UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA

Evolução de *Cereus hildmannianus* (Cactaceae) no sul do Brasil

GISLAINE ANGÉLICA RODRIGUES SILVA

Ribeirão Preto, SP

2013

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA

Evolução de *Cereus hildmannianus* (Cactaceae) no sul do Brasil

GISLAINE ANGÉLICA RODRIGUES SILVA

Tese de Doutorado apresentada ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências.

Área de Concentração: Genética

Orientadora: Profª. Drª. Maura Helena Manfrin

Ribeirão Preto, SP

2013

Autorizo a divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Silva, Gislaine Angélica Rodrigues

Evolução de *Cereus hildmannianus* (Cactaceae) no sul do Brasil. Gislaine Angélica Rodrigues Silva; orientadora: Maura Helena Manfrin. Ribeirão Preto, 2013.

129 f.

Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto-USP. Área de Concentração: Genética.

1. Filogeografia. 2. *Cereus hildmannianus*. 3. DNA cloroplastidial. 4. gene nuclear *PhyC*. 5. Especiação.

Este trabalho foi realizado com auxílio financeiro da CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior através do Programa de Demanda Social do Departamento de Genética da FMRP-USP e do Programa Institucional de Bolsas de Doutorado Sanduíche no Exterior (PDSE) – Processo BEX: 9815/11-2.

FOLHA DE APROVAÇÃO

Gislaine Angélica Rodrigues Silva

Título da Tese: Evolução de *Cereus hildmannianus* (Cactaceae) no sul do Brasil

Tese de Doutorado apresentada ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências.
Área de Concentração: Genética

Aprovado em:

Banca Examinadora:

Prof.(a) Dr.(a)_____
Instituição:_____Assinatura:_____

Prof.(a) Dr.(a)_____
Instituição:_____Assinatura:_____

Prof.(a) Dr.(a)_____
Instituição:_____Assinatura:_____

Prof.(a) Dr.(a)_____
Instituição:_____Assinatura:_____

Prof.(a) Dr.(a)_____
Instituição:_____Assinatura:_____

Dedico à minha mãe Rute, minha irmã Aline e à minha sobrinha *Duda*.

Agradecimentos

À Profª. Drª. Maura Helena Manfrin (FFCLRP/USP) pela oportunidade de trabalhar com tantas questões interessantes em um só projeto e me proporcionar um aprendizado tanto intelectual como pessoal.

Ao Prof. Dr. Fábio de Melo Sene (FMRP/USP) pela oportunidade de participar do grupo de pesquisa do Laboratório de Genética Evolutiva.

Ao Prof. Dr. Evandro Marsola de Moraes por toda a orientação nas etapas de coleta e adequação do processamento do material.

Ao Prof. Dr. Alexandre Antonelli pela colaboração durante o estágio *sanduíche* realizado na Universidade de Gotemburgo, Suécia.

Ao Prof. Dr. Ademilson Espencer Egea Soares (FMRP/USP) pelo excelente trabalho que realiza na coordenação do Departamento de Genética.

Aos técnicos Paulo Ricardo Epifânio e Mendelson Mazucato (FMRP/USP) pela disponibilidade que sempre tiveram em me ajudar.

Aos colegas e aos amigos do Laboratório de Genética Evolutiva: Érica C.C. Silva-Bernardi, Cíntia G. Santos, Luis E. M. Bizzo, Mateus H. Santos, Rafael F. Fransak, Camila K. B. Santos, Rogério P. Mateus, Fernando F. Franco, Thaís C. Lavagnini, Camila Borgonove, Rafaela Rossetti, Natácia Evangelista de Lima e Dora Y. Barrios, pelos animados *coffee break*, pela amizade, apoio e aprendizado.

Ao Mateus Henrique Santos pela amizade, auxílio com as análises e discussões durante todo o desenvolvimento do meu projeto e pelo apoio profissional.

À Dora Y. Barrios e Thaís C. Lavagnini pela amizade e compreensão, por compartilharemtantos assuntos interessantes e agradáveis que só acrescentam em minha vida profissional e pessoal.

Ao Thomas, à Karine, ao Filipe, à Rose e ao casal Ivonete e Georgio pela adorável companhia e por todo o conforto que me deram na Suécia.

À Jaqueline Souto, Michele Kirner e Ana Rita Baptistella por fazerem me sentir em casa, em família, desde que cheguei a Ribeirão Preto.

Aos amigos que "reconheci" durante o doutorado: Amanda, Juli, Thiago Ribers, Ivan, Mauro, Ludmila, Juliana, Camila e Sara pelos vários momentos de risadas, festas e a companhia agradável durante esses anos.

À Mayara Marcon, Fernanda Errero, Araceli Lima, Caroline Emanuelle, Vanessa Falchett e Luciana Edling, amigos que mesmo de longe, conseguem se manter tão presentes na minha vida.

À minha família, pelo apoio que recebi.

À CAPES, pelas bolsas concedidas.

Ao CNPq, FAEPA, USP e ao Jardim Botânico de Gotemburgo (Suécia) pelo auxílio.

Ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, seus funcionários e secretárias pela ajuda com as pendências burocráticas e pela disponibilização dos recursos e espaço físico para a realização deste trabalho.

A curiosidade brotou da necessidade de conhecer para viver.

Unamuno

Silva, G. A. R. Evolução de *Cereus hildmannianus* (Cactaceae) no sul do Brasil. 2013. 129f. Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto-USP.

Resumo

Há controvérsia sobre os processos responsáveis pela atual distribuição de Florestas Tropicais Sazonalmente Secas (FTSS) na América do Sul. Este tipo de vegetação compreende uma grande proporção de todas as espécies neotropicais. Entender o que modela a sua distribuição pode fornecer novas perspectivas para a evolução deste bioma e contribuirpara os aspectos de sua conservação. O trabalho avaliou a evolução deste bioma no sul do Brasil, onde as FTSS e as Florestas Tropicais (FT) são amplamente intercaladas. Para isso, foi reconstruídaa história filogeográfica do cacto, *Cereus hildmannianus,* uma espécie característica e abundante das FTSS. Métodos de datação molecular, estrutura populacional e filogeografia foram realizadas para avaliar os eventos histórico-demográficospor meio de uma amostragem densa que compreendeu 24 populações e cerca de 150 amostrasde, pelo menos, uma dentre as seis regiões genômicas nuclear e cloroplastidiais selecionadas. A partir disso, foi investigado um possível cenário da dinâmica populacional de *C. hildmannianus.* Os resultados indicam uma separação da espécie em dois grupos principais (Φ_{ST}: 0,788) com eventos de expansão populacional: umem regiões costeiras e o outro no interior do sul do Brasil, concondante com a distribuição dos núcleos das FTSS. O tempo do ancestral comum mais recente de *C. hildmannianus*, há 2,56 milhões de anos, remete a especiação deste ao período pré-Glacial. Os resultados do padrão de distribuição de *C. hildmannianus* foram concordantes com as áreas de endemismo para outros táxons das FTSS. Os eventos de dispersão e de vicariância entre as FTSS e as FT podem estar associados às mudanças paleoclimáticas durante os períodos glaciais do Quaternário, promovendo eventos de retração/expansão nestas florestas. A compreensão desses padrões na história biogeográfica de populações naturais podem auxiliar futuros planos de conservação deste bioma, na América do Sul.

Silva, G. A. R. Evolution of *Cereus hildmannianus* (Cactaceae) in southern Brazil.2013. 129f. Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto-USP.

Abstract

There is controversy about the processes responsible for the current distribution of Seasonally Dry Tropical Forests (SDTF) in South America. This vegetation type comprises a large proportion of all Neotropical species. Understanding what shapes your distribution may provide new insights into the evolution of this ecosystem and contribute to aspects of conservation. The study evaluated the evolution of this biome in southern Brazil, where SDTF and Rainforests are widely interspersed. For this, we reconstructed the phylogeographic history of the cactus, *Cereus hildmannianus*, a kind of characteristic and abundant SDTF. Molecular dating methods, population structure and phylogeography were performed to evaluate the historical and demographic events through a dense sampling which comprised 24 populations and about 150 samples of at least one among the six nuclear and chloroplast genomic regions selected. From this, we investigated a possible scenario of population dynamics of *C. hildmannianus*. The results indicate a separation of the species into two main groups (ΦST: 0.788) with events of population expansion: one in coastal regions and the other inside the south of Brazil, concondante with the distribution of the nuclei of SDTF. The time of the most recent common ancestor of *C. hildmannianus*were 2.56 million years ago, this speciation refers to the pre-Glacial. The results of the distribution pattern *C. hildmannianus* were consistent with areas of endemism for other taxa of SDTF. The events of dispersal and vicariance between SDTF and Rainforests may be related to paleoclimatic changes during glacial periods of the Quaternary, promoting events shrinkage /expansion in these forests. Understanding these patterns in the biogeographic history of natural populations may aid future conservation plans this biome in South America.

Sumário

I. INTRODUÇÃO 14
1. CONSIDERAÇÕES INICIAIS 14
2. OS TÁXONS E OS EVENTOS HISTÓRICOS 14
3. O MODELO BIOLÓGICO *Cereus hildmannianus* (CACTACEAE) 18
4. MODELAGEM DE NICHO ECOLÓGICO E PALEOMODELAGEM 20
5. FILOGEOGRAFIA 21
6. MARCADORES MOLECULARES 22
II. OBJETIVOS 25
III. MATERIAL E MÉTODOS 26
1. AMOSTRAS 26
2. OBTENÇÃO DAS SEQUÊNCIAS 26
2.1. Extração de DNA 26
2.2. Isolamento, amplificação e sequenciamento 27
3. ANÁLISES ESTATÍSTICAS DOS DADOS 32
3.1. Alinhamento, caracterização de *indels* e modelo de substituição nucleotídica 32
3.2. Índices de diversidade genética 32
3.3. Análises demográficas 32
3.4. Filogeografia e estrutura populacional 35
3.5. Relógio molecular e tempo de divergência 37
3.6. Análise biogeográfica 39
3.7. Modelagem de nicho ecológico e paleomodelagem 39
IV. RESULTADOS 41
1. MATERIAL BIOLÓGICO 41
2. ANÁLISE DE DADOS 41
2.1. Índices de diversidade genética 45
2.2. Análises demográficas 45
2.4. Relógio molecular e tempo de divergência 59
2.5. Análise biogeográfica 62
2.6. Modelagem de nicho ecológico e paleomodelagem 66
V. DISCUSSÃO 68

1. A VARIABILIDADE GENÉTICA EM *Cereus hildmannianus* 68
2. TEMPO DE DIVERGÊNCIA MOLECULAR 71
3. HISTÓRIA POPULACIONAL 74
4. INCONGRUÊNCIA ENTRE OS MARCADORES 79
5. CENÁRIO DA DINÂMICA POPULACIONAL 80
VI. CONCLUSÕES 83
VII. REFERÊNCIAS BIBLIOGRÁFICAS 84
APÊNDICE 1 99

I. INTRODUÇÃO

1. CONSIDERAÇÕES INICIAIS

No processo de evolução das espécies estão envolvidos aspectos ecológicos e históricos. Em termos históricos, as influências dos ciclos glaciais e interglaciais do Pleistoceno, que começaram há aproximadamente 2,5 milhões de anos (Hewitt, 2011) tem sido muito discutido como um fator influente na genética das populações e distribuição geográfica atual de grupos animais e vegetais (Hoorn et al., 2011; Rull, 2011; Thomé et al., 2010; Carnaval et al., 2009; Pennington et al., 2009; Rull, 2008; Pennington et al., 2000; Prado e Gibbs, 1993).

Essa distribuição de táxons tem sido investigada na América do Sul. A América do Sul é caracterizada pela combinação de diversidade climática e ambiental que permitem uma grande diversificação da sua cobertura florestal, englobando desde as florestas úmidas até regiões semi-áridas (Hueck, 1972). Entre as duas principais florestas úmidas da América do Sul, a Floresta Amazônica e a Floresta Atlântica, há uma faixa de vegetação sazonalmente seca e aberta, orientada na direção nordeste-sudoeste, conhecida como Floresta Tropical Sazonalmente Seca (FTSS) (Ab´Saber, 2003; Hueck, 1972).

2. OS TÁXONS E OS EVENTOS HISTÓRICOS

A dinâmica das espécies e populações na faixa de vegetação sazonalmente seca e aberta tem sido proposta como consequência das mudanças climáticas globais (Pennington et al., 2000; Prado e Gibbs, 1993). A hipótese mais tradicional sugere que, durante as condições frias e secas dos períodos glaciais nos Neotrópicos, foi favorecida a expansão geográfica da vegetação xerofítica ou aberta sobre as florestas úmidas, que foram reduzidas para locais chamados de refúgios do Quaternário (Haffer, 1969) e durante os períodos interglaciais, com condições quentes e úmidas, a situação seria revertida (Pennington et al., 2000; Ab´Saber, 1977). Diversos estudos apoiam as suposições citadas acima, incluindo estudos geomorfológicos (Ab´Saber, 1977; Wang et al., 2004), palinológicos (de Oliveira et al., 1999; Ledru, 1993), espeleológicos (Auler et

al., 2004; Wang et al., 2004), paleomodelagem (Carnaval et al., 2009; Carnaval e Moritz, 2008), e estudos biogeográficos (Carnaval e Bates, 2007; Moraes et al., 2009; Pennington et al., 2000; Quijada-Mascareñas et al., 2007).

Neste sentido, a relevância das alterações paleoclimáticas durante o Quaternário, na origem e distribuição das espécies, tem sido muito discutida na literatura (Hoorn et al., 2011; Rull, 2011; Pennington et al., 2009, 2004; Colinvaux et al., 2001). Entretanto, em plantas, Pennington e colaboradores (2004) observaram um alto grau de endemismo na América do Sul, nas famílias Polygonaceae, Leguminosae, Anacardiaceae e Cactaceae, cujas origens são do final do Terciário, indicando que as mudanças no Pleistoceno podem não ter sido um fator direcional de especiação em plantas. Além disso, uma maior diversificação de gramíneas e suculentas (Arakaki et al., 2011) e a especiação em táxons das FTSS (Pennington et al., 2009; Caetano et al., 2008) são datadas ao pré-Pleistoceno.

As FTSS neotropicais ocorrem como áreas disjuntas, conhecidas por núcleos, espalhados por toda a região neotropical (Figura 1), sendo que esta distribuição fragmentada tem persistido ao longo do tempo (Pennington et al., 2009). Uma das considerações para explicar a atual distribuição fragmentada das Florestas Tropicais Sazonalmente Secas (FTSS) nos Neotrópicos, tem sido em relação às flutuações climáticas do Quaternário (Pennington et al., 2000). Dentre os núcleos das FTSS, o Domínio da Caatinga, no nordeste brasileiro, o núcleo de Missiones, nos sistemas dos rios da bacia do Paraná-Paraguai, e formações nos vales Andinos e Caribe, atualmente isolados, apresentam conexões florísticas sugerindo interconexão no passado, durante períodos frios e secos do Pleistoceno, formando uma unidade fitogeográfica coesiva denominada o Arco Pleistocênico, como parte das FTSS (Prado e Gibbs, 1993; Pennington et al., 2000). Além disso, de acordo com Pennington e colaboradores (2000), em épocas de baixa precipitação, é plausível que espécies da FTSS possam ter aumentado sua extensão, por exemplo, em áreas de cerrado, no Brasil Central. Entretanto, um trabalho com reconstruções paleoclimáticas das FTSS sugere uma antiga e fraca extensão, sendo que uma maior expansão ocorreu mais recentemente, durante o Holoceno, contrastando com a hipótese do Arco Pleistocênico (Werneck et at. 2011).

Recentemente, Pennington e colaboradores (2009) propuseram que, dentre as características que diferenciam as FTSS como um bioma, a ausência de um gradiente de diversidade latitudinal quando comparada às florestas tropicais deveriam ser consideradas em estudos de evolução e macroecologia, uma vez que as FTSS tem ampla

distribuição, porém fragmentada em núcleos, enquanto as florestas tropicais tem expansão equatorial. Além disso, a alta diversidade de espécie entre os núcleos das FTSS, contrastando o Arco-Pleistocênico sugerido por Prado e Gibbs (1993), sugere que essas espécies são ecologicamente confinadas (Linares-Palomino et al., 2010) o que reflete na idade dos eventos cladogenéticos dos táxons, a qual tem sido estimada em sua maioria, para o final do Terciário.

Uma grande diversidade e densidade de espécies de cactos ocorrem na diagonal aberta da América do Sul, incluindo os Domínios da Caatinga, Cerrado e Chaco (Prado e Gibbs, 1993). Além destas regiões, domínios adjacentes incluindo florestas, também contêm cactos (Taylor e Zappi, 2004) os quais são manchas de vegetação adaptada a condições xéricas, sendo encontrados em afloramentos rochosos ou solos arenosos (Pennington et al., 2000, Taylor, 2000). Esta distribuição é considerada remanescente da retração da vegetação xeromórfica em períodos interglaciais (Ledru et al., 1998; Behling & Hooghiemstra, 2000), em que supostamente ocorreu uma expansão das formações abertas sobre as grandes massas florestais. Entretanto, estas mudanças também estão relacionadas às condições topográficas do relevo sul-americano.

Após o predomínio de vegetação aberta no final Pleistoceno (23.000 – 11.000 anos atrás), as mudanças climáticas favoráveis à expansão de vegetação de florestas alteraram a paisagem no início do Holoceno. A Floresta Atlântica, adjacente à diagonal aberta da América do Sul, expandiu em direção norte/sul, promovendo a dispersão de vários táxons neotropicais relacionados a essa vegetação, de mode semelhante ao gênero *Passiflora*, na região Sul do Brasil (Mader et al., 2009).

Outro exemplo de táxon como potencial marcador de eventos históricos, tais comoas alterações climáticas adjacentes a diagonal de vegetação aberta, são as araucárias, na região Sudeste e Sul do Brasil (Behling, 2002; Behling et al., 2004). A ampla área de distribuição do cerrado no sudeste brasileiro durante no início do Holoceno foi substituída por Floresta Semi-decídua, devido a condições climáticas mais úmidas, além de Floresta Araucária em vales de rios (Behling, 1998, 2002), refletindo nos isolados de cerrado atuais. Além das araucárias, dados palinológicos demonstraram a presença de isolados de campos subtropicais dentro do domínio da Floresta Atlântica, no núcleo Curucutu, Parque Estadual da Serra do Mar (Sudeste do Brasil), sugerindo também uma antiga e ampla distribuição dessa vegetação de campo que foi naturalmente substituída, no fim do Holoceno, por vegetação de floresta, provavelmente devido à presença de maior umidade (Pessenda et al., 2009).

Essas alterações históricas na distribuição de áreas adjacentes têm sido inferidas por meio de dados florísticos, palinológicos e por isótopos de carbono, os quais mostram interconexões passadas e fragmentações relacionadas com alterações paleoclimáticas (Prado e Gibbs, 1993; Ledru et al., 1998, Prado, 2000; Pennington et al., 2000, Taylor e Zappi, 2004; Pessenda et al., 2009).

Neste contexto, as cactáceas, que são marcadores de áreas secas tanto no passado como no presente (Prado e Gibbs, 1993; Taylor e Zappi, 2004) são bons modelos biológicos para o estudo das alterações na paisagem devido a mudanças climáticas. As Cactaceae representam uma das famílias mais abundantes das FTSS, e por isso, têm sido citadas como exemplo, para essas hipóteses descritas acima sobre a história da distribuição da vegetação sazonalmente seca e aberta da América do Sul.

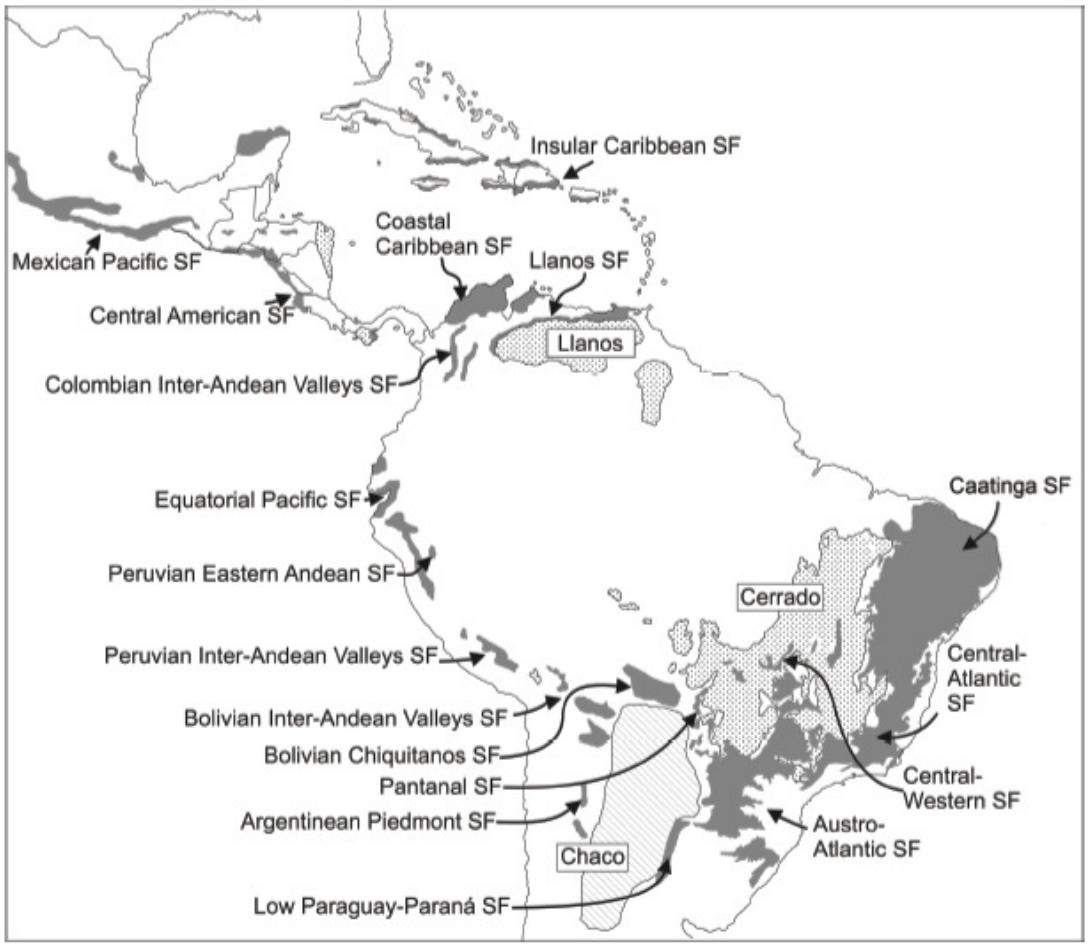


Figura 1. Mapa das áreas de ocorrência do bioma das Florestas Tropicais Sazonalmente Secasna América do Sul, representadas por seus núcleos em cinza escuro. Os três maiores núcleos dessas florestas são: a Caatinga no nordeste do Brasil; Missiones, na bacia dos rios Paraná-Paraguai; e sudoeste da Bolívia e noroeste da Argentina, além de toda a região dos vales andinos. Os outros biomas secos são mostrados em cinza claro, sendo o Chaco e as Savanas (Cerrado e Llanos). Mapa de Linares-Palomino e colaboradores (2011).

3. O MODELO BIOLÓGICO *Cereus hildmannianus* (CACTACEAE)

As Cactaceae compreendem 1.500-1.800 espécies em aproximadamente 100 gêneros (Barthlott e Hunt, 1993). A família Cactaceae é representada por plantas perenes, geralmente suculentas e altamente especializadas, apresentando um alto grau de evolução paralela em morfologia vegetativa e na estrutura floral (Barthlott e Hunt 1993). Os cactos constituem uma família de plantas predominantemente neotropical (Castro, 2008), com exceção da espécie epifítica *Rhipsalis baccifera* (J. S. Muell), que ocorre no Continente Africano, Madagascar e Sri Lanka (Barthlott, 1983; Barthlott e Taylor, 1995). Os cactos têm uma ampla gama de distribuição, ocorrendo desde o Sul da Patagonia na Argentina ao Canadá em vários habitats, incluindo desertos, extensões litorais arenosas, campos arbustivos, florestas secas decíduas, estepes alpinos, e até mesmo florestas tropicais (Barthlott e Hunt, 1993).

Os principais centros de diversidade conhecidos são as regiões áridas, como no México e sudeste dos Estados Unidos, os Andes na América do Sul e o leste do Brasil (Nyffeler, 2002; Ortega-Baes e Godinez-Alvarez, 2006), sendo o México o país com maior riqueza de espécies e endemismos. Embora a alta diversidade e densidade sejam encontradas em regiões áridas, os cactos podem crescer em diferentes habitats, desenvolvendo um importante papel devido a numerosas interações biológicas estabelecidas com outras plantas e animais (Ortega-Baes e Godinez-Alvarez, 2006).

Relações de parentesco e inferências sobre a história biogeográfica dos cactos têm sido estabelecidas com base em sequência de regiões gênicas cloroplastidiais e nucleares. A presença de espécies de *Rhipsalis* no Continente Africano permitiu alguns autores sugerir a origem das Cactaceae ao final do Cretáceo, há 65 - 90 milhões de anos, seguido pela quebra do supercontinente Gondwana (Nyffeler, 2002), embora o padrão de distribuição de *Rhipsalis baccifera* é considerado resultado da recente dispersão de longa distância por pássaros (Barthlott e Hunt, 1993). Trabalhos utilizando marcadores moleculares sugerem que as Cactaceae tiveram sua origem por volta de 30 milhões de anos atrás (Arakaki et al., 2011; Hershkovitz e Zimmer, 1997).

A Cactaceae compreende as subfamílias Pereskioideae, Opuntioideae e Cactoideae. Recentemente, o gênero *Maihuenia* foi removido de Pereskioideae e tem sido classificado como Maihuenioideae, uma nova subfamília (Wallace, 1995). As espécies de *Pereskia*, que são interpretadas como espécies relictuais, têm sido usadas como modelo de condição ancestral, por apresentarem características como presença de

folhas, sistema fotossintetizante C3, em contraposição às demais cactáceas, cujo sistema é CAM (Edwards et al., 2005). O estudo realizado por Nyffeler (2002) demonstra que as espécies de *Pereskia* não são parafiléticas ou monofiléticas, mas sugere um clado consistindo de Maihuenia, Opuntioideae e Cactoideae irmão de *Pereskia*. Porém, Edward e colaboradores (2005), sugerem uma divisão basal em Cactaceae, sendo *Pereskia* parafilética. A região geográfica de origem as subfamílias Opuntioideae e Cactoideae foram inferidas na metade sul da América do Sul, possivelmente na região central dos Andes (Edwards et al., 2005).

Relações filogenéticas (Figura 2) inferidas por meio de sequências de DNA, para a subfamília Cactoideae sugerem que o clado denominado BCT, compreendendo as tribos Browningieae, Cereeae e Trichocereeae, é derivado dentro desta subfamília (Nyffeler, 2002) e tem sua origem geográfica na região dos Andes (Ritz et al., 2007). Este clado compreende cerca de 30 gêneros e 400 espécies, sendo a maioria cactos colunares e globulares da América do Sul (Barthlott e Hunt, 1993; Nyfffeler, 2002).

Pertencente a tribo Cereeae, o gênero *Cereus* (tribo Cereeae, subfamília Cactoideae) compreende quatro subgêneros distribuídos na América do Sul, dentre os quais, o subgênero *Cereus* é encontrado em vários biomas, como Mata Atlântica, Caatinga, Cerrado e Campos rupestres (Taylor e Zappi, 2004). Neste gênero, *C. hildmannianus*é uma planta ereta, colunar, arbórea ou arbustiva, rupícola, de 8 a15 metros de altura. Seus cladódios, geralmente verde, são articulados com constrições de crescimento características. Os frutos são carnosos, oval-alongado com sementes obovadas. Sua floração pode ser registrada principalmente nos meses de outubro a janeiro e sua frutificação se estende pelo mesmo período (Taylor e Zappi, 2004; Bruxel e Jasper, 2005).

Toda a área de ocorrência de *Cereus hildmannianus* está associada principalmente ao bioma da FTSS, no núcleo de Missiones (Paraná-Paraguai) e na Floresta Sazonal Atlântico Sul (Figura 1). *Cereus hildmannianus* é encontrado em florestas úmidas e semiúmidas no sul do Brasil, com distribuição no leste, sudeste e sul do Brasil, centro e sudeste da América do Sul (Paraguai, Uruguai, Argentina) e leste do Chaco, associado a afloramentos rochosos ou solos que apresentam baixa umidade (Taylor e Zappi, 2004) e também na planície costeira do Rio Grande do Sul (Bauer e Waechter, 2006; Bruxel e Jasper, 2005; Gonçalves e Waechter, 2003). Muitas das populações de *C. hildmannianus* na região sul do Brasil são encontradas em matas de galeria, o que pode ser importante como corredor para dispersão.

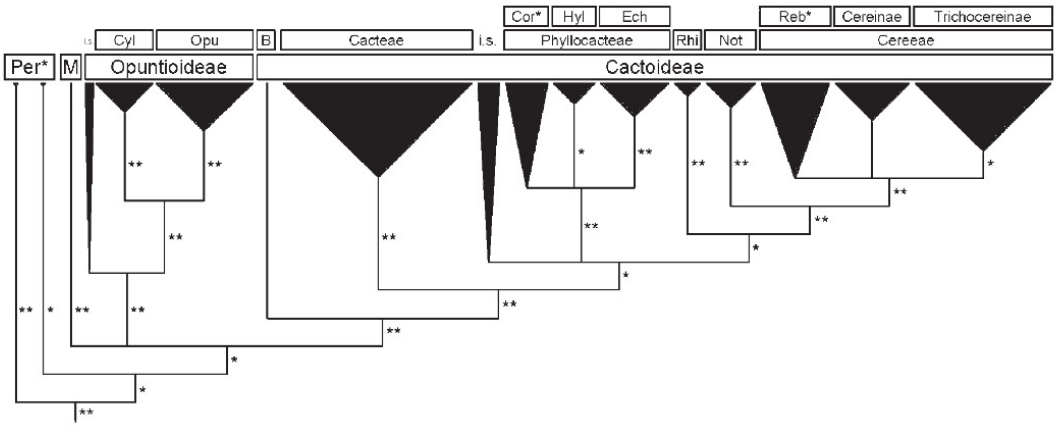


Figura 2. Diagrama da filogenia dos maiores clados em Cactaceae inferidos por análises filogenéticas de vários marcadores moleculares. Os triângulos indicam os subclados que estão representando subtribos, tribos ou subfamílias. Figura de Nyffeler e Eggli, 2010.

4. MODELAGEM DE NICHO ECOLÓGICO E PALEOMODELAGEM

A avaliação de dados macroecológicos, quer sob a forma de sensoriamento remoto ou interpolação de dados mensurados tem proporcionado o aumento de muitos estudos biológicos com informações do meio ambiente (Kozak et al., 2008). A modelagem parte de pontos previamente conhecidos para um determinado táxon e prediz uma distribuição potencial a partir da projeção das mesmas condições ambientais identificadas.

Os modelos de distribuição atual da espécie e os modelos condizentes com as alterações paleoclimáticas têm sido gerados para diversos táxons (Murray-Smith et al., 2008; Carnaval e Bates, 2007; Carnaval e Moritz, 2008), os quais permitem predições sobre como os eventos de flutuações climáticas influenciaram as expansões e retrações da vegetação e na dinâmica dos táxons associados a elas. Alguns autores assumem um modelo de nicho conservado quanto às relações abióticas e avaliam parâmetros de dimensões ecológicas e geográficas da distribuição de uma espécie para predições de futuros cenários geográficos, tanto em animais como em plantas (Cordellier e Pfenninger, 2009; Skov e Svenning, 2004; Thomas et al., 2004; Thuiller et al., 2006).

Para que as predições da modelagem de nicho ecológico sejam mais robustas deve-se considerar que a distribuição das espécies é determinada pelo ambiente e pelo conservadorismo de nicho (Waltari e Guralnick, 2009).

As críticas quanto à modelagem começam pela definição da teoria do nicho, sendo necessária uma interpretação cautelosa sobre o reconhecimento do nicho da

espécie em estudo, uma vez que a modelagem se baseia em pontos de ocorrência em um sub-espaço do nicho da espécie (De Marco Júnior e Siqueira, 2009). Além disso, a ampla variedade de técnicas para a realização da análise deve ser considerada, quanto menor a quantidade de dados, mais simples deve ser o modelo utilizado (De Marco Júnior e Siqueira, 2009).

Entretanto, apesar das críticas, a aplicação da modelagem de nicho ecológico para inferir a provável distribuição passada ou áreas de estabilidade para um táxon tem sido descritas na literatura. Em geral, os trabalhos utilizam a modelagem para complementaroutras análises, como por exemplo, análises filogeográficas e biogeográficas. Para anfíbios e lagartos foi sugerido refúgio, ou seja, áreas de estabilidade, em áreas de mata atlântica (Carnaval et al., 2009; Carnaval e Moritz, 2008), em mamíferos, os refúgios foram condizentes com os resultados das análises de filogeografia (Waltari et al., 2007) e para as FTSS, corroboraram áreas de estabilidade para a distribuição de espécies inferidas em táxons desta floresta (Werneck et al., 2011).

5. FILOGEOGRAFIA

A hipótese de alterações paleoclimáticas na distribuição da vegetação xerófita devido às mudanças na paisagem na América do Sul pode ser averiguada por meio de análises filogeográficas. Estudos filogeográficos associam genealogias de genes, incorporando um elemento temporal, com a distribuição geográfica das populações, permitindo a inferência de processos históricos envolvidos na determinação dos padrões atuais de distribuição e diversificação de populações e espécies (Avise, 2000). A distribuição espacial e temporal de alelos é influenciada por eventos recorrentes, tais como a mutação, o fluxo gênico, a deriva genética e a seleção natural e eventos históricos, como a expansão da área de ocorrência e a divisão populacional (Avise, 2000; Avise et al., 1987), sendo cada evento responsável por uma padrão filogeográfico distinto.

A abordagem filogeográfica tradicional conhecida por Análise Filogeográfica de Clado Aninhado (NCPA, *Nested Clade Phylogeographic Analysis*) converte uma árvore de haplótipos em uma série hierárquica de ramos ou clados utilizando um algoritmo de parcimônia estatística (Templeton et al., 1995). Embora a NCPA tenha sido a primeira

abordagem filogeográfica a utilizar a estatística, o termo "filogeografia estatística" foi introduzido por (Knowles e Maddison, 2002).

As críticas à NCPA quanto à discriminação dos eventos recorrentes e históricos se referem principalmente, as inferências falso positivas mais comuns, sendo elas, fluxo gênico restrito com isolamento por distância e expansão populacional (Panchall e Beaumont, 2007). Entretanto, Templeton (2004b, 2008) mostrou que a ocorrência de falsos positivos são menores em populações reais que em simuladas e sugeriu que a validação cruzada realizadas com o uso de mais de um marcador, pode minimizar esses equívocos, além de outras metodologias que podem ser testadas conjuntamente (Garrick et al., 2008).

A filogeografia estatística estima a inferência demográfica considerando a estocasticidade dos processos genéticos, ou seja, diferentes cenários históricos foram testados por simulações de coalescência gerando distribuições nulas (Knowles e Maddison, 2002). Esse método calcula a probabilidade dos parâmetros inferidos para a população utilizando a função de verossimilhança (Nielsen e Beaumont, 2009) ou inferência bayesiana (Beaumont e Rannala, 2004; Nielsen e Beaumont, 2009) por meio da abordagem das Cadeias de Markov Monte Carlo.

As inferências filogeográficas são baseadas na teoria da coalescência, que identifica o processo genealógico de uma amostra de genes em uma população até o ancestral comum mais recente (MRCA, *Most Recent Common Ancestor*) da amostra (Kingman 1982). A teoria da coalescência parte do presente para inferir o passado utilizando uma abordagem probabilística (Fu e Li, 1999).

6. MARCADORES MOLECULARES

Uma condição fundamental para estudos como esse é a prospecção de regiões genômicas com variação intraespecífica suficiente para traçar a genealogia dos diferentes haplótipos. A escolha de DNA cloroplastidial (cpDNA) para estudos filogeográficos em plantas tem sido realizada em detrimento ao genoma mitocondrial (mtDNA), o marcador molecular tradicional para estudos filogeográficos em animais. Isto ocorre devido às taxas evolutivas baixa do mtDNA das plantas, diferentemente das taxas encontradas no mtDNA animal, o que não gera variação suficiente para análises filogeográficas (Avise, 2009). O genoma do cloroplasto é haplóide, sendo uma única e

não-recombinante unidade de herança, com diferentes taxas de mutação (Schaal et al., 1998). A organização da molécula de cpDNA de plantas terrestres consiste de uma região de *Large Single Copy* (LSC) e *Small Single Copy* (SSC), separadas por duas regiões idênticas ou não idênticas de repetições invertidas, IR_A e IR_B. As repetições invertidas são estáveis e evoluem duas a três vezes mais lentamente quando comparadas com a região de cópia única, sendo a região de LSC a que apresenta os maiores níveis de variação (Ravi et al., 2008).

Além disso, o tamanho efetivo populacional do cpDNA e do mtDNA, é menor em relação ao DNA nuclear, o que reflete no tempo de coalescência dos haplótipos de cpDNA de até quatro vezes mais rápido, pois o DNA nuclear é diplóide e pode ser transmitido através dos dois sexos (Avise, 2009; Templeton, 2006).

Estas características do cpDNA permitem o seu uso em estudos populacionais e filogeográficos, sendo que Shaw et al. (2005, 2007) demonstraram que regiões nãocodificantes seriam uma potencial fonte de variação para estes estudos.O desenvolvimento de *primers* para cpDNA em regiões codificantes conservada ao lado de regiões não codificantes, como espaçadores intergênicos e íntrons é bastante utilizado. Isto permitiu o desenvolvimento de *primers* universais para o estudo em diversos grupos de plantas (Grivet et al., 2001, Borsch e Quandt, 2009).

Entretanto, marcadores nucleares de baixa ou única cópia têm sido utilizados para estudar casos de especiação envolvendo hibridação e subsequente alopoliploidização em plantas, além de contornar os problemas com a baixa taxa mutacional dos cpDNA encontrada para alguns táxons, sendo importante para os estudos populacionais (Zimmer e Wen, 2012). A maioria das árvores de espécies tem usado os genes do fitocromopara níveis taxonômicos inferiores, como o *PhyC* (Zimmer e Wen, 2012).

Com base no exposto, neste trabalho foirealizada uma análise da estrutura de populações e história demográfica do cacto *C. hildmannianus.* Estes estudos proporcionaram informações sobre a variação genética dentro e entre as populações, bem como a avaliação de barreiras ao fluxo gênico, que é o evento inicial no processo de diferenciação biológica de populações e de espécies, além de inferir eventos históricos associados às mudanças paleoclimáticas. Estas informações, quando comparadas com outros grupos biológicos, ajudam a avaliar os padrões contrastantes de expansão entre espécies, tanto para o núcleo de Missiones, como para os domínios

fitogeográficos da Floresta Atlântica, de Araucária e dos Campos Sulinos, no sul do Brasil.

II. OBJETIVOS

Este projeto tem como objetivo descrever a história evolutiva do cacto *Cereus hildmannianus*, tendo como metas:

- Analisar o nível de variação nucleotídica em regiões gênica e não codificantes (segmentos intergênicos e íntrons) para *C. hildmannianus*;
- Definir a estrutura populacional para *C. hildmannianus*;
- Realizar uma análise filogeográfica;
- Datar os eventos de diversificação em C. *hildmannianus* usando um relógio molecular relaxado e aspectos biogeográficos;
- Inferir eventos históricos e, associar tais eventos com a história do cacto *C. hildmannianus*;
- Inferir a provável distribuição de *C. hildmannianus* na região do núcleo de Missiones e planície costeira no Sul do Brasil, para o presente e passado, por meio de modelagem de nicho ecológico e paleomodelagem.

III. MATERIAL E MÉTODOS

1. AMOSTRAS

Para a obtenção de amostras de espécimes de cactos, foram coletados segmentos das pontas das raízes e armazenados em sacos plásticos. A coleta foi realizada observando uma distância mínima de dez metros entre os indivíduos, com o objetivo de diminuir a probabilidade de coleta de clones. Cada amostra foi processada manualmente para a retirada do excesso de terra e do súber e, então, acondicionada em tubos criogênicos e armazenada em botijão de nitrogênio líquido até o final da coleta. No laboratório o material foi armazenado em temperaturas de -80ºC. O material testemunho foi preparado a partir de cladódios prensados e secos em estufa, montados como exsicatas e armazenados no herbário SPFR da Universidade de São Paulo, Ribeirão Preto-SP. Além do registro fotográfico, características de valores taxonômicos e ambientais foram registradas para uma melhor averiguação da classificação taxonômica. Inicialmente os pontos de coleta foram definidos a partir de registros dos cadernos de coleta para *Drosophila* cactófilas, pertencentes ao Laboratório de Genética Evolutiva-USP-RP. As demais localidades foram exploradas nas saídas de campo, considerando a distribuição da espécie descrita por Taylor e Zappi (2004) para a região Sul e Sudeste do Brasil. As populações naturais de *C. hildmannianus* coletadas estão indicadas na Tabela 1 e o mapa com toda a distribuição destas populações está representado na Figura 3.

2. OBTENÇÃO DAS SEQUÊNCIAS

2.1. Extração de DNA

O procedimento para a extração do DNA utilizou, aproximadamente, 100 mg de tecido de raiz processado. Esse material foi previamente macerado em nitrogênio líquido (-196°C), cuja temperatura propicia a quebra mecânica da parede celular e inibe a degradação enzimática do DNA por metabólitos secundários. A extração do DNA foi realizada com auxílio do conjunto de reagentes Dneasy Plant Mini Kit Qiagen, com modificações no protocolo, que incluíram a adição de 600 uL de Buffer AP1 e aumento de um minuto em todas as etapas de centrifugação. A otimização da extração foi realizada considerando a alta concentração de polissacarídeos presente nos tecidos de cactáceas (Helsen et al., 2007; Ritz et

al., 2007; Nyffeler, 2002). O DNA extraído foi submetido à corrida eletroforética em gel de agarose 0,6%, corado com brometo de etídio e visualizado através de um transiluminador com luz ultravioleta, para teste de quantidade e qualidade do produto isolado, utilizando como padrão o peso molecular do marcador Low DNA Mass Ladder (Invitrogen). O DNA também foi quantificado com o aparelho NanoDrop® ND-1000 Spectrophotometer.

Tabela 1. Locais e datas de coletas de amostras de indivíduos das populações de *C. hildmannianus.*

		Coordenadas geográficas	
Localidade	**Data**	**Latitude**	**Longitude**
Itatiba-SP	11/04/2009	23º0'20,88"	46º50'20,04"
Caçapava do Sul-RS	24/05/2009	30º49'9,2"	53º30'46,9"
Jaguari-RS	25/05/2009	29º22'59,9"	54º44'38,9"
Santiago-RS	25/05/2009	29º02'7,7"	55º04'6,1"
Santa Maria-RS	26/05/2009	29º46'46,4"	53º46'39,3"
Viamão-RS	27/05/2009	30º14'04,6"	50º57'38,7"
Mostardas-RS	28/05/2009	31º06'36,9"	50º55'16,1"
Penha-SC	29/09/2009	26º46'31,2"	48º35'53,2"
Laguna-SC	29/09/2009	28º29'15,1"	48º46'43,8"
Florianópolis-SC	30/09/2009	27º37'11,5"	48º26'54,7"
Cianorte - PR	31/10/2009	23º39'21,8"	52º30'26,2"
Guarapuava-PR	28/04/2011	25º17'42"	51º53'7,1"
Cantagalo-PR	28/04/2011	25º25'0"	52º04'14,9"
Mangueirinha-PR	28/04/2011	25º46'27,2"	52º06'55,6"
Arroio do Sal-RS	17/07/2011	29º32'27"	49º55'26"
Osório-RS	18/07/2011	29º57'31,1"	50º13'30,9"
Barra do Ribeiro-RS	19/07/2011	30º16'19,4"	51º24'55,1"
Pouso Novo-RS	20/07/2011	29º13'03"	52º10'10,5"
Jacutinga-RS	20/07/2011	27º43'52,5"	52º31'11,9"
Serrana-SP	19/09/2011	21º15'33"	47º34'28"
Sengés-PR	01/11/2011	24º07'	49º23'
Jacarezinho-PR	02/11/2011	23º14'14,1"	50º02'19,4"
Piratininga-SP	02/11/2011	22º26'9,2"	49º07'57,8"
Itu-SP	11/11/2011	23º15'5,4"	47º13'22,6"

2.2. Isolamento, amplificação e sequenciamento

A escolha das regiões de DNA cloroplastidial para o presente estudo considerou valores elevados para os PICs (*potentially informative characters*), isto é, o número de substituições nucleotídicas, *indels* e inversões descritas na literatura como potenciais para estudos filogenéticos e filogeográficos em angiospermas (Shaw et al., 2007, Shaw et al.,

2005). Da região LSC do DNA cloroplastidial, foram testados 16 pares de *primers*, flanqueando as seguintes regiões: *trnH-psbA, trnS-trnG-trnG* (região intergênica *trnS-trnG* e íntron *trnG*), *trnT-trnL, trnL-trnL-trnF* (íntron *trnL* e região intergênica *trnL-trnF*), *psbJ-petA, atpI-atpH, 3'rps16-5'trnK, 3'trnk-matk, trnQ-5'rps16, psbD-trnT* e *petL-psbE* (Figura 4, Tabela 2). Para a região nuclear, o exon 1 do gene *PhyC* foi escolhido por ser um gene com baixo número de cópias e já ter sido utilizado em trabalhos com Cactaceae apresentando caracteres informativos (Arakaki et al., 2011; Helsen et al., 2009; Edwards et al., 2005). Os *primers* utilizados para cada região amplificada estão listados na Tabela 2, a amplificação das regiões *trnS-trnG* e *3'trnK-matK* foi realizada em duas etapas utilizando *primers* internos.

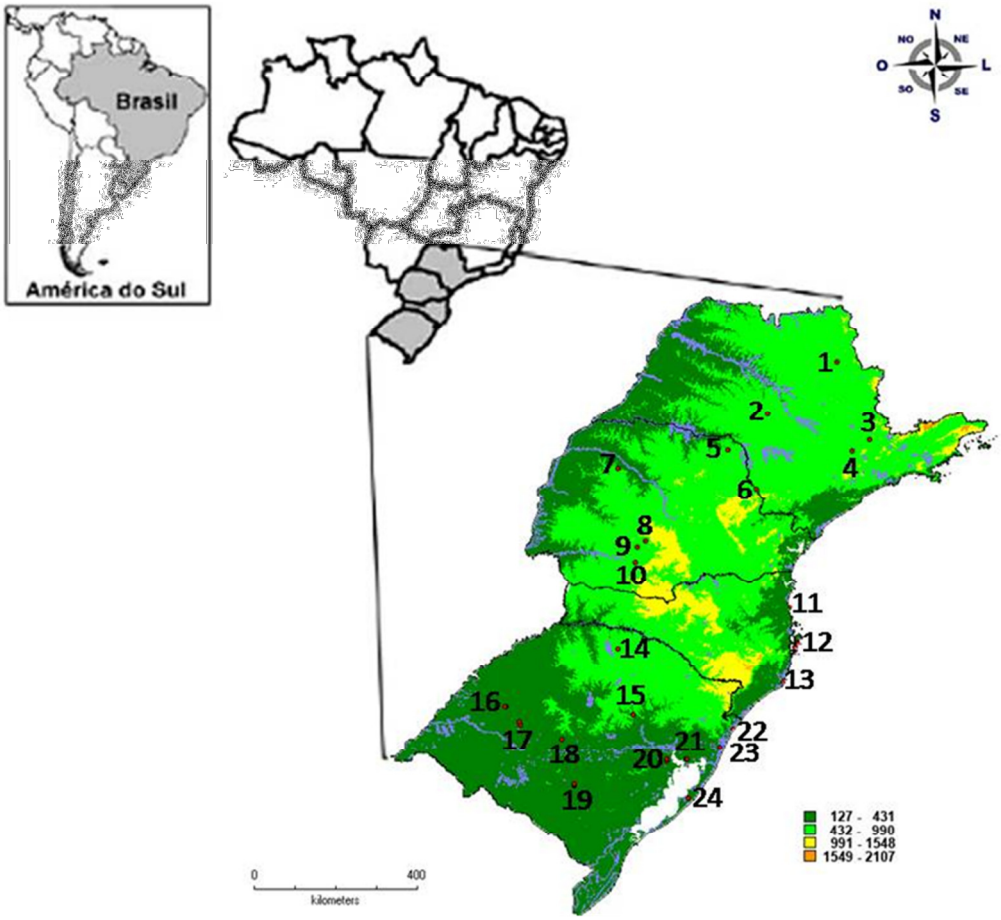


Figura 3. Mapa com indicação dos pontos de amostragem para *Cereus hildmannianus*. A legenda mostra as altitudes em metros relacionados às cores, obtidas e editadas em DIVA-GIS (Hijmans et al., 2005). Localidades: 1 – Serrana-SP; 2 – Piratininga-SP; 3 – Itatiba-SP; 4 – Itu-SP; 5 – Jacarezinho-PR; 6 – Sengés-PR; 7 – Cianorte-PR; 8 – Guarapuava-PR; 9 – Cantagalo-PR; 10 – Manguerinha-PR; 11 – Penha-SC; 12 – Florianópolis-SC; 13 – Laguna-SC; 14 – Jacutinga-RS; 15 – Pouso Novo-RS; 16 – Santiago-RS; 17 – Jaguari-RS; 18 – Santa Maria-RS; 19 – Caçapava do Sul-RS; 20 – Barra do Ribeiro-RS; 21 – Viamão-RS; 22 – Arroio do Sal; 23 – Osório-RS; 24 – Mostardas-RS.

As reações em cadeia da polimerase (PCR – *Polymerase Chain Reaction*) foram realizadas inicialmente, seguindo o protocolo proposto por Shaw et al. (2007). Entretanto, para otimizá-las, diferentes reações e condições de PCR foram testadas, utilizando um termociclador de gradiente Master Cycler Gradient (Eppendorf). O volume final de cada reação foi de 30 µl, contendo 1 µL de DNA genômico (10 – 60 ng/µl), 1Xtampão de reação (Invitrogen), 200µM de dNTP, 0,1µM de cada *primer*, $MgCl_2$ concentrado a 3 mM, 1,25 unidades de GoTaq®Flexi DNA Polymerase (Promega) e água livre de nuclease. Somente para as regiões *trnH-psbA* foi utilizado $MgCl_2$ a 2 mM e 0,4 unidades de Taq polimerase, e para o íntron *trnG* e *trnS-trnG* foram utilizados $MgCl_2$a 1,5 mM, sendo a concentração da Taq polimerase de 0,4 unidades para o íntron *trnG* e 1 unidade para *trnS-trnG*. As condições gerais de PCR foram de 95°C de temperatura de denaturação iniciala 2 mim, seguido por 35 ciclos de denaturação de 30s a 95°C, *annealing* (pareamento dos *primers*) de 1 mim a 50°C e extensão de 2 mim a 72°C, com uma extensão final de 5 mim a 72°C.Variações deste protocolo foram realizadas para a temperatura de *annealing* do gene *PhyC* (54°C). Para a região *trnS-trnG-trnG*, as condições foram 80°C de temperatura de denaturação iniciala 5 mim, seguido por 40 ciclos de denaturação de 1 mim a 95°C, *annealing* (pareamento dos *primers*) de 1 mim a 62°C e extensão de 5 mim a 65°C, com uma extensão final de 5 mim a 65°C. Após a padronização, as reações de amplificação foram realizadas em um termociclador Gene Amp PCR® System 9700 (Applied Biosystems). Os segmentos amplificados foram purificados com o kit *GFXTM PCR and Gel Band Purification (GE Healthcare)*, seguindo o protocolo do kit.

As reações de sequenciamento foram realizadas em ambas as fitas, *forward* e *reverse*. Para cada fita, foram utilizados 50 ng de produto de PCR purificado, 10 µM de *primer*, 1,5 µL de tampão de sequenciamento 5X (kit BigDye) e 1 µL de BigDye® Terminator v 3.1 Cycle Sequencing kit (Applied Biosystems), completadas para um volume final de 10 µLcom água livre de nuclease. Essas reações de sequenciamento foram realizadas em um termociclador Gene Amp PCR® System 9700 (Applied Biosystems), com as seguintes condições: uma etapa de 1 mim a 96°C, seguido por 39 ciclos de 15 s a 96°C, 15 s a 50°C (correspondente à temperatura de *annealing* de cada par de *primer*), e uma etapa final de 4 mim a 60°C. Para a precipitação das amostras, foram adicionados às reações 80 µL de isopropanol 65 %, que permaneceram por 15 mim no escuro à temperatura ambiente, seguidos de centrifugação a 4000 rpm a 4°C por 45 mim. Este procedimento foi finalizado com o descarte do sobrenadante. Por duas vezes, 200 µL de etanol 70 % gelado foram adicionados, seguido de centrifugação a 4000 rpm a 4°C por 10 mim e descarte do

sobrenadante. Finalizado com uma etapa de *spin* reverso a 1000 rpm a 4°C por 1 mim. As amostras precipitadas foram deixadas a 37°C por 1 hora. Após a precipitação, o sequenciamento foi realizado em ambas as fitas, no sequenciador automático ABI 3730 XL (Applied Biosystems).

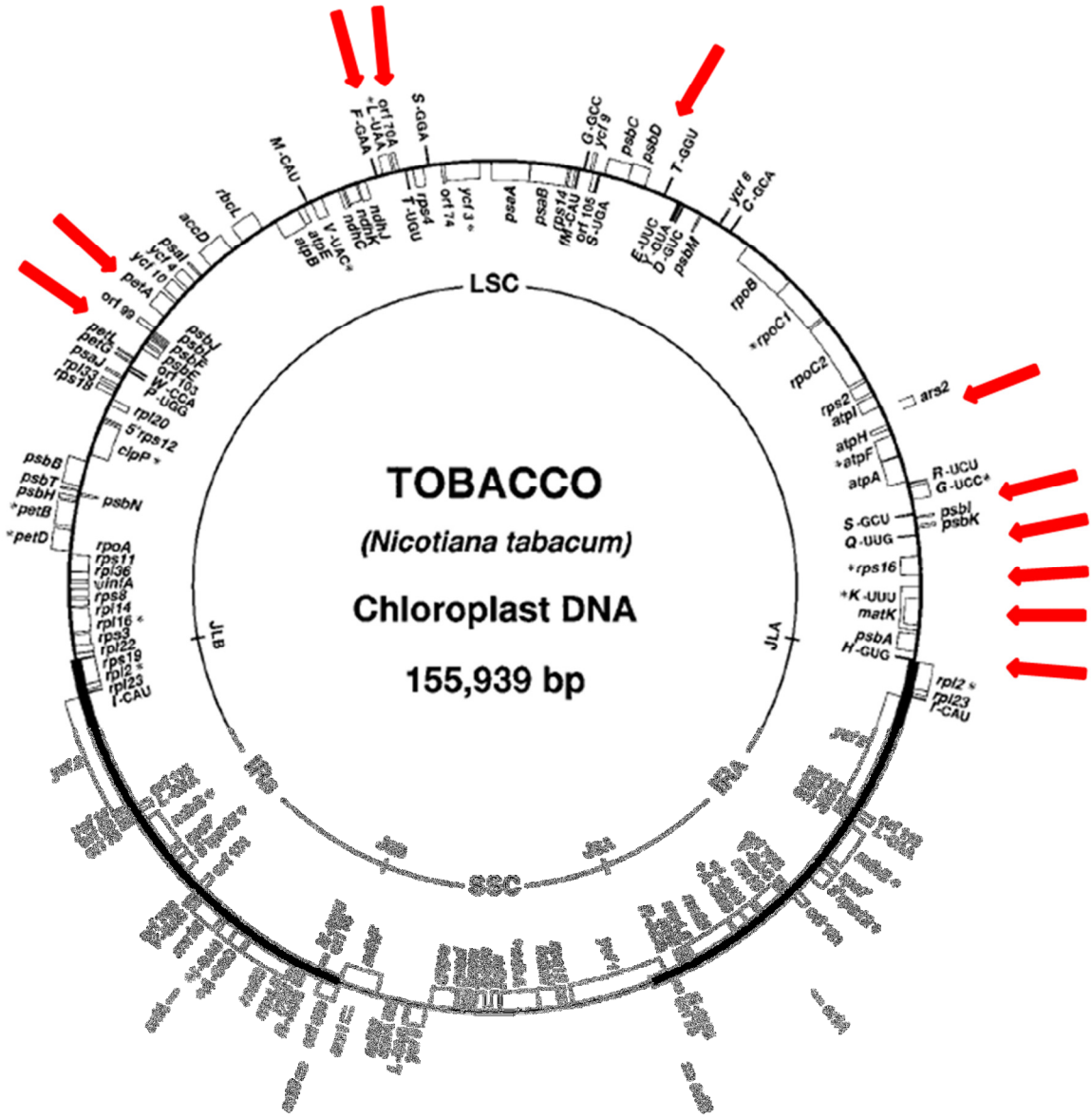


Figura 4. Mapa genético do cpDNA de tabaco (Wakasugi et al., 1998). As setas indicam as regiões cloroplastidiais utilizadas nos indivíduos de *C. hildmannianus*.

Tabela 2. Sequências dos pares de *primers* utilizados na amplificação das regiões intergênicas de DNA cloroplastidial e nuclear em *C. hildmannianus*.

Segmento	*Primers*	Referência
Nuclear		
phyC	**phyC-R** TCCTCCACTTGACCACCTCT **phyC-F** AGCTGGGGCTTTCAAATCTT	Helsen et al. (2009)
cpDNA		
trnH-psbA	**trnH** $^{(GUG)}$ ACT GCC TTG ATC CAC TTG GC **psbA** CGA AGC TCC ATC TAC AAA TGG	Hamilton (1999)
íntron *trnG*	**5'trnG2G** GCG GGT ATA GTT TAG TGG TAA AA **trnG** $^{(UUC)}$ GAA TCG AAC CCG CAT CGT TAG	Shaw et al. (2005) Shaw et al. (2007)
trnS-trnG	**SG Rev 2** TCC GCT CAT TAG CTC TCC TC **trnS**$^{(GCU)}$ AAC TCG TAC AAC GGA TTA GCA ATC **SG Fwd 2** CAC CCA TGG TTC CCA TTA GA	Bonatelli (2010) Shaw et al. (2007) Bonatelli (2010)
trnT-trnL	**5'trnG2S** TTT TAC CAC TAA ACT ATA CCC GC **5'trnL**^{UAA}**R (TabB)** TCT ACC GAT TTC GCC ATA TC **trnT**^{UGU}**F (TabA)** CAT TAC AAA TGC GAT GCT CT	Shaw et al. (2005) Taberlet et al. (1991)
trnL-trnF	**trnF**^{GAA} **(TabF)** ATT TGA ACT GGT GAC ACG AG **3´trnL**^{UAA}**R (TabE)** GGT TCA AGT CCC TCT ATC CC	Taberlet et al. (1991)
íntron *trnL*	**trnL5'**^{UAA}**F (TabC)** CGA AAT CGG TAG ACG CTA CG **3'trnL**^{UAA}**R (TabD)** GGG GAT AGA GGG ACT TGA AC	Taberlet et al. (1991)
psbJ-petA	**psbJ** ATA GGT ACT GTA RCY GGT ATT **petA** AAC ART TYG ARA AGG TTC AAT T	Shaw et al. (2007)
atpI-atpH	**atpI** TAT TTA CAA GYG GTA TTC AAG CT **atpH** CCA AYC CAG CAG CAA TAA C	Shaw et al. (2007)
3'rps16–5'trnK	**rpS16x2F2** AAA GTG GGT TTT TAT GAT CC **trnK**$^{(UUU)}$**x1** TTA AAA GCC GAG TAC TCT ACC	Shaw et al. (2007)
3'trnk-matk	**matK50-Fdi** GTT TTG ACT GTA TCG CAC TAT GTA TC **trnK-41R** ATG GAT TTT TG**G** G**G**A GTA AT**A** AGA C **ACmatk500F** TTC TTC TTT GCA TTT ATT ACG **trnK-71R** CTA ATG GGA TGT CCT AAT AC	Demaio et al. (2011) Nyffeler (2002)* Müller e Borsch (2005) Nyffeler (2002)
petL-psbE	**petL** AGT AGA AAA CCG AAA TAA CTA GTT A **psbE** TAT CGA ATA CTG GTA ATA ATA TCA GC	Shaw et al. (2007)
trnQ-5'rps16	**trnQ(UUG)** GCG TGG CCA AGY GGT AAG GC **rpS16x1** GTT GCT TTY TAC CAC ATC GTT T	Shaw et al. (2007)
psbD-trnT	**psbD** CTC CGT ARC CAG TCA TCC ATA **trnT(GGU)-R** CCC TTT TAA CTC AGT GGT AG	Shaw et al. (2007)

*: modificado de Nyffeler (2002).

3. ANÁLISES ESTATÍSTICAS DOS DADOS

3.1. Alinhamento, caracterização de *indels* e modelo de substituição nucleotídica

Os cromatogramas de cada região amplificada foram visualizados e editados com o auxílio do programa Chromas Lite v.2.0. O alinhamento das sequências *forward* e *reverse* foi realizado com auxílio do programa CLUSTAL W v.1.8 (Thompson et al., 1994). A busca por sequências semelhantes às encontradas neste trabalho foi realizada no banco de dados do *NCBI GenBank*, no programa BLAST (Altschul et al.,1990).

A caracterização de *indels* seguiu o modelo de codificação simples de *indel* (Simmons e Ochoterena, 2000), com o auxílio do programa SeqState v.1.4.1 (Müller, 2005).

O modelo de substituição nucleotídica que melhor explica a evolução das sequências foi selecionado de acordo com o critério de informação de Akaike (*Akaike Information Criterion* - AIC) (Posada and Crandall, 1998), implementado no programa jModelTest (Posada, 2008).

3.2. Índices de diversidade genética

As sequências foram analisadas por meio de diversidade nucleotídica (π), que corresponde ao número médio de diferenças nucleotídicas por sítio entre duas sequências, diversidade haplotípica (h), que corresponde à probabilidade de dois haplótipos serem diferentes, composição nucleotídica. Para o gene nuclear *PhyC* e a região cloroplastidial concatenada (*trnQ-5'rps16 e psbJ-petA*), as diversidades nucleotídica (π) e haplotípica (h) foram estimadas para cada população.

Para cada região de cpDNA, o nível de variação foi estimado por meio do índice PIC (potentially informative characters; Shaw et al., 2005), definido como a soma de substituições, *indels* e inversões (NS + ID + IV), considerando substituições e inversões dentro de *indels* como características independentes e L o comprimento da sequência. Os valores de PIC dividido pelo comprimento das sequências (PIC/L) e multiplicado por 100 permitiram o cálculo da percentagem de variabilidade para cada região.

3.3. Análises demográficas

A história demográfica foi inferida pelos seguintes testes de neutralidade baseados em polimorfismos de sequências: os testes de D de Tajima (Tajima, 1989) e F_S de Fu (Fu, 1997), que consideram o modelo de sítios infinitos sem recombinação, calculados com o programa

Arlequin v.3.5 (Excoffier e Lischer, 2010), o R_2, que considera ocorrência de recombinação intragênica (Ramos-Onsins e Rozas, 2002), sendo calculado com o auxílio do programa DnaSP v.5.1 (Librado e Rozas, 2009). Todos estes testes foram realizados por meio de simulações baseadas na coalescência, com 10000 réplicas. As sequências foram testadas quanto a eventos de recombinação com o auxílio do DnaSP v.5.1 (Librado e Rozas, 2009).

Valores de Θ foram calculados com duas diferentes medidas de polimorfismos de sequências, uma definida pelo número de sítios segregantes (S) e a outra pela média de diferenças entre pares de sequências amostradas aleatoriamente (Π) (Tajima, 1989). Ao comparar as duas estimativas de Θ é possível detectar assinaturas de seleção natural ou evento demográfico. O teste D de Tajima (Tajima, 1989) calcula o valor de D como sendo a diferença entre as estimativas de $\Theta(\Pi)$ e $\Theta(S)$ considerada que sob hipótese nula, ou seja, ausência de seleção, essa diferença seja aproximadamente zero e significativo. Por outro lado, se o desvio de D for positivo ($D > 0$) e significativo há evidência para seleção estabilizadora ou gargalo populacional, enquanto valores negativos ($D < 0$) e significativos sugerem expansão populacional ou seleção purificadora.

O teste F_S de Fu (1997) estima a probabilidade de observar uma amostra aleatória de um número de alelos igual ou menor que o valor observado considerando o nível de diversidade observada e o pressuposto de que todos os alelos são seletivamente neutros. O valor de F_S tende a ser negativo quando há um excesso de mutações recentes ou alelos raros sugerindo expansão populacional recente ou efeito carona, por outro lado, um valor positivo sugere um gargalo populacional recente ou seleção devido ao baixo número de alelos. Este teste é mais sensível para detectar expansão populacional e efeito carona, quando comparado ao D de Tajima, entretanto os valores de F_S devem ser considerados como significantes com um valor de $p<0,02$, para a rejeição da hipótese nula de neutralidade.

O teste R_2 considera a diferença entre o número de mutações únicas e o número médio de diferenças nucleotídicas, sendo o número esperado de sequências únicas em uma genealogia, após um crescimento populacional recente, k/2, sendo k o número médio de diferenças nucleotídicas. Consequentemente, baixos valores de R_2 são esperados sob este cenário demográfico (Ramos-Onsins e Rozas, 2002).

Dentre os testes de neutralidade, o F_S de Fu e o R_2 possuem maior poder estatístico para avaliar tamanho populacional constante contra expansão. Porém, o teste F_S de Fu tem poder estatístico superior para amostras populacionais maiores, enquanto o R_2 é melhor para

amostras populacionais menores e com recombinação intragênica (Ramos-Onsins e Rozas, 2002).

A análise de *Mismatch distribution*, implementada no programa Arlequin v.3.5 (Excoffier e Lischer, 2010), pode inferir a ocorrência ou não de expansão demográfica e espacial. Este modelo considera três parâmetros: $\Theta = \Theta_0$, uma população estável inicial; $\Theta = \Theta_1$, pode crescer ou contrair rapidamente a um novo tamanho; τ (Tau), o crescimento ocorre a unidades de tempo mutacional antes do presente (Rogers, 1995; Rogers e Harpending, 1992). Esse modelo é a forma mais simples de crescimento e é chamado de modelo de expansão súbita. Este teste compara a distribuição das diferenças nucleotídicas entre pares de sequências com a distribuição esperada sob uma condição de expansão populacional, o resultado dessa comparação gera uma curva unimodal para expansão ou bimodal para equilíbrio populacional (Rogers e Harpending, 1992). A expansão populacional pode ser validada por valores de *Raggedness* e SSD (*Sum of Square Deviations - SSD*) obtidos da *Mismatch distribution*. As populações em expansão mostram baixos valores de *Raggedness* e valores de p de *SSD* maiores que 0,05, enquanto, as populações em equilíbrio apresentam um índice *Raggedness* elevado e *SSD* com valores de p menores que 0,05 (Schneider e Excoffier, 1999). Estes valores foram gerados por 10000 simulações de coalescência utilizando a abordagem de *bootstrap* paramétrico, o qual resultou em um intervalo de confiança estatisticamente significativo.

Para corroborar as análises demográficas citadas acima, regiões cloroplastidiais concatenadas para representar uma maior variabilidade, foram utilizadas em para uma análise *Bayesian Skyline Plot* (BSP), que estima a variação do tamanho efetivo populacional em relação ao tempo (Drummond et al., 2005). A análise foi realizada com o auxílio do programa BEAST v.1.7.2 (Drummond e Rambaut, 2007), que utiliza uma abordagem coalescente. Os parâmetros dessa análise foram selecionados em BEAUTI v.1.7.2 (Drummond e Rambaut, 2007), com dois arquivos de entrada devido à codificação de *indels* realizados no SeqState v 1.4.1 (Müller, 2005). Tais parâmetros foram: HKY como modelo de substituição nucleotídica, *tree prior* coalescente *Bayesian Skyline* e a taxa de $0{,}0011e^{-6}$ substituições sinônimas por milhão de ano sugerida como a menor taxa para sequências de DNA cloroplastidial (Wolfe et al., 1987). Uma análise com 100 milhões de gerações utilizando cadeias de Markov Monte de Carlo (MCMC), realizada em BEAST v.1.7.1, foi utilizada no programa TRACER v.1.5 (Drummond e Rambaut, 2007), com um *burn in* de 10 milhões para plotar o gráfico da BSP.

3.4. Filogeografia e estrutura populacional

A Análise Filogeográfica de Clado Aninhado (NCPA - *Nested Clade Phylogeographic Analysis*; Templeton, 1998; 2003) testa a hipótese nula de não associação entre haplótipos e geografia. Inicialmente, uma rede de haplótipos não enraizada é gerada por parcimônia estatística, com 95% de significância, no programa TCS v.1.21 (Clement et al., 2000). A rede gerada representa os haplótipos conectados por passos mutacionais, permitindo determinar a relação genealógica entre eles e a inferência do haplótipo ancestral nas populações analisadas, usando os critérios descritos em Castelloe e Templeton (1994).

A construção de clados aninhados, com base na rede de haplótipos, seguiu os critérios estabelecidos em Templeton e Sing (1993) e Templeton et al., (1987). Para criar o primeiro nível de aninhamento, foi considerado um passo mutacional a partir de haplótipos presentes na ponta da rede, considerados os mais recentes, em direção aos haplótipos de interior, os mais ancestrais, formando clados. O segundo e terceiro níveis de aninhamentos foram definidos considerando um passo mutacional a partir dos clados aninhados resultantes do primeiro e do segundo aninhamentos, respectivamente. As ambiguidades no aninhamento dos clados foram solucionadas considerando os critérios descritos em Crandall e Templeton (1993), em que: os haplótipos devem ser conectados a haplótipos mais freqüentes; os haplótipos devem ser conectados a haplótipos do interior ao invés de haplótipos de ponta; os haplótipos devem ser conectados a outros haplótipos da mesma população ou região geográfica. A hierarquia dos clados contém informações temporais, mesmo para as árvores não enraizadas, pois a teoria da coalescência prediz que os haplótipos de ponta são mais recentes do que os haplótipos de interior (Castelloe e Templeton, 1994).

Para as inferências demográficas usando a NCPA são necessárias as informações geográficas, ou seja, a distribuição espacial dos haplótipos, a qual é quantificada por dois valores: distância de clado (Dc), que mede a amplitude geográfica e estima o centro geográfico da distribuição de um determinado clado e distância de clado aninhado (Dn), o qual quantifica a distância entre haplótipos ou clados evolutivamente próximos, ou seja, aninhados dentro de um clado de alto nível (Templeton et al., 1995). A hipótese nula de que haplótipos ou clados não apresentam associação geográfica foi testada por meio de 10000 permutações randômicas, recalculando os valores de Dc e Dn (Templeton et al., 1995), com o programa GeoDis v2.6 (Posada et al., 2000). Desta forma, a distribuição dessas distâncias sob a hipótese de não associação geográfica foi simulada e contrastada com as distâncias observadas, para inferir as distâncias estatísticas significativas (Templeton, 2004a). As associações encontradas entre a localidade geográfica e o haplótipo foram interpretadas com

o auxílio da chave de inferência de Templeton de 2011, disponível em http://darwin.uvigo.es/software/geodis.html, permitindo identificar eventos históricos e padrões de fluxo gênico recorrente (Templeton, 1998; Templeton et al., 1995). Embora a chave de inferência seja usada para discriminar entre fluxo gênico, expansão da população, eventos de colonização a longa distância e fragmentação, a análise de clados aninhados não trata esses padrões como mutuamente exclusivos, mas procura por múltiplos padrões de sobreposição dentro do mesmo conjunto de dados (Templeton, 2001).

A análise de variância molecular (AMOVA) foi realizada para estimar índices de estrutura genética, utilizando a informação sobre o conteúdo de haplótipos, assim como as suas frequências, sendo as diferenças haplotípicas representadas por uma matriz de distâncias Euclidianas ao quadrado (Excoffier et al., 1992). A significância dos componentes de covariâncias, associados a níveis hierárquicos de diferenciação, foram testados com o programa Arlequin v.3.5 (Excoffier e Lischer, 2010) com 16000 permutações utilizando o modelo de substituição de Tamura e Nei. As percentagens de componentes de covariância geraram os índices de Φ_{ST} para os haplótipos dentro das populações em relação aos haplótipos totais da espécie, Φ_{SC} da diversidade molecular de haplótipos dentro das populações em relação aos pares de haplótipos da região e Φ_{CT} correlação de haplótipos dentro de um grupo de populações em relação aos pares de haplótipos do total da espécie (Excoffier et al., 1992). Primeiramente, uma análise de variância molecular foi realizada sem subdivisão populacional para avaliar o índice de estruturação global, e então, testes com diferentes agrupamentos de populações foram realizadas de acordo com a distância geográfica.

O teste de Mantel foi realizado para testar as associações entre distâncias geográfica e genética (Smouse et al., 1986; Mantel, 1967). Para isso, uma matriz de distância geográfica entre pares de populações foi calculada a partir das coordenadas geográficas destas localidades com o auxílio de um conversor disponível em http://jan.ucc.nau.edu/~cvm/latlongdist.html. A matriz de distância genética foi gerada com os valores de F_{ST} calculados pelo programa Arlequin v.3.5 (Excoffier e Lischer, 2010), no qual o teste de Mantel foi gerado utilizando o modelo de substituição nucleotídica de Tamura e Nei, com 16000 permutações.

Estimativas de fluxo gênico entre grupos de populações foram calculadas com o programa MIGRATE v.3.2.1 (Beerli e Felsenstein, 1999). Este método utiliza a abordagem da Máxima Verossimilhança e a teoria da coalescência para estimar alguns parâmetros populacionais como taxa de imigração efetiva escalada por mutação (M) e tamanho

populacional (ϴ) (Beerli e Felsenstein, 1999). O parâmetro de fluxo gênico pode ser representado por M ou pelo número efetivo de imigrantes por geração xNm (x equivale a 1 para DNA cloroplastidial e 4 para DNA nuclear, N é o tamanho efetivo populacional e m a taxa de migração) estimado a partir de valores de ϴ multiplicado por M. Estimativas de fluxo gênico foram feitas considerando as populações agrupadas em quatro grupos a partir da estrutura populacional inferida pela AMOVA, considerando a distância geográfica, sendo esses: (A) São Paulo e Paraná, (B) região oeste do Rio Grande do Sul, abrangendo as cidades de Jacutinga, Santiago, Jaguari e Santa Maria, (C) região leste do Rio Grande do Sul, com Caçapava do Sul, Barra do Ribeiro, Pouso Novo e Viamão, (D) litoral do Rio Grande do Sul com Mostardas, Osório, Arroio do Sal,e Santa Catarina, com Laguna e Florianópolis. Foram realizadas cinco análises no MIGRATE, as quais utilizaram 10 cadeias curtas gerando 10000 árvores, três cadeias longas gerando 1000000 de árvores com 10000 de *burn-in*. Os valores de ϴ para a primeira análise foram obtidos dos valores de F_{ST}, e nas demais análises, os valores de ϴ foram substituídos pelos valores gerados nas análises anteriores.

3.5. Relógio molecular e tempo de divergência

As análises de tempo de divergência seguiram uma abordagem bayesiana, optando-se por um relógio molecular relaxado, com o auxílio do programa BEAST v.1.7.2 (Drummond e Rambaut, 2007), e consistiram de uma primeira calibração interespecífica dentro de eudicotiledôneas, com enfoque no clado BCT das Cactaceae, e uma segunda calibração intraespecífica em *C. hildmannianus*.

Primeiramente, sequências de *C. hildmannianus* de quatro regiões cloroplastidiais (*trnS-trnG, 3'trnK-matK, trnQ-5'rps16 e psbJ-petA*) foram concatenadas e incluídas em uma supermatriz para datação das Cactaceae (Lendel et al., dados a publicar). A escolha destas regiões intergênicas foi realizada considerando os dados para os outros táxons presentes nesta supermatriz. A primeira calibração seguiu todos os parâmetros utilizados em Lendel et al. (dados a publicar). A partir desta primeira calibração, o tempo do ancestral comum mais recente (TMRCA) de *C. hildmannianus* foi utilizada nas inferências dos tempos de divergência entre as populações desta espécie.

A segunda calibração utilizou as duas regiões cloroplastidiais com maior variabilidade intrapopulacional (*trnQ-5'rps16 e psbJ-petA*), em um total de 85 indivíduos representando toda a distribuição da espécie. Os parâmetros para a datação foram selecionados no programa BEAUTI v.1.7.2 (Drummond e Rambaut, 2007), e devido à presença de *indels*, dois arquivos

de entrada foram montados: um com tipo de dados de DNA e o outro com dados binários (*Standard*) obtidos com o auxílio do programa SeqState v 1.4.1 (Müller, 2005). O modelo de substituição nucleotídica HKY foi utilizado por ser mais semelhante ao modelo de substituição nucleotídica avaliado pelo jModelTest (Posada, 2008). Tendo em vista que estas análises referem-se à história populacional, o *prior* escolhido para a árvore foi o coalescente com crescimento em expansão. A data obtida para *C. hildmannianus* foi usada como um *root-height prior* com distribuição normal e o desvio padrão foi de 0,05, com um desvio padrão baixo, o erro da datação fica associado à matriz de dados e não à primeira calibração.

Os dados foram testados quanto ao modelo de relógio molecular por meio de uma inferência Bayesiana, no BEAST v.1.7.2 (Drummond e Rambaut, 2007). As análises foram realizadas sob um modelo de relógio estrito e de um relógio relaxado exponencial e lognormal, uma vez que cada modelo assume uma diferente taxa de evolução nos ramos da árvore, sendo essa taxa constante no relógio estrito, enquanto para o relógio relaxado as taxas são consideradas independentes nos diferentes ramos, sob uma distribuição exponencial ou lognormal (Drummond et al., 2006). A taxa para o relógio molecular utilizou um *prior* uniforme com valores para regiões cloroplastidiais entre 1.e-9 e 1.e-11.

A escolha do relógio molecular mais adequado para os dados foi realizada por meio do fator Bayes. Este fator representa a razão da probabilidade marginal entre os modelos de relógio molecular testados, considerando que quando o fator Bayes for maior que 20 ou o lnBayes for maior que 2,99, então o modelo com maior probabilidade marginal é favorecido (Hon et al., 2008; Drummond et al., 2006). Desta forma, para cada relógio molecular uma análise de cadeia de Markov Monte Carlo (MCMC) de 1 milhão de gerações foi realizada no BEAST v.1.7.2 (Drummond e Rambaut, 2007), e ao final, estas foram comparadas no programa TRACER v.1.5 (Drummond e Rambaut, 2007) para o cálculo do fator Bayes. O modelo favorecido foi utilizado nas análises finais, sendo estas cinco análises independentes com 20 milhões de gerações cada, com uma árvore amostrada a cada 1000 gerações. Os valores efetivos das amostras e as densidades posteriores foram observados para todos os parâmetros no TRACER. As corridas independentes foram combinadas no LogCombiner v.1.7.2 (Drummond e Rambaut, 2007) para cada conjunto de dados, com um *burn in* de 2 a 5 milhões, e uma árvore consenso foi gerada no TreeAnnotator v.1.7.2 (Drummond e Rambaut, 2007). Todas as análises utilizando o pacote computacional Beast v.1.7.2 foram realizadas com o auxílio computacional do Albiorix *cluster* (*Department of Plant and Environmental Sciences – Gothenburg University, Sweden*).

3.6. Análise biogeográfica

Uma distribuição ancestral geográfica hipotética das populações de *C. hildmannianus* foi elaborada por metodologia Bayesiana, implementada em S-DIVA, no programa RASP v.2.0 (Yu et al., 2010). A probabilidade de cada região geográfica ancestral é determinada para cada nó medido sobre todas as árvores amostradas derivados das cadeias de Markov (Chaves, et al., 2011). As árvores obtidas da segunda calibração foram testadas quanto à distribuição ancestral referente a quatro grandes áreas: (A) São Paulo e Paraná, (B) região oeste do Rio Grande do Sul, abrangendo as cidades de Jacutinga, Santiago, Jaguari e Santa Maria, (C) região leste do Rio Grande do Sul, com Caçapava do Sul, Barra do Ribeiro, Pouso Novo e Viamão, (D) litoral do Rio Grande do Sul e Santa Catarina, com Mostardas, Osório, Arroio do Sal, Laguna e Florianópolis. Estas áreas foram divididas considerando a proximidade geográfica, mesmo princípio aplicado para as análises filogeográficas. As inferências resultaram da combinação de duas corridas independentes e foram verificadas quanto à distância entre elas, validadas para valores menores que 0,01. As definições foram de ampla distribuição para os virtuais *outgroups*, com o número máximo de quatro áreas e o modelo de substituição nucleotídica utilizado foi o F81, obtido em jModelTest (Posada, 2008).

Uma segunda análise foi realizada no RASP para testar a probabilidade da distribuição das populações em relação ao tipo de solo. Para isso, um mapa de solos da EMBRAPA disponível em http://sosgisbr.com/2012/06/04/mapa-de-solos-do-brasil-2011-embrapafoi utilizado para classificar o solo das regiões coletadas neste trabalho: (A) latossolo vermelho-amarelo - Serrana-SP e Itatiba-SP; (B) latossolo vermelho - Serrana-SP; (C) argilossolo vemelho-amarelo - Piratininga-SP, Itu-SP, Penha-SC, Laguna-SC e Viamão-RS; (D) argilossolo vermelho - Jacarezinho-PR e Jaguari-RS; (E) cambissolo húmico - Sengés-PR; (F) nitossolo vermelho - Cianorte-PR; (G) latossolo bruno - Guarapuava-PR, Cantagalo-PR e Mangueirinha-PR; (H) neossolo quartzarênico - Florianópolis-SC e Arroio do Sal-RS; (I) cambissolo háplico - Jacutinga-RS; (J) neossolo litólico - Pouso Novo-RS, Santiago-RS e Caçapava do Sul-RS; (K) planossolo háplico - Santa Maria-RS e Barra do Ribeiro-RS; (L) planossolo hidromórfico - Osório-RS e Mostardas-RS.

3.7. Modelagem de nicho ecológico e paleomodelagem

A modelagem de nicho ecológico e paleomodelagem foram realizadas a partir de dados de coordenadas geográficas obtidas de herbários: ICN-UFRS e PACA (Herbarium Anchieta-RS), do site http://splink.cria.org.br, MBM, UPCB, ESA, HSJRP, SinBiota, SPSF e

dos locais coletados no presente trabalho (Apêndice 1 – Tabela 1). Esta análise foi realizada com o auxílio do algoritmo MaxEnt (*Maximum Entropy Species Distribution Modeling* v.3.3.3e) (Phillips et al., 2006). Foi definido um valor de 20% de aleatoriedade para o teste com 61 réplicas, o qual foi escolhido considerando o número de pontos amostrados. Com o intuito de investigar as alterações climáticas em dois cenários, foram realizadas projeções da área de ocorrência para o presente, aproximadamente do ano de 1950 a 2000, com mapas de resolução espacial de 30 arco-segundos e as projeções de ocorrência do passado, aproximadamente há 21000 anos no último máximo glacial (Last Glacial Maximum)pelos modelos gerados pela *Commmunity Climate System Model* (CCSM) e *Model for Interdisciplinary Research On Climate* – (MIROC), ambos com resolução de 2,5 arco-minutos. As camadas de mapas utilizadas têm as seguintes variáveis bioclimáticas: altitude, temperatura média anual, temperatura de sazonalidade, variação da temperatura anual, média do trimestre mais úmido, média do trimestre mais seco, média do trimestre mais quente, média do trimestre mais frio, precipitação anual, precipitação de sazonalidade, precipitação do trimestre mais chuvoso, precipitação do trimestre mais seco, precipitação do trimestre mais quente, precipitação do trimestre mais frio. A altitude foi retirada das análises para o passado, devido à diminuição do nível do mar (Clapperton et al., 1993). As camadas estão disponíveis em http://www.worlclim.org. Para a avaliação do modelo, foi utilizado um parâmetro independente do limiar de corte que mede o desempenho da predição dos modelos definida a área sob a curva (*area under curve* – AUC) (Manel et al., 2001). Os valores de AUC têm como significados: 0,5 e 0,7 - baixa precisão; 0,7 e 0,9 - aplicável; e > 0,9 - excelente precisão (Swets, 1988). Os mapas de distribuição foram editados com o auxílio do programa DIVA-GIS (Hijmans et al., 2005) e o limiar de corte, que indica a probabilidade mínima de ocorrência da espécie, foi obtido da média de todas as réplicas geradas no MaxEnt.

IV. RESULTADOS

1. MATERIAL BIOLÓGICO

Para a realização deste trabalho foram analisados 157 indivíduos de 24 populações de *C. hildmannianus*. Um espécime de cada localidade foi utilizado como material testemunho e armazenado como exsicata no herbário SPFR (Tabela 3).

2. ANÁLISE DE DADOS

As reações de isolamento de DNA geraram amostras com 30 - 50 ng/µL. A partir destas amostras, foram sequenciadas regiões de um gene nuclear e 13 sequências não codificantes de cpDNA. O número de indivíduos a partir das quais foram obtidas as sequências, suas respectivas populações, o número de haplótipos e os seus comprimentos em pares de base estão na Tabela 4. As sequências com variação haplotípica entre as 14 avaliadas foram: o gene nuclear *PhyC* com cinco haplótipos; a região intrônica cloroplastidial *trnL* com dois haplótipos; os espaçadores cloroplastidiais *petL-psbE* com 2 haplótipos, *atpI-atpH* com 3 haplótipos, *trnQ-5'rps16* com seis haplótipos e *psbJ-petA* com 12 haplótipos. Embora a amostra de indivíduos e populações analisada para cada sequência seja diferente, a escolha desta considerou a distância geográfica e representatividade da distribuição de *C. hildmannianus*.

As sequências das regiões cloroplastidiais *trnQ-5'rps16* e *psbJ-petA* apresentaram *indels*, os quais foram codificados manualmente ou com o auxílio do programa SeqState v.1.4.1 (Müller, 2005) para a realização das análises populacionais.

O resultado do programa jModelTest (Posada, 2008) para definição do modelo que melhor explica a substituição nucleotídicadas sequências de DNA cloroplastidial foi o F81, enquanto para o gene nuclear *PhyC* foi o TPM1uf+I+G. De acordo com a disponibilidade do programa utilizado para as inferências populacionais, foram selecionados os modelos de substituições nucleotídicas com valores mais próximos aos modelos citados acima.

Tabela 3. Pontos de coletas das diferentes localidades e número de indivíduos amostrados de populações naturais de *C. hildmannianus* com a identificação do material testemunho.

Localidade	Coordenadas geográficas		Número de espécimes coletados	Identificação*	Data
	Latitude	Longitude			
Itatiba-SP	23º0'20,88"	46º50'20,04"	15	SPFR 11888	11/4/2009
Caçapava do Sul-RS	30º49'26,1"	53º30'52,3"	14	SPFR 12343	24/05/2009
Jaguari-RS	29º26'55,4"	54º43'50,3"	9	SPFR 12345	25/05/2009
Santiago-RS	29º02'13,2"	55º03'8,8"	7	SPFR 12353	25/05/2009
Santa Maria-RS	29º46'46,4"	53º46'39,3"	4	-	26/05/2009
Viamão-RS	30º14'45,7"	50º58'20,8"	9	SPFR 12349	27/05/2009
Mostardas-RS	31º06'36,9"	50º55'16,1"	8	SPFR 12333	28/05/2009
Penha-SC	26º46'31,2"	48º35'53,2"	2	SPFR 12351	29/09/2009
Laguna-SC	28º29'15,1"	48º46'43,8"	3	SPFR 12346	29/09/2009
Florianópolis-SC	27º34'49,9"	48º27'25,8"	11	SPFR 12344	30/09/2009
Cianorte - PR	23º39'21,8"	52º30'26,2"	8	SPFR 12342	31/10/2009
Guarapuava-PR	25º17'42"	51º53'7,1"	6	SPFR 12989	28/04/2011
Cantagalo-PR	25º25'0"	52º04'14,9"	7	SPFR 12987	28/04/2011
Mangueirinha-PR	25º46'27,2"	52º06'55,6"	8	SPFR 12988	28/04/2011
Arroio do Sal-RS	29º32'27"	49º55'26"	4	SPFR 13083	17/07/2011
Osório-RS	29º57'31,1"	50º13'30,9"	10	SPFR 13089	18/07/2011
Barra do Ribeiro-RS	30º16'19,4"	51º24'55,1"	5	SPFR 13085	19/07/2011
Pouso Novo-RS	29º13'03"	52º10'10,5"	7	SPFR 13086	20/07/2011
Jacutinga-RS	27º43'52,5"	52º31'11,9"	1	-	20/07/2011
Serrana-SP	21º15'33"	47º34'28"	10	SPFR 13140	19/09/2011
Sengés-PR	24º07'	49º23'	3	SPFR 13388	1/11/2011
Jacarezinho-PR	23º14'14,1"	50º02'19,4"	2	SPFR 13389	2/11/2011
Piratininga-SP	22º26'9,2"	49º07'57,8"	2	SPFR 13390	2/11/2011
Itu-SP	23º15'5,4"	47º13'22,6"	2	SPFR 13391	11/11/2011

* SPFR: Herbário São Paulo Faculdade Ribeirão - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto - USP.

Tabela 4. Descrição do número de indivíduos analisados por localidade com o número e comprimento em pares de base dos haplótipos obtidos para as regiões do gene nuclear e cloroplastidiais.

Segmento	Nº indivíduos/ Nº população	Localidades	Nº haplótipos	bp*
Nuclear				
PhyC	56/24	Todas as localidades amostradas.	5	~ 998
cpDNA				
trnH-psbA	8/8	Itatiba-SP, Cianorte-PR, Santiago-RS, Caçapava do Sul-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	1	~ 360
íntron *trnG*	5/5	Cianorte-PR, Caçapava do Sul-RS, Penha-SC, Laguna-SC, Mostardas-RS	1	~ 476
trnS-trnG	8/8	Itatiba-SP, Cianorte-PR, Santiago-RS, Caçapava do Sul-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	1	~ 970
trnT-trnL	8/8	Itatiba-SP, Cianorte-PR, Santiago-RS, Caçapava do Sul-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	1	~ 348
trnL-trnF	7/7	Itatiba-SP, Cianorte-PR, Caçapava do Sul-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	1	~ 413
íntron *trnL*	18/14	Itatiba-SP, Cianorte-PR, Guarapuava-PR, Cantagalo-PR, Mangueirinha-PR, Jaguari-RS, Santa Maria-RS, Viamão-RS, Santiago-RS, Caçapava do Sul-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	2	~ 645
3'trnK-matK	29/24	Todas as localidades amostradas.	1	~ 1224
psbD-trnT	12/9	Itatiba-SP, Cianorte-PR, Caçapava do Sul-RS, Jaguari-RS, Santiago-RS, Santa Maria-RS, Mostardas-RS, Penha-SC, Laguna-SC	1	~ 668
psbJ-petA	149/24	Todas as localidades amostradas.	12	~ 512

atpI-atpH	29/12	Itatiba-SP, Cianorte-PR, Santa Maria-RS, Caçapava do Sul-RS, Jaguari-RS, Viamão-RS, Pouso Novo-RS, Jacutinga-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	3	~ 591
3'rps16-5'trnK	8/8	Itatiba-SP, Cianorte-PR, Santiago-RS, Caçapava do Sul-RS, Penha-SC, Florianópolis-SC, Laguna-SC, Mostardas-RS	1	~ 221
petL-psbE	16/13	Itatiba-SP, Cianorte-PR, Guarapuava-PR, Cantagalo-PR, Mangueirinha-PR, Jaguari-RS, Santa Maria-RS, Viamão-RS, Caçapava do Sul-RS, Santiago-RS, Florianópolis-SC, Laguna-SC, Mostardas-RS	2	~ 531
trnQ-5'rps16	93/23	Todas as localidades amostradas, com exceção de Penha-SC.	6	~ 541

* comprimento das sequências em número de pares de bases (pb) obtidas após sequenciamento.

2.1. Índices de diversidade genética

Dentre as 14 regiões analisadas, somente seis apresentaram variação populacional em *C. hildmannianus*: o gene nuclear *PhyC* e cinco sequências de DNA cloroplastidiais (íntron *trnL, petL-psbE,atpI-atpH, trnQ-5'rps16* e *psbJ-petA*), a relação dos haplótipos (com *indel* codificado) e suas localidades estão no Apêndice 2 – Tabela 1. Para essas seis regiões foi observada variação interpopulacional, enquanto variação intrapopulacional foi encontrada nas sequências do gene *PhyC* e cloroplastidiais intergênicas *atpI-atpH,trnQ-5'rps16* e *psbJ-petA*. Os dados obtidos para composição nucleotídica, índice de diversidade nucleotídica (π) e diversidade haplotípica (h) estão na Tabela 5. A região *psbJ-petA* apresentou o maior índice de diversidade nucleotídica e diversidade haplotípica. A composição nucleotídica das sequências mostrou uma maior porcentagem de bases AT, tanto para o DNA nuclear como para o DNA cloroplastidial.

A percentagem de variabilidade das sequências foi calculada somente para o DNA cloroplastidial. Os valores de PIC/L foram (1+0+0)/645 para o íntron *trnL*, (1+0+0)/531 para o *petL-psbE*, (2+0+0)/591 para o *atpI-atpH*, (6+2+0)/541 para o *trnQ-5'rps16* e (5+82+0)/512 para o *psbJ-petA* (Tabela 5), com a percentagem de variabilidade de 0,15%, 0,18%, 0,33%, 1,47% e 16%, respectivamente. Desta forma, a maior percentagem de variabilidade foi encontrada para as duas regiões intergênicas *psbJ-petA* e *trnQ-5'rps16.*

Os índices de diversidade nucleotídica (π) e diversidade haplotípica (h) para cada população podem ser visualizados na Tabela 6. Para o gene nuclear *PhyC*, as populações com maiores índices de diversidade nucleotídica e haplotípica estão nas localidades de Cianorte e Mangueirinha no Paraná, Laguna em Santa Catarina e Jaguari, Barra do Ribeiro, Mostardas e Osório no Rio Grande do Sul. Essas localidades representam tanto o grupo de SP/PR + Oeste-PR como o grupo Leste-RS + Litoral. Para as sequências cloroplastidiais concatenadas *trnQ-rps16/psbJ-peta*, as populações de Caçapava do Sul e Barra do Ribeiro, do grupo Leste-RS, apresentam os maiores índices de diversidade nucleotídica.

2.2. Análises demográficas

Os valores calculados com os testes de neutralidade D de Tajima e F_S de Fu não foram significativos, não sugerindo eventos populacionais tais como expansão. Embora para o concatenado *trnQ-5'rps16/psbJ-petA* há valores significantes quando analisados somente os clados gerados na NCPA: clado 1.2 foi negativo e significante com a análise de F_S de Fu; clado 2.3 foi negativo e significante com as análises de F_S de Fu e D de Tajima (Tabela 7).

No presente trabalho, valores de R_2 foram significativos, portanto, condizentes com um cenário de crescimento populacional. As sequências não apresentaram evento de recombinação quando analisadas no programa DnaSP.

A soma dos desvios quadrados e o índice de *Raggedness* gerados pelas diferenças nucleotídicas entre pares de sequências na análise de *Mismatch distribution* permitiram testar a hipótese de expansão populacional. Os valores desta análise foram baixos e não significativos para as regiões não codificantes cloroplastidiais e o gene nuclear (Tabela 7), sugerindo a hipótese de expansão populacional (Schneider e Excoffier, 1999; Harpending, 1994). Assinaturas de expansão espacial e demográfica foram sugeridas pelas análises da maioria das regiões testadas, para as quais os valores de p de (*SSD*) foram maiores que 0,05.

Tabela 5. Índices de diversidade nucleotídica (π), diversidade haplotípica (h), composição nucleotídica e número de caracteres potencialmente informativos (PIC/L) para o seis segmentos do gene nuclear e das regiões intergênicas de DNA cloroplastidial *C. hildmannianus*.

Sequências	Diversidade nucleotídica (π)	Diversidade haplotípica (h)	Composição nucleotídica (%)	PIC/L
Nuclear				
PhyC	0,000563 +/- 0,000519	0,5071 +/- 0,0619	C : 17.94; T : 28.78; A : 28.08; G : 25.20	-
cpDNA				
íntron *trnL*	0,001044 +/- 0,001019	0,5294 +/- 0,0404	C: 11,80; T: 29,01; A: 43,04; G: 16,15	1/645
petL-psbE	0,001035 +/- 0,001021	0,5250 +/- 0,0546	C : 16,42; T : 34,54; A : 34,37; G : 14,66	1/531
atpI-atpH	0,001447 +/- 0,001400	0,4631 +/- 0,0798	C: 20,07; T: 32,87; A: 36,77; G: 10,29	2/591
trnQ-5'rps16	0,001506 +/- 0,001251	0,5482 +/- 0,0348	C : 12,24; T : 38,52; A : 36,99; G : 12,25	8/541
psbJ-petA	0,003216 +/- 0,002177	0,7105 +/- 0,0288	C: 15,46; T: 38,02; A: 34,50; G: 12,02	87/512

Entretanto, para as regiões cloroplastidiais *petL-psbE* e *íntron trnL* e a nuclear *PhyC*, foi detectado somente expansão demográfica, enquanto para *psbJ-petA* somente expansão espacial. Os baixos índices de *Raggedness* e os gráficos unimodais sugeriram expansão populacional para todas as regiões, enquanto para *trnQ-5'rps16, psbJ-petA* e o concatenado *trnQ-5'rps16/psbJ-petA* apresentaram uma curva com tendência bimodal (Figura 5. E, F e G).

Os clados 1.2 e 2.3 da NCPA para o concatenado *trnQ-5'rps16/psbJ-petA* indicaram expansão populacional e demográfica.

Osdados obtidos das sequências concatenadas *trnQ-5'rps16/psbJ-petA* por meio da *Bayesian Skyline Plot* (BSP) geraram uma distribuição posterior com valores do tamanho efetivo amostral superiores a 200. O gráfico gerado por esta análise demonstrou que ao longo da história de *C. hildmannianus*, as populações estavam em equilíbrio demográfico, e mais recentemente, passaram por uma expansão (Figura 6).

Tabela 6. Resultados dos índices de diversidade haplotípica (h) e diversidade nucleotídica (π) estimados para a região concatenada cloroplastidial *trnQ-5'rps16/psbJ-petA* e o gene nuclear *PhyC* para cada população de *C. hildmannianus*.

Localidade	***PhyC***		**Concatenado (*trnQ-5'rps16/psbJ-petA*)**	
	π	**h**	**π**	**h**
Serrana-SP	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	1.0000 +/- 0.0000
Piratininga-SP	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000
Itatiba-SP	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.000329 +/- 0.000449	0.6667 +/- 0.1598
Itu-SP	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	1.0000 +/- 0.0000
Jacarezinho-PR	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	1.0000 +/- 0.0000
Sengés-PR	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	1.0000 +/- 0.0000
Cianorte - PR	0.001010 +/- 0.001428	1.0000 +/- 0.5000	0.000577 +/- 0.000620	0.4643 +/- 0.2000
Guarapuava-PR	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.000590 +/- 0.000649	0.5333 +/- 0.1721
Cantagalo-PR	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.000326 +/- 0.000446	0.5238 +/- 0.2086
Mangueirinha-PR	0.001009 +/- 0.001426	1.0000 +/- 0.5000	0.000278 +/- 0.000394	0.2500 +/- 0.1802
Penha-SC	0.0000 +/- 0.0000	0.0000 +/- 0.0000	-	-
Laguna-SC	0.001009 +/- 0.001426	1.0000 +/- 0.5000	0.0000 +/- 0.0000	0.0000 +/- 0.0000
Florianópolis-SC	0.000432 +/- 0.000492	0.4286 +/- 0.1687	0.000786 +/- 0.000787	0.6071 +/- 0.1640
Jacutinga-RS	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	1.0000 +/- 0.0000
Jaguari-RS	0.001009 +/- 0.00142	1.0000 +/- 0.5000	0.000727 +/- 0.000906	0.6667 +/- 0.3143
Santa Maria-RS	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000
Santiago-RS	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000
Pouso Novo-RS	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000
Caçapava do Sul-RS	0.000673 +/- 0.000755	0.6667 +/- 0.2041	0.004908 +/- 0.004168	1.0000 +/- 0.2722
Barra do Ribeiro-RS	0.001009 +/- 0.001426	1.0000 +/- 0.5000	0.001855 +/- 0.001528	0.7000 +/- 0.2184
Viamão-RS	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.000579 +/- 0.000717	0.5000 +/- 0.2652
Mostardas-RS	0.001009 +/- 0.001426	1.0000 +/- 0.5000	0.0000 +/- 0.0000	0.0000 +/- 0.0000
Arroio do Sal-RS	0.0000 +/- 0.0000	0.0000 +/- 0.0000	0.0000 +/- 0.0000	1.0000 +/- 0.0000
Osório-RS	0.001009 +/- 0.001426	1.0000 +/- 0.5000	0.0000 +/- 0.0000	0.0000 +/- 0.0000

Tabela 7. Valores calculados pelos testes de neutralidade D de Tajima, Fs de Fu, R2 e *Mismatch distribution* em amostras de *C. hildmannianus* para o gene nuclear *PhyC* e as regiões cloroplastidiais não codificantes *atpI-atpH, petL-psbE,*íntron *trnL, trnQ-5'rps16, psbJ-petA*. Para a análise de *Mismatch distribution* gerado os índices de *Raggedness* e a soma dos desvios quadrados (*SSD: Sum of Square Deviations*).

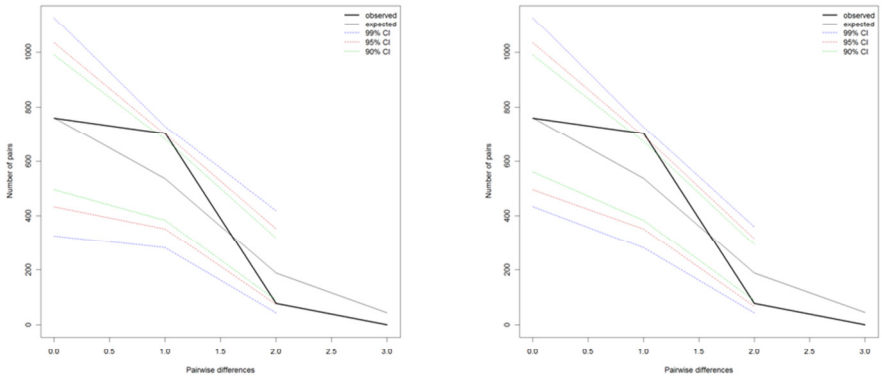
Marcadores	**D de Tajima**	**Fs de Fu**	**R2**	***Mismatch distribution***	
				Demográfica	**Espacial**
Nuclear					
*PhyC***	-0,79594	-1,60169	0,16120*	*SSD*: 0,01690 (*SSD*) p-valor: 0,06380 *Raggedness*: 0,16860 (*Raggedness*) p-valor: 0,06820	*SSD*: 0,01690 (*SSD*) p-valor: 0,00020 *Raggedness*: 0,16860 (*Raggedness*) p-valor: 0,07640
cpDNA					
atpI-atpH	0,01237	0,08871	0,16157*	*SSD*: 0,00934 (*SSD*) p-valor: 0,23610 *Raggedness*: 0,15048 (*Raggedness*) p-valor: 0,31940	*SSD*: 0,00933 (*SSD*) p-valor: 0,09250 *Raggedness*: 0,15048 (*Raggedness*) p-valor: 0,33350
petL-psbE	1,4737	1,33347	0,16173*	*SSD*: 0,02802 (SSD) p-valor: 0,12470 *Raggedness*: 0,27813 (*Raggedness*) p-valor: 0,08930	*SSD*: 0,02802 (*SSD*) p-valor: 0,02540 *Raggedness*: 0,27813 (*Raggedness*) p-valor: 0,08650
íntron trnL	1,54756	1,42756	0,16082*	*SSD*: 0,02923 (*SSD*) p-valor: 0,06620 *Raggedness*: 0,28374 (*Raggedness*) p-valor: 0,04970	*SSD*: 0,02922 (*SSD*) p-valor: 0,01370 *Raggedness*: 0,28374 (*Raggedness*) p-valor: 0,05530
trnQ-5'rps16	-0,86116	0,15863	0,16117*	*SSD*: 0,09681	*SSD*: 0,06671

				(*SSD*) p-valor: 0,12480 *Raggedness*: 0,36608 (*Raggedness*) p-valor: 0,09020	(*SSD*) p-valor: 0,16420 *Raggedness*: 0,36608 (*Raggedness*) p-valor: 0,25440
psbJ-petA	0,74549	12,84485	0,16159*	*SSD*: 0,60683 (*SSD*) p-valor: 0,00000 *Raggedness*: 0,14570 (*Raggedness*) p-valor: 0,99940	*SSD*: 0,06766 (*SSD*) p-valor: 0,21360 *Raggedness*: 0,14570 (*Raggedness*) p-valor: 0,43320
Concatenado (*trnQ-5'rps16/psbJ-petA*)	-0,14795	-4,20408	0,16159*	*SSD*: 0,05654 (*SSD*) p-valor: 0,17010 *Raggedness*: 0,10628 (*Raggedness*) p-valor: 0,20210	*SSD*: 0,04397 (*SSD*) p-valor: 0,25560 *Raggedness*: 0,10628 (*Raggedness*) p-valor: 0,53880
Concatenado *(trnQ-rpS16/psbJ-petA)* clado-1.2	-0,83549	-2,31961*	0,16174*	*SSD*: 0,23052 (*SSD*) p-valor: 0,14580 *Raggedness*: 0,21458 (*Raggedness*) p-valor: 0,21280	*SSD*: 0,00412 (*SSD*) p-valor: 0,25060 *Raggedness*: 0,21458 (*Raggedness*) p-valor: 0,41680
Concatenado *(trnQ-rpS16/psbJ-petA)* clado-2.3	-2,026*	-5,79811*	0,16106*	*SSD*: 0,00326 (*SSD*) p-valor: 0,33810 *Raggedness*: 0,11471 (*Raggedness*) p-valor: 0,54550	*SSD*: 0,00309 (*SSD*) p-valor: 0,36340 *Raggedness*: 0,11471 (*Raggedness*) p-valor: 0,70260

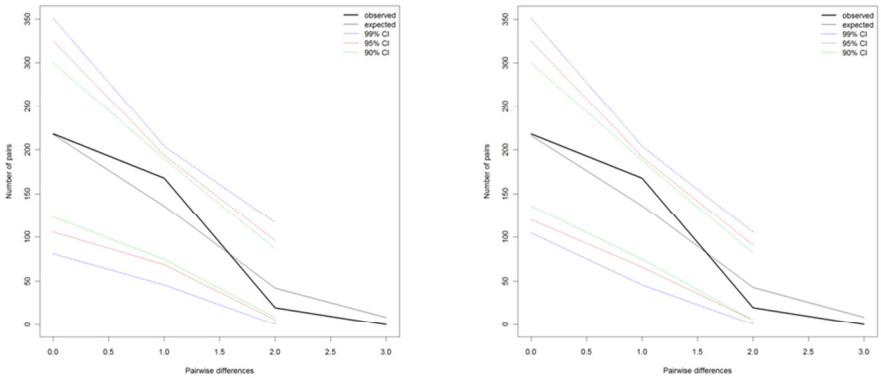
* Testes de neutralidade significativos com p<0,05 para D de Tajima e R_2 e p<0,02 para Fs de Fu.
** Livre de recombinação.

Demográfica *Espacial*

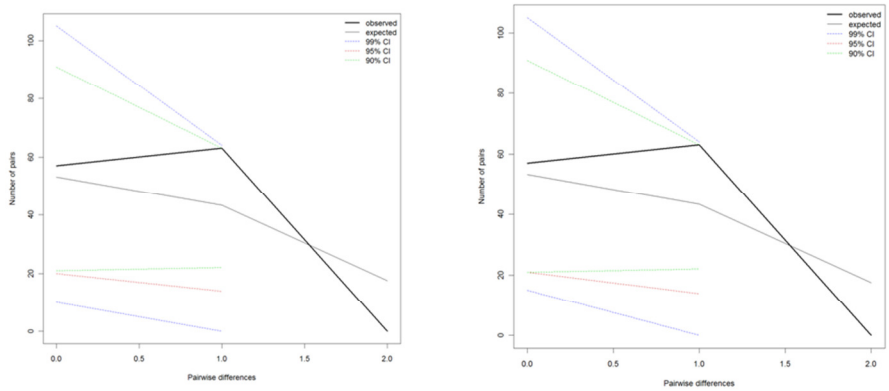
(A)



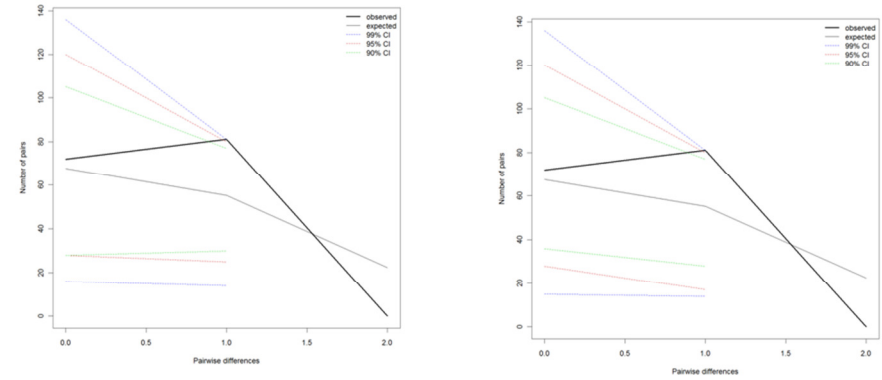
(B)



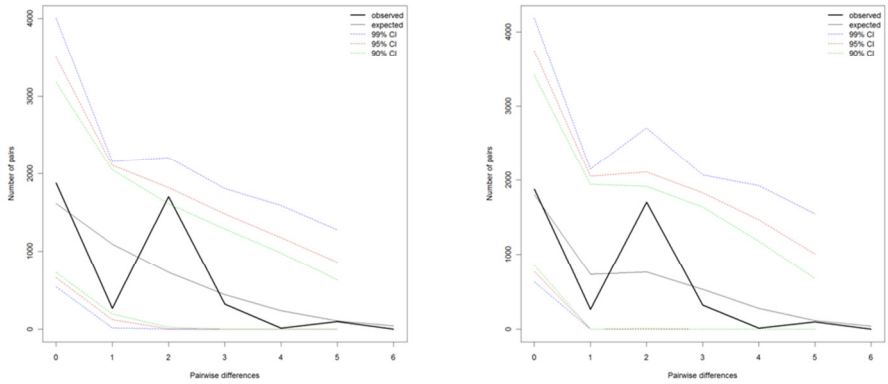
(C)



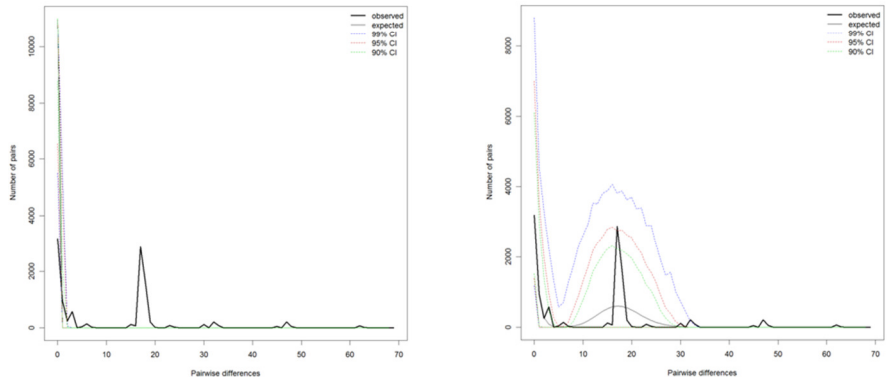
(D)



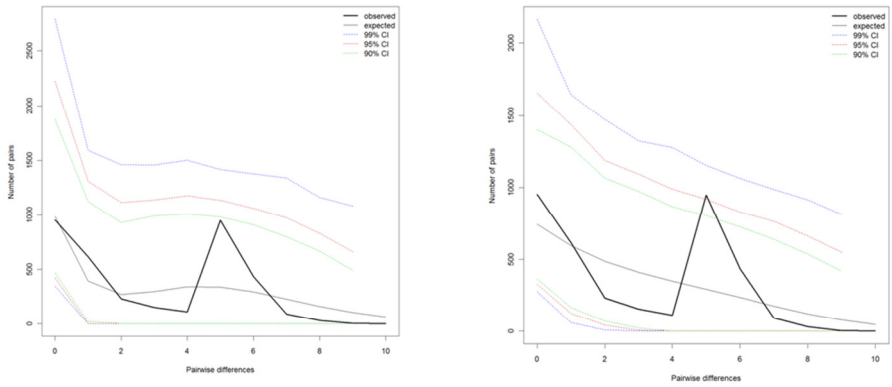
(E)


observed
expected
99% CI
95% CI
90% CI
Number of pairs
Pairwise differences

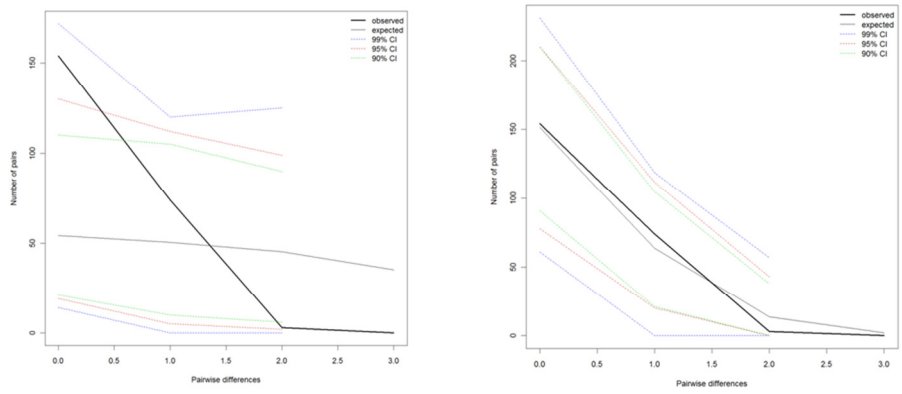
(F)


observed
expected
99% CI
95% CI
90% CI
Number of pairs
Pairwise differences

(G)


observed
expected
99% CI
95% CI
90% CI
Number of pairs
Pairwise differences

(H)


observed
expected
99% CI
95% CI
90% CI
Number of pairs
Pairwise differences

(I)

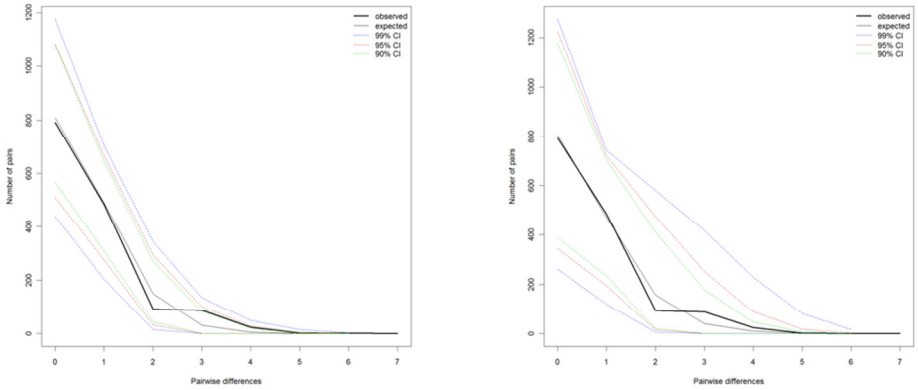


Figura 5. Gráficos gerados pela análise *Mismatch distribuition* referente às amostras de *C. hildmannianus*. Para cada região sequenciada foi gerado um gráfico de expansão demográfica (coluna esquerda) e um gráfico de expansão espacial (coluna direita), sendo (A) gene nuclear *PhyC*, (B) região intergênica cloroplastidial *atpI-atpH*, (C) região intergênica cloroplastidial *petL-psbE*, (D) região intrônica cloroplastidial *trnL*, (E) região intergênica cloroplastidial *trnQ-5'rps16*, (F) região intergênica cloroplastidial *psbJ-petA*e (G) região intergênica cloroplastidial concatenada *trnQ-5'rps16/psbJ-petA,* a qual foi analisada para os clados 1.2 (H) e 2.3 (I), de acordo com a NCPA. Em cada gráfico, a linha em cor preta representa as diferenças nucleotídicas entre pares de sequências observadas, enquanto a linha em cor cinza representa a distribuição esperada sob uma condição de expansão populacional. O intervalo de confiança é mostrado pelas linhas em azul para 99%, em vermelho para 95% e em verde para 90%.

2.3. Filogeografia e estrutura populacional

As Figuras 7, 8, 9 e 10 mostram os resultados da análise NCPA. O Apêndice 2 apresenta os indivíduos e as localidades para cada haplótipo gerado no TCS.

A Figura 7 mostra para o gene nuclear *PhyC* um clado de primeiro nível de aninhamento, sendo H1 e H2 os haplótipos mais amostrados. O haplótipo H1 é sugerido como o mais antigo segundo critérios estabelecidos em Castelloe e Templeton (1994) e a sua distribuição abrange todas as localidades amostradas, o H2, o segundo mais amostrado, está presente em quase todas as regiões, com exceção do estado de São Paulo.

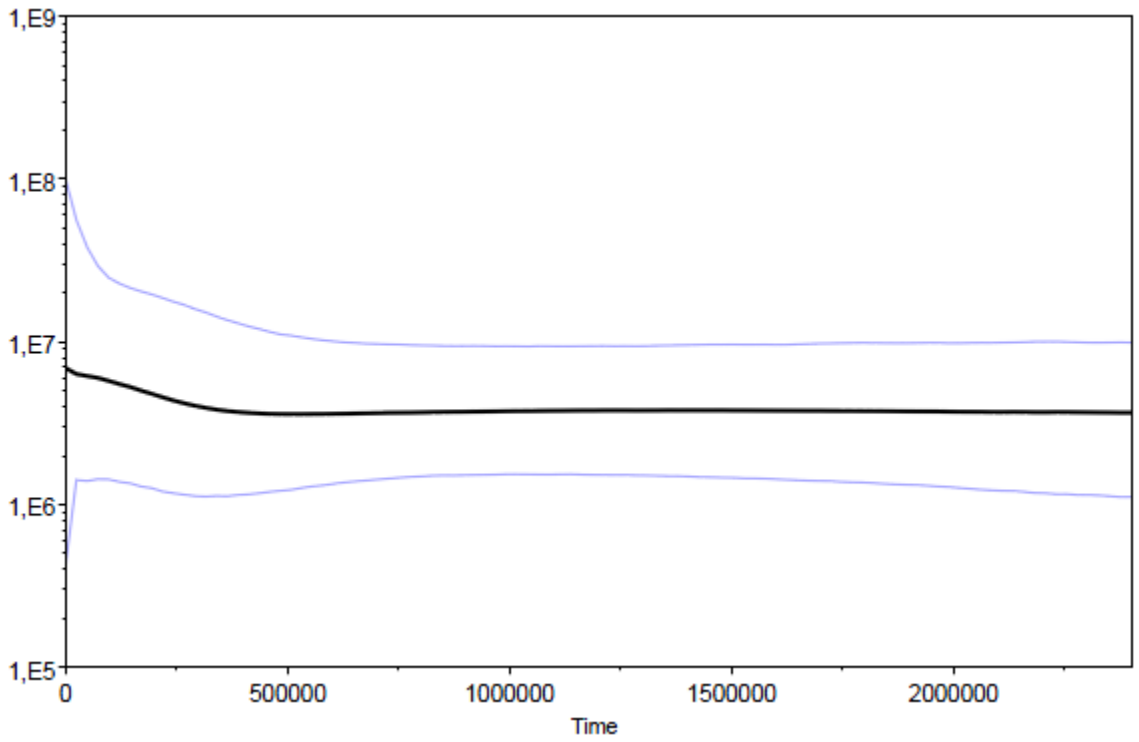


Figura 6. Gráfico gerado por *Bayesian Skyline Plot* (BSP) para as sequências *trnQ-5'rps16/psbJ-petA* concatenadas de *C. hildmannianus*. O tamanho populacional (eixo y) foi plotado em relação ao tempo em milhões de anos (eixo x) gerando a curva em preto, sendo as linhas em azul, o intervalo de confiança de 95%.

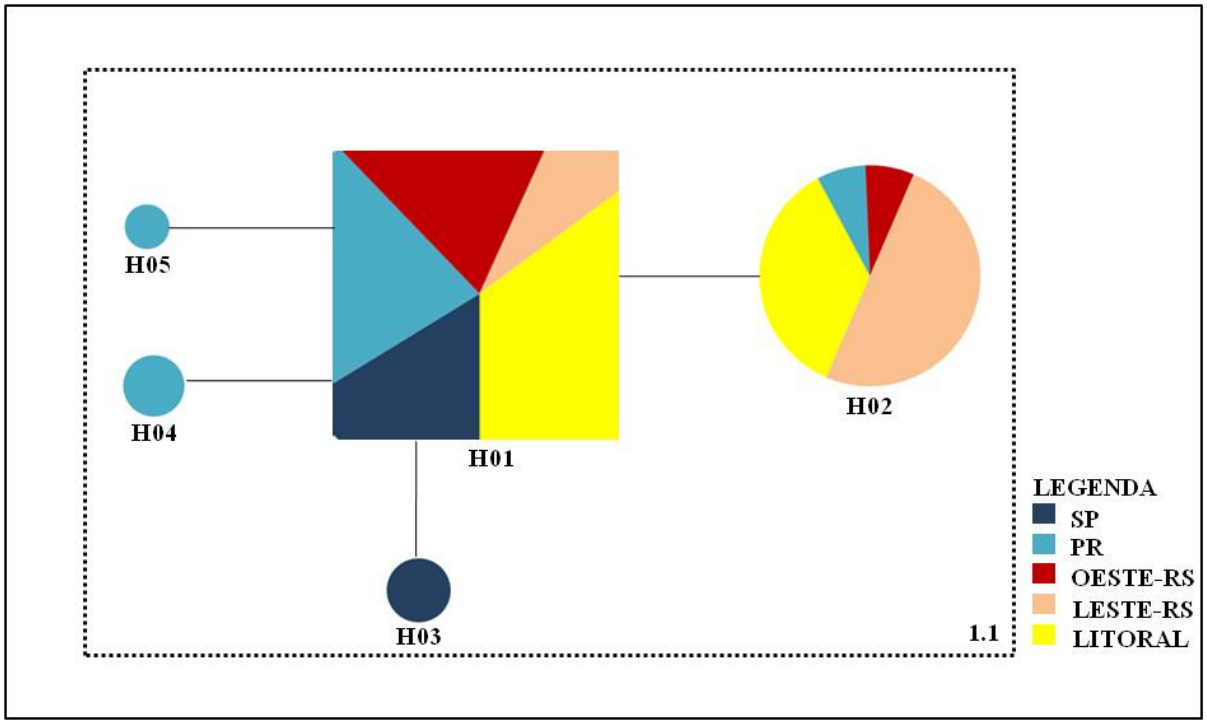


Figura 7. Rede não enraizada de haplótipos do gene nuclear *PhyC* em *C. hildmannianus*. Os haplótipos (H) estão numerados e representados por círculos de diferentes tamanhos, os maiores possuem mais indivíduos por haplótipo que os círculos menores, sendo o haplótipo mais antigo representado por um quadrado. As cores representam as localidades geográficas amostradas. O clado está numerado e organizado em um nível, cada linha da rede entre os haplótipos representa um passo mutacional e o primeiro nível está indicado por linhas pontilhadas.

A Figura 8 apresenta dois níveis de aninhamento para a região intergênica cloroplastidial *trnQ-5'rps16*. Os clados 1.1 e 1.3 abrangem a maior parte das regiões amostradas, com exceção de todo o litoral, sendo o haplótipo H01 o mais antigo, pertencente ao clado 1.3. Por outro lado, a distribuição do clado 1.2, o qual contém o haplótipo H02, o segundo mais amostrado, ocorrendo em todo o litoral e no estado do Rio Grande do Sul.

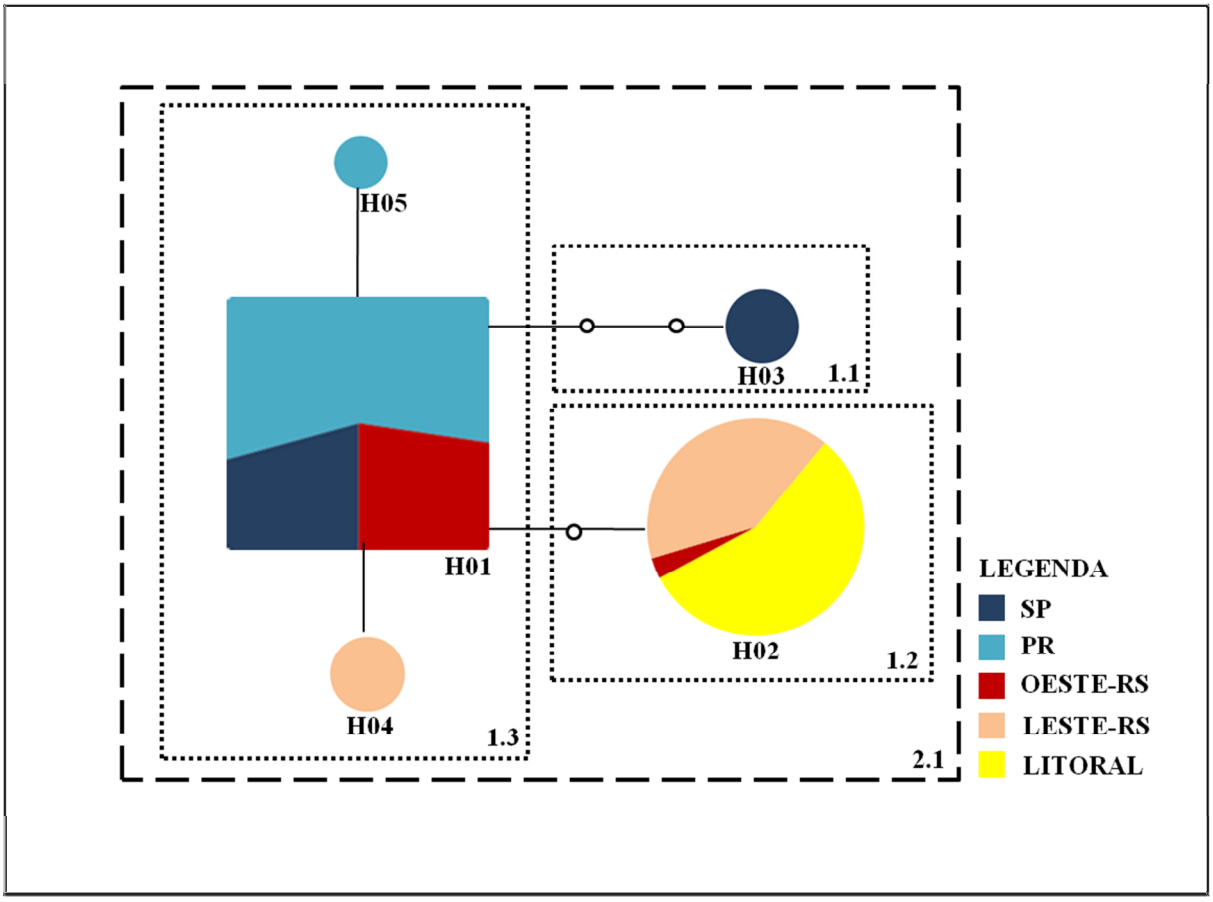


Figura 8. Rede não enraizada de haplótipos da região intergênica cloroplastidial *trnQ-5'rps16* em *C. hildmannianus.* Os haplótipos (H) estão numerados e representados por círculos de diferentes tamanhos, os maiores possuem mais indivíduos por haplótipo que os círculos menores, sendo o haplótipo mais antigo representado por um quadrado e os haplótipos hipotéticos por círculos vazios. As cores representam as localidades geográficas amostradas. Os clados estão numerados e organizados em níveis, cada linha da rede entre os haplótipos representa um passo mutacional. O primeiro nível está indicado por linhas pontilhadas e o segundo nível por traços.

A rede haplotípica para a região intergênica cloroplastidial *psbJ-petA* (Figura 9) tem três níveis de aninhamento, com o haplótipo H01 mais amostrado, pertencente ao clado 1.6, mais antigo e a sua distribuição está ausente somente no litoral. O haplótipo H02 é o segundo mais amostrado e está aninhado no clado 1.1, o qual se distribui nas localidades do litoral e leste do estado do Rio Grande do Sul.

Para as regiões intergênicas cloroplastidiais concatenadas, *trnQ-5'rps16* e *psbJ-petA* (Figura 10) os três níveis de aninhamento obtidos indicaram novamente a presença de dois haplótipos com maior distribuição: o H01 mais antigo nos estados de São Paulo, Paraná e oeste do estado do Rio Grande do Sul e o H02, no litoral e leste do estado do Rio Grande do Sul. Oito haplótipos estão ligados ao H01 por poucos passos mutacionais e suas distribuições ocorrem de forma conjunta nas populações com o haplótipo ancestral, e também em populações isoladas. Os haplótipos derivados do H02 por um passo mutacional estão distribuídos nas populações do litoral de forma conjunta. O H01 e H02 estão conectados por sete passos mutacionais, intermediários são encontrados nas populações de Jaguari, Caçapava do Sul, Barra do Ribeiro e Viamão, ou seja, no interior do Rio Grande do Sul, tanto na parte oeste como na leste. Um destes haplótipos intermediários possui outro derivado por três passos mutacionais que está na região de Pouso Novo, na Serra Gaúcha.

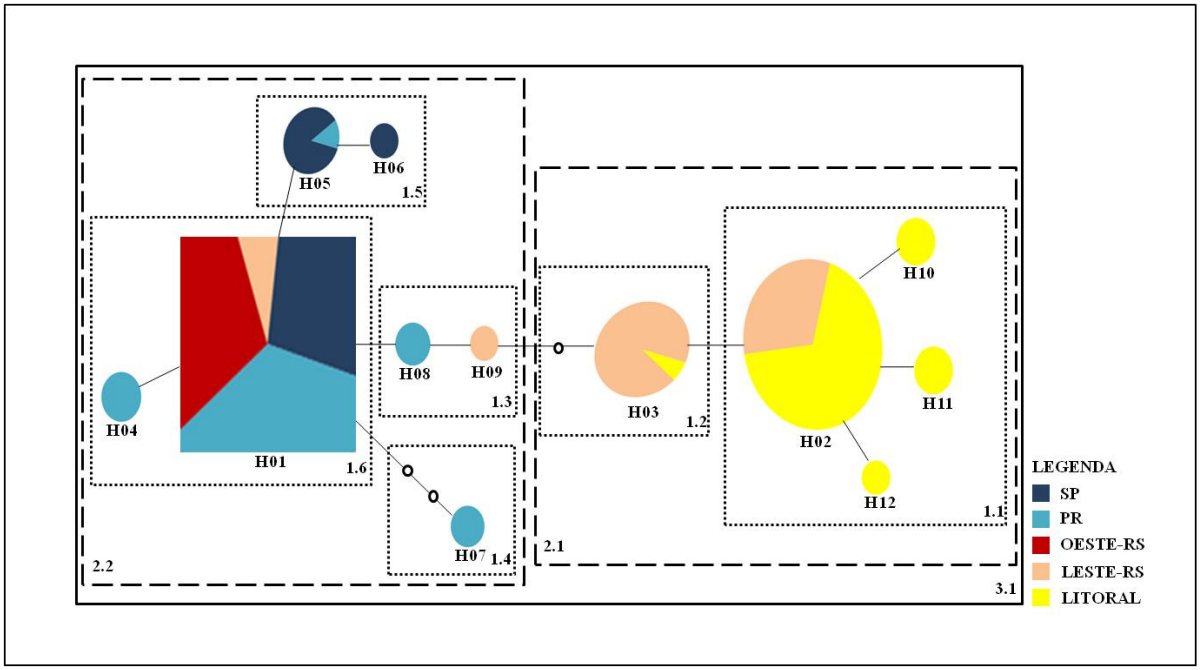


Figura 9. Rede não enraizada de haplótipos da região intergênica cloroplastidial *psbJ-petA* em *C. hildmannianus.* Os haplótipos (H) estão numerados e representados por círculos de diferentes tamanhos, os maiores possuem mais indivíduos por haplótipo que os círculos menores, sendo o haplótipo mais antigo representado por um quadrado e os haplótipos hipotéticos por círculos vazios. As cores representam as localidades geográficas amostradas. Os clados estão numerados e organizados em níveis, cada linha da rede entre os haplótipos representa um passo mutacional. O primeiro nível está indicado por linhas pontilhadas, o segundo nível por traços e o terceiro nível por linha contínua.

A Tabela 8 apresenta os resultados obtidos após a utilização da chave de inferência do Templeton, 2011 (disponível em http://darwin.uvigo.es/software/geodis.html) aplicada aos clados com resultados significativos para associação geográfica das redes haplotípicas

(Figuras 7, 8, 9 e 10). Os valores de Dc e Dn estão indicados no Apêndice 3. Os resultados mais freqüentes, entre os diferentes marcadores moleculares, foram fluxo gênico restrito com isolamento por distância e alguns eventos de expansão, porém, para o *trnQ-5'rps16* e a região concatenada *trnQ-5'rps16/psbJ-petA*, outras análises demográficas foram necessárias para continuar a chave de inferência (Tabela 8). Para tanto, em primeiro lugar foi analisado o cenário de fragmentação no passado seguida por expansão de área na alternativa 2 – passo13, sugerindo também, um teste de contato secundário implementado em Templeton (2001). Entretanto, este teste requer como pré-requisito a detecção estatisticamente significante de evento de fragmentação entre as linhagens (Templeton, 2001) e, portanto, não foi aplicado no presente trabalho. Em segundo lugar, a alternativa 1 foi analisada, e os três tipos de movimentos sugeridos pela chave de inferência foram discriminados utilizando o passo 21, o qual necessita de outro método para validar o crescimento populacional, como por exemplo, a *Mismatch distribution* (Tabela 7).

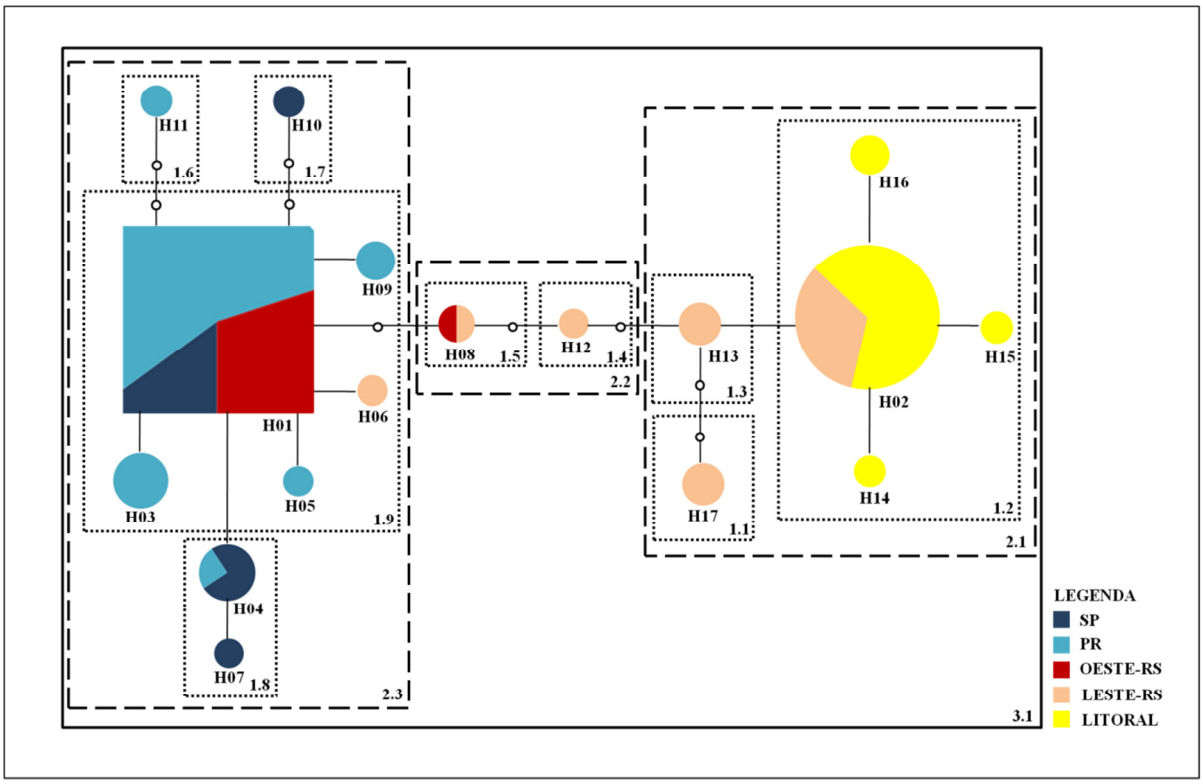


Figura 10. Rede não enraizada de haplótipos das regiõesintergênicas cloroplastidiais concatenadas *trnQ-5'rps16*e *psbJ-petA* em *C. hildmannianus.* Os haplótipos (H) estão numerados e representados por círculos de diferentes tamanhos, os maiores possuem mais indivíduos por haplótipo que os círculos menores, sendo o haplótipo mais antigo representado por um quadrado e os haplótipos hipotéticos por círculos vazios. As cores representam as localidades geográficas amostradas. Os clados estão numerados e organizados em níveis, cada linha da rede entre os haplótipos representa um passo mutacional. O primeiro nível está indicado por linhas pontilhadas, o segundo nível por traços e o terceiro nível por linha contínua.

Tabela 8. Resultados significativos do agrupamento de clados não enraizados obtidos pela chave de inferência de Templeton (2011) na análise das redes haplotípicas do gene nuclear *PhyC* e regiões intergênicas cloroplastidiais *trnQ-5'rps16* e *psbJ-petA*.

Sequência	Clado	Passos da chave de inferência
Nuclear		
PhyC	1.1	1-2-3-5-6-7 (Sim): Fluxo gênico restrito/dispersão, mas com alguma dispersão a longa distância.
cpDNA		
trnQ-5'rps16	2.1	1-2-3-5-6-13 (Sim): 1) Colonização de longa distância, maior extensão no passado com subsequente extinção em algumas áreas geográficas intermediárias, ou maior expansão de área no passado, todas as quais podem, eventualmente, ser relacionadas com subsequente fragmentação ou 2) Fragmentação no passado seguida por expansão de área.
psbJ-petA	1.1	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância.
	1.6	1-2-11-12 (Não): Expansão de área contígua.
	2.1	1-2-11-12 (Não): Expansão de área contígua.
	2.2	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância.
	3.1	1-2-11-12 (Não): Expansão de área contígua.
Concatenado (*trnQ-5'rps16* - *psbJ-petA*)	1.2	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância.
	2.3	1-2-11-12 (Não): Expansão de área contígua.
	3.1	1-2-3-5-6-13 (Sim): 1) Colonização de longa distância, maior extensão no passado com subsequente extinção em algumas áreas geográficas intermediárias, ou maior expansão de área no passado, todas as quais podem, eventualmente, ser relacionadas com subsequente fragmentação ou 2) Fragmentação no passado seguida por expansão de área.

A primeira e segunda AMOVA identificaram alta estruturação para todos os marcadores cloroplastidiais, exceto para o *petL-psbE,* cujo valor de p não foi significativo (Tabela 9). Dentre os agrupamentos testados de acordo com critérios geográficos na segunda AMOVA, foram realizadas duas análises com maior estruturação para a maioria dos marcadores.

A primeira estruturação considerou as populações divididas em dois grupos: o primeiro grupo com as populações dos estados de São Paulo, Paraná e região Oeste do Rio Grande do Sul, enquanto o outro grupo foi representado por populações do Leste do Rio Grande do Sul e região costeira deste estado e de Santa Catarina. A percentagem de variação entre os dois grupos variou de 67,67 a 87,63 % para os marcadores cloroplastidiais, enquanto a variação dentro das populações foi baixa, de 11,73 a 19,09% com valores de ΦCT, ΦSC e

ΦST significativos (Tabela 9), corroborando a estruturação populacional de acordo com a distribuição geográfica dos grupos. Entretanto, o gene nuclear *PhyC* apresentou baixo valor de estruturação entre os grupos (21,18 %) e uma maior variação dentro das populações (60,73).

A segunda estruturação, para a região cloroplastidial concatenada *trnQ-5'rps16/psbJ-petA*, considerou as populações divididas em quatro grupos: São Paulo e Paraná; Oeste-RS; Leste-RS; Litoral. Essa estruturação apresentou 74 % de variação entre os grupos e a variação dentro dos grupos foi de 17,69 %. A estruturação com esses quatro grupos foi utilizada para os testes do MIGRATE.

O coeficiente de correlação entre as distâncias genéticas e geográficas nas populações naturais de *C. hildmannianus* foram baixos, porém significativos para o gene nuclear *PhyC* (r = 0,2032, p = 0,01) e para as regiões cloroplastidiais, *petL-psbE* (r = 0,2541, p = 0,02), íntron *trnL* (r = 0,2570, p = 0,02), *trnQ-5'rps16* (r = 0,2234, p = 0,00), *psbJ-petA* (r = 0,1636, p = 0,02) e o concatenado *trnQ-5'rps16/psbJ-petA* (r = 0,20, p = 0,01), exceto para *atpI-atpH* (r = 0,1033, p = 0,15).

As estimativas de fluxo gênico inferidas com o programa MIGRATE e representadas por valores de taxa de imigração efetiva escalada por mutação (M) são mostradas na Tabela 10. Os resultados obtidos com o gene *PhyC* indicam que o grupo de SP/PR não recebe migrantes, porém doa uma alta taxa de migrantes para todos os grupos. A região Oeste-RS recebe migrantes de todos os grupos, com SP/PR sendo o grupo que doa a maior taxa de migrantes. Para o marcador cloroplastidial concatenado *trnQ-5'rps16-psbJ-petA*, SP/PR também não recebe migrantes e doa migrantes para todos, com exceção do Litoral. A região Oeste-RS novamente recebe migrantes de todos os grupos, com altas taxas de migração de SP/PR e Litoral, sendo este último, o grupo que recebe migrante somente do Leste-RS.

2.4. Relógio molecular e tempo de divergência

As análises bayesianas realizadas com o auxílio do programa BEAST para a primeira calibração estimaram a idade do ancestral comum mais recente (MRCA) de *C. hildmannianus* em 2,56 milhões de anos atrás, com a credibilidade da probabilidade *a posteriori* de 0,96, os valores do tamanho efetivo amostral ficaram entre 100 e 200. O intervalo de confiança de 95% mostrou a extensão da data do MRCA de 4,87 a 0,63 milhões de anos atrás. A topologia da árvore da datação dentro de eudicotiledôneas está apresentada na Figura 11.

Tabela 9. Resultados da análise de variância molecular (AMOVA) para o gene nuclear *PhyC* e as regiões cloroplastidiais não codificantes *atpI-atpH*, *petL-psbE,* íntron *trnL*, *trnQ-5'rps16*, *psbJ-petA* de populações de *C.hildmannianus*.

Marcadores	AMOVA global	Entre grupos		Entre populações dentro de grupos		Dentro das populações		Índices de fixação
	Índice de Φ_{ST}	%	variância	%	variância	%	variância	
		Primeira estruturação com dois grupos: (1) SP/PR + Oeste-RS; (2) Leste-RS e Litoral						
Nuclear								
PhyC	0,336	18,09	0,055	21,18	0,066	60,73	0,187	Φ_{SC}: 0,258; Φ_{ST}: 0,392; Φ_{CT}: 0,180
cpDNA								
atpI-atpH	0,698	87,63	0,459	N	N	15,93	0,083	Φ_{SC}: N; Φ_{ST}: 0,840; Φ_{CT}: 0,876
petL-psbE	1*	100	0,500	0	0	0	0	Φ_{SC}: 0 *; Φ_{ST}: 1 *; Φ_{CT}: 1
íntron trnL	1	100	0,500	0	0	0	0	Φ_{SC}: 0 *; Φ_{ST}: 1; Φ_{CT}: 1
trnQ-5'rps16	0,821	67,67	0,395	20,60	0,120	11,73	0,068	Φ_{SC}: 0,637; Φ_{ST}: 0,882; Φ_{CT}: 0,676
psbJ-petA	0,717	68,78	0,766	12,13	0,135	19,09	0,212	Φ_{SC}: 0,388; Φ_{ST}: 0,809; Φ_{CT}: 0,687
Concatenado *(trnQ-5'rps16/psbJ-petA)*	0,788	73,75	1,127	13,1	0,2	13,15	0,2	Φ_{SC}: 0,499; Φ_{ST}: 0,868; Φ_{CT}: 0,737
		Segunda estruturação com quatro grupos: (1) SP/PR; (2) Oeste-RS; (3) Leste-RS; (4) Litoral						
Concatenado *(trnQ-5'rps16/psbJ-petA)*	0,788	74,15	1,124	8,16	0,123	17,69	0,268	Φ_{SC}: 0,315; Φ_{ST}: 0,823; Φ_{CT}: 0,741

* Valores não significativos para $p<0.05$.

N: Valores negativos.

%: percentagem de variação

Tabela 10. Estimativas de fluxo gênico representadas por M para os marcadores moleculares cloroplastidiais e nuclear entre grupos com estruturação populacional em *C. hildmannianus*. Cada grupo com seta seguida por *i* mostra a taxa de migrantes deste para o grupo *i*.

Grupo, *i*	**Θ [xNemu]**	**M** SP/PR → *i*	**M** Oeste-RS → *i*	**M** Leste-RS → *i*	**M** Litoral → *i*
Nuclear					
PhyC					
SP/PR	0,0009	-	0	0	0
Oeste-RS	4,05e+9	5,82e+9	-	0,020	0,028
Leste-RS	0,000001	6,43e+8	0	-	0
Litoral	0,000004	5,97e+8	0	0	-
cpDNA					
Concatenado *(trnQ-5'rps16 / psbJ-petA)*					
SP/PR	0,0016	-	0	0	0
Oeste-RS	0,000006	9,98e+6	-	0,018	1,10e+7
Leste-RS	0,0006	6,22e+5	0	-	0
Litoral	0,0003	0	0	3,83e+5	-

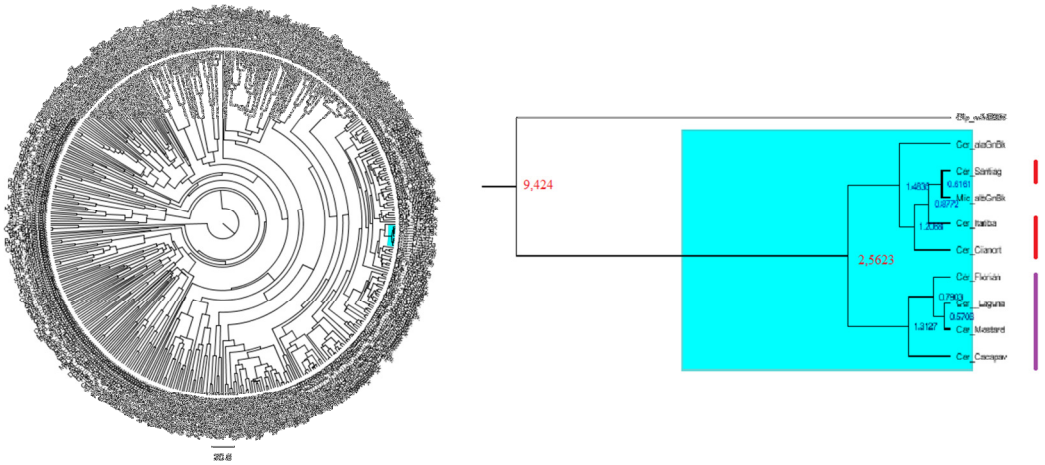


Figura 11. Resultado combinado da filogenia e calibração de eudicotiledôneas por análise bayesiana. À esquerda, a representação de toda a topologia, e ressaltado em azul, os ramos referentes à *C. hildmannianus*. À direita, uma visão ampliada dos ramos mostrando em vermelho as datas obtidas com alta probabilidade *a posteriori*. Os traços representam os indivíduos das localidades amostradas no presente trabalho, da planície costeira e Leste-RS em roxo, e em vermelho, regiões do Oeste-RS e PR. Os demais táxons foram obtidos do *GenBank*.

Os dados gerados na segunda calibração mostraram 11 eventos cladogenéticos intraespecíficos, com probabilidade posterior de 0,95 a 1 e os valores do tamanho efetivo amostral foram superiores a 200. O MRCA com 2,74 milhões de anos atrás foi dividido em dois grandes clados e um terceiro clado menor (Figura 12). O clado com 2,18 milhões de anos atrás compõem as populações do Leste-RS e Litoral, com quatro ramos datados: 1,72 milhões de anos atrás entre as populações do clado; 532 mil anos atrás para os indivíduos de Pouso Novo-RS; 783 e 295 mil anos atrás entre as populações de Florianópolis-SC. O segundo grande clado (2,1 milhões de anos atrás) engloba as populações de SP, PR, Oeste-RS e Leste-RS com três ramos datados: 1 milhão de anos atrás nas populações de Guarapuava-PR, Cantagalo-PR e Cianorte-PR; 1,13 milhões de anos atrás Cantagalo-PR, Itu-SP e Itatiba-SP. O terceiro clado é representado por um ramo datado em 1,12 milhões de anos atrás, para os indivíduos de Jaguari-RS e Barra do Ribeiro-RS. A extensão dos valores para cada evento datado (intervalo de confiança de 95%) podem ser visualizadas na Figura 12. Esta calibração foi realizada com o relógio molecular relaxado lognormal, pois foi o modelo com maior probabilidade marginal (ln -1970,958) entre os relógios testados (exponencial: ln -1980,302 e estrito: -210,243).

2.5. Análise biogeográfica

Os resultados obtidos da primeira análise em RASP mostraram três eventos de dispersão e vicariância concordantes com a distribuição geográfica: um na área entre São Paulo e Paraná com o leste do Rio Grande do Sul; e eventos entre o leste do Rio Grande do Sul com o oeste e com o litoral (Figura 13). A topologia da árvore demonstra que as regiões de São Paulo e Paraná e interior (oeste e leste) do Rio Grande do Sul podem ser mais ancestrais em relação ao litoral, embora muitas áreas do litoral tenham idade antiga. Na segunda análise, os eventos de vicariância e dispersão não são condizentes com os tipos de solo (Figura 14).

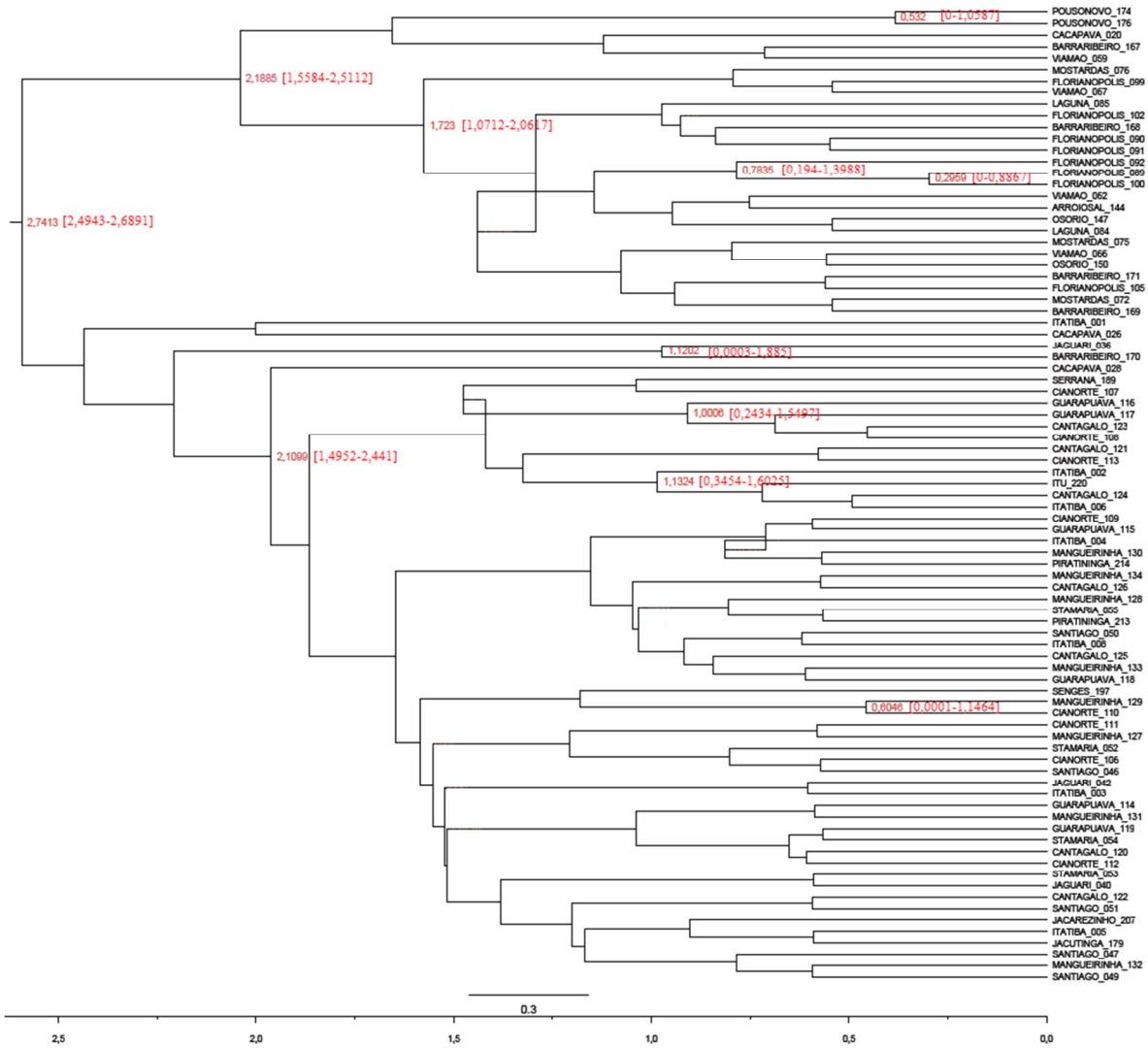


Figura 12. Resultado combinado da calibração intraespecífica de *C. hildmannianus* por análise bayesiana. As idades dos nós com probabilidade > 0,95 estão indicadas juntamente com o respectivo intervalo de confiança de 95% mostrado dentro de colchetes.

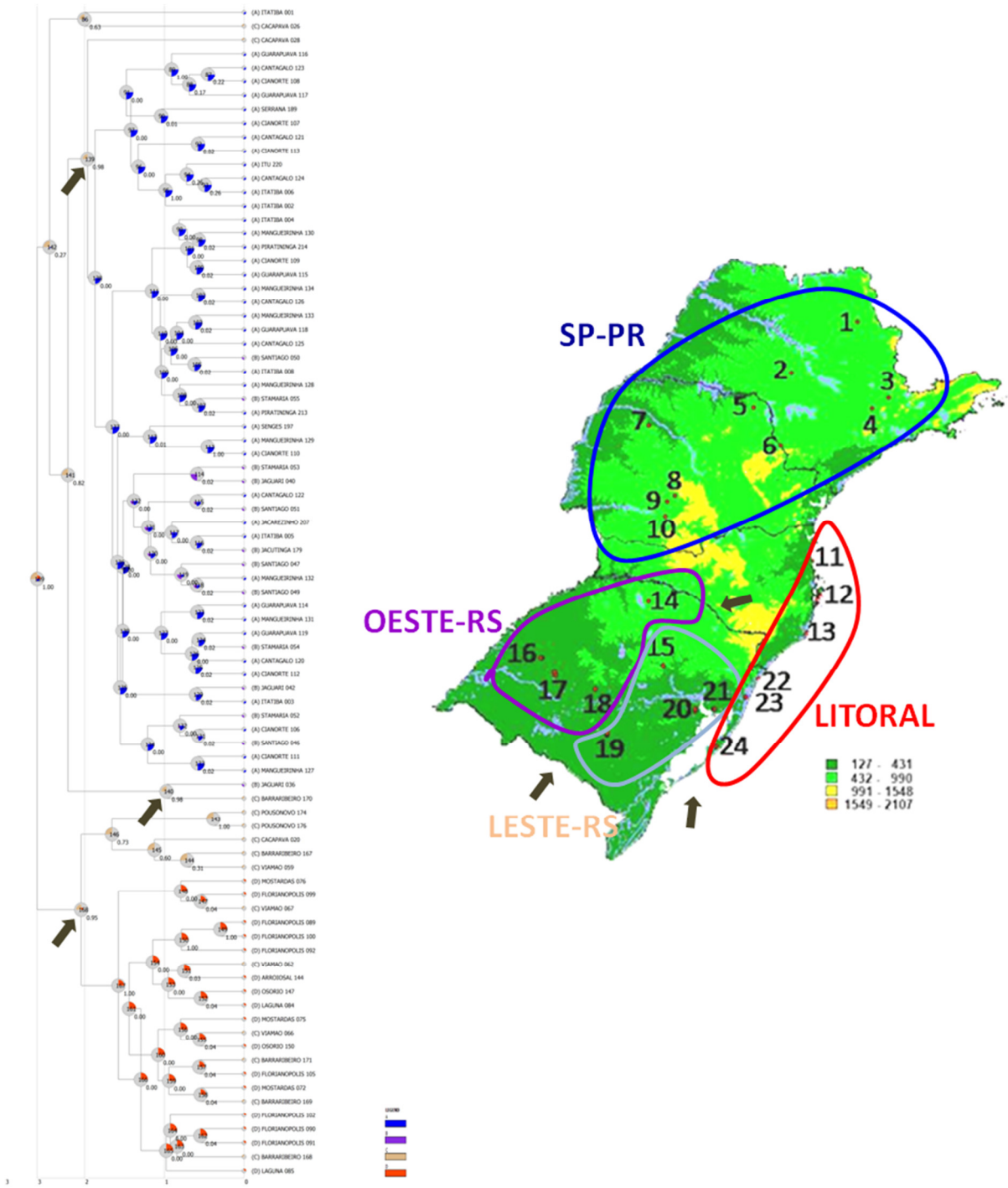


Figura 13. Árvore obtida em RASP à esquerda, com as probabilidades para cada nó com as regiões geográficas relacionadas aos eventos de dispersão e vicariância. À direita, mapa mostrando as quatro grandes áreas geográficas e os pontos de coleta. As setas mostram os eventos de dispersão e vicariância com probabilidades maiores que 0,95. Localidades: 1 - Serrana-SP; 2 - Piratininga-SP; 3 - Itatiba-SP; 4 - Itu-SP; 5 - Jacarezinho-PR; 6 - Sengés-PR; 7 - Cianorte-PR; 8 - Guarapuava-PR; 9 - Cantagalo-PR; 10 - Manguerinha-PR; 11 - Penha-SC; 12 - Florianópolis-SC; 13 - Laguna-SC; 14 - Jacutinga-RS; 15 - Pouso Novo-RS; 16 - Santiago-RS; 17 - Jaguari-RS; 18 - Santa Maria-RS; 19 - Caçapava do Sul-RS; 20 - Barra do Ribeiro-RS; 21 - Viamão-RS; 22 - Arroio do Sal; 23 - Osório-RS; 24 - Mostardas-RS.

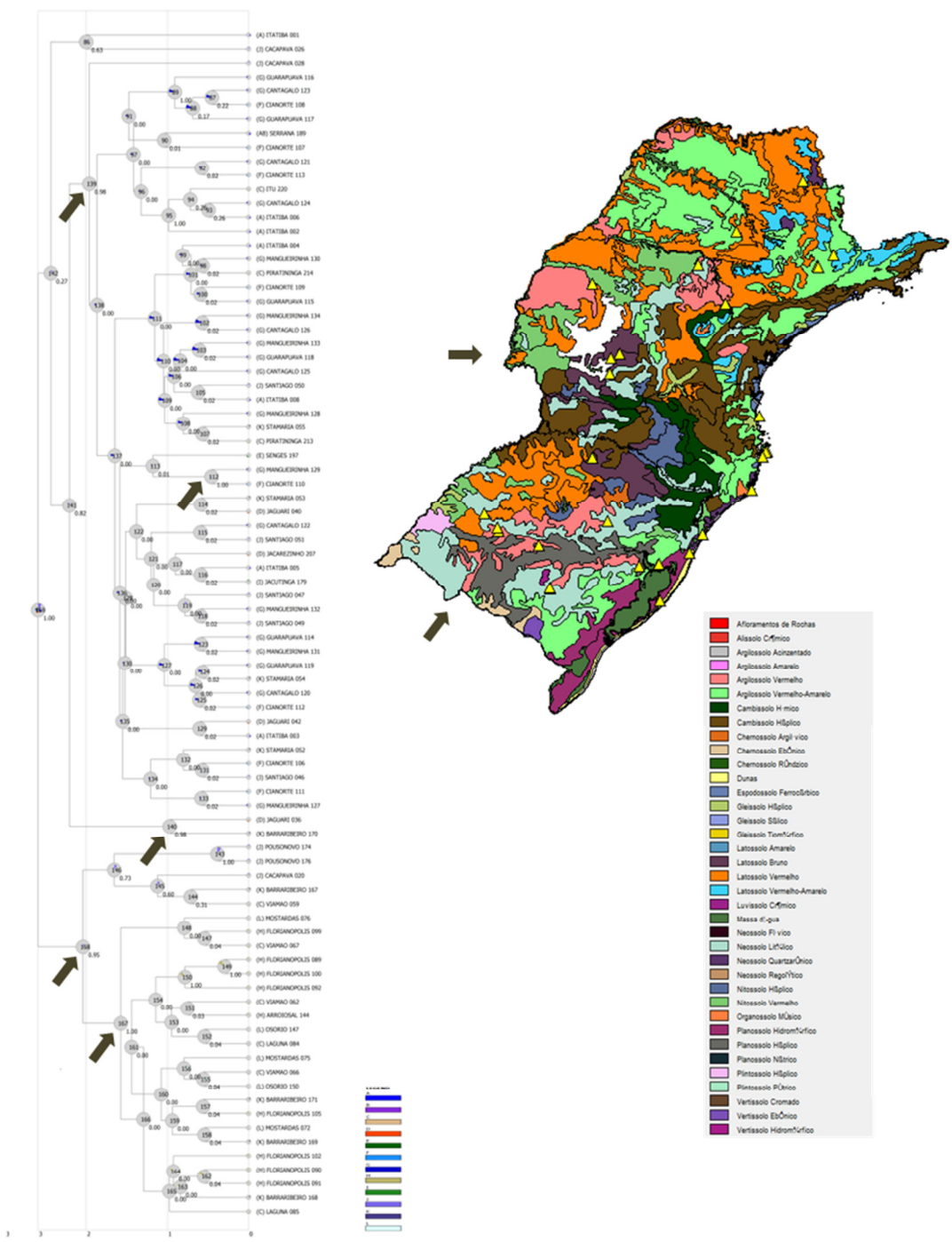


Figura 14. Resultado da análise em RASP em relação aos diferentes tipos de solo. À esquerda, a árvore com as probabilidades para cada nó e o tipo de solo que melhor explica os eventos de dispersão e vicariância. À direita, mapa mostrando os tipos de solo e os pontos de coleta em amarelo. As setas mostram os eventos com probabilidades maiores que 0,95.

2.6. Modelagem de nicho ecológico e paleomodelagem

A distribuição de *C. hildmannianus* foi gerada pelo modelo de nicho ecológico para o presente e paleomodelagem para o passado (Figura 15). A validação do modelo foi realizada por meio dos valores de AUC das três análises, sendo de 0,959 para o presente, 0,953 para o modelo MIROC e 0,939 para o modelo CCSM, os quais indicam uma alta precisão dos modelos (Swets, 1988). O modelo gerado pela atual distribuição da espécie corrobora com as informações obtidas nos herbários e com as localidades previamente propostas para as coletas deste trabalho, com exceção de dois pontos de herbário ausentes no modelo previsto. As distribuições geradas para o passado indicam a probabilidade de uma maior área de ocorrência da espécie, incluindo as áreas do interior do estado de SC, principalmente a oeste, sendo uma área de conectividade para a região Sul do Brasil. Provavelmente as diferenças na probabilidade da distribuição prevista para o passado, entre os modelos MIROC e CCSM, são decorrentes dos parâmetros utilizados para iniciar as simulações, enquanto o CCSM inicia de uma condição fria acoplado às simulações do LGM, o modelo MIROC começa a partir de condições modernas (Weber et al., 2007).

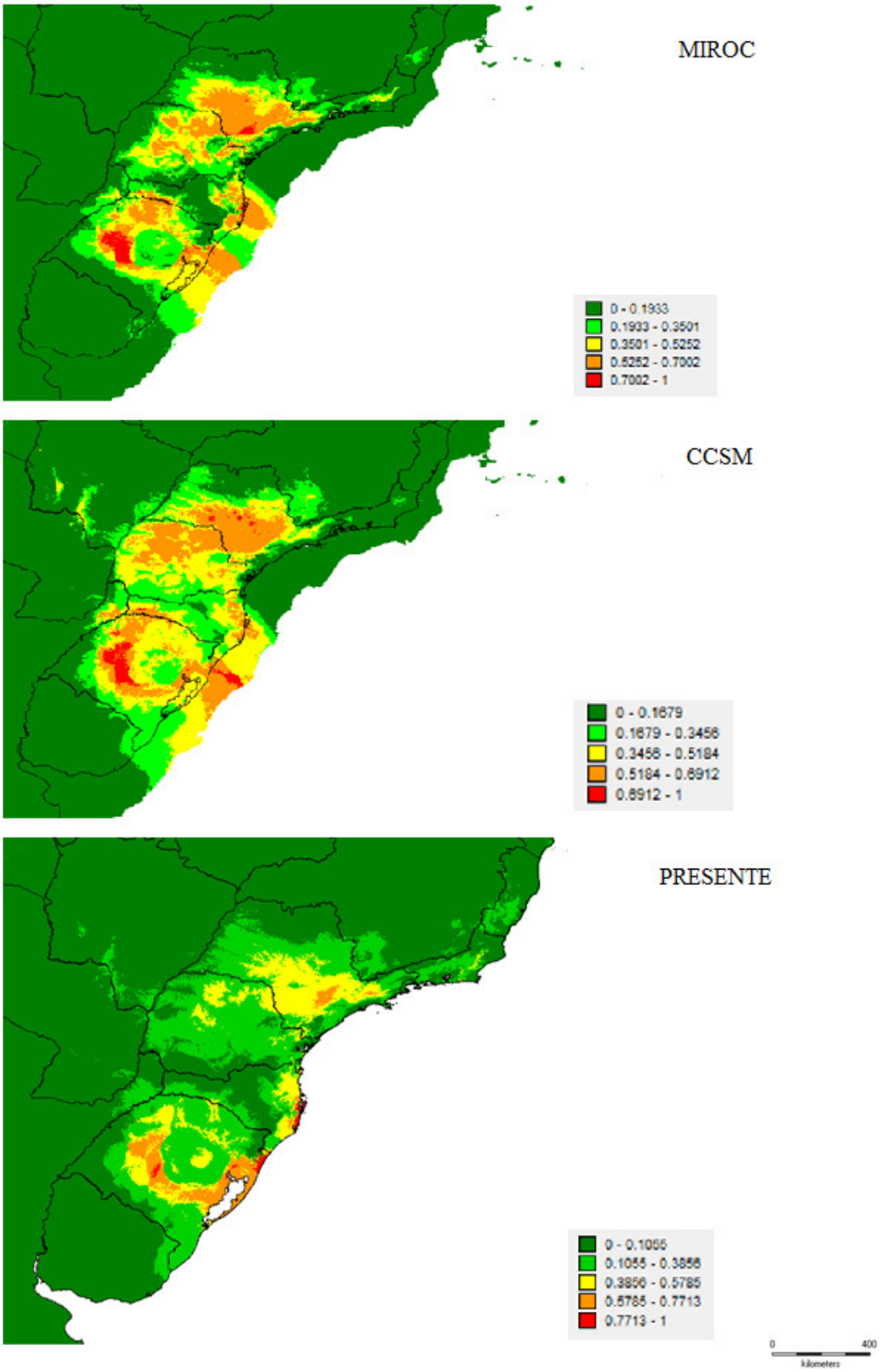


Figura 15. Modelos de nicho ecológico gerado no programa MaxEnt para o Passado (MIROC e CCSM) e Presente em *C. hildmannianus*. As áreas em verde escuro representam o limiar de corte para cada modelo testado, os quais indicam os lugares em que não há probabilidade de ocorrência da espécie. A partir deste limiar, a probabilidade de ocorrência da espécie aumenta até 1, representado pela cor vermelha.

V. DISCUSSÃO

O presente estudo foi realizado com análises de populações de *C. hildmannianus*, uma espécie da família Cactaceae. Os resultados foram obtidos a partir de sequências gênicase mostraram que a história da espécie é complexa, com múltiplos eventos que incluíram principalmente expansões de áreas, em diferentes escalas temporais, e fluxo gênico restrito geograficamente. A estrutura para a espécie divide suas populações em dois grandes grupos, com área de contato e eventos homogeneizados dentro de cada grupo.

1. A VARIABILIDADE GENÉTICA EM *Cereus hildmannianus*

Para trabalhos filogeográficos e de estruturação populacional são necessários marcadores com variação neste nível de análise. Para animais já existe muita informação consolidada na literatura com o gene *COI* (Chase et al., 2007). Para plantas, por outro lado, informações deste nível ainda são incipientes (Zimmer e Wen, 2012; Borsch e Quandt, 2009; Chase et al., 2007; Shaw et al. 2007). Então, para o desenvolvimento deste trabalho, inicialmente definimos as regiões a serem utilizadas fazendo uma prospecção de regiões gênicas com variabilidade intra e interpopulacional. Dentre as 14 regiões avaliadas a partir de marcador molecular no presente trabalho, seis apresentaram variação populacional, sendo cinco marcadores cloroplastidiais e um marcador nuclear (Tabela 4).

Vários trabalhos têm sido realizados para a procura de regiões com maior variabilidade em estudos filogenéticos, filogeográficos e DNA *barcoding* (Borsch e Quandt, 2009; CBOL, 2009; Chase et al., 2007; Shaw et al., 2005; Shaw et al. 2007; Ebert e Peakall, 2009). A seguir, as regiões cloroplastidiais utilizadas neste trabalho foram comparadas com alguns trabalhos, os quais utilizaram os *primers* descritos em Shaw e colaboradores (2005; 2007), para a maioria deles. A comparação será feita com os trabalhos de Shaw, pois dentre os artigos publicados (para mais detalhes, ver Ebert e Peakall, 2009), foram os que mais incluíram diferentes famílias de plantas para testar conjuntos de *primers*.

- Região intergênica cloroplastidial *atpI-atpH*. Sequências de 488 bp com uma variabilidade de 0,82 % em *Podostemum ceratophyllum* (Podostemaceae)

(Fehrmann et al., 2012), comparada a 0,33 % para *C. hildmannianus*. Em trabalhos de filogeografia de *Cardamine scutata* (Brassicaceae), uma taxa de 1,9% foi obtida (Lihov et al., 2010). Somente dois sítios polimórficos foram encontrados em 1104 bp nas sequências de *Fagus engleriana* (Fagaceae) (Lei et al., 2012).

- Região intergênica cloroplastidial *trnL-trnL-trnF*. O íntron *trnL* é sugerido comoapropriado para estudos de DNA *barcoding* (Taberlet et al., 2007) e descrito como fonte de variabilidade quando analisado concatenadamente com a região *trnL-trnL-trnF* (Shaw et al., 2005). Mas de acordo com os valores de PICs, o íntron *trnL* não foi considerado apropriado para estudos filogeográficos (Shaw et al., 2005). Apesar da baixa taxa de evolução, há análises filogeográficas com este íntron em *Lychophora ericoides* (Asteraceae) (Collevatti et al., 2009) e no presente trabalho.Já a região *trnL-trnF* apresentou seis sítios polimórficos e dois *gaps* na filogeografia de *Embothrium coccineum* (Proteaceae) (Russell, et al., 2011), região que não apresentou variabilidade em *C. hildmannianus.*
- Região intergênica cloroplastidial *psbD-trnT*. Para *Lobelia rhynchopetalum* (Campanulaceae), 468 bp foram alinhados com um total de 9,2 % de variação (Geleta e Bryngelsson, 2012). Um estudo filogeográfico em *Podostemum ceratophyllum* (Podostemaceae) encontrou 0,41 % de variação em 739 bp (Fehrmann et al., 2012), sendo que para *C. hildmannianus*, não houve variabilidade em 668 bp.
- Região cloroplastidial *trnS-trnS-trnG*. Fehrmann e colaboradores (2012) obtiveram 569 bp com 0,18 % de variação para *Podostemum ceratophyllum* (Podostemaceae). Em *C. hildmannianus* não houve variabilidade em 476 bp da região do íntron *trnS-trnS* e em 970 bp da região *trnS-trnG*.
- Região intergênica cloroplastidial *matK-trnK*. Altas taxas de variabilidade apontam o *matK*, combinado a outras sequências, como uma das melhores regiões para DNA *barcoding* (Yesson et al., 2011; CBOL, 2009; Starr et al., 2009; Chase et al., 2007), porém pode haver problemas de especificidade dos *primers* nas diferentes famílias de plantas (Chase et al., 2007). Para *C. hildmannianus* não foi encontrada variabilidade nos 1224 bpsequenciados,

embora, em *Lobelia rhynchopetalum* (Campanulaceae), os autores encontraram 8,5 % em 307 bp (Geleta e Bryngelsson, 2012).

- Região intergênica cloroplastidial *trnH-psbA*. Esta região foi sugerida como uma das mais variáveis por Shaw e colaboradores (2005; 2007) e como potencial DNA *barcoding* (Shaw et al., 2007; Kress et al., 2005). Críticas a esse marcador estão relacionadas ao grande número de deleções e inserções e à alta ambiguidade entre os táxons, o que pode enviesar as inferências entre eles. Além disso, críticas também existem devido a diferenças no tamanho das sequências, que variam entre 300 a >1000 bp, devido a cópias de rpl22 e rps16 em algumas espécies (Chase et al., 2007). Em *C. hildmannianus*, não foram encontradas inserções ou deleções, nem variação entre as populações nas análises dos fragmentos de aproximadamente 360 bp (Tabela 4). Por outro lado, 24 sítios variáveis, para fragmentos de 315 bp, foram utilizados para inferir a filogeografia de *Oxyriadigyna* (Polygonaceae) (Allen et al., 2012). Este marcador também permitiu estudar a história populacional em *Encelia farinosa* (Asteraceae) com 10 haplótipos obtidos de aproximadamente 450 bp (Fehlberg e Ranker, 2009).
- Região intergênica cloroplastidial *trnQ-5'rps16*. Em *Juglans mandshurica* (Juglandaceae), fragmentos de 605 bp apresentaram uma única substituição nucleotídica e foram concatenados com mais seis marcadores para o estudo filogeográfico (Bai et al., 2010). Já em *Psammosilene tunicoides* (Caryophyllaceae), aproximadamente 805 bp apresentaram cinco substituições nucleotídicas e 2 a 27 bp com *indels* (Zhang et al., 2011).
- Região intergênica cloroplastidial *trnL-trnT*. Allen e colaboradores (2012) obtiveram 68 sítios variáveis em 910 bp desta região *Oxyria digyna* (Polygonaceae). Em *Lobelia rhynchopetalum* (Campanulaceae), 9 % de variabilidade foi descrita para 525 bp (Geleta e Bryngelsson, 2012). No presente trabalho, as sequências obtidas não apresentaram variação em 348 bp.
- Região intergênica cloroplastidial *psbJ-petA*. Este marcador, quando concatenado com *atpI-atpH* têm sido informativos em estudos filogeográficos, sendo encontradas também, regiões de cpSSR e *indels* (Sebastiani et al., 2004; Provan et al., 2004, Shaw et al., 2007). Em *C. hildmannianus*, a região *psbJ-*

petA foi a região com maior variabilidade dentre os marcadores testados no presente trabalho (Tabela 5).

As comparações realizadas mostram disparidades ou até mesmo ausência de variabilidade das regiões não codificantes de DNA cloroplastidial descritas como as melhores alternativas, ou seja, demonstram que as generalizaçõespodem não refletir a utilidade destes marcadores, pois desvios dessas taxas são comuns (Shaw, 2007). Desta forma, a avaliação das regiões cloroplastidiais para um estudo filogeográfico, mostrou que somente cinco das 13 regiões propostas em Shaw (2005, 2007) apresentam variação.

Embora não seja o escopo deste trabalho fazer afirmações sobre as melhores regiões cloroplastidiais a serem utilizadas em estudos filogeográficos ou filogenéticos e DNA *barcode*, as considerações sobre essas regiões cloroplastidiais em *C. hildmannianus* podem ser úteis em investigações de DNA *barcode* em *Cereus*, uma vez que as cactáceas possuem vários táxon ameaçados de extinção (Zappi et al., 2011), como consequência, entre outras, de coleta ilegal, por sua importância ornamental e para coleções particulares, sendo útil para a rápida identificação de material apreendido ilegalmente (Yesson et al., 2011), principalmente na ausência de flores (Anderson 2001; Hunt et al., 2006).

2. TEMPO DE DIVERGÊNCIA MOLECULAR

A família das cactáceas tem sua origem associada ao Terciário médio (Hershkovitz e Zimmer, 1997; Arakaki et al., 2011) e os padrões de divergência de sequências dentro de Cactaceae parecem refletir uma rápida radiação (Arakaki et al., 2011; Butterworth e Edwards, 2008; Nyffeler, 2002) no final do Terciário, entre 10 a 5 milhões de anos atrás (Arakaki et al., 2011). Um recente trabalho de datação para Cactaceae utilizou uma supermatriz com sequências cloroplastidiais de diversos táxons de eudicotiledôneas para construir uma filogenia, calibrando então essa filogenia a partir de fósseis (Lendel et al., dados a publicar). Essa calibração estimou o tempo de diversificação do ancestral comum mais recente (TMRCA) de *C. hildmannianus* em 2,56 milhões de anos atrás (Figura 11), o que remete a especiação deste ao final do período Terciário.

Concordando com esse período, a diversificação de alguns táxons das FTSS foi estimada como sendo pré-pleistocênica por Pennington e colaboradores (2004), que questionaramo papel das mudanças climáticas do Pleistoceno como fator crucial na especiação em plantas pertencentes a essas florestasna América do Sul. Outro exemplo são as populações de *Astronium urundeuva*, confinadas às FTSS, que se mostraramestruturadas em três principais grupos: nordeste, centro do Brasil e na região da Bolívia, Argentina e Paraguai, com a divergência entre as populações dessa distribuição datadas no pré-Pleistoceno (Caetano et al., 2008). Recentemente, Pennington e colaboradores (2009) observaram que espécies endêmicas confinadas a um único núcleo de FTSS possuem datações estimadas no pré-Pleistoceno.

Portanto, tendo em vista que a família das cactáceascaracterizam as FTSS neotropicais, as datações e filogenias dentro deste grupo (Arakaki et al., 2011; Edwards et al., 2005; Lendel et al., dados a publicar; presente trabalho) corroboram com o padrão filogenético e biogeográfico sugerido para as demais espécies representantes dessas florestas (Pennington et al., 2009; Pennington et al., 2004; Caetano et al., 2008).

Embora não haja um consenso sobre a diversificação neotrópica (Hoorn et al., 2011; Rull, 2011; Pennington et al., 2009), se Terciária ou Quaternária, Rull (2008) sugere um processo de especiação contínua e apesar das influências paleogeográficas e paleoclimáticas interagirem no complexo processo de diversificação, uma alta quantidade de especiação foi encontrada no período de 2,6 milhões de anos atrás, quando as mudanças climáticas começaram a ser um forte fator ambiental.

Com o TMRCA estimado em 2,56 milhões de anos atrás, no início dos períodos glaciais, as mudanças paleoclimáticas durante o Pleistoceno podem ter sido um importante fator na dinâmica populacional, portanto, na história evolutiva de *C. hildmannianus.* O evento demográfico inferido pela análise BSP sugere uma relativa estabilidade em *C. hildmannianus*, seguida por expansão (Figura 6). Considerando a biogeografia histórica e a filogenia das tribos na qual esta espécie está classificada (Edwards et al., 2005; Ritz et al., 2007; Nyffeler, 2002) e a atual distribuição (Taylor e Zappi, 2004), a região Andina e o núcleo de Missiones das FTSS foram provavelmente o local da diversificação de *C. hildmannianus.*

Esta estabilidade e manutenção das populações têm sido encontradas em táxons relacionados às FTSS, mais especificamente, aos núcleos desta floresta, como decorrentes da atual hipótese para as FTSS, conhecida como conservadorismo de nicho (*niche conservatism*) (Hughes e al, 2013; Pennington et al., 2009). Isto significa que as

espécies em geral estão confinadas ou concentradas em núcleos destas florestas, e que é mais provável que os táxons das FTSS se dispersem para um local distante do que as espécies de biomas adjacentes entrem nestas florestas (Pennington et al., 2009). Talvez isso explique a ocorrência de *C. hildmannianus* em toda a região sul do Brasil, que inclui parte do núcleo de Missiones, das FTSS. Esse cacto tem sua distribuição descrita para as florestas úmidas e semiúmidas no sul do Brasil, sempre associados a rochas e solos secos (Bauer e Waechter, 2006; Bruxel e Jasper, 2005; Gonçalves e Waechter, 2003; Taylor e Zappi, 2004, presente trabalho).

Os cactos são conhecidos como plantas CAM, as quais possuem um sistema de fotossíntese alternativo que possibilita a maximização do uso eficiente da água. O final do Mioceno foi marcado por clima frio e seco, com a predominância de plantas C4, como as gramíneas (Arakaki et al., 2011; Edwards et al., 2010), e também, maior diversificação de plantas adaptadas a habitats secos (Arakaki et al., 2011; Antonelli et al., 2010). Neste período, um fenômeno global de redução da precipitação, aumentando a aridez, juntamente com a diminuição dos níveis de gás carbônico (CO_2) propiciaram um ambiente excelente para as plantas C4 e CAM (Arakaki et al., 2011). De fato, baixos valores de CO_2 parecem estar sincronizados com a intensificação das glaciações no Hemisfério Norte e Sul, além da amplitude dos ciclos glaciais e interglaciais no período de 2,5 milhões de anos atrás (Tripati et al., 2009).

Para as populações, os tempos de divergência mostrados na Figura 12 foram associados à reconstrução da expansão geográfica, no programa RASP (Figuras 13 e 14). As datas calculadas para os nós de cada ramificação representam uma média. Assim, considerando o intervalo de confiança entre elas, há sobreposição de praticamente todos os eventos. Os tempos de divergências para as populações mostram 11 eventos cladogenéticos, sendo o primeiro evento entre dois grandes grupos populacionais: SP/PR + Oeste-RS e Leste-RS + Litoral, apenas as populações de Itatiba em São Paulo, e as populações de Caçapava do Sul, Barra do Ribeiro e Jaguari no Rio Grande do Sul, aparecem em ambos os grupos (Figura 12). Essa primeira divergência foi corroborada por eventos de vicariância e dispersão na análise em RASP (Figura 13). Dentro do grupo de SP/PR + Oeste-RS, o haplótipo H09 parece ser mais derivado, pois o intervalo de confiança não se sobrepõe com a divergência mais antiga deste grupo, porém não há indícios de vicariância ou dispersão, uma vez que ele ocorre na localidade de Cianorte-PR, juntamente com haplótipos mais antigos. Para o grupo Leste-RS e Litoral, um evento com menor sobreposição com a primeira diversificação deste grupo,

é sugerido para os haplótipos H15 e H16. Os haplótipos H16 e H 17 são mais derivados e estão presentes nas localidades de Pouso Novo-RS e Florianópolis-SC. Disjunção e dispersão entre a região Leste-RS com o Litoral foi bem suportada na análise em RASP (Figura 13). Para os tipos de solo, as análises biogeográficas não indicaram relação com eventos de vicariância e dispersão (Figura 14).

3. HISTÓRIA POPULACIONAL

A dinâmica das populações naturais de *C. hildmannianus* inferidas por DNA cloroplastidial está claramente relacionada com a distribuição geográfica. Embora compartilhem haplótipos, as populações de São Paulo, Paraná e oeste do Rio Grande do Sul (SP/PR + Oeste-RS) apresentam variabilidade genética diferente da encontrada nas populações do leste do Rio Grande do Sul e litorâneas (Leste-RS + Litoral). A diferença se deve principalmente às localidades de Santa Maria (Oeste-RS) e Caçapava do Sul (Leste-RS), no interior do Rio Grande do Sul. Os locais de ocorrência sugerem que as populações desse cacto têm sua distribuição associada ao núcleo de Missiones das FTSS, afloramentos rochosos nos Pampas e em enclaves da Floresta Atlântica.

A estrutura populacional de *C. hildmannianus* inferida para o marcador cloroplastidial concatenado *trnQ-5'rps16-psbJ-petA,* na análise AMOVA, é condizente com a separação geográfica entre os grupos de SP/PR + Oeste-RS e Leste-RS + Litoral, com sobreposição de haplótipos destes grupos nas localidades de Santa Maria e Caçapava do Sul, no interior do Rio Grande do Sul. Dentre os grupos, altosíndices de diversidade nucleotídica populacional foram encontrados para as cidades de Caçapava do Sul e Barra do Ribeiro, no interior do Rio Grande do Sul, as quais pertencem ao grupo Leste-RS (Tabela 9). Para o grupo de SP/PR + Oeste-RS, os índices foram menores, mas se destacaram as cidades de Cianorte e Guarapuava no Paraná e Jaguari no Rio Grande do Sul. De acordo com essa estruturação, a percentagem de variação dentro das populações foi baixa (11,73 a 19,09 %) em relação à variação entre os dois grupos (67,67 a 87,63 %). Para o gene nuclear *PhyC*, a percentagem de variação foi maior dentro dos grupos e não entre eles, sugerindo ausência de estruturação das populações; em relação à diversidade nucleotídica, índices igualmente altos são encontrados nas localidades do estado do Paraná (Cianorte e Mangueirinha), Santa

Catarina (Laguna) e Rio Grande do Sul (Jaguari, Barra do Ribeiro, e as cidades litorâneas de Mostardas e Osório) (Tabela 6).

Um padrão de separação geográfica e estrutura populacional semelhantes aos citados acima foi descrito para as espécies cactófilas de *Drosophila* (Barrios-Leal, 2013; de Brito et al., 2002). Estas espécies de moscas cactófilas têm sua distribuição associadacom as FTSS, pelo fato de utilizarem tecidos de cactos em decomposição como fonte de alimento para o desenvolvimento de suas larvas (Pereira et al., 1983).

De acordo com informações de variabilidade haplotípica de genes nucleares e mitocondriais, populações de *Drosophila meridionalis* estão altamente estruturadas, divididas em dois grupos: o primeiro se refere às localidades do interior dos estados de São Paulo, Paraná e Rio Grande do Sul; e o segundo grupo está nas localidades litorâneas do Rio Grande do Sul, Santa Catarina e São Paulo (Barrios-Leal, 2013). Inferências populacionais com o gene mitocondrial *COI* sugerem que populações da espécie *Drosophila antonietae* estão também estruturadas em dois grupos: São Paulo e Paraná, e a parte sul no interior e litoral do Rio Grande do Sul (de Brito et al., 2002) e litoral de Santa Catarina (Morales et al., 2005). Apesar da não discriminação entre fragmentação passada e fluxo gênico restrito por ausência de análise populacional entre as duas regiões, os autores observaram um compartilhamento de haplótipos nas populações vizinhas dessas áreas amostradas e sugeriram que poderia ter havido um contínuo na distribuição destas espécies (Manfrin e Sene, 2006, de Brito et al. 2002). Essa congruência de padrões filogeográficos entre grupos co-distribuídos pode ser resultante de eventos históricos e geomorfológicos associados ao sul do Brasil.

Um compartilhamento da diversidade genética cloroplastidial foi observado no presente trabalho entre os dois grupos de *C. hildmannianus* e pode ser visualizado pelo padrão filogeográfico obtido na análise NCPA. Os clados de maior nível mostram dois grandes grupos de haplótipos, SP/PR + Oeste e Leste-RS + Litoral. Para o *trnQ-5'rps16*, compartilhamento é visto entre as localidades do Paraná, Oeste e Leste do Rio Grande do Sul entre os clados 1.2 e 1.3, para o *psbJ-petA*, compartilhamento de haplótipos ocorre nas localidades do leste do Rio Grande do Sul, entre os clados 2.1 e 2.2, para o concatenado *trnQ-5'rps16-psbJ-petA* o compartilhamento de haplótipos é observado nas localidades do interior do Rio Grande do Sul, com o clado 2.2 conectando os grupos (Figuras 8, 9 e 10).

Para as regiões de cpDNA, as populações amostradas com maiores índices de diversidade genética foram Caçapava do Sul e Barra do Ribeiro, no Rio Grande do Sul

(Tabela 6). Para o DNA nuclear, os maiores índices foram encontrados nas populações de Cianorte e Mangueirinha no Paraná, Jaguari, Barra do Ribeiro, Osório e Mostardas no Rio Grande do Sul e Laguna em Santa Catarina (Tabela 6). Dentre os dois grandes grupos, a distribuição das populações com menores valores de diversidade nucleotídica, estão localizadas no grupo de SP/PR + Oeste-RS (Tabela 6). Os maiores índices de diversidade genética podem ser evidência de populações mais antigas. Embora as áreas de SP/PR + Oeste-RS apresentem baixos índices de diversidade genética, quando comparados ao Leste-RS, ele está no núcleo de Missiones das FTSS e sua distribuição coincide com outros táxons destas florestas, como por exemplo, em insetos (Barrios-Leal, 2013; Franco e Manfrin, 2013; de Brito et al., 2002) e plantas (Särkinen et al., 2011; Pennington et al., 2009; Caetano et al., 2008, Prado e Gibbs, 1993). O que sugere que apesar destas populações apresentarem menores índices de diversidade, a distribuição delas pode ser antiga assim como as populações do leste do Rio Grande do Sul.

Desta forma, a combinação da distância genética com a localização geográfica nos dois grupos em *C. hildmannianus* resulta em um padrão filogeográfico de baixa divergência genética com linhagens (grupos) parcialmente separadas. Este padrão pode ser consequência de populações historicamente ligadas, e se aplica a linhagens com ampla distribuição geográfica (populações ancestrais) e algumas linhagens restritas a certas localidades (derivadas) apresentando atualmente um fluxo gênico limitado ou ausente (Carvalho e Almeida, 2010). A baixa divergência genética em relação à geografia também foi encontrada pela análise de Mantel. Apesar dos valores significativos deste teste, a baixa correlação positiva encontrada (0,1033 a 0,2570) pode ser um indicativo da ampla distribuição geográfica das populações ancestrais em relação às populações derivadas. Além disso, os marcadores cloroplastidiais em *C. hildmannianus* mostram fluxo gênico limitado geograficamente: no primeiro grupo há migrantes de SP/PR para o Oeste-RS e Leste-RS; no segundo grupo há migrantes entre o Leste-RS, Oeste-RS e Litoral (Tabela 10). Entretanto para o marcador nuclear, a delimitação entre os grupos ocorre quanto à direção da migração, pois hámigrantes de SP/PR para o Litoral (Tabela 10).

Um trabalho com mamífero, o morcego *Desmodus rotundus* na América do Sul, apresentou semelhante discordância entre os padrões filogeográficos obtidos de diferentes marcadores. Para o marcador mitocondrial, cuja herança e ploidia são as mesmas da molécula de cloroplasto, o padrão filogeográfico mostra que as populações

estão separadas há mais tempo que quando analisadas com o marcador nuclear (Martins et al., 2009). A análise filogeográfica realizada com variabilidade haplotípicado DNA mitocondrial exibe clados com alta divergência condizente com alopatria entre Floresta Atlântica e Amazônica, definida pela diagonal seca de vegetação aberta. Além da existência de uma linhagem basal, ligando Floresta Atlântica Sul e Floresta Atlântica Norte, sugerindo polimorfismo ancestral com expansão demográfica associado ao refúgio de São Paulo (Pleistoceno), ou maior estruturação genética não discriminada na análise, com zona de contato entre os clados divergentes. Embora o marcador nuclear não exiba um cenário histórico compatível com os refúgios, no Pleistoceno, por ser encontrado mais distribuído geograficamente.

Padrões múltiplos com sobreposição entre o processo recorrente de fluxo gênicocom o evento histórico de expansão populacional foram sugeridos em *C. hildmannianus* pela análise NCPA (Tabela8). De acordo com o marcador cloroplastidial, esses padrões sugeriram diferentes cenários para as populações dentro do grupo Leste-RS + Litoral, e também, para as populações dentro do grupo de SP/PR + Oeste. Isso foi observado, principalmente, devido às diferentes variabilidades observadas para cada marcador o qual gerou diferentes aninhamentosde haplótipos presentes nas populações das áreas intermediárias aos dois grandes grupos, como para as localidades de Jaguari, Caçapava do Sul e Barra do Ribeiro, no Rio Grande do Sul (Figuras 9, 10 e Apêndice 2 - Tabela 1). Embora este resultado seja válido, uma vez que a NCPA não trata esses padrões como mutuamente exclusivos para o mesmo conjunto de dados (Templeton, 2001), a validação da NCPA foi realizada considerando os resultados dos testes de neutralidade, para confirmar as expansões de área, e o teste de correlação da distância geográfica com a distância genética para o fluxo gênico restrito com isolamento por distância, além da taxa de fluxo gênico.

A hipótese de expansão populacional incluindo todos os indivíduos amostrados no presente trabalho, representados pelos clados 1.1 para o gene nuclear *PhyC*, o clado 2.1 para o marcador cloroplastidial *trnQ-5'rps16,* o clado 3.1 para o marcador cloroplastidial *psbJ-petA* e o clado 3.1 para o concatenado *trnQ-5'rps16-psbJ-petA* (Tabela 8) foi corroborada por meio dos testes de neutralidade R_2 e *Mismatch distribution* (Tabela 7 e Figura 5).

As assinaturas de expansão podem ser observadas de diferentes formas na diversidade molecular. Sob expansão demográfica são esperadas genealogias de genes em forma de estrela ou estrutura em pente, como consequência de mutações em ramos

terminais explicando a baixa frequência de alelos raros, bem como o baixo desequilíbrio de ligação entre locos estreitamente ligados, além de um padrão unimodal obtido pelas análises de distribuição das diferenças nucleotídicas entre pares de sequências (Excoffier et al., 2009). Para o gene nuclear *PhyC* e os marcadores cloroplastidiais *atpI-atpH*, *petL-psbE*, e íntron *trnL*, expansão foi detectada no teste R_2 e na análise *Mismatch distribution*, sendo esta última representada por um padrão unimodal, indicando expansão demográfica quando considerado todas as populações amostradas (Figura 5 - A, B, C e D). Por outro lado, as regiões intergênicas cloroplastidiais *trnQ-5'rps16* e *psbJ-petA* exibiram uma curva com tendência bimodal na análise da *Mismatch distribution*, apesar do teste R_2 detectar expansão (Figura 5 - E, F e G). Então, como marcadores concatenados, essas regiões cloroplastidiais foram avaliadas com a análise BSP e mostraram concordância com a expansão populacional demográfica para todos os indivíduos amostrados de *C. hildmannianus* (Figura 6).

Na assinatura de expansão espacial há uma diminuição constante da diversidade genética ao longo do eixo de expansão em habitats lineares. A população ocorre inicialmente em uma área restrita e a distância do espalhamento da população aumenta ao longo do tempo e do espaço. Isto resulta, geralmente, em uma população subdividida considerando que a escolha dos parceiros para o acasalamento está relacionada com a proximidade geográfica (Excoffier e Lischer, 2010). De fato, foi observado ausência de expansão espacial para o gene nuclear *PhyC*, o qual possui baixa subdivisão populacional em *C. hildmannianus*, para os cloroplastidiais, a análise AMOVA não detectou valores significativos de estruturação populacional em *petL-psbE*, já o íntron *trnL* tem um valor de Φ_{ST} igual a 1, mas não há variação dentro das populações (Tabela 9). Embora a curva bimodal gerada pelo gráfico da *Mismatch distribution* indique que as populações estejam em equilíbrio, as curvas com tendência bimodal encontrada para os marcadores *trnQ-5'rps16* e *psbJ-petA* (Figura 5 - E, F e G) podem refletir um padrão de expansão espacial antiga nas populações, quando o número de migrantes entre os demes é baixo (Excoffier et al., 2009).

Para uma melhor compreensão da expansão espacialdas populações do cacto *C. hildmannianus*, por meio dos marcadores cloroplastidiais *trnQ-5'rps16* e *psbJ-petA*, os resultados das análises geradas pelo MIGRATE foram considerados para essas regiões concatenadas (Tabela 10). Apesar dos elevados índices de taxa de migração inferidos com o programa MIGRATE, é interessante notar que há fluxo gênico restrito entre as populações, com ausência de trocas de migrantes entre as populações de São Paulo e

Paraná em relação às populações litorâneas. Essa restrita troca de migrantes reflete na estruturação populacional encontrada em *C. hildmannianus*, e pode explicar também a não significância dos valores dos testes de D de Tajima e F_S de Fu encontradas no presente trabalho (Excoffier et al., 2009). De fato, quando consideramos o clado 1.2 representando as populações do Leste e Litoral e o clado 2.3 com as populações de São Paulo, Paraná e oeste do Rio Grande do Sul para a região cloroplastidial concatenada *trnQ-5'rps16-psbJ-petA* (Figura 10), os testes D de Tajima, F_S de Fu, além do teste R_2 e curva unimodal para a análise *Mismatch distribution* sugerem expansão populacional nestes grupos (Tabela 7 e Figura 5).

Com base no exposto, o provável cenário filogeográfico para *C. hildmannianus* sugere maior extensão de área no passado (passo 13 e 21 - chave de inferência na Tabela 8) como efeito de movimentos de longa distância, talvez com subsequente fragmentação. Em plantas, o movimento de dispersão de longa distância é realístico quando eventos climáticos extremos e múltiplos vetores de dispersão, como o vento, a água, os pássaros, os morcegos e os animais de grande porte são considerados para uma única espécie de planta, ou seja, quando não existe relação específica entre dispersor e a planta (Nathan, 2006). Para Cactacea, diferentes dispersores são considerados, tais como morcegos, lagartos e pássaros, sendo que em *C. hildmannianus*, os principais dispersores são os pássaros (Zappi et al., 2011).

4. INCONGRUÊNCIA ENTRE OS MARCADORES

Neste trabalho, incongruência foi observada entre o gene nuclear *PhyC* com os marcadores cloroplastidiais da região intrônica *trnL*, e regiões não codificantes *petL-psbE, atpI-atpH, trnQ-5'rps16* e *psbJ-petA*. Para o gene nuclear, as populações não mostraram estruturação entre os dois grupos populacionais definidos neste trabalho (Tabela 9), embora, ausência de estruturação também tenha sido observada com as análises das regiões cloroplastidiais *petL-psbE* e íntron *trnL*. Os padrões de fluxo gênico também foram diferentes (Tabela 10), porém a diferença está entre os grupos SP/PR e Litoral, os quais não trocam migrantes quando analisados pelas regiões cloroplastidiais concatenadas *trnQ-5'rps16* e *psbJ-petA*, enquanto considerando o gene nuclear *PhyC*, SP/PR doa migrantes para o Litoral.

Os diferentes padrões definidos de estruturação e filogeografia nas populações de *C. hildmannianus* quando analisados o gene nuclear *PhyC* e as regiões não codificantes de DNA cloroplastidial podem ser decorrentes do modo de herança entre essas moléculas. O modo de herança biparental do DNA nuclear tem um efeito duplo no tamanho efetivo populacional comparado ao DNA cloroplastidial, portanto é esperado que o tempo médio de coalescência dos genes nucleares seja mais antigo (Avise, 2009). As relações filogenéticas com o marcador *PhyC* em *Opuntia* (Cactaceae) tiveram baixo suporte filogenético e não apresentaram associação entre geografia e variabilidade, em contraposição ao marcador cloroplastidial *trnT-trnL*, sugerindo incongruências entre esses marcadores (Helsen et al., 2009).

5. CENÁRIO DA DINÂMICA POPULACIONAL

De acordo com as datações e os dados biogeográficos para os segmentos concatenados *trnQ-5'rps16-psbJ-petA*, nota-se uma distribuição antiga para a maioria das populações (Figura 12). No entanto, uma análise biogeográfica dessas áreas revelou algumas pistas sobre aatual estrutura populacional desta espécie. Eventos de dispersão e vicariância foram bem suportados na região de São Paulo e Paraná em relação às populações do leste do Rio Grande do Sul, a área leste também mostrou eventos de vicariância e dispersão tanto com a parte oeste do Rio Grande do Sul, como também em relação à parte litorânea (Figura 13).

Os eventos de vicariância e expansão ocorreram nos limites da área Leste-RS (Figura 13), cujas populações possuem maiores índices de diversidade genética, sendo suas distribuições coincidentes com duas áreas geomorfológicas do estado do Rio Grande do Sul, a Depressão Central e o Escudo Sul-Rio Grandense. Um mapeamento de bioma com espécies das FTSS (Särkinen et al., 2011) mostraram que além do núcleo de Missiones, na bacia do rio Paraná-Paraguai, uma parte do interior do Rio Grande do Sul, nas proximidades da parte Leste-RS, possui alta ocorrência de espécies endêmicas para das FTSS, e então consideraram essa região como mais um núcleo desta floresta no sul do Brasil. Além disso, estimativas de espécies endêmicas para conservação de Cactaceae no Brasil, sugeriram dois *hotspots* de conservação para a região sul, um em afloramentos rochosos no Pampa e outro na Mata Atlântica do sul do Brasil (Zappi et al., 2011).

A parte Leste-RS está delimitada pela Planície Costeira e, por isso, é de se esperar que a vegetação desta área esteja em contato com táxons da Floresta Atlântica. Concordante a isso, estudos com o gênero *Passiflora*, de Mata Atlântica, mostram que *Passiflora alata* está na Planície Costeira e na Depressão Central, sendo esta última, composta por Floresta Sazonal Decídua e zona de Tensão Ecológica (Koehler-Santos et al., 2006). Os cenários filogeográficos em *Passiflora actinia* e *Passiflora elegans* parecem concordar com os dados biogeográficos descritos para a Floresta Atlântica no sul do Brasil, para os quais vários táxons estão distribuídos ao longo de uma via migratória (Mader et al., 2009). Essa via migratória foi descrita por Rambo (1950) como a "Porta de Torres", estreita planície no RS entre o Planalto Meridional e o Oceano Atlântico, que teria sido importante rota para espécies tropicais do norte e centro do Brasilem direção a parte mais oriental do Rio Grande do Sul.

Além desses aspectos, no litoral dos estados de Santa Catarina e Rio Grande do Sul, uma zona de contato secundário tem sido descrita tanto para espécies de Mata Atlântica, como *Passiflora actinia* e *Passiflora elegans* (Lorenz-Lemke et al., 2005), como para espécies relacionadas à vegetação das FTSS, tais como *Drosophila serido* e *Drosophila antonietae* (Manfrin e Sene, 2006). Em ambos os trabalhos citados acima, os autores sugerem as oscilações climáticas do Quaternário para explicar a dinâmica da vegetação nesta região. Trigo e colaboradores (2008) encontraram uma zona de hibridação interespecífica para mamíferos do gênero *Leopardus* no estado do RS, a qual ocorre no interior do Rio Grande do Sul, tanto à oeste, quanto ao leste, entre *L. trigrinus* e *L. geoffroyi*. Para esses mamíferos, os resultados permitiram associar a zona híbrida como consequência de processo histórico de crescimento populacional, em resposta a mudanças climáticas.

Dado o exposto, um cenário plausível para *C. hildmannianus* é de que a distribuição desta espécie é antiga, considerando tanto os eventos demográficos como as datações moleculares destes, e pode ter sido favorecida tanto pelo aumento de plantas C4 devido às baixas condições de CO_2 no final do Terciário, as quais propiciam um ambiente favorável para as plantas suculentas, como de condições climáticas favoráveis de clima frio e seco nos períodos glaciais. Atualmente, há pequenas populações isoladas entre as duas maiores áreas de ocorrência, exibindo um limitado fluxo gênico, o que resulta em grupos altamente estruturados, com a distribuição de haplótipos mais recentes restritos às bordas destes grupos. Além disso, o aumento da umidade e da temperatura nos períodos interglaciais podem ter interferido na dinâmica entre expansão

e retração dessas florestas, na qual as Florestas Tropicais expandiram em áreas abertas, isolando a vegetação de áreas abertas e secas no sul do Brasil (Behling, 2002).

VI. CONCLUSÕES

Os resultados obtidos no presente trabalho nos permitem concluir que:

- Os múltiplos eventos na história de *C. hildmannianus* incluíram expansão de área e vicariância em diferentes escalas temporais e proporcionaram fluxo gênico restrito entre as populações, gerando estruturação populacional em dois principais grupos: o grupo do interior de SP, PR e oeste do RS e o grupo da parte leste do RS e Planície Costeira;

- A incongruência entre os marcadores nuclear e cloroplastidiais na história evolutiva de *C. hildmannianus* sugeriram um polimorfismo ancestral entre os dois grupos estruturados, sendo que a diferença estaria relacionada ao tempo de coalescência e ao tamanho efetivo populacional desses marcadores;

- O tempo do ancestral comum mais recente em *C. hildmannianus* inferido ao final do Terciário corrobora com a especiação estimada para outros táxons, que assim como este cacto, estão relacionados às Florestas Tropicais Sazonalmente Secas, na região neotropical.

VII. REFERÊNCIAS BIBLIOGRÁFICAS

Ab'Saber, A. 1977. Espaços ocupados pela expansão dos climas secos na América do Sul, por ocasião dos períodos glaciais quaternários. *Paleoclimas* 3:1-19.

Ab´Saber, A N. 2003. Os Domínios de Natureza no Brasil: potencialidades paisagísticas. São Paulo. Atelie Editorial. 159 pp.

Allen G A, Marr K L, McCormick L J, Hebda J R. 2012. The impact of Pleistocene climate change on an ancient arctic–alpine plant: multiple lineages of disparate history in *Oxyria digyna*. Ecol Evol: 2(3):649-65.

Altschul S F, Gish W, Miller W, Myers E W, Lipman D J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-10.

Arakaki M, Christin P, Nyffeler R, Lendel A, Eggli U, Ogburn R M, Spriggs E, Moore M J, Edwards E. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *PNAS.* 108(20): 8379-8384.

Anderson E F. 2001. The Cactus Family. Timber Press Inc., Portland, Oregon.

Auler A S, Wang X, Edwards R L, Cheng H, Cristalli P S, Smart P L, Richards D A. 2004. Palaeoenvironments in semi-arid northeastern Brazil inferred from high precision mass spectrometric speleothem and travertine ages and the dynamics of South American rainforests. *Speleog. Evol. Karst Aquif.* 2: 1-4.

Avise JC. 2000. Phylogeography: history and formation of species. Harvard University Press. 447p.

Avise JC. 2009. Phylogeography: retrospect and prospect. *Journal of Biogeography,* 36: 3-15.

Avise JC, Arnold J, Ball R, Bermingham E, Lamb T, Neigel JE, Reeb CA e Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18: 489–522.

Bai W-N, Liao W-J, Zhang D-Y. 2010. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in atemperate walnut tree from East Asia. New Phytologist. 188: 892–901.

Barrios-Leal D Y. 2013. História demográfica e estrutura de populações para a espécie cactófila Drosophila meridionalis. Tese de Mestrado, apresentada à Faculdade de Medicina de Ribeirão/USP. Área de Concentração: Genética.

Barthlott W. 1983. Biogeography and evolution in neo and palaeotropical Rhipsalinae. *In* K. Kubitzki [ed.], Dispersal and distribution, 241–248. Verlag Paul Parey, Hamburg, Germany.

Barthlott W, Hunt D R. 1993. Cactaceae. In K. Kubitzki, J. G. Rohwer, and V. Bittrich [eds.], The families and genera of vascular plants. *Springer Verlag, Berlin, Germany.* 2, 161–197.

Barthlott W, Taylor N P. 1995. Notes towards a monograph of Rhipsalideae (Cactaceae). Bradleya 13: 43–79.

Bauer D, Waechter J L. 2006. Sinopse taxonômica de Cactaceae epifíticas no Rio Grande do Sul, Brasil. *Acta bot. bras*. 20(1): 225-239.

Beaumont M A, Rannala B. 2004. The Bayesian revolution in genetics. Nature Reviews. Genetics, v. 5, n. 4, p. 251-61.

Beerli P, Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics.* 152 (2):763.

Behling H. 1998. Late Quaternary vegetational and climatic changes in Brazil. *Review of Palaeobotany and Palynology.* 99: 143-156.

Behling H. 2002. South and southeast Brazilian grassland during Late Quaternary times: *a synthesis. Palaeo* 177: 19-27.

Behling H, Hooghiemstra H. 2000. Holocene Amazon rainforest – savanna dynamics and climatic implications: high-resolution pollen record from Laguna Loma Linda in eastern Colombia. *J. Quat. Sci*. 15: 687-695.

Behling H, Pillar V D, Orlóci L, Bauermann S G. 2004. Late Quaternary Araucaria forest, grassland (Campos), fire and climate dynamics, studied by high-resolution pollen, charcoal and multivariate analysis of the Cambará do Sul core in southern Brazil. *Palaeogeography, Palaeoclimatology, Palaeoecology.* 203: 277-297.

Bonatelli I A S. 2010. Análise da estrutura filogeográfica do grupo *Pilosocereusaurisetus* (Cactaceae) utilizando marcadores moleculares do genoma do cloroplasto (cpDNA). *Dissertação (Mestrado em Genética e Evolução) - Universidade Federal de São Carlos, Fundação de Amparo à Pesquisa do Estado de São Paulo.*

Borsch T, Quandt D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolutio.* 282: 169-199.

Brito A R, Manfrin M H, Sene F M. 2002. Nested cladistic analysis of brasilian populations of *Drosophila serido*. *Mol. Phylog. Evol.* 22: 131-143.

Bruxel J, Jasper A. 2005. A família Cactaceae na Bacia Hidrográfica do Rio Taquari, RS, Brasil. *Acta bot*. 19 (1): 71-79.

Butterworth C, Edwards E J. 2008. Investigating *Pereskia* and the earliest divergences in Cactaceae. *Haseltonia* 14: 46–53.

Caetano S, Prado D, Pennington R, Beck S, Oliveira-Filho A, Spichiger R, Naciri Y. 2008. The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Molecular Ecology* 17 (13):3147-3159.

Carnaval A, Bates J. 2007. Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in northeastern Brazil. *Evolution.* 61 (12): 2942-2957.

Carnaval A C, Hickerson M J, Haddad C F B. 2009. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. Science, v. 323, n. 6, p. 785-789.

Carnaval A, Moritz C. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. Journal of Biogeography 35 (7):1187-1201.

Carvalho C J B, Almeida E A B. 2010. Biogeografia da América do Sul: padrões e processos. Editora Roca LTDA.

Castelloe J, Templeton A R. 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. *Mol. Phylog. Evol.* 3:102–113.

Castro J P. 2008. Números cromossômicos em espécies de Cactaceae ocorrentes no nordeste do Brasil. Dissertação (Mestrado em Agronomia) - Universidade Federal da Paraíba-Centro de Ciências Agrárias, Areia.

CBOL Plant Working Group. 2009. A DNA barcode for land plants. Proceedings of the National Academy of Sciences of the United States of America, 106, 12794–12797.

Chaves J A, Weir J T, Smith T B. 2011. Diversification in Adelomyia hummingbirds follows Andean uplift. *Molecular Ecology.* 20, 4564-4576.

Chase M W, Cowan R S, Hollingsworth P M, Berg C, Madriñán S, Petersen G, Seberg O, Jørgsensen T, Cameron K, Carine M, Pedersen N, Hedderson T A J, Conrad F, Salazar G A, Richardson J E, Hollingsworth M L, Barraclough T G, Kelly L,

Wilkinson M. 2007. A proposal for a standardised protocol to barcode all land plants. TAXON 56 (2): 295–299.

Clapperton C. 1993. Quaternary geology and geomorphology of South America. Vol. 779: Amsterdam Elsevier.

Clement M, Posada D, Crandall K A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology.* 9 (10): 1657-1660.

Collevatti R, Rabelo S, Vieira R. 2009. Phylogeography and disjunct distribution in *Lychnophora ericoides* (Asteraceae), an endangered cerrado shrub species. *Annals of Botany.* 104: 655-664.

Colinvaux P A, Irion G, Räsänen M E, Bush M B, Nunes de Mello J A S. 2001. A paradigm to be discarded: geological and paleontological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation. *Amazoniana.* 16: 609-646.

Cordellier M e Pfenninger M. 2009. Inferring the past to predict the future: climate modelling predictions and phylogeography for the freshwater gastropod Radix balthica (Pulmonata, Basommatophora). *Molecular Ecology*. 18, 534-544.

Crandall K A, Templeton A R. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics.* 134: 959-969.

Demaio P H, Barfuss M H J, Kiesling R, Till W, Chiapella J O. 2011. Molecular phylogeny of *Gymnocalycium* (Cactaceae): Assessment of alternative infrageneric systems, a new Subgenus, and trends in the evolution of the genus. *American Journal of Botany.* 98(11): 1841-1854.

De Marco Júnior, Siqueira M F. 2009. Como determinar a distribuição potencial de espécies sob uma abordagem conservacionista? Megadiversidade, 5: 1-2.

De Oliveira P E, Barreto F A M, Suguio K. 1999. Late Pleistocene/Holocene climatic and vegetational history of the Brazilian caatinga: the fossil dunes of the middle São Francisco River. Palaeogeography, Palaeoclimatology, Palaeoecology, v. 152, p. 319-337.

Drummond A J, Rambaut A. 2007. "BEAST: Bayesian evolutionary analysis by sampling trees." *BMC Evolutionary Biology.***7**, 214.

Drummond A J, Phillips S Y W Ho M J e Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 4(5):699-710.

Drummond A J, Rambaut A, Shapiro B, Pybus O G. 2005. Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Mol. Biol. Evol* . 22(5):1185–1192.

Ebert D, Peakall R. 2009. Invited Technical Review. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Molecular Ecology Resources:* 9: 673–690.

Edward E J, Nyffeler R e Donogue M J. 2005. Basal cactus phylogeny: implications of *Pereskia* (cactaceae) paraphyly for the transition to the cactus life form. *American Journal of Botany.* 92(7): 1177-1188.

Excoffier L, Smouse P E, Quattro J M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 131:479-491.

Excoffier L, Foll M, Petit R. 2009. Genetic Consequences of Range Expansion. Annu. *Review of Ecology, Evolution, and Systematics* 40:481–501.

Excoffier L, Lischer H E L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564-567.

Fehlberg S D, Ranker T A. 2009. Evolutionary history and phylogeography of *Encelia farinosa* (Asteraceae) from the Sonoran, Mojave, and Peninsular Deserts. *Molecular Phylogenetics and Evolution.* 50(2): 326-35.

Fehrmann S, Philbrick C T, Halliburton A R. 2012. Intraspecific Variation in *Podostemum ceratophyllum* (Podostemaceae): Evidence of Refugia And Colonization Since The Last Glacial Maximum. *American Journal of Botany.* 99(1): 145–151.

Franco F F, Manfrin M H. 2013. Recent demographic history of cactophilic *Drosophila* species can be related to Quaternary palaeoclimatic changes in South America. Journal of Biogeography, p. 1-13.

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

Fu Y, Li W. 1999. Coalescencing into the 21st Century: An Overview and Prospects of Coalescent Theory. *Theory in Population Biology* 56:1-10.

Garrick RC, Dyer RJ, Beheregaray LB e Sunnucks P. 2008. Babies and bathwater: a comment on the premature obituary for nested clade phylogeographic analysis. *Mol. Ecol.* 17: 1401–1403.

Geleta M, Bryngelsson T. 2012. Population Genetic Analysis of Lobelia rhynchopetalum Hemsl. (Campanulaceae) Using DNA Sequences fromITS and Eight Chloroplast DNA Regions. *The ScientificWorld Journal*. 10p.

Gonçalves C N, Waechter J L. 2003. Aspectos florísticos e ecológicos de epífitos vasculares sobre figueiras isoladas no norte da planície costeira no Rio Grande do Sul. *Acta bot. Brás*. 17(1): 89-100.

Grivet D, Heinze B, Vendramin G G, Petit R J. 2001. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Resource.* 1: 345-349.

Haffer J. 1969. Speciation in Amazonian Forest Birds. Science, v. 165, n. 3889, p. 131-136.

Hamilton M B. 1999. Four primers pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology.* 8: 513-525.

Harpending H C. 1994. Signature of Ancient Population Growth in a Low-Resolution Mitochondrial DNA Mismatch Distribution. *Human Biology.* 66(4): 591-600.

Helsen P, Verdyck P, Tye A, Desender K, Van Houtte N, Van Dongen S. 2007. Isolation and characterization of polymorphic microsatellite markers in Galapagos prickly pear (Opuntia) cactus species. *Molecular Ecology Notes.* 7: 454-456.

Helsen P, Robert A. Browne R A, David J. Anderson D J, Verdyck P, Van Dongen S. 2009. Galápagos' *Opuntia* (prickly pear) cacti: extensive morphological diversity, low genetic variability. *Biological Journal of the Linnean Society*, 96, 451–461.

Hershkovitz M A, Zimmer E A. 1997. On the evolutionary origins of the cacti. *Taxon.* 46(2): 217-232.

Hewitt G M. 2011. Quaternary phylogeography: the roots of hybrid zones. Genetica, v. 139, n. 5, p. 617-38.

Hijmans R J, Guarino L, Bussink C, Barrantes I, Rojas E. 2005. DIVA-GIS, Version 5.2. http://www.diva-gis.org (ultimo acesso em 30 de Maio de 2011).

Hon C, Lam T, Shi Z, Drummond A J, Yip C, Zeng F, Lam P, Leung F C. 2008. Evidence of the Recombinant Origin of a Bat Severe Acute Respiratory

Syndrome (SARS)-Like Coronavirus and Its Implications on the Direct Ancestor of SARS Coronavirus. *Journal of Virology.* 82(4): 1819–1826.

Hoorn C, Wesselingh F P, Steege H T, Bermudez M A, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson C L, Figueiredo J P, Jaramillo C, Riff D, Negri F R, Hooghiemstra H, Lundberg J, Stadler T, Sarkinen T, Antonelli A. 2011. Origins of Biodiversity–Response. *Science.* 331: 399-400.

Hueck K. 1972. As Florestas da América do Sul. Trad. Hans Reichardt. Editora Universidade de Brasília/Editora Polígono. São Paulo, Brazil.

Hughes C E, Pennington R T, Antonelli A. 2013. Neotropical Plant Evolution: Assembling the Big Picture. *Botanical Journal of the Linnean Society*, 171, 1–18.

Hunt D, Taylor N, Charles G. 2006. The New Cactus Lexicon, Volumes I and II. DH Books, Milborne Port.

Kingman J F C. 1982. The coalescent. Stochastic Processes and their Applications, 13(3): 235-248.

Kress W J, Wurdack K J, Zimmer E A, Weigt L A, Janzen D H. 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. U.S.A. 102: 8369–8374.

Koehler-Santos P, Lorenz-Lemke A P, Salzano F M. e Freitas, L.B. 2006. Ecological-evolutionary relationships in *Passiflora alata* from Rio Grande do Sul, Brasil. *Braz. J. Biol.* 66(3): 809-816.

Kozak K H, Graham C H, Wiens J J. 2008. Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution.* 23: 141–148.

Knowles L, Maddison W. 2002. Statistical phylogeography. Molecular Ecology, p. 2623-2635.

Lorenz-Lemke A P, Muschner V C, Bonatto S L, Cervi A C, Salzano F M, Freitas L B. 2005. Phylogeographic Inferences Concerning Evolution of Brazilian Passiflora actínia and P. elegans (Passifloraceae) Based on ITS (nrDNA) Variation. Annals of Botany 95: 799–806.

Ledru, M.-P. 1993. Late Quaternary Environmental and Climatic Changes in Central Brazil. Quaternary Research, 39, 90-98.

Ledru M P, Salgado-Labouriau M L, Lorscheitter M L. 1998. Vegetation dynamics in southern and central Brazil during the last 10,000 yr B.P. *Rev. Palaeobot. Palyn.* 99: 131-142.

Lei M, Wang Q, Wu Z-J, Pujol J L, Li D-Z, Zhang Z-Y. 2012. Molecular phylogeography of *Fagus engleriana* (Fagaceae) in subtropical China: limited admixture among multiple refugia. *Tree Genetics & Genomes.* 8(6): 1203-1212.

Librado P, Rozas J. 2009. DnaSPv.5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 25: 1451-1452.

Lihov J, Kudoh H, Marhold M. 2010. Genetic Structure And Phylogeography of a Temperate-Boreal Herb, *Cardamine scutata* (Brassicaceae), In Northeast Asia Inferred From Aflps And Cpdna Haplotypes. *American Journal of Botany.* 97(6): 1058–1070.

Linares-Palomino R, Oliveira-Filho A T, Pennington R T. 2010. Neotropical seasonally dry forests: diversity, endemism and biogeography of woody plants. In *Seasonally Dry Tropical Forests: Biology and Conservation*, ed. R Dirzo, H Mooney, G Ceballos, H Young.Washington, DC: Island Press.

Linares-Palomino R, Oliveira-Filho A T, Pennington R T. 2011. Neotropical seasonally dry forests: Diversity, endemism, and biogeography of woody plants. In Seasonally Dry Tropical Forests: Ecology and Conservation. Edited by: Dirzo R, Mooney H, Ceballos G, Young H. Island Press; 3-21.

Mader G, Lorenz-Lemke A P, Cervi A C, Freitas L B. 2009. Novas ocorrências e distribuição do gênero *Passiflora* L. no Rio Grande do Sul, Brasil. *Braz. J. Biol.* 7(4): 364-367.

Manfrin M H, Sene F M. 2006. Cactophilic *Drosophila* in South America: a model for evolutionary studies. *Genetica.* 126: 57-75.

Manel S, Williams H C, Ormerod S J. 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology.* 38: 921–931.

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer research* 27 (2 Part 1):209.

Martins F M, Templeton A R, Pavanl A C, Kohlbach B C, Morgante J S. 2009. Phylogeography of the common vampire bat (Desmodus rotundus): Marked population structure, Neotropical Pleistocene vicariance and incongruence between nuclear and mtDNA markers. *BMC Evolutionary Biology.* 9: 294-307.

Moraes E M, Yotoko K, Manfrin M, Solferini V, Sene F. 2009. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence

timing in dry vegetation enclaves in eastern Brazil. *Journal of Biogeography* 36 (11):2136-2147.

Morales A C. 2005. Análise histórico-evolutiva de populações das espécies cactofílicas *Drosophila serido* e *Drosophila antonietae* (Díptera, Drosophilidae). Tese de Doutorado. Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

Murray-Smith C, Brummitt N A, Oliveira-Filho A T, Bachman S, Moat J, Nic Lughadha E M, Lucas E J. 2008. Plant Diversity Hotspots in the Atlantic Coastal Forests of Brasil. *Conservation Biology.* 23(1): 151-163.

Müller K. 2005. SeqState - primer design and sequence statistics for phylogenetic DNA data sets. Applied Bioinformatics. 4:65-69.

Müller K, Borsch T. 2005. Phylogenetics of Amaranthaceae based on *matK/trnK* sequence data – Evidence from parsimony, likelihood, and Bayesian analyses. Ann. Missouri Bot. *Gard*. 92: 66-102.

Nathan R. 2006. Long-Distance Dispersal of Plants. Science, v 313.

Nielsen R, Beaumont M A. 2009. Statistical inferences in phylogeography. Molecular Ecology, 18(6) 1034-1047.

Nyffeler R. 2002. Phylogenetic Relationships In The Cactus Family (Cactaceae) Based On Evidence From Trnk/ Matk And TrnL-TrnF Sequences. *American Journal Of Botany.* 89(2): 312–326.

Ortega-Baes P, Godínez-Alvarez H. 2006. Global Diversity and Conservation Priorities in the Cactaceae. Biodiversity and Conservation, 15, 817–827.

Panchal M, Beaumont M A. 2007. The automation and evaluation of nested clade phylogeographic analysis. Evolution; International Journal of Organic Evolution, 61(6) 1466-80.

Pennington R T, Prado D E, Pendry C A. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. J. *Biogeogr*. 27: 261-273.

Pennington R T, Lavin M, Oliveira-Filho A. 2009. Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests. Annu. Rev. Ecol. Evol. Syst. 40:437–57.

Pennington R T, Lavin M, Prado D E, Pendry C A, Pell S K, Butterworth C A. 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions Royal Society Lond*. 359: 515-538.

Pereira M A Q R, Vilela C R, Sene F M. 1983. Notes on breeding and feeding sites of some species of the repleta group of the genus *Drosophila* (Diptera, Drosophilidae). Ciencia e Cultura, 35, 1313-1319.

Pessenda L C R, De Oliveira P E, Mofatto M, Medeiros V B, Garcia R J F, Aravena R, Bendassoli J A, Leite A Z, Saad A R, Etchebehere M L. 2009. The evolution of a tropical rainforest/grassland mosaic in southeastern Brazil since 28,000 14C yr BP based on carbon isotopes and pollen records. *Quaternary Research.* 71: 437–452.

Phillips S, Anderson R, Schapire R. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling.* 190 (3-4):231-259.

Posada D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*. 25(7): 1253–1256.

Posada D, Crandall K A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 14: 817-818.

Posada D, Crandal K A, Templeton A R. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology.* 9(4): 487-488.

Prado D E. 2000. Seasonally Dry Forests of Tropical South America: From Forgotten Ecosystems to a New Phytogeographic Unit. *Edinburgh Journal of Botany*. 57(3): 437–461.

Prado D E, Gibbs P E. 1993. Patterns of species distributions in the dry seasonal forests of South America. Ann. Miss. *Bot. Gard*. 80: 902-927.

Provan J, Biss P M, McMeel D, Mathews S. 2004. Universal primers for the amplification of chloroplast microsatellites in grasses (Poaceae). *Molecular Ecology Notes.* 4: 262-264.

Quijada-Mascareñas J, Ferguson J, Pook C, Salomão M, Thorpe R, Wüster W. 2007. Phylogeographic patterns of trans-Amazonian vicariants and Amazonian biogeography: the Neotropical rattlesnake (Crotalus durissus complex) as an example. *Journal of Biogeography* 34 (8):1296-1312.

Rambo B. 1950. A porta de Torres. *Anais Botânicos do Herbário Barbosa Rodrigues,* 2: 9-20.

Ramos-Onsins S E, Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol*. 19 (12): 2092–2100.

Ravi V, Khurana J, Tyagi A, Khurana P. 2008. An update on chloroplast genomes. *Plant Systematics and Evolution.* 271: 101-122.

Ritz C M, Martins L, Mecklenburg R, Goremykin V, Hellwig F. 2007. The molecular phylogeny of *Rebutia* (cactaceae) and its allies demonstrates the influence of paleogeography on the evolution of South American mountain cacti. *American journal of Botany*. Rep. 94(8): 1321–1332.

Rogers A R. 1995. Genetic evidence for a Pleistocene population explosion. *Evolution*. 49(4): 608-615.

Rogers A, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution.* 9(3): 552-569.

Rull V. 2008. Speciation timing and neotropical biodiversity: the Tertiary–Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology.* 17: 2722–2729.

Rull V. 2011. Origins of Biodiversity. *Science* 331: 399-400.

Russell R V, SoutoA C P, PremoliA A C. 2011. Multiple Pleistocene refugia in the widespread Patagonian tree *Embothrium coccineum* (Proteaceae). *Australian Journal of Botany.* 59: 299–314.

Särkinen T, Iganci J R V, Linares-Palomino R, Simon M F, Prado D E. 2011. Forgotten forests - issues and prospects in biome mapping using Seasonally Dry Tropical Forests as a case study. *BMC Ecology*. 11:27.

Schaal B, Da Hayworth K, Rauscher J, Smith W. 1998. Phylogeographic studies in plants: problems and prospects. *Molecular Ecology,* 7: 465-474.

Schneider S, Excoffier L. 1999. Estimation of Past Demographic Parameters From the Distribution of Pairwise Differences When the Mutation Rates Vary Among Sites: Application to Human Mitochondrial DNA. *Genetics*. 152:1079–1089.

Sebastiani F, Carnevale S, Vendramin G G. 2004. A new set of mono- and dinucleotide chloroplast microsatellites in Fagaceae. *Molecular Ecology Notes.* 4: 259-261.

Shaw J, Lickey E, Beck J, Farmer S, Liu W, Miller J, Siripun K, Winder C, Schilling E, Small R. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany.* 92: 142-166.

Shaw J, Lickey E, Schilling E, Small R. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in

angiosperms: the tortoise and the hare III. *American Journal of Botany.* 94: 275-288.

Simmons M P, Ochoterena H. 2000. Gaps as characters in sequence based phylogenetic analysis. *Syst Biol.* 49: 369–381.

Skov F, Svenning J C. 2004. Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography.* 27: 366– 380.

Smouse P, Long J, Sokal R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Biology* 35 (4): 627.

Starr J R, Naczi F C, Chouinard B N. 2009. Plant DNA barcodes and species resolution in sedges (*Carex*, Cyperaceae). *Molecular Ecology Resources.* 9(1): 151–163.

Swets J A. 1988. Measuring the accuracy of diagnostic systems. *Science.* 240: 1285–1293.

Taberlet P, Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann & Willerslev, E. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acid. Res. 35: e14.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal *primers* for amplification of three non-oding regions of chloroplast DNA. *Plant Molecular Biology.* 17: 1105-1109.

Tajima F. 1989. Statistical Method for testing the neutral mutation hypothesis by DNA polimorphism. *Genetics*. 123: 584-595.

Taylor N P. 2000. Taxonomy and phytogeography of the Cactaceae of eastern Brazil. Tese de Doutorado. The Open University, England.

Taylor N, Zappi D. 2004. Cacti of Eastern Brasil. *Royal Botanic Gardens*, Kew. 511p.

Templeton A R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*. 7: 381–397.

Templeton A R. 2001. Using phylogeographic analyses of gene trees to test species status and processes. *Molecular Ecology.* 10: 779–791.

Templeton A R. 2003. A maximum likelihood framework for cross validation of phylogeographic hypotheses. In: Wasser SP, ed. Evolutionary theory and processes: modern horizons. *Boston, MA: Kluwer*. 209–230.

Templeton A R. 2004a. Using haplotype trees for phylogeographic and species in fish population. *Environmental Biology of Fishes.* 69:7–20.

Templeton A R. 2004b. Statistical phylogeography: methods of evaluating and minimizing inference errors. Molecular Ecology, 13(4): 789-809.

Templeton A R. 2006. Population Genetics and Microevolutionary Theory. St. Louis Missouri: John Wiley &Sons., Inc., Wiley-Liss, p. 716.

Templeton A R. 2008. Nested clade analysis: an extensively validated method for strong phylogeographic inference. Molecular Ecology, 17, p. 1877-1880.

Templeton A R, Boerwinkle, Sing C F. 1987. A Cladistic Analysis of Phenotypic Associations With Haplotypes Inferred From Restriction Endonuclease Mapping. I. Basic Theory and an Analysis of Alcohol Dehydrogenase Activity in *Drosophila*. Genetics Society of America.

Templeton A R, Routman E, Phillips C .1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics*. 140: 767–782.

Templeton A R, Sing C F. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics.* 134: 659–669.

Thomé M, Zamudio K, Giovanelli J, Haddad C, Baldissera Jr F, Alexandrino J. 2010. Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution* 55 (3):1018-1031.

Thomas C D, Cameron A, Green R E *et al*. 2004. Extinction risk from climate change. *Nature*. 427: 145–148.

Thompson J D, Higgins D G, Gibson T J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. 22: 4673-4680.

Thuiller W, Lavorel S, Sykes M T, Araujo M B. 2006. Using nichebased modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distribution.*, 12: 49–60.

Trigo T C, Freitas T R O, Kunzler G, Cardoso L, Silva C R, Johnson W E, O'Brien J O, Bonatto L, Eizirik E. 2008. Inter-species hybridization among Neotropical cats of the genus *Leopardus*, and evidence for na introgressive hybrid zone between *L. geoffroyi* and *L. tigrinus* in southern Brasil. *Molecular Ecology*. 17: 4317-4333.

Tripati A K, Roberts C D, Eagle R A. 2009. Coupling of CO_2 and ice sheet stability over major climate transitions of the last 20 million years. Science 326:1394–1397.

Wallace R S. 1995. A family-wide phylogeny, subfamilial and tribal relationships, and suggestions for taxonomic realignments. *IOS Bulletin.* 6(1): 13.

Waltari E, et al. 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoS ONE, San Francisco, 2:7 e563.

Waltari E, Guralnick R P. 2009. Ecological niche mode-ling of montane mammals in the Great Basin, North America: examining past and present connectivity of species across basins and ranges. Journal of Biogeography, Oxford, 36: 148-161.

Wakasugi T, Sugita M, Tsudzuki T, Sugiura M. 1998. Updated Gene Map of Tobacco Chloroplast DNA. *Plant Molecular Biology Reporter* 16**:** 231–241.

Wang X, Auler A S, Edwards R L, Cheng H, Cristalli P S, Smart P L, Richards D A, Shen C C. 2004. Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. *Nature.* 432: 740-743.

Weber S L, Drijfhout S S, Abe-Ouchi A, Crucifix M, Eby M, Ganopolski, A, Murakami S, Otto-Bliesner B, Peltier W R. 2007. The modern and glacial overturning circulation in the Atlantic ocean in PMIP coupled model simulations. *Climate of the Past.* 3: 51–64.

Werneck F P, Costa G C, Colli G R, Prado D E, Sites Jr J W. 2011. Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. Global Ecology and Biogeography 20, 272–288.

Wolfe K H, Li W, Sharp P M. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad.Sci. USA*. 84: 9054-9058.

Yesson C, Barcenas R T B, Hernadndez H C M, Ruiz- Maqueda M L, Rodriguez V M, Prado A, Hawkins A J. 2011. DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting. *Molecular Ecology Resources.* 11: 775–783.

Yu Y, Harris A J, He X-J. 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution.* 56: 848–850.

Zappi D, et al. 2011. Plano de Ação Nacional Para a Conservação das Cactáceas. Brasília: Instituto Chico Mendes de Conservação da Biodiversidade, ICMBIO. 112p.

Zhang Q, Zhao Y, Gong X. 2011. Genetic variation and phylogeography of *Psammosilene tunicoides* (Caryophyllaceae), a narrowly distributed and endemic species in south-western China. *Australian Journal of Botany*, 59, 450–459.

Zimmer E A, Wen J. 2012. Using nuclear gene data for plant phylogenetics: Progress and prospects. Molecular Phylogenetics and Evolution 65: 774–785.

APÊNDICE 1

Tabela 1. Coordenadas geográficas em decimais obtidas de infomações de banco de dados, de herbáros e de coleta do presente trabalho utilizadas para modelagem de nicho ecológico e paleomodelagem.

Espécie - Longitude, Latitude	Local	Dados*
Cereus hildmannianus, -51.023333,-30.081111	Viamão-RS	ICN - UFRS
Cereus hildmannianus,-55.131111,-29.550278	São Francisco de Assis-RS	ICN - UFRS
Cereus hildmannianus,-51.325,-30.113889	Guaiba-RS	ICN - UFRS
Cereus hildmannianus,-54.071667,-29.3525	Quevedos-RS	PACA - RS
Cereus hildmannianus,-53.491389,-30.512222	Caçapava do Sul-RS	PACA - RS
Cereus hildmannianus,-54.178889,-29.620556	São Pedro do Sul-RS	PACA - RS
Cereus hildmannianus,-54.071667,-29.3525	Quevedos-RS	PACA - RS
Cereus hildmannianus,-51.147222,-29.760278	São Leopoldo-RS	PACA - RS
Cereus hildmannianus,-55.611944,-28.576389	Santo Antônio das Missões-RS	PACA - RS
Cereus hildmannianus,-51.555278,-27.451389	Zortéa-SC	PACA - RS
Cereus hildmannianus,-49.4636,-24.1128	Sengés-PR	MBM
Cereus hildmannianus,-52.3831,-24.0456	Campo Mourão-PR	MBM
Cereus hildmannianus,-49.2731,-25.4278	Curitiba-PR	MBM
Cereus hildmannianus,-51.4581,-25.3953	Guarapuava-PR	MBM
Cereus hildmannianus,-50.1619,-25.095	Ponta Grossa-PR	UPCB
Cereus hildmannianus,-47.64,-22.72	Piracicaba-SP	ESA
Cereus hildmannianus,-52.196583,-22.486278	Teodoro Sampaio-SP	ESA
Cereus hildmannianus,-49.22,-21.46	Novo Horizonte-SP	HSJRP
Cereus hildmannianus,-46.88,-23.18	Jundiaí-SP	MBM
Cereus hildmannianus,-48.516,-22.0892	Bocaina-SP	SinBiota
Cereus hildmannianus,-48.5284,-22.0834	Bocaina-SP	SinBiota

Cereus hildmannianus,-48.4672,-22.0216	Boa Esperança do Sul-SP	SinBiota
Cereus hildmannianus,-49.079444,-22.238889	Bauru-SP	SPSF
Cereus hildmannianus,-49.07,-24.062222	Itapeva-SP	SPSF
Cereus hildmannianus,-46.8389,-23.0058	Itatiba - SP	Coleta
Cereus hildmannianus,-53.514528,-30.823917	Caçapava do Sul - RS (I)	Coleta
Cereus hildmannianus,-53.514028,-30.821139	Caçapava do Sul - RS (II)	Coleta
Cereus hildmannianus,-53.513028,-30.819222	Caçapava do Sul - RS (III)	Coleta
Cereus hildmannianus,-53.509778,-30.811139	Caçapava do Sul - RS (IV)	Coleta
Cereus hildmannianus,-53.513194,-30.797222	Caçapava do Sul - RS (V)	Coleta
Cereus hildmannianus,-54.730639,-29.448722	Jaguari - RS (I)	Coleta
Cereus hildmannianus,-54.744139,-29.383306	Jaguari - RS (II)	Coleta
Cereus hildmannianus,-55.052444,-29.037	Santiago - RS (I)	Coleta
Cereus hildmannianus,-55.068361,-29.035472	Santiago - RS (II)	Coleta
Cereus hildmannianus,-53.777583,-29.779556	Santa Maria -RS	Coleta
Cereus hildmannianus,-50.972444,-30.246028	Região de Viamão - RS (I)	Coleta
Cereus hildmannianus,-50.96075,-30.234611	Região de Viamão - RS (II)	Coleta
Cereus hildmannianus,-50.952972,-30.230083	Região de Viamão - RS (III)	Coleta
Cereus hildmannianus,-50.921139,-31.11025	Mostardas - RS	Coleta
Cereus hildmannianus,-48.598111,-26.775333	Penha - SC	Coleta
Cereus hildmannianus,-48.457167,-27.580528	Florianópolis - SC (Costa da Lagoa)	Coleta
Cereus hildmannianus,-48.422889,-27.574306	Florianópolis - SC (Praia da Barra)	Coleta
Cereus hildmannianus,-48.448528,-27.619861	Florianópolis - SC (Praia da Joaquina)	Coleta
Cereus hildmannianus,-48.505028,-27.724667	Florianópolis - SC (Praia da Armação)	Coleta
Cereus hildmannianus,-48.507139,-27.705222	Florianópolis - SC (Morro das Pedras)	Coleta
Cereus hildmannianus,-48.778833,-28.487528	Laguna - SC	Coleta
Cereus hildmannianus,-52.507278,-23.656056	Cianorte - PR	Coleta

Cereus hildmannianus,-51.885306,-25.295	Guarapuava - PR (Rio do poço)	Coleta
Cereus hildmannianus,-52.070806,-25.416667	Guarapuava - PR (Cantagalo)	Coleta
Cereus hildmannianus,-52.115444,-25.774222	Segredo - PR	Coleta
Cereus hildmannianus,-49.923889,-29.540833	Arroio do Sal - RS	Coleta
Cereus hildmannianus,-50.22525,-29.958639	Osório - RS	Coleta
Cereus hildmannianus,-51.415306,-30.272056	Barra do Ribeiro (Ponto I)	Coleta
Cereus hildmannianus,-51.406333,-30.236333	Barra do Ribeiro (Ponto II)	Coleta
Cereus hildmannianus,-52.169583,-29.2175	Pouso Novo - RS	Coleta
Cereus hildmannianus,-52.519972,-27.73125	Jacutinga - RS	Coleta
Cereus hildmannianus,-47.574444,-21.259167	Serrana - SP	Coleta
Cereus hildmannianus,-49.383333,-24.116667	Sengés - PR	Coleta
Cereus hildmannianus,-50.038722,-23.23725	Jacarezinho - PR	Coleta
Cereus hildmannianus,-49.132722,-22.435889	Piratininga - SP	Coleta
Cereus hildmannianus,-47.222944,-23.2515	Itu - SP (Salto I)	Coleta
Cereus hildmannianus,-47.218361,-23.266639	Itu - SP (Salto II)	Coleta

*: Dados do herbário: ICN-UFRS e PACA (Herbarium Anchieta-RS), dados do site http://splink.cria.org.br: MBM, UPCB, ESA, HSJRP, SinBiota, SPSF e dados de coletas realizadas pelo Laboratório de Genética Evolutiva-FMRP-USP.

APÊNDICE 2

Tabela 1. Haplótipos com *indels*codificados, número de indivíduos e suas localidades amostradas.

phyC	H01	37	ITATIBA-SP, CAÇAPAVA DO SUL-RS, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, MOSTARDAS-RS, PENHA-SC, LAGUNA-SC, FLORIANÓPOLIS-SC, CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR, MANGUEIRINHA-PR, ARROIO DO SAL-RS, OSÓRIO-RS, BARRA DO RIBEIRO-RS, JACUTINGA-RS, JACAREZINHO-PR, PIRATININGA-PR, ITU-SP.
	H02	14	CAÇAPAVA DO SUL-RS, JAGUARI-RS, VIAMÃO-RS, MOSTARDAS-RS, LAGUNA-SC, FLORIANÓPOLIS-SC, MANGUEIRINHA-PR, OSÓRIO-RS, BARRA DO RIBEIRO-RS, POUSO NOVO-RS.
	H03	2	SERRANA-SP.
	H04	2	SENGÉS-PR.
	H05	1	CIANORTE-PR.
íntron *trnL*	H01	9	ITATIBA-SP, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR, MANGUEIRINHA-PR.
	H02	9	CAÇAPAVA DO SUL-RS, VIAMÃO-RS, MOSTARDAS-RS, PENHA-SC, LAGUNA-SC, FLORIANÓPOLIS-SC.
petL-psbE	H01	9	ITATIBA-SP, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR, MANGUEIRINHA-PR.
	H02	7	CAÇAPAVA DO SUL-RS, VIAMÃO-RS, MOSTARDASRS, LAGUNA-SC, FLORIANÓPOLIS-SC.
atpI-atpH	H01	8	ITATIBA-SP, CAÇAPAVA DO SUL-RS, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, CIANORTE-PR, JACUTINGA-RS.
	H02	20	CAÇAPAVA DO SUL-RS, VIAMÃO-RS, MOSTARDAS-RS, PENHA-SC, LAGUNA-SC, FLORIANÓPOLIS-SC, POUSO NOVO-RS.
	H03	1	CIANORTE-PR.

	H01	54	ITATIBA-SP, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR, MANGUEIRINHA-PR, JACUTINGA-RS, SERRANA-SP, SENGÉS-PR, JACAREZINHO-PR, PIRATININGA-PR, ITU-SP.
trnQ-5'rps16	H02	32	CAÇAPAVA DO SUL-RS, JAGUARI-RS, VIAMÃO-RS, MOSTARDAS-RS, LAGUNA-SC, FLORIANÓPOLIS-SC, ARROIO DO SAL-RS, OSÓRIO-RS, BARRA DO RIBEIRO-RS.
	H03	3	SERRANA-SP.
	H04	3	CAÇAPAVA DO SUL-RS, POUSO NOVO-RS.
	H05	1	SENGÉS-PR.
	H01	69	ITATIBA-SP, CAÇAPAVA DO SUL-RS, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR, MANGUEIRINHA-PR, BARRA DO RIBEIRO-RS, JACUTINGA-RS, SERRANA-SP, JACAREZINHO-PR, PIRATININGA-PR, ITU-SP.
	H02	41	CAÇAPAVA DO SUL-RS, VIAMÃO-RS, MOSTARDAS-RS, PENHA-SC, LAGUNA-SC, FLORIANÓPOLIS-SC, ARROIO DO SAL-RS, OSÓRIO-RS, BARRA DO RIBEIRO-RS, POUSO NOVO-RS.
	H03	15	CAÇAPAVA DO SUL-RS, VIAMÃO-RS, OSÓRIO-RS, BARRA DO RIBEIRO-RS, POUSO NOVO-RS.
	H04	4	CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR.
psbJ-petA	H05	7	ITATIBA-SP, CANTAGALO-PR, ITU-SP.
	H06	1	ITATIBA-SP.
	H07	2	SENGÉS-PR.
	H08	2	CIANORTE-PR, MANGUEIRINHA-PR.
	H09	1	CAÇAPAVA DO SUL-RS.
	H10	3	ARROIO DO SAL-RS.
	H11	3	FLORIANÓPOLIS-SC.
	H12	1	FLORIANÓPOLIS-SC.
Concatenado (*trnQ-5'rps16 - psbJ-petA*)	H01	40	ITATIBA-SP, JAGUARI-RS, SANTIAGO-RS, SANTA MARIA-RS, CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR, MANGUEIRINHA-PR, JACUTINGA-RS, PIRATININGA-PR.

H02	18	VIAMÃO-RS, MOSTARDAS-RS, LAGUNA-SC, FLORIANÓPOLIS-SC, OSÓRIO-RS, BARRA DO RIBEIRO-RS.
H03	4	CIANORTE-PR, GUARAPUAVA-PR, CANTAGALO-PR.
H04	4	ITATIBA-SP, CANTAGALO-PR, ITU-SP.
H05	1	JACAREZINHO-PR.
H06	1	CAÇAPAVA-RS.
H07	1	ITATIBA-SP
H08	2	JAGUARI-RS, BARRA DO RIBEIRO-RS.
H09	2	CIANORTE-PR, MANGUEIRINHA-PR.
H10	1	SERRANA-SP.
H11	1	SENGÉS-PR.
H12	1	CAÇAPAVA DO SUL-RS.
H13	3	CAÇAPAVA DO SUL-RS, VIAMÃO-RS, BARRA DO RIBEIRO-RS.
H14	1	ARROIO DO SAL-RS.
H15	1	FLORIANÓPOLIS-SC.
H16	2	FLORIANÓPOLIS-SC.
H17	2	POUSO NOVO-RS.

APÊNDICE 3. Valores de Dc e Dn para cada clado gerados pelo programa Geodis 2.6.

```
Differentiating population structure from history - Geodis  2.6
(c) Copyright, 1999-2009 David Posada and Alan Templeton
Contact: David Posada, University of Vigo, Spain (dposada@uvigo.es)
```

DNA cloroplastidial *psbJ-petA*

```
Input file: C:\Users\Gislaine\Documents\1 Doc\Suecia\seqs\GEODIS_PSBJPETA_CODIFICADO.txt

psbJ-petA codificado - cpDNA
Wed Jun 27 14:43:17 CEST 2012

PERMUTATION ANALYSIS OF Clade 1-1
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =    70.5041
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.1108
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-1       -28 52'16"    -49 48'30"
               I       -29 02'55"    -50 02'24"
              II       -27 35'05"    -48 26'01"
             III       -27 34'49"    -48 27'25"
              IV       -29 32'27"    -49 55'26"

CLADE I (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      216.7010       0.9914      0.0086*
      NESTED CLADE      216.1448       0.9913      0.0087*

CLADE II (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        2.2447       0.0089*    0.9919
      NESTED CLADE      196.2919       0.4927      0.5081

CLADE III (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      195.0912       0.5784      0.5051

CLADE IV (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       0.0067*    1.0000
      NESTED CLADE       75.2248       0.0004*    1.0000

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      215.7390       1.0000       0.0000*
      NESTED CLADE       71.9103       0.9836       0.0164
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0127 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 1-3
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =     3.0000
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     1.0000
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-3       -27 25'06"    -52 50'43"
               I       -30 47'50"    -53 30'47"
              II       -24 42'54"    -52 18'41"

CLADE I (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      380.9227       0.6724      0.6594
```

```
CLADE II (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      119.3126       0.3318      1.0000
      NESTED CLADE      307.9198       0.3318      1.0000
NO INTERIOR/TIP CLADES EXIST IN THIS GROUP
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 1-5
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =     0.3810
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     1.0000
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-5       -23 23'17"    -47 34'20"
               I       -23 00'20"    -46 50'20"
              II       -23 24'20"    -47 36'19"

CLADE I (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE       86.0914       0.8685      0.8752

CLADE II (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       97.9394       0.8752      0.8685
      NESTED CLADE       94.8231       0.8752      0.8685

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       97.9394       0.8752       0.8685
      NESTED CLADE        8.7316       0.8752       0.8685
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 1-6
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =    14.6164
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.5542
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-6       -26 23'35"    -51 56'09"
               I       -26 26'58"    -51 55'52"
              II       -24 58'59"    -52 03'31"

CLADE I (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      404.7478       0.9517      0.0489
      NESTED CLADE      404.6125       0.9490      0.0516

CLADE II (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       65.0327       0.0100*    0.9906
      NESTED CLADE      158.0212       0.0104*    0.9902

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE     -339.7152       0.0075*     0.9931
      NESTED CLADE     -246.5913       0.0110*     0.9896
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 2-1
```

```
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =    30.3647
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0157
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 2-1       -29 14'42"    -50 29'29"
             1-1       -28 52'16"    -49 48'30"
             1-2       -30 22'44"    -52 33'46"

CLADE 1-1 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      205.2429       0.0215*    0.9786
      NESTED CLADE      219.5441       0.1870      0.8130

CLADE 1-2 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      123.0903       0.0052*    0.9948
      NESTED CLADE      247.2143       0.8265      0.1735

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      -82.1526       0.0372       0.9628
      NESTED CLADE       27.6701       0.8230       0.1770
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 2-2
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   132.7100
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0021
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 2-2       -26 02'49"    -51 27'34"
             1-3       -27 25'06"    -52 50'43"
             1-4       -24 07'00"    -49 22'60"
             1-5       -23 23'17"    -47 34'20"
             1-6       -26 23'35"    -51 56'09"

CLADE 1-3 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      340.3655       0.7297      0.2723
      NESTED CLADE      350.3504       0.3283      0.6737

CLADE 1-4 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       0.0654      1.0000
      NESTED CLADE      299.3535       0.2897      0.7104

CLADE 1-5 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       94.4458       0.0004*    0.9996
      NESTED CLADE      500.5985       0.9583      0.0417

CLADE 1-6 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      395.1340       0.2462      0.7538
      NESTED CLADE      400.9009       0.4358      0.5642

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      317.4155       0.9998       0.0002*
      NESTED CLADE      -61.4440       0.1602       0.8398
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0127 (Dunn-Sidak correction)
*****************************************************************
```

```
PERMUTATION ANALYSIS OF Clade 3-1
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   136.0241
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0000
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 3-1       -27 35'23"    -50 59'33"
             2-1       -29 14'42"    -50 29'29"
             2-2       -26 02'49"    -51 27'34"

CLADE 2-1 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      226.4047       0.0000*    1.0000
      NESTED CLADE      301.1625       0.0000*    1.0000

CLADE 2-2 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      402.4854       0.9753      0.0247*
      NESTED CLADE      445.2561       1.0000      0.0000*

NO INTERIOR/TIP CLADES EXIST IN THIS GROUP
****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
****************************************************************
** ANALYSIS FINISHED **
It took 6.6070 seconds.
Differentiating population structure from history - Geodis  2.6
(c) Copyright, 1999-2009 David Posada and Alan Templeton
Contact: David Posada, University of Vigo, Spain (dposada@uvigo.es)
```

DNA cloroplastidial *trnQ-5'rps16*

```
Input file: C:\Users\Gislaine\Documents\1 Doc\Suecia\seqs\GEODIS_TRNQ-RPS16.txt
trnQ-rpS16 cpDNA

Mon Jul 02 14:54:44 CEST 2012
PERMUTATION ANALYSIS OF Clade 1-3
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   144.4528
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0004
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-3       -26 06'02"    -51 37'49"
               I       -24 07'00"    -49 22'60"
              II       -29 44'39"    -52 37'02"
             III       -23 14'14"    -50 02'19"
              IV       -25 58'28"    -51 43'37"

CLADE I (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      315.6735       0.5478      0.4866

CLADE II (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       96.9120       0.1422      0.8578
      NESTED CLADE      417.9178       0.7283      0.2717

CLADE III (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      356.3587       0.6056      0.4132
CLADE IV (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      353.3492       0.1687      0.8313
      NESTED CLADE      356.8147       0.2624      0.7376

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
```

```
      WITHIN CLADE      295.2020       0.8219       0.1781
      NESTED CLADE      -28.3424       0.2618       0.7382
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0127 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 2-1
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   159.2926
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0000
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 2-1       -27 26'19"    -51 01'54"
I(1-3)       -26 06'02"    -51 37'49"
         II(1.2)       -29 20'35"    -50 31'32"
        III(1.1)       -21 15'33"    -47 34'28"

CLADE I(1-3) (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      361.2990       0.3018      0.6982
      NESTED CLADE      401.9211       0.9852      0.0148*

CLADE II(1.2) (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      213.1574       0.0000*    1.0000
      NESTED CLADE      314.6223       0.0071*    0.9929

CLADE III(1.1) (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       0.0021*    1.0000
      NESTED CLADE      770.3838       1.0000      0.0000*

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      166.4122       1.0000       0.0000*
      NESTED CLADE       48.2335       0.9021       0.0979
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0170 (Dunn-Sidak correction)
*****************************************************************
** ANALYSIS FINISHED **
It took 2.9990 seconds.
Differentiating population structure from history - Geodis  2.6
(c) Copyright, 1999-2009 David Posada and Alan Templeton
Contact: David Posada, University of Vigo, Spain (dposada@uvigo.es)
```

DNA cloroplastidialconcatenado *trnQ-5'rps16* e *psbJ-petA*

```
Input file: C:\Users\Gislaine\Documents\1 Doc\Suecia\seqs\1 concatenado\Geodis_JAQ16.txt
Concatenado_psbjtrnq  - cpDNA

Wed Jul 04 11:37:18 CEST 2012
PERMUTATION ANALYSIS OF Clade 1-2
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =    33.9167
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.2661
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-2       -28 58'15"    -49 32'18"
               I       -29 04'34"    -49 37'50"
              II       -29 32'27"    -49 55'26"
             III       -27 34'49"    -48 27'25"
              IV       -27 36'24"    -48 27'04"

CLADE I (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      179.3032       0.9315      0.0699
      NESTED CLADE      178.5173       0.9013      0.1001
```

```
CLADE II (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE       73.5304       0.0448      1.0000

CLADE III (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      187.2237       0.5864      0.5897

CLADE IV (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        1.9836       0.1141      0.9188
      NESTED CLADE      185.1415       0.5316      0.5013

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      178.3114       0.9939       0.0063*
      NESTED CLADE       20.7580       0.8339       0.1663
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0127 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 1-8
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =     0.8333
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     1.0000
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-8       -23 22'52"    -47 33'32"
               I       -23 00'20"    -46 50'20"
              II       -23 25'07"    -47 37'51"

CLADE I (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE       84.5278       0.8055      0.8044

CLADE II (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      100.5909       0.8044      0.8055
      NESTED CLADE       93.8706       0.8044      0.8055

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      100.5909       0.8044       0.8055
      NESTED CLADE        9.3428       0.8044       0.8055
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 1-9
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   108.0500
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0049
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:

GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 1-9       -26 27'55"    -52 23'41"
               I       -26 42'54"    -52 36'20"
              II       -23 14'14"    -50 02'19"
             III       -24 58'59"    -52 03'31"
              IV       -30 47'50"    -53 30'47"
               V       -24 42'54"    -52 18'41"

CLADE I (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      325.8211       0.3636      0.6364
```

```
      NESTED CLADE      327.4159       0.3463      0.6537

CLADE II (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      430.1000       0.8471      0.1746

CLADE III (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       65.0327       0.0319      0.9715
      NESTED CLADE      171.0898       0.0879      0.9155

CLADE IV (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      493.4110       0.8795      0.1411

CLADE V (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      119.3126       0.4823      0.5764
      NESTED CLADE      197.0032       0.2581      0.8006

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      263.4766       0.9218       0.0782
      NESTED CLADE       77.1814       0.7654       0.2346
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0102 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 2-1
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =    46.3295
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0043
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:
GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 2-1       -29 15'13"    -50 11'42"
             1.1       -29 13'03"    -52 10'10"
             1.2       -28 58'15"    -49 32'18"
             1.3       -30 30'16"    -52 09'45"

CLADE 1.1 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       0.0620      1.0000
      NESTED CLADE      191.4500       0.4985      0.5043

CLADE 1.2 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      168.3254       0.0431      0.9569
      NESTED CLADE      177.4966       0.1335      0.8665

CLADE 1.3 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      121.8665       0.4067      0.5951
      NESTED CLADE      239.9254       0.8727      0.1291

PART III.  TEST OF INTERIOR VS. TIP CLADES:

 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      -32.4318       0.5449       0.4551
      NESTED CLADE       61.2660       0.8881       0.1119
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0170 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 2-2
   BASED ON 10000 RESAMPLES
PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =     3.0000
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     1.0000
```

```
PART II.  GEOGRAPHIC DISTANCE ANALYSIS:
GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 2-2       -29 56'32"    -53 58'57"
             1.4       -30 47'50"    -53 30'47"
             1.5       -29 35'09"    -54 10'41"

CLADE 1.4 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      105.1067       0.6594      0.6724

CLADE 1.5 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       92.4029       1.0000      0.3318
      NESTED CLADE      117.1107       0.6724      0.6594
NO INTERIOR/TIP CLADES EXIST IN THIS GROUP
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0253 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 2.3
   BASED ON 10000 RESAMPLES

PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   133.7679
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0006

PART II.  GEOGRAPHIC DISTANCE ANALYSIS:
GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 2.3       -25 39'55"    -51 23'57"
             1.6       -24 07'00"    -49 22'60"
             1.7       -21 15'33"    -47 34'28"
             1.8       -23 22'52"    -47 33'32"
             1.9       -26 27'55"    -52 23'41"

CLADE 1.6 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      266.1875       0.5657      0.4529

CLADE 1.7 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      625.5958       1.0000      0.0186

CLADE 1.8 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE       93.0213       0.0234      0.9767
      NESTED CLADE      474.3756       0.8852      0.1149

CLADE 1.9 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      332.3701       0.0048*    0.9952
      NESTED CLADE      346.3000       0.0267      0.9733

PART III.  TEST OF INTERIOR VS. TIP CLADES:
 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      265.9264       0.8993       0.1007
      NESTED CLADE     -119.9373       0.0419       0.9581
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0127 (Dunn-Sidak correction)
*****************************************************************

PERMUTATION ANALYSIS OF Clade 3.1
   BASED ON 10000 RESAMPLES

PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   129.8754
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0000

PART II.  GEOGRAPHIC DISTANCE ANALYSIS:
GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       Clade 3.1       -27 26'19"    -51 01'54"
```

```
            2.1       -29 15'13"    -50 11'42"
            2.2       -29 56'32"    -53 58'57"
            2.3       -25 39'55"    -51 23'57"

CLADE 2.1 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      189.3167       0.0000*    1.0000
      NESTED CLADE      294.0690       0.0028*    0.9972

CLADE 2.2 (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      113.5801       0.1088      0.8912
      NESTED CLADE      417.6981       0.7525      0.2475

CLADE 2.3 (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      372.1352       0.5287      0.4713
      NESTED CLADE      431.0501       1.0000      0.0000*

PART III.  TEST OF INTERIOR VS. TIP CLADES:
 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE     -198.3587       0.2411       0.7589
      NESTED CLADE       31.7515       0.6972       0.3028
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0170 (Dunn-Sidak correction)
*****************************************************************
** ANALYSIS FINISHED **
It took 4.3050 seconds.
Differentiating population structure from history - Geodis  2.6
(c) Copyright, 1999-2009 David Posada and Alan Templeton
Contact: David Posada, University of Vigo, Spain (dposada@uvigo.es)
```

DNA nuclear *PhyC*

```
Input file: C:\Users\Gislaine\Documents\1 Doc\Suecia\seqs\GEODIS_PHYC.txt
PHYC - cpDNA

Wed Jun 27 14:48:59 CEST 2012

PERMUTATION ANALYSIS OF clade 1.1
   BASED ON 10000 RESAMPLES

PART I.  PERMUTATIONAL CONTINGENCY TEST:
  OBSERVED CHI-SQUARE STATISTIC =   173.1892
  THE PROBABILITY OF A RANDOM CHI-SQUARE BEING GREATER THAN
    OR EQUAL TO THE OBSERVED CHI-SQUARE =     0.0004

PART II.  GEOGRAPHIC DISTANCE ANALYSIS:
GEOGRAPHICAL CENTERS    LATITUDE      LONGITUDE
       clade 1.1       -27 11'01"    -50 40'05"
               I       -21 15'33"    -47 34'28"
II       -29 20'58"    -51 18'53"
             III       -23 39'21"    -52 30'26"
              IV       -24 07'00"    -49 22'60"
               V       -26 56'44"    -50 36'38"

CLADE I (Tip)
TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       0.0184      1.0000
      NESTED CLADE      728.8687       1.0000      0.0004*

CLADE II (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      214.6166       0.0003*    0.9997
      NESTED CLADE      332.6558       0.1606      0.8394

CLADE III (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       1.0000      1.0000
      NESTED CLADE      433.1549       0.6358      0.3971

CLADE IV (Tip)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE        0.0000       0.0167      1.0000
      NESTED CLADE      364.1864       0.5406      0.4600
```

```
CLADE V (Interior)
 TYPE OF DISTANCE       DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      356.3885       0.3217      0.6783
      NESTED CLADE      356.3196       0.2572      0.7428

PART III.  TEST OF INTERIOR VS. TIP CLADES:
 TYPE OF DISTANCE   I-T DISTANCE      PROB.<=      PROB.>=
      WITHIN CLADE      198.2499       0.9995       0.0005*
      NESTED CLADE      -26.6513       0.2572       0.7428
*****************************************************************
*NOTE: TESTS ARE REGARDED AS SIGNIFICANT AT THE 5% LEVEL TREEWISE
IF THEIR PROBABILITY IS LESS THAN 0.0102 (Dunn-Sidak correction)
*****************************************************************
** ANALYSIS FINISHED **
It took 1.7210 seconds.
```

APÊNDICE 4

TECHNICAL NOTE

Analysis of genetic variability in noncoding chloroplast DNA regions of *Cereus hildmannianus* (CACTACEAE)

Silva, GAR[1]; Moraes, EM[2]; Manfrin, MH[1,3]
[1]Laboratório de Genética Evolutiva, Departamento de Genética, Programa de Pós-Graduação em Genética, FMRP-USP. [2]Campus Sorocaba, Universidade Federal de São Carlos. [3]Laboratório de Genética Evolutiva, Departamento de Biologia, FFCLRP-USP.

Abstract

Phylogeographical, studies of cacti species may help in understanding the evolutionary history of these species and add information about the dynamics of the dry areas in South America. However, a fundamental condition for such studies is to found sufficiently variable DNA regions that allow tracking the genealogical relationships among lineages. Universal primers were described in the literature and sequence amplification of noncoding chloroplast DNA (cpDNA) regions have generally shown to provide adequate phylogenetic information on both species or population-level studies. Herein, our objective was to evaluate the genetic variation through of 13 noncoding cpDNA regions of *Cereus hildmannianus* (Cactaceae) that occurs in the southeast and south of Brazil. The sequence choice was *trnH-psbA, trnS-trnG-trnG* (região intergênica *trnS-trnG* e íntron *trnG*), *trnT-trnL, trnL-trnL-trnF* (íntron *trnL* e região intergênica *trnL-trnF*), *psbJ-petA, atpI-atpH, 3'rps16-5'trnK, 3'trnk-matk, trnQ-5'rps16, psbD-trnT* e *petL-psbE*. Only five of these noncoding regions, showed genetic variability that could be useful for population studies in *C. hildmannianus*.

Keywords: phylogeography, chloroplast DNA, *Cereus hildmannianus*, Cactaceae, variability.

O gênero *Cereus* (tribo Cereeae, subfamília Cactoideae) compreende quatro subgêneros distribuídos na América do Sul, dentre os quais, o subgênero *Cereus* é encontrado em vários tipos de vegetação, como Mata Atlântica, caatinga, cerrado e campos rupestres (Taylor e Zappi, 2004). Entre as espécies do gênero, *Cereus hildmannianus* ocorre em manchas de afloramentos rochosos ou solos bem drenados no interior de florestas úmidas e semiúmidas no Leste, Sudeste e Sul do Brasil, Centro e Sudeste da América do Sul (Paraguai, Uruguai, Argentina) e leste do Chaco (Taylor e Zappi, 2004).

A hipótese de alterações paleoclimáticas atuando na distribuição da vegetação xerófita na América do Sul e, portanto, na história evolutiva de táxons associados a essa vegetação, pode ser averiguada por meio de análises filogeográficas. Estudos

filogeográficos associam genealogias de genes com a distribuição geográfica das populações, permitindo a inferência de processos históricos e recorrentes envolvidos na determinação dos padrões atuais de distribuição e diversificação de populações e espécies (Avise, 2000).

Uma condição fundamental para estudos como esse é a prospecção de regiões genômicas com variação suficiente para traçar a genealogia dos diferentes haplótipos. A escolha de DNA cloroplastidial (cpDNA) para estudos filogeográficos em plantas tem sido realizada em detrimento ao genoma mitocondrial (mtDNA), o marcador molecular tradicional para estudos filogeográficos em animais. Isto ocorre devido ao mtDNA de plantas evoluir lentamente em sua seqüência nucleotídica primária (Avise, 2009), diferentemente das taxas encontradas no mtDNA animal. Essa lenta evolução de seqüências de mtDNA em plantas gera variação insuficiente para análises filogeográficas (Avise, 2009). Além do mtDNA em plantas apresentar baixos índices de substituição nucleotídica há ainda a possibilidade de heteroplasmia devido à recombinação intramolecular (Schaal et al., 1998).

O genoma do cloroplasto é haplóide, sendo uma única e não-recombinante unidade de herança, com regiões possuindo diferentes taxas de mutação (Schaal et al., 1998). A organização da molécula de cpDNA de plantas terrestres consiste de uma região de *large single copy* (LSC) e *small single copy* (SSC), separadas por duas regiões idênticas ou não idênticas de repetições invertidas, IR_A e IR_B. As repetições invertidas são estáveis e evoluem duas a três vezes mais lentamente quando comparadas com a região de cópia única, sendo a região de *large single copy* a que apresenta os maiores níveis de variação (Ravi et al., 2008).

Estas características do cpDNA permite o seu uso em estudos populacionais e filogeográficos, auxiliando no entendimento da história evolutiva de um táxon. As regiões de cpDNA a serem analisadas são consideradas de acordo com a taxa de mutação e Shaw et al. (2005, 2007) demonstraram que regiões não codificantes seriam uma potencial fonte de variação para estudos intra e interespecíficos. O desenvolvimento de *primers* para o cpDNA tem sido realizado utilizando a condição conservada de regiões codificantes vizinhas a regiões não codificantes, como espaçadores intergênicos e íntrons. Isto permitiu o desenvolvimento de *primers* universais para o estudo em diversos grupos de plantas (Grivet et al., 2001, Borsch e Quandt, 2009).

O objetivo do presente trabalho foi analisar o nível de variação nucleotídica em regiões não codificantes (segmentos intergênicos e íntrons) descritos na literatura como potenciais para estudos filogenéticos e filogeográficos em angiospermas (Shaw et al., 2005, 2007). A escolha das regiões não codificantes investigadas no presente estudo levou em consideração a análise realizada por Shaw et al. (2005, 2007) sobre o número de Caracteres Potencialmente Informativos PICs (Potentially Informative Characters) em cada região, uma medida do número de substituições nucleotídicas, *indels* e inversões.

Da região LSC do DNA cloroplastidial, foram testados 16 pares de *primers*, flanqueando as seguintes regiões: *trnH-psbA, trnS-trnG-trnG* (região intergênica *trnS-trnG* e íntron *trnG*), *trnT-trnL, trnL-trnL-trnF* (íntron *trnL* e região intergênica *trnL-trnF*), *psbJ-petA, atpI-atpH, 3'rps16-5'trnK, 3'trnk-matk, trnQ-5'rps16, psbD-trnT* e *petL-psbE*. Os *primers* utilizados para cada região amplificada estão listados na Tabela 1, a amplificação das regiões *trnS-trnG* e *3'trnK-matK* foi realizada em duas etapas utilizando *primers* internos.

O DNA genômico de cada indivíduo analisado foi isolado a partir de tecido de raiz utilizando o *Dneasy Plant Mini Kit*. A concentração de DNA em cada amostra foi avaliada com o uso de um NanoDrop ND-1000 Spectrophotometer e a qualidade do isolamento foi verificado em gel de agarose 1%. Em uma análise inicial o nível de variação nas 13 regiões estudas do cpDNA foi investigada de cinco a oito indivíduos provenientes das localidades Itatiba-SP, Caçapava do Sul-RS, Santiago-RS, Mostardas-RS, Penha-SC, Laguna-SC, Florianópolis-SC e Cianorte-PR, posteriormente, amostras de diferentes localidades foram adicionadas (Tabela 1).

As reações em cadeia da polimerase (PCR – *Polymerase Chain Reaction*) foram realizadas inicialmente, seguindo o protocolo proposto por Shaw et al. (2007). Entretanto, para otimizá-las, diferentes reações e condições de PCR foram testadas, utilizando um termociclador de gradiente Master Cycler Gradient (Eppendorf). O volume final de cada reação foi de 30 µl, contendo 1 µL de DNA genômico (10 – 60 ng/µl), 1X tampão de reação (Invitrogen), 200µM de dNTP, 0,1 µM de cada *primer*, $MgCl_2$ concentrado a 3 mM, 1,25 unidades de GoTaq® Flexi DNA Polymerase (Promega) e água livre de nuclease. Somente para as regiões *trnH-psbA* foi utilizado $MgCl_2$ a 2 mM e 0,4 unidades de Taq polimerase, e para o íntron *trnG* e *trnS-trnG* foram utilizados $MgCl_2$ a 1,5 mM, sendo a concentração da Taq polimerase de 0,4 unidades para o íntron *trnG* e 1 unidade para *trnS-trnG*.

As condições gerais de PCR foram de 95°C de temperatura de denaturação inicial a 2 mim, seguido por 35 ciclos de denaturação de 30 s a 95°C, *annealing* (pareamento dos *primers*) de 1 mim a 50°C e extensão de 2 mim a 72°C, com uma extensão final de 5 mim a 72°C. Variações deste protocolo foram realizadas para a temperatura de *annealing* do gene *PhyC* (54°C). Para a região *trnS-trnG-trnG*, as condições foram 80°C de temperatura de denaturação inicial a 5 mim, seguido por 40 ciclos de denaturação de 1 mim a 95°C, *annealing* (pareamento dos *primers*) de 1 mim a 62°C e extensão de 5 mim a 65°C, com uma extensão final de 5 mim a 65°C. Após a padronização, as reações de amplificação foram realizadas em um termociclador Gene Amp PCR® System 9700 (Applied Biosystems). Os segmentos amplificados foram purificados com o kit *GFXTM PCR and Gel Band Purification (GE Healthcare)*, seguindo o protocolo do kit.

As reações de sequenciamento foram realizadas em ambas as fitas, *forward* e *reverse*. Para cada fita, foram utilizados 50 ng de produto de PCR purificado, 10 µM de *primer*, 1,5 µL de tampão de sequenciamento 5X (kit BigDye) e 1 µL de BigDye ® Terminator v 3. 1 Cycle Sequencing kit (Applied Biosystems), completadas para um volume final de 10 µL com água livre de nuclease. Essas reações de sequenciamento foram realizadas em um termociclador Gene Amp PCR® System 9700 (Applied Biosystems), com as seguintes condições: uma etapa de 1 mim a 96°C, seguido por 39 ciclos de 15 s a 96°C, 15 s a 50°C (correspondente à temperatura de *annealing* de cada par de *primer*), e uma etapa final de 4 mim a 60°C.

Para a precipitação das amostras, foram adicionados às reações 80 µL de isopropanol 65 %, que permaneceram por 15 mim no escuro à temperatura ambiente, seguidos de centrifugação a 4000 rpm a 4°C por 45 mim. Este procedimento foi finalizado com o descarte do sobrenadante. Por duas vezes, 200 µL de etanol 70 % gelado foram adicionados, seguido de centrifugação a 4000 rpm a 4°C por 10 mim e descarte do sobrenadante. Finalizado com uma etapa de *spin* reverso a 1000 rpm a 4°C por 1 mim. As amostras precipitadas foram deixadas a 37°C por 1 hora. Após a precipitação, o sequenciamento foi realizado em ambas as fitas, no sequenciador automático ABI 3730 XL (Applied Biosystems).

Os cromatogramas de cada região amplificada foram visualizados e editados com o auxílio do programa Chromas Lite v.2.0. O alinhamento das sequências *forward* e *reverse* foi realizado com auxílio do programa CLUSTAL W v.1.8 (Thompson et al.,

1994). A busca por sequências semelhantes às encontradas neste trabalho foi realizada no banco de dados do *NCBI GenBank*, no programa BLAST (Altschul et al., 1990).

As sequências foram analisadas por meio de diversidade nucleotídica (π), que corresponde ao número médio de diferenças nucleotídicas por sítio entre duas sequências, diversidade haplotípica (h), que corresponde à probabilidade de dois haplótipos serem diferentes, composição nucleotídica. Para cada região de cpDNA, o nível de variação foi estimado por meio do índice PIC (potentially informative characters; Shaw et al., 2005), definido como a soma de substituições, *indels* e inversões (NS + ID + IV), considerando substituições e inversões dentro de *indels* como características independentes e L o comprimento da sequência. Os valores de PIC dividido pelo comprimento das sequências (PIC/L) e multiplicado por 100 permitiram o cálculo da percentagem de variabilidade para cada região.

As reações de isolamento de DNA foram realizadas para 157 indivíduos de 24 populações de *C. hildmannianus* e geraram amostras com 30 - 50 ng/µL. Um espécime de cada localidade foi utilizado como material testemunho e armazenado como exsicata no herbário São Paulo Faculdade Ribeirão, com as seguintes identificações: SPFR 11888; SPFR 12333; SPFR 12342; SPFR 12343; SPFR 12344; SPFR 12345; SPFR 12346; SPFR 12349; SPFR 12351; SPFR 12353; SPFR 12987; SPFR 12988; SPFR 12989; SPFR 13083; SPFR 13085; SPFR 13086; SPFR 13089; SPFR 13140; SPFR 13388; SPFR 13389; SPFR 13390; SPFR 13391. O número de indivíduos a partir das quais foram obtidas as sequências, suas respectivas populações, o número de haplótipos e os seus comprimentos em pares de base estão na Tabela 1.

Dentre as 13 regiões analisadas, somente cinco sequências de DNA cloroplastidiais apresentaram variação populacional em *C. hildmannianus*: íntron *trnL*; *petL-psbE*; *atpI-atpH*; *trnQ-5'rps16* e *psbJ-petA*. Para essas cinco regiões foi observada variação interpopulacional, enquanto variação intrapopulacional foi encontrada nas sequências cloroplastidiais intergênicas *atpI-atpH, trnQ-5'rps16* e *psbJ-petA*. Os dados obtidos para composição nucleotídica, índice de diversidade nucleotídica (π) e diversidade haplotípica (h) estão na Tabela 2. A região *psbJ-petA* apresentou o maior índice de diversidade nucleotídica e diversidade haplotípica. A composição nucleotídica das sequências mostrou uma maior porcentagem de bases AT para o DNA cloroplastidial.

A percentagem de variabilidade das sequências foi calculada somente para o DNA cloroplastidial. Os valores de PIC/L foram (1+0+0)/645 para o íntron *trnL*,

(1+0+0)/531 para o *petL-psbE*, (2+0+0)/591 para o *atpI-atpH*, (6+2+0)/541 para o *trnQ-5'rps16* e (5+82+0)/512 para o *psbJ-petA* (Tabela 2), com a percentagem de variabilidade de 0,15%, 0,18%, 0,33%, 1,47% e 16%, respectivamente. Desta forma, a maior percentagem de variabilidade foi encontrada para as duas regiões intergênicas *psbJ-petA* e *trnQ-5'rps16.*

Para trabalhos filogeográficos e de estruturação populacional são necessários marcadores com variação neste nível de análise. Para animais já existe muita informação consolidada na literatura com o gene *COI* (Chase et al., 2007). Para plantas, por outro lado, informações deste nível ainda são incipientes (Zimmer e Wen, 2012; Borsch and Quandt, 2009; Chase et al., 2007; Shaw et al., 2007). Alem dos estudos filogeográficos, vários trabalhos têm sido realizados para a procura de regiões com maior variabilidade em estudos filogenéticos e de DNA *barcoding* (Borsch e Quandt, 2009; Janzen, 2009; Chase et al., 2007, Shaw et al., 2005; Shaw et al., 2007; Ebert e Peakall, 2009). A seguir, as regiões cloroplastidiais utilizadas neste trabalho foram comparadas com alguns trabalhos, os quais utilizaram os *primers* descritos em Shaw e colaboradores (2005, 2007), para a maioria deles. A comparação foi realizada com os trabalhos de Shaw, pois dentre os artigos publicados (para mais detalhes, ver Ebert e Peakall, 2009), foram os que mais incluíram diferentes famílias de plantas para testar conjuntos de diferentes pares de *primers*.

Região intergênica cloroplastidial atpI-atpH. Sequências de 488 bp com uma variabilidade de 0,82 % em *Podostemum ceratophyllum* (Podostemaceae) (Fehrmann et al., 2012), comparada a 0,33 % para *C. hildmannianus*.Em trabalhos de filogeografia de *Cardamine scutata* (Brassicaceae), uma taxa de 1,9 % foi obtida (Lihov et al., 2010). Somente dois sítios polimórficos foram encontrados em 1104 bp nas sequências de *Fagusengleriana* (Fagaceae) (Lei et al., 2012).

Região intergênica cloroplastidial trnL-trnL-trnF. O íntron *trnL* é sugerido como apropriado para estudos de DNA *barcoding* (Taberlet et al., 2007) e descrito como fonte de variabilidade quando analisado concatenadamente com a região *trnL-trnL-trnF* (Shaw et al., 2005). Mas de acordo com os valores de PICs, o íntron *trnL* não foi considerado apropriado para estudos filogeográficos (Shaw et al., 2005). Apesar da baixa taxa de evolução, há análises filogeográficas com este íntron em *Lychophora ericoides* (Asteraceae) (Collevatti et al., 2009) e no presente trabalho. Já a região *trnL-trnF* apresentou seis sítios polimórficos e dois *gaps* na filogeografia de *Embothrium*

coccineum (Proteaceae) (Russell, et al., 2011), região que não apresentou variabilidade em *C. hildmannianus.*

Região intergênica cloroplastidial psbD-trnT. Para *Lobelia rhynchopetalum* (Campanulaceae), 468bp foram alinhados com um total de 9,2 % de variação (Geleta e Bryngelsson, 2012). Um estudo filogeográfico em *Podostemum ceratophyllum* (Podostemaceae) encontrou 0,41 % de variação em 739 bp (Fehrmann et al., 2012), sendo que para *C. hildmannianus*, não houve variabilidade em 668 bp.

Região cloroplastidial trnS-trnS-trnG. Fehrmann e colaboradores (2012) obtiveram 569 bp com 0,18 % de variação para *Podostemum ceratophyllum* (Podostemaceae). Em *C. hildmannianus* não houve variabilidade em 476 bp da região do íntron *trnS-trnS* e em 970 bp da região *trnS-trnG*.

Região intergênica cloroplastidial matK-trnK. Altas taxas de variabilidade apontam o *matK*, combinado a outras sequências, como uma das melhores regiões para DNA *barcoding* (Yesson et al., 2011; Janzen, 2009; Starr et al., 2009; Chase et al., 2007), porém pode haver problemas de especificidade dos *primers* nas diferentes famílias de plantas (Chase et al., 2007). Para *C. hildmannianus* não foi encontrada variabilidade nos 1224 bp sequenciados, embora, em *Lobelia rhynchopetalum* (Campanulaceae), os autores encontraram 8,5 % em 307 bp (Geleta e Bryngelsson, 2012).

Região intergênica cloroplastidial trnH-psbA. Esta região foi sugerida como uma das mais variáveis por Shaw (2005; 2007) e como potencial DNA *barcoding* (Shaw et al., 2007; Kress et al., 2005). Críticas a esse marcador estão relacionadas ao grande número de deleções e inserções e à alta ambiguidade entre os táxons, o que pode enviesar as inferências entre eles. Além disso, críticas também existem devido a diferenças no tamanho das sequências, que variam entre 300 a >1000 bp, devido a cópias de rpl22 e rps16 em algumas espécies (Chase et al., 2007). Em *C. hildmannianus*, não foram encontradas inserções ou deleções, nem variação entre as populações nas análises dos fragmentos de aproximadamente 360 bp (Tabela 1). Por outro lado, 24 sítios variáveis, para fragmentos de 315 bp, foram utilizados para inferir a filogeografia de *Oxyriadigyna* (Polygonaceae) (Allen et al., 2012). Este marcador também permitiu estudar a história populacional em *Encelia farinosa* (Asteraceae) com 10 haplótipos obtidos de aproximadamente 450 bp (Fehlberg e Ranker, 2009).

Região intergênica cloroplastidial trnQ-5'rps16. Em *Juglans mandshurica* (Juglandaceae), fragmentos de 605 bp apresentaram uma única substituição nucleotídica

e foram concatenados com mais seis marcadores para o estudo filogeográfico (Bai et al., 2010). Já em *Psammosilene tunicoides* (Caryophyllaceae), aproximadamente 805 bp apresentaram cinco substituições nucleotídicas e 2 a 27 bp com *indels* (Zhang et al., 2011).

Região intergênica cloroplastidial trnL-trnT. Allen et al. (2012) obtiveram 68 sítios variáveis em 910 bp desta região para *Oxyria digyna* (Polygonaceae). Em *Lobelia rhynchopetalum* (Campanulaceae), 9 % de variabilidade foi descrita para 525 bp (Geleta e Bryngelsson, 2012). No presente trabalho, as sequências obtidas não apresentaram variação em 348 bp.

Região intergênica cloroplastidial psbJ-petA. Este marcador, quando concatenado com *atpI-atpH* têm sido informativos em estudos filogeográficos, sendo encontradas regiões de cpSSR e *indels* (Sebastiani et al., 2004; Provan et al., 2004, Shaw et al., 2007). Em *C. hildmannianus*, a região *psbJ-petA* foi a região com maior variabilidade dentre os marcadores testados no presente trabalho (Tabela 2).

As comparações realizadas mostram disparidades ou até mesmo ausência de variabilidade das regiões não codificantes de DNA cloroplastidial descritas como as melhores alternativas, ou seja, demonstram que as generalizações podem não refletir a utilidade destes marcadores, pois desvios dessas taxas são comuns (Shaw, 2007). Desta foram, a avaliação das regiões cloroplastidiais para um estudo filogeográfico, mostrou que somente cinco das 13 regiões propostas em Shaw (2005, 2007) são úteis. Embora não seja o escopo deste trabalho fazer afirmações sobre as melhores regiões cloroplastidiais a serem utilizadas em estudos filogeográficos ou filogenéticos e DNA *barcode*, as considerações sobre essas regiões cloroplastidiais em *C. hildmannianus* podem ser úteis nestas investigações para outros táxons em *Cereus*.

REFERÊNCIAS BIBLIOGRÁFICAS

Allen G A, Marr K L, McCormick L J, Hebda J R. 2012. The impact of Pleistocene climate change on an ancient arctic–alpine plant: multiple lineages of disparate history in *Oxyria digyna*. Ecol Evol: 2(3):649-65.

Altschul S F, Gish W, Miller W, Myers E W, Lipman D J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-10.

Avise JC. 2000. Phylogeography: history and formation of species. Harvard University Press. 447p.

Avise JC. 2009. Phylogeography: retrospect and prospect. *Journal of Biogeography,* 36: 3-15.

Bai W-N, Liao W-J, Zhang D-Y. 2010. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in atemperate walnut tree from East Asia. New Phytologist. 188: 892–901.

Bonatelli I A S. 2010. Análise da estrutura filogeográfica do grupo *Pilosocereus aurisetus* (Cactaceae) utilizando marcadores moleculares do genoma do cloroplasto (cpDNA). *Dissertação (Mestrado em Genética e Evolução) - Universidade Federal de São Carlos, Fundação de Amparo à Pesquisa do Estado de São Paulo.*

Borsch T, Quandt D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolutio.* 282: 169-199.

Chase M W, Cowan R S, Hollingsworth P M, Berg C, Madriñán S, Petersen G, Seberg O, Jørgsensen T, Cameron K, Carine M, Pedersen N, Hedderson T A J, Conrad F, Salazar G A, Richardson J E, Hollingsworth M L, Barraclough T G, Kelly L, Wilkinson M. 2007. A proposal for a standardised protocol to barcode all land plants. TAXON 56 (2): 295–299.

Collevatti R, Rabelo S, Vieira R. 2009. Phylogeography and disjunct distribution in *Lychnophora ericoides* (Asteraceae), an endangered cerrado shrub species. *Annals of Botany.* 104: 655-664.

Demaio P H, Barfuss M H J, Kiesling R, Till W, Chiapella J O. 2011. Molecular phylogeny of *Gymnocalycium* (Cactaceae): Assessment of alternative infrageneric systems, a new Subgenus, and trends in the evolution of the genus. *American Journal of Botany.* 98(11): 1841-1854.

Ebert D, Peakall R. 2009. Invited Technical Review. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Molecular Ecology Resources:* 9: 673–690.

Fehlberg S D, Ranker T A. 2009. Evolutionary history and phylogeography of *Encelia farinosa* (Asteraceae) from the Sonoran, Mojave, and Peninsular Deserts. *Molecular Phylogenetics and Evolution.* 50(2): 326-35.

Fehrmann S, Philbrick C T, Halliburton A R. 2012. Intraspecific Variation in *Podostemum ceratophyllum* (Podostemaceae): Evidence of Refugia And Colonization Since The Last Glacial Maximum. *American Journal of Botany.* 99(1): 145–151.

Geleta M, Bryngelsson T. 2012. Population Genetic Analysis of Lobelia rhynchopetalum Hemsl. (Campanulaceae) Using DNA Sequences fromITS and Eight Chloroplast DNA Regions. *The ScientificWorld Journal*. 10p.

Grivet D, Heinze B, Vendramin G G, Petit R J. 2001. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Resource.* 1: 345-349.

Hamilton M B. 1999. Four primers pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology.* 8: 513-525.

Janzen D H. 2009. A DNA barcode for land plants. *CBOL Plant Working Group.* PNAS (106): 31.

Kress W J, Wurdack K J, Zimmer E A, Weigt L A, Janzen D H. 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. U.S.A. 102: 8369–8374.

Lei M, Wang Q, Wu Z-J, Pujol J L, Li D-Z, Zhang Z-Y. 2012. Molecular phylogeography of *Fagus engleriana* (Fagaceae) in subtropical China: limited admixture among multiple refugia. *Tree Genetics & Genomes.* 8(6): 1203-1212.

Lihov J, Kudoh H, Marhold M. 2010. Genetic Structure And Phylogeography of a Temperate-Boreal Herb, *Cardamine scutata* (Brassicaceae), In Northeast Asia Inferred From Aflps And Cpdna Haplotypes. *American Journal of Botany.* 97(6): 1058–1070

Müller K, Borsch T. 2005. Phylogenetics of Amaranthaceae based on *matK/trnK* sequence data – Evidence from parsimony, likelihood, and Bayesian analyses. Ann. Missouri Bot. *Gard*. 92: 66-102.

Nyffeler R. 2002. Phylogenetic Relationships In The Cactus Family (Cactaceae) Based On Evidence From Trnk/ Matk And TrnL-TrnF Sequences. *American Journal Of Botany.* 89(2): 312–326.

Provan J, Biss P M, McMeel D, Mathews S. 2004. Universal primers for the amplification of chloroplast microsatellites in grasses (Poaceae). *Molecular Ecology Notes.* 4: 262-264.

Ravi V, Khurana J, Tyagi A, Khurana P. 2008. An update on chloroplast genomes. *Plant Systematics and Evolution.* 271: 101-122.

Russell R V, Souto A C P, Premoli A A C. 2011. Multiple Pleistocene refugia in the widespread Patagonian tree *Embothrium coccineum* (Proteaceae). *Australian Journal of Botany.* 59: 299–314.

Schaal B, Da Hayworth K, Rauscher J, Smith W. 1998. Phylogeographic studies in plants: problems and prospects. *Molecular Ecology,* 7: 465-474.

Sebastiani F, Carnevale S, Vendramin G G. 2004. A new set of mono- and dinucleotide chloroplast microsatellites in Fagaceae. *Molecular Ecology Notes.* 4: 259-261.

Shaw J, Lickey E, Beck J, Farmer S, Liu W, Miller J, Siripun K, Winder C, Schilling E, Small R. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany.* 92: 142-166.

Shaw J, Lickey E, Schilling E, Small R. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany.* 94: 275-288.

Starr J R, Naczi F C, Chouinard B N. 2009 . Plant DNA barcodes and species resolution in sedges (*Carex*, Cyperaceae). *Molecular Ecology Resources.* 9(1): 151–163.

Taberlet P, Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann & Willerslev, E. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acid. Res. 35: e14.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal *primers* for amplification of three non-oding regions of chloroplast DNA. *Plant Molecular Biology.* 17: 1105-1109.

Taylor N, Zappi D. 2004. Cacti of Eastern Brasil. *Royal Botanic Gardens*, Kew. 511p.

Thompson J D, Higgins D G, Gibson T J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. 22: 4673-4680.

Yesson C, Barcenas R T B, Hernadndez H C M, Ruiz- Maqueda M L , Rodriguez V M, Prado A, Hawkins A J. 2011. DNA barcodes for Mexican Cactaceae, plants

under pressure from wild collecting. *Molecular Ecology Resources.* 11: 775–783.

Zhang Q, Zhao Y, Gong X. 2011. Genetic variation and phylogeography of *Psammosilene tunicoides* (Caryophyllaceae), a narrowly distributed and endemic species in south-western China. *Australian Journal of Botany*, 59, 450–459.

Tabela 1. Sequências dos pares de *primers* utilizados na amplificação das regiões intergênicas de DNA cloroplastidial em *C. hildmannianus* e descrição do número de indivíduos analisados por localidade, com o número e comprimento em pares de base dos haplótipos.

Segmento	*Primers*	Referência	Nº indivíduos / Nº população	Localidades	Nº haplótipos	(bp)**
trnH-psbA	**trnH** $^{(GUG)}$ **psbA**	Hamilton (1999)	8/8	3, 7, 11, 12, 13, 16, 19, 24	1	~ 360
íntron *trnG*	**5'trnG2G** **trnG** $^{(UUC)}$	Shaw et al. (2005) Shaw et al. (2007)	5/5	7, 11, 13, 19, 24	1	~ 476
trnS-trnG	**SG Rev 2** **trnS**$^{(GCU)}$ **SG Fwd 2**	Bonatelli (2010) Shaw et al. (2007) Bonatelli (2010)	8/8	3, 7, 16, 19, 11, 12, 13, 24	1	~ 970
trnT-trnL	**5'trnG2S** **5'trnL**^{UAA}**R (TabB)** **trnT**^{UGU}**F (TabA)**	Shaw et al. (2005) Taberlet et al. (1991)	8/8	3, 7, 11, 12, 13, 16, 19, 24	1	~ 348
trnL-trnF	**trnF**^{GAA} **(TabF)** **3´trnL**^{UAA}**R (TabE)**	Taberlet et al. (1991)	7/7	3, 7, 11, 12, 13, 19, 24	1	~ 413
íntron *trnL*	**trnL5'**^{UAA}**F (TabC)** **3'trnL**^{UAA}**R (TabD)**	Taberlet et al. (1991)	18/14	3, 7, 8, 9, 10, 11, 12, 13, 16, 17, 18, 19, 21, 24	2	~ 645
psbJ-petA	**psbJ** **petA**	Shaw et al. (2007)	149/24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24	12	~ 512
atpI-atpH	**atpI** **atpH**	Shaw et al. (2007)	29/12	3, 7, 11, 12, 13, 14, 15, 18, 19, 17, 21, 24	3	~ 591
3'rps16–5'trnK	**rpS16x2F2** **trnK**$^{(UUU)}$**x1**	Shaw et al. (2007)	8/8	3, 7, 11, 12, 13, 16, 19, 24	1	~ 221

3'trnk-matk	**matK50-Fdi**	Demaio et al. (2011)	29/24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24	1	~ 1224
	trnK-41R	Nyffeler (2002)*				
	ACmatk500F	Müller e Borsch (2005)				
	trnK-71R	Nyffeler (2002)				
petL-psbE	**petL** **psbE**	Shaw et al. (2007)	16/13	12, 13, 16, 19, 24	2	~ 531
trnQ-5'rps16	**trnQ(UUG)** **rpS16x1**	Shaw et al. (2007)	93/23	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24	6	~ 541
psbD-trnT	**psbD** **trnT(GGU)-R**	Shaw et al. (2007)	12/9	3, 7, 19, 17, 16, 18, 24, 11	1	~ 668

*: modificado de Nyffeler (2002).

**: comprimento das sequências em número de pares de bases (pb) obtidas após sequenciamento.

Localidades: 1 – Serrana-SP; 2 – Piratininga-SP; 3 – Itatiba-SP; 4 – Itu-SP; 5 – Jacarezinho-PR; 6 – Sengés-PR; 7 – Cianorte-PR; 8 – Guarapuava-PR; 9 – Cantagalo-PR; 10 – Manguerinha-PR; 11 – Penha-SC; 12 – Florianópolis-SC; 13 – Laguna-SC; 14 – Jacutinga-RS; 15 – Pouso Novo-RS; 16 – Santiago-RS; 17 – Jaguari-RS; 18 – Santa Maria-RS; 19 – Caçapava do Sul-RS; 20 – Barra do Ribeiro-RS; 21 – Viamão-RS; 22 – Arroio do Sal; 23 – Osório-RS; 24 – Mostardas-RS.

Tabela 2. Índices de diversidade nucleotídica (π), diversidade haplotípica (h), composição nucleotídica e número de caracteres potencialmente informativos (PIC/L) para os cinco segmentos das regiões intergênicas de DNA cloroplastidial em *C. hildmannianus*.

Sequências	**Diversidade nucleotídica (π)**	**Diversidade haplotípica (h)**	**Composição nucleotídica (%)**	**PIC/L**
íntron *trnL*	0,001044 +/- 0,001019	0,5294 +/- 0,0404	C: 11,80; T: 29,01; A: 43,04; G: 16,15	1/645
petL-psbE	0,001035 +/- 0,001021	0,5250 +/- 0,0546	C : 16,42; T : 34,54; A : 34,37; G : 14,66	1/531
atpI-atpH	0,001447 +/- 0,001400	0,4631 +/- 0,0798	C: 20,07; T: 32,87; A: 36,77; G: 10,29	2/591
trnQ-5'rps16	0,001506 +/- 0,001251	0,5482 +/- 0,0348	C : 12,24; T : 38,52; A : 36,99; G : 12,25	8/541
psbJ-petA	0,003216 +/- 0,002177	0,7105 +/- 0,0288	C: 15,46; T: 38,02; A: 34,50; G: 12,02	87/512