



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde

Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

**Estudo de estrutura genética da espécie de Cactácea colunar
Pilosocereus machrisii utilizando DNA microssatélite**

Manolo Fernandez Perez

**SÃO CARLOS - SP
- 2012 -**

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Biológicas e da Saúde

Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

**“Estudo de estrutura genética da espécie de Cactácea colunar
Pilosocereus machrisii utilizando DNA microssatélite”**

Manolo Fernandez Perez

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

Orientação: Prof. Dr. Evandro Marsola de Moraes

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

F363ee

Perez, Manolo Fernandez.

Estudo de estrutura genética da espécie de Cactácea colunar *Pilosocereus machrisii* utilizando DNA microsatélite / Manolo Fernandez Perez. -- São Carlos : UFSCar, 2012. 65 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2012.

1. Genética e evolução. 2. *Pilosocereus machrisii*. 3. Estrutura populacional. 4. Campos rupestres. 5. Microsatélite. I. Título.

CDD: 575 (20^a)

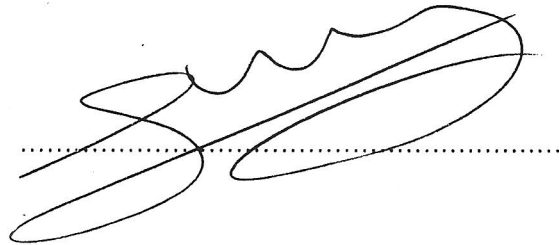
UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

“Estudo de estrutura genética da espécie de Cactácea colunar
Pilosocereus marchisii utilizando DNA microssatélite”

Dissertação de Mestrado de
Manolo Fernandez Perez

Banca Examinadora

Prof. Dr. Evandro Marsola de Moraes



Profa. Dra. Maria Imaculada Zucchi



Profa. Dra. Vera Nisaka Solferini



Dedico a todos aqueles que
contribuíram de alguma forma para a
realização desse trabalho, em especial
aos meus pais e irmãos.

Este trabalho foi realizado com apoio financeiro da FAPESP (Proc. 05-55200/8) e do CNPq (Proc. 471963/2007-0) e com bolsa de Mestrado da FAPESP (2010/02665-1).

AGRADECIMENTOS

Agradeço aos meus pais Miguel e Elaine pelo carinho, amor, compreensão e apoio incondicional em todas as etapas de minha vida.

Aos meus irmãos Paola e Santiago por estarem presentes nessa caminhada, me ajudando em tudo o que fosse possível.

Aos meus avós, tios e primos, que sempre estiveram ao meu lado me apoiando em todos os momentos.

À minha namorada Ana Cláudia por me ajudar na redação e nas figuras desse trabalho, por discutir muitos dos assuntos aqui apresentados e por sempre estar ao meu lado nos momentos bons e ruins, me dando força e coragem.

Ao Prof. Dr. Evandro Marsola de Moraes pela orientação e amizade. Por confiar no meu trabalho e por seus valiosos ensinamentos e paciência em todos os momentos.

Ao Prof. Dr. Fernando de Faria Franco pelas contribuições e conselhos importantes para o desenvolvimento desse trabalho.

Aos colaboradores Nigel Taylos, Daniela Zappi e Marlon Machado, pela imensa ajuda nos trabalhos desenvolvidos no Laboratório de Diversidade Genética e Evolução e pela ajuda com a taxonomia e as coletas das Cactáceas.

Ao Sr. Gerardus Hubertus Olsthoorn pela ajuda com muitas informações importantes sobre os cactos e pelas amostras gentilmente cedidas.

Às técnicas de laboratório Heidi e Fabrícia por toda a ajuda e o enorme auxílio na realização dos experimentos.

À minha colega de Graduação, de Pós-Graduação e de laboratório Isabel, pelas discussões e pela ajuda inestimável em todos esses anos.

Aos colegas do laboratório de Diversidade Genética e Biologia Evolutiva do campus Sorocaba da Universidade Federal de São Carlos, pela ajuda em tudo que fosse preciso.

Aos verdadeiros amigos, em especial ao meu quase irmão Eduardo Moreno Stucchi, por sempre me apoiarem e trazerem muita alegria para minha vida.

A todo o pessoal da UFSCar Sorocaba, principalmente os técnicos e amigos da copinha e do futebol, pelas conversas e momentos de descontração, bem como pelos conselhos e companheirismo.

Ao colega de Wesley Cardoso Generoso, uma grande amizade que encontrei em São Carlos, que me ajudou muito nas atividades da Pós-Graduação e sempre esteve disposto a resolver todo e qualquer problema que eu tivesse.

Aos colegas e professores do Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular (PPGGev), pelo acolhimento e auxílio no que fosse preciso.

À Universidade Federal de São Carlos (UFSCar) e ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular (PPGGev), pela infra-estrutura oferecida para a realização do projeto.

À FAPESP pelo apoio financeiro ao projeto e a concessão da bolsa de Mestrado.

RESUMO

Pilosocereus machrisii é um cacto colunar com distribuição naturalmente fragmentada no centro e leste do Brasil, restrito aos enclaves de vegetação de campos rupestres ou de afloramentos rochosos no domínio Cerrado. Essas características tornam a espécie um modelo biológico apropriado para a realização de estudos de estrutura populacional, uma vez que suas populações pequenas e isoladas podem estar altamente vulneráveis a efeitos de deriva genética e endogamia. Foram estimados os níveis de diversidade genética e estruturação populacional a partir de dez locos marcadores microssatélite em doze populações de ocorrência natural de *P. machrisii*, cobrindo a maior parte de sua distribuição. Os níveis de diversidade genética dentro das populações apresentaram resultados relativamente similares nas localidades amostradas, exceto pelas amostras do município de Brotas-SP, com índices menores que os observados em outras populações. Desvios significativos em relação às proporções do Equilíbrio de Hardy-Weinberg foram detectados para um ou dois locos em oito populações diferentes. Alelos privados em alta frequência foram detectados, assim como altos níveis de diferenciação genética, com um valor de F_{ST} total de 0,357. A estrutura genética das populações, verificada a partir da análise com os programas STRUCTURE e SAMOVA apresentou uma estruturação mais provável em quatro grupos principais e níveis secundários de estruturação dentro desses grupos. Níveis significativos de correlação entre as distâncias genéticas e geográficas entre pares de populações foram detectados para o conjunto total de dados, no entanto não houve correlação quando os agrupamentos genéticos foram analisados separadamente. Além disso, foi possível verificar as relações genéticas entre as diferentes populações a partir de uma análise de coordenadas principais (PCA) e de um dendrograma construído a partir do método UPGMA. A estrutura genética observada em *P. machrisii* aponta para um importante efeito de eventos históricos que causaram a fragmentação das

populações, com posterior isolamento geográfico e aquisição de elevada diferenciação genética.

Palavras-chave: *Pilosocereus machrisii*, estrutura populacional, campos rupestres, microsátélites, distribuição fragmentada.

ABSTRACT

Pilosocereus machrisii is a columnar cacti with natural fragmented distribution, restrict to *campos rupestres* vegetation patches or rocky outcrops on the Cerrado domain, in Central and eastern Brazil. These features makes *P. machrisii* an appropriate biological model for population genetic studies aiming to gather information on population structure, since small and isolated populations are often vulnerable to genetic drift and high levels of inbreeding. Genetic diversity and population structure were assessed for ten microsatellite loci in 12 *P. machrisii* populations, covering the majority of the species distribution. Genetic diversity levels were relatively similar on sampled populations, except for Brotas-SP samples, that showed lower levels than those observed for other locations. Significant Hardy Weinberg Equilibrium departures were detected in eight different populations, for one or two loci. High levels of genetic differentiation and private alleles with elevated frequencies were detected, with total F_{ST} value of 0.357. The genetic structure of *P. machrisii* was analyzed with STRUCTURE and SAMOVA softwares, and populations were grouped in four main clusters, with secondary structure on two of these clusters. Significant levels of correlation between genetic and geographic distances were observed for the whole dataset, however, when the genetic groups were analyzed separately, no correlation was verified. Furthermore, the genetic relationship between the populations was observed with a principal coordinate analysis (PCA) and an UPGMA dendrogram. The genetic structure observed in *P. machrisii* suggests that historical factors caused populations fragmentation, with subsequent geographic isolation and high genetic differentiation between them.

Keywords: *Pilosocereus machrisii*, population structure, *campos rupestres*, microsatellites, patchy distribution.

LISTA DE FIGURAS

- Figura 1.** Características morfológicas de *P. machrisii*. a) Indivíduo em afloramento rochoso na população de Uruaçu-GO. b) Indivíduo com a flor aberta na localidade do município de Alto Paraíso de Goiás-GO. c) Fruto maduro de um indivíduo da localidade Cristalina-GO. Fotos: Arquivo Pessoal. 17
- Figura 2.** Mapa da região leste do Brasil apresentando a distribuição de *P. machrisii*. A área sombreada corresponde às populações amostradas e à distribuição total sugerida para a espécie (ZAPPI, 1994; HUNT et al., 2006). 18
- Figura 3.** Resultado gráficos da análise para identificação do valor mais provável de K a partir da estatística ΔK (EVANNO et al., 2005), calculado por meio de 10 replicações..... 39
- Figura 4.** Esquema dos resultados obtidos a partir da análise *bayesiana* hierárquica do programa STRUCTURE (PRITCHARD et al., 2000). Cada indivíduo é representado por uma barra vertical, apresentando a proporção de seu genoma pertencente a cada um dos grupos estabelecidos,. Linhas pretas separam os indivíduos de diferentes populações. As abreviações utilizadas para as populações estão representadas na Tabela 1. 42
- Figura 5.** Mapa de elevação (medida em metros) da região central e leste do Brasil apresentando a proporção média do genoma dos indivíduos em cada população atribuída aos grupos determinados pelo programa STRUCTURE (PRITCHARD et al., 2000; K=4). As cores representadas seguem o esquema apresentado na Figura 4. 43
- Figura 6.** Representação da autocorrelação espacial da diferenciação genética apresentando os coeficientes médios de correlação entre pares de populações, plotados em relação às suas distâncias geográficas subdivididas em 12 classes. Barras indicam o intervalo de confiança do coeficiente de parentesco ($IC_{95\%}$) e linhas tracejadas indicam os valores críticos de rejeição ($VC_{95\%}$) da hipótese nula de ausência de estruturação genética espacial. 46
- Figura 7.** Dendrograma UPGMA construído a partir da distância genética D_C (CAVALLI-SFORZA e EDWARDS, 1967). Árvore baseada em 500 *bootstraps* nos indivíduos das populações. Os agrupamentos hierárquicos obtidos pela análise com o programa STRUCTURE (PRITCHARD et al., 2000) estão representados ao lado direito dos ramos. Coeficiente de Correlação Cofenética=0,82 ($P<0,005$). 47
- Figura 8.** PCA a partir da distância genética D_C (CAVALLI-SFORZA e EDWARDS, 1967) mostrando a) divergência entre todos os indivíduos (57,41% da variação total nos dois eixos representados) e b) a divergência todas populações (60,69% da variação total nos dois eixos representados). As elipses representam os grupos revelados pelo STRUCTURE (PRITCHARD et al., 2000) com K=4. 48

LISTA DE TABELAS

- Tabela 1.** Localidades onde foram realizadas coletas de material biológico de indivíduos da espécie *P. machrisii*.25
- Tabela 2.** Relação dos 10 locos de DNA microsatellite utilizados para as genotipagens em *Pilosocereus machrisii*. Sequências dos iniciadores *forward* (F) e *reverse* (R), unidades repetitivas, variação no tamanho dos alelos e temperaturas de hibridização padronizadas (*Ta*).27
- Tabela 3.** Parâmetros de variabilidade genética estimados ao nível populacional em *P. machrisii*. *A*, número médio de alelos por loco; *H_O*, heterozigosidade média observada; *H_E*, heterozigosidade média esperada, estimada pelo método de Nei (1987); *A_R*, riqueza alélica; *F_{IS}*, valor do coeficiente de endocruzamento; EHW, locos com desvios significativos em relação às proporções do Equilíbrio de Hardy-Weinberg.37
- Tabela 4.** Alelos exclusivos de DNA microsatélite com frequências acima de 10% em populações de *P. machrisii*.38
- Tabela 5.** Análise de variância molecular para microsatélites em *P. machrisii* usando diferentes níveis hierárquicos. As populações foram agrupadas de acordo com os resultados da análise com os programas STRUCTURE (PRITCHARD et al., 2000) ou SAMOVA (DUPANLOUP et al., 2002).40
- Tabela 6.** Análise de variância molecular para microsatélites em *P. machrisii* para os indivíduos pertencentes aos grupos formados com K=4 pela análise do programa STRUCTURE (PRITCHARD et al., 2000).41
- Tabela 7.** Resultados da análise de detecção de migrantes. São apresentadas as populações amostradas e as prováveis populações-fonte de acordo com os programas STRUCTURE (PRITCHARD et al., 2009) e GENECLASS (Piry et al., 2004).44
- Tabela 8.** Matriz das distâncias geográficas par-a-par (em quilômetros, diagonal inferior) e dos valores par-a-par do parâmetro *F_{ST}* (diagonal superior) para as populações de *P.machrisii*.45

SUMÁRIO

1. INTRODUÇÃO	13
2. MATERIAL E MÉTODOS	25
2.1 Amostras e extração de DNA	25
2.2 Locos de DNA microssatélite	26
2.3 Análise dos dados	28
2.3.1 Índices de Diversidade	28
2.3.2 Diferenciação Populacional	29
2.3.3 Estrutura Populacional	30
2.3.4 Detecção de Migrantes.....	32
2.3.5 Relação Entre as Distâncias Genéticas e Geográficas.....	33
2.3.6 Relações Genéticas entre as Populações	34
3. RESULTADOS	36
3.1 Índices de Diversidade.....	36
3.2 Diferenciação Populacional	37
3.3 Estrutura Populacional	38
3.4 Detecção de Migrantes.....	43
3.5 Relação Entre as Distâncias Genéticas e Geográficas	44
3.6 Relações Genéticas entre as Populações.....	46
4. DISCUSSÃO	49
5. CONCLUSÃO.....	56
6. REFERÊNCIAS BIBLIOGRÁFICAS	57
ANEXO.....	65

1. INTRODUÇÃO

A família Cactaceae apresenta cerca de 1400 espécies distribuídas em 124 gêneros (HUNT et al., 2006), com distribuição praticamente restrita ao continente americano, com exceção de uma espécie do gênero *Rhipsalis* que ocorre no sul da África, Madascar e Sri Lanka (CALVENTE et al., 2011).

De acordo com dados moleculares, as espécies dessa família possuem origem monofilética (WALLACE e COTA, 1996; NYFFELER, 2002) e apresentam uma série de características que podem ser relacionadas às condições secas presentes nos habitats em que ocorrem, dentre as quais podem ser enumeradas: (1) capacidade de produzir tricomas em suas raízes, o que possibilita absorção imediata da precipitação, (2) cutícula espessa na epiderme, (3) tecidos modificados que possibilitam a estocagem de líquidos, (4) modificação de folhas em espinhos, (4) caule com pigmentos fotossintéticos e (5) o metabolismo ácido das crassuláceas (CAM), que permite a absorção de gás carbônico utilizado para a fotossíntese durante o período noturno, minimizando a perda de água (NYFFELER, 2002; ARAKAKI et al., 2011; ZAPPI et al., 2011).

Os principais centros de diversidade de espécies da família localizam-se no sudeste dos Estados Unidos e México, bem como na região dos Andes, em particular no Peru e na Bolívia. O leste do Brasil abriga o terceiro centro mais importante em diversidade de Cactaceae, contendo cerca de 130 espécies distribuídas principalmente nos limites do Domínio Caatinga e associadas a campos rupestres localizados no Domínio do Cerrado (ZAPPI et al., 2011).

O gênero *Pilosocereus* é considerado um dos três gêneros de Cactaceae mais especiosos no Brasil, compreendendo cerca de 40 espécies, divididas em grupos taxonômicos informais com base em caracteres morfológicos e distribuição geográfica (ZAPPI, 1994; HUNT et al., 2006). O gênero ocorre desde o Domínio da Caatinga na região nordeste até as

regiões Centro-Oeste e Sudeste do território brasileiro, onde as espécies ocorrem de maneira disjunta e associadas a áreas de vegetação xérica. O maior centro de diversidade do gênero é a região da Cadeia do Espinhaço, no estado de Minas Gerais (ZAPPI, 1994), onde as espécies ocorrem geralmente associadas à vegetação de campo rupestre (GIULIETTI et al., 1997).

As espécies do gênero são subdivididas em dois subgêneros, *Gounellea* e *Pilosocereus*. O primeiro deles é composto por somente duas espécies (*P. gounellei* e *P. tuberculatus*) e o segundo é subdividido em cinco grupos taxonômicos informais (*P. ARRABIDAE*, *P. PENTAEDROPHORUS*, *P. LEUCOCEPHALUS*, *P. AURISSETUS* e *P. PIAUHIENSIS*).

O grupo *P. AURISSETUS* foi descrito originalmente para abrigar as espécies *P. machrisii*, *P. vilaboensis*, *P. aureispinus* e *P. aurisetus* (a última contendo as subespécies *P. aurisetus aurisetus* e *P. aurisetus aurilanatus*). Esse agrupamento foi realizado de acordo com características morfológicas de flores e frutos, ramificação predominantemente no nível do solo, além de ocorrência restrita a afloramentos rochosos (ZAPPI, 1994) associados aos enclaves de vegetação de campo rupestre (GIULIETTI et al., 1997).

Pilosocereus machrisii tem como localidade tipo a população do município de Uruaçu-GO e possui a distribuição mais ampla entre as espécies que compõem o grupo *P. AURISSETUS*, com um grande número de populações distribuídas principalmente na região central e leste do Brasil. Em uma revisão recente da família Cactaceae (HUNT et al., 2006), essa espécie sofreu algumas modificações taxonômicas. Três táxons considerados sinônimas (ZAPPI, 1994; TAYLOR e ZAPPI, 2004) foram elevadas à categoria de espécie (*P. parvus*, *P. pusillibuccatus* e *P. jauruensis*). Essas populações apresentam diferenças morfológicas importantes em relação à *P. machrisii* e ocupam áreas restritas localizadas nas bordas da sua distribuição.

Outras populações atualmente classificadas como *P. machrisii* também apresentam diferenças morfológicas significativas, sendo motivo de discussões em relação à sua posição taxonômica. Essa instabilidade pode ser associada em grande parte à elevada plasticidade fenotípica em diversas características utilizadas na taxonomia de Cactaceae (APPLEQUIST e WALLACE, 2002), assim como a descrição de diversas espécies por amadores, os quais adotam somente o conceito tipológico de espécie (TAYLOR e ZAPPI, 2004).

Próximo ao município de Posse, no nordeste do estado de Goiás, Braun e Esteves Pereira (2002) descreveram a espécie *P. goianus*, a qual apresenta características morfológicas singulares, principalmente na coloração das flores e frutos. Além disso, possui ocorrência associada com afloramentos de rochas calcárias do grupo Bambuí, diferentemente das outras populações de *P. machrisii* que, segundo esses autores ocorre somente associado a rochas areníticas e quartizíticas. Taylor e Zappi (2004) propuseram a sinonimização dessa espécie com *P. machrisii*, em contraponto ao que foi proposto por Braun e Esteves Pereira (2002). Posteriormente, Hunt et al. (2006) propuseram que *P. goianus* seria um táxon distinto e altamente relacionado com *P. albissimus*, uma espécie pertencente ao grupo *P. PENTAEDROPHORUS* com ocorrência no município de Unaí-MG.

As populações de *Pilosocereus* localizadas na região da Chapada dos Veadeiros, entre o norte do estado de Goiás e sul do Tocantins foram descritas inicialmente como *P. lindanus* por Braun e Esteves Pereira (1987a) e posteriormente sinonimizadas com *P. machrisii* por Zappi (1994).

Na região de Cristalina-GO, Braun e Esteves Pereira (1987b) descreveram as populações como *P. cristalinensis*, as quais posteriormente foram reduzidas ao nível de subespécie (*P. machrisii* subsp. *cristalinensis*) pelos mesmos autores (BRAUN e ESTEVES PEREIRA, 2002). No entanto, de acordo com outros autores (ZAPPI, 1994; TAYLOR e

ZAPPI, 2004; HUNT et al., 2006), essas populações ainda são consideradas sinônimas de *P. machrisii*.

De acordo com as populações atualmente consideradas como pertencentes à espécie, *P. machrisii* é um cacto colunar com hábito arbustivo que possui cladódios eretos ramificados somente na altura do solo (Figura 1a). A epiderme apresenta coloração variando de verde escuro a verde-azulado e espinhos marrons ou avermelhados (ZAPPI, 1994; HUNT et al., 2006). As flores possuem antese noturna, apresentando coloração de verde-claro a vermelho-escuro e marrom (Figura 1b). Essas características florais são compatíveis com a síndrome de polinização por morcegos, sendo que aves e abelhas também são consideradas como possíveis polinizadores (ZAPPI, 1994; MUNGUÍA-ROSAS et al., 2010). Os frutos apresentam cores entre vermelho-escuro e marrom e possuem polpa esbranquiçada, onde ficam imersas sementes pequenas de coloração marrom-escuro ou preta (Figura 1c), características que provavelmente estão associadas à dispersão por pássaros e morcegos (ZAPPI, 1994). Estudos com outras espécies do gênero *Pilosocereus* apontam que pássaros (GOMES et al., 2008) e formigas (MUNGUÍA-ROSAS et al., 2009) são dispersores importantes, de forma que eles também podem contribuir para a síndrome de dispersão de *P. machrisii*.

A reprodução dos indivíduos também pode ocorrer de maneira clonal, por meio de pedaços do cladódio que se destacam e caem no solo, gerando um novo indivíduo geneticamente idêntico ao anterior (ZAPPI, 1994).

A espécie possui um conjunto diplóide de 22 cromossomos (JESUS et al., 2009), não apresentando evidências de poliploidia, a qual é bastante frequente na família Cactaceae (GORELICK, 2002).



Figura 1. Características morfológicas de *P. machrisii*. a) Indivíduo em afloramento rochoso na população de Uruaçu-GO. b) Indivíduo com a flor aberta na localidade do município de Alto Paraíso de Goiás-GO. c) Fruto maduro de um indivíduo da localidade Cristalina-GO. Fotos: Arquivo Pessoal.

A ocorrência natural de *P. machrisii* está localizada nas regiões Nordeste, Sudeste e Centro-oeste do Brasil (Figura 2). Sua distribuição é naturalmente fragmentada, ocorrendo sobre afloramentos rochosos de composição quartzítica, arenítica ou calcária, associados aos enclaves de vegetação de campo rupestre (GIULIETTI et al., 1997), localidades consideradas áreas prioritárias para conservação de espécies de Cactaceae no território brasileiro (ZAPPI et al., 2011).

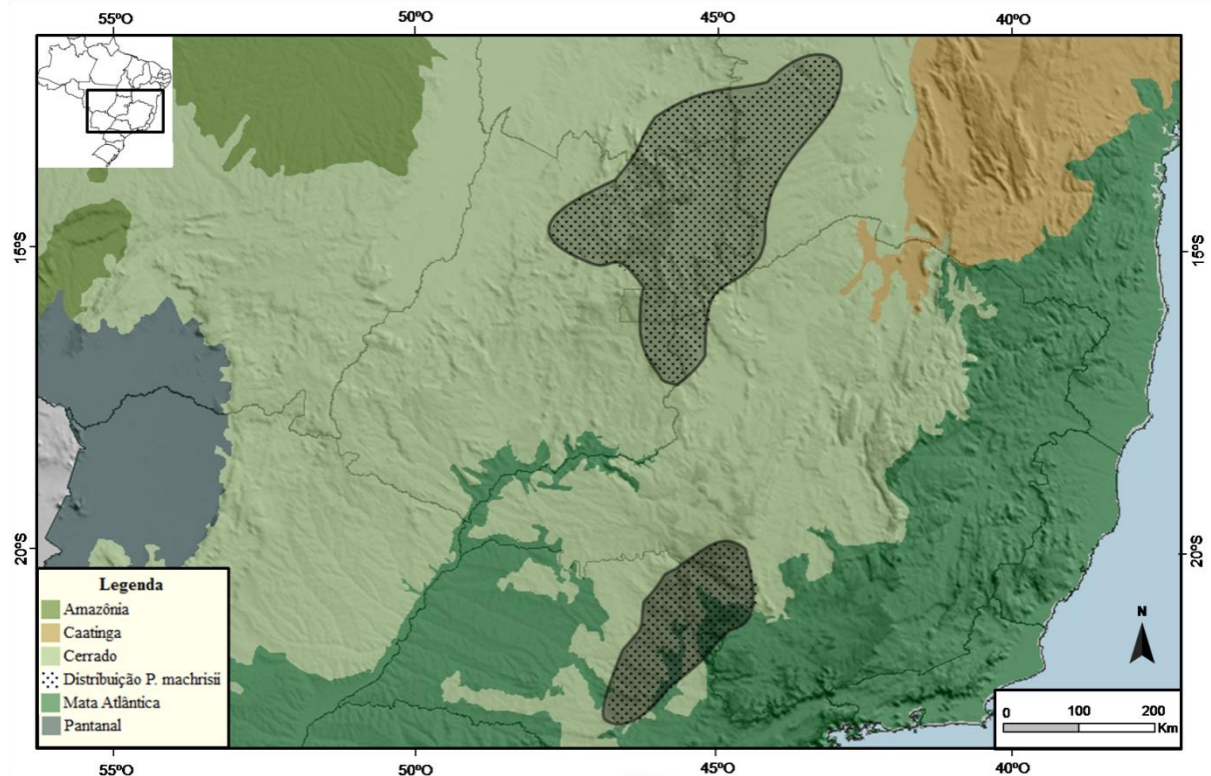


Figura 2. Mapa da região leste do Brasil apresentando a distribuição de *P. machrisii*. A área sombreada corresponde às populações amostradas e à distribuição total sugerida para a espécie (ZAPPI, 1994; HUNT et al., 2006).

Devido ao padrão de distribuição disjunta, *P. machrisii* pode ser considerado um modelo biológico apropriado para estudos de estrutura populacional que permitam obter informações sobre mudanças históricas na área de distribuição dos enclaves de vegetação xérica no domínio do Cerrado e desenvolver estratégias de conservação para os táxons desses ambientes. Muitas espécies de animais e vegetais apresentam padrões semelhantes de distribuição em outras regiões do globo (HAMILTON e ECKERT, 2007; HOLSINGER e WEIR, 2009). Esta descontinuidade na distribuição geográfica pode favorecer a diferenciação das populações ao longo do tempo por meio da deriva genética ou seleção natural diversificadora, tornando-as estruturadas geneticamente e causando a redução da variabilidade genética dentro das populações e o aumento dessa variabilidade entre elas (HOLSINGER e WEIR, 2009).

Apesar do padrão de distribuição geográfica ser um fator importante, a estrutura genética de uma espécie não é resultante apenas de fatores contemporâneos, como deriva genética, fluxo gênico e seleção natural (SLATKIN, 1987), mas também pela ação de fatores históricos, como eventos de fragmentação (vicariância), expansão, gargalo populacional e colonização a longa distância (HEWITT, 2000). Estudos que permitem entender os níveis de estruturação genética de uma espécie, bem como as causas históricas e atuais responsáveis por tal estruturação, são fundamentais para o conhecimento de suas histórias evolutivas e para a definição de estratégias de conservação. Espécies com padrão de distribuição disjunta e fragmentada são importantes modelos para esse tipo de estudo, uma vez que o fluxo gênico restrito entre suas populações pode resultar em baixa diversidade genética, aumentando o risco de extinção das populações (FRANKHAM, 2005). No entanto, as elevadas taxas de diferenciação genética resultantes desse padrão de distribuição podem resultar em linhagens evolutivas distintas com grande potencial evolutivo, podendo proporcionar o surgimento de novas unidades taxonômicas (LESICA e ALLENDORF, 1995; DAVIS e SHAW, 2001).

Moraes e Sene (2007) estudaram a variação em locos de DNA microssatélite na espécie cactofílica *Drosophila gouveai*, a qual utiliza *P. machrisii* como hospedeiro para criação das larvas e, portanto, possui uma distribuição geográfica que sobrepõe em grande parte a distribuição de *P. machrisii*. Esses autores inferiram que a estruturação genética das populações foi formada predominantemente por fatores demográficos históricos, sendo que fatores atuais como o fluxo gênico possuem menor importância. Em outro estudo com *D. gouveai* utilizando sequências de DNA mitocondrial, Moraes et al. (2009) confirmaram essas conclusões e apresentaram evidências da ocorrência de eventos cíclicos de expansão populacional a partir do norte e em direção ao sul da distribuição da espécie durante o Pleistoceno médio e posterior isolamento das populações.

Esses resultados são compatíveis com a hipótese de uma distribuição mais ampla das áreas de vegetação xerófita nas regiões central e leste do Brasil, ocasionada por eventos de mudanças climáticas durante o Pleistoceno (PRADO e GIBBS, 1993; PENNINGTON et al., 2000, 2004). A partir de um estudo sobre a estrutura genética de *Astronium urundeuva*, uma espécie arbórea que possui distribuição disjunta associada a uma unidade vegetacional denominada floresta tropical sazonalmente seca (SDTF; PRADO e GIBBS, 1993; PENNINGTON et al., 2000), Caetano et al. (2008) também obtiveram resultados que apontam para uma distribuição mais ampla dessa espécie no passado. Levando-se em conta a sobreposição parcial na distribuição de *A. urundeuva* e *D. gouveai* e a ocorrência de ambas as espécies em áreas de vegetação xerófita, as alterações históricas verificadas em suas distribuições geográficas podem ter sido causadas pelos mesmos eventos paleoclimáticos no leste da América do Sul (MORAES et al., 2009). Da mesma forma, esses eventos históricos podem ter causado alterações demográficas em *P. machrisii*, contribuindo para moldar a estrutura genética atual da espécie.

A utilização de marcadores moleculares é uma ferramenta importante para estimar os parâmetros de estrutura populacional, ocorrência de fluxo gênico em populações naturais e realizar inferências sobre a história evolutiva dos táxons estudados. Dentre esses marcadores, um dos mais utilizados são os locos de marcadores microssatélites (SELKOE e TOONEN, 2006; SUNNUCKS, 2000; SCHLÖTTERER, 2004), os quais têm sido amplamente utilizados em estudos populacionais e ecológicos para estimar parâmetros como estrutura populacional, taxa de migração, tamanho populacional efetivo, ocorrência de flutuações históricas no tamanho populacional e grau de parentesco (OLIVEIRA et al., 2006).

Os microssatélites consistem em seqüências pequenas de nucleotídeos (1 a 6 pares de bases) repetidas em série, ocorrendo em alta frequência e bem distribuídas ao longo do genoma (TÓTH et al., 2000). As mutações em regiões de microssatélites provocam alterações

no número de repetições em taxas significativamente mais altas que as observadas para substituições de bases, variando entre 10^{-2} a 10^{-6} nucleotídeos por loco por geração (SCHLÖTTERER, 2000). Geralmente se encontram localizados em regiões não-codificadoras do genoma (HANCOCK, 1995), sendo assumidos na maioria das vezes como seletivamente neutros, ou seja, a variação existente não sofre efeito da seleção natural (ELLEGRÉN, 2004). Outra característica importante é o fato desses marcadores serem codominantes, possibilitando diferenciar os indivíduos heterozigotos, o que torna os microssatélites mais informativos que outros marcadores dominantes (SELKOE e TOONEN, 2006).

Os diferentes alelos para um loco de marcadores microssatélites são identificados por diferenças no tamanho de produtos de PCR revelados por análise eletroforética, sendo assumido que alelos do mesmo tamanho são idênticos por origem. No entanto, é importante considerar a ocorrência de homoplasias, ou seja, a existência de alelos iguais em estado (i.e. possuem o mesmo tamanho ou número de repetições), mas que não são iguais por descendência (ESTOUP et al., 2002).

O modelo de mutação *Stepwise mutation model* (SMM), desenvolvido por Ohta & Kimura (1973) é um dos mais utilizados em estudos com microssatélites, uma vez que neste modelo é assumido que novos alelos são formados pela adição ou diminuição de apenas uma repetição em uma taxa fixa, independente do número total de repetições. A ocorrência de homoplasias nos locos de DNA microssatélite é prevista pelo modelo SMM, uma vez que as mutações podem tanto aumentar como diminuir o tamanho de alelos diferentes, alelos com o mesmo número de repetições podem ser gerados independentemente.

O efeito homoplásico é potencializado pelo fato dos locos de DNA microssatélite sofrerem restrições quanto ao número máximo de repetições (POLLOCK et al., 1998) e estarem sujeitos a mutações que gerem novos alelos a partir de diferenças de mais de uma repetição (SELKOE e TOONEN, 2006).

Kruglyak et al. (1998) defendem que a restrição no número de repetições é causada por mutações de ponto que ocorrem dentro das regiões repetitivas. Segundo esses autores, as mutações de ponto restringem o tamanho dos alelos por gerarem um balanço entre o efeito de mutações do tipo *stepwise*, relacionadas com o maior número de repetições e as mutações de ponto, que interrompem as longas repetições em série.

Adams et al. (2004) encontraram pouco efeito da homoplasia em análises populacionais, verificando uma ocorrência maior desses eventos em microsatélites interrompidos (i.e. possuem a sequência de unidades de repetição interrompidas por um ou mais nucleotídeos) e compostos (dois ou mais tipos de unidades repetitivas) em comparação com microsatélites perfeitos. Apesar dessa maior ocorrência de homoplasia em locos não-perfeitos, Adams et al. (2004) sugerem que estes locos sejam utilizados no caso de pouca disponibilidade de marcadores perfeitos, uma vez que a utilização de um número maior de locos é um fator que diminui os efeitos da homoplasia.

Grande parte dos estudos sobre estrutura genética populacional tem utilizado o parâmetro F_{ST} , da estatística- F de Wright (WRIGHT, 1951), para descrever os níveis gerais da distribuição da variação genética entre as populações. No entanto, apesar da fácil aplicação e eficiência para medir a estruturação genética, esse parâmetro não permite distinguir de maneira direta a contribuição relativa dos fatores históricos e contemporâneos que levaram à estrutura observada entre as populações (NIELSEN e WAKELEY, 2001; PEARSE e CRANDALL, 2004).

Atualmente, uma variedade de métodos analíticos que utilizam estatísticas *Bayesianas* e de máxima verossimilhança está expandindo a eficiência e os tipos de informação que podem ser obtidas a partir de dados genéticos, permitindo refinar as inferências sobre a importância relativa de fatores históricos e contemporâneos relacionados com a estruturação genética das populações (PEARSE e CRANDALL, 2004). Acompanhando o

desenvolvimento desses métodos, há uma variedade de novos programas computacionais voltados para a análise de dados genéticos, tornando necessária a publicação de guias gerais de uso desses programas e discussões sobre as premissas assumidas em cada método (PEARSE e CRANDALL, 2004; EXCOFFIER e HECKEL, 2006).

O objetivo do presente estudo foi analisar a estrutura populacional e o nível de diversidade genética da espécie *P. machrisii*, por meio do marcador molecular de DNA microssatélite. Os resultados das análises permitiram realizar inferências sobre a história demográfica e evolutiva dessas populações, verificando sinais sobre alterações históricas em sua distribuição e no nível de conexão entre elas.

2. MATERIAL E MÉTODOS

2.1 Amostras e extração de DNA

Foram coletados tecidos de segmentos da raiz de 8 a 30 indivíduos pertencentes a 12 populações (Tabela 1), abrangendo a maior parte da distribuição de *P. machrisii*. A distância mínima de 10 metros entre cada indivíduo foi guardada, de forma a diminuir a probabilidade de amostragem de clones. O material obtido foi armazenado em nitrogênio líquido, sendo posteriormente depositado em um freezer -80°C.

Tabela 1. Localidades onde foram realizadas coletas de material biológico de indivíduos da espécie *P. machrisii*.

Município	Coordenadas Geográficas	Número de indivíduos coletados
Brotas-SP (BRO)	S 22°17'25.5'' W 47°57'57.9''	10
Morro do Forno - Altinópolis-SP (FOR)	S 21°05'43.8'' W 47°20'03.2''	21
Furnas-MG (FUR)	S 20°38'53.4'' W 46°19'16.4''	21
Delfinópolis-MG (DEL)	S 20°21'44.5'' W 46°44'59.4''	30
Formosa-GO (FMS)	S 15°14'00.3'' W 47°09'10.8''	22
Alto Paraíso1-GO (APA1)	S 14°12'34.3'' W 47°29'15.2''	20
Alto Paraíso2-GO (APA2)	S 13°55'33.6'' W 47°25'47.5''	25
Aurora do Tocantins-TO (ART)	S 12°36'34.6'' W 46°24'10.1''	8
Uruaçu-GO (URU)	S 14°33'27.4'' W 49°09'23.3''	20
Cristalina-GO (CRI)	S 16°47'43.6'' W 47°34'10.7''	25
Posse-GO (POS)	S 14°15'16.7'' W 46°21'49.2''	18
Barreiras-BA (BAR)	S 12°05'45.2'' W 45°02'57.5''	21
Total		241

2.2 Locos de DNA microssatélite

A genotipagem dos indivíduos foi realizada a partir da utilização de 10 dos 11 locos de DNA microssatélite previamente descritos por Perez et al. (2011) (Tabela 2). O loco *Pmac135*, descrito nesse mesmo estudo, foi descartado das análises por apresentar ampliações consistentes somente nas populações localizadas na região Sul da distribuição da espécie (BRO, FOR, FUR e DEL).

Para os locos utilizados nas análises, o procedimento de genotipagem dos indivíduos foi realizado por meio da marcação de produtos de PCR com moléculas de fluorocromo, utilizando o sistema *primer tail* (SCHUELKE, 2000). Neste procedimento a sequência de DNA do iniciador universal M13 é unida à extremidade 5' do iniciador *forward* específico de cada locos. Após os primeiros ciclos de PCR, são gerados fragmentos com a sequência complementar ao iniciador universal, o que possibilita que uma nova sequência M13 marcada com um fluorocromo específico (NED, de cor preta ou 6-FAM, de cor azul) hibridize com os fragmentos gerados, possibilitando o reconhecimento do tamanho do fragmento quando analisado com um seqüenciador automático.

As reações foram realizadas em um volume de 10 μ L, contendo 1 μ L de DNA (5-50ng/ μ L), 1X de tampão padrão para PCR (Promega), 0,2 mM de cada dNTP, 0,25 μ M de cada um dos 3 iniciadores, 0,5 U de Taq DNA polimerase (Promega) e 1,5 mM de MgCl₂. As condições físicas empregadas seguiram dois protocolos diferentes. Para os iniciadores com temperatura de hibridização igual ou abaixo de 54°C os parâmetros do termocilador utilizado (Mastercycler Gradient, Eppendorf) foram 94°C por 2 minutos; 35x (94°C por 40 segundos, temperatura de hibridização específica por 40 segundos, 72°C por 40 segundos); 72°C por 10 minutos. Para aqueles iniciadores com temperatura maior que 54°C, foi realizado um procedimento de *touchdown* para as temperaturas de hibridização, no qual 25 ciclos foram

efetuados na primeira temperatura, seguidos de um decréscimo de 2°C por ciclo até que a segunda temperatura fosse alcançada e mantida em oito ciclos finais.

Tabela 2. Relação dos 10 locos de DNA microsatellite utilizados para as genotipagens em *Pilosocereus machrisii*. Sequências dos iniciadores *forward* (F) e *reverse* (R), unidades repetitivas, variação no tamanho dos alelos e temperaturas de hibridização padronizadas (*Ta*).

Loco	Unidade Repetitiva	Sequência do iniciador (5'-3')	<i>Ta</i> (°C)	Varição dos alelos (pb)*
<i>Pmac082</i>	(GAG) ₅ (GAA) ₂ (GAG) ₂ GAA(GAG) ₂	F: GTAAAAGAGGAGGATGGAGAGG R: CTTCTTCTTGCTAGGTTTCTCG	56	99 - 114
<i>Pmac084</i>	(AG) ₉ CG(AG) ₂	F: CATAAATTGCAGAAATGAGGAC R: AGGTAAACCGCTAACTCGATG	60	88 - 108
<i>Pmac085</i>	(AG) ₆ AC(AG) ₃ AC (AG) ₄	F: CCCCTCACTTCTCCAATCTC R: TCCCTCCTTCTTTTGTGTTT	54	90 - 100
<i>Pmac101</i>	(TC) ₁₅ TATG (TA) ₃	F: TGATAGCTGCAACGATGTC R: CATTGTGTTTGTGTTTGTGTTT	52	108 - 144
<i>Pmac102</i>	(AG) ₉	F: TCTATAAGTGCCGATGGATGC R: CACACCTCACTCCCAACCTC	60	118 - 144
<i>Pmac108</i>	(AG) ₁₄ (TG) ₇	F: TGAATGTCTGCCATTAGAAAGC R: TAATCGCCCCAAAGAAGAAC	52	117 - 165
<i>Pmac128</i>	(TC) ₅ TT(TC) ₁₀ (AC) ₁₀	F: GTGTTGATTGTACTCTTCAG R: CTAACCCTTTGTATACATGC	60	126 - 154
<i>Pmac130</i>	(AG) ₇ CA(AG) ₁₂	F: GAGGTGCCAATAAATCG R: TGTCACGCAATCTTGAACC	58	131 - 161
<i>Pmac146</i>	(AG) ₂₀	F: ACCCGACATCCCCTTGTAG R: TAGTCTGAAACGGAGCAAGG	64	131 - 173
<i>Pmac149</i>	(TC) ₁₉	F: TTCATCCTGCTTTTGAAGTTTG R: TGATGGATTAGGATTGACCTG	62	156 - 178

*A variação alélica apresentada considera os 18pb do iniciador universal M13, assim como apresentado em Perez et al., (2011).

As reações de amplificação dos loco *Pmac101* e *Pmac108* foram modificadas em algumas populações, pois apresentaram altos índices de amplificações inespecíficas. Nesses casos, foi adotada a modificação sugerida por Arruda et al. (2010), que consiste em adicionar o iniciador marcado com fluorocromo somente nos oito últimos ciclos da PCR, visando diminuir as amplificações inespecíficas.

Os produtos de PCR marcados foram enviados ao Centro de Estudos do Genoma Humano (CEGH) do Instituto de Biociências da Universidade de São Paulo ou ao Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias (FCAV) do

Perez, MF

campus Jaboticabal da Universidade Estadual Paulista, onde foram genotipadas em um seqüenciador *ABI PRISM 3730 DNA Analyzer* (Applied Biosystems) utilizando *LIZ 500* (Applied Biosystems) como padrão de tamanho. As genotipagens foram analisadas com o programa *GeneMarker v1.91* (SoftGenetics).

2.3 Análise dos dados

2.3.1 Índices de Diversidade

A presença de indivíduos com genótipos idênticos foi testada a partir do método implementado no programa GIMLET 1.3.3 (VALIÈRE, 2001), uma vez que os indivíduos de *P. machrisii* podem apresentar propagação vegetativa, o que resultaria na análise do mesmo genótipo mais de uma vez.

A diversidade genética dentro das populações foi estimada utilizando o programa FSTAT (GOUDET, 1995), por meio da estimativa das frequências alélicas, número de alelos por loco (A), heterozigosidade observada (H_o) e esperada (H_e). O parâmetro H_e foi estimado para cada população por meio do método não viesado de Nei (1987).

Além disso, a riqueza alélica (A_R), uma medida do número de alelos estimada independentemente do tamanho amostral também foi calculada. Essa estimativa adota o menor tamanho amostral entre as populações analisadas, permitindo comparações menos viesadas entre amostras com tamanhos diferentes (EL MOUSADIK e PETIT, 1996). As localidades BRO e ART foram desconsideradas dessa análise por apresentarem tamanhos amostrais muito menores que as outras populações.

Para verificar a significância estatística das diferenças nos níveis diversidade genética entre as populações foram comparadas as estimativas de heterozigosidade esperada (H_e) e de número de alelos por loco (A), por meio do teste de Wilcoxon (*Wilcoxon Sign Rank Test*,

Wilcoxon, 1945), esses testes foram realizados no programa R (R DEVELOPMENT CORE TEAM, 2012).

O nível de endogamia dentro das populações foi estimado usando o programa FSTAT (GOUDET, 1995), a partir dos desvios médios em relação às proporções do Equilíbrio de Hardy-Weinberg (EHW) e pelo cálculo do coeficiente de endocruzamento (F_{IS}) da estatística- F de Wright (WRIGHT, 1951).

A ocorrência de desequilíbrio de ligação entre pares de locos também foi testada em cada uma das populações com o programa Arlequin 3.11 (EXCOFFIER et al., 2005), por meio de 10000 permutações. A significância estatística dos testes foi realizada após a aplicação de correção seqüencial de Bonferroni para múltiplos testes com $\alpha=0,05$ (RICE, 1989).

2.3.2 Diferenciação Populacional

A diferenciação entre as populações para os marcadores microssatélites foi avaliada por meio do índice de fixação F_{ST} , calculado a partir do estimador θ (WEIR e COCKERHAM, 1948) com o programa FSTAT (GOUDET, 1995). Esse parâmetro pode ser definido como a proporção de diversidade genética devido às diferenças alélicas entre as populações ou como a correlação entre os alelos de uma população em relação a todas as populações (HOLSINGER e WEIR, 2009). O nível de diferenciação populacional também foi evidenciado a partir de estimativas do número e frequência de alelos exclusivos em cada uma das populações por meio do programa GENALEX6 (PEAKALL e SMOUSE, 2006). Somente alelos com frequência maior que 10% foram considerados nessa análise.

2.3.3 Estrutura Populacional

A estrutura populacional em *P. machrisii* foi investigada a partir da análise *bayesiana* implementada no software STRUCTURE 2.3 (PRITCHARD et al., 2000; HUBISZ et al., 2009). Esta análise tem como objetivo inferir o número mais provável de grupos de indivíduos intercruzantes (K) na amostra total. O delineamento dos grupos é realizado descartando a informação sobre a localização geográfica dos indivíduos e com base apenas em seus genótipos, buscando identificar agrupamentos populacionais com os menores índices de desequilíbrio de ligação e maximizando as proporções do EHW. A quantidade mais provável de grupos genéticos (K) foi estimada a partir dos resultados de dez simulações simultâneas independentes, permutados com o programa CLUMPP (JAKOBSSON e ROSENBERG, 2007). Para cada uma das simulações, 1.100.000 interações foram realizadas, sendo as 100.000 interações iniciais descartadas (burn-in), com número de grupos variando de um a 13, de acordo com a quantidade de populações amostradas. Foi utilizado o modelo que admite mistura de genótipos entre as populações e frequências alélicas correlacionadas. A representação gráfica dos resultados foi realizada com o programa DISTRUCT (ROSENBERG, 2004). Os resultados foram analisados a partir da estatística ΔK proposta por Evanno et al. (2005), que torna mais objetiva a interpretação do número provável de grupos em comparação com o método proposto originalmente (PRITCHARD et al., 2000). Dado que a análise dos resultados por meio da estatística ΔK apresenta a tendência de capturar somente a estruturação primária presente nos dados (EVANNO et al., 2005), os agrupamentos formados com valores maiores de K também foram avaliados.

A distribuição da variabilidade genética dentro e entre os grupos formados pela análise *bayesiana* no programa STRUCTURE foi investigada pela análise de variância molecular (AMOVA) hierárquica a partir dos valores de F_{ST} , com o auxílio do programa Arlequin 3.11 (EXCOFFIER et al., 2005). O nível principal de estruturação genética é determinado pelos

agrupamentos em que a variação genética entre as populações de grupos diferentes é maior que a variação dentro desses grupos.

A estruturação genética de *P. machrisii* também foi analisada por meio do método *bayesiano* implementado no programa SAMOVA (DUPANLOUP et al. 2002), o qual considera a variação genética e a localização geográfica das populações para definir grupos de populações similares geneticamente e próximas geograficamente. Esse método busca maximizar a diferenciação genética entre grupos (F_{CT}), sem utilizar premissas em relação ao EHW ou desequilíbrio de ligação entre os locos (DUPANLOUP et al., 2002). Os agrupamentos geográficos foram determinados por meio de 100 simulações a partir dos valores de F_{ST} . O número de grupos (K) foi obtido testando valores crescentes, até que fosse observada uma maior variação genética entre grupos em relação à variação dentro de cada grupo.

A partir dos agrupamentos formados pela estruturação principal sugerida pelas análises com STRUCTURE e SAMOVA, foi realizada uma segunda análise com o programa STRUCTURE para inferir a presença de subestruturação genética, de acordo com o procedimento apresentado por Vähä et al. (2007). Nesse procedimento os indivíduos de cada um dos grupos genéticos determinados pela estatística ΔK são retirados do conjunto de dados e analisados separadamente, até que nenhuma estruturação genética seja detectada.

2.3.4 Detecção de Migrantes

A presença de indivíduos migrantes nas populações analisadas foi inferida a partir do método *bayesiano* implementado no programa STRUCTURE 2.3 (PRITCHARD et al., 2000; HUBISZ et al., 2009). Nessa análise, as informações geográficas foram incorporadas e o número de grupos foi fixado de acordo com o número de populações amostradas ($K=12$). Esta análise estima a probabilidade posterior dos indivíduos pertencerem à população onde foram coletados ou de serem migrantes ou descendentes de migrantes até a segunda geração. A análise identifica migrantes a partir do estabelecimento de uma probabilidade mínima (0,5) do genótipo de um determinado indivíduo pertencer à população onde este foi amostrado. Indivíduos com valores abaixo desse valor de corte foram considerados migrantes ou descendentes de migrantes e a provável população-fonte foi determinada pelo maior valor de probabilidade do genótipo desse indivíduo entre as populações amostradas.

Uma segunda estatística proposta por Cornuet et al. (1999) e implementada no programa GENECLASS2 (PIRY et al., 2004) também foi realizada para complementar os resultados obtidos na análise anterior. Diferentemente do procedimento realizada pelo programa STRUCTURE, a análise proposta por Cornuet et al. (1999) permite detectar potenciais migrantes nas populações analisadas provenientes de populações não amostradas (MANEL et al. 2002).

A análise com o programa GENECLASS foi realizada em duas etapas, primeiramente, somente a probabilidade de um indivíduo ser migrante de uma população não amostrada foi estimada, por meio dos valores de verossimilhança de um determinado indivíduo pertencer à população onde foi amostrado (L_{home}). Esse método é considerado apropriado quando não há certeza de que todas as populações foram amostradas (PAETKAU et al., 2004; PIRY et al., 2004). Além disso, a probabilidade de um indivíduo ser migrante de uma população amostrada foi estimada por meio da razão entre a verossimilhança de um indivíduo pertencer

à população onde foi amostrado e o valor máximo de verossimilhança desse indivíduo entre as populações amostradas ($L_{\text{home}}/L_{\text{max}}$). Essa estatística foi calculada por ser considerada mais informativa quando todas as populações são amostradas (PAETKAU et al., 2004). Para a realização das duas análises, foram utilizados o critério *bayesiano* de Rannala e Mountain (1997) e o método de Paetkau et al., (2004) a partir de 10.000 reamostragens. O parâmetro *alpha*, que determina a probabilidade de ocorrência de erros do tipo I, foi fixado em 0,01, de acordo com o procedimento sugerido por Paetkau et al. (2004) para dados simulados.

2.3.5 Relação Entre as Distâncias Genéticas e Geográficas

A correlação entre as distâncias geográficas e genéticas foi estimada a partir de um teste de Mantel (MANTEL, 1967). As distâncias genéticas foram estimadas pelo parâmetro F_{ST} para cada um dos pares de população, linearizados de acordo com o método proposto por Slatkin (1995) pela fórmula $F_{ST}/(1-F_{ST})$. As probabilidades dos valores foram obtidas por meio de 999 permutações. Todas as estimativas referentes ao teste de Mantel foram realizadas com o programa GENALEX6 (PEAKALL e SMOUSE, 2006). A análise foi realizada primeiramente com todo o conjunto de dados e posteriormente considerando os agrupamentos obtidos pelo programa STRUCTURE.

Com o objetivo de inferir em quais escalas de distância geográfica esta associação foi mais importante, a autocorrelação espacial (geográfica) da diferenciação genética também foi estimada de acordo com o método implementado no programa GENALEX6 (PEAKALL e SMOUSE, 2006). Nessa análise também foram utilizados os valores de F_{ST} linearizados para cada par de populações. Esse método estima a similaridade genética entre pares de populações dentro de diferentes classes de distância geográfica, permitindo revelar a possível associação da diferenciação genética com a distância geográfica em diferentes escalas geográficas (SMOUSE e PEAKALL, 1999). Foram estabelecidas 12 classes de distâncias (0-121; 121-

170; 170-229; 229-302; 302-354; 354-522; 522-681; 681-741; 741-799; 799-934 e 934-1184, distâncias medidas em quilômetros), selecionadas de forma a igualar o número de pares de populações em cada uma das classes.

2.3.6 Relações Genéticas entre as Populações

A relação de similaridade genética entre as populações foi analisada através da distância genética D_c “chord distance” (CAVALLI-SFORZA e EDWARDS, 1967) entre pares de populações. De acordo com Laval et al. (2002), esse método de distância é apropriado para estimar a diferenciação entre populações ou espécies altamente relacionadas, com pouco tempo de divergência. Outra característica importante é o fato das distâncias genéticas estimadas por esse método serem livres da influência causada por alterações no tamanho populacional ou diferenças na taxa de mutação entre os locos, dando ênfase apenas ao número e frequência de alelos compartilhados entre populações (CHAPUIS e ESTOUP, 2007). Além disso, a utilização de locos de microssatélites compostos ou interrompidos, os quais frequentemente violam as premissas dos modelos mutacionais propostos especificamente para microssatélites (ESTOUP et al., 2002), realçam a importância de utilizar um método de distância como o D_c , que não considera diferenças nas taxas de mutação entre os locos.

Os valores de distância obtidos foram utilizados na construção de um dendrograma pelo método UPGMA (Unweighted Pair Group Method with Arithmetic Mean; SNEATH e SOKAL, 1973) e na realização de uma análise de coordenadas principais (PCA).

A construção do dendrograma foi realizada como o programa POPULATIONS (LANGELLA, 1999; Figura 6), com a consistência dos ramos estimada a partir de 1.000 pseudoreplicações de *bootstrap* nos indivíduos. A fim de avaliar a confiabilidade da árvore

obtida em representar a matriz de distâncias, o coeficiente de correlação cofenética foi estimado com o programa NTSYSpc2.1 (ROHLF, 2000), a partir de 1000 permutações.

Para complementar as relações entre as populações observadas com a análise *bayesiana* de estruturação genética hierárquica e com a árvore UPGMA, duas análises de coordenadas principais (PCA) foram realizadas com o programa GENALEX6 (PEAKALL e SMOUSE, 2006), utilizando a distância D_c (CAVALLI-SFORZA e EDWARDS, 1967). Primeiramente foram consideradas as distância calculadas entre todos os indivíduos, para verificar o nível de estruturação principal. A segunda análise foi realizada com as distâncias médias entre populações, facilitando a visualização dos outros níveis hierárquicos.

3. RESULTADOS

3.1 Índices de Diversidade

A análise de detecção de genótipos idênticos não resultou em nenhum par de indivíduos identificados como possíveis clones, com isso todos os indivíduos foram mantidos nas análises posteriores.

Os níveis de diversidade obtidos foram relativamente similares nas populações analisadas (Tabela 3), com exceção das amostras populacionais de BRO ($H_E=0,302$) e BAR ($H_E=0,351$; $A_R=2,7$), que apresentaram índices menores. As populações restantes apresentaram valores de heterozigosidade esperada variando de 0,428 em FUR a 0,570 em APA2. Os valores de riqueza alélica variaram de 3,2 em FOR a 4,3 alelos em APA2. No entanto, as diferenças entre os índices de diversidade entre BAR e as outras populações somente foram estatisticamente significativas em relação a POS (para H_E) e em relação a APA2 e URU (para A), de acordo com teste de Wilcoxon ($P<0,05$). No caso de BRO, as estimativas de H_E foram estatisticamente diferentes de APA2 e POS, enquanto as estimativas de número de alelos por loco (A) foram significativamente diferentes em relação a todas as populações, exceto FMS, BAR e ART.

Desvios significativos em relação às proporções do EHW foram observados em sete populações após a correção de Bonferroni (RICE, 1989; $\alpha=0,05$). Desse total, cinco populações apresentaram desvios em apenas um ou dois locos (DEL, APA1, APA2, URU e CRI) e as outras duas não apresentaram desvios em nenhum loco específico (ART e POS). O loco *Pmac146* apresentou desvios em oito populações. As populações com seus valores do coeficiente de endogamia (F_{IS}) e os locos e que apresentaram estimativas significativas estão representados na Tabela 3.

Foram realizados no total 540 testes de desequilíbrio de ligação, seis pares de locos (1,11% do total) apresentaram resultados significativos após a correção de Bonferroni

(RICE, 1989; $\alpha=0,05$), sendo somente dois testes significativos entre o mesmo par de locos (*Pmac101* e *Pmac146*) e três estatisticamente significativos para a mesma população (POS).

Tabela 3. Parâmetros de variabilidade genética estimados ao nível populacional em *P. machrisii*. *A*, número médio de alelos por loco; H_O , heterozigiosidade média observada; H_E , heterozigiosidade média esperada, estimada pelo método de Nei (1987); A_R , riqueza alélica; F_{IS} , valor do coeficiente de endocruzamento; EHW, locos com desvios significativos em relação às proporções do Equilíbrio de Hardy-Weinberg.

População	N	A	H_O	H_E	A_R	F_{IS}	EHW
Brotas-SP (BRO)	10	2,1	0,250	0,302	-	0,221	-
Altinópolis-SP (FOR)	21	3,9	0,379	0,443	3,2	0,169	<i>Pmac146</i>
Furnas-MG (FUR)	21	4,0	0,367	0,428	3,3	0,167	<i>Pmac146</i>
Delfinópolis-MG (DEL)	30	4,5	0,413	0,500	3,5	0,190*	<i>Pmac101</i> <i>Pmac146</i>
Formosa-GO (FMS)	22	4,1	0,489	0,516	3,4	0,084	-
Alto Paraíso1-GO (APA1)	20	5,2	0,392	0,510	4,2	0,259*	<i>Pmac146</i>
Alto Paraíso2-GO (APA2)	25	5,5	0,400	0,570	4,3	0,323*	<i>Pmac84</i> <i>Pmac146</i>
Aurora do Tocantins (ART)	8	4,0	0,417	0,513	-	0,280*	-
Uruaçu-GO (URU)	20	4,4	0,411	0,497	3,6	0,201*	<i>Pmac101</i> <i>Pmac146</i>
Cristalina-GO (CRI)	25	4,5	0,350	0,470	3,6	0,281*	<i>Pmac128</i> <i>Pmac146</i>
Posse-GO (POS)	18	4,5	0,436	0,559	3,8	0,252*	-
Barreiras-BA (BAR)	21	3,2	0,300	0,351	2,7	0,170	<i>Pmac146</i>
Média	20,08	4,16	0,384	0,472	3,56	0,206*	

*Testes significativos após correção de Bonferroni (RICE, 1989; $\alpha=0,05$)

3.2 Diferenciação Populacional

O resultado obtido para a estimativa de F_{ST} apontou uma alta e significativa diferenciação entre as populações ($\theta=0,357$; IC95%: 0,272-0,458).

Alelos exclusivos com elevada frequência foram verificados em seis populações. A população ART apresentou o maior número (oito) de alelos exclusivos com frequência acima de 10%, enquanto as populações BRO e CRI apresentaram um alelo cada e as populações POS, URU e BAR apresentaram três desses alelos (Tabela 4).

Tabela 4. Alelos exclusivos de DNA microssatélite com frequências acima de 10% em populações de *P. machrisii*.

População (Loco)	Alelo	Frequência	População (Loco)	Alelo	Frequência
BRO (<i>Pmac108</i>)	117	0,650	URU (<i>Pmac108</i>)	161	0,500
ART (<i>Pmac102</i>)	132	0,167	URU (<i>Pmac149</i>)	176	0,275
ART (<i>Pmac102</i>)	140	0,250	CRI (<i>Pmac82</i>)	105	0,200
ART (<i>Pmac102</i>)	142	0,167	POS (<i>Pmac85</i>)	94	0,250
ART (<i>Pmac102</i>)	144	0,250	POS (<i>Pmac108</i>)	145	0,115
ART (<i>Pmac108</i>)	143	0,333	POS (<i>Pmac108</i>)	151	0,192
ART (<i>Pmac130</i>)	153	0,250	BAR (<i>Pmac82</i>)	99	0,238
ART (<i>Pmac130</i>)	161	0,188	BAR (<i>Pmac128</i>)	126	0,238
ART (<i>Pmac149</i>)	178	0,125	BAR (<i>Pmac130</i>)	131	0,167
URU (<i>Pmac108</i>)	159	0,278			

3.3 Estrutura Populacional

A análise *bayesiana* do programa STRUCTURE a partir do cálculo ΔK (Figura 3) apontou para a estruturação dos indivíduos em dois grupos principais ($K=2$). O primeiro grupo foi formado por populações localizadas no sul da distribuição da espécie (BRO, FUR, DEL e FOR). Um segundo agrupamento foi formado pelas populações CRI, FMS, POS, APA1, APA2, URU, BAR e ART, localizadas nas porções central e norte da distribuição da espécie. Foi observado um segundo pico na curva dos valores de ΔK para os resultados da estruturação em quatro grupos ($K=4$; Figura 3). Neste resultado, um primeiro grupo, denominado Sudeste, foi composto por indivíduos de populações da região Sudeste do Brasil (BRO, FUR, DEL e FOR). Os indivíduos das populações FMS, APA1, APA2 e ART, localizadas no Brasil Central (norte de Goiás e sul do Tocantins), formaram um segundo

grupo, denominado Central. Um terceiro grupo foi formado pelos indivíduos da população URU, localizada também no Brasil Central. As amostras da região noroeste do Estado de Goiás, pertencentes a duas populações localizadas no leste (CRI) e nordeste (POS) do Estado de Goiás, em conjunto com a população BAR no oeste do Estado da Bahia, formaram o quarto grupo, denominado Nordeste.

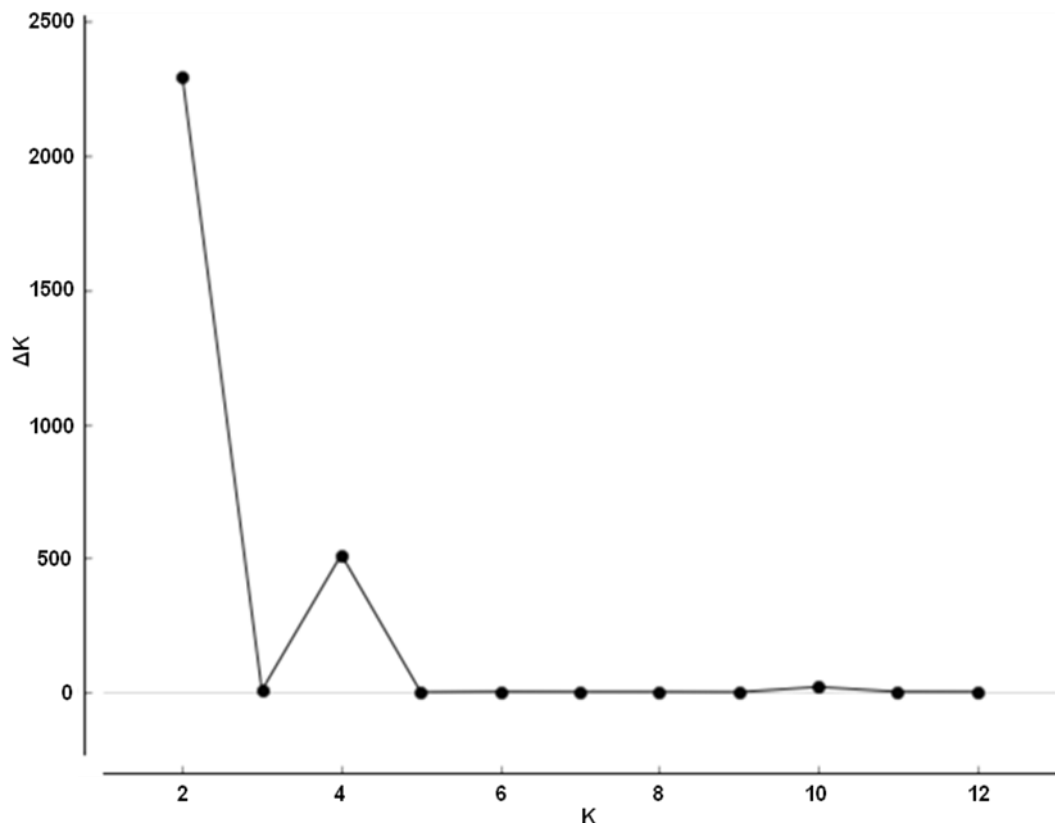


Figura 3. Resultado gráfico da análise para identificação do valor mais provável de K a partir da estatística ΔK (EVANNO et al., 2005), calculado por meio de 10 replicações.

A AMOVA hierarquizada considerando esses dois valores de K (K=2 e K=4), assim como o valor intermediário entre eles (K=3), demonstrou que somente a organização das populações em quatro grupos apresentou maior variação entre grupos em relação àquela observada dentro dos grupos (Tabela 5), evidenciando que esse agrupamento explicou melhor a estrutura populacional principal.

Os resultados da SAMOVA também indicaram que o menor número de grupos em que a variação entre os grupos de populações foi maior que a variação dentro dos grupos foi K=4.

Nesse resultado, as populações foram reunidas nos mesmos grupos apontados pelo programa STRUCTURE. Para $K=2$, o resultado da SAMOVA formou um agrupamento diferente daquele do programa STRUCTURE, com um dos grupos contendo somente BAR e o outro contendo todas as outras populações. Esse agrupamento apresentou um índice de variação entre as populações maior que o agrupamento sugerido pelo STRUCTURE (Tabela 5), embora o valor de F_{SG} para essa análise não foi significativo. Os resultados para as análises AMOVA e SAMOVA estão representados na Tabela 5.

Tabela 5. Análise de variância molecular para microssatélites em *P. machrisii* usando diferentes níveis hierárquicos. As populações foram agrupadas de acordo com os resultados da análise com os programas STRUCTURE (PRITCHARD et al., 2000) ou SAMOVA (DUPANLOUP et al., 2002).

Fonte de Variação	% de variação	Estatística F
Dois Grupos (SAMOVA)		
{BRO, FUR, DEL, FOR, FMS, APA1, APA2, ART, URU, CRI e POS} {BAR}		
Entre Grupos	19,56	$F_{GT}=0,196^{NS}$
Entre pop. mesmo grupo	27,21	$F_{SG}=0,338$
Dois Grupos (STRUCTURE)		
{BRO, FUR, DEL e FOR} {FMS, APA1, APA2, ART, URU, CRI, POS e BAR}		
Entre Grupos	15,98	$F_{GT}=0,160$
Entre pop. mesmo grupo	23,78	$F_{SG}=0,283$
Três Grupos (SAMOVA e STRUCTURE)		
{BRO, FUR, DEL e FOR} {FMS, APA1, APA2, ART e URU} {CRI, POS e BAR}		
Entre Grupos	20,76	$F_{GT}=0,208$
Entre pop. mesmo grupo	19,41	$F_{SG}=0,245$
Quatro Grupos (SAMOVA e STRUCTURE)		
{BRO, FUR, DEL e FOR} {FMS, APA1, APA2 e ART} {URU} {CRI, POS e BAR}		
Entre Grupos	22,69	$F_{GT}=0,227$
Entre pop. mesmo grupo	17,00	$F_{SG}=0,220$

^{NS} Testes não significativos ($P>0,01$).

As análises AMOVA realizadas separadamente para cada um dos grupos formados em $K=4$ com mais de uma população (Tabela 6) demonstrou que o grupo Nordeste apresentou o

maior valor de diferenciação entre suas populações. O índice apresentado por essas populações ($F_{ST}=0,321$) foi maior que aquele obtido entre grupos para $K=4$ ($F_{GT}=0,227$), evidenciando a elevada diferenciação entre as populações desse grupo.

Tabela 6. Análise de variância molecular para microssatélites em *P. machrisii* para os indivíduos pertencentes aos grupos formados com $K=4$ pela análise do programa STRUCTURE (PRITCHARD et al., 2000).

Fonte de Variação	% de variação	Estatística F^*
Grupo Sudeste {BRO, FUR, DEL e FOR}		
Entre Populações	22,06	$F_{ST}=0,221$
Grupo Central {FMS, APA1, APA2 e ART}		
Entre Populações	12,54	$F_{ST}=0,125$
Grupo Nordeste {CRI, POS e BAR}		
Entre Populações	32,06	$F_{ST}=0,321$

*Todos os testes foram significativos ($P<0,01$).

As análises de segundo nível com o programa STRUCTURE para os indivíduos de cada um dos grupos separadamente, realizado a partir dos agrupamentos formados com $K=4$, apresentou uma subestrutura adicional somente para Sudeste e Nordeste, nos quais os indivíduos foram reunidos em três subgrupos ($K=3$). No Grupo Sudeste, os indivíduos das duas populações amostradas no Estado de Minas Gerais (FUR e DEL) foram reunidos, enquanto os indivíduos das duas populações amostradas no Estado de São Paulo (FOR e BRO) formaram dois subgrupos separados. No Grupo Nordeste, os indivíduos das populações CRI, POS e BAR foram separados em três subgrupos. Este resultado reflete a alta diferenciação entre as três populações do subgrupo Nordeste. Os resultados gráficos da análise do STRUCTURE para valores de $K=2-4$, bem como os resultados das análises de segundo nível com $K=4$ estão representados na Figura 4. A proporção do genótipo dos

indivíduos atribuída a cada um desses grupos em cada população analisada está representado na Figura 5.

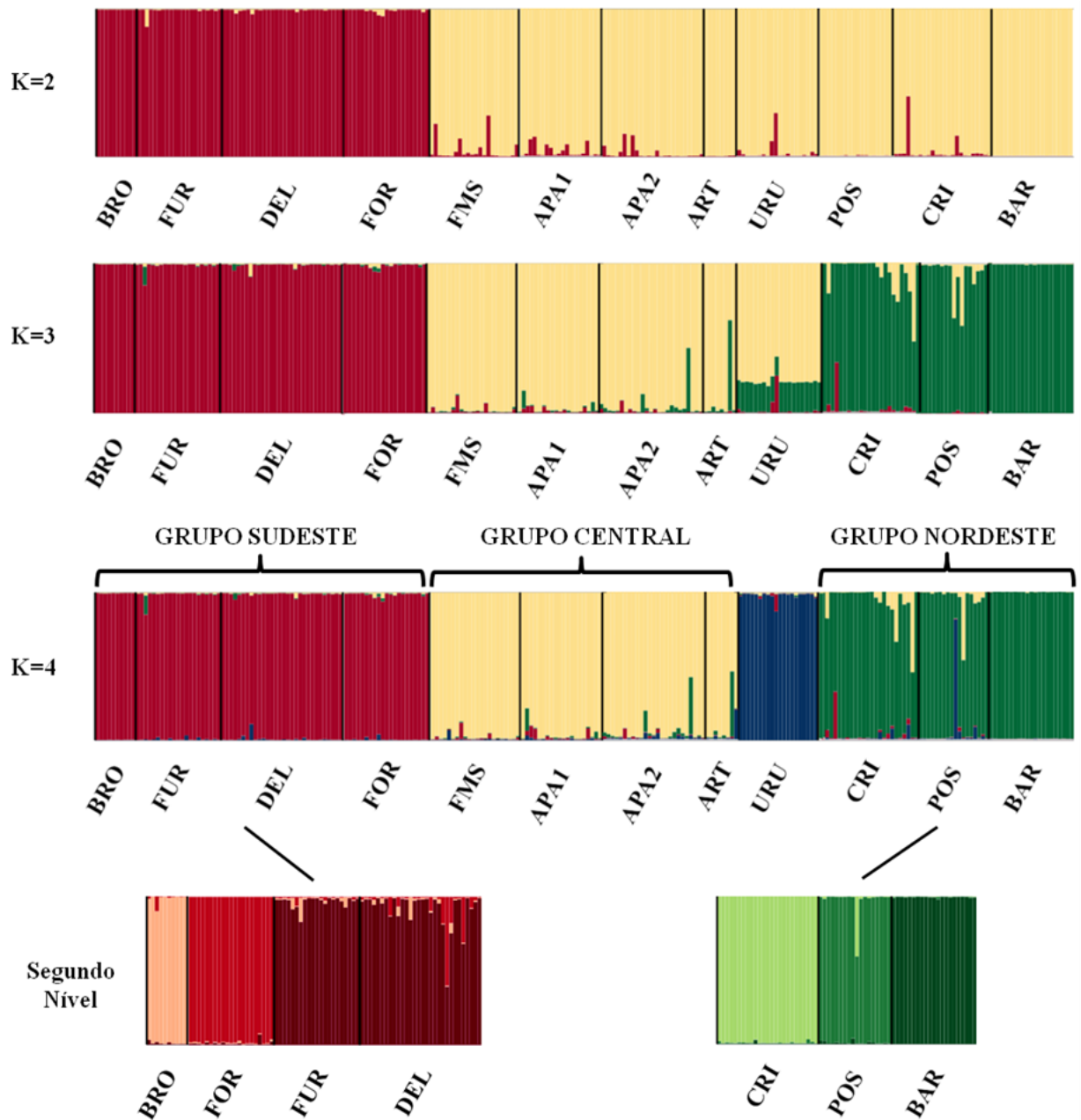


Figura 4. Esquema dos resultados obtidos a partir da análise bayesiana hierárquica do programa STRUCTURE (PRITCHARD et al., 2000). Cada indivíduo é representado por uma barra vertical, apresentando a proporção de seu genoma pertencente a cada um dos grupos estabelecidos,. Linhas pretas separam os indivíduos de diferentes populações. As abreviações utilizadas paraas populações estão representadas na Tabela 1.

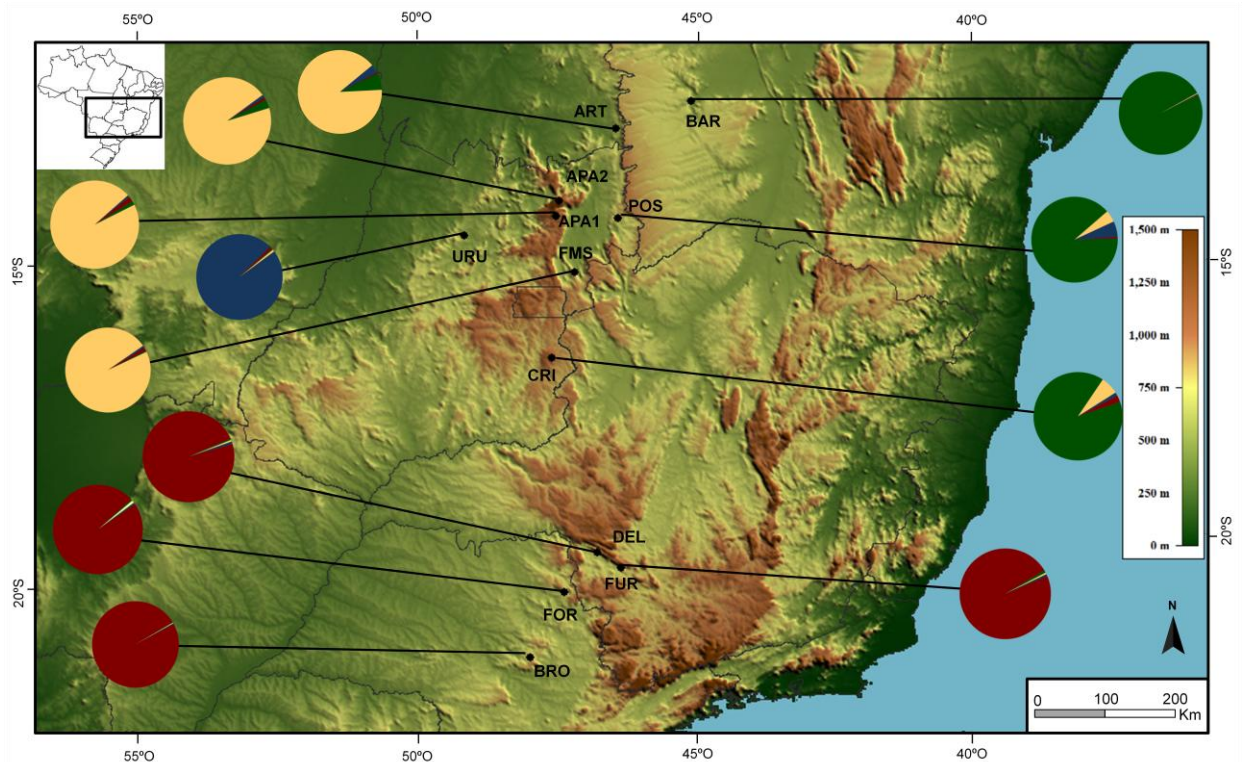


Figura 5. Mapa de elevação (medida em metros) da região central e leste do Brasil apresentando a proporção média do genoma dos indivíduos em cada população atribuída aos grupos determinados pelo programa STRUCTURE (PRITCHARD et al., 2000; K=4). As cores representadas seguem o esquema apresentado na Figura 4.

3.4 Detecção de Migrantes

As análises de identificação de migrantes realizadas com o programa STRUCTURE resultaram na detecção de um indivíduo na população POS (S16A10) como potencial migrante de URU e um indivíduo na população CRI (S6A24) como provável migrante de APA1. Os resultados apresentados pelo programa GENECLASS2 confirmaram esses resultados para os mesmos indivíduos, além de apontarem cinco indivíduos adicionais como potenciais migrantes de outras populações amostradas e outros cinco prováveis migrantes de populações não amostradas (Tabela 7).

Tabela 7. Resultados da análise de detecção de migrantes. São apresentadas as populações amostradas e as prováveis populações-fonte de acordo com os programas STRUCTURE (PRITCHARD et al., 2009) e GENECLASS (Piry et al., 2004).

Indivíduo migrante	População amostrada	GENECLASS	STRUCTURE
		Provável pop-fonte	Provável pop-fonte
S16A10	POS	URU	URU
S6A24	CRI	APA1	APA1
S4A1	FUR	DEL	-
S5A19	DEL	FUR	-
S17A18	APA1	APA2	-
S34A24	APA2	APA1	-
S33A6	ART	POS	-
S2A3	BRO	?	-
S5A10	DEL	?	-
S17A2	APA1	?	-
S18A3	URU	?	-
S22A15	BAR	?	-

- inferência de migração não apontada pelo programa STRUCTURE; ? População fonte não amostrada.

3.5 Relação Entre as Distâncias Genéticas e Geográficas

As estimativas de F_{ST} par a par foram significativas entre todos os pares de populações (Tabela 8). O maior valor de F_{ST} foi observado entre as populações BAR e BRO, as quais ocorrem nos extremos norte e sul da distribuição da espécie. O menor valor de F_{ST} foi observado entre as populações FUR e DEL, que estão separadas por apenas 55 km.

O resultado do teste de correlação de distância genética e geográfica entre todas as populações apontou associação significativa entre as duas variáveis ($r=0,676$; $P < 0,001$), indicando que 45,7% (r^2) da variação genética observada entre as populações pode ser explicada pelas distancias geográficas entre elas. Entretanto, quando os grupos genéticos obtidos para $K=4$ com mais de uma população foram analisados separadamente, não foram

verificadas correlações significativas (grupo Sudeste: $r=0,628$ com $P=0,180$; grupo Central: $r=0,163$ com $P=0,360$; grupo Nordeste $r=-0,971$ com $P=0,344$).

Tabela 8. Matriz das distâncias geográficas par-a-par (em quilômetros, diagonal inferior) e dos valores par-a-par do parâmetro F_{ST} (diagonal superior) para as populações de *P.machrisii*.

População	BRO	FUR	DEL	CRI	FOR	FMS	POS	APA1	URU	BAR	APA2	ART
BRO	-	0,304	0,249	0,509	0,349	0,420	0,484	0,419	0,471	0,626	0,419	0,475
FUR	256	-	0,087	0,429	0,255	0,322	0,396	0,304	0,386	0,543	0,341	0,355
DEL	256	55	-	0,368	0,199	0,274	0,368	0,275	0,343	0,501	0,313	0,296
CRI	621	448	406	-	0,466	0,358	0,235	0,315	0,362	0,372	0,309	0,341
ALT	156	116	102	479	-	0,346	0,437	0,376	0,416	0,555	0,369	0,394
FOR	798	608	572	179	652	-	0,278	0,151	0,310	0,463	0,126	0,129
POS	918	711	680	311	768	138	-	0,302	0,306	0,273	0,252	0,245
API	909	727	689	288	766	119	121	-	0,314	0,488	0,122	0,155
URU	878	741	694	301	752	228	303	184	-	0,470	0,267	0,307
BAR	1184	961	937	589	1030	416	279	353	522	-	0,422	0,484
APA2	941	757	719	319	797	148	121	32	199	329	-	0,123
ART	1083	880	848	466	934	287	169	196	353	171	167	-

*Todos os testes foram significativos ($P<0,001$).

A análise de autocorrelação espacial apontou para uma correlação positiva e significativa entre distâncias genéticas e geográficas nas classes de 0-121 km, de 121-170 km e de 229-302 km (Figura 6), indicando que os indivíduos nessas classes são mais relacionados que o esperado somente por chance. Foi observada correlação negativa nas classes de 681-741 km e de 934-1184 km.

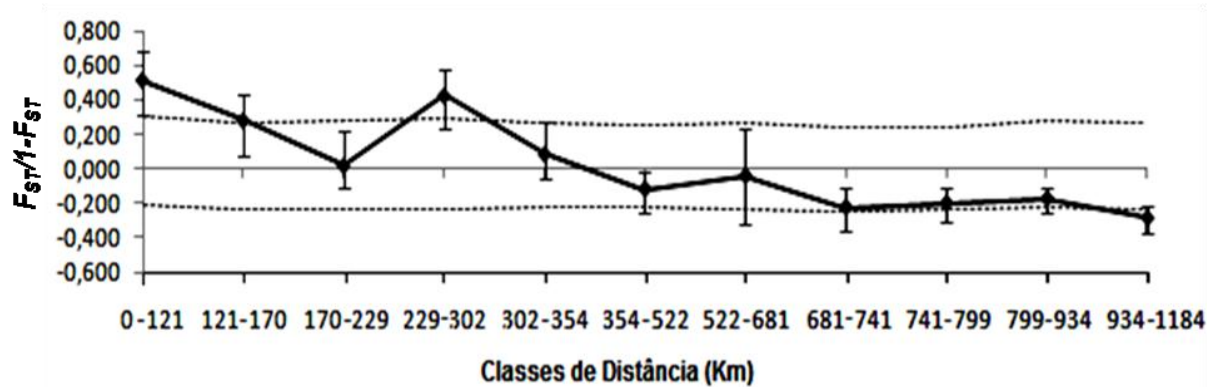


Figura 6. Representação da autocorrelação espacial da diferenciação genética apresentando os coeficientes médios de correlação entre pares de populações, plotados em relação às suas distancias geográficas subdivididas em 12 classes. Barras indicam o intervalo de confiança do coeficiente de parentesco (IC_{95%}) e linhas tracejadas indicam os valores críticos de rejeição (VC_{95%}) da hipótese nula de ausência de estruturação genética espacial.

3.6 Relações Genéticas entre as Populações

A partir do dendrograma UPGMA contruído a partir da matriz de distância D_c (Figura 7) foi possível observar os agrupamentos de populações de *P. machrisii* formados com base nas similaridades genéticas. Os agrupamentos formados foram compatíveis com o resultado de $K=4$ observado na análise de estruturação genética realizada com o programa STRUCTURE. As populações foram agrupadas em dois ramos principais com alto suporte de *bootstrap*. Um desses ramos agrupou as mesmas populações do grupo Sudeste revelado pela análise de estrutura populacional. Este agrupamento apresentou a mesma estruturação observada anteriormente, com as populações FUR e DEL demonstrando alta similaridade entre elas. O segundo ramo principal foi formado por dois grupos de populações altamente sustentado pelos valores de *bootstrap*. Um desses grupos foi formado pela população URU em um ramo e as populações do subgrupo Central em outro ramo. Dentro do segundo grupo, as populações do subgrupo Nordeste (CRI, POS e BAR) ficaram reunidas em um ramo com elevado suporte (80%). Dentro desse último ramo, por sua vez, pode ser observada uma maior similaridade entre BAR e POS em relação a CRI.

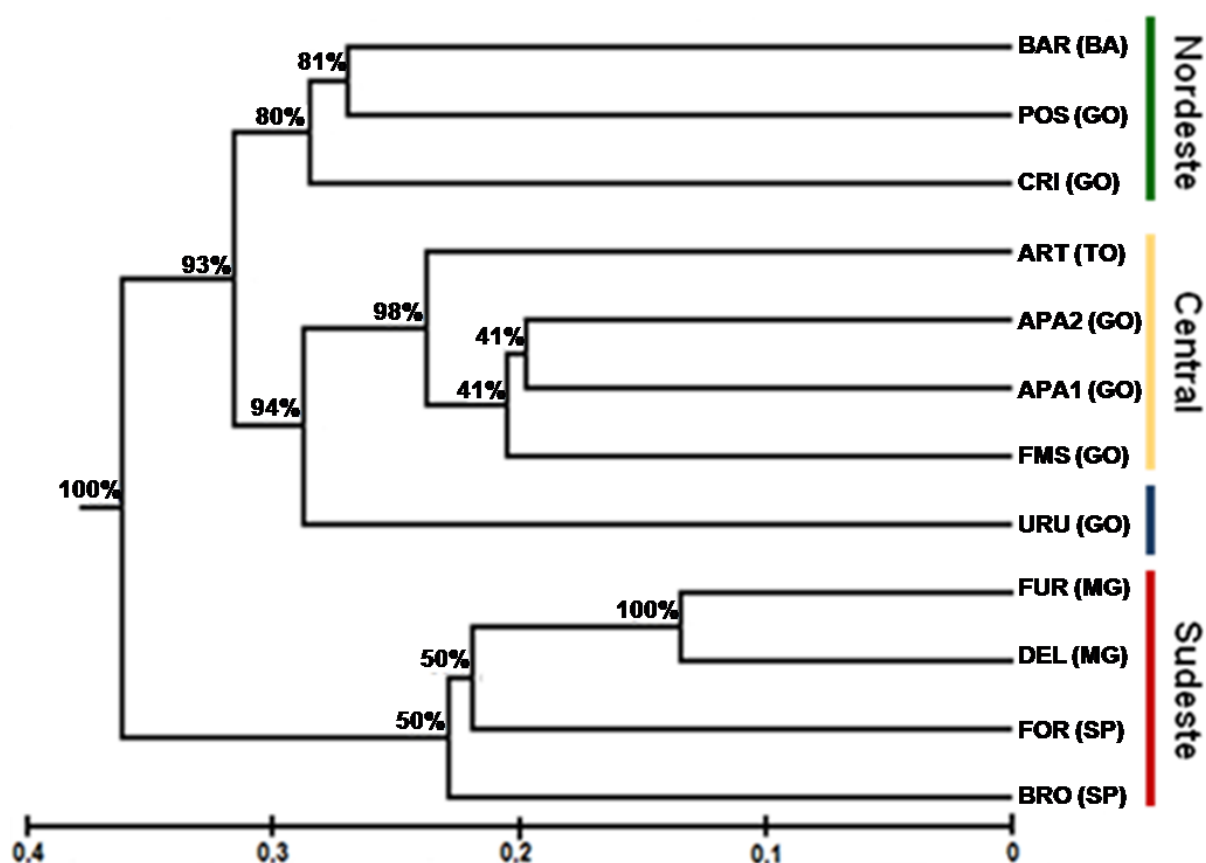


Figura 7. Dendrograma UPGMA construído a partir da distância genética D_c (CAVALLI-SFORZA e EDWARDS, 1967). Árvore baseada em 500 *bootstraps* nos indivíduos das populações. Os agrupamentos hierárquicos obtidos pela análise com o programa STRUCTURE (PRITCHARD et al., 2000) estão representados ao lado direito dos ramos. Coeficiente de Correlação Cofenética=0,82 ($P<0,005$).

A PCA utilizando as distâncias entre todos os indivíduos apresentou 57,41% da variação total explicada pelos dois primeiros eixos (eixo 1=34,9% e eixo 2=22,51%). Os resultados foram concordantes com as análises anteriores, ordenando os indivíduos dos grupos Sudeste, Central, URU e Nordeste em regiões distintas do gráfico (Figura 8a). Quando foram consideradas as distâncias entre as populações, os dois primeiros eixos principais explicaram 60,69% da variação total (eixo 1=38,9% e eixo 2=21,79%) e foi possível verificar as relações entre as populações de forma mais objetiva. Novamente foi possível verificar a grande diferenciação entre as populações do grupo Nordeste (Figura 8b).

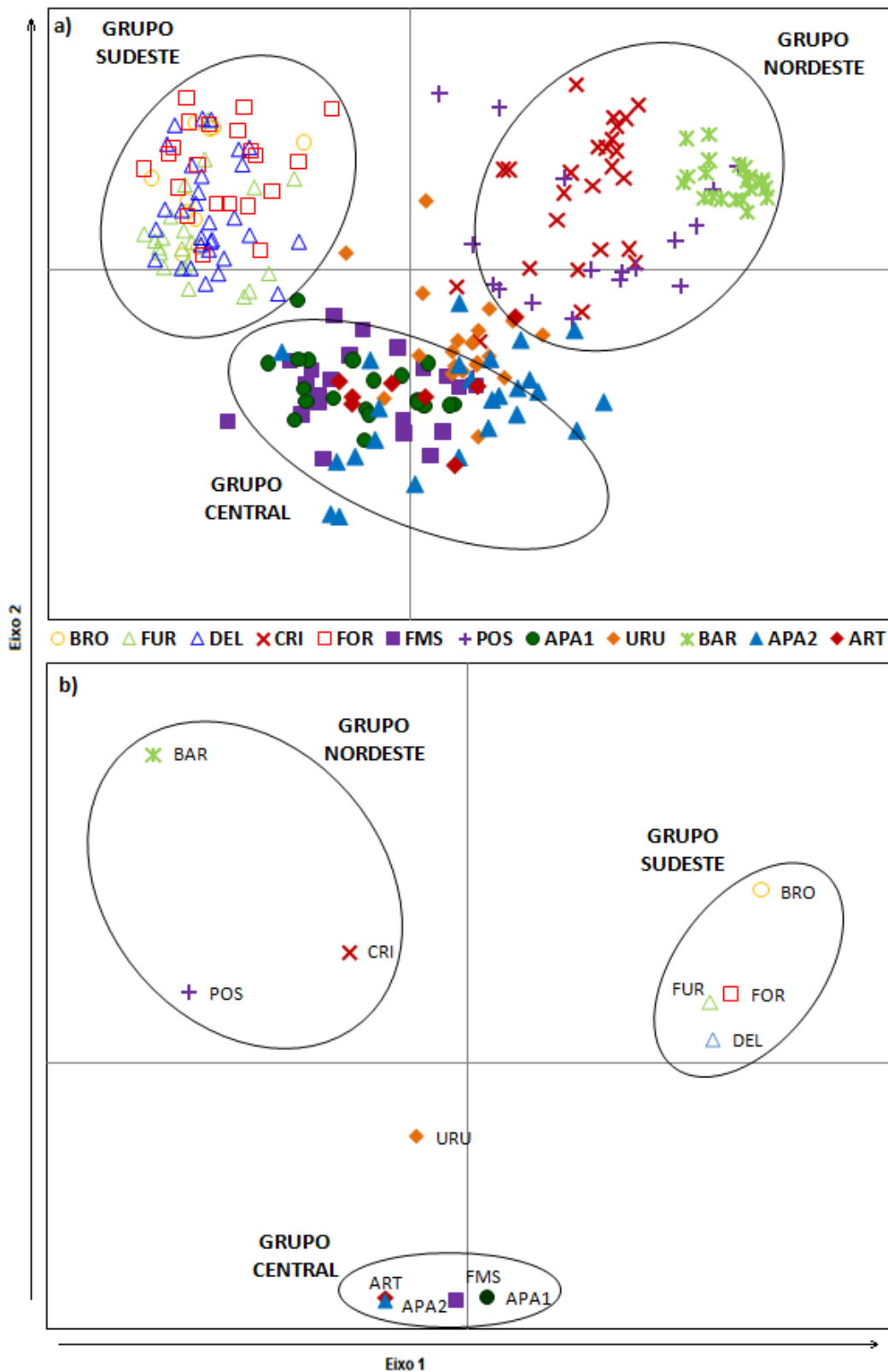


Figura 8. PCA a partir da distância genética D_C (CAVALLI-SFORZA e EDWARDS, 1967) mostrando a) divergência entre todos os indivíduos (57,41% da variação total nos dois eixos representados) e b) a divergência todas populações (60,69% da variação total nos dois eixos representados). As elipses representam os grupos revelados pelo STRUCTURE (PRITCHARD et al., 2000) com $K=4$.

4. DISCUSSÃO

Os resultados obtidos indicaram elevada diferenciação genética e níveis limitados de fluxo gênico recente entre as populações de *P. machrisii*, verificados a partir do alto valor do parâmetro F_{ST} e a ocorrência de alelos exclusivos em alta frequência, assim como esperado inicialmente de acordo com a distribuição fragmentada da espécie. Os testes realizados para detectar possíveis migrantes corroboram esses resultados, uma vez que somente dois entre os 241 indivíduos analisados foram apontados como migrantes pelos dois métodos utilizados. Os migrantes identificados somente com o programa GENECLASS devem ser analisados com cuidado, uma vez que esse método apresenta menor resolução quando comparado aos resultados do método implementado pelo programa STRUCTURE (MANEL et al., 2002).

Os testes de desequilíbrio de ligação não apontaram nenhum par de locos específico ligados entre si, uma vez que somente seis testes foram significativos em um total de 540 e em apenas um par de locos ocorreram testes significativos em mais de uma população (*Pmac101* e *Pmac146*). Outro resultado importante dessa análise apontou testes significativos em três pares de locos diferentes para a população POS. Esse resultado pode representar a ocorrência de mistura nessa população, o que é concordante a alta diversidade genética nessa população e com o teste de detecção de migrantes, que evidenciou a presença de um indivíduo proveniente de URU nessa população.

A dificuldade em obter amplificações em algumas populações, principalmente para o loco *Pmac135*, que foi descartado das análises, pode estar relacionada com a acentuada diferenciação populacional observada, uma vez que os iniciadores para os locos de marcadores microssatélites foram desenhados a partir da sequência de um indivíduo da população BRO (PEREZ et al., 2011). A ocorrência dessa população está localizada no extremo sul da distribuição da espécie e apresenta elevada diferenciação em relação às populações do grupo Norte. Com isso, a maior dificuldade de amplificação em indivíduos de

populações geneticamente distantes da população BRO deve ser resultado da ocorrência de mutações nas regiões de ligação dos iniciadores utilizados.

Os índices de diversidade genética foram similares em todas as populações, exceto para as localidades amostradas nos extremos norte e sul da ocorrência da espécie (BRO e BAR), onde foram aparentemente menores. No entanto, os testes estatísticos realizados para a população BAR não demonstraram diferenças significativas entre os níveis de diversidade genética entre essa e as outras populações, uma vez que um número muito pequeno de testes foram significativos. No caso de BRO, as estimativas de número de alelos por loco apresentaram diferenças significativas em relação a diversas populações, indicando que essa população pode ter um nível menor de diversidade. Índices menores de diversidade genética nos extremos da distribuição são esperados, em decorrência do fluxo gênico restrito ou do maior tempo de isolamento em relação às populações no centro da distribuição (LESICA e ALLENDORF, 1995). Outro fator que poderia ter influenciado essas estimativas seria o tamanho populacional reduzido nessas localidades, o que acentuaria os efeitos da deriva genética. Essa hipótese é ainda mais provável em BRO, uma vez que essa localidade é um morro de arenito isolado, com pequenas dimensões (< 5 ha). Uma terceira hipótese seria a colonização mais recente dessas áreas, com ocorrência de efeito fundador, que levaria à diminuição da diversidade nessas localidades (HEWITT, 2000).

Moraes e Sene (2007) também observaram índices mais baixos de diversidade genética utilizando locos de DNA microssatélite em duas populações de *D. gouveai* com ocorrência próxima à região de BRO. Os autores associaram os índices verificados ao tamanho reduzido dessas áreas ou a um isolamento mais antigo dessas populações. Em outro estudo utilizando as mesmas populações e sequências do DNA mitocondrial, Moraes et al. (2009) confirmaram que essas populações apresentam um isolamento mais antigo, com alta divergência em relação às outras populações da espécie. É importante destacar que a espécie

D. gouveai, utilizada como modelo para esses dois trabalhos, apresenta distribuição muito semelhante à *P. machrisii*, pois seu ciclo de vida está relacionado com a presença de cactos colunares, nos quais as fêmeas realizam a ovoposição em tecidos necrosados.

Desvios em relação ao esperado pelo EHW foram detectados para diversas populações de *P. machrisii*. Os resultados obtidos podem apontar para um efeito de endogamia nas populações. No entanto, outros fatores podem estar associados a esses valores, uma vez que cada uma das populações apresentou no máximo dois locos com desvios em relação às proporções do EHW e os efeitos endogâmicos deveriam afetar todos os locos. A deficiência verificada de heterozigotos pode estar relacionada à presença de alelos nulos ou a seleção natural agindo sobre regiões próximas a esses locos. No caso do loco *Pmac146*, que apresentou deficiência significativa de heterozigotos em oito populações diferentes, Perez et al. (2011) sugeriram a presença de alelos nulos para explicar os desvios observados para esse marcador em duas populações de *P. machrisii* (DEL e FUR). Outra possível explicação para esses desvios pode ser o baixo tamanho amostral utilizado em relação ao elevado polimorfismo verificado no loco *Pmac146*.

Os testes de Mantel e de autocorrelação espacial resultaram em consideráveis correlações para as distâncias genéticas e geográficas entre as populações, principalmente para as classes de distância menores. No entanto, ao analisar os agrupamentos principais separadamente (Sudeste, Central, URU e Nordeste), não foram verificadas correlações significativas, indicando que o resultado anterior foi viesado pelos elevados valores de diferenciação genética e distância geográfica entre as populações desses grupos. A ausência de correlação significativa dentro de cada grupo sugere que a similaridade genética entre as populações não está sendo determinada predominantemente por fluxo gênico atual, mas por fatores históricos. Esse resultado foi corroborado pelas baixas estimativas de migração detectadas.

Embora a análise de estrutura populacional por meio da estatística ΔK com o programa STRUCTURE tenha resultado na separação das populações em dois grupos principais, os resultados da AMOVA hierárquica mostraram que a maior parte da variação genética ocorreu dentro desses grupos. Esse resultado mostra que o agrupamento das populações em dois grupos não reflete a estrutura populacional predominante nos dados obtidos. A partir dos resultados para um número maior de grupos, foi possível inferir que o agrupamento que reflete melhor a estruturação da variação no conjunto de populações analisadas foi aquele em que as populações foram reunidas em quatro grupos (Sudeste, Central, URU e Nordeste). Os resultados obtidos para a PCA e a árvore UPGMA foram congruentes com a estruturação sugerida pelo STRUCTURE, representando as relações de similaridade genética entre as populações, com a formação de grupos regionais. Os dois métodos também revelaram que as populações do subgrupo Nordeste (CRI, POS e BAR) exibem um nível de diferenciação comparável ao observado entre diferentes grupos.

A possível existência de táxons distintos entre as populações amostradas pode ter contribuído para os resultados de elevada estruturação e fluxo gênico restrito, uma vez que a análise conjunta desses possíveis táxons incluiria os efeitos do isolamento reprodutivo entre espécies diferentes. Braun e Esteves Pereira (1987a, 1987b, 2002) descreveram três táxons com ocorrência no Planalto Central (*P. goianus*, *P. cristalinensis* e *P. lindanus*), onde parte das amostras deste estudo foi obtida, e que atualmente são considerados sinônimos de *P. machrisii*. *Pilosocereus goianus* apresenta distribuição associada exclusivamente a rochas calcárias (BRAUN e ESTEVES PEREIRA, 2002) e pode ter sido representada em ART, a única população com ocorrência sobre esse tipo de rocha. As outras populações amostradas apresentaram ocorrência sobre afloramentos de arenito ou quartzito. O morfotipo descrito originalmente como *P. cristalinensis* foi amostrado na localidade CRI, a qual apresentou indivíduos com grande diferenciação genética, inclusive em relação às outras populações do

subgrupo Nordeste (BAR e POS). A alta diferenciação dessa população também foi verificada por Moraes et al. (2005) em um estudo com isoenzimas, o qual revelou que a população CRI apresenta grande divergência e relação a outras populações de *P. machrisii* e está mais relacionada com uma população de *P. vilaboensis* distante 177 km de sua localização. As populações FMS, APA1, APA2 e ART estão na região descrita como área de ocorrência do morfotipo *P. lindanus*. Essas populações foram reunidas no grupo Central pela análise do programa STRUCTURE, assim como pela PCA e o dendrograma UPGMA.

Esses resultados podem sugerir a presença de mais de uma linhagem evolutiva distinta dentro do conjunto de populações atualmente classificadas como *P. machrisii*. Essas incertezas taxonômicas reforçam a necessidade de realização de uma abordagem filogenética para os táxons do Grupo *P. AURISSETUS*, a fim de que as relações entre essas espécies e suas diferentes populações possam ser esclarecidas de forma mais apropriada.

Em conjunto, os resultados apontam que a distribuição atual de *P. machrisii* foi provavelmente formada por eventos de vicariância a partir de uma conexão maior entre as populações no passado. Essa hipótese é corroborada pelos índices muito baixos de fluxo gênico atual entre as populações, verificados a partir da ocorrência de alelos exclusivos em frequências altas e a detecção de apenas dois migrantes em um total de 241 indivíduos. A alta diferenciação observada entre as populações e a ausência de correlação entre as distâncias genéticas e geográficas dentro dos dois agrupamentos principais também são indicativos de prováveis relações históricas entre as populações, com o fluxo gênico atual apresentando pequena importância na estruturação da diversidade genética. Os resultados obtidos concordam com estudos em outras espécies que ocorrem associadas a afloramentos rochosos, os quais também sugerem uma distribuição mais ampla no passado, com posterior retração e isolamento das populações, como *D. gouveai* (MORAES e SENE, 2007; MORAES et al., 2009) e *Lychnophora ericoides* (COLLEVATTI et al., 2009). Diversas espécies que

apresentam ocorrência disjunta em ambientes com condição xérica e associadas a uma unidade vegetacional denominada floresta tropical sazonalmente seca (SDTF) também são relacionadas com distribuições mais contínuas em períodos anteriores (PRADO e GIBBS, 1993; PENNINGTON et al., 2000, 2004; CAETANO et al., 2008). O padrão de ocorrência disjunta dessas espécies tem sido relacionado a mudanças climáticas que ocorreram durante o Quaternário e teriam afetado os padrões de distribuição dessas espécies (CAETANO et al., 2008; WERNECK, 2011).

De acordo com essa hipótese de distribuição mais ampla no passado, as populações localizadas no Norte e no Sul da distribuição da espécie, que são separadas por uma grande extensão geográfica e possuem grande diferenciação genética teriam sido as primeiras a se dividirem. Posteriormente, as populações do Norte teriam se separado, formando a população do município de Uruaçu e os grupos Nordeste e Central. Por fim, as populações do grupo Nordeste se dividiram e as populações do grupo Sudeste teriam sido as últimas a se dividir, de forma que FUR e DEL apresentam uma grande semelhança genética, uma vez que seus indivíduos foram classificados no mesmo grupo genético pela análise de segundo nível do programa STRUCTURE.

A partir dos padrões observados no presente estudo, estratégias de conservação para *P. machrisii*, assim como para outras espécies que ocupam as mesmas áreas podem ser traçadas mais apropriadamente. A alta diferenciação entre populações de diferentes localidades deve ser levada em consideração, assim como as taxas muito baixas de migração. O fato da ocorrência atual dessas populações representar um possível relictos de uma distribuição mais ampla no passado também deve ser considerado. De acordo com essas características, a estratégia mais importante para a conservação da variação genética nessa espécie seria a conservação do maior número possível de populações de diferentes regiões geográficas. A translocação de indivíduos para efeitos de reintrodução e aumento da variabilidade genética

deveria levar em conta essa elevada diferenciação, bem como as situações de incertezas taxonômicas.

5. CONCLUSÃO

A espécie *P. machrisii* apresenta níveis de diversidade genética similares entre suas populações, exceto para aquelas localizadas nos limites da distribuição, que possuem menor diversidade, provavelmente associada com menores tamanhos populacionais ou maior tempo de isolamento dessas populações.

Foram observados altos índices de diferenciação populacional, com baixos níveis de fluxo gênico, em concordância com a distribuição fragmentada da espécie. Esse padrão disjunto provavelmente representa um relictos de uma distribuição mais ampla, assim como sugerida para diversas espécies de áreas secas que ocorrem na mesma região.

As incertezas e a instabilidade taxonômica de *P. machrisii* e do grupo *P. AURISSETUS* torna necessária a realização de uma filogenia para esclarecer as relações entre essas espécies, bem como de outras espécies do gênero *Pilosocereus* na região leste do Brasil.

Além dos resultados obtidos a partir dos marcadores de DNA microsatélite, sequências do DNA do cloroplasto estão em fase final de análise em outro projeto em desenvolvimento pelo nosso grupo de pesquisa. A comparação desses dois conjuntos de dados pode auxiliar no esclarecimento das relações das populações atualmente classificadas como *P. machrisii*.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, R.I., BROWN, K.M.; HAMILTON, M.B. The impact of microsatellite electromorph size homoplasy on multilocus population structure estimates in a tropical tree (*Corythophora alta*) and an anadromous fish (*Morone saxatilis*). **Mol. Ecol. Notes**, v.19, p. 2579-2588, 2004.
- APPLEQUIST, W.L.; WALLACE, R.S. Deletions in the plastid *trnT-trnL* intergenic spacer define clades within Cactaceae subfamily Cactoideae. **Plant Syst. Evol.**, v. 231, p. 153–162, 2002.
- ARAKAKI, M., CHRISTIN, P.A., LENDEL, A., NYFFELER, R., EGGLI, U., OGBURN, R.M., SPRIGGS, E., MOORE, M.; EDWARDS, E.J. Recent and contemporaneous radiations of the worlds succulent plant lineages. **Proc. Nat. Acad. Scien. USA**, v. 108, p. 8379-8384, 2011.
- ARRUDA, M.P., GONÇALVES, E.C., SCHNEIDER, M.P.C.; SILVA, A.L.C.; MORIELLE-VERSUTE, E. An alternative genotyping method using dye-labeled universal primer to reduce unspecific amplifications. **Molec. Biol. Rep.**, v. 37, p. 2031-2036, 2010.
- BRAUN, P.J.; ESTEVES PEREIRA, E. *Pilosocereus lindaianus* Braun; Esteves. spec. nov. **Kakt. And. Sukk.**, v. 38, n.5, p. 132, 1987a.
- BRAUN, P.J.; ESTEVES PEREIRA, E. *Pilosocereus cristalinensis* Braun; Esteves. spec. nov. **Kakt. And. Sukk.**, v. 38, n.5, p. 132, 1987a.
- BRAUN, P.J.; ESTEVES PEREIRA, E. *Pilosocereus goianus* Braun ; Esteves. **Brit. Cact. Succ. J.**, v. 20, n. 2, 93-103, p. 2002.
- CAETANO, S., PRADO D., PENNINGTON, R.T., BECK, S., OLIVEIRA-FILHO, A., SPICHIGER, R.; NACIRI, Y. The history of Seasonally Dry Tropical Forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). **Molec. Ecol.**, v. 13, p. 3147-3159, 2008.
- CALVENTE, A., ZAPPI, D.C., FOREST, F.; LOHMANN L.G. Molecular phylogeny of tribe Rhipsalideae (Cactaceae) and taxonomic implications for Schlumbergera and Hatiora. **Mol. Phyl. Evol.**, v. 58, p. 456-468, 2011.

- CAVALLI-SFORZA, L.L.; EDWARDS, A.W.F. Phylogenetic analysis: models and estimation procedures. **Amer. J. Hum. Genet.**, v. 19, p. 233-257, 1967.
- CHAPUIS, M.P.; ESTOUP, A. Microsatellite null alleles and estimation of population differentiation. **Mol. Biol. Evol.**, v. 24, p. 621-631, 2007.
- COLLEVATTI, R.G.; RABELO, S.G.; VIEIRA, R.F. Phylogeography and disjunct distribution in *Lychnophora ericoides* (Asteraceae), an endangered cerrado shrub species. **Ann. Bot.**, v. 104, p. 655-664, 2009.
- CORNUET, J.M.; PIRY, S.; LUIKART, G.; ESTOUP A.; SOLIGNAC, M. New methods employing multilocus genotypes to select or exclude population as origins of individuals. **Genetics**, v. 153, p. 1989–2000, 1999.
- DAVIS, M.B ; SHAW, R.G. Range shifts and adaptive responses to Quaternary climate change. **Science**, v. 292, p. 673–679, 2001.
- DUPANLOUP, I., SCHNEIDER, S.; EXCOFFIER, L. A simulated annealing approach to define the genetic structure of populations. **Mol. Ecol.**, v. 11, n. 12, p. 2571-81, 2002.
- EXCOFFIER, L.; HECKEL, G. Computer programs for population genetics data analysis: a survival guide. **Nat. Rev. Genet.**, v. 7, p. 745–758, 2006.
- EL MOUSADIK, A.; PETIT, R.J. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. **Theor. App. Genet.**, v. 92, p. 832–839, 1996.
- ELLEGREN H. Microsatellites: simple sequences with complex evolution. **Nat. Rev. Genet.**, v. 5, p. 435–445, 2004.
- ESTOUP, A., JARNE, P. ; CORNUET, J.M. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. **Mol. Ecol.**, v 11, n.9, p. 1591-1604, 2002.
- EVANNO, G., REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Mol. Ecol.**, v. 14, p. 2611–2620, 2005.
- EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. **Evol. Bioinf. Onl.**, v. 1, p. 47-50, 2005.
- FRANKHAM R. Genetics and extinction. **Biol. Conserv.**, v. 126, p. 131–140, 2005.

- GIULIETTI, A.M.; PIRANI, J. R .; HARLEY, R.M. Espinhaço Range Region, Eastern Brazil. In: DAVIS, S.D.; HEYWOOD, V.H.; HERRERA-MACBRYDE, O.; VILLALOBOS, J.; HAMILTON A.C. (eds.). **Centres of plant diversity: A guide and strategy for their conservation**. Cambridge, Massachusetts: The Americas. IUCN Publication Unity, 1997. v. 3, p. 397-404.
- GOMES, V.S.M.; CORREIA, M.C.R.; LIMA, H.A.; ALVES, M.A.S. Potential role of frugivorous birds (Passeriformes) on seed dispersal of six plant species of a restinga habitat, southeastern Brazil. **Rev. Biol. Trop.**, v. 56, p. 205-216, 2008.
- GORELICK, R. Evolution of cacti is largely driven by genetic drift, not selection. **Bradleya**, v. 27, p. 37-48. 2009.
- GOUDET, J. FSTAT (version 1.2): a computer program to calculate F-statistics. **J. Hered.**, v. 86, p. 485-486, 1995.
- HAMILTON, J.A.; ECKERT, C.G. Population genetic consequences of geographic disjunction: a prairie plant isolated on Great Lakes alvars. **Mol. Ecol.**, v. 16, p. 1649-1660, 2007.
- HANCOCK, J.M. The contribution of slippage-like processes to genome evolution. **J. Mol. Evol.**, v. 41, p. 1038–1047, 1995.
- HEWITT, G. The genetic legacy of quaternary ice ages. **Nature**, v. 405, p. 907–913, 2000.
- HOLSINGER, K.E.; WEIR, B. S. Genetics in geographically structured populations: defining, estimating and interpreting FST. **Nat. Rev. Genet.**, v. 10, p. 639-50, 2009.
- HUBISZ, M.J.; FALUSH, D.; STEPHENS, M.; PRITCHARD, J.K. Inferring weak population structure with the assistance of sample group information. **Mol. Ecol. Res.**, v. 9, p. 1322–1332, 2009.
- HUNT, D.; TAYLOR, N. P.; CHARLES, G. **The New Cactus Lexicon: Text**. Milborne Port: dh Books, 2006.
- 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.

- JESUS, A.A.; MORAES, E.M.; ORTOLANI, F.A. Análise citogenética do grupo de espécies relacionadas a *Pilosocereus aurisetus* (Cactaceae). **XII Congresso de Iniciação Científica da Universidade Federal de São Carlos**. São Carlos: UFSCar, 2009.
- KRUGLYAK, S.; DURRETT, R.T.; SCHUG, M.D.; AQUADRO, C.F. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. **Proc. Natl Acad. Sci. USA**, v. 95, p. 10774–10778, 1998.
- LANGELLA, O. **Populations version 1.2.30**. Disponível em: <<http://bioinformatics.org/~tryphon/populations/>>, 1999.
- LAVAL, G.; SANCRISTOBAL, M.; CHEVALET, C. Measuring genetic distances between breeds: use of some distances in various short term evolution models. **Genet. Sel. Evol.**, v. 34, p.481–507, 2002.
- LESICA, P.; ALLENDORF, F.W. When are peripheral populations valuable for conservation? **Cons. Biol.**, v. 9, p. 753–760, 1995.
- MANEL, S.; BERTHIER, P.; LUIKART, G. Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. **Conserv. Biol.**, v. 16, p. 650–659, 2002.
- MANTEL, N.A. The detection of disease clustering and a generalized regression approach. **Canc. Res.**, v. 27, p. 209 – 220, 1967.
- MORAES, E.M.; ABREU, A.G.; ANDRADE, S.C.S.; SENE, F.M.; SOLFERINI, V.N. Population genetic structure of two columnar cacti with a patchy distribution in eastern Brazil. **Genetica.**, v. 125, p. 311-323, 2005.
- MORAES, E.M.; SENE, F.M. Microsatellite and morphometric variation in *Drosophila gouveai*: the relative importance of historical and current factors in shaping the genetic population structure. **J. Zoolog. Syst. Evol. Res.**, v. 45, p. 336-344, 2007.
- MORAES, E.M.; YOTOKO, K.S.C.; MANFRIN, M.H.; SOLFERINI, V.N.; SENE, F.M. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. **J. Biog.**, v. 36, p. 2136–2147, 2009.
- MUNGUÍA-ROSAS, M.A.; JÁCOME-FLORES, M.E.; SOSA V.; QUIROZ-CERÓN, M.L. Removal of *Pilosocereus leucocephalus* (Cactaceae, tribe Cereeae) seeds by ants and their potential role as primary seed dispersers. **J. Arid Envir.**, v. 73, p. 578-581, 2009.

- MUNGUÍA-ROSAS, M.A.; SOSA, V.J.; JÁCOME-FLORES, M.E. Pollination system of the *Pilosocereus leucocephalus* columnar cactus (tribe Cereeae) in eastern Mexico. **Plant Biol (Stuttg)**, v. 12, n. 4, p. 578-586, 2010.
- NEI, M. **Molecular Evolutionary Genetics**. New York: Columbia University Press, 1987.
- NIELSEN, R.; WAKELEY, J. Distinguishing migration from isolation: A markov chain monte carlo approach. **Genetics**, v. 158, p. 885–896, 2001.
- NYFFELER, R. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from *trnK/matK* and *trnL-trnF* sequences. **Amer. J. Bot.**, v. 89, p. 312–326, 2002.
- OLIVEIRA, E.J.; PÁDUA, J.G.; ZUCCHI, M.I.; VENCOVSKY, R.; VIEIRA, M.L.C. Origin, evolution and genome distribution of microsatellites. **Genet. Molec. Biol.**, v. 29, n. 2, p. 294-307, 2006.
- OHTA, T.; KIMURA, M. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. **Genet. Res.**, v. 22, p. 201–204. 1973.
- PAETKAU, D.; SLADE, R.; BURDEN, M.; ESTOUP, A. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. **Mol. Ecol.**, v. 13, p. 55–65, 2004.
- PEAKALL, R.; SMOUSE, P.E. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. **Molec. Ecol. Not.**, v. 6, p. 288–295, 2006.
- PEARSE, D.E.; CRANDALL, K.A. Beyond F_{ST} : Analysis of population genetic data for conservation. **Conserv. Genet.**, v. 5, p. 585–602, 2004.
- PENNINGTON, R.T., PRADO, D.E.; PENDRY, C.A. Neotropical seasonally dry forests and Quaternary vegetation changes. **J. Biog.**, v. 27, p. 261–273, 2000.
- PENNINGTON, R.T.; LAVIN, M.; PRADO, D.E.; PENDRY, C.A.; PELL, S.K.; BUTTERWORTH, C.A. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. **Phil. Trans. R. Soc. Lond.**, v. 359, p. 515-538, 2004.

- PEREZ, M.F.; TÉO, M.F.; ZAPPI, D.C.; TAYLOR, N.P.; MORAES, E.M. Isolation, characterization, and cross-species amplification of polymorphic microsatellite markers for *Pilosocereus machrisii* (Cactaceae). **Amer. J. Bot.**, v. 98, n.8, p. e204–e206, 2011.
- PIRY, S.; ALAPETITE, A.; CORNUET, J.M.; PAETKAU, D.; BAUDOUIN, L.; ESTOUP, A. GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. **J. Hered.**, v. 95, p. 536-539, 2004.
- POLLOCK, D.D.; BERGMAN, A.; FELDMAN, M.W.; GOLDSTEIN, D.B. Microsatellite behavior with range constraints: parameter estimation and improved distances for use in phylogenetic reconstruction. **Theor. Popul. Biol.**, v. 53, p. 256–271, 1998.
- PRADO, D.E.; GIBBS, P.E. Patterns of species distributions in the dry seasonal forests of South America. **Ann. Miss. Bot. Gard.**, v. 80, p. 902-927, 1993.
- PRITCHARD, J.K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, p. 945–959, 2000.
- R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<http://www.R-project.org/>>, 2012.
- RANNALA, B.; MOUNTAIN, J.L. Detecting immigration by using multilocus genotypes. **Proc. Natl. Acad. Sci. USA**, v. 94, p. 9197-9221, 1997.
- RICE, W. Analyzing tables of statistical tests. **Evolution**, v. 43, p. 223–225, 1989.
- ROHLF, F.J. NTSYSpc numerical taxonomy and multivariate analysis system v.2.1. Setauket: Exeter Software, 2000.
- ROSENBERG, N.A. DISTRUCT: a program for the graphical display of population structure. **Mol. Ecol. Notes**, v. 4, p. 137–138, 2004.
- SCHLÖTTERER, C. Evolutionary dynamics of microsatellite DNA. **Chromosoma**, v. 109, p. 365–371, 2000.
- SCHUELKE, M. An economic method for the fluorescent labeling of PCR fragments. **Nat. Biotech.**, v 18, p. 233-234, 2000.
- SELKOE, K.A.; TOONEN, R.J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. **Ecol. Lett.**, v. 9, p. 615–629, 2006.

- SLATKIN, M. Gene flow and the geographic structure of natural populations. **Genetics**, v. 139, p. 787–792, 1987.
- SLATKIN, M. A measure of population subdivision based on microsatellite allele frequencies. **Genetics**, v. 139, p. 457-462, 1995.
- SMOUSE, P.E.; PEAKALL, R. Spatial autocorrelation analysis of multi-allele and multi-locus genetic microstructure. **Heredity**, v. 82, p. 561-573, 1999.
- SNEATH, P.H.A.; SOKAL, R.R. **Numerical taxonomy**. San Francisco: W. H. Freeman and Company, 1973.
- SUNNUCKS, P. Efficient genetic markers for population biology. **Trends Ecol. Evol.**, v. 15, p. 199–203, 2000.
- TAYLOR, N.P.; ZAPPI, D.C. **The Cacti of Eastern Brazil**. Kew: Royal Botanic Gardens, 2004.
- TÓTH, G.; GÁSPÁRI, Z.; JURKA, J. Microsatellites in different eukaryotic genomes: survey and analysis. **Gen. Res.**, v. 10, p. 967–981, 2000.
- VÄHÄ, J.P.; ERKINARO, J.; NIEMELA, E.; PRIMMER, C.R. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. **Mol. Ecol.**, v. 16, p. 2638–2654, 2007.
- VALIÈRE, N. GIMLET: a computer program for analysing genetic individual identification data. **Molec. Ecol. Notes.**, v. 2, p. 377–379, 2002.
- WALLACE, R.S.; COTA, J.H. An intron loss in the chloroplast gene *rpoC1* support: a monophyletic origin for the subfamily Cactoideae of the Cactaceae. **Curr. Genet.**, v. 29, p.275-281, 1996.
- WEIR, B.S.; COCKERHAM, C.C. Estimating *F*-statistics for the analysis of population structure. **Evolution**, v. 38, p. 1358–1370, 1984.
- WERNECK, F.P. The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. **Quat. Scien. Rev.**, v. 30, p. 1630-1648, 2011.
- WILCOXON, F. Individual comparisons by ranking methods. **Biomet. Bull.**, v. 1, n. 6, p. 80–83, 1945.
- WRIGHT, S. The genetical structure of populations. **Ann. Eugen.**, v. 15, p. 323–353, 1951.

ZAPPI, D.C. **Pilosocereus (Cactaceae):** The genus in Brazil. Kew: Royal Botanic Gardens, 1994.

ZAPPI, D.C.; TAYLOR, N.P.; SILVA, S.R.; MACHADO, M.; MORAES, E.M.; CALVENTE, A.; CRUZ, B.; CORREIA, D.; LAROCCA, J.; ASSIS, J.G.; MEIADO, M.; SANTOS, M.R.; BELLINTANI, M.; COELHO, P.; NAHOUM, P.I.; RESENDE, S. **Plano de ação nacional para a conservação das Cactáceas:** Série Espécies Ameaçadas. 1. ed. Brasília: Instituto Chico Mendes de Conservação da Biodiversidade, 2011.

ANEXO

**ISOLATION, CHARACTERIZATION, AND CROSS-SPECIES
AMPLIFICATION OF POLYMORPHIC MICROSATELLITE MARKERS
FOR *PILOSOCEREUS MACHRSII* (CACTACEAE)¹**

MANOLO F. PEREZ², MARIANA F. TÉO², DANIELA C. ZAPPI³, NIGEL P. TAYLOR³, AND
EVANDRO M. MORAES^{2,4}

²Campus Sorocaba, Universidade Federal de São Carlos, Rod. João Leme dos Santos km 110 18052780, Sorocaba, Brazil;

³Royal Botanic Gardens, Richmond, Surrey, TW9 3AE, Kew, United Kingdom

- *Premise of the study:* Microsatellite primers were developed in *Pilosocereus machrsii*, a columnar cactus with a patchy distribution in eastern tropical South America, to assess its genetic diversity and population structure.
- *Methods and Results:* Eleven microsatellite markers were developed, of which one was monomorphic among 51 individuals from two populations. The number of alleles per polymorphic locus ranged from two to eight, and the total number of alleles was 57. From the 11 isolated loci, nine were successfully amplified in the other four *Pilosocereus* species.
- *Conclusions:* The results showed that these markers will be useful for studies of genetic diversity, patterns of gene flow, and population genetic structure in *P. machrsii*, as well as across other congeneric species.

Key words: *campos rupestres*; cross-amplification; microsatellites; *Pilosocereus machrsii*.

Eastern tropical South America is an important diversity center for Cactaceae, harboring about 10% of the 1300 recognized species (Taylor and Zappi, 2004), most of which are endemic and with at least 65 taxa of conservation concern (Taylor and Zappi, 2004; IUCN, 2010). In the semiarid *caatinga* biome in northeastern Brazil (Andrade-Lima, 1981), the cacti populations have broad and relatively contiguous distribution. Outside this biome, cactus populations have a markedly patchy distribution, generally occurring in highlands, in enclaves of *campo rupestre* vegetation (Giulietti et al., 1997) and other rock formations, under locally xeric conditions and often with a high number of endemic plants. In spite of their high diversity and endemism in eastern Brazil, there are few reports assessing the genetic diversity, evolutionary history, and conservation issues of Cactaceae in this region (Moraes et al., 2005).

The genus *Pilosocereus*, among the three most speciose genera of cacti in Brazil (Zappi, 1994) spans throughout the *caatinga* biome, occurring also further south and west, within the *campo rupestre*. Among its species, *Pilosocereus machrsii* (E. Y. Dawson) Backeb. is a columnar cactus with a patchy distribution, restricted to summits of isolated hills or rocky outcrops in the southern part of its range, while its northern populations

are mainly associated with the dry *campo rupestre* vegetation. These features make *P. machrsii* an appropriate biological model for population genetic studies to gather information on historical range shifts and to develop comprehensive conservation plans for the restricted ecosystems where it occurs. Here we describe the isolation and evaluation of 11 microsatellite loci for *P. machrsii* and test the transferability of those markers to other related species belonging to *Pilosocereus*. These markers will be used to assess patterns of historic and recent gene flow, genetic diversity, and population structure across the species distribution, contributing to the knowledge of the species and the conservation of its peculiar, fragmented habitat.

METHODS AND RESULTS

For the isolation of microsatellite markers, total genomic DNA was isolated from root tissue of a single individual using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Extracted DNA was digested with *Xmn*I, and the microsatellite loci were isolated through the enrichment protocol described by Glenn and Schable (2005), using the oligonucleotides probes (GA)₁₀, (CA)₁₀, and (TC)₁₀. Amplified fragments were cloned into plasmids using the TOPO TA cloning kit (Invitrogen, Carlsbad, California, USA). A total of 273 bacterial colonies containing inserts were isolated, from which 94 were sequenced with the universal M13 primer using the BigDye Terminator version 3.1 and visualized with an ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster City, California, USA). Nineteen inserts showed clear microsatellite motifs and contained sufficient flanking regions for primer design. BLAST similarity analysis in GenBank showed that none of the resultant sequences were reported within the database. All primer pairs were assayed in individual PCRs, and products were run on 1.5% agarose gels. The forward primer of each pair exhibiting successful amplification was tailed with an M13 sequence (5'-TGTA AAA-CGACG CCA GT-3') to its 5' to enable use of a third M13 primer fluorescently labeled for automatic genotyping, as outlined by Schuelke (2000). Eleven primer pairs exhibited high-quality PCR products of the expected size, 10 of them showed dinucleotide and one a trinucleotide repeat motif.

The 11 loci were characterized on 51 *P. machrsii* individuals from two natural populations, at Delfinópolis (DEL) and Furnas (FUR), at the southeastern

¹Manuscript received 20 January 2011; revision accepted 23 March 2011.

The authors thank Heidi Utsunomiya, Mercival Francisco, Karina Martins, Fabricia Madia, Carolina Nascimento, and Isabel Bonatelli for laboratory assistance and Marlon Machado for assistance with the field sampling. This work was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo—FAPESP (grants no. 05/55200-8 to EMM, 08/08421-7 to MFP and 10/05690-7 to MFT) and Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq—Brasil (grants no. 471963/2007-0 and 500903/2009-3 to EMM).

⁴Author for correspondence: emarsola@ufscar.br

TABLE 1. Characterization of 10 polymorphic and one monomorphic microsatellite loci for *Pilosocereus machrisii*. GenBank accession numbers (below loci names), repeat motifs, forward (F) and reverse (R) primer sequences, optimized annealing temperatures (T_a), and size ranges of PCR products.

Locus (GenBank accession)	Repeat motif	Primer Sequence (5'-3')	T_a (°C)	Size Range (bp)
<i>Pmac082</i> (HQ667127)	(GAG) ₅ (GAA) ₂ (GAG) ₂ GAA(GAG) ₂	F: GTAAAAGAGGAGGATGGAGAGG R: CTTCTTCTTGCTAGGTTTCTCG	56/54	108
<i>Pmac084</i> (HQ667128)	(AG) ₉ CG(AG) ₂	F: CATAAATTGCAGAAATGAGGAC R: AGGTAAACCGCTAACTCGATG	60/54	88–106
<i>Pmac085</i> (HQ667129)	(AG) ₆ AC(AG) ₃ AC(AG) ₄	F: CCCTCACTTCTCCAATCTC R: TCCCTCCTTCTTTGTGTTC	54	90–98
<i>Pmac101</i> (HQ667130)	(TC) ₁₅ TATG (TA) ₃	F: TGATAGCTGCAACGATGTC R: CATTTGTTTGTGTGTTTCTTCC	52	108–114
<i>Pmac102</i> (HQ667131)	(AG) ₉	F: TCTATAAGTGCCGATGGATGC R: CACACCTCACTCCCAACCTC	60/54	188–120
<i>Pmac108</i> (HQ667132)	(AG) ₁₄ (TG) ₇	F: TGAATGTCTGCCATTAGAAAGC R: TAATCGCCCCAAGAAGAAC	52	125–137
<i>Pmac128</i> (HQ667133)	(TC) ₅ TT(TC) ₁₀ (AC) ₁₀	F: GTGTTGATGTACTCTTTCAG R: CTAACCCCTTGTATACATGC	60/54	138–154
<i>Pmac130</i> (HQ667134)	(AG) ₇ CA(AG) ₁₂	F: GAGGTGCCAATAAATCG R: TGTCCACGCAATCTTGAACC	58/54	141–149
<i>Pmac135</i> (HQ667135)	(TC) ₅ TG(TC) ₁₂	F: ACCAGAATGAGCTCAGCTGTAG R: CCTAGCTAGCAGAATCAGAGTGAAGAC	52	126–140
<i>Pmac146</i> (HQ667136)	(AG) ₂₀	F: ACCCGACATCCCCTTGTAG R: TAGTCTGAAACGGAGCAAGG	64/54	143–157
<i>Pmac149</i> (HQ667137)	(TC) ₁₉	F: TTCATCCTGCTTTTGAAGTTTG R: TGATGGATTAGGATTGACCTG	62/54	162–174

TABLE 2. Results of initial primer screening in two *Pilosocereus machrisii* populations (sample sizes in parentheses). Number of alleles (A), observed (H_o), and expected (H_e) heterozygosity, and Weir and Cockerham's (1984) fixation index (F) followed by statistical significance for the Hardy–Weinberg equilibrium test (HWE) are given for each marker and population.

Locus	A	H_o	H_e	F
Population DEL ^a (N = 30)				
<i>Pmac082</i>	1	0	0	—
<i>Pmac084</i>	5	0.400	0.648	0.387*
<i>Pmac085</i>	2	0.133	0.127	-0.054
<i>Pmac101</i>	4	0.367	0.664	0.452*
<i>Pmac102</i>	2	0.100	0.097	-0.036
<i>Pmac108</i>	6	0.867	0.826	-0.050
<i>Pmac128</i>	4	0.633	0.657	0.037
<i>Pmac130</i>	5	0.500	0.645	0.228
<i>Pmac135</i>	7	0.733	0.656	-0.120
<i>Pmac146</i>	8	0.367	0.705	0.484*
<i>Pmac149</i>	8	0.767	0.718	-0.069
Population FUR ^a (N = 21)				
<i>Pmac082</i>	1	0	0	—
<i>Pmac084</i>	5	0.381	0.592	0.362
<i>Pmac085</i>	1	0	0	—
<i>Pmac101</i>	3	0.333	0.484	0.317
<i>Pmac102</i>	2	0.238	0.215	-0.111
<i>Pmac108</i>	8	0.714	0.747	0.045
<i>Pmac128</i>	5	0.524	0.597	0.125
<i>Pmac130</i>	2	0.190	0.176	-0.081
<i>Pmac135</i>	5	0.333	0.540	0.389
<i>Pmac146</i>	8	0.381	0.789	0.523*
<i>Pmac149</i>	5	0.905	0.782	-0.162

* Significant departures from HWE after Bonferroni correction ($\alpha = 0.05$).

^a Vouchers of *P. machrisii* populations were deposited at the non-indexed (*Index Herbariorum*) Universidade Federal de São Carlos, Campus Sorocaba Herbarium (HUFS): DEL population (HUFS636), FUR population (HUFS637).

part of the species range (geographical coordinates: DEL 20°21'44.5"S, 46°44'59.4"W and FUR 20°38'53.4"S, 46°19'16.4"W). Amplification was carried out in a total volume of 10 μ L using 1 \times *Taq* Buffer (Fermentas, Burlington, Ontario, Canada) with (NH₄)₂SO₄ (750 mM Tris-HCl, pH 8.8, 200 mM (NH₄)₂SO₄, and 0.1% Tween 20), 1.5 mM MgCl₂, 200 μ M dNTPs, 0.25 μ M of M13 tailed forward primer, 0.25 μ M reverse primer, and 0.25 μ M M13 fluorescently labeled primer (either 6-FAM or NED, Applied Biosystems), 0.5 U of *Taq* DNA Polymerase (Fermentas) and 5–50 ng of template DNA. PCR was performed on a Mastercycler Gradient thermocycler (Eppendorf AG, Hamburg, Germany) with the following conditions: 94°C for 2 min; followed by 30 cycles of 94°C for 40 s, specific annealing temperature for 40 s (see Table 1 for details), 72°C for 40 s; and a final cycle of 72°C for 10 min. For the loci with two annealing temperatures (*Pmac082*, *Pmac102*, *Pmac128*, *Pmac130*, *Pmac146*, and *Pmac149*; Table 1), the same conditions were employed, except for a touch-down procedure, on which 25 cycles were performed on the first annealing temperature and then the temperature was decreased 2°C per cycle until reaching the second temperature, followed by eight final cycles using this last annealing temperature. PCR products were subjected to fragment analysis on an ABI PRISM 3730 DNA Analyzer (Applied Biosystems) using 500 LIZ (Applied Biosystems) as the size standard. Alleles were sized using GeneMarker v. 1.91 (SoftGenetics, State College, Pennsylvania, USA). The observed (H_o) and expected heterozygosity (H_e), number of alleles per locus (A), deviation from Hardy–Weinberg equilibrium (HWE), fixation index (F), and occurrence of linkage disequilibrium were assessed by using Genepop version 4.0.11 (Rousset, 2008). Genotyping errors due to occurrence of null alleles were tested with the software Micro-Checker v. 2.2.3 (van Oosterhout et al., 2004).

From the 11 loci tested (Table 1), *Pmac082* was monomorphic in both populations analyzed, while *Pmac085* was monomorphic in FUR. No significant linkage disequilibrium was detected among loci after Bonferroni correction ($\alpha = 0.05$). The number of alleles per loci ranged from 1 to 8, with a total of 57 alleles on the 51 individuals scored. Expected heterozygosity ranged from 0.097 to 0.826, and observed heterozygosity ranged from 0.1 to 0.905. Significant departures from Hardy–Weinberg equilibrium after Bonferroni correction ($\alpha = 0.05$) were observed at one locus (*Pmac146*) in FUR and at three loci (*Pmac084*, *Pmac101*, and *Pmac146*) in DEL (Table 2). Micro-Checker pointed to the presence of null alleles ($P < 0.05$) in those three loci, but the observed HWE departures could be also due to inbreeding, as those populations have limited size.

We tested the 11 primer pairs described here for cross-species amplification in the other four species belonging to *Pilosocereus* [*P. jauruensis* (Buining & Brederoo) P. J. Braun (voucher specimen HUFS638, Brazil: Mato Grosso do

Sul, Aquidauana); *P. aurisetus* (Werderm.) Byles & G. D. Rowley (HUFS639, Brazil: Minas Gerais, Joaquim Felício; HUFS640, Brazil: Minas Gerais, Grão Mogol); *P. vilaboensis* (Diers & E. Esteves Pereira) P. J. Braun (HUFS641, Brazil: Goiás, Pirenópolis), and *P. aureispinus* (Buining & Brederoo) F. Ritter (HUFS642, Brazil: Bahia, Ibotirama)], using two individuals from each species. All the loci successfully amplified in the tested species, except for *Pmac101* and *Pmac108*, which did not amplify on any of these other species, and *Pmac085*, which failed to amplify on *P. aureispinus*.

CONCLUSIONS

The newly developed microsatellite markers showed high levels of polymorphism and will be useful for describing the genetic diversity and population structure across the distribution range of *P. machrisii*. These markers will also be used for understanding patterns of gene flow and to provide information about changes in historical distribution and about connectivity among populations. Furthermore, the successful cross-species amplification will allow the study of similar issues on other *Pilosocereus* species.

LITERATURE CITED

- ANDRADE-LIMA, D. 1981. The caatingas dominium. *Revista Brasileira de Botânica* 4: 149–153.
- GIULIETTI, A. M., J. R. PIRANI, AND R. M. HARLEY. 1997. Espinhaço range region, eastern Brazil. In S. D. Davis, V. H. Heywood, O. Herrera-Macbride, J. Villa-Lobos, and A. C. Hamilton [eds.], *Centres of plant diversity. A guide and strategy for their conservation*, vol. 3, 397–404. The Americas. IUCN Publication Unity, Cambridge, Massachusetts, USA.
- GLENN, T. C., AND N. A. SCHABLE. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* 395: 202–222.
- IUCN. 2010. IUCN Red List of threatened species. Version 2010.4. Website <http://www.iucnredlist.org/> [accessed on 11 January 2011].
- MORAES, E. M., A. G. ABREU, S. C. S. ANDRADE, F. M. SENE, AND V. N. SOLFERINI. 2005. Population genetic structure of two columnar cacti with a patchy distribution in eastern Brazil. *Genetica* 125: 311–323.
- ROUSSET, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Notes* 8: 103–106.
- SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.
- TAYLOR, N. P., AND D. C. ZAPPI. 2004. The cacti of eastern Brazil. Royal Botanic Gardens, Kew, England.
- VAN OOSTERHOUT, C. V., W. F. HUTCHINSON, D. P. M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology* 4: 535–538.
- WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
- ZAPPI, D. C. 1994. *Pilosocereus* (Cactaceae), the genus in Brazil. *Succulent Plant Research* 3: 1–160.