

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

Isabel Aparecida da Silva Bonatelli

Análise da estrutura filogeográfica das espécies do grupo *PILOSOCEREUS*
AURISSETUS (Cactaceae) utilizando marcadores moleculares do genoma do
cloroplasto (cpDNA)

São Carlos

2012

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

Isabel Aparecida da Silva Bonatelli

Análise da estrutura filogeográfica das espécies do grupo *PILOSOCEREUS*
AURISSETUS (Cactaceae) utilizando marcadores moleculares do genoma do
cloroplasto (cpDNA)

Dissertação apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular, como parte dos requisitos para obtenção do título de mestre em Genética Evolutiva e Biologia Molecular.

Orientação: Prof. Dr. Evandro Marsola de Moraes

São Carlos

2012

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

B699ae

Bonatelli, Isabel Aparecida da Silva.

Análise da estrutura filogeográfica das espécies do grupo *Pilosocereus Aurisetus* (Cactaceae) utilizando marcadores moleculares do genoma do cloroplasto (cpDNA) / Isabel Aparecida da Silva Bonatelli. -- São Carlos : UFSCar, 2012. 70 f.

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

1. Filogenia. 2. Filogeografia. 3. Genomas. 4. *Pilosocereus aurisetus*. 5. Cactaceae. I. Título.

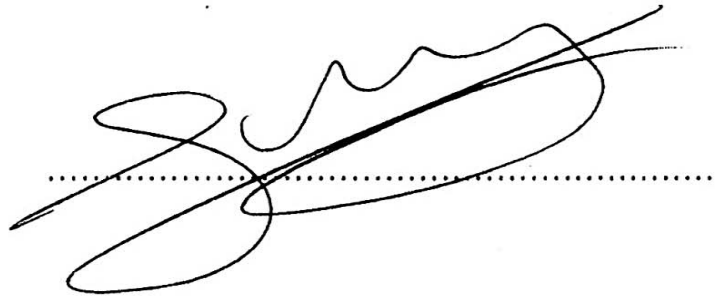
CDD: 575.0012 (20ª)

**“Análise da estrutura filogeográfica das espécies do grupo
PILOSOCEREUS AURISSETUS (Cactaceae) utilizando marcadores
moleculares do genoma do cloroplasto (cpDNA)”**

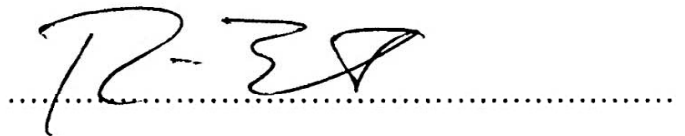
Dissertação de Mestrado de
Isabel Aparecida da Silva Bonatelli

Banca Examinadora

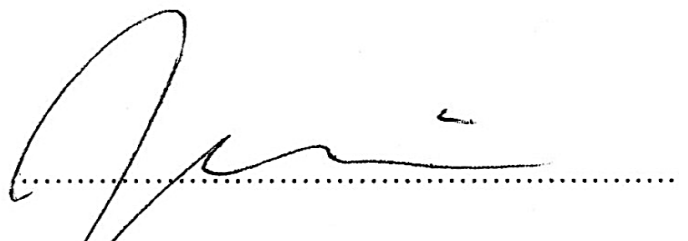
Prof. Dr. Evandro Marsola de Moraes



Prof. Dr. Reinaldo Alves de Brito



Prof. Dr. João Miguel de Barros Alexandrino



Este trabalho foi realizado com auxílio financeiro da Fundação de Amparo a Pesquisa do estado de São Paulo (FAPESP) através dos processos 2005/55200-8 e 2010/02186-6.

Dedico este trabalho aos meus pais, Theresinha e José António.

Agradecimentos

Ao professor Dr. Evandro Marsola de Moraes pela oportunidade, confiança, pelos ensinamentos diários e dedicação na orientação deste trabalho.

Ao professor Dr. Fernando de Faria Franco pela frequente disponibilidade em ajudar, pelos conselhos e ensinamentos, em especial pelo auxílio com a análise NCPA.

Aos meus pais Theresinha da Silva Bonatelli e José António Bonatelli por sempre acreditarem, pelo carinho, incentivo e sábios conselhos.

Ao meu namorado Danilo Trabuco do Amaral por todo auxílio na realização desse trabalho, pelo carinho, dedicação e presença constante na vida e na profissão.

Às técnicas e amigas Heide Utsunomiya e Fabrícia Madia pelo carinho, disposição e auxílio essencial em todas as etapas deste trabalho.

Ao meu companheiro de laboratório e Pós-Graduação, Manolo Perez, pelas opiniões, discussões e auxílio em tudo que fosse preciso.

Aos especialistas em Cactaceae, Daniela Zappi, Nigel Taylor e Marlon Machado por todo auxílio prestado na identificação das espécies, coletas e em diversas atividades do laboratório de Diversidade Genética e Evolução.

Aos colegas do laboratório de Diversidade Genética e Evolução pela amizade, apoio e por tornarem o dia-a-dia da pesquisa mais divertido.

Aos amigos Caroline Tocchet, Aline Silveira e Manolo Perez pelo auxílio com os mapas produzidos ao longo desse projeto.

Ao Sr. Gerardus Hubertus Olsthoorn pelas amostras gentilmente cedidas.

Ao colega de Pós-Graduação Wesley Generoso pela amizade, disponibilidade em ajudar e todo auxílio com as atividades da Pós-Graduação.

Aos professores e técnicos da UFSCar – Sorocaba pela amizade e colaboração com as atividades do presente projeto.

Aos professores e colegas da PPGGEV por complementarem minha formação acadêmica e contribuírem com minha formação científica.

Aos amigos que me acompanharam ao longo da realização desse projeto, por trazerem motivação para cada dia de trabalho.

A FAPESP pelo financiamento do projeto e pela concessão da bolsa de mestrado.

Ao CNPq e FAPESP pelo auxílio de manutenção do Laboratório de Diversidade Genética e Evolução da UFSCar - Campus Sorocaba.

RESUMO

PILOSOCEREUS AURISSETUS é um grupo taxonômico composto por oito espécies de cactos colunares que ocorrem associadas a afloramentos rochosos em ambientes xéricos. Assim como outras espécies xerófitas encontradas fora do domínio da Caatinga, as populações das espécies do grupo *P. AURISSETUS* exibem um padrão de distribuição descontínua dentro do domínio Cerrado no centro e leste do Brasil. Esse padrão de distribuição pode ser resultado de eventos de dispersão a longa distância ou de fragmentação de uma distribuição mais extensa no passado. O objetivo principal desse trabalho foi identificar eventos demográficos que possivelmente moldaram a história evolutiva do grupo e resultaram no padrão biogeográfico atual. No presente trabalho foram realizadas análises filogenéticas e inferências filogeográficas a partir da variação nucleotídica de sequências do genoma plastidial. As análises filogenéticas demonstraram que a maior parte das espécies do grupo não apresenta monofilia recíproca de seus haplótipos e que o grupo abrange três espécies bastante divergentes das demais: *P. jauruensis*, *P. aureispinus* e *P. bohlei*. Adicionalmente, as populações da espécie *P. machrisii* formaram dois clados distintos e bem suportados na filogenia, os quais concordam parcialmente com a distribuição geográfica da espécie (norte e centro-sul). As análises populacionais e as inferências filogeográficas permitiram inferir a origem da diversificação do grupo e os possíveis eventos de expansão e fragmentação que determinaram o caminho evolutivo das populações. Os níveis de variabilidade genética e estruturação observados, juntamente com o padrão atual de distribuição das populações e informações sobre a biologia das espécies, sugerem que eventos de fragmentação foram os principais responsáveis pelo padrão biogeográfico do grupo, os quais devem estar relacionados às oscilações paleoclimáticas do Quaternário.

Palavras-chave: Filogeografia. Filogenia. Genoma plastidial. *Pilosocereus*. Cactaceae.

ABSTRACT

PILOSOCEREUS AURISSETUS is a taxonomic group composed by eight columnar cacti species occurring on rock outcrops in xeric environments. As other xerophitic species found outside the Caatinga domain, the species populations of *P. AURISSETUS* group exhibit a disjunct distribution pattern within the Cerrado domain in central and eastern Brazil. Such a distribution pattern may have resulted from long-distance dispersal events or fragmentation of a more extensive distribution in the past. The main goal of this work was identify demographic events that possibly shaped the evolutionary history of the group and resulted in the current biogeographic pattern. In the present work, phylogenetic analyses and phylogeographical inferences were implemented using the nucleotide variation of plastid genomic sequences. The phylogenetic analyses showed the absence of reciprocal monophyly for great part of the species and the occurring of quite divergent species from the others of the group: *P. jauruensis*, *P. aureispinus* e *P. bohlei*. In addition, *P. machrisii* populations formed two distinct clades highly supported by the phylogeny which partly agree with the geographic distribution of the species (north and central-south). Population analyses and phylogeographical inferences allowed the identification of the diversification origin of the group and the possible expansion and fragmentation events which determined the evolutionary pathway of the populations. The genetic variability level and the genetic structure observed alongside the current distribution pattern of the populations and biological information about the species suggest that fragmentation events were the main responsible for the biogeographical pattern of the group, which should be related to the palioclimatic oscillations of the Quaternary.

Keywords: Phylogeography. Phylogeny. Plastid genome. *Pilosocereus*, Cactaceae.

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Estudos filogeográficos em plantas.....	1
1.2. Genoma Plastidial.....	3
1.3. Cactaceae no Brasil	4
1.4. Grupo <i>PILOSOCEREUS AURISSETUS</i>	5
2. OBJETIVOS	10
3. MATERIAL E MÉTODOS	11
3.1. Material biológico e extração de DNA	11
3.2. Amplificação e sequenciamento de regiões do genoma do cloroplasto	14
3.3. Análise da variação nucleotídica	15
3.4. Filogenia do grupo <i>P. AURISSETUS</i>	16
3.5. Rede de haplótipos e Nested Clade Phylogeographic Analysis (NCPA)	17
3.6. <i>Mismatch distribution</i> e testes de neutralidade	18
3.7. Estruturação populacional	19
4. RESULTADOS	20
4.1. Análise da variação nucleotídica	20
4.2. Análises Filogenéticas	23
4.3. Relação entre os haplótipos e Inferências Filogeográficas	25
4.4. <i>Mismatch distribution</i> e testes de neutralidade	33
4.5. Estruturação populacional	35
5. DISCUSSÃO	37
5.1. Variabilidade do cpDNA e Análises Filogenéticas	37
5.2. Estruturação genética	41
5.3. História Filogeográfica.....	42
5.4. Implicações para a conservação das espécies	47
6. CONCLUSÕES.....	48
7. REFERÊNCIAS BIBLIOGRÁFICAS	49
APÊNDICE 1	57
APÊNDICE 2	58
ANEXO.....	59

1. INTRODUÇÃO

1.1. Estudos filogeográficos em plantas

A Filogeografia é uma disciplina associada à Biogeografia que investiga eventos históricos e contemporâneos nas populações a partir de inferências genealógicas analisadas em uma escala de tempo e espaço (AVISE, 2009). O entendimento de eventos geológicos e ambientais bem como de fatores geográficos que influenciaram a história natural das espécies e que moldaram sua evolução estão entre os principais objetivos dos estudos filogeográficos (KNOWLES, 2009). Em plantas, esses estudos filogeográficos têm sido utilizados para relacionar a distribuição de linhagens genealógicas com eventos históricos como expansão e contração de populações, fragmentação e migração que ocorreram durante as oscilações climáticas do Quaternário (RAMOS et al., 2007; CAETANO et al., 2008).

Neste contexto, a origem da distribuição disjunta em espécies vegetais é um tema recorrente (GAUDEUL, 2006; COLLEVATTI et al., 2009). Esse padrão de distribuição pode ser resultado de fragmentação de uma população ancestral mais amplamente distribuída ou de dispersão a longa distância. A distribuição disjunta de algumas espécies arbóreas da floresta tropical sazonalmente seca da América Central e América do Sul pode ser explicada pela primeira hipótese, estando relacionada com a existência de populações relictuais das eras interglaciais (PRADO; GIBBS, 1993; PENNINGTON et al., 2004). Por outro lado, eventos de dispersão a longa distância são a possível causa da ocorrência de populações fragmentadas de *Symphonia globulifera* (DICK et al., 2003) e *Ceiba pentandra* (DICK et al., 2007), espécies arbóreas das florestas tropicais úmidas. No entanto, ainda são poucos os estudos envolvendo espécies neotropicais com distribuição descontínua, principalmente aquelas restritas a áreas secas.

Um desses estudos foi realizado para a espécie *Lychnophora ericoides* (Asteraceae), um arbusto endêmico do cerrado rupestre (COLLEVATTI et al., 2009). O cerrado rupestre é uma formação do domínio cerrado presente em elevadas altitudes caracterizado por manchas de vegetação xérica associadas a afloramentos rochosos (GIULLIETI et al., 1997). Esses ambientes ocorrem de forma restrita em planaltos e topos de montanha e abrigam um grande número de espécies endêmicas adaptadas ao ambiente seco. A distribuição

descontínua de *L. ericoides* foi associada a eventos de vicariância, uma vez que foi observada grande diferenciação e pouco fluxo gênico entre populações do leste e do centro do Brasil. Adicionalmente, a datação do ancestral comum mais recente (T_{MRCA}) foi inferida para o Pleistoceno médio, concordando com os ciclos sucessivos de expansão e contração das populações devido às oscilações climáticas do período.

Espécies animais podem apresentar uma distribuição disjunta devido à associação com espécies vegetais que apresentam o mesmo padrão de distribuição. A espécie *Drosophila gouveai*, por exemplo, possui uma distribuição estreitamente relacionada a cactos do gênero *Pilosocereus*, uma vez que as fêmeas depositam seus ovos em tecido necrosado desses cactos e as larvas se alimentam exclusivamente desse tecido. Dessa forma, a distribuição desse drosofilídeo reflete a distribuição descontínua de *Pilosocereus* (MORAES et al., 2009). A ocorrência de ciclos de expansão e contração das populações durante o Pleistoceno médio explicam que a distribuição de *D. gouveai* é consistente com as alterações palioclimáticas que influenciaram a distribuição da vegetação seca do leste do Brasil (MORAES et al., 2009).

Em alguns casos o padrão de distribuição disjunta é observado não apenas para a espécie, mas para todo o seu habitat, como é o caso da Floresta Tropical Sazonalmente Seca (*Seasonally Dry Tropical Forest* – SDTF) no leste da América do Sul. Atualmente essa formação vegetal ocorre bem definida no domínio da Caatinga e em enclaves no Cerrado e Chaco. A espécie arbórea *Astronium urundeuva* (Anacardiaceae), restrita a SDTF foi utilizada como modelo para o entendimento da distribuição dessa vegetação. O estudo revelou que a espécie apresenta uma estruturação em três grupos regionais (Norte, Centro e Sudoeste), sugerindo eventos de vicariância na formação SDTF e contato secundário das populações no centro da distribuição, área de transição entre Caatinga e Cerrado (CAETANO et al., 2008).

As espécies de bromélias *Alcantarea imperialis* e *A. geniculata* também apresentam uma distribuição naturalmente descontínua devido à ocorrência de seu habitat. Essas espécies se desenvolvem em afloramentos rochosos isolados (*inselbergs*) na mata atlântica, os quais apresentam um padrão de distribuição comumente descrito como “ilhas terrestres”. Estudos com essas espécies indicaram uma baixa variabilidade intraespecífica para o genoma plastidial, evidenciado pelo fluxo de sementes restrito entre populações

(BARBARÁ et al., 2008). A manutenção de um baixo fluxo gênico entre populações é atribuída à dificuldade do polinizador (morcegos) em percorrer diferentes *inselbergs*. Dessa forma, a distribuição descontínua de algumas espécies pode acarretar uma perda de variabilidade genética acompanhada de um aumento na diferenciação entre populações.

Uma particularidade dos estudos filogeográficos de plantas é a possibilidade da ocorrência de isolamento genético abaixo do nível de espécie e de compartilhamento de haplótipos entre espécies diferentes, um padrão originado por eventos de hibridização ou retenção de polimorfismo ancestral. Dessa forma, a inclusão de espécies relacionadas à espécie de interesse pode ser necessária, uma vez que os limites taxonômicos das espécies podem não concordar com os limites genéticos. Neste sentido todo o potencial de fluxo gênico deve ser considerado (SCHAAL et al., 1998).

1.2. Genoma Plastidial

As plantas apresentam dois genomas citoplasmáticos: mitocondrial e plastidial. Ao contrário do mtDNA dos animais, o mtDNA das plantas possui um tamanho relativamente grande e muito variável entre grupos (200 a 2.400 kb), extensiva recombinação intramolecular e uma baixa taxa de substituição nucleotídica (100 vezes menor que dos animais), sendo pouco adequado para estudos filogeográficos (AVISE, 2009). O cpDNA, por sua vez, é similar ao mtDNA de animais em sua estrutura circular, modo de replicação e padrão de herança não – Mendeliano, mas diferente no que se refere a propriedades evolutivas e moleculares. O genoma plastidial apresenta um tamanho relativamente pequeno, variando entre 120 a 217 kb em plantas fotossintetizantes, é composto por uma região de cópia única longa (LSC) e uma região de cópia única pequena (SSC), as quais são separadas por duas regiões de repetições invertidas (IR_A e IR_B). Apesar do cpDNA também evoluir em um ritmo lento quando comparado ao mtDNA animal, sua taxa de evolução é aproximadamente três a quatro vezes maior que a do mtDNA de plantas. Além disso, o cpDNA apresenta uma estrutura estável com pouca recombinação intramolecular permitindo a retenção de sinais históricos ao longo da evolução das populações (AVISE, 2009; RAVI, 2008). Dessa forma sequências do genoma plastidial são os marcadores moleculares mais comumente utilizados em estudos filogeográficos de plantas.

O cpDNA é herdado maternalmente na maioria das angiospermas (CORRIVEAU; COLEMAN, 1988). Dessa forma, o fluxo de genes maternos ocorre por meio do movimento de sementes, sendo geralmente mais restrito do que o fluxo de genes nucleares, os quais são transportados pelo movimento de pólen e sementes. Como o cpDNA possui um menor tamanho efetivo quando comparado ao DNA nuclear, devido a sua haploidia e herança uniparental, é esperado que esse marcador apresente maior estruturação espacial quando comparado a marcadores nucleares (PETIT et al., 2005). Para estudos abaixo do nível de espécie, as regiões não codificadoras do cpDNA são as mais utilizadas. A frequência de uso dessas regiões é decorrente de sua menor exposição à seleção natural, o que permite um maior acúmulo de variação nucleotídica. Além disso, a baixa influência da seleção nessas regiões indica que a variabilidade observada retém sinais de eventos históricos como expansão e subdivisão de populações e eventos recorrentes como fluxo gênico e mutação.

1.3. Cactaceae no Brasil

A família Cactaceae abriga mais de 1300 espécies e está distribuída ao longo do continente americano ocorrendo do sul da Patagônia ao Canadá (HUNT et al., 2006). O Brasil é um dos principais centros de diversidade e endemismo da família juntamente com México, Estados Unidos, Peru e Bolívia. Entre as 227 espécies reconhecidas até o momento no Brasil, 176 são endêmicas (78%). O leste do país concentra grande parte dessa diversidade abrangendo 130 espécies, sendo 117 endêmicas do Brasil, das quais 88 são endêmicas da região leste (ZAPPI et al., 2010). As espécies ocorrem principalmente ao longo do domínio da Caatinga, onde apresentam uma distribuição ampla e relativamente homogênea, e em enclaves de vegetação seca e afloramentos rochosos no Cerrado (campos rupestres), apresentando distribuição disjunta e restrita nesse bioma (TAYLOR; ZAPPI, 2004).

Como outras famílias de plantas, a classificação de Cactaceae é principalmente baseada em características morfológicas e, muitas vezes, as diferenças entre espécies para essas características são pouco perceptíveis. Uma desvantagem desse tipo de classificação é a plasticidade fenotípica que as espécies podem apresentar. Dessa forma, vários

pesquisadores preferem se utilizar da localização geográfica como principal parâmetro para a identificação taxonômica (HELSEN et al., 2009)

A ocorrência de populações de Cactaceae fora do domínio da Caatinga, juntamente com evidências de climas mais secos no passado para o Cerrado do centro e sudeste do Brasil (BEHLING, 2003) indicam que a distribuição atual dessas populações representa fragmentos remanescentes de uma distribuição mais ampla no passado (PENNINGTON et al., 2004). Dados palinológicos que sugerem a expansão da vegetação seca em áreas atualmente ocupadas por vegetação úmida corroboram a hipótese de expansão e contração da distribuição dessas espécies (BEHLING; HOOGHMESTRA, 2000). Dessa forma, espécies restritas a habitats xéricos provavelmente experimentaram alterações em sua distribuição envolvendo ciclos de aumento do fluxo gênico e posterior isolamento, os quais influenciaram a estrutura genética e a história evolutiva das populações (MORAES et al., 2009).

Atualmente, muitas populações remanescentes apresentam limitado potencial de dispersão por estarem separadas por ambientes inadequados ao fluxo gênico, o que favorece a retenção de sinais genéticos que podem revelar alterações na sua distribuição ao longo do tempo (MORAES et al., 2009). No entanto, os estudos genéticos envolvendo espécies de Cactaceae no Brasil priorizam o entendimento de relações filogenéticas, não investigando eventos históricos e distribuição da variabilidade genética. Apenas para algumas espécies dos gêneros *Facheiroa*, *Praecereus*, *Pilosocereus* e *Melocactus* foram realizadas análises sobre a diversidade genética e estrutura populacional a partir de marcadores isoenzimáticos (MORAES et al., 2005; LAMBERT et al., 2006a, 2006b).

1.4. Grupo *PILOSOCEREUS AURISSETUS*

Pilosocereus é um gênero neotropical de cactos colunares subdividido em grupos taxonômicos de espécies morfologicamente semelhantes (ZAPPI, 1994). Dentre eles, o grupo de espécies relacionadas à *P. AURISSETUS* é composto por oito espécies: *P. aurisetus*, *P. machrisii*, *P. vilaboensis*, *P. aureispinus*, *P. jauruensis*, *P. bohlei*, *P. parvus* e *P. pusillibaccatus* (Figura 1). Essas espécies apresentam grande variação morfológica ao longo da distribuição

geográfica resultando em uma história taxonômica instável, com várias espécies sendo inseridas ou removidas de sinonímia e novas espécies sendo descritas (ZAPPI, 1994; TAYLOR; ZAPPI, 2004; HUNT et al., 2006). As espécies *P. jauruensis*, *P. parvus* e *P. pusillibaccatus*, por exemplo, foram consideradas sinonímias de *P. machrisii* na revisão taxonômica realizada por Zappi (1994) e posteriormente elevadas ao status de espécie novamente por Hunt et al. (2006). Populações de *P. machrisii* da região de Cristalina-GO, Alto Paraíso de Goiás-GO e de Posse/Formosa-GO, foram originalmente descritas como *P. cristalinensis* (BRAUN; ESTEVES PEREIRA, 1987a), *P. lindaianus* (BRAUN; ESTEVES PEREIRA, 1987b, 1988) e *P. goianus* (BRAUN; ESTEVES PEREIRA 2002), respectivamente. Atualmente esses táxons são considerados sinonímias de *P. machrisii* (ZAPPI, 1994; TAYLOR; ZAPPI, 2004). Populações com ocorrência na Cadeia do Espinhaço, nas localidades de Engenheiro Dolabela-MG e Itamarandiba-MG foram inicialmente descritas como *P. supthutianus* (BRAUN; ESTEVES PEREIRA, 1985) e *P. werdermannianus* subespécie *densilanatus* (RITTER, 1979), respectivamente, e posteriormente sinonimizadas com *P. aurisetus* por Zappi (1994).

Dentre as características compartilhadas pelo grupo destacam-se a ramificação de cladódios na altura do solo, tamanho relativamente pequeno (raramente ultrapassando 3,5m), epiderme variando de verde escuro a verde-azulado e espinhos numerosos diferenciados em centrais e radiais. As flores apresentam antese noturna e coloração bem variada entre vermelho escuro, rosa e branco e são polinizadas por morcegos e beija-flores. Os frutos possuem coloração de vermelho-escuro a azulado e contém uma polpa funicular esbranquiçada ou avermelhada com sementes marrom-escuras ou pretas, as quais são dispersas por pássaros, morcegos e formigas (ZAPPI, 1994). *Pilosocereus aureispinus* apresenta algumas características peculiares em suas sementes como a forma alongada e células da testa cônicas que podem estar relacionadas à dispersão por formigas, as quais são abundantes no Cerrado. O grupo *P. AURISSETUS* também apresenta reprodução clonal por meio de fragmentos de cladódio que se destacam e originam um novo indivíduo por brotamento. Não existem relatos sobre ocorrência de hibridização nessas espécies e a existência de flores auto-incompatíveis foi observada em outras espécies do gênero *Pilosocereus* (LOCATELLI et al., 1997; NASSAR et al., 1997).

O grupo *P. AURISSETUS* apresenta uma distribuição bastante disjunta e muitas vezes restrita. As populações ocorrem em elevadas altitudes, em enclaves de campos rupestres no centro e leste do Brasil, principalmente associadas a afloramentos rochosos areníticos, quartzíticos ou de rocha calcária.

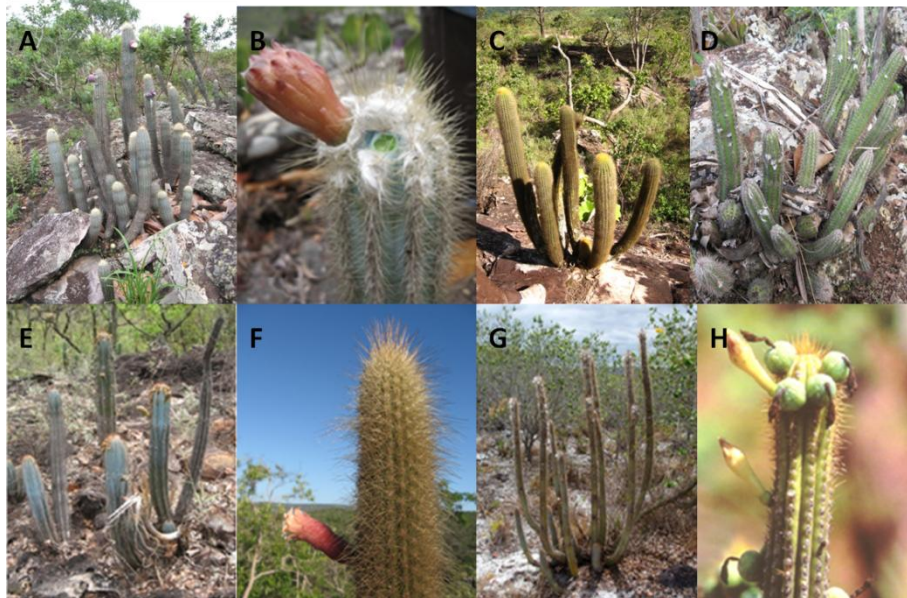


Figura 1. Espécies que compõem o grupo *PILOSOCEREUS AURISSETUS*. (A) *P. machrisii*, (B) *P. aurisetus*, (C) *P. vilaboensis*, (D) *P. parvus*, (E) *P. jauruensis*, (F) *P. aureispinus*, (G) *P. bohlei* e (H) *P. pusillibaccatus*.

A única ocorrência de sobreposição de espécies acontece na localidade de Posse-GO, onde estão localizadas a espécie *P. parvus* e uma população de *P. machrisii*. No entanto, essas populações ocupam diferentes afloramentos rochosos, sendo possível afirmar que as espécies do grupo não ocorrem em simpatria, como foi observado em estudos de campo anteriores (TAYLOR; ZAPPI 2004). Com base nesse cenário de distribuição, o padrão de especiação do grupo é considerado como predominantemente alopátrico (ZAPPI, 1994).

Em 2006, as espécies *P. vilaboensis* e *P. parvus* foram consideradas pela INTERNATIONAL UNION FOR CONSERVATION OF NATURE - IUCN como vulneráveis a extinção. *P. aurisetus aurilanatus* foi considerada em perigo de extinção pela IUCN (HUNT et al., 2006) e também está na lista oficial de espécies brasileiras ameaçadas de extinção (BRASIL, 2008). Atualmente, a única espécie encontrada na lista vermelha da IUCN é *P. aureispinus*, categorizada como em deficiência de dados (INTERNATIONAL UNION FOR

CONSERVATION OF NATURE, 2012). De fato, essa espécie bem como *P. bohlei* apresenta uma distribuição muito restrita. A comprovação da ocorrência limitada dessas espécies certamente acarretaria um status nas categorias de risco. O endemismo natural e a vulnerabilidade ambiental a qual essas espécies estão sujeitas, bem como algumas atividades humanas como queimadas e produção de carvão vegetal que ocorrem no norte de Minas Gerais, devem influenciar o atual estado de conservação das mesmas.

A espécie *P. machrisii* é a mais amplamente distribuída do grupo, ocorrendo no norte do estado de São Paulo, sul de Minas Gerais, centro (Planalto Central), leste e norte de Goiás (Chapada dos Veadeiros) e oeste da Bahia. *Pilosocereus aurisetus* apresenta distribuição restrita à região da Cadeia do Espinhaço em Minas Gerais, sendo dividida em duas subespécies: *P. aurisetus aurilanatus* localizada no oeste da distribuição, no município de Joaquim Felício, e *P. aurisetus aurisetus* distribuída em grande parte da Cadeia do Espinhaço no estado de Minas Gerais. A forma de *P. aurisetus* localizada no extremo leste da distribuição (Itamarandiba) é reconhecida por alguns pesquisadores pelo epíteto *P. densilanatus*, pois foi reconhecida como a espécie *P. werdermannianus* var. *densilanatus* no passado (RITTER, 1979).

Pilosocereus jauruensis ocorre no centro e norte do estado de Mato Grosso do Sul ao longo da Serra de Maracajú. *Pilosocereus vilaboensis* ocorre no centro de Goiás, com poucas populações conhecidas na Serra dos Pirineus e Serra Dourada. *Pilosocereus aureispinus*, *P. bohlei*, *P. parvus* e *P. pusillibaccatos* são espécies microendêmicas, possuindo uma ou poucas populações muito próximas. *Pilosocereus aureispinus* e *P. bohlei* ocorrem na região central do estado da Bahia, *P. parvus* na divisa entre Goiás e Bahia e *P. pusillibaccatus* no sul do estado do Piauí (ZAPPI, 1994; HUNT et al., 2006). Localmente, essas espécies são distribuídas em manchas de vegetação xérica associadas a formações rochosas devido à descontinuidade do seu habitat.

A distribuição disjunta das espécies do grupo *P. AURISSETUS* pode ser resultado de eventos de vicariância relacionados à contração da distribuição das espécies no passado ou de dispersão a longa distância e colonização de novas áreas adequadas. A primeira hipótese resultaria em grande diferenciação populacional e ausência ou grande restrição de fluxo

gênico, enquanto que para a segunda hipótese seria esperado pouca diferenciação populacional e sinais de ocorrência de fluxo gênico.

No presente estudo, análises filogenéticas e filogeográficas foram usadas para entender as relações evolutivas entre as espécies, bem como os eventos biogeográficos e demográficos que influenciaram o padrão atual de distribuição do grupo *P. AURISSETUS* e, possivelmente outros táxons relacionados aos mesmos ambientes. A hipótese de trabalho é que a distribuição disjunta do grupo é um remanescente de uma distribuição mais ampla no passado.

2. OBJETIVOS

O objetivo inicial do presente projeto foi estudar a variabilidade genética e a estrutura filogeográfica da espécie *P. machrisii* a partir de sequências nucleotídicas do cpDNA. No entanto, o reconhecimento por estudos conduzidos pelo nosso grupo de pesquisa sobre a estreita relação genética entre as espécies do grupo *P. AURISSETUS* revelou a necessidade de o presente trabalho envolver todas as espécies desse grupo taxonômico e não apenas a espécie *P. machrisii*. As análises genéticas iniciais mostraram que todas as espécies provavelmente compartilham uma história evolutiva comum e recente, refletindo no compartilhamento de linhagens gênicas. Dessa forma, os objetivos específicos do projeto inicial foram ampliados para todas as espécies do grupo *P. AURISSETUS*, com exceção da espécie *P. pusillibaccatus*, para a qual não houve amostra disponível. Os objetivos específicos desse estudo foram: (1) estudar a distribuição geográfica das linhagens gênicas por meio da construção de uma rede de haplótipos; (2) verificar o nível e o padrão de distribuição da variação genética dentro e entre populações; (3) verificar sinais na variação genética que indiquem alterações históricas na distribuição e conexão das populações, (4) contribuir com dados a respeito do estado de conservação das espécies.

3.MATERIAL E MÉTODOS

3.1. Material biológico e extração de DNA

As coletas necessárias para esse trabalho foram realizadas no período de outubro de 2007 a julho de 2011, durante o desenvolvimento de um projeto mais amplo ao qual este está associado (“Estudo filogeográfico do grupo de espécies relacionadas à *P. AURISSETUS* (Cactaceae) utilizando DNA microssatélite nuclear e sequências nucleotídicas do genoma do cloroplasto”) realizado no laboratório de Diversidade Genética e Evolução da UFSCar-Sorocaba. Foi estabelecida uma distância mínima de 10 metros entre os indivíduos coletados a fim de se evitar a coleta de clones. A amostra proveniente de Campos Belos e algumas amostras de Aurora do Tocantins e Alto Paraíso de Goiás foram gentilmente cedidas pelo Sr. Gerardus Hubertus Olsthoorn. As localidades de coleta são mostradas na figura 2.

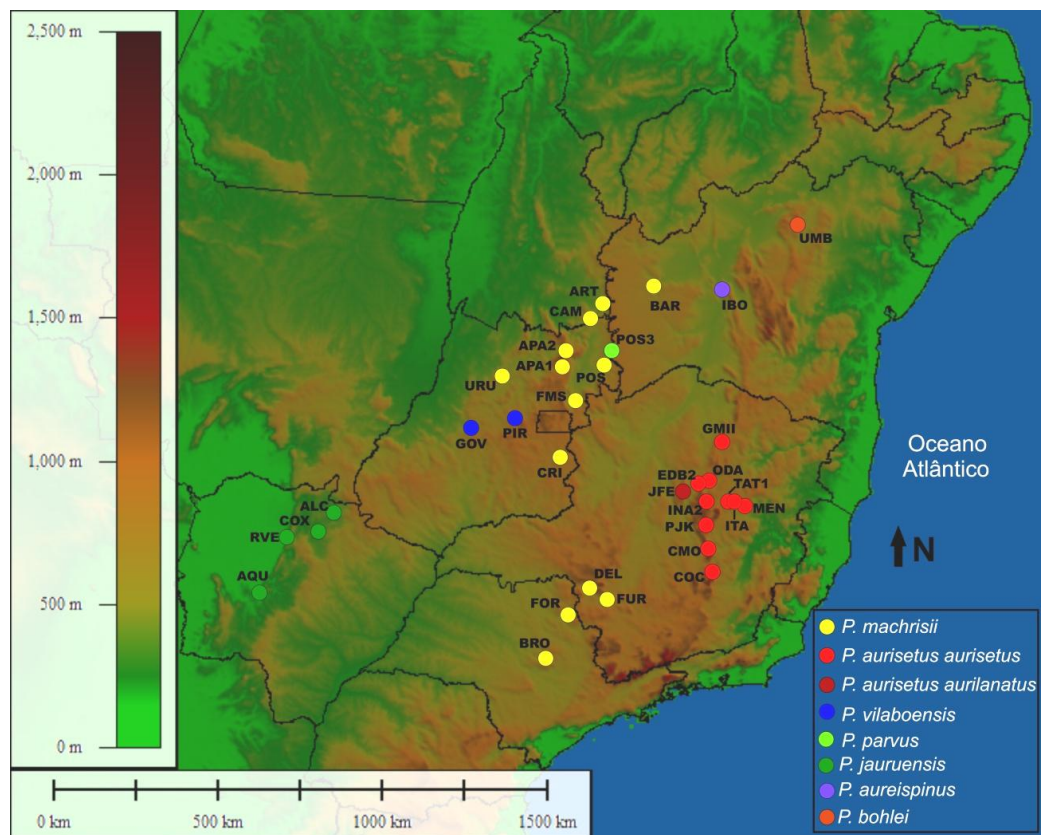


Figura 2. Mapa de elevação do leste do Brasil indicando a localização das populações amostradas no presente estudo. Os códigos das populações são mostrados na tabela 1.

A identificação taxonômica das espécies foi realizada segundo as descrições da localização geográfica e das características morfológicas, contidas nos trabalhos de Zappi (1994) e Hunt et al. (2006), e confirmada por análise de material fotográfico pela Dra. Daniela Zappi (National Parks Board, Gardens by the Bay and Singapore Botanic Gardens, Singapore).

No total foram analisados 180 indivíduos provenientes de 13 populações de *P. machrisii*, 11 de *P. aurisetus* (incluindo as subespécies *P. a. aurisetus* e *P. a. aurilanatus*) quatro de *P. jauruensis*, duas de *P. vilaboensis* e uma população das seguintes espécies: *P. aureispinus*, *P. bohlei* e *P. parvus* (Tabela 1).

O DNA utilizado neste projeto foi extraído no Laboratório de Diversidade Genética e Evolução a partir de maceração em nitrogênio líquido para rompimento da parede celular, e o kit comercial Qiagen - Dneasy Plant Mini Kit (Qiagen, Hilden, Germany) com todas as centrifugações em velocidade máxima (13.500 rpm).

Tabela 1. Populações das espécies do grupo *P. aurisetus* e número de indivíduos analisados no trabalho

Espécie	Localidade	Código	Data coleta	Coordenadas Geográficas	Altitude	N	
<i>P. machrisii</i>	Brotas-SP	BRO	13/10/2007	22°17'25.5"S/47°57'57.9"W	802m	5	
	Furnas-MG	FUR	27/10/2007	20°38'53.4"S/46°19'16.4"W	789m	5	
	Delfinópolis-MG	DEL	28/10/2007	20°21'44.5"S/46°44'59.4"W	1082m	8	
	Cristalina-GO	CRI	16/12/2007	16°47'43.6"S/47°34'10.7"W	1126m	5	
	Altinópolis-SP	FOR	30/08/2008	21°05'43.8"S/47°20'03.2"W	780m	6	
	Formosa-GO	FMS	07/01/2009	15°14'00.3"S/47°09'10.8"W	797m	6	
	Posse-GO	POS	07/01/2009	14°15'16.7"S/46°21'49.2"W	794m	5	
	Alto Paraíso de Goiás-GO 1	APA1	08/01/2009	14°12'34.3"S/47°29'15.2"W	1204m	10	
	Alto Paraíso de Goiás-GO 2	APA2	01/06/2010	13°55'33.6"S/47°25'47.5"W	1461m	11	
	Uruaçu-GO	URU	10/01/2009	14°33'27.4"S/49°09'23.3"W	588m	6	
	Barreiras-BA	BAR	01/02/2009	12°05'45.2"S/45°02'57.5"W	687m	5	
	Aurora do Tocantins-TO	ART	30/05/2010	12°36'34.6"S/46°24'10.1"W	501m	10	
	Campos Belos-GO	CAM	Nov/2006	12°59'22.0"S/46°43'44.0"W	753m	1	
	<i>P. aurisetus</i>	Joaquim Felício-MG	JFE	18/01/2008	17°41'25.5"S/44°11'54.8"W	968m	5
		Olhos D'Água MG	ODA	19/01/2008	17°26'19.8"S/43°37'20.5"W	724m	5
Mendanha-MG		MEN	20/01/2008	18°07'08.1"S/42°32'11.7"W	859m	5	
Presidente Kubitschek-MG		PJK	21/01/2008	18°37'52.4"S/43°33'04.3"W	1208m	5	
Cardeal Mota-MG		CMO	23/01/2008	19°16'01.4"S/43°30'57.4"W	1251 m	5	
Grão Mogol-MG		GMII	26/01/2010	16°19'59.2"S/43°07'48.0"W	917m	5	
Itamarandiba-MG		ITA	28/01/2010	18°00'07.0"S/42°56'09.5"W	1034m	5	
Cocais-MG		COC	29/01/2010	19°52'43.8"S/43°27'52.3"W	867m	5	
Engenheiro Dolabela-MG		EDB2	12/07/2011	17°27'05.1"S/43°47'37.2"W	688m	5	
Tromba D'Anta (Itamarandiba-MG)		TAT1	13/07/2011	18°01'00.6"S/42°57'49.5"W	1118m	5	
Inhaí-MG		INA2	17/07/2011	17°58'47.2"S/43°36'21.0"W	798m	2	
<i>P. jauruensis</i>	Aquidauana-MS	AQU	17/09/2009	20°21'41.2"S/55°41'36.2"W	292m	5	
	Rio Verde do Mato Grosso-MS	RVE	20/09/2009	18°50'27.6"S/54°59'29.9"W	474m	5	
	Alcinópolis-MS	ALC	21/09/2009	18°08'53.9"S/53°40'49.0"W	605m	5	
	Coxim-MS	COX	17/02/2011	18°30'36.0"S/ 54°45'00.0"W	-	5	
<i>P. vilaboensis</i>	Pirenópolis-GO	PIR	12/01/2009	15°48'24.4"S/48°50'46.0"W	1284m	6	
	Goiás Velho -GO	GOV	11/01/2009	16°00'04.1"S/50°03'03.3"W	765m	5	
<i>P. aureispinus</i>	Ibotirama-BA	IBO	30/01/2009	12°05'21.0"S/43°09'28.5"W	502m	5	
<i>P. parvus</i>	Posse - GO	POS3	07/01/2011	13°57'56.1"S/46°11'31.2"W	925m	4	
<i>P. bohlei</i>	Umburanas-BA	UMB	11/01/2011	10°21'54.3"S/41°11'46.6"W	960m	5	

3.2. Amplificação e sequenciamento de regiões do genoma do cloroplasto

Foram realizadas reações de amplificação de três regiões do genoma do cloroplasto, o espaçador intergênico *trnT-trnL* e as porções *3'trnS-trnG* e *5'trnS-trnG* do espaçador *trnS-trnG* (Figura 3). As PCRs para algumas populações de *P. aurisetus*, *P. vilaboensis* e *P. jauruensis* foram realizadas por colaboradores do Laboratório de “Diversidade Genética e Evolução”.

A região *trnT-trnL* foi amplificada a partir de iniciadores desenvolvidos por Taberlet et al. (1991) em um volume de 20 µl contendo 1X tampão de reação [100 mM Tris-HCl (pH 8,8), 500 mM KCl, 0,8% Nonidet P40], 3 mM de MgCl₂, 200µM de cada dNTP, 0,1µM de cada iniciador, 1,25U de Taq DNA polymerase (Fermentas, Burlington, Ontario, Canadá) e 1µl de DNA (5-40ng). As regiões *3'trnS-trnG* e *5'trnS-trnG* estão contidas no espaçador *trnS-trnG* e apresentaram sobreposição de 138 pb. Para a amplificação de *3'trnS-trnG* foram utilizados os iniciadores *5'trnG2S* (SHAW et al., 2005) e *SGFwd2* (BONATELLI et al., 2012) e, para *5'trnS-trnG*, os iniciadores *trnS^{GCU}* (SHAW et al., 2005) e *SGRev2* (BONATELLI et al., 2012). As concentrações dos reagentes foram as mesmas das utilizadas para *trnT-trnL*, com exceção do MgCl₂ que foi usado a 1,5 mM e a Taq polymerase que foi usada com 1U. As condições físicas das reações foram: Desnaturação inicial de 80°C por 5 min, 30-40 ciclos de Desnaturação de 95°C por 1 min., Hibridização de 50°C (*trnT-trnL*) e 62°C (*3'trnS-trnG* e *5'trnS-trnG*) por 1 min e Extensão de 65°C por 5 min, seguidos de Extensão Final de 65°C por 5 min. Todas as amplificações foram realizadas em um termociclador Eppendorf Mastercycler[®] gradient (Eppendorf AG, Hamburg, Germany).

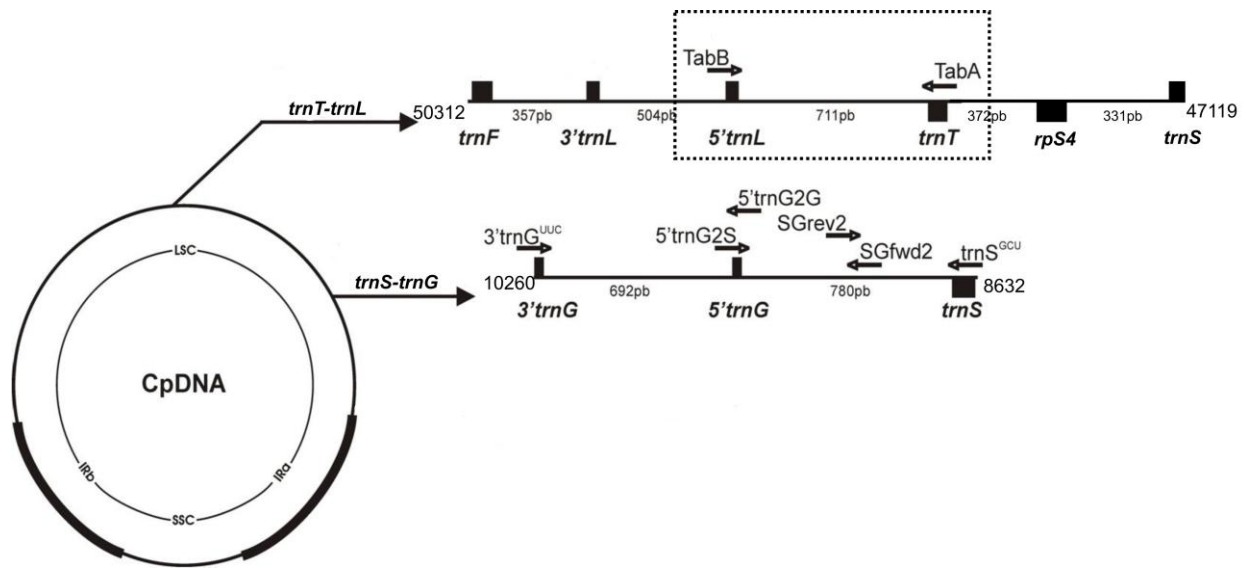


Figura 3. Mapa de escala dos segmentos do genoma do cloroplasto e os *primers* utilizados no estudo com o grupo *P. AURISSETUS* (redesenhado a partir de WAKASUGI et al., 1998 e SHAW et al., 2005 com base no genoma de *Nicotiana*). SSC e LSC são, respectivamente, as regiões de cópia única pequena e grande. IRA e IRB são as regiões invertidas do cpDNA. Todos os segmentos estudados estão localizados na grande região de cópia única. Os números ao lado das regiões indicam a posição no genoma de *Nicotiana*.

Os produtos de PCR foram conferidos em gel de agarose 1,5% corado com brometo de etídio e purificados com *ExoSAP - IT* ou *illustra GFX PCR DNA and Gel Band Purification kit* (GE Healthcare, Piscataway, New Jersey, USA). Foram sequenciadas as duas fitas de DNA usando os mesmos iniciadores da amplificação. Os sequenciamentos foram realizados no Centro de Estudos do Genoma Humano (CEGH – USP) e no Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista (UNESP) - Campus Jaboticabal com o sequenciador ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster City, California, USA) e o kit de sequenciamento Big Dye terminator 3.1 (Applied Biosystems, Foster City, California, USA).

3.3. Análise da variação nucleotídica

Os cromatogramas das duas fitas de DNA foram combinados no software ChromasPro 1.5 (Technelysium Pty Ltd., Tewantin, Australia), permitindo a resolução de inconsistências entre os cromatogramas. As seqüências obtidas foram alinhadas a partir do

software CLUSTALW (THOMPSON et al., 1994). Variações do tipo inserção-deleção (*indels*) de mesmo tamanho e limites foram analisadas como um único evento mutacional e convertidos em um sítio variável para os indivíduos que possuíam os *indels*, conforme Amico e Nickrent (2009). As sequências das três regiões do cpDNA de cada indivíduo foram então concatenadas para serem utilizadas nas análises. As análises da variação nucleotídica foram realizadas com o software DNAsp (ROZAS et al., 2003) a partir do qual foram calculados o número de sítios polimórficos, o número de haplótipos (Nh), a variabilidade nucleotídica (π), o valor de theta por sequência (θ) e a diversidade haplotípica (h) das espécies.

3.4. Filogenia do grupo *P. AURISSETUS*

As relações filogenéticas entre os haplótipos foram estimadas a partir de uma árvore de Máxima Verossimilhança implementada no programa MEGA v4 (TAMURA et al., 2007), usando a espécie *Arrajadoa rhodantha* (Cactaceae) como grupo externo. O modelo de evolução mais adequado foi identificado pelo critério de AIC (Akaike Information Criterion) a partir do software jModelTest 1.1 (POSADA, 2008). A árvore MV foi inferida usando o algoritmo de busca heurística CNI (Close-Neighbor-Interchange) e o modelo de evolução General time-reversible com taxa de distribuição gama (GTR+G). O suporte de *bootstrap* dos ramos foi estimado a partir de 5000 replicações. Inferências filogenéticas também foram realizadas a partir de uma abordagem Bayesiana usando a Cadeia Markoviana de Monte Carlo (MCMC), com o software BEAST v1.6.1 (DRUMMOND et al., 2010). Foram realizadas duas corridas independentes com 10^6 gerações, as quais foram amostradas a cada 1000 gerações. Os parâmetros utilizados foram o modelo de evolução GTR+G, selecionado pelo jModelTest, com quatro categorias de gama e um relógio molecular relaxado. Após a análise, 25% das gerações iniciais foram excluídas das análises e as topologias remanescentes foram usadas para estimar a credibilidade dos clados usando o programa TreeAnnotator, encontrado no software BEAST. O suporte estatístico dos ramos foi determinado pela probabilidade Bayesiana posterior.

3.5. Rede de haplótipos e Nested Clade Phylogeographic Analysis (NCPA)

A relação entre os haplótipos obtidos foi estimada a partir de uma rede de parcimônia estatística construída com o software TCS 1.21 (CLEMENT et al., 2000) e as conexões estabelecidas com limite de 95%. As reticulações presentes foram resolvidas segundo os critérios geográfico e topológico descritos em Crandall e Templeton (1993). No primeiro critério predominam as relações entre os haplótipos originados pela mesma população, enquanto no segundo predominam as relações com os haplótipos de interior, ou seja, há maior probabilidade dos haplótipos estarem conectados a haplótipos centrais (antigos) do que a haplótipos das extremidades (recentes).

As inferências filogeográficas foram estimadas a partir da Análise Filogeográfica de Clados Aninhados (*Nested Clade Phylogeographic Analysis*; TEMPLETON et al., 1995, 2004). Para essa análise foram excluídas as populações de *P. jauruensis*, *P. bohlei* e *P. parvus*, devido a grande divergência entre essas espécies e as demais do grupo. A análise NCPA é baseada no aninhamento de clados definidos na rede de haplótipos e na realização de testes estatísticos de associação entre os clados e sua distribuição geográfica. O aninhamento dos clados foi obtido a partir do agrupamento hierárquico dos haplótipos amostrados ou não amostrados (representados por pequenos círculos na rede de haplótipos) segundo os critérios de Templeton et al. (1987, 1995). Os agrupamentos formados foram definidos como clados, os quais foram separados em níveis. Clados de primeiro nível foram formados por haplótipos, clados de segundo nível foram formados por clados de primeiro nível e assim sucessivamente foram formados clados abrangentes, até toda a rede estar contida em um único clado.

Os testes estatísticos de associação entre clados e distribuição geográfica foram realizados por meio dos índices de distância do clado (Dc) e de distância de agrupamento (Dn). O primeiro mede a amplitude da distribuição geográfica de cada haplótipo e o segundo, a distância média de um haplótipo em relação ao centro geográfico de outros haplótipos que são relacionados evolutivamente, ou seja, compartilham um mesmo clado em um nível de agrupamento superior. Os índices Dc e Dn bem como sua significância foram calculados com o programa GeoDis 2.6 com 5000 permutações (POSADA et al., 2000),

partindo da hipótese nula de que os haplótipos estão distribuídos aleatoriamente ao longo das localidades (POSADA et al., 2000). Para a interpretação dos resultados e identificação dos eventos históricos envolvidos nos padrões observados foi utilizada uma chave de inferências descrita em Templeton et al. (1995, 2004) segundo a versão 2011 encontrada no programa GeoDis 2.6 (disponível em: <http://darwin.uvigo.es/software/geodis.html>). Essa versão da chave de inferências é mais conservadora do que as anteriores por implementar a correção de múltiplos testes de Dunn-Sidak. Dessa forma, os resultados foram considerados significativos ao nível de 5% quando a probabilidade de ocorrência foi menor que o valor crítico de 0,0170.

3.6. *Mismatch distribution* e testes de neutralidade

A história demográfica do grupo *P. AURISSETUS* foi inferida a partir de uma análise de *Mismatch distribution* (ROGERS; HARPENDING, 1992) utilizando o modelo de expansão espacial das populações com 1000 replicações. As espécies *P. jauruensis*, *P. aureispinus* e *P. bohlei* foram excluídas por apresentarem muitas diferenças em relação às demais. Essa análise permite a identificação de alterações demográficas nas populações a partir do número médio de diferenças par a par entre as sequências de DNA dos indivíduos. Populações em equilíbrio demográfico apresentam um padrão gráfico multimodal, enquanto que populações que sofreram expansão recente devem apresentar uma distribuição unimodal (ROGERS; HARPENDING, 1992). A significância da análise foi obtida pelo teste SSD (*Sum of Squared Differences*).

Testes de neutralidade de F_s (FU, 1997) e D (TAJIMA, 1989) também foram utilizados para testar o equilíbrio populacional. O teste D (TAJIMA, 1989) compara duas estimativas independentes a partir do parâmetro de mutação θ ($\theta=2Neu$). A primeira estimativa corresponde ao número de sítios polimórficos, enquanto a segunda estimativa corresponde a média da distribuição dos pares de diferenças entre as sequências. Segundo Tajima (1989), a diferença (D) entre essas estimativas pode detectar desvios da neutralidade, sendo que valores negativos sugerem seleção positiva (efeito carona), negativa ou expansão populacional. O teste F_s (FU, 1997) compara o número de alelos observados com o número

de alelos esperados em uma amostra aleatória, permitindo a identificação de alelos raros, ou seja, mutações recentes. Valores de F_s tendem a ser negativos mediante um excesso de mutações recentes, indicando expansão populacional. Essas análises foram realizadas com o software Arlequin ver. 3.11 (EXCOFFIER et al., 2005).

3.7. Estruturação populacional

A existência de estruturação genética dentro do grupo *P. AURISSETUS* foi estimada a partir de uma análise espacial da variância molecular implementada no programa SAMOVA 1.0 (DUPANLOUP et al., 2002). Esta análise agrupa populações geograficamente homogêneas e altamente diferentes dos demais grupos, a partir de uma definição a priori do número de grupos (K) e de uma análise de variância molecular (AMOVA), resultando em valores de F_{ST} , F_{SC} e F_{CT} (EXCOFFIER et al., 1992). A partir da comparação dos valores de F_{CT} (o nível de diferenciação entre grupos) e F_{SC} (o nível de diferenciação dentro dos grupos) é possível identificar um número ótimo de agrupamentos. Foram realizados testes de SAMOVA com valores de k variando entre dois e 32 utilizando 100 simulações com a finalidade de observar os diferentes agrupamentos possíveis e identificar grupos nos quais os valores de F_{CT} fossem maiores que os de F_{SC} . Para a espécie *P. machrisii* foi realizada uma AMOVA a partir dos clados regionais de haplótipos identificados na análise filogenética e na rede de haplótipos. Esta análise foi realizada no programa Arlequin ver. 3.11 (EXCOFFIER et al., 2005) utilizando 1000 replicações e o método de distância pareada dos haplótipos (“pairwise differences”).

4. RESULTADOS

4.1. Análise da variação nucleotídica

Os alinhamentos dos espaçadores intergênicos *trnT-trnL*, *3'trnS-trnG* e *5'trnS-trnG* geraram 311 pb, 690 pb, 612 pb, respectivamente, totalizando um segmento concatenado de 1475 pb. Quando todos os 180 indivíduos de todas as espécies foram comparados 46 sítios variáveis foram encontrados, 43 com duas variantes e três com três variantes. Também foram observados 15 indels, dos quais três repetições mononucleotídicas foram excluídas das análises devido a ambigüidade dos dados e 10 indels foram considerados parcimoniosamente informativos (Tabela 2).

Tabela 2. Caracterização da variação nucleotídica dos espaçadores plastidiais *trnT-trnL*, *3'trnS-trnG* e *5'trnS-trnG* para as espécies do grupo *P. AURISSETUS*. Número de sequências analisadas (N), número de populações analisadas (p), número de haplótipos (Nh), haplótipos observados (Hap), número de substituições (S), número de indels (I), variabilidade nucleotídica (π), diversidade haplotípica (h), theta por sequência (θ) e nível taxonômico da variação são indicados (Nt).

	N/p	Nh	Hap	S	I	π	h	θ	Nt
<i>P. machrisii</i>	83/13	9	H4-H12	9	2	0,00242	0,860	2,204	Pw, Pa
<i>P. aurisetus</i>	52/11	3	H1-H3	5	1	0,00091	0,336	1,328	Pa
<i>P. vilaboensis</i>	11/2	4	H16-H19	4	2	0,00173	0,709	2,049	Pw, Pa
<i>P. jauruensis</i>	20/4	2	H20, H21	1	0	0,00027	0,395	0,282	Pa
<i>P. aureispinus</i>	5/1	2	H14, H15	2	0	0,00054	0,400	0,960	Pw
<i>P. parvus</i>	4/1	1	H13	0	0	0	0	0	—
<i>P. bohlei</i>	5/1	2	H22, H23	2	0	0,00054	0,400	0,960	Pw

Pw: variação dentro da população; Pa: variação entre populações

A maior parte da variação nucleotídica foi encontrada no nível interespecífico. Dentre a variação nucleotídica intraespecífica observada estão nove substituições e dois indels (10 pb e 30 pb) em *P. machrisii*, cinco substituições e um indel (10 pb) em *P. aurisetus*, quatro substituições e dois indels (1 pb) em *P. vilaboensis*, uma substituição em *P. jauruensis*, duas substituições em *P. aureispinus* e duas substituições em *P. bohlei*.

No total foram observados 23 haplótipos (Tabela 3), nove em *P. machrisii*, três em *P. aurisetus*, quatro em *P. vilaboensis*, dois em *P. jauruensis*, *P. aureispinus* e *P. bohlei* e apenas um em *P. parvus*. Variação dentro de populações foi identificada apenas nas populações ART de *P. machrisii*, GOV e PIR de *P. vilaboensis*, IBO de *P. aureispinus* e UMB de *P. bohlei* (Tabela 2).

No geral as espécies apresentaram baixa variabilidade nucleotídica e diversidade haplotípica ($\pi=0,00027$ a $0,00242$, $h=0,336$ a $0,860$). *Pilosocereus machrisii*, a espécie mais amplamente distribuída do grupo e, conseqüentemente, a mais amostrada em número de populações, apresentou maior variabilidade genética ($\pi=0,00242$ e $h=0,860$). No entanto, um aumento no número de populações amostradas não conferiu necessariamente um aumento na variabilidade observada. *Pilosocereus vilaboensis*, por exemplo, com apenas duas populações analisadas apresentou valores de variabilidade superiores aos de *P. aurisetus* ($h=0,336$), espécie amostrada para 11 populações. A diversidade haplotípica observada para *P. aurisetus* também foi menor do que o observado para espécies com distribuição bastante restrita, limitadas a uma população (*P. aureispinus*, *P. parvus* e *P. bohlei*, $h=0,400$).

4.2. Análises Filogenéticas

As árvores ML e Bayesiana apresentaram topologias semelhantes, sendo que foram observados padrões monofiléticos apenas para os haplótipos das espécies mais divergentes (Figura 4). Nas duas análises os haplótipos das espécies *P. bohlei* e *P. aureispinus* se mostraram os mais divergentes do grupo.

A existência de dois clados envolvendo os haplótipos de *P. machrisii* foi bem suportada, concordando parcialmente com a distribuição geográfica da espécie (norte e centro-sul), com exceção de haplótipos encontrados em ART (H8), URU (H9) e APA2 (H9), os quais não apresentaram relação filogenética bem resolvida com os demais haplótipos do grupo. A relação estreita entre os haplótipos de APA1 (H5) e CRI/FMS (H4) também foi bem suportada. Os haplótipos encontrados para *P. vilaboensis* foram agrupados em dois clados distintos de acordo com a população de ocorrência.

A principal discordância observada entre as topologias geradas foi a posição dos haplótipos de *P. jauruensis*. A árvore ML agrupou esses haplótipos em um ramo bem divergente, mas com baixo suporte, dentro de um clado envolvendo haplótipos de *P. machrisii*. A árvore Bayesiana, por sua vez, agrupou os haplótipos de *P. jauruensis* em um clado com alto suporte juntamente com haplótipos de *P. aurisetus* (Figura 4). Além disso, na topologia Bayesiana os haplótipos H2 e H3 de *P. aurisetus* definiram um clado, enquanto na árvore de ML essa relação apresentou um baixo suporte.

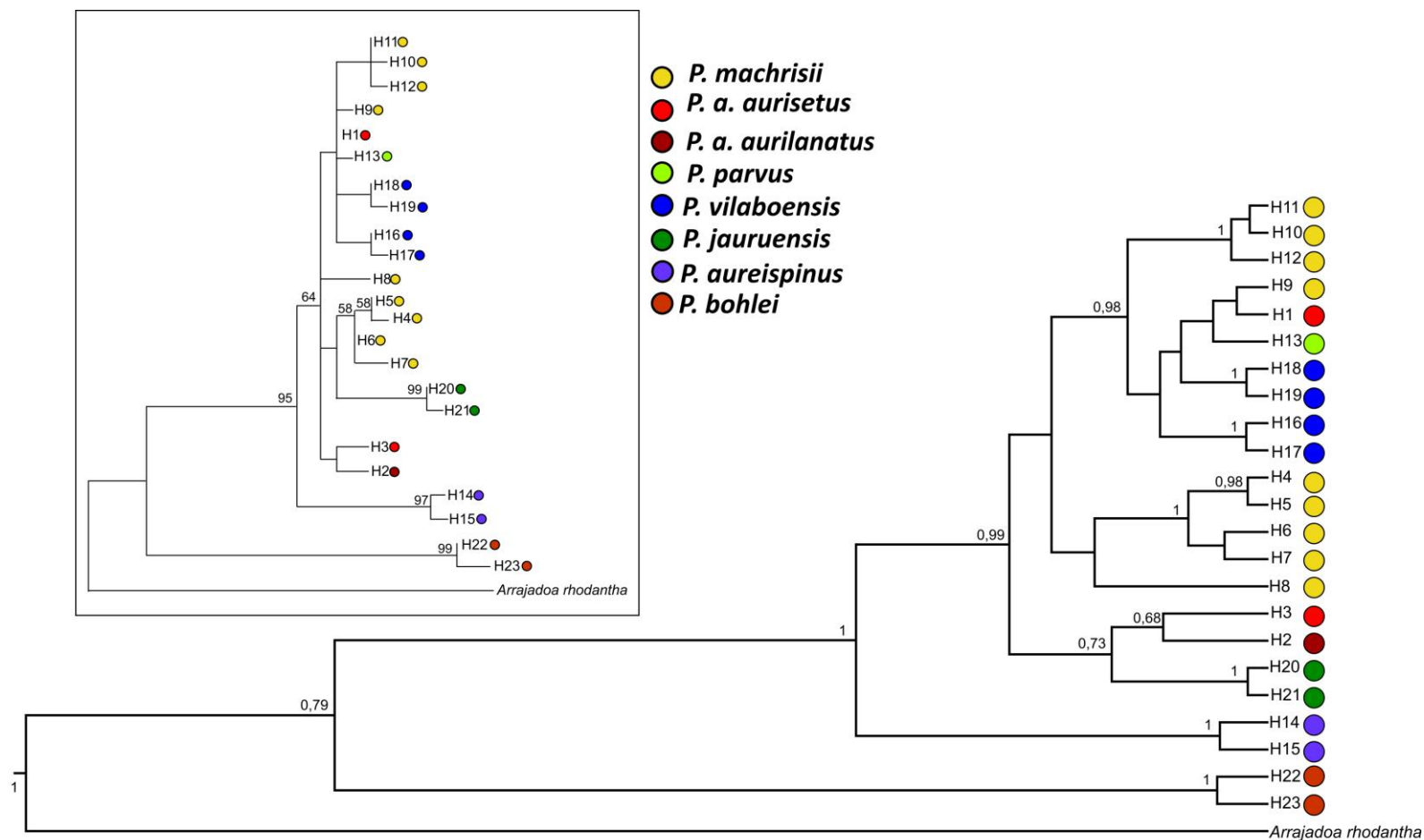


Figura 4. Relações filogenéticas entre as espécies do grupo *P. AURISSETUS* inferidas a partir de uma análise Bayesiana e Máxima Verossimilhança (*insert box*) utilizando sequências do cpDNA. A espécie de Cactaceae *Arrajadoa rhodantha* foi utilizada como grupo externo. Os índices dos nós representam os valores de probabilidade posterior na árvore Bayesiana e os valores de *bootstrap* (>50%) na árvore de Máxima Verossimilhança.

4.3. Relação entre os haplótipos e Inferências Filogeográficas

A relação entre os haplótipos revelou que o haplótipo H3 foi o mais frequente no grupo, sendo o único haplótipo encontrado em nove populações de *P. aurisetus* da região da Cadeia do Espinhaço em Minas Gerais (ODA, MEN, PJK, CMO, ITA, COC, EDB2, TAT1 e INA2). O haplótipo mais interno (H1), o qual foi considerado o mais antigo do grupo, foi encontrado na população GMII de *P. aurisetus*. Este haplótipo está relacionado a outros onze haplótipos do grupo com um número máximo de quatro passos mutacionais, com exceção de *P. aureispinus* e *P. bohlei* que estão separados por dez e 26 mutações distintas, respectivamente (Figura 5).

A maior parte das populações analisadas apresentou haplótipos únicos e exclusivos. Apenas as populações de ART (*P. machrisii*), PIR e GOV (*P. vilaboensis*), IBO (*P. aureispinus*) e UMB (*P. bohlei*) apresentaram dois haplótipos. Compartilhamento de haplótipos foi observado entre algumas populações de *P. machrisii*: BRO, DEL e FOR (H6), CRI e FMS (H4), URU e APA2 (H9) POS e CAM (H10) e entre as populações AQU, RVE e COX (H20) de *P. jauruensis*. Não foi observado compartilhamento de haplótipos entre espécies. A distância geográfica das populações que compartilharam haplótipos variou de 101 km a 32 km.

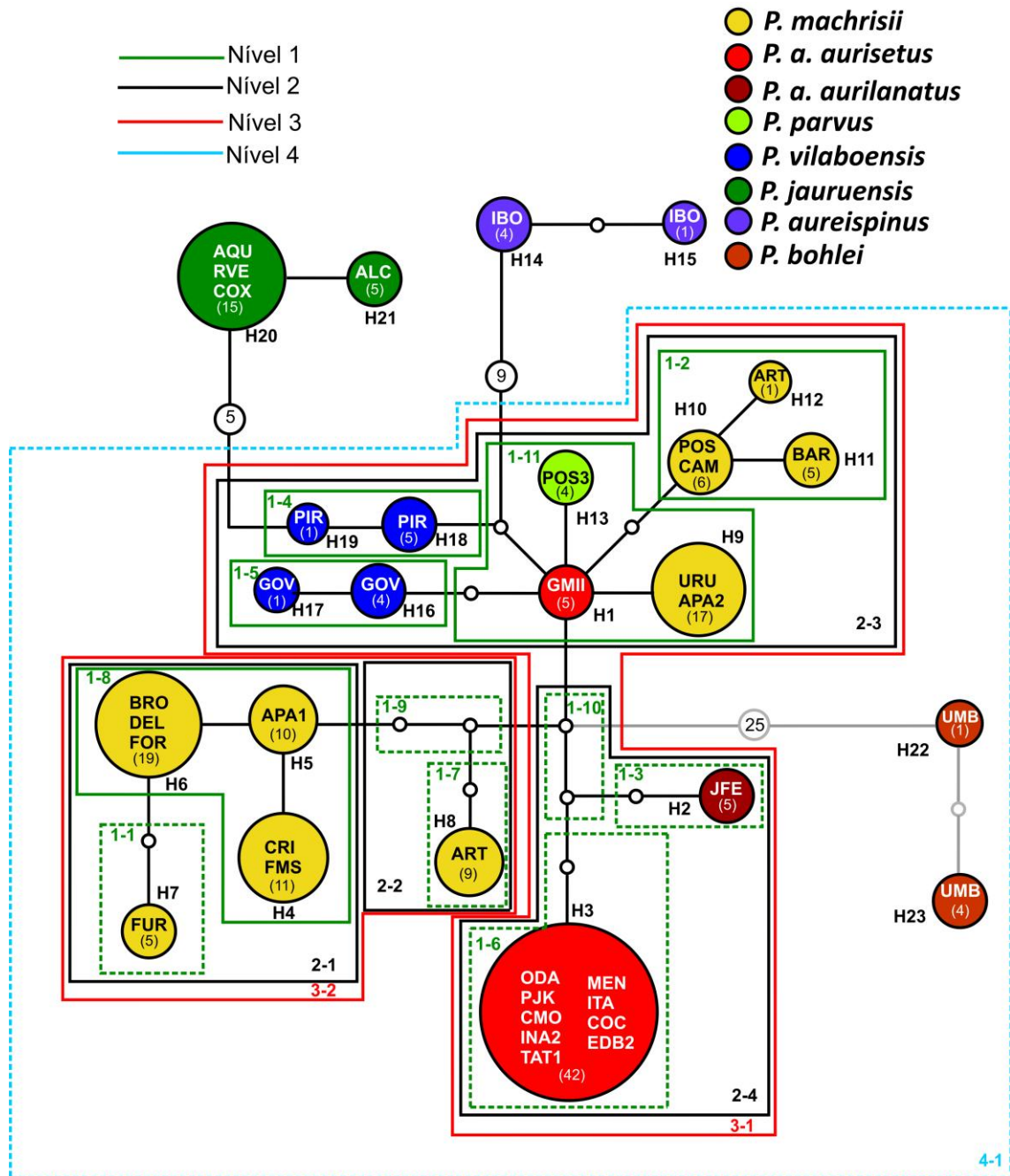


Figura 5. Rede de haplótipos elaborada a partir do segmento concatenado das regiões *tnT-trnL*, *3'trnS-trnG* e *5'trnS-trnG* (1475pb). O tamanho dos círculos é proporcional à frequência dos haplótipos na amostra. Os números entre parênteses indicam o número de indivíduos portadores do haplótipo. Os pequenos círculos vazios indicam haplótipos não observados e o número dentro deles a quantidade desses haplótipos. A linha entre os haplótipos representa um único passo mutacional. As cores representam as diferentes espécies, conforme legenda ao lado. Os ramos em cinza representam relações fora do limite de 95% de parcimônia. A figura também mostra os quatro níveis hierárquicos dos clados definidos na análise NCPA. Linhas pontilhadas indicam clados com resultados não significativos segundo a NCPA. Os códigos de localidades são os mesmos apresentados na Tabela 1.

Os haplótipos de *P. machrisii* foram separados em dois clados que concordaram parcialmente com a distribuição geográfica da espécie. Em um clado foram agrupadas a população da Bahia (BAR) e as populações do norte de Goiás e sul do Tocantins. O outro clado foi formado por populações do leste de Goiás, sul de Minas Gerais e São Paulo. Dessa forma definiu-se um haplogrupo norte (POS, CAM, BAR, URU, APA2 e ART) e um haplogrupo centro-sul (CRI, FMS, APA1, FUR, BRO, DEL, FOR e ART) com sobreposição na região de Aurora do Tocantins e Alto Paraíso de Goiás. O haplótipo mais frequente na população ART (H8) foi também o mais divergente dentro da espécie *P. machrisii* e se mostrou mais relacionado aos haplótipos de distribuição sul (Figura 5).

Os haplótipos com maior divergência em relação ao conjunto de haplótipos foram encontrados em *P. jauruensis* (H20 e H21), *P. aureispinus* (H14 e H15) e *P. bohlei* (H22 e H23). Apesar da grande distância genética, os haplótipos de *P. jauruensis* se mostraram mais relacionados aos haplótipos H18 e H19 de *P. vilaboensis* (seis passos mutacionais de diferença) do que aos demais haplótipos do grupo, sugerindo que populações de *P. jauruensis* se originaram a partir de populações de *P. vilaboensis*.

Os três haplótipos encontrados para *P. aurisetus*, estão ligados por quatro passos mutacionais. Dentre eles foi encontrado um haplótipo exclusivo (H2) da subespécie *P. aurisetus aurilanatus* (JFE). O grande compartilhamento de haplótipos entre as populações da Cadeia do Espinhaço envolveu populações definidas como pertencentes a subespécie *P. aurisetus aurisetus*. Dentre elas estão as populações ITA e EDB2, as quais já foram classificadas como espécies diferentes: *P. werdermannianus densilanatus* (RITTER, 1979) e *P. supthutianus* (BRAUN; ESTEVES PEREIRA, 1985), respectivamente. Atualmente a população ITA continua sendo reconhecida como uma forma diferenciada de *P. aurisetus* (subespécie informal *P. a. densilanatus*).

Foram encontrados dois haplótipos para cada população de *P. vilaboensis* (GOV=H16 e H17, PIR=H18 e H19), os quais não se relacionaram diretamente. Esses haplótipos se mostraram mais próximos a haplótipos de *P. machrisii* (H9) e *P. parvus* (H13) do que entre si.

A rede de haplótipos obtida, excluindo-se os haplótipos mais divergentes (H14, H15, H20, H21, H22 e H23) foi utilizada para a realização da análise NCPA (Figura 5). Nove clados

definidos a partir da rede de haplótipos mostraram associação geográfica significativa conforme apontado pelos valores de qui-quadrado, Dc e Dn (APÊNDICE 1). No entanto dois desses clados apresentaram resultados inconclusivos (Tabela 4).

Para os clados 1-2 e 1-8, os quais são formados por haplótipos de populações de *P. machrisii* que ocorrem ao longo da maior parte da distribuição da espécie, foram detectados três possíveis eventos para explicar o padrão filogeográfico observado: expansão contígua, colonização a longa distância ou fragmentação passada. A inferência sobre o evento de expansão está relacionada aos valores de Dc significativamente pequenos nos haplótipos de interior que compõem esses clados. De acordo com as relações genealógicas dos haplótipos e sua distribuição geográfica é possível inferir a direção dos eventos de expansão, os quais provavelmente se originaram a partir de populações do norte e leste do estado de Goiás (Planalto Central) em direção ao oeste da Bahia e ao sul, na divisa entre Minas Gerais e São Paulo (Figura 6).

Os eventos de colonização a longa distância e fragmentação passada foram inferidos devido ao fato dos haplótipos 5 (clado 1-8) e 11 (clado 1-2) apresentarem valores significativamente altos de Dn e estarem separados dos demais haplótipos por áreas que não foram amostradas.

Tabela 4. Inferências filogeográficas dos clados com associação geográfica significativa na análise NCPA

Clado	Percurso na Chave de inferências (TEMPLETON et al. 1995, 2004) para identificação de eventos demográficos
1-2	1-19-20-2-11 (Sim): Expansão – 12-13-14 (Sim): Amostragem inadequada para distinguir entre Expansão Contígua, Colonização a longa distância ou Fragmentação passada
1-8	1-19-20-2-11 (Sim): Expansão – 12-13-14 (Sim): Amostragem inadequada para distinguir entre Expansão Contígua, Colonização a longa distância ou Fragmentação passada
1-11	1-19-20-2-9-11(Sim): Expansão – 12-13-21 (Não): Amostragem inadequada para distinguir entre colonização a longa distância e expansão gradual passada seguida de fragmentação ou ampla distribuição seguida de extinção em áreas intermediárias.
2-1	1-19-20 (Não): Inconclusivo devido a amostragem inadequada
2-3	1-2-3-5-6-7 (Sim): Fluxo gênico restrito com alguma dispersão a longa distância
2-4	1-19 (Não): Fragmentação alopátrica
3-1	1-19-20 (Não): Inconclusivo. Amostragem inadequada
3-2	1-19-20-2-11 (Sim): Expansão – 12-13 (Sim): Colonização a longa distância, Expansão passada, ampla distribuição seguida de extinção em áreas intermediárias ou fragmentação seguida de expansão - 21 (Não): Expansão gradual passada seguida de fragmentação ou ampla distribuição seguida de extinção em áreas intermediárias.
4-1	1-2: Inconclusivo. I-T não pode ser determinado.

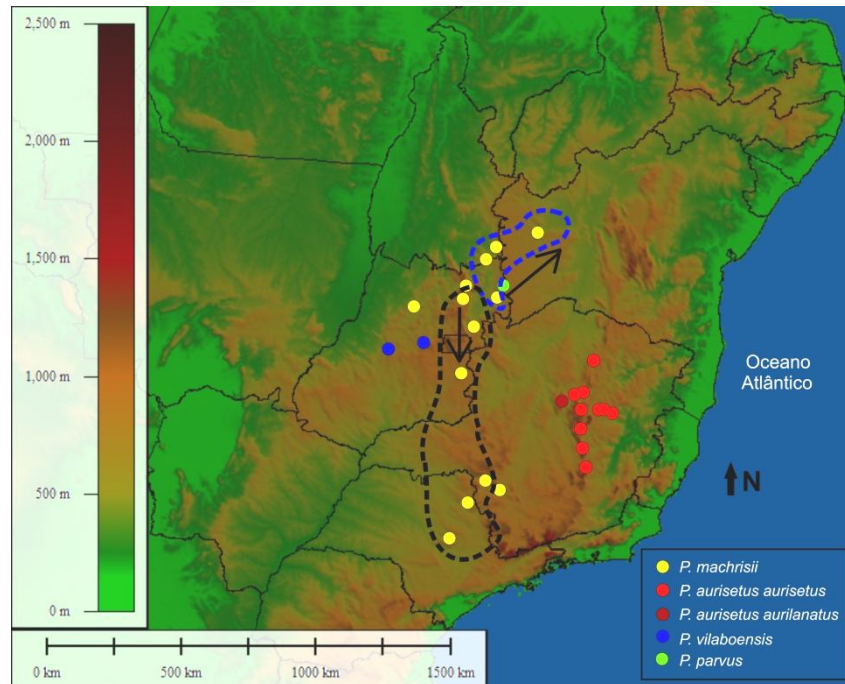


Figura 6. Mapa do leste do Brasil mostrando a distribuição das populações contidas nos clados 1-2 (azul) e 1-8 (preto) e os possíveis eventos de expansão detectados pela NCPA. As setas indicam a direção das expansões.

Eventos de colonização a longa distância e expansão gradual passada seguida de fragmentação ou ampla distribuição seguida de extinção em áreas intermediárias foram propostos para o clado 1-11 (Figura 7), composto por haplótipos de *P. machrisii* e *P. parvus* do norte de Goiás e *P. aurisetus* de Grão Mogol – MG. Esse resultado foi inferido devido aos baixos valores de D_c para os haplótipos de extremidade e aos valores de D_n não significativos para o haplótipo de interior, juntamente com a ocorrência de pouca sobreposição geográfica dentro do clado. Os eventos de fragmentação e de isolamento por distância não puderam ser distinguidos, pois foi considerada a possibilidade da ocorrência das espécies em áreas intermediárias não amostradas.

Os eventos de expansão foram propostos devido a existência de haplótipos internos com valores de D_c significativamente pequenos nesse clado. A expansão inferida para o clado 1.11 concorda com o possível evento de expansão que deu origem a todas as linhagens do grupo *P. AURISSETUS*, uma vez que esse clado abriga o haplótipo mais interno de toda genealogia (H1). Dessa forma, os eventos de expansão provavelmente ocorreram a partir de populações do norte do estado de Minas Gerais. Além disso, o haplótipo I

apresentou altos valores de D_n e está separado dos demais haplótipos relacionados por áreas que foram amostradas, mas que apresentaram haplótipos aninhados em outros clados. Esse resultado sugere a ocorrência de eventos de colonização a longa distância, extinção em áreas intermediárias de uma distribuição mais ampla no passado ou fragmentação seguida de expansão.

O resultado para o clado 2-1 foi inconclusivo devido à amostragem inadequada, uma vez que a espécie *P. machrisii* pode estar presente em áreas intermediárias não amostradas.

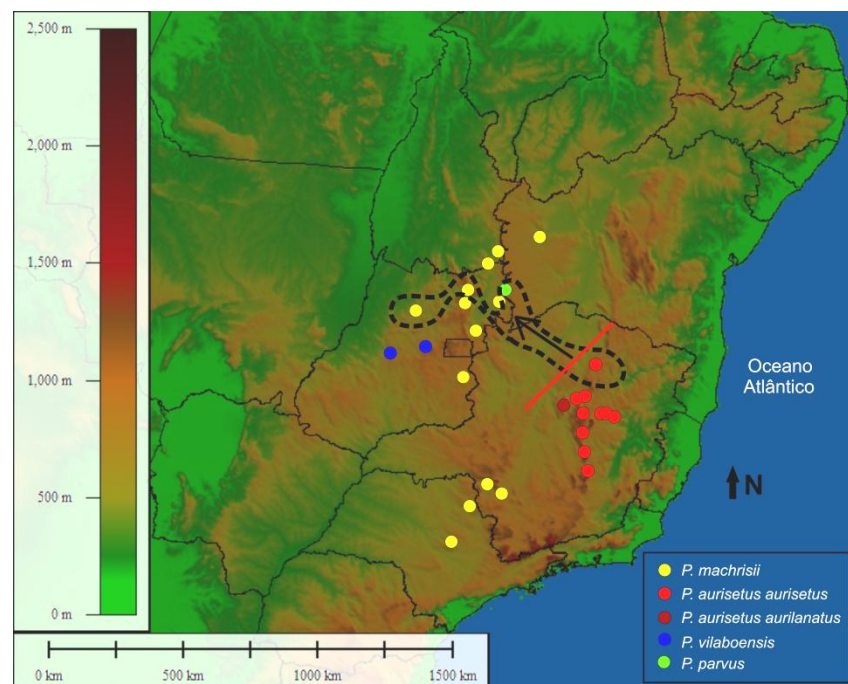


Figura 7. Mapa do leste do Brasil mostrando a distribuição das populações contidas no clado 1-11 e os possíveis eventos de expansão e fragmentação detectados pela NCPA. As setas indicam a direção das expansões. Linhas vermelhas representam possíveis eventos de fragmentação detectados pela NCPA.

O clado 2-3 (composto por haplótipos de populações de *P. machrisii*, *P. aurisetus*, *P. vilaboensis* e *P. parvus*), apresentou padrões possivelmente originados por fluxo gênico restrito, acompanhado de alguma dispersão a longa distância (Figura 8). Esse resultado é decorrente dos valores significativamente altos de D_n para o clado 1-5 que está separado dos demais clados relacionados por áreas amostradas, mas que não apresentaram haplótipos no clado 2-3.

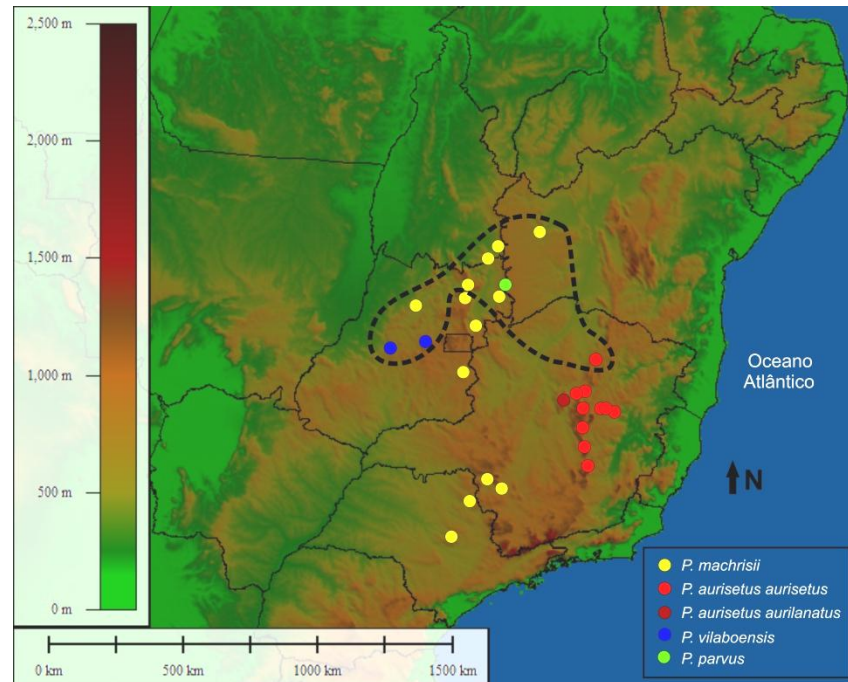


Figura 8. Mapa do leste do Brasil mostrando a distribuição das populações contidas no clado 2-3.

Para o clado 2-4, composto por haplótipos de *P. aurisetus* (Figura 9), foi proposto um padrão de fragmentação alopátrica evidenciado pela distância de quatro mutações entre os clados 1-3 (*P. aurisetus aurilanatus*) e 1-6 (*P. aurisetus aurisetus*). Apesar do resultado para o clado 3-1 ser inconclusivo devido à amostragem inadequada, a distância genética entre os haplótipos do clado 2-3 e 2-4 (quatro mutações distintas) também sugere a ocorrência de fragmentação alopátrica.

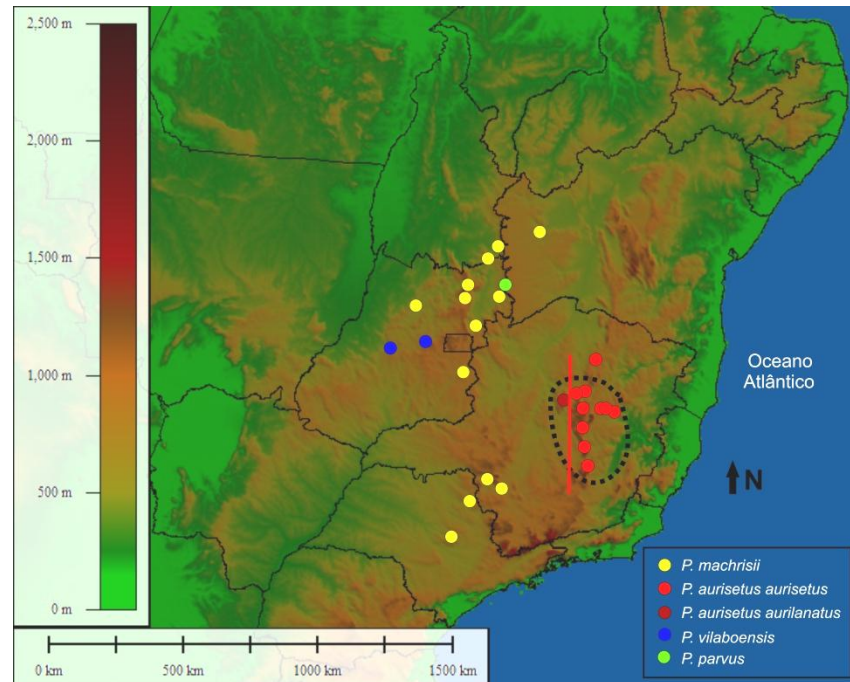


Figura 9. Mapa do leste do Brasil mostrando a distribuição das populações contidas no clado 2-4. Linhas vermelhas representam possíveis eventos de fragmentação detectados pela NCPA.

Para o clado 3-2, composto exclusivamente por populações de *P. machrisii* (Figura 10), foram propostos eventos de expansão gradual passada, extinção em áreas intermediárias de uma distribuição mais ampla no passado ou fragmentação seguida de expansão. Esse resultado foi observado devido à existência de haplótipos internos no clado 3-2 com valores de D_c significativamente pequenos. Além disso, o clado 2-2 apresentou altos valores de D_n e está separado dos demais haplótipos relacionados por áreas que foram amostradas, mas que não apresentaram haplótipos no clado 3-2. A expansão detectada para o clado 3-2, está relacionada a expansão a partir de populações do Planalto Central observada para o clado 1-8, uma vez que o clado 3-2 aloca o clado 1-8 e outros dois haplótipos de *P. machrisii*.

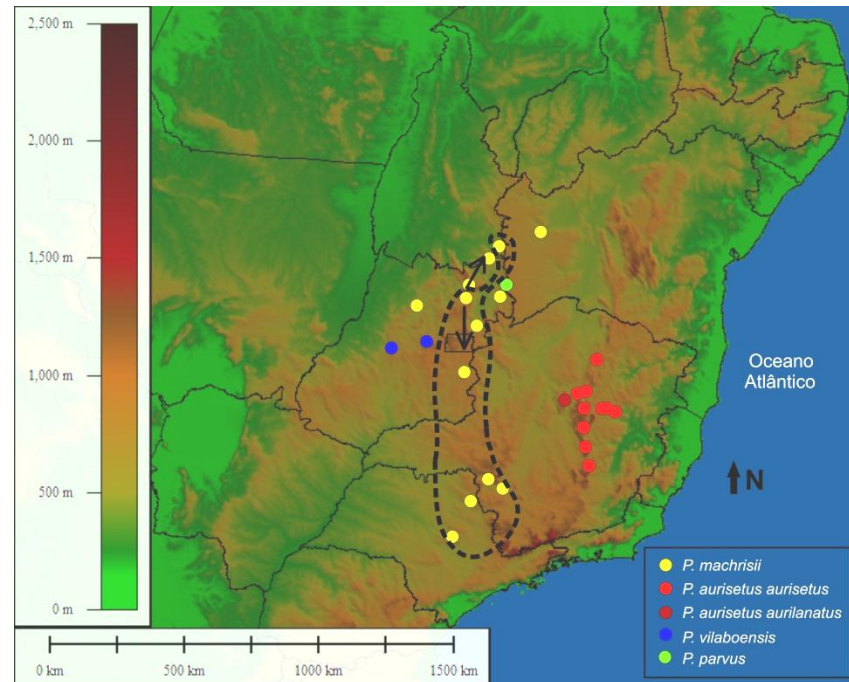


Figura 10. Mapa do leste do Brasil mostrando a distribuição das populações contidas no clado 3-2 e os possíveis eventos de expansão detectados pela NCPA. As setas indicam a direção das expansões.

4.4. *Mismatch distribution* e testes de neutralidade

O padrão de distribuição observado na análise de *Mismatch distribution* para as espécies do grupo *P. AURISSETUS* com exceção das mais divergentes (*P. aureispinus*, *P. bohlei* e *P. jauruensis*) foi representado por uma curva unimodal (Figura 11A). Embora esse resultado seja esperado quando as populações experimentaram um evento de expansão, foi verificado um desvio significativo entre a curva observada e a curva simulada para uma população que sofreu expansão espacial ($SSD = 0,039$; $P_{SSD} = 0,043$). Quando essa mesma análise foi realizada com um número igual de indivíduos (cinco) para as populações, com exceção das populações de CAM, INA e POS3 as quais possuíam menos amostras disponíveis, a curva observada não desviou significativamente da curva esperada sob um cenário de expansão ($SSD = 0,033$; $P_{SSD} = 0,114$). Este segundo resultado sugere que a falha na detecção de expansão populacional sobre o conjunto de haplótipos mais relacionados do grupo *P. AURISSETUS* pode ter sido simplesmente causada pela inclusão de tamanhos amostrais muito diferentes na análise.

Sinais de expansão também foram identificados para os clados 1-8 ($SSD = 0,048$; $P_{SSD} = 0,162$, Figura 11C) e 1-2 ($SSD = 0,041$; $P_{SSD} = 0,034$; Figura 11B). No entanto, para o clado 1-2 os dados observados desviaram significativamente do modelo de expansão. Esse resultado provavelmente está relacionado ao baixo poder estatístico da análise, decorrente da baixa variação observada entre os haplótipos envolvidos.

Apesar dos resultados para o clado 3-2 não terem desviado significativamente do modelo de expansão ($SSD = 0,020$; $P_{SSD} = 0,275$), a curva obtida apresentou um padrão multimodal, contrariando a hipótese de expansão. No entanto, esse resultado pode ser encontrado em populações que expandiram há um espaço de tempo suficientemente grande que permitiu a diversificação de linhagens após a expansão. Um resultado que apoia esta possibilidade foi a inferência de expansão passada para esse clado na análise NCPA.

Os resultados para os testes de neutralidade D (TAJIMA, 1989) e F_s (FU, 1997), por sua vez, não foram significativos para nenhum dos agrupamentos testados.

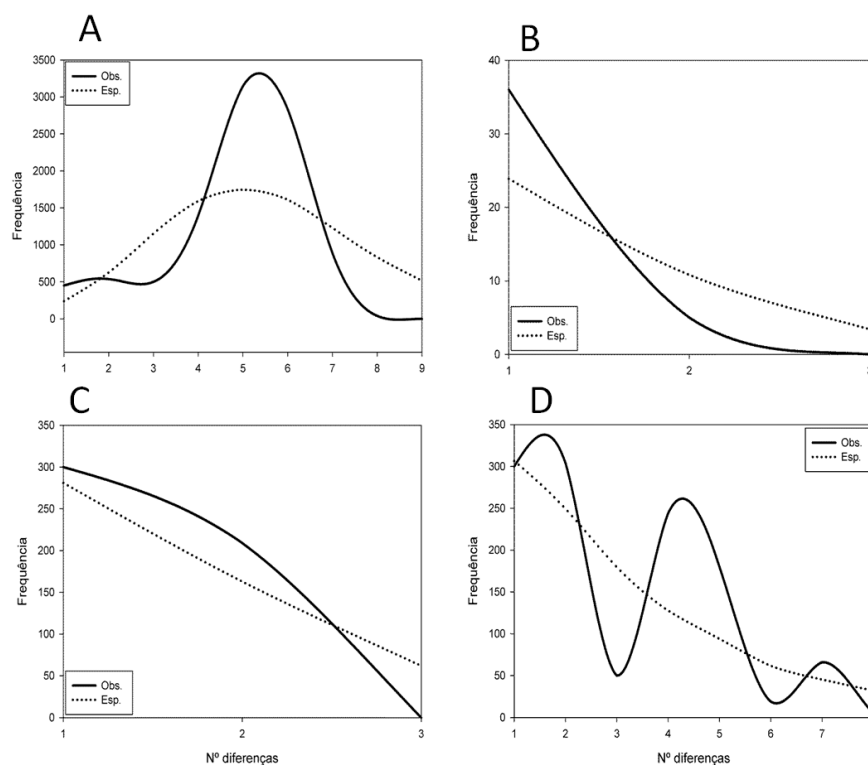


Figura 11. Curvas geradas pela análise *Mismatch distribution* da distribuição do número de diferenças entre pares de sequências para (A) Grupo *P. AURISSETUS* excluindo as espécies mais divergentes e os clados (B) 1-2, (C) 1-8 e (D) 3-2 definidos na NCPA. Linhas contínuas representam a distribuição das diferenças observadas enquanto linhas pontilhadas representam a distribuição esperada segundo o modelo de expansão espacial.

4.5. Estruturação populacional

A análise de SAMOVA foi realizada para identificar os níveis de estruturação genética associada à distribuição geográfica das populações. Essa análise identificou claramente que um aumento no valor de k resultou em um aumento no valor de F_{CT} , o qual se estabilizou a partir de um valor de $k=20$ (Figura 12).

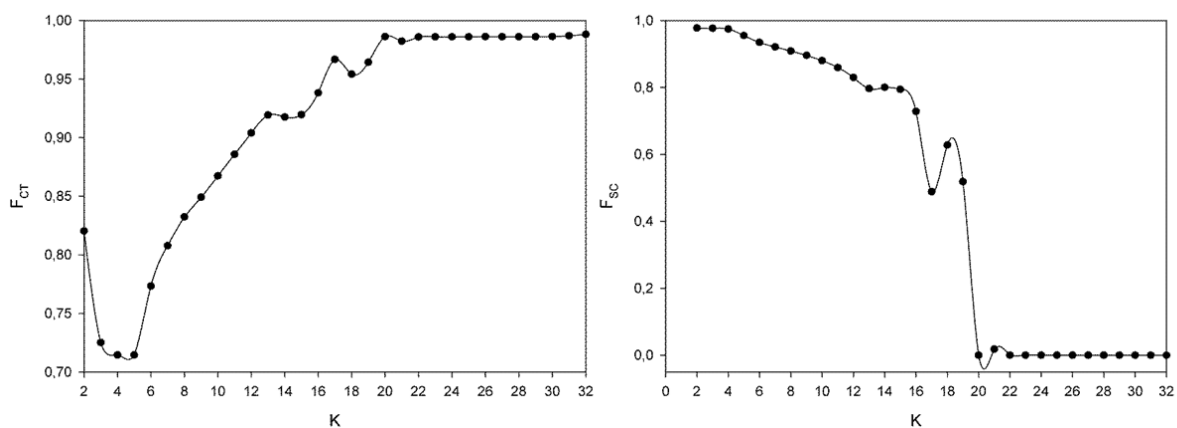


Figura 12. Representação gráfica dos valores de F_{CT} e F_{SC} para os diferentes números de agrupamentos (k) definidos na análise SAMOVA

O número de grupos que melhor explicou a estruturação das populações foi 11, valor no qual a diferenciação entre grupos superou a diferenciação dentro dos grupos. Segundo a configuração proposta por esses agrupamentos a espécie *P. machrisii* foi separada em um grupo para a distribuição sul: (BRO DEL FOR e FUR), um grupo central: (CRI FMS APA1) e dois grupos para a distribuição norte: (ART) (POS CAM URU BAR APA2 GMII POS3) sendo que este último agrupamento incluiu a população GMII de *P. aurisetus* e a população POS3 de *P. parvus* (Figura 13). As demais populações de *P. aurisetus* formaram dois grupos: (ODA MEN PJK CMO ITA COC EDB2 TAT1 INA2) e (JFE) localizados a leste e a oeste da Cadeia do Espinhaço, respectivamente. A espécie *P. jauruensis* formou um único grupo (AQU RVE COX e ALC). As duas populações de *P. vilaboensis* foram definidas em grupos isolados: (PIR) e (GOV), bem como as espécies *P. aureispinus* (IBO) e *P. bohlei* (UMB).

Este resultado mostra uma grande estruturação das espécies a nível populacional, uma vez que a condição de estabilidade dos valores de F_{CT} somente foi atingida com a formação de um número elevado de grupos (APÊNDICE 2), ou seja, os valores de F_{CT} somente

se estabilizaram quando os grupos foram formados por um menor número de populações ou por populações isoladas. Além disso, apesar da diferenciação observada entre os 11 grupos identificados ser grande, a estruturação dentro dos grupos permaneceu elevada ($F_{SC}=0.859$). Todos os valores F_{CT} e F_{SC} observados foram significativos ($P<0,01$).

Para a espécie *P. machrisii* a AMOVA envolvendo os dois clados identificados na análise filogenética resultou em uma divergência entre grupos de 73,37% e dentro de grupos de 23,63%. Para os haplogrupos norte e centro-sul identificados na rede de haplótipos, a divergência entre grupos foi menor, 61,40% e, conseqüentemente, maior dentro dos grupos 38,60% do que a observada segundo os agrupamentos da árvore filogenética. Esse resultado é decorrente da existência de haplótipos divergentes nas populações ART (haplogrupo norte e centro-sul), APA2 e URU (haplogrupo norte) que não foram agrupados nos clados filogenéticos por não apresentarem relação filogenética bem suportada com os demais haplótipos do grupo. Apesar disso, esses agrupamentos demonstraram uma clara estruturação regional de *P. machrisii* em dois grupos.

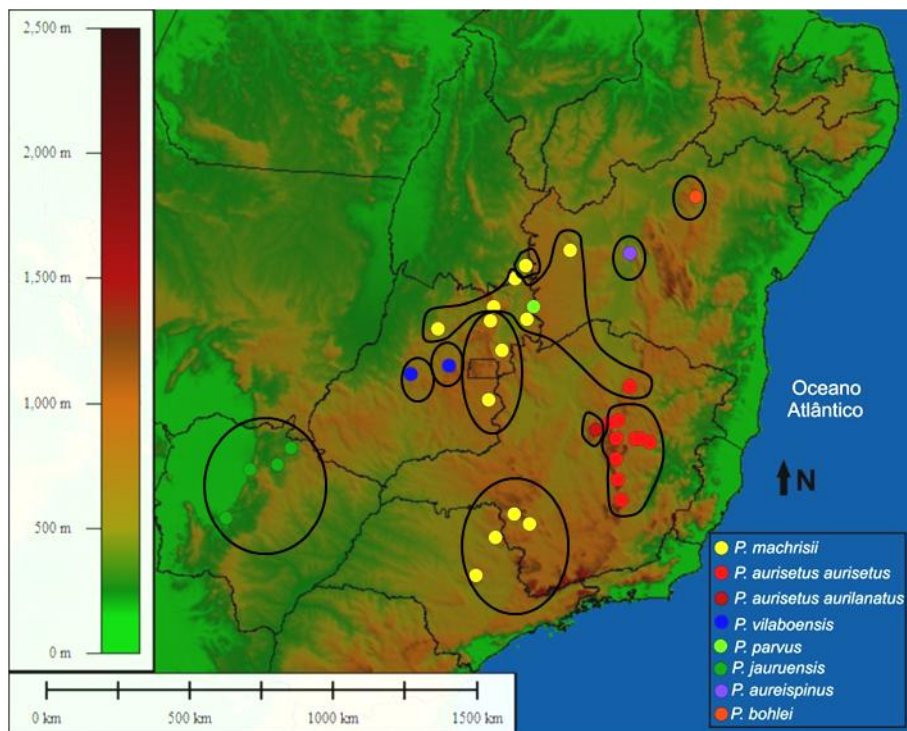


Figura 13. Mapa do leste do Brasil mostrando os agrupamentos gerados na análise SAMOVA.

5. DISCUSSÃO

5.1. Variabilidade do cpDNA e Análises Filogenéticas

A maior parte da variação intraespecífica observada no presente estudo se refere a diferenças entre populações, pois um pequeno número de populações apresentou variabilidade intrapopulacional (ART, GOV, PIR, IBO e UMB). Algumas espécies como *P. aurisetus*, *P. jauruensis* e *P. parvus* não apresentaram variação dentro de populações. Esse nível de variação em marcadores plastidiais é comumente observado em espécies neotropicais com distribuição descontínua como bromélias, cactos e *Petunia* (Solanaceae) (BARBARÁ et al., 2008; HELSEN et al., 2009; LORENZ-LEMKE et al., 2010). Espécies com distribuição bastante fragmentada comumente apresentam baixos níveis de variação genética dentro de populações e grande diferenciação populacional (YOUNG et al., 1996; JUMP; PEÑUELAS, 2006; ARTYUOKOVA et al., 2009), provavelmente um efeito do tamanho populacional reduzido e do limitado movimento de sementes entre populações. Nesse sentido, as espécies de *Pilosocereus* aqui estudadas se mostraram muito estruturadas no nível populacional, apresentando, na maior parte dos casos, um único haplótipo por população.

Os valores de diversidade nucleotídica e haplotípica encontrados foram equivalentes aos observados para regiões do cpDNA (ARTYUOKOVA et al., 2009; NOVAES et al., 2010) para espécies com distribuição bastante fragmentada como *Microbiota decussata* (Cupressaceae), endêmica das montanhas Sikhote-Alin da Rússia e espécies com distribuição mais ampla associadas a áreas secas como *Plathymentia reticulata*, de ocorrência no Cerrado e na Mata Atlântica (NOVAES et al., 2010). Por outro lado, os valores são inferiores aos encontrados para espécies com distribuição descontínua no Cerrado como *Hymenaea stigonocarpa* (RAMOS et al., 2007) e *Lychnophora ericoides* (COLLEVATTI et al., 2009).

Apesar da grande estruturação populacional do cpDNA ser uma observação comum em estudos filogeográficos (BARBARÁ et al., 2008; RAMOS et al., 2009; LORENZ-LEMKE et al., 2010), a estruturação encontrada nas espécies do grupo *P. AURISSETUS* parece estar bastante relacionada ao padrão de distribuição e a história de diversificação recente do grupo.

Algumas observações como a ocorrência de várias populações com baixa variabilidade, elevada divergência entre haplótipos da mesma espécie e a similaridade de haplótipos de diferentes espécies corroboram essa afirmação. Além disso, as regiões do cpDNA utilizadas nesse estudo estão entre as mais utilizadas em estudos filogeográficos de angiospermas e foram as regiões mais variáveis entre oito regiões analisadas até o momento para o grupo de espécies estudado (BONATELLI et al., 2012).

No grupo *P. AURISSETUS* a ausência de monofilia para a maior parte das espécies deve estar relacionada à diversificação recente no grupo a qual ainda não permitiu que as espécies atingissem a condição de monofilia recíproca. No entanto os clados observados concordam parcialmente com a distribuição geográfica das populações, mostrando que alguns sinais genéticos dessa história de diversificação recente já foram incorporados à variação observada. Dois clados envolvendo apenas populações de *P. machrisii* foram bem suportados, sugerindo um maior tempo de isolamento entre essas populações. O primeiro clado foi formado por quatro populações da distribuição norte da espécie (POS, CAM, ART e BAR) e o segundo por sete populações da distribuição centro-sul (APA1, CRI, FMS, BRO, DEL, FOR e FUR). Devido a maior relação evolutiva observada entre as populações do leste de Goiás e as populações do sul de Minas Gerais e norte de São Paulo é possível inferir que eventos de fragmentação isolaram primeiramente as populações do leste e norte de Goiás e em seguida, as populações do leste de Goiás e do sudeste do país.

Embora várias populações de *P. machrisii* tenham sido descritas como espécies diferentes e linhagens haplotípicas divergentes tenham sido observadas no Planalto Central, a distribuição e as relações genealógicas entre os haplótipos encontrados não suportaram a existência de táxons distintos de *P. machrisii* nessa região. Por exemplo, embora as populações de Alto Paraíso de Goiás e Cristalina correspondam às localidades de ocorrência de duas espécies sinonimizadas com *P. machrisii* (*P. lindaianus* e *P. cristalinensis*), os haplótipos dessas populações (H4 e H5) foram bastante similares e agrupados em um clado bem suportado nas análises filogenéticas. Em Alto Paraíso de Goiás, onde as populações foram descritas como *P. lindaianus* (BRAUN; ESTEVES PEREIRA, 1987b, 1988), a estreita relação entre os haplótipos dessa região com haplótipos de outras localidades de *P. machrisii* (APA1 com CRI e APA2 com URU) não suportou a possibilidade da existência de um táxon distinto de *P. machrisii* na localidade. Outro exemplo é encontrado na população de *P.*

machrisii de Cristalina, localidade que corresponde a ocorrência da espécie anteriormente descrita como *P. cristalinensis* (BRAUN; ESTEVES PEREIRA, 1987a). No entanto, a população CRI compartilha seu haplótipo com FMS, localidade na qual a espécie *P. machrisii* já foi descrita como *P. goianus* (BRAUN; ESTEVES PEREIRA, 2002). A ocorrência de *P. goianus* também foi descrita para a região de Posse e Campos Belos. As populações dessas localidades analisadas no presente estudo apresentaram o mesmo haplótipo alocado no clado norte de *P. machrisii*, o qual se mostrou bastante relacionado aos demais haplótipos do clado. Dessa forma, apesar das populações de *P. machrisii* definirem dois clados, a história da instabilidade taxonômica da espécie não pode ser evocada para explicar o padrão observado. No entanto, deve-se levar em conta que se os eventos de especiação nessa região forem muito recentes ou se ocorreu hibridização e introgressão entre esses supostos táxons, a distribuição e as relações genealógicas entre os haplótipos nessas populações podem não revelar uma possível diversidade taxonômica. Adicionalmente, pouco se sabe sobre o potencial de plasticidade fenotípica dessas espécies, o que pode ter influenciado a classificação taxonômica das populações.

Dois haplótipos de *P. aurisetus* foram agrupados em um clado bem suportado pela árvore Bayesiana, os quais corresponderam as subespécies *P. a. aurisetus* e *P. a. aurilanatus*. Enquanto *P. a. aurisetus* ocorre ao longo da porção mineira da Cadeia do Espinhaço, *P. a. aurilanatus* ocupa a Serra do Cabral, um maciço relativamente pequeno e isolado no lado oeste da Cadeia do Espinhaço. Algumas características morfológicas que corroboram a diferenciação entre essas populações são o caule mais alto e robusto e a abundância de cerdas nas auréolas observadas em *P. a. aurilanatus*. Essa concordância entre a variação nos níveis molecular e morfológico aponta para uma fragmentação relativamente antiga entre as populações dos dois táxons, que desde então permaneceram isolados, apesar da pequena distância geográfica que os separam.

De acordo com a análise filogenética Bayesiana, os haplótipos das espécies *P. aurisetus* e *P. jauruensis* apresentaram uma relação mais estreita entre si do que com os demais haplótipos do grupo. Esse resultado representa uma discordância em relação ao observado na rede de haplótipos. Apesar dos haplótipos de *P. jauruensis* se apresentarem bastante divergentes aos demais do grupo, na rede de haplótipos eles se mostraram mais relacionados a haplótipos de *P. vilaboensis*. O resultado discrepante observado na árvore

filogenética deve ser resultado da divergência encontrada para os haplótipos de *P. aurisetus* e *P. jauruensis* em relação aos demais haplótipos do grupo, gerando uma tendência no algoritmo Bayesiano de agrupá-los. Dessa forma, esses haplótipos definiram um clado mais externo na topologia da análise Bayesiana, depois dos haplótipos de *P. bohlei* e *P. aureispinus*.

As três espécies apontadas pela filogenia e pela genealogia como as mais divergentes do grupo apresentam algumas características morfológicas peculiares. *Pilosocereus aureispinus*, por exemplo, apresenta flores com coloração castanho escura ou rosada, pequenas sementes alongadas com células da testa cônicas e com a região do hilo e micrópila estreita. *Pilosocereus jauruensis* pode ser diferenciado das demais espécies pelo caule pequeno e flores com forma mais alongada. *Pilosocereus bohlei*, por sua vez, apresenta diferença morfológica principalmente em relação aos frutos, os quais exibem coloração verde ou verde-azulada.

Devido a grande divergência filogenética e genealógica observada para *P. aureispinus* e *P. bohlei*, possivelmente essas espécies se isolaram há mais tempo ou possuem origens independentes, não participando da diversificação das demais espécies do grupo. Dessa forma, a presença dessas espécies no grupo *P. AURISSETUS* é suportada apenas por características morfológicas compartilhadas definidas pela sua taxonomia.

Pilosocereus jauruensis, por sua vez, apesar de ser uma espécie distantemente relacionada às demais, apresenta maior relação com *P. vilaboensis* segundo os dados de genealogia dos haplótipos. De acordo com a distribuição geográfica dessas espécies, as populações de *P. jauruensis* podem ter se originado da diferenciação de populações de *P. vilaboensis* que se expandiram a partir do Planalto Central em direção a Serra de Maracajú, no Mato Grosso do Sul. Dessa forma, a origem da espécie *P. jauruensis* estaria associada à diversificação do grupo *P. AURISSETUS*. Essa hipótese é suportada pela análise filogeográfica no drosófilideo cactofílico *D. gouveai* (MORAES et al., 2009). Esse estudo apontou para a possibilidade das populações dessa espécie na Serra de Maracajú terem se originado a partir da expansão de populações no Planalto Central, incluindo as localidades de ocorrência de *P. vilaboensis*.

5.2. Estruturação genética

Os 11 grupos propostos na análise SAMOVA concordam com a distribuição dos haplótipos na genealogia e apontam para uma subestruturação nos haplogrupos de *P. machrisii*. No geral, os grupos foram formados sem sobreposição de espécies, com exceção do grupo norte de *P. machrisii* que incluiu a população GMII de *P. aurisetus* e a população POS3 de *P. parvus*. Essa similaridade genética observada entre a população de GMII e as populações do clado norte de *P. machrisii* pode ser explicada por um evento de fragmentação recente dessas populações. Outro aspecto que chama a atenção nessa relação é o fato da população GMII (Grão Mogol-MG) ter sido classificada como *P. a. supthutianus* pelos taxonomistas do grupo (N.P. Taylor e M.C. Machado, com. pes.). Esse táxon foi descrito originalmente para a localidade de Engenheiro Dolabela-MG, na parte oeste da Cadeia do Espinhaço, o qual foi posteriormente colocado em sinonímia com *P. aurisetus*. Porém, na localidade Engenheiro Dolabela foi encontrado apenas o haplótipo H3, comum a todas as outras populações de *P. a. aurisetus* e divergente em quatro passos mutacionais do haplótipo encontrado em Grão Mogol. Dessa forma, esses dados não suportam a distinção taxonômica da população de Engenheiro Dolabela-MG e lança dúvidas se a população de Grão Mogol-MG pertence a espécie *P. aurisetus*. Por outro lado, a similaridade genética da população de *P. parvus* e das populações do norte de *P. machrisii* concorda com a história taxonômica da espécie, uma vez que *P. parvus* já foi considerada sinonímia de *P. machrisii* (ZAPPI, 1994). No entanto, a distância genética entre as populações de *P. parvus* e *P. machrisii* que ocorrem na mesma localidade, sugere que essas espécies são realmente táxons isolados.

Para a espécie *P. machrisii*, a SAMOVA apontou uma subestrutura para os haplogrupos norte e centro-sul. O haplogrupo norte foi separado em dois agrupamentos devido a grande divergência entre os haplótipos da população ART e os das demais populações. Da mesma forma, o haplogrupo centro-sul foi dividido em dois grupos. Nesse caso, a proximidade geográfica das populações CRI, FMS e APA1 permitiu que essas populações formassem um grupo distinto das populações do sul de Minas Gerais e norte de São Paulo (BRO DEL FOR e FUR).

As populações de *P. aurisetus* do leste e oeste da Cadeia do Espinhaço formaram grupos isolados, corroborando a hipótese de diferenciação regional da espécie observada na filogenia e na NCPA.

A grande divergência observada entre as espécies *P. aureispinus* (IBO), *P. bohlei* (UMB) e o restante do grupo resultou na formação de grupos isolados para essas espécies. O mesmo foi observado para a espécie *P. jauruensis* e para as populações de *P. vilaboensis*. Dessa forma, os agrupamentos gerados concordam parcialmente com os limites taxonômicos das espécies, exceto para as situações envolvendo as populações de *P. aurisetus* de Engenheiro Dolabela e Grão Mogol em Minas Gerais e a população de *P. parvus* de Posse.

5.3. História Filogeográfica

O padrão de especiação do grupo *P. AURISSETUS* foi considerado como predominantemente alopátrico (ZAPPI, 1994), em concordância com o observado para vários organismos de ocorrência em áreas secas no Cerrado, como lagartos, roedores e plantas do gênero *Hymenaea* (NOGUEIRA et al., 2011; ALMEIDA et al., 2007; RAMOS et al., 2009), por exemplo. No entanto, a análise de NCPA sugeriu outros possíveis eventos influenciando a história filogeográfica do grupo como expansões e colonização a longa distância. A ocorrência de algumas dessas expansões foi confirmada apenas com a análise de *Mismatch distribution*, uma vez que os testes de neutralidade não foram significativos. Esse resultado possivelmente está relacionado ao baixo número de sítios polimórficos encontrados nas sequências analisadas, o que acarreta um menor poder estatístico na análise (EXCOFFIER et al., 2009). A maior parte das inferências da NCPA apontaram para eventos de fragmentação. Além disso, as características dos dispersores, a instabilidade dos habitats intermediários e a elevada diferenciação entre populações próximas da mesma espécie (p. ex. *P. a. aurilanatus* e *P. a. aurisetus*), juntamente com a acentuada estruturação genética observada, indicam que a colonização a longa distância pode ser considerada um evento improvável no grupo.

As populações da distribuição norte e centro-sul de *P. machrisii* (clados 1-2, 1-8 e 3-2) apresentaram inferências em comum quanto a eventos de fragmentação e expansão. Apesar dos eventos de expansão de *P. machrisii* detectados para os clados 1-8 e 1-2 envolverem poucos haplótipos, resultando em menor poder estatístico para serem revelados pelas análises empregadas, esses eventos também foram inferidos para o clado de maior nível no qual o clado 1-8 está alocado (clado 3-2). No caso do clado 1.11, a mesma limitação foi observada, no entanto os eventos de expansão não foram confirmados nos níveis mais elevados. Apesar disso, o evento de expansão inferido para esse clado é possivelmente o evento inicial de diversificação do grupo *P. AURISSETUS*, devido a relação genealógica que o haplótipo H1 (clado 1.11) apresenta com todos os demais haplótipos observados.

O resultado inconclusivo observado para o clado 2-1, composto por populações de *P. machrisii* do leste de Goiás e sudeste do Brasil, foi atribuído a uma amostragem inadequada. Esse resultado é decorrente da possibilidade de ocorrência das espécies em áreas intermediárias que não foram amostradas, como alguns afloramentos rochosos isolados que ocorrem entre o norte de São Paulo e sul de Minas Gerais.

Eventos de fluxo gênico restrito foram detectados apenas entre as populações da distribuição norte de *P. machrisii* e para as espécies *P. parvus* e *P. vilaboensis* (clado 2-3). Esse resultado é indicativo de diversificação com fluxo gênico nas linhagens envolvidas. No entanto, essa inferência foi fortemente influenciada pela presença de haplótipos de *P. vilaboensis* em clados distintos (1-4 e 1-5), uma vez que para os demais clados (1-2 e 1-11) que compõem o clado 2-3 foram propostos eventos de fragmentação.

A divergência entre populações do leste e oeste da Cadeia do Espinhaço pode ser explicada pelo evento de fragmentação alopátrica detectado para o clado 2-4. Adicionalmente, a presença do haplótipo ancestral do grupo na população do extremo norte da Cadeia do Espinhaço e sua distância genética em relação aos demais haplótipos da espécie sugere que o isolamento das linhagens de *P. aurisetus* foi anterior a diversificação das demais espécies do grupo.

A pouca divergência observada entre os haplótipos do grupo *P. AURISSETUS*, com exceção de três espécies (*P. jauruensis*, *P. aureispinus* e *P. bohlei*), bem como a ausência de monofilia recíproca na maior parte dos clados sugerem uma diversificação recente do grupo. Além disso, a ocorrência de vários haplótipos conectados a um haplótipo central resultando

em uma distribuição *mismatch* unimodal (Figura 11A) é indicativa de ocorrência de expansão populacional. Esses dois resultados em conjunto, além da indicação de expansão e fragmentação em vários subclados pela NCPA, sugere que a diversificação desse grupo de espécies, excluindo as duas espécies mais divergentes (*P. aureispinus* e *P. bohlei*) está associada a um ou mais eventos importantes de expansão a partir do norte de Minas Gerais. Esta expansão propiciou a ocupação de áreas no Planalto Central, a partir de onde eventos posteriores de expansão ocuparam o restante da distribuição do grupo no leste de Goiás e também na divisa entre Minas Gerais e São Paulo (Serra da Canastra e morros isolados). Em seguida esta distribuição teria sofrido fragmentação e permitido a divergência de grupos de populações, conforme sugerido pela acentuada estrutura filogeográfica observada na rede de haplótipos. Neste cenário, *P. jauruensis* parece ter permanecido isolado há mais tempo que outros grupos de populações ou ter se originado de uma expansão mais antiga e, conforme sugerem as relações entre os haplótipos dessa espécie com aqueles de *P. vilaboensis*, possivelmente *P. jauruensis* se originou a partir de populações ocupando o centro de Goiás.

Um estudo filogeográfico com a espécie de Asteraceae *Wunderlichia mirabilis* também identificou rotas iniciais de expansão a partir da Cadeia do Espinhaço. Esta espécie apresenta sobreposição parcial com as espécies do grupo *P. AURISSETUS*, uma vez que ocorre associada a ambientes xéricos e rupestres na porção mineira da Cadeia do Espinhaço e no estado de Goiás (FERES et al., 2009).

As direções dos eventos de expansão a partir do Planalto Central (norte-sul) observadas para o grupo *P. AURISSETUS* concordam com expansões observadas para a espécie cactofílica *D. gouveai* (de BRITO et al., 2002; MORAES et al., 2009). Dentre as espécies de cactos hospedeiros estão *P. machrisii* e *P. vilaboensis*. Moraes et al. (2009) encontraram evidências de eventos de expansão para dois subclados de *D. gouveai*, cuja distribuição atual acompanha os topos de montanhas nordeste, centro-oeste e sudeste do Brasil. Considerando a sobreposição parcial das distribuições do grupo *P. AURISSETUS* e da mosca *D. gouveai* é possível que as expansões possam ter seguido as mesmas rotas propostas por Moraes et al. (2009). Expansões no sentido norte-sul também foram observadas para a espécie de anuro *Hypsiboas albopunctatus* que ocorre em áreas abertas no Cerrado e em áreas adjacentes de florestas na América do Sul. Neste caso as expansões

ocorreram a partir do Planalto Central em direção a Mata Atlântica do sudeste e sul do país (PRADO et al., 2011). No geral, essas expansões concordam com a alteração da distribuição da vegetação de áreas abertas durante os períodos secos do Pleistoceno, quando o Cerrado expandiu para o sudeste do Brasil (BEHLING; HOOGHIEMSTRA, 2000).

O padrão filogeográfico atual do grupo *P. AURISSETUS*, também poderia ser explicado por um contato secundário no norte de Goiás de linhagens que divergiram em alopatria na região do Planalto Central e no sul de Minas Gerais e norte de São Paulo. A ocorrência de contato secundário nessa região foi proposta para a espécie *A. urundeuva* como consequência de expansões após a fragmentação das populações em um grupo norte (Caatinga) e um grupo sul (Chaco) durante períodos úmidos do Pleistoceno. No entanto a diversidade haplotípica encontrada nas populações de *P. machrisii* no norte de Goiás é consistente com a existência de populações grandes e estáveis no passado, permitindo o surgimento e manutenção de linhagens diferenciadas. Além disso, a região citada está no vale do Rio Paranã, uma área relativamente grande e de estabilidade para a vegetação seca, além de ser um dos principais centros de endemismo do Cerrado, onde várias espécies se divergiram por isolamento geográfico e expandiram sua distribuição (WERNECK, 2011). Dessa forma, a diversidade encontrada na região parece estar mais relacionada a retenção de polimorfismo ancestral do que a um contato secundário de linhagens independentes.

A ausência de diversidade nas populações de *P. aurisetus* na Cadeia do Espinhaço é um resultado difícil de ser explicado, em função da provável estabilidade histórica das condições xéricas dessa área, conforme demonstrado pelos elevados níveis de endemismo e diversidade genética (WERNECK, 2011). Este resultado pode ser atribuído a ocorrência de um gargalo populacional nessas populações. No entanto, esse evento ainda seria incompatível com a história da região.

O compartilhamento de haplótipos entre as populações de *P. machrisii* pode ser explicado pela retenção de polimorfismo ancestral (*incomplete lineage sorting*) ou fluxo gênico. Para as populações de Goiás, o compartilhamento de haplótipos observado entre CRI e FMS e URU e APA2 pode ser resultado de isolamento mais recente ou fluxo gênico, uma vez que essas populações estão bem próximas geograficamente e possuem haplótipos mais derivados na genealogia. Para as populações do sul (BRO DEL e FOR), o compartilhamento de

haplótipos deve estar relacionado a retenção de polimorfismo ancestral, uma vez que a população FUR, geograficamente muito próxima a população de DEL, apresenta um haplótipo diferente. O compartilhamento de haplótipos entre populações de *P. jauruensis* também parece ser resultado de retenção de polimorfismo ancestral, uma vez que populações bem distantes geograficamente apresentaram o mesmo haplótipo. Eventos de expansão recentes também podem ser responsáveis pelo compartilhamento de haplótipos observado.

A distribuição restrita das espécies do grupo *P. AURISSETUS* em ambientes xéricos associados a afloramentos rochosos sugere que as oscilações demográficas que as espécies sofreram devem estar relacionadas aos ciclos paleoclimáticos de retração e expansão da vegetação seca. Estudos com espécies restritas a esses ambientes apresentaram dados que suportam a possibilidade de uma distribuição mais ampla da vegetação seca no Quaternário, concordando com a hipótese da existência de um “arco Pleistocênico” (WERNECK; COLLI, 2006; CAETANO et al., 2008; MORAES et al., 2009). Segundo essa hipótese os domínios da Caatinga, Cerrado e Chaco formaram um contínuo de vegetação xérica durante o Último Máximo Glacial (LGM). Dessa forma, os enclaves de vegetação seca, como as SDTFs, observados ao longo do Cerrado seriam remanescentes de uma distribuição mais ampla no passado.

Dentre as evidências para a hipótese de um arco Pleistocênico estão estudos com organismos endêmicos da SDTF como lagartos da espécie *Lygodactylus klugei* (WERNECK; COLLI, 2006) e espécies vegetais como a árvore *Astronium urundeuva* (CAETANO et al. 2008). A presença dessas espécies em enclaves de SDTF ao longo do Cerrado e do Chaco e sua relação genética com populações mais amplamente distribuídas na Caatinga indicam que essas espécies possuíam uma distribuição mais ampla e contígua no passado.

5.4. Implicações para a conservação das espécies

O entendimento de eventos históricos na dinâmica populacional é necessário para prever o comportamento das espécies em diferentes cenários ambientais. Atualmente, são raros os estudos envolvendo análises da variabilidade genética e limites geográficos da distribuição das espécies. Conforme apresentado no presente estudo, algumas espécies do grupo *P. AURISSETUS* exibem distribuição conhecida bastante restrita, o que demanda maiores esforços para investigar a extensão da ocorrência das espécies e de sua diversidade genética. As observações de campo durante o desenvolvimento desse projeto mostram que as populações possuem geralmente menos de 100 indivíduos, embora esse dado necessite de estudos de censo para ser confiável.

Estudos sobre diversidade genética, limites de distribuição e censo populacional são necessários para a definição do estado de conservação das espécies e de possíveis estratégias de conservação para o táxon e para o ambiente como um todo, dependendo da extensão da ameaça. Os dados aqui apresentados sugerem uma grande estruturação e isolamento das populações, além da ocorrência de baixa diversidade genética para algumas espécies. Esses resultados podem estar relacionados a diversificação recente do grupo, motivo pelo qual muitas espécies não apresentam padrões monofiléticos em sua filogenia

As populações que apresentaram maior diversidade e provavelmente mantiveram um tamanho populacional maior e estável ao longo do tempo foram encontradas na região do Planalto central, entre o norte de Goiás e sul de Tocantins. Esses dados associados a dados históricos sobre a estabilidade ambiental apontam essa área como prioritária para conservação. Mediante alterações ambientais essa região seria mais estável e permitiria a manutenção de maior biodiversidade. Estudos utilizando outros marcadores moleculares e dados sobre a biologia e distribuição dessas espécies serão necessários para a identificação do seu real estado de conservação.

6. CONCLUSÕES

Os resultados aqui apresentados sugerem uma história evolutiva complexa para o grupo *P. AURISSETUS*. As inferências filogenéticas elaboradas a partir de sequências do cpDNA mostraram a ausência de monofilia recíproca em grande parte das espécies do grupo sugerindo eventos de diversificação recentes. As espécies *P. bohlei* e *P. aureispinus* se mostraram bastante distantes filogeneticamente e provavelmente não participaram da diversificação das demais espécies do grupo. Para a espécie *P. machrisii* foram identificados dois clados filogenéticos que mostraram importante associação com a distribuição geográfica da espécie. A divergência observada na genealogia dos haplótipos e o padrão atual de distribuição das espécies sugerem que a diversificação do grupo se originou a partir de eventos independentes de expansão e fragmentação da distribuição das populações os quais teriam se originado no norte de Minas Gerais. De forma geral, as alterações demográficas detectadas para o grupo *P. AURISSETUS* podem estar relacionadas com as alterações na distribuição da vegetação seca durante o Pleistoceno na América do Sul. Dessa forma, eventos de fragmentação e isolamento das populações seguido de expansões para novas áreas seriam os principais responsáveis pelo padrão biogeográfico atual das espécies. A combinação dos dados aqui apresentados com informações de marcadores moleculares mais variáveis e com padrão de herança biparental poderão melhor elucidar a história filogeográfica do grupo.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, F.C.; BONVICINO, C.R.; CORDEIRO-ESTRELA, P. Phylogeny and temporal diversification of *Calomys* (Rodentia, Sigmodontinae): implications for the biogeography of an endemic genus of the open/dry biomes of South America. **Molecular Phylogenetics and Evolution**, v. 42, p. 449–466. 2007.

AMICO, G.C.; NICKRENT, D.I. Population structure and phylogeography of the mistletoes *Tristerix corymbosus* and *T. aphyllus* (Loranthaceae) using chloroplast DNA sequence variation. **American Journal of Botany**, v. 96, p. 1571–1580. 2009.

ARTYUKOVA, E.V. et al. Plastid DNA variation in highly fragmented populations of *Microbiota decussata* Kom. (Cupressaceae), an endemic to Sikhote Alin Mountains. **Genetica**, v. 137, p. 201-212. 2009.

AVISE, J. C. Phylogeography: retrospect and prospect. **Journal of Biogeography**, v. 36 p. 3-15. 2009.

BARBARÁ, T. et al. Within-population spatial genetic structure in four naturally fragmented species of a neotropical inselberg radiation, *Alcantarea imperialis*, *A. geniculata*, *A. glaziouana* and *A. regina* (Bromeliaceae). **Heredity**, v. 101, p. 285–296. 2008.

BEHLING, H.; HOOGHMSTRA H. Holocene Amazon rainforest – savanna dynamics and climate implications high-resolution pollen record from Laguna Loma Linda in eastern Colombia. **Journal of Quaternary Science**, v. 15, p. 687-695. 2000.

BEHLING, H. Late glacial and Holocene vegetation, climate and firehistory inferred from Lagoa Nova in the southeastern Brazilian lowland. **Vegetation History Archaeobotany**, v. 12, p. 263–270. 2003.

BONATELLI I.A.S. et al. Applicability of plastid DNA regions for intra- and interspecific studies in closely related cacti species. 15p. 2012.

BRASIL. Ministério do Meio Ambiente. **Lista Oficial das Espécies da Flora Brasileira Ameaçadas de Extinção**. 2008. Disponível em: <<http://www.mma.gov.br>>. Acesso em 15 de janeiro de 2012.

BRAUN, P.J.; ESTEVES PEREIRA, E. *Pilosocereus supthutianus* P. J. Braun. Eine neue Art aus dem zentralen Minas Gerais, Brasilien. **Kakteen und andere Sukkulente**, v. 36, p. 100-103. 1985.

BRAUN, P.J.; ESTEVES PEREIRA E. *Pilosocereus lindaianus* Braun e Esteves Pereira spec. nov. **Kakteen und andere Sukkulente**, v. 38, p. 132. 1987a.

BRAUN, P.J.; ESTEVES, PEREIRA E. *Pilosocereus cristalinensis* Braun e Esteves Pereira spec. nov. **Kakteen und andere Sukkulente**, v. 38, p. 163-167. 1987b.

BRAUN, P.J.; ESTEVES PEREIRA, E. *Pilosocereus lindaianus* Braun e Esteves Eine neue Art aus dem nordöstlichen Goiás/Brasilien **Kakteen und andere Sukkulente**, v. 39, p. 222-228. 1988.

BRAUN, P.J.; ESTEVES PEREIRA, E. *Pilosocereus goianus* Braun e Esteves spec nov. **British Cactus and Succulent Journal**, v. 20, p. 93-103. 2002.

CAETANO, S. et al. The history of Seasonally Dry Tropical Forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). **Molecular Ecology**, v. 17, p. 3147–3159. 2008.

CLEMENT, M.; POSADA, D; CRANDALL, K. TCS: a computer program to estimate gene genealogies. **Molecular Ecology**, v. 9, p. 1657-1660. 2000.

COLLEVATTI, R.G.; RABELO, S.G.; VIEIRA, R.F. Phylogeography and disjunct distribution in *Lychnophora ericoides* (Asteraceae), an endangered cerrado shrub species. **Annals of Botany**, v. 104, p. 655–664. 2009.

CORRIVEAU, J.L.; COLEMAN, A.W. Rapid screening method to detect potential biparental inheritance of plastid DNA and results over 200 angiosperm species. **American Journal of Botany**, v. 75, p. 1443-1458. 1988.

- CRANDALL, K.A.; TEMPLETON, A.R. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. **Genetics**, v. 134, p. 959–969. 1993.
- DE BRITO, R.A.; MANFRIN, M.H.; SENE, F.M. Nested cladistic analysis of Brazilian populations of *Drosophila serido*. **Molecular Phylogenetics and Evolution**, v. 22, p. 131–143. 2002.
- DICK, C.W.; ABDUL-SALIM, K.; BERMINGHAM, E. Molecular systematic reveals cryptic Tertiary diversification of a widespread tropical rainforest tree. **American Naturalist**, v. 162, p. 691–703. 2003.
- DICK, C.W. et al. Extreme longdistance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. **Molecular Ecology**, v. 16, p. 3039–3049. 2007.
- DRUMMOND, A.J.; RAMBAUT, A.; SUCHARD, M.A. BEAST version 1.6.1 [computer program]. Disponível em: <<http://beast.bio.ed.ac.uk>>. Acesso em 21 de dezembro de 2011.
- DUPANLOUP, I.; SCHNEIDER, S.; EXCOFFIER L. A simulated annealing approach to define the genetic structure of populations. **Molecular Ecology**, v. 11, p. 2571–2581. 2002.
- EXCOFFIER, L.; SMOUSE, P.E.; QUATTRO, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial restriction data. **Genetics**, v. 131, p. 479–491. 1992.
- EXCOFFIER, L., LAVAL, G.; SCHNEIDER, S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. **Evolutionary Bioinformatics Online**, v. 1, p. 47–50. 2005.
- EXCOFFIER, L.; FOLL, M.; PETIT R.J. Genetic consequences of range expansions. **Annual Review in Ecology, Evolution and Systematics**, v. 40, p. 481–501. 2009.
- FERES, F. et al. Phylogeographic studies of Brazilian “campo-rupestre” species: *Wunderlichia mirabilis* Riedel ex Baker (Asteraceae). **Biotemas**, v. 22, p. 17–26. 2009.

FU, Y-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. **Genetics**, v. 147, p. 915-925. 1997.

GAUDEUL, M. Disjunct distribution of *Hypericum nummularium* L. (Hypericaceae): molecular data suggest bidirectional colonization from a single refugium rather than survival in distinct refugia. **Biological Journal of the Linnean Society**, v. 87, p. 437–447. 2006.

GIULIETTI, A.M.; PIRANI, J. R.; HARLEY R.M. Espinhaço Range Region, Eastern Brazil. In: DAVIS, S. D.; HEYWOOD, V. H.; HERRERA-MACBRYDE, O.; VILLA-LOBOS, J.; HAMILTON A. C. **Centres of plant diversity. A guide and strategy for their conservation**. Cambridge, Massachusetts, USA: The Americas. IUCN Publication Unity, 1997. p. 397-404.

HELSEN, P. et al. Galápagos' *Opuntia* (prickly pear) cacti: extensive morphological diversity, low genetic variability **Biological Journal of the Linnean Society**, v. 96, p. 451–461. 2009.

HUNT, D.; TAYLOR, N. P.; CHARLES, G. **The New Cactus Lexicon**. 1. ed. Milborne Port: Dh Books, 2006. 373 p.

INTERNATIONAL UNION FOR CONSERVATION OF NATURE. IUCN Red List of Threatened Species. Version 2012. Disponível em: <www.iucnredlist.org>. Acesso em 8 de fevereiro de 2012.

JUMP, A.S.; PEÑUELAS, J. Genetic effects of habitat fragmentation in wind-pollinated tree. **Proc Natl Acad Sci USA**, v. 103, p. 8096–8100. 2006.

KNOWLES, L.L. Statistical Phylogeography. **Annual Review of Ecology, Evolution, and Systematics**, v. 40, p. 593-612. 2009.

LAMBERT, S.M.; BORBA, E.L.; MACHADO, M. Allozyme diversity and morphometrics of the endangered *Melocactus glaucescens* (Cactaceae), and investigation of the putative hybrid origin of *Melocactus* × *albicephalus* (*Melocactus ernestii* × *M. glaucescens*) in north-eastern Brazil. **Plant Species Biology**, v. 21, p. 93-108. 2006a.

LAMBERT, S.M. et al. Allozyme diversity and morphometrics of *Melocactus paucispinus* (Cactaceae) and evidence for hybridization with *M. concinnus* in the Chapada Diamantina, north-eastern Brazil. **Annals of Botany**, v. 97, p. 389-403. 2006b.

- LOCATELLI, E.; MACHADO, I.C.; MEDEIROS, P. Floral biology and bat pollination in *Pilosocereus catingicola* (Cactaceae) in northeastern Brazil. **Bradleya**, v. 15, p. 28–34. 1997.
- LORENZ-LEMKE, A.P. et al. Diversification of plant species in a subtropical region of eastern South American highlands: a phylogeographic perspective on native *Petunia* (Solanaceae). **Molecular Ecology**, v. 19, p. 5240–5251. 2010.
- MORAES, E. M. et al. Genetic variability and population structure of columnar cacti in dry regions of Brazil. **Genetica**, v. 125, p. 311-323. 2005.
- MORAES, E.M. et al. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. **Journal of Biogeography**, v.36, p. 2136-2147. 2009.
- NASSAR, J.M.; RAMIREZ N.; LINARES O. Comparative pollination biology of Venezuelan columnar cacti and the role of nectar-feeding bats in their sexual reproduction. **American Journal of Botany**, v. 84, p. 918–927. 1997.
- NOGUEIRA, C. et al. Vicariance and endemism in a Neotropical savanna hotspot: distribution patterns of Cerrado squamate reptiles. **Journal of Biogeography**, v.38, p. 1907–1922. 2011.
- NOVAES, R.M.L. et al. Phylogeography of *Plathymentia reticulata* (Leguminosae) reveals patterns of recent range expansion towards northeastern Brazil and southern Cerrados in Eastern Tropical South America. **Molecular Ecology**, v.19, p. 985-998. 2010.
- PENNINGTON, R.T. et al. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. **Philosophical Transactions of The Royal Society of London Series B-Biological Sciences**, v. 359, p. 515-538. 2004.
- PETIT, R.J. et al. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. **Molecular Ecology**, v. 14, p. 689–701. 2005.
- POSADA, D. jModelTest: Phylogenetic Model Averaging. **Molecular Biology and Evolution**, v. 25, p. 1253-1256. 2008.

- POSADA, D.; CRANDALL, K.A; TEMPLETON, A.R. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. **Molecular Ecology**, v.9, p. 487–488. 2000.
- PRADO, D.; GIBBS, P.E. Patterns of species distributions in the dry seasonal forests of South America. **Annals of the Missouri Botanical Garden**, v. 80, p. 902-927. 1993.
- PRADO, C.P.A.; HADDAD, C.F.B.; ZAMUDIO, K.R. Cryptic lineages and Pleistocene population expansion in a Brazilian Cerrado frog. **Molecular Ecology**, v. 21, p. 921-941. 2011.
- RAMOS, A.C.S. et al. Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of Quaternary climate changes in the Brazilian Cerrado. **Annals of Botany**, v. 100, p. 1219–1228. 2007.
- RAMOS, A.C.S.; LEMOS-FILHO, J.P.; LOVATO, M.B. Phylogeographical Structure of the Neotropical Forest Tree *Hymenaea courbaril* (Leguminosae: Caesalpinioideae) and Its Relationship with the Vicariant *Hymenaea stigonocarpa* from Cerrado. **Journal of Heredity**, v. 100, p. 206-216. 2009.
- RAVI, V. et al. An update on chloroplast genomes. **Plant Systematics and Evolution**, v. 271, p. 101–122. 2008.
- RITTER, F. 1979. Kakteen in Südamerika, v. 1. F. Ritter Selbstverlag, Spangenberg.
- ROGERS, A.R.; HARPENDING H. Population growth makes waves in the distribution of pairwise genetic differences. **Molecular Biology and Evolution**, v. 9, p. 552-569. 1992.
- ROZAS, J. et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods. **Bioinformatics**, v.19, p. 2496-2497. 2003.
- SCHAAL, B.A et al. Phylogeographic studies in plants: problems and prospects. **Molecular Ecology**, v. 7, p. 465–474. 1998.
- SHAW, J. et al. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. **American Journal of Botany**, v. 92, p. 142-166. 2005.

- TABERLET, P. et al. Universal primers for the amplification of three non-coding regions of chloroplastic DNA. **Plant Molecular Biology**, v.17, p. 1105–1109. 1991.
- TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, p. 585-595. 1989.
- TAMURA, K. et al. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology and Evolution**, v. 24, p. 1596-1599. 2007.
- TAYLOR, N.P.; Zappi D.C. **The Cacti of Eastern Brazil**. 1 ed. Richmond, Surrey: Royal Botanic Garden, Kew. 2004, 499 p.
- TEMPLETON, A.R. Statistical phylogeography: methods of evaluating and minimizing inference errors. **Molecular Ecology**, v. 13, p. 789-80. 2004.
- TEMPLETON, A.R; BOERWINKLE, E.; SING, C.F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping .1 Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. **Genetics**, v. 117, p. 343–351. 1987.
- TEMPLETON, A.R; ROUTMAN, E.; PHILLIPS, C.A. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. **Genetics**, v.140, p.767–782. 1995.
- THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific, gaps penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p. 4673-80. 1994.
- WERNECK, F. P. The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. **Quaternary Science Reviews**, v. 30, p. 1630-1648. 2011.
- WERNECK, F.P.; COLLI, G.R. The lizard assemblage from seasonally dry tropical forest enclaves in the Cerrado biome, Brazil, and its association with the Pleistocenic Arc. **Journal of Biogeography**, v. 33, p. 1983–1992. 2006.

YOUNG, A.; BOYLE, T.; BROWN, T. The population genetic consequences of habitat fragmentation for plants. **Trends in Ecology and Evolution**, v. 11, p. 413–418. 1996.

ZAPPI, D.C. *Pilosocereus* (Cactaceae). **The genus in Brazil**. Milborne Port: D. HUNT & N. TAYLOR Royal Botanic Gardens, Kew, 1994. 160p.

ZAPPI, D. C. et al. **Catálogo de Plantas e Fungos do Brasil**. (Eds.: WALTER, B. M. T & ZAPPI, D.C.), Jardim Botânico do Rio de Janeiro, v.1, p. 822 – 832. 2010.

APÊNDICE 1. Resultados da *Nested Clade Phylogeographic Analysis* para as quatro espécies do grupo *P. AURISSETUS* analisadas a partir de haplótipos do cpDNA. Os valores de I-T indicam a diferença entre os clados de interior e extremidade, L S sobrescritos indicam os valores significativamente altos (L) ou baixos (S). As linhas sombreadas indicam os clados de interior.

Haplótipos			Clado Nível 1			Clado Nível 2			Clado Nível 3			Clado Nível 4
Clado	Dc	Dn	Clado	Dc	Dn	Clado	Dc	Dn	Clado	Dc	Dn	Clado
4	89,56 ^S	258,38 ^S										
5	0,00 ^S	458,01 ^L										
6	88,58 ^S	328,74										
I-T	-23,13	102,68 ^L	1-8	326,83	327,19							
7			1-1	0,00 ^S	240,07							
			I-T	326,83 ^L	87,12	2-1	314,75 ^S	324,51 ^S				
8						2-2	0,00	588,13 ^L				
10	72,86 ^S	102,79				I-T	-314,74 ^S	263,62 ^L	3-2	370,28	444,42 ^L	
11	0,00 ^S	156,99 ^L										
12	0,00	60,56										
I-T	72,86	-38,13	1-2	108,41 ^S	179,88							
1	0,00 ^S	401,76 ^L										
9	99,36 ^S	210,10										
13	0,00 ^S	86,74 ^S										
I-T	-70,97	226,27 ^L	1-11	227,62	216,04							
18												
19			1-4	0,00 ^S	260,44							
16												
17			1-5	0,00 ^S	379,75 ^L							
			I-T	160,90 ^L	-17,78	2-3	228,03 ^S	385,39 ^L				
3			1-3	0,00 ^S	103,04							
2			1-6	85,78	85,70							
			I-T	-	-	2-4	87,44 ^S	277,81 ^S				
						I-T	-140,59 ^S	-107,57 ^S	3-1	313,24 ^S	325,05 ^S	
									I-T	-	-	4-1

APÊNDICE 2. Valores de F_{CT} e F_{SC} obtidos na análise de SAMOVA segundo o número de agrupamentos sugerido (k).

k	F_{CT}	F_{SC}
2	0.820	0.977
3	0.725	0.977
4	0.714	0.975
5	0.714	0.955
6	0.773	0.934
7	0.808	0.921
8	0.832	0.909
9	0.849	0.896
10	0.867	0.880
11	0.886	0.859
12	0.904	0.829
13	0.919	0.796
14	0.917	0.800
15	0.919	0.794
16	0.938	0.728
17	0.967	0.488
18	0.954	0.628
19	0.964	0.518
20	0.986	0,000
21	0.982	0.018
22	0.986	0,000
23	0.986	0,000
24	0.986	0,000
25	0.986	0,000
26	0.986	0,000
27	0.986	0,000
28	0.986	0,000
29	0.986	0,000
30	0.986	0,000
31	0.986	0,000
32	0.988	0,000

ANEXO. Manuscrito elaborado relacionado ao presente trabalho

Applicability of plastid DNA regions for intra- and interspecific studies in closely related cacti species

Isabel A. S. Bonatelli¹, Daniela C. Zappi^{2,3}, Nigel P. Taylor³, and Evandro M. Moraes¹,

¹Departamento de Biologia, Universidade Federal de São Carlos, Rodovia João Leme dos Santos km 110, 18052780, Sorocaba, São Paulo, Brazil; ²Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, United Kingdom; ³National Parks Board, Gardens by the Bay and Singapore Botanic Gardens, 1 Cluny Road, Singapore 259569.

Abstract

Despite the fact that plastid DNA having been widely explored as the marker of choice for phylogeny and phylogeography, little is known about its utility in closely related species studies. The slow evolutionary rates inherent in chloroplast DNA may make it difficult to perform lower level taxonomic analyses, particularly to the population level.

We characterized the nucleotide variation and investigated the utility of eight noncoding cpDNA regions in closely related species of the *PILOSOCEREUS AURISSETUS* group (Cactaceae). The plastid intergenic spacers *5'trnS-trnG*, *3'trnS-trnG* and *trnT-trnL* were pointed as the most variable regions and fared best for lower level taxonomy approaches, mainly when they were used together. The analysis suggests that using an adequate combination of regions alongside indels as an additional character improves the phylogenetic signal of cpDNA.

The study of intra- and interspecific variation presents an inherent challenge: the choice of the appropriate molecular markers for the relevant taxonomic level. In plants, noncoding chloroplast DNA (cpDNA) regions have been explored for lower level taxonomic studies because they are under less selective pressure than coding regions and show an amount of polymorphic sites useful for phylogenetic studies (Pleines *et al.*, 2009). However, the phylogenetic signal of different noncoding cpDNA regions can vary extensively among taxonomic groups (Small *et al.*, 1998; Shaw *et al.*, 2005) and the evolutionary rates of different regions are poorly understood. As pointed by Shaw *et al.* (2005), the choice of DNA

markers for phylogenetic studies may involve guesswork, as such information remains scarce for many taxa.

The cpDNA has a maternal inheritance in most species, biparental in some (Zhang, 2010), and paternal in others (Testolin and Cipriani, 1997). The maternally inherited cpDNA has an interesting role because this allows comparisons between seed gene flow and pollen gene flow as well as the identification of hybridization events when comparing to nuclear, biparentally inherited DNA (Wang *et al.*, 2011).

When it comes to intramolecular rearrangement, cpDNA shows a much more stable structure than plant mitochondrial DNA (mtDNA). However, the substitution rate of the plastid genome is three to four times higher than plant mtDNA (Avice, 2009). Much of the variation observed in plastid noncoding regions is related to insertion-deletion mutations (indels) and should be treated carefully, since these are prone to homoplasy and can be a result of PCR artifacts, thus providing an erroneous phylogenetic signal.

In recent years, more efforts to identify plastid regions useful as molecular markers for plant phylogeny have been made. Shaw *et al.* (2005) evaluated the variation in 21 plastid regions of 29 species from 9 families of seed plants using only noncoding DNA in order to find useful markers for low taxonomic levels. In a subsequent work, Shaw *et al.* (2007) analyzed the most variable regions compared from whole genomes (*Atropa* vs. *Nicotiana*, Solanaceae; *Lotus* vs. *Medicago*, Fabaceae; *Saccharum* vs. *Oryza*, Poaceae), which allowed the identification of additional high variation regions.

Pilosocereus is a genus of neotropical columnar cacti subdivided in taxonomic groups of morphologically similar species (Zappi, 1994), among them the *PILOSOCEREUS AURISSETUS* group. The species belonging to this group exhibit substantial morphological variation across their range, leading to an unstable taxonomic history, with several species being placed in and removed from synonymy (Zappi, 1994; Taylor and Zappi, 2004; Hunt *et al.*, 2006). Currently, this group consists of eight columnar species: *P. aureispinus* (Buining and Brederoo) Ritter; *P. aurisetus* (Werdermann) Byles and Rowley; *P. bohlei* Hofacker; *P. jauruensis* (Buining and Brederoo) Braun; *P. machrisii* (Dawson) Backeberg; *P. parvus* (Diers and Esteves Pereira) Braun; *P. pusillibaccatus* Braun and Esteves Pereira; and *P. vilaboensis* (Diers and Esteves Pereira) Braun, which show a disjunct distribution pattern occurring in enclaves of dry vegetation, mostly on rock outcrops within the Cerrado domain in eastern and central Brazil. (Zappi, 1994). Here we characterize the variability of noncoding plastid

regions and investigate their utility for intra- and interspecific studies in closely related species of the *P. AURISSETUS* group. We intend to point to promising candidate cpDNA regions to guide future phylogenetic and population studies in cacti species.

Seven populations, comprising two to 10 individuals per population were analyzed across the distribution of the *P. AURISSETUS* group, in eastern and central Brazil: *P. machrisii* from Altinópolis-SP (FOR), Delfinópolis-SP (DEL), Alto Paraíso de Goiás-GO (APA), and Cristalina-GO (CRI) localities; *P. aurisetus* from Cardeal Mota-MG (CMO); *P. vilaboensis* from Pirenópolis-GO (PIR); and *P. aureispinus* from Ibotirama-BA (IBO).

The genomic DNA was extracted from root tissue using Qiagen - Dneasy Plant Mini Kit (Qiagen, Hilden, Germany) with all centrifugations in maximum speed (13,200 rpm). Eight regions of cpDNA were selected based on the availability of PCR primers and previous reports on phylogenetic utility, including the intron *trnL* and the intergenic spacers *trnS-trnG*, *trnH-psbA*, *trnT-trnL*, *trnL-trnF*, *3'rps16-5'trnK* and *atpI-atpH*. All these regions are located in the cpDNA Large Single Copy (LSC) region.

The *trnS-trnG* was amplified in two halves overlapping in 104 bp by using previously reported primers in combination with primers designed in this study (primers *SGRev2* and *SGFwd2*). The region 5' *trnS-trnG* was amplified by using primers *trnS^{GCU}* (Shaw *et al.* 2005) and *SGRev2* (5' TCC GCT CAT TAG CTC TCC TC 3'), while 3'*trnS-trnG* was amplified by using primers *5'trnG2S* (Shaw *et al.*, 2005) and *SGFwd2* (5' CAC CCA TGG TTC CCA TTA GA 3'). PCR was performed in 25 µL containing 1X reaction buffer [100 mM Tris-HCl (pH 8,8), 500 mM KCl, 0,8% Nonidet P40], 1.5 mM MgCl₂, 200µM of each dNTP, 0.1µM of each primer, 1U of Taq DNA polymerase (Fermentas, Burlington, Ontario, Canada) and 1µl of template DNA (5-40ng). The *trnH-psbA* region was amplified using primers developed by Hamilton (1999), 2 mM of MgCl₂, 0.4µM of each primer and 0,5 U of Taq DNA polymerase. The intron *trnL* and the intergenic spacers *trnT-trnL* and *trnL-trnF* were amplified with Taberlet *et al.* (1991) primers. A final concentration of 3 mM MgCl₂ and 1,25U Taq polymerase were used for *trnT-trnL* and 2mM MgCl₂ for *trnL* and *trnL-trnF*. The regions *3'rps16-5'trnK* and *atpI-atpH* were amplified using primers of Shaw *et al.* (2007), 3mM MgCl₂ and 0.8 U of Taq DNA polymerase. All amplifications were performed using an Eppendorf Mastercycler® gradient (Eppendorf AG, Hamburg, Germany) according to the PCR conditions described by Shaw *et al.* (2005, 2007) with the following annealing temperatures: 62°C for *3'trnS-trnG*, *5'trnG2S* and *trnL*,

53°C for *trnH-psbA*, 50°C for *trnT-trnL*, 60°C for *trnL-trnF*, 48°C for *3'rps16-5'trnK* and 56°C for *atpI-atpH*.

The PCR products were run in 1.5% agarose gels stained with ethidium bromide and purified by ExoSAP - IT or *illustra GFX PCR DNA and Gel Band Purification kit* (GE Healthcare, Piscataway, New Jersey USA). Both strands were sequenced using the same primers as for the amplification. The samples were sequenced with an ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster City, California, USA) and prepared using Big Dye terminator version 3.1 cycle sequencing kit (Applied Biosystems, Foster City, California, USA).

Forward and reverse DNA sequences were combined in Chromas 1.5 software (Technelysium Pty Ltd., Tewantin, Australia) and sequences were aligned through CLUSTALW (Thompson *et al.*, 1994). Indels were manually coded as a single mutation when they shared identical boundaries and length.

Nucleotide analyses were carried out in DNAsp (Rozas *et al.*, 2003). The total number of substitutions (S) and indels (Id), the number of parsimoniously informative substitutions and indels, the number of sequences analyzed (N) and the number of haplotypes (h) were estimated. The proportion of variable sites (% variability) was calculated as $(S+Id/L) \times 100$, where S = number of substitution sites, Id = number of indels sites, and L = total sequence length. Based on our results, the three most variable regions (*5'trnS-trnG*, *3'trnS-trnG* and *trnT-trnL*) were used separately as well as in combination to construct parsimony haplotypes networks implemented in TCS 1.21 (Clement *et al.*, 2000). The networks were implemented by using five sequences of each population, which covered all variation found in our total sample. We also implemented a saturation test in DAMBE 5.2.32 software for each region (Xia and Xie, 2001).

We analyzed a total of 3809 bp obtained from 270 sequences (Table 1), including 40 sequences of *5'trnS-trnG* region (GenBank accession numbers JN035420-JN035459), 38 sequences of *3'trnS-trnG* region (JN035381-JN035418), 48 sequences of *trnH-psbA* region (JN035466-JN035513), 44 sequences of *trnT-trnL* region (JN035570- JN035613), 24 sequences of *trnL* region (JN035515-JN035538), 31 sequences of *trnL-trnF* region (JN035539-JN035569), 17 of *3'rps16-5'trnK* (JN035368-JN035372, JN035374-JN035379, JQ323547-JQ323552), and 28 of *atpI- atpH* (JN035462, JN035464, JQ323521-JQ323546).

Table 1 - Characterization of nucleotide variation of the plastid regions analyzed for the major sample of *PILOSOCEREUS AURISSETUS* group. The number of analyzed sequences (N), haplotypes (h), substitution sites (S), parsimoniously informative substitution sites (S informative), indels (Id), and parsimoniously informative indels (Id informative) are shown.

	<i>5'trnS-trnG</i>	<i>3'trnS-trnG</i>	<i>trnH-psbA</i>	<i>trnT-trnL</i>	<i>trnL</i>	<i>trnL-trnF</i>	<i>rps16-trnK</i>	<i>atpI-atpH</i>
Length (bp)	555	649	296	293	618	378	425	595
N/h	40/4	38/5	48/3	44/5	24/3	31/2	17/3	28/5
% variability	1,26	1,85	0,68	1,7	0,32	0,26	0,70	1,0
S/S informative	4/3	12/11	0	3/2	1/1	0	1/1	1/1
Id/Id informative	3/3	0	2/2	2/1	1/1	1/1	2/2	5/3
% GC	30,8	40,2	26,1	32,0	27,6	35,2	23,8	32,9
Taxonomic level variation	Sp/Pw	S/Pw/Pa	Sp/Pa	Sp/Pa	Sp	Sp	Sp	Sp/Pa

Notes: Sp = interspecies variation; Pw = within-population variation; Pa = among-population variation

The saturation test for all regions provided no significant results. The percentage of variability found in each region analyzed ranged from 0.26% in the *trnL-trnF* region to 1.85% in the *3'trnS-trnG* region. The main source of variability was nucleotide substitutions, despite the fact that the analyzed regions also presented indels, with the exception of *3'trnS-trnG* (Table 1). A total of eight variation sites were found in the *5'trnS-trnG* region consisting of four substitutions, two indels (9 bp and 10 bp) and two mononucleotide repeat variations (ranging from 1bp to 4bp), one of which was excluded from the analysis because of its ambiguous alignment. *3'trnS-trnG* showed a more variable nucleotide content than its adjacent region (*5'trnS-trnG*). Twelve variable sites were found in *3'trnS-trnG*, all of them substitutions. The variation in *trnH-psbA* region was characterized by two indels (11 bp and 5 bp). The intergenic spacer *trnT-trnL* showed three substitutions, one indel (10 bp) and a mononucleotide repeat (1bp). We found a substitution and an indel (10 bp) for *trnL* intron and only an indel (12 bp) for *trnL-trnF*. The *3'rps16-5'trnK* region showed one substitution and two indels (11 bp and 5 bp). Six polymorphic sites were observed in *atpI-atpH* (595 bp), one substitution, two small indels (11 bp and 12 bp), and three large indels (144 bp, 149 bp and 356 bp) which could be noted in agarose gels. In spite of these overlapped indels, they could be coded following the simple indel coding procedure of Simmons and Ochoterena (2000) and three of them were parsimoniously informative.

Five regions (*5'trnS-trnG*, *3'trnS-trnG*, *trnH-psbA*, *trnT-trnL* and *atpI-atpH*) showed variation in both population and species levels, while three regions (*trnL*, *trnL-trnF* and *3'rps16-5'trnK*) were variable only at interspecies level (Table 1). Although *3'trnS-trnG* was

the most variable region, the largest part of the variation was caused by differences between *P. aureispinus* and the remaining species. Variation within-population has been only observed for *5'trnS-trnG* and *3'trnS-trnG* in *P. aureispinus*, while among-population variation has been found for *3'trnS-trnG*, *trnH-psbA*, *trnT-trnL* and *atpI-atpH* in *P. machrisii*.

We selected the most variable regions (*5'trnS-trnG*, *3'trnS-trnG* and *trnT-trnL*) to construct the haplotype networks (Figure 1). When the regions were used separately they provided different topologies and no region was able to show the phylogenetic distinctiveness for all species. The phylogenetic resolution of the networks could be increased by including indels in the analysis. The most informative topologies were provided by *3'trnS-trnG* and *trnT-trnL* (Figure 1) regions. The *3'trnS-trnG* haplotypes distinguished the central-western (APA and CRI) from the southernmost (FOR and DEL) populations of *P. machrisii*, while *trnT-trnL* haplotypes distinguished two nearby *P. machrisii* populations in the central-western region (APA and CRI). However, the topology of *3'trnS-trnG* network provided a non-monophyletic grouping for *P. machrisii* haplotypes, as the same haplotype was shared by *P. aurisetus* and the southernmost populations of *P. machrisii* (FOR and DEL).

A higher informative topology was obtained by concatenating the three most variable regions (*5'trnS-trnG/3'trnS-trnG/trnT-trnL*). In this analysis we found seven haplotypes that resolved the species in distinct branches (Figure 1). *Pilosocereus machrisii* showed three haplotypes of which only one was shared by two nearby populations (FOR and DEL).

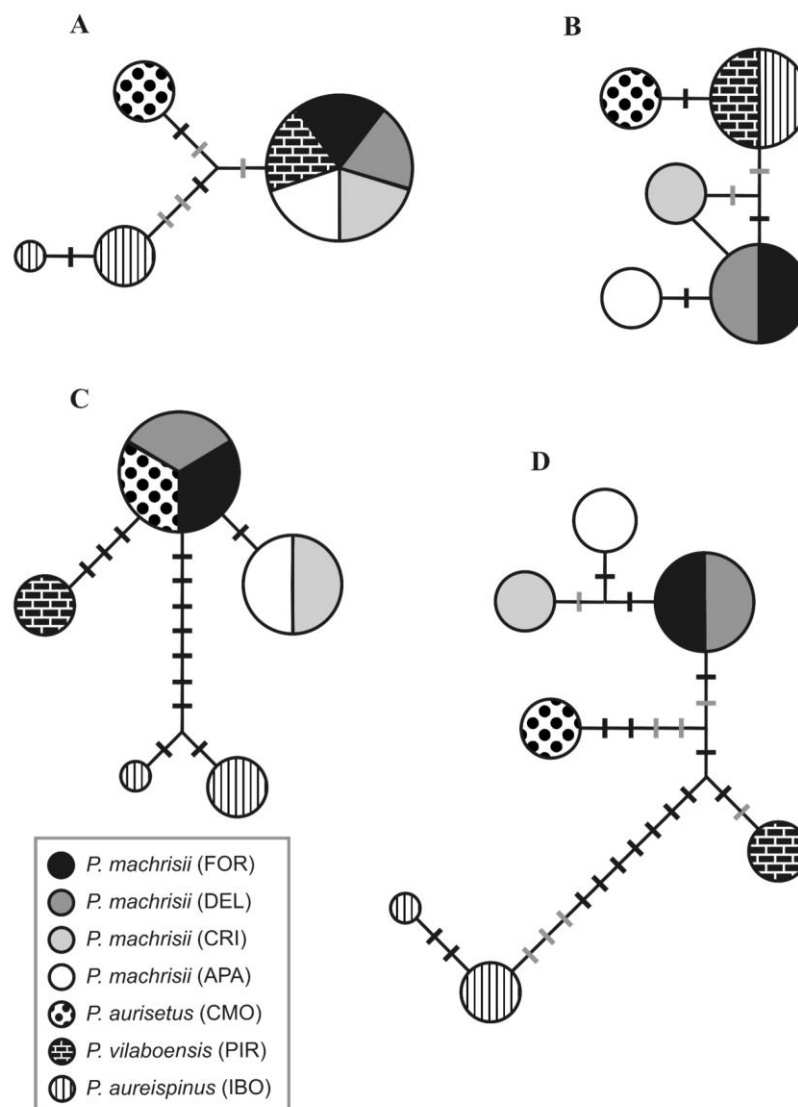


Figure 4 - Statistical parsimony network of the *P. AURISSETUS* species group based on noncoding cpDNA spacers 5' *trnS-trnG* (A), 3' *trnS-trnG* (B), *trnT-trnL* (C) and on the combination of these three regions (D). The circles sizes are proportional to the frequency of the haplotypes in the sample. Each species/populations was assigned according to the fill patterns and colors shown in the insert box. Black and grey crossed lines represent differences due to indels or substitutions, respectively.

In this study we showed that some widely used plastid regions, such as the intergenic spacers *trnH-psbA* and *trnL-trnF*, along with the *trnL* intron can be very conservative in related species. The *trnH-psbA* region has shown high occurrence of indels even in closely related species, as first observed by Aldrich *et al.* (1988) and corroborated by Shaw *et al.* (2005). According to those observations, all variations observed for *trnH-psbA* in our study were due to indels. *trnL* and *trnL-trnF* regions showed exiguous variation, while their adjacent region *trnT-trnL* seems to be more informative as noted by previous studies (see

Shaw *et al.*, 2005). We observed that the *trnS-trnG* intergenic spacer had the most variable sites in the 3' region portion. For *atpI-atpH*, we found the occurrence of large indels, an observation also noted by Shaw *et al.* (2007) in *Magnolia* (Magnoliaceae).

There is no consensus regarding the contribution of indels to cpDNA polymorphism, despite recent studies have showed that such source of variation may occur more frequently than substitutions (Ingvarsson *et al.*, 2003). As noted by Ravi *et al.* (2008), plastid genomes are generally rich in AT-content, as found in the species studied here (Table 1). This feature facilitates the occurrence of DNA secondary structure formation, which in turn contributes to the occurrence of indels, mainly the larger ones. Mononucleotide repeats of A or T also increase the chance of indels in these regions due to polymerase slippages in DNA replication. In our analysis indels were important in resolving haplotype relationships and as they are a common feature in cpDNA, we advocate that they should be used in phylogenetic analyses.

In the *Pilosocereus* species studied here, most intraspecific variation observed was due to differences among populations, while within population variation was only observed in *P. aureispinus*. Such pattern of variation is commonly observed in phylogeographic studies of plant species with a patchy distribution in the Neotropics, such as bromeliads, cacti, and *Petunia* (Solanaceae) (Barbará *et al.*, 2008; Helsen *et al.*, 2009; Lorenz-Lemke *et al.*, 2010). Probably, the lack of intrapopulation variation derives more from the distribution pattern and demographic history of these Neotropical species than from the molecular marker.

Our data reflect the difficulty of finding the appropriate molecular markers for intra- and interspecific studies in closely related plant species. No single region was able to provide a complete taxonomic resolution within the analyzed species when used separately, even the *trnH-psbA* region (data not shown), which has been elected as a barcode for plant species (Kress *et al.*, 2005).

Previous studies have shown that combining different cpDNA regions is more useful for differentiating species as well as increases the resolution and support for phylogenetic topologies (Wortley *et al.*, 2005; Miller *et al.*, 2009). In our study, the network topology provided by the three concatenated regions (5'*trnS-trnG*/3'*trnS-trnG/trnT-trnL*) agreed with morphological data that suggest *P. aureispinus* as the more differentiated species within the *P. AURISSETUS* group (Zappi, 1994). Some divergent morphological characters found in *P.*

aureispinus are stems with horizontal rings of bristles in the flowering zones (vs. longitudinally organized pseudocephalia) narrow (vs. infundibuliform) flower tube, relatively small fruits, and conic testa cells with coarse cuticular folds (vs. flat testa cells with less coarse cuticular folds) (Zappi, 1994). Furthermore, by using those three regions in combination differences among *P. machrisii* populations were revealed more effectively. However, even adopting such procedure we were unable to find within-population variation in most populations, an important source of genetic signals of historic demographic events for phylogeographic analyses.

Herein we have pointed to the potential of some cpDNA regions to help us resolve the relationships among closely related species and to perform population genetics studies in Cactaceae. Thus, based on the molecular data shown in this study, we suggest the use of the concatenated regions *5'trnS-trnG/3'trnS-trnG/trnT-trnL* in future studies with the *PILOSOCEREUS AURISSETUS* group, which certainly have an interesting evolutionary history to tell us.

Acknowledgments The authors thank H. Utsunomiya, F. Madia and M. Perez for laboratory assistance, M. Machado for assistance with the field sampling and F. F. Franco for valuable comments on the manuscript. This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP and Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq – Brasil.

References

- Aldrich J, Crnheey BW, Merlin E, Christopherson L (1988) The role of insertions/deletions in the evolution of the intergenic region between *psbA* and *trnH* in the chloroplast genome. *Curr Genet* 14: 137–146
- Avisé JC (2009) Phylogeography: retrospect and prospect. *J Biogeogr* 36: 3–15
- Barbará T, Lexer C, Martinelli G, Mayo S, Fay MF, Heuertz M (2008) Within-population spatial genetic structure in four naturally fragmented species of a neotropical inselberg radiation, *Alcantarea imperialis*, *A. geniculata*, *A. glaziouana* and *A. regina* (Bromeliaceae). *Heredity* 101: 285–296

-
- Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1660
- Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Mol Ecol* 8: 521–523
- Helsen P, Browne RA, Anderson DJ, Verdyck P, Dongen SV (2009) Galápagos' *Opuntia* (prickly pear) cacti: extensive morphological diversity, low genetic variability. *Biol J Linn Soc* 96: 451–461
- Hunt D, Taylor NP, Charles G (2006) *The New Cactus Lexicon, Atlas & Text*. dh Books, Milborne Port
- Ingvarsson PK, Ribstein S, Taylor DR (2003). Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. *Mol Biol Evol* 20: 1737–1740
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA* 102: 8369–8374
- Lorenz-Lemke AP, Togni PD, Mader G, Kriedt RA, Stehmann JR, Salzano FM, Bonatto SL, Freitas LB (2010) Diversification of plant species in a subtropical region of eastern South American highlands: a phylogeographic perspective on native *Petunia* (Solanaceae). *Mol Ecol* 19: 5240–5251
- Miller JS, Kamath A, Levin RA (2009) Do multiple tortoises equal a hare? The utility of nine noncoding plastid regions for species-level phylogenetics in tribe Lycieae (Solanaceae). *Syst Bot* 34: 796–804.
- Pleines T, Jakob SS, Blattner FR (2009) Application of non-coding DNA regions in intraspecific analyses. *Plant Syst Evol* 282: 281–294
- Ravi V, Khurana JP, Tyagi AK, Khurana P (2008) Review Article An update on chloroplast genomes. *Plant Syst Evol* 271: 101–122
- Rozas J, Sánchez-Delbarrio JC, Messeguer X, Rozas R (2003) DNAsp, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497

-
- Shaw J, Lickey E, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analyses. *Am J Bot* 92: 142–166
- Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am J Bot* 94: 275–288
- Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Syst Biol* 49 : 369 – 381
- Small RL, Ryborn JA, Cronn RC, Seelanan T, Wendel JF (1998) The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogenetic reconstruction in a recently diverged plant group. *Am J Bot* 85: 1301–1315
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol* 17: 1105–1109
- Taylor NP, Zappi DC (2004) *Cacti of Eastern Brazil*. Royal Botanic Gardens, Kew, England.
- Testolin R, Cipriani G (1997) Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in the genus *Actinidia*. *Theor Appl Genet* 94: 897–903.
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific, gaps penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-80
- Wang J, Wu YX, Ren GP, Guo QH, Liu JQ, Lascoux M (2011) Genetic Differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. *PLoS One* 6: e26530
- Wortley AH, Rudall PJ, Harris DJ, Scotland RW (2005). How much data are needed to resolve a difficult phylogeny? Case study in Lamiales. *Syst Biol* 54: 697–709

Xia X, Xie Z (2001) DAMBE: data analyses in molecular biology and evolution. *J Hered* 92: 371–373

Zappi DC (1994) *Pilosocereus* (Cactaceae) the genus in Brazil. *Succulent Plant Research* 3: 1-160

Zhang Q (2010) Why does biparental plastid inheritance revive in angiosperms? *J Plant Res* 123: 201–20