Microsatellite markers developed in the stingless bee *Melipona fasciculata* by next-generation sequencing and an exploratory analysis of geographic genetic variation

Geice R Silva [1], **Isis G B Souza** [2], **Bruno A Souza** [3], **Fábia M Pereira** [3], **Maria Teresa R Lopes** [3], **Paul Bentzen** [4], **Fabio Mendonça Diniz** [Corresp. 3, 5]

[1] Department of Animal Science, Universidade Federal do Piauí, Teresina, Piauí, Brazil

[2] Northeast Biotechnology Network - RENORBIO, Universidade Federal do Piauí, Teresina, Piauí, Brazil

[3] Laboratory of Molecular Biology & Biotechnology, Embrapa Meio-Norte, Teresina, Piauí, Brazil

[4] Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada

[5] Embrapa Caprinos e Ovinos, Sobral, Ceará, Brazil

Corresponding Author: Fabio Mendonça Diniz
Email address: fabio.diniz@embrapa.br

Background. Native meliponines are currently threatened by increased human impacts. The assessment of their genetic variation by microsatellite DNA markers can assist in the conservation of populations and help in the planning and establishment of efficient management strategies. Next generation sequencing has proven to be useful for identifying microsatellite loci from the large amounts of sequence data generated.

Methods. The purpose of this study was to develop the first set of microsatellite markers for *Melipona fasciculata*, selected from partial genome assembly of Illumina paired-end reads. Contigs were created from the resulting paired-end sequence data and these were analyzed with specialized software to extract those reads that contained microsatellite loci. Primer pairs were designed for each detected locus at their flanking regions. Bee samples were genotyped from two different locations for markers characterization and validation.

Results. A total of 17 microsatellite loci displayed polymorphism from two different populations of Northeastern Brazil. Mean H_E and H_O heterozygosities were 0.453 and 0.536, respectively. PIC across all loci ranged from 0.108 to 0.714. A genetic diversity analysis revealed high values for population differentiation estimates (F_{ST} = 0.194, R_{ST} = 0.230, and D_{est} = 0.162). PCoA and Bayesian clustering showed a separation of the species into two distinct clusters.

Discussion. The Illumina paired-end sequencing system provided a large number of microsatellite loci from the *M. fasciculata* genome. From the genotyped data this study was able to reveal high F_{ST} and R_{ST} estimates and suggest the existence of genetic structure. These microsatellite markers have demonstrated strong potential for population-level genetic studies and can be used effectively as a molecular tool. Moreover, the exploratory analysis of the genetic diversity in *M. fasciculata* provides provisional evidence of significant population differentiation between the two studied populations.

Microsatellite markers developed in the stingless bee *Melipona fasciculata* by next-generation sequencing and an exploratory analysis of geographic genetic variation

Geice Ribeiro da Silva [1,*], Isis Gomes de Brito Souza [2,*], Bruno de Almeida Souza [3], Fabia de Mello Pereira [3], Maria Teresa do Rego Lopes [3], Paul Bentzen [4], Fabio Mendonça Diniz [3,5]

[1] Universidade Federal do Piauí, Campus Universitário Ministro Petrônio Portella, Ininga, Teresina, PI, 64049-550, Brazil.
[2] Universidade Federal do Piauí, Northeast Biotechnology Network - RENORBIO, Teresina, PI, 64049-550, Brazil.
[3] EMBRAPA Meio-Norte, Laboratory of Molecular Biology & Biotechnology, CP: 01, CEP: 64.006-220, Teresina, PI, Brazil.
[4] Department of Biology, Dalhousie University, Halifax, NS, B3H 4J1, Canada.
[5] Embrapa Caprinos e Ovinos, Caixa Postal: 145, CEP: 62010-970, Sobral, CE, Brazil.

.

* Corresponding author:
Fabio Mendonça Diniz [3,5]

Email address: fabio.diniz@embrapa.br

[*] These authors contributed equally to this work

Abstract

Background. Native meliponines are currently threatened by increased human impacts. The assessment of their genetic variation by microsatellite DNA markers can assist in the conservation of populations and help in the planning and establishment of efficient management strategies. Next generation sequencing has proven to be useful for identifying microsatellite loci from the large amounts of sequence data generated.
Methods. The purpose of this study was to develop the first set of microsatellite markers for *Melipona fasciculata*, selected from partial genome assembly of Illumina paired-end reads. Contigs were created from the resulting paired-end sequence data and these were analyzed with specialized software to extract those reads that contained microsatellite loci. Primer pairs were designed for each detected locus at their flanking regions. Bee samples were genotyped from two different locations for markers characterization and validation.
Results. A total of 17 microsatellite loci displayed polymorphism from two different populations of Northeastern Brazil. Mean H_E and H_O heterozygosities were 0.453 and 0.536, respectively. PIC across all loci ranged from 0.108 to 0.714. A genetic diversity analysis revealed high values for population differentiation estimates (F_{ST} = 0.194, R_{ST} = 0.230, and D_{est} = 0.162). PCoA and Bayesian clustering showed a separation of the species into two distinct clusters.
Discussion. The Illumina paired-end sequencing system provided a large number of microsatellite loci from the *M. fasciculata* genome. From the genotyped data this study was able to reveal high F_{ST} and R_{ST} estimates and suggest the existence of genetic structure. These microsatellite markers have demonstrated strong potential for population-level genetic studies and can be used effectively as a molecular tool. Moreover, the exploratory analysis of the genetic diversity in *M. fasciculata* provides provisional evidence of significant population differentiation between the two studied populations.

Introduction

Stingless bees (Hymenoptera: Apidae: Meliponini) are quite a diverse group of bees regarded for their great economic and ecological importance. For instance, beekeeping provides a sustainable source of income under a low-cost investment for smallholder farming communities; and these native bees support an efficient pollination service in both natural and agricultural systems (*Heard 1999; Slaa et al., 2006; Garibaldi et al., 2013*).

Currently, native meliponines are threatened by increased human impacts such as destruction of native vegetation and consequent landscape transformation (*Brown & Paxton, 2009; Winfree et al., 2009; Potts et al., 2010; Roulston & Goodell, 2011*). Anthropogenic disturbances or intervention may negatively affect the existence of small populations of native stingless bees, leading to the risk of local extinction (*Silva et al., 2014*). Therefore, a clear understanding of the genetic variation and population structure of meliponine bees can contribute to the development of effective conservation strategies to secure the continued survival of these original populations and the species itself.

The *Melipona* (Melikerria) *fasciculata* Smith, 1854 (Hymenoptera, Apidae), popularly known as "uruçu-cinzenta" or "tiúba", is a native stingless bee species that can be found in the neotropical region of Brazil, within the states of Pará, Tocantins, Maranhão, Piauí and Mato Grosso (*Silveira, Melo & Almeida, 2002*). Apart from its role as a pollinator in most ecosystems and crops (*Cortopassi-Laurino et al. 2006; Nunes-Silva et al., 2013*), a great interest in the species has emerged because (*i*) stingless beekeeping is relatively easy, as long as flowering plants are available, and (*ii*) its production of honey and geopropolis with antioxidant potential (*Oliveira et al., 2012; Dutra et al., 2014*) and anti-inflammatory effect (*Liberio et al., 2011*).

Microsatellites, stretches of short DNA sequences tandemly repeated, have become the markers of choice for high-resolution assessment of genetic variation and population structure studies, most importantly, due to their abundance throughout the eukaryote genome and their hypervariability (Goldstein & Schlötterer, *1999; Wan et al., 2004*). Emerging technologies in DNA sequencing (i.e. next generation sequencing - NGS) have proven to be useful for identifying microsatellite loci from the large amounts of sequence data they generate with much less effort and low cost, therefore, challenging traditional approaches for their development (*Mardis, 2008; Zalapa et al., 2012; Park et al., 2013; Souza et al., 2015*).

In this paper, we describe the first set of microsatellite markers developed for *Melipona fasciculata*, selected from partial genome assembly of Illumina paired-end reads. An exploratory analysis of its geographic genetic variation is also performed to characterize and validate these polymorphic markers.

Materials and Methods

Bee materials and genomic DNA isolation. Genomic DNA was extracted from each adult worker thorax (n = 50) using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instruction. Bees were collected from hives originally from the Northeast region of Brazil, in the states of Piauí (Elesbão Veloso city; 6°11'56.2"S 42°11'43.8"W) and Maranhão (São Bento city; 2°42'30.6"S 44°50'18.9"W). The extracted DNA samples were electrophoresed on 0.8 % agarose gel to test for overall quantity and quality of the DNA yield.

Library preparation and NGSequencing. A single individual with the highest quality DNA yield was selected for sequencing. DNA was quantified using a PicoGreen protocol and was run using a Perkin Elmer Fusion DNA Quantifier (Perkin Elmer, Waltham, Massachusetts). An Illumina paired-end library was created using 1 ng of genomic DNA, following the standard

protocol of the Illumina Nextera XT Library Preparation kit (Illumina Inc., San Diego, CA). DNA was tagged and fragmented by the Nextera XT transposome, followed by limited-cycle PCR amplification, AMPure XP magnetic-bead purification (Agencourt Bioscience Corporation, Beverly, MA) and the Illumina Nextera XT bead-based normalization protocol. The DNA library was sequenced using a MiSeq Benchtop Sequencer (Illumina Inc., San Diego, CA). Contigs were created from the resulting paired-end sequence data (reads) using CLC Genomics Workbench 7.0.4 (Qiagen).

Microsatellite loci search and primer design. All these contigs were subsequently added directly into MSATCOMMANDER 0.8.2 (*Faircloth, 2008*) for detection of possible microsatellite loci with at least four repeats, except for dinucleotides (six repeats), and designing of primer pairs for each detected locus at their flanking regions. Long mononucleotide repeats were ignored for marker development. Primer design was performed with the Primer3 (*Rozen & Skaletsky, 2000*).

Microsatellite-PCR amplification for primer validation and genotyping. Genomic DNAs from 5 individuals were initially used to validate all designed primer pairs using polymerase chain reactions (PCRs). Reactions were performed in a 10 µL total volume containing at least 10 ng of genomic DNA, with 1.25 to 1.5X buffer, 2 to 2.5 mM $MgCl_2$, 10 mM dNTP mix, 0.25 mM of each primer and 0.25 units of Taq DNA polymerase (Thermo Scientific Inc.) or HotStar Taq DNA Polymerase (Qiagen). All amplifications were run in a Veriti 96-well Thermal Cycler (Applied Biosystems) using the PCR temperature profile indicated in Table 1. The amplification products were screened by silver nitrate detection on denatured 6% polyacrylamide gels. Additional bee samples were genotyped from two different locations (Piauí and Maranhão), 25 individuals each, to obtain baseline allele frequency information.

Data analysis. The genotyped data was initially analyzed using Micro-Checker 2.2.3 (*Van Oosterhout et al., 2004*) to test for the presence of null alleles, large alleles dropout and scoring errors by stuttering. Observed and expected heterozygosities (H_O and H_E), the number of alleles (Na), and the polymorphic information content (PIC) were determined using CERVUS 3.0 (*Kalinowski, Taper & Marshal, 2007*). Allelic richness (A_R) as a measure of the number of alleles per locus independent of population size was calculated by FSTAT version 2.9.3.2 (*Goudet, 1995*). Deviations from Hardy–Weinberg equilibrium (HWE) and tests for linkage disequilibrium were conducted using Genepop software (*Raymond & Rousset, 1995*)[26]. The Bonferroni correction (*Rice, 1989*) was applied when multiple pairwise tests were performed to assess the significance ($P < 0.05$).

The genetic diversity for each locus was evaluated by ARLEQUIN ver. 3.5 (*Excoffier & Lischer, 2010*), which determined the value of θ (F_{ST}) for the whole sample set.

A Bayesian grouping admixture model was used to infer possible population structuring using the software STRUCTURE v2.3.3 (*Pritchard, Stephens & Donnelly, 2000*). The program was set up for 1,000,000 Markov chain Monte Carlo repetitions after an initial burn-in of 500,000 steps. The estimate of the best K was calculated based on 5 replications for each K (from 1 to 6) as described by Evanno et al. (*Evanno, Regnaut & Goudet, 2005*) using Structure Harvester v.0.6.92 (*Earl & vonHoldt, 2012*). The program CLUMPP v.1.1.2 (*Jakobsson & Rosenberg, 2007*) was used to align the five repetitions of the best K. The program DISTRUCT v.1.1 (*Noah, 2004*) was used to graphically display the results produced by CLUMPP. Population structure was also analyzed using principal coordinate analysis (PCoA), R_{ST} (*Slatkin, 1995*), a measure of genetic differentiation analogous to F_{ST}, and the D_{est} estimator of actual differentiation (*Jost, 2008*) as implemented in GenAlEx v.6.5 (*Peakall & Smouse, 2012*).

Table 1 Characteristics of 18 microsatellite markers developed for *Melipona fasciculata*.

Results and Discussion

Genome survey and assembly, SSR loci identification and primer design. A total of 2,669,884 sequencing reads were generated from a 300 bp insert library using Illumina paired-end sequencing technology. These reads were assembled into 47,087 contigs. These contigs were then screened for microsatellite motifs using MSATCOMMANDER 0.8.2 (*Faircloth, 2008*). Search results revealed 9,954 contigs (11.3 % of total contigs) containing 11,869 microsatellite sequences. As expected, mononucleotide (6,444) and dinucleotide (3,225) repeats were the most frequent, followed by tri- (734) and tetranucleotide (574) loci. We focused only on tri- and tetra-repeat microsatellites with more than 10 repeats mainly because they are easier to score on polyacrylamide gels than dinucleotides, as well as longer repeats generally show higher mutation rates (*Petit et al. 2005*). From these, 37 microsatellite loci were selected for primer designing and validation.

Polymorphism and validation of genomic microsatellite markers. The Micro-Checker analysis of the entire dataset revealed null alleles for loci Mfsc3 and Mfsc11, at lower frequencies than 0.2 (Table 2). Null allele frequencies below 0.2 are acceptable in most microsatellite data sets (*Dakin & Avise, 2004*). After dividing dataset into two groups, one corresponding to samples collected in Piauí state and another composed of only individuals from Maranhão, Micro-Checker analysis revealed nulls only for Mfsc3 in the Piauí population after Bonferroni's correction, which may be a possible cause of its deviation from HWE. No pairwise combination of microsatellite loci showed significant linkage disequilibrium indicating that none of the loci were physically linked.

PCR products of expected size based on sequence data from the partial genome assembly with clear and consistent bands were obtained from 18 primer pairs, out of 37 tested on 50 individual bees from the two surveyed groups in Northeastern Brazil (Table 2). Seventeen microsatellite loci were polymorphic across the entire data set. However, few have revealed a monomorphic binding pattern at a population-level analysis (Mfsc24 and Mfsc30 in Piauí and Mfsc3, Mfsc7, Mfsc13, Mfsc31 and Mfsc32 in Maranhão). Nevertheless, it is still expected that these loci may become polymorphic once again when additional individuals are sampled.

Table 2 Variability of 17 microsatellite loci and genetic diversity estimates in *Melipona fasciculata*.

The genotyping of the entire dataset has revealed 70 alleles, ranging from 1, for locus Mfsc34 to 10, for locus Mfsc17, with an average of 3.9 ± 2.7 alleles per locus (Table 2). This result was of similar magnitude to that found in other species within the same genus such as *M. rufiventris* (*Lopes et al., 2009*), *M. seminigra merrillae* (*Francini et al., 2009*), *M. interrupta manaosensis* (*Francini et al., 2010*), *M. mondury* (*Lopes et al., 2010*) and *M. yucatanica* (*May-Itzá et al., 2010*). The size of alleles in the least polymorphic locus (H_E and PIC), Mfsc31, ranged from 260 to 263 bp, while for the most polymorphic locus, Mfsc27, alleles varied from 205 to 227. These two loci (Mfsc31 and Mfsc27) were composed of trinucleotide motifs. As shown in Table 2, the level of polymorphism of each locus was also evaluated by the allelic richness (A_R) and the polymorphic information content (*PIC*). The values of allelic richness varied from 2 to 9.1 (average of 3.9 ± 2.4), while *PIC* values ranged between 0.108 and 0.714. Mean PIC (0.372 ±

0.198) characterize all microsatellite loci as reasonably informative markers (*Botstein, White & Skolnick, 1980*). Overall mean observed and expected heterozygosity was estimated to be 0.536 and 0.453, respectively. These estimates were higher when compared to most heterozygosities found for *Melipona* species, exception made for *M. subnitida* (*Souza et al., 2015*). It is noteworthy that low levels of heterozygosity are known to occur in social Hymenoptera compared to other insects (*Graur, 1985*). Nine microsatellite loci exhibited significant probabilities ($P < 0.05$) of departure from Hardy-Weinberg equilibrium, likely due to the mixing of individuals from populations of different allelic frequencies (*Templeton, 2006*).

Exploratory analysis of the genetic diversity in *Melipona fasciculata*. Genetic diversity between *M. fasciculata* populations, as measured by the mean number of alleles per microsatellite locus, mean allelic richness, heterozygosity and PIC, was characterized by a slightly higher level of genetic variability from samples collected in Maranhão when compared to those sampled in Piauí (Table 2).

The high F_{ST} (0.194) and R_{ST} (0.230) estimates found in *M. fasciculata* suggest the existence of genetic structure. However, additional genetic surveys should be carried out to confirm this observation. F_{IS} estimates for most loci were negative, indicating a trend towards an excess of observed heterozygosity. Similarly high F_{ST} estimates were reported in wild populations of *M. rufiventris* (*Tavares et al., 2007*) and *M. beecheii* (*Quezada-Euan, 2007*), with values of 0.25 and 0.280, respectively. D_{est}, which is a measure based on the proportion of alleles that are unique to a subpopulation (*Jost, 2008*), provides further evidence of population differentiation ($D_{est} = 0.162$). Low rates of dispersion and short flight distance, less than 2000 m, might have contributed to the levels of population differentiation (*Silveira, Melo & Almeida, 2002; Araújo et al., 2004; Duarte, Gaiotto & Costa, 2014*).

All these estimates reflect that gene exchange is not prevailing in these two *M. fasciculata* populations. Therefore, concerns should be addressed to the question of up to what level the destruction of native semiarid vegetation is influencing genetic drift or reduction of gene flow among *M. fasciculata* populations currently restricted to the remaining fragments of native forests.

The scatter-plot of the principal coordinate analysis (PCoA) showed a clear separation of the species in two distinct clusters of stingless bees, coincidental with the origin of the individuals. This analysis further confirmed the genetic differentiation of the two populations (Fig. 1A). The analysis of microsatellite variation using the admixture model of STRUCTURE, at the first level of sub-population separation ($K = 2$), has also revealed two distinct clusters (Fig. 1B). These clusters represent each sampling population separately, except for very few individuals mostly located in Piauí that appears to be admixed.

Figure 1 A. Principal coordinates analysis of microsatellite variation in the stingless bee *Melipona fasciculata*. B. Genetic structure inferred using Bayesian analysis in the program STRUCTURE. Each individual is represented by a vertical line.

Conclusion

Overall, analyses provide provisional evidence of significant population differentiation between Maranhão and Piauí, in *M. fasciculata*. However, the data generated by this study should be further investigated using the same microsatellites markers, but larger sample size and more widespread sampling throughout the distribution of the species. Given all these considerations, the eighteen isolated microsatellite loci have demonstrated strong potential for population-level

genetic studies and can be used effectively as molecular tool to aid in the conservation of the species. Moreover, the results obtained from this exploratory analysis of the genetic diversity in *M. fasciculata*, with these molecular markers, support their use for conducting population genetics and landscape genetics studies.

Acknowledgments

Sequencing was conducted on a MiSeq DNA sequencer purchased with a bequest from Elizabeth Ann Nielsen to the Marine Gene Probe Laboratory. The authors are grateful to Mr. José Maria Vieira-Neto (*in memoriam*) for help in field collection.

References

Araújo, E. D., Costa, M., Chaud-Netto, J. & Fowler, H.G. Body size and flight distance in stingless bees (Hymenoptera: Meliponini): inference of flight range and possible ecological implications. *Braz. J. Biol.* **64**(3b), 563−568 (2004).

Botstein, D., White, R. L. & Skolnick, M. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**, 314−331 (1980).

Brown, M. J. F. & Paxton, R. J. The conservation of bees: a global perspective. *Apidologie* **40**(3), 410–416 (2009).

Cortopassi-Laurino, M., Imperatriz-Fonseca, V. L., Roubik, D. W., Dollin, A., Heard, T., Aguilar, I. B., Venturieri, G. C., Eardley, C. & Nogueira-Neto, P. Global Meliponiculture: challenges and opportunities. *Apidologie* **37**, 1−18 (2006).

Dakin, E. E. & Avise, J. C. Microsatellite null alleles in parentage analysis. *Heredity* **93**, 504–509 (2004).

Duarte, O. M. P., Gaiotto, F. A. & Costa, M. A. Genetic differentiation in the stingless bee, *Scaptotrigona xanthotricha* Moure, 1950 (Apidae, Meliponini): a species with wide geographic distribution in the Atlantic rainforest. *J. Hered.* **105**(4), 477–484 (2014).

Dutra, R. P., Abreu, B. V. B., Cunha, M. S., Batista, M. C. A., Torres, L. M. B., Nascimento, F. R. F., Ribeiro, M. N. S. & Guerra, R. N. M. Phenolic acids, hydrolyzable tannins, and antioxidant activity of geopropolis from the stingless bee *Melipona fasciculata* Smith. *J. Agri. Food Chem.* **62**, 2549–2557 (2014).

Earl, D.A. & vonHoldt, B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).

Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).

Excoffier, L. & Lischer, H. E. L. Arlequin suite v 35: a new series of programs to perform population genetic analysis under Linux and Windows. *Mol. Ecol. Res.* **10**, 564–567 (2010).

Faircloth, B. Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Res.* **8**, 92–94 (2008).

Francini, I. B., Sforça, D. A., Sousa, A. C. B., Campos, T., Cidade, F. W., Zucchi, M. I., Souza, A. P., Nunes-Silva, C. G. & Carvalho-Zilse, G. A. Microsatellite loci for an endemic stingless bee *Melipona seminigra merrillae* (Apidae, Meliponini) from Amazon. *Conserv. Genet. Resour.* **1**, 487–490 (2009).

Francini, I. B., Sousa, A. C. B., Sforça, D. A., Costa-Pinto, M. F. F., Campos, T., Nunes-Silva, C. G., Zucchi, M. I., Souza, A. P. & Carvalho-Zilse, G. A. Isolation and characterization of microsatellite loci in the stingless bee *Melipona interrupta manaosensis* (Apidae: Meliponini). *Conserv. Genet. Resour.* **2**, 27–30 (2010).

Garibaldi, L. A, Steffan-Dewenter, I., Winfree, R., Aizen, M. A., Bommarco, R. & Cunningham, S. A. Wild pollinators enhance fruit set of crops regardless of honeybee abundance. *Science* **339**(6127), 1608–1611 (2013).

Goldstein, D. B. & Schlötterer, C. Microsatellites: evolution and applications Oxford University Press, Oxford (1999).

Goudet, J. FSTAT Version 2932: a computer program to calculate F-statistics. *J. Hered.* **86**, 485–486 (1995).

Graur, D. Gene diversity in Hymenoptera. *Evolution* **39**, 190–199 (1985).

Heard, T. A. The role of stingless bees in crop pollination. *Annu. Rev. Entomol.* **44**(131), 183–206 (1999).

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* **23**, 1801–1806 (2007).

Jost, L. GST and its relatives do not measure differentiation *Mol. Ecol.* **17**(18), 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. *Mol. Ecol.* **16**, 1099–1106 (2007).

Liberio, S. A., Pereira, A. L. A., Dutra, R. P., Reis, A. S., Araújo, M. J. A. M., Mattar, N.S., Silva, L. A., Ribeiro, M. N. S., Nascimento, F. R. F., Guerra, R. N. M. & Monteiro-Neto, V. Antimicrobial activity against oral pathogens and immunomodulatory effects and toxicity of geopropolis produced by the stingless bee *Melipona fasciculata* Smith. *BMC Complement. Altern. Med.* **11**, 108 (2011).

Lopes, D. M. & Silva, F. O., Fernandes-Salomão, T. M., Campos, L. A. O., Tavares, M. G. A Scientific note on the characterization of microsatellite loci for *Melipona mondury* (Hymenoptera: Apidae). *Apidologie* **41**, 138–140 (2010).

Lopes, D. M., Silva, F. O., Salomão, T. M. F., Campos, L. A. O. & Tavares, M. G. Microsatellite loci for the stingless bee *Melipona rufiventris* (Hymenoptera, Apidae). *Mol. Ecol. Res.* **9**, 923–925 (2009).

Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**, 133–141 (2008).

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. *Conserv. Genet.* **11**, 2079–2084 (2010)

Noah, A. R. Distruct: a program for the graphical display of population structure. *Mol. Ecol. Notes* **4**, 137–138 (2004).

Nunes-Silva, P., Hrncir, M., Silva, C. I., Roldão, Y. S. O. & Imperatriz-Fonseca, V. L. Stingless bees, *Melipona fasciculata*, as efficient pollinators of eggplant (*Solanum melongena*) in greenhouses. *Apidologie* **44**(5), 537–546 (2013).

Oliveira, P. S., Müller, R. C. S., Dantas, K. D. G. F., Alves, C. N., Vasconcelos, M. A. M. D. & Venturieri, G. C. Phenolic acids, flavonoids and antioxidant activity in honey of *Melipona fasciculata*, *M flavolineata* (Apidae, Meliponini) and *Apis mellifera* (Apidae, Apini) from the Amazon. *Quim. Nova* **35**, 1728–1732 (2012).

Park, I., Kim, J., Lee, J., Kim, S., Cho, O., Yang, K., Ahn, J., Nahm, S. & Kim, H. Development of SSR markers by next-generation sequencing of Korean landraces of chamoe (*Cucumis melo* var. *makuwa*). *Mol. Biol. Rep.* **40**(12), 6855–6862 (2013).

Peakall, R. & Smouse, P. E. GenAlEx 65: Genetic analysis in Excel Population genetic software for teaching and research: An update. *Bioinformatics* **28**, 2537–2539 (2012).

316 Petit, R. J., Deguilloux, M. F., Chat, J., Grivet, D., Garnier-Géré, P. & Vendramin, G. G.
317 Standardizing for microsatellite length in comparisons of genetic diversity. *Mol Ecol* **14** 885–
318 890 (2005).
319 Potts, S. G, Biesmeijer, J. C, Kremen, C., Neumann, P., Schweiger, O. & Kunin, W. E. Global
320 pollinator declines: Trends, impacts and drivers. *Trends Ecol Evol* **25**, 345–353 (2010).
321 Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus
322 genotype data. *Genetics* **155**(2), 945–959 (2000).
323 Quezada-Euan, J. J. G., Paxton, R. J., Palmer, K. A., May-Itza, W. J., Tay, W. T. & Oldroyd, B. P.
324 Morphological and molecular characters reveal differentiation in a Neotropical social bee,
325 *Melipona beecheii* (Apidae: Meliponini). *Apidologie* **38**, 1–12 (2007).
326 Raymond, M. & Rousset, F. An exact test for population differentiation *Evolution* **49**, 1280–1283
327 (1995).
328 Rice, W. R. Analyzing tables of statistical tests. *Evolution* **43**, 223–225 (1989).
329 Roulston, T. H. & Goodell, K. The role of resources and risks in regulating wild bee populations.
330 *Annu. Rev. Entomol.* **56**, 293–312 (2011).
331 Rozen, S. & Skaletsky, H. J. Primer3 on the WWW for general users and for biologist
332 programmers In: Krawetz, S. & Misener, S. (Eds), Bioinformatics Methods and Protocols:
333 Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365–386 (2000).
334 Silva, G. R., Pereira, F. M., Souza, B. A., Lopes, M. T. R., Campelo, J. E. G. & Diniz, F. M.
335 Bioecological and behavioral genetic aspects involved in the conservation of the stingless bee
336 *Melipona subnitida* Ducke (Apidae, Meliponini) and the use of molecular tools in studies of
337 diversity. *Arq Inst Biol* **81**(3), 299–308 (2014).
338 Silveira, F. A., Melo, G. A. R. & Almeida, E. A. B. Abelhas do Brasil: Sistemática e Identificação,
339 first edition, Author´s edition Belo Horizonte (Brazil), 253 p. (2002).
340 Slaa, E. J., Sánchez-Chaves, L. A. Malagodi-Braga, K. S., & Hofstede, F. E. Stingless bees in
341 applied pollination, practice and perspectives. *Apidologie* **37**(2), 293–315 (2006).
342 Slatkin, M. A measure of population subdivision based on microsatellite allele frequencies.
343 *Genetics* **139**, 457–462 (1995).
344 Souza, I. G. B., Paterson, I., McBride, M. C., Souza, B. A., Pereira, F. M., Lopes, M. T. R.,
345 Bentzen, P. & Diniz, F. M. Isolation and characterization of 23 microsatellite loci in the
346 stingless bee *Melipona subnitida* using next-generation sequencing. *Conserv. Genet. Resour.*
347 **7**(1), 239–241 (2015).
348 Tavares, M. G., Dias, L. A. S., Borges, A. A., Lopes, D. M., Busse, A. H. P., Costa, R. G.,
349 Fernandes-Salomão, T. M. & Campos, L. A. O. Genetics divergence between populations of
350 the stingless bee uruçu-amarela (*Melipona rufiventris* group, Hymenoptera, Meliponini): Is
351 there a new *Melipona* species in the Brazilian state of Minas Gerais? *Gen. Mol. Biol.* **30**, 667–
352 675 (2007).
353 Templeton, A. R. Population Genetics and Microevolutionary Theory. John Wiley & Sons,
354 Hoboken, New Jersey, 705 p. (2006)
355 Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. Micro-Checker: software for
356 identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–
357 538 (2004).
358 Wan, Q. H., Wu, H., Fujihara, T. & Fang, S. G. Which genetic marker for which conservation
359 genetics issue? *Electrophoresis* **25**(14), 2165–2176 (2004).
360 Winfree, R., Aguilar, R., Vazquez, D. P., LeBuhn, G. & Aizen, M. A. A meta-analysis of bees'
361 responses to anthropogenic disturbance. *Ecology* **90**, 2068–2076 (2009).
362 Zalapa, J. E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R. &
363 Simon P. Using next-generation sequencing approaches to isolate simple sequence repeat
364 (SSR) loci in the plant sciences. *Am. J. Bot.* **99**(2), 193–208 (2012).

Table 1(on next page)

Characteristics of 18 microsatellite markers developed for *Melipona fasciculata*.

Table 1 Characteristics of 18 microsatellite markers developed for *Melipona fasciculata*.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3382v1 | CC BY 4.0 Open Access | rec: 31 Oct 2017, publ: 31 Oct 2017

Table 1 Characteristics of 18 microsatellite markers developed for *Melipona fasciculata*.

Locus	Primer sequence (5'-3')	Motifs repeats	Ta (°C)	Na	Size range (bp)	Reaction profile	GenBank accession no.
Mfsc3	F: GAGCGAGAGGGAGCAAGATA R: TAGTAACGTAATTCTGCGCGAT	$(AGAT)_{14}$	50	2	090-094	PCR_{STD}	KT730150
Mfsc7	F: TCTACCCATCTCTGTTTCTCTCC R: TCGCAGTTTCGTTGATTTTG	$(ATCT)_{10}$	50	2	232-236	PCR_{STD}	KT730151
Mfsc10	F: AGTAGAACGATTTCGAGAAC R: ACGAAGCCGTATGCTAA	$(CTTT)_{10}$	43	2	148-160	$PCR_{STD+HOT1}$	KT730152
Mfsc11	F: GGAAGGACGAGAGAATTCAAGA R: ATAGTCGTTTGTCGCGAGTGTA	$(CTTT)_{10}$	50	5	170-186	PCR_{STD}	KT730153
Mfsc13	F: GCAGTAACGGTAGCAGTGGTG R: ACTCCTTTCTCCTTCTCGGTCT	$(ACG)_{16}$	52	2	189-207	PCR_{STD}	KT730154
Mfsc14	F: AGTTGCAGCGTTTGTGAAATC R: GTGGGTTCGGAGATGTGTATAAG	$(AGT)_{17}$	47-57	2	116-122	$PCR_{TCD+HOT1}$	KT730155
Mfsc17	F: ATTTTCTCAGTAAGCGAGTCCG R: CGACCTTGTTCGTATAATAGCA	$(ATT)_{17}$	50	10	142-187	$PCR_{STD+HOT1}$	KT730156
Mfsc22	F: GTGACAATAATAGGAGGGAAATCG R: GAAGCTGGTACAGGTATCGGAG	$(GAT)_{14}$	58-48	2	231-234	$PCR_{TCD+HOT1}$	KT730157
Mfsc23	F: ATTCGGCATCGGCGTTAT R: TTAGAGAAAGTTGTTGGACCCG	$(CGT)_{14}$	48	3	243-261	PCR_{STD}	KT730158
Mfsc24	F: GTAGAGGAGTAGTAACAGCAA R: CGAGTCCCGGTTAGC	$(AGC)_{14}$	48	4	165-189	PCR_{STD}	KT730159
Mfsc27	F: CGTCTCCACCGTCTTCTATTTC R: GCGTGTCCTCTCTTTCTCTCTC	$(AGC)_{13}$	50	6	205-227	PCR_{STD}	KT730160
Mfsc28	F: ATGATTCTCGCTTTCGTCGT R: GTGAGGAGACGCTGGATTTC	$(AGC)_{13}$	52-62	5	160-184	$PCR_{TCD+HOT2}$	KT730161
Mfsc30	F: TCTATAAGCGCCAGAGAGGAAG R: TTTCAGGGATGCGCC	$(ACG)_{12}$	50	3	186-192	PCR_{STD}	KT730162
Mfsc31	F: TGTGGTCGCGGTTGC R: TCGCCGCTCGGAACT	$(AAG)_{12}$	50	2	260-263	PCR_{STD}	KT730163
Mfsc32	F: GTTATCGTTATCGTCATCGTCGT R: CCGTGAGCGAACTCGAAC	$(CGT)_{12}$	47-57	3	105-108	$PCR_{TCD+HOT1}$	KT730164
Mfsc34	F: AACTTTGAGGACGCACGAG R: CACTTCTTGTTCGACTTGGTTG	$(ACGA)_{12}$	53	1	109	$PCR_{STD+HOT1}$	KT730165
Mfsc36	F: CGCCTACACCTAGAACCAAAA R: ACGTACACCGATGGCGTT	$(AAAG)_{13}$	55	9	097-109	PCR_{STD}	KT730166
Mfsc37	F: GAAGGAAGGAAAGAGGCCG R: CCATTGCTACCCGTACCTCC	$(AAAG)_{10}$	55	8	103-119	PCR_{STD}	KT730167

Ta: Annealing temperature; Na: Number of alleles; **PCR_{STD}**: Standard PCR [94°C-1 min; 40 cycles (94°C-30 sec; Ta-30 sec; 72°C-30 sec); 72°C-3 min]; **$PCR_{TCD+HOT1}$**: Touchdown PCR with Hotstart Taq DNA polymerase {95°C-15 min; 10 cycles [94°C-30 sec; Ta1-30 sec; 72°C-30 sec]; 10 cycles [94°C-30 sec; Ta2(-1°C/cycles)-30 sec; 72°C-30 sec]; 25 cycles [94°C-30 sec; Ta1-30 sec; 72°C-30 sec); 72°-10 min]}; **$PCR_{TCD+HOT2}$**: Touchdown PCR with Hotstart Taq DNA polymerase = {95°C-15 min; 10 cycles [94°C-30 sec; Ta1 (-1,0°C/cycle)-30 sec;72°C-30 sec]; 25 cycles [94°C-30 sec; Ta2-30 sec; 72°C-30 sec]; 72°C-10 min.

Table 2(on next page)

Variability of microsatellite loci and genetic diversity estimates.

Table 2 Variability of 17 microsatellite loci and genetic diversity estimates in *Melipona fasciculata*.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3382v1 | CC BY 4.0 Open Access | rec: 31 Oct 2017, publ: 31 Oct 2017

Table 2 Variability of 17 microsatellite loci and genetic diversity estimates in *Melipona fasciculata*.

Locus	All Individuals *(n = 50)*						**Piauí** *(n = 25)*						**Maranhão** *(n = 25)*						F_{ST}	R_{ST}	D_{est}
	A_R	H_O	H_E	PIC	pHWE	Null	A	H_O	H_E	PIC	pHWE	Null	A	Ho	HE	PIC	pHWE	Null			
Mfsc3	2.0	0.000	0.186	0.167	0.000*	0.156	2	0.000	0.423	0.325	0.000*	0.290	1	0.000	0.000	0.000	Mono	0.000	0.321	0.424	0.089
Mfsc7	2.0	0.353	0.295	0.248	0.556	-0.048	2	0.522	0.394	0.311	0.267	-0.098	1	0.000	0.000	0.000	Mono	0.000	0.189	0.396	0.079
Mfsc10	2.0	0.617	0.431	0.336	0.002*	-0.133	2	0.500	0.384	0.305	0.272	-0.091	2	0.720	0.470	0.355	0.008	-0.177	0.016	0.067	0.012
Mfsc11	4.9	0.326	0.646	0.570	0.000*	0.191	2	0.381	0.316	0.261	1.000	-0.056	4	0.280	0.409	0.376	0.01	0.086	0.601	0.158	0.873
Mfsc13	2.0	0.306	0.262	0.226	0.575	-0.037	2	0.625	0.439	0.337	0.055	-0.137	1	0.000	0.000	0.000	Mono	0.000	0.309	0.084	0.122
Mfsc14	2.0	0.378	0.504	0.373	0.184	0.079	2	0.357	0.516	0.374	0.318	0.094	2	0.391	0.507	0.730	0.403	0.070	-0.038	0.258	-0.039
Mfsc17	9.1	0.732	0.610	0.550	0.013	-0.081	5	0.611	0.571	0.501	0.038	-0.036	7	0.826	0.642	0.572	0.047	-0.122	0.003	0.103	0.004
Mfsc22	2.0	1.000	0.505	0.375	0.000*	-0.333	2	1.000	0.512	0.375	0.000*	-0.333	2	1.000	0.510	0.375	0.000*	-0.333	0.000	0.083	0.000
Mfsc23	3.0	0.864	0.574	0.494	0.000*	-0.189	3	0.762	0.547	0.469	0.056	-0.149	2	0.957	0.510	0.375	0.000*	-0.305	0.161	-0.019	0.207
Mfsc24	3.9	0.370	0.335	0.304	0.395	-0.029	1	0.000	0.000	0.000	Mono	0.000	4	0.680	0.528	0.457	0.281	-0.107	0.253	0.094	0.132
Mfsc27	5.9	0.979	0.762	0.714	0.000*	-0.129	3	1.000	0.532	0.406	0.000*	-0.315	5	0.960	0.731	0.668	0.008	-0.142	0.290	0.017	0.684
Mfsc28	4.8	0.732	0.556	0.509	0.142	-0.118	2	0.375	0.315	0.258	1.000	-0.054	4	0.960	0.626	0.546	0.000*	-0.215	0.182	0.351	0.209
Mfsc30	2.9	0.156	0.186	0.173	0.017	0.024	1	0.000	0.000	0.000	Mono	0.000	3	0.292	0.324	0.286	0.031	0.019	0.125	0.035	0.030
Mfsc31	2.0	0.073	0.116	0.108	0.120	0.037	2	0.188	0.272	0.229	0.302	0.060	1	0.000	0.000	0.000	Mono	0.000	0.164	0.333	0.023
Mfsc32	2.9	0.362	0.334	0.288	0.004*	-0.024	3	0.773	0.538	0.427	0.000*	-0.162	1	0.000	0.000	0.000	Mono	0.000	0.399	0.062	0.226
Mfsc36	8.3	1.000	0.746	0.700	0.000*	-0.151	3	1.000	0.613	0.516	0.000*	-0.252	9	1.000	0.817	0.775	0.000*	-0.111	0.051	0.209	0.137
Mfsc37	7.4	0.872	0.664	0.601	0.000*	-0.130	2	0.727	0.474	0.356	0.017	-0.181	8	1.000	0.783	0.734	0.000*	-0.132	0.082	0.131	0.151
Mean	3.9	0.536	0.453	0.396	-	-	2.3	0.519	0.403	0.321	-	-	3.4	0.533	0.403	0.367	-	-	0.194	0.230	0.162

A_R: Allelic richness; **A**: Number of alleles within population; H_O: Observed heterozygosity; H_E: Expected heterozygosity; **PIC**: Polymorphic Information Content; **pHWE**: probabilities of departure from Hardy-Weinberg equilibrium; **Null**: Null alleles frequency. *Locus that deviated significantly from HWE after Bonferroni correction (adjusted critical $P < 0.0029$).

Figure 1

Figure with PCA and Structure analyses

Figure 1 A. Principal coordinates analysis of microsatellite variation in the stingless bee Melipona fasciculata. B. Genetic structure inferred using Bayesian analysis in the program STRUCTURE. Each individual is represented by a vertical line.