

WILLIAN SILVA BARROS

**FILOGEOGRAFIA MOLECULAR DE POAIA [*Psychotria
ipecacuanha* (Brot.) STOKES]**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2004

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

B277f
2004

Barros, Willian Silva, 1978-
Filogeografia molecular de poaia (*Psychotria
ipecacuanha* (Brot.) Stokes) / Willian Silva Barros. -
Viçosa : UFV, 2004.
x, 74f. : il. ; 29cm.

Inclui anexos.

Orientador: Luiz Orlando de Oliveira.
Dissertação (mestrado) - Universidade Federal de
Viçosa.

Referências bibliográficas: f. 52-63.

1. *Psychotria ipecacuanha* - Genética molecular. 2.
Psychotria ipecacuanha - População. 3. *Psychotria
ipecacuanha* - Filogenia. 4. *Psychotria ipecacuanha* -
Distribuição geográfica. 5. Evolução (Biologia).
I. Universidade Federal de Viçosa. II. Título.

CDD 20.ed. 583.520487328

WILLIAN SILVA BARROS

**FILOGEOGRAFIA MOLECULAR DE POAIA [*Psychotria
ipecacuanha* (Brot.) STOKES]**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 01 de março de 2004

Cosme Damião Cruz
(Conselheiro)

Jorge Abdala Dergam dos Santos
(Conselheiro)

Cláudio Horst Bruckner

Luiz Antônio dos Santos Dias

Luiz Orlando de Oliveira
(Orientador)

A minha filha Luana Fonseca Barros

Aos meus pais Vera Lúcia Silva Barros e Wilson Júlio de Barros

Aos meus irmãos Wallace Silva Barros e Vanessa Aparecida Silva Barros

Dedico

AGRADECIMENTOS

A Deus, pelas oportunidades que obtive.

A minha família, pela tolerância, pela paciência e pelo incentivo.

À Universidade Federal de Viçosa, em especial ao Programa de Genética e Melhoramento pela oportunidade de realizar e concluir este curso.

À CAPES, pelos 18 meses de bolsa concedidos.

Ao coordenador do Programa de Genética e Melhoramento, Cláudio Horst Bruckner, e às funcionárias Rita de Cássia Rosado Cruz e Maria da Conceição Lopes, pela amizade, pela dedicação e apoio ao longo de todo o curso.

Ao meu orientador Luiz Orlando de Oliveira, pela paciência, dedicação e amizade presentes em todas as etapas de minha formação.

Ao professor e amigo Pedro Crescêncio Souza Carneiro, por ter me estimulado a ingressar no Programa de Genética e Melhoramento.

Ao conselheiro Cosme Damião Cruz, pelos ensinamentos, pela amizade, pelo incentivo, pelas sugestões e por proporcionar um excelente convívio no Laboratório de Bioinformática.

Aos conselheiros Jorge Abdala Dergam dos Santos e Maurílio Alves Moreira, pelas sugestões e apoio.

Ao pesquisador Luiz Antônio dos Santos Dias, pela amizade e pelo incentivo durante o curso.

Aos Professores José Marcelo Soriano Viana, Aluizio Borém, Tuneo Sedyama, Paulo Roberto Cecon, Cosme Damião Cruz, Everaldo Gonçalves de Barros e Elza Fernandes de Araújo, pelos ensinamentos em suas disciplinas.

A todos os amigos e colegas dos Programas de Genética e Melhoramento, pelo convívio agradável durante a realização desse curso.

Aos amigos Eugênio Eduardo de Oliveira, Héder Henrique Ribeiro Faria, Mauro Henrique Pinheiro, Carlos Lásaro Pereira de Melo, Luciano da Costa e Silva, Lauro José Moreira Guimarães, Edgar Paulino de Sousa, Rodrigo Vieira da Silva, pelo convívio e demonstração de companheirismo desde 1998.

Ao amigo Rodrigo Barros Rocha, pela convivência, pelas palavras de incentivo, pelo companheirismo e pelos obstáculos vencidos juntos.

Aos amigos Márcio Mendes, Francismar, Cândida, Márcio Antônio, Taís Cristina, Andréia, Taís, Cíntia, Mariana, Gustavo, Wagner, Ricardo, James e Sidmar do Laboratório de Seqüenciamento, pela amizade, pelo apoio ao longo do trabalho.

Ao Professor Sérgio Hermínio Brommonschenkel e a Francis, Júlio e Klaus, pelo apoio na conclusão do trabalho.

A Roberta dos Santos Silva, pela indispensável ajuda na conclusão do trabalho.

Ao IEF, por permitir a realização dos estudos na Unidade de Conservação Parque Estadual do Rio Doce.

BIOGRAFIA

WILLIAN SILVA BARROS, filho de Vera Lúcia Silva Barros e Wilson Júlio de Barros, nasceu em 26 de outubro de 1978, em Belo Horizonte, Minas Gerais. Casado com Tatiana Chagas Fonseca Barros, pai de Luana Fonseca Barros.

Em 1996, formou-se em Técnico em Contabilidade pela Escola Estadual Padre José Maria de Man, em Coronel Fabriciano, MG. Em 1997, formou-se em Técnico em Eletrônica pelo Instituto Municipal de Educação Técnica de Timóteo (IMETT), Timóteo, MG. Em 1998, ingressou na Universidade Federal de Viçosa, Viçosa, MG, no curso de Agronomia, concluindo-o em setembro de 2002.

Iniciou o curso de Mestrado em Genética e Melhoramento em setembro de 2002, na Universidade Federal de Viçosa, defendendo tese no dia primeiro de março de 2004.

CONTEÚDO

RESUMO	ix
ABSTRACT	x
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	4
2.1. A Espécie em Estudo	4
2.2. DNA de Organelas	6
2.3. Filogeografia	9
2.3.1. Análise Filogeográficas	11
2.3.1.1. Estimação da Rede de Haplótipos	11
2.3.1.2. Análises dos Clados Aninhados ou “Nested Clade Analysis” e Teste por Associação Geográfica	13
3. MATERIAL E MÉTODOS	16
3.1. Material Genético	16
3.2. Extração e Quantificação do DNA Total	18
3.3. Reações de Amplificação Via PCR	19
3.4. Eletroforese em Gel de Agarose	21
3.5. Purificação e Quantificação	21
3.6. Seqüenciamento do DNA	22
3.7. Análise Filogeográfica	22
3.7.1. Estimação da Rede de Haplótipos	22
3.7.2. Análises dos Clados Aninhados ou “Nested Clade Analysis” e Teste por Associação Geográfica	23
4. RESULTADOS E DISCUSSÃO	25

5. CONCLUSÕES	51
6. REFERÊNCIAS BIBLIOGRÁFICAS	52
ANEXOS I	64
ANEXOS II	66

RESUMO

BARROS, Willian Silva, M.S., Universidade Federal de Viçosa, março de 2004.
Filogeografia Molecular de poaia [*Psychotria ipecacuanha* (brot.) stokes].
Orientador: Luiz Orlando de Oliveira. Conselheiros: Cosme Damião Cruz, Jorge Abdala Dergam dos Santos e Maurílio Alves Moreira.

Psychotria ipecacuanha é uma espécie medicinal da família Rubiaceae, conhecida popularmente como poaia e que se desenvolve em agregados perenes, denominados reboleiras, em áreas úmidas e sombrias do sub-bosque das matas tropicais do continente americano. O presente trabalho visou investigar a filogeografia da poaia. Para isto, seqüenciou-se o gene cloroplástico *trnT-trnL* de 40 indivíduos originários de 15 populações naturais. Nove destas populações foram identificadas na região Sudeste e seis na região Centro-Oeste do Brasil. Os dados de seqüenciamento possibilitaram a obtenção de um total de 959 bases. O alinhamento destas seqüências mostrou a existência de 13 sítios polimórficos entre elas, que diferenciaram oito haplótipos. Todos os indivíduos amostrados na região Centro-Oeste possuíam um único haplótipo, enquanto cada um dos indivíduos da região Sudeste foi alocado em um dos sete haplótipos restantes. A análise dos dados via “nested clade analyses” possibilitou a obtenção de uma rede única e se inferiu que os componentes desta rede estão agrupados em nível de cladograma total, de acordo com a região geográfica de procedência. A separação entre as populações do Sudeste e do Centro-Oeste segue um padrão que sugere fragmentação alopátrica, enquanto certas populações restritas ao Sudeste estão sofrendo restrição ao fluxo gênico por isolamento à distância. Na elaboração de um programa de conservação genética, as populações do Parque Estadual do Rio Doce, Itaperuna e Guaraciaba podem ser indicadas como prioritárias, pois além de apresentarem variação no número de haplótipos dentro da população, possuem também cinco haplótipos privados. Além disso, as três populações detêm juntas, 75% da variabilidade das linhagens do gene *trnT-trnL*.

ABSTRACT

BARROS, Willian Silva, M.S., Universidade Federal de Viçosa, March de 2004. **Molecular Phylogeography of poaia [*Psychotria ipecacuanha* (brot.) stokes]**. Advisor: Luiz Orlando de Oliveira. Committee Members: Cosme Damião Cruz, Jorge Abdala Dergam dos Santos and Maurílio Alves Moreira.

Psychotria ipecacuanha is a medicinal species of the Rubiaceae family, popularly known as poaia, which grows in perennial aggregates denominated clumps, in moist and shady areas of the tropical forests of the American continent. The present work aimed at an investigation of the phylogeography of *Psychotria ipecacuanha*. The sequence of the chloroplast gene *trnT-trnL* from 40 individuals of 15 natural populations was obtained for this purpose. Nine of these populations were identified in the southeast and six in the central-western area of Brazil. A total of 959 bases were obtained by the sequence data. These aligned sequences showed the existence of 13 polymorphic sites among them, which differentiated eight haplotypes. All individuals sampled in the central-western area had a single haplotype, while the individuals of the southeast area were each allocated in one of the seven remaining haplotypes. The analysis of the data via “nested clade analyses” made the achievement of a single net possible and it was inferred that the components of this net are grouped at the level of a total cladogram in agreement with the geographical area of origin. The separation among the populations of the southeast and of the center-west follows a pattern that suggests allopatric fragmentation, while certain populations limited to the southeast suffer restriction of gene flow through isolation by distance. In the elaboration of a program of genetic conservation, the populations from the Rio Doce State Park, from Itaperuna, and from Guaraciaba can be indicated as priority because besides presenting variation within the haplotype number of the population, they also have five private haplotypes. Furthermore, these three populations together hold 75% of the variability of the lineages of the *trnT-trnL* gene.

1. INTRODUÇÃO

Durante toda a sua história, a humanidade tem utilizado espécies vegetais para suprir inúmeras das suas necessidades diárias, sendo elas provedoras de alimentos, vestuários, abrigo, medicamentos, artesanatos, óleos, lubrificantes, borracha, resinas, ceras, perfumes, corantes, fibras etc. No entanto, estima-se que somente 10% das espécies vegetais tenham sido avaliadas com mais critérios no que se refere ao seu potencial de utilização na agricultura e na medicina. Portanto novas culturas e novos medicamentos ainda estão por ser descobertos (Prance, 1997).

Além disso, as regras que governam os padrões de distribuição das espécies sobre a superfície terrestre ainda são pouco entendidas. No entanto, já tem sido demonstrado claramente que algumas áreas de menor extensão podem conter maior biodiversidade, enquanto áreas de extensão muito maior podem ser bastante pobres em sua biodiversidade (Wilson, 1992). As florestas tropicais úmidas, por exemplo, embora estejam presentes em somente 7% da superfície do planeta, podem conter cerca de 50% das espécies (Prance, 1997).

Para proteger a biodiversidade das florestas e de seus fragmentos, é fundamental a utilização de estratégias de preservação e conservação eficientes. Vários estudos de genética de populações têm servido de suporte para o desenvolvimento dos critérios de conservação (Hoelzel et al., 1993; Menotti-Raymond e O'Brien, 1993; Gilpin, 1996). Estes

estudos são importantes na medida em que se permite quantificar a variabilidade genética das populações e sua organização no tempo e no espaço, quantificar ainda o fluxo gênico e a estrutura de acasalamento e conhecer a história evolutiva, a colonização, a expansão e a retração das populações. Além disso, os estudos genéticos possibilitaram detectar os efeitos da redução do tamanho efetivo da população e do isolamento entre populações, sendo possível inferir sobre a viabilidade de populações remanescentes (Leite, 2002).

A importância de se conhecerem os padrões da distribuição espacial e temporal da diversidade genética de uma dada unidade taxonômica está em permitir, por exemplo, que estratégias de conservação possam proteger a real diversidade representada pela espécie e, ao mesmo tempo, considerar as possíveis variações geográficas e o potencial evolucionário do táxon em estudo. A identificação das linhagens genealógicas permite também que seja estimado quanto da variabilidade genética existente está a salvo em uma unidade de conservação já implantada ou se haveria necessidade de expandir tal área ou estabelecer novas unidades.

O presente trabalho visou estudar a filogeografia molecular de *Psychotria ipecacuanha*, espécie popularmente conhecida por poaia ou ipeca. A poaia é reconhecida mundialmente como planta medicinal por sua ação expectorante, emética e amebicida (Pinto, 1976). Suas propriedades farmacológicas são atribuídas aos alcalóides emetina, cefelina e psicotrina, encontrados em suas raízes secas (Bruneton, 1995). Os indivíduos de poaia se desenvolvem em agregados perenes, denominados reboleiras, em áreas de sub-bosque das matas tropicais da região neotropical do continente americano que recebem baixos níveis de radiação luminosa (Velo, 1947).

Esse estudo teve como objetivo: (1) Identificar quais são as diferentes linhagens genealógicas do fluxo gênico materno, que constituem as populações de *Psychotria ipecacuanha* incluídas neste estudo. (2) Estudar a existência de variabilidade de

linhagens genealógicas dentro de população. (3) Verificar quais dessas populações possuem haplótipos privados ou raros. (4) Verificar a existência de algum evento na história dessas populações, que possa ser detectado pelo teste de análise de clados aninhados (“nested clade analysis”), seguindo a “teoria da coalescência”. (5) Averiguar qual a implicação desses estudos para a montagem de um programa de conservação genética da espécie.

2. REVISÃO DE LITERATURA

2.1. A Espécie em Estudo

A *Psychotria ipecacuanha*, conhecida no Brasil vulgarmente como poaia, poaia-roxa, ipeca ou ipecacuanha (Figura 1), é um subarbusto perene de altura entre 20 a 60 cm e de crescimento muito lento. Sua raiz tem sabor amargo e cheiro nauseante, possuindo anéis salientes próximos uns dos outros, envolvendo-a completamente (Bruneton, 1995).

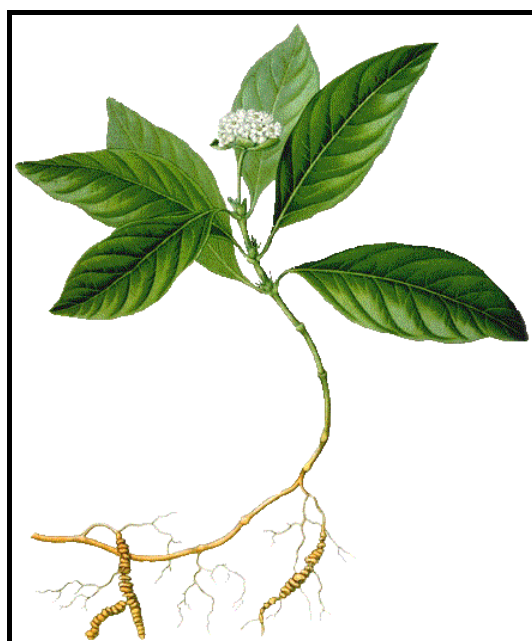


Figura 1 - Planta adulta de poaia (*Psychotria ipecacuanha*)

A poaia é encontrada em agregados, conhecidos popularmente como reboleira, que crescem naturalmente em lugares úmidos e sombrios, no sub-bosque de florestas tropicais. A propagação natural da poaia ocorre por multiplicação vegetativa, principalmente a partir do rizoma. Então, cada reboleira pode ser representada por um único indivíduo, pois a propagação sexuada dessa espécie é muito restrita (Rossi, 2003). Suas raízes são mundialmente reconhecidas por suas propriedades expectorante, amebicida e “vomitiva”, propriedades devidas ao acúmulo, em suas raízes, dos alcalóides isoquinolínicos, emetina e cefelina (Trease e Evans, 1989). Atualmente, populações silvestres de poaia podem ser encontradas em três grandes regiões disjuntas (Figura 2): a) florestas da América Central (Nicarágua, Costa Rica e Panamá) e norte da América do Sul (Colômbia), b) sul da Floresta Amazônica (nos Estados de Rondônia e Mato Grosso) e c) na Mata Atlântica (principalmente nos estados da Bahia, Minas Gerais, Espírito Santo e Rio de Janeiro) (Skorupa e Assis, 1998; Oliveira e Martins, 1998; Assis e Giuliett, 1999).



Figura 2 - Regiões de ocorrência de populações nativas da poaia (destaque em vermelho). Fonte: Assis adaptado por Oliveira e Martins (1998).

No passado, as raízes de poaia atingiram grande importância medicinal, sendo a demanda do mercado externo atendida, basicamente, por meio de intenso processo de extrativismo (Addor, 1945; Oliveira e Martins, 1998). Atualmente, a poaia consta na lista das espécies ameaçadas de extinção no Estado de Minas Gerais (Mendonça e Lins, 2000). Estudo de erosão genética permitiu a identificação dos principais fatores de risco e a formulação de recomendações específicas para a proteção e manejo de suas populações na Floresta Atlântica (Oliveira e Martins, 2002). Porém, a implementação de ações efetivas de conservação depende da diversidade genética e do entendimento dos processos ecológicos e históricos que moldaram a atual distribuição da diversidade genética da poaia, e é precisamente a caracterização desses processos que pode servir de guia para a elaboração e implementação de programas de conservação.

2.2. DNA de Organelas

Nos últimos anos, o DNA de organelas tem sido importante em diversos trabalhos relacionados a estudos evolutivos. Pelo fato de apresentar normalmente herança uniparental, permite analisar a estrutura populacional (e.g. Comes e Abbott, 1998; Levy e Neal, 1999; Matocq, 2002), discernir diferentes padrões de migração (e.g. Foster et al., 1996), estabelecer ou examinar relações filogenéticas (e.g. Maguire et al., 1997; Kim e Jansen, 1998; Schnabel e Wendel, 1998; Garcia-Moreno et al., 1999; Salomone et al., 2002) e comparar taxas evolutivas entre linhagens (Wolfe et al., 1987). De acordo com McCauley (1994), os marcadores genéticos de genomas de organelas são poderosas ferramentas para a detecção de diferenciação populacional porque são mais sensíveis a efeitos de deriva genética que os genomas nucleares. Isso se deve ao fato de o tamanho

efetivo da população para genoma de organela ser reduzido, se comparado com o do genoma nuclear.

Alguns trabalhos têm sido desenvolvidos com DNA de organelas, seja analisando o modo de herança em determinadas espécies (Hachtel, 1980), seja demonstrando sua importância em análise de estrutura genética, fluxo gênico e de biogeografia de populações, para a tomada de decisões relativas à conservação de espécies (Barrowclough et al., 1999). Pares de iniciadores (primers) de regiões conservadas de DNA de cloroplasto (cpDNA) são atualmente disponíveis para gerar mais ferramentas que possam fornecer auxílio para estudos de evolução e de identificação genética entre indivíduos (Taberlet et al., 1991).

O cloroplasto é uma organela celular que apresenta seu próprio genoma, cuja transmissão segue “um modelo não-mendeliano”. Contém um único tipo de cromossomo, embora em múltiplas cópias, no qual estão contidos genes necessários para sua função. O genoma do cloroplasto de plantas terrestres é uma molécula circular de 120 a 217 Kb (Palmer, 1985, 1987), caracterizada por dois segmentos invertidos e repetidos que separam o restante da molécula em pequena e larga região de cópia única. Atualmente existe, em estudos de genética de populações de plantas, crescente interesse no uso do genoma de cloroplasto (e.g. Soltis et al., 1989; McCauley, 1994; Comes e Abbott, 1998; Provan et al., 1999; Caron et al., 2000; Gugerli et al., 2001).

Quando o cpDNA é herdado matematicamente, como na maioria dos casos das Angiospermas (Corriveau e Coleman, 1988; Sears, 1980), a dispersão é promovida pelas sementes e propágulos vegetativos sendo a taxa de migração potencialmente muito menor que a de migração dos genes nucleares (McCauley, 1995).

De acordo com Dumolin-Lapèque et al. (1997), quando o cpDNA apresenta herança materna, é possível um estudo direto de dispersão mediada por semente, possibilitando, portanto, inferir rotas de colonização.

O genoma de cloroplasto é efetivamente haplóide, enquanto o nuclear, diplóide ou poliplóide. Assim, em espécies monóicas, o tamanho efetivo da população em genomas nucleares é no mínimo duas vezes o genoma de cloroplasto (Schaal et al., 1998). Como o efeito da deriva genética é maior à medida que o tamanho efetivo da população diminui, a deriva genética terá impacto maior na formação da estrutura populacional do genoma de cloroplasto que dos genes nucleares (Birky et al., 1983). Assim, a árvore de genes baseada no cpDNA revela mais provavelmente, padrões de diferenciação populacional e refletem, por certo, menos polimorfismo ancestral que árvores de genes oriunda do DNA nuclear (Schaal et al., 1998).

Devido ao seu modo clonal de transmissão e à baixa taxa de mutação (Wolfe et al., 1987), as variações encontradas em suas seqüências possibilitam revelar eventos históricos, como rotas de migração e afunilamento de populações (Dumolin-Lapègue et al., 1997).

As variações na seqüência do DNA de cloroplasto são amplamente utilizadas para investigar relações inter-específicas entre angiospermas e outras plantas (Taberlet et al., 1991), existindo porém, uma série de limitações na utilização dessa molécula. Assim, as regiões não codificantes exibem alta freqüência de mutações (Clegg et al., 1991), podendo ser, portanto, usadas para estudos evolutivos em espécies de relações muito próximas ou para análises intra-específicas. De acordo com Palmer et al. (1989), as mutações em cpDNA são basicamente de dois tipos: mutações pontuais e rearranjos. A maioria das mutações é pontual, sendo comum também deleções e/ou inserções de bases em regiões não-codificantes. Apesar das variações nas seqüências de cpDNA serem poucas e menores que em mtDNA (Wolfe et al., 1987), aquelas em nível intra-específico têm sido encontradas e demonstradas em um crescente número de espécies (Soltis et al., 1991; Soltis et al., 1992; Ferris et al., 1993; Leite, 2002).

2.3. Filogeografia

Muitas vezes, uma nova palavra ou frase envolvendo conceitos torna-se parte do léxico de funcionamento de uma disciplina ou áreas de conhecimento, sendo exemplo o caso da “biodiversidade”, aceita no tratado ecológico, no qual promove a noção de variabilidade biológica como um recurso apreciado (Wilson, 1988). Contudo, a palavra filogeografia foi introduzida no meio científico mais por necessidade do que um novo modelo ou um tratado. A filogeografia é considerada uma ciência multidisciplinar, que integra a genética molecular, genética de populações, etologia, demografia, filogenia, paleontologia e geologia (Avise, 2000). Pode ser definida como um campo de estudo interessado nos princípios e processos que governam as distribuições geográficas de linhagens genealógicas, incluindo aquelas em nível intra-específico (Avise et al., 1987). A análise da distribuição espacial da genealogia de genes forma a base da filogeografia (Schaal et al., 1998), na qual os indivíduos são genotipados e relacionados com padrões de distribuição geográfica (Weir, 1996). A ecogeografia e a filogeografia são subáreas da biogeografia. A filogeografia se coloca num contexto temporal com perspectivas mais amplas do que a ecogeografia tradicional, que enfatiza o papel de pressões ecológicas contemporâneas, moldando as distribuições no espaço das características dos organismos. Porém, ecogeografia e filogeografia não são mutuamente exclusivas, servindo a filogeografia para expandir e equilibrar as perspectivas ecogeográficas (Thorpe et al., 1995).

A filogeografia é uma ferramenta poderosa para estudos evolutivos, além de permitir a detecção de áreas que apresentam indivíduos com os mesmos ou diferentes padrões genéticos (Avise, 1998; Leite, 2002). Assim, oferece maneiras de examinar a

história da mudança genética entre populações, distinguindo padrões biogeográficos de variação genética causados por fluxo gênico, daqueles provocados por ancestrais comuns (Schaal et al., 1998). Além disso, possibilita a detecção de eventos históricos e recentes de fluxo gênico, bem como o discernimento do papel do fluxo gênico na estruturação de populações de plantas, superando as limitações básicas da genética clássica de população e sistemática, uma vez que resume a influência da história e da troca gênica, oferecendo a possibilidade de discernir relações de diferenciação fenotípica e isolamento genético (Schaal et al., 1998). Os estudos filogeográficos permitem que a dinâmica espacial e temporal de organismos estudados possa ser alcançada a partir da comparação entre a identificação da genealogia das variantes e suas distribuições geográficas (Dumolin-Lapègue et al., 1997).

A filogeografia tem sido utilizada com o intuito de analisar a estrutura genética de populações de plantas e animais. Tendo como base a interpretação de padrões de congruência ou falta de congruência entre a distribuição geográfica dos haplótipos e suas relações genealógicas (Schaal et al., 1998), a filogeografia vem fazendo parte de alguns trabalhos, seja para fornecer suporte para decisões com fins conservacionistas (e.g. Barrowclough et al., 1999; Collevatti et al., 2003), seja para confirmar a divisão de gênero das espécies, verificar as relações filogenéticas e reconstruir a história dos possíveis eventos de dispersão (e.g. Lessios et al., 1999), para verificar o possível papel histórico de rios como barreiras geográficas na diversificação de espécies (Gascon et al., 2000; Leite, 2002) ou para inferir rotas de colonização de espécies florestais (e.g. Dumolin-Lapègue et al., 1997). Embora os estudos filogeográficos em animais tenham sido mais numerosos que em vegetais, Schaal et al. (1998) não consideram que a causa principal da escassez desse tipo de estudo em plantas ocorra porque a filogeografia seja menos aplicável ou útil no reino vegetal.

2.3.1. Análises Filogeográficas

2.3.1.1. Estimação da Rede de Haplótipos

Templeton et al. (1992, 1995) apresentaram uma explicação detalhada para a construção de redes de haplótipos, usando o método da parcimônia do modelo local-finito da evolução do DNA. Este método, além de estimar a rede de haplótipos, permite calcular o número máximo de conexões parcimoniosas entre pares de seqüências com 95% de probabilidade em todos os acoplamentos dos haplótipos dentro da rede. O “software” TCS (versão 1.13), desenvolvido por Clement et al. (2000), pode ser usado para a construção das redes de haplótipos.

Um exemplo da rede de haplótipos foi apresentado na Figura 3, em que a barra de ligação entre os haplótipos representa uma conexão parcimoniosa, correspondendo a um único evento de mutação (Figura 3A). De acordo com Templeton et al. (1988), os nodos intermediários são deduzidos entre os dois pares de seqüências diferentes mais próximos da rede que diferem entre si em dois ou mais eventos de mutação (Figura 3B). O haplótipo I mostrado na Figura 3C apresenta uma multifurcação, conectando-se com outros oito haplótipos. Com base na teoria da coalescência, esse haplótipo, em média, é o mais antigo numa escala evolutiva, em relação aos haplótipos que se encontram nas extremidades da rede.

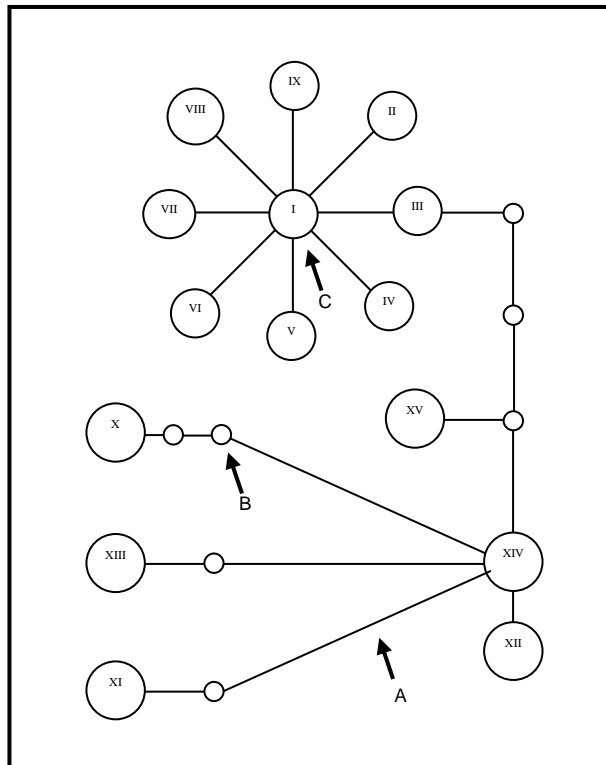


Figura 3 - Exemplo demonstrativo de uma rede de 15 haplótipos gerada pelo “software” TCS (Versão 1.13) de dados de seqüência de mtDNA de Lagarto (*Lacerta shreiberi*). Os números em romano nomeiam os haplótipos. A - Uma barra de conexão parcimoniosa entre os haplótipos mais próximos correspondendo a um único evento de mutação; B - Um nodo intermediário que não foi amostrado no momento, isto é, haplótipos são deduzidos entre os pares de seqüências mais próximos; C - Uma multifurcação derivada do haplótipo I. Fonte: Adaptado de Paulo et al. (2002).

2.3.1.2. Análises dos Clados Aninhados ou “Nested Clade Analysis” e Teste por Associação Geográfica

Depois que a genealogia da rede de haplótipos foi estimada com 95% de probabilidade, os clados de níveis inferiores são aninhados em clados de nível superior até envolver toda a rede de haplótipos, dando origem ao cladograma total. Um exemplo

da rede aninhada é demonstrado na Figura 4. Todos os aninhamentos construídos seguiram as regras propostas por Templeton et al. (1987, 1995) e Templeton e Sing (1993), sendo no primeiro passo, os haplótipos considerados clados de nível zero. Então, o primeiro passo foi aninhar os haplótipos da extremidade separados por um único evento de mutação, agrupados com linha cheia para formar clados de primeiro nível (Figura 4). Em seguida, foram agrupados os haplótipos dos interiores separados entre si por um evento de mutação, em relação aos haplótipos das extremidades. Depois, os clados de primeiro nível foram aninhados em clados de segundo nível, agrupados com linha tracejada (Figura 4).

Uma vez determinados os aninhamentos dos clados, pode ser feito um teste simples para associação geográfica, tratando, então, cada local da amostra (população) como uma variável, através de uma análise exata de contingência permutacional pelo teste de χ^2 (Roff e Bentzen, 1989; Hudson et al., 1992; Templeton et al., 1995). Esse teste de contingência pode ser realizado pelo “software” GEODIS desenvolvido por Posada et al. (2000) nos clados com mais de um haplótipo, seguindo o algoritmo dado por Templeton e Sing (1993). Isso poderá evidenciar interpretações filogeográficas mais ajustadas, que permitirão distinguir entre a história e a estrutura da população (Templeton e Sing 1993; Templeton et al., 1995; Crandall e Templeton, 1996; Templeton, 1998).

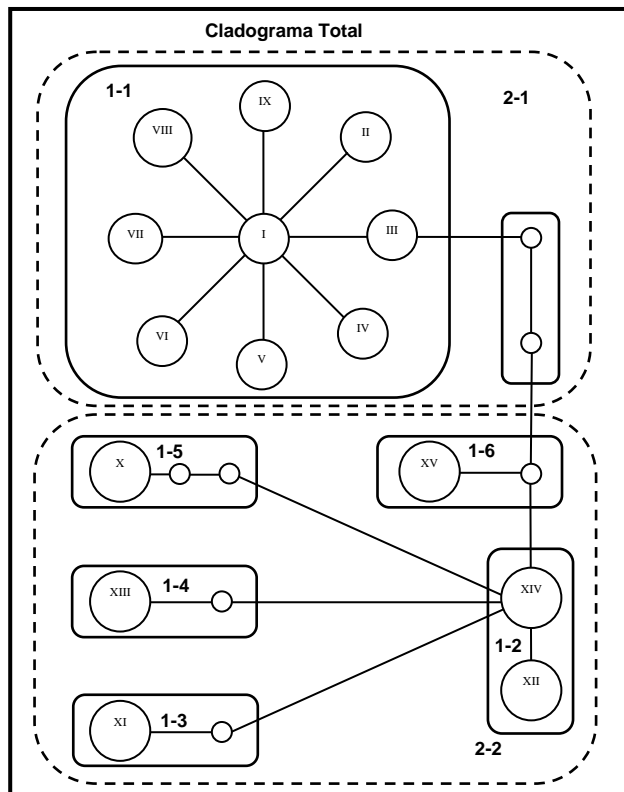


Figura 4 - Exemplo demonstrativo do aninhamento dos 15 haplótipos de mtDNA de Lagarto (*Lacerta shreiberi*) deduzido de acordo com as regras propostas por Templeton et al. (1987, 1995) e Templeton e Sing (1993). Os haplótipos são considerados clados de nível zero. Os clados de primeiro nível (1-1 até 1-6) foram aninhados por linhas cheias e os clados de segundo nível (2-1 e 2-2), por linhas tracejadas. O cladograma total foi a união dos dois clados de segundo nível agrupados por um retângulo. Fonte: Paulo et al. (2002).

Esse teste de associação geográfica usa duas medidas de distâncias relativas. A primeira é a distância entre um determinado clado e um outro clado específico “X”, simbolizada por $D_C(x)$. Essa distância mede como os indivíduos são difundidos geograficamente, amparados pela distância dos haplótipos originados do clado X (veja Templeton et al., 1995, para explicação adicional em termos de operações matemáticas). A segunda distância estimada é a distância entre um clado aninhado e o clado X, simbolizada por $D_N(X)$. Esta distância não mede como o clado X é difundido

geograficamente, mas como os indivíduos estão distantes do haplótipo do clado X, em relação a todos os indivíduos que estão ligados aos haplótipos do clado Y. As distribuições dessas duas medidas são determinadas, recalculando-se ambas com no mínimo 1.000 permutações aleatórias para se fazerem inferências estatísticas com o nível de 5% de significância (Edgington, 1986).

O procedimento randômico permite ao investigador testar as distâncias significativamente grandes e pequenas, em D_C e D_N , para cada clado dentro de um grupo aninhado de clados, que considera a hipótese nula de nenhuma associação geográfica dentro do clado aninhado (Templeton et al., 1995). Uma estatística das diferenças das médias das distâncias geográficas dos clados interiores com os clados das extremidades também é calculada dentro de cada categoria aninhada. Para o cálculo dessas médias, cada distância do clado é ponderada pelo número de cópias naquele clado relativo focado para o número total de cópias no clado aninhado. Esta extremidade vs. interior corresponde ao contraste novo vs. velho, podendo estender-se ao raro vs. comum (Crandall e Templeton, 1993). Assim, fica possível deduzir interpretações filogeográficas de valores significativos para D_C e D_N , usando a chave de inferências proposta por Templeton (2004).

3. MATERIAL E MÉTODOS

3.1. Material Genético

Expedições de coleta de germoplasma empreendidas entre 1995 e 2000 permitiram a montagem de um banco de germoplasma com indivíduos que representam mais de 70 reboleiras de *Psychotria ipecacuanha* (Martins, 2000; Oliveira e Martins, 2002). Esses indivíduos, na sua maioria, são provenientes de reboleiras originárias do Sudeste do Brasil e cobrem toda a área de ocorrência da espécie nessa região. De forma preliminar, uma população foi definida como um conjunto de reboleiras, que podem ser encontradas em um mesmo fragmento florestal. No banco de germoplasma, cada reboleira está representada por um único indivíduo que foi transplantado do campo e é mantido em vaso, sob condições de sombreamento. O banco de germoplasma encontra-se mantido no “campus” da UFV. A posição de cada população localizada no Sudeste brasileiro foi georeferenciada por meio de coordenadas geográficas (GPS II Plus, Garmin Corp. USA), tomadas durante as expedições de coleta e, assim, a localização original de cada indivíduo foi conhecida. Porém, essas coordenadas geográficas não serão apresentadas neste estudo, mas estão disponíveis para consulta, se requisitadas ao curador do banco de germoplama.

O presente estudo filogeográfico utilizou um total de 40 indivíduos originários de 15 populações naturais de *Psychotria ipecacuanha*. A caracterização desse material quanto à sua região de origem e localização é apresentada na Tabela 1. Foram utilizados 26 indivíduos de nove populações do Sudeste brasileiro, amostrados no banco de germoplasma, e 14 indivíduos de seis populações provenientes da Região Centro-Oeste. As posições geográficas dos indivíduos localizados no Centro-Oeste não foram georeferenciadas no momento da coleta, por isso as coordenadas geográficas foram estimadas de acordo com as informações fornecidas pelos coletores.

Tabela 1 - Identificação dos locais onde foram amostrados indivíduos de *Psychotria ipecacuanha*, separados pelas Regiões Sudeste (SE) e Centro-Oeste (CO), com o respectivo código da população e sua localização

Região	Código	Localização
SE	UNA	UNA, BA
	PRD	Parque Estadual do Rio Doce, Marliéria, MG
	PTN	Ponte Nova, MG
	GUA	Guaraciaba, MG
	VRB	Visconde do Rio Branco, MG
	CAR	Carangola, MG
	IRU	Irupi, ES
	ITA	Itaperuna, RJ
	CON	Conceição do Macabu, RJ
CO	BBU	Barra do Bugres, MT
	FRA	Fazenda Raizama, Barra do Bugres, MT
	AIN	Área 1, Barra do Bugres, MT
	MAN	Manilha, Barra do Bugres, MT
	CUR	Curvelândia, MT
	SAC	Salto do Céu, MT

3.2. Extração e Quantificação do DNA Total

De cada indivíduo, foram coletadas três folhas expandidas para obtenção do DNA genômico total. O DNA genômico foi extraído de material foliar fresco de cada indivíduo, seguindo-se o método descrito por Doyle e Doyle (1990). Na extração, foram utilizados em torno de 50 mg de tecido de cada um dos indivíduos. O material foi macerado na presença de nitrogênio líquido, objetivando romper as paredes e membranas celulares do tecido. Logo em seguida, ao tecido recém-macerado foram adicionados 600 µl de tampão de extração contendo: CTAB 2%; NaCl 1,4 M; EDTA 20 mM, pH 8,0; Tris-HCl 100 mM, pH 8,0; PVP 4% e β-mercaptoetanol 2,5%, previamente aquecidos, visando a solubilização de membranas lipoprotéicas, a desnaturação de proteínas e a proteção do DNA contra a ação de enzimas de degradação. Depois de homogeneizado, o material foi incubado em banho-maria a uma temperatura de 65 °C, em torno de 60 minutos, agitado a cada 10 minutos, visando a solubilização e homogeneização da suspensão. Em capela de exaustão, foi feita a primeira extração com solvente orgânico, adicionando-se 700 µl de CAI (clorofórmio - álcool isoamílico 24:1), a fim de separar lipídios, proteínas e polissacarídeos dos ácidos nucleicos (DNA e RNA). Os tubos foram agitados até que a emulsão ficasse homogênea e, em seguida, foram centrifugados a 13.000 rpm, em uma microcentrífuga por cinco minutos, visando a separação da fase orgânica (inferior), enquanto DNA, RNA e alguns polissacarídeos ficaram retidos na fase aquosa (superior). Transferiu-se toda a fase superior (aquosa) de cada tubo para um novo tubo correspondente. À fase aquosa no novo tubo correspondente, foram adicionados novamente 700 µl de CAI, agitando-se e centrifugando-se por mais cinco minutos a 13.000 rpm. Foram transferidos 250 µl do sobrenadante para tubos novos e foram adicionados 20 µl de acetato de amônia (7,5 M). Em seguida, foram acrescentados 146 µl de isopropanol, tendo os tubos sido mantidos em repouso durante uma hora, a -20 °C,

para precipitação do DNA e RNA. Após, foram centrifugados a 13.000 rpm por cinco minutos para sedimentar o “pellet” (precipitado). Descartou-se o sobrenadante e o precipitado foi lavado duas vezes com 500 µl de etanol 70% e uma vez com 500 µl etanol 95%. Deixou-se o tubo contendo o “pellet” secar à temperatura ambiente. Depois, o precipitado de DNA foi diluído em 40 µl do tampão TE (Tris-HCl 10 mM, pH 8,0 e EDTA 1 mM, pH 8,0) e mantido a –20 °C, até uso posterior.

A quantidade de DNA total foi estimada, comparando-se com DNA de quantidade conhecida de DNA do fago λ, por eletroforese em gel de agarose 0,8%, contendo 0,2 µg/mL de brometo de etídio.

3.3. Reações de Amplificação Via PCR

A amplificação da região *trnT-trnL* do genoma de cloroplasto foi conduzida com os pares de iniciadores A + D, descritos por Taberlet et al. (1991) e cujas seqüências estão representadas no Quadro 1. Nas amostras que falharam na amplificação ou cujos produtos possuíam baixa qualidade com o uso dessa combinação de iniciadores (A + D), a estratégia de amplificação foi redesenhada, de tal modo que o fragmento foi amplificado, usando-se duas reações independentes. A primeira reação foi conduzida com a combinação A + B, que tem como alvo, amplificar a região espaçadora intergênica entre o exon 5' de *trnT* (UGU) e *trnL* (UAA) (Taberlet et al., 1991). Paralelamente, uma segunda reação foi conduzida com a combinação C + D e se amplificou o íntron *trnL* (UAA) (Taberlet et al., 1991). Essas reações independentes resultaram em dois fragmentos, cuja seqüência de bases contém uma informação similar àquela obtida do fragmento originário da reação conduzida com os iniciadores A + D.

Quadro 1 - Seqüências de quatro iniciadores universais desenhados para amplificar segmentos conservados do gene *trnT-trnL* do genoma cloroplastídico de plantas superiores

Iniciador	Seqüências (5' → 3') – 20 bases
A – forward	CATTACAAATGCGATGCTCT
B – reverse	TCTACCGATTTGCCATATC
C – forward	CGAAATCGGTAGACGCTACG
D – reverse	GGGGATAGAGGGACTTGAAC

Fonte: Taberlet et al. (1991).

As reações de amplificação foram conduzidas no termociclador da APPLIED BIOSYSTEMS, modelo 9.600, com um programa do tipo touch down, com período inicial de desnaturação a 95 °C por cinco minutos, seguido por sete ciclos (94 °C por um minuto, 58 °C por um minuto, 72 °C por dois minutos) e 33 ciclos (94 °C por um minuto, 52° C por um minuto, 72 °C por dois minutos) e um período adicional de polimerização a 72 °C por cinco minutos. As reações foram realizadas em volume final de 25 µl, contendo 2,5 mmoles de cada dNTP, 0,25 U de *Taq* DNA polimerase, 12,5 pmoles de cada iniciador, 50 ng DNA molde, 1,25 µl dedimetil-sulfóxido (DMSO) e 2,5 µl do tampão “soltis buffer” contendo: KCl 0,5 mM; Tris-HCl 0,2 mM, pH 8,0; MgCl₂ 15 µM e 0,1% de Tween-20. Cada reação de amplificação foi replicada quadro vezes, para obter uma quantidade final do produto de PCR igual ou superior a 50 ng/µl, necessários para o seqüenciamento desse fragmento.

3.4. Eletroforese em Gel de Agarose

Para visualizar os fragmentos amplificados, foi utilizada eletroforese em gel de agarose 1%, contendo 0,2 µg/mL de brometo de etídio. Então, foram aplicados no gel de agarose, 5 µl do produto da reação de PCR com 3 µl de corante tipo IV (0,25% de azul de bromofenol e 40% de sacarose) para cada amostra e submetido a uma tensão constante de 80 Volts em tampão TEB 1X (Tris-borato 0,09 M e EDTA 0,002 M) por duas horas. Após a corrida, os géis foram visualizados por meio de luz ultravioleta e suas imagens foram digitalizadas no sistema de vídeo Eagleye™ (Stratagene®).

3.5. Purificação e Quantificação

Antes da etapa de seqüenciamento, os produtos da reação de amplificação foram purificados, utilizando-se colunas da “QIAquick” (QIAGEN), para retirar iniciadores e dNTPs não incorporados durante a reação. Para quantificar o DNA purificado, visando utilizar 40 ng/µl a 50 ng/µl de DNA para as reações de seqüenciamento, foi utilizada eletroforese em gel de agarose 1% (contendo 0,2 µg/mL de brometo de etídio). Em seguida, foram aplicados no gel de agarose, 1 µl de DNA purificado e 2 µl de corante tipo IV (0,25% de azul de bromofenol e 40% de sacarose) e 9 µl de água para cada amostra. Essa solução foi submetida a uma tensão constante de 80 Volts em tampão TEB 1X (Tris-borato 0,09 M e EDTA 0,002 M), por duas horas. Após a corrida, os fragmentos amplificados contidos nos géis foram visualizados por meio de luz ultravioleta, sendo suas imagens digitalizadas no sistema de vídeo Eagleye™ (Stratagene®).

3.6. Seqüenciamento do DNA

As reações de seqüenciamento dos fragmentos amplificados foram realizadas por seqüenciadores automatizados, utilizando-se os iniciadores A, B, C ou D, dependendo da região desejada. O seqüenciamento foi conduzido nas dependências da Agrogenética e do Laboratório de Genômica, localizados no BIOAGRO (Viçosa, MG), utilizando os seqüenciadores automatizados “ABI Prism 377 Genetic Analyzer” da APPLIED BIOSYSTEMS e “MegaBace DNA Analysis System 500” da AMERSHAM BIOSCIENCES CORP, respectivamente. As seqüências obtidas dos indivíduos foram importadas para o aplicativo computacional SEQUENCHER versão 4.0.5 e submetidas a correções manuais e edição. Finalmente, as seqüências completas dos indivíduos foram alinhadas com a introdução de “gaps” (quebra da continuidade da seqüência de DNA), de forma a compensar “indels” (inserção ou deleção de bases). As extremidades das seqüências foram eliminadas de maneira a obter uma região comum para todas as 40 seqüências dos indivíduos.

3.7. Análise Filogeográfica

3.7.1. Estimação da Rede de Haplótipos

O haplótipo completo do genoma cloroplastídico de um dado indivíduo é resultante da junção das seqüências nucleotídicas dos segmentos AB e CD. Para evidenciar a relação de parentesco entre os pares de seqüências dos indivíduos, foi utilizado o programa TCS (Versão 1.13), desenvolvido por Clement et al. (2000). Então, foi obtida uma rede de haplótipos, que estimou a genealogia entre as 40 seqüências obtidas dos indivíduos. Esse programa aplica a estatística da parcimônia pelo algoritmo desenvolvido

por Templeton et al. (1992). Quando o programa foi alimentado com as seqüências dos indivíduos, o número máximo de conexões parcimoniosas entre os pares de seqüências com 95% de probabilidade para todos os acoplamentos dos haplótipos dentro da rede foi estimado. Na estimação da rede de haplótipos, os “indels” foram considerados como quinto caráter (as bases A, T, C e G são os outros caracteres), indiferentemente do tamanho do “gap”, sendo cada “indel” considerado um único evento de mutação.

3.7.2. Análises dos Clados Aninhados ou “Nested Clade Analysis” e Teste por Associação Geográfica

Com o algoritmo descrito por Templeton et al. (1987) e Templeton e Sing (1993), foi possível aninhar passo a passo, os clados de acordo com seu nível, até envolver toda a rede de haplótipos, gerando o cladograma total. O aninhamento começa pelas extremidades da rede de haplótipos, separados por um evento de mutação. Em seguida, agrupam-se os haplótipos interiores conectados a outros haplótipos também por um evento de mutação, obtendo-se, assim, os clados de primeiro nível, pois os haplótipos são considerados clados de nível zero. Portanto, os clados inferiores foram aninhados dentro dos clados superiores, respeitando as regras descritas por Templeton et al. (1987). O aninhamento final ocorre quando envolve a rede completa, formando o cladograma total.

Foi executada uma análise exata de contingência aninhada para testar associação geográfica de uma categoria aninhada com uma rede de haplótipos, pelo programa GEODIS (Posada et al., 2000). Nesse caso, a localização das amostras foi tratada como uma variável e um teste exato de contingência permutacional executado para cada nível de clado que apresentava variação no número de haplótipos ou clados. Uma estatística

de χ^2 foi calculada de uma tabela de contingência, na qual as linhas representam a distribuição dos clados e as colunas, as coordenadas geográficas (Posada et al., 2000). Essa análise permitiu testar a hipótese nula de que não há associação entre os clados com a localização geográfica da população de origem, contrastando-se a hipótese alternativa de que os clados estejam agrupados de acordo com a localização geográfica da população de origem. Em cada análise, um nível de contingência pelo teste de χ^2 foi calculado e a significância, determinada por 1.000 permutações aleatórias, preservando-se os valores marginais. A teoria foi descrita por Hudson et al. (1992) e o teste de contingência aninhada, por Templeton e Sing (1993). A significância dos valores do χ^2 e o nível de divergência foram providos através do procedimento “Monte Carlo”, usando-se o algoritmo de Roff e Bentzen (Roff e Bentzen, 1989).

4. RESULTADOS E DISCUSSÃO

Foram amostrados 40 indivíduos de *Psychotria ipecacuanha*, representando nove populações do Sudeste e seis do Centro-Oeste. A região-alvo do genoma de cloroplasto seqüenciada foi o gene *trnT-trnL*. Os alinhamentos das seqüências dos indivíduos foram realizados com a introdução de “gaps” (quebra da continuidade da seqüência de DNA) para compensar a presença de “indels” (inserção ou deleção de bases), sendo eliminados o início e o final de cada seqüência, tanto para o segmento AB, quanto para o segmento CD, a fim de manter uma região comum a todas as 40 seqüências dos indivíduos,

A amplificação do fragmento *trnT-trnL* do DNA genômico de cloroplasto produziu duas seqüências por indivíduo, porque os segmentos não se sobrepunham um ao outro. O segmento AB corresponde ao espaço intergênico entre o exon 5' de *trnT* (UGU) e *trnL* (UAA) e permitiu a obtenção de uma seqüência de 549 bases comuns a todos os indivíduos estudados. O segmento CD corresponde ao íntron localizado entre o *trnL* (UAA) 5'exon e o *trnL* (UAA) 3' exon e permitiu a obtenção de uma seqüência de 410 bases comuns a todos os indivíduos. O haplótipo completo foi a junção das seqüências originadas dos segmentos AB e CD, totalizando uma extensão de 959 bases. As seqüências nucleotídicas dos indivíduos divergentes foram alocadas em oito haplótipos. A seqüência completa de cada haplótipo está mostrada na Figura 5.

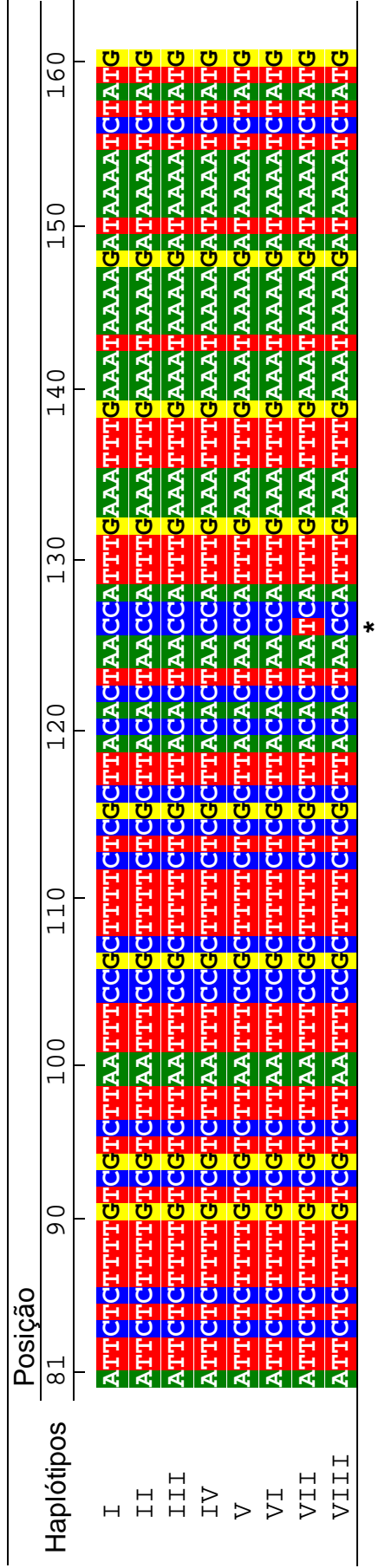
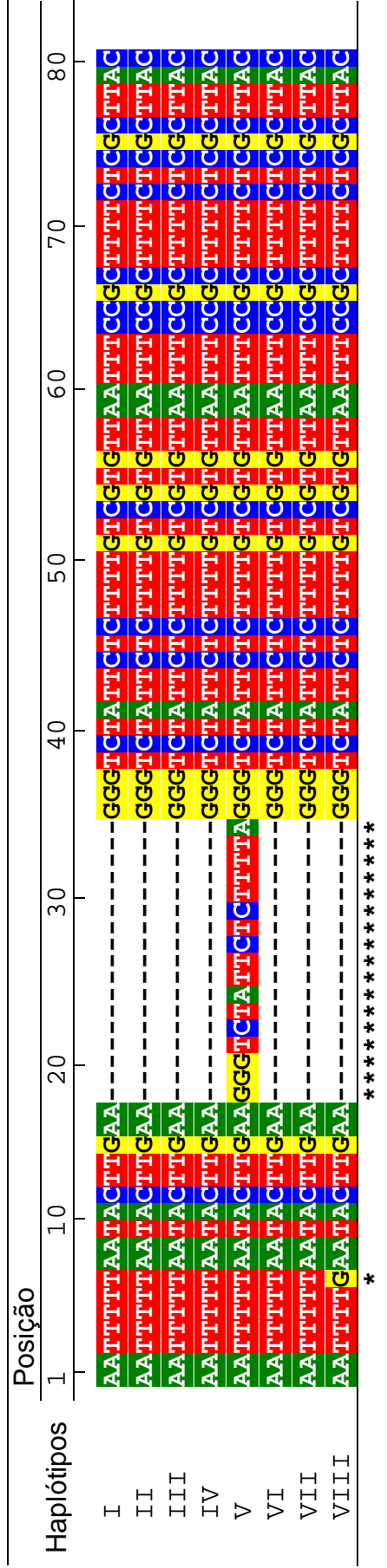


Figura 5 - Alinhamento entre seqüências nucleotídicas de oito haplótipos de cloroplasto de *Psychotria ipecacuanha*. As seqüências foram alinhadas no aplicativo computacional SEQUENCHER, versão 4.0.5. As posições que diferiram estão destacadas com asterisco.

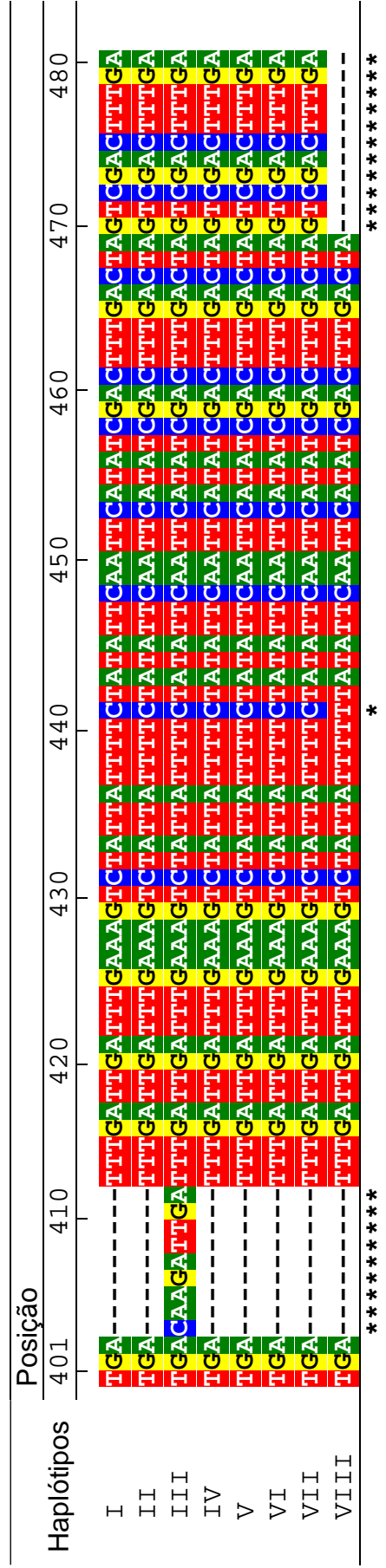
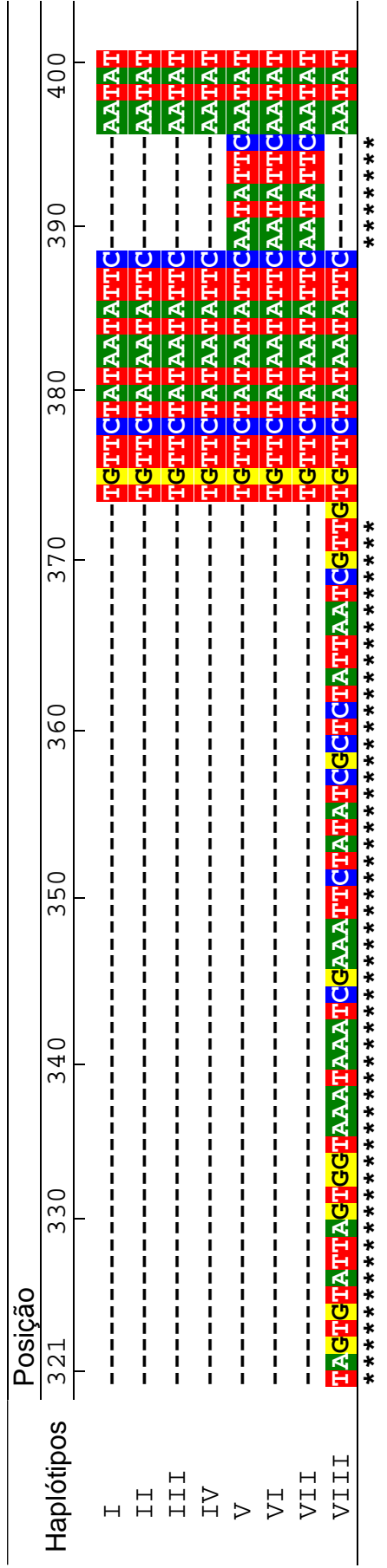


Figura 5 - Continuação

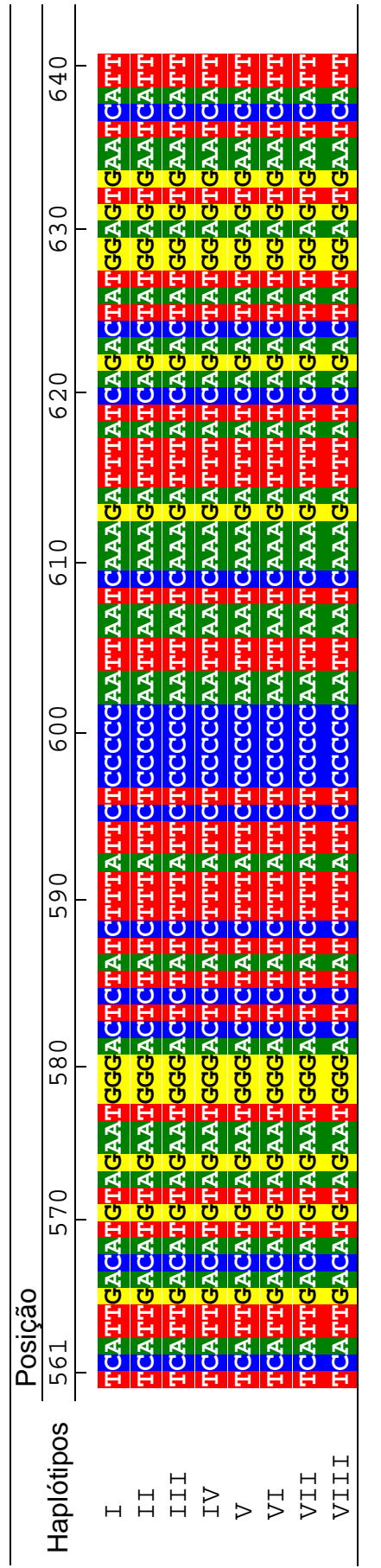
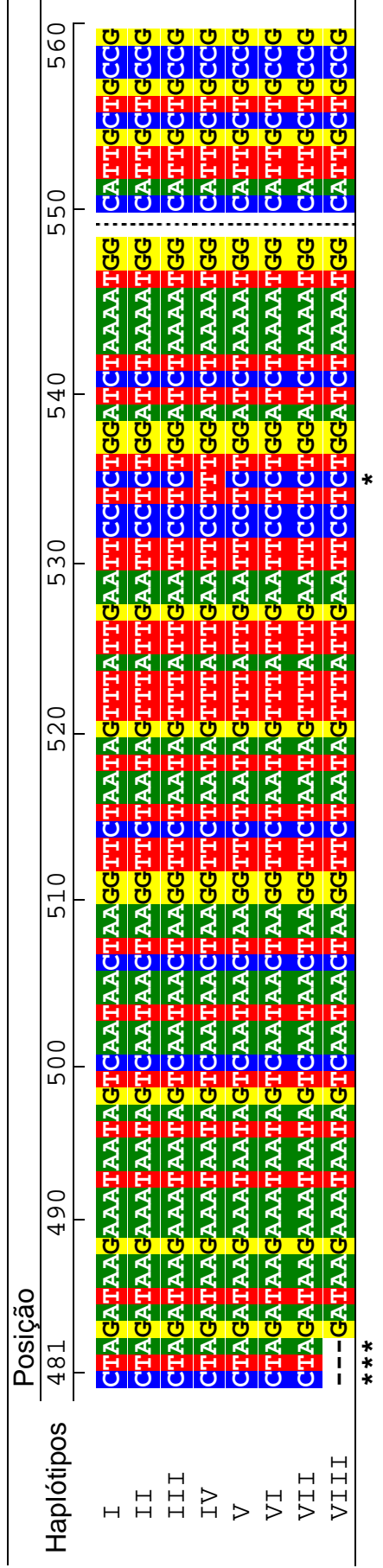
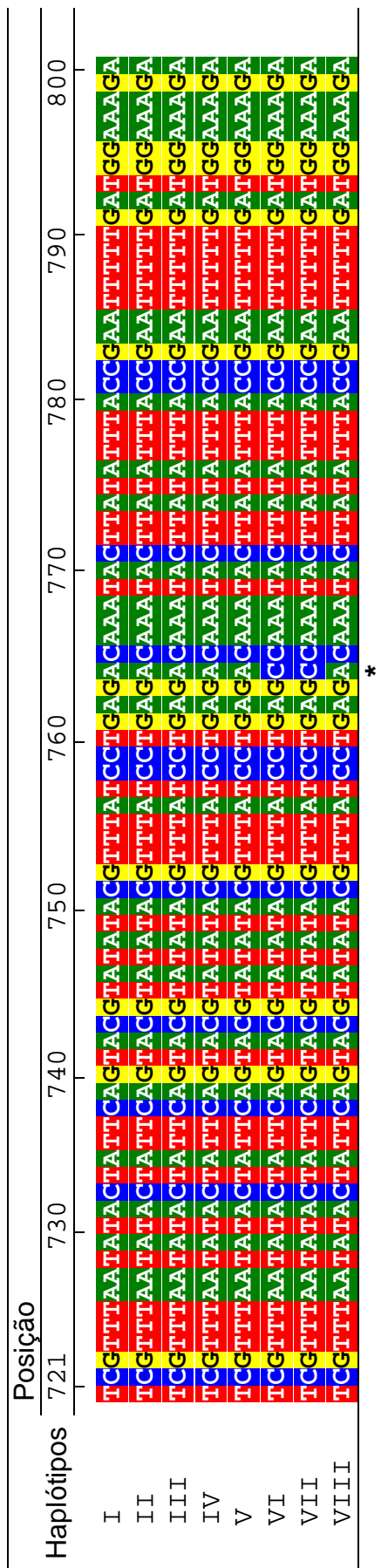
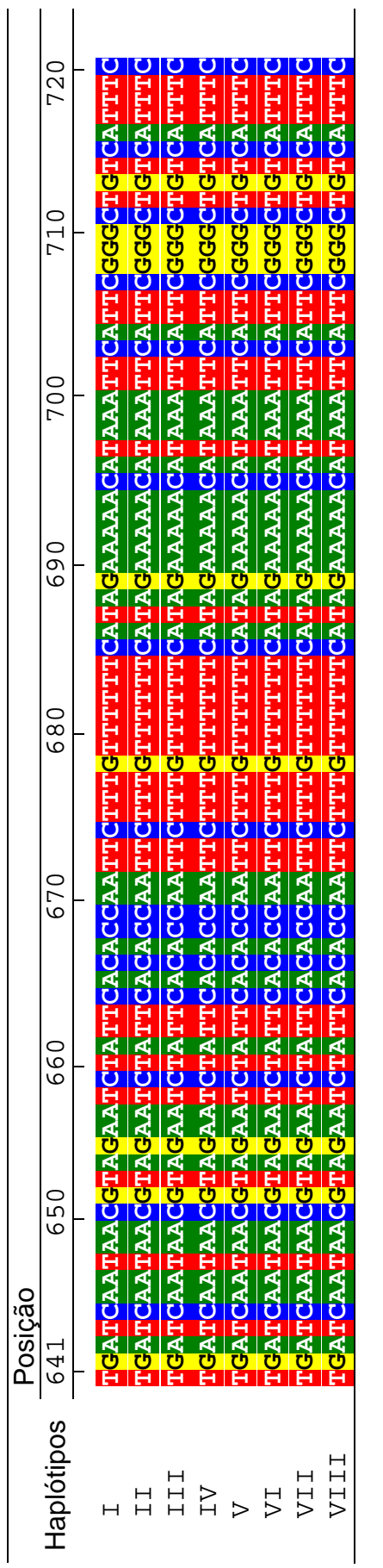


Figura 5 - Continuação



*

Figura 5 - Continuação

As seqüências divergentes entre os oitos haplótipos, tendo o haplótipo I como seqüência consenso, são mostradas na Figura 6. O alinhamento das 40 seqüências dos indivíduos revelou a presença de 13 sítios polimórficos que foram caracterizados por substituições e “indels” (Figura 6). As seqüências dos indivíduos referentes ao segmento AB forneceram nove sítios polimórficos. Três deles foram caracterizados como substituição do tipo transição de bases e uma substituição do tipo transversão, tendo sido os outros cinco “indels”, uma extensão de 17, 213, 7, 9 e 14 bases, respectivamente. As seqüências dos indivíduos referentes ao segmento CD apresentaram cinco polimorfismos, que se caracterizaram como substituição de bases, sendo porém, somente uma das substituições do tipo transição. Um total de oito haplótipos pôde ser distinguido entre as 40 seqüências dos indivíduos de *Psychotria ipecacuanha* investigados.

Haplótipos	Posição																
	7	18	34	126	161	373	389	395	404	412	441	470	483	535	764	802	804
I	T	-----	-----	C	-----	-----	-----	-----	-----	-----	C	GTCGACTTTGACTA	C	A	G	C	T
II	.	-----	-----	.	-----	-----	-----	-----	-----	-----	.	GTCGACTTTGACTA	C
III	.	-----	-----	.	-----	-----	-----	CAAGATTGA	-----	-----	.	GTCGACTTTGACTA
IV	.	GGGTCTATTCCTTTTA	-----	.	-----	-----	-----	-----	-----	-----	.	GTCGACTTTGACTA	T
V	.	-----	-----	.	-----	-----	AAATATTC	-----	-----	-----	.	GTCGACTTTGACTA
VI	.	-----	-----	.	-----	-----	AAATATTC	-----	-----	-----	.	GTCGACTTTGACTA	.	C	.	.	.
VII	.	-----	-----	T	-----	-----	AAATATTC	-----	-----	-----	.	GTCGACTTTGACTA	.	C	.	.	.
VIII	G	-----	-----	.	ATC...TTG	-----	-----	-----	-----	-----	T	-----	.	.	T	A	.
Substituição	TV	-----	-----	TS	-----	-----	-----	-----	-----	-----	TS	-----	TS	TV	TV	TV	TS
Indels		17			213		7		9			14					
		Segmento AB															
		Segmento CD															

Figura 6 - Sequências alinhadas do fragmento *tnrT-trnL* do cpDNA de *Psychotria ipecacuanha* discriminando oito haplótipos entre 40 sequências obtidas. Somente as posições que diferem do haplótipo I são demonstradas. Cada ponto nesta figura indica que não houve variação de base nitrogenada em relação ao haplótipo I. Os travessões indicam os “gaps” para compensar a presença de “indels”. O haplótipo completo é resultado da união das seqüências desses dois segmentos. As seqüências dos segmentos AB e CD resultaram em 549 e 410 bases, respectivamente, totalizando um fragmento com extensão de 959 bases. O alinhamento revelou 13 sítios polimórficos, sendo oito substituições de base e cinco “indels” com extensão de 17, 213, 7, 9 e 14 bases. Os tipos de substituições foram classificados como sendo transverso (TV) e transição (TS).

Na Região Sudeste, as seqüências dos indivíduos foram alocadas nos haplótipos de I a VII e distribuídas de acordo com sua população de origem. Entretanto, no Centro-Oeste, entre as 14 seqüências dos indivíduos investigados, foi detectado apenas o haplótipo VIII (Tabela 2).

O haplótipo I foi exclusivo da Região Sudeste e o mais freqüente, com 16 seqüências entre as 26 obtidas dessa região. Esse haplótipo também foi encontrado em quase todas as populações analisadas, menos na população de IRU, por isso foi tomado como consenso na Figura 5. O haplótipo VIII, além de exclusivo da Região Centro-Oeste, foi o único encontrado nessa região entre as 14 seqüências dos indivíduos investigados.

Os haplótipos II, III, IV e VII foram exclusivos das populações GUA, IRU e ITA. Esses haplótipos podem ser considerados raros ou de baixa freqüência, porque cada um deles foi encontrado em apenas um único indivíduo. Portanto, tais populações devem ter prioridade em um trabalho de conservação genética, uma vez que possuem haplótipos endêmicos em áreas de risco de erosão genética.

As populações PRD, GUA e ITA foram as que apresentaram maior número de haplótipos por população. Foram amostrados oito indivíduos na população PRD, na qual, as seqüências dos indivíduos constituíram os haplótipos I, V e VI, na proporção 2:2:4. Dos dois indivíduos amostrados na população GUA, foi possível identificar dois haplótipos: I e II. Por sua vez, na população ITA, as seqüências dos seis indivíduos foram alocadas nos haplótipos I, IV e VII, na proporção 4:1:1 (Tabela 2).

A variação no número de haplótipos por população é muito importante na detecção de evidências de eventos na história das populações de poaia, pois pode revelar se esses haplótipos foram derivados de algum outro encontrado na população ou de haplótipos de outras populações.

Tabela 2 – Distribuição das 40 seqüências dos indivíduos de *Psychotria ipecacuanha* entre os oito haplótipos encontrados. Os indivíduos foram separados de acordo com a sua região de origem (Sudeste (SE) ou Centro-Oeste (CO)) e por sua população correspondente ao local onde foi amostrado (listados na Tabela 1). N é o número total de indivíduos analisados por população. H/P indica o número de haplótipos encontrados por população

Região	Código	N	Haplótipos								H/P	
			I	II	III	IV	V	VI	VII	VIII		
SE	UNA	2	2									1
	PTN	1	1									1
	VRB	3	3									1
	CAR	1	1									1
	CON	2	2									1
	GUA	2	1	1								2
	IRU	1			1							1
	PRD	8	2				2	4				3
	ITA	6	4				1		1			3
Total SE	—	26	16	1	1	1	2	4	1		7	
CO	BBU	3									3	1
	FRA	1									1	1
	AIN	2									2	1
	MAN	2									2	1
	CUR	4									4	1
	SAC	2									2	1
Total CO	—	14									14	1
Total Geral		40	16	1	1	1	2	4	1	14		

A estimação da rede de haplótipos foi gerada de acordo com o método proposto por Templeton et al. (1992), usando-se o aplicativo computacional TCS (Versão 1.13). Os “indels” foram considerados como o quinto caráter manifestado. Assim, cada “indel” foi codificado, independentemente de sua extensão, considerando-se somente um caráter. O número máximo estimado de conexões parcimoniosas entre pares de seqüências dos indivíduos com 95% de probabilidade para todos os acoplamentos dos haplótipos dentro da rede foi de 11 eventos de mutação. Assim, entre dois haplótipos vizinhos, em relação às seqüências dos seus indivíduos, a divergência genética máxima aceitável é no máximo de 11 eventos de mutação para se manterem agrupados numa mesma rede. Caso contrário, as seqüências desses indivíduos seriam alocadas em redes diferentes. Esse número máximo de conexões parcimoniosas para manter dois haplótipos vizinhos agrupados numa mesma rede está relacionado ao tamanho da seqüência do haplótipo completo, pois, quanto maior o número de bases nas seqüências dos indivíduos, maior o número máximo de conexões parcimoniosas entre dois haplótipos vizinhos.

Como demonstrado na Figura 7, foi obtida uma única rede contendo os oito haplótipos, pois o número máximo de conexões parcimoniosas entre dois haplótipos vizinhos mais divergentes (haplótipo I e o VIII) foi de seis eventos de mutação. Os haplótipos detectados não mostraram ambigüidade quanto à sua origem, ou seja, não apresentaram homoplasia. Cada barra de ligação entre um haplótipo e outro (Figura 7A) representa uma conexão parcimoniosa e resulta de um único evento de mutação,

No processo de estimação da rede de haplótipos, foi necessário o uso de seis nodos interiores (Figura 7B), para os quais, nenhuma das 40 seqüências dos indivíduos pôde ser alocada. De acordo com Templeton et al. (1988), estes nodos são deduzidos entre dois haplótipos mais próximos da rede, que diferem entre si em dois ou mais eventos de mutação. Esses haplótipos intermediários foram extintos ou estão com uma

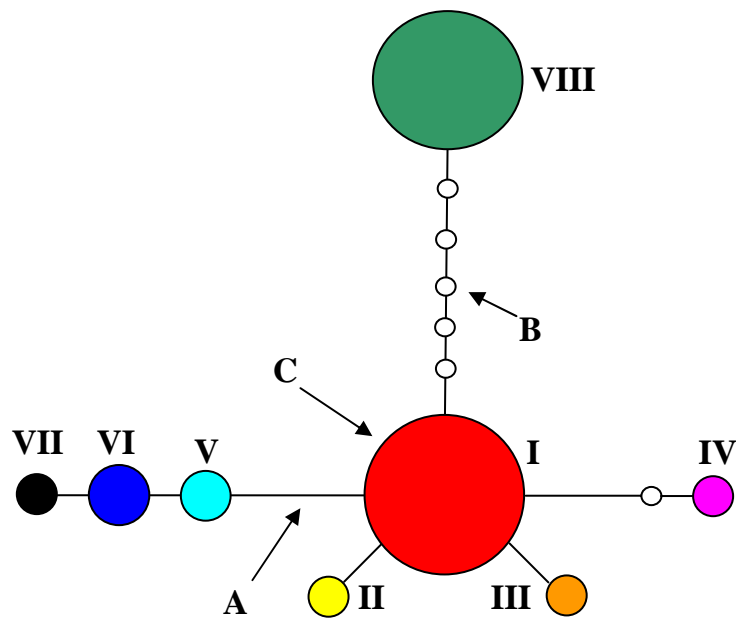


Figura 7 - Rede de haplótipos baseada em seqüências do gene *trnT-trnL* do genoma de cloroplasto de *Psychotria ipecacuanha*. Os números em romanos representam os haplótipos descritos na Tabela 4. A rede de haplótipos foi obtida pelo aplicativo computacional TCS (Versão 1.13), desenvolvido por Clement et al. (2000). A área de cada círculo representa um haplótipo particular e é proporcional ao número de indivíduos alocados neste haplótipo. A - Barra de conexão parcimoniosa entre haplótipos com uma probabilidade superior ou igual a 95%, cada uma correspondendo a um único evento de mutação. B - Nodo interior da rede que não foi amostrado no momento ou foi extinto; tendo sido esses haplótipos deduzidos entre dois ou mais haplótipos intermediários. C - Haplótipo que fica no interior da rede, apresentando uma multifurcação.

freqüência muito baixa na população e raramente serão amostrados. Como exemplo, foi verificada a existência de um haplótipo intermediário que conecta os haplótipos I e IV, que estão separados por dois eventos de mutação e de cinco haplótipos intermediários, que conectam os haplótipos I e VIII, diferenciados por seis eventos de mutação.

As áreas dos círculos dos haplótipos correspondem às suas freqüências. Os haplótipos II, III, IV, VII e VIII estão nas extremidades, enquanto os haplótipos I, V e VI, no interior da rede, sendo, o haplótipo I, o mais central, apresentando uma multifurcação com cinco conexões parcimoniosas (Figura 7C).

Considerando apenas as seqüências dos indivíduos referentes ao segmento AB, os haplótipos I e II não foram separados, assim com os haplótipos V e VI. No entanto, o seqüenciamento do segmento AB discriminou o haplótipo I/II dos haplótipos III e V/VI em um único evento de mutação, discriminou ainda o haplótipo IV em dois eventos e o haplótipo VIII em quatro eventos de mutação. Foi possível também discriminar o haplótipo VII do haplótipo V/VI em um único evento de mutação (Figura 8). Considerando apenas as seqüências dos indivíduos referentes ao segmento CD, às seqüências dos haplótipos I, III, IV e V foram idênticas, como as seqüências dos haplótipos VI e VII. Porém, as seqüências dos indivíduos referentes a esse segmento discriminaram o haplótipo I/III/IV/V dos haplótipos II e VI/VII, em apenas um evento de mutação e o haplótipo VII, em dois eventos (Figura 9).

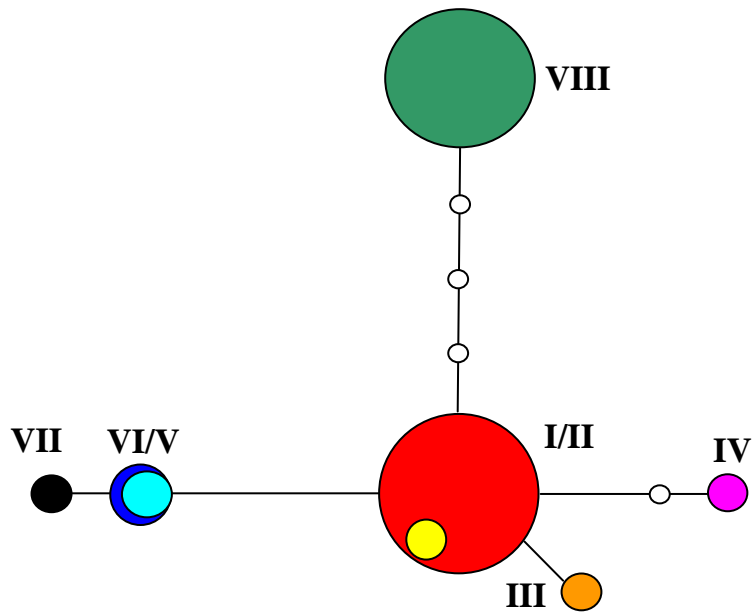


Figura 8 - Rede de haplótipos obtida exclusivamente com informações provenientes das seqüências dos indivíduos da região do segmento AB do gene *trnT-trnL*, do genoma de cloroplasto de *Psychotria ipecacuanha*.

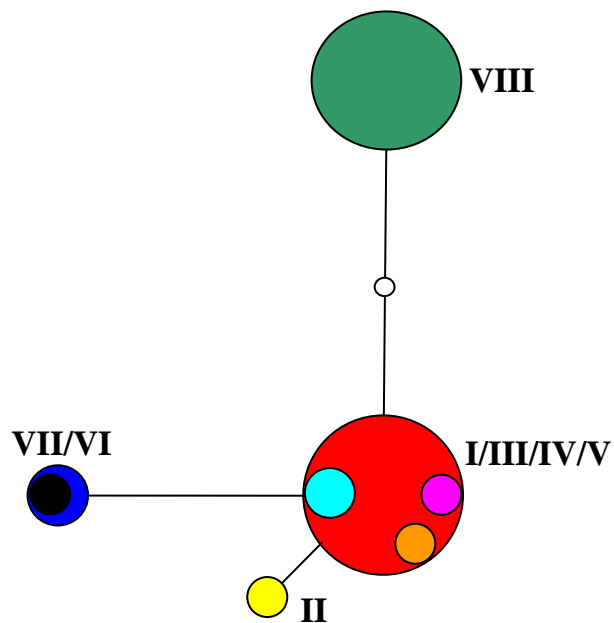


Figura 9 - Rede de haplótipos obtida exclusivamente com informações provenientes das seqüências dos indivíduos da região do segmento CD do gene *trnT-trnL*, do genoma de cloroplasto de *Psychotria ipecacuanha*.

Comparando as seqüências dos indivíduos em relação ao segmento AB, foi possível diferenciar o haplótipo I/III/IV/V da Figura 9, em quatro haplótipos e o haplótipo VI/VII, em dois. No entanto, as seqüências dos indivíduos em relação ao segmento CD distinguiram o haplótipo I/II da Figura 8, em dois haplótipos, assim como o haplótipo V/VI. As seqüências dos indivíduos referentes ao segmento AB discriminaram seis dos oitos haplótipos e as seqüências dos indivíduos referentes ao segmento CD, quatro. Porém, as seqüências dos indivíduos na região de cada segmento apresentaram estrutura semelhante em suas redes de haplótipos, isto é, um haplótipo central conectado a todos os outros haplótipos das extremidades (Figura 8 e 9). Provavelmente, este fato tenha ocorrido em razão de as seqüências dos indivíduos referentes aos segmentos AB e CD estarem fisicamente ligadas. Assim, a pressão de seleção deve ser bem similar nas seqüências dos indivíduos na região dos segmentos AB e CD.

As análises seguintes têm por base a rede de haplótipos mostrada na Figura 7. Após obter essa rede de haplótipos, foram feitos aninhamentos de acordo com as regras propostas por Templeton et al. (1987) e Templeton e Sing (1993). Para a construção de clados de primeiro nível, todos os haplótipos são considerados clados de nível zero. Então, pelo algoritmo de aninhamento, foram delimitados com linha cheia, os haplótipos das extremidades da rede, separados por um único evento de mutação dando origem aos clados 1-1, 1-3 e 1-6. Logo em seguida, os haplótipos internos foram agrupados também por um evento de mutação. No caso de uma multifurcação, como ocorrida no haplótipo I, foram aninhados todos os outros haplótipos separados do central (haplótipo I) por um evento de mutação. Então, os clados de primeiro nível foram aninhados em clados de segundo nível. Como demonstrado na Figura 10, as linhas cheias representam os aninhamentos dos clados de primeiro nível e as linhas tracejadas indicam o

aninhamento dos clados de segundo nível. O cladograma total foi delimitado por um retângulo de linha cheia.

O clado 1-1 foi formado pelos haplótipos VI e VII; o clado 1-2 pelos haplótipos I, II, III e V, o clado 1-3 apenas pelo haplótipo IV; o clados 1-4 e 1-5, por haplótipos deduzidos e o clado 1-6, somente pelo haplótipo VIII.

Nas extremidades da rede dos clados de primeiro nível, estão os clados 1-1, 1-3 e 1-6 e no interior da rede, os clados 1-2, 1-4 e 1-5. O clado 1-2 é considerado central na rede de haplótipos, porque apresenta uma mutifurcação com os clados 1-1, 1-3 e 1-4.

O clado 2-1 é formado por sete haplótipos e o clado 2-2, apenas por um. O clado 2-1 contém 26 seqüências dos indivíduos das nove populações amostradas, enquanto o clado 2-2 contém 14 seqüências dos indivíduos das seis populações investigadas.

Após obter a rede de clados aninhados, foi realizado inicialmente um teste simples de associação geográfica, tratando cada local da população como uma variável. Esse teste de contingência exata foi realizado pelo “software” GEODIS (Versão 2.0) desenvolvido por Posada et al. (2000), nos clados com mais de um haplótipo e mais de um local (população), de acordo com o algoritmo descrito por Templeton e Sing (1993). Os clados 1-1, 1-2, 2-1 e o cladograma total apresentaram variação no número de clados nos níveis inferiores e também variação no número de população. Então, a hipótese nula testada foi a de que não há associação entre os clados e a localização geográfica da população de origem. Já a hipótese alternativa foi a de que os clados estão agrupados de acordo com a localização geográfica da população de origem.

A primeira análise efetuada no aplicativo computacional GEODIS está apresentada na Tabela 3. Nessa análise, rejeitou-se a hipótese nula somente no cladograma total com o nível de 5% de probabilidade, pelo teste de χ^2 , proposto por Roff e Bentzen (1989). A aceitação de hipótese alternativa implica que os clados de segundo nível não sejam derivados de uma única linhagem evolutiva (Templeton, 2001).

Tabela 3: Análise exata de contingência de associações geográficas nos clados que tiveram variação da Figura 10, obtida pelo aplicativo computacional GEODIS (Versão 2.0)

Clados	χ^2	<i>P</i>
Clado 1-1	5,00	0,2350
Clado 1-2	38,13	0,0760
Clado 2-1	11,27	0,7140
Cladograma Total	40,00	0,0000*

*Nível de significância de 0,05.

No arquivo de saída do GEODIS (Versão 2.0), também foram extraídas as medidas relativas de distância geográfica com rede de clados aninhados com suas respectivas probabilidades (Tabela 4), sob hipótese nula de nenhuma associação geográfica com um particular nível de clado. De acordo com Templeton et al. (1995), a primeira distância simbolizada por D_C , (distância do clado), mede como os indivíduos estão difundidos e geograficamente amparados por um particular nível de clado. A segunda, indicada por D_N , (distância do clado aninhado), mede a distância média de todos os indivíduos com um clado de nível superior. Estas duas distâncias são recalculadas com no mínimo 1.000

permutações aleatórias para se ter uma inferência com o nível de significância de 5% de probabilidade (Edgington, 1986). Esse procedimento de permutação randômica permitiu testar as distâncias D_C e D_N com probabilidades significativamente grande (+) ou pequena (-). Foram calculados também os contrastes das distâncias dos clados interiores com os das extremidades dentro de cada nível de clado aninhado, isso quando possível. No clado 1-1, as distâncias D_N foram significativamente grande (+) no clado VII e significativamente pequena (-) no contraste (I-E), no nível de 5% de probabilidade. No clado 1-2, apenas a distância D_N do clado II foi significativamente pequena (-) no nível de 5% de probabilidade. No cladograma 2-1, nenhuma das distâncias D_C e D_N foram significativas no nível de 5% de probabilidade. Entretanto, no cladograma total, todas as distâncias D_C e D_N foram significativas, porém a distância D_N do clado 2-2 foi a única significativamente grande (+) (Tabela 4).

A chave de inferências com base nos padrões esperados de associação geográfica proposta por Templeton (2004) foi utilizada de acordo D_C , D_N e I-E significativos de cada clado. Na Tabela 5, nenhuma evidência de padrão filogeográfico foi detectada no clado 1-1. Entretanto, o clado 1-2 apresentou fluxo de gene restrito com isolamento através da distância que é comum em espécies de propagação vegetativa. No caso do cladograma total, foi possível evidenciar que os clados 2-1 e 2-2 foram separados por fragmentação alopátrica.

Tabela 4 - Resultado da análise da distância geográfica aninhada de populações de *Psychotria ipecacuanha* obtidas da análise feita no aplicativo computacional GEODIS (Versão 2.0). Nos clados, são informados os níveis dos clados/dispersão de haplótipos (D_C), a deposição de clados com níveis mais altos de aninhamento (D_N) e os contrastes do interior vs. extremidade (I-E), quando disponíveis. A probabilidade significativa (probabilidade inferior ou igual a 0,05 está indicada por asterisco) de valores de distância grande (+) ou pequena (-) foi baseada em comparações com hipótese nula, em que H_0 corresponde a nenhuma associação geográfica com os clados. Os resultados foram gerados por 1.000 permutações dos dados mostrados no Anexo I

	Posição	D_C	P	D_N	P
Clado 1-1					
Clado VI	Interior	0,0000	(-) 0,2350	26,5716	(-) 0,2350
Clado VII	Extremidade	0,0000	(+) 1,0000	79,9103	(+) 0,0000*
Clado I-E	—	0,0000	(-) 0,2350	-53,3387	(-) 0,0000*
Clado 1-2					
Clado I	Interior	181,0483	(+) 0,1460	178,1115	(+) 0,0690
Clado II	Extremidade	0,0000	(-) 0,6850	41,4239	(-) 0,0000*
Clado III	Extremidade	0,0000	(-) 0,7080	88,7299	(+) 0,3910
Clado V	Interior	0,0000	(-) 0,0850	66,4098	(-) 0,0830
Clado I-E	—	160,9318	(+) 0,2320	100,6233	(+) 0,0730
Clado 2-1					
Clado 1-1	Extremidade	39,9063	(-) 0,1640	66,8364	(-) 0,0690
Clado 1-2	Interior	152,4825	(-) 0,4320	150,8018	(+) 0,4400
Clado 1-3	Extremidade	0,0000	(-) 0,7780	78,6815	(-) 0,3770
Clado I-E	—	119,2272	(+) 0,3120	81,9912	(+) 0,1300
Cladograma total					
Clado 2-1	Extremidade	143,2466	(-) 0,0000*	715,2905	(-) 0,0000*
Clado 2-2	Extremidade	61,4703	(-) 0,0010*	1044,4366	(+) 0,0000*

Tabela 5 - Inferências demográficas de análise de distância geográfica aninhada, de acordo com a chave de inferências para análise de árvore de haplótipos aninhados com distância geográfica (Templeton, 2004), de acordo com a chave de inferências apresentada do Anexo II

Clado	Ordenamento das Inferências	Evidência
Clado 1-1	1(Não) → 2(Não) → 11(Não) → 17(Não)	Resultado inconclusivo
Clado 1-2	1(Não) → 2(Não) → 11(Não) → 17(Não) → 4(Não)	Fluxo de gene restringido por isolamento através da distância
Cladograma Total	1(Sim) → 19(Não)	Fragmentação alopátrica

No mapa da Figura 11 é mostrada a distribuição geográfica das populações de *Psychotria ipecacuanha* amostradas neste estudo. A rede de haplótipos aninhados foi incluída como referência na distribuição dos oito haplótipos entre as 15 populações estudadas.

Todas as 14 seqüências dos indivíduos das seis populações do Centro-Oeste foram alocadas no haplótipo VIII, o qual foi aninhado no clado 1-6, enquanto no Sudeste, as 26 seqüências dos indivíduos foram alocadas nos haplótipos I a VII, aninhados nos clados 1-1, 1-2 e 1-3.

As seqüências dos indivíduos das populações UNA, VRB e CON foram todas alocadas no haplótipo I, não tendo havido variação no número de haplótipos dentro das mesma. Nas seqüências das populações de PTN e CAR, foi possível amostrar um único indivíduo, sendo sua seqüência alocada no haplótipo I. A seqüência do indivíduo da população de IRU foi alocada no haplótipo III que está conectado ao haplótipo I, por um único evento de mutação. O haplótipo III é considerado privativo pelo fato de se localizar em uma das extremidades da rede e também por ter apresentado frequência baixa neste estudo. As populações que apresentaram variação no número de haplótipos foram PRD, ITA e GUA.

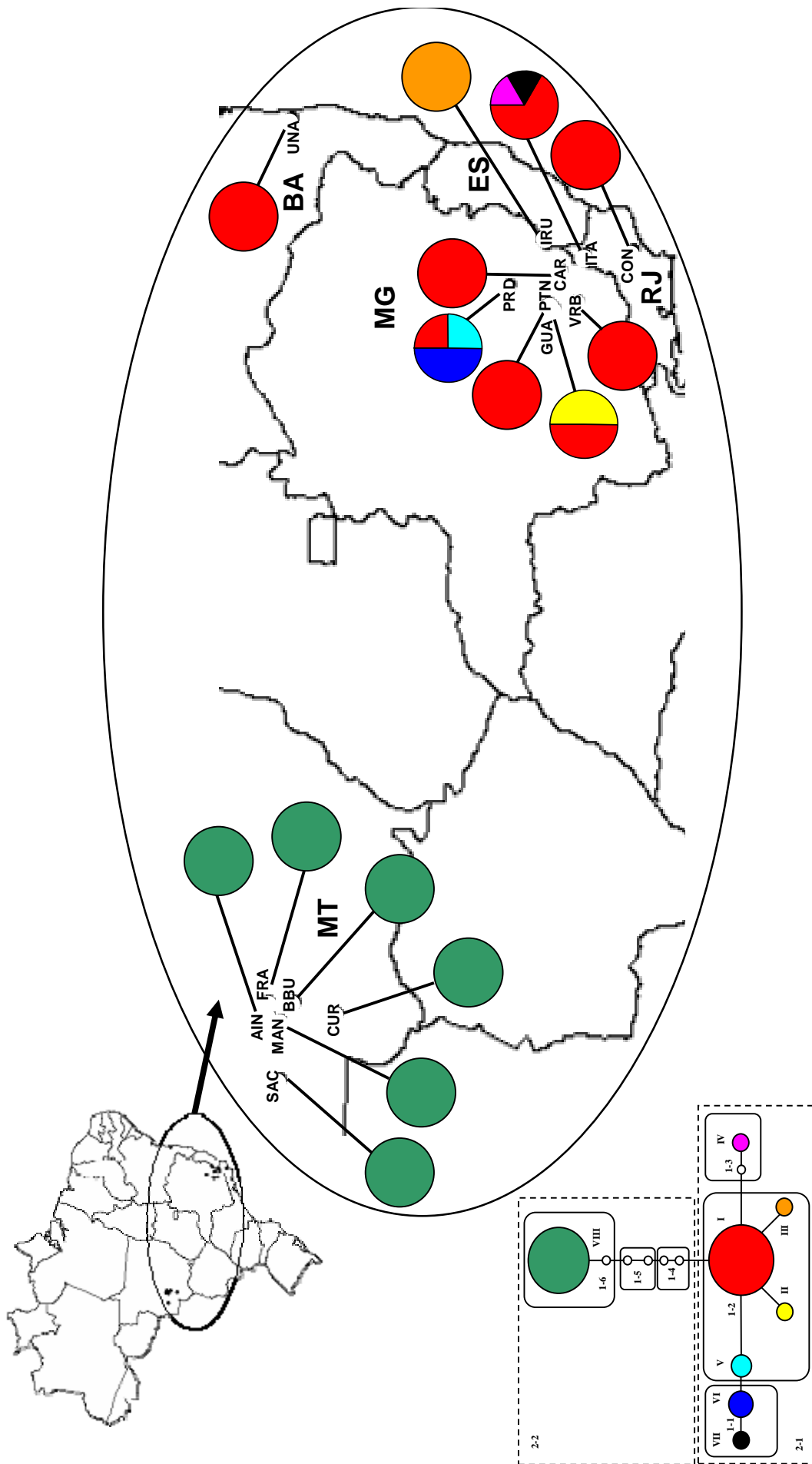


Figura 11 - Distribuição geográfica de oito haplótipos de cpDNA de *Psychotria ipecacuanha* amostrados nas Regiões Sudeste e Centro Oeste. Os códigos correspondem aos locais das populações e estão listados na Tabela 1). As figuras em formato de círculo representam a frequência dos haplótipos na população. A rede de haplótipos aninhados foi colocada como referência.

Na população PRD, as seqüências de seus indivíduos foram alocadas nos haplótipos I, V e VI, que são os haplótipos internos da rede e estão interligados entre si pelo haplótipo V, em apenas um único evento de mutação. Os haplótipos da população PRD foram aninhados em dois clados de primeiro nível: o 1-1 e o 1-2. O clado 1-1 é formado pelo haplótipo VI da população PRD e pelo haplótipo VII da população de ITA. Nesse clado, não foi possível detectar nenhuma associação geográfica, mesmo apresentando D_N 's significativas. Na população de ITA as seqüências de seus indivíduos foram alocadas nos haplótipos I, IV e VII, tendo sido aninhados nos clados de primeiro nível 1-1, 1-2 e 1-3.

Considerando a Região Sudeste, os haplótipos que apresentaram mais eventos de mutação entre si foram os haplótipos IV e VII, que foram detectados na população ITA. Uma hipótese seria a de que o haplótipo VII teria migrado da população PRD, onde foram detectados os haplótipos ancestrais V e VI. Na população de GUA, foram detectados os haplótipos I e II, que estão separados em um único evento de mutação e foram aninhados no clado 1-2.

O clado 1-2 foi formado pelos haplótipos I, II, III e V. Cada população do Sudeste contribuiu com pelo menos uma seqüência de dado indivíduo alocados em um desses quatro haplótipos. No clado 1-2, o padrão filogeográfico detectado foi o fluxo gênico restrito com isolamento através da distância que é comum em espécies de propagação vegetativa. O resultado da Tabela 5 para o clado 1-2 está coerente com o que se conhece da poaia, visto que Rossi (2003), trabalhando com a biologia reprodutiva da planta, observou que as populações possuem reprodução sexuada restrita e que a propagação vegetativa demonstra ser mais eficiente.

Considerando o fluxo gênico restrito por distância, uma linhagem de gene ancestral pode ter sido difundida em toda a área de ocorrência natural de uma espécie, sendo as linhagens de genes descendentes restritas em algumas áreas e consideradas privativas

(Avise, 2000). No caso da poaia, considerando o clado 1-2, o haplótipo I foi encontrado em quase todas as populações investigadas do Sudeste e os haplótipos II, III e V foram privativos das populações de GUA, IRU e PRD. Com base na teoria da coalescência, os haplótipos privativos são considerados em média, os mais novos evolutivamente. Em geral, esses haplótipos possuem baixa frequência na população e estão localizados nas extremidades da rede. Ao contrário, os haplótipos localizados no interior da rede são considerados ancestrais e em média são os mais antigos e, portanto, espera-se que sejam os mais frequentes na população (Avise, 2000). Observando o clado 1-2 da rede de haplótipos aninhados de poaia, o haplótipo I é o mais frequente e está no interior da rede. Os haplótipos II e III, além de serem privativos, estão nas extremidades e exibiram baixa frequência neste estudo.

Conforme o modelo de restrição do fluxo gênico por isolamento de distância e seguindo a teoria da coalescência, Brito (2002), trabalhando com *Drosophila buzzatii*, observou também, em suas populações localizadas na América do Sul, que o haplótipo posicionado no interior da rede foi o mais frequente e estava mais difundido entre as populações estudadas. Porém, os haplótipos das extremidades foram menos frequentes e privativos de algumas áreas.

Considerando o nível de cladograma total, foi possível separar em clados de segundo nível, as populações, de acordo com a região de origem. Os haplótipos do clado 2-1 foram originados das seqüências de indivíduos das populações do Sudeste e os do clado 2-2, do Centro-Oeste. Como foi demonstrado na Tabela 5, a evidência filogeográfica no nível de cladograma total foi uma fragmentação alopátrica. De acordo com Avise (2000), para existir uma evidência de fragmentação alopátrica, as linhagens de genes têm que estar separadas por vários eventos de mutação. Além disso, essas linhagens de genes estão separadas por algum tipo de barreira geográfica. Assim, a fragmentação alopátrica evidenciada em poaia entre os clados 2-1 e 2-2 foi satisfatória,

porque existe uma barreira geográfica entre essas duas regiões. Essa barreira está relacionada com presença do cerrado, que apresenta condições desfavoráveis para o desenvolvimento da *Psychotria ipecacuanha*. Os clados de segundo nível foram separados por cinco haplótipos intermediários, que, ao longo da expansão e colonização da poaia, podem ter sido extintos.

Lima (2002), trabalhando com DNA genômico total de poaia pela técnica de RAPD (“Random Amplification of Polymorphic DNA”), não detectou associação da divergência genética entre os indivíduos do seu trabalho e sua distribuição geográfica, mesmo entre os indivíduos das Regiões Centro-Oeste e Sudeste. Entretanto, Martins (2000), estudando também *Psychotria ipecacuanha* e utilizando a técnica de RAPD, conseguiu separar o indivíduo do Centro-Oeste, dos indivíduos do Sudeste. A metodologia apresentada neste estudo, além de separar as linhagens de genes do Sudeste da linhagem de gene do Centro-Oeste, permitiu também, identificar parte da história das populações de poaia, em diferentes níveis de clado aninhados, associadas à sua distribuição atual. A restrição do fluxo gênico por distância geográfica evidenciada nas populações do Sudeste é considerada evento recente, ao contrário da fragmentação alopátrica entre as populações do Sudeste e do Centro-Oeste, que é considerada evento antigo (Avice, 2000).

Este trabalho também pode servir como suporte na elaboração de um programa de conservação de *Psychotria ipecacuanha*, porque, além de identificar os haplótipos privativos entre as populações da poaia, possibilitou detectar as populações que apresentam variabilidade no número de haplótipos.

5. CONCLUSÕES

Os resultados obtidos neste trabalho permitiram chegar às seguintes conclusões principais:

- Identificação e diferenciação das oito linhagens genealógicas do fluxo gênico materno, sendo sete da região Sudeste e apenas uma da região Centro-Oeste.
- As populações UNA, PTN, CAR, VRB e CON foram alocados no haplótipo I.
- As populações PRD, ITA e GUA exibiram variabilidade no número de linhagens genealógicas dentro da população.
- Os haplótipos V e VI foram privativos da população PRD; os haplótipos IV e VII foram privativos da população ITA e haplótipo II da população GUA.
- A nível de cladograma total as distribuições das linhagens dos cladogramas 2-1 e 2-2 evidenciada pelo “nested clade analysis” foi uma fragmentação alopátrica.
- Na região Sudeste apenas no clado de primeiro nível 1-2 é que foi possível evidenciar uma restrição do fluxo gênico por isolamento através da distância.
- Na elaboração de um programa de conservação genética as populações de PRD, ITA e GUA podem ser indicadas como prioridade, pois além de apresentarem variação no número de haplótipos dentro da população, possuem também cinco haplótipos privativos. Além disso, as três populações detêm juntas 75% da variabilidade das linhagens de gene *trnT-trnL* de cloroplasto.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ADDOR, A.A. Considerações acerca da poaia. **Bol. Minist. Agric.**, v.34, n. 2, p. 1-18, 1945.
- ASSIS, M.C.; GIULIETTI, A.M. Diferenciação morfológica e anatômica em populações de “ipecacuanha” - *Psychotria ipecacuanha* (Brot.) Stokes (Rubiaceae). **Rev. bras. Bot.**, v.22, n. 2, p.205-216, 1999.
- AVISE, J.C. The history and purview of Phylogeography: A personal reflection. **Mol. Ecol.**, v.7, n. 4, p. 371-379, 1998.
- AVISE, J.C. **Phylogeography: the history and formation of species**. Cambridge: Harvard Univ. Press, 2000. 447 p.
- AVISE, J.C.; ARNOLD, J.; BALL, R.M.; BERMINGHAM, E.; LAMB, T.; NEIGEL, J. E.; REEB, C.A.; SAUNDERS, N.C. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. **Annu. Rev. Ecol. Syst.**, v.18, p. 489-522, 1987.

- BARROWCLOUGH, G.F.; GUTIÉRREZ, R.J.; GROTH, J.G. Phylogeography of spotted owl (*Strix occidentalis*) populations based on mitochondrial DNA sequences: gene flow, genetic structure, and a novel biogeographic pattern. **Evolution**, v.53, n. 3, p. 919-931, 1999.
- BIRKY, C.W.; MARUYAMA, T.; FUERST, P. An approach to population genetic and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. **Genetics**, v.103, n. 3, p. 513-527, 1983.
- BRITO, R.A.; MANFRIN, M.H.; SENE, F.M. Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. **Genet. Mol. Biol.**, v.25, n. 2, p. 161-171, 2002.
- BRUNETON, J. **Pharmacognosy, phytochemistry, medicinal plants**. Paris: Lavoisier, 1995. 915 p.
- CARON, H.; DUMAS, S.; MARQUE, G.; MESSIER, C.; BANDO, E.; PETIT R.J.; KREMER, A. Spatial and temporal distribution of cpDNA polymorphism in a tropical tree species. **Mol. Ecol.**, v.9, n. 8, p. 1089-1098, 2000.
- CLEGG, M.T. LEARN, G.H.; GOLENBERG, E.M. Molecular evolution of chloroplast DNA. In: SELANDER, R.K.; CLARK, A. G.; WHITTAM, T.S. (Eds.). **Evolution at the molecular level**. Sunderland: Sinauer Associates, 1991. p. 135-149.
- CLEMENT, M.; POSADA, D.; CRANDALL, K.A. TCS: a computer program to estimate gene genealogies. **Mol. Ecol.**, v.9, n. 10, p. 1657-1659, 2000.

- COLLEVATTI, R.G.; GRATTAPAGLIA, D.; HAY, J.D. Evidences for multiple maternal lineages of *Caryocar brasiliense* populations in the Brazilian Cerrado based on the analysis of chloroplast DNA sequences and microsatellite haplotype variation. **Mol. Ecol.**, v.12, n. 1, p. 105-115, 2003.
- COMES, H.P.; ABBOTT, R.J. The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). **Evolution**, v.52, n. 2, p. 390-402, 1998.
- CORRIVEAU, J.L.; COLEMAN, A.W. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. **Am. J. Bot.**, 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, n. 3, p. 959-969, 1993.
- CRANDALL, K.A.; TEMPLETON, A.R. Applications of intraspecific phylogenetics. In: HARVEY, P.H.; BROWN, A.J.L.; SMITH, J.M.; NEE, S. (Eds.). **New uses for new phylogenies**. Oxford: Oxford University Press, 1996. p. 81-99.
- DOYLE, J.J.; DOYLE, J.L. Isolation of plant DNA from fresh tissue. **Focus**, v.12, n. 27, p. 13-15, 1990.

DUMOLIN-LAPÈGUE, S.; DEMESURE, B.; FINESCHI, S.; CORRE, V.L.; PETIT, R.J.
Phylogeographic structure of white oaks throughout the european continent.
Genetics, v.146, n. 4, p. 1475-1487, 1997.

EDGINGTON, E. S. Randomization tests. In: KOTZ, S.; JOHNSON, N.L. (Eds.).
Encyclopedia of statistical sciences 7. New York: Wiley, 1986. p.530-538.

FERRIS, C.; OLIVER, R.P.; DAVY, A.J.; Hewitt, G.M. Native oak chloroplasts reveal an
ancient divide across Europe. **Mol. Ecol.**, v.2, n. 3, p. 337-344, 1993.

FOSTER, P.; HARDING, R.; TORRONI, A.; BANDELT, H.-J. Origin and evolution of
Native American mtDNA variation: A reappraisal. **Am. J. Hum. Genet.**, v.59, n. 4, p.
935-945, 1996.

GARCÍA-MORENO, J.; ARCTANDER, P.; FJELDSÅ, J. A case of rapid diversification in
the Neotropics. Phylogenetic relationships of *Cranioleuca* spinetails (aves,
Furnariidae). **Mol. Phyl. Evol.**, v.12, n. 3, p. 273-281, 1999.

GASCON, C.; MALCOLM, J.R.; PATTON, J.L.; SILVA, M.N.F.; BOGART, J.P.;
LOUGHEED, S.C.; PERES, C.A.; NECKEL, S.; BOAG, P.T. Riverine barriers and the
geographic distribution of Amazonian species. **Proc. Natl. Acad. Sci.**, v.97, n. 25, p.
13672-13677, 2000.

GILPIN, M. Forty-eight parrots and the origin of population viability analysis. **Conserv.**
Biol., v.10, n. 6, p. 1491-1493, 1996.

GUGERLI, F.; SENN, J.; ANZIDEI, M.; MADAGHIELE, A.; BÜCHLER, U.; SPERINSEN, C.; VENDRAMIN, G.G. Chloroplast microsatellites and mitochondrial *nad1* intron 2 sequences indicate congruent phylogenetic relationships among Swiss stone pine (*Pinus cembra*), Siberian stone pine (*Pinus sibirica*), and Siberian dwarf pine (*Pinus pumila*). **Mol. Ecol.**, v.10, n. 6, p. 1489-1497, 2001.

HACHTEL, W. Maternal inheritance of chloroplast DNA in some *Oenothera* species. **J. Hered.**, v.71, n. 3, p. 191-194, 1980.

HOELZEL, A.R.; HALLEY, J.; O'BRIEN S.J.; CAMPAGNA, C.; ARNBOM, T.; LE BOEUF, B.; RALLS, K.; DOVER, G.A. Elephant seal genetic variation and the use of simulation models to investigate historical population bottlenecks. **J. Hered.**, v.84, n. 6, p. 4430-449, 1993.

HUDSON, R.R.; BOSS, D.D.; KAPLAN, N.L. A statistical test for detecting geographic subdivision. **Mol. Biol. Evol.**, v.9, n. 1, p. 138-151, 1992.

KIM, K.-J.; JANSEN, R.K. A chloroplast DNA phylogeny of lilacs (*Syringa*, Oleaceae): plastome groups show a strong correlation with crossing groups. **Am. J. Bot.**, v.85, n. 9, p. 1338-1351, 1998.

LEITE, S.A.A. **Filogeografia de *Caryocar villosum* (Aubl.) Pers. (CARYOCARACEAE) no Baixo Rio Negro (AM) e subsídios para a sua conservação.** Brasília, DF: UCB, 2002. 64 f. Dissertação (Mestrado em Ciências Genômicas e Biotecnologia) – Universidade Católica de Brasília, Brasília.

- LESSIOS, H.A.; KESSING, B.D; ROBERTONS, D.R.; PAULAY, G. Phylogeography of the pantropical sea urchin *Eucidaris* in relation to land barriers and ocean currents. **Evolution**, v.53, n. 3, p. 806-817, 1999.
- LEVY, F.; NEAL, C.L. Spatial and temporal genetic structure in chloroplast and allozyme markers in *Phacelia dubia* implicate genetic drift. **Heredity**, v.82, n. 4, p. 422-431, 1999.
- LIMA, P.S. G. **Divergência Genética e efeito do Nitrogênio Total no Crescimento *in vitro* de ipeca [*Psychotria ipecacuanha* (Brot.) Stokes]**. Lavras, MG: UFLA, 2002. 83 f. Tese (Mestrado em Genética e Melhoramento de Plantas). Universidade Federal de Lavras, Lavras.
- MAGUIRE, T.L.; CONRAN, J.G.; COLLINS, G.G.; SEDGLEY, M. Molecular analysis of interspecific and intergeneric relationships of Banksias using RAPDs and non-coding chloroplast DNA sequences. **Theor. Appl. Genet.**, v.95, p. 253-260, 1997.
- MARTINS, E.R. **Conservação da Poaia (*Psychotria ipecacuanha*): coleta, ecogeografia, variabilidade genética e caracterização reprodutiva**. Campos dos Goytacazes, RJ: UENF, 2000. 109 f. Tese (Doutorado em Produção Vegetal). Universidade Estadual do Norte Fluminense, Campos dos Goytacazes.
- MATOCQ, M.D. Phylogeographical and regional history of the dusky-footed woodrat, *Neotoma fuscipes*. **Mol. Ecol.**, v.11, n. 2, p. 229-242, 2002

McCAULEY, D.E. Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: Implications for studies of gene flow in plants. **Proc. Natl. Acad. Sci.**, v.91, n. 17, p. 8127-8131, 1994.

McCAULEY, D.E. The use of chloroplast DNA polymorphism in studies of gene flow in plants. **Trends Ecol. Evol.**, v.10, n. 5, p. 198-202, 1995.

MEDONÇA, M.P.; LINS, L.V. **Lista vermelha das espécies ameaçadas de extinção da flora de Minas Gerais**. Belo Horizonte: Fundação Biodiversitas, 2000. 157 p. (Fundação Zôo-Botânica de Belo Horizonte).

MENOTTI-RAYMOND, M; O'BRIEN, S.J. Dating the genetic bottle-neck of the African cheetah. **Proc. Natl. Acad. Sci.**, v.90, n. 8, p. 3172 - 3176, 1993.

OLIVEIRA, L.O.; MARTINS, E.R. **Os desafios das plantas medicinais brasileiras: I - o caso da poaia (*Cephaelis ipecacuanha*)**. Campos dos Goytacazes, RJ: UENF/FENORTE, 1998. 73 p.

OLIVEIRA, L.O.; MARTINS, E.R. A quantitative assessment of genetic erosion in ipecac (*Psychotria ipecacuanha*). **Genet. Res. Crop Evol.**, v.49, n. 6, p. 607-617, 2002.

PALMER, J.D. Comparative organization of chloroplast genomes. **Annu. Rev. Genet.**, v.19, p. 325-354, 1985.

- PALMER, J. D. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. **Amer. Natur.**, v.130, p. 6-29, 1987.
- PALMER, J.D.; JANSEN, R.K.; MICHAELS, H.J.; CHASE, M.W.; MANHART, J.R. Chloroplast DNA variation and plant phylogeny. **Ann. Missouri Bot. Gard.**, v.75, p. 1180-1206, 1989.
- PAULO, O.S.; JORDAN, W.C.; BRUFORD, M.W.; NICHOLS, R.A. Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian Lizard. **Mol. Ecol.**, v.11, n. 4, p. 809-819, 2002.
- PINTO, C.M.D. A ipecacuanha. In: Simposio Internacional sobre plantas de interesse econômico de la Flora Amazônica, 1976, Belém: IICA, 1976. p. 109-119.
- POSADA, D.; CRANDALL, K.A.; TEMPLETON, A.R. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. **Mol. Ecol.**, v.9, n. 4, p. 487-488, 2000.
- PRANCE, G.T. The conservation of botanical diversity. In: MAXTED, N.; FORD-LLOYD, B.V.; HAWKES, J.G.; CHAPMAN, H. (Eds.). **Plant genetic conservation**. London: Boundary Row, 1997. p. 3-11.
- PROVAN, J.; SORANZO, N.; WILSON, N.J.; GOLDSTEIN, D.G.; POWELL, W. A low mutation rate for chloroplast microsatellites. **Genetics**, v.153, n. 2, p. 943-947, 1999.

ROFF, D.A.; BENTZEN, P. The statistical analysis of mitochondrial DNA polymorphisms: Chi-square and the problem of small samples. **Mol. Biol. Evol.**, v.6, n. 5, p. 539-545, 1989.

ROSSI, A.A.B. **Biologia reprodutiva de *Psychotria ipecacuanha* (Brot.) STOKES.** Viçosa, MG: UFV, 2003. 61 f. Dissertação (Mestrado em Botânica) - Universidade Federal de Viçosa, Viçosa.

SALOMONE, N.; EMERSON, B.C.; HEWITT, G.M.; BERNINI, F. Phylogenetic relationships among the Canary Island Steganacaridae (Acari, Oribatida) inferred from mitochondrial DNA sequence data. **Mol. Ecol.**, v.11, n. 1, p. 79-89, 2002.

SCHAAL, B.A.; HAYWORTH, D.A.; OLSEN, K.M.; RAUSCHER, J.T.; SMITH, W.A. Phylogeographic studies in plants: problems and prospects. **Mol. Ecol.**, v.7, n. 4, p. 465-474, 1998.

SCHNABEL, A.; WENDEL, J.F. Cladistic biogeography of *Gleditsia* (Leguminosae) based on *NDHF* and *RPL16* chloroplast gene sequences **Am. J. Bot.**, v.85, n. 12, p. 1753-1765, 1998.

SEARS, B.B. Elimination of plastids during spermatogenesis and fertilization in the plant kingdom. **Plasmid**, v.4, n. 3, p. 233-255, 1980.

- SKORUPA, L.A.; M.C. ASSIS. Collecting and conserving Ipecac (*Psychotria ipecacuanha*, Rubiaceae) germplasm in Brazil. **Econ. Bot.**, v.52, n. 2, p. 209-210, 1998.
- SOLTIS, D.E.; MAYER, M.; SOLTIS, P.S.; EDGERTON, M. Chloroplast-DNA variation in *Tellima grandiflora* (Saxifragaceae). **Am. J. Bot.**, v.78, p. 1379-1390, 1991.
- SOLTIS, D.E.; SOLTIS, P.S.; RANKER, T.A.; NESS, B.D. Chloroplast DNA Variation in a Wild Plant, *Tolmiea menziesii*. **Genetics**, v.121, n. 4, p. 819-826, 1989.
- TABERLET, P.; GIELLY, L.; PAUTOU G.; BOUVET, J. Universal primers for amplification of three non-coding regions of chloroplast DNA. **Plant Mol. Biol.**, v.17, n. 5, p. 1105-1109, 1991.
- TEMPLETON, A.R. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. **Mol. Ecol.**, v.7, n. 4, p. 381-397, 1998.
- TEMPLETON, A. R. Using phylogeographic analyses of gene trees to test species status and processes. **Mol. Ecol.**, v.10, n. 3, p. 779-791, 2001.
- TEMPLETON, A.R. Statistical phylogeography: methods of evaluating and minimizing inference errors. **Mol. Ecol.**, 2004. Disponível em: <<http://www.blackwell-synergy.com/link/doi/10.1046/j.1365-294X.2003.02041.x/full>>. Acesso em: 25 jan. 2004.

TEMPLETON, A.R.; BOERWINKLE, E.; SING, C.F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and analysis of alcohol dehydrogenase activity in *Drosophila*. **Genetics**, v.117, n. 2, p. 343-351, 1987.

TEMPLETON, A.R.; CRANDALL, K.A.; SING, C.F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. **Genetics**, v.132, n. 2, p. 619-633, 1992.

TEMPLETON, A.R.; SING, C.F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. **Genetics**, v.134, n. 2, p. 659-669, 1993.

TEMPLETON A.R.; SING, C.S.; KESSLING, A.; HUMPHRIES, S. A cladistic analysis of phenotype associations with haplotype inferred from restriction endonuclease mapping. II. The analysis of natural populations. **Genetics**, v.120, n. 4, p. 1145-1154, 1988.

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, n. 2, p. 767-782, 1995.

TREASE, G.E.; EVANS, W.C. **Pharmacognosy**. London: Baillière Tindal, 1989. p. 450.

THORPE, R. S.; MALHOTRA, A.; BLAC, H.; DALTRY, J. C.; WÜSTER, W. Relating geographic pattern to phylogenetic process. **Phil. Trans. Roy. Soc. Lond. B**, v.349, p. 61-68, 1995.

VELOSO, H.P. As condições ecológicas da *Cephaelis ipecacuanha* Rich. **Mem. Inst. Oswaldo Cruz**, v.45, p. 361-372, 1947.

WEIR, B.S. **Genetic data analysis II**. Sunderland: Sinauer Associates, 1996. p. 454.

WILSON, E.O. (Ed.). **Biodiversity**. Washington: National Academy Press, 1988. p. 521.

WILSON, E.O. **The diversity of life**. London: Penguin Books Ltd., 1992. p. 424.

WOLFE, K.H.; LI, W.-H.; SHARP, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. **Proc. Natl. Acad. Sci.**, v.84, n. 24, p. 9054-9058, 1987.

ANEXO I

Psychotria ipecacuanha cpDNA

15

1 Una

2 39 22 00 S 15 20 00 W

2 Marliéria

8 42 60 00 S 19 65 00 W

3 Visconde do Rio Branco

3 42 91 00 S 21 01 00 W

4 Itaperuna

6 41 99 00 S 21 18 00 W

5 Ponte Nova

1 42 96 00 S 20 44 00 W

6 Guaraciaba

2 43 06 00 S 20 57 00 W

7 Carangola

1 42 31 00 S 20 68 00 W

8 Irupi

1 41 65 00 S 20 31 00 W

9 Conceição do Macabu

2 41 86 00 S 22 15 00 W

10 Barra do Bugres

3 57 16 00 S 15 26 00 W

11 Barra dos Bugres (Fazenda Raizama)

1 57 16 00 S 14 73 00 W

12 Barra dos Bugres (Área Indígena)

2 57 48 00 S 14 71 00 W

13 Barra dos Bugres (Manilha)

2 57 57 00 S 14 98 00 W

14 Salto do Céu

2 58 76 00 S 14 93 00 W

15 Curvelândia

4 57 41 00 S 16 11 00 W

4

Clade 1-1

2

VI VII

0 1

2

2 4

4 0

0 1

Clade 1-2

4

I II III V

0 1 1 0

9

1 2 3 4 5 6 7 8 9

2 2 3 4 1 1 1 0 2

0 0 0 0 0 0 0 1 0

0 0 0 0 0 1 0 0 0

0 2 0 0 0 0 0 0 0

Clade 2-1

3

1-1 1-2 1-3

1 0 1

9

1 2 3 4 5 6 7 8 9

0 4 0 1 0 0 0 0 0

2 4 3 4 1 2 1 1 2

0 0 0 1 0 0 0 0 0

Total Cladogram

2

2-1 2-2

1 1

15

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

2 8 3 6 1 2 1 1 2 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 3 1 2 2 2 4

END

ANEXO II

Inference Key for the Nested Haplotype Tree Analysis of Geographical Distances. (TEMPLETON, A.R. Statistical phylogeography: methods of evaluating and minimizing inference errors. **Mol. Ecol.**, 2004. Disponível em: <<http://www.blackwell-synergy.com/link/doi/10.1046/j.1365-294X.2003.02041.x/full>>. Acesso em: 25 jan. 2004.)

Start with haplotypes nested within a one-step clade and work up to clades nested within the total tree. If the tree is not rooted through an outgroup or if none of the clades nested at the total tree level have the sum of the outgroup probabilities of their haplotypes greater than or equal to 0.95, regard all clades nested at the total tree level as tips. When rooting is deemed reliable, interiors should also refer to the older clades in a nesting category, and tips to their evolutionary descendants.

This key is applied only if there are some significant values for D_C , D_N or I-T within the nesting clade. If there are no statistically significant distances within the clade, the null hypothesis of no geographical association of haplotypes cannot be rejected (either panmixia in sexual populations, extensive dispersal in nonsexual populations, small sample size or inadequate geographical sampling). In that case, move on to another clade at the same or higher level.

1. Are all clades within the nesting clade found in separate areas with no overlap?

- NO – Go to step 2.
- YES - Go to step 19.

2. Is at least one of the following conditions satisfied?

- a. The D_C 's for one or more tips are significantly small and the D_C 's for one or more of the interiors are significantly large or non-significant.

- b. The D_C 's for one or more tips are significantly small or non-significant and the D_C 's for some but not all of the interiors are significantly small.
 - c. The D_C 's for one or more interiors are significantly large and the D_C 's for the tips are either significantly small or non-significant.
 - d. The I-T D_C is significantly large.
- NO - Go to step 11.
 - YES - Go to step 3.
 - Tip/Interior Status Cannot be Determined - **Inconclusive Outcome.**

3. Is at least one of the following conditions satisfied?

- a. Some D_N and/or I-T D_N values are significantly reversed from the D_C values.
 - b. One or more tip clades show significantly large D_N 's.
 - c. One or more interior clades show significantly small D_N 's.
 - d. I-T has a significantly small D_N with the corresponding D_C value non-significant.
- NO - Go to step 4.
 - YES - Go to step 5.

4. Are both of the following conditions satisfied?

- a. The clades (or 2 or more subsets of them) with significantly small D_C values have ranges that are completely or mostly non-overlapping with the other clades in the nested group (particularly interiors).
- b. The pattern of restricted ranges represents a break or reversal from lower level trends within the nested series (applicable to higher-level clades only).

NO - Restricted Gene Flow with Isolation by Distance (Restricted Dispersal by Distance in Non-sexual species). This inference is strengthened if the clades with restricted distributions are found in diverse locations, if the union of their ranges roughly

corresponds to the range of one or more clades (usually interiors) within the same nested group (applicable only to nesting clades with many clade members or to the highest level clades regardless of number), and if the D_C values increase and become more geographically widespread with increasing clade level within a nested series (applicable to lower level clades only).

- YES - Go to step 9.

5. Are both of the following conditions satisfied?

a. The clades (or 2 or more subsets of them) with significantly small D_C values have ranges that are completely or mostly non-overlapping with the other clades in the nested group (particularly interiors).

b. The pattern of restricted ranges represents a break or reversal from lower level trends within the nested series (applicable to higher-level clades only).

- NO - Go to step 6.

- YES - Go to step 15.

6. Are either of the following conditions satisfied?

a. Clades (or haplotypes within them) with significant reversals or significant D_N values without significant D_C values define two or more geographically concordant subsets.

b. Clades (or haplotypes within them) with significant reversals or significant D_N values without significant D_C values are geographically concordant with other haplotypes/clades showing similar distance patterns?

- No - Go to step 7.

- YES - Go to step 13.

- **TOO FEW CLADES (< 2) TO DETERMINE CONCORDANCE – Insufficient Genetic Resolution to Discriminate between Range Expansion/Colonization and Restricted Dispersal/Gene Flow** - Proceed to step 7 to determine if the geographical sampling is sufficient to discriminate between short versus long distance movement.

7. Are the clades with significantly large D_N 's (or tip clades in general when D_N for I-T is significantly small) separated from the other clades by intermediate geographical areas that were sampled?

- **NO** - Go to step 8.

- **YES - Restricted Gene Flow/Dispersal but with some Long Distance Dispersal.**

8. Is the species absent in the non-sampled areas?

- **NO - Sampling Design Inadequate to Discriminate between Isolation by Distance (Short Distance Movements) versus Long Distance Dispersal.**

YES - Restricted Gene Flow/Dispersal but with some Long Distance Dispersal over Intermediate Areas not Occupied by the Species; or Past Gene Flow Followed by Extinction of Intermediate Populations.

9. Are the different geographical clade ranges identified in step 4 separated by areas that have not been sampled?

- **NO - Allopatric Fragmentation.** (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps).

- **YES** - Go to step 10.

10. Is the species absent in the non-sampled areas?

- NO - **Geographical Sampling Scheme Inadequate to Discriminate Between Fragmentation and Isolation by Distance.**
- YES - **Allopatric Fragmentation.** (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps).

11. Is at least one of the following conditions satisfied?

- a. The D_C value(s) for some tip clade(s) is/are significantly large.
 - b. The D_C value(s) for all interior(s) is/are significantly small.
 - c. The I-T D_C is significantly small.
- NO - Go to step 17
 - YES - **Range Expansion**, go to step 12.

12. Are the D_N and/or I-T D_N values significantly reversed from the D_C values?

- NO - **Contiguous Range Expansion.**
- YES - Go to step 13.

13. Are the clades with significantly large D_N 's (or tip clades in general when D_N for I-T is significantly small) separated from the geographical center of the other clades by intermediate geographical areas that were sampled?

- NO - Go to step 14.
- YES - **Long Distance Colonization Possibly Coupled with Subsequent Fragmentation** (subsequent fragmentation is indicated if the clades displaying restricted but at least partially non-overlapping distributions are mutationally connected to one

another by a larger than average number of steps) or Past Fragmentation Followed by Range Expansion. To see if secondary contact is involved, perform the supplementary tests given in Templeton, *Molecular Ecology* 10: 779-791, 2001. To discriminate the type of movement leading to this pattern, go to step 21.

14. Is the species present in the intermediate geographical areas that were not sampled?

YES - Sampling Design Inadequate to Discriminate between Contiguous Range Expansion, Long Distance Colonization, and Past Fragmentation.

- **NO - Long Distance Colonization and/or Past Fragmentation** (not necessarily mutually exclusive). If inferred at a high clade level, fragmentation rather than colonization is inferred if the clades displaying restricted but at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps. If the branch lengths are short, a colonization event is inferred, perhaps associated with recent fragmentation. To discriminate the type of movement leading to this pattern, go to step 21.

15. Are the different geographical clade ranges identified in step 5 separated by areas that have not been sampled?

- **NO - Past Fragmentation and/or Long Distance Colonization** (not necessarily mutually exclusive). If inferred at a high clade level, fragmentation rather than colonization is inferred if the clades displaying restricted but at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps. If the branch lengths are short, a colonization event is inferred, perhaps associated with recent fragmentation. To discriminate the type of movement leading to this pattern, go to step 21.

- **YES - Go to step 16.**

16. Is the species present in the intermediate geographical areas that were not sampled?

- YES - Go to step 18.
- NO - **Allopatric Fragmentation.** If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps.

17. Are either of the following conditions satisfied?

- a. The D_N values for tip or some (but not all) interior clades are significantly small.
 - b. The D_N for one or more interior clades is/are significantly large.
 - c. The I-T D_n value is significantly large.
- NO - **Inconclusive Outcome.**
 - YES - Go to step 4.

18. Are the clades found in the different geographical locations separated by a branch length with a larger than average number of mutational steps.

- NO - **Geographical Sampling Scheme Inadequate to Discriminate Between Fragmentation, Range Expansion, and Isolation by Distance.**
- YES - **Geographical Sampling Scheme Inadequate to Discriminate Between Fragmentation and Isolation by Distance.**

19. Is the species present in the areas between the separated clades?

- NO – **Allopatric Fragmentation.** If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps.
- YES - Go to step 20.

20. Was the species sampled in the areas between the separated clades?

- **NO – Inadequate Geographical Sampling.**
- YES – Go to step 2.

21. Are all of the following true?

- a. Is it biologically realistic that the organism could have undergone long-distance movement?
 - b. Are the nested haplotypes that mark a potential long-distance colonization event within a clade that shows evidence of population growth by other methods (such as mismatch distributions)?
 - c. At the level of the entire cladogram, does the clade not inferred to have produced long-distance colonization not show evidence of past population growth with other methods?
- YES – **Long-distance movement.**
 - NO – **Insufficient evidence to discriminate between long-distance movements of the organism and the combined effects of gradual movement during a past range expansion and fragmentation.**