 10) Lave duas vezes da seguinte maneira:
 - A - Adicione 45 μ L de LNW1 a cada poço.
 - B - Agite a 1800 rpm por 5 minutos.
 - C - Coloque em um suporte magnético e espere até que o líquido esteja limpo (~ 2 minutos).
 - D - Remova e descarte todo o sobrenadante de cada poço.
- 11) Adicionar 30 μ L de NaOH 0,1 N a cada poço;
- 12) Agite a 1800 rpm por 5 minutos;
- 13) Durante a eluição de 5 minutos, nomeie uma nova placa de PCR de 96 poços SGP para placa de armazenamento;
- 14) Adicione 30 μ L de LNS1 a cada poço da placa SGP. Deixe em repouso;
- 15) Após a eluição de 5 minutos, certifique-se de que todas as amostras na placa *midi* sejam ressuspensas. Se isso não acontecer, ressuspender da seguinte forma:
 - A - pipetar para misturar ou bater levemente a placa na bancada;
 - B - Agite a 1800 rpm por 5 minutos.
- 16) Coloque em um suporte magnético e espere até que o líquido esteja limpo (~ 2 minutos);
- 17) Transfira o sobrenadante da placa *midi* para a placa SGP.
- 18) Centrifugue a 1000 \times g por 1 minuto.

PONTO DE PARADA

Se você estiver parando, sele a placa e armazene entre -25°C e -15°C por até sete dias.

1.5 Agrupamento das bibliotecas**Consumíveis:**

Adesivo para vedar a placa; tubos Eppendorf “*LoBind*” para microcentrifugação; microtubos de PCR (fita de 8 tubos).

Preparação:

- 1) Para preparar a corrida de sequenciamento, comece a descongelar os reagentes de acordo com as instruções do fabricante do MiSeq;
- 2) Se a placa SGP foi armazenada congelada entre -25°C e -15°C, descongele à temperatura ambiente. Pipete para misturar.

Procedimento:

- 1) Centrifugue a $1000 \times g$ a 20°C durante 1 minuto;
- 2) Nomeie um novo tubo Eppendorf PAL (*Pooled Amplicon Library*);
- 3) Transfira 5 μL de cada biblioteca da placa SGP para o tubo PAL. Inverta para misturar;
- 4) Dilua as bibliotecas agrupadas à concentração da corrida de sequenciamento;
- 5) Armazene as bibliotecas agrupadas não utilizadas no tubo PAL e na placa SGP entre -25°C e -15°C por até 7 dias.

3.4 Referências

ANDERSEN, J. C.; MILLS, N.J. iMSAT: a novel approach to the development of microsatellite loci using barcoded Illumina libraries. *BMC Genomics*, v.15, p.858, 2014.

ANDREWS S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

AVISE, J.C. **Molecular markers, natural history e evolution**. New York: Chapman e Hall, Inc., 511p. 5^a ed. Garland Science, New York. 725p. 1994.

CASTELO AT, MARTINS W, GAO GR. TROLL—tandem repeat occurrence locator. *Bioinformatics*, v.18, n.4, p.634–636, 2002.

DINIZ, F. M.; IYENGAR, A.; LIMA, P. S. C.; MACLEAN, N.; BENTZEN, P. Application of a double-enrichment procedure for microsatellite isolation and the use of tailed primers for high throughput genotyping. *Genetics and Molecular Biology*, v. 30, p. 380-384, 2007.

Du, L., Li, Y., Xiuyue, L., Yue, B. MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. *Journal of Heredity*, v.104, n.1, p.154-157, 2013.

EWING, B. AND GREEN, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research*, v.8, p.186-194, 1998.

EWING, B., HILLIER, L., WENDL, M.C. AND GREEN, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research*. v.8, p.175-185, 1998.

FAIRCLOTH B. Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, v.8, p.92-94, 2008.

FARRER RA, KEMEN E, JONES JD, STUDHOLME DJ. De novo assembly of the *Pseudomonas syringae* pv. *syringae* B728a genome using Illumina/Solexa short sequence reads. *FEMS. Microbiol. Lett.* v.291, p.103-111, 2009.

FEDURCO, M.; ROMIEU, A.; WILLIAMS, S. et al. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Research*, v. 34, p.22, 2006.

GRABHERR , M. G. , B. J. HAAS , M. YASSOUR , J. Z. LEVIN , D. A. THOMPSON , I. AMIT , X. ADICONIS. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, v.29, p.644 – 652, 2011.

JI P, LIU G, XU J, WANG X, et al. (2012). Characterization of common carp transcriptome: sequencing, de novo assembly, annotation and comparative genomics. **PLoS One** **7**: e35152.

LI Q & WAN J.M. SSRHunter: development of a local searching software for SSR sites. **Yi Chuan**, v.27, p.808-810, 2005.

Mayer C (2006) Phobos 3.3.11. Disponível em: http://www.rub.de/spezzoo/cm/cm_phobos.htm

PEREIRA, V. A.; FORTE, J. M.; ARRUDA-JUNIOR, J. P. V.; DINIZ, FÁBIO M. ; [MAGGIONI, R.](#) ; SALMITO-VANDERLEY, C. S. B. . Microsatellite discovery and characterization in the West Atlantic sea cucumber *Holothuria grisea* (Selenka 1867). **JOURNAL OF GENETICS**, 2018. Aceito para publicação. Disponível em: <https://www.ias.ac.in/public/Resources/General/jgen/JGEN-D-18-00073.pdf>

PERRY JC, ROWE L. Rapid microsatellite development for water striders by next-generation sequencing. **Journal of Heredity**. v.102, p.125-129, 2011.

ROZEN S.; SKALETSKY H.J. 2000: Primer 3 on the www for general users and for biologist programmers. In: Krawetz S. & Misener S. (eds) *Bioinformatics methods and protocols: methods in molecular biology*. **Human Press**, Totowa, p 365–386, 2000.

SANTOS, M. F.; [SOUZA, I. G. B.](#) ; GOMES, S. O. ; [SILVA, G. R.](#) ; BENTZEN, PAUL ; DINIZ, F. M. . Isolation and characterization of microsatellite markers in the spiny lobster, *Panulirus echinatus* Smith, 1869 (Decapoda: Palinuridae) by Illumina MiSeq sequencing. **Journal of Genetics**, p. e25-e30, 2018.

SILVA, GR ; SOUZA, IGD ; PEREIRA, FD ; SOUZA, BD ; LOPES, MTD ; BENTZEN, P ; DINIZ, F.M. Genome-wide discovery and characterization of microsatellite markers from *Melipona fasciculata* (Hymenoptera: Apidae), cross-amplification and a snapshot assessment of the genetic diversity in two stingless bee populations. **European Journal of Entomology**. v.115, p.614-619, 2018.

SIMPSON JT, WONG K, JACKMAN SD, SCHEIN JE, JONES SJ, BIROL I. ABySS: a parallel assembler for short read sequence data. **Genome Research**. v.19, p.1117–1123, 2009.

SOUZA W, CARVALHO B. Rqc: Quality Control Tool for High-Throughput Sequencing Data. R package version 1.2.0; 2014.

SOUZA, G.A. et al. Diversidade genética estimada com marcadores ISSR em populações brasileiras de *Zabrotes subfasciatus*. **Pesquisa Agropecuária Brasileira**, v.43, n.7, p.843-49, 2008.

SOUZA, IGB; PATERSON, I; MCBRIDE, MC; SOUZA, BA; PEREIRA, FM; LOPES, MTR; BENTZEN, P; DINIZ, F.M. Isolation and characterization of 23 microsatellite loci in the stingless bee *Melipona subnitida* using next-generation sequencing. **Conservation Genetics Resources**. v.7, p.239-241, 2015.

TEMNYKH S, DECLERCK G, LUKASHOVA A, LIPOVICH L, CARTINHO S, MCCOUCH S. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. **Genome Research**, v.11, n.8, p.1441–1452, 2001.

THIEL , T. , W. MICHALEK , R. K. VARSHNEY , AND A. GRANER. Exploiting EST database for the development and characterization of gene-derived SSR markers in barley (*Hordeum vulgare* L.). **Theoretical and Applied Genetics**, v.106, p.411 – 422, 2003.

TURCATTI, G.; ROMIEU, A.; FEDURCO, M. et al. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. **Nucleic Acids Research**, v. 36, p.25, 2008.

YU JN, WON C, JUN J, LIM Y, KWAK M. Fast and cost-effective mining of microsatellite markers using NGS technology: an example of a Korean water deer *Hydropotes inermis argyropus*. **PLoS One**. v.6, p.26-33, 2011.

ZHANG J, KOBERT K, FLOURI T, STAMATAKIS A. PEAR: a fast and accurate Illumina Paired-end read mergeR. **Bioinformatics**, v.30, p.614-620, 2014.

ZHAO, W-G. et al. Molecular genetic diversity and population structure in *Lycium* accessions using SSR markers. **Comptes Rendus Biologies**, v.333, n.11-12, p.793-800, 2010.

CAPÍTULO 4

Mineração do genoma de espécies de abelhas do gênero *Melipona* para a busca de sequências simples repetidas: abundância, composição e variação do comprimento

Artigo: “Mining of the genome of bees of the genus Melipona for the search for repeated simple sequences: abundance, composition and length variation”, a ser submetido

4.1 Resumo

Para estudos de genética populacional, os microssatélites se tornaram um dos marcadores mais utilizados devido o grande número de informações por eles gerados. No entanto, o conhecimento sobre a distribuição e frequência dos principais tipos de microssatélites se torna essencial para se selecionar aqueles que atendam aos requisitos de um efetivo marcador molecular. Com isso tem-se na tecnologia de sequenciamento de Nova Geração (*Next-generation*) uma grande vantagem, quando se compara com outros métodos, uma vez que, a longo prazo, as análises envolvem baixo custo e gera grande número de *reads* em um curto período de tempo. Baseado nisso, por meio do emprego do método de sequenciamento de Nova Geração de emparelhamento de extremidades Illumina, essa pesquisa teve como foco, identificar e compreender a função, evolução e a dinâmica dos microssatélites, em três espécies de abelhas-sem-ferrão (*Melipona fasciculata*, *Melipona subnitida* e *Melipona rufiventris*) de forma a auxiliar no processo de desenvolvimento dos marcadores. Para isso empregou-se na análise o mais recente software de detecção de microssatélites, denominado MSDB (*Microsatellite Search and Building Database*). Por meio desse estudo, foi possível gerar informações que podem ser importantes na compreensão dos microssatélites presentes no genoma dessas espécies, podendo inclusive auxiliar na seleção e desenvolvimento de marcadores que poderão ser aplicados em estudos genéticos, não só para essas espécies como para outras abelhas.

Palavras-chave: Frequência de Microssatélites, Illumina, Sequenciamento *Next-generation*.

4.2 Introdução

O conhecimento da estrutura genética de diversas espécies de abelhas pode ser obtido com o uso de marcadores moleculares microssatélites, os quais representam poderosa ferramenta na averiguação de parâmetros genéticos.

Esses microssatélites, também denominados repetições de sequência simples (SSRs) ou repetições curtas *em tandem* (STR), são sequências especiais de DNA com unidades repetidas de um a seis nucleotídeos em genomas eucarióticos, sendo classificadas pelo *motif* repetido, podendo ser mono, di, tri, tetra, penta e hexanucleotídeos e, também, pelo tipo de sequência repetitiva, como microssatélites perfeitos, imperfeitos, interrompidos ou compostos (LI et al., 2002).

No geral, têm sido a classe de marcadores moleculares mais amplamente utilizados em estudos genéticos, com aplicações em vários campos da genética, incluindo conservação genética, genética populacional e testes de paternidade. Esta gama de aplicações deve-se ao fato de que os marcadores de microssatélites são codominantes e multialélicos, altamente reprodutíveis, têm alta resolução e são baseados na reação em cadeia da polimerase (PCR) (OLIVEIRA et al., 2006; ELLEGREN, 2004). Outro fator que torna esses marcadores tão valiosos é seu maior nível de transferibilidade para espécies relacionadas, e muitas vezes podem ser usados como marcadores de âncora para mapeamento comparativo e estudos evolutivos (VARSHNEY; GRANER; SORRELLS, 2005). Além disso, a frequência, densidade e distribuição de microssatélites variam amplamente em genomas de diferentes espécies.

Porém, para se isolar e empregar os microssatélites como marcadores é necessário que se conheça previamente sua localização no genoma do organismo em estudo, de forma a delimitar as regiões flangeadoras que servirão como referência para o anelamento dos *primers* para a progressão da reação de PCR (*Polymerase Chain Reaction*). Para isso, existem diferentes protocolos de isolamento de microssatélites, desde os mais tradicionais aos mais modernos. Com o avanço das tecnologias e consequentemente o surgimento do sequenciamento do genoma em larga escala abriram-se novas vias para executar varreduras genômicas para um grande número de microssatélites.

No Brasil as abelhas-sem-ferrão se destacam pela importância econômica, principalmente na produção de mel e propólis, além de serem cruciais na polinização da vegetação nativa. Porém, esse importante recurso natural têm sofrido ameaças devido os constantes desmatamentos e uso desordenado de agrotóxicos. Isso tem provocado queda na variabilidade

genética de várias populações, seja pela fixação de genes ou pela deriva genética, o que vem a dificultar a adaptação a possíveis mudanças ambientais.

Por isso torna-se necessário fazer uma caracterização genética populacional mais eficiente das espécies desse grupo de abelhas, a fim de auxiliar no processo de conservação e manejo das populações, de forma a estimular práticas de meliponicultura (criação racional de abelhas-sem-ferrão) que garantam a manutenção desse importante recurso natural (SILVA et al., 2014).

Assim, por meio da técnica de sequenciamento de emparelhamento de extremidades Illumina (Next-generation), nesse estudo foi realizado uma busca e caracterização dos microssatélites no genoma das abelhas *Melipona fasciculata*, *Melipona subnitida* e *Melipona rufiventris*, com o propósito de compreender a função, evolução e a dinâmica dos microssatélites em abelhas-sem-ferrão.

4.3 Material e Métodos

Foram coletadas abelhas operárias de uma população de cada espécie de abelha: *Melipona fasciculata*, *Melipona subnitida* e *Melipona rufiventris* proveniente do meliponário da EMBRAPA Meio-Norte, localizado na cidade Teresina, Estado do Piauí, Brasil. Previamente as abelhas foram armazenadas em etanol absoluto a -20°C, antes do início dos procedimentos. No Laboratório de Biologia Molecular e Biotecnologia da Embrapa Meio-Norte o DNA genômico total foi extraído a partir do tórax de cinco indivíduos de cada espécie, seguindo recomendações de Sambrook e Russell (2001).

No Laboratório *Marine Gene Probe* do departamento de Biologia na Dalhousie University (Halifax-Canadá) um único indivíduo da espécie *M. fasciculata* e outro da *M. subnitida* com o maior rendimento e qualidade de DNA foram selecionados para o sequenciamento. Já a espécie *M. rufiventris* teve seu DNA sequenciado no NPDM – Núcleo de Pesquisa e Desenvolvimento de Medicamentos da Universidade Federal do Ceará-UFC.

Para a obtenção dos *reads* usando 1 ng de DNA genômico, uma biblioteca de extremidades emparelhadas da Illumina foram produzidas conforme especificações do kit de preparação da Illumina Nextera (Illumina Inc.). O sequenciamento foi conduzido usando o sequenciador Miseq Benchtop (Illumina Inc.). Os *contigs* foram criados a partir dos dados da sequência emparelhado pelas extremidades obtidas por meio do CLC *Genomics Workbench* 7.0.4 (Qiagen).

A fim de selecionar os *motifs* de microssatélites, adotou-se o software MSDB (*Microsatellite Search Building Database*) v.2.4.3 (DU et al., 2013). Foram feitas análises

para mononucleotídeos, dinucleotídeos, trinucleotídeos, tetranucleotídeos, pentanucleotídeos e hexanucleotídeos com número mínimo de repetições de 14, sete, cinco, quatro, três e três, respectivamente. Desses resultados gerou-se arquivos em planilhas no Excel, nos quais foram obtidos o número de *motifs*, o comprimento total, média de cada *motif*, abundância relativa (Nº de *motifs* / tamanho total das sequências analisadas em Mega bases, Mb) e densidade (comprimento total do *motif* em bp / tamanho total das sequências analisadas em Mb).

Na análise foram distinguidos as diferentes classes de microssatélites presentes nos *reads*, obtidos do genoma das três espécies (*M. fasciculata*, *M. subnitida* e *M. rufiventris*), conforme determinado no manual do software MSDB (<http://msdb.biosv.com/>) (DU et al., 2013), a partir dos *contigs* obtidos no formato FASTA. Para tanto, foram considerados os microssatélites puros, puros interrompidos, compostos, compostos interrompidos, complexos e complexos interrompidos (CHAMBERS; MACAVOY, 2000; BACHMANN; BAREI; TOMIUK, 2004).

4.4 Resultados e Discussões

A partir do sequenciamento do genoma de três espécies de abelhas do gênero *Melipona*, obteve-se uma plataforma com 137.313 *contigs* para a espécie *M. rufiventris* (Mruf), 47.081 *contigs* para *M. fasciculata* (Mfsc) e 141.412 *contigs* para *M. subnitida* (Msub). Um total de 3.495 SSR foi identificado em Mruf, 11.869 em Mfsc e 93.588 em Msub (**Tabela 1**). O comprimento total, a densidade, a abundância e a frequência das seis classes de SSR perfeitas estão resumidos na Tabela 1.

Quando se analisa as unidades repetitivas básicas de Mfsc e Msub, os mononucleotídeos foram os mais abundantes, seguidos do di- e trinucleotídeos, com frequência para Mfsc de 54,3; 27,2 e 6,2% e abundância relativa de 302,20, 151,2 e 34,4; e para Msub de 93,8; 5,5 e 0,4% e abundância relativa de 1.301,8, 77,0 e 5,7, respectivamente.

Os microssatélites mais abundantes para Mruf foram os pentanucleotídeos com frequência de 66,2% e abundância relativa de 221, sendo que os mononucleotídeos tiveram o menor percentual, com 1,7 e 5,5% de abundância (Tabela 1). De acordo com Sharma et al. (2007) a prevalência de repetições mononucleotídicas é comum em genomas eucarióticos, sobre as outras classes de nucleotídeos.

Tabela 1. Cálculo da contagem total (N), comprimento total (bp), densidade (SSR em bp/Mb), abundância (N° de SSR/Mb) e frequência (%) dos microssatélites do genoma das abelhas *Melipona rufiventris* (*Mruf*), *M. fasciculata* (*Mfsc*) e *M. subnitida* (*Msub*). Mono – Mononucleotídeos, Di – Dinucleotídeos, Tetra – Tetranucleotídeos, Penta – Pentanucleotídeos e Hexa – Hexanucleotídeos

		Tipos de microssatélites						
		Mono	Di	Tri	Tetra	Penta	Hexa	Total
<i>Mruf</i>	Contagem total (N)	58,0	562,0	252,0	246,0	2312,0	65,0	3495,0
	Comprimento total (bp)	886,0	10436,0	4773,0	4468,0	40125,0	1200,0	61888,0
	Densidade (SSR em bp/Mb)	84,7	997,4	456,2	427,0	3835,0	114,7	5915,0
	Abundância (N° de SSR/Mb)	5,5	53,7	24,1	23,5	221,0	6,2	334,0
	Frequência (%)	1,7	16,1	7,2	7,0	66,2	1,9	100,0
<i>Mfsc</i>	Contagem total (N)	6444,0	3225,0	734,0	574,0	658,0	234,0	11869,0
	Comprimento total (bp)	105844,0	68118,0	15369,0	12048,0	12100,0	4848,0	218327,0
	Densidade (SSR em bp/Mb)	4963,7	3194,5	720,8	565,0	567,5	227,4	10238,8
	Abundância (N° de SSR/Mb)	302,2	151,2	34,4	26,9	30,9	11,0	556,6
	Frequência (%)	54,3	27,2	6,2	4,8	5,5	2,0	100,0
<i>Msub</i>	Contagem total (N)	87762,0	5192,0	384,0	163,0	54,0	33,0	93588,0
	Comprimento total (bp)	1366226,0	152418,0	14196,0	6836,0	2270,0	1506,0	1543452,0
	Densidade (SSR em bp/Mb)	20265,1	2260,8	210,6	101,4	33,7	22,3	22893,9
	Abundância (N° de SSR/Mb)	1301,8	77,0	5,7	2,4	0,8	0,5	1388,2
	Frequência (%)	93,8	5,5	0,4	0,2	0,1	0,0	100,0

Quando se analisa os diferentes *motifs* de microssatélites nas três espécies verifica-se que as repetições dinucleotídicas, de *motifs* (AT)₃ em *Melipona rufiventris*, (AC)₃ em *Melipona fasciculata* e (TC)₇ em *Melipona subnitida* foram os mais abundantes neste estudo, com abundância relativa de 744,25; 366,03 e 4,79 SSR/Mb, respectivamente. Entretanto quando se analisa os menos abundantes é possível identificar que para a espécie *Msub*, os quatro *motifs* pentanucleotídeos foram os que apresentaram valores mais baixos, seguidos dos quatro *motifs* tetranucleotídeos da espécie *Mfsc* e quatro *motifs* hexanucleotídeos da espécie *Mruf* (Tabela 2).

De acordo com Karaoglu-Lee & Meyer (2005) as repetições dinucleotídicas, como marcadores moleculares, são mais importantes do que os outros microssatélites por apresentarem taxas de mutação mais elevadas. No entanto, as outras repetições, como as tri e tetranucleotídicas por ter facilidade de visualização e *scoring* em gel, têm sido preferidas em estudos populacionais.

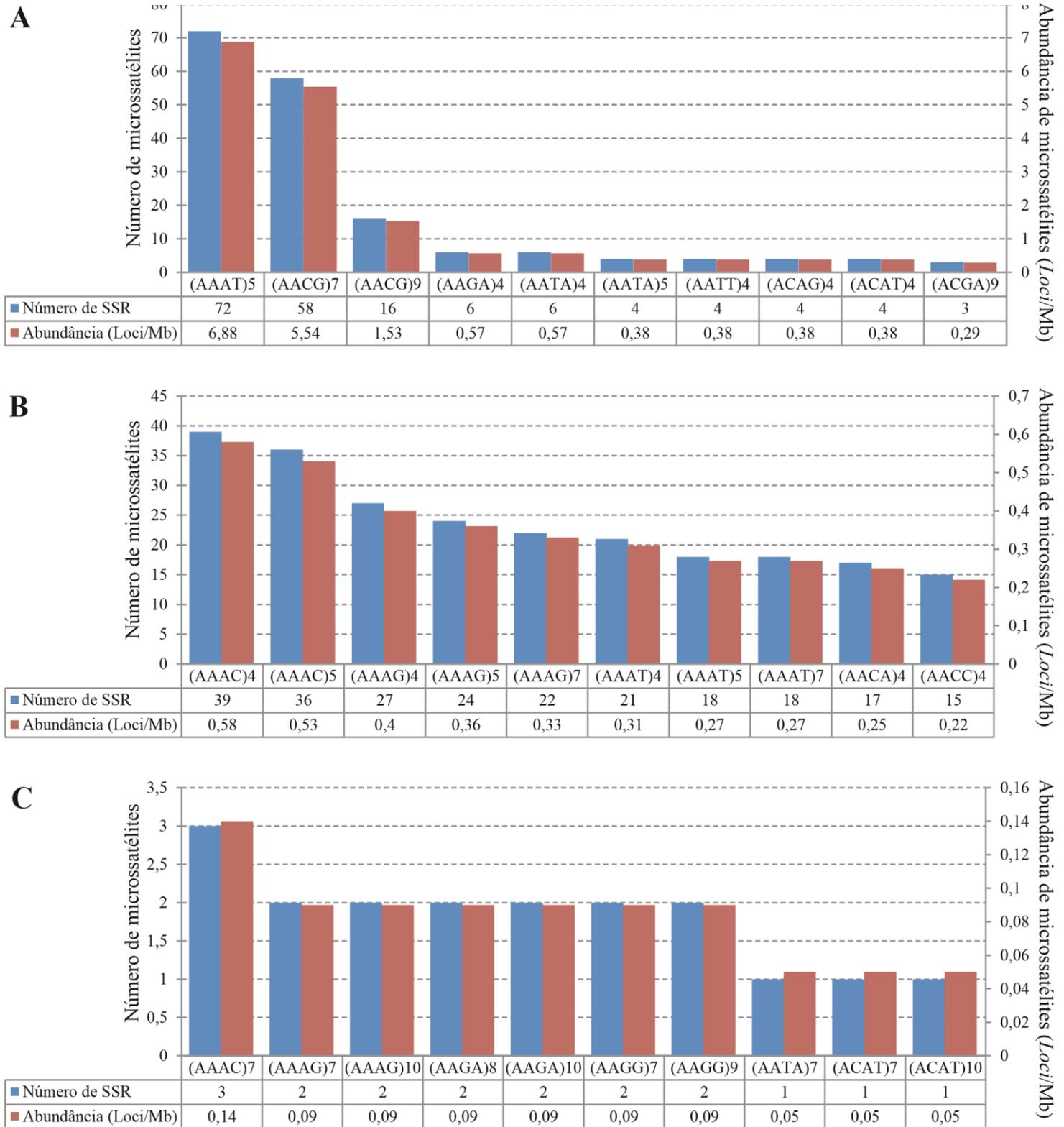
A distribuição dos dez *motifs* mais frequentes de repetições tetranucleotídeos, encontrados neste estudo, estão resumidas na Figura 1. Sendo que o destaque vai para os dez *motifs* da espécie *M. rufiventris*, que apresentaram um maior número de SSR, assim como

uma maior relação de abundância, quando se compara com os valores das outras duas espécies. Os *motifs* de microssatélites mais frequentes foram: AAAT para Mruf e AAAC para Msub e Mfsc. Os tetranucleotídeos compreendem uma das classes de microssatélites mais indicadas como marcadores genéticos, devido a fácil distinção no gel de poliacrilamida, assim como pela grande frequência nas regiões neutras, teoricamente livres de seleção e pela tendência de serem mais informativas e mais estáveis sob as condições da PCR, do que, por exemplo, as repetições di- ou tri-nucleotídeos (GUO et al., 2013).

Tabela 2. Descrição dos quatro primeiros *motifs* com maior número (No) e abundância (N° SSR/Mb) dos microssatélites dinucleotídeos (Di), trinucleotídeos (Tri), tetranucleotídeos (Tetra), pentanucleotídeos (Penta) e hexanucleotídeos (Hexa) das espécies de abelhas-sem-ferão *Melipona rufiventris*, *M. fasciculata* e *M. subnitida*.

Tipo	<i>Melipona rufiventris</i>			<i>Melipona fasciculata</i>			<i>Melipona subnitida</i>		
	<i>Motifs</i>	No	Abundância	<i>Motifs</i>	No	Abundância	<i>Motifs</i>	No	Abundância
Di	(AT)3	7787	744,25	(AC)3	7805	366,03	(TC)7	323	4,79
	(TA)3	4058	387,85	(GT)3	6943	325,60	(CT)7	313	4,64
	(GA)3	2758	263,6	(AT)3	6902	323,68	(TA)7	297	4,41
	(TC)3	2205	210,74	(CA)3	6536	306,52	(GA)7	285	4,23
Tri	(TAT)4	142	13,57	(CAA)8	7	0,33	(TGT)4	206	3,06
	(CTC)4	140	13,38	(AAT)7	6	0,28	(GAA)4	162	2,40
	(ATA)4	139	13,28	(TAT)7	5	0,23	(ACA)4	156	2,31
	(GAG)4	136	13,00	(TCG)8	5	0,23	(AAC)4	136	2,02
Tetra	(ATCT)4	72	6,88	(GAAA)14	3	0,14	(GAAA)4	39	0,58
	(TAGA)4	58	5,54	(AAGA)10	2	0,09	(CTTT)4	36	0,53
	(AAGA)4	16	1,53	(CGAA)8	2	0,09	(TGTT)4	27	0,40
	(GACA)5	6	0,57	(GAAA)7	2	0,09	(AAGA)4	24	0,36
Penta	(ATTTT)3	151	14,43	(CTTTT)3	23	1,08	(GAAAA)7	3	0,04
	(TTATT)3	151	14,43	(GAAAA)3	15	0,70	(AAAAG)7	2	0,03
	(TATTT)5	135	12,90	(TGTTT)3	13	0,61	(AACGA)7	2	0,03
	(ATTTA)3	134	12,81	(GTTTT)3	12	0,56	(CGAGA)7	2	0,03
Hexa	(TTAGGT)3	9	0,86	(CTTTTT)3	9	0,42	(CTTTTT)3	29	0,43
	(AAGTAA)3	7	0,67	(GAGAAA)3	7	0,33	(GTTTTT)3	22	0,33
	(ATTAAA)3	6	0,57	(ATTTTT)3	4	0,19	(GAAAAA)3	14	0,21
	(AATTAA)3	5	0,48	(GAAAAA)3	4	0,19	(GTTTGT)3	14	0,21

Figura 1. Os dez *motifs* de microssatélites tetranucleotídeos mais frequentes e abundantes encontrados no genoma sequenciado das abelhas-sem-ferrão *Melipona rufiventris* (A), *Melipona subnitida* (B) e *Melipona fasciculata* (C).



Os microssatélites formam grupos em seis classes quanto à natureza de sua repetição: composto, complexo, composto interrompido, complexo interrompido, puro interrompido e puros. No geral, analisando os valores na tabela 3, de todos os microssatélites avaliados para as três espécies, a maioria, foram perfeitos, seguido dos imperfeitos (puros interrompidos, compostos interrompidos, e complexos interrompidos) e por último dos microssatélites compostos puros (complexos e compostos).

Em *M. rufiventris* os microssatélites puros foram os mais abundantes (3.219), seguidos por composto (52), composto interrompido (25), complexo interrompido (18), complexo (15) e puro interrompido (7). O comprimento total, frequência e densidade foi de 56.505, 96,49 e 307,660 para a espécie, respectivamente. As abelhas *M. fasciculata* e *M. subnitida* tiveram uma quantidade muito maior de microssatélites puros (11.234 e 82.659, respectivamente). Quando se analisa individualmente os valores de *M. subnitida* verifica-se que essa espécie obteve maiores valores para todos os tipos de microssatélite, exceto para o tipo complexo, que no caso, foi atribuído o maior valor para espécie *M. rufiventris*.

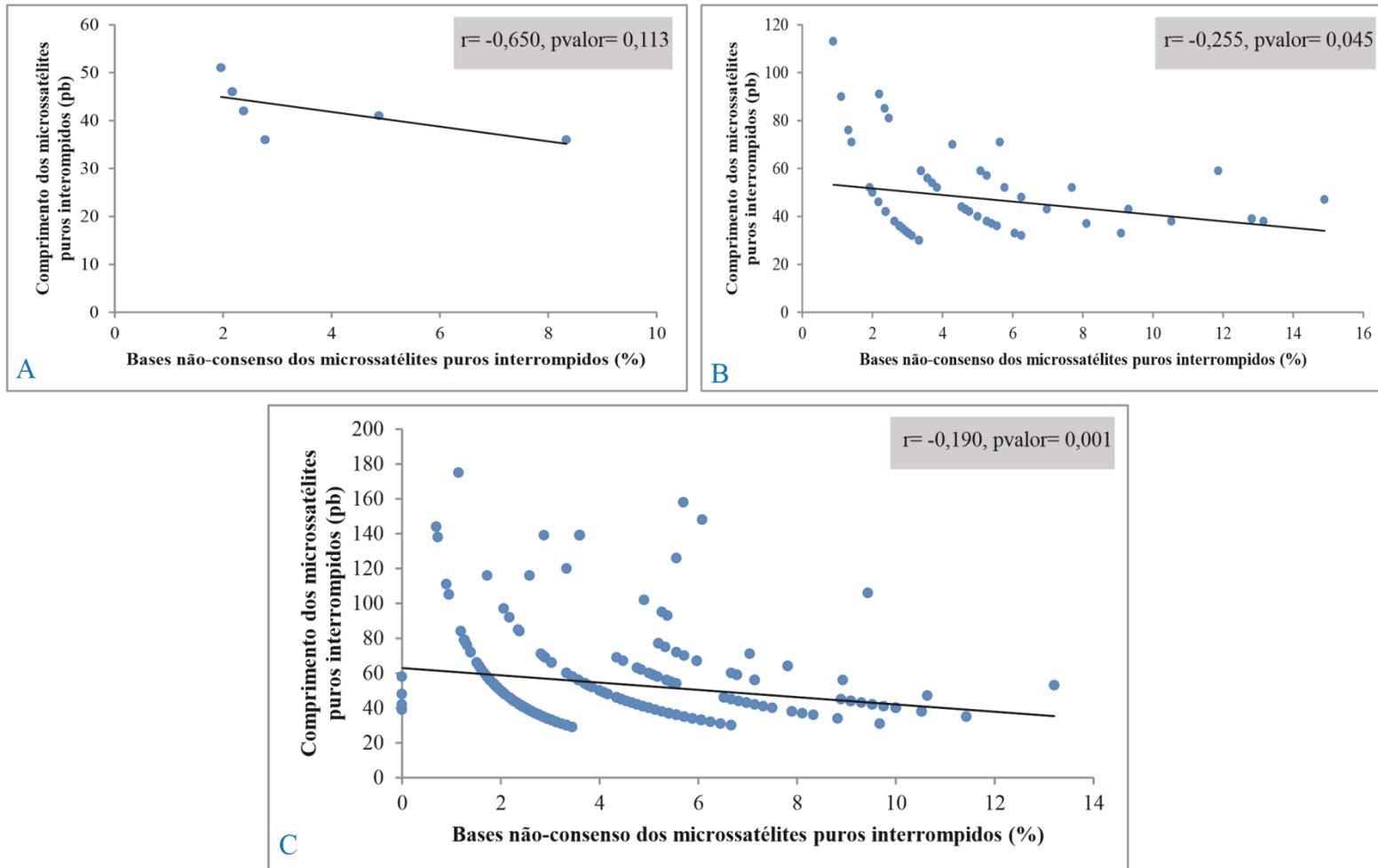
Apesar de serem considerados em sua maioria marcadores evolutivamente neutros, sabe-se que os microssatélites puros tem papel fundamental nas conformações estruturais do DNA, que são essenciais para alguns processos metabólicos, bem como para a regulação da expressão genética e para organização do centrômero, telômero e participação no enovelamento ou compactação dos cromossomos (KUMAR et al. 2010; LI et al. 2002).

O aumento no número de bases não-consenso, ou de imperfeições nos microssatélites, tenderam a provocar uma redução no tamanho dos microssatélites (M_{ruf} : $r = -0,650$, $p_{valor} = 0,113$; M_{fsc} : $r = -0,255$, $p_{valor} = 0,045$; M_{sub} : $r = -0,190$, $p_{valor} = 0,001$), o que mostra que quanto menos interrupção nos microssatélites maior será a possibilidade de mutações (Figura 2).

Tabela 3. Descrição do número (No), comprimento total, frequência, densidade e abundância dos tipos de microssatélites: composto, complexo, composto interrompido, puro e puro interrompido nas espécies de abelhas-sem-ferrão *M. rufiventris*, *M. fasciculata* e *M. subnitida*.

Espécie	Parâmetros	Tipos de microssatélites					
		composto	complexo	composto int	complexo int	puro int	puro
<i>Melipona rufiventris</i>	No	52,000	15,000	25,000	18,000	7,000	3219,000
	comprimento total (bp)	2022,000	925,000	1048,000	1342,000	298,000	56505,000
	Frequência (%)	1,559	0,450	0,749	0,540	0,210	96,493
	Densidade (SSR em bp/Mb)	4,970	1,430	2,390	1,720	0,670	307,660
	Abundância (N° de SSR/Mb)	193,250	88,410	100,160	128,260	28,480	5400,500
<i>Melipona fasciculata</i>	No	80,000	6,000	127,000	23,000	62,000	11234,000
	comprimento total (bp)	3390,000	385,000	5831,000	1822,000	2971,000	204860,000
	Frequência (%)	0,694	0,052	1,101	0,199	0,538	97,416
	Densidade (SSR em bp/Mb)	158,980	18,060	273,460	85,450	139,330	9607,280
	Abundância (N° de SSR/Mb)	3,750	0,280	5,960	1,080	2,910	526,840
<i>Melipona subnitida</i>	No	249,000	8,000	382,000	62,000	521,000	82659,000
	comprimento total (bp)	10482,000	551,000	17127,000	4325,000	25571,000	1446162,000
	Frequência (%)	0,697	0,037	1,139	0,288	1,700	96,140
	Densidade (SSR em bp/Mb)	155,480	8,170	254,040	64,150	379,290	21450,770
	Abundância (N° de SSR/Mb)	3,690	0,120	5,670	0,920	7,730	1226,070

Figura 2. Correlação entre Bases não-consenso e comprimento dos microssatélites puros interrompidos nas espécies de abelhas-sem-ferrão *Melipona rufiventris* (A), *M. fasciculata* (B) e *M. subnitida* (C).



4.5 Conclusões

Utilizando o sequenciamento de Nova Geração, foi possível identificar um grande volume de microssatélites nas espécies *Melipona rufiventris*, *Melipona fasciculata* e *Melipona subnitida*, com destaque para os mono (exceto Mruf) e dinucleotídeos. Além disso, os microssatélites puros foram os mais frequentes em ambas as espécies, com potencial para serem utilizados para a síntese de *primers* que servirão para amplificar marcadores moleculares para o posterior uso em análises genéticas para a distinção das populações de abelhas.

4.6 Referências

- BACHMANN, L.; BAREI, P.; TOMIUK, J. Allelic variation , fragment length analyses and population genetic models : a case study on Drosophila microsatellites. **Journal of Zoological Systematics and Evolutionary Research**, v. 42, p. 215–222, 2004.
- CHAMBERS, G. K.; MACAVOY, E. S. Microsatellites: Consensus and controversy. **Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology**, v. 126, n. 4, p. 455–476, 2000.
- DU, L. et al. MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. **Journal of Heredity**, v. 104, n. 1, p. 154–157, 2013.
- ELLEGREN, H. Microsatellites: Simple sequences with complex evolution. **Nature Reviews Genetics**, v. 5, n. 6, p. 435–445, 2004.
- GUO, X. Z. et al. Development of twenty-one polymorphic tetranucleotide microsatellite loci for 53 *Schizothorax o'connori* and their conservation application. **Biochemical Systematics and Ecology**, v. 51, p. 259–263, 2013.
- KARAOGLU, H.; LEE, C.M.Y.; MEYER, W. Survey of Simple Sequence Repeats in Completed Fungal Genomes. **Molecular Biology and Evolution**, v.22, p.639-649, 2005.
- KUMAR, R. P. et al. Repeat performance: how do genome packaging and regulation depend on simple sequence repeats? **Bio Essays**. v.32, p.165-174, 2010.
- LI, Y. C. et al. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. **Molecular ecology**. v.11, p.2453-2465, 2002.
- OLIVEIRA, E. J. et al. Origin , evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, v. 29, n.2, p. 294–307, 2006.
- SAMBROOK, J.; RUSSELL, D.W. **Molecular Cloning: A Laboratory Manual**. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 2001.

SILVA, G.R. et al. New molecular evidence for fragmentation between two distant populations of the threatened stingless bee *Melipona subnitida* Ducke (Hymenoptera, Apidae, Meliponini). **Journal of Hymenoptera Research** v.38, p.1-9, 2014.

SHARMA, P. C.; GROVER, A.; KAHL, G. Mining microsatellites in eukaryotic genomes. **Trends in Biotechnology**, v. 25, n. 11, p. 490–498, 2007.

VARSHNEY, R. K.; GRANER, A.; SORRELLS, M. E. Genic microsatellite markers in plants: Features and applications. **Trends in Biotechnology**, v. 23, n. 1, p. 48–55, 2005.

CAPÍTULO 5

Estudo da diversidade genética de *Melipona rufiventris* (Hymenoptera: Apidae) no semiárido nordestino usando microssatélites heteroespecífico

Artigo “Snapshot of the genetic diversity of Melipona rufiventris (Hymenoptera: Apidae) in semiarid Northeastern Brazil using heterospecific microsatellites”, Submetido a revista Journal of Apicultural Research no dia 14 de Fev de 2019.

5.1 Resumo

A *Melipona rufiventris* é parte do grupo de abelhas-sem-ferrão nativas do Brasil, onde a ação humana vem levando a espécie a se enquadrar na lista vermelha de animais ameaçados de extinção. Contudo, existe um grande interesse no desenvolvimento de marcadores genéticos que possibilitará a quantificação da diversidade genética nestas abelhas e, portanto, ajudar em mecanismos de conservação. Dessa forma, este estudo teve por objetivo avaliar o potencial de utilização de *loci* de microssatélites desenvolvidos para as espécies *Melipona fasciculata* e *Melipona subnitida*, a fim de estudar a diversidade genética e a estrutura de populações de *M. rufiventris* no Estado do Piauí. Para estimar o nível de variabilidade e estruturação genética de populações de *M. rufiventris*, analisou-se o DNA extraído de abelhas operárias coletadas em três localidades do Piauí. Foram utilizados nessa pesquisa, 37 *loci* de microssatélites heteroespecíficos onde se obteve nove polimórficos, um monomórfico (Mfsc13) e 27 com resultados insatisfatórios ou não mostrando nenhuma amplificação. Nas três populações foi encontrada heterozigosidade média esperada de 0,43. Valores de PIC de 0,37, revelaram um moderado nível de eficiência dos marcadores. A análise de coordenadas principais (PCoA) e a análise de agrupamento bayesiano mostraram uma clara separação das espécies em dois grupos genéticos distintos de meliponíneos, confirmando assim uma diferença genética molecular significativa entre as três populações. Os resultados revelaram existência de grau intermediário de variabilidade genética, com estreita ligação entre as populações amostradas, porém, merecendo atenção em termos de sua preservação.

Palavras-chave: Conservação, Marcadores genéticos, Meliponíneos, Primers Heterólogos Transferabilidade.

5.2 Introdução

As abelhas-sem-ferrão pertencem à família Apidae na ordem Hymenoptera e são membros da tribo Meliponini, que inclui todos os gêneros de abelhas-sem-ferrão existentes. Essas abelhas eussociais são distribuídas em regiões tropicais e subtropicais do mundo. Existem mais de 500 espécies de abelhas no mundo, descritas em 32 gêneros com mais de 100 espécies novas a serem caracterizadas. Mais de 200 espécies são nativas do Brasil, com a maior diversidade de espécies concentradas na Bacia Amazônica (RASMUSSEN; CAMERON, 2010; MICHENER, 2013).

Elas são polinizadoras da flora nativa local e cultivada (CASTRO et al., 2006; FREITAS & NUNES-SILVA, 2012), portanto, de grande importância ecológica, adaptados para sobreviver em uma variedade de habitats, como diferentes ambientes florestais naturais, Cerrados, zonas úmidas, áreas protegidas, fazendas, casas de madeira, entre outros (CHUTTONG et al., 2016; LEE; DUWAL; LEE, 2016). A intensificação da agricultura para o aumento da produção de alimentos e consequente perda de habitat tem sido reconhecida como o principal fator para o declínio das populações de abelhas-sem-ferrão, especialmente devido ao uso inadequado de fertilizantes e pesticidas (NARJES; LIPPERT, 2016; POTTS et al., 2010).

A espécie *Melipona rufiventris* Lepeletier, 1836 faz parte do grupo das abelhas-sem-ferrão nativas do Brasil, vulgarmente conhecidas como urucu-amarela, cujas populações sofreram grandes perdas nos últimos anos pela crescente destruição da vegetação semiárida nativa e pela coleta predatória de mel. Está atualmente incluído na lista vermelha de fauna brasileira ameaçada (ICMBio, 2016; ROULSTON E GOODELL 2011).

Além dessas limitações impostas pelas perturbações antropogênicas do habitat natural, que reduzem drasticamente muitas populações, estudos mostram que as rainhas virgens de *M. rufiventris* se acasalam com apenas um macho durante um vôo nupcial. Assim, todas as fêmeas compartilharão o mesmo pai, herdando os cromossomos masculinos. A variabilidade genética entre seus descendentes deve ser, portanto, baixa (COLONELLO; HARTFELDER, 2005).

Assim, o conhecimento da variabilidade genética e diversidade em *M. rufiventris* e suas populações, como eles são estruturados, é essencial para uma interpretação correta da dinâmica populacional e também pode ser usado para prever o potencial impacto de perturbações para o habitat e para a formulação de estratégias de manejo adequadas para a conservação de espécies (LOPES et al., 2009).

Marcadores moleculares podem ser usados para medir a diversidade e variabilidade genética dentro e entre populações e, portanto, abordando questões de relevância para a conservação em espécies de abelhas-sem-ferrão ameaçadas.

Os microssatélites, ou sequências de DNA de repetição curta (STR), são exemplos de tais marcadores que ocorrem em todo o genoma de muitos organismos, incluindo insetos (RODERICK, 1996). Suas unidades curtas de repetição de núcleo variam de dois a seis pares de bases de comprimento repetidas em uma matriz geralmente de 100 pb (JARNE & LAGODA, 1996). Os marcadores de microssatélites de DNA são altamente polimórficos e, portanto, adequados para avaliar a variabilidade genética e a diversidade presentes nas populações de *M. rufiventris*. No entanto, o desenvolvimento de marcadores microssatélites específicos para espécies pode ser desafiador devido aos altos custos de desenvolvimento e ao esforço significativo necessário para projetar conjuntos de *primers*. Uma abordagem alternativa viável é simplesmente testar *primers* descritos para microssatélites em espécies intimamente relacionadas, que são denominadas amplificação de microssatélites de espécies cruzadas ou amplificação heteróloga (CARVALHO-ZILSE & KERR, 2006; VIANA et al. 2011).

Este estudo fornece análise da diversidade genética e estrutura das populações de *Melipona rufiventris* no semiárido brasileiro usando *primers* heterólogos projetados para duas outras espécies de abelhas-sem-ferrão, *Melipona fasciculata* e *Melipona subnitida*, um ponto de partida para avaliar mudanças genéticas espaciais e temporais de populações de *Melipona rufiventris* no Brasil.

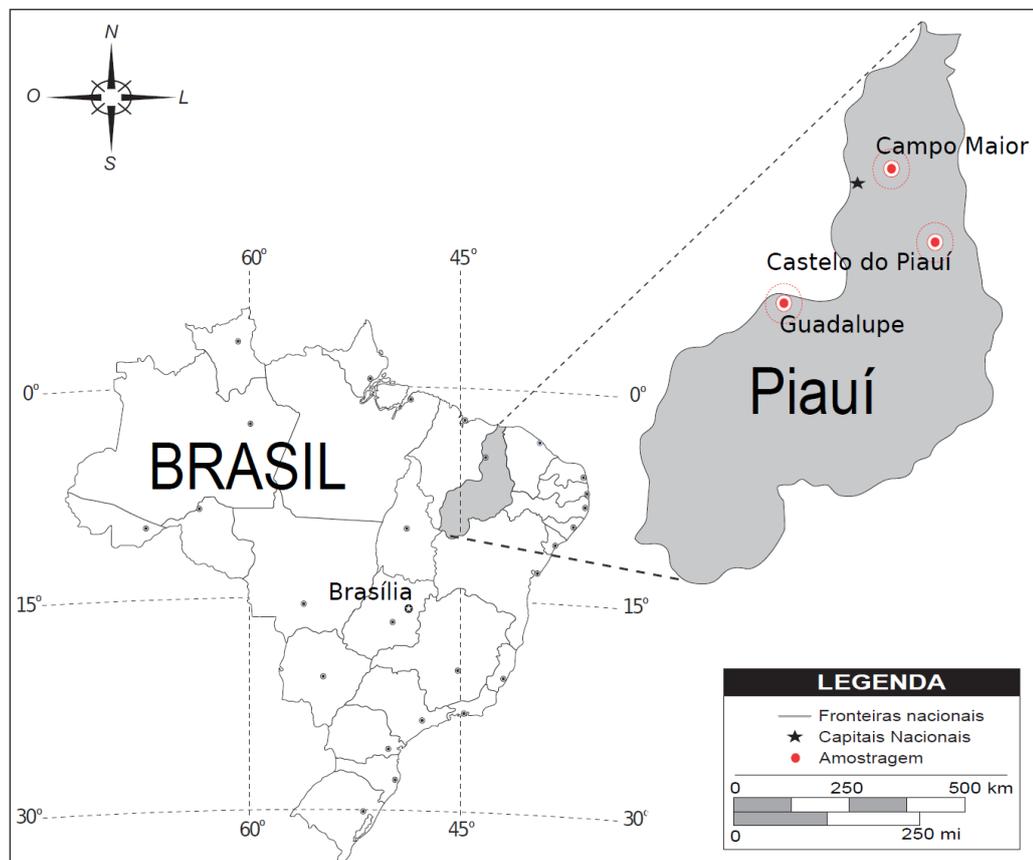
5.3 Material e Métodos

5.3.1 Amostras de abelha e isolamento de DNA genômico

Abelhas operárias foram coletadas aleatoriamente em colônias distribuídas em três localidades, da seguinte maneira: 25 ninhos em Campo Maior (CAM; coordenadas: 4 ° 49'19 "S, 42 ° 09'52" W), sete ninhos em Castelo do Piauí (CAP ; coordenadas: 5 ° 23'15 "S, 41 ° 31'17" W), e seis ninhos em Guadalupe (GUA; coordenadas: 6 ° 47'30 "S, 43 ° 34'14" W), todos no estados do Piauí-Brasil (Figura 1). Todas as amostras foram levadas para o Laboratório de Biologia Molecular e Biotecnologia da Embrapa Meio-Norte e armazenadas a -20 ° C até nova utilização.

O DNA genômico foi extraído do tórax de cada abelha adulta usando o protocolo HotSHOT padrão, conforme descrito por Truett et al. (2000). O tampão de lise alcalina foi aquecido a 95 ° C durante 60 min; depois, as amostras foram arrefecidas a 4 ° C e o pH ajustado a 5 com 40 mM de Tris-HCl. A qualidade (razão 260/280) e a quantidade (ng/ μL) de DNA foram determinadas usando o espectrofotômetro NanoDrop 2000 (Thermo Scientific).

Figura 1. Mapa do Brasil mostrando a região onde os ninhos de *Melipona rufiventris* (Hymenoptera: Apidae) foram amostrados no semiárido.



5.3.2 Teste e genotipagem de marcadores microssatélites heteroespecíficos

Inicialmente, a transferibilidade cruzada de *primers* microssatélites heterólogos foi rastreada no genoma de *M. rufiventris* para encontrar *loci* polimórficos. A amplificação por PCR foi realizada utilizando-se 19 pares de *primers* microssatélites desenvolvidos a partir de *M. subnitida* (SOUZA et al., 2014) e 18 pares de *M. fasciculata* (números de acesso do Genbank KT730150 a KT730167).

As reações de PCR foram realizadas em um volume total de 10 µL contendo aproximadamente 10-50 ng de DNA, 1 × tampão PCR (40 mM Tris-HCl; 100 mM KCl), 0,2-0,25 µM de *primer* (F e R), 2,5 a 3,0 mM de MgCl₂, 0,5-1,0 µM de cada dNTP, 0,25-0,75 U de DNA polimerase Taq da Invitrogen. Todas as amplificações foram realizadas utilizando um termociclador Gradient VERITI (Applied Biosystem). As condições de PCR foram as seguintes: perfil (PCR1) 94°C por 5 min, 30 ciclos de 94°C por 40 seg, temperatura de anelamento (Ta°C) por 30 seg, 72°C por 40 seg e um ciclo de 72°C durante 7 min. Para otimizar a PCR para alguns *primers*, um perfil de PCR alternativo (PCR2) foi usado com base no seguinte: 94 ° C por 1 min, 40 ciclos de 94 ° C por 30 seg, Ta ° C por 30 seg, 72 ° C por 30 seg e um ciclo de 72 ° C por 3 min (Tabela 1).

Os marcadores microssatélites foram rastreados pela coloração por nitrato de prata em géis de poliacrilamida a 6% desnaturados de acordo com BENBOUZA et al. (2006). Um *ladder* de 10 pb (*Life Technologies*) foi usada como marcador de peso molecular. Amplificações de *loci* microssatélites foram consideradas bem sucedidas quando o gel apresentou uma ou duas bandas claras e consistentes, com tamanho de produto próximo ao das espécies originais.

Tabela 1. Marcadores, sequência de *primers* e parâmetros experimentais utilizados para a amplificação de *loci* de microssatélites em *Melipona rufiventris* (*Hymenoptera: Apidae*) e seu número de acesso ao GenBank.

<i>Loco</i>	Sequência dos <i>primers</i> (5'-3')	Tipo de repetição	Ta (°C)	Tamanho do fragmento (bp)	Perfil da reação	Nº de acesso do GeneBank
<i>Msub2</i>	F:GCCCCAAGATGGTATGCCG R:ACGAGGCGGATTCAACGAG	(ACG) ₁₄	60	172-177	PCR ₁	KM494946
<i>Msub3</i>	F: CTCGGCGCACAAATTCGAG R: GGTTATTTTCGCCGGCAAGC	(CGTT) ₁₁	60	132-136	PCR ₁	KM494947
<i>Msub18</i>	F: TCCCGATTTCCACCGATCC R: GCCGACCTCTTCGACGG	(ACG) ₁₈	60	142-160	PCR ₁	KM494953
<i>Msub31</i>	F: TTACCGTCTGTGCTACTGATCC R:TGTCTGTCTGTCTGTCTATCTTTCTG	(AGAT) ₁₄	60	134-150	PCR ₁	KM494956
<i>Msub38</i>	F:AATACTCTGTTTCTTCCAGGGG R: CTGAAATTGCTTTTCGTGCC	(AAAG) ₁₅	60	110-135	PCR ₁	KM494958
<i>Msub46</i>	F: CACTGTTTCTCCAGTTGCTGTC R: GTTTCGTTTCGCGTGATTC	(AAAG) ₁₂	60	113-132	PCR ₁	KM494960
<i>Msub48</i>	F: AAAGAGCGTAGGACTTCCACAG R: CATCCATCTATCCGTACATCCA	(GGAT) ₁₀	58	115-119	PCR ₁	KM494961
<i>Msub51</i>	F: GGCGTTACAAAGGGGAGAA R: AGTTGACAGCGTTTCTACCTC	(AGAA) ₉	60	148-152	PCR ₁	KM494962
<i>Mfsc11</i>	F: GGAAGGACGAGAGAATTCAAGA R: ATAGTCGTTTGTGCGGAGTGTA	(CTT) ₁₃	50	142-168	PCR ₂	KT730153
<i>Mfsc13</i>	F: GCAGTAACGGTAGCAGTGGTG R: ACTCCTTTCTCCTTCTCGGTCT	(ACG) ₁₆	52	157	PCR ₁	KT730154

Ta: Temperatura de Anelamento; Perfil da Reação: (PCR₁ = [94°C-5 min; 30 ciclos (94°C-40 seg; Ta-30 seg; 72°C-40 seg); 72°C-7 min], PCR₂ = [94°C-1 min; 40 ciclos (94°C-30 seg; Ta-30 seg; 72°C-30 seg); 72°C-3 min].

5.4 Análise estatística

Os dados genotipados foram analisados usando o Microchecker 2.2.3 (Van Oosterhout et al., 2004) para testar a presença de alelos nulos ou possíveis inconsistências de pontuação. O número de alelos (N_a), heterozigosidades observadas e esperadas (H_o e H_e) e o conteúdo de informação polimórfica (PIC) foram determinados usando CERVUS 3.0.3 (KALINOWSKI et al., 2007). A riqueza alélica (A_R) foi calculada pelo FSTAT versão 2.9.3.2 (GOUDET, 2002). Testes para Equilíbrio de Hardy-Weinberg (HWE) foi conduzido usando o software baseado na web GENEPOP (RAYMOND & ROUSSET, 1995). Os P valores corrigidos por Bonferroni foram usados para avaliar a significância ($P < 0,05$).

Um modelo de mistura de agrupamento bayesiano foi utilizado para inferir a possível estrutura populacional utilizando o software STRUCTURE v2.3.3 (PRITCHARD et al. 2000). O programa foi criado para 1.000.000 repetições de Monte Carlo da cadeia de Markov (MCMC) após uma queima inicial de 500.000 passos. A estimativa do melhor K foi calculada com base em cinco repetições para cada K (de 1 a 6), conforme descrito por Evanno et al. (2005) usando STRUCTURE HARVESTER v.0.6.92 (EARL & VONHOLDT, 2012). O programa CLUMPP v.1.1.2 (JAKOBSSON & ROSENBERG, 2007) foi utilizado para alinhar as cinco repetições do melhor K. O programa DISTRUCT v.1.1 (NOAH, 2004) foi usado para exibir graficamente os resultados produzidos por CLUMPP. A estrutura populacional também foi analisada por meio da análise de coordenadas principais (PCoA), conforme implementado no GENALEX v.6.5 (PEAKALL & SMOUSE, 2012).

5.5 Resultado e Discussão

Após o teste de PCR e otimização de 37 potenciais pares de *primers* heterólogos de microssatélites desenvolvidos a partir das espécies *M. subnitida* e *M. fasciculata*, apenas nove marcadores (24,3%) revelaram polimorfismo em *M. rufiventris*. Um *primer* monomórfico (Mfsc13) amplificou bandas claras e consistentes, no entanto, apenas uma banda foi detectada nas populações estudadas (Tabela 1). Todos os outros pares de *primers* (73%) produziram resultados insatisfatórios com forte padrão de bandas não específicas ou não apresentaram amplificação, embora *M. subnitida* e *M. fasciculata* sejam congêneres de *M. rufiventris*. Segundo Primmer & Merilä (2000) e Oliveira et al. (2006) o sucesso da amplificação declina de acordo com a distância genética entre os táxons, ou seja, a proximidade filogenética é o principal fator no sucesso da transferência para uso de *primers* heterólogos. No entanto, outros fatores, como o tamanho e a complexidade do genoma e a localização do microssatélite, seja na região codificadora ou não, também influenciam no processo de transferibilidade.

Diversos estudos sobre diversidade genética foram conduzidos em populações de abelhas usando marcadores microssatélites transferidos (SHAO et al., 2004, CARVALHO-ZILSE e KERR, 2006; VIANA et al., 2011; KOSER et al., 2014). Essa alternativa parece ser eficaz quando marcadores microssatélites específicos para espécies não estão disponíveis para as espécies-alvo.

Quatro *loci*, Mfsc11, Msub31, Msub38 e Msub51 desviaram do HWE ($P > 0,05$) (Tabela 2), geralmente devido à presença de alelos nulos ou uma grande quantidade de dados

faltantes. Esses *loci* mostraram excesso heterozigoto e as possíveis causas são as pontuações artefactuais de bandas não específicas e *stutters* (OOSTERHOUT et al., 2004), bem como o pequeno tamanho efetivo da população.

A análise de todo o conjunto de dados para o nível de polimorfismo de cada *locus* revelou que nove microssatélites apresentaram a riqueza alélica (A_R) entre 2 e 6,6, com uma média de 3,2, enquanto os valores de PIC variaram entre 0,12 e 0,67, com uma média de 0,37, portanto, moderadamente informativos. Valores de PIC acima de 0,5 são considerados muito informativos; medianamente informativo entre 0,25 e 0,5 e pouco informativo menor que 0,25 (BOTSTEIN et al., 1980). A heterozigosidade observada (H_O) variou de 0,00 a 0,85 com média de 0,47, enquanto a heterozigosidade esperada (H_E) apresentou valores entre 0,14 e 0,72, com média de 0,43. O *locus* Msub31 foi o mais polimórfico. Msub51 foi o único locus que apresentou evidências de alelos nulos (Tabela 2).

Tabela 2. Variabilidade de nove *loci* de microssatélites e estimativa da diversidade genética em *Melipona rufiventris* (Hymenoptera: Apidae).

Loco	A_R	H_O	H_E	PIC	pHWE	Nulos
Msub2	2,0	0,29	0,25	0,22	1,000	-0,033
Msub3	2,0	0,58	0,42	0,33	0,0334	-0,119
Mfsc11	3,8	0,63	0,51	0,44	0,0006*	-0,087
Msub18	3,8	0,53	0,45	0,39	0,0105	-0,059
Msub31	6,6	0,85	0,72	0,67	0,0000*	-0,079
Msub38	4,9	0,58	0,72	0,65	0,0000*	0,075
Msub46	2,0	0,50	0,42	0,33	0,6226	-0,061
Msub48	2,0	0,24	0,22	0,19	1	-0,024
Msub51	2,0	0,00	0,14	0,12	0,0010*	0,117
Média	3,2	0,47	0,43	0,37	–	–

AR: riqueza alélica; Ho: Heterozigotos observados; He: heterozigotos esperados; PIC: conteúdo de informação polimórfica; pHWE: probabilidade de equilíbrio de Hardy-Weinberg; i: sem informação; *: 5% de significância (correção de Bonferroni <0,005)

A média foi a mesma encontrada por Lopes et al. (2010) para *M. rufiventris* (H_E = 0,43) e superiores aos encontrados em estudos com outras espécies de meliponíneos, tais como: 0,38 para *M. mondury* e 0,35 para *M. mandacaia* (MIRANDA et al., 2012) e 0,105 para *M. mondury* e 0,189 para *M. quadrifasciata* (KOSER et al., 2014). A diversidade genética encontrada em *M. rufiventris*, embora superior à maioria dos estudos supracitados, é considerada baixa, e este reduzido valor está associado a aspectos genéticos, biológicos e ambientais, como é o caso de interferências antropogênicas que contribuem para a redução de fatores genéticos como a variabilidade, ocorridos por meio da fragmentação do habitat e ações

predatórias da coleta de mel, que reduzem o número de colônias em uma dada localidade (LOPES et al., 2010; TAVARES et al., 2013).

Quando o conjunto de dados foi dividido em três populações (Campo Maior, Castelo do Piauí e Guadalupe), observou-se que Campo Maior apresentou a maior média em relação ao número de alelos/população (3,0), variando de 1 a 7, quando comparado para Castelo do Piauí e Guadalupe. O desvio do equilíbrio de Hardy-Weinberg foi observado para os mesmos *loci* da população, Mfsc11, Msub18, Msub31 e Msub38, após a correção de Bonferroni. Para as outras populações, nenhum *locus* mostrou significância para o HWE (Tabela 3).

Tabela 3. Variabilidade de nove *loci* microssatélites e estimativas de diversidade genética em *Melipona rufiventris* (Hymenoptera: Apidae) para as três populações analisadas no Estado do Piauí.

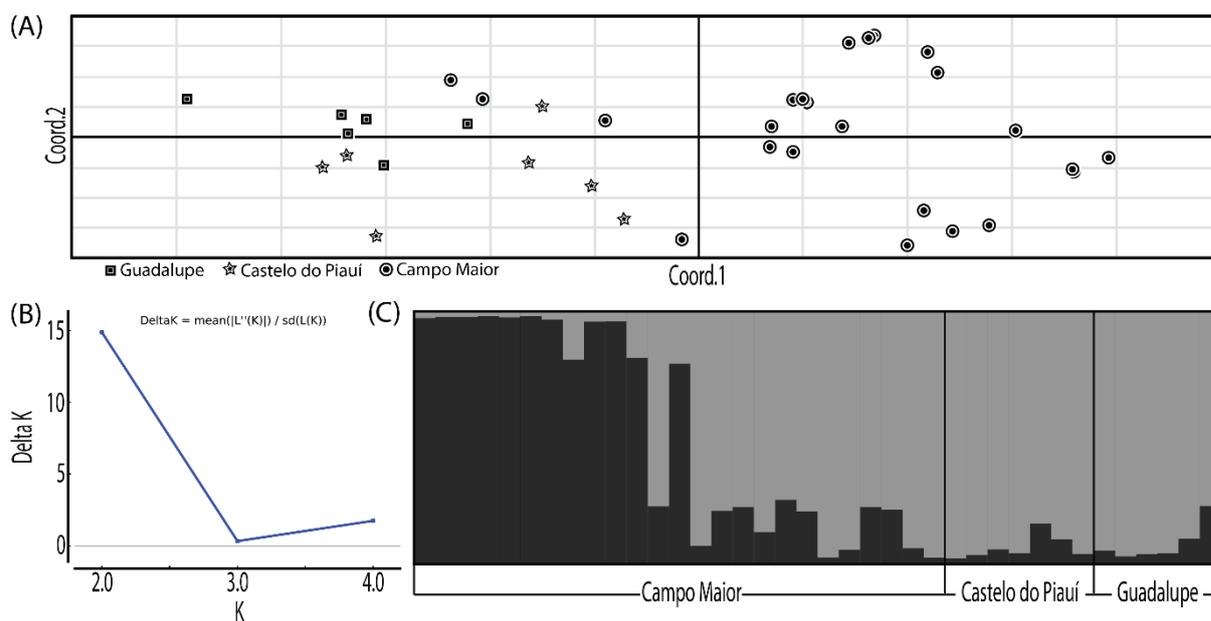
Locos	Campo Maior-PI (n=25)					Castelo do Piauí-PI (n=7)					Guadalupe-PI (n=6)				
	A	H _O	H _E	PIC	pHWE	A	H _O	H _E	PIC	pHWE	A	H _O	H _E	PIC	pHWE
Msub2	2	0,360	0,301	0,252	0,556	2	0,250	0,250	0,195	s.i	1	0,000	0,000	0,000	s.i
Msub3	2	0,640	0,444	0,341	0,057	2	0,333	0,333	0,239	s.i	2	0,333	0,333	0,239	s.i
Mfsc11	2	0,818	0,495	0,367	0,0017*	1	0,000	0,000	0,000	s.i	3	0,667	0,667	0,535	0,309
Msub18	4	0,750	0,557	0,466	0,0005*	1	0,000	0,000	0,000	s.i	1	0,000	0,000	0,000	s.i
Msub31	7	1,000	0,79	0,741	0,000*	2	0,500	0,409	0,305	1,000	2	0,500	0,429	0,239	1,000
Msub38	5	0,421	0,679	0,606	0,0003*	2	1,000	0,600	0,375	0,398	2	1,000	0,571	0,535	0,314
Msub46	2	0,571	0,455	0,346	0,344	1	0,000	0,000	0,000	s.i	0	0,000	0,000	0,000	s.i
Msub48	2	0,240	0,216	0,189	1	2	0,400	0,356	0,269	1,000	1	0,000	0,000	0,000	s.i
Msub51	1	0,000	0	0	0	2	0,000	0,485	0,346	0,030	1	0,000	0,000	0,000	s.i
Média	3	0,533	0,437	0,367	–	1,6	0,275	0,270	0,192	–	1,4	0,277	0,222	0,172	–

A: Número de alelos na população; Ho: heterozigotos observados; He: heterozigotos esperados; PIC: conteúdo de informação polimórfica; pHWE: probabilidade de equilíbrio de Hardy-Weinberg; si: sem informação; *: Significância de 5% (correção de Bonferroni <0,005).

Ao estimar os valores da estatística F , com base nesses marcadores moleculares, o valor de F_{is} encontrado foi de $-0,177$, o que indica um baixo nível de endogamia, enquanto o valor de F_{st} total foi de $0,151$ e o R_{st} de $0,288$ foi considerado alto, sugerindo uma estruturação genética entre as populações analisadas. A análise dos valores de F_{st} considerou a classificação de Nei (1978) onde os valores de $F_{st} < 0,05$ são considerados baixos, entre $0,05$ e $0,15$ moderados e os valores acima de $0,15$ são altos, indicando baixa, média e alta estrutura populacional, respectivamente.

A análise de coordenadas principais (PCoA) mostrou uma separação das espécies em dois grupos com algum grau de intersecção, confirmando assim uma diferença genética molecular significativa entre as três populações (Figura 2A). Na análise bayesiana, o valor ótimo determinou ser $K = 2$, o que revela que exista provavelmente dois agrupamentos genéticos distintos (Figura 2B-C).

Figura 2. Gráfico de dispersão da análise de coordenadas principais (PCoA) usando locos microssatélites de *M. rufiventris*; (B) Determinação do melhor número de *clusters* a partir da análise STRUCTURE; (C) Gráfico de barras da estrutura da população Inferida usando o programa baseado em modelo de mistura de agrupamento bayesiano STRUCTURE ($K = 2$). Indivíduos são representados por cada barra.



5.6 Conclusão

Pelo que foi apresentado nos resultados, a transferência dos marcadores microssatélites desenvolvidos para *Melipona subnitida* demonstraram maior sucesso para estudos populacionais com a espécie *M. rufiventris*. Dessa forma, foi possível detectar a presença de dois grupos genéticos distintos, sendo que a maior diversidade genética foi observada em Campo Maior-PI, com o *locus* Msub31 apresentando-se como o mais polimórfico.

5.7 Referências

- BENBOUZA H. et al. Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels. **Biotechnology, Agronomy, Society and Environment**. v.10, n.2, p.77-81, 2006.
- BOTSTEIN, D., WHITE, R. L. & SKOLNICK, M. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**. v.32, p.314-331, 1980.
- CARVALHO-ZILSE G. A. & KERR W.E. Utilização de marcadores microssatélites para estudos populacionais em *Melipona scutellaris* (Apidae, Meliponini). **Magistra**, v.18, p.213-220, 2006.
- CASTRO M. S., KOEDAM D., & CONTRERA F. A. L., “**Stingless bees,**” in **Bees as Pollinators in Brazil: Assessing the Status and Suggesting Best Practices**, V. L. Imperatriz-Fonseca, Ed., p. 112, Holos Editora, Ribeirão Preto, Brazil, 2006.
- COLONELLO, NA & HARTFELDER, K. She's my girl - male accessory gland products and their function in the reproductive biology of social bees. **Apidologie**, v.36, p.231-244, 2005.
- CHUTTONG, B. et al. Physicochemical profiles of stingless bee (Apidae: Meliponini) honey from South East Asia (Thailand). **Food Chemistry**, v. 192, p. 149–155, 2016.
- EARL, D.A. & VONHOLDT, B.M. Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**, v.4, p.359–361, 2012.
- EVANNO, G., REGNAUT, S. & GOUDET, J. Detecting the number of clusters of individuals using the software structure: a simulation study. **Molecule Ecology**, n.14, p.2611–2620, 2005.
- FREITAS B. M. & NUNES-SILVA P., “**Polinização agrícola e sua importância no Brasil,**” in **Polinizadores no Brasil**. Contribuição e Perspectivas Para a Biodiversidade, Uso Sustentável, Conservação e Serviços Ambientais, V. L. Imperatriz-Fonseca, Ed., 2012.
- GOUDET J . **FSTAT**: a program to estimate and test gene diversities and fixation indices. Version 2.9.3.2. 2002. Available at <http://www.unil.ch/izea/software/fstat.html>.

ICMBIO-MMA (Instituto Chico Mendes, Ministério do Meio Ambiente) (2016) Executive summary of Brazil red book of threatened species of fauna. Instituto Chico Mendes de Conservação da Biodiversidade. Disponível em: http://www.icmbio.gov.br/portal/images/stories/comunicacao/publicacoes/publicacoesdiversas/dcom_sumario_executivo_livro_vermelho_ed_2016.pdf. Acesso 19 de dezembro de 2016.

JAKOBSSON, M. & ROSENBERG, N.A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. **Bioinformatics**, v.23, p.1801–1806, 2007.

JARNE P, LAGODA P.J.L. Microsatellites, from molecules to populations and back. **Trends Ecol Evol**, v.11, p.424–429, 1996.

KALINOWSKI, S. T., TAPER, M. L. & MARSHAL, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. **Molecular Ecology**. 16, p.1099–1106, 2007.

KOSER, J.R., FRANCISCO, F.O., MORETTO, G. Genetic Variability of Stingless Bees *Melipona mondury* Smith and *Melipona quadrifasciata* Lepeletier (Hymenoptera: Apidae) from a Meliponary. **Sociobiology**, v.61, n.3, p.313-314, 2014.

LEE, Y.; KIM, S.; LEE, W. Notes on the subfamily greenideinae (Hemiptera: Aplanidae) from Laos, with descriptions of two new species. **Journal of Pacific Entomology**, v.19, p.735-742, 2016.

LOPES, D. M. et al. Microsatellite loci for the stingless bee *Melipona rufiventris* (Hymenoptera: Apidae). **Molecular Ecology Resources**, v. 9, n. 3, p. 923–925, 2009.

LOPES, D. M. et al. Comparative study on the use of specific and heterologous microsatellite primers in the stingless bees *Melipona rufiventris* and *M. mondury* (Hymenoptera, Apidae). **Genetics and Molecular Biology**, v. 33, n. 2, p. 390–393, 2010.

MICHENER, C D. **The Meliponini**. Chapter 1, In: Pot-Honey A legacy of stingless bees. Vit, P et al.(eds.) Springer Science + Business, New York. pp. 3-17, 2013.

MIRANDA, E.A. et al. Genetic Diversity of *Melipona mandacaia* SMITH 1863 (Hymenoptera, Apidae), an Endemic Bee Species from Brazilian Caatinga, Using ISSR. **Psyche** p.1-6, 2012.

NARJES, M. E. & LIPPERT, C. Longan fruit farmers' demand for policies aimed at conserving native pollinating bees in Northern Thailand. **Ecosystem Services**, v. 18, n. June 2011, p. 58–67, 2016.

NEI, M. Estimation of average heterozygosity and genetic distance from small number of individuals. **Genetics**, v. 89, p.583-590, 1978.

NOAH, A. R. Distruct: a program for the graphical display of population structure. **Molecular Ecology**, v. 4, p.137–138, 2004.

- OLIVEIRA, E.J. et al. Origin, Evolution and Genome distribution of microsatellites. **Genetics and Molecular Biology**, v.29, n.2, p.294-307, 2006.
- OOSTERHOUT C.V. et al. Micro-Checker: software for identification and correcting genotyping errors in microsatellites data. **Molecular Ecology**, v.4, p.535- 538, 2004.
- PEAKALL, R., S. GenAlEx 6.5: genetic analysis in Excel Population genetic software for teaching and research, an update, **Bioinformatics**, v.28, p.2537-2539, 2012.
- POTTS, S.G. et al. Global pollinator declines: trends, impacts and drivers. **Trends in Ecology and Evolution**, v.25, n.6, p. 345–353, 2010.
- PRIMMER, C.R.; MERILÄ, J. A low rate of cross-species microsatellite amplification success in Ranid frogs. **Conservation Genetics**, v.3, p. 445-449, 2000.
- PRITCHARD J.K., STEPHENS, M., DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**. v.155, p.945-959, 2000.
- RASMUSSEN, C. & CAMERON, S.A. Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. **Biological Journal of the Linnean Society**., v.99, p. 206–232, 2010.
- RAYMOND M, ROUSSET F. **GENEPOP (version 1.2)** Population genetics software for exact tests, 1995.
- ROULSTON T.H. & GOODELL K. The role of resources and risks in regulating wild bee populations. **Annual Review of Entomology**, v.56, p.293–312, 2011.
- RODERICK, G.K. Geographic structure of insect populations: gene flow, phylogeography, and their uses. **Annual Review of Entomology**, v. 41, p.325–352, 1996.
- SHAO, Z.Y.; MAO, H.X.; FU, W.J.; ONO, M.; WANG, D.S.; BONIZZOMI, M.; ZHANG, Y.P. Genetic structure of Asian Population of *Bombus ignitus* (Hymenoptera: Apidae). **Journal of Heredity**, v.95, n.1, p.46-52, 2004.
- SOUZA, I. G. B. et al. Isolation and characterization of 23 microsatellite loci in the stingless bee *Melipona subnitida* using next-generation sequencing. **Conservation Genetics Resources**, v. 7, n. 1, p. 239–241, 2014.
- TAVARES, M.G. et al. “Genetic divergence between *Melipona quadrifasciata* Lepelletier (Hymenoptera, Apidae) populations.” **Genetics and Molecular Biology**, v.36, n.1, p.111–117, 2013.
- TRUETT, G. E. et al. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and Tris (HotSHOT). **BioTechniques** v.29, p.52-54, 2000.
- VIANA, M. V. C. et al. Transferability of microsatellite primers developed for stingless bees to four other species of the genus *Melipona*. **Genetics and Molecular Research**, v. 10, n. 4, p. 3942–3947, 2011.

VIANA, M.V.C. et al. Transferability of microsatellite primers developed for stingless bees to four other species of the genus *Melipona*. **Genetics and Molecular Research** v.10, n.4, p.3942-3947, 2011.

CAPÍTULO 6

Desenvolvimento de marcadores microssatélite por meio do sequenciamento Illumina MiSeq para uso em genética de populações de abelha-sem-ferrão urucu-amarela (*Melipona rufiventris*, Hymenoptera, Meliponini)

*Artigo: “Development of microsatellite markers by Illumina MiSeq Sequencing for use in population genetics of the stingless bee urucu-amarela (*Melipona rufiventris* group, Hymenoptera, Meliponini). Submetido a revista Scientific Reports no dia 19 de Dez. de 2018.*

6.1 Resumo

A espécie *Melipona rufiventris* Lepeletier 1836 é uma abelha-sem-ferrão nativa que faz parte de um complexo de espécies conhecido como “grupo rufiventris”, que são abelhas operárias de reconhecimento muito difícil. As populações deste grupo estão enfrentando um grave declínio, levando ao risco de extinção local e, portanto, sua conservação deve ser tratada como uma questão importante. Este estudo descreve o primeiro conjunto de marcadores microssatélites tri- e tetranucleotídeos usando a tecnologia de sequenciamento de Nova Geração para uso na identificação da diversidade genética e estrutura populacional do grupo "*rufiventris*". Um total de 16 *loci* de microssatélites exibiram polimorfismo. Uma análise de todo o conjunto de dados ($n = 50$) detectou 63 alelos em todos os *loci*, variando de 2 a 7 com uma média de 3,9 alelos / *locus*. A análise da diversidade genética revelou valores elevados para estimativas de diferenciação populacional ($F_{ST} = 0,252$, $R_{ST} = 0,317$ e $D_{EST} = 0,284$) nos biomas Mata Atlântica, Cerrado e Caatinga. Uma evidência adicional para a divergência genética entre as populações também foi apresentada no grupo "*rufiventris*", e que estas devem ser tratadas como unidades de conservação separadas ou até como espécies separadas. Estes marcadores microssatélites demonstraram um forte potencial para avaliar a diversidade populacional neste grupo de abelhas-sem-ferrão ameaçado.

Palavras-chave: Conservação, Diferenciação Genética, Genética de Populações Marcadores Moleculares, Plataforma Illumina, Tecnologia SNG.

6.2 Introdução

As abelhas são consideradas os principais polinizadores em ambientes naturais e agrícolas. Este serviço ecossistêmico é essencial para sustentar a diversidade de espécies e a produção de culturas. A tribo Meliponini pertence a esse grupo de insetos eusociais, sendo caracterizados por seus ferrões atrofiados, e encontrados na maioria das regiões tropicais ou subtropicais do mundo (SILVEIRA; MELO; ALMEIDA, 2002; VOSSLER, 2012).

A espécie *Melipona rufiventris* Lepeletier 1836 é uma abelha-sem-ferrão nativa, distribuída principalmente nos biomas Mata Atlântica, Caatinga e Cerrado do Brasil (MELO, 2003; TAVARES et al., 2008). É uma espécie politípica, parte de um grupo conhecido como 'grupo *rufiventris*', devido à sua morfologia similar, que é de difícil reconhecimento em abelhas operárias (COSTA et al., 2005) e, portanto, mostra a necessidade de dados biológicos adicionais para uma melhor identificação do status taxonômico dessas espécies. Este grupo nativo de abelhas-sem-ferrão tem capacidade de voo limitada e, quando associado a barreiras geográficas, pode acarretar baixas taxas de dispersão entre as populações. Além disso, o acasalamento da rainha de *M. rufiventris* ocorre com um único macho e, como resultado, pode-se gerar variabilidade genética intracolônial consideravelmente baixa (LOPES et al., 2009).

Algumas das abelhas endêmicas brasileiras estão na lista nacional oficial de espécies da fauna ameaçadas de extinção (ICMBIO, 2016). A espécie *M. rufiventris* é uma dessas abelhas que se torna muito rara e está listada como ameaçada de extinção (EN), devido aos efeitos adversos do desmatamento, perda de habitat e coleta predatória de mel (CAMPOS et al., 1988; COSTA et al., 2005; TAVARES et al., 2007). Portanto, sua conservação deve ser tratada como uma questão importante, visando o manejo adequado desse importante recurso genético e a implementação de estratégias de conservação eficientes, assegurando serviços de polinização tanto na agricultura comercial quanto nos ecossistemas naturais (GARIBALDI et al., 2013).

Os marcadores microssatélites (SSR) têm emergido como uma das ferramentas mais eficazes para determinar a divergência genética entre populações e a estimativa da estrutura populacional e da diversidade genética em diferentes táxons (MARIETTE et al., 2002; MAY-ITZÁ et al., 2010; LOPES et al., 2010), o que é essencial para o desenvolvimento de estratégias eficientes de conservação.

O sequenciamento de Nova Geração tem sido utilizado recentemente para o isolamento e desenvolvimento de marcadores microssatélites (SSR) em espécies de *Melipona*

(SOUZA et. al., 2015; SILVA et al. 2018). A tecnologia tem a vantagem sobre os métodos tradicionais usados para a descoberta de marcadores SSR (ZANE & PATARNELLO et. al., 2002), pois gera grandes quantidades de dados de sequenciamento e, conseqüentemente, grande número de marcadores em uma abordagem mais simples e rápida, evitando a construção de bibliotecas de DNA enriquecidas com microssatélites, procedimento desgastante e trabalhoso (ABDELKRIM et al., 2009).

Portanto, este estudo descreve o primeiro conjunto de marcadores microssatélites tri- e tetranucleotídeos, desenvolvidos por meio de sequenciamento de alto rendimento da plataforma Illumina, para uso na identificação da diversidade genética e estrutura populacional no grupo *rufiventris*.

6.3 Material e Métodos

6.3.1 Amostragem e Extração de DNA

Um total de 50 abelhas operárias foram coletadas aleatoriamente nos estados de Minas Gerais (n = 20), Goiás (n = 15) e Piauí (n = 15), Brasil (Figura 1, Tabela 1). O DNA genômico total foi extraído do tórax de cada abelha usando o Kit Promega Extract ME Genomic DNA (DNA-Gdansk), seguindo o protocolo do fabricante para o tecido animal. Amostras de DNA foram submetidas à eletroforese em gel de agarose a 0,8 % para testar a quantidade e a qualidade geral do rendimento de DNA.

Figura 1. Locais de amostragem de espécimes do "grupo *rufiventris*" nos biomas continentais brasileiros. Mata Atlântica (□), Caatinga (■) e Cerrado (▲). Este mapa do Brasil utilizado para a elaboração da Figura 1, e o shapefile para gerá-lo, foi extraído do banco de dados de domínio público do Instituto Brasileiro de Geografia e Estatística - IBGE (<http://mapas.ibge.gov.br/bases-e-referenciais.html>), e foi modificado usando o Inkscape 0.91 (<https://inkscape.org/en/>).

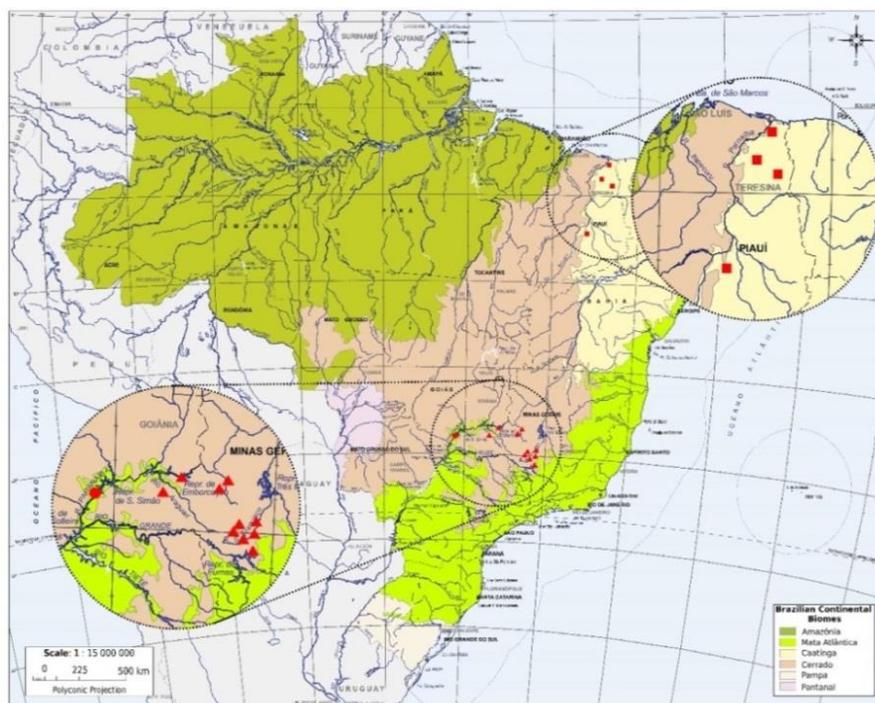


Tabela 1. Localização e número de colônias do grupo *rufiventris* amostradas nos biomas continentais brasileiros.

Localidades	Estado	Bioma	Latitude/Longitude	Número de colônias
Guimarânia	Minas Gerais	Cerrado	-18°50'/-46°47'	4
Araguari	Minas Gerais	Cerrado	-18°38'/-48°11'	3
Patos de Minas	Minas Gerais	Cerrado	-18°34'/-46°31'	1
Uberlândia	Minas Gerais	Cerrado	-18°55'/-48°16'	1
Águas do Paraíso	Minas Gerais	Cerrado	-20°19'/-45°27'	1
Iguatama	Minas Gerais	Cerrado	-20°10'/-45°42'	2
Tapiraí	Minas Gerais	Cerrado	-19°53'/-46°01'	1
Bambuí	Minas Gerais	Cerrado	-20°00'/-45°58'	1
Arcos	Minas Gerais	Cerrado	-20°16'/-45°32'	4
Lagoa da Prata	Minas Gerais	Cerrado	-20°01'/-45°32'	1
Córrego Danta	Minas Gerais	Cerrado	-19°49'/-45°54'	1
Caçu	Goiás	Mata Atlântica	-18°33'/-51°07'	15
Murici dos Portelas	Piauí	Caatinga	-03°19'/-42°05'	4
Campo Maior	Piauí	Caatinga	-04°49'/-42°10'	5
Castelo do Piauí	Piauí	Caatinga	-05°19'/-41°33'	3
Guadalupe	Piauí	Caatinga	-06°47'/-43°34'	3
Total	-	-	-	50

6.3.2 Preparação de biblioteca e sequenciamento de alto rendimento

Utilizou-se uma única amostra de DNA individual de aproximadamente 1 µg para preparar a biblioteca genômica para sequenciamento seguindo o protocolo padrão do kit de preparação da biblioteca Illumina Nextera XT (Illumina Inc., San Diego, CA). A biblioteca de DNA foi sequenciada utilizando um sequenciador MiSeq Benchtop (Illumina Inc., San Diego, CA), tendo como alvo fragmentos de 500 pb com *reads* de 2 x 250 pb na configuração do sequenciamento de extremidade emparelhada. As versões da Illumina foram combinadas para produzir *contigs* com o CLC *Genomics Workbench* 7.0.4 (Qiagen).

6.3.3 Mineração SSR e desenho de primer

Contigs foram adicionados diretamente no Msatcommander 0.8.2 (FAIRCLOTH, 2008) para detecção de possíveis *locus* microssatélites com pelo menos seis repetições para tri- e tetranucleotídeos. *Primers*, *forward* e *reverse*, foram projetados para cada sequência de repetição curta em tandem em suas regiões flanqueadoras. Repetições longas de mononucleídeos [(A) *n*, (G) *n*, (C) *n*, (T) *n*, *n* > 5] entre locais de emparelhamento de iniciadores foram evitadas para o desenvolvimento de marcadores. O desenho do *primer* foi realizado com o software *Primer 3* (ROZEN & SKALETSKY, 2000).

6.3.4 Amplificação por PCR e validação de SSRs selecionados

O DNA genômico dos 50 indivíduos descritos na Tabela 1 foram utilizados para validar todos os pares de *primers* por meio de reações em cadeia de polimerase (PCRs) e para obter uma linha base de informações da frequência alélica dos marcadores. As reações foram realizadas num volume total de 10 µl contendo pelo menos 20 ng de DNA genômico, com tampão 1 X, MgCl₂ 2 a 3 mM, 1,0 mM do *mix* de dNTP, 0,25 mM de cada *primer* e 0,75 unidades de Taq DNA-polimerase (*Thermo Scientific Inc.*). Todas as reações foram corridas em um Termociclador Veriti de 96-*well* (*Applied Biosystems*) utilizando na programação, o perfil de temperaturas de PCR indicado na Tabela 2. Os produtos amplificados foram detectados em géis de poliacrilamida desnaturados a 6% sob coloração em nitrato de prata.

6.3.5 Análise estatística

O *software* MicroChecker 2.2.3 (VAN OOSTERHOUT et al., 2004) foi usado para detectar alelos nulos e problemas de genotipagem. A heterozigosidade observada (H_O), heterozigosidade esperada (H_E) e riqueza alélica (AR) foram calculadas utilizando o software Fstat v.2.9.3.2 (GOUDET, 1995). O Cervus 3.0 (KALINOWSKI et al., 2007) foi usado para

calcular o conteúdo de informação polimórfica (PIC) de cada *locus* de microssatélites. Um modelo bayesiano de agrupamento de mistura implementado no *software* Structure 2.3.4 (PRITCHARD; STEPHENS & DONNELLY, 2000) foi usado para identificar grupos geneticamente homogêneos dentro dos dados genotipados. A estimativa do melhor K, número de grupos que mais se ajustou aos dados, foi calculado baseado em cinco replicações para cada K (de 1 a 11), conforme descrito por Evanno et al. (2005) utilizando *Structure* Harvester v.0.6.92 (EARL & VONHOLDT 2012). O programa foi organizado para 1.000.000 repetições de Cadeia de Markov de Monte Carlo (MCMC) seguido por descarte de 500.000 iterações. A estrutura populacional também foi analisada por meio da análise de coordenadas principais (PCoA), R_{ST} (SLATKIN, 1995), uma medida de diferenciação genética análogo ao F_{ST} , e o estimador D_{EST} de diferenciação real (JOST, 2008), como implementado no GenAIEx v.6.5 (PEAKALL & SMOUSE 2012).

Tabela 2. Características e amplificação de 16 marcadores microssatélites polimórficos desenvolvidos para *Melipona rufiventris* (Hymenoptera: Apidae) (N = 50).

Locus	Sequência dos primers (5'-3')	Tipo de motivo	PCR perfil	Ta (°C)	Tamanho (bp)	N _A	Nº de acesso no GenBank
<i>Mruf 1</i>	F: CAGTCGCCCAAGTAAATACG R: CTTATGAAACGAACCACAAGCC	(AGCC) ₈	PCR _{STD1}	54	142-166	6	MK133898
<i>Mruf 4</i>	F: GTTACGTTGGCAGGAGAGC R: AACTTGATTATTAGCGCGTGA	(TGCT) ₈	PCR _{STD1}	52	146-166	6	MK133901
<i>Mruf 5</i>	F: AGTGAAATCCGAGAGTGGGTT R: TCTCCACCGTCTTTTGTCTT	(AGAA) ₉	PCR _{STD1}	51	154-170	4	MK133902
<i>Mruf 6</i>	F: GTGCCTCGTTACCACCTTCTC R: TTAAAAGTGCACGGGGA	(CT) ₉ N ₁₃ (GTCT) ₆	PCR _{STD1}	54	106-110	2	MK133903
<i>Mruf 8</i>	F: CATCGTCCTCCCGTGAATATAG R: TGCTTTTCCTTCCACGACC	(GCTG) ₁₀	PCR _{STD1}	50	102-122	5	MK133905
<i>Mruf 9</i>	F: TATACTTACGAGAGCGCACGAG R: TATTTTCTACGGTCCCACCTCG	(AACG) ₉	PCR _{STD1}	55	118-140	7	MK133906
<i>Mruf 11</i>	F: TGTGACGTTTTGGACGTAATTC R: CGCTTCCTTTGATCTCTCGAT	(TTTA) ₁₃	PCR _{STD2}	48	112-124	4	MK133908
<i>Mruf 13</i>	F: GCTAGGGGACCTTCTTCTTCTT R: GTGATAAGGCGGAGTGTAATC	(TTC) ₁₂	PCR _{STD1}	50	100-106	2	MK133910
<i>Mruf 15</i>	F: AAGTGGGGAGCTAATAAGGGAG R: TGCAGAGGAGCAGTACAGAGAG	(TTC) ₁₂	PCR _{STD3}	51	157-172	4	MK133912
<i>Mruf 17</i>	F: GTCGAGGACGACTACACAACAA R: CTCACCGCACACAGGGTT	(ACG) ₁₅	PCR _{STD1}	52	161-173	3	MK133914
<i>Mruf 18</i>	F: AAGCGGACAAGCAGATCACT R: ACTGTATGTCGTTCCCTCGTCCT	(AAT) ₉	PCR _{STD1}	52	108-126	6	MK133915
<i>Mruf 19</i>	F: CACTGTCTTGTATTTAGACGCAATC R: GGTCGGGGACTTTAGTGTTTTA	(TTG) ₁₄	PCR _{STD1}	55	125-134	2	MK133916
<i>Mruf 20</i>	F: CGGGTAGTATTAAGGGAATTGA R: TGTGTCAGGAAGAAAAGCAA	(ATT) ₉	PCR _{STD1}	54	184-193	2	MK133917
<i>Mruf 21</i>	F: CTACCGAGAGTAGCGACGACAT R: TCAGTTCTCAATGTTGCAGGC	(ACG) ₁₁	PCR _{STD1}	54	150-162	2	MK133918
<i>Mruf 22</i>	F: CGACTTCGCGTGGTGCTAC R: AGAGGTTTCGGCGGCTTC	(ACG) ₉	PCR _{STD1}	54	125-131	2	MK133919
<i>Mruf 25</i>	F: AACAAGAGCAAAGTAACGACGA R: GAAGGAACAAGTCGAAACCAAC	(ACG) ₇ N ₁₁ (GAA) ₁₃	PCR _{STD3}	51	137-155	6	MK133922

Ta: temperatura de anelamento; NA: Número de alelos; PCR_{STD1}: (94 ° C-5 min; 30 ciclos (94 ° C-40 seg; Ta-30 seg; 72 ° C-40 seg); 72 ° C-7 min); PCR_{STD2}: (94 ° C-5 min; 45 ciclos (94 ° C-40 seg; Ta-30 seg; 72 ° C-40 seg); 72 ° C-7 min); PCR_{STD3}: (94 ° C-5 min; 40 ciclos (94 ° C-40 seg; Ta-30 seg; 72 ° C-40 seg); 72 ° C-7 min).

6.4 Resultados e Discussão

6.4.1 Montagem de sequência e mineração SSR

O sequenciamento de Illumina MiSeq resultou em 54.555.929 *reads*, que foram montadas em um total de 137.313 sequências de *contigs*. Os *reads* mínimos e máximos foram de 200 e 13.505 bases, respectivamente, com um tamanho médio de 397 bases. O programa Msatcommander 0.8.2 identificou 9.745 (7,1%) *contigs* com *loci* de microsatélites consistindo de SSRs di- a hexa-nucleotídeos com pelo menos seis repetições.

Para facilidade de imagem e pontuação, apenas tri- e tetranucleotídeos foram examinados. Destes potenciais marcadores microsatélites, 25 *loci* foram selecionados aleatoriamente para o delineamento e validação de *primers* em *M. rufiventris*. Todas as sequências de microsatélites isoladas e validadas neste estudo foram depositadas no NCBI no banco de dados do GenBank sob os números de acesso MK133898-MK133922.

6.4.2 Validação SSR

Entre todos os 25 pares de *primers* projetados, 16 *loci* (64%) foram amplificados com sucesso, produzindo bandas de PCR consistentes e específicas do tamanho esperado. Possíveis erros de complementação ou mutações em ambos os sítios de anelamento dos *primers* em cada *locus* poderiam ter resultado nas falhas de amplificação.

Análise de todo o conjunto de dados ($n = 50$) revelou 63 alelos em todos os 16 *loci*, variando de 2 (Mruf 6, 13, 19, 20, 21, 22) a 7 (Mruf 9) com uma média de 3,9 alelos/*locus* (Tabela 2). Esse resultado corrobora com o encontrado em pesquisas anteriores com espécies intimamente relacionadas como *Melipona seminigra* (FRANCINI et al., 2009) ($N_A = 3,7$), *Melipona interrupta manaosensis* (FRANCINI et al., 2010) ($N_A = 2,8$), *Melipona mondury* (LOPES et al., 2010) ($N_A = 3,0$), *Melipona yucatanica* (MAY-ITZÁ et al. 2010) ($N_A = 2,6$ a 3,6) e *Melipona fasciculata* (SILVA et al. 2018) ($N_A = 3,9$). Além disso, esses novos *loci* microsatélites foram mais polimórficos do que os marcadores SSR desenvolvidos a partir de um estudo anterior por métodos tradicionais de clonagem nas mesmas espécies (LOPES et al., 2009) ($N_A = 2,6$).

Todos os *loci* foram observados segregando independentemente um do outro, sem evidência de desequilíbrio de ligação ($P > 0,05$). O polimorfismo dos *loci* de microsatélites também foi avaliado separadamente em amostras coletadas de três biomas distintos, por riqueza alélica (A_R), heterozigosidades (H) e conteúdo de informação polimórfica (PIC) (Tabela 3). A A_R média, as heterozigosidades observadas e esperadas (H_O e H_E) e PIC foram

baixas a moderadas, respectivamente 2,4, 0,514, 0,425, 0,345 para Minas Gerais no bioma Cerrado, 2,7, 0,547, 0,479, 0,400 para Goiás na Mata Atlântica, e 2,7, 0,563, 0,502, 0,412 para o Piauí na Caatinga. Apesar do fato de que baixos níveis de diversidade genética podem ocorrer em Hymenoptera sociais em comparação com outros insetos (METCALF et al., 1975; GRAUR, 1985), a redução populacional recente também pode ter contribuído para esse cenário alarmante, diminuindo a variabilidade genética nessas populações do ‘grupo *rufiventris*’ (COSTA et al., 2005). As estimativas de heterozigosidade foram de magnitude semelhante quando comparadas àquelas encontradas para *Melipona fasciculata* (SILVA et al., 2018), mas um pouco menor para *M. subnitida* (SOUZA et al., 2015).

De 48 combinações (16 *loci* e 3 populações) de todas as amostras, 10 desvios significativos da HWE após a correção sequencial de Bonferroni (pvalor <0,0031) foram detectados, principalmente para um excesso de heterozigotos, o que poderia ser atribuído a gargalos populacionais (CORNUET & LUIKART 1996), provavelmente associado a desmatamento extensivo para novos usos da terra e expansão urbana (POTTS et al., 2010). Além disso, alelos nulos, endogamia ou seleção a favor ou contra um certo alelo também poderiam ter desempenhado um papel nas saídas do HWE (SELKOE & TOONEN 2006). A evidência de alelos nulos foi encontrada apenas para Mruf1 em Goiás e Piauí, nas frequências estimadas de 0,216 e 0,218, respectivamente, e para os *loci* Mruf9 (0,234) e Mruf18 (0,261) em Minas Gerais.

6.4.3 Divergência genética entre populações no grupo ‘*rufiventris*’

Em nosso estudo, estimativas de F_{ST} (0,252), R_{ST} (0,317) e D_{EST} (0,284) obtidas de *locus* de microssatélites sugerem a existência de diferenciação genética entre populações/colônias no grupo ‘*rufiventris*’ (Tabela 3). Embora essas estimativas sejam obtidas através de diferentes métodos computacionais, elas foram de magnitude similar, indicando diferenças genéticas significativas entre as amostras coletadas nos três biomas contrastantes. Valores de $F_{ST} > 0,25$ indicam diferenciação genética muito alta, que pode ser causada por seleção natural, deriva genética e / ou migração (WRIGHT, 1978). Nossos resultados estão em consonância com observações anteriores em *M. rufiventris* (TAVARES et al., 2007) ($F_{ST} = 0,250$) e em *M. beecheii* (QUEZADA-EUAN et al. 2007) ($F_{ST} = 0,280$). Essa alta diferenciação pode refletir isolamento parcial com fluxo de genes reduzido entre colônias localizadas em diferentes biomas, resultando em alta endogamia dentro das localidades (NASCIMENTO, 2008). Uma baixa taxa de migração entre populações de *M. rufiventris*,

0,055 abelhas por geração, foi detectada a partir de dados de microssatélites coletados em Minas Gerais (TAVARES et al., 2007). Além disso, pesquisas anteriores sugerem que a distância máxima de vôo por operárias de *M. rufiventris* é de aproximadamente 1000 m (FREITAS et al., 2015), o que pode atuar como restrição ao fluxo gênico e, assim, contribuir para o isolamento parcial.

O agrupamento baseado em modelo de mistura (análise de estrutura) reconheceu três populações genéticas distintas. A estatística *ad hoc* ΔK , usada para inferir o número real de clusters (K) que retém a estrutura principal do conjunto de dados, revelou o melhor K no terceiro nível de separação da subpopulação (K = 3), com um sinal forte ($\Delta K > 400$), como visto na Figura 2A. Esses agrupamentos representam um nível de organização biologicamente significativo nas colônias pesquisadas neste estudo, considerando três biomas distintos (Fig. 2B). A alta estrutura genética da população sugere baixos níveis de fluxo gênico entre colônias de biomas distintos que podem não ser suficientes para contrabalançar a deriva genética (BEEBEE, 2004).

A análise de coordenadas principais (PCoA) da distância genotípica codominante foi conduzida para obter informações adicionais sobre as relações entre as populações do grupo "*rufiventris*". A análise corroborou as atribuições populacionais inferidas por Structure e os cálculos F_{ST} pareados (Fig. 2C). As colônias de Goiás, no bioma Mata Atlântica, divergem ao longo dos dois eixos PCoA, enquanto as colônias mineiras e piauienses, embora claramente distintas, apresentaram muito pouca sobreposição na distribuição dos indivíduos. Todas as colônias, no entanto, foram claramente agrupadas em três grupos diferentes formando agrupamentos genéticos separados e, portanto, sugerindo alta diferenciação genética.

De fato, indivíduos de colônias que ocorrem na Mata Atlântica mostram diferenças morfológicas que não são descritas para abelhas em outras regiões, e de acordo com Melo (2003) essas duas formas devem ser tratadas como duas espécies diferentes. A variedade encontrada no bioma Cerrado (Minas Gerais) deve ser denominada *Melipona rufiventris* Lepeletier 1836, enquanto a forma encontrada na Mata Atlântica (Goiás) deve ser denominada *M. mondury* Smith 1863 (MELO, 2003; BARNI et al., 2007). Nosso estudo também sugere que existe uma terceira forma na Caatinga (Piauí), em concordância com resultados de pesquisas anteriores (PIRES, 2010) que indicam uma nova espécie no '*grupo rufiventris*', com uma distribuição que vai do Noroeste de Minas Gerais até o Maranhão, incluindo o estado do Piauí.

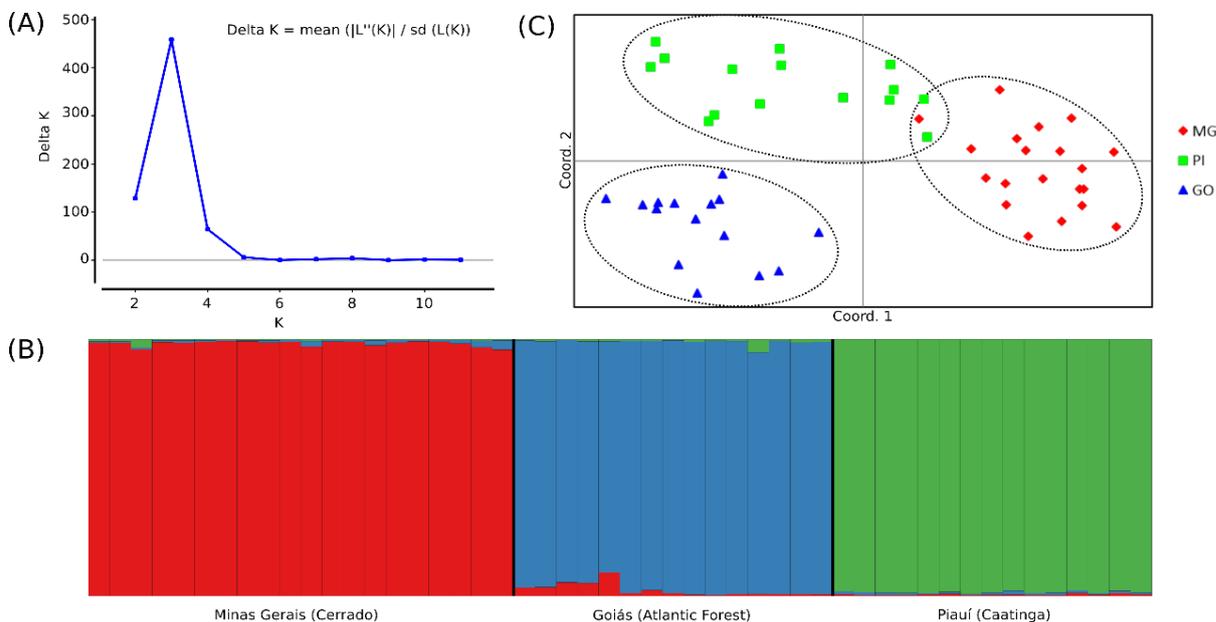
Tabela 3. Variabilidade de 16 *loci* de microssatélites e estimativas de diversidade genética no grupo "*Melipona rufiventris*" (*Hymenoptera: Apidae*).

<i>Loco</i>	Minas Gerais (<i>n</i> =20)					Goiás (<i>n</i> =15)					Piauí (<i>n</i> =15)					F_{ST}	R_{ST}	D_{EST}
	A_R	H_O/H_E	PIC	pHWE	Nulo	A_R	H_O/H_E	PIC	pHWE	Nulo	A_R	H_O/H_E	PIC	pHWE	Nulo			
<i>Mruf</i> 1	1,0	0,000/0,000	0,000	-	0,000	3,5	0,286/0,664	0,569	0,001*	0,216	4,6	0,357/0,762	0,689	0,001*	0,218	0,288 [¶]	0,279 [¶]	0,333 [¶]
<i>Mruf</i> 4	2,9	0,500/0,472	0,410	0,664	-0,027	3,6	0,786/0,643	0,562	0,473	-0,102	3,0	0,429/0,593	0,501	0,224	0,091	0,329 [¶]	0,676 [¶]	0,622 [¶]
<i>Mruf</i> 5	2,0	0,077/0,212	0,183	0,118	0,106	3,0	0,615/0,625	0,532	0,025	-0,009	3,0	0,429/0,569	0,485	0,355	0,077	0,287 [¶]	0,428 [¶]	0,357 [¶]
<i>Mruf</i> 6	2,0	0,917/0,518	0,373	0,014	-0,281	1,0	0,000/0,000	0,000	-	0,000	2,0	1,000/0,520	0,375	0,001*	-0,333	0,270 [¶]	0,270 [¶]	0,229 [¶]
<i>Mruf</i> 8	2,0	0,400/0,337	0,269	1,000	-0,061	2,6	0,385/0,335	0,290	1,000	-0,047	3,8	0,500/0,540	0,482	0,166	0,011	0,433 [¶]	0,363 [¶]	0,508 [¶]
<i>Mruf</i> 9	3,9	0,294/0,709	0,630	0,000*	0,234	4,8	0,800/0,766	0,701	0,351	-0,035	3,9	0,556/0,732	0,631	0,036	0,080	0,058	0,160	0,193 [¶]
<i>Mruf</i> 11	4,0	1,000/0,779	0,696	0,010	-0,147	3,8	0,786/0,717	0,633	0,000*	-0,056	4,0	0,667/0,752	0,657	0,050	0,025	0,030	-0,014	0,083
<i>Mruf</i> 13	2,0	0,933/0,515	0,374	0,002*	-0,291	2,0	0,818/0,506	0,367	0,066	-0,226	2,0	1,000/0,524	0,375	0,003*	-0,333	0,002	0,002	0,002
<i>Mruf</i> 15	3,1	0,714/0,521	0,433	0,430	-0,141	2,7	0,308/0,283	0,255	1,000	-0,028	2,8	0,400/0,416	0,347	0,223	-0,004	0,053	0,085	0,035 [¶]
<i>Mruf</i> 17	1,0	0,000/0,000	0,000	-	0,000	1,0	0,000/0,000	0,000	-	0,000	2,0	0,462/0,492	0,361	1,000	0,008	0,850 [¶]	0,771 [¶]	0,847 [¶]
<i>Mruf</i> 18	2,4	0,053/0,437	0,354	0,000*	0,261	4,0	0,833/0,775	0,695	0,005	-0,052	2,7	0,917/0,562	0,432	0,013	-0,246	0,309 [¶]	0,333 [¶]	0,613 [¶]
<i>Mruf</i> 19	2,0	0,850/0,501	0,369	0,002*	-0,243	2,0	1,000/0,517	0,375	0,000*	-0,333	2,0	1,000/0,517	0,375	0,000*	-0,333	0,005	0,005	0,005
<i>Mruf</i> 20	2,0	0,500/0,386	0,305	0,526	-0,091	2,0	0,308/0,271	0,226	1,000	-0,038	1,0	0,000/0,000	0,000	-	0,000	0,106	0,106	0,033
<i>Mruf</i> 21	2,0	0,375/0,315	0,258	1,000	-0,054	2,0	0,636/0,455	0,340	0,481	-0,141	2,0	0,250/0,233	0,195	1,000	-0,026	0,019	0,019	0,012
<i>Mruf</i> 22	2,0	0,800/0,513	0,375	0,022	-0,200	2,0	0,692/0,471	0,350	0,208	-0,165	2,0	0,333/0,287	0,239	1,000	-0,044	0,112 [¶]	0,112 [¶]	0,083 [¶]
<i>Mruf</i> 25	3,5	0,813/0,579	0,498	0,206	-0,162	3,0	0,500/0,633	0,511	0,322	0,059	3,0	0,700/0,532	0,442	0,503	-0,130	0,278 [¶]	0,649 [¶]	0,491 [¶]
Mean	2,4	0,514/0,425	0,345	-		2,7	0,547/0,479	0,400	-		2,7	0,563/0,502	0,412	-		0,252 [¶]	0,317 [¶]	0,284 [¶]

AR: riqueza alélica; HO: heterozigidade observada; HE: Heterozigose esperada; PIC: Conteúdo de Informação Polimórfica; pHWE: probabilidades de partida do equilíbrio de Hardy-Weinberg; Nulo: frequência de alelos nulos. * Locus que se desviou significativamente do HWE após a correção de Bonferroni ($P < \text{crítico crítico ajustado} < 0,0029$). [¶]Significativo ($P < 0,05$). Frequências de alelos nulos negativos são um artefato de software e podem ser interpretadas como zero.

Figura 2. (A) Determinação do melhor número de *clusters* genéticos da análise STRUCTURE para *loci* de microssatélites em colônias do grupo *Melipona rufiventris*; K é o número de *clusters*. (B) Gráfico de barras da estrutura populacional inferida usando o programa baseado em modelo de mistura de agrupamento bayesiano STRUCTURE (K = 2). (C) Gráfico de dispersão da análise de coordenadas principais (PCoA) usando *loci* SSR.

◆ Minas Gerais (Cerrado); ▲ Goiás (Mata Atlântica); ■ Piauí (Caatinga).



6.5 Conclusão

Os resultados encontrados aqui mostraram que o conjunto de 16 marcadores moleculares descritos, que são os primeiros *loci* de microssatélites para o "*grupo rufiventris*" usando a tecnologia NGS, tem grande potencial para estudos genéticos em nível populacional e, conseqüentemente, acrescentará informações valiosas para auxiliar em estudos de conservação da espécie. Além do mais, foi detectado a presença de alto grau de diferenciação genética no grupo "*rufiventris*", representando colônias de três biomas distintos, e que estas devem ser tratadas como unidades de conservação separadas ou mesmo como espécies separadas.

6.6 Referências

ABDELKRIM, J. et al. Fast, cost-effective development of species-specific microsatellite markers by genome sequencing. **Biotechniques**. v.46, p.185–191, 2009.

BARNI, G. S. et al. Mitochondrial genome differences between the stingless bees *Melipona rufiventris* and *Melipona mondury* (Apidae: Meliponini). **Genetics and Molecular Research**. v.6, p.8–14, 2007.

BEEBEE, T. J. C. & ROWE, G. Conservation genetics in Introduction to molecular ecology. p.199–222, **Oxford University Press**, 2004.

CAMPOS, L. A. O. *Melipona rufiventris* Lepeletier, 1836 in Livro vermelho das espécies ameaçadas de extinção da fauna de Minas Gerais ed. Machado, A. B. *et al.*) 608 p. **Fundação Biodiversitas**, 1988.

CORNUET, J.M. & LUIKART, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. **Genetics**, v.144, 2001–2014, 1996.

COSTA, R.G. et al. Isoenzyme Variation in *Melipona rufiventris* (Hymenoptera: Apidae, Meliponina) in Minas Gerais State, Brazil. **Biochemical Genetics**. n.43, p. 49–58, 2005.

EARL, D. A. & VONHOLDT, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**. v.4, p.359–361, 2012.

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.

FAIRCLOTH, B. Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. **Molecular Ecology Research**. v.8, p.92–94, 2008.

FRANCINI, I. B. et al. Microsatellite *loci* for an endemic stingless bee *Melipona seminigra merrillae* (Apidae, Meliponini) from Amazon. **Conservation Genetics Resources**, v.1, p.487–490 2009.

FRANCINI, I. B. et al. Isolation and characterization of microsatellite *loci* in the stingless bee *Melipona interrupta manaosensis* (Apidae: Meliponini). **Conservation Genetics Resources**. v.2, p.27–30, 2010.

FREITAS, P. V. D. X. et al. Avaliação da área de voo da abelha *Melipona rufiventris rufiventris* por meio do teste de captura e recaptura, **Congresso Estadual de Iniciação Científica do IF Goiano**, p.1–2, 2015.

GARIBALDI, L. A. et al. Wild pollinators enhance fruit set of crops regardless of honeybee abundance. **Science**. v.339, p.1608–1611, 2013.

GOUDET, J. FStat Version 2.9.3.2: a computer program to calculate F-statistics. **Journal of Heredity**, v.86, p.485–486, 1995.

GRAUR, D. Gene diversity in Hymenoptera. **Evolution**. v.39, p.190–199, 1985.

ICMBIO. **Livro Vermelho da Fauna Brasileira: Ameaçada de Extinção** (Portarias MMA nº 444/2014 e 445/2014). <http://www.icmbio.gov.br/portal/faunabrasileira?id=6706:portarias-fauna-ameacada>, 2015.

JOST, L. GST and its relatives do not measure differentiation. **Molecular Ecology**, v.17, p.4015–4026, 2008.

KALINOWSKI, S. T., TAPER, M. L. & MARSHAL, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. **Molecular Ecology**. v.16, p.1099–1106, 2007.

LOPES, D. M. et al. Comparative study on the use of specific and heterologous microsatellite primers in the stingless bees *Melipona rufiventris* and *M. mondury* (Hymenoptera, Apidae). **Genetics and Molecular Biology**, v. 33, p.390–393, 2010a.

LOPES, D. M. et al. Microsatellite *loci* for the stingless bee *Melipona rufiventris* (Hymenoptera: Apidae). **Molecular Ecology Research**, n.9, p.923–925, 2009.

LOPES, D. M. et al. Scientific note on the characterization of microsatellite locus for *Melipona mondury* (Hymenoptera: Apidae). **Apidologie**. v.41, p.138–140 2010b.

MARIETTE, S. et al. Sampling within the genome for measuring within-population diversity: Tradeoffs between markers. **Molecular Ecology** v.11, p.1145–1156, 2002.

MAY-ITZÁ, W. J., QUEZADA-EUÁN, J. J. G., MEDINA, L. A. M., ENRIQUEZ, E. & DE LA RÚA, P. Morphometric and genetic differentiation in isolated populations of the endangered Mesoamerican stingless bee *Melipona yucatanica* (Hymenoptera: Apoidea) suggest the existence of a two species complex. **Conservation Genetics**. v.11, p.2079–2084, 2010.

MELO, G. A. R. Notas sobre melipóíneos neotropicais (Hymenoptera, Apidae), com a descrição de três novas espécies, **Apoidea Neotropica**, p.85–92, 2003.

METCALF, R. L., MARLIN, J. C. & WHITT, G. S. Low levels of genetic heterozygosity in Hymenoptera. **Nature**. v.257, p.792–794, 1975.

NASCIMENTO, M. A. Variabilidade genética de *Melipona quadrifasciata* (Hymenoptera: Apidae) no estado de Minas Gerais com marcadores ISSR. **Master's Dissertation in Cellular and Structural Biology**, Universidade Federal de Viçosa, 2008.

PEAKALL, R. & SMOUSE, P. E. GENALEX 65: Genetic analysis in Excel Population genetic software for teaching and research: An update. **Bioinformatics**. v.28, p.2537–2539, 2012.

PIRES, C.V. Filogenia de espécies *Melipona* do complexo *rufiventris*, com base em sequências de DNA mitocondrial in Master's Dissertation in Genetics and Breeding, **Universidade Federal de Viçosa**, 2010.

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

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

QUEZADA-EUAN, J. J. G. et al. Morphological and molecular characters reveal differentiation in a Neotropical social bee, *Melipona beecheii* (Apidae: Meliponini). **Apidologie**. v.38, p.1–12, 2007.

ROZEN, S. & SKALETSKY, H. J. **Primer 3 on the www for general users and for biologist programmers in Bioinformatics methods and protocols: methods in molecular biology** (ed. Krawetz, S. & Misener, S.) 365–386 (Human Press, 2000).

SELKOE, K. A. & TOONEN, R. J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite marker. **Ecology Letters**, v.9, p.615–629, 2006.

SILVA, G. R. et al. Genome-wide discovery and characterization of microsatellite markers from *Melipona fasciculata* (Hymenoptera: Apidae), cross-amplification and a snapshot assessment of the genetic diversity in two stingless bee populations. **European Journal of Entomology**, v.115, p.614–619, 2018.

SILVEIRA, F. A., MELO, G. A. R. & ALMEIDA, E. A. B. **Abelhas brasileiras: sistemática e identificação**, 253 p. Fundação Araucaria, 2002.

SINGH, R. S. & RHOMBERG, L. R. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. **Genetics**. v.115, p.313–322, 1987.

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

SOUZA, I. G. B. et al. Isolation and characterization of 23 microsatellite *loci* in the stingless bee *Melipona subnitida* using next-generation sequencing. **Conservation Genetics Resources**. v.7, p.239–241, 2015.

TAVARES, M. G. et al. Genetics divergence between populations of the stingless bee uruçum-amarela (*Melipona rufiventris* group, Hymenoptera, Meliponini): Is there a new *Melipona* species in the Brazilian state of Minas Gerais? **Genetics and Molecular Biology**, v.30, p.667–675 2007.

TAVARES, M. G. et al. Uso de biologia molecular como subsidio para preservação de abelhas: espécies do grupo *rufiventris* (Meliponini) em Minas Gerais in **Insetos Sociais: da biologia a aplicação**, ed. Vilela, p.128–141, UFV, 2008.

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

VOSSLER, F. G. Flower visits, nesting and nest defense behaviour of stingless bees (Apidae: Meliponini): suitability of the bee species for meliponiculture in the Argentinean Chaco region. **Apidologie**. 43, p.139–161, 2012.

WRIGHT, S. Evolution and the genetics of populations. Variability within and among natural populations. 580 p. **The University of Chicago Press**, 1978.

ZANE, L., BARGELLONI, L. & PATARNELLO, T. Strategies for microsatellite isolation: a review. **Molecular Ecology**, v.11, p.1–16, 2002.

7 CONSIDERAÇÕES FINAIS

Mediante todas as análises realizadas, ao final deste presente estudo de tese conclui-se que:

- a) Os protocolos tradicionais que utilizam os solventes orgânicos para a purificação ainda são bastante eficazes para extração de DNA de abelhas, principalmente do gênero *Melipona*. Apesar da presença de impurezas, o Chelex e HOTSHOT também demonstraram serem úteis para procedimento comuns da biologia molecular que utilizam como base as reações de PCR;
- b) Com o sequenciamento de Nova Geração em baixa cobertura (Illumina) foi possível identificar grande abundância de sequências de microssatélites nas abelhas *Melipona rufiventris*, *M. subnitida* e *M. fasciculata*, com potencial para serem utilizados como marcadores moleculares em abelhas-sem-ferrão;
- c) A partir dessa tecnologia foi possível obter mais 16 marcadores microssatélites polimórficos específicos para *M. rufiventris*, dos quais incluem os do tipo tri e tetranucleotídeos.
- d) Dessa forma, utilizando marcadores microssatélites heteroespecíficos e específicos desenvolvidos a partir dessa tecnologia foi possível verificar a eficácia na avaliação da diversidade genética e estrutura populacional das populações de *Melipona rufiventris*.