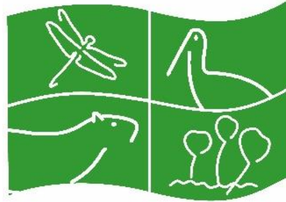

Programa de Pós-Graduação em Ecologia e Conservação

Centro de Ciências Biológicas e da Saúde

Universidade Federal de Mato Grosso Do Sul

Ecologia Molecular de Leguminosae e Cactaceae no Cerrado e Chaco: Fenologia, Filogeografia e Genética de Populações

Wellington Santos Fava

Campo Grande

novembro 2015

Ecologia Molecular de Leguminosae e Cactaceae no Cerrado e Chaco: Fenologia, Filogeografia e Genética de Populações

Wellington Santos Fava

Tese apresentada como requisito para a obtenção do título de **Doutor em Ecologia**, pelo Programa de Pós-Graduação em Ecologia e Conservação, Universidade Federal de Mato Grosso do Sul.

Orientadora: Aline Pedroso Lorenz Lemke

Banca avaliadora

Dra. Andreia Carina Turchetto Zolet
Universidade Federal do Rio Grande do Sul (UFRGS)

Dra. Ângela Lúcia Bagnatori Sartori
Universidade Federal de Mato Grosso do Sul (UFMS)

Dra. Camila Martini Zanella
Universidade Federal do Rio Grande do Sul (UFRGS)

Dra. Gecele Matos Paggi
Universidade Federal de Mato Grosso do Sul (UFMS)

Dra. Mariana Pires de Campos Telles
Universidade Federal de Goias (UFG)

Aos meus pais
(*in memoriam*)

“Much in evolution makes even more sense in the light of historical genealogy.”
John C. Avise, 2004 – *Molecular Markers, Natural History, and Evolution*

Agradecimentos

Agradeço a colaboração recebida na realização deste trabalho:

Dra. Aline Pedroso Lorenz Lemke, pela orientação e incentivo no desenvolvimento deste projeto, confiando no meu trabalho e o apoiando incondicionalmente.

Dra. Maria Ligia Rodrigues Macedo pelo apoio e confiança neste projeto.

Programa de Pós-Graduação em Ecologia e Conservação da Universidade Federal de Mato Grosso do Sul (UFMS), em especial à Rosilene Bejarano pelo apoio logístico e solicitude.

Dra. Andreia Carina Turchetto Zolet, Dra. Ângela Lúcia Bagnatori Sartori, Dra. Camila Martini Zanella, Dra. Gecele Matos Paggi e Dra. Mariana Pires de Campos Telles, por terem aceitado participar da banca examinadora.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de doutorado.

Laboratório de Evolução e Biodiversidade (LEBio/UFMS), pelo apoio técnico, e a todos os seus integrantes, pela amizade construída.

Dra. Mariana Pires de Campos Telles e Thais Castro da Universidade Federal de Goiás pelo sequenciamento das amostras de leguminosas.

Herbário da Universidade Federal de Mato Grosso do Sul (CGMS) pelo auxílio logístico para a realização dos Capítulos 1 e 2.

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - Edital MCT/CNPq/FNDCT/FAPs/MEC/CAPES/PRO-CENTRO-OESTE n° 31/2010 – e Rede

GENPAC (Genética Geográfica e Planejamento Regional para Conservação de Recursos Naturais no Cerrado) pelo auxílio financeiro e logístico para a realização do Capítulo 2.

CNPq e Ação Transversal nº 6/2011 – Casadinho/Procad (Angiospermas do Chaco Brasileiro: Sistemática, Diversidade, Fenologia e Adaptações) pelo auxílio financeiro para a realização dos Capítulos 3 e 4.

Dra. Anete Pereira de Souza (Universidade de Campinas), Dra. Fernanda Bered (Universidade Federal do Rio Grande do Sul – UFGRS) e Dra. Camila M. Zanella (UFGRS) pelo apoio no ensino e desenvolvimento de técnicas em biologia molecular necessárias para a realização dos capítulos 3 e 4.

Na caminhada para a realização deste trabalho, não estive sozinho. Senti a presença atuante de amigos, ouvi palavras animadoras de incentivo, de indicações e de críticas, que me ajudaram a seguir em frente. Não me passaram despercebidas companhias silenciosas, mas solidárias, e aqueles que dispuseram seu tempo, muitas vezes, simplesmente para ficar comigo, tornando meu trabalho ainda mais prazeroso e menos solitário. Gostaria de nomeá-los um a um, mas temo não incluir os que preferiram se manter no anonimato, torcendo por mim. Então prefiro dizer ***"muito obrigado"*** *a vocês todos que me estenderam a mão ou a força do seu espírito e do seu amor.*

Índice

Resumo Geral 09

General Abstract 11

Introdução Geral 13

Objetivo Geral 48

Chapter 1: Climate-driven and long term phenology of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) across the Cerrado 50

Chapter 2: Combining ecological niche modelling and phylogeography to reveal the demographic history of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) in the Cerrado 76

Chapter 3: Development and characterization of microsatellite markers for *Echinopsis rhodotricha* and cross-amplification in other species of Cactaceae 145

Chapter 4: Genetic variation in *Echinopsis rhodotricha* (Cactaceae) populations of Brazilian Chaco revealed by microsatellite markers 166

Conclusão Geral 189

Ecologia Molecular de Leguminosae e Cactaceae no Cerrado e Chaco: Fenologia, Filogeografia e Genética de Populações

No Brasil, o Cerrado e o Chaco, juntamente com a Caatinga, fazem parte da diagonal de formações vegetacionais abertas e dividem duas grandes florestas tropicais: a Amazônica e a Atlântica. Em geral, essas formações apresentam secas sazonais, com vegetação adaptada, distribuições complexas e, quando comparada às florestas tropicais úmidas, apresentam poucos estudos relacionados aos mais variados temas, como fertilidade do solo, biodiversidade, dinâmica populacional, estrutura genética e conservação.

Estudos que avaliem a distribuição geográfica da variabilidade genética de espécies dessas regiões abertas, sobretudo daquelas espécies filogeneticamente relacionadas que compartilham atributos ecológicos e traços comuns de história evolutiva, podem fornecer informações importantes sobre as respostas demográficas das populações adaptadas ao estresse hídrico durante oscilações climáticas passadas. Entretanto, por vezes a ampla distribuição geográfica das espécies dificulta uma correta avaliação dos aspectos ecológicos e evolutivos das populações, devido principalmente à dificuldade em se acessar as populações naturais ao longo de toda sua extensão. Uma saída é o uso de coleções biológicas, e em especial as exsicatas de plantas, pois permite a obtenção de amostras de uma grande variedade taxonômica e geográfica, sendo especialmente útil em estudos de espécies com ampla distribuição.

Para melhor compreender os processos relacionados a origem, evolução e manutenção da biodiversidade dessas formações abertas brasileiras, estudamos os padrões fenológicos (por meio de registros de herbários) e filogeográficos (aliados com modelagem de nicho ambiental) de duas espécies arbóreas do Cerrado (*Leptolobium dasycarpum* e *L. elegans*, Leguminosae - Papilionoideae), e também avaliamos a estrutura genética populacional de uma espécie de cacto (*Echinopsis rhodotricha*, Cactaceae — Trichocereeae) presente no Chaco brasileiro.

As duas espécies de Leguminosae apresentaram uma tendência para a floração e frutificação tardia com o aumento da chuva, porém, a temperatura parece ter influenciado apenas a floração de *L. dasycarpum*, levando à um aumento da data de floração.

As análises filogeográficas baseadas em regiões do DNA cloroplastidial e nuclear revelaram o compartilhamento de haplótipos entre *L. dasycarpum* e *L. elegans*. As populações são fortemente estruturadas, mas sem a formação de grupos geográficos diferenciados. As análises demográficas indicaram estabilidade populacional prolongada, corroborada pela modelagem de nicho ambiental das espécies.

Por meio dos marcadores microssatélites desenvolvidos para *Echinopsis rhodotricha* vimos que, na região do Chaco brasileiro, essa espécie apresenta baixa estruturação populacional e índices de diversidade genética semelhantes a outras espécies de Cactaceae. Além disso, a amplificação desses marcadores em outras espécies de Cactaceae foi bem-sucedida, sobretudo nos exemplares pertencentes à tribo Trichocereeae, contribuindo para a realização de estudos de estrutura populacional, diversidade genética e fluxo gênico em populações de outras espécies de cactos que ocorrem na região.

Molecular Ecology of Leguminosae and Cactaceae in the Cerrado and Chaco: Phenology, Phylogeography and Population Genetic

In Brazil, the Cerrado and Chaco, along with the Caatinga, are part of the diagonal of open formations and divide two large tropical forests: the Amazon and the Atlantic. In general, these formations have seasonal droughts, with adapted vegetation, complex distributions and, compared to tropical rainforests, have few studies related to various topics such as soil fertility, biodiversity, population dynamics, genetic structure and conservation.

Studies assessing the geographic distribution of genetic variability of species of open regions, especially those closed related species that share ecological attributes and common traits of evolutionary history, could provide important information on demographic responses of populations adapted to water stress during past climate fluctuations. However, sometimes the wide geographic distribution of species hampers a correct assessment of the ecological and evolutionary aspects of the population, mainly due to the difficulty in accessing natural populations along its entire distribution. One solution is the use of biological collections, especially the herbarium specimens, since it allows obtaining samples from a large taxonomic and geographical range, being especially useful in studies of species with broad distribution.

To better understand the processes related to the origin, evolution and maintenance of biodiversity of these Brazilian open formations, we studied the phenological patterns (using herbarium records) and phylogeographic (allied with environmental niche modeling) of two tree species of the Cerrado (*Leptolobium dasycarpum* and *L. elegans*, Leguminosae — Papilionoideae), and we also investigated the populational genetic structure of one species of cactus (*Echinopsis rhodotricha*, Cactaceae — Trichocereeae) present in the Brazilian Chaco.

The two Leguminosae species tended to have delayed flowering and fruiting with increasing of the rainfall, however, the temperature seems to have influenced only the flowering of *L. dasycarpum*, leading to an increase in flowering date.

The phylogeographic analyzes based on regions of the chloroplast and nuclear DNA revealed the haplotype sharing between *L. dasycarpum* and *L. elegans*. Populations are highly structured, but without the formation of different geographical groups. Demographic analysis showed prolonged population stability, supported by environmental niche modeling species.

Using microsatellite markers developed for *Echinopsis rhodotricha* we saw that in the Brazilian Chaco region, this species has low populational structure and genetic diversity rates similar to other Cactaceae species. Moreover, amplification of these markers in other Cactaceae species was successful, especially in specimens belonging to Trichocereeae tribe, contributing to the realization of population structure studies, genetic diversity and genetic flow in populations of other species of cactus occurring in the region.

Cerrado e Chaco: uma breve introdução

Na América do Sul, e em especial no Brasil, há uma diagonal de formações vegetacionais abertas que divide duas grandes florestas tropicais: a Amazônica e a Atlântica. Essa diagonal é composta por três formações: (i) as Florestas Tropicais Sazonalmente Secas (com a maior área no nordeste brasileiro — Caatinga), (ii) as savanas do Brasil central (Cerrado), e (iii) o Chaco, que ocorre em sua maioria na Argentina, Paraguai e Bolívia. No Brasil, o Chaco é representado por uma faixa estreita paralela ao Rio Paraguai, na cidade de Porto Murtinho, Mato Grosso do Sul (Figura 1). Esta região, que compreende ca. 20.000 km^2, possui aproximadamente 7,800 km^2 de vegetação natural, com cerca de 6.700 km^2 de área de sanava estépica (ZEE-MS 2015).

De maneira geral, a Caatinga e o Cerrado ocorrem sob as mesmas condições climáticas (Mooney et al. 1995, Mayle 2004), enquanto que o Chaco está frequentemente sujeito à invernos mais rigorosos (Pennington et al. 2000a). Em comum, essas três formações apresentam secas sazonais, com vegetação adaptada, biotas únicas, distribuições complexas (geralmente em forma de mosaico) e, quando comparada às florestas tropicais úmidas, apresentam poucos estudos relacionados à diversos temas, como por exemplo, fertilidade do solo, ciclos biogeoquímicos, biodiversidade, dinâmica populacional e conservação (Mooney et al. 1995, Furley & Metcalfe 2007). Entretanto, apesar de suas semelhanças, estas formações respondem diferentemente às mudanças climáticas e ambientais e devem ser considerados separadamente em estudos biogeográficos (Pennington et al. 2000a).

O Cerrado é a maior savana brasileira e ocupa aproximadamente 22% da área do país, correspondendo cerca de 2 milhões km^2. Atualmente é uma das regiões mais ameaçadas da América do Sul, devido à rápida expansão da agricultura, onde mais de 50% do Cerrado está convertido em áreas de pastagem e plantio de culturas (Oliveira & Marquis 2002), e apenas 2,2% está legalmente protegido (Klink & Machado 2005). Sendo assim, aumentar o

conhecimento sobre os padrões de distribuição da biodiversidade ao longo do Cerrado é essencial para identificar áreas prioritárias para proteção.

De maneira geral, a biodiversidade do Cerrado, incluindo todas as suas fisionomias, está estimada em 160.000 espécies de plantas, animais e fungos (Oliveira & Marquis 2002). O cerrado é considerado um “hotspot” de biodiversidade (Myers et al. 2000), com 4.400 espécies de plantas endêmicas, representando 1,5% de todas as espécies de plantas vasculares do mundo.

O Cerrado está circundado pela Floresta Amazônica ao norte, pela Mata Atlântica ao sul e sudeste e pela Caatinga ao nordeste (Figura 1). Também está incluída no Cerrado a parte ocidental não inundada do Pantanal (Motta et al. 2002). Desse modo, o Cerrado desempenha um papel ecológico importante, uma vez que ele serve de corredor para as espécies que ocorrem nas formações adjacentes (Oliveira & Marquis 2002).

Durante seu processo evolutivo, a distribuição do Cerrado sofreu expansões e contrações em resposta às flutuações climáticas. Durante os períodos mais secos, o Cerrado expandiu sobre áreas de floresta (Ab’Saber 1963) e durante os períodos mais úmidos as florestas expandiam sobre o Cerrado, exceto em áreas com deficiência de nutrientes ou que apresentavam déficit de água (Resende 1976).

O tempo de diversificação das linhagens lenhosas do Cerrado datam entre o Mioceno tardio e o Pleistoceno (9,8–0,4 Ma — milhões de anos atrás), sendo que a maioria das linhagens não possuem mais de 4 Ma (Simon et al. 2009). O registro polínico mais antigo de vegetação típica do Cerrado (e.g., gêneros *Byrsonima* Rich. ex Kunth — Malpighiaceae, *Schefflera* J.R. Forst. & G. Forst. — Araliaceae e *Curatella* Loefl. — Dilleniaceae) é de 32.000 ybp (anos antes do presente) estando localizado no Planalto Central Brasileiro. Já a vegetação que se assemelha às formações atuais não ocorreu ao norte do Brasil e na região central antes de 10.000 e 7.000 ybp, respectivamente. A presença desse tipo de vegetação em ambas as regiões é provavelmente uma consequência de um aumento progressivo na

sazonalidade, ao mesmo tempo que a temperatura também aumentava, e sugere uma dinâmica na história do Cerrado (Ledru 2002, Ledru et al. 2006).

A teoria dos refúgios pleistocênicos diz que as flutuações climáticas do Quaternário foram as principais responsáveis pelas altas taxas de especiação nos Neotrópicos por meio de ciclos sucessivos de expansão e contração da vegetação (Haffer 1969). Essa hipótese sugere que as savanas sul-americanas alcançaram sua máxima extensão para dentro da Bacia Amazônica durante o Último Máximo Glacial (c. 21 mil anos atrás), com as maiores conexões com as savanas do norte e do sul por meio de três principais corredores: (1) ao longo das encostas orientais dos Andes; (2) seguindo um corredor de baixa precipitação através da Amazônia Central; ou (3) ao longo da costa oriental da Floresta Atlântica (Sarmiento 1983, Webb 1991, Silva & Bates 2002). Contudo, a magnitude da influência dos ciclos do Quaternário na diversificação da biota neotropical ainda é bastante controversa (Moritz et al. 2000, Edwards et al. 2010, Hoorn et al. 2010). De maneira geral, a estabilidade (permanência de determinada formação vegetacional em um mesmo local) ao longo do Quaternário seria a responsável por uma alta diversidade genética e forte estrutura filogeográfica entre os refúgios ao contrário de regiões instáveis onde as populações teriam assinaturas genéticas marcadas pelas repetidas mudanças distribucionais (Hewitt 2004). Áreas recentemente colonizadas, por exemplo, em geral apresentam baixa diversidade genética e linhagens derivadas das regiões-fonte (refúgios).

A história biogeográfica da vegetação Neotropical é um produto de interações complexas entre processos históricos e biológicos (Burnham & Graham 1999), mas os padrões gerais de fauna e flora não estão bem estabelecidos, especialmente para as formações abertas. A integração de perspectivas biogeográficas e filogeográficas (baseadas em análises moleculares) fornecem dados que revelam padrões como espécies crípticas, linhagens distintas de espécies conhecidas e áreas de refúgios históricos (Moritz et al. 2009). Estes

padrões podem auxiliar na delimitação de espécies e conservação da biodiversidade dessas regiões (Moritz & Faith 1998, Davis et al. 2008, Riddle et al. 2008).

Vegetações savânicas são dominantes nos planaltos mais antigos do Cerrado, enquanto nas depressões, além de formações típicas do Cerrado, são encontradas comunidades de plantas mais heterogêneas, como os campos úmidos, florestas de galeria, manchas de florestas decíduas e semi-decíduas (Cole 1986, Silva 1997). Estima-se que, dentro do Cerrado, as formações abertas são mais antigas do que as florestas (Werneck 2011). Além disso, é possível que durante o Pleistoceno tardio tenha se formado uma área de refúgio de vegetação típica do Cerrado nos planaltos mais altos e espacialmente contínuos na região central de Goiás e da Chapada dos Guimarães em Mato Grosso (Ab'Saber 1983).

Distribuição espacial, riqueza de espécies, endemismos, diversidade genética e, consequentemente, padrões biogeográficos, podem ser fortemente moldados pela estabilidade do habitat (Graham et al. 2006, Carnaval & Moritz 2008). A modelagem da paleodistribuição surgiu como um método alternativo capaz de produzir modelos confiáveis da dinâmica do habitat em escala de tempo geológico recente (Pleistoceno tardio). Com esse método é possível identificar a presença e extensão de áreas de estabilidade (potenciais refúgios) através da combinação de modelos climáticos atuais e do passado, os quais podem ser validados por meio de evidências geológicas, paleo-ambientais (registro de pólen) e dados de diversidade genética (Hugall et al. 2002, Carnaval et al. 2009).

Métodos avançados em análise de dados moleculares, como padrões filogeográficos baseados em teoria da coalescência e em múltiplos, e independentes, marcadores moleculares são capazes de revelar complexas histórias evolutivas com altos níveis de migração, fluxo gênico e contato secundário entre populações de formações distintas (Avise 2009).

O Chaco compreende uma das poucas áreas no mundo onde a transição entre a zona tropical e a zona temperada não ocorre na forma de um deserto, mas sim como florestas

semiáridas e bosques (Morello 1967). Seu relevo é plano por toda a sua extensão e seus solos, compactos e com drenagem dificultada, são derivados da acumulação de sedimentos aluviais durante o Quaternário (Prado 1993). Há uma predominância de solos salinos, devido às intrusões oceânicas passadas nas planícies Chaco-Pampeanas (Pennington et al. 2000a), e o clima na região está sujeito a geadas no período seco, alagamentos e temperaturas do ar de até 49°C no período úmido (Pennington et al. 2004).

Árvores do gênero *Schinopsis* Engl. (Anacardiaceae), junto com *Aspidosperma quebracho-blanco* Schltdl. (Apocynaceae), *Tabebuia nodosa* (Griseb.) Griseb. (Bignoniaceae), e várias espécies de *Vachellia* Wight & Arn. (Fabaceae) e *Bulnesia* Gay (Zygophyllaceae) compõem as florestas do Chaco. Além disso, existe uma camada descontínua de arbustos, consistindo principalmente de espécies de Leguminosae (Mimosoideae) com espinhos, e de um estrato geralmente esparso de Bromeliaceae e Cactaceae, com poucas espécies de gramíneas (Pennington et al. 2000a).

No Brasil, somente as florestas próximas da cidade de Porto Murtinho, Mato Grosso do Sul, poderiam ser aceitas como vegetação chaquenha, uma vez que tanto suas espécies dominantes como as demais são elementos de Chaco sensu stricto (Prado et al. 1992). Entretanto, grande parte dessa área vem sendo desmatada para o plantio de pastagens, uma vez que a principal atividade econômica da região é a pecuária de corte (Silva et al. 2008, 2011). Além disso, tanto a porção brasileira do Chaco como outras áreas do Pantanal, não possuem Unidades de Conservação com tamanho suficiente para preservar as espécies que o Chaco abriga (Silva et al. 2011).

Grupo de estudo: Leguminosae

Leguminosae é a terceira maior família de Angiospermas (Mabberley 1997), com aproximadamente 730 gêneros e mais de 19.000 espécies (Lewis et al. 2005). No que se refere à importância econômica, as leguminosas perdem apenas para Poaceae, possuindo espécies utilizadas como alimento, óleo, fonte de fibras, combustível, madeira, fármacos e produtos químicos, além de variedades cultivadas na horticultura e enriquecimento do solo (Wojciechowski 2003).

Ecologicamente a família é importante em diversos ecossistemas, sobretudo os membros da subfamília Papilionoideae que estão presentes, e geralmente dominantes, em praticamente todo tipo de formação vegetacional da Terra, de florestas tropicais úmidas à desertos e tundras alpinas, e desempenham um papel vital na fixação de nitrogênio atmosférico (Sprent & McKey 1994, Sprent 2009).

A origem das primeiras leguminosas ocorreu no Paleoceno tardio, há cerca de 56 Ma (Herendeen 2001, Herendeen & Wing 2001, Wing et al. 2004). Todas as três subfamílias tradicionalmente reconhecidas — Caesalpinioideae, Mimosoideae e Papilionoideae — (Polhill et al. 1981) possuem registros fósseis por volta de 50 a 55 Ma (Herendeen & Dilcher 1992).

A subfamília Papilionoideae possui 440 gêneros e cerca de 12.000 espécies (Polhill 1981). Algumas linhagens, marcadas por flores não-papilionadas e com simetria radial, têm sido alvo de recentes estudos filogenéticos e moleculares (Lavin et al. 2001, Mansano et al. 2004, McMahon & Hufford 2004, Boatwright et al. 2008). Tais estudos vêm contradizendo a ideia original (e.g., Polhill 1981) de que flores radiais ou não-papilionadas definem grupos primitivos de Papilionoideae (e.g., Pennington et al. 2000b, Cardoso et al. 2012). Estes gêneros considerados mais basais estão reunidos nas tribos Swartzieae e Sophoreae. Swartzieae possui 15 gêneros principalmente neotropicais (Polhill 1994) e tem sido considerada como uma transição entre Caesalpinioideae e Papilionoideae (Polhill 1994,

Ireland et al. 2000). Já Sophoreae possui 40 gêneros mundialmente distribuídos. Ela foi descrita como uma tribo de conveniência (Polhill 1994), sendo caracterizada por flores relativamente simples com estames livres e folhas pinadas não especializadas.

Leptolobium Vogel (Papilionoideae: Sophoreae) compreende 13 espécies (uma ainda em fase de descrição) de árvores e arbustos (Rodrigues & Tozzi 2008, 2010, 2012), e possui uma taxonomia bem resolvida em dois sub-clados: (i) a sessão *Mesitis* (*Acosmium* sect. *Mesitis*, Yakovlev 1969) que possui duas espécies endêmicas do Brasil, sendo *Leptolobium bijugum* (Spreng.) Vogel encontrada na restinga e *L. brachystachyum* (Benth.) Sch. Rodr. & A.M.G. Azevedo nos campos rupestres ao norte de Minas Gerais (Cardoso et al. 2012); e (ii) a sessão *Leptolobium* (*Acosmium* sect. *Leptolobium*, Yakovlev 1969) que compreende 11 espécies, sendo seis restritas a florestas úmidas e o restante a savanas (Rodrigues & Tozzi 2010, 2012).

Análises filogenéticas (Figura 2) baseadas na morfologia e em marcadores de DNA nuclear (gene 5.8S do DNA ribossomal e o espaçador interno transcrito, ITS) e de cloroplasto (região codificante da enzima maturase kinase — matK e a região não-codificante do RNA mensageiro (intron) do gene trnL) revelaram que das cinco espécies de *Leptolobium* que ocorrem no Cerrado, três (*L. dasycarpum* Vogel, *L. elegans* Vogel e *L. parvifolium* (Harms) Sch. Rodr. & Azevedo) ficam em um clado composto predominantemente por espécies de florestas úmidas. Este padrão de espécies de savana serem derivadas de um ancestral que habitava florestas úmidas é um indicativo de evolução de nicho na origem das linhagens de savana de *Leptolobium* (Cardoso et al. 2012).

Leptolobium dasycarpum é uma arvore ou arvoreta (1–10 m de altura), com tronco geralmente tortuoso, com profundas fissuras longitudinais. Ocorre na Bolívia e no Brasil, no Cerrado, em diferentes fitofisionomias de savana e em áreas de transição a Caatinga (Figura 3). No Brasil é conhecida como romã-brava, pau-pratudo, perobinha, perobinha-de-chapada, unha-de-anta, amargozinho, amorgosa e cascudinho. É a espécie do gênero que apresenta a

distribuição mais ampla no Brasil: Bahia, Ceará, Distrito Federal, Goiás, Maranhão, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Piauí, Rondônia, São Paulo, Tocantins (Rodrigues & Tozzi 2012). *Leptolobium dasycarpum* é filogeneticamente próxima de *L. parvifolium* e *L. elegans*, e difere da última principalmente pelos pulvínulos menores, tubo do cálice maior, ovário tomentoso e por apresentar brácteas e bractéolas geralmente maiores.

Leptolobium elegans é uma arvore ou arvoreta (1,5–18 m de altura), tronco geralmente tortuoso, suberoso, com profundas fissuras longitudinais. Ocorre na Argentina, Paraguai e Brasil (Figura 3). No Cerrado pode ser encontrada em diferentes fitofisionomias de savana e em florestas baixas semideciduais. Conhecida no Brasil como amendoim-falso, chapadinha, sucupira-branca, genciana, cascudinho, perobinha-do-campo, leptolóbio, óleo-balsamo-do-campo e no Paraguai como incienco-del-campo (Rodrigues & Tozzi 2012). Tem propriedades medicinais (Mello et al. 2007) e é utilizada como fonte de madeira, como forrageira e apícola (Lorenzi 1992, Pott & Pott 1994). No Brasil ocorre nos Estados de Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, São Paulo e Paraná. No Pantanal, geralmente ocorre em fitofisionomias de Cerrado, de campos a savanas, em áreas pouco ou não alagáveis, ou em caronais. Foi descrita para mais de uma das sub-regiões pantaneiras, especialmente Nhecolândia e Paiaguás (Pott & Pott 1994). Esta espécie apresenta uma história taxonômica complexa, discutida detalhadamente por Rodrigues & Tozzi (2008). Foi descrita por Vogel (1837), que a considerou semelhante a *Leptolobium dasycarpum*, e reconheceu duas variedades: *L. elegans* var. *grandifolia* (espécimes pubescentes) e *L. elegans* var. *parvifolia* (espécimes glabros). Mohlenbrock (1963) as considerou duas espécies distintas, respectivamente *Sweetia elegans* e *Sweetia subelegans*. Yakovlev (1969) sinonimizou *L. elegans* (*S. elegans*) parte como *Acosmium dasycarpum*, parte como *Acosmium subelegans*. Posteriormente, *L. elegans* (*S. elegans*) deixou de ser sinônimo de *A. dasycarpum* (Bridgewater & Stirton 1997). Rodrigues & Tozzi (2008) reestabeleceram

Leptolobium elegans como nome da espécie, incluindo tanto indivíduos glabros quanto esparsamente pubescentes.

Grupo de estudo: Cactaceae

Os cactos estão entre as plantas mais características e conspícuas de regiões quentes e áridas do Novo Mundo, ocorrendo desde o sul da Argentina até o Canadá. Ocupam vários habitats, sobretudo os mais quentes e secos, mas também em florestas tropicais úmidas (Barthlott & Hunt 1993).

Os centros de diversidade dos cactos são as regiões áridas da América do Norte e do Sul, principalmente o sudoeste dos Estados Unidos e México, leste do Brasil e as encostas leste e oeste dos Andes (Nyffeler 2002). Enquanto no Brasil os cactos da região nordeste são bem conhecidos e estudados, o mesmo não ocorre para as demais regiões, onde se presume que a família Cactaceae apresenta poucos táxons nas áreas mais úmidas (Eggli 2002).

Cactaceae compreende cerca de 1.500–1.800 espécies em pouco mais de 100 gêneros (Barthllott & Hunt 1993) e é caracterizada principalmente pela presença de ramos curtos modificados em aréolas, meristema apical organizado em quatro zonas distintas e ovários "submersos" no receptáculo, que por sua vez são cobertos por brácteas e aréolas (Nyffeler 2002). Essa família é geralmente classificada em três subfamílias: Pereskioideae, Opuntioideae e Cactoideae (Barthlott & Hunt 1993). O gênero *Maihuenia* (Phil. ex F.A.C. Weber) Phil. ex K. Schum., tradicionalmente classificado na subfamília Pereskioideae, foi considerado como uma subfamília separada após análises moleculares (Anderson 2001).

Os domínios morfoclimáticos da Caatinga e do Chaco possuem uma alta densidade e diversidade de espécies de cactos, assim como as costas caribenhas da Colômbia e Venezuela (Hueck 1972). Domínios adjacentes, incluindo florestas, também contém populações isoladas de cactos principalmente em áreas rochosas ou associadas com vegetação em solo arenoso. Acredita-se que estas populações de cactos sejam um

remanescente da retração da vegetação xeromórfica durante os períodos interglaciais (Manfrin & Sene 2006). Ab'Saber (1977), em um estudo baseado em critérios climático, geomorfológicos, fitogeográficos e ecológicos, sugeriu que a Caatinga e o Chaco estiveram conectados durante o Pleistoceno. Contudo, baseado na composição florística, Prado & Gibbs (1993) sugeriram que estes dois domínios não foram conectados, pois o Chaco é derivado da vegetação seca do extremo sul da América do Sul.

Os cactos são extremamente diversos, porém com distribuição quase exclusivamente no Novo Mundo, sugerindo uma possível origem entre 90 e 65 Ma, o que permitiria um tempo máximo para diversificação e a separação espacial da África e América do Sul (Nyffeler 2002). Outros estudos sugerem uma origem mais recente devido à pouca divergência entre sequências moleculares entre os maiores clados de Cactaceae (Hershkovitz & Zimmer 1997, Nyffeler 2002). Além disso, não há registros fósseis relevantes para cactos ou seus parentes mais próximos, o que tem dificultado a investigação do tempo de divergência dentro do grupo (Ocampo & Columbus 2010).

A subfamília Cactoideae apresenta o maior número de espécies e as mais variadas formas de vida e habitats (Taylor 2000, Terrazas & Arias 2003), e está dividida em nove tribos, das quais Cereeae, Rhipsalideae, Trichocereeae, Pachycereeae e Hylocereeae ocorrem no Brasil (Taylor & Zappi 2004). Eggli (2002) destacou a falta de informação para Cactaceae no sudoeste do Brasil em comparação as demais regiões e elaborou uma lista com 33 espécies nativas para os estados de Mato Grosso e Mato Grosso do Sul. Destas espécies, 27 pertencem à subfamília Cactoideae, e seis gêneros à tribo Trichocereeae: *Arthrocereus* A. Berger, *Cleistocactus* Lem., *Discocactus* Pfeiff., *Echinopsis* Zucc., *Gymnocalycium* Pfeiff. ex Mittler e *Harrisia* Britton, sendo que todos ainda são pouco estudados, sobretudo no estado de Mato Grosso do Sul.

Echinopsis é um dos maiores e menos compreendidos gêneros de Trichocereeae, composto por 100–150 espécies distribuídas em várias regiões da América do Sul que

exibem uma ampla variedade de formas (Anderson 2001, Schlumpberger & Renner 2012). A filogenia molecular, feita com marcadores não-codificantes do DNA cloroplastidial (trnS–trnG, região trnL e intron rpl16) do grupo indica que *Echinopsis* não é monofilético e que o modo de polinização e crescimento (hábito) não foram características evolutivamente informativas. Este gênero ocorre desde o sudeste do Brasil até o litoral do Chile e o norte do Equador, com um centro de diversidade de espécies no leste dos Andes e norte da Argentina e Bolívia (Schlumpberger & Renner 2012).

No estado de Mato Grosso do Sul, este gênero é representado por *E. calochlora* K. Schum., endêmica das morrarias do Maciço do Urucum e Morraria do Mutum-Jacadigo, no município de Corumbá, *E. calochlora* spp. *glaetzleana* P.J. Braun & Esteves, encontrada na Serra da Boquena, região sudoeste do estado, e por *E. rhodotricha* K. Schum., que ocorre na região sudoeste em remanescentes de Chaco arborizado, no município de Porto Murtinho, estendendo-se pelo Paraguai, Uruguai e Argentina (Eggli 2002). A relação filogenética entre essas duas espécies e, mesmo entre *E. rhodotricha* e o restante do gênero *Echinopsis*, é desconhecida, uma vez que esta espécie não foi incluída na filogenia recém-publicada do grupo (Schlumpberger & Renner 2012).

Echinopsis rhodotricha possui grandes flores laterais, brancas, com antese noturna, de aroma agradável e que duram apenas uma noite (Eggli 2002). Essas características se encaixam em duas síndromes de polinização (Faegri & Van der Pijl 1971): falenofilia/esfingofilia (mariposas/esfingídeos) e cantarofilia (besouros). Entretanto, Gomes & Araujo (2015) constataram que as flores desta espécie permanecem abertas até a manhã do dia seguinte após a antese, o que favorece a participação de polinizadores noturnos e diurnos e sugere um sistema de polinização mais generalista. Não há dados sobre os sistemas reprodutivos de *E. rhodotricha*, entretanto, já foi observada auto-incompatibilidade em outras espécies do gênero (*E. chiloensis* (Colla) Friedrich & G.D. Rowley, Ossa & Medel

2011; *E. terscheckii* (Parm. ex Pfeiff.) Friedrich & G.D. Rowley, Ortega-Baes et al. 2011; *E. schickendantzii* F.A.C. Weber, Alonso-Pedano & Ortega-Baes 2012).

Uso de exsicatas para estudos ecológicos e genéticos

Estudos que visam compreender as relações filogenéticas e os níveis de diversidade genética usam várias técnicas para obtenção de dados moleculares, como o sequenciamento de DNA por exemplo. Em comum, os diferentes métodos requerem que a extração do DNA seja segura, rápida, relativamente barata e que produza um DNA de alta pureza e qualidade (Cota-Sánchez et al. 2006). Para tanto, tecido fresco ou congelado é ideal, mas diversos estudos têm demostrado que é possível obter DNA de qualidade a partir de espécimes de plantas preservados em herbários por mais de 100 anos (Soltis & Soltis 1993, Taylor & Swann 1994, Golenberg 1999). Coleções biológicas são facilmente acessadas e possuem amostras de uma grande variedade taxonômica, e muitas vezes representantes de espécies ameaçadas de extinção (Cota-Sánchez et al. 2006).

Particularmente na filogeografia de plantas, as exsicatas são usadas para acessar a totalidade da distribuição das espécies, especialmente daquelas com ocorrência em grandes áreas, como é o caso de *Acacia senegal* (L.) Willd. (Leguminosae, Mimosoideae; Odee et al. 2012), nas savanas da África subsaariana e de *Cordia alliodora* (Ruiz & Pav.) Oken (Boraginaceae; Rymer et al. 2013) nos Neotrópicos. Além disso, os espécimes de herbário se mostraram úteis na compreensão de como fatores históricos mediados pelo homem (e.g., colonização da América do Norte) influenciaram a filogeografia de *Ambrosia artemisiifolia* L. (Asteraceae; Martin et al. 2014), uma espécie invasora vinda da Europa.

Estima-se que nos 3.400 herbários do mundo há cerca de 350 milhões de espécimes (Thiers 2015). Estas coleções contém a maioria das espécies de plantas descritas, incluindo aquelas endêmicas com distribuição restrita, espécies ainda não descritas e aquelas já extintas, ou de populações impactadas pelo homem (Bebber et al. 2010, Joppa et al. 2011).

Apesar do grande número de amostras nos herbários e do crescente número de estudos com esse tipo de material, apenas uma pequena porção destes espécimes tem sido usada para pesquisas moleculares. Isso se deve ao fato de que o sucesso de extração e amplificação de DNA de herbário (hDNA) muitas vezes é baixo. Desse modo, os métodos para utilização de hDNA devem ser aprimorados, sobretudo quando é mais rápido e barato obter amostras de herbário de várias localidades (principalmente quando se trabalha com espécies de ampla distribuição) do que de populações naturais (Särkinen et al. 2012), especialmente quando as espécies estão se tornando extintas ou extremamente raras na natureza.

Espécimes de herbário também apresentam informações sobre a fenologia das espécies e que podem ser usadas para detectar mudanças na periodicidade de eventos cíclicos, como a floração e frutificação (Gallagher et al. 2009). Como essas mudanças geralmente estão associadas aos fatores ambientais (e.g., temperatura e precipitação), mudanças na fenologia das espécies estão sendo utilizadas para analisar o impacto das mudanças climáticas sobre plantas e animais (Lavoie & Lachance 2006).

A maioria dos trabalhos analisam as variações fenológicas de populações de plantas restritas à um determinado local e em um curto espaço de tempo, principalmente devido ao alto custo logístico para se amostrar toda a distribuição de uma espécie por longos períodos (Primack et al. 2004, Bolmgren & Lönnberg 2005). Desse modo, a falta de observações fenológicas de longa duração pode ser suprida com o uso de exsicatas, pois para muitas espécies os herbários possuem um grande número de espécimes coletados ao longo de toda sua distribuição (Lavoie & Lachance 2006).

O número de trabalhos mostrando que os padrões fenológicos oriundos de exsicatas coincidem com os dados de coleta de campo têm aumentado significativamente (Primack et al. 2004, Gallagher et al. 2009, Gaira et al. 2011, Zalamea et al. 2011, Diskin et al. 2012, Lavoie 2013), permitindo então que sejam feitas análises mais abrangentes do comportamento fenológico de espécies e de comunidades. Além disso, a análise fenológica

em larga escala (e.g., continental) permite uma investigação da dinâmica das espécies e sua interação com as mudanças climáticas (Zalamea et al. 2011). Essas mudanças podem alterar a interação entre espécies, como a assincronia entre a oferta de recursos florais e a presença de polinizadores e dispersores (Visser & Both 2005, Hegland et al. 2009), possivelmente levando a um aumento do risco de extinção e perda de serviços ecossistêmicos (Both et al. 2006, Memmott et al. 2007).

Filogeografia e marcadores moleculares

A filogeografia tem sido utilizada para examinar a distribuição geográfica da variabilidade genética de populações de uma única ou de várias espécies (Avise 1998, 2009). Novos métodos estatísticos têm ganhado força entre os filogeógrafos ao permitir que sejam feitas estimativas da história demográfica das populações e dos processos microevolutivos que governam essa estruturação (Knowles & Maddison 2002, Knowles 2009). Nesse contexto, os métodos baseados na teoria da coalescência garantem um arcabouço metodológico para testar hipóteses sobre a história demográfica de populações, assim como o tempo e os padrões de divergência (Kuhner 2008).

A teoria da coalescência, desenvolvida por Kingman (1982), é uma ferramenta de modelagem poderosa para a genética de populações e filogeografia. Os estados dos alelos de todas as cópias de genes homólogos em uma população são determinados pela história genealógica e mutacional destas cópias. A abordagem coalescente baseia-se na constatação de que a árvore genealógica é geralmente mais fácil de modelar para trás no tempo, e de que as mutações seletivamente neutras podem ser sobrepostas posteriormente (Norborg 2007). Desse modo, a teoria da coalescência contribuiu com a fusão conceitual e analítica das áreas da genética de populações e da filogenética, sendo a árvore de genes o foco de estudo em ambas as áreas. No entanto, a filogenia tradicional busca estimar uma árvore e usá-la para deduzir as relações evolutivas, enquanto a genética de populações tradicional observa a

árvore como um resultado aleatório de um processo genético populacional (Nielsen & Beaumont 2009).

Muitos padrões filogeográficos são explicados pelas alterações paleoclimáticas ocorridas no Pleistosceno. Em muitas espécies de árvores das regiões tropicais observa-se uma forte estruturação filogeográfica, provavelmente relacionadas às contrações florestais ocorridas nos períodos glaciais, tanto em populações de formações florestais (Dauby et al. 2010, Koffi et al. 2011) como de formações savânicas (Allal et al. 2011). Em contraste, padrões filogeográficos de espécies de transição floresta-savana ainda são escassos (Dupas et al. 2014).

Devido à dificuldade logística na amostragem de populações naturais, muitos estudos em filogeografia ficam limitados a um número relativamente pequeno de amostras com uma distribuição muito restrita (Barbosa et al. 2015). Além do poder estatístico ser menor quando se tem uma amostragem pequena, as estimativas de padrões geográficos podem ser fortemente enviesadas, levando a uma interpretação errada dos processos ecológicos e evolutivos responsáveis por tais padrões e, consequentemente, a uma estratégia de conservação ou manejo ineficiente (Blakney et al. 2014, Pan et al. 2014, Zhihao & Mingli 2014).

Assim como os métodos baseados em coalescência, a modelagem de distribuição de espécies também tem sido usada para se tentar explicar os processos de estruturação populacional. A modelagem de nicho ecológico (MNE; Warren 2012) inclui diferentes métodos capazes de identificar o nicho ambiental e a distribuição potencial de espécies e/ou comunidades (Ferrier & Guisan 2006, Svenning et al. 2011, Werneck et al. 2012a). Isso é feito por meio de associações estatísticas entre as localidades onde uma espécie (ou um conjunto de espécies) vive e as condições ambientais desses lugares, ou também através do cálculo das tolerâncias fisiológicas de uma espécie baseada em princípios biofísicos (Kearney & Porter 2009).

Como a MNE gera informações acerca das preferências e tolerâncias de uma espécie e, por sua vez, é capaz de estimar a distribuição potencial atual, do passado (e.g., no último período glacial) e também do futuro (Scoble & Lowe 2010, Chan et al. 2011, Svenning et al. 2011), ela se torna uma ferramenta importante para se entender os processos que estruturam a variabilidade genética de espécies e populações através de diferentes paisagens (Knowles 2009, Chan et al. 2011).

A MNE é uma fonte independente de dados que pode ser usada para avaliar ou desenvolver hipóteses filogeográficas (Richards et al. 2007, Alvarado-Serrano & Knowles 2014) tais como: identificar locais potenciais de populações antigas (e.g., Swenson 2006, Knowles et al. 2007, Morgan et al. 2011), caracterizar as preferências e tolerâncias ambientais de espécies (e.g., Stockman & Bond 2007, Wooten et al. 2010), avaliar a adaptação de populações às condições ambientais locais (e.g., Fournier-Level et al. 2011, Banta et al. 2012) e testar se a divergência de nicho acompanha a divergência de espécies (e.g., McCormack et al. 2010, Kalkvik et al. 2012).

Recentemente, Alvarado-Serrano & Knowles (2014) revisaram os diferentes usos da MNE em estudos filogeográficos destacando as vantagens e limitações dos modelos utilizados. Dentre os principais usos da MNE destacados pelos autores, o mais utilizado em conjunto com análises filogeográficas é a sobreposição dos padrões genéticos encontrados com a distribuição potencial das espécies, principalmente àqueles modelos do passado (último máximo glacial ou último período interglacial).

Por serem reconhecidos como simplificações de uma realidade complexa e difícil de ser compreendida, os resultados dos modelos devem ser interpretados e aplicados com cautela, considerando suas limitações, como por exemplo, a incapacidade de modelar interações bióticas e os padrões de dispersão (Tôrres & Vercillo 2012). De toda forma, a modelagem de distribuição de espécies é considerada uma das ferramentas mais avançadas

para estudos de distribuição em larga escala (Franklin 2010), sendo úteis para orientar tomadas de decisão e implementação de medidas de gestão por parte do governo.

Estudos de filogeografia de espécies da América do Sul têm contribuído para um melhor entendimento dos processos que geraram a diversificação das espécies nas regiões tropicais (Turchetto-Zolet et al. 2013). Esses trabalhos buscam elucidar os processos, tanto ao nível de espécies como populações, responsáveis pela diversidade genética, assim como os modos de dispersão, tempo de divergência, extinções, área de refúgio, entre outros. Além disso, a filogeografia comparativa tem sido útil para se entender como as espécies e/ou populações de uma mesma formação/região geográfica responderam aos eventos climáticos e geomorfológicos do passado (Ramos et al. 2009).

Embora a biodiversidade da América do Sul esteja entre as maiores do mundo (Beheregaray 2008, Sérsic et al. 2011) o número de estudos filogeográficos ainda não condiz com essa alta diversidade. Contudo, segundo revisão de Turchetto-Zolet et al. (2013) estão surgindo cada vez mais trabalhos nessa área e, em 2012, esse número chegou a 214 artigos. Entre os trabalhos com plantas (um total de 36 artigos) as Angiospermas constituem a maioria (34 artigos) e as famílias com maior representatividade são Leguminosae (17%), Nothofagaceae (12%) e Asteraceae (9%). Esse viés taxonômico pode estar relacionado ao fator de não haver marcadores moleculares específicos para estudos populacionais, principalmente para famílias de plantas que não possuem um apelo comercial ou científico evidente.

Estudos filogeográficos podem ser baseados em informações do DNA nuclear, mitocondrial e cloroplastidial. Entretanto, os marcadores que são herdados uniparentalmente (como os de cloroplasto em plantas) são preferidos porque são mais propensos a reter informações sobre histórias de migração do que os marcadores nucleares (Weising et al. 2005). Genomas plastidiais possuem certo número de propriedades peculiares que os distinguem dos genomas nucleares: (i) há apenas um cromossomo circular presente em

cópias múltiplas, em vez de vários cromossomas lineares; (ii) o genoma é haplóide e, portanto, tem um tamanho efetivo de população menor, resultando em forte diferenciação de populações fragmentadas por deriva genética; (iii) normalmente não há recombinação e polimorfismos de sequência individual podem ser combinados em haplótipos; e (iv) o genoma possui herança materna (Weising et al. 2005). Contudo, uso de apenas um tipo de marcador para inferir relações filogeográficas ou filogenéticas em algumas espécies pode não ser suficiente (Avise & Wollenberg 1997), correndo o risco de não se representar as corretas relações entre os táxons (Edwards & Bensch 2009, Toews & Brelsford 2012).

Desde 1987, ano em que o termo filogeografia foi cunhado por Avise et al. (1987), até março de 2015, 12 artigos foram publicados com enfoque em filogeografia de espécies arbóreas do Cerrado, sendo a família Leguminosae a mais representativa, com quatro estudos (Tabela 1). Três dos trabalhos com leguminosas do Cerrado utilizaram apenas marcadores plastidiais em suas análises, sendo que apenas um utilizarou uma combinação de marcadores plastidiais e nucleares (Tabela 1). Em comum, as espécies apresentaram sinais genéticos de expansão recente a partir de regiões centrais e ao leste do Cerrado, principalmente a partir dos estados de Goiás e Minas Gerais. O Mato Grosso do Sul, que possui 61% (ca. 217.000 km^2 de área) de seu território formado pelo Cerrado (Sano et al. 2007), além de áreas de transição Cerrado-Pantanal e Cerrado-Chaco, é pouco amostrado ou ausente nestes estudos.

Como citado anteriormente, ainda há uma escassez de estudos com abordagem filogeográfica ou molecular com espécies de cactos. Até março de 2015, 12 artigos foram publicados com enfoque molecular, sendo que a maioria (cerca de 42%) tiveram um enfoque em genética de populações ou filogeografia (Tabela 2). Desses trabalhos, 71% eram de espécies brasileiras, mas nenhum analisou a diversidade genética ou padrões filogeográficos de espécies ocorrentes no Chaco.

Com o objetivo de melhor compreender os processos relacionados a origem, evolução e manutenção da biodiversidade dessas formações abertas brasileiras, estudamos a

ecologia molecular de duas espécies arbóreas do Cerrado (*Leptolobium dasycarpum* e *L. elegans*, Leguminosae — Papilionoideae) por meio da análise dos padrões fenológicos e filogeográficos aliados com a modelagem de nicho ambiental. Além disso, avaliamos a estrutura genética populacional de *Echinopsis rhodotricha* (Cactaceae, Trichocereeae) que ocorre na estreita porção do Chaco brasileiro.

Literatura citada

Ab'Saber AN. 1963. Contribuição a geomorfologia da área dos cerrados. In: Simpósio sobre o Cerrado (Ed. MG Ferri), pp. 119–128. Editora da Universidade de São Paulo, São Paulo.

Ab'Saber AN. 1977. Os domínios morfoclimáticos da América do Sul. Geomorfologia, Universidade de São Paulo 52: 1–22.

Ab'Saber AN. 1983. O domínio dos cerrados: introdução ao conhecimento. Revista do Serviço Público 111: 41–55.

Allal F, Sanou H, Millet L, Vaillant A, Camus-Kulandaivelu L, Logossa ZA, Lefèvre F & Bouvet J-M. 2011. Past climate changes explain the phylogeography of *Vitellaria paradoxa* over Africa. Heredity 107: 174–186.

Alonso-Pedano M & Ortega-Baes P. 2012. Generalized and complementary pollination system in the Andean cactus *Echinopsis schickendantzii*. Plant Systematics and Evolution 298: 1671–1677.

Alvarado-Serrano DF & Knowles LL. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. Molecular Ecology Resources 14: 233–248.

Anderson EF. 2001. The cactus family. Timber Press, Portland, Oregon, USA, 777 pp.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489–522.

Avise JC & Wollenberg K. 1997. Phylogenetics and the origin of species. Proceedings of the National Academy of Sciences of the United States of America 94: 7748–7755.

Avise JC. 1998. The history and purview of phylogeography: a personal reflection. Molecular Ecology 7: 371–379.

Avise JC 2009. Phylogeography: retrospect and prospect. Journal of Biogeography 36: 3–15.

Banta JA, Ehrenreich IM, Gerard S, Chou L, Wilczek A, Schmitt J, Kover PX & Purugganan MD. 2012. Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in *Arabidopsis thaliana*. Ecology Letters 15: 769–777.

Barbosa ACDOF, Collevatti RG, Chaves LJ, Guedes LBS, Diniz-Filho JAF & de Campos Telles MP. 2015. Range-wide genetic differentiation of *Eugenia dysenterica* (Myrtaceae) populations in Cerrado. Biochemical Systematics and Ecology 59 288–296.

Barthlott W & Hunt DR. 1993. Cactaceae. In: The families and genera of vascular plants, vol. 2 (Eds. K Kubitzki, JG Rohwer & V Bittrich), pp. 161–197. Springer Verlag, Berlin, Germany.

Bebber DP, Carine MA, Wood JRI, Wortley AH, Harris DJ, Prance GT, Davidse G, Paige J, Pennington TD, Robson NKB & Scotland RW. 2010. Herbaria are a major frontier for species discovery. Proceedings of the National Academy of Sciences of the United States of America 107: 22169–22171.

Beheregaray LB. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Molecular Ecology 17: 3754–3774.

Blakney JR, Loxterman JL & Keeley ER. 2014. Range-wide comparisons of northern leatherside chub populations reveal historical and contemporary patterns of genetic variation. Conservation Genetics 15: 757–770.

Boatwright JS, Savolainen V, Van Wyk, B-E, Schutte-Vlok, AL, Forest F & van der Bank M. 2008. Systematic position of the anomalous genus *Cadia* and the phylogeny of the tribe Podalyrieae (Fabaceae). Systematic Botany 33: 133–147.

Bolmgren K & Lönnberg K. 2005. Herbarium data reveal an association between fleshy fruit type and earlier flowering time. International Journal of Plant Sciences 166: 663–670.

Bonatelli IAS, Zappi DC, Taylor NP & Moraes EM. 2013. Usefulness of cpDNA markers for phylogenetic and phylogeographic analyses of closely-related cactus species. Genetics and Molecular Research 12: 4579–4585.

Bonatelli IAS, Perez MF, Peterson AT, Taylor NP, Zappi DC, Machado MC, Koch I, Pires AH & Moraes EM. 2014. Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. Molecular Ecology 23: 3044–3063.

Both C, Bouwhuis S, Lessells CM & Visser ME. 2006 Climate change and population declines in a long-distance migratory bird. Nature 441: 81– 83.

Bridgewater SGM & Stirton CH. 1997. A morphological and biogeographic study of the *Acosmium dasycarpum* complex (Leguminosae: Papilionoideae: Sophoreae). Kew Bulletin 52: 471–475.

Burnham RJ & Graham A. 1999. The history of Neotropical vegetation: new developments and status. Annals of the Missouri Botanical Garden 86: 546–589.

Caetano S, Prado D, Pennington RT, 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: 3147–3159.

Cardoso DBOS, Lima HC, Rodrigues RS, Queiroz LP, Pennington RT & Lavin M. 2012. The *Bowdichia* clade of Genistoid legumes: Phylogenetic analysis of combined molecular and morphological data and a recircumscription of *Diplotropis*. Taxon 61: 1074–1087.

Carnaval AC & Moritz C, 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. Journal of Biogeography 35: 1187–1201.

Carnaval AC, Hickerson MJ, Haddad CF, Rodrigues MT & Moritz C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science 323: 785–789.

Chan LM, Brown JL & Yoder AD. 2011. Integrating statistical genetic and geospatial methods brings new power to phylogeography. Molecular Phylogenetics and Evolution 59: 523–537.

Cole M. 1986. The Savannas: Biogeography and Geobotany. Academic Press, London, 438 pp.

Collevatti RG, Grattapaglia D & Hay JD. 2003. Evidences for multiple maternal lineages of *Caryocar brasiliense* populations in the Cerrado based on the analysis of chloroplast DNA sequences and microsatellite haplotype variation. Molecular Ecology 12: 105–115.

Collevatti RG, Terribile LC, Lima-Ribeiro MS, Nabout JC, Oliveira G, Rangel TF, Rabelo SG & Diniz-Filho JA. 2012a. A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species. Molecular Ecology 21: 5845–5863.

Collevatti RG, Lima-Ribeiro MS, Souza-Neto AC, Franco AA & Terribile LV. 2012b. Recovering the demographical history of a Cerrado tree species *Caryocar brasiliense*: coupling ecological niche modeling and coalescent analyses. Natureza Conservação 10: 169–176.

Collevatti RG, de Castro TG, de Souza Lima J & de Campos Telles MP. 2012c. Phylogeography of *Tibouchina papyrus* (Pohl) Toledo (Melastomataceae), an endangered tree species from rocky savannas, suggests bidirectional expansion due to climate cooling in the Pleistocene. Ecology and Evolution 2: 1024–1035.

Collevatti RG, Lima-Ribeiro MS, Terribile LC, Guedes LB, Rosa FF & Telles MP. 2014. Recovering species demographic history from multi-model inference: the case of a Neotropical savanna tree species. BMC Evolutionary Biology 14: 213.

Cornejo-Romero A, Medina-Sánchez J, Hernández-Hernández T, Rendón-Aguilar B, Valverde PL, Zavala-Hurtado A, Rivas-Arancibia SP, Pérez-Hernández MA, López-Ortega G, Jiménez-Sierra & Vargas-Mendoza CF. 2014. Quaternary origin and genetic divergence of the endemic cactus *Mammillaria pectinifera* in a changing landscape in the Tehuacán Valley, Mexico. Genetics and Molecular Research 13: 73–88.

Cota-Sánchez JH, Remarchuk K & Ubayasena, K. 2006. Ready-to-use DNA extracted with a ctab method adapted for herbarium specimens and mucilaginous plant tissue. Plant Molecular Biology Reporter 24: 161–167.

Dauby G, Duminil J, Heuertz M & Hardy OJ. 2010. Chloroplast DNA polymorphism and phylogeography of a Central African tree species widespread in mature rainforests: *Greenwayodendron suaveolens* (Annonaceae). Tropical Plant Biology 3: 4–13.

Davis EB, Koo MS, Conroy C, Patton JL & Moritz C. 2008. The California Hotspots Project: identifying regions of rapid diversification of mammals. Molecular Ecology 17: 120–138.

Diskin E, Proctor H, Jebb M, Sparks T & Donnelly A. 2012. The phenology of *Rubus fruticosusin* Ireland: herbarium specimens provide evidence for the response of phenophases to temperature, with implications for climate warming. International Journal of Biometeorology 56:1103–1111.

Dupas S, Ru B, Branca A, Faure N, Gigot G, Campagne P, Sezonlin M, Ndemah R, Ong'amo G, Calatayud PA & Silvain JF. 2014. Phylogeography in continuous space: coupling species

distribution models and circuit theory to assess the effect of contiguous migration at different climatic periods on genetic differentiation in *Busseola fusca* (Lepidoptera: Noctuidae). Molecular Ecology 23: 2313–2325.

Edwards SV & Bensch S. 2009. Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. Molecular Ecology 18: 2930–2933.

Edwards EJ, Osborne CP, Strömberg CA, Smith SA, Bond WJ, Christin PA, Cousins AB, Duvall MR, Fox DL, Freckleton RP, Ghannoum O, Hartwell J, Huang Y, Janis CM, Keeley JE, Kellogg EA, Knapp AK, Leakey AD, Nelson DM, Saarela JM, Sage RF, Sala OE, Salamin N, Still CJ & Tipple B. 2010. The origins of C_4 grasslands: integrating evolutionary and ecosystem science. Science 328: 587–591.

Eggli U. 2002. Synopsis of the Cactaceae of Mato Grosso, Brazil. Haseltonia 9: 146–166.

Faegri K &. van der Pijl EL. 1971. The principles of Pollination Ecology. 2ª ed. Pergamon Press, Oxford, Oxfordshire, Reino Unido, 291 pp.

Faria-Tavares JS, Martin PG, Mangolin CA, de Oliveira-Collet SA & Machado MFPS. 2013. Genetic relationships among accessions of mandacaru (*Cereus* spp.: Cactaceae) using amplified fragment length polymorphisms (AFLP). Biochemical Systematics and Ecology 48: 12–19.

Ferrier S & Guisan A. 2006. Spatial modelling of biodiversity at the community level. Journal of Applied Ecology 43: 393–404.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J & Wilczek AM. 2011. A map of local adaptation in *Arabidopsis thaliana*. Science 334: 86–89.

Franck AR, Cochrane BJ & Garey JR. 2013. Phylogeny, biogeography, and infrageneric classification of *Harrisia* (Cactaceae). Systematic Botany 38: 210–223.

Franklin J. 2010. Mapping Species distributions: spatial inference and prediction. Cambridge University Press, Cambridge, 340 pp.

Furley PA & Metcalfe SE. 2007. Dynamic changes in savanna and seasonally dry vegetation through time. Progress in Physical Geography 31: 633–642.

Gaira KS, Dhar U & Belwal OK. 2011. Potential of herbarium records to sequence phonological pattern: a case study of *Aconitum heterophyllum* in the Himalaya. Biodiversity and Conservation 20: 2201–2210.

Gallagher RV, Hughes L & Leishman MR. 2009. Phenological trends among Australian alpine species: using herbarium records to identify climate-change indicators. Australian Journal of Botany 57: 1–9.

Garcia MG, Silva RS, Carniello MA, Veldman JW, Rossi AAB & de Oliveira LO. 2011. Molecular evidence of cryptic speciation, historical range expansion, and recent intraspecific hybridization in the Neotropical seasonal forest tree *Cedrela fissilis* (Meliaceae). Molecular Phylogenetics and Evolution 61: 639–649.

Golenberg EM. 1999. Isolation, identification, and authentication of DNA sequences derived from fossil material. In: Fossil Plants and Spores: Modern Techniques (Eds TP Jones & NP Rowe), pp 156–160. The Geological Society, London.

Gomes VG & Araujo AC. 2015. Cacti species from the Brazilian Chaco: floral and fruit traits. Gaia Scientia 9: 1–8.

Graham CH, Moritz C & Williams SE. 2006. Habitat history improves prediction of biodiversity in rainforest fauna. Proceedings of the National Academy of Sciences of the United States of America 103: 632–636.

Haffer J. 1969. Speciation in Amazonian forest birds. Science 168: 131–137.

Hegland SJ, Nielsen A, Lázaro A, Bjerknes A-L & Totland Ø. 2009. How does climate warming affect plant-pollinator interactions? Ecology Letters 12: 184–195.

Herendeen PS & Dilcher DL. 1992. Advances in legume systematics, Part 4. The fossil record. Royal Botanic Gardens, Kew, UK, 336 pp.

Herendeen PS. 2001. The fossil record of the Leguminosae: Recent advances. In: Legumes down under: The Fourth International Legume Conference, Abstracts, pp. 34–35. Australian National University, Canberra, Australia.

Herendeen PS & Wing S. 2001. Papilionoid legume fruits and leaves from the Paleocene of northwestern Wyoming. Botany 2001 Abstracts, published by Botanical Society of America (http://www.botany2001.org/).

Hershkovitz MA & Zimmer EA. 1997. On the evolutionary origins of the cacti. Taxon 46: 217–232.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical transactions of the Royal Society of London B 359: 183–195

Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T & Antonelli A. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330: 927–931.

Hueck K, 1972. As florestas da América do Sul: ecologia, composição e importância econômica. São Paulo: Polígono, 466 pp.

Hugall A, Moritz C, Moussall A & Stanisic J. 2002. Reconciling paleodistribution models and comparative phylogeography in the wet tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). Proceedings of the National Academy of Sciences of the United States of America 99: 6112–6117.

Ireland H, Pennington RT & Preston J. 2000. Molecular systematics of the Swartzieae. In: Advances in legume systematics: Part 9 (Eds. PS Hereendeen & A Bruneau), pp. 217–231. Royal Botanic Gardens, Kew.

Joppa LN, Roberts DL & Pimm SL. 2011. How many species of flowering plants are there? Proceedings of the Royal Society B: Biological Sciences 278: 554–559.

Junqueira KP, Faleiro FG, Junqueira NTV, Bellon G, Lima CD & Souza LD. 2010. Diversidade genética de pitayas nativas do cerrado com base em marcadores RAPD. Revista Brasileira de Fruticultura 32: 819–824.

Kalkvik HM, Stout IJ, Doonan TJ & Parkinson CL. 2012. Investigating niche and lineage diversification in widely distributed taxa: phylogeography and ecological niche modeling of the *Peromyscus maniculatus* species group. Ecography 35: 54–64.

Kearney M & Porter W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecology Letters 12: 334–350.

Kingman JFC. 1982. The coalescent. Stochastic Processes and their Applications 13: 235–248.

Klink CA & Machado RB. 2005 Conservation of the Cerrado. Conservation Biology 19: 707–713.

Knowles LL & Maddison WP.2002. Statistical phylogeography. Molecular Ecology 11: 2623–2635.

Knowles LL, Carstens BC & Keat ML. 2007. Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. Current Biology 17: 940–946.

Knowles LL. 2009. Statistical phylogeography. Annual Review of Ecology, Evolution, and Systematics 40: 593–612.

Koffi KG, Hardy OJ, Doumenge C, Cruaud C & Heuertz M. 2011. Diversity gradients and phylogeographic patterns in *Santiria trimera* (Burseraceae), a widespread African tree typical of mature rainforests. American Journal of Botany 98: 254–264.

Kuhner MK. 2008. Coalescent genealogy samplers: windows into population history. Trends in Ecology and Evolution 24: 86–93.

Lavin M, Pennington RT, Klitgaard BB, Sprent JI, Lima HC & Gasson P.E. 2001. The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. American Journal of Botany 88: 503–533.

Lavoie C & Lachance D. 2006. A new herbarium-based method for reconstructing the phenology of plant species across large areas. American Journal of Botany 93: 512–516.

Lavoie C. 2013. Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. Perspectives in Plant Ecology, Evolution and Systematics 15: 68–76.

Ledru M-P. 2002. Late Quaternary history and evolution of the Cerrados as Revealed by Palynological records. In: The Cerrados of Brazil. Ecology and natural history of a neotropical savanna (Eds. PS Oliveira & RJ Marquis), pp. 33–50. Columbia University Press, New York.

Ledru M-P, Ceccantini G, Gouveia SEM, López-Sáez JA, Pessenda LCR & Ribeiro AS. 2006. Millenial-scale climatic and vegetation changes in a northern Cerrado (Northeast, Brazil) since the Last Glacial Maximum. Quaternary Science Reviews 25: 1110–1126.

Lewis GP, Schrire B, Mackinder B & Lock M. 2005. Legumes of the World. Royal Botanic Gardens, Kew, pp. 592.

Lorenzi H. 1992. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Nova Odessa: Editora Plantarum, pp. 352.

Mabberley DJ. 1997. The Plant-Book. A portable dictionary of the vascular plants. 2ª ed. Cambridge University Press, Cambridge, New York, 858 pp.

Majure LC, Judd WS, Soltis PS & Soltis DE. 2012. Cytogeography of the *Humifusa* clade of *Opuntia* ss Mill. 1754 (Cactaceae, Opuntioideae, Opuntieae): correlations with pleistocene refugia and morphological traits in a polyploid complex. Comparative Cytogenetics 6: 53–77.

Manfrin MH & Sene FM. 2006. Cactophilic *Drosophila* in South America: a model for evolutionary studies. Genetica 126: 57–75.

Mansano VF, Bittrich V, Tozzi AMGA & Souza AP. 2004. Composition of the *Lecointea* clade (Leguminosae, Papilionoideae, Swartzieae), a re-evaluation based on combined evidence from morphology and molecular data. Taxon 53: 1007–1018.

Martin MD, Zimmer EA, Olsen MT, Foote AD, Gilbert MTP & Brush GS. 2014. Herbarium specimens reveal a historical shift in phylogeographic structure of common ragweed during native range disturbance. Molecular Ecology 23: 1701–1716.

Mayle FE. 2004. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. Journal of Quaternary Science 19: 713–720.

McCormack JE, Zellmer AJ & Knowles LL. 2010. Does niche divergence accompany allopatric divergence in *Aphelcoma* jays as predicted under ecological speciation? Insights from tests with niche models. Evolution 64: 1231–1244.

McMahon M & Hufford L. 2004. Phylogeny of *Amorpheae* (Fabaceae: Papilionoideae). American Journal of Botany 91: 1219–1230.

Mello FB, Langeloh A & Mello JRB. 2007 Toxicidade Pré-Clínica de Fitoterápico Contendo *Passiflora alata*, *Erythrina mulungu*, *Leptolobium elegans* e *Adonis vernalis*. Latin American Journal of Pharmacy 26: 191–200.

Memmott J, Craze PG, Waser NM & Price MV. 2007 Global warming and the disruption of plant-pollinator interactions. Ecology Letters 10: 710–717.

Mohlenbrock RH. 1963. A revision of the leguminous genus *Sweetia*. Webbia 17: 223–263.

Mooney HA, Bullock SH & Medina E. 1995. Introduction. In: Seasonally Dry Tropical Forests (Eds. SH Bullock, HA Mooney & E Medina), pp. 1–8. Cambridge University Press, Cambridge.

Moraes EM, Abreu AG, Andrade SC, Sene FM & Solferini VN. 2005. Population genetic structure of two columnar cacti with a patchy distribution in eastern Brazil. Genetica 125: 311–323.

Moraes EM, Perez MF, Téo MF, Zappi DC, Taylor NP & Machado MC. 2012. Cross-species amplification of microsatellites reveals incongruence in the molecular variation and taxonomic limits of the *Pilosocereus aurisetus* group (Cactaceae). Genetica 140: 277–285.

Morello J. 1967. Bases para el estudio fitoecológico de los grandes espacios (el Chaco argentino). Cienca e Investigacion 23: 252–267.

Moreno NC, Amarilla LD, Las Peñas ML & Bernardello G. 2015. Molecular cytogenetic insights into the evolution of the epiphytic genus *Lepismium* (Cactaceae) and related genera. Botanical Journal of the Linnean Society 177: 263–277.

Morgan K, O'Loughlin SM, Chen B, Linton Y-M, Thongwat D, Somboon P, Fong MY, Butlin R, Verity R, Prakash A, Htun PT, Hlaing T, Nambanya S, Socheat D, Dinh TH & Walton C. 2011. Comparative phylogeography reveals a shared impact of pleistocene environmental change in shaping genetic diversity within nine *Anopheles* mosquito species across the Indo-Burma biodiversity hotspot. Molecular Ecology 20: 4533–4549.

Moritz C & Faith DP. 1998. Comparative phylogeography and the identification of genetically divergent areas for conservation. Molecular Ecology 7: 419–429.

Moritz C, Patton JL, Schneider CJ & Smith TB. 2000. Diversification of rainforest faunas: an integrated molecular approach. Annual Review of Ecology and Systematics 31: 533–563.

Moritz C, Hoskin CJ, Mackenzie JB, Phillips BL, Tonione M, Silva N, VanDerWal J, Williams SE & Graham CH. 2009. Identification and dynamics of a cryptic suture zone in tropical rainforest. Proceedings of the Royal Society of London Series B: Biological Sciences 276: 1235–1244.

Motta PEF, Curi N, Franzmeier DP. 2002. Relation of soils and geomorphic surfaces in the Cerrado. In: The Cerrados of Brazil. Ecology and natural history of a neotropical savanna (Eds. PS Oliveira & RJ Marquis), pp. 13–32. Columbia University Press, New York.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA & Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature: 403: 853–858.

Nason JD, Hamrick JL & Fleming TH. 2002. Historical vicariance and postglacial colonization effects on the evolution of genetic structure in *Lophocereus*, a Sonoran Desert columnar cactus. Evolution 56: 2214–2226.

Nielsen R & Beaumont MA. 2009. Statistical inferences in phylogeography. Molecular Ecology 18: 1034–1047.

Norborg M. 2007. Coalescent theory. In: Handbook of statistical genetics (Eds. DJ Balding, M Bishop & C Cannings), pp. 843–877. John Wiley & Sons, England.

Novaes RML, de Lemos-Filho JP, Ribeiro RA & Lovato MB. 2010. Phylogeography of *Plathymenia reticulata* (Leguminosae) reveals patterns of recent range expansion towards northeastern Brazil and southern Cerrados in Eastern Tropical South America. Molecular Ecology 19: 985–998.

Novaes RML, Ribeiro RA, Lemos-Filho JP & Lovato MB. 2013. Concordance between phylogeographical and biogeographical patterns in the Cerrado: diversification of the endemic tree *Dalbergia miscolobium* (Fabaceae). PLoS ONE 8: e82198.

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: 312–326.

Ocampo G & Columbus T. 2010. Molecular phylogenetics of suborder Cactineae (Caryophyllales), including insights into photosynthetic diversification and historical biogeography. American Journal of Botany 97: 1–21.

Odee DW, Telford A, Wilson J, Gaye A & Cavers S. 2012. Plio-Pleistocene history and phylogeography of *Acacia senegal* in dry woodlands and savannahs of sub-Saharan tropical Africa: evidence of early colonisation and recent range expansion. Heredity 109: 372–382.

Oliveira FIC, Bordallo PN, Castro ACR & Correia D. 2013. Genetic diversity of spineless *Cereus jamacaru* accessions using morphological and molecular markers. Genetics and Molecular Research 12: 4586–4594.

Oliveira PS & Marquis RJ. 2002. The Cerrados of Brazil. Ecology and natural history of a neotropical savanna. Columbia University Press, New York, 409 pp.

Ortega-Baes P, Saraiva M, Suhring S, Godínez-Alvarez H & Zamar M. 2011. Reproductive biology of *Echinopsis terscheckii* (Cactaceae): the role of nocturnal and diurnal pollinators. Plant Biology 13: 33–40.

Ossa CG & Medel R. 2011. Notes on the floral biology and pollination syndrome of *Echinopsis chiloensis* (Colla) Friedrich & G.D. Rowley (Cactaceae) in a population of semiarid Chile. Gayana Botanica 68: 213–219.

Pan T, Wang H, Hu C-C, Shi W-B, Zhao K, Huang X, Zhang B-W. 2014. Range-wide phylogeography and conservation genetics of a narrowly endemic stream salamander, *Pachyhynobius shangchengensis* (Caudata, Hynobiidae): implications for conservation. Genetics and Molecular Research 13: 2873–2885.

Pennington RT, Prado DE & Pendry CA. 2000a. Neotropical seasonally dry forests and Quaternary vegetation changes. Journal of Biogeography 27: 261–273.

Pennington RT, Klitgaard BB, Ireland H & Lavin M. 2000b. New insights into floral evolution of basal Papilionoideae from molecular phylogenies. In: Advances in legume systematics: Part 9 (Eds. PS Hereendeen & A Bruneau), pp. 233–248. Royal Botanic Gardens, Kew.

Pennington RT, Lavin M, Prado DE, Pendry CA, Pell SK & Butterworth CA. 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. Philosophical Transactions of the Royal Society B: Biological Sciences 359: 515–538.

Polhill RM. 1981. Papilionoideae. In: Advances in legume systematics, Part 1 (Eds. RM Polhill & PH Raven), pp. 191–208. Royal Botanic Gardens, Kew.

Polhill RM, Raven PH & Stirton CH. 1981. Evolution and systematics of the Leguminosae. In: Advances in legume systematics, Part 1 (Eds. RM Polhill & PH Raven), pp. 1–26. Royal Botanic Gardens, Kew.

Polhill RM. 1994. Classification of the Leguminosae. In: Phytochemical dictionary of the Leguminosae, vol. 1: Plants and their constituents (Eds. FA Bisby, J Buckingham & JB Harborne), pp 35–48. Chapman and Hall, London.

Pott A & Pott VJ. 1994. Plantas do pantanal. Embrapa-CPAP, Corumbá, 320 pp.

Prado DE, Gibbs PE, Pott A & Pott VJ. 1992. The Chaco-Pantanal transition in southern Mato Grosso, Brazil. In: Nature and Dynamics of Forest-Savanna Boundaries (Eds. PA Furley, J Proctor & JA Ratter), pp. 451–470. Chapman & Hall, Londres, Greater London, Reino Unido.

Prado DE. 1993. What is the Gran Chaco vegetation in South America? 1. A review. Contribution to the study of flora and vegetation of the Chaco. V. Candollea 48: 145–172.

Prado DE. & Gibbs PE. 1993. Patterns of species distributions in the dry seasonal forests of South America. Annals of the Missouri Botanical Garden 80: 902–927.

Primack D, Imbres C, Primack RB, Miller-Rushing AJ & Del Tredici P. 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. American Journal of Botany 91: 1260–1264.

Ramos ACS, Lemos-Filho JP, Ribeiro RA, Santos FR & Lovato MB. 2007. Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of Quaternary climate changes in the Cerrado. Annals of Botany 100: 1219–1228.

Ramos ACS, de Lemos-Filho JP & Lovato MB. 2009. Phylogeographical structure of the neotropical forest tree *Hymenaea courbaril* (Leguminosae: Caesalpinioideae) and its

relationship with the vicariant *Hymenaea stigonocarpa* from Cerrado. Journal of Heredity 100: 206–216.

Resende M. 1976. Mineralogy, Chemistry, morphology and geomorphology of some soils of the central plateau of Brazil. Tese de doutorado, Purdue University, West Lafayette, USA.

Richards CL, Carstens BC & Knowles LL. 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. Journal of Biogeography 34: 1833–1845.

Riddle BR, Dawson MN, Hadly EA, Hafner DJ, Hickerson MJ, Mantooth SJ & Yoder AD. 2008. The role of molecular genetics in sculpting the future of integrative biogeography. Progress in Physical Geography 32: 173–202.

Rodrigues RS & Tozzi AMGA. 2008. Reinstatement of the name *Leptolobium* Vogel (Leguminosae, Papilionoideae, Sophoreae). Taxon 57: 980–984.

Rodrigues RS & Tozzi AMGA. 2010. Two new species of *Leptolobium* (Leguminosae, Papilionoideae) from Brazil. Brittonia 62: 20–25.

Rodrigues RS & Tozzi AMGA. 2012. Revisão taxonômica de *Leptolobium* (Papilionoideae, Leguminosae). Acta Botanica Brasilica 26: 146–164.

Rymer PD, Dick CW, Vendramin GG, Buonamici A & Boshier D. 2013. Recent phylogeographic structure in a widespread 'weedy' Neotropical tree species, *Cordia alliodora* (Boraginaceae). Journal of Biogeography 40: 693–706.

Särkinen T, Staats M, Richardson JE, Cowan RS & Bakker FT. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. PLoS ONE 7: e43808.

Sano EE, Rosa R, Brito JLS & Ferreira LG. 2007. Mapeamento de cobertura vegetal do bioma Cerrado. Embrapa Cerrados, Planaltina, 30 pp.

Sarmiento G. 1983. The savannas of tropical America. In: Ecosystems of the world (Ed. F Bouliére), pp. 245–288. Elsevier, Amsterdam.

Schlumpberger BO & Renner SS. 2012. Molecular phylogenetics of *Echinopsis* (Cactaceae): Polyphyly at all levels and convergent evolution of pollination modes and growth forms. American Journal of Botany 99: 1335–1349.

Scoble J & Lowe AJ. 2010. A case for incorporating phylogeography and landscape genetics into species distribution modelling approaches to improve climate adaptation and conservation planning. Diversity and Distributions 16: 343–353.

Sersic AN, Cosacov A, Cocucci AA, Johnson LA, Pozner R, Avila LJ, Sites Jr JW & Morando M. 2011. Emerging phylogeographical patterns of plants and terrestrial vertebrates from Patagonia. Biological Journal of the Linnean Society 103: 475–494.

Silva JMC & Bates JM. 2002 Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. BioScience 52: 225–233.

Silva JMC. 1997. Endemic bird species and conservation in the cerrado region, South America. Biodiversity and Conservation 6: 435–450.

Silva JSV, Pott A, Abdon MM, Pott VJ, Santos EKR. 2011. Projeto GeoMS: cobertura vegetal e uso da terra do Estado de Mato Grosso do Sul. 1a ed.. Embrapa Informação Agropecuária, Campinas, 64 pp.

Silva MP, Mauro RA, Abdon MM, Silva EJSV. 2008. Estado de conservação do Chaco (Savana Estépica) brasileiro. In: IX Simpósio Nacional Cerrado e II Simpósio Internacional Savanas Tropicais, Brasília, Distrito Federal, Brasil.

Simon MF, Grether R, Queiroz LP, Skema C & Pennington RT. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. Proceedings of the National Academy of Sciences the United States of America 106: 20359–20364.

Soltis PS & Soltis PE. 1993. Ancient DNA: Prospects and limitations. New Zealand J Bot 31: 203–209.

Sprent JI. 2009. Legume nodulation: a global perspective. Wiley-Blackwell, United Kingdom, 200 pp.

Sprent JI & McKey D. 1994. Advances in Legume Systematics, Part 5, The Nitrogen Factor. Royal Botanic Gardens, Kew, 256 pp.

Stockman AK & Bond JE. 2007. Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. Molecular Ecology 16: 3374–3392.

Svenning JC, Flojgaard C, Marske KA, Nogues-Bravo D & Normand S. 2011. Applications of species distribution modeling to paleobiology. Quaternary Science Reviews 30: 2930–2947.

Swenson NG. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. Journal of Evolutionary Biology 19: 717–725.

Taylor JW & Swann EC. 1994. DNA from herbarium specimens. In: Ancient DNA (Eds. B Hermann & S Hummel), pp. 166–181, Springer-Verlag, New York.

Taylor NP & Zappi DC. 2004. Cacti of eastern Brazil. Royal Botanic Gardens, Kew, 499 pp.

Taylor NP. 2000. Cactaceae of eastern Brazil. Tese de doutorado, Royal Botanic Gardens, Kew. 414 pp.

Telles MPC, Dobrovolski R, Souza KDS, de Souza Lima J, Collevatti RG, Soares TN, Chaves LJ & Diniz-Filho JAF. 2014. Disentangling landscape effects on population genetic structure of a Neotropical savanna tree. Natureza & Conservação 12: 65–70.

Terrazas T & Arias S. 2003. Comparative stem anatomy in the subfamily Cactoideae. The Botanical Review 68: 444–473.

Thiers B. 2015. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. Disponpivel em: http://sweetgum.nybg.org/ih/. Acesso em 20 de outubro de 2015.

Toews DPL, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Molecular Ecology 21: 3907–3930.

Tôrres NM & Vercillo UE. 2012. Como ferramentas de modelagem de distribuição de espécies podem subsidiar ações de governo. Natureza & Conservação 10: 228–230.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F & Palma-Silva C. 2013. Phylogeographical patterns shed light on evolutionary process in South America. Molecular Ecology 22: 1193–1213.

Vieira FDA, Novaes RML, Fajardo CG, Santos RMD, Almeida HDS, Carvalho DD & Lovato MB. 2015. Holocene southward expansion in seasonally dry tropical forests in South America: phylogeography of *Ficus bonijesulapensis* (Moraceae). Botanical Journal of the Linnean Society 177: 189–201.

Visser ME & Both C. 2005 Shifts in phenology due to global climate change: the need for a yardstick. Proceedings of the Royal Society B: Biological Sciences 272: 2561–2569.

Vogel T. 1837. Caesalpinieis Brasiliae. Linnaea 11: 381–416.

Warren DL. 2012. In defense of 'niche modeling'. Trends in Ecology & Evolution: 27: 497–500.

Webb SD. 1991. Ecogeography and the Great American Interchange. Paleobiology 17: 266–280.

Weising K, Nybom H, Pfenninger M, Wolff K, & Kahl G. 2005. DNA fingerprinting in plants: principles, methods, and applications. CRC press. Boca Raton FL. 470 pp.

Werneck FP. 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. Quaternary Science Reviews 30: 1630–1648.

Werneck FP, Nogueira C, Colli GR, Sites JW & Costa GC. 2012a. Climatic stability in the Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. Journal of Biogeography 39: 1695–1706.

Wing SL, Herrera F & Jaramillo C. 2004. A Paleocene flora from the Cerrajón Formation, Guajíra Peninsula, northeastern Colombia. VII International Organization of Paleobotany Conference Abstracts, pp. 146–147. Museo Egidio Feruglio, Trelew, Argentina.

Wojciechowski MF. 2003. Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective In: Advances in Legume Systematics, Part 10, Higher Level Systematics (Eds. BB Klitgaard & A Bruneau), pp. 5–35. Royal Botanic Gardens, Kew.

Wooten JA, Camp CD & Rissler LJ. 2010. Genetic diversity in a narrowly endemic, recently described dusky salamander, *Desmognathus folkertsi*, from the southern Appalachian Mountains. Conservation Genetics 11: 835–854.

Yakovlev GP. 1969. A review of *Sweetia* and *Acosmium*. Notes from the Royal Botanic Garden, Edinburgh 29: 347–355.

Zalamea P, Munoz F, Stevenson PR, Paine CET, Sarmiento C, Sabatier D & Heuret P. 2011. Continental-scale patterns of *Cecropia* reproductive phenology: evidence from herbarium specimens. Proceedings of the Royal Society B: Biological Sciences 278: 2437–2445.

ZEE-MS. 2015. Zoneamento Ecológico Econômico do Estado de Mato Grosso do Sul: Elementos para construção da sustentabilidade do território sul-mato-grossene.Segunda Aproximação. Campo Grande, MS.

Zhihao S & Mingli Z. 2014. A range wide geographic pattern of genetic diversity and population structure of *Hexinia polydichotoma* (Asteraceae) in Tarim Basin and adjacent areas. Biochemical Systematics and Ecology 56: 49–59.

Tabela 1. Estudos filogeográficos realizados com espécies arbóreas do Cerrado (levantamento realizado em 20 de março de 2015 na base de dados Web of Science® - Institute of Scientific Information, Thomson Scientific).

Família	Espécie	Marcador*	Referência
Anacardiaceae	*Astronium urundeuva*	cpDNA SSR	Caetano et al. 2008
Bignoniaceae	*Tabebuia aurea*	SSR	Collevatti et al. 2014
	Tabebuia impetiginosa	cpDNA ncDNA	Collevatti et al. 2012a
Caryocaraceae	*Caryocar brasiliense*	cpDNA SSR	Collevatti et al. 2003
	Caryocar brasiliense	cpDNA	Collevatti et al. 2012b
Leguminosae	*Plathymenia reticulata*	cpDNA	Novaes et al. 2010
	Hymenaea courbaril *Hymenaea stigonocarpa*	cpDNA	Ramos et al. 2009
	Hymenaea stigonocarpa	cpDNA	Ramos et al. 2007
	Dalbergia miscolobium	cpDNA ncDNA	Novaes et al. 2013
Melastomataceae	*Tibouchina papyrus*	cpDNA SSR	Collevatti et al. 2012c
Meliaceae	*Cedrela fissilis*	cpDNA ncDNA	Garcia et al. 2011
Moraceae	*Ficus bonijesulapensis*	cpDNA	Vieira et al. 2015

* cpDNA = sequências de DNA plastidial; ncDNA = sequências de DNA nuclear; SSR = seqüências simples repetidas (Simple Sequence Repeat).

Tabela 2. Estudos moleculares realizados com cactáceas (levantamento realizado em 4 de março de 2015 na base de dados Web of Science® - Institute of Scientific Information, Thomson Scientific).

Subfamília	Espécie	Tema (marcadores moleculares*)	Referência
Cactoideae	*Lophocereus schottii*	Genética de populações Filogeografia (aloenzimas)	Nason et al. 2002
	Harrisia spp.	Filogenia Molecular Biogeografia	Franck et al. 2013
	Lepismium spp. *Rhipsalis* spp.	Citogenética	Moreno et al. 2015
	Mammillaria pectinifera	Filogeografia (cpDNA)	Cornejo-Romero et al. 2014
	Cereus spp.	Diversidade Genética (AFLP)	Faria-Tavares et al. 2013
	Cereus jamacaru	Diversidade Genética (RAPD, ISSR)	Oliveira et al. 2013
	Pilosocereus aurisetus	Filogeografia (cpDNA, ncDNA, SSR)	Bonatelli et al. 2014
	Pilosocereus spp.	Filogeografia (cpDNA)	Bonatelli et al. 2013
	Selenicereus setaceus *Selenicereus megalanthus* *Hylocereus undatus* *Harrisia* sp.	Diversidade Genética (RAPD)	Junqueira et al. 2010
	Praecereus euchlorus *Pilosocereus* spp. *Facheiroa squamosa*	Genética de populações (aloenzima)	Moraes et al. 2005
	Pilosocereus aurisetus	Amplificação heteróloga de SSRs	Moraes et al. 2012
Opuntioideae	*Opuntia* spp.	Citogenética e variação geográfica	Majure et al. 2012

* AFLP = Polimorfismo de comprimento de fragmento amplificado (Amplified Fragment Length Polymorphism); RAPD = DNA polimórfico amplificado ao acaso (Random Amplified Polymorphic DNA); ISSR/SSR = seqüências simples repetidas internas (Inter Simple Sequence Repeat); SSR = seqüências simples repetidas (Simple Sequence Repeat); cpDNA = DNA cloroplastidial; ncDNA = DNA nuclear.

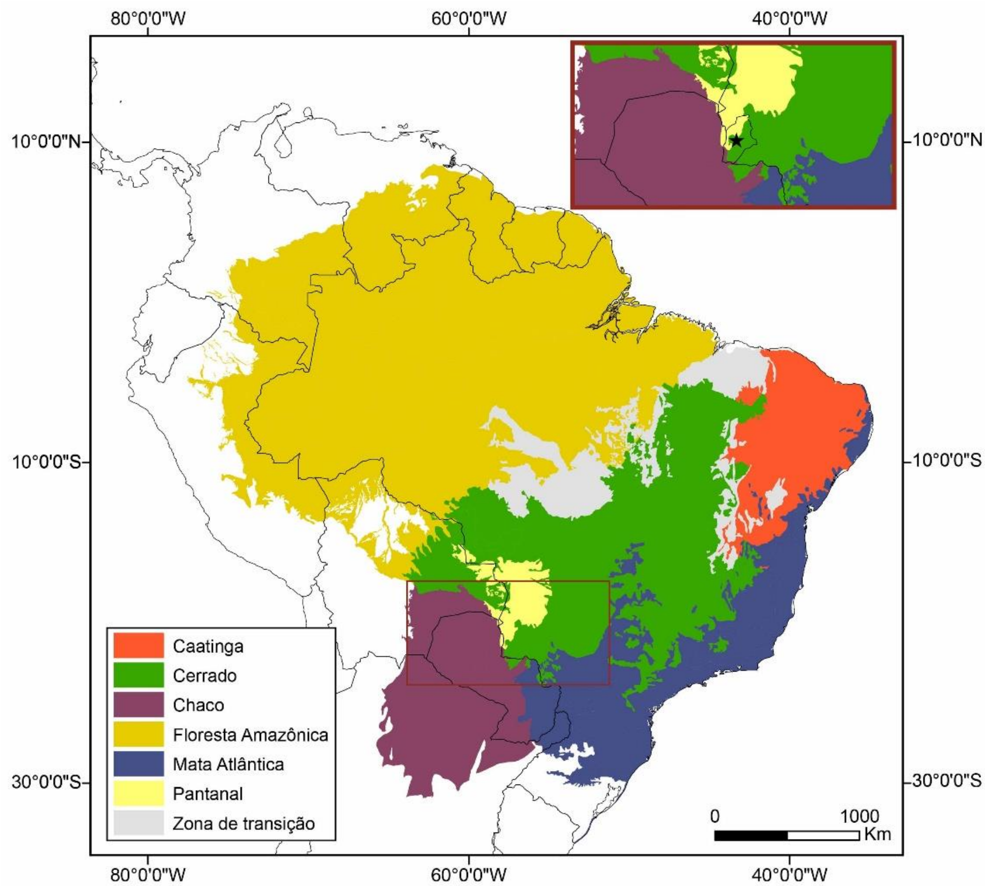

Figura 1. Localização do Cerrado em relação às demais formações vegetacionais da América do Sul. Em destaque a localização do município de Porto Murtinho, Mato Grosso do Sul, onde se localiza a porção brasileira do Chaco. Dados geográficos disponíveis em http://mapas.mma.gov.br/.

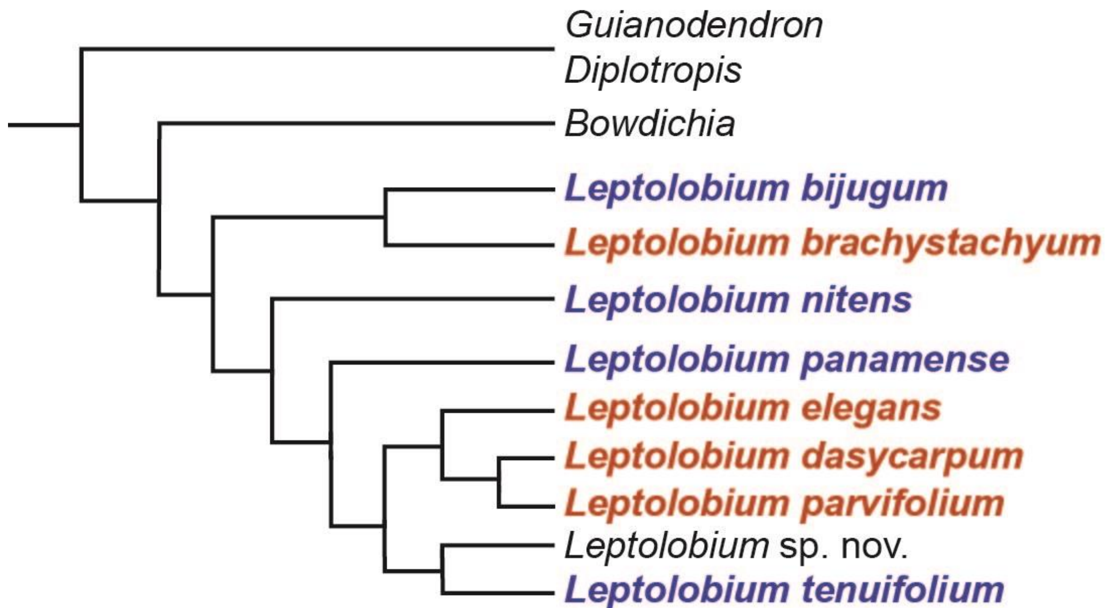

Figura 2. Representação da relação entre as espécies do gênero *Leptolobium* (Leguminosae, Papilionoideae) a partir da combinação de dados morfológicos e moleculares (DNA nuclear: gene 5,8S e ITS; DNA cloroplastidial: gene matk e intron trnL) (modificado a partir de Cardoso *et al.* 2012). As espécies que ocorrem em florestas úmidas estão representadas na cor azul, e as que ocorrem no Cerrado em vermelho.

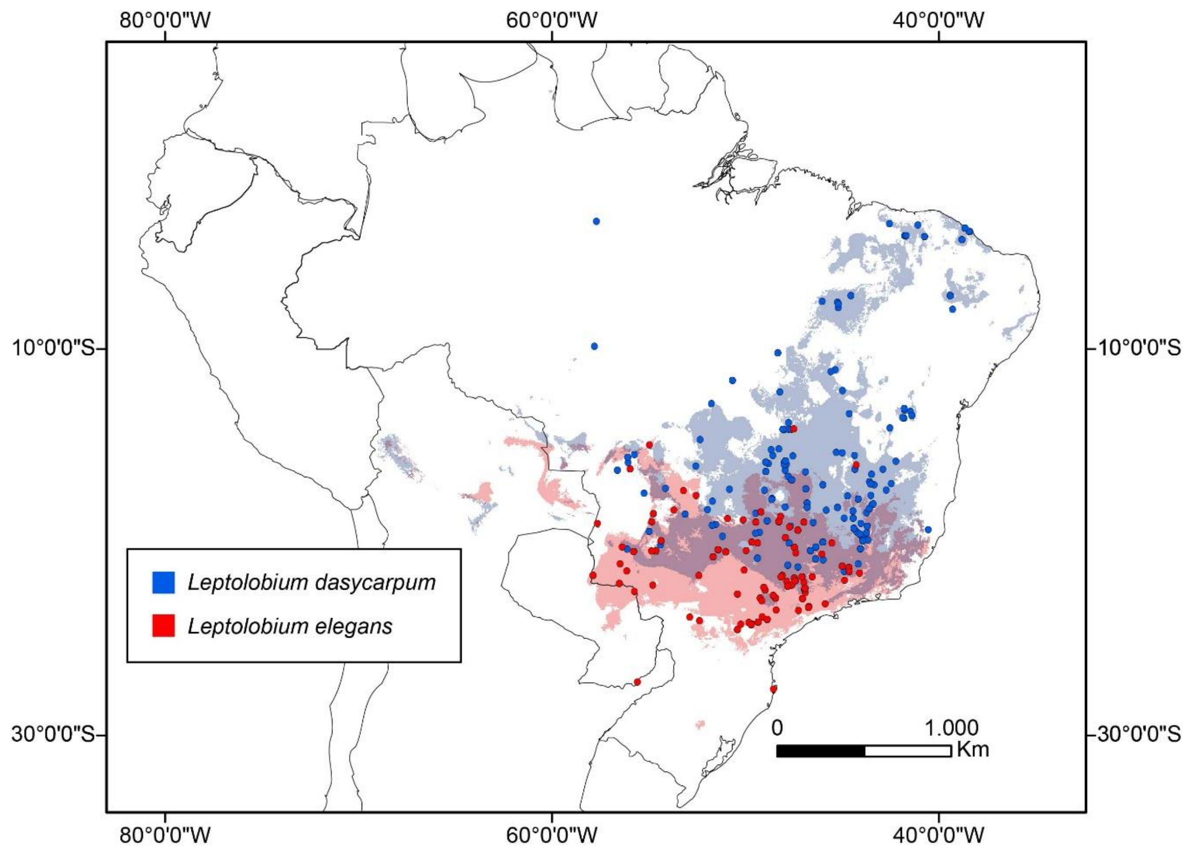

Figura 3. Distribuição de *Leptolobium dasycarpum* e *L. elegans* (Leguminosae, Papilionoideae) de acordo com registros de herbário. As áreas coloridas se referem à distribuição potencial das espécies para o presente (1950-2000) modeladas pelo algoritmo de Máxima Entropia (geradas pelos programas Maxent 3.3.3 e ArcGIS 10.1).

Objetivos

O objetivo geral da presente tese foi estudar os padrões fenológicos, filogeográficos e populacionais de espécies vegetais do Cerrado e Chaco, contribuindo para o melhor entendimento dos padrões e processos relacionados a origem, evolução e manutenção da biodiversidade dessas formações, dando ênfase nas seguintes abordagens:

1. Leguminosas arbóreas:

1.1. Analisar a fenologia reprodutiva de *Leptolobium dasycarpum* e *L. elegans*, ao longo de toda a sua distribuição, por meio de espécimes de herbário.

1.2. Estudar a estrutura genética e filogeográfica de populações de leguminosas arbóreas ao longo de sua distribuição no Cerrado através da análise dos seguintes marcadores moleculares: espaçadores intergênicos plastidiais (psbA–trnH e trnL–trnF) e espaçadores internos transcritos dos genes ribossomais nucleares (ITS).

1.3. Relacionar a estrutura genética e o padrão filogeográfico a eventos históricos, flutuações climáticas do Terciário/Quaternário e padrões geográficos para contribuir para a melhor compreensão dos efeitos de eventos históricos nas espécies do Cerrado.

1.4. Analisar a distribuição potencial dessas espécies, por meio de modelagem de nicho ecológico, gerando subsídios para uma melhor compreensão da história evolutiva do Cerrado.

2. Cactáceas

2.1. Desenvolver marcadores de microssatélites para espécies de cactos ocorrentes no Chaco brasileiro, gerando ferramentas moleculares para estudos de genética de populações e conservação.

2.2. Descrever a distribuição da variabilidade genética populacional das espécies com marcadores de SSR, correlacionando os dados genéticos com as características do habitat das espécies.

Chapter 1

Article to be submitted to Journal of Vegetation Science, ISSN: 1654-1103

Climate-driven and long term phenology of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) across the Cerrado

Resumo

Estudos fenológicos de plantas geralmente são de curta duração e abrangem uma pequena área geográfica. Análises em larga escala, na qual o clima pode variar, podem aprimorar nosso conhecimento acerca da dinâmica das espécies, sua ecologia e evolução. Registros fenológicos de espécies com ampla distribuição geográfica podem ser obtidos pelas exsicatas de herbário, as quais têm sido coletadas em abundância por vezes ao longo de toda a distribuição. Nós usamos espécimes de herbário coletados entre 1961 e 2012 para analisar os padrões fenológicos em duas espécies de *Leptolobium* (Leguminosae, Papilionoideae) em função de variáveis climáticas históricas. Encontramos uma tendência para a floração e frutificação tardia com o aumento da chuva em ambas as espécies (menos de 0,5 d/mm). Todavia, a temperatura parece ter influenciado apenas a floração de *L. dasycarpum*, levando à um aumento da data de floração (c. 7,5 d/°C). Apesar das limitações de se usar dados históricos, espécimes de herbário muitas vezes representam a única forma de registro e de caracterização abrangente dos padrões fenológicos de espécies com ampla distribuição, oferecendo dessa forma um complemento promissor aos estudos de caso em escala local, assim como um método efetivo para se detectar os efeitos da mudança climática nas épocas de floração e frutificação das espécies.

Abstract

Plant phenological studies are generally of short duration and cover a small geographic area. Large-scale analyses, in which the climate can vary, improve our knowledge about the dynamics of the species, their ecology and evolution. Phenological records of species with broad geographic distribution can be obtained by herbarium herbarium specimens, which have been collected in abundance at times throughout the distribution. Here we used herbarium specimens collected between 1961 and 2012 to examine the phenological trends in two *Leptolobium* species (*L. dasycarpum* and *L. elegans*) due to historic climatic variables. We found a trend towards later flowering and fruiting dates with increasing rainfall in both species (less than 0.5 d/mm). Notwithstanding, the temperature seems to have influenced only the flowering of *L. dasycarpum*, leading to a flowering advancement (c. 7.5 d/°C). Despite the limitations of using historical data, herbarium specimens often represent the only means of characterizing large-scale patterns in plant phenology, and offers a promising complement to local intensive field studies, as well as an effective method for detecting the effects of climate change on species flowering and fruiting season.

Introduction

The study of the timing of recurring biological events, as plant flowering and fruiting, and their interactions with the abiotic forces are the main themes of phenology. With the exception of agricultural researches, phenology has been regarded as a domain of natural history (Sparks & Menzel 2002), but this view started to change in the 1990s when it began to be used as a tool to investigate the impacts of climate change on the biota (Keatley & Hudson 2010).

Seasonal phases of the plant life play an important role in many ecological interactions. Pollination, competitive and predation require overlapping of interactive life stages (Johansson et al. 2015). Likewise, changes in phenological events across trophic levels may lead to an increase in extinction risks and loss of ecosystem services (Both et al. 2006, Memmott et al. 2007). These changes have been reported in many systems across the world and comprise one of the most well-documented ecological footprints of global warming (Parmesan 2006). Therefore, investigating the fluctuation in phenology in relation to variation in climate, life history, and phylogeny is a key not only for ecology but also for the plant evolution.

Phenology is mainly based on localized, short-term (generally 1–3 years) studies. Evidences that plant phenology changes due to climatic shifts rely on long-term datasets, usually more than 20 years. These studies are concentrated in the northern hemisphere, particularly in Europe (Menzel & Fabian 1999, Fitter & Fitter 2002, Primack et al. 2004, Miller-Rushing et al. 2006, Inouye 2008), with few studies in the southern (Gallagher et al. 2009, Zalamea et al. 2011). Data collected specifically for the study of climate-induced phenological changes are not easily available (Sparks & Carey 1995). Therefore, the lack of long-term phenological observations can be filled by specimen-based records of biological collections.

Herbarium specimens are a valuable source of phenological data since they include information on date and locality of collection and species' life phase. Many studies have been using these historical records (Primack et al. 2004, Bolmgren & Lönnberg 2005, Lavoie & Lachance 2006, Miller-Rushing et al. 2006, Bowers 2007, Gallagher et al. 2009) as they may represent a substitute for field observation. Herbarium specimens are also likely to have been collected when flowering and fruiting phenophases are near their peak, which are preferable for a suitable analysis (Miller-Rushing et al. 2008). Besides, by studying phenology at a wide geographic scale over which climate varies (i.e., continental or across a biome) we may improve our understanding of species dynamics and their ecological interactions.

Phenological patterns and environmental variables are closely related, and understanding this relationship, under a climate change scenario, is a crucial issue regarding the species and community dynamics. In tropical plant species, the periodicity of the phenophases is generally controlled by the variation in precipitation, soil water availability, temperature, irradiance and day length (Borchert et al. 2004, Borchert et al. 2005, Wright & Calderón 2006, Zimmerman et al. 2007). Moreover, the effects in the interaction between the species (e.g. pollination and dispersion) produced by the accelerate climate change can only be understood if the individual species responses are known (Robbirt et al. 2011).

In this study, we evaluated the reproductive phenological response to temperature and rainfall in two *Leptolobium* species (Leguminosae, Papilionoideae) throughout their distribution ranges in the Cerrado. We aim to test whether herbarium records could be used to detect long-term changes in flowering and fruiting times and the responses of the species to changes in climatic variables. This is the first attempt to use herbarium records to analyze the phenology of species that are widely distributed in the savanna vegetation of Brazil.

Methods

Studied species

Leptolobium Vogel is a Neotropical genus of Leguminosae that occurs from Mexico to northern Argentina, which comprises 12 species (eleven are found in Brazil). It is characterized by arboreal or shrubby habitat, flowers with white, actinomorphic or slightly zygomorphic corollas, 10 free stamens, a stipitate ovary with many ovules, indehiscent (samara-like or nut-like) fruit, compressed seeds, and a bulbose hypocotyl-radicle axis (Rodrigues & Tozzi 2012). Of the species occurring in Brazil, *Leptolobium dasycarpum* Vogel exhibit a wide distribution in Cerrado (Brazilian Savanna) whereas *L. elegans* Vogel is confined to the southern areas and in the transition between Cerrado and Pantanal wetlands (Figure 1). These two species are phylogenetically closed related (Cardoso et al. 2012), present similar morphologies and habitats (see Chapter 2) and therefore represent ideal models to investigate climate effects on the flowering and fruiting periods.

Climatic data

We monthly recorded averages of total precipitation (mm) and temperatures (°C) from the Instituto Nacional de Meteorologia/Banco de Dados Meteorológicos para Ensino e Pesquisa (INMET/BDMEP) to calculate the average monthly precipitation and temperature for each specimen in each year. For each location indicated in the specimen (exsiccate) label, we gathered the climatic data of the nearest meteorological base.

In order to assess the climatic changes across the distribution areas of the species over the 52 years of collected data (1961–2012), we calculated the precipitation and temperature anomalies by subtracting the 52-years average from yearly averages (Diskin et al. 2012, Calinger et al. 2013). These climatic data were also obtained from the INMET/BDMEP database.

Phenological data

We analyzed 631 specimens of *Leptolobium dasycarpum* and *L. elegans* collected in the Cerrado, and adjacent areas, in order to check for the presence of flowers and/or fruits.

We gathered 363 specimens, 246 for *L. dasycarpum* and 117 for *L. elegans* (Supplement S1).

We count the number of specimens with the presence of four reproductive phenophases: (i) floral buds (which indicate the beginning of the flowering period), (ii) opened flowers (which can indicate the flowering peak), (iii) immature fruits (beginning of the fruiting period), and (iv) mature fruits (beginning of the dispersal period). For each reproductive phenophase we calculate the Julian Date (*D*), i.e., the day of the year (Day 1 = January 1st) in which a specimen presented a given floral characteristic.

In the southern hemisphere the predominant flowering season falls in the turn of the year (spring-summer flowering season), which may cause a wrong Euclidean distance between observations. For example, when converting two observations from, e.g., 31 December 2010 and 1 January 2011, we arrive at the Julian Date 365 and 1, respectively, although these two observations have a Euclidian distance of just one. To prevent this problem, we performed a linear transformation by adding a constant (x = 365) to the portion of the data that fell between 1 and 183 (the mid point of the year) in order to preserve the Euclidian distance between the observations (Gallagher et al. 2009).

Statistical analyses

In order to evaluate if there any the correlation between the climatic variables, we firstly performed a Shapiro-Wilk's W-test for residual normality, and then made a Pearson Correlation between average annual precipitation and temperature for the whole 52-years period.

To determine the effect of precipitation and temperature on flowering and fruiting date, or a species' phenological responsiveness (ρ_x, *day* $°C^{-1}$, for temperature; *day* mm^{-1} for rainfall, Calinger et al. 2013), we regressed *D* against the average precipitation and temperature of each specimen's month of flowering and fruiting and the 3 months prior (P_{4i} and T_{4i}). That is,

$D_i = b_x + \rho_x(P4i \text{ or } T4i)$

where D_i is the phenophase date of specimen i, P_{4i} or T_{4i} is the average precipitation or temperature of the month of D_i and 3 months prior, and ρ_x is the slope, or climatic effect size, of this relationship.

Results

Climatic data

Climate variables taken for the study period (1961-2012) was not correlated (Shapiro-Wilk's W-test = 0.9704; p = 0.2187; Pearson (r) = -0.2469; p = 0.0777; r^2 = 0.0609). The 52-years temperature anomalies (relative to 1961-2012) indicated a recent trend towards warmer temperatures (Figure 2A). Positive temperature anomalies (mean 0.58°C; standard deviation (s.d.) ± 0.25) was evident in the last 19 years of the study period (1994–2012) and in three previous years (1986–1988; mean 0.31; s.d. ± 0.30). The mean rainfall anomalies (Figure 2B) showed no significant trend changes over the 52 years.

Phenological data

In general, the two species were continuously collected over the 52-years, but in differing intensities (Figure 3). An average of 6.0 (s.d. ± 5.7) and 3.1 (s.d. ± 3.4) specimens for *L. dasycarpum* and *L. elegans*, respectively, were collected per year. The earliest and latest collection day of the year was 9 and 363 for *L. dasycarpum* and 9 and 357 for *L. elegans*.

Both species have a similar flowering (presence of buds and open flowers) and fruiting (presence of immature and ripe fruit) season, peaking in the second half of the year (Figure 4). *Leptolobium dasycarpum* presents flowers and fruits mainly between October and January, whereas *L. elegans* mostly between October and November.

The flowering of *Leptolobium dasycarpum* showed a significant and positive phenological responsiveness to the climatic variables, but the temperature led to a major flowering advancement (c. 7.5 d/°C; Figure 5A, C) than the rainfall (c. 0.4 d/mm; Figure 5B, D). On the other hand, only the rainfall produced a significant and positive phenological responsiveness on *L. dasycarpum* fruiting (Figure 6B, D). Both immature and mature fruit set had an advancement of less than 0.5 d/mm.

The phenology of flowers and fruits of *Leptolobium elegans* seems to be influenced only by the rainfall. Buds and opened flowers showed a significant and positive phenological responsiveness to the rainfall, with an advancement of c. 0.4 d/mm (Figure 7). In relation to fruiting, *L. elegans* presented only the early fruiting phase (represented by the presence of immature fruits) influenced by the rainfall, with an advancement of c. 0.2 d/mm (Figure 8).

Discussion

This study assessed the patterns of reproductive phenological responsiveness of two species with broad distribution in the Cerrado by the analysis of herbarium specimens over a 52-years period. Both species showed significant advancement of flowering and/or fruiting phenology with the increasing of rainfall. The temperature influenced only the flowering of *L. dasycarpum*, making it blooms one week later. Even though precipitation is less strongly correlated with plant traits than temperature (Moles et al. 2014), it still can plays a role in life cycles shaping, as we demonstrated here. In spite of significant, the precipitation influence was considerably lower than that presented by temperature on the phenology of the phenology of *L. dasycarpum*. An explanation is that the high levels of precipitation seasonality in the tropics (Vázquez & Stevens 2004) do not mean an obligatory high variability in water access for plants. This is due deep root systems that allow many plants to access groundwater, and thus maintain photosynthesis and reproductive activities even during dry periods (Nepstad et al. 1994, Decker et al. 2013).

Herbarium collections were originally created for taxonomical purposes, but now their applicability for genetic, ecological, and environmental studies is increasingly common. The main relevance of herbarium specimens is that they represent a unique opportunity to obtain data from past time periods (Primack & Miller-Rushing 2009, Vellend et al. 2013). In regard to phenological studies, the use of herbarium specimens was proposed as a complement to field-based observations. In contrast with *in situ* studies, which are, normally, developed in a limited area, herbarium collections cover a large spatial extent and allow a broad view of the variation in phenology throughout the range of a species (Borchert 1996, Boulter et al. 2006). As suggested before (Bolmgren & Lönnberg 2005, Miller-Rushing et al. 2006), this study also supports the potential use of herbarium specimens for phenological studies without the need to perform local case studies.

While our results showed significant climatic responses in *Leptolobium* phenology, some problems with the herbarium-based approach should be considered. One of the issues is the collector preference bias, in which the collection dates could be non-randomly distributed throughout the year. However, Boulter et al. (2006) demonstrated that the total number of specimens collected in Australia in a month was randomly distributed across the course of the year. Indeed, Robbirt et al. (2011) also showed that collector bias is not a problem when the herbarium data are subjected to carefully controlled selection criteria in the phenological analyses.

With the increasing of phenological studies involving historical data, analyses number and refinement are also rising. The most used interpretation of phenological data relies on correlation analysis (Parmesan & Yohe 2003, Sparks & Tryjanowski 2005), usually by simple linear regression. As used in this study, the simple linear regression can determine whether there has been a change in the time of a phenophase, indicated by a significant estimative of the slope. Thus, this analysis is robust and has a role to play in phenology, especially when new methods of analysis (such as GAMLSS — Generalised Additive Model

for Location, Scale and Shape — approach, Hudson et al. 2010) are not presented in a clear and replicable way.

Shifts in phenology as a consequence of climate change are emerging from the scientific literature and becoming widely accepted. The intimate relationship between seasonal flowering and climatic conditions, coupled with ease of observation, makes the monitoring of flowering events a reliable and cost effective method for the detection of change in biological systems and an important tool in global change research (MacGillivray et al. 2010). At the same time, the long-term data sets required to determine the nature and magnitude of climatic impacts are very limited, and the use of herbarium collections and photographic images have been providing valuable information on species distribution and description of reproductive patterns, especially at large geographical scales (Zalamea et al. 2011).

In a recent review, Lavoie (2013) found 382 studies (from 1993 up to February 2012) with original data that used herbarium specimens to document biogeographical patterns or environmental changes. Most studies are less than 10 years old, and only 1.4% of the herbarium specimens worldwide have been used to answer biogeographical or environmental questions. In relation to Brazil, the number of herbarium specimens is impressive: currently there are 386 collections and subcollections that holds more than seven million records (five million georeferenced), more than 600,000 records with online images, and almost 80,000 records of species in redlists (Species Link Project, available at http://specieslink.org.br). Despite these numbers, our study is the first to show how herbarium data can provide enough information on phenology of Cerrado plants. As we have seen, and highlighted by Lavoie (2013), there are no longer any excuses for not exploiting this invaluable resource in environmental studies, since herbarium specimens can be considered excellent tools to investigate past, present and future patterns of plant phenology.

Acknowledgments

Coordination for the Improvement of Higher Education Personnel (CAPES) for W. S. Fava's Ph.D fellowship. Herbarium curators, especially the CGMS Herbarium (Federal University of Mato Grosso do Sul, Brazil) who granted access to their collections.

References

Bolmgren K & Lönnberg K. 2005. Herbarium data reveal an association between fleshy fruit type and earlier flowering time. International Journal of Plant Science 166: 663–670.

Borchert R. 1996. Phenology and flowering periodicity of Neotropical dry forest species: evidence from herbarium collections. Journal of Tropical Ecology 12: 65–80.

Borchert R, Meyer SA, Felger RS & Porter-Bolland L. 2004. Environmental control of flowering periodicity in Costa Rican and Mexican tropical dry forests. Global Ecology and Biogeography 13: 409–425.

Borchert R, Renner S, Calle Z, Navarrete D, Tye A, Gautier L, Spichiger R & Von Hildebrand P. 2005. Photoperiodic induction of synchronous flowering near the Equator. Nature 433: 627–629.

Both C, Bouwhuis S, Lessells CM & Visser ME. 2006. Climate change and population declines in a long-distance migratory bird. Nature 441: 81–83.

Boulter SL, Kitching RL & Howlett BG. 2006. Family, visitors and the weather: patterns of flowering in tropical rain forests of northern Australia. Journal of Ecology 94: 369–382.

Bowers JE. 2007. Has climatic warming altered spring flowering date of Sonoran Desert shrubs? The Southwestern Naturalist 52: 347–355.

Calinger KM, Queenborough S & Curtis PS. 2013. Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. Ecology Letters 16: 1037–1044.

Cardoso D, de Lima HC, Rodrigues RS, de Queiroz LP, Pennington RT & Lavin M. 2012. The *Bowdichia* clade of Genistoid legumes: Phylogenetic analysis of combined molecular and morphological data and a recircumscription of Diplotropis. Taxon 61: 1074–1087.

Decker M, Pitman AJ & Evans JP. 2013. Groundwater constraints on simulated transpiration variability over South-eastern Australian Forests. Journal of Hydrometeorology 14: 543–559.

Diskin E, Proctor H, Jebb M, Sparks T & Donnely A. 2012. The phenology of *Rubus fruticosus* in Ireland: herbarium specimens provide evidence for the response of phenophases to temperature, with implications for climate warming. International Journal of Biometeorology 56: 1103–1111.

Fitter AH & Fitter RSR. 2002. Rapid changes in flowering time in British plants. Science 296: 1689–1691.

Gallagher RV, Hughes L & Leishman MR. 2009. Phenological trends among Australian alpine species: using herbarium records to identify climate-change indicators. Australian Journal of Botany 57: 1–9.

Hudson IL. 2010. Interdisciplinary approaches: towards new statistical methods for phenological studies. Climatic Change 100: 143–171.

Inouye DW. 2008. Effects of climate change on phenology, frost damage, and floral abundance of montane wildlife flowers. Ecology 89: 353–362

Johansson J, Nilson J & Jonzén N. 2015. Phenological change and ecological interactions: an introduction. Oikos 124: 1–3.

Keatley MR & Hudson IL. 2010. Introduction and Overview. In: Phenological research: methods for environmental and climate change analysis (Eds. IL Hudson & MR Keatley), pp. 1–22. Springer Science & Business Media.

Lavoie C. 2013. Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. Perspectives in Plant Ecology, Evolution and Systematics 15: 68–76.

Lavoie C & Lachance D. 2006. A new herbarium-based method for reconstructing the phenology of plant species across large areas. American Journal of Botany 93: 512–516.

MacGillivray F, Hudson IL & Lowe AJ. 2010. Herbarium collections and photographic images: alternative data sources for phenological research. In: Phenological research: methods for environmental and climate change analysis (Eds. IL Hudson & MR Keatley), pp. 425–461. Springer Science & Business Media.

Memmott J, Craze PG, Waser NM & Price M. V. 2007. Global warming and the disruption of plant–pollinator interactions. Ecology Letters 10: 710–717.

Menzel A & Fabian P. 1999. Growing season extended in Europe. Nature 397: 659.

Miller-Rushing AJ, Primack RB, Primack D & Mukunda S. 2006. Photographs and herbarium specimens as tools to document phenological changes in response to global warming. American Journal of Botany 93: 1667–1674.

Miller-Rushing AJ, Inouye DW & Primack RB. 2008. How well do first flowering dates measure plant responses to climate change? The effects of population size and sampling frequency. Journal of Ecology 96: 1289–1296.

Moles AT, Perkins SE, Laffan SW, Flores-Moreno H, Awasthy M, Tindall ML, Sack L, Pitman A, Kattge J, Aarssen LW, Anand M, Bahn M, Blonder B, Cavender-Bares J, Cornelissen JHC, Cornwell WK, Díaz S, Dickie JB, Freschet GT, Griffiths JG, Gutierrez AG, Hemmings FA, Hickler T, Hitchcock TD, Keighery M, Kleyer M, Kurokawa H, Leishman MR, Liu K, Niinemets Ü, Onipchenko V, Onoda Y, Penuelas J, Pillar VD, Reich PB, Shiodera S, Siefert A, Sosinski EE, Soudzilovskaia NA, Swaine EK, Swenson NG, van Bodegom PM, Warman L, Weiher E, Wright IJ, Zhang H, Zobel M & Bonser SP. 2014. Which is a better predictor of plant traits: temperature or precipitation? Journal of Vegetation Science 25: 1167–1180.

Nepstad DC, Carvalho CR, Davidson EA, Jipp PH, Lefebvre PA, Negreiros GH, Silva ED, Stone TA, Trumbore SE & Vieira S. 1994. The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures. Nature 372: 666–669.

Parmesan C & Yohe G 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421: 37–42.

Parmesan C. 2006. Ecological and evolutionary responses to recent climate change. Annual Review of Ecology, Evolution, and Systematics 37: 637–669.

Primack D, Imbres C, Primack RB, Miller-Rushing AJ & Del Tredici P. 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. American Journal of Botany 91: 1260–1264.

Primack RB & Miller-Rushing AJ. 2009. The role of botanical gardens in climate change research. New Phytologist 182: 303–313.

Robbirt KM, Davy AJ, Hutchings MJ & Roberts DL. 2011. Validation of biological collections as a source of phenological data for use in climate change studies: a case study with the orchid *Ophrys sphegodes*. Journal of Ecology 99: 235–241.

Rodrigues RS & Tozzi AMGA. 2012. Revisão taxonômica de *Leptolobium* (Papilionoideae, Leguminosae). Acta Botanica Brasilica 26: 146–164.

Sparks TH & Menzel A. 2002. Observed changes in seasons: an overview. International Journal of Climatology 22: 1715–1725.

Sparks TH & Tryjanowski P. 2005. The detection of climate change impacts: Some methodological considerations. International Journal of Climatology 25: 271–277.

Sparks TH & Carey PD. 1995. The responses of species to climate over two centuries: an analysis of the Marsham phenological record 1736-1947. Journal of Ecology 83: 321–329.

Vázquez DP & Stevens RD. 2004. The latitudinal gradient in niche breadth: concepts and evidence. The American Naturalist 164: E1–E19.

Vellend M, Brown CD, Kharouba HM, Mccune JL & Myers-Smith IH. 2013. Historical ecology: Using unconventional data sources to test for effects of global climate change. American Journal of Botany 100: 1294–1305.

Wright SJ & Calderón O. 2006. Seasonal, El Niño and longer term changes in flower and seed production in a moist tropical forest. Ecology Letters 9: 35–44.

Zalamea P, Munoz F, Stevenson PR, Paine CET, Sarmiento C, Sebatier D & Heuret P. 2011. Continental-scale patterns of *Cecropia* reproductive phenology: evidence from herbarium specimens. Proceedings of the Royal Society B 278: 2437–2445.

Zimmerman JK, Wright SJ, Calderón O, Pagan MA & Paton S. 2007. Flowering and fruiting phenologies of seasonal and aseasonal Neotropical forests: the role of annual changes in irradiance. Journal of Tropical Ecology 23: 225–231.

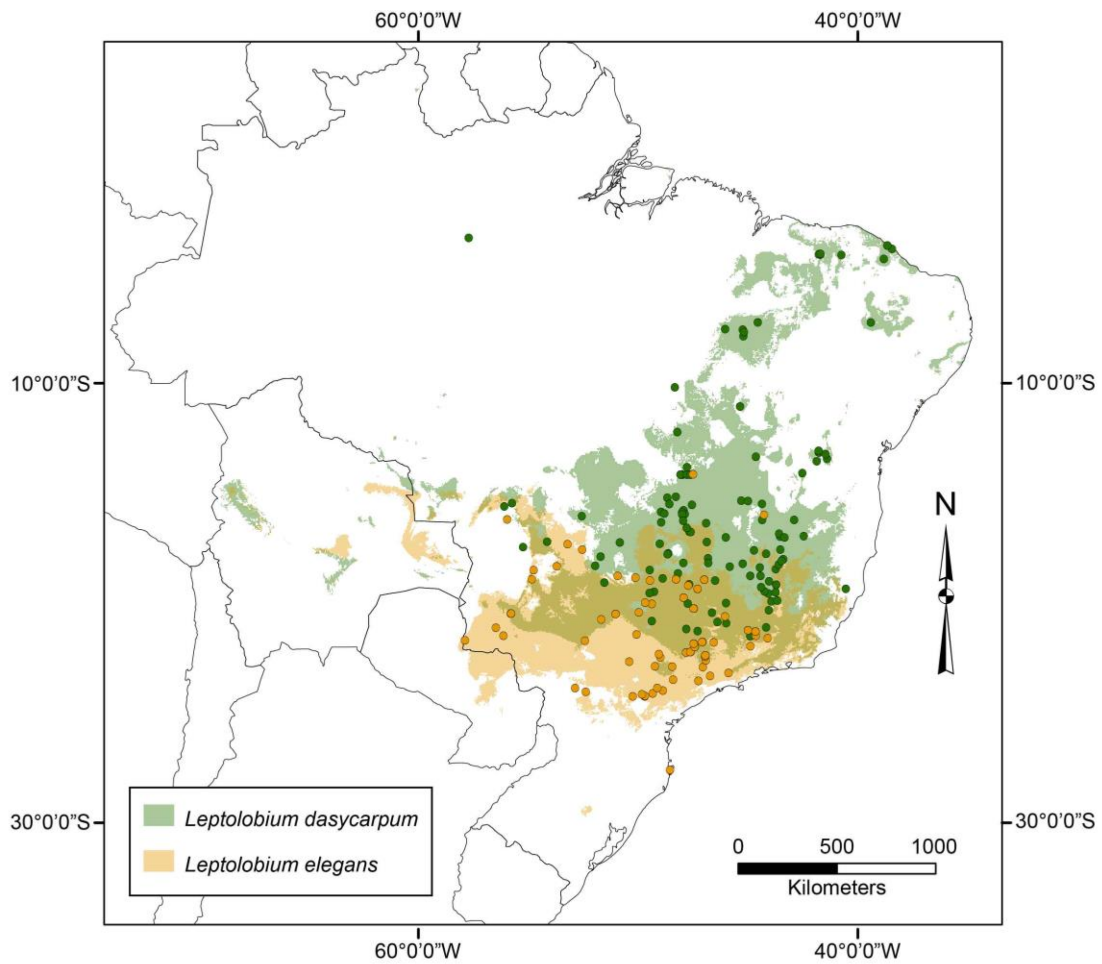

Figure 1. Distribution of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) according to herbarium records. The colored areas refer to the potential distribution of the species for the current time (1950-2000; see page 96, Chapter 2). Each point represent at least one specimen analyzed: *L. dasycarpum* (green points) N = 246 and *L. elegans* (orange points) N = 117.

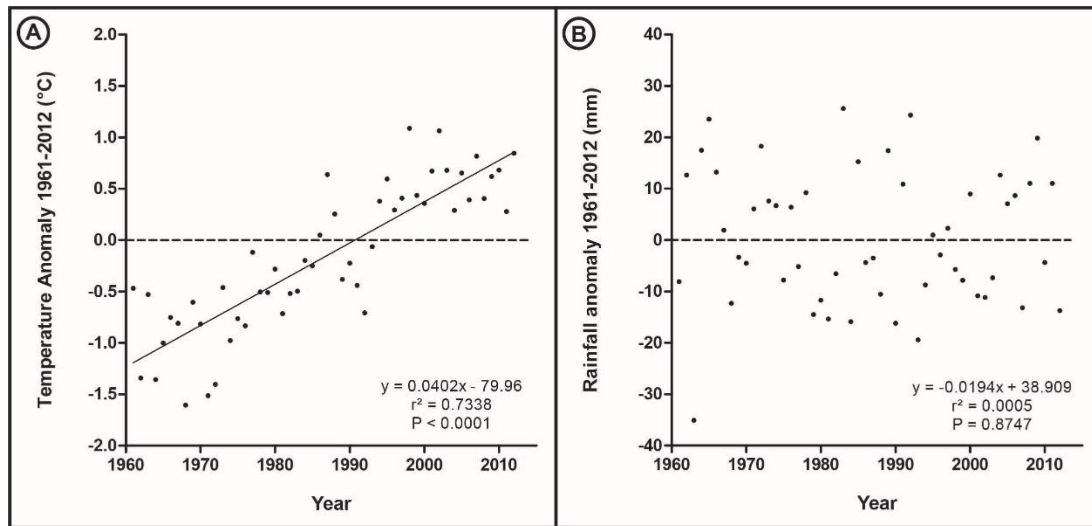

Figure 2. Annual climatic anomalies relative to the 52-years mean (shown as the horizontal dashed line). The trend line facilitate comparison of temperature (A) and rainfall (B) anomalies with the entire study period averages (1961–2012).

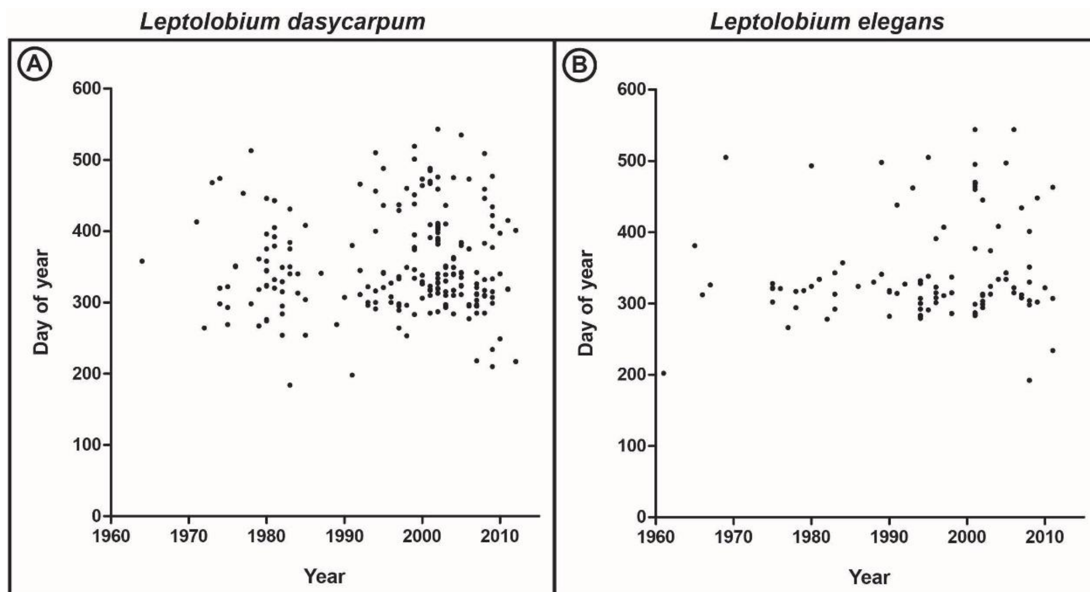

Figure 3. Collection date (day of year, or Julian Date) of the analyzed *Leptolobium dasycarpum* (A) and *L. elegans* (B) specimens. Each point represents one specimen (N = 246 for *L. dasycarpum* and N = 117 for *L. elegans*). Note that the Julian Date was linear transformed (as suggested by Gallagher et al. 2009).

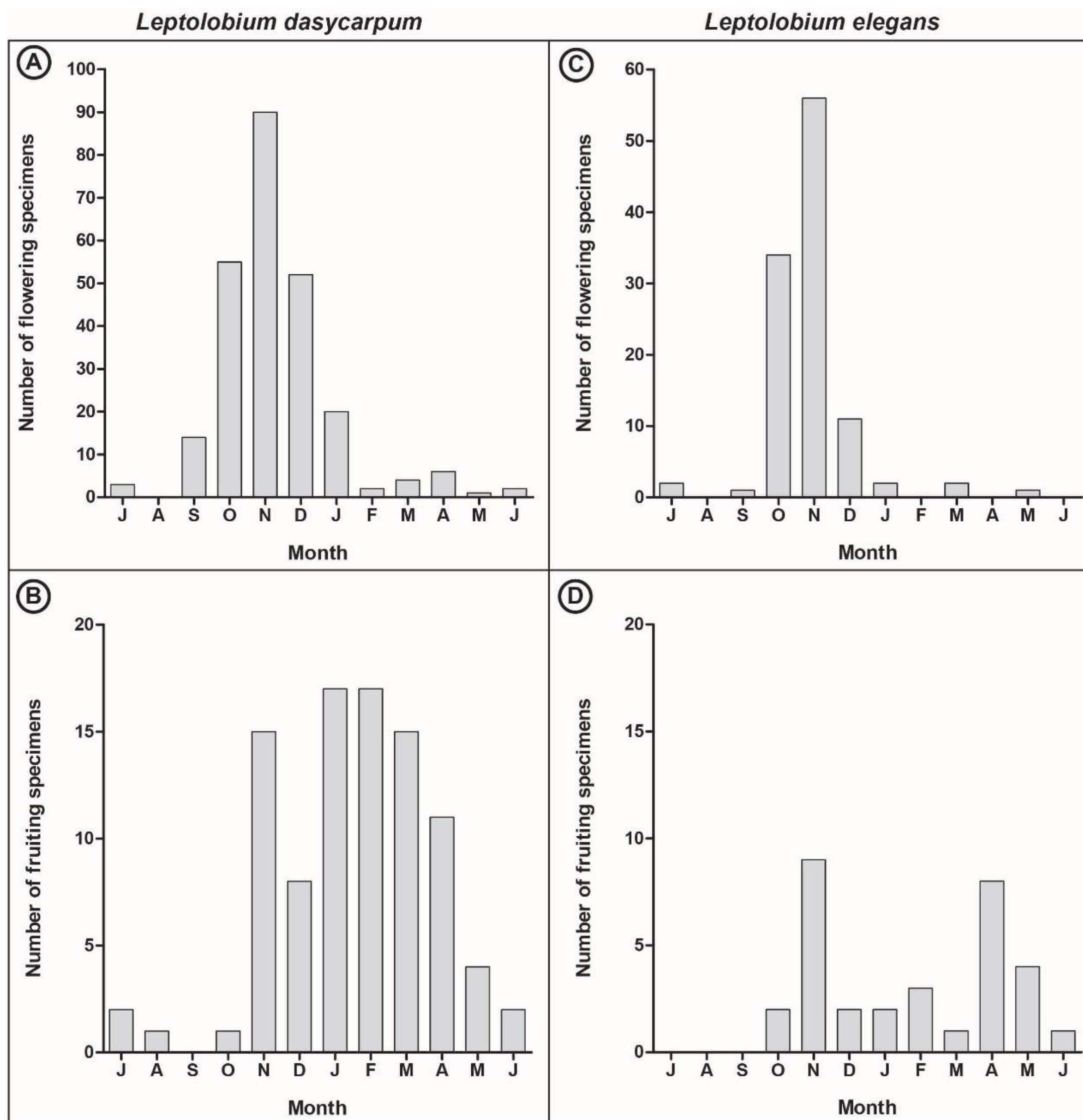

Figure 4. Reproductive phenology of *Leptolobium dasycarpum* (A and B) and *L. elegans* (C and D) based on herbarium samples collected between 1961 and 2012. The graphics show the total number of specimens exhibiting flowers (A and C) and fruits (B and D).

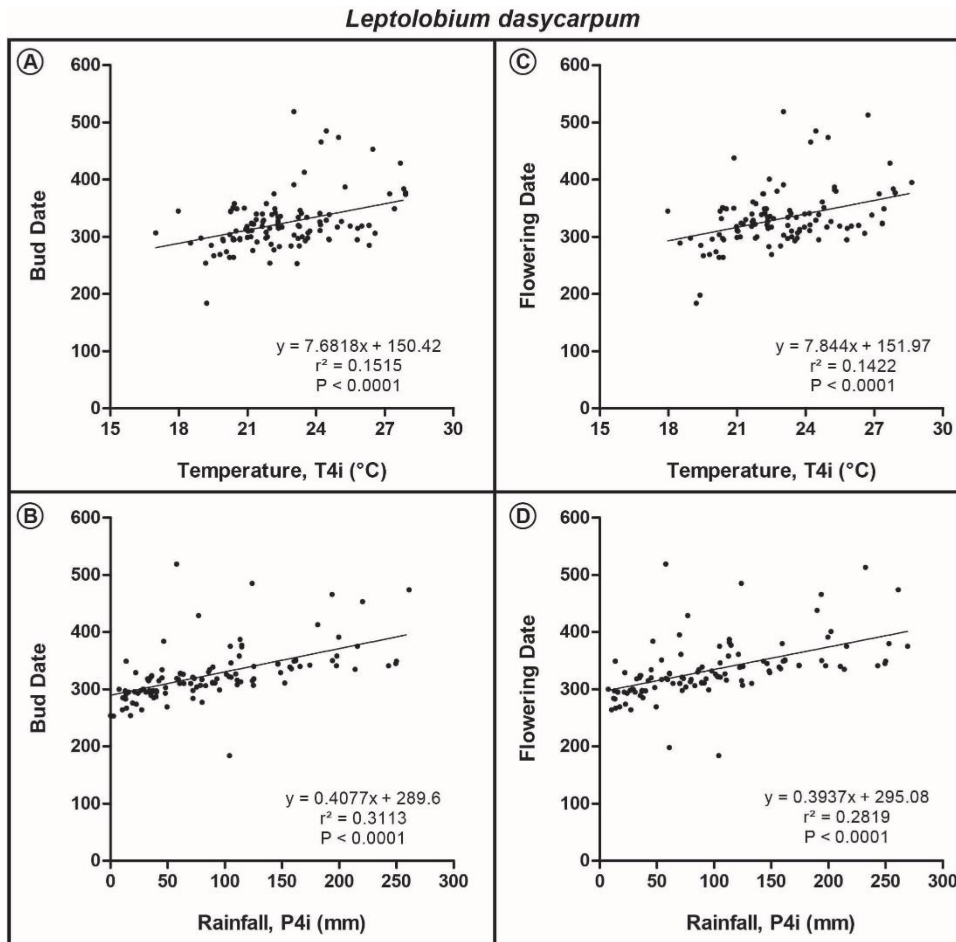

Figure 5. Flowering responsiveness of *Leptolobium dasycarpum* (Leguminosae, Papilionoideae). The graphics show the regression of the day of the year for buds (A and B) and opened flowers (C and D) against the average temperature (A and C) and precipitation (B and D) for each specimen's month of flowering and the 3 months prior (T_{4i} and P_{4i}). Note that the Julian Date was linear transformed (as suggested by Gallagher et al. 2009).

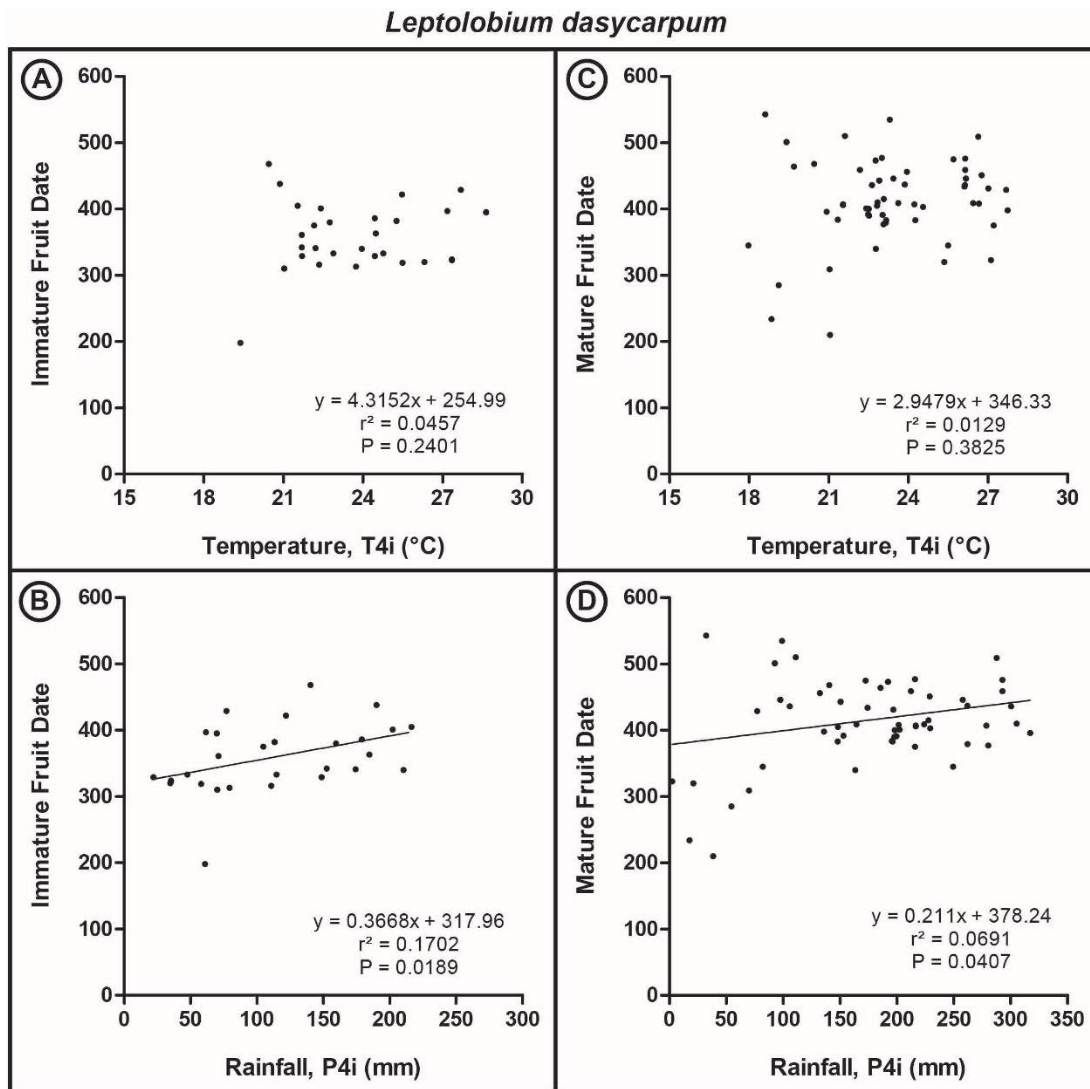

Figure 6. Fruiting responsiveness of *Leptolobium dasycarpum* (Leguminosae, Papilionoideae). The graphics show the regression of the day of the year for immature (A and B) and mature fruits (C and D) against the average temperature (A and C) and precipitation (B and D) for each specimen's month of flowering and the 3 months prior (T_{4i} and P_{4i}). Note that the Julian Date was linear transformed (as suggested by Gallagher et al. 2009).

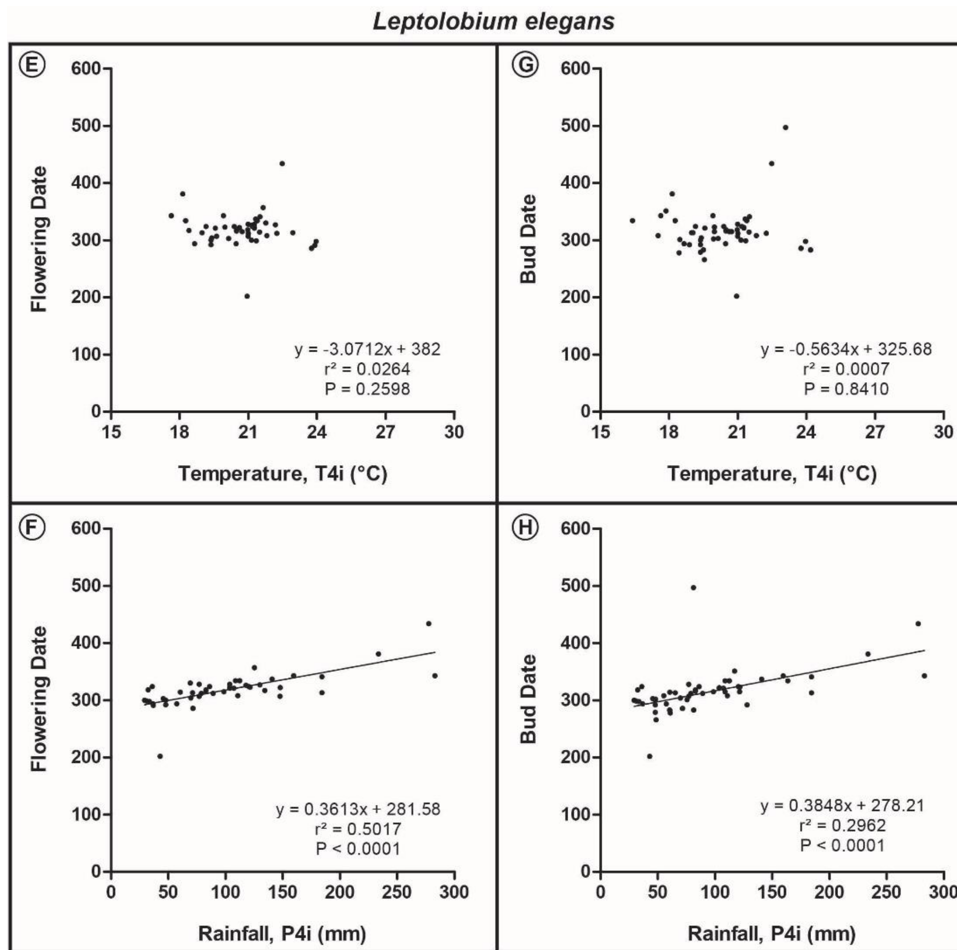

Figure 7. Flowering responsiveness of *Leptolobium elegans* (Leguminosae, Papilionoideae). The graphics show the regression of the day of the year for buds and opened flowers against the average temperature and precipitation for each specimen's month of flowering and the 3 months prior (T_{4i} and P_{4i}). Note that the Julian Date was linear transformed (as suggested by Gallagher et al. 2009).

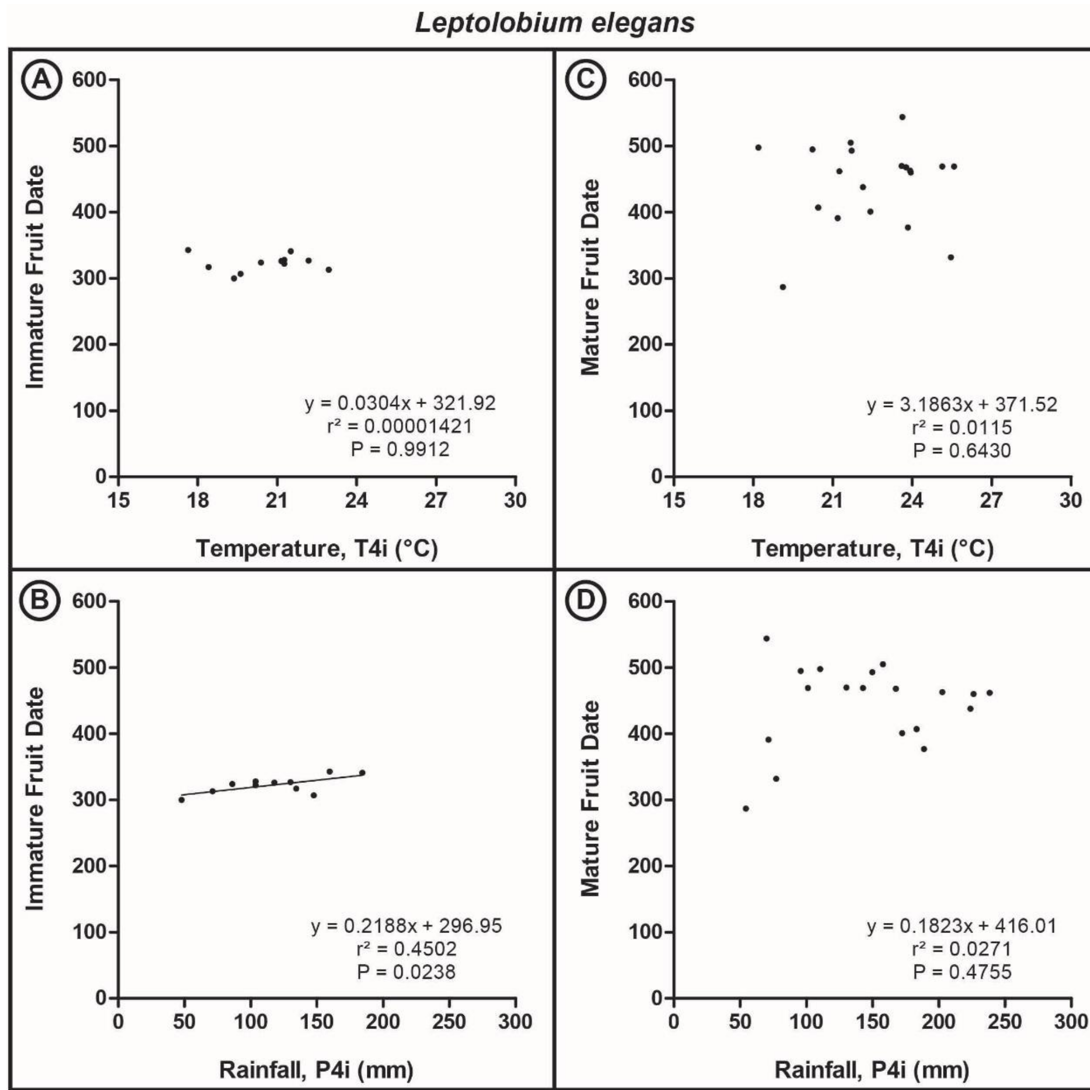

Figure 8. Fruiting responsiveness of *Leptolobium elegans* (Leguminosae, Papilionoideae). The graphics show the regression of the day of the year for immature (A and B) and mature fruits (C and D) against the average temperature (A and C) and precipitation (B and D) for each specimen's month of flowering and the 3 months prior (T_{4i} and P_{4i}). Note that the Julian Date was linear transformed (as suggested by Gallagher et al. 2009).

Chapter 1

SUPPORTING INFORMATION

Supplement S1

Leptolobium dasycarpum and *L. elegans* exsiccatae using in the phenological analyses.

Leptolobium dasycarpum

BHCB: 4397, 4794, 5284, 10715, 10784, 17389, 17390, 22510, 24594, 24748, 25609, 33630, 43312, 43706, 49491, 49917, 50943, 51493, 54356, 62968, 66753, 69669, 84739, 87350, 96282, 96831, 96976, 108961, 120638, 122115, 132249, 132320, 133933, 133943, 133952, 138740, 138781

CGMS: 330, 14251, 33495, 33643, 33644, 34705, 34706

EAC: 2514, 3122, 4397, 8247, 9113, 27298, 30869, 31507, 33252, 34716, 35529, 44295, 44709

ESA: 41052, 49037, 60275, 60944, 70990, 87224, 88974, 89287, 89383, 89608, 102754, 102860, 103372, 109118, 110860

FLOR: 28039, 28771

FUEL: 34819, 39442

FURB: 3014

HCF: 6135

HST: 8855, 13310, 16202

HUEG: 5564, 5565, 5566

HUFU: 0, 3657, 8070, 34850, 35709, 36140, 36141, 43723, 43858, 43962, 45088, 45765, 45766, 45949, 49218, 49366, 50759, 50773, 54350, 55703

IAC: 50782, 51129, 51801, 52949

IPA: 15149

SPSF: 15090, 22049, 36747

UB: 110, 381, 1586, 2150, 2394, 3265, 4174, 4899, 6203

UEC: 7410, 7411, 7412, 7413, 7414, 7415, 22904, 23496, 24991, 25899, 25912, 25932, 32208, 32402, 32434, 33288, 36319, 39005, 39018, 47759, 51790, 51976, 65006, 65012, 66434, 68018, 77116, 84381, 84409, 90957, 92386, 95957, 95962, 103189, 115865, 119224, 119225, 119464, 119474, 119478, 121060, 121568, 123498, 123550, 123558, 123560, 123571, 123576, 123577, 123581, 123585, 123593, 123599, 123602, 123615, 123617, 123659, 123660, 123661, 123662, 123664, 123666, 123667, 123973, 124616, 124881, 124887, 125127, 125989, 126290, 126400, 126401, 127448, 127460, 127696, 131025, 147426, 147427, 149107, 149188, 149189, 151012, 151017, 151020, 151026, 151027, 155640, 156321, 156338, 156359, 156361, 156365, 156368, 156369, 156525, 156984, 159810, 159811, 161085, 163121

R: 178767

UFRN: 6269, 9238

Leptolobium elegans

BHCB: 25621, 32446, 33629, 39640, 59226, 84079, 124710

CGMS: 7464, 7465, 17344, 17991, 18185, 19062, 24116, 28682, 33253, 33556

ESA: 9261, 20775, 20888, 27895, 33950, 79355, 93921, 103145, 115563, 115992

FLOR: 26072, 28037

FUEL: 6436, 13922, 32044

FURB: 4872

HCF: 536, 6420, 9910

HUEG: 5299

HUFU: 9031, 10333, 11611, 29149, 29179, 29631, 30837, 31383, 34765, 44952, 45645, 49362, 50078, 50243, 53930, 54893, 60957

IAC: 22345, 29495, 47206

MBML: 28872

SPSF: 30609, 30878, 32289, 37117, 41195, 41201, 42155, 42705

UB: 1444, 4301

UEC: 303, 440, 1085, 1110, 1226, 7418, 7419, 7420, 7421, 9183, 9186, 23034, 25689, 25898, 34165, 34358, 38727, 41257, 43208, 44743, 55601, 59310, 63008, 68333, 68761, 74437, 75063, 77905, 84377, 84380, 84383, 84389, 84393, 84394, 84398, 84401, 84402, 87306, 87307, 98842, 98845, 98857, 104813, 107981, 119463, 119465, 119471, 119476, 119477, 119479, 119480, 121293, 121299, 123658, 126285, 126324, 147979, 153895, 156364, 156366, 156367, 159803, Glaucia 12, Glaucia 5

R: 130470, 131758, 160634, 171548, 171549

UFRN: 2292, 6305

Herbaria names and their corresponding abbreviations.

BHCB
Herbário da Universidade Federal de Minas Gerais
CGMS
Herbário da Universidade Federal de Mato Grosso do Sul
EAC
Herbário Prisco Bezerra
ESA
Herbário da Escola Superior de Agricultura Luiz de Queiroz
FLOR
Herbário do Departamento de Botânica da Universidade Federal de Santa Catarina
FUEL
Herbário da Universidade Estadual de Londrina
FURB
Herbário Dr. Roberto Miguel Klein
HCF
Herbário da Universidade Tecnológica Federal do Paraná
HST
Herbário Sérgio Tavares
HUEG
Herbário da Universidade Estadual de Goiás
HUFU
Herbário Uberlandense
IAC
Herbário do Instituto Agronômico de Campinas
IPA
Herbário Dárdano de Andrade Lima
MBML
Herbário Mello Leitão
SPSF
Herbário Dom Bento Pickel
UB
Herbário da Universidade de Brasília
UEC
Herbário da Universidade Estadual de Campinas
R
Herbário do Museu Nacional Rio de Janeiro
UFRN
Herbário da Universidade Federal do Rio Grande do Norte

Chapter 2

Article to be submitted to Journal of Biogeography, ISSN: 1365-2699

Combining ecological niche modelling and phylogeography to reveal the demographic history of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) in the Cerrado

Resumo

Nós investigamos a história demográfica de *Leptolobium dasycarpum* e *L. elegans* (Leguminosae, Papilionoideae) unindo modelagem de distribuição potencial e análises filogeográficas. Amostramos populações das espécies ao longo de suas distribuições no Cerrado e analisamos os polimorfismos nos genomas cloroplastidial (espaçadores intergênicos psbA–trnH e trnL–trnF) e nuclear (ITS). A análise de relacionamento filogenético do cpDNA baseada no método de Median-Joining mostrou que *L. elegans* não tem haplótipos exclusivos, e compartilha seus três haplótipos com *L. dasycarpum*. Para o ITS, as espécies compartilham três dos 19 haplótipos (*L. dasycarpum* possui 16 e *L. elegans* seis haplótipos). As análises de coalescência mostraram que as espécies tiveram um tamanho populacional constante com fluxo gênico praticamente inexistente entre as populações, mas a análise de Extended Bayesian Skyline Plot indicou que ambas as espécies experimentaram uma expansão demográfica histórica, com evidências de declínio seguido de expansão populacional. A modelagem de nicho ambiental sugeriu que ambas as espécies experimentaram uma redução de suas distribuições durante o Último Máximo Glacial (LGM), e que elas apresentam uma moderada sobreposição de nicho no tempo atual, LGM e no último período interglacial. Esta sobreposição também é observada nas características morfológicas, uma vez que as espécies apresentam folíolos com tamanhos próximos. A divergência recente das linhagens de cpDNA, combinada com uma alta diferenciação populacional sem estruturação geográfica pode ter sido causada pelo efeito fundador, combinada com modos de polinização e de dispersão de curto alcance. Em relação ao ITS, as populações estão estáveis por um longo tempo e, aparentemente, elas não foram afetadas pela última glaciação, o que é evidenciado pela modelagem de nicho ambiental das espécies e do Cerrado como um todo. A parte central do Cerrado, devido à sua maior diversidade genética, pode representar uma área de refúgio, enquanto as populações nas porções mais periféricas seria um reflexo do efeito fundador.

Abstract

We investigated the demographic history of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) combining potential distribution modeling and phylogeographic analyzes. We sampled populations of the species throughout their distributions in the Cerrado and analyze polymorphisms in chloroplast genomes (intergenic spacers psbA–trnH and trnL–trnF) and nuclear (ITS). Phylogenetic relationship analysis of cpDNA based on Median-Joining method showed that *L. elegans* has no unique haplotypes, and shares its three haplotypes with *L. dasycarpum*. For the ITS, the species share three of the 19 haplotypes (*L. dasycarpum* has 16 and *L. elegans* six haplotypes). Analyses of coalescence showed that the species had a constant population size with negligible gene flow among populations, but the Extended Bayesian Skyline Plot analysis indicated that both species experienced a historic demographic expansion, with evidences of decline followed by population expansion. The environmental niche modeling suggested that both species experienced a reduction of its distributions during the Last Glacial Maximum (LGM), and they present a moderate niche overlap at the present time, LGM and the last interglacial period. This overlap is also observed in morphological characteristics, since the species present leaflets with similar sizes. The recent divergence of the lineages of cpDNA, combined with a high population differentiation without geographic structure may have been caused by the founder effect, combined with ways of pollination and short-range dispersion. Regarding the ITS, populations are stable for a long time and apparently they were not affected by the last ice age, which is evidenced by environmental niche modeling of species and the Cerrado as a whole. The central part of the Cerrado, because of their greater genetic diversity may represent a refuge area, while populations in the more peripheral portions would be a reflection of the founder effect.

Introduction

The Cerrado is the largest and most threatened savanna-like ecosystem in South America. It covers approximately 22% of Brazil's territory (ca. 2 million km^2), lying between the Amazon and Atlantic forests. Along with the Caatinga (northeastern Brazil) and the chaquenian province (Argentina, Bolívia and Paraguay), it forms a corridor of open vegetation seasonally stressed by drought, present unique biotas and have complex mosaic-type distributions (Prado & Gibbs 1993). The high number of species (ca. 160,000 species of plants, animals, and fungi) and endemic plants (ca. 4,400 species) made the Cerrado an important biodiversity hotspot (Myers et al. 2000). However, nearly half of the region is currently under direct human use, and about 35% of its total cover had been converted into planted pastures and crops (Oliveira & Marquis 2002).

The biogeographical history of Cerrado during the Quaternary period was spatially complex, as well as the factors that determine its distribution, like rainfall, dry season duration, soil fertility and fire (Eiten 1972, Rizzini 1997, Oliveira-Filho & Ratter 2002). Although there are few palynological records for the region, in general, pollen data suggest that the last glacial period, colder and drier than present-day conditions, caused the expansion of open vegetation and reduction in forests ranges (Behling 2002). At the last glacial maximum (LGM; ca. 20,000 years ago), the vegetation of the Cerrado was replaced by subtropical grassland, which apparently expanded further north (Behling & Lichte 1997, Behling 1998, Behling & Hooghiemstra 2001). The warmer and wetter conditions of early Holocene led to the grasslands replacement by different forms of Cerrado, by semideciduous forests in regions with short dry seasons and by rainforests in regions without significant dry periods (Behling 1995). Areas of historical stability were important refugia for vegetation during Quaternary climatic fluctuations. As postulated by Ab' Sáber (1983) and corroborated by Werneck et al. (2012), the higher plateaus (up to 1700 m) in central and

north-eastern Cerrado remained stable during the late Pleistocene, and probably formed a single large refuge.

Phylogeographical studies, i.e., studies concerned with the principles and processes governing the geographic distributions of genealogical lineages (Avise 2000), may provide valuable insights into the historical processes underlying diversification of the Cerrado species. Changes in the vegetation distribution associated with climatic instability, have been considered important factors to genetic diversity and population differentiation of the Cerrado plant species, leading to a geographical differentiation of populations into well-structured genetic groups (Ramos et al. 2007, 2009, Novaes et al. 2010, Viana e Souza & Lovato 2010, Collevatti et al. 2012, Novaes et al. 2013, Collevatti et al. 2014). Thus, the xeric regions of South America (e.g., Cerrado, Chaco, Caatinga, Pampas and Orinoco) represent an important source of information regarding the demographic responses of populations adapted to drought stress during climatic oscillations (Turchetto-Zolet et al. 2013). Therefore, comparative phylogeographic studies of closely related species sharing ecological attributes and common life-history can be useful to understand the effects of past climatic changes (Ramos et al. 2009).

Uniparentally inherited organellar markers, such as the cloroplastid DNA (cpDNA), are broadly used in phylogeographical studies because they are more likely to retain information about past migration histories than are nuclear markers (Weising et al. 2005). However, the use of only maternal-inherited molecular markers sometimes is not sufficient to distinguish between common patterns found in closely related species, like hybridization and incomplete lineage sorting, making it necessary to use other markers, such as nuclear DNA. The internal transcribed spacer (ITS) region of nuclear ribosomal DNA (nrDNA) is the most widely explored molecular marker in plants (Sang et al. 1995, Soltis et al. 1998), and in some genus the high mutation rate of the ITS region is enough to reveal intraspecific polymorphisms and enable its use in phylogeographical studies (Kropf et al. 2002, Parker et

al. 2004, Lorenz-Lemke et al. 2005, Cubas et al. 2006, Javadi et al. 2007, Ramdhani et al. 2010, Collevatti et al. 2012, Odee et al. 2012, Turchetto-Zolet et al. 2012, Novaes et al. 2013).

A methodological tool that have been supporting the interpretation of genetic-geographic patterns is the Environmental (or Ecological) Niche Models (ENM). Known as a set of different methods that aim to identify the environmental niche and potential distribution of species (Svenning et al. 2011, Warren 2012, Alvarado-Serrano & Knowles 2014), ENM provides a rigorous assessment of spatially explicit scenarios during different time periods and offers the ability to identify processes operating at the population level (Premoli et al. 2012). Therefore, one of the most common applications of ENMs is the identification of suitable areas of potential distribution of the species along different periods of time.

Here, we combined genetic and ecological niche modeling data to elucidate the spatial genetic structure of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae), two closely related species with different distribution patterns along the Cerrado.

Methods

Studied species

Leptolobium Vogel (Leguminosae, Papilionoideae) is a neotropical genus which comprises 12 species (11 are found in Brazil), characterized by its arboreal or shrubby habitat, non-papilionate white flowers with actinomorphic or slightly zygomorphic symmetry, 10 free stamens, a stipitate ovary with many ovules, indehiscent fruit (samara-like or nut-like), compressed seeds, and a bulbose hypocotyl-radicle axis (Rodrigues & Tozzi 2012). *Leptolobium dasycarpum* Vogel exhibits a wide distribution in Cerrado (Brazilian Savanna)

whereas *L. elegans* Vogel is confined to the southern areas, in different vegetation types of Cerrado and low semideciduous forests (Figure 1A).

Sampling strategy

We used two approaches to access the entire distribution of the studied species. First we collected leaflets during field collection performed between 2010 and 2013 (deposited in CGMS Herbarium). The leaflets were preserved in 95% ethanol prior to DNA isolation. Second, in order to complete the wide species distribution range, herbarium specimens were also sampled. The specimen leaflets were stored in silica gel until the DNA isolation. Information regarding geographic location, voucher specimens and sample size are given in Table S1 (Support Information).

DNA isolation, amplification and sequencing

Total DNA was extracted from the collected leaflets following the protocol of Doyle & Doyle (1987) with some modifications (Protocol 1, Support Information). We extracted DNA from the herbarium specimens using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA). Quality of DNA was assessed by visualization in 1% agarose gel. IMAGEJ 1.45 (Schneider et al. 2012, available at http://imagej.nih.gov/ij/), an image software, was used to quantify the genomic DNA by comparison with a 100-bp DNA Ladder (Promega, Fitchburg, WI, USA).

The screening for cpDNA amplification and polymorphism in *Leptolobium dasycarpum* and *L. elegans* was conducted on a subset of individuals (n = 30) using 11 universal primers pairs for non-coding cpDNA regions previously described (Table S2, Support Information). In addition, the nuclear ribosomal ITS region (ITS1, 5.8S and ITS2) was tested for amplification and polymorphism using two primers pairs: 75/92 (Desfeux & Lejeune 1996) and 18Sa/26Sa (Beira-M & Lavin 1999). Only the specific primers for

Leguminosae (18Sa/26Sa) showed satisfactory amplification. Two non-coding cpDNA regions (the psbA–trnH and trnL–trnF intergenic spacers) revealed polymorphisms in the analyzed individuals and were therefore selected for a large-scale survey of haplotype variation in the studies species. The trnV–trnM intergenic region also showed polymorphism, but the amplification success was lower than the other cpDNA markers, making it difficult to be used in all individuals.

Amplifications were performed in 25 µl reactions containing ca. 10 ng of genomic DNA, 1X Taq buffer with 2.0 mM $MgCl_2$, 0.2 mM of dNTP set, 0.2 µM of each primer, 1U GoTaq® Flexi DNA Polymerase (Promega, Madison, WI, USA) and UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, Thermo Fisher Scientific Inc.). The amplifications were carried out in a Veriti® (Applied Biosystems®, California, USA) thermal cycler with the following conditions: I. psbA–trnH = 95°C for 2 min (1 cycle); 95°C for 1 min, 50°C for 1 min, 72°C for 1 min (35 cycles) and 72°C for 5 min (1 cycle). II. trnL–trnF = 95°C for 2 min (1 cycle); 95°C for 1 min, 57°C for 1 min, 72°C for 1 min (30 cycles) and 72°C for 5 min (1 cycle). III. ITS = 95°C for 3 min (1 cycle); 95°C for 1 min, 50°C for 1 min, 72°C for 1.5 min (30 cycles) and 72°C for 5 min (1 cycle).

Following the amplifications, the products were verified in 1% agarose gels and purified using ExoSAP-IT® (Affymetrix, Inc., Santa Clara, CA) following the manufacturer's instructions. PCR products were double-strand sequenced using the DYEnamic ET dye terminator sequencing kit (GE Healthcare, Buckinghamshire, England, UK), and analyzed in ABI3100 (Applied Biosystems®, California, USA) automated sequencer.

Sequence analysis

We use CHROMAS 2.0 (Technelysium Pty Ltd, Queenland, Australia) to verify and edit forward and reverse reads, and the sequences were aligned using CLUSTAL-W (Thompson

et al. 1994) implemented in the MEGA 6 software (Tamura et al. 2013). The alignments were checked and edited by hand. All polymorphisms detected were validated by visually checking the original electropherograms.

Polymorphisms at mononucleotide microsatellites regions (Poly-A/T) were not considered for further analyses due to ambiguous alignment and higher mutation rates. Indels with more than 1 bp were coded as one mutational event.

Ambiguous nucleotide positions at nrDNA sequences, identified by double peaks, were classified as heterozygous when the weakest signal was at least 25% of the strength of the strongest signal (Fuertes Aguilar et al. 1999, Fuertes Aguilar & Nieto Feliner 2003). To reconstruct the nrDNA haplotypes, we used PHASE 2.1 software (Stephens et al. 2001), which uses a Bayesian statistical method to infer haplotypes of individuals that have more than one heterozygous site based on genotypic data. It was run under default conditions, allowing for multiallelic loci (-d option) and for 10,000 iterations. The output of PHASE was transformed through the web tool SEQPHASE (Flot 2010).

The analyses were performed with two datasets: cpDNA regions concatenated (psbA–trnH and trnL–trnF) and nrDNA (ITS). All sequences generated in this study will be deposited in GenBank.

The haplotypes were defined by DNASP 5.10 (Librado & Rozas 2009). The haplotype diversity (h) and nucleotide diversity (π) were estimated in ARLEQUIN 3.5 (Excoffier & Lischer 2010). Genealogical relationships among haplotypes were estimated with two network approaches, the median-joining method (Bandelt et al. 1999), implemented in NETWORK 4.2.0.1 (Fluxus Technology Ltd. at www.Fluxus-engineering.com), and the statistical parsimony (Templeton et al. 1992), implemented in TCS (Clement et al. 2000).

To investigate the population structure (Φ_{ST} values) of both species, analyses of molecular variance (AMOVA) using pairwise differences were performed in ARLEQUIN.

A spatial analysis of molecular variance (SAMOVA) was employed using the SAMOVA 1.0 software (Dupanloup et al. 2002) to identify spatial boundaries among the sampled localities. One thousand simulated annealing processes were employed to optimally allocate all samples into geographic groups (considering K from 2 to 10). The optimal number of groups was selected according to the highest F_{CT} value (differentiation among groups). A bayesian analysis of population structure, implemented in BAPS 6 (Corander et al. 2008) was also employed to analyze the population genetic structure. This method is based on the Markov chain Monte Carlo simulation approach that group together populations into variable, user-defined, number (K) of clusters. The optimal K cluster population partition is characterized by the highest marginal log-likelihood.

Demographic analyses and divergence times

The population expansion hypothesis was tested with different approaches. The pairwise mismatch distribution (Rogers & Harpending 1992) i.e., a histogram of genetic differences between pairs of individuals within a sample, was simulated under the sudden-demographic expansion and the spatial-demographic expansion models. We computed two groups of neutrality tests: (i) Tajima's D (1989) and Fu and Li's (1993) F* and D*, considering the frequency of mutations (segregating sites); and (ii) Fu's (1997) F_S, based on the haplotype distribution. All the tests were performed using ARLEQUIN and DNASP.

We also tested if differentiation is an effect of isolation by distance (Wright 1943). Pairwise F_{ST} among pairs of populations were estimated and correlated with a geographical distance matrix by a Mantel test (Mantel 1967) using 10,000 random permutations. Spatial autocorrelation analyses quantified the genetic distance mean between pairs of individuals that fell into distance class y (Ay). Analyses were performed using 10 equal distance classes (with unequal sample sizes) and 10,000 replicates to identify distance classes where average genetic distances were significantly larger or smaller than random expectations. The

definition of distances classes, in terms of the total number of classes and their upper and lower limits, is somewhat arbitrary and depends on the spatial distribution of the populations. A "rule of thumb" suggests about four to five classes for 20 populations (Diniz-Filho et al. 2013). These two analyses were performed using ALLELES IN SPACE 1.0 (Miller 2005).

The demographical parameter θ ($\theta = 4\mu N_e$ for diploid genome, where N_e is the effective population size and μ is the mutation rate in mutations per generation), were calculated based on a Bayesian estimation using Markov chain Monte Carlo (MCMC) approach (Beerli & Felsenstein 2001) implemented in LAMARC 2.1.10 software (Kuhner 2006). We also explored changes in effective population size by estimating the demographic parameter g (exponential growth rate), where $\theta_t = \theta_{now}\exp(-gt)$ and t is time in mutational unit. To access historical genetic connectivity, we estimated the number of migrants per generation from scaled migration rate, $M = 4N_e m/\theta$, where m is the migration rate. Because of the low sample size of some populations, demographic parameters were not estimated for all populations (see Tables S8 to S13, Supporting Information). The analyses were run with one first long chain with different number of recorded steps depending on the species and molecular marker used (Table S3, Supporting Information). Most probable estimates (MPE) were obtained, that is, the highest points on the posterior probability curve for a given parameter, which is the best solution found by a Bayesian run, and also the credibility interval of each parameter (Kuhner & Smith 2007).

To estimate the diversification time of the *L. dasycarpum* and *L. elegans* lineages we used a Bayesian approach implemented in BEAST 1.8 (Drummond et al. 2012) using *Leptolobium parvifolium* (Harms) Sch.Rodr. & A.M.G. Azevedo as outgroup. Sequence from nrDNA of *L. parvifolium* was obtained from Cardoso et al. (2012) (Accession number JX124499). The cpDNA sequences were obtained in this study. The settings used were a Yule tree prior with a random starting tree, a substitution model given by JMODELTEST 2 (Darriba et al. 2012), and the uncorrelated log-normal relaxed clock. We used universal

mutation rates estimated by Wolfe et al. (1987) for chloroplast noncoding regions (1.1–2.9 x 10^{-9} substitutions/nucleotide/year) and for ITS (5.8–31.5 x 10^{-9} substitutions/nucleotide/year). The software TRACER 1.5 (Rambaut et al. 2014) was used to check for convergence of Monte Carlo Markov Chains (MCMC) and adequate effective sample sizes (≥ 200). The first 10% of generations had been deleted as burn-in. The final joint sample was used to estimate the maximum clade credibility tree using the program TREEANNOTATOR, which is part of the BEAST 1.8 package. Statistical support for the clades was determined by assessing the Bayesian posterior probability.

We also used the Extended Bayesian Skyline Plot method (EBSP; Heled & Drummond 2008) implemented in BEAST 1.8 to analyze the population size dynamics through time, without a priori specification of the number of groups. The best fit substitution model for each genetic region was determined by JMODELTEST 2. A random initial tree, the linear model and the lognormal relaxed molecular clock were set. The lengths (from 50,000 to 100,000) of the MCMC chains for BEAST analysis were set to achieve ESS ≥ 200 in order to avoid autocorrelation of parameters sampling and assure proper mixing.

Environmental niche models

To estimate *L. dasycarpum* and *L. elegans* species distribution models (SDMs) along Quaternary climatic fluctuations, we implemented the maximum entropy machine-learning algorithm MAXENT (Phillips et al. 2006), which estimates an index of relative habitat suitability, based on presence-only data. The occurrence dataset (167 points for *L. dasycarpum* and 87 for *L. elegans*; see Figure S1; Supporting Information) included our field records and georeferenced herbarium specimens. The herbarium data were verified with geoLoc tool (SpeciesLink) for those specimens that lacked geographic coordinates of the collection site. When it was not possible to obtain the coordinates of the collection site, we used the municipality's coordinates.

We use data from BIOCLIM, consisting of 19 bioclimate variables (Table S4) available in Worldclim (Hijmans et al. 2005; http://worldclim.org). These variables are derived from the monthly temperature and rainfall values in order to generate more biologically meaningful variables. The bioclimatic variables represent annual trends (e.g., mean annual temperature, annual precipitation) seasonality (e.g., annual range in temperature and precipitation) and extreme or limiting environmental factors (e.g., temperature of the coldest and warmest month). The scenario created for current time (1950 to 2000) was extrapolated to the climatic conditions of the past during the Last Glacial Maximum (about 22,000 years ago), through the scenarios MIROC (Model for Interdisciplinary Research on Climate; Hasumi & Emori 2004), CCSM (Community Climate System Model; Shin et al. 2003) and MPI-ESM (Max-Planck-Institute Earth System Model; Giorgetta et al. 2013), all in 2.5 minutes resolution. It was also created a scenario for the Last Interglacial period (LIG; ca. 120,000 to 140,000 years ago; Otto-Bliesner et al. 2008), with a 30 arc-seconds resolution. Bioclimatic layers were cropped to span the South American region, a larger spatial range than the Cerrado distribution. We then developed three model for SDM of *L. dasycarpum* and *L. elegans*: (i) present (1950 to 2000); (ii) LGM, overlapping CCSM, MIROC and MPI scenarios; and (iii) LIG.

Model validation

To statistically evaluate model performance, we used the area under the curve (AUC) of the receiver operating characteristic (ROC) plot, which is designed to evaluate the specificity (absence of commission error) and sensitivity (absence of omission error) of a diagnostic test (Fielding & Bell 1997). The AUC provides a threshold-independent measure of model performance as compared to null expectations: an AUC of 0.50 indicates model performance no better than random, while higher AUC values indicate better models (Hanley & McNeil 1982). We use the recommendations of Swets (1988) and Araújo et al. (2005) to interpreting

range values as: Excellent if AUC>0.90; Good if 0.80<AUC<0.90; Acceptable if 0.70<AUC<0.80; Bad if 0.60<AUC<0.70; Invalid if 0.50 <AUC<0.60.

We also validated our models using a set of comparative similarity measures and statistical tests that permit quantitative comparison of Environmental Niche Models (ENM). We use ENMTOOLS (Warren et al. 2010) to quantify niche similarity between *L. dasycarpum* and *L. elegans* using the Schoener's (1968) D index. As niche overlap indices ideally range linearly from 0 (no overlap) to 1 (identical potential distributions), we arranged the results in five overlapping classes (sensu Rödder & Engler 2011): (i) 0–0.2 = no or very limited overlap; (ii) 0.2–0.4 = low overlap, (iii) 0.4–0.6 = moderate overlap; (iv) 0.6–0.8 = high overlap; and (v) 0.8–1.0 = very high overlap. In addition to quantifying niche similarity, we implemented a quantitative test of niche similarity that ask whether the ENMs generated from two species are identical, or, at the other extreme, are merely more similar than expected by chance. This hypothesis is addressed with the niche Identity Test, a null model also implemented in ENMTOOLS.

Morphological analysis

There is great intraspecific variation along the occurrence of *Leptolobium dasycarpum* and *L. elegans* and, in some cases, they are poorly differentiated by the morphology (Rodrigues & Tozzi 2012). We conducted morphometric analysis of both species using herbarium specimens (218 for *L. dasycarpum* and 177 for *L. elegans*; Supplement S1; Supporting Information).

We scanned the specimens with a HP Scanjet G4050 scanner (resolution of 300 dpi) and used the program IMAGEJ (Abràmoff et al. 2004) to measure the area, perimeter, length and width of the blade of the leaflets, the leaflet shape, and the petiolule length of all leaflets which were quite visible in exsiccates, irrespective of position. The shape of the leaflets was quantified using the Dissection Index (Kincaid & Schneider 1983), defined as the ratio of

perimeter to the square root of area. These characters were selected since they are vegetative features that help distinguish the two species (Rodrigues & Tozzi 2012). Reproductive characters were not used in the analysis due to the small number of flowers and fruits present in the specimens.

We calculated the mean value of each character, by specimen, using all leaflets in good conditions. To test if the two species share morphological characteristics, we use the R software (R Development Core Team 2015) to perform a Principal Component Analysis (PCA) (Pearson 1991) with Spearman correlation (Spearman 1904).

Results

Patterns of variability in cpDNA and ITS sequences

Amplification of the chloroplast intergenic spacers psbA–trnH and trnL–trnF generated fragments of 410 and 454 bp (base pairs), respectively. The combined data presented 864 bp (27.6% GC content) with six polymorphic sites in *Leptolobium dasycarpum* and just one in *L. elegans* sequences. For ITS region the alignment presented 685 bp, with 12 and five polymorphic sites. Our database was composed by 280 sequences (individuals) for the cpDNA region (159 *L. dasycarpum* sequences and 121 *L. elegans*), and 271 sequences for the ITS region (148 *L. dasycarpum* and 123 *L. elegans*), comprising 39 populations of *L. dasycarpum* (Figure 1B; Table S1) and 27 populations of *L. elegans* (Figure 1C; Table S1).

We found nine cpDNA haplotypes, all present in *L. dasycarpum. Leptolobium elegans* did not show exclusive haplotypes, sharing its three haplotypes (H03, H04 and H05) with *L. dasycarpum*. The haplotype and nucleotide diversity of the cpDNA region for *L. dasycarpum* was 0.7898 and 0.002049, respectivally. *Leptolobium elegans* presented a haplotype and nucleotide diversity of 0.2335 and 0.000530, respectivelly. The haplotype diversity of the ranged from 0 to 1.0 in both species, and the nucleotide diversity ranged from 0 to 0.003497 in the 39 populations of *L. dasycarpum*, and from 0 to 0.001297 in the

27 populations of *L. elegans* (Tables S5 and S6; Supporting Information. The majority of populations, both *L. dasycarpum* (74%) and *L. elegans* (92%), were monomorphic, presenting only one cpDNA haplotype.

For the ITS sequences, we found a total of 19 haplotypes, 16 in *L. dasycarpum* and six in *L. elegans* populations. The species shared three haplotypes (Hn04, Hn07 and Hn09) and most of the populations were polymorphic, presenting two up to seven haplotypes. Along the Cerrado, central and southern populations presented the highest number of haplotypes. *Leptolobium dasycarpum* populations 9 (Campo Grande/MS) and 34 (Serranópolis/GO) presented seven and five haplotypes, respectively (Figure 3), while *L. elegans* populations 64 (Selvíria/MS) and 65 (Taunay/MS) presented four (Figure 4). The haplotype and nucleotide diversity of the ITS region for *L. dasycarpum* was 0.5822 and 0.003144, respectivally. *Leptolobium elegans* presented a haplotype and nucleotide diversity of 0.5487 and 0.001204, respectivally. The haplotype diversity of the ITS region ranged from 0 to 1.0 in both species, and the nucleotide diversity ranged from 0 to 0.008165 in the 30 populations of *L. dasycarpum*, and from 0 to 0.004323 in the 26 populations of *L. elegans* (Table S5 and S6).

The genetic groups identified were not consistent with the geographic population locations in both species (Figures S6 and S7, Supporting Information). For *L. dasycarpum*, the SAMOVA and BAPS analysis revealed, respectively, two and eight groups for cpDNA, and four and seven groups for the ITS region. Similarly, for *L. elegans* we find four and three groups for the cpDNA (SAMOVA and BAPS, respectively), and two and five groups for the ITS sequences.

The two different network approaches used to infer the relationships among the cpDNA and nrDNA haplotypes achieved similar topologies (only the median-joining networks are shown in Figures 2C, 3B and 4B, respectively). Regarding to cpDNA, H04 was the most frequent haplotype, occurring in 16 of the 36 populations of *L. dasycarpum* (Figure

2A) and in almost all *L. elegans* population (except population 62) (Figure 2B). For the ITS region, the most frequent haplotype for *L. dasycarpum* was Hn01 (Figure 3A) occurring in 23 of 30 populations, as for *L. elegans* Hn07 was the most abundant, being present in 22 of 26 populations (Figure 4A).

Analysis of molecular variance showed a high differentiation among populations for chloroplast region for *L. dasycarpum* ($\Phi_{ST} = 0.672^{p<0.001}$) and *L. elegans* ($\Phi_{ST} = 0.752^{p<0.001}$). On the other hand, the ITS region showed lower levels of population structure, presenting a Φ_{ST} value of $0.246^{p<0.001}$ and $0.203^{p<0.001}$ for *L. dasycarpum* and *L. elegans*, respectively.

Demographic history

Mantel tests show different patterns of spatial genetic structure within the same species according to the analyzed region. For cpDNA, there was a significant correlation between genetic and geographic distance in *L. dasycarpum* (r = 0.367; $p < 0.001$) but not for *L. elegans* (r = 0.012; p = 0.333). By contrast, the Mantel tests with ITS region indicated significant correlation for *L. elegans* (r = 0.55; p = 0.040) rather than *L. dasycarpum* (r = 0.016; p = 0.312). Likewise, spatial autocorrelations illustrated that pairwise genetic distances on the cpDNA region of *L. dasycarpum* were significantly higher than random expectations as geographic distances became larger than 1500km (Figure S3, Supporting Information). Regarding to *L. elegans*, although there were some distance classes that showed significant deviations from the mean value, the overall distogram illustrated no monotonically increasing values as geographical distances increased (Figure S3).

The spatial autocorrelation analyses of the ITS region showed similar patterns to those found in Mantel test with no (or minimal) relation between spatial and genetic distances in *L. dasycarpum* populations. *Leptolobium elegans* showed distance classes that deviated from the mean pairwise genetic distance value as the geographical distance

increases. However, in distances classes above 1000 km, the genetic distances were significantly smaller than the average (Figure S3).

Tajima's (1989) D, Fu & Li (1993) F* and D*, and Fu's (1997) F_S neutrality tests showed non-significant values for cpDNA and ITS regions in both species (Table S7, Supporting Information). A single wave in the mismatch distribution was found in *L. dasycarpum* and *L. elegans* using cpDNA sequences. For the ITS region, we gathered a multimodal distribution for *L. dasycarpum* and a unimodal for *L. elegans* (Figure S2, Supporting Information).

The evolution model inferred for concatenated cpDNA region, for both species, was F81. For the ITS region, the TrN+G substitution model was used for *L. dasycarpum* sequences, and the TrN model for *L. elegans*. Coalescent analysis with LAMARC showed low values of θ for all populations and overall populations of both species using the two genetic regions. We also found positive growth parameter *g* values, but the 95% confidence intervals encompassed zero (Tables S8 and S9, Supporting Information), which indicates that no population growth occurred during most of the history of these species. Finally, gene flow among all population pairs for the two species, considering the two genetic regions, was negligible (< 1.0 migrant per generation; Tables S10–S13, Supporting Information). Meanwhile, the EBSP indicated that both species experienced a historical demographic expansion, with evidences of slight population decline followed by recent population expansion (Figures 5 and 6).

The coalescence estimates of the divergence time of *L. dasycarpum* and *L. elegans* lineages showed a lack of a strong genetic structure between the haplotypes, both regarding the cpDNA (Figure S4) as the ITS region (Figure S5). Besides the low node supports, it is not clear whether the haplotype grouping reflects on its geographical distribution. Furthermore, the divergence time found for the groups was very recent, ranging from slightly more than 100 kya to about 750 kya.

Species distribution modeling

The Maximum Entropy Modeling revealed that the potential distribution of *L. dasycarpum* and *L. elegans* in the current time (1950–2000) and in the Last Interglacial Period (ca. 130 kya) were quite similar, with some retraction in the boundaries (Figure 7). In the Last Glacial Maximum period, both species shrank their distributions had been confined in the central-east portion of Brazil (Figure 7). The AUC values for all modeling scenarios were above 0.90 (data not shown).

It was also evident that the two species share their distributions, i.e. have an ecological niche overlap. The modelling of niche similarity between *L. dasycarpum* and *L. elegans* produced moderate Schoener's D indexes (0.4–0.6), supported by the null models of the Identity Tests (Figure 8). Therefore, Cerrado and Seasonally Dry Tropical Forests (SDTF) areas in the central-east Brazil may be considered as potential refuges for the legume-trees here studied in the glacial periods.

Morphological traits

The PCA analysis (Figure 9) showed that, in most analyzed individuals, there are clear differences in vegetative characters of *L. dasycarpum* and *L. elegans*, but some of them had overlapping features. The first two components amounted 88.36% of the variation, with the first component explaining 69.73%. The leaflet area (47.87%), width (47.60%), perimeter (47.55%) and length (46.64%) and were the variables with the higher weight in Principal Component 1. The Dissection index (-14.21%) and the length of the petiolule (-28.30%) had negligible contributions in PC1. This means that, among the vegetative characters, the size of the leaflet is the characteristic that best distinguish the species.

Discussion

Genetic diversity

Leptolobium elegans showed no unique haplotypes of cpDNA, and of its six ITS haplotypes, three were shared with *L. dasycarpum*, mainly among nearby populations in the sympatric areas (Figures 3 and 4), suchs as populations 9 and 48, 32 and 64, 36 and 65 (Table S5). This sharing pattern was also detected in ecological niche modeling (Figure 8) and leaf morphology (Figure 9). Sharing of polymorphisms across species boundaries was often explained as the result of secondary contact and introgressive hybridization (Ferris et al. 1983). In fact, in the populations 36 (*L. dasycarpum*) and 65 (*L. elegans*) (Taunay/MS) we found potential hydrides among the studied species: the leaf morphology of the specimens was characteristic of *L. elegans*, but the ovarian hairiness was more typical of *L. dasycarpum*.

In both *Leptolobium* species, we find morphological, geographical and genetic evidences of hybridization in the contact regions, but as *L. elegans* did not show unique cpDNA haplotypes, this further reinforces the idea of incomplete lineage sorting. Together with the retention of ancestral polymorphism, the incomplete lineage sorting and have been investigated as alternative hypotheses for haplotype sharing among closely related species (Moran & Kornfield 1993, Avise 2000, Knowles & Carstens 2007). Lineage sorting is the elimination of ancestral polymorphism in a way that sister species become monophyletic. A polymorphic ancestor with large effective population size and multiple divergent haplotypes may give rise to sister species, two or more, which inherit the same polymorphisms.

Phylogeographic patterns and palaeoclimatic models

The genetic diversity found in populations of both species was not geographically structured. Both BAPS and SAMOVA analyze formed groups of broadly distribution and/or between populations distant from each other (Figures S6 and S7). The presence of a region with high genetic diversity may be associated with the maintenance of one large population for a long period of time, or the existence of glacial refuges (Widmer & Lexer 2001), whereas areas with low genetic variation are interpreted as recently colonized populations, via founder

effect, as is evidenced by the high number of monomorphic populations (Figure 2, 3 and 4). The eastern region of Brazil (i.e., South of Bahia and north of Minas Gerais states) has been characterized in many studies as the greatest genetic diversity region for legume tree species such as *Hymenaea courbaril* L. (Ramos et al. 2009), *Plathymenia reticulata* Benth. (Novaes et al. 2010), *Dalbergia nigra* (Vell.) Allemão ex Benth. (Ribeiro et al. 2011) and *Schizolobium parahyba* (Vell.) S.F. Blake (Turchetto-Zolet et al. 2012). On the other hand, the midwestern and southeastern regions of Brazil (i.e., Goiás, west of Minas Gerais and mainly Mato Grosso do Sul states) were poorly sampled and, as we have seen in this study, these regions also have a high genetic diversity, and therefore, coulb be a possible center of origin of *L. dasycarpum* and *L. elegans* since this region presented the highest gene diversity for these species.

The AMOVA of the cpDNA region showed a high genetic structure among populations of both studied species, and ecological and life history factors may have effects on this structure. Self-fertilizing breeding system, pollination by sedentary animals and limited seed dispersal are some of the factors that influence plant population genetic structure (Avise 2004). Comparing the three legume subfamilies, self-incompatibility is proportionately less frequent in Papilionoideae (Arroyo 1981). On the other hand, *Leptolobium dasycarpum* is self-incompatible (Facharo & Menino in prep.), therefore its congeneric *L. elegans* may also has the same reproductive mechanism. Likewise, the species present a massive flowering, providing many flowers in a relatively short period (see Chapter 1). In some Leguminosae species — e.g., *Inga* spp. Mill. — the pollinators are thought to show territorial or sedentary behavior (Cruz-Neto et al. 2015), which may result in pollen flow between flowers of the same individual and, in self-incompatible species, this may result in a low fruit set. Furthermore, the two species studied here are anemochoric with limited dispersion (Lorenzi 1998, Alves et al. 2013), which may explain the presence of many monomorphic populations (only one haplotype present).

Similarly to our findings, most studies of legumes species that used cpDNA (e.g., psbC–trnS, trnL intron, trnL–trnF, trnS–trnG) and ITS markers found high levels of differentiation among populations, and were phylogeographically structured, which is not true for the *Leptolobium* species. Other legume species, e.g., *Hymenaea stigonocarpa* Mart. ex Hayne (Caesalpinioideae), *Plathymenia reticulata* (Mimosoideae), and *Dalbergia miscolobium* Benth. (Papilionoideae), formed major groups of populations subdivided longitudinally and exhibited signs of recent range expansion of northern resources towards the southern Cerrado (Ramos et al. 2007, Novaes et al. 2010, 2013). Although the *Leptolobium* species had high genetic diversity, we did not find clear signs of geographic expansion, but the region of the Cerrado with the higher genetic diversity seems to be the same for most studies.

The central, central western and central eastern regions of the Cerrado (which comprises, respectively, the states of Goiás, Mato Grosso and Mato Grosso do Sul, Minas Gerais and São Paulo) formed the main region with the greatest diversity of both cpDNA and ITS for *Leptolobium dasycarpum* and *L. elegans*. As the demographic analyzes of the studied species were not in accordance with a recent population expansion, it may indicate that they might have been more stable during their life histories. In fact, the ENM for both species demonstrated that during the LGM the potential distribution of the species shrank, and the region with the higher index of habitat suitability coincided with the principal Cerrado refugium area, as proposed by Werneck et al. (2012).

Species associated with open vegetation domains seemed to show variable responses to climatic oscillations, with a tendency to expand, to maintain or to shrink their geographical distribution ranges during glacial cycles (Turchetto-Zolet et al. 2013). Our results showed that both species studied were not deeply affected by glaciations, without geographically structured groups of populations and a probable hybridization zone in the

contact areas. Based on that, we reinforce the important role that the central and estern portion of Cerrado in maintaining the genetic diversity of widespread tree species.

Acknowledgments

Coordination for the Improvement of Higher Education Personnel (CAPES) for WS Fava's Ph.D fellowship. National Counsel of Technological and Scientific Development (CNPq) — Edital MCT/CNPq/FNDCT/FAPs/MEC/CAPES/PRO-CENTRO-OESTE n° 31/2010 — and Rede GENPAC (Genética Geográfica e Planejamento Regional para Conservação de Recursos Naturais no Cerrado) for financial and logistical support. MPC Telles and T Castro from Federal University of Goiás for the sequencing of samples.

References

Ab' Sáber AN. 1983. O domínio dos cerrados: introdução ao conhecimento. Revista do Serviço Público 111: 41–55.

Abràmoff MD, Magalhães PJ & Ram SJ. 2004. Image processing with ImageJ. Biophotonics International 11: 36–42.

Alvarado-Serrano DF & Knowles LL. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. Molecular Ecology Resources 14: 233–248.

Alves HR, Júnior JP, Lopes S, Silva PP, Peppe FB & Schiavini I. 2013. Fitossociologia e grupos ecológicos da comunidade lenhosa em um remanescente de cerradão em Uberlândia, MG. Caminhos de Geografia 14: 236–245.

Araújo MB, Pearson RG, Thuiller W & Erhard M. 2005. Validation of species-climate impact models under climate change. Global Change Biology 11: 1504–1513.

Arroyo MTK. 1981. Breeding systems and pollination biology in Leguminosae. In: Advances in legume systematics (Eds, RM Polhill & PH Raven), pp. 723–769. Royal Botanic Gardens, Kew.

Avise JC. 2000. Phylogeography: the history and formation of species. Harvard University Press, London, 447 pp.

Avise JC. 2004. Molecular markers, natural history and evolution. Sinauer Associates, Inc. USA, 694 pp.

Bandelt HJ, Forster P & Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16: 37–48.

Beerli P & Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in subpopulations by using a coalescent approach. Proceedings of the National Academy of Sciences of the USA 98: 4563–4568.

Behling H & Hooghiemstra H. 2001. Neotropical savanna environments in space and time: late Quaternary interhemispheric comparison. In: Interhemispheric climate linkages (Ed. V Markraf), pp. 307–323. Academic Press, New York.

Behling H & Lichte M. 1997. Evidence of dry and cold climatic conditions at glacial times in tropical southeastern Brazil. Quaternary Research 48: 348–358.

Behling H. 1995. A high resolution Holocene pollen record from Lago do Pires, SE Brazil: vegetation, climate and fire history. Journal of Paleolimnology 14: 253–268.

Behling H. 1998. Late quaternary vegetational and climatic changes in Brazil. Review of Paleobotany and Palynology 99: 143–156.

Behling H. 2002. South and southeast Brazilian grasslands during late quaternary times: a synthesis. Palaeogeography Palaeoclimatology Palaeoecology 177: 19–27.

Beyra-M A & Lavin M. 1999. Monograph of *Pictetia* (Leguminosae-Papilionoideae) and review of the Aeschynomeneae. Systematic Botany Monographs 56: 1–93.

Cardoso D, de Lima HC, Rodrigues RS, de Queiroz LP, Pennington RT & Lavin M. 2012. The Bowdichia clade of Genistoid legumes: Phylogenetic analysis of combined molecular and morphological data and a recircumscription of *Diplotropis*. Taxon 61: 1074–1087.

Clement M, Posada D & Crandall K. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657–1660.

Collevatti RG, Estolano R, Ribeiro ML, Rabelo SG, Lima EJ & Munhoz CB. 2014. High genetic diversity and contrasting fine-scale spatial genetic structure in four seasonally dry tropical forest tree species. Plant Systematics and Evolution 300: 1671–1681.

Collevatti RG, Terribile LC, Lima-Ribeiro MS, Nabout JC, Oliveira G, Rangel TF, Rabelo SG & Diniz-Filho JA. 2012. A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species. Molecular Ecology 21: 5845–5863.

Corander J, Siren J & Arjas E. 2008. Bayesian spatial modeling of genetic population structure. Computational Statistics 23: 111–129.

Cruz-Neto O, Machado IC, Galetto L & Lopes AV. 2015. The influence of nectar production and floral visitors on the female reproductive success of *Inga* (Fabaceae): a field experiment. Botanical Journal of the Linnean Society 177: 230–245.

Cubas P, Pardo C & Tahiri H. 2006. Morphological convergence or lineage sorting? The case of *Cytisus purgans* auct. (Leguminosae). Taxon 55: 695–704.

Darriba D, Taboada GL, Doallo R & Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9: 772.

Desfeux C & Lejeune B. 1996. Systematics of Euromediterranean *Silene* (Caryophyllaceae): evidence from a phylogenetic analysis using ITS sequences. Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie 319: 351–358.

Diniz-Filho JA, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, Telles MPC, Rangel TF & Bini LM. 2013. Mantel test in population genetics. Genetics and Molecular Biology 36: 475–485.

Doyle JJ & Doyle JL. 1987. Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

Drummond, AJ, Suchard MA, Xie D & Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29: 1969–1973.

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. Molecular Ecology 11: 2571–2581.

Eiten G. 1972. The Cerrado vegetation of Brazil. Botanical Review 38: 201–345.

Excoffier L & Lisxher HEL. 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.

Fachardo ALS & Menino GCO. Competition for nectar between pollinators and nectar robbers, and the breeding system of *Leptolobium dasycarpum* (Vogel) Fabaceae in a Cerrado fragment. In preparation.

Ferris SD, Sage RD, Huang CM, Nielsen JT, Ritte U & Wilson AC. 1983. Flow of mitochondrial DNA across a species boundary. Proceedings of the National Academy of Sciences 80: 2290–2294.

Fielding AH & Bell JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/ absence models. Environmental Conservation 24: 38–49.

Flot J-F. 2010. SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments Molecular Ecology Ressources 10: 162–166.

Fu YX & Li WH. 1993. Statistical tests of neutrality of mutations. Genetics 133: 693–709.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915–925.

Fuertes Aguilar J & Nieto Feliner G. 2003. Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). Molecular Phylogenetics and Evolution 28: 430–447.

Fuertes Aguilar J, Rosselló JÁ & Nieto Feliner G. 1999. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Molecular Ecology 8: 1341–1346.

Giorgetta MA, Jungclaus J, Reick CH, Legutke S, Bader J, Böttinger M, Brovkin V, Crueger T, Esch M, Fieg K, Glushak K, Gayler V, Haak H, Hollweg H-D, Ilyina T, Kinne S, Kornblueh L, Matei D, Mauritsen T, Mikolajewicz U, Mueller W, Notz D, Pithan F, Raddatz T, Rast S, Redler R, Roeckner E, Schmidt H, Schnur R, Segschneider J, Six KD, Stockhause M, Timmerck C, Wegner J, Widmann H, Wieners K-H, Claussen M, Marotzke J & Stevens B. 2013. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. Journal of Advances in Modeling Earth Systems 5: 572–597.

Hanley JA & McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143: 29–36.

Hasumi H & Emori S. 2004. K-1 coupled gcm (MIROC) description. Center for Climate System Research, University of Tokyo, Tokyo.

Heled J & Drummond AJ. 2008. Bayesian inference of population size history from multiple loci. BMC Evolutionary Biology 8: 289.

Hijmans RJ, Cameron SE, Parra JL, Jones PG & Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965–1978.

Javadi F, Wojciechowski MF & Yamaguchi H. 2007. Geographical diversification of the genus *Cicer* (Leguminosae: Papilionoideae) inferred from molecular phylogenetic analyses

of chloroplast and nuclear DNA sequences. Botanical Journal of the Linnean Society 154: 175–186.

Kincaid DT & Schneider RB. 1983. Quantification of leaf shape with a microcomputer and Fourier transform. Canadian Journal of Botany 61: 2333–2342.

Kropf M, Kadereit JW & Comes HP. 2002. Late Quaternary distributional stasis in the submediterranean mountain plant *Anthyllis montana* L. (Fabaceae) inferred from ITS sequences and amplified fragment length polymorphism markers. Molecular Ecology 11: 447–463.

Kuhner MK & Smith LP. 2007. Comparing likelihood and Bayesian coalescent estimation of population parameters. Genetics 175: 155–165.

Kuhner MK. 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22: 768–770.

Knowles LL & Carstens BC. 2007. Delimiting species without monophyletic gene trees. Systematic Biology 56: 887–895.

Librado P & Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.

Lorenzi H. 1998. Árvores Brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. vol. 2. Nova Odessa, Brazil: Instituto Plantarum de Estudos da Flora Ltda. 352pp.

Lorenz-Lemke AP, Muschner VC, Bonatto SL, Cervi AC, Salzano FM & Freitas LB 2005. Phylogeographic inferences concerning evolution of Brazilian *Passiflora actinia* and *P. elegans* (Passifloraceae) based on ITS (nrDNA) variation. Annals of Botany 95: 799–806.

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209–220.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA & Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature: 403: 853–858.

Miller MP. 2005. Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. Journal of Heredity 96: 722–724.

Moran P & Kornfield I. 1993. Retention of an ancestral polymorphism in the Mbuna species flock (Teleostei: Cichlidae) of Lake Malawi. Molecular Biology and Evolution, 10: 1015–1015.

Novaes RML, de Lemos Filho JP, Ribeiro RA & Lovato MB. 2010. Phylogeography of *Plathymenia reticulata* (Leguminosae) reveals patterns of recent range expansion towards northeastern Brazil and southern Cerrados in Eastern Tropical South America. Molecular Ecology 19: 985–998.

Novaes RML, Ribeiro RA, Lemos-Filho JP, Lovato MB. 2013. Concordance between phylogeographical and biogeographical patterns in the Cerrado: diversification of the endemic tree *Dalbergia miscolobium* (Fabaceae). PLoS ONE 8: e82198.

Odee DW, Telford A, Wilson J, Gaye A & Cavers S. 2012. Plio-Pleistocene history and phylogeography of *Acacia senegal* in dry woodlands and savannahs of sub-Saharan tropical Africa: evidence of early colonisation and recent range expansion. Heredity, 109: 372–382.

Oliveira PS & Marquis RJ. 2002. The cerrados of Brazil: ecology and natural history of a neotropical savanna. Columbia University Press. New York. 409pp.

Oliveira-Filho AT & Ratter JA. 2002. Vegetation physiognomies and woody flora of the Cerrado biome. In: The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna (Eds PS Oliveira & RJ Marquis), pp. 91–120. Columbia University Press, New York.

Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH & Hu A. 2006. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. Science 311: 1751–1753.

Parker MA, Doyle JL & Doyle JJ. 2004. Comparative phylogeography of *Amphicarpaea* legumes and their root-nodule symbionts in Japan and North America. Journal of Biogeography 31: 425–434.

Person K. 1901. On lines and planes of closest fit to system of points in space. Philiosophical Magazine 2: 559–572.

Phillips SJ, Anderson RP & Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231–259.

Prado DE & Gibbs PE. 1993. Patterns of species distributions in the dry seasonal forest South America. Annals of the Missouri Botanical Garden 80: 902–927.

Premoli AC, Quiroga MP, Mathiasen P & Kitzberger T. 2012. Ecological niche modeling meets phylogeography to unravel hidden past history of key forest genera in plant geography: *Podocarpus* and *Nothofagus*. Natureza & Conservação 10: 160–168.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rambaut A, Suchard MA, Xie D & Drummond AJ. 2014. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.

Ramdhani S, Cowling RM & Barker NP. 2010. Phylogeography of *Schotia* (Fabaceae): recent evolutionary processes in an ancient thicket biome lineage. International Journal of Plant Sciences 171: 626–640.

Ramos ACS, Lemos-Filho JP, Ribeiro RA, Santos FR & Lovato MB. 2007. Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of Quaternary climate changes in the Cerrado. Annals of Botany 100: 1219–1228.

Ramos ACS, Lemos-Filho JP & Lovato MB. 2009. Phylogeographical structure of the neotropical forest tree *Hymenaea courbaril* (Leguminosae: Caesalpinioideae) and its relationship with the vicariant *Hymenaea stigonocarpa* from Cerrado. Journal of Heredity 100: 206–216.

Ribeiro RA, Lemos-Filho JP, Ramos ACS & Lovato MB. 2011. Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. Heredity 106: 46–57.

Rizzini CT. 1997. Tratado de fitogeografia do brasil: aspectos ecologicos, sociologicos e floristicos, 2nd ed. Âmbito Cultural, Rio de Janeiro, 747pp.

Rödder D & Engler JO. 2011. Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. Global Ecology and Biogeography 20: 915–927.

Rodrigues RS & Tozzi AMGDA. 2012. A taxonomic revision of *Leptolobium* (Papilionoideae, Leguminosae). Acta Botanica Brasilica 26: 146–164.

Rogers AR & Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution 9: 552–569.

Sang T, Crawford DJ & Stuessy TF. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proceedings of the National Academy of Sciences, USA 92: 6813–6817.

Schneider CA, Rasband WS & Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9: 671–675.

Schoener TW. 1968. The Anolis lizards of Bimini: resource partitioning in a complex fauna. Ecology 49: 704–726.

Shin SI, Liu Z, Otto-Bliesner B, Brady E, Kutzbach J & Harrison S. 2003. A simulation of the Last Glacial Maximum climate using the NCAR-CCSM. Climate Dynamics 20: 127–151.

Soltis DE, Soltis PS & Doyle JJ. 1998. Molecular Systematics of Plants II: DNA Sequencing. Kluwer Academic Publishers, Boston, 585pp.

Spearman C. 1904. The proof and measurement of association between two things. The American Journal of Psychology 15: 72–101.

Stephens M, Smith NJ & Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68: 978–989.

Svenning JC, Flojgaard C, Marske KA, Nogues-Bravo D & Normand S. 2011. Applications of species distribution modeling to paleobiology. Quaternary Science Reviews 30: 2930–2947.

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

Tajima F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Tamura K, Stecher G, Peterson D, Filipski A & Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30: 2725–2729.

Templeton AR, Crandall KA & Sing CF. 1992. A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. III. Cladogram estimation. Genetics 132: 619–633.

Thompson JD, Higgins DG & Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673–4680.

Turchetto-Zolet AC, Cruz F, Vendramin GG, Simon MF, Salgueiro F, Margis-Pinheiro M & Margis R. 2012. Large-scale phylogeography of the disjunct Neotropical tree species *Schizolobium parahyba* (Fabaceae-Caesalpinioideae). Molecular Phylogenetics and Evolution 65: 174–182.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F & Palma-Silva C. 2013. Phylogeographical patterns shed light on evolutionary process in South America. Molecular Ecology 22: 1193–1213.

Viana e Souza HA & Lovato MB. 2010. Genetic diversity and structure of the critically endangered tree *Dimorphandra wilsonii* and of the widespread in the Cerrado *Dimorphandra mollis*: implications for conservation. Biochemical Systematics and Ecology 38: 49–56.

Warren DL, Glor RE & Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33: 607–611.

Warren DL. 2012. In defense of ‘niche modeling’. Trends in Ecology & Evolution 27: 497–500.

Weising K, Nybom H, Wolff K & Kahl G. 2005. Applications of DNA Fingerprinting in Plant Sciences. In: DNA fingerprinting in plants: principles, methods, and applications (Ed. K Weising), pp. 270–276. CRC press. Boca Raton.

Werneck FP, Nogueira C, Colli GR, Sites JW & Costa GC. 2012. Climatic stability in the Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. Journal of Biogeography 39: 1695–1706.

Widmer A, Lexer C. 2001. Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. Trends in Ecology and Evolution 16: 267–269.

Wolfe KH, Li WH & Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of the National Academy of Sciences, USA 84: 9054–9058.

Wright S. 1943. Isolation by distance. Genetics 28: 114–138

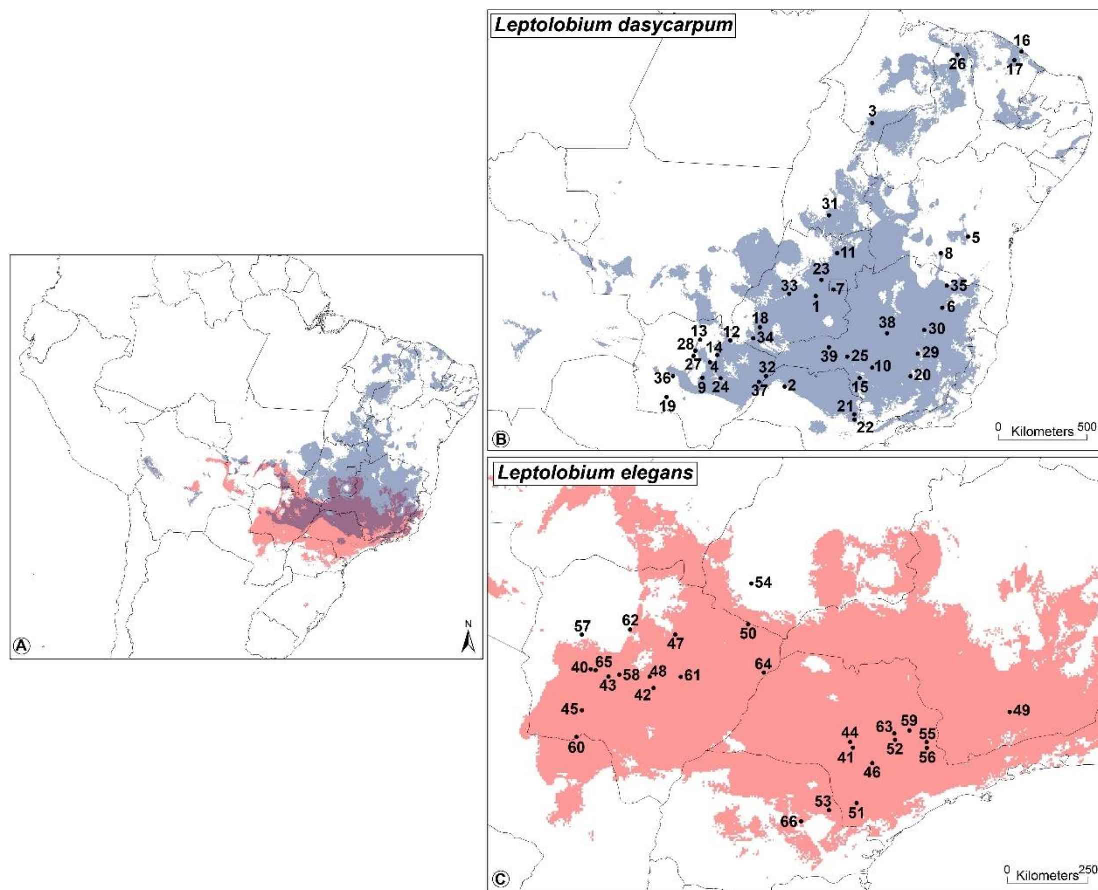

Figure 1. Geographical distribution of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae). A: Potential distribution of *L. dasycarpum* (blue area) and *L. elegans* (red area) for the current time (1950–2000) based on the maximum entropy model. Sampled populations of *L. dasycarpum* (B) and *L. elegans* (C) (more details can be obtained in Table S1).

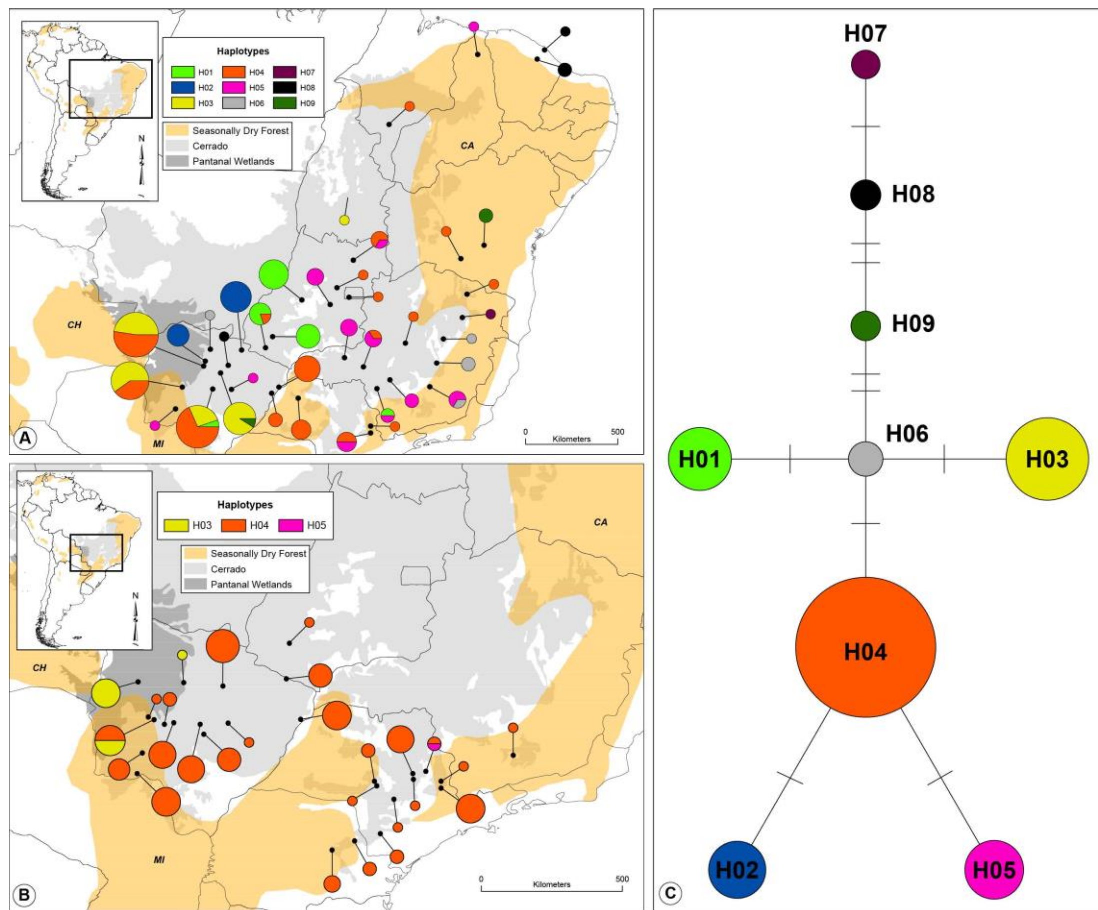

Figure 2. Distribution of *Leptolobium dasycarpum* (A) and *L. elegans* (B) cpDNA haplotypes; median-joining network (C). Circles sizes are proportional to sampling. Seasonally dry forests areas were obtained from Collevatti et al. 2012.

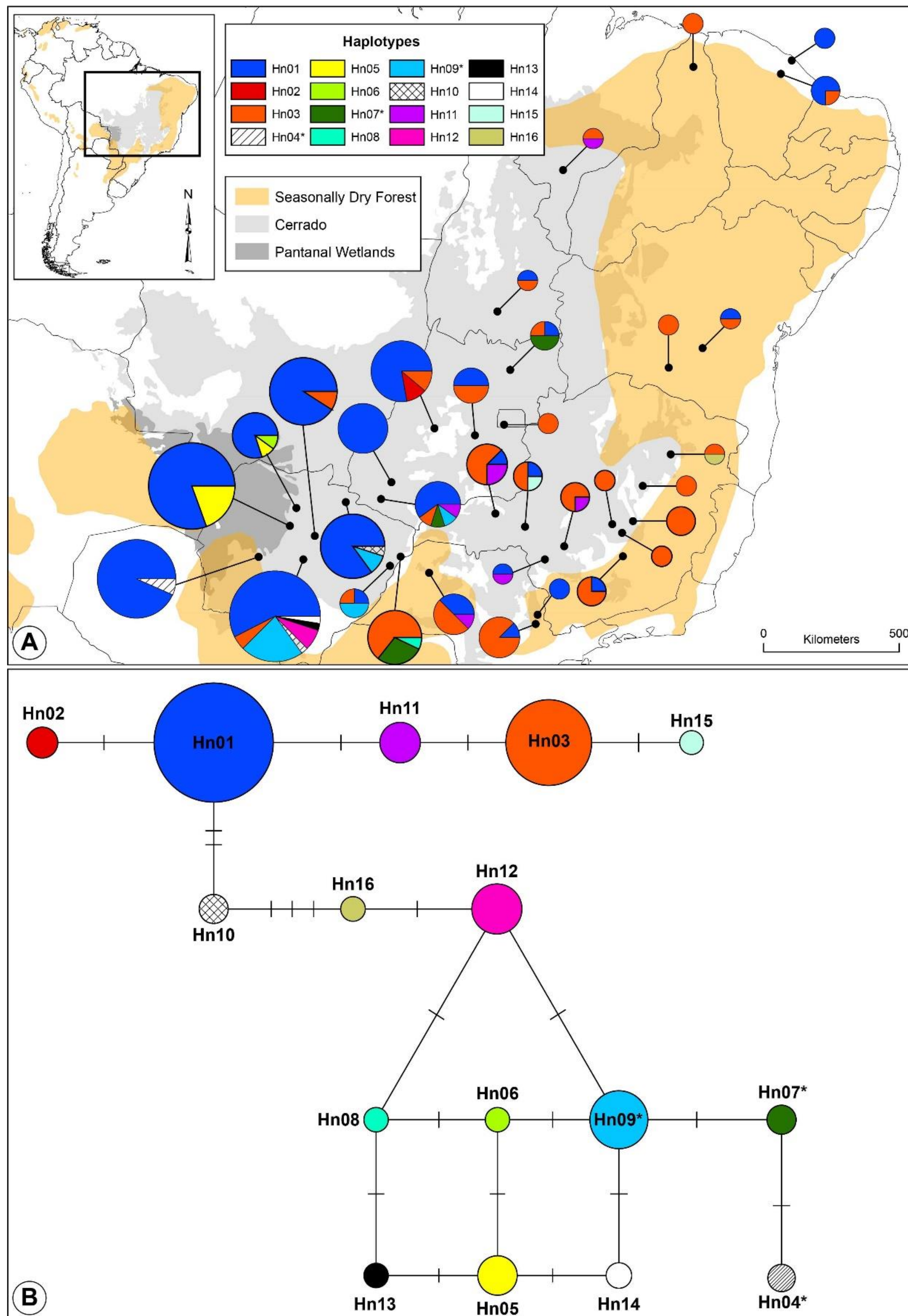

Figure 3. Distribution of *Leptolobium dasycarpum* nrDNA haplotypes (A); median-joining network (B). Circles sizes are proportional to sampling. Seasonally dry forests areas were obtained from Collevatti et al. 2012. * Haplotypes shared between the two species.

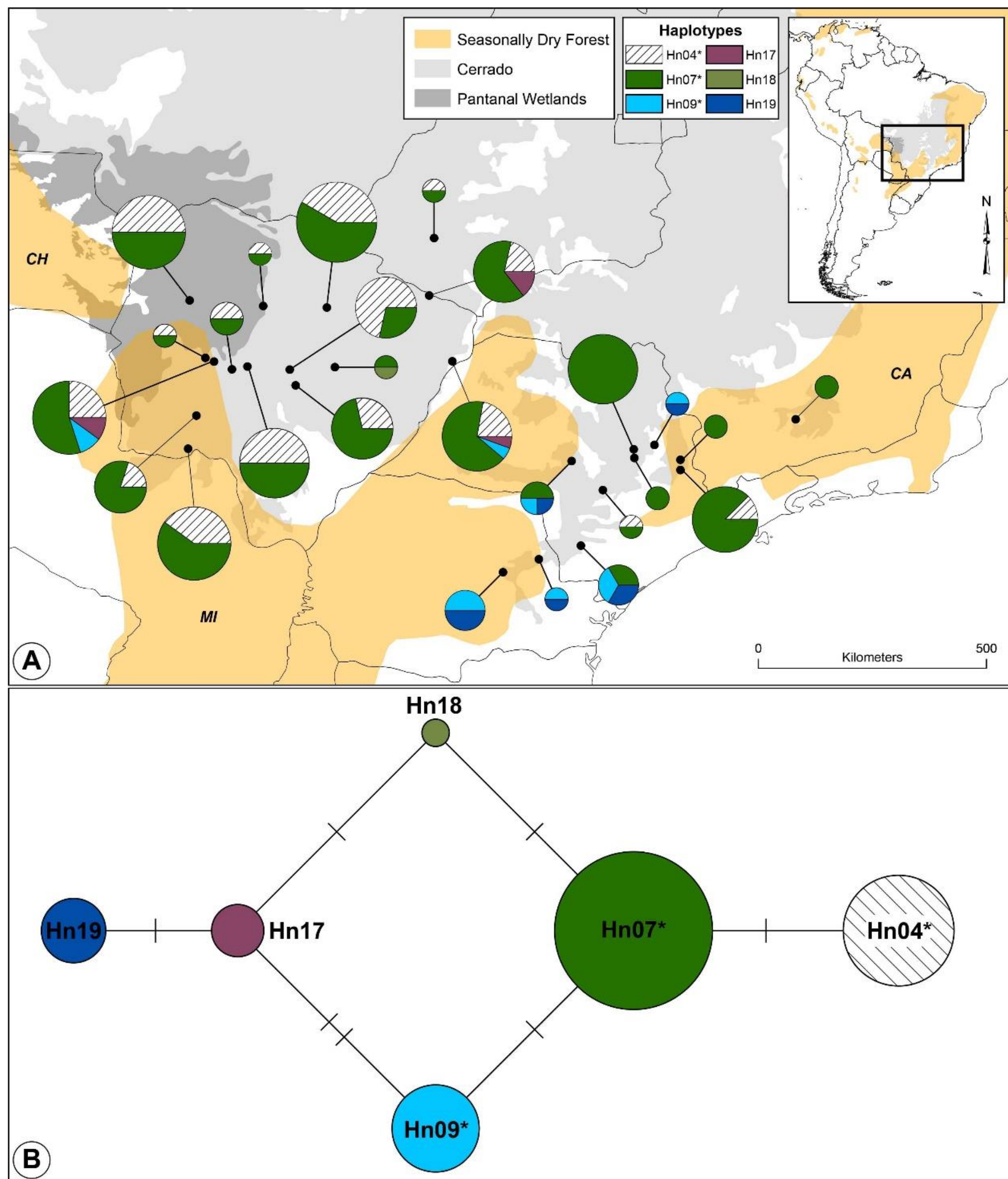

Figure 4. Distribution of *Leptolobium elegans* nrDNA haplotypes (A); median-joining network (B). Circles sizes are proportional to sampling. Seasonally dry forests areas were obtained from Collevatti et al. 2012. * Haplotypes shared between the two species.

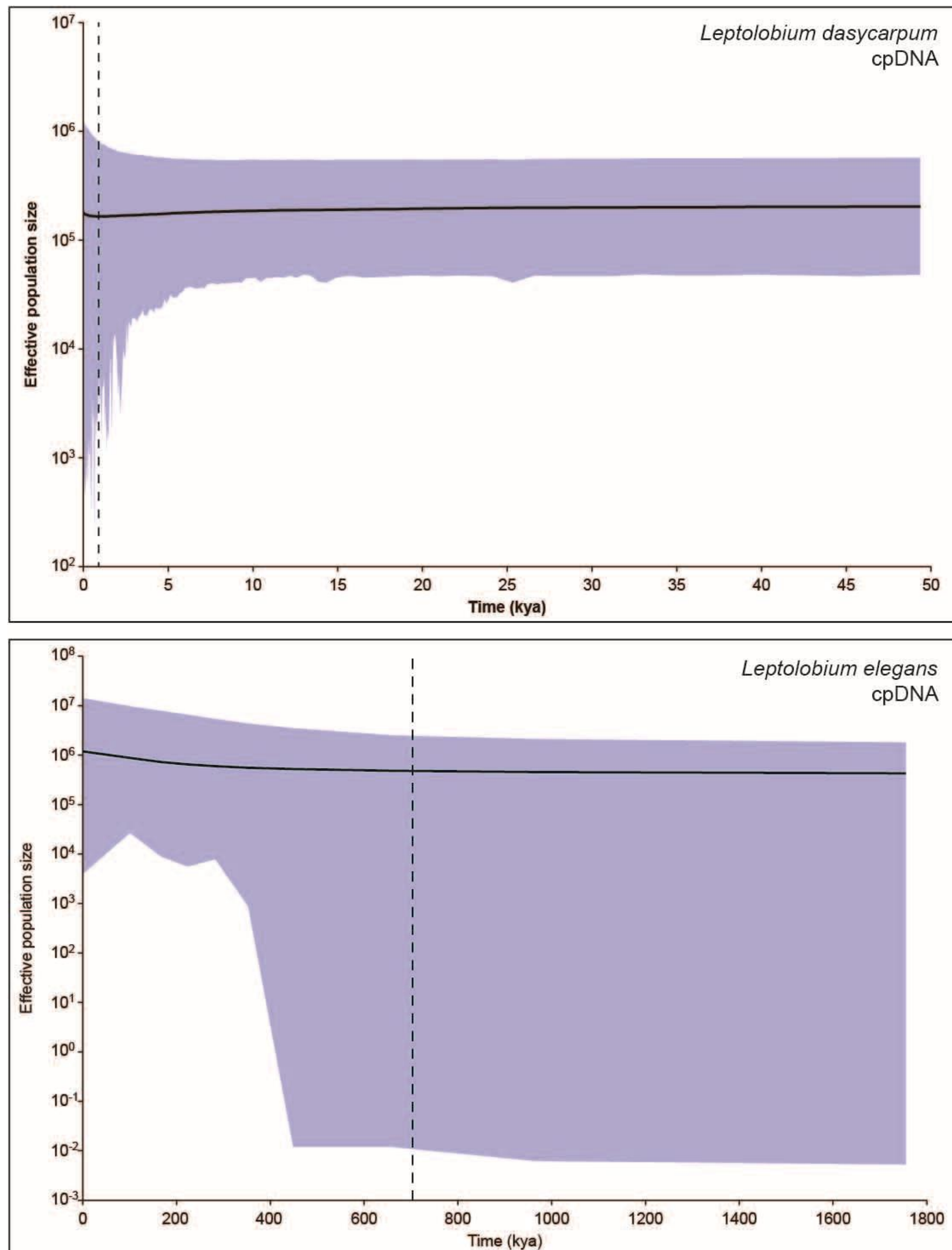

Figure 5. Extended Bayesian Skyline Plot of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) estimated with cpDNA sequences. The horizontal line shows the median estimate of the EBSP and the blue area shows the upper and lower 95% highest posterior density limits. The dotted vertical line illustrates the approximate time when the demographic expansion started. The Y-axis is in natural logarithm scale.

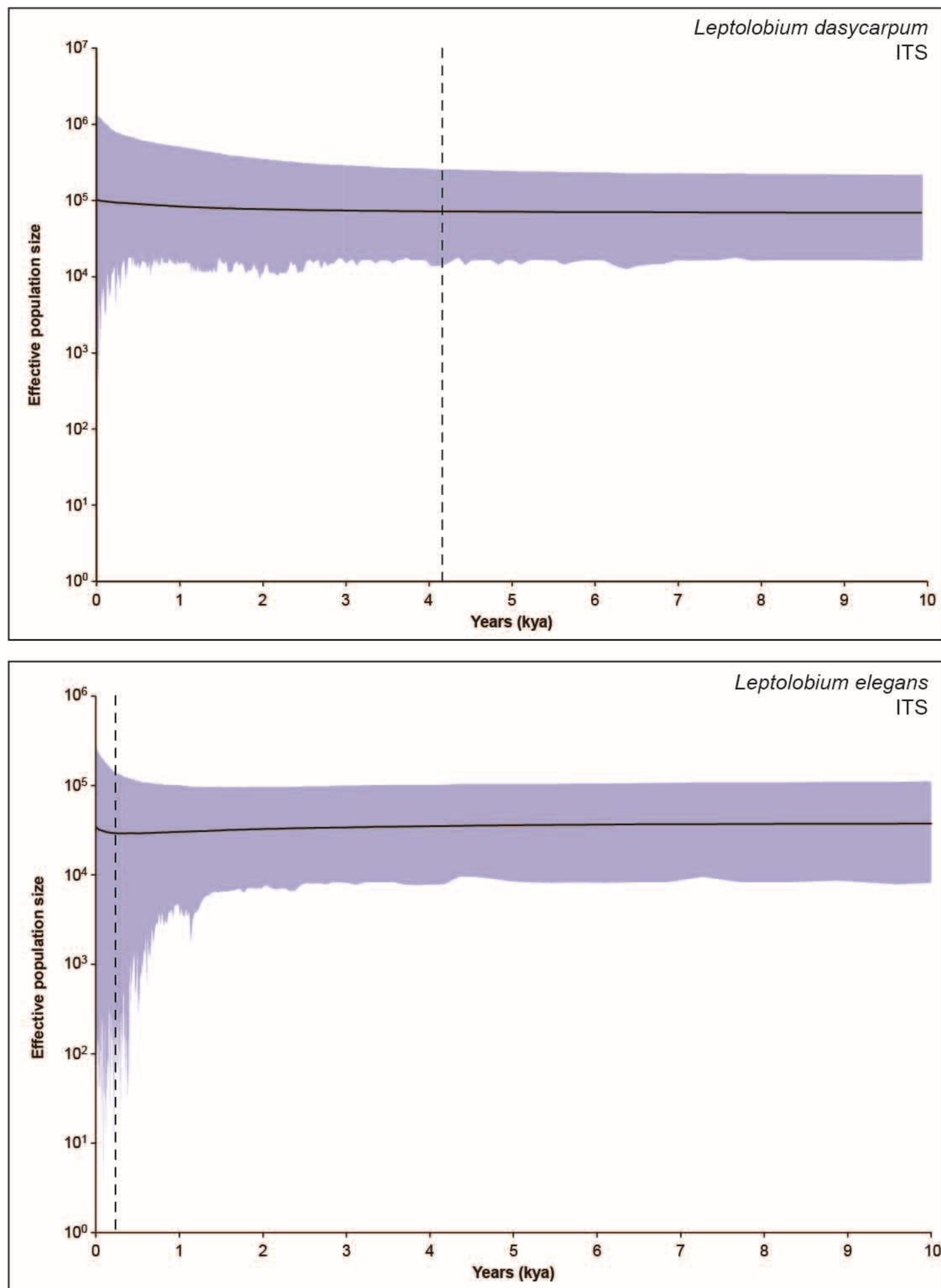

Figure 6. Extended Bayesian Skyline Plot of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) estimated with ITS sequences. The horizontal line shows the median estimate of the EBSP and the blue area shows the upper and lower 95% highest posterior density limits. The dotted vertical line illustrates the approximate time when the demographic expansion started. The Y-axis is in natural logarithm scale.

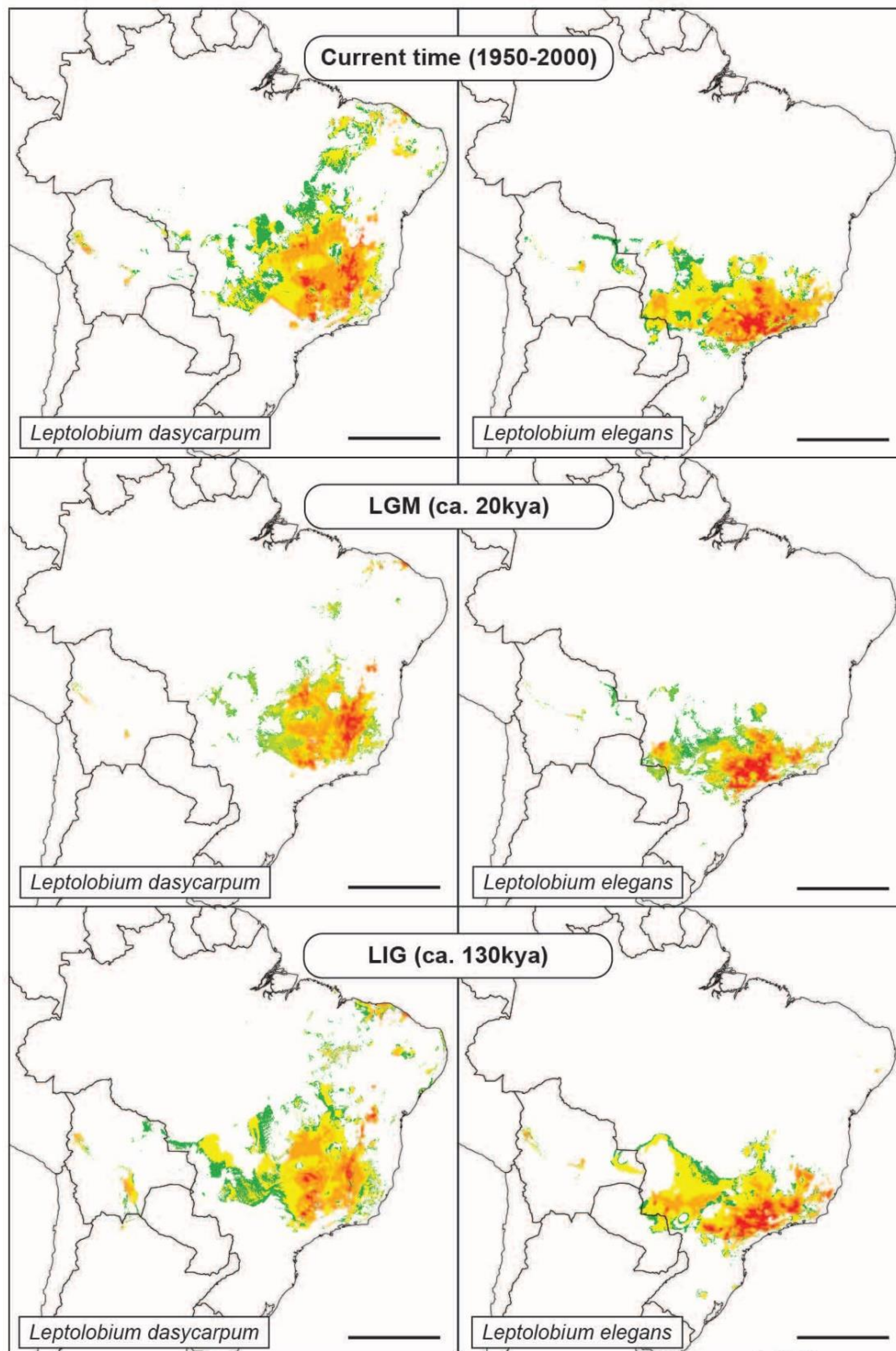

Figure 7. Modeled ranges of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) under definitions for the current (1950-2000), the Last Glacial Maximum (LGM), and the Last Interglacial (LIG) climatic scenarios. Warmer colors (red/yellow) of the logistic output correspond to regions with a higher index of relative habitat suitability, based on presence-only data using the maximum entropy algorithm. Scale bars in the lower right corner represent 500km.

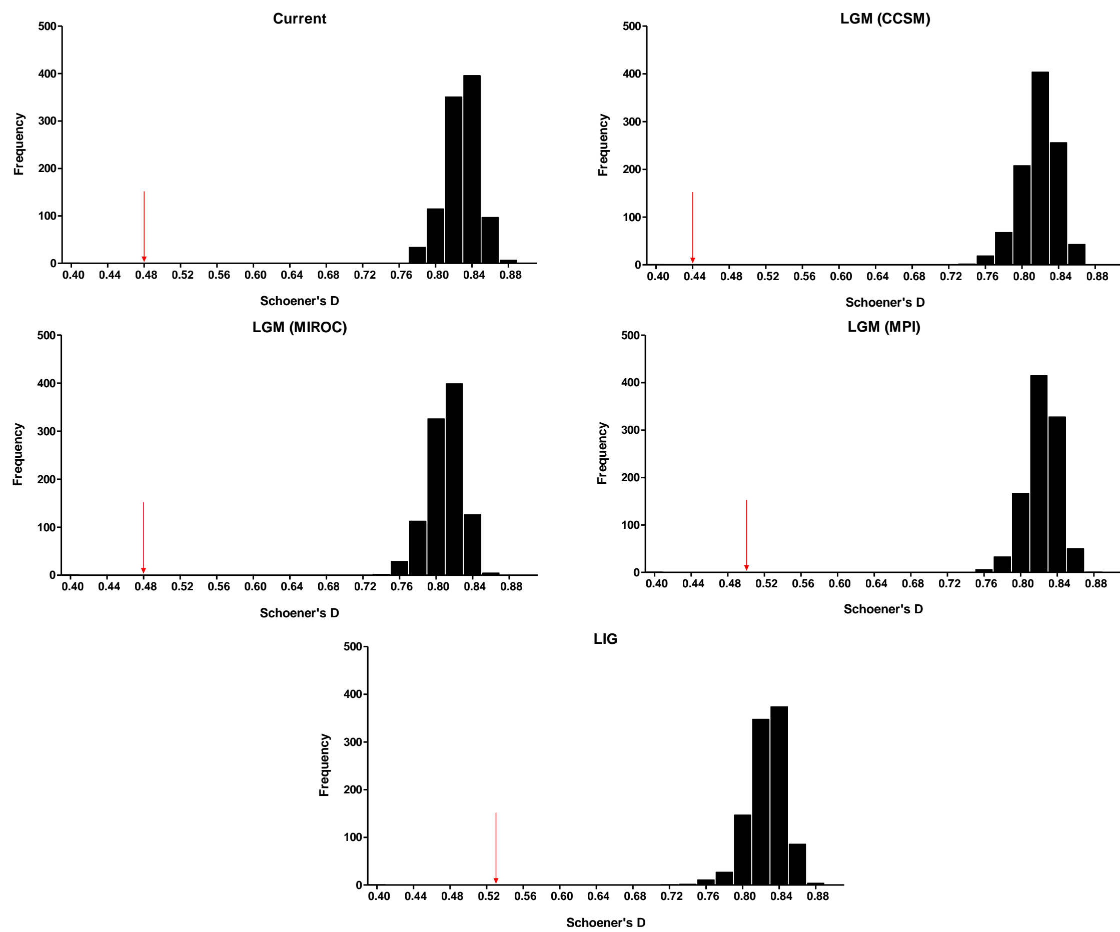

Figure 8. Ecological niche overlap between *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) based on five climatic scenarios: current time, three models for the Last Glacial Maximum (CCSM, MIROC and MPI), and the Last Interglacial period (LIG). The similarity score for the ecological niche modeling built for each climatic scenarios of the two species (red arrow) is much lower than expected based on the null hypothesis of niche equivalency (black bars), indicating that the two species' environmental niches are not equivalent.

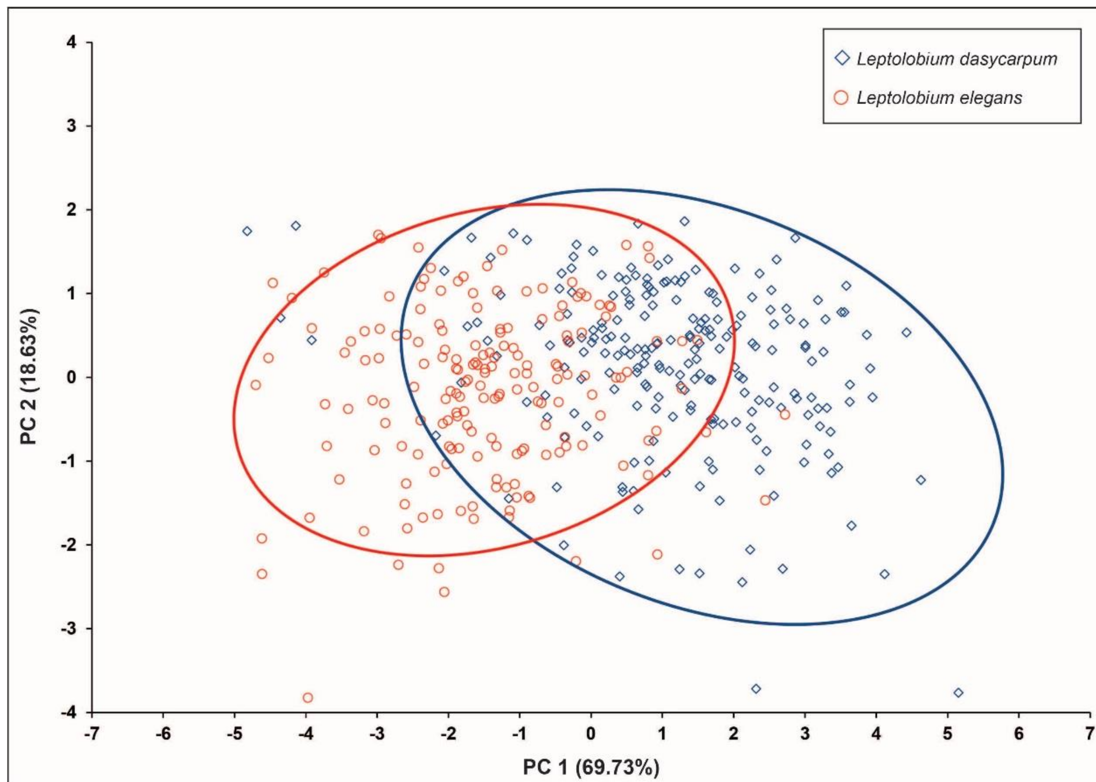

Figure 9. Principal Component Analysis of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) leaflets traits: area, perimeter, length and width of the blade, leaflet shape, and petiolule length.

Chapter 2

SUPPORTING INFORMATION

Table S1. Location, sample size and voucher numbers of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) populations used in the genetic analyses in this study. The names of the herbaria are given in Chapter 1, Supplement S1. The numbers to the left of the sampling locations represent the populations on Figure 1.

Sampling places*	Latitude	Longitude	Voucher	Sample size
Leptolobium dasycarpum				
1. Anápolis / GO	-16.3267	-48.9111	HUEG 5564	3
2. Araçatuba / SP	-20.8861	-50.4724	CGMS 37112	4
3. Balsas / MA	-7.5328	-46.0358	HST 16202	1
4. Bandeirantes / MS	-19.6568	-54.2613	CGMS 34703	11
5. Barra da Estiva / BA	-13.4317	-41.4056	UEC 123660	2
6. Berilo / MG	-16.9517	-42.4656	BHCB 69669	1
7. Brasília / DF	-15.9686	-47.9961	UEC 123577	1
8. Caetité / BA	-14.0833	-42.5333	UEC 131025	1
9. Campo Grande / MS	-20.4413	-54.6446	CGMS 36829	19
10. Capitólio / MG	-20.0000	-46.0011	HUFU 49366	2
11. Colinas do Sul / GO	-14.1522	-47.7842	UEC 123577	3
12. Costa Rica / MS	-18.5394	-53.2331	CGMS 34704	10
13. Coxim / MS	-18.5067	-54.7600	UEC 119473	1
14. Crato / CE	-19.2856	-53.8887	HST 8855	1
15. Delfinópolis / MG	-20.4344	-46.6311	UEC 156984	2
16. Eusébio / CE	-3.8900	-38.4506	EAC 35529	1
17. Fortaleza / CE	-4.3411	-38.8111	EAC 31507	2
18.Jataí / GO	-17.8797	-51.7270	CGMS 34705	6
19. Matozinhos / MG	-21.4157	-56.4584	BHCB 132249	1
20. Moeda / MG	-20.3328	-44.0528	BHCB 133933	3
21. Mogi Guaçu / SP	-22.3722	-46.9422	UEC 156365	1
22. Mogi Mirim / SP	-22.4344	-46.9425	CGMS 33643	4
23. Padre Bernardo / GO	-15.5000	-48.5833	UEC 7411	1
24. Paraopeba / MG	-20.4635	-53.7266	UEC 159811	1
25. Perdizes / MG	-19.3528	-47.2911	BHCB 24748	3
26. Piracuruca / PI	-4.1047	-41.7156	UEC 124887	1
27. Rio Negro / MS	-19.3309	-55.0914	CGMS 33646	22
28. Rio Verde de Mato Grosso / MS	-19.0727	-55.0094	CGMS 33645	5
29. Santana do Riacho / MG	-19.1689	-43.7144	BHCB 18913	2
30. São Gonçalo do Rio Preto / MG	-18.0039	-43.3908	BHCB 51493	1
31. São Valério da Natividade / TO	-12.2195	-48.2195	UB 6203	1
32. Selvíria / MS	-20.3481	-51.4135	CGMS 33648	7
33. Serra Dourada / GO	-16.0953	-50.2985	CGMS (Fava 049)	9
34. Serranópolis / GO	-18.4368	-52.0659	CGMS 34706	5
35. Taiobeiras / MG	-15.8078	-42.2331	BHCB 14895	1
36. Taunay / MS	-20.3535	-56.1281	CGMS 33495	15
37. Três Lagoas / MS	-20.6481	-51.7689	CGMS (Fava 048)	2
38. Três Marias / MG	-18.2064	-45.2417	BHCB 10784	1
39. Uberlândia / MG	-18.9186	-48.2111	HUFU 8070	3

* Brazilian States. BA = Bahia state; CE = Ceará; DF = Federal District (Goiás state); GO = Goiás; MA = Maranhão; MG = Minas Gerais; MS = Mato Grosso do Sul; PI = Piauí; SP = São Paulo; TO = Tocantins.

Table S1. (continued)

Sampling places*	Latitude	Longitude	Voucher	Sample size
Leptolobium elegans				
40. Agachi / MS	-20.2973	-56.2442	CGMS 36840	1
41. Agudos / SP	-22.4689	-48.9878	UEC 87307	1
42. Anhanduí / MS	-20.8191	-54.5006	CGMS (Fava 050)	6
43. Aquidauana / MS	-20.5019	-55.7547	CGMS 36857	2
44. Bauru / SP	-22.3150	-49.0608	FUEL 32044	2
45. Bonito / MS	-21.0745	-56.3803	CGMS (Fava 037)	5
46. Botucatu / SP	-22.8858	-48.44500	BHCB 124710	1
47. Camapuã / MS	-19.0087	-53.9063	CGMS 36849	12
48. Campo Grande / MS	-20.5068	-54.6152	CGMS (Fava 051)	8
49. Carrancas / MG	-21.4878	-44.6427	UEC 126324	1
50. Cassilândia / MS	-19.0497	-51.8712	CGMS 36853	6
51. Itapeva / SP	-23.9819	-48.8758	FUEL 21904	2
52. Itirapina / SP	-22.2528	-47.8227	UEC 147979	1
53. Jaguariaíva / PR	-24.2508	-49.7058	FLOR 28037	2
54. Jataí / GO	-17.9158	-51.7727	CGMS 36850	1
55. Mogi Guaçu / SP	-22.3719	-46.9419	UEC 121299	1
56. Mogi Mirim / SP	-22.4318	-46.9432	CGMS (Fava 026)	9
57. Nhecolândia / MS	-19.1449	-56.5918	CGMS (Fava 025)	9
58. Piraputanga / MS	-20.4403	-55.6238	CGMS 36837	8
59. Pirassununga / SP	-21.9911	-47.4258	UEC 84394	2
60. Porto Murtinho / MS	-22.0719	-56.6229	CGMS (Fava 052)	9
61. Ribas do Rio Pardo / MS	-20.4635	-53.7266	CGMS (Fava 054)	1
62. Rio Verde de Mato Grosso / MS	-19.1760	-55.1398	CGMS (Fava 029)	1
63. São Carlos / SP	-22.0738	-47.8411	CGMS (Fava 028)	2
64. Selvíria / MS	-20.3480	-51.4120	CGMS (Fava 030)	15
65. Taunay / MS	-20.3260	-56.1088	CGMS (Fava 053)	10
66. Tibagi / PR	-24.5089	-50.4138	BHCB 38448	3

* Brazilian States. GO = Goiás state; MG = Minas Gerais; MS = Mato Grosso do Sul; PR = Paraná; SP = São Paulo.

Table S2. Primer pairs tested for chloroplast and nuclear DNA amplification.

Primer	5' – 3' sequence	Reference
psbA	CGAAGCTCCATCTACAAATGG	Hamilton 1999
trnH	ACTGCCTTGATCCACTTGGC	
trnS	GCCGCTTTAGTCCACTCAGC	Hamilton 1999
trnG	GAACGAATCACACTTTTACCAC	
trnL (e)	GGTTCAAGTCCCTCTATCCC	Taberlet et al. 1991
trnF (f)	ATTTGAACTGGTGACACGAG	
trnV	GCTATACGGGCTCGAACC	Huang et al. 2002
trnM	TACCTACTATTGGATTTGAACC	
psbC1	GGTCGTGACCAAGAAACCAC	Demesure et al. 1995
trnS1	GGTTCGAATCCCTCTCTCTC	
Intron trnL (c)	CGAAATCGGTAGACGCTACG	Taberlet et al. 1991
Intron trnL (d)	GGGGATAGAGGGACTTGAAC	
trnD	ACCAATTGAACTACAATCCC	Demesure et al. 1995
trnT	CTACCACTGAGTTAAAAGGG	
trnY	AGGACATCTCTCTTTCAAGGAG	Shaw et al. 2005
trnE	CCGAGCTGGATTTGAACCA	
trnCF	CCAGTTCRAATCYGGGTG	Shaw et al. 2005
ycf6R	GCCCAAGCRAGACTTACTATATCCAT	
ycf6F	ATGGATATAGTAAGTCTYGCTTGGGC	Shaw et al. 2005
psbMR	ATGGAAGTAAATATTCTYGCATTTATTGCT	
psbMF	AGCAATAAATGCRAGAATATTTACTTCCAT	Shaw et al. 2005
trnDR	GGGATTGTAGYTCAATTGGT	
ITS 75	TATGCTTAAACTCAGCGGG	Desfeux & Lejeune 1996
ITS 92	AAGGTTTCCGTAGGTGAAC	
ITS 18Sa	GTCCACTGAACCTTATCATTTAGAGG	Beyra-M. & Lavin 1999
ITS 26Sa	GCCGTTACTAAGGGAATCCTTGTTAG	

References of Supplementary Table S2

Beyra-M A & Lavin M. 1999. Monograph of *Pictetia* (Leguminosae-Papilionoideae) and review of the *Aeschynomeneae*. Systematic Botany Monographs 56: 1–93.

Demesure B, Sodzi N & Petit RJ. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Molecular Ecology 4: 129–134.

Desfeux C & Lejeune B. 1996. Systematics of Euromediterranean *Silene* (Caryophyllaceae): evidence from a phylogenetic analysis using ITS sequences. Comptes rendus de l'Academie des sciences 319: 351–358.

Hamilton MB. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Molecular ecology 8: 521–523.

Huang SS, Hwang SY & Lin TP. 2002. Spatial pattern of chloroplast DNA variation of *Cyclobalanopsis glauca* in Taiwan and East Asia. Molecular Ecology 11: 2349–2358.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling E & Small RL. 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.

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

Table 3. Sampling strategies of the demographical parameter analyzes of *Leptolobium dasycarpum* and *L. elegans* implemented in LAMARC 2.1.10 software and calculated based on a Bayesian estimation using Markov chain Monte Carlo (MCMC) approach.

	Sampling strategy	
	cpDNA: *Leptolobium dasycarpum* & *L. elegans*	
Initial Chains	Number of initial chains	1
	Number of recorded parameters set	50,000
	Interval between recorded itens	100
	Initial burn-in	10,000
Final Chains	Number of final chains	1
	Number of recorded parameters set	100,000
	Interval between recorded itens	100
	Final burn-in	10,000
	ITS: *Leptolobium dasycarpum*	
Initial Chains	Number of initial chains	1
	Number of recorded parameters set	200,000
	Interval between recorded itens	100
	Initial burn-in	50,000
Final Chains	Number of final chains	1
	Number of recorded parameters set	250,000
	Interval between recorded itens	100
	Final burn-in	50,000
	ITS: *Leptolobium elegans*	
Initial Chains	Number of initial chains	1
	Number of recorded parameters set	120,000
	Interval between recorded itens	100
	Initial burn-in	30,000
Final Chains	Number of final chains	1
	Number of recorded parameters set	200,000
	Interval between recorded itens	100
	Final burn-in	30,000

cpDNA: non-coding region of the chloroplastidial DNA ; ITS: Internal Transcribed Spacer of the ribosomal nuclear DNA.

Table S4. Bioclimatic variables used in the environmental niche modelling of *Leptolobium dasycarpum* and *L. elegans*, derived from the monthly temperature and rainfall values available at WordClim database (worldclim.org/bioclim). A quarter is a period of three months (1/4 of the year).

Bioclim code	Bioclimatic variables
BIO 1	Annual Mean Temperature
BIO 2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO 3	Isothermality (BIO2/BIO7) (* 100)
BIO 4	Temperature Seasonality (standard deviation *100)
BIO 5	Maximum Temperature of Warmest Month
BIO 6	Minimum Temperature of Coldest Month
BIO 7	Temperature Annual Range (BIO5-BIO6)
BIO 8	Mean Temperature of Wettest Quarter
BIO 9	Mean Temperature of Driest Quarter
BIO 10	Mean Temperature of Warmest Quarter
BIO 11	Mean Temperature of Coldest Quarter
BIO 12	Annual Precipitation
BIO 13	Precipitation of Wettest Month
BIO 14	Precipitation of Driest Month
BIO 15	Precipitation Seasonality (Coefficient of Variation)
BIO 16	Precipitation of Wettest Quarter
BIO 17	Precipitation of Driest Quarter
BIO 18	Precipitation of Warmest Quarter
BIO 19	Precipitation of Coldest Quarter

Table S5. Haplotype diversity (h), nucleotide diversity (π) and haplotypes of *Leptolobium dasycarpum* populations for cpDNA (*psbA-trnH* + *trnL-trnF*) and nrDNA (ITS).

Sampling places	cpDNA	nrDNA	***h*** cpDNA / nrDNA	***π*** cpDNA / nrDNA
Anápolis / GO	H05	Hn01 / Hn03	0.0000 (0.6000)	0.000000 (0.001729)
Araçatuba / SP	H04	Hn01 / Hn03 / Hn11	0.0000 (0.6786)	0.000000 (0.001595)
Balsas / MA	H04	Hn03 / Hn11	1.0000 (1.0000)	0.000000 (0.001441)
Bandeirantes / MS	H03 / H09	Hn01 / Hn03	0.1818 (0.1732)	0.000637 (0.000499)
Barra da Estiva / BA	H09	Hn01 / Hn03	0.0000 (1.0000)	0.000000 (0.002882)
Berilo / MG	H07	Hn03 / Hn16	1.0000 (1.0000)	0.000000 (0.007205)
Brasília / DF	H04	Hn03	1.0000 (0.0000)	0.000000 (0.000000)
Caetité / BA	H01	Hn03	1.0000 (0.0000)	0.000000 (0.000000)
Campo Grande / MS	H01 / H03 / H04	Hn01 / Hn03 / Hn09 / Hn10 / Hn12 / Hn13 / Hn14	0.4854 (0.6244)	0.001133 (0.004742)
Capitólio / MG	H05	Hn03 / Hn11	0.0000 (0.5000)	0.000000 (0.000720)
Colinas do Sul / GO	H04 / H05	Hn01 / Hn03 / Hn07	0.6667 (0.8333)	0.000777 (0.008165)
Costa Rica / MS	H02	Hn01 / Hn09 / Hn10	0.0000 (0.2789)	0.000000 (0.002139)
Coxim / MS	H06	-	1.0000 (-)	0.000000 (-)
Crato / CE	H08	-	1.0000 (-)	0.000000 (-)
Delfinópolis / MG	H01 / H05	Hn01 / Hn11	1.0000 (1.0000)	0.003497 (0.001441)
Eusébio / CE	H08	Hn01	1.0000 (0.0000)	0.000000 (0.000000)
Fortaleza / CE	H08	Hn01 / Hn03	0.0000 (0.5000)	0.000000 (0.001441)
Jataí / GO	H01	Hn01	0.0000 (0.0000)	0.000000 (0.000000)
Matozinhos / MG	H05	Hn03	1.0000 (0.0000)	0.000000 (0.000000)
Moeda / MG	H05 / H06	Hn01 / Hn03	0.6667 (0.5000)	0.001554 (0.001441)
Mogi Guaçu / SP	H04	Hn01	1.0000 (0.0000)	0.000000 (0.000000)
Mogi Mirim / SP	H04 / H05	Hn01 / Hn03	0.6667 (0.2500)	0.000777 (0.000720)

Table S5. (continued)

Sampling places	cpDNA	nrDNA	*h* cpDNA / nrDNA	*π* cpDNA / nrDNA
Padre Bernardo / GO	H04	-	1.0000 (-)	0.000000 (-)
Paraopeba / MG	H05	Hn03	1.0000 (0.0000)	0.000000 (0.000000)
Perdizes / MG	H04 / H05	Hn01 / Hn03 / Hn15	0.6667 (0.8333)	0.000777 (0.002161)
Piracuruca / PI	H05	Hn03	1.0000 (0.0000)	0.000000 (0.000000)
Rio Negro / MS	H03 / H04	Hn01 / Hn05	0.5238 (0.3222)	0.001222 (0.002321)
Rio Verde de Mato Grosso / MS	H02	Hn01 / Hn05 / Hn06	0.0000 (0.3778)	0.000000 (0.002850)
Santana do Riacho / MG	H06	Hn03	0.0000 (0.0000)	0.000000 (0.000000)
São Gonçalo do Rio Preto / MG	H06	Hn03	1.0000 (0.0000)	0.000000 (0.000000)
São Valério da Natividade / TO	H03	Hn01 / Hn03	1.0000 (1.0000)	0.000000 (0.002882)
Selvíria / MS	H04	Hn03 / Hn07 / Hn08	0.0000 (0.5385)	0.000000 (0.005463)
Serra Dourada / GO	H01	Hn01 / Hn02 / Hn03	0.0000 (0.3922)	0.000000 (0.000904)
Serranópolis / GO	H01 / H04	Hn01 / Hn03 / Hn07 / Hn09 / Hn11	0.4000 (0.6667)	0.000933 (0.004547)
Taiobeiras / MG	H04	-	1.0000 (-)	0.000000 (-)
Taunay / MS	H03 / H04	Hn01 / Hn04	0.5143 (0.1287)	0.001200 (0.001669)
Três Lagoas / MS	H04	Hn01 / Hn03 / Hn09	0.0000 (0.8333)	0.000000 (0.007205)
Três Marias / MG	H04	-	1.0000 (-)	0.000000 (-)
Uberlândia / MG	H05	Hn01 / Hn03 / Hn11	0.0000 (0.6071)	0.000000 (0.001132)

Table S6. Haplotype diversity (h), nucleotide diversity (π) and haplotypes of *Leptolobium elegans* populations for cpDNA (*psbA-trnH* + *trnL-trnF*) and nrDNA (ITS).

Sampling places	cpDNA	nrDNA	***h*** cpDNA / nrDNA	***π*** cpDNA / nrDNA
Agachi / MS	H04	Hn04 / Hn07	1.0000 (1.0000)	0.000000 (0.001441)
Agudos / SP	H04	-	1.0000 (-)	0.000000 (-)
Anhanduí / MS	H04	Hn04 / Hn07	0.0000 (0.4396)	0.000000 (0.000633)
Aquidauana / MS	H04	Hn04 / Hn07	0.0000 (0.6667)	0.000000 (0.000961)
Bauru / SP	H04	Hn07 / Hn09 / Hn19	0.0000 (0.8333)	0.000000 (0.003122)
Bonito / MS	H04	Hn04 / Hn07	0.0000 (0.3556)	0.000000 (0.000512)
Botucatu / SP	H04	Hn04 / Hn07	1.0000 (1.0000)	0.000000 (0.001441)
Camapuã / MS	H04	Hn04 / Hn07	0.0000 (0.5072)	0.000000 (0.000731)
Campo Grande / MS	H04	Hn04 / Hn07	0.0000 (0.4396)	0.000000 (0.000633)
Carrancas / MG	H04	Hn07	1.0000 (0.0000)	0.000000 (0.000000)
Cassilândia / MS	H04	Hn04 / Hn07 / Hn17	0.0000 (0.5604)	0.000000 (0.001663)
Itapeva / SP	H04	Hn07 / Hn09 / Hn19	0.0000 (0.8000)	0.000000 (0.003074)
Itirapina / SP	H04	Hn07	1.0000 (0.0000)	0.000000 (0.000000)
Jaguariaíva / PR	H04	Hn09 / Hn19	0.0000 (1.0000)	0.000000 (0.004323)
Jataí / GO	H04	Hn04 / Hn07	1.0000 (1.0000)	0.000000 (0.001441)
Mogi Guaçu / SP	H04	Hn07	1.0000 (0.0000)	0.000000 (0.000000)
Mogi Mirim / SP	H04	Hn04 / Hn07	0.0000 (0.2333)	0.000000 (0.000336)

Table S6. (continued)

Sampling places	cpDNA	nrDNA	*h* cpDNA / nrDNA	*π* cpDNA / nrDNA
Nhecolândia / MS	H04	Hn04 / Hn07	0.0000 (0.5263)	0.000000 (0.000758)
Piraputanga / MS	H04	Hn04 / Hn07	0.0000 (0.5294)	0.000000 (0.000763)
Pirassununga / SP	H04 / H05	Hn09 / Hn19	1.0000 (1.0000)	0.001166 (0.004323)
Porto Murtinho / MS	H04	Hn04 / Hn07	0.0000 (0.5053)	0.000000 (0.000728)
Ribas do Rio Pardo / MS	H04	Hn07 / Hn18	1.0000 (1.0000)	0.000000 (0.002882)
Rio Verde de Mato Grosso / MS	H03	Hn04 / Hn07	1.0000 (1.0000)	0.000000 (0.001441)
São Carlos / SP	H04	Hn07	0.0000 (0.0000)	0.000000 (0.000000)
Selvíria / MS	H04	Hn04 / Hn07 / Hn09 / Hn17	0.0000 (0.5294)	0.000000 (0.001149)
Taunay / MS	H03 / H04	Hn04 / Hn07 / Hn09 / Hn17	0.5556 (0.6474)	0.001297 (0.001600)
Tibagi / PR	H04	Hn09 / Hn19	0.0000 (0.6000)	0.000000 (0.002594)

Table S7. Summary of demographic tests performed for *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae). The values were not statistically significant.

Parameter	*Leptolobium dasycarpum*	*Leptolobium elegans*
	cpDNA	
Tajima's D	-0.14367	0.19423
Fu & Li's F'	0.78932	0.46008
Fu & Li's D'	1.08285	0.48087
Fu's F_S	-0.09380	0.64192
	ITS	
Tajima's D	0.81951	0.46790
Fu & Li's F'	0.69129	0.87359
Fu & Li's D'	0.58652	0.86895
Fu's F_S	-2.46229	-0.24052

Table S8. Demographical parameters for *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) populations based in coalescence analysis of the cpDNA region. θ = mutation parameter; *g* = exponential growth rate; 95% CI = confidence interval.

	Population	θ	95% CI	*g*	95% CI
Leptolobium dasycarpum	Campo Grande / MS	0.0004680	0.0000674 to 0.002643	822.03	-457.48 to 982.98
	Costa Rica / MS	0.0000188	0.0000104 to 0.000472	875.06	-467.62 to 976.58
	Jataí / GO	0.0000181	0.0000106 to 0.000692	657.79	-467.75 to 977.03
	Mogi Mirim / SP	0.0000140	0.0000104 to 0.001867	871.12	-467.60 to 976.15
	Rio Negro / MS	0.0001100	0.0000151 to 0.000827	867.14	-469.46 to 974.85
	Rio Verde de Mato Grosso / MS	0.0000243	0.0000106 to 0.000962	868.45	-469.63 to 978.18
	Selvíria / MS	0.0000142	0.0000107 to 0.000947	727.44	-466.69 to 976.59
	Serra Dourada / GO	0.0000622	0.0000108 to 0.000623	880.85	-468.91 to 978.94
	Serranópolis / GO	0.0002670	0.0000169 to 0.003488	874.75	-463.30 to 982.47
	Taunay / MS	0.0002200	0.0000178 to 0.001094	693.00	-467.70 to 977.48
	Overall	0.0014140	0.0004630 to 0.003063	370.37	-461.05 to 967.81
Leptolobium elegans	Anhanduí / MS	0.0000211	0.0000106 to 0.001379	532.73	-470.07 to 979.62
	Bonito / MS	0.0000245	0.0000103 to 0.001687	849.10	-469.89 to 974.64
	Camapuã / MS	0.0000795	0.0000110 to 0.000909	772.54	-469.59 to 972.52
	Campo Grande / MS	0.0000131	0.0000105 to 0.000917	704.16	-465.28 to 977.07
	Cassilândia / MS	0.0000537	0.0000108 to 0.001303	763.45	-467.27 to 978.32
	Mogi Mirim / SP	0.0000186	0.0000103 to 0.001054	743.25	-464.77 to 975.50
	Nhecolândia / MS	0.0000258	0.0000106 to 0.000430	618.66	-467.22 to 978.24
	Piraputanga / MS	0.0000138	0.0000104 to 0.000923	858.59	-471.43 to 977.04
	Porto Murtinho / MS	0.0000357	0.0000107 to 0.001241	544.37	-469.39 to 976.79
	São Carlos / SP	0.0000258	0.0000102 to 0.000792	470.57	-471.94 to 973.28
	Selvíria / MS	0.0000361	0.0000104 to 0.000682	873.08	-470.58 to 976.82
	Taunay / MS	0.0002300	0.0000176 to 0.002067	861.98	-470.39 to 977.60
	Overall	0.0002110	0.0000152 to 0.000858	819.54	-466.38 to 967.38

Table S9. Demographical parameters for *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae) populations based in coalescence analysis of the ITS region. θ = mutation parameter; *g* = exponential growth rate; 95% CI = confidence interval.

	Population	θ	95% CI	*g*	95% CI
Leptolobium dasycarpum	Anápolis / GO	0.0005540	0.0000186 to 0.049550	914.18	-452.16 to 982.25
	Araçatuba / SP	0.0005100	0.0000656 to 0.024016	899.47	-449.28 to 981.86
	Bandeirantes / MS	0.0000143	0.0000108 to 0.001576	898.06	-456.80 to 977.30
	Campo Grande / MS	0.0032220	0.0009660 to 0.007316	804.04	-453.56 to 981.42
	Capitólio / MG	0.0003750	0.0000182 to 0.079469	850.25	-451.28 to 977.97
	Colinas do Sul / GO	0.0009690	0.0000737 to 3.481169	687.40	-460.97 to 976.24
	Costa Rica / MS	0.0005810	0.0001130 to 0.002984	894.60	-444.52 to 977.92
	Fortaleza / CE	0.0002370	0.0000205 to 3.030233	904.35	-453.23 to 979.70
	Jataí / GO	0.0000120	0.0000103 to 0.000457	770.51	-462.98 to 976.70
	Moeda / MG	0.0003640	0.0000328 to 1.014680	816.27	-454.66 to 975.77
	Mogi Mirim / SP	0.0001980	0.0000213 to 0.001572	917.32	-451.19 to 980.95
	Perdizes / MG	0.0009810	0.0001080 to 5.539105	862.26	-454.02 to 971.69
	Rio Negro / MS	0.0003520	0.0000123 to 0.000959	460.37	-463.68 to 974.33
	Rio Verde de Mato Grosso / MS	0.0009560	0.0001570 to 0.019802	703.56	-453.93 to 978.78
	Santana do Riacho / MG	0.0000143	0.0000102 to 0.001010	682.27	-459.41 to 978.54
	Selvíria / MS	0.0008210	0.0000853 to 0.003298	816.74	-452.83 to 980.65
	Serra Dourada / GO	0.0003260	0.0000348 to 0.001912	910.42	-447.01 to 979.78
	Serranópolis / GO	0.0026560	0.0005170 to 4.513353	897.22	-452.14 to 981.01
	Taunay / MS	0.0003960	0.0000384 to 0.000882	889.44	-454.85 to 973.83
	Três Lagoas / MS	0.0009890	0.0001400 to 4.723254	869.94	-457.58 to 979.45
	Uberlândia / MG	0.0008480	0.0000635 to 0.004726	820.96	-448.44 to 979.95
	Overall	0.0048760	0.0026240 to 0.008689	198.88	-351.71 to 937.18

Table S9. (continued)

	Population	θ	95% CI	*g*	95% CI
Leptolobium elegans	*Anhanduí / MS*	0.0003710	0.0000216 to 0.002011	895.54	-455.91 to 978.61
	Aquidauana / MS	0.0005090	0.0000332 to 1.853273	897.95	-459.38 to 977.95
	Bauru / SP	0.0018270	0.0001280 to 6.530590	766.93	-462.48 to 981.08
	Bonito / MS	0.0002630	0.0000225 to 0.001887	904.18	-455.31 to 981.40
	Camapuã / MS	0.0002960	0.0000123 to 0.001012	907.87	-460.49 to 980.11
	Cassilândia / MS	0.0009900	0.0000564 to 0.004981	892.22	-457.52 to 979.47
	Campo Grande / MS	0.0001800	0.0000186 to 0.000979	467.65	-457.12 to 973.81
	Itapeva / SP	0.0010570	0.0001240 to 0.013237	892.25	-451.36 to 980.22
	Nhecolândia / MS	0.0001320	0.0000160 to 0.001299	896.86	-460.06 to 974.32
	Mogi Mirim / SP	0.0003180	0.0000196 to 0.002146	869.66	-455.94 to 978.65
	Piraputanga / MS	0.0002460	0.0000130 to 0.001863	891.81	-459.62 to 981.36
	Porto Murtinho / MS	0.0002230	0.0000205 to 0.001192	893.02	-460.35 to 978.35
	São Carlos / SP	0.0000128	0.0000100 to 0.000876	713.82	-459.43 to 973.84
	Tibagi / PR	0.0005360	0.0000185 to 0.002207	784.61	-455.77 to 977.72
	Selvíria / MS	0.0016480	0.0002550 to 0.007657	906.23	-446.44 to 987.40
	Taunay / MS	0.0016360	0.0003220 to 0.007443	903.94	-448.67 to 983.36
	Overall	0.0011660	0.0004400 to 0.003377	160.96	-466.17 to 956.08

Table S10. Migration rates (4Nm) among *Leptolobium dasycarpum* (Leguminosae, Papilionoideae) populations based in coalescence analyses of the cpDNA region. The migration takes place toward the source populations (line) to the host populations (column).

		Host									
	Population	Campo Grande / MS	Costa Rica / MS	Jataí / GO	Mogi Mirim / SP	Rio Negro / MS	Rio Verde de Mato Grosso / MS	Selvíria / MS	Serra Dourada / GO	Serranópolis / GO	Taunay / MS
Source	Campo Grande / MS	-	0.00102	0.00125	0.00128	0.01014	0.00210	0.00130	0.00240	0.02252	0.02017
	Costa Rica / MS	0.01897	-	0.00120	0.00070	0.00377	0.00225	0.00108	0.00566	0.00784	0.01669
	Jataí / GO	0.04310	0.00170	-	0.00120	0.00108	0.00077	0.00129	0.00578	0.02474	0.00343
	Mogi Mirim / SP	0.04067	0.00101	0.00134	-	0.01014	0.00196	0.00117	0.00057	0.02461	0.02028
	Rio Negro / MS	0.04328	0.00061	0.00147	0.00128	-	0.00100	0.00119	0.00484	0.02407	0.02019
	Rio Verde de Mato Grosso / MS	0.01017	0.00174	0.00081	0.00097	0.00770	-	0.00088	0.00081	0.02002	0.01244
	Selvíria / MS	0.04329	0.00099	0.00161	0.00128	0.01010	0.00115	-	0.00121	0.02446	0.02030
	Serra Dourada / GO	0.04327	0.00166	0.00166	0.00118	0.00724	0.00196	0.00045	-	0.02468	0.00735
	Serranópolis / GO	0.04338	0.00081	0.00167	0.00123	0.00934	0.00189	0.00129	0.00574	-	0.01489
	Taunay / MS	0.04329	0.00112	0.00117	0.00118	0.01012	0.00129	0.00124	0.00082	0.02242	-

Table S11. Migration rates (4Nm) among *Leptolobium dasycarpum* (Leguminosae, Papilionoideae) populations based in coalescence analyses of the ITS region. The migration takes place toward the source populations (line) to the host populations (column).

	Population	Host																				
		Anápolis / GO	Araçatuba / SP	Bandeirantes / MS	Campo Grande / MS	Capitólio / MG	Colinas do Sul / GO	Costa Rica / MS	Fortaleza / CE	Jataí / GO	Moeda / MG	Mogi Mirim / SP	Perdizes / MG	Rio Negro / MS	Rio Verde de Mato Grosso / MS	Santana do Riacho / MG	Selvíria / MS	Serra Dourada / GO	Serranópolis / GO	Taunay / MS	Três Lagoas / MS	Uberlândia / MG
Source	Anápolis / GO	-	0.045	0.001	0.279	0.030	0.075	0.054	0.016	0.001	0.028	0.016	0.089	0.032	0.009	0.001	0.061	0.030	0.225	0.003	0.091	0.079
	Araçatuba / SP	0.042	-	0.001	0.152	0.032	0.070	0.017	0.022	0.001	0.033	0.017	0.092	0.011	0.026	0.000	0.052	0.030	0.249	0.030	0.084	0.080
	Bandeirantes / MS	0.043	0.041	-	0.199	0.006	0.046	0.051	0.020	0.001	0.033	0.018	0.082	0.033	0.088	0.001	0.059	0.031	0.247	0.037	0.072	0.079
	Campo Grande / MS	0.045	0.028	0.001	-	0.021	0.090	0.055	0.007	0.000	0.004	0.017	0.092	0.020	0.089	0.000	0.061	0.004	0.248	0.028	0.090	0.006
	Capitólio / MG	0.052	0.046	0.001	0.108	-	0.082	0.009	0.022	0.000	0.033	0.016	0.091	0.013	0.051	0.001	0.075	0.030	0.241	0.035	0.063	0.079
	Colinas do Sul / GO	0.004	0.005	0.001	0.301	0.021	-	0.052	0.022	0.000	0.027	0.012	0.089	0.022	0.014	0.000	0.069	0.021	0.240	0.035	0.092	0.008
	Costa Rica / MS	0.031	0.036	0.000	0.304	0.013	0.070	-	0.021	0.001	0.029	0.015	0.075	0.033	0.089	0.001	0.015	0.028	0.249	0.037	0.092	0.076
	Fortaleza / CE	0.046	0.047	0.001	0.267	0.017	0.090	0.042	-	0.001	0.032	0.015	0.091	0.032	0.069	0.000	0.077	0.030	0.241	0.033	0.087	0.043
	Jataí / GO	0.050	0.041	0.001	0.283	0.008	0.086	0.052	0.020	-	0.031	0.016	0.084	0.027	0.089	0.001	0.006	0.029	0.246	0.033	0.088	0.079
	Moeda / MG	0.046	0.040	0.001	0.301	0.031	0.080	0.004	0.017	0.000	-	0.017	0.092	0.033	0.089	0.001	0.076	0.021	0.209	0.028	0.088	0.074
	Mogi Mirim / SP	0.045	0.044	0.001	0.156	0.035	0.082	0.006	0.022	0.000	0.032	-	0.091	0.008	0.061	0.001	0.071	0.024	0.248	0.016	0.079	0.078
	Perdizes / MG	0.039	0.042	0.001	0.175	0.034	0.089	0.012	0.019	0.000	0.029	0.015	-	0.018	0.072	0.001	0.077	0.019	0.140	0.004	0.069	0.014
	Rio Negro / MS	0.051	0.046	0.001	0.301	0.032	0.008	0.044	0.008	0.001	0.023	0.016	0.032	-	0.086	0.001	0.042	0.028	0.206	0.030	0.079	0.043
	Rio Verde de Mato Grosso / MS	0.050	0.048	0.001	0.280	0.027	0.090	0.054	0.018	0.001	0.024	0.010	0.091	0.025	-	0.000	0.017	0.007	0.249	0.030	0.078	0.076
	Santana do Riacho / MG	0.051	0.048	0.001	0.276	0.035	0.091	0.006	0.017	0.001	0.034	0.018	0.091	0.014	0.025	-	0.056	0.025	0.246	0.006	0.092	0.079
	Selvíria / MS	0.051	0.012	0.001	0.263	0.032	0.091	0.047	0.022	0.000	0.033	0.011	0.058	0.008	0.019	0.001	-	0.027	0.249	0.037	0.093	0.071
	Serra Dourada / GO	0.046	0.047	0.001	0.255	0.035	0.084	0.054	0.022	0.001	0.030	0.018	0.091	0.033	0.069	0.000	0.018	-	0.249	0.037	0.084	0.076
	Serranópolis / GO	0.051	0.027	0.001	0.299	0.032	0.039	0.053	0.004	0.001	0.034	0.013	0.075	0.019	0.016	0.000	0.058	0.014	-	0.013	0.088	0.075
	Taunay / MS	0.049	0.036	0.001	0.275	0.030	0.083	0.022	0.021	0.001	0.034	0.017	0.079	0.015	0.086	0.001	0.063	0.030	0.249	-	0.090	0.079
	Três Lagoas / MS	0.035	0.006	0.001	0.300	0.035	0.084	0.054	0.021	0.000	0.024	0.002	0.073	0.024	0.088	0.001	0.067	0.019	0.241	0.028	-	0.046
	Uberlândia / MG	0.052	0.048	0.001	0.051	0.035	0.088	0.024	0.016	0.000	0.030	0.018	0.090	0.002	0.052	0.001	0.077	0.027	0.250	0.029	0.090	-

Table S12. Migration rates (4Nm) among *Leptolobium elegans* (Leguminosae, Papilionoideae) populations based in coalescence analyses of the cpDNA region. The migration takes place toward the source populations (line) to the host populations (column).

		Host											
	Population	Anhanduí / MS	Bonito / MS	Camapuã / MS	Campo Grande / MS	Cassilândia / MS	Mogi Mirim / SP	Nhecolândia / MS	Piraputanga / MS	Porto Murtinho / MS	São Carlos / SP	Selvíria / MS	Taunay / MS
Source	Anhanduí / MS	-	0.00225	0.00704	0.00121	0.00363	0.00137	0.00236	0.00103	0.00317	0.00203	0.00327	0.02114
	Bonito / MS	0.00188	-	0.00682	0.00069	0.00350	0.00164	0.00110	0.00118	0.00270	0.00234	0.00310	0.01838
	Camapuã / MS	0.00192	0.00204	-	0.00085	0.00468	0.00153	0.00193	0.00106	0.00288	0.00193	0.00323	0.01948
	Campo Grande / MS	0.00134	0.00222	0.00724	-	0.00418	0.00170	0.00196	0.00126	0.00326	0.00231	0.00329	0.02031
	Cassilândia / MS	0.00184	0.00225	0.00732	0.00104	-	0.00170	0.00052	0.00123	0.00291	0.00181	0.00333	0.01854
	Mogi Mirim / SP	0.00149	0.00223	0.00733	0.00081	0.00436	-	0.00210	0.00119	0.00215	0.00236	0.00330	0.01904
	Nhecolândia / MS	0.00052	0.00178	0.00528	0.00083	0.00454	0.00171	-	0.00062	0.00303	0.00098	0.00286	0.02137
	Piraputanga / MS	0.00178	0.00212	0.00576	0.00114	0.00467	0.00150	0.00226	-	0.00264	0.00221	0.00316	0.01906
	Porto Murtinho / MS	0.00090	0.00158	0.00702	0.00120	0.00396	0.00170	0.00188	0.00126	-	0.00231	0.00332	0.01640
	São Carlos / SP	0.00190	0.00160	0.00668	0.00121	0.00471	0.00170	0.00187	0.00125	0.00273	-	0.00249	0.02100
	Selvíria / MS	0.00193	0.00225	0.00564	0.00094	0.00479	0.00159	0.00087	0.00121	0.00328	0.00184	-	0.02101
	Taunay / MS	0.00162	0.00160	0.00479	0.00075	0.00472	0.00142	0.00237	0.00091	0.00143	0.00217	0.00329	-

Table S13. Migration rates (4Nm) among *Leptolobium elegans* (Leguminosae, Papilionoideae) populations based in coalescence analyses of the ITS region. The migration takes place toward the source populations (line) to the host populations (column).

		Host															
	Population	Anhanduí / MS	Aquidauana / MS	Bauru / SP	Bonito / MS	Camapuã / MS	Cassilândia / MS	Campo Grande / MS	Itapeva / SP	Nhecolândia / MS	Mogi Mirim / SP	Piraputanga / MS	Porto Murtinho / MS	São Carlos / SP	Tibagi / PR	Selvíria / MS	Taunay / MS
Source	Anhanduí / MS	-	0.0424	0.1457	0.0235	0.0275	0.0845	0.0166	0.0837	0.0117	0.0294	0.0226	0.0206	0.0012	0.0040	0.1515	0.1478
	Aquidauana / MS	0.0344	-	0.1659	0.0243	0.0275	0.0819	0.0165	0.0867	0.0115	0.0185	0.0227	0.0171	0.0002	0.0039	0.1514	0.1398
	Bauru / SP	0.0285	0.0152	-	0.0125	0.0065	0.0591	0.0167	0.0913	0.0122	0.0129	0.0052	0.0020	0.0004	0.0394	0.1518	0.1372
	Bonito / MS	0.0322	0.0473	0.1693	-	0.0265	0.0779	0.0167	0.0619	0.0097	0.0296	0.0201	0.0207	0.0010	0.0256	0.1529	0.1524
	Camapuã / MS	0.0345	0.0472	0.1457	0.0204	-	0.0884	0.0167	0.0940	0.0123	0.0296	0.0227	0.0200	0.0010	0.0494	0.1420	0.1517
	Cassilândia / MS	0.0343	0.0473	0.1591	0.0092	0.0266	-	0.0148	0.0973	0.0095	0.0216	0.0214	0.0190	0.0012	0.0470	0.1537	0.1445
	Campo Grande / MS	0.0345	0.0438	0.0919	0.0243	0.0276	0.0914	-	0.0844	0.0105	0.0296	0.0212	0.0206	0.0002	0.0109	0.1529	0.1518
	Itapeva / SP	0.0275	0.0457	0.1706	0.0117	0.0185	0.0910	0.0033	-	0.0117	0.0115	0.0019	0.0071	0.0006	0.0495	0.1525	0.1532
	Nhecolândia / MS	0.0344	0.0463	0.1658	0.0242	0.0254	0.0882	0.0167	0.0104	-	0.0246	0.0228	0.0207	0.0004	0.0369	0.1518	0.1492
	Mogi Mirim / SP	0.0344	0.0475	0.1696	0.0229	0.0210	0.0915	0.0168	0.0973	0.0109	-	0.0221	0.0182	0.0011	0.0068	0.1288	0.1208
	Piraputanga / MS	0.0345	0.0474	0.1600	0.0245	0.0275	0.0837	0.0163	0.0251	0.0118	0.0295	-	0.0182	0.0009	0.0208	0.1528	0.1526
	Porto Murtinho / MS	0.0318	0.0471	0.1614	0.0244	0.0232	0.0918	0.0153	0.0917	0.0116	0.0282	0.0219	-	0.0012	0.0238	0.1518	0.1392
	São Carlos / SP	0.0344	0.0473	0.1695	0.0220	0.0273	0.0915	0.0151	0.0957	0.0116	0.0289	0.0142	0.0182	-	0.0109	0.1388	0.1421
	Tibagi / PR	0.0093	0.0095	0.1716	0.0152	0.0131	0.0877	0.0056	0.0997	0.0025	0.0065	0.0019	0.0016	0.0007	-	0.1538	0.1522
	Selvíria / MS	0.0307	0.0475	0.1647	0.0236	0.0195	0.0894	0.0167	0.0833	0.0106	0.0292	0.0227	0.0206	0.0012	0.0242	-	0.1512
	Taunay / MS	0.0342	0.0403	0.1493	0.0202	0.0275	0.0926	0.0167	0.0904	0.0112	0.0227	0.0104	0.0180	0.0012	0.0497	0.1544	-

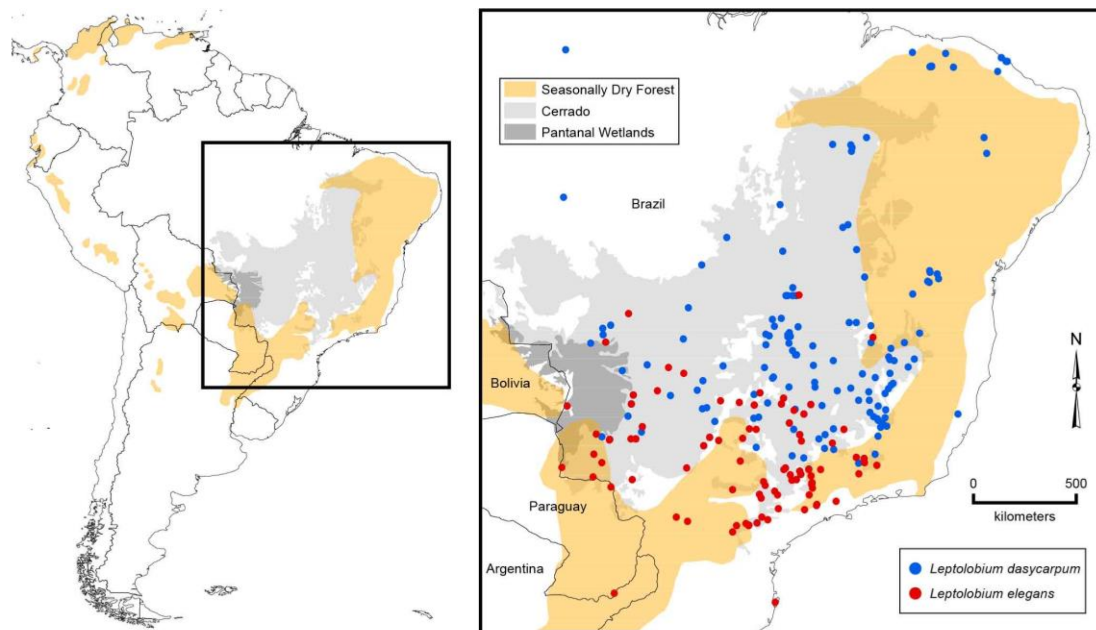

Figure S1. Presence points of *Leptolobium dasycarpum* (N = 167) and *L. elegans* (N = 87), obtained from field collections and herbarium specimens, used in the environmental niche modeling. The distribution of seasonally dry forest (orange areas) in South America was obtained from Collevatti et al. 2012[1].

[1] Collevatti RG, Terribile LC, Lima-Ribeiro MS, Nabout JC, Oliveira G, Rangel TF, Rabelo SG & Diniz-Filho JA. 2012. A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species. Molecular ecology 21: 5845–5863.

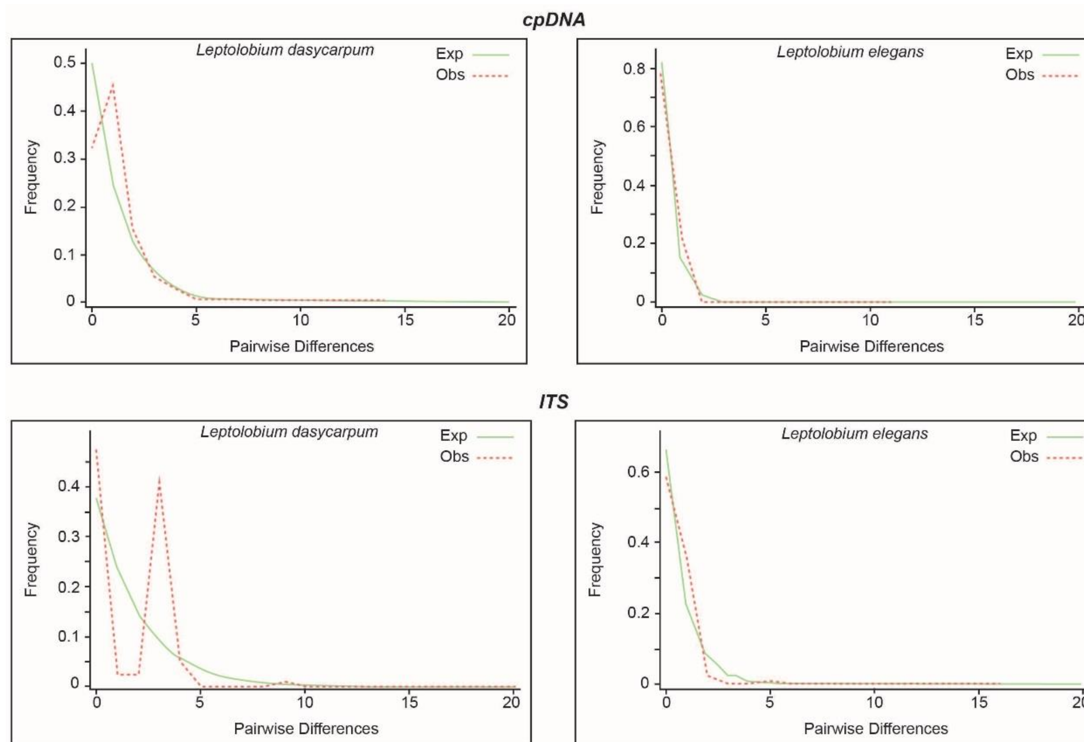

Figure S2. Mismatch distributions of cpDNA and nrDNA regions for *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae). The graphics shows a histogram of genetic differences between pairs of individuals within a sample. Lines represent the expected (Exp; continuous line) and observed (Obs; dashed line) pairwise mismatch distributions.

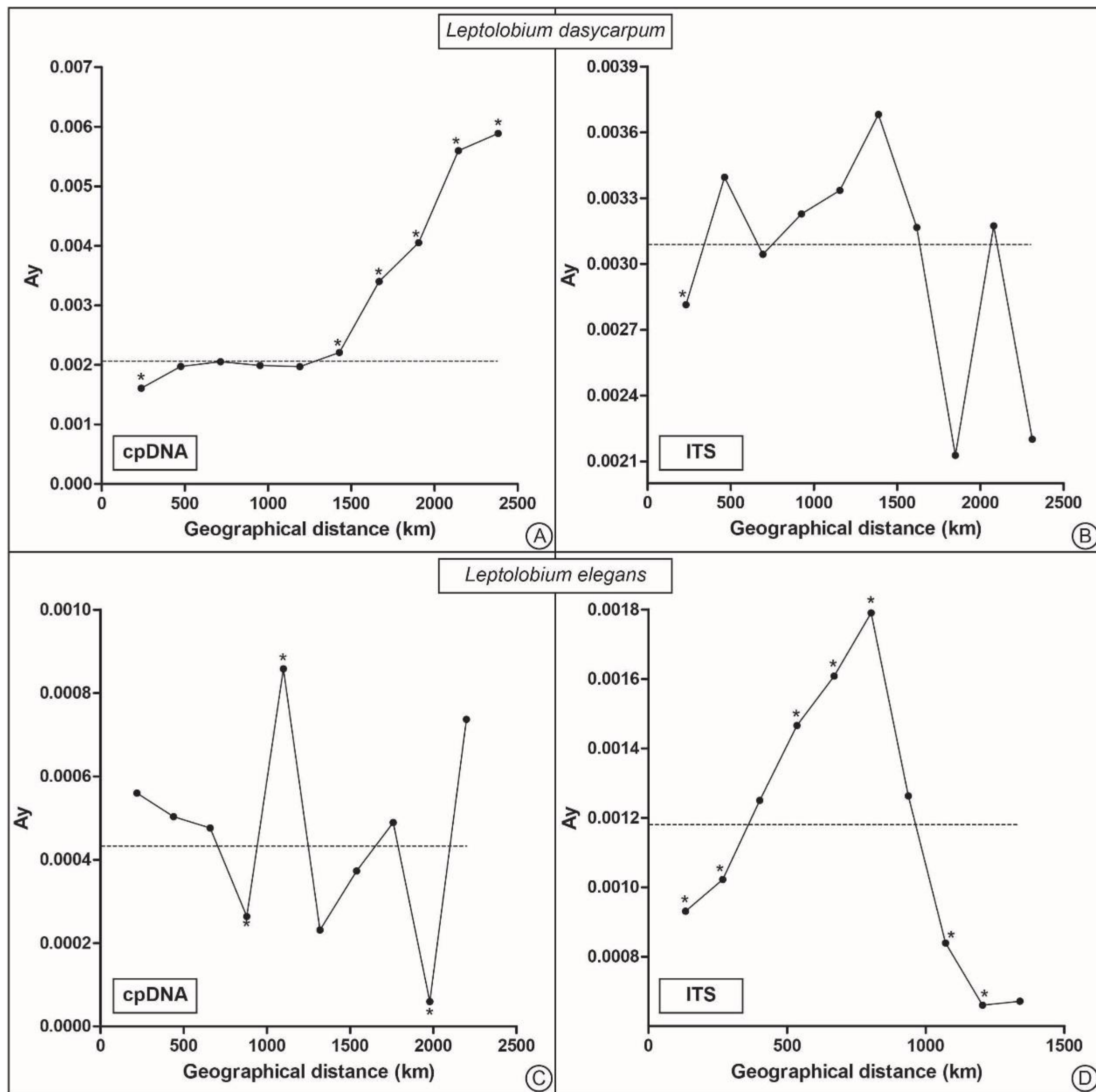

Figure S3. Results of Spatial Autocorrelation analyses for *Leptolobium* dasycarpum (top) and *L. elegans* (down) basead on cpDNA (A and C) and ITS sequences (B and D). Distance classes that showed significantly larger or smaller values at α = 0.05 level than average (indicated by horizontal dashed lines) are marked with asterisks.

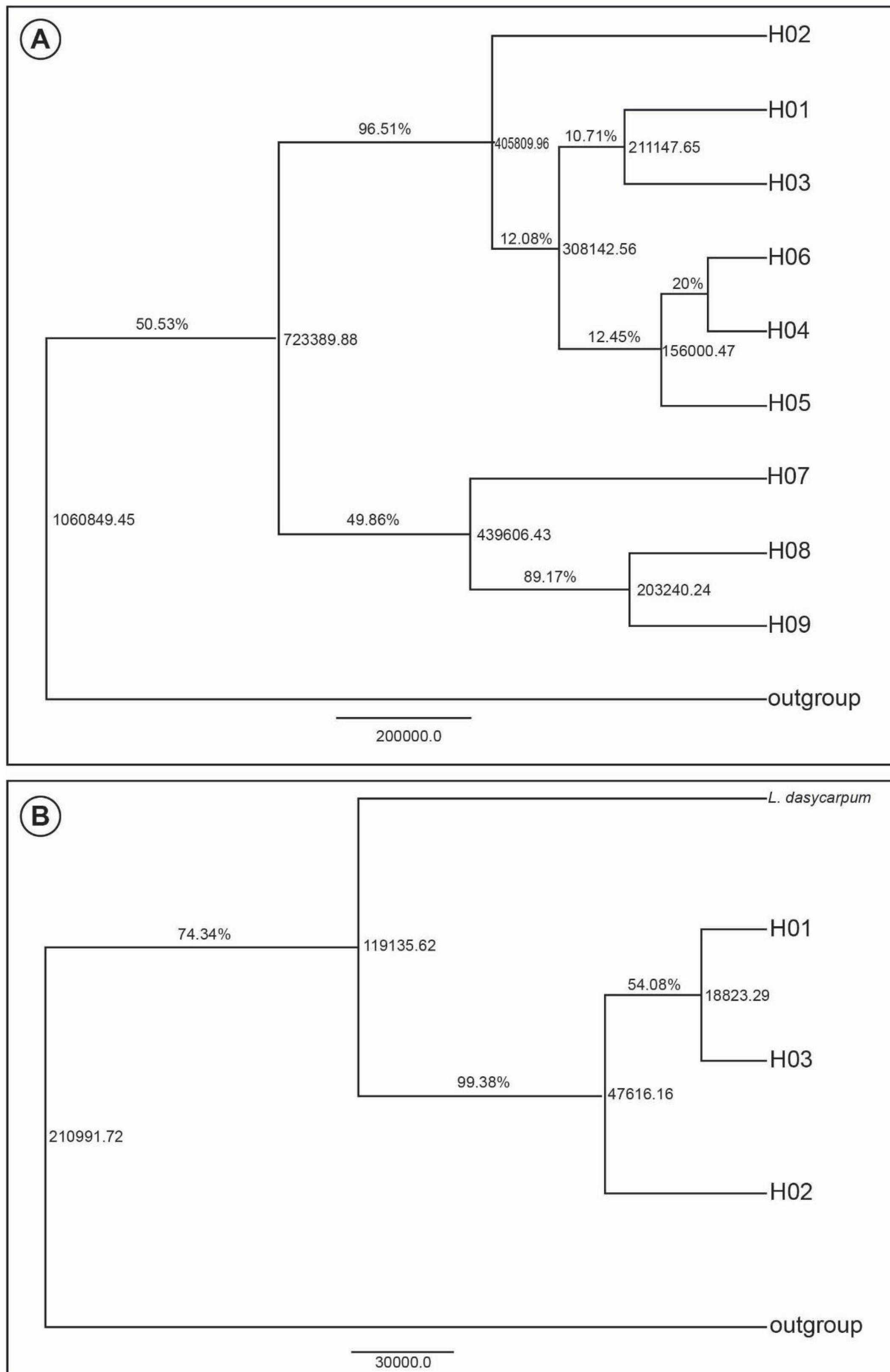

Figure S4. Phylogenetic relationships between the haplotypes of *Leptolobium dasycarpum* (A) and *L. elegans* (B) based on cpDNA. Numbers in each node and branch indicate mean node age and posterior probability, respectively.

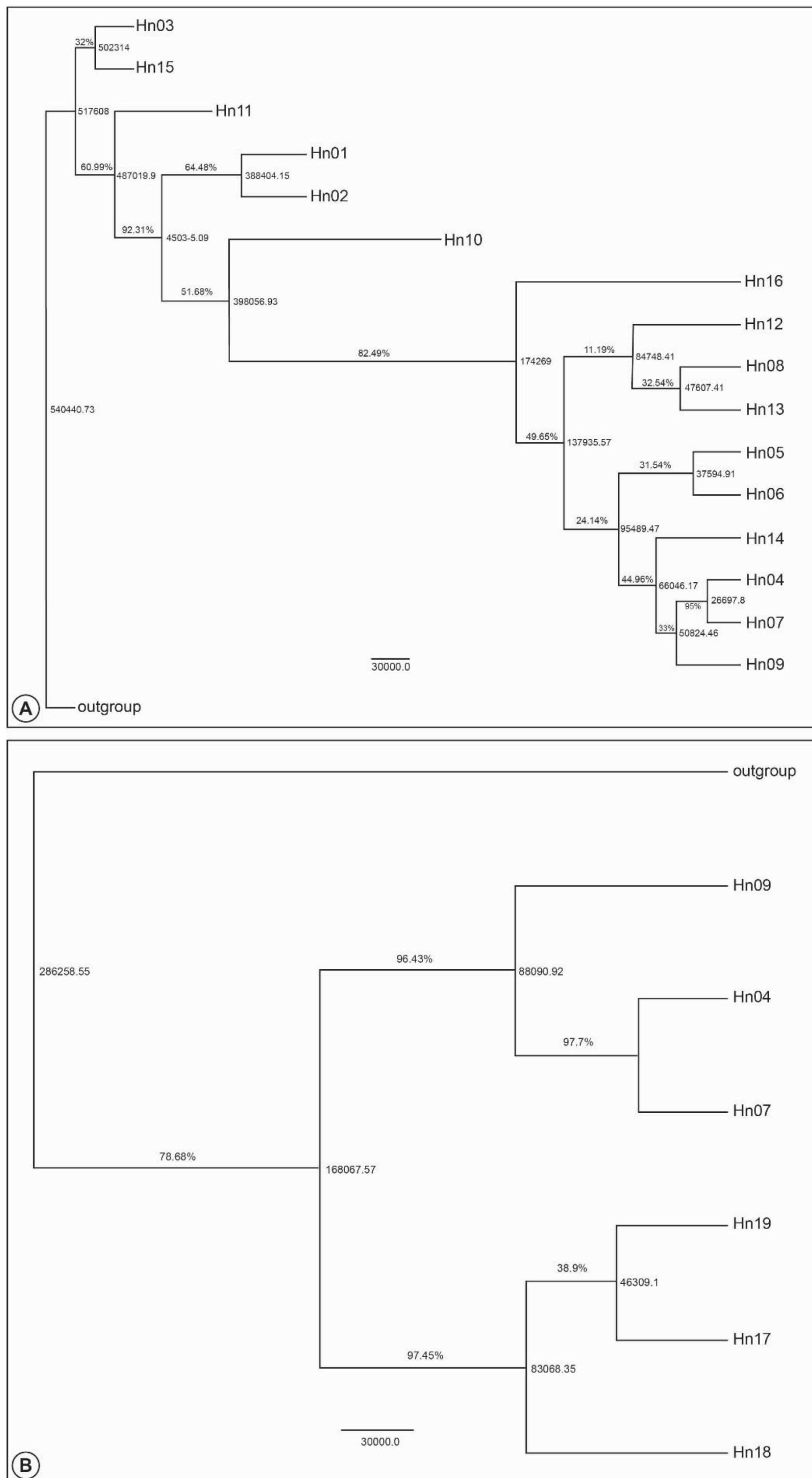

Figure S5. Phylogenetic relationships between the haplotypes of *Leptolobium dasycarpum* (A) and *L. elegans* (B) based on ITS region. Numbers in each node and branch indicate mean node age and posterior probability, respectively.

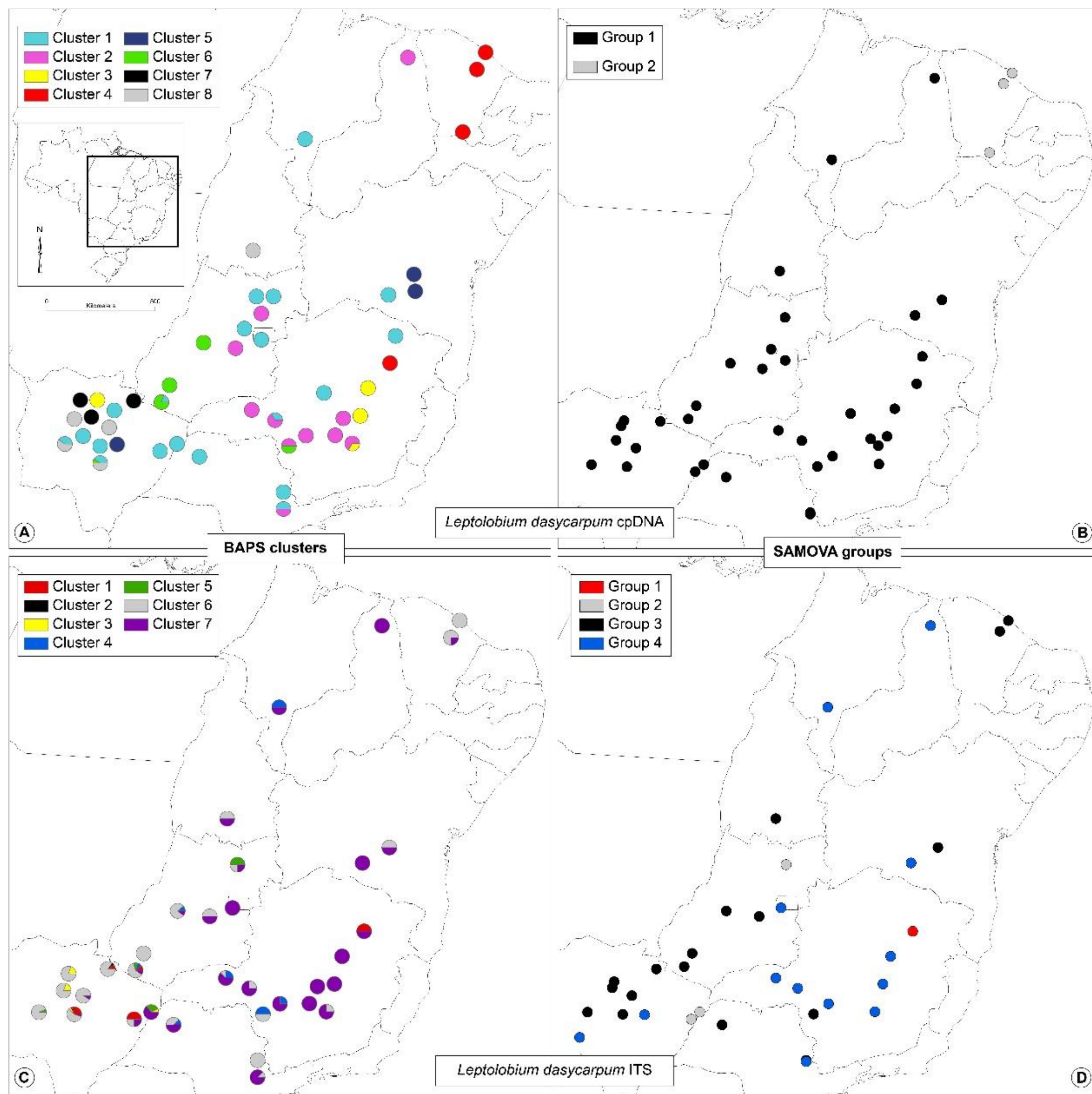

Figure S6. Geographical representation of the BAPS individual clusters (A and C) and SAMOVA population groups (B and D) of *Leptolobium dasycarpum* based on cpDNA (A and B) and ITS (C and D) regions.

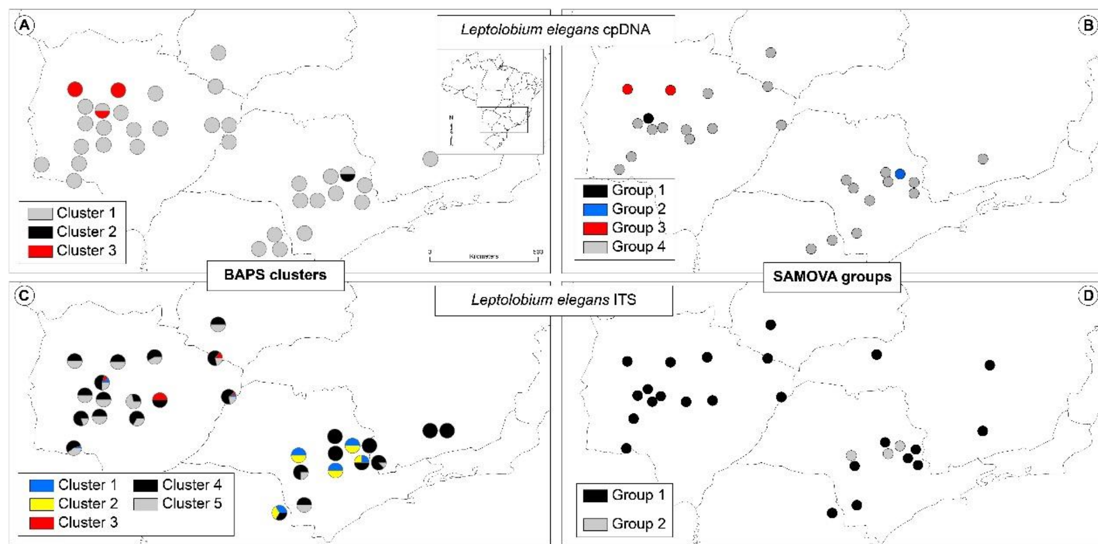

Figure S7. Geographical representation of the BAPS individual clusters (A and C) and SAMOVA population groups (B and D) of *Leptolobium elegans* based on cpDNA (A and B) and ITS (C and D) regions.

Protocol 1. DNA isolation protocol modified from Doyle & Doyle (1987)[2].

Reagents and chemicals required

EXTRACTION BUFFER:

0.5% β- mercaptoethanol

2% Cetyl Trimethylammonium Bromide (CTAB)

0.02 M Ethylenediaminetetraacetic Acid (EDTA)

2% Polyvinylpyrrolidone (PVP)

1.4 M Sodium Chloride (NaCl)

100 Mm TRIS-HCl pH 8.0

β- mercaptoethanol

Chloroform

Ethanol 70%

Isopropanol

Phenol or SEVAG (chloroform: isoamyl alcohol (24:1) (v/v))

Proteinase K (10 mg/ml)

Sodium acetate 3M

RESUSPENSION BUFFER:

Ultrapure water or TE Buffer (10 mM Tris-Cl (pH 8.0) and 1 mM EDTA)

[2] Doyle JJ & Doyle JL. 1987. Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

Procedures

1. Grind ca. 100 mg of fresh or dry leaves in liquid nitrogen or using a mixer mill.
2. Transfer the material to a 1.5 ml microfuge tube and add:

   600 µl of EXTRACTION BUFFER

   60 µl of β- mercaptoethanol

   6 µl of Proteinase K
3. Vortex the sample for 5 sec to homogenize.
4. Incubate at 65ºC for 30 min (for fresh material) or 1 hr (for herbarium specimens).
5. Add 600 µl of SEVAG (or 300 µl of Chloroform and 300 µl of Phenol).
6. Centrifuge at 14,000 rpm for 15 min.
7. Transfer the supernatant to a new 1.5ml tube containing:

   Equal supernatant volume of Isopropanol

   1/10 supernatant volume of Sodium Acetate 3M
8. Gently mix the solution by inversion and place the tubes on freezer (-20°C) overnight.
9. Centrifuge at 14,000 rpm for 20 min. The DNA will be visible as a small white pellet.
10. Discard the supernatant and wash the pellet twice with 70% Ethanol.
11. Dry the pellet and dissolve it in 100 µl of RESUSPENSION BUFFER.
12. Put the samples at 2-8°C overnight to allow the correct resuspension of the DNA.
13. Store DNA at -20°C until use.

Supplement S1. List of the herbarium samples used for the morphometric analysis of *Leptolobium dasycarpum* and *L. elegans* (Leguminosae, Papilionoideae). The specimens are separated by state, and the numbers in parentheses represent the number of leaflets measured in each voucher. For the herbaria names, see Chapter 1, Supplement S1.

Leptolobium dasycarpum (2184)

Bahia (73): FUEL 35305 (14), HST 13763 (10), UB 2150 (14), UEC 123661 (14), UEC 155640 (6), UEC 7411 (15)

Ceará (170): EAC 13993 (9), EAC 2514 (4), EAC 27298 (4), EAC 30560 (5), EAC 30869 (3), EAC 3122 (4), EAC 31507 (6), EAC 34716 (23), EAC 4397 (5), EAC 44295 (10), EAC 6306 (15), EAC 6566 (9), EAC 7865 (6), EAC 7893 (6), EAC 8247 (5), FUEL 34819 (9), HST 8855 (20), IPA 21482 (5), IPA 21493 (11), IPA 23278 (7), UFRN 6269 (2), UFRN 6295 (2)

Espírito Santo (5): MBML 28872 (5)

Federal District (94): EAC 33252 (2), UB 2317 (10), UB 2394 (15), UB 2640 (5), UB 2720 (4), UEC 119224 (10), UEC 126400 (4), UEC 127448 (9), UEC 161085 (6), UEC 23496 (9), UEC 32434 (7), UEC 55735 (3), UEC 66434 (10)

Goiás (187): BHCB 96282 (9), CGMS 14251 (12), CGMS 34705 (8), CGMS 34706 (5), ESA 49037 (13), IAC 20866 (8), SPSF 15090 (9), UB 110 (8), UB 1586 (18), UB 4899 (8), UEC 103189 (11), UEC 123550 (13), UEC 123571 (11), UEC 123585 (6), UEC 123602 (8), UEC 126401 (11), UEC 127460 (4), UEC 159809 (5), UEC 7409 (12), UEC 7412 (8)

Maranhão (9): HST 16202 (5), UB 3349 (4)

Minas Gerais (812): BHCB 36593 (24), BHCB 87350 (13), BHCB 1008 (14), BHCB 10715 (8), BHCB 10784 (7), BHCB 108961 (9), BHCB 120638 (11), BHCB 122115 (3), BHCB 129878 (1), BHCB 132249 (9), BHCB 133933 (8), BHCB 133943 (14), BHCB 133952 (3), BHCB 138781 (8), BHCB 14895 (11), BHCB 17100 (10), BHCB 18913 (9), BHCB 22510 (11), BHCB 24594 (8), BHCB 24748 (9), BHCB 25609 (9), BHCB 33630 (9), BHCB 43706 (10), BHCB 4397 (11), BHCB 4662 (7), BHCB 4794 (9), BHCB 49491 (10), BHCB 51493 (5), BHCB 54356 (9), BHCB 58469 (7), BHCB 58471 (10), BHCB 62968 (9), BHCB 63473 (7), BHCB 65790 (9), BHCB 65793 (9), BHCB 69669 (10), BHCB 8825 (11), BHCB 8886 (5), ESA 110860 (9), FLOR 28039 (7), FLOR 28771 (6), FUEL 39442 (18), FURB 3014 (5), HCF 6135 (10), HUFU 0000 (14), HUFU 2778 (9), HUFU 32718 (10), HUFU 34850 (11), HUFU 3657 (9), HUFU 45088 (22), HUFU 45765 (10), HUFU 45766 (11), HUFU 50243 (12), HUFU 50759 (10), HUFU 54350 (15), HUFU 54893 (9), HUFU 635 (12), IAC 50782 (9), IPA 29655 (13), UB 27030 (10), UB 3265 (12), UB 381 (11), UB 98 (3), UEC 119478 (17), UEC 123560 (10), UEC 123615 (18), UEC 123659 (9), UEC 123664 (11), UEC 123973 (11), UEC 126290 (9), UEC 149107 (6), UEC 156337 (5), UEC 159810 (7), UEC 159811 (13), UEC 22904 (7), UEC 2460 (15), UEC 52381 (7), UEC 7416 (12), UEC 95962 (7), UEC 95987 (10), UFRJ 123004 (17), UFRJ 64240 (8)

Mato Grosso (62): BHCB 95802 (7), IPA 17752 (14), UEC 156984 (5), UEC 51976 (10), UEC 9185 (12), UEC 95207 (14)

Mato Grosso do Sul (383): CGMS 13939 (9), CGMS 14481 (11), CGMS 14946 (5), CGMS 17437 (7), CGMS 2073 (19), CGMS 23163 (7), CGMS 24924 (19), CGMS 2513 (5), CGMS 33495 (24), CGMS 33645 (8), CGMS 33646 (17), CGMS 33647 (16), CGMS 33648 (4), CGMS 34703 (9), CGMS 34704 (7), CGMS 36829 (19), CGMS 36830 (18), CGMS 36831 (31), CGMS 36832 (21), CGMS 4118 (9), CPAP 330 (12), IAC 45566 (14), UB 3 (17), UEC 119469 (9), UEC 156360 (10), UEC 156362 (7), UEC 156363 (7), UEC 25912 (9), UEC 48251 (13), UEC 69026 (10), UEC 79286 (10)

Pernambuco (14): IPA 12952 (14)

Piauí (102): EAC 9064 (7), EAC 9113 (16), HST 13310 (17), HST 13311 (26), UEC 75067 (11), UEC 77116 (8), UFRN 9238 (17)

São Paulo (266): CGMS 33643 (11), CGMS 33644 (8), ESA 10728 (9), ESA 88974 (16), IAC 51129 (15), IAC 51801 (13), IAC 52949 (11), SPSF 22049 (7), SPSF 2732 (3), SPSF 3391 (8), SPSF 36747 (9), SPSF 3759 (2), UEC 149188 (11), UEC 151013 (6), UEC 151017 (15), UEC 151027 (9), UEC 156359 (7), UEC 156368 (6), UEC 163121 (6), UEC 24991 (9), UEC 25899 (12), UEC 25932 (10), UEC 26033 (11), UEC 7413 (8), UEC 7415 (15), UEC 84409 (16), UEC 91457 (6), UEC 92386 (7)

Tocantins (7): UB 6203 (7)

Leptolobium elegans (3340)

Goiás (15): UEC 119476 (15)

Minas Gerais (412): BHCB 15237 (13), BHCB 25621 (5), BHCB 33629 (8), BHCB 34662 (15), BHCB 59226 (18), FUEL 13922 (29), HUFU 29149 (15), HUFU 43723 (11), HUFU 44952 (17), HUFU 49362 (19), HUFU 50129 (16), HUFU 552 (28), HUFU 9031 (20), UB 550 (7), UEC 1170 (23), UEC 119477 (21), UEC 1226 (21), UEC 25689 (41), UEC 34358 (21), UEC 43208 (16), UEC 440 (28), UEC 77905 (20)

Mato Grosso (23): UEC 119463 (23)

Mato Grosso do Sul (1022): CGMS 2347 (8), CGMS 2477 (8), CGMS 3440 (26), CGMS 4583 (8), CGMS 683 (12), CGMS 8689 (7), BHCB 32446 (30), BHCB 84079 (30), CGMS 10000 (30), CGMS 15418 (7), CGMS 17344 (6), CGMS 17991 (16), CGMS 18043 (12), CGMS 18185 (5), CGMS 19062 (8), CGMS 24723 (6), CGMS 24724 (16), CGMS 27328 (12), CGMS 28682 (14), CGMS 31328 (21), CGMS 33253 (7), CGMS 33556 (8), CGMS 392 (21), CGMS 528 (16), CGMS 533 (20), CGMS 599 (2), CGMS 7464 (9), CGMS 7465 (9), CGMS 9993 (13), CGMS 9994 (21), CGMS 33496 (22), FLOR 24013 (31), UB 114 (16), UEC 1085 (19), UEC 1110 (30), UEC 153895 (25), UEC 156366 (15), UEC 48138 (35), UEC 48267 (20), UEC 63476 (38), CGMS (363)

Pantanal (44): UEC 140397 (11), CGMS (33)

Paraná (433): BHCB 38448 (35), BHCB 39640 (29), BHCB 61318 (26), BHCB 61319 (18), CGMS 5944 (21), FLOR 27649 (39), FUEL 6007 (18), FURB 4872 (9), FURB 8609 (1), HCF 536 (31), HCF6420 (33), HCF9910 (37), UEC 53510 (35), UEC 59310 (19), UEC 75695 (30), UEC 83294 (20), UFRJ 111059 (12), UFRJ 130470 (20)

Santa Catarina (38): FLOR 26072 (38)

São Paulo (1293): BHCB 124710 (35), BHCB 44520 (21), CGMS 24116 (3), ESA 10888 (10), ESA 13278 (9), ESA 20775 (18), ESA 20888 (14), ESA 25856 (7), ESA 40653 (30),

ESA 93921 (31), FUEL 13836 (7), FUEL 21904 (28), FUEL 32044 (12), FUEL 6436 (14), IAC 29495 (17), IAC 47206 (14), IPA 16256 (15), IPA 60841 (17), UB 4301 (5), UEC 104813 (18), UEC 119470 (21), UEC 13185 (25), UEC 147979 (25), UEC 151019 (10), UEC 1533 (12), UEC 156364 (31), UEC 163118 (29), UEC 25898 (23), UEC 303 (49), UEC 34165 (22), UEC 38727 (18), UEC 41257 (30), UEC 44045 (15), UEC 44743 (37), UEC 44759 (23), UEC 44761 (28), UEC 46103 (21), UEC 63676 (36), UEC 63738 (25), UEC 68333 (21), UEC 7417 (23), UEC 7418 (28), UEC 7419 (16), UEC 7420 (24), UEC 7421 (14), UEC 75063 (12), UEC 77836 (36), UEC 82468 (19), UEC 84377 (13), UEC 84379 (22), UEC 84380 (10), UEC 84393 (14), UEC 84398 (9), UEC 84401 (26), UEC 84402 (13), UEC 87306 (18), UEC 87307 (19), UEC 9183 (25), UEC 98842 (19), UEC 98845 (15), UEC 99939 (22), UEC 99951 (36), CGMS (34)

Argentina (15): UEC 170194 (15)

Paraguay (45): FUEL 48022 (45)

Chapter 3

Article to be submitted to Biochemical Systematics and Ecology, ISSN: 0305-1978

Development and characterization of microsatellite markers for *Echinopsis rhodotricha* and cross-amplification in other species of Cactaceae

Resumo

O Chaco brasileiro, extremo nordeste do amplo domínio chaquenho, possui 12 espécies de Cactaceae. Apesar de ser uma considerada uma região de alta prioridade para conservação, faltam informações básicas a respeito da diversidade de espécies, bem como detalhes de suas distribuições geográficas, taxonomia, ecologia e diversidade genética. Este trabalho apresenta o isolamento e a caracterização de loci de microssatélites de *Echinopsis rhodotricha* (Trichocereeae) e a amplificação heteróloga em outras 13 espécies de cactos. Desenvolvemos 12 loci de microssatélites a partir de uma biblioteca genômica enriquecida, oito desses loci foram polimórficos e caracterizados em 48 indivíduos de três populações de *E. rhodotricha*. Os *loci* apresentaram uma média de 3,2 alelos por locus e a heterozigosidade esperada e observada variou de 0,00 à 0,83 e de 0,00 à 0,66, respectivamente. Cinco *loci* mostraram desvios significativos do equilíbrio de Hardy-Weinberg e exibiram sinais de presença de alelos nulos. A amplificação dos loci em outras espécies de Cactaceae foi bem-sucedida, sobretudo nos exemplares pertencentes à tribo Trichocereeae. O desenvolvimento destes marcadores de microssatélites contribuirá para a realização de estudos de estrutura populacional, diversidade genética e fluxo gênico nas populações de *E. rhodotricha*, bem como de outras espécies do gênero.

Abstract

There are 12 species of Cactaceae in the Brazilian Chaco, in the extreme northeast of the Chaquenian Region. Despite being considered a high priority region for conservation, lack basic information about the total diversity of species and details about their geographical distribution, taxonomy, ecology and genetic diversity. This study reports the characterization of microsatellite loci isolated from *Echinopsis rhodotricha* (Trichocereeae), and the cross-amplification in other 13 cacti species. Twelve microsatellite loci were developed from an enriched genomic library, and eight of these were polymorphic and characterized in 48 individuals from three *E. rhodotricha* populations. The loci showed a mean of 3.2 alleles per locus and overall levels of expected and observed heterozygosities ranging from 0.00 to 0.83 and 0.00 to 0.66, respectively. Five loci showed significant departures from the Hardy-Weinberg equilibrium and also exhibited signs of null allele. Cross-amplification in other Cactaceae species was successful, mainly in the Trichocereeae tribe. The development of these microsatellite markers will contribute to investigations of the population structure, genetic diversity, and gene flow in the *E. rhodotricha* populations as well as in other genus species.

Introduction

The Chaco or “Gran Chaco” is a vegetation type that covers a wide plain (ca. 800,000 km^2) in Argentina, Bolivia, Paraguay and Brazil. The Brazilian Chaco (Figure 1) is restricted to a narrow strip parallel to the river Paraguay in Mato Grosso do Sul state (Prado 1993), occupying a transitional area, among the Pantanal wetlands and the Brazilian Savanna (Cerrado). This region, which comprises ca. 20,000 km^2, has ca. 7,800 km^2 of natural vegetation, with a predominance of stepic savanna with ca. 6,700 km^2 (ZEE-MS 2015). Its vegetation is usually associated with saline soils, and includes a great diversity of Leguminosae, Bromeliaceae and Cactaceae species (Pennington et al. 2000, Silva et al. 2000). In Brazil, Cactaceae is represented by 260 species (187 endemics) and only twelve of them are reported to the Brazilian Chaco (Eggli 2002, Freitas et al. 2013, Taylor et al. 2015, Zappi et al. 2015). Despite being considered a high priority region for conservation (Tálamo & Caziani 2003), there are no conservation units in this area mainly because of the lack of basic information about the total diversity of species, as well as details of their geographical distribution, taxonomic, ecological and genetic aspects.

Echinopsis Zucc. (Trichocereeae) is one of the largest Cactaceae genus of South America that comprises 128 poorly understood species. Plants vary from large-treelike to small-globose, and is obviously complex and in need of DNA analysis and extensive field research (Anderson 2001). *Echinopsis rhodotricha* K. Schum. occurs in Chaco formations and is locally threatened by deforestation for planting pastures (Silva et al. 2011). Sometimes *E. rhodotricha* is found within temporary water puddles, since the soil of the Chaco is very compressed (Prado 1993) and hampers the drainage of water.

A major role of conservation is to preserve genetic variation in all biodiversity levels. Genetic diversity is one of three forms of biodiversity recognized by the International Union for Conservation of Nature (IUCN) as deserving conservation, along with species and ecosystem diversity (McNeely et al. 1990). Conservation genetics is a discrete discipline

focusing on the consequences arising from reduction of once-large, outbreeding, populations to small units where stochastic factors and the effects of inbreeding are paramount (Frankham et al. 2004). The field of conservation genetics also includes the use of molecular genetic analyses to elucidate aspects of specie' biology relevant to conservation management, such as the deleterious effects of inbreeding on reproduction and survival (inbreeding depression); loss of genetic diversity and ability to evolve in response to environmental change (loss of evolutionary potential); fragmentation of populations and reduction in gene flow; resolution of taxonomic uncertainties; definition of management units within species; use of molecular genetic analyses in forensics and elucidation of aspects of species biology important to conservation (Frankham et al. 2004).

Molecular markers has been helpful in defining alleles and studying genetic flow, population structure, paternity, inheritability, genetic maps and conservation genetics (Frankham et al. 2004). Simple sequence repeat markers (SSRs), commonly referred as microsatellite markers, are tandem repeats of short DNA motifs, typically 1–6 bases in length. Microsatellites have advantages over other methods to measure DNA variation as they are highly variable, individual genotypes can be directly inferred, and individuals can be typed following non-invasive sampling. They have the disadvantage that the primers must be developed for each species, although primers from closely related species will often work (Frankham et al. 2004, Barbará et al. 2007).

The use of SSR-markers in genetic variation studies of Cactaceae species is relatively new. In a survey in the Web of Science® (Institute of Scientific Information, Thomson Scientific) using the key-words 'Cactaceae' and 'microsatellite', we found 28 studies (Table 1) that characterized new microsatellite markers (68%) and/or described the genetic diversity of cacti species using cross-amplification loci (32%). Considering all studies, only five of them were developed in Brazil, mainly with *Pilosocereus* genus and using heterologous microsatellite markers (Table 1). This represents ca. 0.02% of all

Brazilian species. Until now, there are no studies in the Brazilian Chaco region, neither with *Echinopsis* genus. Here we describe 12 new microsatellite loci for *E. rhodotricha*, and their cross-species amplification across 13 species within the Cactaceae.

Materials and Methods

To identify and characterize microsatellites, genomic DNA was extracted from root tip fragments of one *Echinopsis rhodotricha* individual following the CTAB protocol of Roy et al. (1992). A microsatellite-enriched library was constructed according to Billotte et al. (1999) with modifications described in Zanella et al. (2012). Sixty positive clones were bi-directionally sequenced with T7 and SP6 primers in an automated sequencer ABI 377 (Perkin-Elmer, Applied Biosystems) using the Big Dye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems). Sequences were aligned and edited using Chromas (Technelysium) and Chromatogram Explorer (Heracle Biosoft). The web software WebSat (Martins et al. 2009) was used to design 20 primers pairs.

The amplifications were performed in a Veriti Thermal Cycler (Applied Biosystems) in 10 μL reactions containing ca. 10 ng of DNA template, 1x Taq buffer, 2 mM $MgCl_2$, 0.2 mM of dNTP set, 4 pmol forward primer, 4 pmol reverse primer, and 0.5 U Taq DNA polymerase (GoTaq; Promega, Madison, WI, USA). A touchdown cycling was used: 95 °C for 3 min, 10 cycles of 94 °C for 30 s, 58 °C decreasing to 48 °C at 1 °C per cycle for 30 s, 72 °C for 30 s, followed by 30 cycles of 94 °C for 30 s, 48 °C for 30 s, and 72 °C for 30 s, and concluding with a 10 min extension at 72 °C. Amplification products were verified by electrophoresis on 2% agarose gel (Promega, Fitchburg, WI, USA) stained with GelRed™ (Biotium, Hayward, CA, USA) and visualized under ultraviolet light. A locus was considered successfully amplified when one band of the expected size was visualized. A 100-bp DNA Ladder (Promega, Fitchburg, WI, USA) was used as a molecular size marker.

To assess the polymorphisms levels, 48 individuals from three *E. rhodotricha* populations (16 per population) were genotyped. All of them were collected in patches of wooded Chaco located in Mato Grosso do Sul State, Brazil (Table 2). For each designed microsatellite marker, the forward primers were synthesized with a 19-bp M13 tail (5'- CAC GAC GTT GTA AAA CGA C -3') following the method of Schuelke (2000).

All PCR amplifications were performed in a Veriti Thermal Cycler (Applied Biosystems) in 10 µL reactions following the protocol described in Faggioni et al. (2014), containing ca. 10 ng of DNA template, 1x Taq buffer, 1.5 mM $MgCl_2$, 0.4 mM dNTP set, 0.1 µM of the forward primer, 0.1 µM of the reverse primer, 0.3 µM of the universal fluorescent M13 primer (6-FAM, NED, PET or VIC), and 0.25 U GoTaq Flexi DNA polymerase (Promega, Fitchburg, WI, USA). We also used the same touchdown cycling program as described above.

The microsatellite alleles genotyping was performed by Macrogen (Seoul, South Korea) using an ABI 3730XL Genetic Analyzer (Applied Biosystems). Fragment size determination was done with Geneious 8.0.3 (Biomatters) and GS500 LIZ as the molecular size standard (Applied Biosystems). The number of alleles, and expected and observed heterozygosities were analyzed with MSA 4.00 software (Dieringer & Schötterer 2003). Genepop 4.1 (Raymond & Rosset 1995) was used to test for departure from the Hardy-Weinberg equilibrium (HWE). The frequency of null alleles, i.e., alleles that, for various reasons, do not appear in a PCR product (Oddou-Muratorio et al. 2009) were calculated following Brookfield (1996) for each population using the software Micro-Checker version 2.2.3 (Van Oosterhout et al. 2004).

To evaluate cross-species amplification, we tested the 12 microsatellites primer pairs described here in 13 cacti species obtained from the Cactus Collection of the National Semi-Arid Institute (Instituto Nacional do Semiárido — INSA, Campina Grande, Paraíba State, Brazil). One individual of each species was analyzed with the same PCR amplifications

conditions described above (without the universal fluorescent M13 primer). Amplification products were visualized in 2% agarose gels, and fragments were sized by comparison with a 100-bp DNA Ladder (Promega, Fitchburg, WI, USA). The cross-species amplifications were considered as successful when the generated fragments had similar sizes to those observed in *E. rhodotricha*.

Results

From the 20 *E. rhodotricha* tested microsatellite markers, twelve were successfully amplified by PCR. Eight of these were polymorphic and used for further characterization of *E. rhodotricha* populations. The primer sequences and product sizes are shown in Table 2.

The number of alleles per locus ranged from 1 to 6 by population, with an average of 3.2, and the expected and observed heterozygosity ranged from 0.00 to 0.83 and from 0.00 to 0.66, respectively (Table 3). Five loci showed significant departures from the HWE due to heterozygote deficiency in the populations (Table 3), and also exhibited signs of null alleles (Table S1).

Cross-species amplification was successful in different degrees (Table 4). All loci amplified in the congeneric species, *Echinopsis calochlora*. EchiA09, EchiC11, EchiD11, and EchiG10 were successful in more than 50% of the species evaluated indicating high potential for use of these loci in population genetic studies with Cactaceae.

Discussion

The eight polymorphic loci developed here are consistent with findings for another studies (mean 9.06 ± 4.00 loci per species; Table 1) where a microsatellite-enriched library were used.

When an allele fails to amplify in a PCR assay it may indicate the presence of null alleles. According to Dabrowski et al. (2015) these null alleles in the data may lead to: (i)

bias of the genetic diversity estimates of the population; (ii) the exclusion of the true parents in parentage analyses; (iii) a reduction in the power to correctly assign individuals to populations; (iv) bias of the estimation of genetic differentiation among populations. All five loci that present null alleles are dinucleotide microsatellite (mostly GA motifs, Table 2), and three of these loci were cross-amplified in only one or two species (Table 4). There is an association between the presence of null alleles and highly variable flanking regions (Chapuis & Estoup 2007). On the other hand, microsatellite unit-repeat length or motif complexity, which are the factors related to the mutation rate of the microsatellite repeat region (Jin et al. 1996, Chakraborty et al. 1997), presented no correlation with the null allele frequency (Li et al. 2003).

The null allele frequency increases as the phylogenetic distance between the species increases as well (Chapuis & Estoup 2007). Furthermore, the level of polymorphism observed in the species in which the loci have been described occasionally cannot be observed in related species, especially with the increase of the evolutionary distance (Rubinsztein et al. 1995, Morin et al. 1998). This could explain the low cross-amplification success among members of the Cacteae tribe (Table 4), for example.

The 12 microsatellite loci reported here are the first set of molecular markers developed for the genus *Echinopsis*, and will be useful for describing the genetic diversity and population structure across the distribution range of *E. rhodotricha*. These markers will also be used for understanding patterns of gene flow and to provide information about changes in historical distribution and about connectivity among cacti populations. Further, this set of markers may be useful in management plans for conserving the endangered congeneric species *E. calochlora*, endemic to iron-rich mountains of the Pantanal wetlands.

Acknowledgments

Coordination for the Improvement of Higher Education Personnel (CAPES) for WS Fava's Ph.D fellowship. National Counsel of Technological and Scientific Development (CNPq), Ação Transversal nº6/2011 — Casadinho/Procad (Angiospermas do Chaco Brasileiro: Sistemática, Diversidade, Fenologia e Adaptações) for financial support. WS Covre for his assistance with field work and DNA extraction. AP Souza from University of Campinas (Unicamp) for the development of microsatellite enriched library. VGN Gomes for providing cacti samples from INSA (National Semi-Arid Institute).

References

Anderson EF. 2001. The Cactus Family. Portland: Timber Press.

Arakaki M, Soltis DE, Soltis OS & Speranza PR. 2010. Characterization of polymorphic microsatellite loci in *Haageocereus* (Trichocereeae, Cactaceae). American Journal of Botany 97: e17–e19.

Arakaki M, Speranza P, Soltis OS & Soltis DE. 2013. Genetic Variability of an Unusual Apomictic Triploid Cactus – *Haageocereus* tenuis Ritter – from the Coast of Central Peru. Journal of Heredity 2013: 104:127–133.

Baker MA & Butterworth CA. 2013. Geographic distribution and taxonomic circumscription of populations within *Coryphantha* section *Robustispina* (Cactaceae). American Journal of Botany 100: 984–997.

Barbará T, Palma-Silva C, Paggi, GM, Bered F, Fay MF & Lexer C. 2007. Cross-species transfer of nuclear microsatellite markers: potential and limitations. Molecular Ecology 16: 3759–3767.

Billotte N, Lagoda PJL, Risterucci AM & Baurens FC. 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54: 277–288.

Bonatelli IAS, Perez MF, Peterson AT, Taylor NP, Zappi DC, Machado MC, Pires AHC & Moraes EM. 2014. Interglacial microrefugia and diversification of a cactus species complex: phylogeography and paleodistributional reconstructions for *Pilosocereus aurisetus* and allies. Molecular Ecology 23: 3044–3063.

Brookfield JFY. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5: 453–455.

Butterworth CA. 2011. Isolation and characterization of 10 polymorphic microsatellite loci in *Coryphantha robustispina* ssp. *robustispina*. Conservation Genetics Resources 3:247–249.

Caruso M, Currò S, Las Casas G, La Malfa S & Gentile A. 2010. Microsatellite markers help to assess genetic diversity among *Opuntia ficus indica* cultivated genotypes and their relation with related species. Plant Systematics and Evolution 290: 85–97.

Chakraborty R, Kimmel M, Stivers DN, Davison LJ & Deka R. 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proceedings of the National Academy of Sciences USA 94: 1041–1046.

Chapuis M & Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24: 621–631.

Chessa I, Erre P, Barbato M, Nieddu G & Ochoa J. 2013. Polymorphic microsatellite DNA markers in *Opuntia* spp. collections. In: VII International Congress on Cactus Pear and Cochineal 995 (pp. 35-41).

Contreras-Negrete G, Ruíz-Durán ME, Cabrera-Toledo D, Casas A, Vargas O & Parra F. 2015. Genetic diversity and structure of wild and managed populations of *Polaskia chende* (Cactaceae) in the Tehuacán-Cuicatlán Valley, Central Mexico: insights from SSR and allozyme markers. Genetic Resources and Crop Evolution 62: 85–101.

Cruse-Sanders JM, Parker KC, Friar EA, Huang DI, Mashayekhi S, Prince LM, Otero-Arnaiz A & Casas A. 2013. Managing diversity: Domestication and gene flow in *Stenocereus stellatus* Riccob. (Cactaceae) in Mexico. Ecology and Evolution 3: 1340–1355.

Dabroeski MJ, Bornelöv S, Kruczyk M, Baltzer N & Komorowski J. 2015. 'True' null allele detection in microsatellite loci: a comparison of methods, assessment of difficulties, and survey of possible improvements. Molecular Ecology Resources 15: 477–488.

Dieringer D & Schlötterer C. 2003. Microsatellite Analyser (MSA): A platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes 3: 167–169.

Eggli U. 2002. Synopsis of the Cactaceae of Mato Grosso, Brazil. Haseltonia, 9: 146–166.

Faggioni GP, Zamudio KR, Souza FL & Prado CPA. 2014. Isolation and characterization of microsatellites markers for two South American frogs (*Leptodactylus bufonius* and *L. chaquensis*) using next generation sequencing. Amphibia-Reptilia 35: 405–412.

Flores CG, Garza OAL, Luz JLL de la & León FJG de. 2014. Development and characterization of 10 microsatellite loci in the giant cardon cactus, *Pachycereus pringlei* (Cactaceae). Applications in Plant Sciences 2: 1300066.

Frankham R, Ballou JD & Briscoe DA. 2004. A primer of conservation genetics. Cambridge: University Press, 200p.

Freitas TG, Souza CS, Aoki C, Arakaki LMM, Stefanello TH, Sartori ALB & Sigrist MR. 2013. Flora of Brazilian humid Chaco: Composition and reproductive phenology. Checklist 9: 973–979.

Hardesty BD, Hughes SL, Rodriguez VM & Hawkins JA. 2008. Characterization of microsatellite loci for the endangered cactus *Echinocactus grusonii*, and their cross-species utilization. Molecular Ecology Resources 8: 164–167.

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.

Hughes SL, Rodriguez VM, Hardesty BD, Luna RTB, Hernández HM, Robson RM & Hawkins JA. 2008. Characterization of microsatellite loci for the critically endangered cactus *Ariocarpus bravoanus*. Molecular Ecology Resources 8: 1068–1070.

Jin L, Macaubas C, Hallmayer J, Kimura A & Mignot E. 1996. Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. Proceedings of the National Academy of Sciences USA 93: 15285–15288.

Li G, Hubert S, Bucklin K, Ribes V & Hedgecock D. 2003. Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas*. Molecular Ecology Notes 3: 228–232.

Martins WS, Lucas DCS, Neves KFS & Bertioli DJ. 2009. WebSat - A web software for microsatellite marker development. Bioinformation 3: 282–283.

McNeely JA, Miller KR, Reid WV, Mittermeier RA & Werner TB. 1990. Conserving the world's biological diversity. IUCN, World Resources Institute, Conservation International, WWF-US, and the World Bank, Washington, DC.

Monteiro ER, Mangolin CA, Neves AF das, Orasmo GR, Silva JGM da & Machado MFPS. 2015. Genetic diversity and structure of populations in *Pilosocereus gounellei* (F.A.C. Weber ex K. Schum.) (Cactaceae) in the Caatinga biome as revealed by heterologous microsatellite primers. Biochemical Systematics and Ecology 58: 7–12.

Moraes EM, Cidade FW, Silva GAR & Machado MC. 2014. Polymorphic microsatellite markers for the rare and endangered cactus *Uebelmannia pectinifera* (Cactaceae) and its congeneric species. Genetics and Molecular Research 13: 10359–10366.

Moraes EM, Perez MF, Téo MF, Zappi DC, Taylor NP & Machado MC. 2012. Cross-species amplification of microsatellites reveals incongruence in the molecular variation and taxonomic limits of the *Pilosocereus aurisetus* group (Cactaceae). Genetica 140: 277–285.

Morin PA, Mahboubi P, Wedel S & Rogers J. 1998. Rapid screening and comparison of human microsatellite markers in baboons: allele size is conserved, but allele number is not. Genomics 53: 12–20.

Oddou-Muratorio S, Vendramin G, Buiteveld J & Fady B. 2009. Population estimators or progeny tests: what is the best method to assess null allele frequencies at SSR loci? Conservation Genetics 10: 1343–1347.

Otero-Arnaiz A, Casas A, Hamrick JL & Cruse-Sanders J. 2005a. Genetic variation and evolution of *Polaskia chichipe* (Cactaceae) under domestication in the Tehuacán Valley, central Mexico. Molecular Ecology 14: 1603–1611.

Otero-Arnaiz A, Casas A & Hamrick JL. 2005b. Direct and indirect estimates of gene flow among wild and managed populations of *Polaskia chichipe*, an endemic columnar cactus in Central Mexico. Molecular Ecology 14: 4313–4322.

Otero-Arnaiz A, Cruse-Sanders J, Casas A & Hamrick JL. 2004. Isolation and characterization of microsatellites in the columnar cactus: *Polaskia chichipe* and cross-species amplification within the Tribe Pachycereeae (Cactaceae). Molecular Ecology Notes 4: 265 –267.

Parra F, Casas A, Peñaloza-Ramírez JM, Cortés-Palomec AC, Rocha-Ramírez V & González-Rodríguez. 2010. Evolution under domestication: ongoing artificial selection and divergence of wild and managed *Stenocereus pruinosus* (Cactaceae) populations in the Tehuacán Valley, Mexico. Annals of Botany 106: 483–496.

Pennington RT, Prado DE & Pendry CA. 2000. Neotropical seasonally dry forest and quaternary vegetation changes. Journal of Biogeography 27: 261–273.

Perez MF, Téo MF, Zappi DC, Taylor NP & Moraes EM. 2011. Isolation, characterization, and cross-species amplification of polymorphic microsatellite markers for *Pilosocereus machrisii* (Cactaceae). American Journal of Botany: e204–e206.

Prado DE, Gibbs PE, Pott A & Pott VJ. 1992. The Chaco – Pantanal transition in southern Mato Grosso, Brazil. In: Nature and dynamics of forest savanna boundaries (Eds. PA Furley & JA Proctor). Chapman & Hill, London. Pp 451-470.

Prado DE. 1993. What is the Gran Chaco vegetation in South America? 1. A review. Contribution to the study of flora and vegetation of the Chaco. V. Candollea 48: 145–172.

Raymond M, Rosset F. 1995. Genepop (version 1.2): Population genetics software for exact test and ecumenicism. Journal of Heredity 86: 248–249.

Roy A, Frascaria N, Mackay J & Bousquet J. 1992. Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. Theoretical and Applied Genetics 85: 173–180.

Rubinsztein DC, Amos W, Leggo J, Goodburn S, Jain S, Li SH, Margolis RL, Ross AC & Ferguson-Smith MA. 1995. Microsatellites evolution – evidence for directionality and variation in rate between species. Nature Genetics 10: 337–343.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18: 233–234.

Schwabe AL, Hubbard AR, Neale JR & McGlaughlin ME. 2013. Microsatellite loci development for rare Colorado *Sclerocactus* (Cactaceae). Conservation Genetics Resources 5: 69–72.

Schwabe AL, Neale JR & McGlaughlin ME. 2015. Examining the genetic integrity of a rare endemic Colorado cactus (*Sclerocactus glaucus*) in the face of hybridization threats from a close and widespread congener (*Sclerocactus parviflorus*). Conservation Genetics 16: 443–457.

Silva JSV, Pott A, Abdon MM, Pott VJ & Santos EKR. 2011. Projeto GeoMS: cobertura vegetal e uso da terra do Estado de Mato Grosso do Sul. 1ª ed., 64 p. Embrapa Informação Agropecuária, Campinas, São Paulo, Brasil.

Silva MP, Mauro R, Mourão G & Coutinho M. 2000. Distribuição e quantificação de classes de vegetação do Pantanal através de levantamento aéreo. Revista Brasileira de Botânica 23: 143–152.

Solórzano S, Cortés-Palomec AC, Ibarra A, Dávila P & Oyama K. 2009. Isolation, characterization and cross-amplification of polymorphic microsatellite loci in the threatened endemic *Mammillaria crucigera* (Cactaceae). Molecular Ecology Resources 9: 156–158.

Solórzano S, Cuevas-Alducin PD, García-Gómez V & Dávila P. 2014. Genetic diversity and conservation of *Mammillaria huitzilopochtli* and *M. supertexta*, two threatened species endemic of the semiarid region of central Mexico. Revista Mexicana de Biodiversidad 85: 565–575.

Tálamo A & Caziani SM. 2003. Variation in wood vegetation among sites with different disturbance histories in the Argentine Chaco. Forest Ecology and Management 184: 79–91.

Taylor N, Santos MR, Larocca J & Zappi D. 2015. Cactaceae. In: Lista de Espécies da Flora do Brasil, Rio de Janeiro: Jardim Botânico do Rio de Janeiro. Available from URL: http://www.floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB70.

Terry M, Pepper AE & Manhart JR. 2006. Development and characterization of microsatellite loci in endangered *Astrophytum asterias* (Cactaceae). Molecular Ecology Notes 6: 865–866.

Terry MK, Pepper AE, Strong AW, Tarin DM, Price DM & Manhart JR. 2012. Genetic structure of a population of the endangered star cactus (*Astrophytum asterias*) in Southern Texas. The Southwestern Naturalist 57: 182–188.

Van Oosterhout C, Hutchinson WF, Wills DPM & Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Zanella CM, Janke A, Paggi GM, GoetzeM, Reis MS & Bered F. 2012. Microsatellites in the endangered species *Dyckia distachya* (Bromeliaceae) and cross-amplification in other Bromeliads. Internation Journal of Molecular Sciences 13: 15859–15866.

Zappi, D.; Taylor, N.; Santos, M.R.; Larocca, J. 2015. Cactaceae. In: Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. Available at floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB70. Accessin: 27 Aug 2015.

Zarroug MB, Baraket G, Zourgui L, Souid S & Hannachi AS. 2015. Genetic diversity and phylogenetic relationship among Tunisian cactus species (*Opuntia*) as revealed by random amplified microsatellite polymorphism markers. Genetics and Molecular Research 14: 1423–1433.

ZEE-MS. 2015. Zoneamento Ecológico Econômico do Estado de Mato Grosso do Sul: Elementos para construção da sustentabilidade do território sul-mato-grossene.Segunda Aproximação. Campo Grande, MS.

Table 1. Summary of the SSRs studies reviewed, including the cacti species name, geographic region, the type of microsatellite used, number of loci and references. (Survey conducted in March 20, 2015 in the Web of Science® - Institute of Scientific Information, Thomson Scientific)

Species	Country	SSR source*	Loci number	References
Ariocarpus bravoanus	Mexico	SSR-EL	8	Hughes et al. 2008
Astrophytum asterias	USA	SSR-EL	7	Terry et al. 2012
Astrophytum asterias	USA, Mexico	SSR-EL	6	Terry et al. 2006
Coryphantha robustispina	USA	SSR-EL	10	Butterworth 2011
Coryphantha section *Robustispina*	USA, Mexico	SSR-EL	10	Baker and Butterworth 2013
Echinocactus grusonii	Mexico	SSR-EL	12	Hardesty et al. 2008
Haageocereus spp.	Peru	SSR-EL	5	Arakaki et al. 2010
Haageocereus tenuis	Peru	SSR-EL	5	Arakaki et al. 2013
Mammillaria crucigera	Mexico	SSR-EL	8	Solórzano et al. 2009
Mammillaria spp.	Mexico	Cross-amplification	11	Solórzano et al. 2014
Opuntia ficus indica (genotypes)	Italy	Cross-amplification	8	Caruso et al. 2010
Opuntia spp. (accessions)	Argentina, Italy	Cross-amplification	13	Chessa et al. 2013
Opuntia spp. (cultivars)	Tunisia	RAMPO	72	Zarroug et al. 2015
Opuntia spp. (varieties)	Equador	SSR-EL	16	Helsen et al. 2007
Pachycereus pringlei	Mexico	PYRO	10	Flores et al. 2014
Pilosocereus aurisetus complex	Brazil	Cross-amplification	10	Bonatelli et al. 2014
Pilosocereus aurisetus Group	Brazil	Cross-amplification	11	Moraes et al. 2012
Pilosocereus gounellei	Brazil	Cross-amplification	5	Monteiro et al. 2015
Pilosocereus machrisii	Brazil	SSR-EL	10	Perez et al. 2011
Polaskia chende	Mexico	Cross-amplification	7	Contreras-Negrete et al. 2015
Polaskia chichipe	Mexico	SSR-EL	5	Otero-Arnaiz et al. 2005a; 2005b
Polaskia chichipe	Mexico	SSR-EL	7	Otero-Arnaiz et al. 2004
Sclerocactus glaucus	USA	Cross-amplification	13	Schwabe et al. 2015
Sclerocactus spp.	USA	SSR-EL	13	Schwabe et al. 2013
Stenocereus pruinosus	Mexico	Cross-amplification	4	Parra et al. 2010
Stenocereus stellatus	Mexico	SSR-EL	5	Cruse-Sanders 2013
Uebelmannia pectinifera	Brazil	SSR-EL	18	Moraes et al. 2014

*RAMPO = random amplified microsatellite polymorphism; PYRO = pyrosequencing; SSR-EL = SSR-enriched library.

Table 2. Description and characterization of twelve microsatellite loci of *Echinopsis rhodotricha*, including locus name, primer sequences, repeat motif, allele size, mean number of alleles (A), and GeneBank accession number.

Locus[a]	Primer sequences (5' - 3')	Repeat motif	Size (bp)	A	GenBank
EchiA01	F: TCAGCAATCAGGTAGTCAATGT R: TCTACTCCTTCTCAGCCAGC	(TG)6	296	1	KJ572954
EchiA03	F: CTGGGGATTAGCAGACCAAG R: ATTTCCAATTAAGGGGCTGG	(ATAA)3	303-310	4	KJ572955
EchiA04	F: GAACAAACCAATACCCCACG R: CCACCCAACACCTCACTTG	(TG)9A(GA)13	231	1	KJ572956
EchiA05	F: ATCATGTGGGTGGTGGATCT R: CCATAACCGAGACCCTCTTTC	(AG)8	161-165	2	KJ572957
EchiA09	F: AGGGTCATTAGAGAAGCGATTT R: CTCGGTTGTTGATTTCATGC	(GA)13	163-185	4	KJ572958
EchiA12	F: AGCAACAGGAACCAGGCTAA R: ACTTTTACTGCCCATCGACC	(GT)16	210-219	3	KJ572959
EchiC11	F: CCATTGCTGTGCTTGTTTGA R: CCCTTGTGTGAATAGTCGGTTT	(AG)8	302	1	KJ572960
EchiD04	F: CACGACAGAGAAGAAAGGGG R: AAAGATGGTCGGTTACTCGC	(GA)18	280-290	5	KJ572961
EchiD10	F: ACCCTCACAAACACCTGTCC R: CACCCACTAAATTGTTCCTGGT	(TG)7	311	2	KJ572962
EchiD11	F: AGAGTGGAAGTGAGGAGGGC R: TGGCAGCAGTAAAGCAGAGA	(GA)7	170-173	2	KJ572963
EchiG10	F: GAGCGTGTGAGAATTGGGAT R: TGTCGGCATTAGGGGTTAAG	(TC)6	281	1	KJ572964
EchiH02	F: ATCAAGATCAACTGGGGTGG R: CACTAGCTGTCGGTGTCCCT	(AG)17	260-278	4	KJ572965

Note: A = Total number of alleles in a sample of 48 individuals.
[a] Polymorphic loci are shown in bold.

Table 3. Characterization of eight microsatellite loci in three populations of *Echinopsis rhodotricha*, including number of individuals (N), number of alleles (A), observed (Ho) and expected (He) heterozygosities. The coordinates of each population are given in decimal degrees.

Locus	Pop 1 (N = 16) -21.68S; -57.77W			Pop 2 (N = 16) -21.67S -57.77W			Pop 3 (N = 16) -21.70S -57.76W		
	A	**Ho**	**He**	**A**	**Ho**	**He**	**A**	**Ho**	**He**
A03	4	0.50	0.58	5	0.62	0.62	2	0.63	0.45
A05	2	0.38	0.32	2	0.21	0.19	1	0.00	0.00[ni]
A09	5	0.00	0.63*	4	0.14	0.71*	4	0.40	0.75*
A12	2	0.00	0.41*	4	0.27	0.39	4	0.57	0.65
D04	6	0.18	0.78*	5	0.21	0.77*	5	0.16	0.82*
D10	2	0.66	0.45	2	0.60	0.43	2	0.57	0.42
D11	2	0.00	0.53*	1	0.00	0.00[ni]	2	0.00	0.42
H02	3	0.50	0.83	4	0.00	0.72*	4	0.20	0.73*
Mean	**3.3**	**0.28**	**0.57**	**3.4**	**0.26**	**0.48**	**3.0**	**0.32**	**0.53**

* Indicates Ho departed significantly from He under HWE ($p < 0.05$).

[ni] = No information.

Table 4. Cross-amplification of twelve microsatellite markers isolated from *Echinopsis rhodotricha* across thirteen species of Cactaceae.

Species	Echi A01	Echi A03	Echi A04	Echi A05	Echi A09	Echi A12	Echi C11	Echi D04	Echi D10	Echi D11	Echi G10	Echi H02	**Total**
CACTOIDEAE													
Cacteae													
Astrophytum myriostigma	-	-	-	-	-	-	-	-	-	-	-	-	**0**
Echinocactus grusonii	-	-	+	-	+	+	+	-	-	-	+	-	**5**
Ferocactus latispinus	-	-	-	-	+	-	-	-	-	-	-	-	**1**
Mammillaria elongata	-	-	-	-	-	-	-	-	-	-	-	-	**0**
Cereeae													
Cereus albicaulis	-	-	+	-	+	-	+	-	-	+	+	-	**5**
Melocactus zehntneri	+	-	+	-	+	-	-	-	-	-	-	-	**3**
Pilosocereus gounellei	-	-	-	+	-	-	+	-	-	+	+	-	**4**
Trichocereeae													
Discocactus sp.	-	-	-	+	+	-	+	-	-	+	+	-	**5**
Echinopsis calochlora	+	+	+	+	+	+	+	+	+	+	+	+	**12**
Facheiroa squamosa	-	-	+	+	+	-	+	+	+	+	+	-	**8**
Harrisia adscendens	-	-	-	-	+	-	-	-	-	-	-	-	**1**
OPUNTIOIDEAE													
Opuntia microdasys	+	-	+	+	+	-	+	-	-	+	+	-	**7**
Tacinga inamoena	-	-	-	+	+	-	-	-	-	+	-	-	**3**
Total	**3**	**1**	**6**	**6**	**10**	**2**	**7**	**2**	**2**	**7**	**7**	**1**	

Note: + = successful PCR amplification; - = PCR failure.

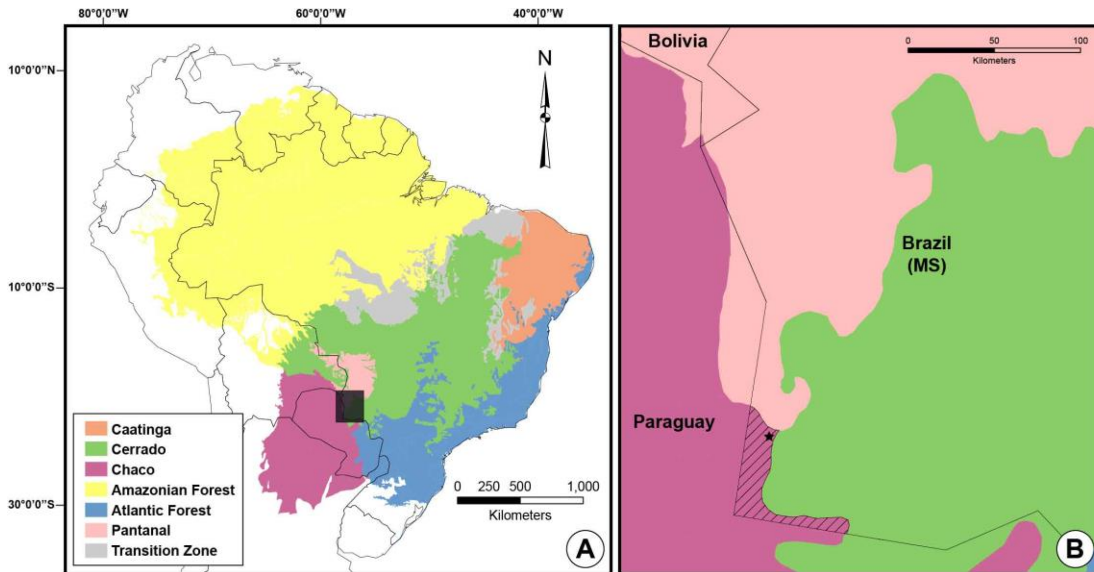

Figure 1. Main phytogeographic formations of South America. A: Locations of the six main Brazilian phytogeographic formations. B: Study area (star) in the Brazilian Chaco portion (shaded area). Map generated by ArcGIS 10.1 software. The maps of phytogeographic regions are available at mapas.mma.gov.br/i3geo/datadownload.htm.

Chapter 3

SUPPORTING INFORMATION

Table S1. Results from Micro-Checker showing possible null alleles in five microsatellite loci developed for *Echinopsis rhodotricha* (Cactaceae, Trichocereeae).

Locus	Null Allele Present	Oosterhout	Chakraborty	Brookfield 1	Brookfield 2
A12	yes	0.2023	0.3278	0.1594	0.6657
D10	no	-0.3784	-0.1812	-0.1321	0.2127
A09	yes	0.3569	0.6021	0.3210	0.7546
H02	yes	0.3680	0.6364	0.3182	0.9215
A03	no	0.0085	0.0068	0.0049	0.4063
A05	no	-0.1084	-0.0541	-0.0178	0.4328
D04	yes	0.3631	0.6114	0.3336	0.6085
D11	yes	0.3703	1.0000	0.2871	0.9213

Chapter 4

Article to be submitted to Biochemical Systematics and Ecology, ISSN: 0305-1978

Genetic variation in *Echinopsis rhodotricha* (Cactaceae) populations of Brazilian Chaco revealed by microsatellite markers

Resumo

Investigamos a diversidade e estrutura genética de populações de *Echinopsis rhodotricha* (Cactaceae) de remanescentes de Chaco arborizado no extremo oeste do Brasil. Seis locos de microssatélites nucleares foram utilizados para analisar 96 indivíduos de seis populações. A diversidade genética mostrou uma riqueza alélica geral de 8,947; heterozigosidade esperada e observada de 0,580 e 0,328, respectivamente. O coeficiente de endogamia foi moderado, variando de 0,255 a 0,568. Todas as populações desviaram significativamente do equilíbrio de Hardy-Weinberg devido a deficiência de heterozigotos. A maior parte da variação genética (93,20%) foi encontrada dentro das populações e não foram detectados sinais de diminuição recente do tamanho populacional. Análises bayesianas revelaram que no Chaco Brasileiro *E. rhodotricha* é composta por dois grupos genéticos, com indivíduos mistos nas seis populações analisadas. Nós apresentamos os primeiros dados sobre a distribuição da variabilidade genética em uma espécie de cacto presente na porção brasileira do Chaco, e que podem ajudar na delimitação e criação de áreas protegidas nesta região.

Abstract

We investigated the genetic diversity and structure of populations of *Echinopsis rhodotricha* (Cactaceae) in wooded remnants of Chaco in the far west of Brazil. Six nuclear microsatellite loci were used to analyze 96 individuals from six populations. The genetic diversity showed an overall allelic richness of 8.947; expected and observed heterozygosity of 0.580 and 0.328, respectively. The inbreeding coefficient was moderate, ranging from 0.255 to 0.568. All populations significantly deviated from Hardy-Weinberg equilibrium due to heterozygote deficiency. Most of the genetic variation (93.20%) was found within populations and signs of recent decrease in population size were not detected. Bayesian analysis revealed that, in the Brazilian Chaco, *E. rhodotricha* comprises two genetic groups, with mixed individuals in the six populations analyzed. We present the first data on the distribution of genetic variation in a species of cactus present in the Brazilian portion of the Chaco, and that can help in the definition and creation of protected areas in the region.

Introduction

In the South American continent we can find open phytogeographical domains occurring across a great variety of environmental conditions, including large climatic, latitudinal and altitudinal ranges. These open domains are organized diagonally, including three tropical/sub-tropical regions: the Seasonally Dry Tropical Forests (with the largest area in northeastern Brazil, Caatinga), the Cerrado savanna (central Brazil), and the Chaco (northeastern Argentina, western Paraguay, and south-eastern Bolivia) (Werneck 2011). Seasonally Dry Tropical Forests and the Cerrado occur under the same climatic conditions (Mooney et al. 1995, Mayle 2004), whereas Chaco is often subject to winter frosts (Pennington et al. 2000). These three regions are seasonally stressed by drought, presenting vegetation adapted to these climatic conditions frequently characterized by complex mosaic-type distributions. Although composed by unique biotas, they have received less research attention than tropical wet forests (Mooney et al. 1995, Furley & Metcalfe 2007).

The Chaco, over lowland alluvial plains of central South America, covers about 840,000 km^2. It is located in northern Argentina, western Paraguay, south-eastern Bolivia (Prado et al. 1992, Prado 1993b, Pennington et al. 2000), and the Brazilian extreme western edge which comprises ca. 20,000 km^2 (ZEE-MS 2015). The region has ca. 7,800 km^2 of natural vegetation, with a predominance of stepic savanna with ca. 6,700 km^2. The climate is marked by strong seasonality, severe summers (up to 48.9 °C) and winter frosts. Considering its extension and occurrence of different environmental gradients, Chaco vegetation shows a pattern of spatial variability including arboreal and savanna-like communities, with few endemic genera but numerous endemic species (Prado 1993a). At the regional level, the Chaco vegetation is determined by rainfall, and locally by the degree of soil drainage and possibility of flooding (Navarro et al. 2006).

Cactaceae is one of the well-represented plant families in open formations. It comprises more than 1600 recognized species (Anderson 2001) and Brazil is considered the

third largest diversity center for this group (Taylor & Zappi 2004). The family presents a great diversity of growth forms, since xerophytic trees or shrubs with conspicuous persistent leaves (*Pereskia* Mill.) to branched (or unbranched) columnar to globular stem succulents (2002). While the cacti of the arid northeastern part of Brazil have become comparatively well-known, this is not the case for southwestern Brazil (Eggli 2002). Twelve species, including globose, epiphytes and columnar cacti were reported in the Brazilian Chaco (Gomes & Araujo 2015). In this region, the combined impacts of livestock and forestry practices have led to a highly fragmented landscape and an impoverished ecological system (Bucher & Huszar 1999), threatening the few remaining areas (Pott & Pott 2003).

Studies concerning floristic, physiological, taxonomic and ecological aspects of cacti in the Brazilian Chaco are scarce (Gomes & Araujo 2015), and there are no studies evaluating its populational genetic diversity. One of the least understood groups of cacti is *Echinopsis* Zucc., which includes 100–150 species that exhibit a great diversity in architecture. The genus occurs from southeastern Brazil to the coast of Chile and to northern Ecuador, with a species diversity center in the eastern Andes of northern Argentina and Bolivia (Schlumpberger & Renner 2012).

Echinopsis rhodotricha K. Schum. occurs only in the central-west region in wooded Brazilian Chaco remnants (Taylor et al. 2015), extending to northern Argentina, Bolivia, Paraguay (Chaco and Oriental regions) and Uruguay (Eggli 2002, Oakley et al. 2013).

The National Plan for Cactaceae Conservation (Silva et al. 2011) highlight that the assessment of the genetic variability levels and its distribution among individuals and populations represent fundamental actions for the understanding of the species' conservation status. Thereby, in this study we aim to assess patterns of genetic diversity and structure in Brazilian Chaco populations of *Echinopsis rhodotricha* (Cactaceae) using nuclear microsatellite markers that were previously constructed based on a microsatellite-enriched library technique. This work is the first to analyze the genetic diversity of a cactus species

that occurs in the Brazilian Chaco, and might serve as a basis for decision-making on the conservation and management of this region.

Methods

Population sampling

We sampled 96 plants of *Echinopsis rhodotricha* in patches of wooded Brazilian Chaco distributed in six populations (Figure 1A): Population 1 (-21.68274, -57.77875), Population 2 (-21.67477, -57.77779), Population 3 (-21.70204, -57.76340), Population 4 (-21.68514, -57.71909), Population 5 (-21.69516, -57.71556), and Population 6 (-21.74608, -57.50941). Root tip fragments from each individual (16 by population) were stored in silica gel for further DNA extraction following the CTAB protocol (Cetyl trimethylammonium bromide) of Roy et al. (1992).

Molecular markers and genotyping assays

For the population genetic analyses of *Echinopsis rhodotricha* we used six nuclear microsatellite markers which were previously isolated for this species (see Chapter 3). For each designed microsatellite marker, the forward primers were synthesized with a 19-bp M13 tail (5'- CAC GAC GTT GTA AAA CGA C -3') following the method of Schuelke (2000).

All amplifications were performed in a Veriti Thermal Cycler (Applied Biosystems) in 10 µL reactions following the protocol described in Faggioni et al. (2014), containing ca. 10 ng of DNA template, 1x Taq buffer, 1.5 mM $MgCl_2$, 0.4 mM dNTP set, 0.1 µM of the forward primer, 0.1 µM of the reverse primer, 0.3 µM of the universal fluorescent M13 primer (6-FAM, NED, PET or VIC), and 0.25 U GoTaq Flexi DNA polymerase (Promega, Fitchburg, WI, USA). A touchdown cycling was used: 95 °C for 3 min, 10 cycles of 94 °C for 30 s, 58 °C decreasing to 48 °C at 1 °C per cycle for 30 s, 72 °C for 30 s, followed by 30

cycles of 94 °C for 30 s, 48 °C for 30 s, and 72 °C for 30 s, and concluding with a 10 min extension at 72 °C. Amplification products were verified by electrophoresis on 2% agarose gel (Promega, Fitchburg, WI, USA) stained with GelRed™ (Biotium, Hayward, CA, USA) and visualized under ultraviolet light. A 100-bp DNA Ladder (Promega, Fitchburg, WI, USA) was used as a molecular size marker.

The microsatellite alleles genotyping was performed by Macrogen (Seoul, South Korea) using an ABI 3730XL Genetic Analyzer (Applied Biosystems). Fragment size determination was done with Geneious 8.0.3 (Biomatters) using GS500 LIZ as the molecular size standard (Applied Biosystems).

Genetic diversity of the sampled loci and populations

In order to characterize the microsatellite loci, the number of alleles (A), expected heterozygosity (H_e), observed heterozygosity (H_o), and the within- and total-population inbreeding coefficients F_{IS} and F_{IT} were calculated for each locus using the computer programs MSA (Dieringer & Schlotterer 2003) and FSTAT (Goudet 1995). In addition, departures from Hardy–Weinberg equilibrium (HWE) for each locus within populations were tested using FSTAT. The frequency of null alleles (i.e., allele nonamplification) was calculated following Brookfield (1996) for each population using the software MICRO-CHECKER version 2.2.3 (Van Oosterhout et al. 2004). Subsequently, each population was characterized using the variance in allele size (Var), H_e, and H_o calculated by MSA and the allelic richness in FSTAT. Departures from HWE for each population were identified using exact tests in GENEPOP (Raymond & Rousset 1995).

Gene flow analyses

To determine if the *Echinopsis rhodotricha* populations experienced recent reductions in effective population size (e.g., genetic bottlenecks), we performed a heterozygosity excess test using the 'sign test' and 'Wilcoxon sign-rank' test in the BOTTLENECK program (Piry et al. 1999). The analyses were carried out both for the 'Infinite Allele Model' (IAM) and for the 'Two-Phased Mutation model' (TPM).

Isolation by distance (Wright 1943) was tested using a regression analysis of Slatkin's (1995) linearized F_{ST} transformation ($F_{ST}/1-F_{ST}$) onto the natural log of geographic distance (Rousset 1997). A Mantel test (Mantel 1967) (10,000 permutations) was performed in GENEPOP to determine the significance of this relationship.

Analyses of Molecular Variance (AMOVA) were performed to estimate the distribution of the inter- and intra- populational genetic variation and the F_{ST} values between pairs of populations. The significance of the F-statistics analyses was tested through 10,000 permutations in ARLEQUIN 3.5 (Excoffier & Lischer 2010).

Effective population sizes and migration rates

We estimated the population parameters Theta ($\theta = 4N_e\mu$, with N_e = effective population size and μ = mutation rate) and the Effective Number of Migrants (N_em) between pairs of *E. rhodotricha* populations. The computations were carried out following a coalescent and maximum-likelihood-based approaches using MIGRATE 2.0.6 (Beerli & Felsenstein 1999). The effective population sizes were estimated from Theta values by assuming a rate of 10^{-3} mutation per gamete per generation (Zhang & Hewitt 2003).

Bayesian genetic structure

In order to obtain additional insights into gene flow patterns and population subdivision in *E. rhodotricha*, we achieved a Bayesian analysis in STRUCTURE version 2 (Pritchard et al.

2000) to determine the most likely number of populations (K) and to estimate admixture proportions (Q) for individuals of each population. The analyses were carried out under the admixture model for independent allele frequencies using a burn-in period of 50,000, run length of 300,000 (based on the diagnostic tools available in STRUCTURE), and 10 iterations per K (from 1 to 8) to confirm stabilization of summary statistics. To determine the most likely number of clusters (K), we used the method proposed by Evanno et al. (2005) and available in STRUCTURE HARVESTER website (Earl & von Holdt 2012; taylor0.biology.ucla.edu/structureHarvester/), which is based on an ad hoc measure of ΔK that evaluates the second-order rate of change of the likelihood function with respect to K.

Results

Genetic diversity

The six microsatellite loci presented up to 12 alleles per locus and gene diversities (H_e) of up to 0.736. Except for locus EchiD10, all others presented significant within-population (F_{IS}) and total-population inbreeding coefficients (F_{IT}) as well as consistent departures from HWE (Table 1).

Overall genetic diversity indices as well for each of the six *E. rhodotricha* populations analyzed are presented in Table 2. Considering all loci, the six *E. rhodotricha* populations analyzed departed significantly from HWE. Loci EchiA09, EchiA12 and EchiD04, presented signs of a null allele presence (data not shown).Allelic richness in the different localities sampled ranged from 2.061 to 2.411, with a variance in allele size of up to 11.401. The expected and observed heterozygosities among the localities ranged from 0.492 to 0.645 and 0.225 to 0.435, respectively. The within-localities inbreeding coefficient (F_{IS}) was relatively high, varying from 0.255 to 0.568, and all localities deviated significantly from HWE, showing a heterozygotes deficit. Regarding the overall genetic diversity indices, we found an allelic richness of 8.947, with variance in allele size of 8.559. The expected

(0.580) and observed (0.328) heterozygosities and the inbreeding coefficient (0.541) also deviated from HWE.

Neither the sign test nor the Wilcoxon sign-rank test for recent population bottlenecks were significant for any of the populations, regardless of the mutation model used (data not shown). Mantel test indicated significant correlation ($r^2 = 0.216$ $^{p\ =\ 0.0019}$) between geographical and genetic distance (Figure 2), thus suggesting the presence of isolation by distance in populations of *E. rhodotricha*. Under isolation by distance, values of the ratios are expected to increase linearly with the logarithm of distance (Rousset 1997).

Genetic structure and gene flow

The AMOVA results showed that most of the genetic variation (93.20%) was attributed to differences within populations, and only 6.80% among populations ($p = 0.06287$). Individual F_{ST} estimated between pairs of *E. rhodotricha* showed not significant fixation index values (all populations combinations reached $F_{ST} = 0.000$). Maximum-likelihood-based estimatives of migration rates (N_em) were low in *E. rhodotricha* populations (Figure 3), showing less than one migrant per generation.

Bayesian genetic structure

Bayesian analyses indicated that the most likely number of groups in the total sample was two (Figure 4). The average admixture proportion for each population of *E. rhodotricha* among the two different genetic clusters is given in Table 3, and the admixture proportions of individual plants in each of the two clusters are graphically represented in Figure 1B.

Discussion

We found relatively low levels of genetic diversity in the *Echinopsis rhodotricha* populations from Brazilian Chaco (Table 2). The average allelic richness (2.3) and the observed heterozygosity (0.328) were lower when compared to other cacti species that occurs in Brazil, e.g., *Pilosocereus aurisetus* (Werderm.) Byles & G.D. Rowley complex, R_S 3.5/ H_o 0.503 (Bonatelli et al. 2014); *Pilosocereus gounellei* (F.A.C.Weber ex K.Schum.), R_S 3.6/ H_o 0.333 (Monteiro et al. 2015); *Pilosocereus machrisii* (E. Y. Dawson) Backeb., R_S 4.4/ H_o 0.403 (Perez et al. 2011); and *Uebelmannia pectinifera* Buining, R_S 3.9/ H_o 0.599 (Moraes et al. 2014). However, the allelic richness, when considering one single population (8.9), was substantially high, indicating that the populations of the Brazilian Chaco, although in the border of the *E. rhodotricha* distribution, can hold high levels of genetic diversity.

All of the six *E. rhodotricha* populations sampled departed from HWE, which may indicate an inbreeding process. In fact, the overall F_{IS} index was higher (0.541) i.e., further inbreeding than expected at random mating populations is occurring. Furthermore, the genetic differentiation indices suggested low levels of genetic differentiation among the populations. This pattern was also confirmed by AMOVA analysis, which showed that only 6.80% of the genetic variability was attributed to differences among localities.

Bayesian analysis revealed that *E. rhodotricha* populations are composed by two genetic groups (Figure 1A): one formed by the populations 1, 2 and 3, and one by the junction of the remaining ones (4, 5 and 6). However, these groups are not completely differentiated, with some mixed individuals within the six populations. Population 6, farther away from the others (ca. 20 km), was similar to populations 4 and 5. This may be due the fact that the *E. rhodotricha* seeds are numerous (more than 1000 per fruit) and very small (up to 1.5 mm), and therefore, may be dispersed by birds, lizards and ants (Gomes & Araujo 2015). Thus, the genetic composition of population 6 may be a consequence of long-distance

dispersal events despite the presence of elevations up to 400 m crossing the chquenian plain where the populations are located (see Figure 1).

The lack of population structure, i.e., low values of F_{ST}, point towards high levels of gene flow among localities, which was not confirmed by the migration rates estimated that showed less than one migrant per generation (Figure 3). However, Morjan & Rieseberg (2004) reviewing the implications of gene flow and selection for the spread of advantageous alleles, found that levels of gene flow between 0.5 and 1 migrants per generation are very common for plant, animal and fungal species. Indeed, in many taxonomic groups, even those with moderate gene flow, there are species with low gene flow levels ($N_em < 1$), particularly selfing plants, snails, amphibians and freshwater fishes (Morjan & Rieseberg 2004).

Although the reproductive system of *E. rhodotricha* is not known, self-incompatibility has been observed in other species of the genus (e.g., *E. chiloensis* (Colla) Friedrich & G.D.Rowley, Ossa & Medel 2011; *E. terscheckii* (Britton & Rose) H. Friedrich & G.D. Rowley, Ortega-Baes et al. 2011; *E. schickendantzii* F.A.C. Weber, Alonso-Pedano & Ortega-Baes 2012). *Echinopsis rhodotricha* have funnel-shaped white flowers that provide pollen and nectar as floral rewards, open at night and remain open until the next morning (Gomes & Araujo 2015). The prolonged anthesis, which favors the participation of diurnal and nocturnal pollinators, suggests a generalist pollination system. Similar results were recorded to *Echinopsis schickendantzii* (Alonso-Pedano & Ortega-Baes 2012) that is pollinated by bees and moths, and to *E. terscheckii* for which moths, bees and birds were identified as floral visitors. The prolonged anthesis may be a strategy to ensure sexual reproduction when the main pollinator (probably nocturnal) is unpredictable (Fleming et al. 2001), and could explain the high rate of inbreeding found, since a wide range of flower visitors would be causing a pollen flow between related individuals.

In conclusion, *Echinopsis rhodotricha* showed low genetic structure and diversity indices similar to others cacti species. The distances among the six populations analyzed

(between 5 and 20 km apart) may be easily overcome by pollinators and dispersers, which thus maintain population's cohesion. Besides, it is possible that the six populations sampled in the Chaco are formed by two (or even one) large populations that are now fragmented but still maintains a gene flow between them. In addition, this study provides the first evidence on the distribution of genetic variability in a cactus species located in the Brazilian Chaco, and can help the delimitation and creation of protected areas in the region, since it is highly impacted by deforestation due to formation of cultivated pasture.

Acknowledgments

Coordination for the Improvement of Higher Education Personnel (CAPES) for WS Fava's Ph.D fellowship. National Counsel of Technological and Scientific Development (CNPq), Ação Transversal nº6/2011 — Casadinho/Procad (Angiospermas do Chaco Brasileiro: Sistemática, Diversidade, Fenologia e Adaptações) for financial support.

References

Alonso-Pedano M & Ortega-Baes P. 2012. Generalized and complementary pollination system in the Andean cactus *Echinopsis schickendantzii*. Plant Systematics and Evolution 298: 1671–1677.

Anderson EF. 2001. The Cactus Family. Portland, Timber Press, 777 pp.

Beerli P & Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152: 763–773.

Bonatelli IA, Perez MF, Peterson AT, Taylor NP, Zappi DC, Machado MC, Koch I, Pires AHC & Moraes EM. 2014. Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. Molecular Ecology 23: 3044–3063.

Brookfield JFY. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5: 453–455.

Bucher EH & Huszar PC. 1999. Sustainable management of the Grand Chaco of South America: ecological promise and economic constraints. Journal of Environmental Management 57: 99–108.

Dieringer D & Schlötterer C. 2003. Microsatellite Analyser (MSA): A platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes 3: 167–169.

Earl DA. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359–361.

Eggli U. 2002. Synopsis of the Cactaceae of Mato Grosso, Brazil. Haseltonia 9: 146–166.

Evanno G, Regnaut S & Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Excoffier L & Lisxher HEL. 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.

Faggioni GP, Zamudio KR, Souza FL & Prado CPA. 2014. Isolation and characterization of microsatellites markers for two South American frogs (*Leptodactylus bufonius* and *L. chaquensis*) using next generation sequencing. Amphibia-Reptilia 35: 405–412.

Fleming TH, Sahley CT, Holland JN, Nason JD & Hamrick JL. 2001. Sonoran desert columnar cacti and the evolution of generalized pollination systems. Ecological Monographs 71: 511–530.

Furley PA & Metcalfe SE. 2007. Dynamic changes in savanna and seasonally dry vegetation through time. Progress in Physical Geography 31: 633–642.

Gomes VG & Araujo AC. 2015. Cacti species from the Brazilian Chaco: floral and fruit traits. Gaia Scientia 9: 1–8.

Goudet J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of Heredity 86: 485–486.

Mantel N. 1967. The detection of disease clustering and generalized regression approach. Cancer Research 27: 209–220.

Mayle FE. 2004. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. Journal of Quaternary Science 19: 713–720.

Monteiro ER, Mangolin CA, das Neves AF, Orasmo GR, da Silva JGM & da Silva MDFP. 2015. Genetic diversity and structure of populations in *Pilosocereus gounellei* (FAC Weber ex K. Schum.) (Cactaceae) in the Caatinga biome as revealed by heterologous microsatellite primers. Biochemical Systematics and Ecology 58: 7–12.

Mooney HA, Bullock SH & Medina E. 1995. Introduction. In: Seasonally Dry Tropical Forests (Eds. SH Bullock, HA Mooney & E Medina), pp.1–8. Cambridge University Press, Cambridge.

Moraes EM, Cidade FW, Silva GAR & Machado MC. 2014. Polymorphic microsatellite markers for the rare and endangered cactus *Uebelmannia pectinifera* (Cactaceae) and its congeneric species. Genetics and Molecular Research 13: 10359–10366.

Morjan CL & Rieseberg LH. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Molecular Ecology 13: 1341–1356.

Navarro G, Molina JÁ & Molas LP. 2006. Classification of the forests of the northern Paraguayan Chaco. Phytocoenologia 36: 473–508.

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: 312–326.

Oakley L, Duarte W & Pin A. 2013. *Echinopsis rhodotricha*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on June 26, 2015.

Ortega-Baes P, Saraiva M, Suhring S, Godínez-Alvarez H & Zamar M. 2011. Reproductive biology of *Echinopsis terscheckii* (Cactaceae): the role of nocturnal and diurnal pollinators. Plant Biology 13: 33–40.

Ossa CG & Medel R. 2011. Notes on the floral biology and pollination syndrome of *Echinopsis chiloensis* (Colla) Friedrich & G.D. Rowley (Cactaceae) in a population of semiarid Chile. Gayana Botanica 68: 213–219.

Pennington RT, Prado DE & Pendry CA. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. Journal of Biogeography 27: 261–273.

Perez MF, Téo MF, Zappi DC, Taylor NP & Moraes EM. 2011. Isolation, characterization, and cross-species amplification of polymorphic microsatellite markers for *Pilosocereus machrisii* (Cactaceae). American Journal of Botany 98: e204–e206.

Piry S, Luikart G & Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90: 502–503.

Pott A & Pott VJ. 2003: Espécies de fragmentos florestais em Mato Grosso do Sul. In: Fragmentação florestal e alternativas de desenvolvimento rural na Região Centro-Oeste (Ed. RB Costa), p. 113–126. UCDB, Campo Grande.

Prado DE, Gibbs PE, Pott A & Pott VJ. 1992. The Chaco-Pantanal transition in southern Mato Grosso, Brazil. In: Nature and Dynamics of Forest-Savanna Boundaries (Eds. PA Furley, J Proctor & JA Ratter), pp. 451–470. Chapman & Hall, Londres, Greater London, Reino Unido.

Prado D. 1993a. What is the Gran Chaco vegetation in South America? I. A review. Contribution to the study of flora and vegetation of the Chaco. Candollea 48: 145–172.

Prado D. 1993b. What is the Gran Chaco vegetation in South America? II. A redefinition. Contribution to the study of flora and vegetation of the Chaco. Candollea 48: 615–629.

Pritchard JK, Stephens M & Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Raymond M & Rosset F. 1995. Genepop (version 1.2): Population genetics software for exact test and ecumenicism. Journal of Heredity 86: 248–249.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219–1228

Roy A, Frascaria N, Mackay J & Bousquet J. 1992. Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. Theoretical and Applied Genetics 85: 173–180.

Schlumpberger BO & Renner SS. 2012. Molecular phylogenetics of *Echinopsis* (Cactaceae): Polyphyly at all levels and convergent evolution of pollination modes and growth forms. American Journal of Botany 99: 1335–1349.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18: 233–234.

Silva SR, Zappi DC, Taylor N & Machado M. 2011. Plano de ação nacional para a conservação das cactáceas. Série Espécies Ameaçadas 24, Instituto Chico Mendes de Conservação da Biodiversidade (ICMBIO), Braspilia, 112 pp.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457–462

Taylor NP & Zappi DC. 2004. Cacti of Eastern Brazil. Kew: Royal Botanic Gardens.

Taylor N, Santos MR, Larocca J & Zappi D. 2015. Cactaceae. In: Lista de Espécies da Flora do Brasil, Rio de Janeiro: Jardim Botânico do Rio de Janeiro. Available from URL: http://www.floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB70.

Van Oosterhout C, Hutchinson WF, Wills DPM & Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Werneck FP. 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. Quaternary Science Reviews 30: 1630–1648.

Wright S. 1943. Isolation by distance. Genetics 38: 114–138.

ZEE-MS. 2015. Zoneamento Ecológico Econômico do Estado de Mato Grosso do Sul: Elementos para construção da sustentabilidade do território sul-mato-grossene.Segunda Aproximação. Campo Grande, MS.

Zhang D-X & Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology 12: 563–584.

Table 1. Characterization of six microsatellite loci in *Echinopsis rhodotricha*, including number of alleles (A), expected (H_e) and observed (H_o) heterozygosity, within-population inbreeding coefficient (F_{IS}), and total-population inbreeding coefficient (F_{IT}).

Locus[a]	A	H_e	H_o	F_{IS}	F_{IT}
EchiA03	10	0.633	0.467	0.230	0.361
EchiA09	12	0.736	0.316	0.655	0.682
EchiA12	10	0.593	0.370	0.412	0.568
EchiD04	9	0.619	0.094	0.797	0.810
EchiD10†	10	0.568	0.625	-0.139*	0.198
EchiD11	5	0.330	0.097	0.580	0.760

[a]For all populations, N = 16 individuals.

* Non-significant (p > 0.05).

† Locus in Hardy Weinberg equilibrium (p > 0.05).

Table 2. Characterization of *Echinopsis rhodotricha* populations with six nuclear microsatellite markers, including the variance in allele size (Var), allelic richness (R_S), expected (H_E) and observed (H_O) heterozygosities, as well as the within-localities inbreeding coefficient (F_{IS}).

Population[a]	Var	R_S	H_e	H_o	F_{IS}
1	11.401	2.166	0.568	0.225	0.568
2	7.032	2.061	0.492	0.295	0.307
3	10.138	2.274	0.591	0.391	0.255
4	5.746	2.330	0.621	0.342	0.373
5	8.718	2.321	0.563	0.282	0.487
6	8.318	2.411	0.645	0.435	0.309
Overall	8.559	8.947	0.580	0.328	0.541

All observed heterozygosities (H_o) and inbreeding coefficient (F_{IS}) departed significantly from Hardy Weinberg Equilibrium ($p < 0.01$).

Table 3. Average admixture proportion (Clusters I and II) in *Echinopsis rhodotricha* populations inferred by Bayesian analysis in STRUCTURE.

Population	Clusters	
	I	II
1	0.054	0.946
2	0.030	0.970
3	0.067	0.933
4	0.906	0.094
5	0.954	0.046
6	0.941	0.059

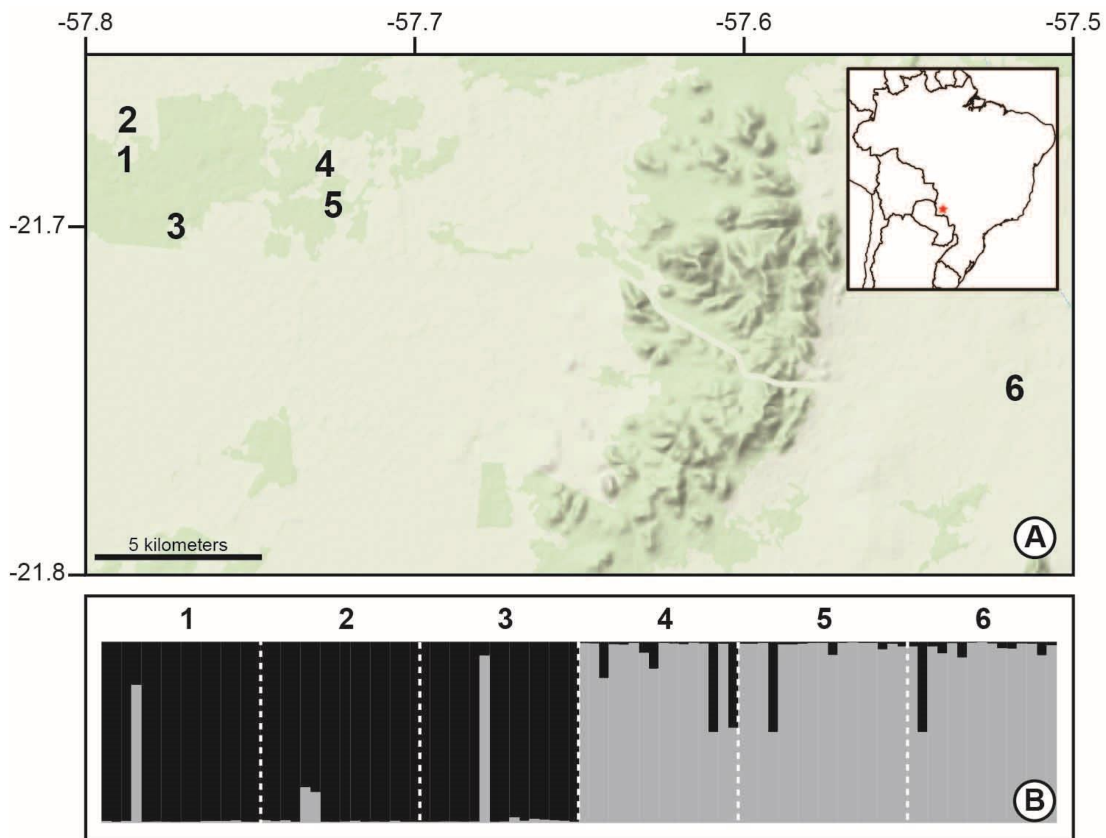

Figure 1. Distribution map (A) of *Echinopsis rhodotricha* (Cactaceae) populations sampled in the Brazilian Chaco. The dark green areas represent regions with Chaco vegetation, and the portion between populations 5 and 6 is a mountain range near the Porto Murtinho municipality. (B) Bayesian admixture proportions for K = 2 genetic clusters, identified by the program STRUCTURE. The genetic clusters are indicated in different colors (Cluster I: gray; Cluster II: black). The numbers and the vertical dashed lines indicate the sampled populations.

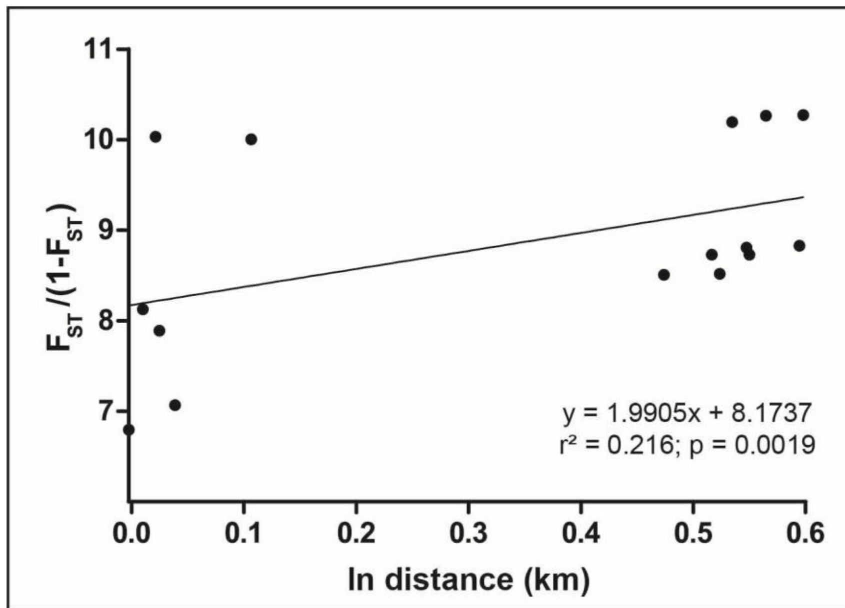

Figure 2. Isolation by distance relationship in *Echinopsis rhodotricha* populations, where pairwise $F_{ST}/(1-F_{ST})$ was regressed over the natural logarithm of distances (in km).

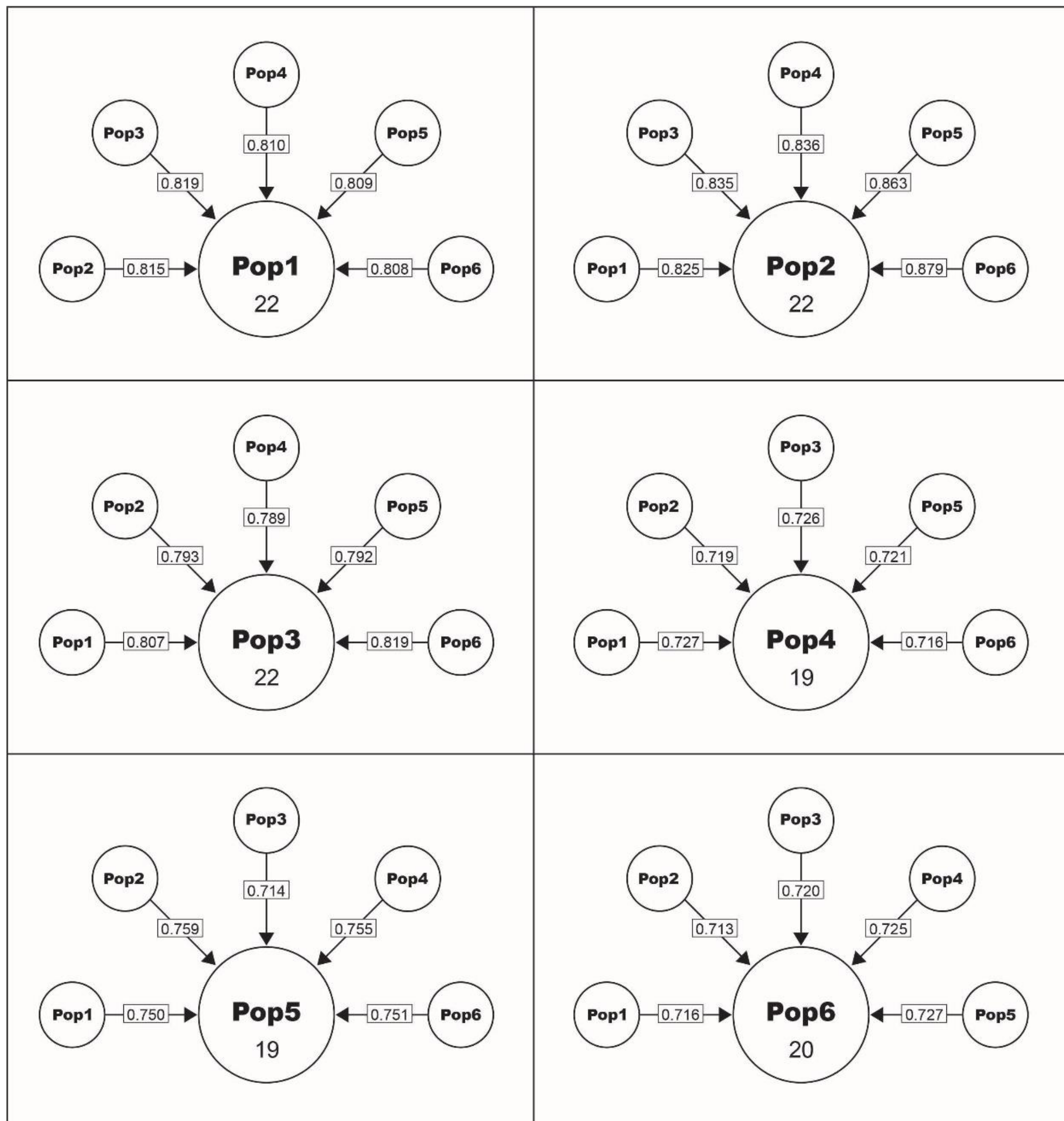

Figure 3. Effective population sizes (N_e, number under the population ID) and number of migrants (N_em) for each *Echinopsis rhodotricha* population sampled in Brazilian Chaco.

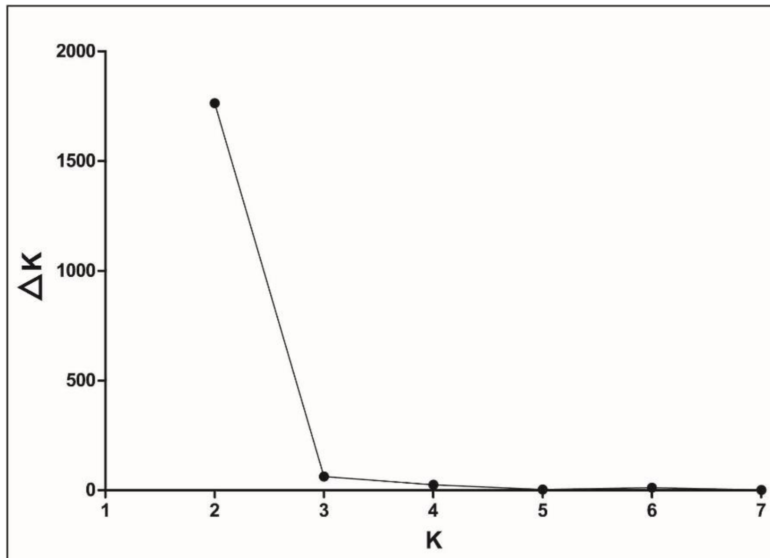

Figure 4. Magnitude of ∆K as a function of K (mean ± SD over 10 replicates), calculated for the Brazilian Chaco population of *Echinopsis rhodotricha* (Cactaceae) using the procedure proposed by Evanno et al. (2005). The modal values of these distributions indicate the true K or the uppermost level of structure is two "genetic clusters".

Conclusão Geral

Considerações finais

Pela presente tese fica evidente a importância em se usar fontes de dados não tradicionais para a realização de estudos ecológicos, moleculares e de modelagem preditiva. O uso de exsicatas de herbário como complemento de amostragem e até mesmo como principal método de coleta mostrou-se eficaz na predição dos padrões fenológicos e de distribuição potencial de espécies arbóreas do Cerrado. Através desses resultados, vimos que *Leptolobium dasycarpum* e *L. elegans* apresentam períodos de floração similares, com picos no início da estação chuvosa, em outubro/novembro. Além disso, *L. dasycarpum*, que possui ampla distribuição no Cerrado, teve sua data de floração alterada com o aumento da temperatura nos últimos 50 anos, demonstrando que análises fenológicas baseadas em espécimes de herbário constituem um método efetivo para se detectar os efeitos das mudanças climáticas na ecologia das espécies.

A modelagem de distribuição potencial de espécies gera resultados úteis tanto para o início de um estudo, mostrando as áreas mais prováveis de ocorrência, como para a discussão de resultados das mais diversas áreas, como a filogeografia, que precisa do perfil histórico da distribuição das espécies para uma adequada interpretação dos dados. Nesse trabalho fomos capazes de predizer não apenas a distribuição de duas espécies de leguminosas arbóreas em diferentes cenários climáticos, como também mensuramos a sobreposição de seus nichos ecológicos. Durante o último período glacial, *Leptolobium dasycarpum* e *L. elegans* tiveram suas distribuições reduzidas, porém não fragmentada, formando aréas de concentração na região central do Cerrado, coincidindo com os locais de maior diversidade genética encontrados.

Ao contrário do que tem sido reportado para leguminosas arbóreas do Cerrado, as populações de *L. dasycarpum* e *L. elegans* não apresentaram sinais de divergência em alopatria e a formação de grupos geograficamente diferenciados. Este padrão foi

corroborado pelas análises de modelagem que mostraram que a distribuição destas espécies, apesar de reduzidas, permaneceram relativamente estáveis na região centro-oeste durante as eras glaciais. Além disso, foi evidenciada a origem recente das linhagens de ambas as espécies e uma provável zona de hibridação interespecífica nas áreas de simpatria. Tradicionalmente, a transição Cerrado–Pantanal tem sido pouco amostrada nos estudos filogeográficos, mas os resultados aqui gerados indicam que a região pode representar um importante reservatório de diversidade genética das espécies arbóreas do Cerrado.

A partir dos marcadores microssatélites desenvolvidos especificamente para a cactácea *Echinopsis rhodotricha*, e testada com sucesso em outras espécies da família, estimamos a estrutura genética de populações desta espécie na estreita porção do Chaco brasileiro. A maior parte da variação genética foi encontrada dentro das populações e não foram detectados sinais de diminuição recente do tamanho populacional. Esse padrão também foi encontrado para *Pilosocereus gounellei* F.A.C.Weber ex K.Schum. (Cactaceae, Cereeae) na Caatinga, fazendo com que os dados aqui gerados possam ser utilizados na comparação da diversidade genética de cactáceas ocorrentes na diagonal seca brasileira. Além disso, é possível que as seis populações amostradas no Chaco sejam formadas por duas grandes populações que agora se encontram fragmentadas, mas que ainda mantém um fluxo gênico entre elas.