



Genoma de abelhas-sem-ferrão em baixa cobertura - WGS (*Whole Genome Sequencing*)*

Isis Gomes de Brito Souza¹; Geice Ribeiro da Silva²; Aline Barbosa Negreiros³; Fábria de Mello Pereira⁴; Bruno Almeida de Souza⁴; Fábio Mendonça Diniz⁵

¹Bolsista de pós-doutorado pela CAPES, estagiária da Embrapa Meio-Norte, isisgomesmd@hotmail.com

²Doutorando pelo Programa Ciência Animal/UFPI, estagiário da Embrapa Meio-Norte. ³Doutoranda em Biotecnologia pelo RENORBIO/UFPI, estagiária da Embrapa Meio-Norte. ⁴Pesquisador da Embrapa Meio-Norte, fabia.pereira@embrapa.br ⁵Pesquisador da Embrapa Caprinos e Ovinos.

A ameaça provocada pela crescente destruição da vegetação nativa e o uso indiscriminado de agrotóxicos às abelhas-sem-ferrão brasileiras têm demandado a realização de estudos populacionais, principalmente no que concerne ao emprego de marcadores moleculares de modo que contribua para o desenvolvimento de estratégias de gestão desse importante recurso genético. Com a identificação de sequências de DNA nas suas extremidades (*paired-endreads* - Illumina), realizou-se o sequenciamento de todo o genoma (*Whole Genome Sequencing* - WGS), em baixa cobertura, das abelhas-sem-ferrão *Melipona rufiventris*, *M. subnitida* e *M. fasciculata*, tendo em vista o desenvolvimento de marcadores e a análise genético-populacional dessas espécies. Para isso, uma biblioteca Illumina *paired-end* foi criada seguindo o protocolo padrão do *kit Nextera DNA library Prep*. Em seguida, as amostras foram sequenciadas por meio do MiSeq Benchtop (Illumina Inc., San Diego, Califórnia). As sequências contíguas (*contigs*) foram montadas a partir das sequências *paired-end* resultantes, usando-se o software CLC Genomics Workbench 7.0.4. Os *contigs* no formato FASTA foram submetidos ao software MSDB (*Microsatellite Search and Building Database*) para a busca de regiões de 2 a 6 repetições. Um total de 137.313, 141.412 e 47.087 *contigs* para *M. rufiventris* (54.555.929 *reads*), *M. subnitida* (1.995.104 *reads*) e *M. fasciculata* (2.669.884 *reads*), respectivamente, resultaram da montagem “*de novo*”. Em todos, o tamanho mínimo dos *contigs* foi de 200 pares de bases, com o máximo sendo de 13.505 bases para *M. rufiventris*, 11.035 para *M. subnitida* e 16.428 para *M. fasciculata*. Os *contigs* tiveram tamanhos médios de 429, 498 e 452 bases para *M. rufiventris*, *M. subnitida* e *M. fasciculata*, respectivamente. Identificaram-se regiões do DNA com repetições em blocos em 10.266 (*M. rufiventris*), 24.128 (*M. subnitida*) e 9.954 (*M. fasciculata*) *contigs*. Por meio do sequenciamento WGS foi possível identificar diversas regiões do DNA com repetições em *tandem* passíveis para o desenho de *primers* nas espécies *M. rufiventris*, *M. subnitida* e *M. fasciculata*. Foram desenhados 50 pares de *primers* para *M. rufiventris*, 52 para *M. subnitida* e 37 para *M. fasciculata* no programa WEBSAT. Estes estão sendo otimizados e validados e posteriormente serão depositados no banco de dados do NCBI (GenBank).

Palavras-chave: Jandaíra, tiúba, urucu-amarela, sequenciamento de alta performance, repetições em *tandem*.

Agradecimentos: Embrapa Meio-Norte, Embrapa Caprinos e Ovinos, Universidade Federal do Piauí (Programa de Pós-Graduação em Ciência Animal) e Rede Nordeste de Biotecnologia - RENORBIO.

* Projeto financiado pela Embrapa, Macroprograma6, código 06.14.01.001.00.00.