

MAGALI GONÇALVES GARCIA

**FILOGEOGRAFIA MOLECULAR DE *Cedrela fissilis* Vell. E *C. odorata* L.
(MELIACEAE) NO BRASIL E BOLÍVIA**

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

VIÇOSA
MINAS GERAIS – BRASIL
2010

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

T

G216f
2010

Garcia, Magali Gonçalves, 1980-
Filogeografia molecular de *Cedrela fissilis* Vell.
e *C. odorata* L. (Meliaceae) no Brasil e Bolívia / Magali
Gonçalves Garcia. – Viçosa, MG, 2010.
xii, 98f. : il. (algumas col.) ; 29cm.

Inclui apêndices.

Texto em português e inglês.

Orientador : Luiz Orlando Oliveira.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. *Cedrela fissilis* - Genética. 2. *Cedrela odorata* -
Genética. 3. Evolução molecular. 4. Filogeografia.
5. *Cedrela fissilis* - Distribuição geográfica.
6. *Cedrela odorata* - Distribuição geográfica.
7. Diversidade das plantas - Conservação. I. Universidade
Federal de Viçosa. II. Título.

CDD 22. ed. 583.78

MAGALI GONÇALVES GARCIA

**FILOGEOGRAFIA MOLECULAR DE *Cedrela fissilis* Vell. E *C. odorata* L.
(MELIACEAE) NO BRASIL E BOLÍVIA**

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

APROVADA: 30 de setembro de 2010.



Prof. Jorge Abdala Dergam dos Santos
(Co-orientador)



Prof. Agostinho Lopes de Souza



Prof.ª Flavia Cristina Pinto Garcia



Prof.ª Telma Nair Santana Pereira



Prof. Luiz Orlando de Oliveira
(Orientador)

Aos meus pais, Dovani e Circe.

Ao meu esposo Cesar.

DEDICO

AGRADECIMENTOS

A Deus, por estar ao meu lado durante esses anos ensinando-me a acertar o caminho, abençoando-me sempre, fortalecendo-me nos momentos de dificuldade, consolando-me nos momentos de tristeza e proporcionando-me momentos de alegria como este. Obrigado Senhor!

A Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento pelo acolhimento e qualificação.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão à bolsa de estudos.

Ao Prof. Dr. Luiz Orlando de Oliveira, pela orientação e ensinamentos recebidos, pelo incentivo e compreensão em todos os momentos.

Aos professores do programa de Programa de Pós-Graduação em Genética e Melhoramento pelo conhecimento compartilhado.

Aos pesquisadores, professores, familiares e colegas que gentilmente coletaram e enviaram folhas e frutos de Cedro para a realização deste trabalho.

Aos meus pais, Dovani e Circe, por me oportunizarem a vida e me ensinarem a vivê-la com dignidade, através de seus exemplos de honestidade, força, perseverança, amor e fé. Pelo incentivo e confiança em todos os momentos.

Ao meu esposo Cesar, pelo companheirismo, incentivo e apoio incondicional em todos os momentos. Por ter me acompanhado e acreditado que eu seria capaz de concluir mais esta etapa.

A toda minha família, em especial ao meu irmão Maycon, que mesmo distantes compartilham comigo o sabor das pequenas conquistas e me incentivam a continuar a caminhada, saibam que nada seria possível sem o apoio de vocês.

Aos amigos, Andressa, Candida, Claudinéia, Cristina, Daiane, Dierlei, Felipe, Frederico, Lucir e Renata, pelo apoio dado a mim e ao meu esposo durante nossa estadia em Viçosa.

Aos colegas de laboratório, Ana Paula, Ana Rossi, Bruna, Camila, Carol, Cassiana, Christina, Denise, Fernanda, Flávia, Giselle, Larissa, Leonardo, Lídia, Lílían, Lucas, Maíra, Márcia, Mariana Bicalho, Mariana Cardoso, Michellia, Roberta, Beatriz, Maria Andréia, Pricila, Rafaella, Roméria, Saulo, Talles e Valquíria pelo apoio nas rotinas de laboratório e pelo companheirismo.

Aos demais colegas de curso pelo convívio.

Aos funcionários do Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), pela convivência.

A todos aqueles que de forma direta ou indireta contribuíram para a realização desse trabalho.

BIOGRAFIA

MAGALI GONÇALVES GARCIA, filha de Dovani Garcia e Circe Gonçalves Garcia, nasceu na cidade de Dourados, Mato Grosso do Sul, em 17 de setembro de 1980.

Concluiu o ensino fundamental em 1994, na Escola Estadual Presidente Vargas em Dourados-MG e o ensino médio na mesma instituição em 1997.

Graduou-se no Curso de Ciências Biológicas da Universidade Federal de Mato Grosso do Sul na modalidade licenciatura em abril de 2002 e na modalidade bacharelado em fevereiro de 2003.

Em março de 2003, iniciou o curso de mestrado no Programa de Pós-graduação em Biotecnologia, área de concentração Agrícola e Florestal na Universidade Federal de Santa Catarina, submetendo-se aos exames finais de defesa de dissertação para obtenção do título de *Magister Scientiae*, em 21 de fevereiro de 2005.

Em maio de 2006, ingressou no curso de doutorado no Programa de Pós-Graduação em Genética e Melhoramento da Universidade Federal de Viçosa, concentrando seus estudos na área de Biologia Molecular de plantas.

SUMÁRIO

RESUMO	ix
ABSTRACT	xi
INTRODUÇÃO GERAL.....	1
REFERÊNCIAS BIBLIOGRÁFICAS.....	4

CAPÍTULO I: Molecular evidences of historical range fragmentation, range expansion and intraspecific hybridization in the Neotropical seasonal forest tree *Cedrela fissilis* (Meliaceae).

ABSTRACT	6
INTRODUCTION	7
MATERIALS AND METHODS	10
Sampling strategy and DNA extraction	10
Assembly of ITS datasets.....	11
Analyses of secondary structures in the ITS region	13
Assembly of the cpNDA dataset.....	14
Phylogenetic analysis.....	15
RESULTS	17
Intraindividual polymorphism for ITS	17
Structure analysis in the ITS region.....	17
Bayesian analysis based on ITS.....	19
Bayesian analysis based on concatenated cpDNA.....	20
DISCUSSION	21

Multiple, functional copies for ITS	21
Incomplete concerted evolution in <i>C. fissilis</i>	22
<i>C. fissilis</i> is not monophyletic.....	22
Association of phylogenetic clades with geographic ranges	23
Genetic connectivity between formerly isolated lineages	24
ACKNOWLEDGMENTS	26
REFERENCES	27

CAPÍTULO II: Filogeografia molecular de *Cedrela fissilis* Vell. e *C. odorata* L. baseada em genes nuclear e cloroplastídico.

RESUMO	47
INTRODUÇÃO	48
MATERIAL E MÉTODOS	53
Estratégia de amostragem e extração de DNA.....	53
Montagem do conjunto de dados de cloroplasto	58
Montagem do conjunto de dados da região de ITS	59
Índices de diversidade nucleotídica	63
Estimativa das redes de haplótipos	63
RESULTADOS	64
Haplótipos da região espaçadora <i>trnT-trnL</i> do genoma de cloroplasto	64
Haplótipos da região de ITS do genoma nuclear	66
Relação genealógica e geográfica entre as linhagens de cpDNA.....	68
Relação genealógica e geográfica entre as linhagens de ITS	73
Compartilhamento de haplótipos entre <i>Cedrela fissilis</i> e <i>Cedrela odorata</i>	77
Índices de diversidade nucleotídica	77
DISCUSSÃO	81

Linhagens ancestrais	81
Linhagens do Atlântico	82
Linhagens de Chiquitano.....	86
Biogeografia molecular.....	87
CONCLUSÕES.....	91
REFERÊNCIAS BIBLIOGRÁFICAS	93

RESUMO

GARCIA, Magali Gonçalves, D.Sc., Universidade Federal de Viçosa, setembro de 2010. **Filogeografia molecular de *Cedrela fissilis* Vell. e *C. odorata* L. (Meliaceae) no Brasil e Bolívia.** Orientador: Luiz Orlando de Oliveira. Co-orientadores: João Augusto Alves Meira Neto e Jorge Abdala Dergam dos Santos.

Embora seja detentora de uma ampla biodiversidade, a flora neotropical tem sido pouco investigada. A perda do patrimônio genético ainda se torna mais grave pelo desconhecimento que temos da abrangência deste patrimônio e do seu valor atual ou potencial. Os fatores que determinam os padrões de distribuição das espécies sobre a superfície terrestre ainda não são totalmente entendidos. Da mesma forma, desconhecemos as normas que regem a distribuição da diversidade dentro de espécies. Abordagens filogenéticas e filogeográficas com base em dados moleculares são ferramentas poderosas para refinar hipóteses biogeográficas que visam explicar as distribuições contemporâneas de populações. No presente estudo, foram utilizadas ferramentas de filogeografia molecular para explorar a história evolutiva das florestas estacionais por meio de uma perspectiva proveniente da espécie *Cedrela fissilis* e de espécies relacionadas, utilizando análise de dados de sequências de DNA dos genomas cloroplastídico e nuclear. Os resultados mostraram que a coesão esperada entre as cópias de rDNA repetidas em *tandem* foi interrompida e a evolução em concerto permaneceu incompleta, resultando no elevado polimorfismo intraindividual observado em *C. fissilis*. Dessa forma, múltiplos parálogos funcionais para a região de ITS coexistem dentro do genoma da maior parte dos espécimes analisados. As análises Bayesianas não sustentaram a monofilia de *C. fissilis*. As sequências nucleotídicas de *Cedrela fissilis* e *Cedrela*

odorata apresentaram alta similaridade e foram agrupadas em uma única rede de haplótipos, sendo que esse agrupamento parece ter-se dado em função da distribuição geográfica. Os resultados sugerem que as populações atuais de *C. fissilis* e *C. odorata* são descendentes de pelo menos duas linhagens ancestrais distintas. A estrutura filogeográfica mostrou uma distribuição desigual da diversidade genética, uma separação evidente e altos níveis de diferenciação genética entre as regiões do Atlântico e de Chiquitano. As populações do Atlântico apresentam níveis de diversidade genética mais elevados do que a região de Chiquitano, sugerindo que a região do Atlântico possa ser o centro de diversidade para a espécie no Brasil. Episódios de isolamento genético associados a eventos de dispersão moldaram a história evolutiva de *C. fissilis* e *C. odorata*. Eventos de expansão recente entre as Regiões do Atlântico e de Chiquitano reconectaram essas regiões e têm favorecido o estabelecimento de zonas de contato secundário entre as populações nas florestas do interior do Paraná-Paraíba (ecorregião NT0150), nas florestas secas de Chiquitano (ecorregião NT0212) e nas florestas secas do Atlântico (ecorregião NT0202).

ABSTRACT

GARCIA, Magali Gonçalves, D.Sc., Universidade Federal de Viçosa, September 2010. **Molecular Phylogeography of *Cedrela fissilis* Vell. and *C. odorata* L. (Meliaceae) in Brazil and Bolivia.** Adviser: Luiz Orlando de Oliveira. Co-Advisers: João Augusto Alves Meira Neto and Jorge Abdala Dergam dos Santos.

Although it holds a wide biodiversity, neotropical flora has been poorly investigated. The loss of genetic patrimony is even more serious by the lack of coverage that we have this patrimony and its current value or future. The rules that govern the distribution patterns of species on land are not yet fully understood. As for the rules that govern the geographical distribution of species, also unaware of the rules governing the distribution of diversity within species. Phylogenetic and phylogeographic approaches based on molecular data are powerful tools to refine biogeographical hypotheses designed to explain contemporary distributions of populations. In this study, we used tools of molecular phylogeography to explore the evolutionary history of seasonal forests through a perspective from the *Cedrela fissilis* and closely related species, using data analysis of DNA sequences of chloroplastic and nuclear genomes. The results showed that the expected cohesion between the tandem repeat of rDNA was interrupted and concerted evolution remains incomplete, resulting in high intraindividual polymorphism observed in *C. fissilis*. Thus, multiple and functional paralog to coexist within the ITS region of the genome of most specimens. The Bayesian analysis is not supported the monophyly of *C. fissilis*. The nucleotide sequences of *Cedrela odorata* and *Cedrela fissilis* showed high similarity and were grouped into a single haplotype network, and this group seems to have been given in terms of geographical distribution. The results

suggest that current populations of *C. fissilis* and *C. odorata* are descended from at least two distinct ancestral lineages. The phylogeographic structure showed an uneven distribution of genetic diversity, a clear separation and high levels of genetic differentiation among Atlantic and Chiquitano ranges. The populations of the Atlantic have levels of genetic diversity higher than the region of Chiquitano, suggesting that the Atlantic region may be the center of diversity for the species in Brazil. Episodes of genetic isolation associated with dispersal events have shaped the evolutionary history of *C. fissilis* and *C. odorata*. Events of recent expansion between the Atlantic and Chiquitano ranges reconnected these regions and has favored the establishment of zones of secondary contact between populations of Paraná-Paraíba interior forests (ecoregion NT0150) Chiquitano dry forests (ecoregion NT0212) and the Atlantic dry forests (ecoregion NT0202).

INTRODUÇÃO GERAL

Cedrela fissilis Vell. e *C. odorata* L. popularmente conhecidas como cedro-branco e cedro-rosa, respectivamente, são espécies de hábito arbóreo pertencentes à Família Meliaceae (Smith, 1960). A área de ocorrência natural de *Cedrela fissilis* e *C. odorata* está restrita às Américas Central e do Sul (Pennington et al., 1981). Segundo a lista de espécies da flora do Brasil de 2010, *C. fissilis* encontra-se distribuída nas regiões Norte (Pará, Amazonas, Acre), Nordeste (Maranhão, Bahia), Centro-Oeste (Mato Grosso, Goiás, Distrito Federal), Sudeste (Minas Gerais, São Paulo, Rio de Janeiro) e Sul (Paraná, Santa Catarina, Rio Grande do Sul), enquanto *C. odorata* é encontrada nas regiões Norte (Amapá, Pará, Amazonas, Acre), Nordeste (Ceará, Bahia), Centro-Oeste (Mato Grosso, Goiás, Distrito Federal) e Sudeste (Espírito Santo, São Paulo, Rio de Janeiro) (Sakuragui et al., 2010). As espécies apresentam ampla distribuição, ocorrendo tanto em áreas de floresta ombrófila quanto em áreas mais secas (Muellner et al., 2009). *Cedrela fissilis* e *C. odorata* são preferencialmente encontradas em áreas de floresta estacional decídua e semidecídua, embora também ocorram de maneira mais restrita em regiões de ecótono entre Cerrado/floresta, em florestas de galeria e floresta ombrófila mista (Carvalho, 1994; Carvalho, 2003).

As características ecológicas apresentadas pelas espécies tais como deciduidade e alternância na condição de heliófita ou esciófita, são fatores que colaboram para a elevada frequência de *Cedrela fissilis* e *C. odorata* nas florestas estacionais (Lorenzi, 2002). Considerando o grupo sucessional, cedro é uma espécie que se comporta como secundária inicial ou secundária tardia. Ocorre tanto na floresta primária, principalmente nas bordas da mata ou clareiras, como na

floresta secundária, porém nunca em formações puras, possivelmente pelos ataques severos da broca-do-cedro e pela necessidade de luz para desenvolver-se, dependendo, portanto, da formação de clareiras (Carvalho, 2003).

A importância econômica das espécies vem do alto valor agregado a sua madeira. Características como a leveza e trabalhabilidade de sua madeira possibilitam seu uso diversificado na confecção de móveis e na construção civil (Smith, 1960; Carvalho, 2003). Além disso, apresenta resistência moderada a fungos e ao apodrecimento (Oliveira et al., 2005). Em função de sua importância econômica *C. fissilis* e *C. odorata* tem sido submetidas à exploração sistemática e sem nenhuma preocupação conservacionista. Adicionalmente, a descaracterização e fragmentação a que tem sido submetida grande parte da vegetação das regiões tropicais tem contribuído para a redução de suas populações naturais (Valera, 1997). Atualmente *C. fissilis* e *C. odorata* integram a lista de espécies ameaçadas (IUCN – Red List of Threatened Species) sendo caracterizadas como “ameaçada” e “vulnerável”, respectivamente (IUCN, 2008). Diante disso, estratégias para a conservação, melhoria e gestão sustentável destes importantes recursos genéticos devem ser estabelecidas, porém, para que medidas eficazes para a conservação possam ser propostas faz-se necessário que as espécies sejam entidades morfológicas e genéticas bem definidas (Valera, 1997).

A diversidade genética é um fator importante para a sobrevivência das populações em ambientes variáveis, e é reconhecida como um componente fundamental da biodiversidade (Mace et al., 1996). O entendimento do processo de divergência ao nível de população é fundamental para o estudo da diversificação evolutiva e do reconhecimento de diferentes ações de conservação. Processos ecológicos e históricos, como deriva genética, fluxo gênico e migração, moldaram a atual distribuição da diversidade genética de uma espécie a compreensão destes

processos é essencial para a elaboração e implementação de programas de conservação e manejo.

Neste contexto, a filogeografia surge como campo de estudo que trata dos princípios e processos que governam a distribuição geográfica de linhagens genealógicas, especialmente no nível intraespecífico e entre espécies altamente relacionadas (Avice, 2000). Como delineamento teórico, a filogeografia permite caracterizar processos de diversificação ao nível de espécie e a formulação de hipóteses de biogeografia histórica. A filogeografia molecular tem sido empregada com sucesso para explorar associação entre redes de haplótipos e geografia em um grande número de espécies vegetais, animais e de microorganismos (Templeton, 2004).

Este estudo se propôs a investigar os padrões filogenéticos e filogeográficos de populações de *Cedrela fissilis* e *C. odorata* com intuito de desenvolver e avaliar hipóteses que contribuam para o entendimento da história evolutiva da Flora Neotropical, especialmente das florestas estacionais.

REFERÊNCIAS BIBLIOGRÁFICAS

- Avice, J.C. 2000. Phylogeography: the history and formation of species. Harvard Univ. Press, Cambridge. 447p.
- Carvalho, P.E.R. 1994. Espécies Florestais Brasileiras. Recomendações silviculturais, potencialidades e uso da madeira. EMBRAPA-CNPQ, Brasília. 640p.
- Carvalho, P.E.R. 2003. Espécies Arbóreas Brasileiras, Vol. 1, Ed: Embrapa Florestas. 1040 p.
- IUCN. 2008. 2008 IUCN Red list of threatened species. Disponível em: <http://www.redlist.org/>.
- Mace, G.M., Smith, T.B., Bruford, M.W., Wayne, R.K. 1996. An overview of the issues. In: Molecular genetic approaches in conservation. Eds. Smith, TB; Wayne, RK. Oxford University Press, New York, 3-12.
- Pennington, T.D., Styles, B.T., Taylor, D.A.H., 1981. Meliaceae. Flora Neotropica, Monograph 28, The New York Botanical Garden, Bronx, New York, USA. 470 p.
- Smith, C. E. JR. 1960. A revision of *Cedrela* (Meliaceae). Fieldiana. 29, 295-341.
- Lorenzi, H. 2002. Árvores Brasileiras - manual de identificação e cultivo de plantas arbóreas nativas do Brasil, Vol. 01, 4ª edição. Nova Odessa, SP: Instituto Plantarum. 384 p.
- Muellner, A.N., Pennington, T.D., Chase, M.W. 2009. Molecular phylogenetics of Neotropical Cedreleae (mahogany family, Meliaceae) based on nuclear and plastid DNA sequences reveal multiple origins of "*Cedrela odorata*". Molecular Phylogenetics and Evolution 52, 461–469.
- Oliveira, J.T.S., Souza, L.C., Della Lucia, R.M., Souza Júnior, W.P. 2005. Influência dos extrativos na resistência ao apodrecimento de seis espécies de madeira. Revista Árvore. 29, 819-826.
- Sakuragui, C.M., Stefano, M.V., Calazans, L.S.B. 2010. Meliaceae in Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. (<http://floradobrasil.jbrj.gov.br/2010/FB009990>).

- Templeton, A.R. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*. 13, 789-809.
- Valera, F.P. 1997. Genetic Resources of *Swietenia* and *Cedrela* in the Neotropics: Proposals for Coordinated Action. Forest Resources Division, Forestry Department, Food and Agriculture Organisation of the United Nations, Rome.

CAPÍTULO I

MOLECULAR EVIDENCES OF HISTORICAL RANGE FRAGMENTATION, RANGE EXPANSION AND INTRASPECIFIC HYBRIDIZATION IN THE NEOTROPICAL SEASONAL FOREST TREE *Cedrela fissilis* (MELIACEAE)

ABSTRACT

Plant molecular phylogeography can lead to a better understanding of the interaction between past climate events, large-scale vegetation shifts and evolutionary history of Neotropical seasonal forests. An endangered, timber tree species, *Cedrela fissilis* is associated with seasonal forests and occurs throughout Central and South America. We sampled *C. fissilis* from 56 sites across the species' range in Brazil and Bolivia and obtained sequence data from the nuclear and chloroplast genomes. Most specimens (120 out of 169) exhibited intraindividual polymorphism for the nuclear internal transcribed spacer (ITS); cloning and an array of complement sequence analyses indicated that the multiple copies of ITS are functional paralogs. – concerted evolution in *C. fissilis* appears to be incomplete. Independent Bayesian analyses using either ITS or cpDNA data revealed two separate phylogenetic lineages within *C. fissilis* that correspond to populations located in separate geographic regions. We argued that past climate change triggered vicariance events that split ancestral populations into at least two refugial areas of seasonal forest that are located to the east and west of the present day Cerrado (savanna-like vegetation). With recent climate amelioration, formerly isolated lineages reconnected and intraspecific hybridization gave rise to intraindividual polymorphism and incomplete concerted evolution in *C. fissilis*.

Keywords: *Cedrela*; Concerted evolution; Intraindividual polymorphism; Internal transcribed spacer (ITS); Neotropics; Phylogeography; Seasonal forests.

1. Introduction

Interest in Neotropical seasonal forests is growing because they are areas of endemism (Prado, 2000; Pennington et al., 2004) and are threatened by high rates of deforestation due to urban expansion, agricultural conversion, and forest fires (Steininger et al., 2001; Fundação SOS Mata Atlântica INPE, 2009); moreover little is known about their evolutionary history. Seasonal forests occur in areas that experience a severe dry season (up to 5 consecutive months), or in areas where precipitation is more uniform, but where winter temperatures ($< 15^{\circ}\text{C}$) cause physiological drought (Veloso et al., 1991; Oliveira-Filho and Fontes, 2000). Although disjunct in their distribution across the Neotropics (Fig. 1) distant blocks of seasonal forest are floristically related (Bigarella et al., 1975; Ab'Saber, 1977). Where seasonal forests do occur, they are often intermingled with patches of other vegetation types, including evergreen forests, savannas, and shrublands (e.g., Killeen et al., 2006; Oliveira-Filho et al., 2006).

Despite geographic proximity to other vegetation types, Neotropical seasonal forests contain many phylogenetically unrelated species that are absent from evergreen forests and savannas in neighboring areas (Prado, 2000; Pennington et al., 2006). To explain these patterns, the 'Pleistocenic Arc' hypothesis (Prado and Gibbs, 1993) suggests that present-day seasonal forests are relicts formed by vicariance of a much larger, single formation that covered South America during the Pleistocene. An alternative hypothesis suggests that rare long-distance dispersal of seasonal forest species might explain the disjunct distribution of floristically-related seasonal forests (Gentry, 1982; Mayle, 2004).

Their disjunct distribution at continental scales, and often patchy distribution at regional scales, suggests that seasonal forests are spatially and temporally

dynamic. Indeed, seasonal forests have both contracted (Carnaval and Moritz, 2008; Carnaval et al., 2009) and expanded (Auler et al., 2004; Mayle, 2004) in response to climatic changes during the Quaternary period (Ledru, 2007). These vegetation shifts were triggered by cycles of alternating cool-dry and warm-wet climates during the last 2 million years of the Pleistocene (Whitmore and Prance, 1987), and possibly earlier (Mayle, 2004; Pennington et al., 2004). The interaction between past climate events, large-scale vegetation shifts, and the evolutionary history of plant lineages characteristic of seasonal forests remains poorly understood.

A tree genus that likely evolved in seasonal forests is *Cedrela* (Meliaceae). *Cedrela* is a Neotropical, monophyletic genus whose diversification started in the Oligocene and Early Miocene and intensified in the Late Miocene and Early Pliocene (Muellner et al., 2010). At present, most of the 17 species of *Cedrela* have restricted distributions in seasonal forests of Central and South America; two species, *C. odorata* and *C. fissilis*, are widespread and occur in both moist evergreen forests and seasonal forests (Muellner et al., 2009). The IUCN Red List (IUCN, 2008) categorizes *C. odorata* as 'vulnerable' and *C. fissilis* as 'endangered' species. Recent molecular studies found three genetically distinct entities in *C. odorata* that were morphologically not separable (Muellner et al., 2009).

Cedrela fissilis Vell. (Meliaceae), commonly known in Brazil as 'cedro branco', is a valuable timber species that has long experienced overharvesting. *Cedrela fissilis* occurs extensively throughout Central and South America, from Costa Rica to northern Argentina (Pennington et al., 1981). In Brazil, *C. fissilis* is associated with seasonal forests, but also occurs in forest-savanna ecotones, gallery forests, and moist evergreen forests that are adjacent to seasonal forests (Carvalho, 1994). Large canopy trees occur at low densities (1-3 trees/ha) in old-growth forests of southern Brazil (Carvalho, 1994); in secondary forests, densities of

C. fissilis are often much higher (Smith, 1960). Important to its biogeography, *Cedrela fissilis* exhibits a number of traits that favor cross-pollination and long-distance gene flow, including: protogynous dichogamy; asynchronous floral anthesis both among inflorescences of the same tree and among individuals; pollination by moths and bees; and wind dispersed seeds that are released from high in the forest canopy (Carvalho, 1994).

Exploring the molecular phylogeography of plants species holds promise for understanding evolutionary history of seasonal forests in the Neotropics (e.g., Caetano et al., 2008; Oliveira et al., 2010). Molecular phylogeography studies in *Carapichea ipecacuanha* (Rubiaceae), an understory shrub with disjunct distribution in seasonal forests, suggested the long-term persistence of the species in the Atlantic range; whilst the colonization of the Amazonian range seem recent, most likely from a single parental source after a strong genetic bottleneck (Oliveira et al., 2010). In *C. ipecacuanha*, the internal transcribed spacer (ITS) exhibited intraindividual polymorphism in a range-dependent manner: intraspecific hybridization in the Atlantic range lead to specimens with multiple, functional copies for ITS and pseudogenes; the Amazonian range, with a single lineage, lacked polymorphisms for ITS (Queiroz et al., submitted). *Cedrela fissilis* and *C. ipecacuanha* are co-distributed throughout much of the range of *C. ipecacuanha* and it is plausible that the two species, although differing in several life history traits and ecological requirements, could have experienced a common history of climatic changes. We anticipated molecular data from *C. fissilis* to be even more informative because of the widespread distribution of the species relative to *C. ipecacuanha*; *C. fissilis* occurs in many areas (Carvalho, 1994) where *C. ipecacuanha* is absent (Oliveira et al., 2010).

In this study we explored the evolutionary history of seasonal forests through molecular analysis of *C. fissilis* nuclear and chloroplast genes of specimens

collected from across the species range in Brazil and Bolivia. We also carried out phylogenetic analyzes that included the other 16 species of *Cedrela*. The questions addressed by this research are: (1) Is there any evidence for range-dependent intraindividual polymorphism and pseudogenes for ITS in *C. fissilis*? (2) Is *C. fissilis* a monophyletic species relative to the other members of the genus? (3) Does *C. fissilis* consist of a single evolutionary lineage, or is there lineage differentiation to a level to indicate cryptic speciation within *C. fissilis*? (4) To what extent is the distribution of genetic diversity in *C. fissilis* associated with geography and vegetation type? (5) Is there evidence for the existence of recently colonized areas and refugia? Addressing these questions with molecular data of *Cedrela* will shed light into the evolutionary history of seasonal forests and contribute to the elaboration and testing of biogeographic hypotheses to better understand the biogeography of an a threatened ecosystem of high conservation importance.

2. Materials and methods

2.1 Sampling strategy and DNA extraction

This study analyzed 169 specimens of *Cedrela fissilis*, collected from 56 sites in Brazil and Bolivia (Fig. 1). The sampling sites were chosen as to cover most of the species geographical range. Most sites were located within seasonal forests (39 of 56 sites) with others located in adjacent vegetation (four in savannas; 13 in moist evergreen forest). To define the vegetation formations in which the species occurred, we consulted the Terrestrial Ecoregions of the World from the World Wildlife Fund (Olson et al., 2001). The habitat classifications for Brazil from the Brazilian Institute of Geography and Statistics (IBGE, 2004) provided details about the vegetation at a local scale.

Most sampling sites were located in seasonal forests on either the east or west of a savanna biome of South America – the open canopy vegetation known as Cerrado (Ecoregion NT0704; Fig. 1). Additional samplings sites located within the Cerrado or in transitional zones provided supplementary specimens. At the western side of the Cerrado, most of sampling sites occurred in two ecoregions: Madeira-Tapajós moist forests (NT0135) and Chiquitano dry forests (NT0212); we now on will refer to those sites collectively as the Chiquitano range. At the eastern side of the Cerrado, our sampling sites were located in five ecoregions: Araucaria moist forests (NT0101), Bahia coastal forests (NT0103), Bahia interior forests (NT0104), Paraná-Paraíba interior forests (NT0150), and Atlantic dry forests (NT0202). We will refer to those sites collectively as the Atlantic range. The Atlantic range had an outlier, located within a block of seasonal forests of northeastern Brazil (Pernambuco coastal forests - Ecoregion NT0151), in an area dominated by scrubland vegetation (Caatinga - Ecoregion NT1304).

Sample sizes per population varied from 1 to 8, depending upon the number of specimens available for sampling within a given population (Appendix A, online supplementary material). Leaf samples were transported to the laboratory while still fresh and then kept at -80 °C; alternatively, they were dried immediately using silica gel and kept at room temperature until the DNA was extracted. Total genomic DNA was extracted following the procedure described in Rossi et al. (2009). Genomic DNA from all specimens was archived in our laboratory, at the Federal University of Viçosa, Brazil.

2.2 Assembly of ITS datasets

The entire internal transcribed spacer (ITS) region of the nuclear 18S-26S ribosomal RNA genes (which included the 5.8S gene) was amplified according to standard PCR protocols (Oliveira et al., 2010; Muellner et al., 2009), using primer

ITS4 described by White et al. (1990) and primer ITS.LEU described by Baum et al. (1998). PCR products were cleaned using ExoSAP IT (USB), at the rate of 9µl reaction to 3µl enzyme. Sequencing was performed by Macrogen Inc., South Korea (www.macrogen.com) using the same primers as in the PCR reactions. Sequences were imported into the program Sequencher 4.8 (Gene Codes) for editing. Complete sequence alignments were performed with the introduction of gaps to compensate for the presence of insertion/deletions (indels).

We selected six specimens whose sequences were associated with electrophoretograms showing multiple overlapping double peaks and cloned the PCR products into the vector pGEM T-Easy (Promega). For each specimen, we resubmitted six clones for sequencing and received usable data from 34 clones. Sequences from the remaining 120 specimens containing multiple overlapping double peaks in both strands were not subjected to cloning and, therefore, were excluded from subsequent DNA analyses.

We set the limits of the two internal transcribed spacer regions (ITS1 and ITS2) and the 5.8S gene in *C. fissilis* informed by the following three databases: Rfam (Gardner et al., 2008), The European Ribosomal Database (Wuyts et al., 2004), and The ITS2 database (Selig et al., 2008).

Subsequently, the flanking regions of the 18S and 28S DNAs were trimmed off to produce an ITS dataset with size ranging from 591 to 639 bases. Size variation was due to the presence of three indels within the ITS1 region: a 36bp indel (in two clones of specimen PLA 75), a 48bp indel (in specimen VNI 38), and a 51pb indel (in specimen AJU 28). The aligned ITS1 subset contained 251 bases, of which 27 were polymorphic (24 substitutions and three indels); the aligned 5.8S subset spanned 156 bases and lacked polymorphic sites; the ITS2 subset was comprised of 232 bases, of which 14 were polymorphic (14 substitutions, no indel). Henceforth, we will refer to this dataset as 'ITS dataset A'. ITS dataset A contained

77 sequences: 20 sequences obtained through direct sequencing of specimens whose electrophoretograms presented no ambiguities, 23 sequences obtained through direct sequencing from specimens whose electrophoretograms contained a single site showing double peaks (coded with IUPAC ambiguity codes: R, Y, K, S, or M), and 34 sequences obtained from cloning.

We also assembled a second ITS dataset, 'ITS dataset B', in which the single ambiguities were resolved; that is, each of the 23 sequence harboring single double peaks gave rise to a pair of new sequences, which were identical to the respective parent sequence except that each sequence of the pair displayed one of the two possible nucleotide codes instead of the formerly ambiguous code. ITS dataset B contained 100 sequences.

2.3 Analyses of secondary structures in the ITS region

We used ITS dataset B to investigate further the intraindividual polymorphism for ITS in *Cedrela fissilis* and whether it was associated to pseudogenization. We inspected the ITS1 and 5.8S gene regions of each sequence for the occurrence of unexpected substitutions within evolutionary conserved motifs; we performed secondary structure analyses in both the 5.8S gene and ITS2 regions and looked for the presence of aberrant structures that may suggest a non-functional copy.

The program RDP3 (Heath et al., 2006) scanned the ITS dataset B for recombination using seven different recombination detection methods. Next, we split the ITS dataset B into three smaller subsets, corresponding to the ITS1, 5.8S, and ITS2 regions. The DnaSP v5 software (Librado and Rozas, 2009) defined haplotypes within each of the three subsets. We used sequence alignments in Sequencher to inspect both the 5.8S gene and the ITS1 region for the presence of unexpected nucleotide substitutions within known conserved angiosperm motifs. In

the 5.8S gene, we inspect the following three conserved 5.8S motifs of angiosperms: M1 (5'-CGATGAAGAACGTAGC-3') and M3 (5'-TTTGAAYGCA-3') of Harpke and Peterson (2008), and M2 (5'-GAATTGCAGAATCC-3') of Jobes and Thien (1997). In the ITS1 region, we inspect the conserved ITS1 angiosperm motif 5'-GGCRY(4–7n)GYGYCAAGGAA-3' of Liu and Schardl (1994).

Both the secondary structure and minimum free energy (ΔG) of haplotypes within the 5.8S and ITS2 subsets were estimated with the online version of the program MFOLD 3.2 – RNAfolding (Mathews et al., 1999; Zuker, 2003), with the standard options of temperature (37°C) and ionic conditions. The secondary structures and the minimum free energies (ΔG) of the ITS2 region were predicted by homology modeling using the secondary structure of an accession of *Cedrela fissilis* (GI219816486) available in the ITS2 database (Selig et al., 2008). Subsequently, the secondary structures for the 5.8S gene and ITS2 region were obtained with the help of the program 4SALE (Seibel et al., 2008) and edited for printing with XRNA 2.0 beta (<http://rna.ucsc.edu/rnacenter/xrna/xrna.html>). MFOLD (Mathews et al., 1999; Zuker, 2003) computed the GC content of 5.8S haplotype and each of the ITS2 haplotypes. We used secondary structure stability of the 5.8S haplotypes and ITS2 haplotypes as evidence for distinguishing functional paralogs from pseudogenes. Substitutions within the conserved ITS1 motif of Liu and Schardl (1994) were regarded as a further evidence for pseudogenization.

2.4 Assembly of the cpDNA dataset

Three regions of the chloroplast genome were investigated. The *trnT–trnL* spacer was amplified using standard PCR protocols (Oliveira et al., 2010; Muellner et al., 2009). The initial amplification used primers A and D, which were described by Taberlet et al. (1991). Samples that yielded no reaction product or ambiguous results were re-examined by two independent amplification reactions. The first was

performed with the primer combination A₂ described by Cronn et al. (2002) and aC (5'-CGTAGCGTCTACCGATTTCG-3'), complementary to the primer C described by Taberlet et al. (1991). This primer pair amplified the intergenic spacer between *trnT* (UGU) and the *trnL* (UAA) 5' exon (Taberlet et al., 1991). The second was performed with primers C and D and amplified the *trnL* (UAA) intron (Taberlet et al., 1991). The *trnS–trnG* intergenic spacer was amplified using primer pair *trnS/trnG*. The *psbB*, *psbT* and *psbN* genes were amplified using primer pair *psbB/psbF*. Primer pairs *trnS/trnG* and *psbB/psbF* were described by Hamilton (1999).

PCR products were cleaned using ExoSAP IT (USB), using a ratio of 9µl reaction to 3µl enzyme. Sequencing was performed by Macrogen Inc., South Korea (www.macrogen.com) using the same primers as in the PCR reactions. Sequences were imported into Sequencher for editing. Complete sequence alignments for each of the three chloroplast regions were performed with the introduction of gaps to compensate for the presence of indels. We refrained from using indels that were bordered by a string of the same type of nucleotide (e.g., polyA) as a source of information in subsequent statistical analyses. The decision to remove those sites from subsequent analyses was intended to eliminate any effect that can arise from experimental error or evolutionary lability associated with this type of indel (see Mast et al., 2001).

Sequencing of the target *trnT–trnL* region produced two overlapping segments, which yielded a total aligned sequence of 1498 bases (GenBank accessions xxxxx to xxxxx), of which 21 were polymorphic (16 substitutions; five indels of 1-bp, 5-bp, 7-bp, and 22-bp, respectively). Sequencing and alignment of the *trnS–trnG* intergenic spacer was 716 bases long (GenBank accessions xxxxx to xxxxx), of which five were polymorphic (four substitutions; one indels of 1-bp). Sequencing and alignment of the *psbB*, *psbT* and *psbN* genes was 689 bases long (GenBank accessions xxxxx to xxxxx), of which one was a substitution. The final

cpDNA dataset resulted from the concatenation of three chloroplast regions, contained 27 sequences, and was 2903 bases long.

2.5 Phylogenetic analyses

Bayesian analyses were carried out independently for the ITS and cpDNA datasets because of the presence of more than one ITS copy per specimen; moreover, we also had an uneven number of samples in each dataset. We conducted two independent phylogenetic analyses, each based on the ITS database A and on an extended version of the ITS dataset A. The extended version contained additional sequences from other members of the genus *Cedrela*, having *Toona ciliata* as an outgroup. Those sequences were obtained from the GenBank (Appendix B, online supplementary material). Before running each analysis, identical sequences belonging to the same taxa were combined into a single terminal. However, the original specimen codes were kept for future reference and were presented in the tree. MrModeltest v2.3 (Nylander, 2004) was fed separately with the two extended versions of the ITS dataset A and the Akaike Information Criteria (Akaike, 1973) indicated GTR+I+G as the best-fit model among 24 models of molecular evolution for both analyses. Using the GTR+I+G model, MrBayes v3.1.2 (Ronquist, 2003) estimated Bayesian phylogeny. Each Bayesian analysis was performed using two simultaneous runs of 10 million generation each, with one cold and three heated chains in each run. Sampling was carried out once every 1000 trees; the first 250 trees were discarded as burn-in samples. For each analysis, 50% majority rule consensus tree of the two independent runs was obtained with posterior probabilities equal to bipartition frequencies.

We also conducted a phylogenetic analysis based on the cpDNA dataset. This time, we did not include additional sequence data from the GenBank because the sequences that were available provided data for only two regions (the *trnS–trnG*

intergenic spacer and *psbB*, *psbT* and *psbN* genes); moreover, alignment analyses had indicated that most of the parsimony informative characters in our cpDNA dataset resided at the *trnT-trnL* region. Again, identical sequences belonging to the same taxa were combined into a single terminal; but the original specimen codes were kept for future reference and were presented in the tree. MrModeltest v2.3 (Nylander, 2004) was fed with the cpDNA dataset and the Akaike Information Criteria (Akaike, 1973) indicated HKY as the best-fit model among 24 models of molecular evolution. Using the HKY model, MrBayes v3.1.2 (Ronquist, 2003) estimated an unrooted Bayesian phylogeny. Bayesian analysis was performed using two simultaneous runs of 2 million generations each, with one cold and three heated chains in each run. Sampling was done once every 1000 trees; the first 250 trees were discarded as burn-in samples. A 50% majority rule consensus tree of the two independent runs was obtained with posterior probabilities equal to bipartition frequencies.

3. Results

3.1 Intraindividual polymorphism for ITS

DNA alignments revealed that the ITS region contained 41 polymorphic sites; all of them were base substitutions and displayed overlapping double peaks. The 41 ambiguous sites were spread along the ITS1 and ITS2 regions; no variability was observed within the 5.8S gene region. Across sequence alignments, each of the 41 polymorphic sites contained at least one ambiguity in at least one specimen. Therefore, a decision to eliminate ambiguous sites from subsequent analyses would reduce the variability within the dataset to zero. On both sides, stretches of clean sequences flanked each site that harbored double peaks (Appendix C, online

supplementary material). The pattern of double peaks in the electrophoretograms suggested that intraindividual polymorphism was present in most of the specimens of *Cedrela fissilis* we had sampled.

3.2 Structure analyses in the ITS region

In *Cedrela fissilis*, the conserved ITS1 angiosperm motif 5'-GGCRY(4–7n)GYGYCAAGGAA-3' of Liu and Schardl (1994) appeared as 5'-GGCGC(GAGCY)GCGCCAAGGAA-3' and contained no further substitutions. The three conserved 5.8S motifs of angiosperm were free of substitutions, and the motif M3 (5'-TTTGAAYGCA-3') of Harpke and Peterson (2008) appeared as 5'-TTTGAACGCA-3'.

No nucleotide substitutions or indels were observed within the 5.8S gene. This single 5.8S haplotype exhibited the predicted secondary structure for a functional 5.8S gene, with the presence of the five conserved helices, CG content of 55.76%, and minimum free energy (ΔG) of -10.20 kcal/mol (Fig. 2). Further alignments with sequences obtained in the GenBank indicated the 5.8S gene of other 17 members of the genus exhibited a sequence identical to the 5.8S haplotype of *Cedrela fissilis*.

Structure analyses of the ITS2 region recovered seven secondary structures that displayed highly similar topologies (Fig. 3). Secondary structures exhibited four helices each, which we numbered I to IV following Coleman (2007). Nucleotide substitutions took place in all four helices, but they were more commonly found in Helix I and Helix III. Helix II exhibited the most conserved secondary structure and harbored a characteristic pyrimidine-pyrimidine bulge (C-U) at the base. Minimum free energy required for the secondary structures varied from -90.0 to -97.4 (mean = -93.99 kcal/mol); the CG content of those sequences ranged from 72.4 to 74.1% (mean= 73.3%).

Two of the seven secondary structures (A_1 and B_1) were recovered with high frequencies, with the remaining five being of rare occurrence (Table 1). The extension of the loop at the end of Helix II distinguished A_1 from B_1 . There was a clear correspondence between geographic range where the specimen was sampled and frequency of a given secondary structure of ITS2: A_1 was the secondary structure prevalent among sequences from the Atlantic range, while B_1 was more frequent among sequences sampled in the Chiquitano range. Only five sequences from the Chiquitano range displayed the most frequent secondary structure of ITS2 among Atlantic sequences (A_1) and they were recovered from clones (SRA113_1, SRA113_3, SRA113_6, PLA75_2, and PLA75_6). Likewise, the six sequences from the Atlantic range that displayed the most frequent secondary structure of ITS2 among Chiquitano sequences (B_1) were also recovered from clones (DOU117_1, DOU117_2, ITA124_1, ITA124_3, ITA124_4, and ITA124_6).

3.3 Bayesian analyses based on ITS

The Bayesian consensus tree for the full extended ITS dataset (Fig. 4) showed the placements of the 77 sequences of *Cedrela fissilis* in relationship to other 26 sequences of *Cedrela*. *Toona ciliata* (FJ462488) was used as outgroup. With strong support from posterior probabilities (PP), the tree displayed a topology that suggested an association between member composition and geography. Overall, the most basal placements contained sequences from Mexican and Central-American species, intermediary positions were taken by sequences from South-American species other than *C. fissilis*; and terminal clades contained sequences from *C. fissilis*. In a strongly supported clade (PP=98%), all 77 sequences of ITS dataset A came together with one sequence of *C. odorata* and two sequences of *C. balansae*. We will henceforth refer to this clade as the *C. fissilis* clade.

To further explore the geographic distributions and phylogenetic relationship of members of the *C. fissilis* clade, we presented an unrooted Bayesian consensus tree using sequences found within that clade in conjunction with a map of Brazil and Bolivia (Fig. 5). The tree recovered two major clades (PP=97%). The West clade (Fig. 5, in green) consisted mostly of sequences from specimens collected in the Chiquitano range: nine sequences from Madeira-Tapajós moist forests (NT0135) and 20 sequences from Chiquitano dry forests (NT0212). Additionally within the West clade, there were two cloned sequences (ITA124_3 and ITA124_5) from a specimen found at the westernmost edge of an ecoregion of Atlantic origin (Paraná-Paraíba interior forests - NT0150), and two sequences from specimens of *C. fissilis* from Paraguay (FJ518893) and Peru (FJ462475). The East clade (Fig. 5, in blue) contained the remaining 46 sequences of *C. fissilis* in addition to two sequences from *C. balansae* from Argentina (FJ462473) and Paraguay (FJ462474) and a sequence of *Cedrela odorata* from Brazil (FJ462471). Sources for most of the sequences in the East clade were specimens collected in ecoregions located along the Atlantic coast. Nevertheless, four sequences obtained through cloning (PLA75_2, PLA75_6, SRA113_1, and SRA113_3) were from two specimens of the Chiquitano range.

3.4 Bayesian analysis based on concatenated cpDNA

The unrooted Bayesian consensus tree for cpDNA and associated posterior probabilities (PP) is presented along with a map of Brazil and Bolivia showing the geographic distribution of 27 specimens of *Cedrela fissilis* (Fig. 6). With high support (PP=100%), the tree recovered two major clades. The West and the East clades brought together sequences from the Chiquitano range (Fig 6, in green) and Atlantic range (Fig. 6, in blue) into almost exclusive groups. The few exceptions were: (i) the placement of specimen SRA113 into the East clade and (ii) the placements of

specimens PAL164 and MAN284 into the West clade. The specimen SRA113 was collected within the core of Chiquitano dry forests (Ecoregion NT0212) in Eastern Bolivia; it showed intraindividual polymorphism for ITS and exhibited paralogs from either Chiquitano or Atlantic origin (Fig. 5). PAL164 was a specimen collected in an isolated block of seasonally dry forests located in Central Brazil that is surrounded by savanna vegetation (Cerrado - NT0704). This disjunct block is part of the Paraná-Paraíba interior forests (NT0150), an ecoregion of Atlantic origin. Meanwhile, MAN284 was a specimen collected in Northern Minas Gerais, in a transition zone involving an enclave of seasonally dry forests (Atlantic dry forests - NT0202), savanna vegetation (Cerrado - NT0704) and scrubland (Caatinga - NT1304).

4. Discussion

4.1 Multiple, functional copies for ITS

In our analyses, *Cedrela fissilis* did not display the expected within-species homogeneity for ITS copies. The nuclear ribosomal DNA (rDNA) consists of tandem units of the genes 18S, 5.8S, and 28S, each of which encodes a ribosomal RNA molecule. Two internal transcribed spacer regions (ITS1 and ITS2) flank each side of the 5.8S gene separating the three ribosomal genes (i.e., 18S - ITS1 - 5.8S - ITS2 - 28S). In plant genomes, it has been proposed that these hundreds to thousands of tandem-repeated copies of rDNA (Rogers and Bendich, 1987) are homogenized to a single sequence through unequal crossing over and high frequency gene conversion, a phenomenon called concerted evolution (Elder and Turner, 1995). When carried to completion, concerted evolution removes paralogous sequences from the genome, thereby eliminating intraindividual polymorphisms; rDNA tandem repeats are presumably homogeneous within species and different between species

(Dover, 1986). In 149 out of 169 specimens of *C. fissilis* displayed ITS sequences with at least one ambiguity; that is, at polymorphic sites, the electrophoretograms displayed overlapping double peaks in both strands; 20 specimens were free of ambiguities. It appears that multiple, paralogous copies for ITS coexist within the genome of most specimens of *C. fissilis* from Brazil and Bolivia.

Intraindividual polymorphism for ITS may be associated with non-functional rDNA copies, or pseudogenes, in addition to functional copies (Bailey et al., 2003; Oliveira et al., 2010). Features such as low CG content, high rates of substitutions and deletions within highly conserved motifs, and aberrant secondary structures are suggestive of paralogs that escaped from functional constraints and became pseudogenes (Bailey et al., 2003; Rosselló et al., 2007). In *C. fissilis*, ITS paralogs exhibited none of those features; we could not identify any pseudogene among the sequences we analyzed. The hypothesis that multiple, functional copies for ITS exist within the genome of some specimens of *C. fissilis* could not be rejected.

4.2 Incomplete concerted evolution in *C. fissilis*

Under certain circumstances, the cohesiveness of the tandem repeats of rDNA is disturbed, concerted evolution remains incomplete, and intraindividual polymorphism such as that we observed in *C. fissilis* emerges and persists through time. In plants, hybridization, polyploidy, and long generation time are factors that contribute to incomplete concerted evolution and likely contributed to intraindividual polymorphism in *C. fissilis*. Both interspecific and intraspecific hybridizations bring together different alleles from the parental sources into a single genome and therefore can give rise to intraindividual polymorphism (Koch et al., 2003; Muir et al., 2001). Long generation time, a common feature of tropical tree species such as *C. fissilis*, is believed to reduce the rates of concerted evolution (Sang et al., 1995). Polyploidization could favor the maintenance of intraindividual polymorphism

because tandem-repeated copies of rDNA in different chromosomes are less vulnerable to the homogenizing effects of concerted evolution (Campbell et al., 1997). Because the number and chromosomal localities of rDNA in *C. fissilis* is unknown; the extent to which polyploidization contributed to the intraindividual polymorphism we uncovered remains speculative. Nonetheless, the high and variable range of chromosome numbers within *Cedrela* (Styles and Vosa, 1971) suggests ancient polyploidy likely contributed to incomplete concerted evolution of *C. fissilis*.

4.3 *C. fissilis* is not monophyletic

Monophyly of *C. fissilis* relative to other species of *Cedrela* was not supported in our Bayesian analysis, because sequences from two species (*Cedrela odorata* and *C. balansae*) came together with sequences of *C. fissilis* and formed a well supported clade (Fig. 4). It is interesting to note that the non-*fissilis* components clustered at the East clade only, and clustering seems to have a geographic component (Fig. 5). The sequence of *Cedrela odorata* (FJ462471, from Brazil) was identical to the sequence of the *C. fissilis* Northeastern Brazil (GUA204); two sequences of *C. balansae* from (Argentina and Paraguay (FJ462473 and FJ462474, respectively) also clustered together with *C. fissilis* sequences of the East clade. Larger sampling within *C. odorata* and *C. balansae* for DNA analysis and morphometric studies is required to determine relationships and delimitations among these three species.

4.4 Association of phylogenetic clades with geographic ranges

For *C. fissilis*, Bayesian analyses using both ITS (Fig. 5) and cpDNA (Fig. 6) recovered trees of highly similar topologies; each tree had two major, well-supported clades suggesting a major population sub-division. With few exceptions (see section

4.5), placement of sequences occurred into almost exclusive clades, which corresponded to geography: the West clade contained sequences from the Chiquitano range, while the East clade brought together sequences from Atlantic range.

The low levels of structuring within the West clade suggested that the Chiquitano range harbored a genetically depauperate lineage. Moreover, the star-like topology we recovered for the West clade can be taken as an indicator of a large population expansion in the past (Marjoram and Donnelly, 1994). The hypothesis that the West clade indeed experienced a recent, rapid expansion is plausible given palinological evidence that the region containing the Chiquitano range endured drastic historic vegetation shifts (Burbridge et al., 2004; Anhufo et al., 2006) and that closed canopy evergreen forest arrived in that region within the last few thousand years (Gosling et al., 2009). We recognized that our sampling within the Chiquitano range was geographically restricted; additional sampling in northern Brazil, Peru and Paraguay is necessary to determine if the low levels of genetic diversity we found holds true for a larger geographic area.

Compared to the West clade, the East clade of *C. fissilis* contained several levels of phylogenetic structures, and thus, harbored higher levels of genetic diversity. The long-term existence of *C. fissilis* in the Atlantic range is plausible given evidence that climate and vegetation stability in southeastern Brazil provided refugial areas during climate cycles that more dramatically affected other regions (Carnaval and Moritz, 2008; Carnaval et al., 2009). The existence of sub-structuring within the East clade was an additional indication that closely related lineages of *C. fissilis* evolved within the Atlantic range. Most likely each of those lineages experienced a certain degree of genetic isolation within distinct refugial areas. To fully differentiate the closely related Atlantic lineages, we would need a much larger dataset.

4.5 Genetic connectivity between formerly isolated lineages

During past climate fluctuations Central Brazil was drier than present-day Cerrado (Whitmore and Prance, 1987; Mayle, 2004; Pennington et al., 2004), and likely inhospitable to *C. fissilis*. Our data are consistent with a scenario of climate-mediated vicariance that split once interconnected *C. fissilis* populations into refugial areas where distinct lineages evolved during periods of genetic isolation. More recently (within the last few thousand years) a warmer wetter trend permitted novel vegetation to assemble in Central Brazil, when forests spread at the expense of savannas which retracted (Mayle, 2004). Currently, seasonal forests permeated the Cerrado in the form of gallery forests and provide otherwise isolated populations of plant (Oliveira-Filho and Ratter, 1995) and animal (Costa, 2003) species with genetic connectivity.

Data from this study provides evidence of relatively recent dispersal events of *C. fissilis* from adjacent ecoregions into the Cerrado. Specimens collected in the Cerrado ecoregion did not form an exclusive group but instead, clustered with specimens from nearby seasonal forests. Despite this general trend exception to this pattern did occur; notable clustering of some Cerrado specimens illustrated the capability of *C. fissilis* to expand deeper through regions of Cerrado vegetation as evidenced by: (i) specimens from distinct edges of the Atlantic range displayed both ITS paralogs and cpDNA that belonged to the Chiquitano lineage; (ii) specimens from within the core of the Chiquitano range exhibited both ITS paralogs and cpDNA that belonged to Atlantic lineages. An easternward expansion brought the Chiquitano lineage towards Central Brazil (disjunct blocks of Paraná-Paraíba interior forests - Ecoregion NT0150) and reached the Atlantic region (Atlantic dry forests - Ecoregion NT0202). Dispersal into both directions allowed both the lineage to reach Ecoregion NT0150 (Paraná-Paraíba interior forests) in southern Brazil and the Atlantic lineage to reach Bolivia and Mato Grosso. We speculate that bidirectional

migrations were possible owing to the Paraná River basin and Chaco, which at the present straddle the Brazilian-Paraguayan-Argentinean border. Because of long distance dispersal routes, the Chiquitano and Atlantic ranges are not completely genetically isolated ranges, even though they are separated by 500-1500 km expanse of Cerrado vegetation.

The pattern of intraindividual polymorphism and incomplete concerted evolution we uncovered for ITS suggest that intraspecific hybridization occurred when allopatric lineages of *C. fissilis* reconnected and established contact zones. Support for a widespread occurrence of intraspecific hybridization within and among ranges came from the fact that the genomes of most specimens combined multiple, functional paralogs for ITS from distinct lineages. We attribute the lack of intraindividual polymorphism for ITS in some specimens as a relict from ancestral populations of seasonal forest refugia that have escaped intra-specific hybridization over time.

The phylogeography we described for *C. fissilis* is highly congruent with geographic patterns of genetic diversity for the understory shrub *Carapichea ipecacuanha* (Oliveira et al., 2010). Expanded molecular studies including even more species with disjunct ranges should help determine if these patterns can be generalized to the evolutionary history of Neotropical seasonal forests by a framework of climate cycles and vegetation shifts that result in periods of geographic separation and genetic isolation that alternate with periods of range expansion and intraspecific hybridization.

Acknowledgments

We would like to express our sincere gratitude to the following people for providing assistance during field trips and sampling: Anderson Luis Moy (ECOPEF),

Alan Sciamarelli (UFGD), Andressa Vasconcelos Flores, Braz Cosenza (UEMG), Cristina Soares de Souza, Ernane R. Martins (UFMG), Gisele Benites Flor (FECLPP), Lourivaldo Amancio de Castro (UNEMAT), James Stahl, Luciene Vieira de Arruda (UEPB), Mauro Parolin (FECILCAM), Osvaldo S. Rossi, Paulo Fermiano da Silva (SEDAN-RO) and Solon Longhi (UFSM). IEF-MG, IAP-PR, and FATMA-SC provided collecting permits. This study was part of MGG Ph.D. Thesis with the support of a CNPq fellowship. LOO received a fellowship from CNPq (PQ 305710/2009-5) and grants from FAPEMIG (PPM-00148-09) and CNPq (471915/2009-2).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi: xxxxxxxx.

References

- Ab'Saber, A.N., 1977. Os domínios morfoclimáticos da América do Sul. Primeira Aproximação. *Geomorfologia*, 53, 1-23.
- Akaike, H., 1973. Information theory and an extension of maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest, pp. 267–281.
- Anhuf, D., Ledru, M-P., Behling, H., Da Cruz, Jr. F.W., Cordeiro, R.C., Van der Hammen, T., Karmann, I., Marengo, J.A., De Oliveira, P.E., Pessenda, L., Siffedine, A., Albuquerque, A.L., Da Silva, Dias, P.L., 2006. Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 239, 510-527.

- Auler, A.S., Wang, X., Edwards., R.L., Cheng, H., Cristalli, P.S., Smart, P.L., Richards, D.A. 2004. Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *J. Quat. Sci.*, 19, 693-701.
- Bailey, C.D., Carr, T.G., Harris, S.A., Hughes, C.E., 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Mol. Phylogenet. Evol.*, 29, 435-455.
- Baum, D.A., Small, R.L., Wendel, J.F., 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Syst. Biol.* 47, 181–207.
- Bigarella, J.J., Andrade-Limam D, Riehs, P.J., 1975. Considerações a respeito das mudanças paleoambientais na distribuição de algumas espécies vegetais e animais no Brasil. *An. Acad. Bras. Ciênc.*, 47, 411-464.
- Burbridge, R.E., Mayle, F.E., Killeen, T.J., 2004. Fifty-thousand-year vegetation and climate history of Noel Kempff Mercado National Park, Bolivian Amazon. *Quaternary Res*, 61, 215-230.
- Caetano, S., Prado, D., Pennington, R.T., Beck, S., Oliveira-Filho, A., Spichiger, R., 2008. The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Mol. Ecol.* 17, 3147-3159.
- Campbell, C.S., Wojciechowski, M.F., Baldwin, L.A., Donoghue, M.J., 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* Agamic complex (Rosaceae). *Mol. Biol. Evol.* 14, 81-90
- Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., Moritz, C., 2009. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. *Science*, 323, 785-789.

- Carnaval, A.C., Moritz, C., 2008. Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *J. Biogeogr.* 35, 1187-1201.
- Carvalho, P.E.R., 1994. *Espécies Florestais Brasileiras. Recomendações Silviculturais, Potencialidades e Uso da Madeira.* EMBRAPA-CNPq, Brasília.
- Coleman, A.W., 2007. Pan-eukaryote ITS2 homologues revealed by RNA secondary structure. *Nucleic Acids Res.* 35, 3322-3329.
- Costa, L.P., 2003. The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *J. Biogeogr.* 30, 71-86.
- Cronn, R.C., Small, R.L., Haselkorn, T., Wendel, J.F., 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *Am. J. Bot.* 89, 707-725.
- Dover, G.A., 1986. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet.* 2, 159-165.
- Elder, J.F., Turner, B.J., 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *Q. Rev. Biol.* 70, 297-320.
- Fundação SOS Mata Atlântica, INPE, 2009. *Atlas dos Remanescentes Florestais da Mata Atlântica. Período de 2005-2008. Relatório Parcial.* Fundação SOS Mata Atlântica and Instituto Nacional de Pesquisas Espaciais, São Paulo.
- Gardner, P.P., Daub, J., Tate, J.G., Nawrocki, E.P., Kolbe, D.L., Lindgreen, S., Wilkinson, A.C., Finn, R.D., Griffiths-Jones, S., Eddy, S.R., Bateman, A., 2008. Rfam: updates to the RNA families database. *Nucleic Acids Res.* 37, D136-D140.
- Gentry, A.H., 1982. Neotropical floristic diversity: phylogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Ann. Mo. Bot. Gard.*, 69, 557-593.
- Gosling, W.D., Mayle, F.E., Tate, N.J., Killeen, T.J., 2009. Differentiation between Neotropical rainforest, dry forest, and savannah ecosystems by their modern

- pollen spectra and implications for the fossil pollen record. *Rev. Palaeobot. Palyno.*, 153, 70-85.
- Hamilton, M.B., 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Mol. Ecol.* 8, 521-523.
- Harpke, D., Peterson, A., 2008. 5.8S motifs for the identification of pseudogenic ITS regions. *Botany* 86, 300-305.
- Heath, L., van der Walt, E., Varsani, A., Martin D.P., 2006. Recombination patterns in aphthoviruses mirror those found in other picornaviruses. *J. Virol.* 80, 11827-11832.
- IBGE, 2004. Mapa de vegetação do Brasil. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.
- IUCN 2008. 2008 IUCN Red list of threatened species. Available from <http://www.redlist.org/>.
- Jobes, D.V., Thien, L.B., 1997. A Conserved motif in the 5.8S ribosomal RNA (rRNA) gene is a useful diagnostic marker for plant internal transcribed spacer (ITS) sequences. *Plant Mol. Biol. Rep.* 15, 326-334.
- Killeen, T.J., Chavez, E., Pena-Claros, M. et al., 2006. The Chiquitano Dry Forest, the transition between humid and dry forest in eastern lowland Bolivia. In: Pennington, R.T., Ratter, J.A., Lewis, G.P. (Eds), *Neotropical Savannas and Seasonally Dry Forests: Diversity, Biogeography and Conservation*, CRC Press, Boca Raton, pp. 213-233.
- Koch, M.A., Dobes, C., Mitchell-Olds, T., 2003. Multiple hybrid formation in natural populations: concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). *Mol. Biol. Evol.* 20, 338-350.
- Ledru, M-P., Salatino, M.L.F., Ceccantini, G., Salatino, A., Pinheiro, F., Pintaud, J-C., 2007. Regional assessment of the impact of climatic change on the

- distribution of a tropical conifer in the lowlands of South America. *Divers. Distrib.* 13, 761-771.
- Librado, P., Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451-1452.
- Liu, J.-S., Schardl, C.L., 1994. A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. *Plant Mol. Biol.* 26, 775-778.
- Marjoram, P., and P. Donnelly, 1994 Pairwise comparisons of mitochondrial-DNA sequences in subdivided populations and implications for early human-evolution. *Genetics*, 136, 673-683.
- Mast, A.A., Kelso, S., Richards, A.J., Lang, D. J., Feller D.M.S., Conti E., 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. *Int. J. Plant Sci.* 162,1381-1400.
- Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structures. *J. Mol. Evol.* 288, 911-940.
- Mayle, F.E., 2004. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. *J. Quat. Sci.*, 19, 713-720.
- Muellner, A.N., Pennington, T.D., Chase M.W., 2009 .Molecular phylogenetics of Neotropical *Cedreleae* (mahogany family, Meliaceae) based on nuclear and plastid DNA sequences reveal multiple origins of "*Cedrela odorata*". *Mol. Phylogenet. Evol.*, 52, 461-469.
- Muellner, A.N., Pennington, T.D., Koecke, A.V., Renner, S.S., 2010. Biogeography of *Cedrela* (Meliaceae, Sapindales) in Central And South America. *Am. J. Botany*, 97, 511-518.

- Muir, G., Fleming, C.C., Schlötterer, C., 2001. Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.. Mol. Biol. Evol. 18, 112–119.
- Nylander, J.A.A., 2004. MrModeltest v2. Computer program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Oliveira, L.O., Rossi, A.A.B, Martins, E.R., Batista, F.R.C., Silva, R.S., 2010. Molecular phylogeography of *Carapichea ipecacuanha*, an amphitropical shrub that occurs in the understory of both semideciduous and evergreen forests. Mol. Ecol. 19, 1410-1422.
- Oliveira-Filho, A.T., Jarenkow, J.A., Rodal M.J.N., 2006. Floristic relationships of seasonally dry forests of eastern South America based on tree species distribution patterns. In: Pennington, R.T., Ratter, J.A., Lewis, G.P. (Eds), Neotropical Savannas and Seasonally Dry Forests: Diversity, Biogeography and Conservation, CRC Press, Boca Raton, pp. 159-192.
- Oliveira-Filho, A.T., Fontes, M.A.L., 2000. Patterns of floristic differentiation among atlantic forests in southeastern Brazil and the influence of climate. Biotropica, 32, 793-810.
- Oliveira-Filho, A.T., Ratter, J.A., 1995. A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. Edinb. J. Bot. 52, 141-194.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Louks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on earth. BioScience, 51, 933-938.
- Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K., Butterworth, C.A., 2004. Historical climate change and speciation: neotropical seasonally dry forest

- plants show patterns of both Tertiary and Quaternary diversification. *Philos. T. Roy. Soc. B*, 359, 515-538.
- Pennington, R.T., Lewis, G.P., Ratter, J.A., 2006. An overview of the plant diversity, biogeography and conservation of neotropical savannas and seasonally dry forests. In: Pennington, R.T., Ratter, J.A., Lewis, G.P. (Eds), *Neotropical Savannas and Seasonally Dry Forests: Diversity, Biogeography and Conservation*, CRC Press, Boca Raton, pp. 1-30.
- Pennington, T.D., Styles, B.T., Taylor, D.A.H., 1981. *Meliaceae*. *Flora Neotropica*, Monograph 28, The New York Botanical Garden, Bronx, New York, USA.
- Prado, D.E., 2000. Seasonally dry forests of tropical South America: from forgotten ecosystems to a new phytogeographic unit. *Edinb. J. Bot.* 57, 437-461.
- Prado, D.E., Gibbs, P.E., 1993. Patterns of species distribution in the dry seasonal forest of South America. *Ann. Mo. Bot. Gard.* 80, 902-927.
- Queiroz, C.S., Batista, F.R.C., Oliveira, L.O., Evolution of the 5.8S gene and internal transcribed spacers in *Carapichea ipecacuanha* (Rubiaceae) within a phylogeographic context. *Mol. Phyl. Evol.* (under review).
- Rogers, S.O., Bendich, A.J., 1987. Ribosomal RNA genes in plants: Variability in copy number and in the intergenic spacer. *Plant Mol. Biol.* 9, 509-520.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572-1574.
- Rosselló, J.A., Lázaro, A., Cosín, R., Molins, A., 2007. A phylogeographic split in *Buxus balearica* (Buxaceae) as evidenced by nuclear ribosomal markers: when ITS paralogues are welcome. *J. Mol. Evol.* 64, 143-157.
- Rossi, A.A.B., Oliveira, L.O., Venturini, B.A., Silva, R.S., 2009. Genetic diversity and geographic differentiation of disjunct Atlantic and Amazonian population of *Carapichea ipecacuanha* (Rubiaceae). *Genetica* 136, 57-67.

- Sang, T., Crawford, D.J., Stuessy, T.F., 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. *Proc. Natl. Acad. Sci. USA* 92, 6813-6817.
- Seibel, P.N., Müller, T., Dandekar, T., Wolf, M., 2008. Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4SALE. *BMC Res. Notes* 1, 91-98.
- Selig, C., Wolf, M., Müller, T., Dandekar, T. and Schultz, J., 2008. The ITS2 Database II: homology modeling RNA structure for molecular systematic Nucleic Acids Res. 36, D377-D380.
- Smith Jr., C. E., 1960. A revision of *Cedrela* (Meliaceae). *Fieldiana* 29, 295-341.
- Steininger, M.K., Tucker, C.J., Townshend, J.R.G., Killeen, T. J., Desch, A., Bell, V., Ersts, P. 2001. Tropical Deforestation in the Bolivian Amazon. *Environmental Conservation*. 28, 127-134.
- Styles, B.T., Vosa, C.G., 1971. Chromosome numbers in the Meliaceae. *Taxon* 20, 485-499.
- Taberlet, P., Gielly, L., Pantaou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17, 1105–1109.
- Veloso, H.P., Rangel Filho, A.L.R., Lima, J.C.A., 1991. *Classificação da Vegetação Brasileira, Adaptada de um Sistema Universal*. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.
- White, T.J., Bruns, T.D., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.) *PCR protocols, a guide to methods and applications*. Academic Press, California, pp. 315-322.

- Whitmore, T.C., Prance, G.T., 1987. Biogeography and Quaternary History in Tropical America. Oxford Science Publications, Oxford.
- Wuyts, J., Perriere, G., Van de Peer, Y., 2004. The European ribosomal RNA database. Nucleic Acids Res. 32, D101-D103.
- Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406-3415.

Table 1

Geographic distribution of the seven secondary structures of ITS2 in *Cedrela fissilis*.

Secondary structure	# haplotypes	Number of sequences	
		Atlantic range	Chiquitano range
A ₁	7	43	5
A ₂	1	3	-
A ₃	1	4	-
A ₄	1	2	-
B ₁	5	6	35
B ₂	1	-	1
B ₃	1	-	1

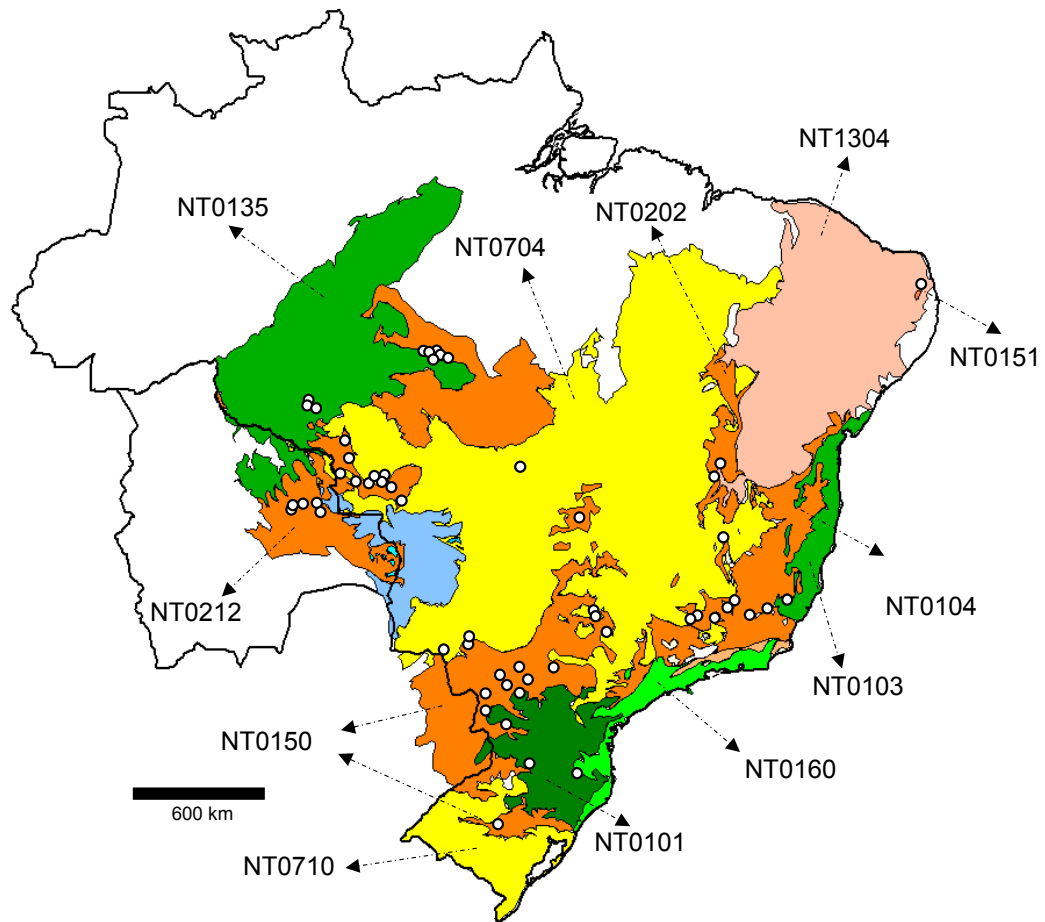


Fig. 1. Locations of populations of *Cedrela fissilis* (white dots) included in this study and their associated vegetation formations. Terrestrial ecoregions (Olson et al., 2001) include: NT0101, Araucaria moist forests; NT0103, Bahia coastal forests; NT0104, Bahia interior forests; NT0135, Madeira-Tapajós moist forests; NT0150, Paraná-Paraíba interior forests; NT0151, Pernambuco coastal forests; NT0202, Atlantic dry forests; NT0212, Chiquitano dry forests; NT0704, Cerrado. Ecoregions NT1304 (Caatinga), NT0710 (Uruguayan savanna), and NT0160 (Serra do Mar coastal forests) are shown for reference purpose. Color code: Orange, seasonal forests; green, moist evergreen forests; yellow, Cerrado; pink; Caatinga.

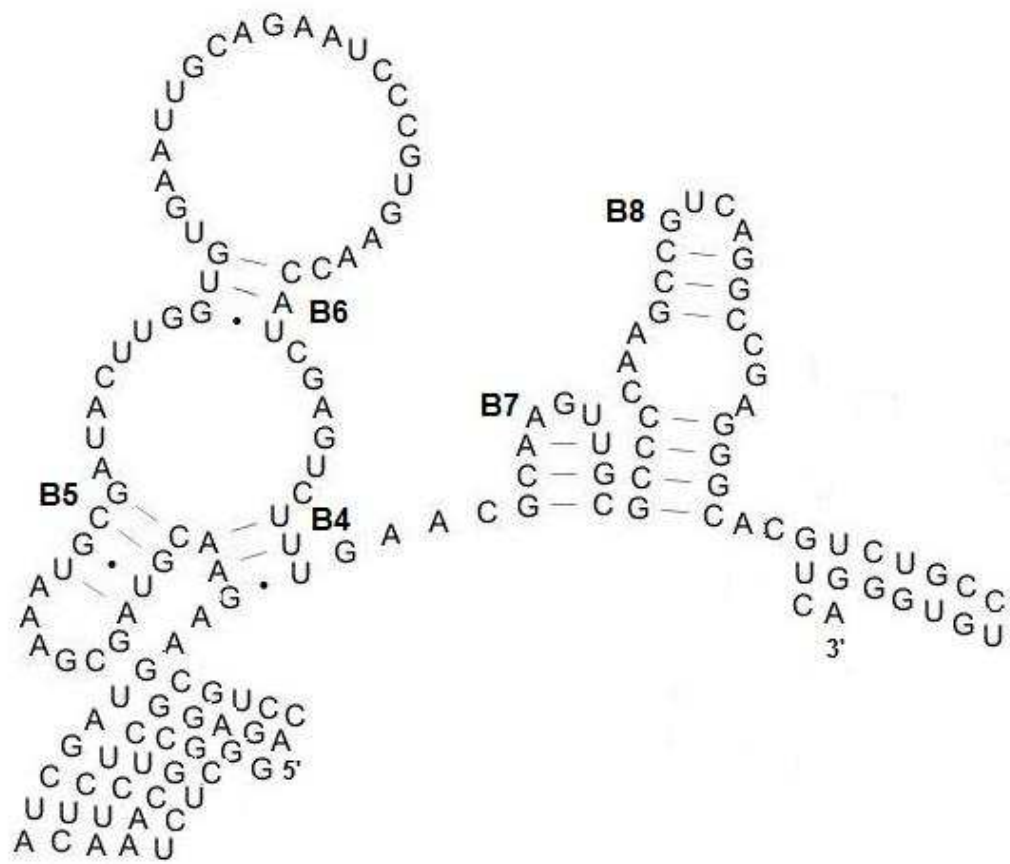


Fig. 2. Secondary structure of the 5.8S gene in *Cedrela fissilis*. Helix are numbered according to Wuyts et al. (2004).

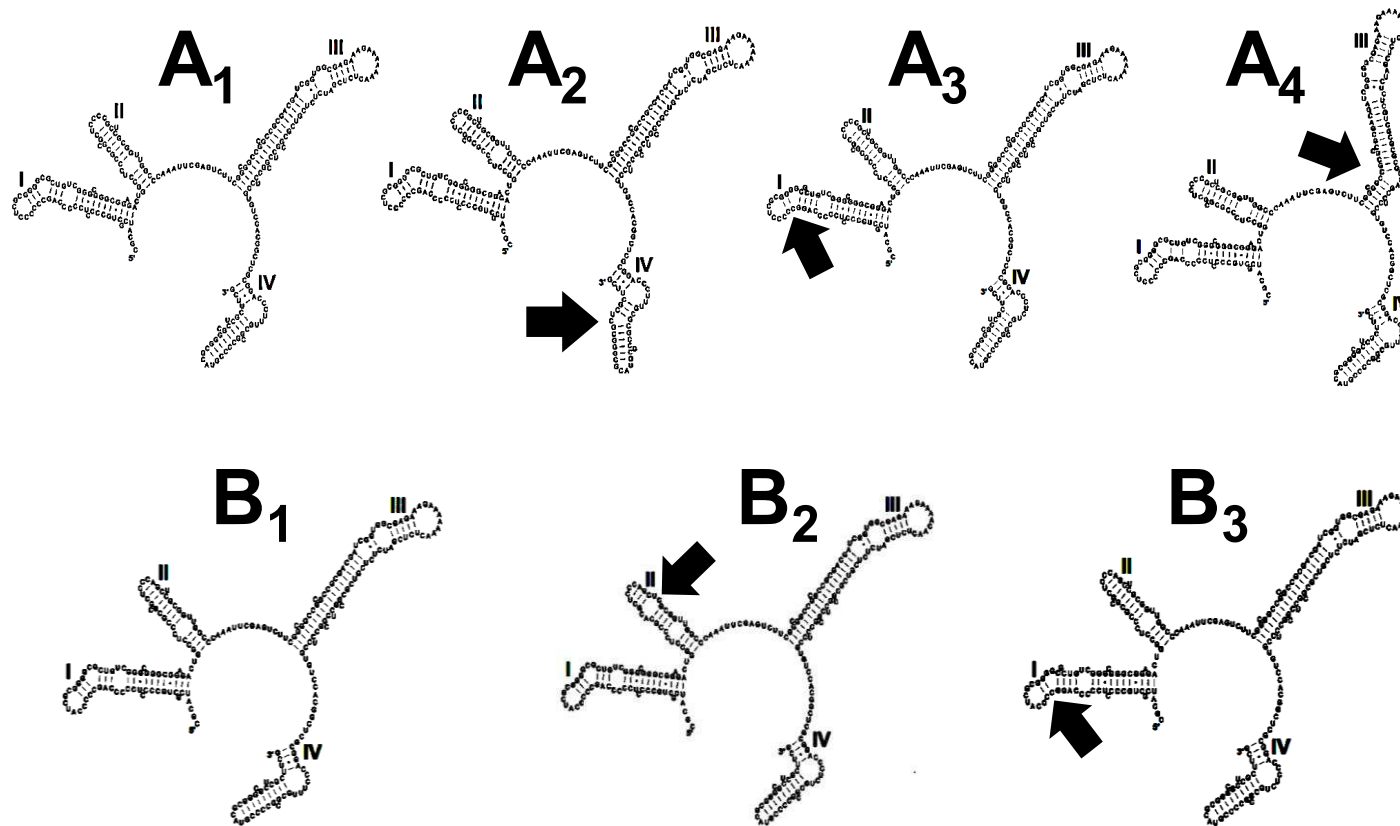


Fig. 3. Secondary structures for the ITS2 of *Cedrela fissilis*. A₁, the most prevalent secondary structure for ITS2 in sequences of the Atlantic range; A₂, A₃, and A₄, rare secondary structures for ITS2 in sequences of the Atlantic range; B₁, the most prevalent secondary structure for ITS2 in sequences of the Chiquitano range; B₂, and B₃, rare secondary structures for ITS2 in sequences of the Chiquitano range. Helices were numbered according to Coleman (2007). Arrows indicate structural changes in rare secondary structures relative to the most prevalent of the series.

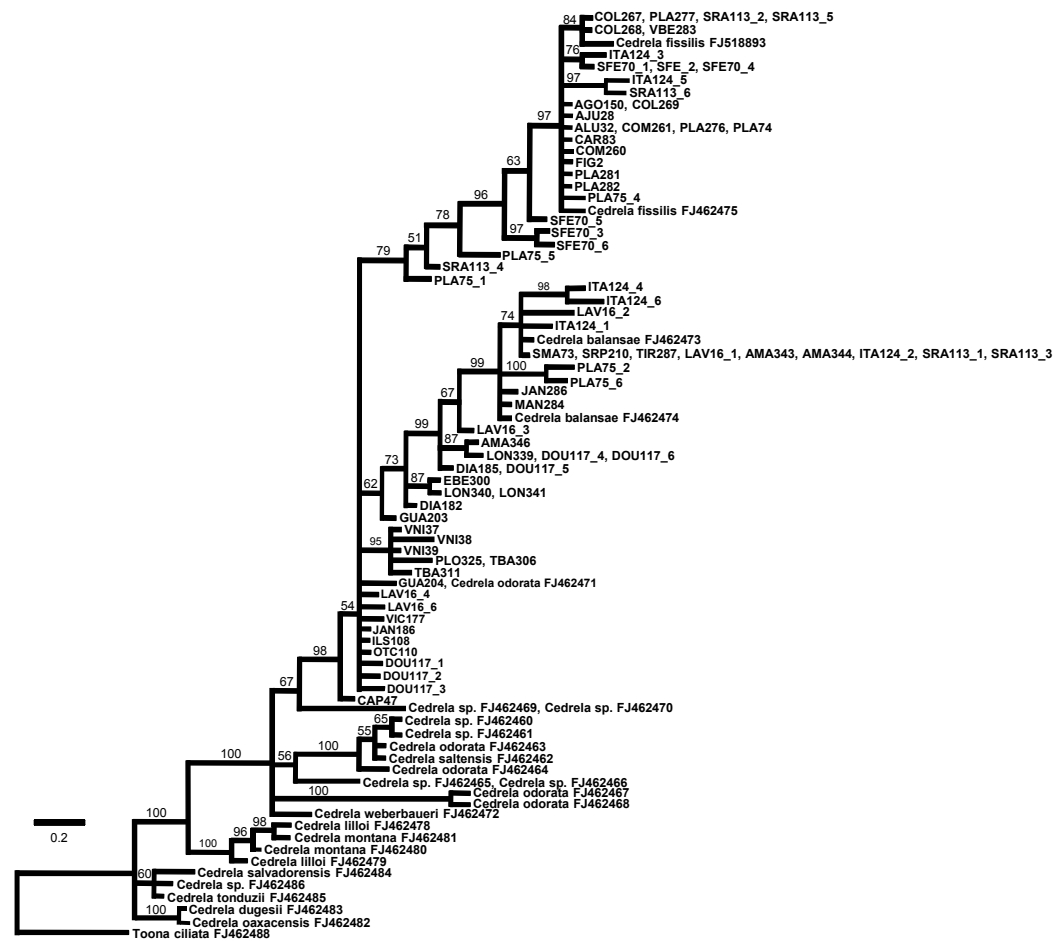


Fig. 4. Bayesian phylogeny (consensus tree) resulting from the ITS dataset for 26 *Cedrela* with *Toona ciliata* as the outgroup. Branch lengths are drawn to scale; nodal support values are given as posterior probabilities (%) above the branches. Scale bar correspond to the expected number of substitutions per site; codes correspond to specimens of *Cedrela fissilis*; GenBank accession numbers are listed in Appendix B (online supplementary material).

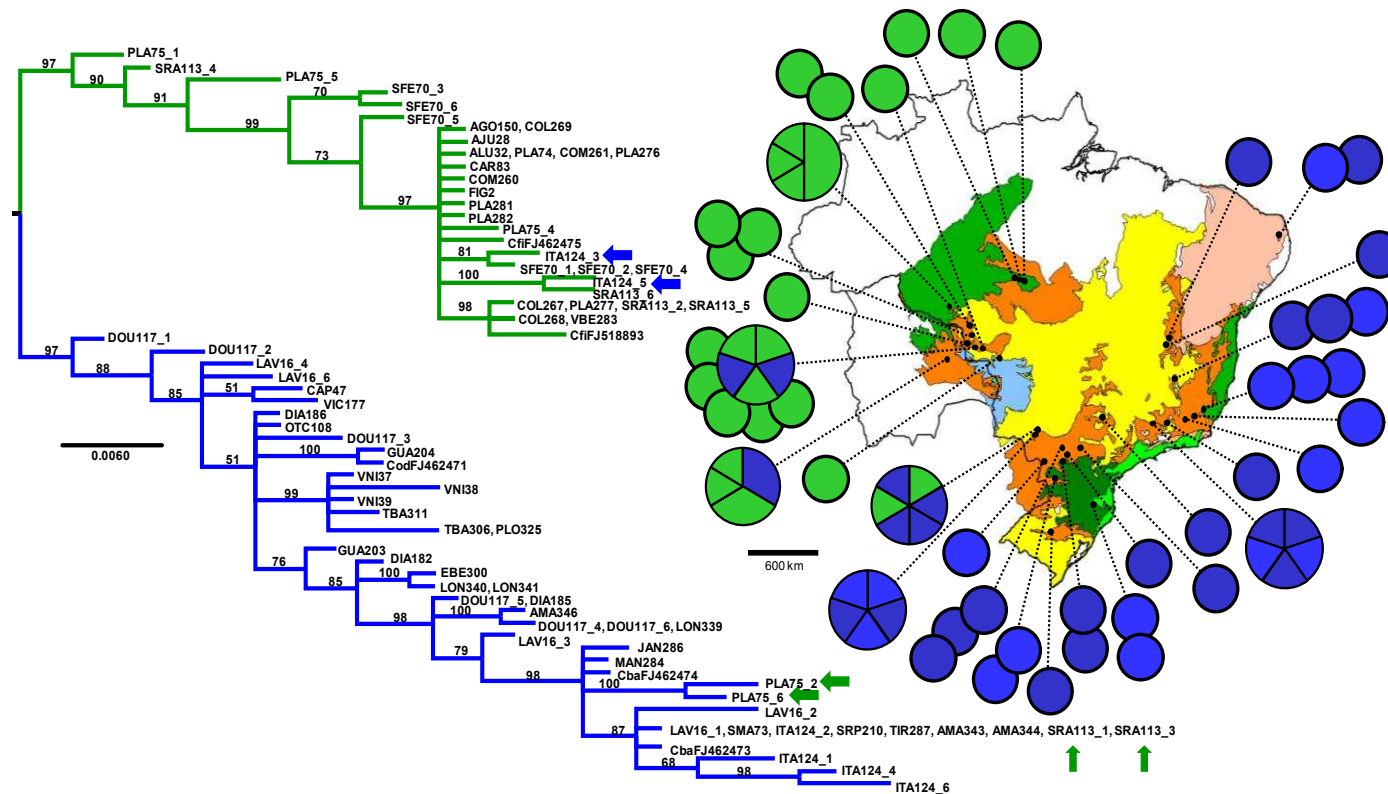


Fig. 5. Correspondence between the unrooted Bayesian phylogeny (consensus tree) resulting from the ITS dataset and the geographic distribution of 49 specimens of *Cedrela fissilis* sampled in Brazil and Bolivia. In the tree, branch lengths are drawn to scale; nodal support values are given as posterior probabilities (%) above the branches; scale bar correspond to the expected number of substitutions per site; codes correspond to specimens of *Cedrela fissilis* listed in Appendix B (online supplementary material); GenBank accession numbers are listed in Appendix B (online supplementary material). In the map, each smaller circle denotes a specimen with no intraindividual polymorphism for ITS (sequences obtained through direct sequencing); each larger circle denotes a specimen harboring intraindividual polymorphism (sequences obtained through cloning - slices represent frequency of a given haplotype). Color code: Blue, Atlantic range; Green, Chiquitano range. Arrows denote unexpected placements (see text).

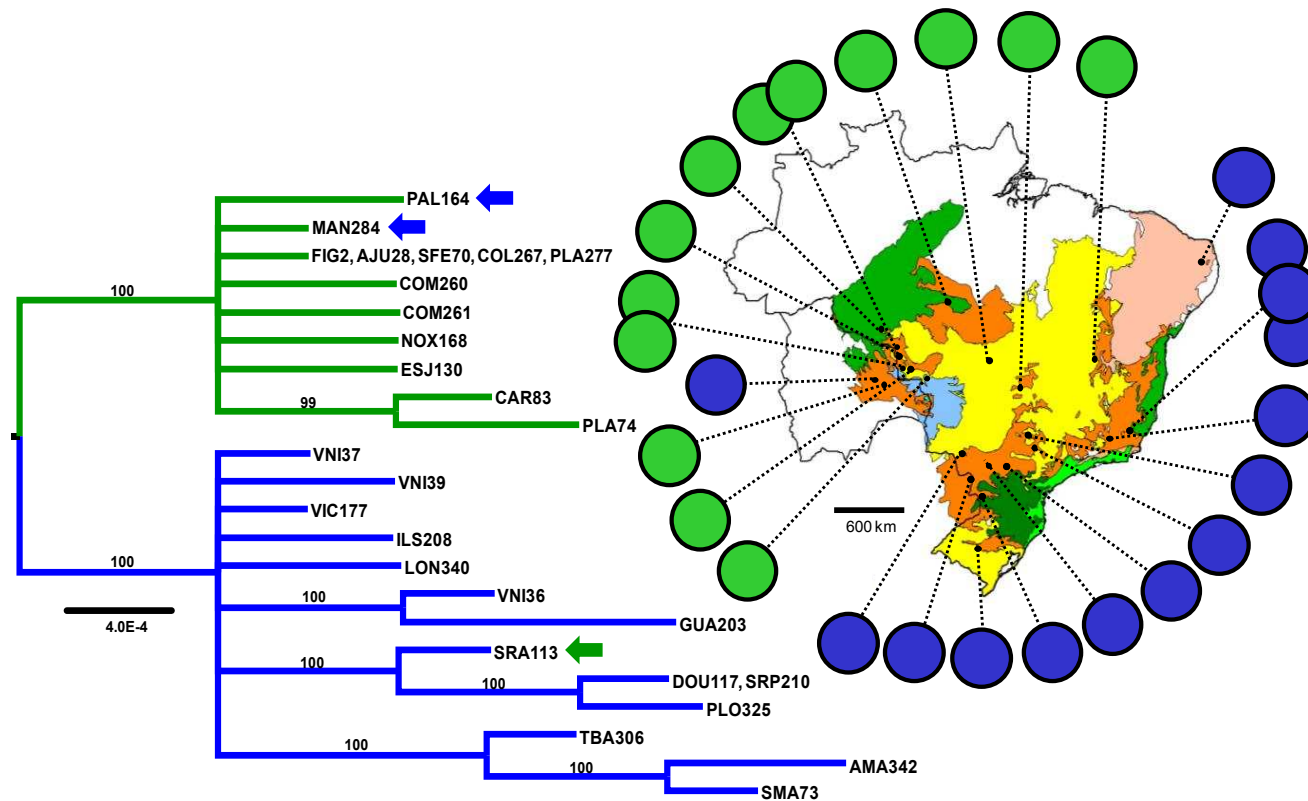


Fig. 6. Correspondence between the unrooted Bayesian phylogeny (consensus tree) resulting from the chloroplastidic regions *trnT-trnL*, *trnS-trnG* and *psbB-psbT-psbN* dataset and the geographic distribution of 27 specimens of *Cedrela fissilis* sampled in Brazil and Bolivia. In the tree, branch lengths are drawn to scale; nodal support values are given as posterior probabilities (%) above the branches; scale bar correspond to the expected number of substitutions per site; codes correspond to specimens of *Cedrela fissilis*; GenBank accession numbers are listed in Appendix B (online supplementary material). In the map, each circle denotes a specimen. Color code: Blue, Atlantic range; Green, Chiquitano range. Arrows denote unexpected placements (see text).

Appendix A, online supplementary material

Terrestrial ecoregions and location of populations of *Cedrela fissilis* sampled for analyses. Color code: Green, moist evergreen forests; orange, seasonal forests; yellow, savanna.

Ecoregion	Population code (collection #)	Locality	Coordinates (lat / lon)
NT0101	CMO (42, 43, 44)	Campo Mourão/PR	-24 02' 17.30"/-52 24' 33.60"
NT0101	OTC (108, 110)	Otacílio Costa/SC	-27 28' 35.45"/-50 07' 10.29"
NT0101	COC (336)	Concórdia/SC	-27 14' 20.86"/-52 00' 38.85"
NT0103	VNI (36, 37, 38, 39)	Venda Nova do Imigrante/ES	-20 20' 05.68"/-41 07' 49.72"
NT0104	LAV (16, 18*)	Lavras/MG	-21 14' 59.78"/-44 59' 22.01"
NT0104	IJA (19)	Ijaci/MG	-21 09' 47.55"/-44 55' 23.82"
NT0104	CAP (47)	Caparaó/MG	-20 31' 45.60"/-41 55' 14.10"
NT0104	ORB (143, 144)	Ouro Branco/MG	-20 30' 17.99"/-43 42' 46.89"
NT0104	ORP (148)	Ouro Preto/MG	-20 23' 18.90"/-43 30' 23.76"
NT0104	VIC (177, 178, 201)	Viçosa/MG	-20 46' 22.62"/-42 52' 26.35"
NT0135	CAC (62, 68, 69)	Cacoal/RO	-11 29' 50.71"/-61 15' 04.45"
NT0135	PBU (64, 71)	Pimenta Bueno/RO	-11 40' 56.40"/-61 32' 25.38"
NT0135	SFE (70)	São Felipe D'oeste/RO	-11 52' 40.92"/-61 31' 16.98"
NT0135	AJU (28, 29, 30)	Júnior - Alta Floresta/MT	-09 54' 00.00"/-55 53' 59.00"
NT0135	ALU (31, 32, 33, 34)	Lucirene - Alta Floresta/MT	-09 52' 23.67"/-56 04' 22.68"
NT0135	AGO (149, 150, 151, 152)	Agrocondor - Alta Floresta/MT	-09 53' 56.97"/-56 09' 16.12"
NT0135	PFI (153, 154, 155)	Ponte frigorífico - Alta Floresta/MT	-09 52' 59.91"/-56 08' 46.25"
NT0135	MOA (157, 158, 159)	Monte Alegre - Alta Floresta/MT	-09 51' 40.59"/-56 11' 05.07"
NT0135	SON (160, 161, 162)	Sol Nascente - Alta Floresta/MT	-10 06' 07.72"/-56 13' 00.49"
NT0150	PAL (163, 164, 165)	Palmeiras de Goiás/GO	-16 47' 23.05"/-49 55' 58.95"
NT0150	DOU (117)	Dourados/MS	-22 11' 56.99"/-54 49' 01.08"
NT0150	ITA (120, 123, 124, 125, 127, 128)	Itaporã/MS	-22 00' 09.68"/-54 42' 52.59"
NT0150	POP (171, 172, 173, 174, 175)	Ponta Porã/MS	-22 32' 58.11"/-55 43' 01.85"
NT0150	SPI (139, 140, 141, 142)	São Pedro do Iguazu/PR	-24 55' 59.58"/-53 50' 52.95"
NT0150	PCA (189, 200)	Porto Camargo - Icaraíma/PR	-23 24' 02.81"/-53 37' 56.58"
NT0150	BRP (192, 193, 195, 197)	BR - Umuarama/PR	-23 39' 08.85"/-53 22' 13.13"
NT0150	EBE (300, 301, 302)	Engenheiro Beltrão/PR	-23 48' 19.16"/-52 15' 41.92"
NT0150	TBA (306, 308, 311, 313)	Três Barras do Paraná/PR	-25 25' 45.56"/-53 07' 37.28"

NT0150	PLO (320, 325, 327, 330)	Palotina/PR	-24 18' 34.29"/-53 54' 32.15"
NT0150	LON (339, 340, 341)	Londrina/PR	-23 09' 43.29"/-51 09' 55.15"
NT0150	AMA (342, 343, 344, 345, 346, 347)	Amaporã/PR	-23 05' 07.95"/-52 47' 46.31"
NT0150	SMA (73, 86, 87, 90)	Santa Maria/RS	-29 43' 03.94"/-53 43' 29.04"
NT0150	CED (207)	Cedral/SP	-20 54' 25.59"/-49 17' 00.72"
NT0150	SRP (210)	São José do Rio Preto/SP	-20 47' 06.37"/-49 19' 12.07"
NT0151	GUA (203, 204)	Guarabira/PB	-06 50' 23.25"/-35 29' 23.73"
NT0202	MAN (284)	Manga/MG	-14 47' 27.96"/-43 56' 19.92"
NT0202	JAN (286)	Januária/MG	-15 11' 22.56"/-44 12' 24.48"
NT0212	RUI (96, 97, 98, 99)	Cármem Ruiz/BO	-16 03' 03.18"/-61 05' 21.95"
NT0212	INP (100, 101, 102, 103, 105)	INPA-Concepción/BO	-16 07' 51.87"/-62 00' 19.71"
NT0212	SRA (112, 113, 114, 115)	Coronacion-San Ramón/BO	-15 59' 59.00"/-61 40' 00.00"
NT0212	ESJ (130, 131, 132, 133)	Estancia San Juan-San Ignacio/BO	-16 22' 16.99"/-60 57' 25.61"
NT0212	LAG (134, 135, 136, 137, 138)	La Lagunilla-Concepción/BO	-16 08' 59.74"/-62 03' 16.71"
NT0212	CAR (80, 83)	Cáceres/MT	-16 04' 35.43"/-57 36' 14.96"
NT0212	COM (260, 261, 262, 263, 264)	Comodoro/MT	-13 41' 26.93"/-59 46' 57.33"
NT0212	COL (266, 267, 268, 269)	Colônia dos mineiros - Nova Lacerda/MT	-09 51' 40.59"/-56 11' 05.07"
NT0212	FIG (1,2,3,4,5,6)	Figueirópolis D'oeste/MT	-15 26' 43.80"/-58 44' 25.73"
NT0212	JUB (7,8,9)	Juba - Lambari D'oeste/MT	-15 18' 55.25"/-57 59' 26.71"
NT0212	MOZ (12)	Mozar - Salto do Céu/MT	-15 07' 48.33"/-58 07' 35.00"
NT0212	PLA (74, 75, 79, 276, 277, 278, 281, 282)	Pontes e Lacerda/MT	-15 13' 52.38"/-59 22' 10.64"
NT0212	RIB (76)	Rio Branco/MT	-15 14' 33.37"/-58 07' 14.89"
NT0212	SOR (13,14, 78, 84, 85)	Soroteca - Curvelândia/MT	-15 36' 23.87"/-57 54' 43.57"
NT0212	VBE (283)	Vila Bela da Santíssima Trindade/MT	-15 01' 39.63"/-59 57' 16.27"
NT0704	DIA (180, 182, 183, 185, 186)	Diamantina/MG	-18 24' 33.60"/-43 29' 24.70"
NT0704	IPO (208)	Itápolis/SP	-21 34' 55.57"/-48 49' 49.31"
NT0704	NOX (166, 168, 169)	Nova Xavantina/MT	-14 40' 19.79"/-52 20' 28.16"
NT0704	TIR (287)	Tiradentes/MG	-21 06' 22.63"/-44 10' 44.71"

Appendix B, online supplementary material

Taxa, origin and GenBank accession numbers for specimens used in ITS phylogenetic analysis. Genbank accession numbers are as in Muellner et al. (2009).

Taxon	Origin	GenBank number ITS
<i>Cedrela balansae</i>	Argentina	FJ462473
<i>Cedrela balansae</i>	Paraguay	FJ462474
<i>Cedrela dugesii</i>	México	FJ462483
<i>Cedrela fissilis</i>	Peru	FJ462475
<i>Cedrela fissilis</i>	Paraguay	FJ518893
<i>Cedrela lilloi</i>	Bolivia	FJ462478
<i>Cedrela lilloi</i>	Peru	FJ462479
<i>Cedrela montana</i>	Equador	FJ462481
<i>Cedrela montana</i>	Peru	FJ462480
<i>Cedrela oaxacensis</i>	México	FJ462482
<i>Cedrela odorata</i>	Guiana Francesa	FJ462463
<i>Cedrela odorata</i>	Equador	FJ462464
<i>Cedrela odorata</i>	Belize	FJ462467
<i>Cedrela odorata</i>	El Salvador	FJ462468
<i>Cedrela odorata</i>	Brasil	FJ462471
<i>Cedrela saltensis</i>	Argentina	FJ462462
<i>Cedrela salvadorensis</i>	Costa Rica	FJ462484
<i>Cedrela sp.</i>	Peru	FJ462460
<i>Cedrela sp.</i>	Peru	FJ462461
<i>Cedrela sp.</i>	Peru	FJ462465
<i>Cedrela sp.</i>	Peru	FJ462466
<i>Cedrela sp.</i>	Peru	FJ462469
<i>Cedrela sp.</i>	Peru	FJ462470
<i>Cedrela sp.</i>	Peru	FJ462486
<i>Cedrela tonduzii</i>	Costa Rica	FJ462485
<i>Cedrela weberbaueri</i>	Peru	FJ462472
<i>Toona ciliata</i>	Australia	FJ462488

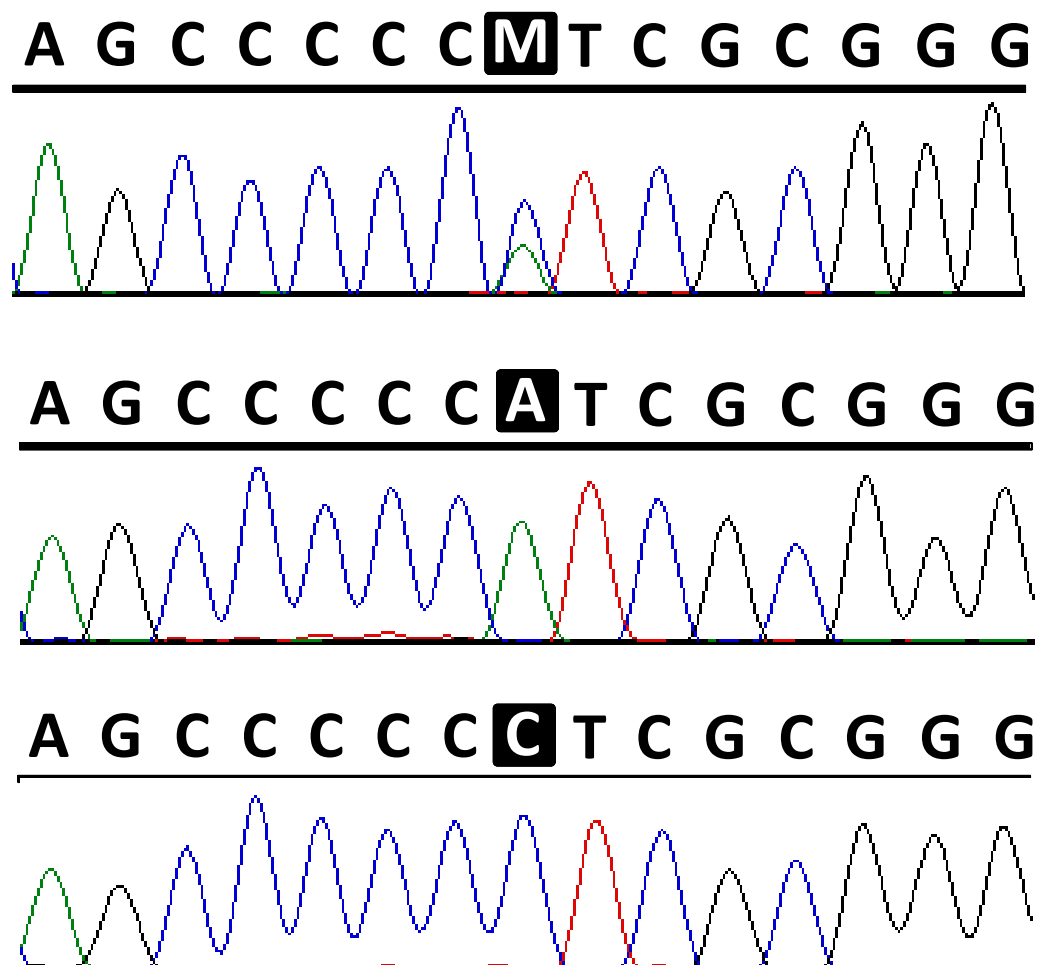


Fig. S1 Partial electropherograms from direct sequences of the ITS-2 spacer from three specimens of *Cedrela fissilis*. Top: Sequence of specimen PLA_281, showing the two overlapping double peaks that resulted in the ambiguous reading M, an indicative of ITS paralogues of same length. Middle and bottom: Sequences of specimens AJU_28 and DIA_182, respectively, showing no double peaks at the corresponding sites harboring M at the top sequence.

CAPÍTULO II

FILOGEOGRAFIA MOLECULAR DE *Cedrela fissilis* Vell. E *Cedrela odorata* L. BASEADA EM GENES NUCLEAR E CLOROPLASTÍDICO

RESUMO

Cedrela fissilis e *Cedrela odorata* são espécies arbóreas reconhecidas pela alta qualidade e valor associado a sua madeira e, em função disso, possuem um longo histórico de exploração. No presente estudo, foram amostrados 186 espécimes de *C. fissilis* e 27 espécimes de *C. odorata* distribuídos em 59 populações para obtenção de dados de sequências de DNA dos genomas cloroplastídico e nuclear. Os resultados mostraram que as sequências nucleotídicas de *C. fissilis* e *C. odorata* apresentaram alta similaridade e foram agrupadas em uma única rede de haplótipos, sendo que esse agrupamento parece ter-se dado em função da distribuição geográfica. As análises filogeográficas sugerem que as populações atuais de *C. fissilis* e *C. odorata* são descendentes de pelo menos duas linhagens ancestrais distintas. A estrutura filogeográfica mostrou uma distribuição desigual da diversidade genética, uma separação evidente e altos níveis de diferenciação genética entre as regiões do Atlântico e de Chiquitano. As populações do Atlântico apresentam níveis de diversidade genética mais elevados do que a região de Chiquitano, sugerindo que a região do Atlântico possa ser o centro de diversidade para a espécie no Brasil. Episódios de isolamento genético associados a eventos de dispersão moldaram a história evolutiva de *C. fissilis* e *C. odorata*. Eventos de expansão recente entre as Regiões do Atlântico e de Chiquitano reconectou essas regiões e tem favorecido o estabelecimento de zonas de contato secundário entre as populações nas florestas do interior do Paraná-Paraíba (ecorregião NT0150), nas florestas secas de Chiquitano (ecorregião NT0212) e nas florestas secas do Atlântico (ecorregião NT0202).

INTRODUÇÃO

As Florestas Estacionais encontram-se distribuídas de forma fragmentada e disjunta na região Neotropical, com blocos de Floresta Estacional se interpondo a outras formações vegetais como a Floresta Ombrófila, o Cerrado, a Caatinga e o Chaco (Olson et al., 2001).

Segundo Veloso et al. (1991) o conceito de Floresta Estacional está condicionado pela dupla estacionalidade climática, com uma estação chuvosa seguida de longo período seco (até 5 meses de estiagem). Além disso, as florestas estacionais podem ocorrer em áreas onde a distribuição das chuvas é mais uniforme, porém o inverno é rigoroso e a temperatura média é inferior a 15° C, o que pode provocar seca fisiológica (Veloso et al., 1991; Oliveira-Filho e Fontes, 2000). As florestas estacionais podem ser classificadas em semidecídua e decídua de acordo com a porcentagem de indivíduos arbóreos que perdem suas folhas nos períodos desfavoráveis, na floresta estacional semidecídua a queda das folhas ocorre em 20-50% dos componentes e na floresta estacional decídua a queda das folhas ocorre em mais de 50% dos componentes (Veloso et al., 1991). Entretanto, alguns autores têm observado que é difícil distinguir os dois tipos de floresta fundamentados unicamente no percentual de deciduidade, principalmente no Brasil central, onde as florestas semidecíduas e decíduas formam um contínuo no qual as características são determinadas por variações locais de umidade e fertilidade do solo (Oliveira Filho e Ratter, 2002).

Alguns autores sugerem que as Florestas Estacionais Neotropicais sofreram contrações (Carnaval e Moritz, 2008; Carnaval et al., 2009) e expansões (Auler et al., 2004; Mayle, 2004) em resposta às mudanças climáticas que ocorreram no

Quaternário (Ledru et al., 2007). Estas alterações de vegetação na região neotropical teriam sido resultantes de perturbações nas condições ecológicas locais e regionais produzidas por ciclos de climas secos/úmidos com resfriamento/aquecimento que aconteceram nos últimos dois milhões de anos do Pleistoceno (Whitmore e Prance, 1987) ou possivelmente ainda em tempos anteriores (Pennington et al., 2004; Mayle, 2004). De acordo com a hipótese do “Arco Pleistocênico”, proposta por Prado e Gibbs (1993), a disjunção atual das florestas estacionais tropicais seriam relíquias formadas por vicariância de uma formação florestal pleistocênica de dimensão muito maior que cobria a América do Sul. Uma explicação alternativa propõe que esta disjunção seria o resultado da ocorrência de raros eventos de dispersão a longa distância de espécies associadas a esse tipo florestal (Gentry, 1982; Mayle, 2004).

O conhecimento sobre as florestas tropicais em todo o mundo, e particularmente nas Américas, é extremamente concentrado em áreas úmidas, como a Amazônia e a Mata Atlântica. E embora as florestas estacionais tropicais recubram grande parte dos trópicos, apenas 14% dos estudos sobre florestas tropicais foram realizados nestas formações florestais (Sánchez-Azofeifa et al., 2005). Segundo Janzen (1988) as florestas estacionais tropicais se constituem no ecossistema mais ameaçado dentre as fitofisionomias tropicais, sendo que na América Latina cerca de 60% destas florestas já foram destruídas. Esse tipo florestal é bastante visado para a extração de madeiras, pois suas espécies mais abundantes têm elevado valor econômico. Além disso, a cobertura vegetal tem sido muito fragmentada, e a maioria dos fragmentos resultantes se encontra sob perturbação em decorrência de fogo e de pecuária extensiva, uma vez que sua ocorrência coincide com solos de origem calcária, altamente férteis, e, portanto, muito cobiçados para atividades agropecuárias (Oliveira-Filho et al., 1994). O estudo da história evolutiva das florestas estacionais tropicais é de grande

importância para o entendimento das ligações históricas entre os fragmentos ainda existentes desta formação florestal, bem como para compreender os processos de dispersão gênica que podem ter ocorrido entre estes fragmentos. O conhecimento dos padrões de distribuição da diversidade genética permite reconhecer áreas com maior diversidade e isto pode resultar no desenvolvimento de propostas eficientes para conservação de populações.

Filogeografia é o estudo dos princípios e processos que governam a distribuição geográfica das linhagens genealógicas, especialmente aquelas no nível intraespecífico e entre espécies altamente relacionadas (Avise, 2000). Os estudos de filogeografia molecular procuram examinar dados moleculares de populações naturais contemporâneas a procura das marcas evolutivas deixadas no genoma por processos históricos, e, deste modo, pode contribuir para desenvolver e aperfeiçoar hipóteses biogeográficas que sejam capazes de explicar as distribuições geográficas atuais destas espécies e populações.

O gênero *Cedrela* compreende 17 espécies neotropicais naturalmente distribuídas pelas Américas Central e do Sul (Muellner et al., 2010). A maior parte das espécies pertencentes ao gênero *Cedrela* apresenta distribuição restrita as florestas tropicais decíduas, entretanto, *Cedrela odorata* e *Cedrela fissilis* apresentam maior amplitude de distribuição, estando presente em áreas de floresta estacional e de floresta ombrófila (Muellner et al., 2009). *C. fissilis* (cedro branco) e *C. odorata* (cedro rosa) são espécies arbóreas reconhecidas pela alta qualidade e valor associado a sua madeira e, em função disso, possuem um longo histórico de exploração, sendo ainda uma “commodity” valiosa na confecção de móveis e na construção civil (Carvalho, 2003). Além da exploração promovida pelo interesse econômico, a perda de habitat decorrente da expansão das cidades e das fronteiras agrícolas tem promovido a redução destas arbóreas tropicais, eliminando ou fragmentando suas populações naturais (Valera, 1997). Como consequência direta

destes fatos, *C. fissilis* e *C. odorata* integram atualmente a lista de espécies ameaçadas (IUCN – Red List of Threatened Species) pertencendo às categorias “ameaçada” e “vulnerável”, respectivamente (IUCN, 2008). Algumas estratégias para a conservação, melhoria e gestão sustentável destes importantes recursos genéticos tem sido propostas para o gênero, como por exemplo, a inclusão de algumas populações de *C. odorata* na lista de espécies do CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora), um acordo internacional entre os governos que procura garantir que o comércio internacional de espécimes de animais e plantas não ameace a sobrevivência dos mesmos. Entretanto, para que medidas eficazes para a conservação das espécies possam ser propostas faz se necessário que as espécies sejam entidades morfológicas e genéticas bem definidas (Valera, 1997).

Algumas populações de *C. odorata* da Costa Rica foram analisadas utilizando marcadores RAPD e apresentaram diferenciação entre as populações submetidas à ambientes com diferentes níveis de precipitação (Gillies et al., 1997). Adicionalmente, Cavers et al. (2003) utilizaram RFLP para analisar a variação intraespecífica de populações mesoamericanas de *C. odorata* e observaram um elevado nível de diferenciação dentro das populações, sendo que as linhagens obtidas foram geograficamente estruturadas em função da umidade presente nos diferentes habitats. Recentes estudos moleculares forneceram evidências de que existam pelo menos três espécies distintas, morfológicamente não diferenciadas sob a denominação de *C. odorata* (Muellner et al., 2009).

Neste estudo, ferramentas de filogeografia molecular foram utilizadas para a análise de dados de sequências de DNA dos genomas cloroplastídico e nuclear das espécies *Cedrela fissilis* e *Cedrela odorata* com intuito de responder as seguintes questões: (i) quantas e quais seriam as linhagens genealógicas que compõem as populações contemporâneas de *Cedrela fissilis* e de *Cedrela odorata*? Houve

concordância entre a história evolutiva revelada pelas análises moleculares de genes cloroplastídicos e de genes nucleares? (ii) as estruturas filogeográficas de *Cedrela fissilis* e de *Cedrela odorata* sugerem a existência de áreas de refúgio e áreas recentemente colonizadas? Onde estariam localizadas estas áreas e quais seriam as possíveis rotas de migração? (iii) o padrão filogeográfico apoiaria a vicariância de acordo com a hipótese do Arco Pleistocênico, a dispersão a longa distância ou uma associação entre vicariância e dispersão para explicar a distribuição das populações atuais?

MATERIAL E MÉTODOS

Estratégia de amostragem e extração de DNA

Neste estudo foram avaliados 186 espécimes de *Cedrela fissilis* e 27 espécimes de *Cedrela odorata* distribuídos em 59 populações (Figura 1). As localidades das populações amostradas foram determinadas de forma a abranger a maior parte da região de ocorrência natural da espécie no Brasil e se concentraram em áreas de floresta estacional ou em formações vegetais adjacentes. Para definir as formações vegetais nas quais as populações de *C. fissilis* e *C. odorata* foram coletadas foi consultado o mapa das Ecorregiões Terrestres do Mundo publicado pelo WWF (World Wildlife Fund) (Olson et al., 2001). A classificação da vegetação do Brasil segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2004) foi utilizada para obtenção de mais detalhes a respeito das formações vegetais em escala local. As localizações geográficas das populações amostradas foram registradas em mapas topográficos (fonte: <http://www.worldwildlife.org>) usando o software de GPS.

De forma geral, a amostragem foi realizada em ambos os lados de uma formação aberta que recobre grande parte do Brasil Central, o Cerrado (ecorregião NT0704). Adicionalmente, algumas populações amostradas foram localizadas dentro do Cerrado ou em áreas de transição. No lado oeste do Cerrado, a maioria das populações amostradas se concentrou em duas ecorregiões: nas florestas úmidas Madeira-Tapajós (NT0135) e nas florestas secas de Chiquitano (NT0212), que serão a partir de agora coletivamente denominadas “Região de Chiquitano”. No lado leste do Cerrado, as populações foram amostradas em cinco ecorregiões: nas florestas úmidas de Araucária (NT0101), nas florestas litorâneas da Bahia

(NT0103), nas florestas do interior da Bahia (NT0104), nas florestas do interior do Paraná-Paraíba (NT0150), e nas florestas secas do Atlântico (NT0202). A partir de agora as populações amostradas nessas áreas serão coletivamente denominadas “Região do Atlântico”. A região do Atlântico incluiu uma população localizada em um bloco de florestas estacionais no Nordeste do Brasil (ecorregião NT0151 – florestas litorâneas de Pernambuco), inserida em uma área dominada por vegetação arbustiva (ecorregião NT1304 - Caatinga). As populações amostradas, suas respectivas localizações geográficas e ecorregiões a que pertencem, bem como o número de espécimes coletados por população estão relacionadas na Tabela 1.

Em cada população, o tamanho amostral variou em função da disponibilidade de árvores encontradas em cada local (1 a 12 espécimes por população). Foram coletados de 3 a 5 folíolos/espécime e o material foliar foi armazenado em freezer -80°C ou desidratado em sílica gel e mantido em temperatura ambiente até que o DNA fosse extraído. O DNA genômico total foi extraído segundo o procedimento descrito por Rossi et al. (2009).

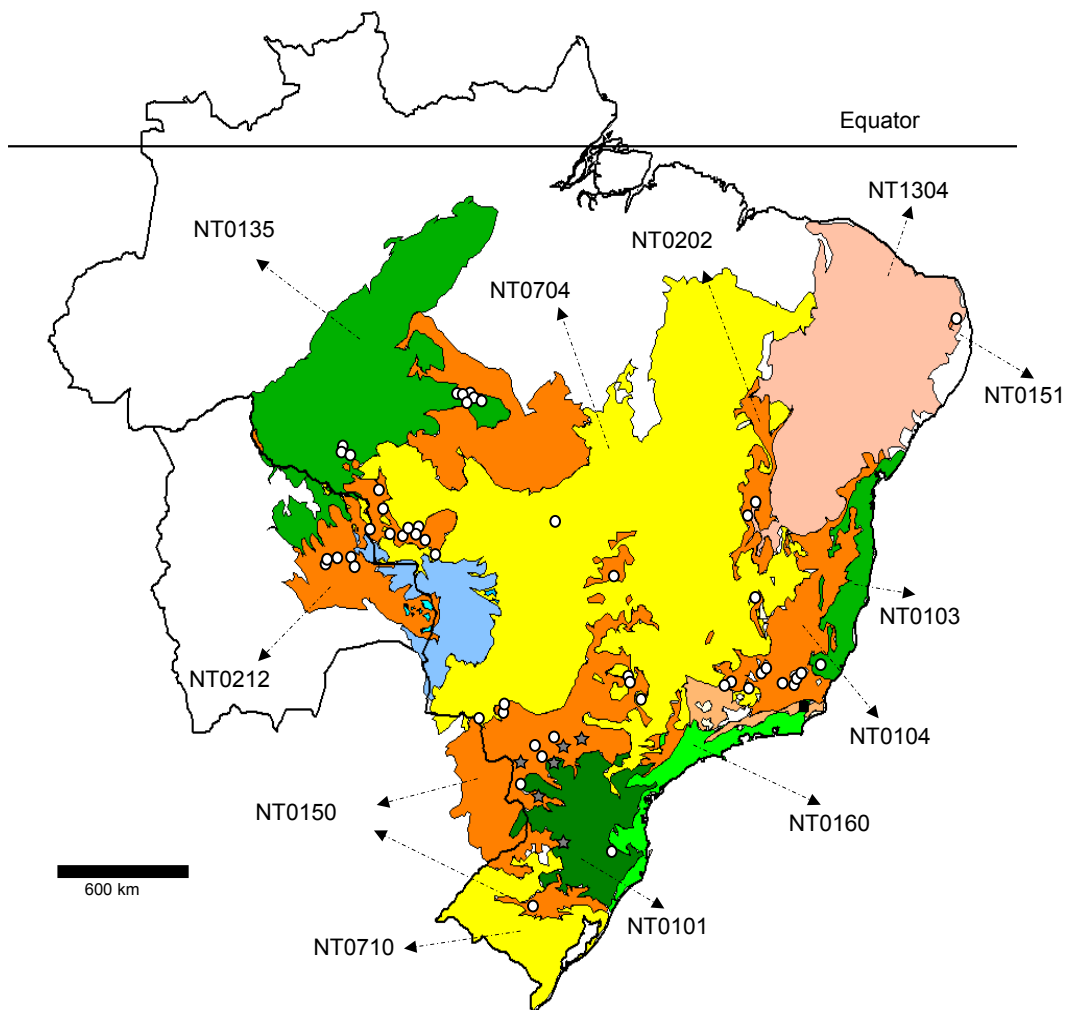


Figura 1. Distribuição geográfica das 59 populações de *Cedrela fissilis* Vell. e *Cedrela odorata* L. e as formações vegetais associadas. Círculos brancos - populações nas quais unicamente *C. fissilis* foi coletada; quadrado preto - população na qual unicamente *C. odorata* foi coletada; estrelas em cinza - populações nas quais *C. fissilis* e *C. odorata* foram coletadas. Ecorregiões terrestres são as seguintes: NT0101, florestas úmidas de Araucaria; NT0103, florestas litorâneas da Bahia; NT0104, florestas do interior da Bahia; NT0135, florestas úmidas do Madeira-Tapajós; NT0150, florestas do interior do Paraná-Paraíba; NT0151, florestas litorâneas de Pernambuco; NT0160, florestas litorâneas da Serra do Mar; NT0202, florestas secas do Atlântico; NT0212, florestas secas de Chiquitano; NT0704, Cerrado; NT0710, savana Uruguaiana e NT1304, Caatinga.

Tabela 1. Populações de *Cedrela fissilis* Vell. e *Cedrela odorata* L. amostradas com seus respectivos códigos, coordenadas geográficas, ecorregiões e número de espécimes utilizados nas análises de cpDNA e ITS.

Populações (código)	Latitude (S)	Longitude (O)	Ecorregiões	Espécie	Tamanho amostral				
					cpDNA		ITS		
REGIÃO DO ATLÂNTICO					94	22	82(30)	26(10)	
Campo Mourão/PR (CMO)	24°02'17.30"	52°24'33.60"	NT0101	C. fi	C. od	3	1	3 (0)	1 (1)
Concórdia/SC (COC)	27°14'20.86"	52°00'38.85"	NT0101	C. fi	C. od	1	2	1 (0)	2 (1)
Otacílio Costa/SC (OTC)	27°28'35.45"	50°07'10.29"	NT0101	C. fi		2		2 (2)	
Venda Nova do Imigrante/ES (VNI)	20°20'05.68"	41°07'49.72"	NT0103	C. fi		4		4 (3)	
Caparaó/MG (CAP)	20°31'45.60"	41°55'14.10"	NT0104	C. fi		1		1 (1)	
Carangola/MG (CAL)	20°44'43.42"	42°01'06.14"	NT0104	C. fi		1		0	
Faria Lemos/MG (FLE)	20°48'39.60"	42°01'11.91"	NT0104	C. fi		2		0	
Ijaci/MG (IJA)	21°09'47.55"	44°55'23.82"	NT0104	C. fi		3		1 (0)	
Lavras/MG (LAV)	21°14'59.78"	44°59'22.01"	NT0104	C. fi		3		2 (1)	
Ouro Branco/MG (ORB)	20°30'17.99"	43°42'46.89"	NT0104	C. fi		2		2 (0)	
Ouro Preto/MG (ORP)	20°23'18.90"	43°30'23.76"	NT0104	C. fi		1		1 (0)	
Viçosa/MG (VIC)	20°46'22.62"	42°52'26.35"	NT0104	C. fi		3		3 (1)	
Palmeiras de Goiás/GO (PAL)	16°47'23.05"	49°55'58.95"	NT0150	C. fi		3		3 (0)	
Cedral/SP (CED)	20°54'25.59"	49°17'00.72"	NT0150	C. fi		1		1 (0)	
São José do Rio Preto/SP (SRP)	20°47'06.37"	49°19'12.07"	NT0150	C. fi		1		1 (1)	
Dourados/MS (DOU)	22°11'56.99"	54°49'01.08"	NT0150	C. fi		5		1 (1)	
Itaporã/MS (ITA)	22°00'09.68"	54°42'52.59"	NT0150	C. fi		8		6 (1)	
Ponta Porã/MS (POP)	22°32'58.11"	55°43'01.85"	NT0150	C. fi		5		5 (0)	
Amaporã/PR (AMA)	23°05'07.95"	52°47'46.31"	NT0150	C. fi		6		6 (3)	
BR - Umuarama/PR (BRP)	23°39'08.85"	53°22'13.13"	NT0150	C. fi		5		4 (0)	
Engenheiro Beltrão/PR (EBE)	23°48'19.16"	52°15'41.92"	NT0150	C. fi	C. od	3	2	3 (1)	2 (1)
Londrina/PR (LON)	23°09'43.29"	51°09'55.15"	NT0150	C. fi	C. od	3	1	3 (3)	1 (1)
Porto Camargo - Icaraíma/PR (PCA)	23°24'02.81"	53°37'56.58"	NT0150	C. fi		2		2 (0)	
Palotina/PR (PLO)	24°18'34.29"	53°54'32.15"	NT0150	C. fi	C. od	4	8	4 (1)	7 (0)
São Pedro do Iguaçu/PR (SPI)	24°55'59.58"	53°50'52.95"	NT0150	C. fi		3		4 (0)	
Três Barras do Paraná/PR (TBA)	25°25'45.56"	53°07'37.28"	NT0150	C. fi	C. od	4	6	4 (2)	7 (1)
Santa Maria/RS (SMA)	29°43'03.94"	53°43'29.04"	NT0150	C. fi		5		4 (1)	
Guarabira/PB (GUA)	06°50'23.25"	35°29'23.73"	NT0151	C. fi		2		2 (2)	

Trajano de Moraes/RJ (TRA)	22°05'12.80"	42°02'06.00"	NT0160		C. od	2	6 (5)
Januária/MG (JAN)	15°11'22.56"	44°12'24.48"	NT0202	C. fi	0	1 (1)	
Manga/MG (MAN)	14°47'27.96"	43°56'19.92"	NT0202	C. fi	1	1 (1)	
Diamantina/MG (DIA)	18°24'33.60"	43°29'24.70"	NT0704	C. fi	5	5 (3)	
Tiradentes/MG (TIR)	21°06'22.63"	44°10'44.71"	NT0704	C. fi	1	1 (1)	
Itápolis/SP (IPO)	21°34'55.57"	48°49'49.31"	NT0704	C. fi	1	1 (0)	
REGIÃO DE CHIQUITANO					87	86 (19)	
Nova Xavantina/MT (NOX)	14°40'19.79"	52°20'28.16"	NT0704	C. fi	4	3 (0)	
Cáceres/MT (CAR)	16°04'35.43"	57°36'14.96"	NT0212	C. fi	3	2 (1)	
Nova Lacerda/MT (COL)	09°51'40.59"	56°11'05.07"	NT0212	C. fi	4	4 (3)	
Comodoro/MT (COM)	13°41'26.93"	59°46'57.33"	NT0212	C. fi	5	5 (2)	
Figueirópolis D'oeste/MT (FIG)	15°26'43.80"	58°44'25.73"	NT0212	C. fi	6	5 (1)	
Juba - Lambari D'oeste/MT (JUB)	15°18'55.25"	57°59'26.71"	NT0212	C. fi	3	3 (0)	
Mozar - Salto do Céu/MT (MOZ)	15°07'48.33"	58°07'35.00"	NT0212	C. fi	1	1 (0)	
Pontes e Lacerda/MT (PLA)	15°13'52.38"	59°22'10.64"	NT0212	C. fi	7	8 (6)	
Rio Branco/MT (RIB)	15°14'33.37"	58°07'14.89"	NT0212	C. fi	1	1 (0)	
Soroteca - Curvelândia/MT (SOR)	15°36'23.87"	57°54'43.57"	NT0212	C. fi	5	5 (0)	
Vila Bela da Santíssima Trindade (VBE)	15°01'39.63"	59°57'16.27"	NT0212	C. fi	0	1 (1)	
San Ramón/Bolívia (SRA)	15°59'59.00"	61°40'00.00"	NT0212	C. fi	3	4 (1)	
Cármén Ruiz/Bolívia (RUI)	16°03'03.18"	61°05'21.95"	NT0212	C. fi	4	4 (0)	
Estancia San Juan /Bolívia (ESJ)	16°22'16.99"	60°57'25.61"	NT0212	C. fi	4	4 (0)	
INPA - Concepción/Bolívia (INP)	16°07'51.87"	62°00'19.71"	NT0212	C. fi	5	5 (0)	
Lagunilla - Concepción/Bolívia (LAG)	16°08'59.74"	62°03'16.71"	NT0212	C. fi	4	5 (0)	
Júnior - Alta Floresta/MT (AJU)	09°54'00.00"	55°53'59.00"	NT0135	C. fi	4	3 (1)	
Agrocondor - Alta Floresta/MT (AGO)	09°53'56.97"	56°09'16.12"	NT0135	C. fi	4	4 (1)	
Lucirene - Alta Floresta/MT (ALU)	09°52'23.67"	56°04'22.68"	NT0135	C. fi	4	4 (1)	
Monte Alegre - Alta Floresta/MT (MOA)	09°51'40.59"	56°11'05.07"	NT0135	C. fi	3	3 (0)	
Ponte do Frigorífico-Alta Floresta/MT (PFI)	09°52'59.91"	56°08'46.25"	NT0135	C. fi	3	3 (0)	
Sol Nascente - Alta Floresta/MT (SON)	10°06'07.72"	56°13'00.49"	NT0135	C. fi	3	3 (0)	
Cacoal/RO (CAC)	11°29'50.71"	61°15'04.45"	NT0135	C. fi	3	3 (0)	
Pimenta Bueno/RO (PBU)	11°40'56.40"	61°32'25.38"	NT0135	C. fi	3	2 (0)	
São Felipe D'oeste/RO (SFE)	11°52'40.92"	61°31'16.98"	NT0135	C. fi	1	1 (1)	
TOTAL					203	194 (59)	

Entre parênteses os espécimes que não apresentaram polimorfismo intraindividual ou foram submetidos à clonagem da região de ITS.

Montagem do conjunto de dados de cloroplasto

A região espaçadora entre os genes de RNA transportador de treonina e de leucina (*trnT-trnL*) do genoma de cloroplasto foi inicialmente amplificada utilizando os iniciadores A (5'-CATTACAAATGCGATGCTTCT-3') e D (5'-GGGGATAGAGGGACTTGAAC-3') descritos por Taberlet et al. (1991). Para as amostras cujas amplificações falharam ou os produtos de PCR possuíam baixa qualidade, uma nova amplificação do fragmento foi realizada utilizando duas reações independentes. A primeira reação foi conduzida com a combinação dos iniciadores A₂ (5'-CAAATGCGATGCTCTAACCT-3') de Cronn et al. (2002) e aC (5'-CGTAGCGTCTACCGATTTCG-3') complementar do iniciador C de Taberlet et al. (1991). Esse conjunto de iniciadores amplificou a região espaçadora intergênica entre o exon 5' de *trnT* (UGU) e *trnL* (UAA) (Taberlet et al., 1991). A segunda reação foi conduzida com os iniciadores C (5'-CGAAATCGGTAGACGCTACG-3') e D e amplificou o íntron *trnL* (UAA) (Taberlet et al., 1991). As reações da polimerase em cadeia (PCR) foram realizadas em um volume final de 25 µL, contendo: Tris-HCl 10 mM (pH8,3); 50 mM KCl; 0,1% de Triton X-100; 2,5 mM MgCl₂; 0,5 µM de cada iniciador; 0,2 mM de cada dNTP; 1,25 U de *Taq* DNA polimerase; 1,75 µl albumina sérica bovina (BSA); aproximadamente 60 ng de DNA e água Milli-Q.

Todas as amplificações foram conduzidas em termociclador GeneAmp PCR System 9700 (Applied Biosystems). Para amplificar a região espaçadora intergênica entre o exon 5' de *trnT-trnL* foi realizado um ciclo inicial de desnaturação a 94 °C por 5 minutos, seguido por 35 ciclos de 94 °C por 1 minuto, 60 °C por 1 minuto, 72 °C por 2 minutos e 1 ciclo de extensão final de 72 °C por 5 minutos. A região do íntron *trnL* foi amplificada com um ciclo inicial de desnaturação a 94 °C por 5 minutos, seguido por 35 ciclos de 94 °C por 1 minuto, 53 °C por 1 minuto, 72 °C por 1 minuto e 1 ciclo de extensão final de 72 °C por 5 minutos.

Os produtos de PCR foram purificados utilizando a enzima ExoSAP IT (USB) na proporção de 9µl de reação para 3µl da enzima, resultando num volume final de 12µl que foram incubados a 37°C por 15 minutos para degradar os remanescentes de iniciadores e nucleotídeos, e em seguida incubados a 80°C por 15 minutos para inativar a ExoSAP. O sequenciamento foi realizado pela companhia Sul Coreana Macrogen (www.macrogen.com) utilizando os mesmos iniciadores das reações de PCR.

As sequências foram importadas para o SEQUENCHER version 4.8 (Gene Codes Corp.) sendo editadas e corrigidas manualmente. O alinhamento completo das sequências foi realizado com a introdução de *gaps* para compensar a presença de *indels*. *Indels* inseridos em uma sequência de nucleotídeos do mesmo tipo (por exemplo, poli A) foram desconsiderados nas análises subsequentes. Esta decisão objetivou retirar da análise qualquer efeito que poderia ser resultado de erro experimental ou da instabilidade evolutiva associada a esse tipo de indel (ver Mast et al., 2001). O conjunto de dados utilizados nas análises do genoma cloroplastídico foi composto por sequências de 203 espécimes de *C. fissilis* e *C. odorata* (Tabela 1). O sequenciamento da região espaçadora *trnT-trnL* produziu dois fragmentos que se sobrepunham e que alinhados resultaram em uma sequência composta por 1500 pares de bases.

Montagem do conjunto de dados da região de ITS

A região completa dos espaçadores internos transcritos (ITS) dos genes nucleares de RNA ribossomal 18S-26S (que inclui o gene 5.8S) foi amplificada com auxílio dos iniciadores ITS4 (5'-TCCTCCGCTTATTGATATGC-3') de White et al. (1990) e ITS.LEU (5'-GTCCACTGAACCTTATCATTTAG-3') descrito por Baum et al. (1998). As reações de PCR foram realizadas em um volume final de 25 µL, contendo: 5µl do Tampão IVB-5X (Phoneutria) com 3,5mM de MgCl₂; 0,5 µM de

cada iniciador; 0,2 mM de cada dNTP; 1,25 U de *Taq* DNA polimerase; 1,25 µl de dimetil-sulfóxido (DMSO); aproximadamente 60 ng de DNA e água Milli-Q.

Todas as amplificações foram conduzidas em termociclador GeneAmp PCR System 9700 (Applied Biosystems). A região de ITS foi amplificada utilizando um ciclo inicial de desnaturação a 94 °C por 4 minutos, seguido por 35 ciclos de 94° C por 1 minuto, 49 °C por 1 minuto, 72 °C por 45 segundos e 1 ciclo de extensão final de 72 °C por 5 minutos. A enzima ExoSAP IT (USB) foi utilizada para purificar os produtos de PCR na proporção de 9µl de reação para 3µl da enzima. O sequenciamento foi realizado pela companhia Sul Coreana Macrogen (www.macrogen.com) utilizando os mesmos iniciadores das reações de PCR.

As sequências de ITS foram importadas para o SEQUENCHER version 4.8 (Gene Codes Corp.) para alinhamento e edição. A inspeção visual dos cromatogramas da região de ITS revelou que a maior parte das sequências apresentou ambiguidades, 170 dos 194 espécimes analisados apresentaram cromatogramas com picos duplos em sítios específicos de ambas as fitas complementares, sendo esses sítios flanqueados em ambos os lados por trechos de sequências limpas (Figura 2). Dentre esses 170 espécimes, 29 apresentaram um único pico duplo e 141 apresentaram múltiplos picos duplos. Seis espécimes que apresentaram múltiplos picos duplos foram aleatoriamente selecionados para investigação das ambiguidades. Para estes seis espécimes, o fragmento de ITS amplificado foi inserido em vetor p-GEM-T Easy (Promega) e utilizado para transformar *Escherichia coli* DH5α por choque térmico segundo Sambrook et al. (1989). A seleção das colônias transformadas foi realizada por marcação branco/azul com X-gal e a extração do DNA plasmidial contendo os fragmentos de interesse foi realizada utilizando o kit de purificação Wizard™ Mini Prep (Promega) conforme especificação do fabricante. Para cada espécime foram reenviados seis clones para sequenciamento e destes, sequências de 34 clones foram úteis para

análise. As sequências advindas dos 135 espécimes restantes, que apresentaram múltiplos picos duplos e não foram submetidas à clonagem, foram excluídas das análises subsequentes.

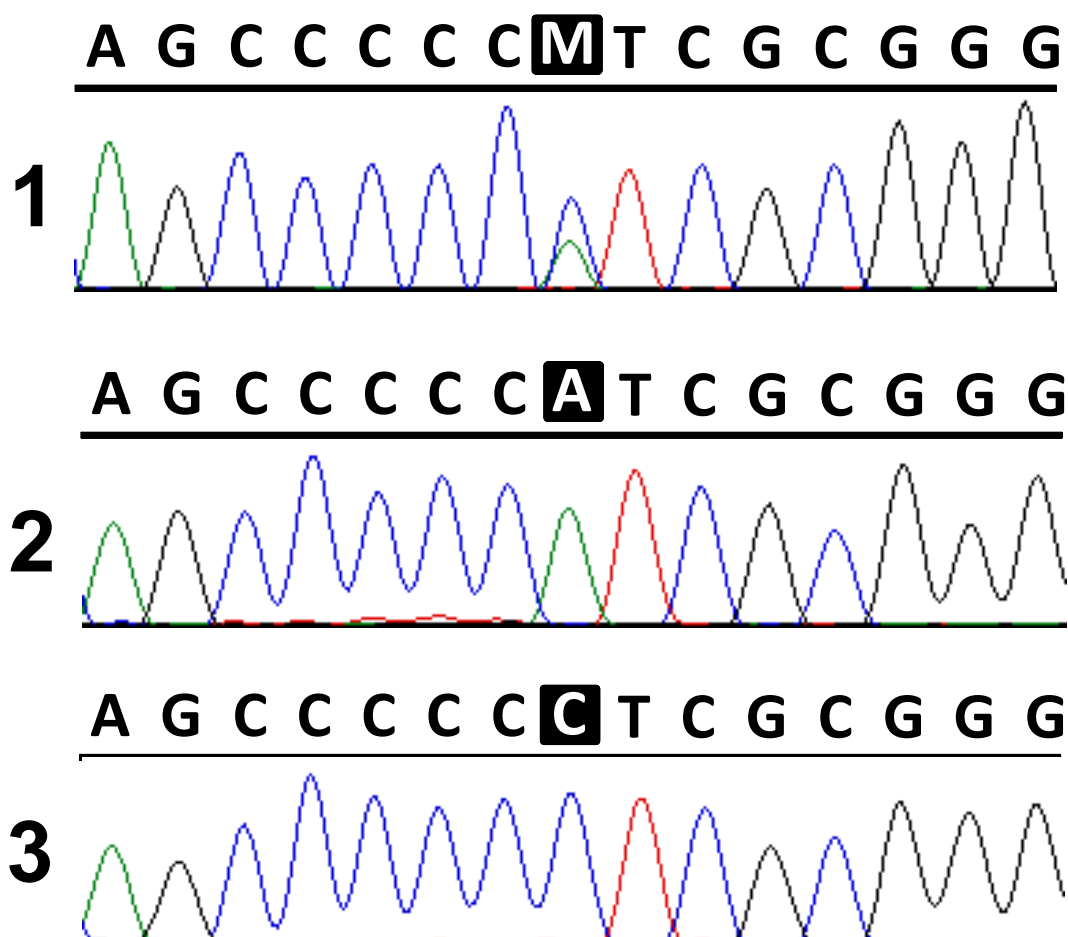


Figura 2. Cromatogramas parciais obtidos a partir de sequenciamento direto da região de ITS2 de três espécimes de *Cedrela fissilis* Vell. 1: sequência de PLA-281 com sobreposição de picos duplos resultando na leitura ambígua M, um indicativo de paralogia em ITS. 2 e 3: sequências de AJU_28 e DIA_182 respectivamente, com ausência de picos duplos no sítio correspondente.

Os limites das duas regiões espaçadoras internas transcritas (ITS1 e ITS2) e do gene 5.8S foram estabelecidos consultando três bancos de dados: “Rfam” (Gardner et al., 2008), “The European Ribosomal Database” (Wuyts e Van de Peer, 2004) e “The ITS2 database” (Selig et al., 2008). Em seguida, as regiões flanqueadoras dos rDNAs 18S e 28S foram removidas e os sítios que apresentaram mais que uma mudança foram desconsiderados nas análises subsequentes, uma vez que o modelo de sítios infinitos de Kimura (1969) foi assumido. Este modelo afirma que o número de sítios em uma sequência de DNA é tão grande que uma mutação recém-surgida ocorrerá em um local diferente dos locais onde as mutações anteriores ocorreram (Kimura, 1969).

O conjunto de dados de ITS apresentou sequências que variaram de 585 a 636pb, sendo essa variação atribuída à presença de três *indels* na região ITS1: um *indel* de 36bp (em dois clones do espécime PLA_75), um *indel* de 48bp (no espécime VNI_38) e um *indel* de 51pb (no espécime AJU_28). Somados esses indels se estendiam por 85 sítios e se sobrepunham a 11 sítios contendo substituições, decidimos então remover as quatro sequências que apresentavam os *indels* e considerar a informação presente nos sítios com substituição.

Para a construção do conjunto de dados utilizado nas análises da região de ITS, as sequências que continham um único sítio com pico duplo tiveram sua ambiguidade resolvida, ou seja, cada uma das 29 sequências que apresentavam um único pico duplo deu origem a duas novas sequências idênticas a sequência original, exceto no sítio ambíguo, onde cada sequência apresentou um dos dois nucleotídeos possíveis para o respectivo código de ambiguidade. Dessa forma, o conjunto de dados foi composto pelas 22 sequências livres de pico duplo, pelas 58 sequências originadas dos 29 espécimes que apresentavam um único sítio com pico duplo e por 32 sequências obtidas pelo sequenciamento dos clones, totalizando 112 sequências advindas de 57 espécimes.

Índices de diversidade nucleotídica

O programa ARLEQUIN v3.0 (Excoffier et al., 2005) foi utilizado para estimar o número de haplótipos (H), a diversidade haplotípica (H_d), a diversidade nucleotídica (π), o número de sítios polimórficos (S) e para realizar os testes de neutralidade pelos métodos D de Tajima e F_s de Fu's a nível de ecorregião e de região de distribuição para os conjuntos de dados de cpDNA e da região de ITS.

Estimativa das redes de haplótipos

O programa DnaSP 5.10 (Librado e Rozas, 2009) foi utilizado para determinação dos haplótipos presentes nos dois conjunto de dados, ou seja, entre as 203 sequências de cpDNA e entre as 112 sequências da região de ITS. A genealogia de genes foi inferida utilizando o método de "*Median Joining (MJ) network*" (Bandelt et al., 1999) como implementado no aplicativo Network 4.5.1.6 (<http://www.fluxus-technology.com>). O aplicativo considerou o *indel* como sendo um quinto estado de caráter e cada *indel*, independente do tamanho, foi considerado como uma mutação única.

RESULTADOS

Haplótipos da região espaçadora trnT-trnL do genoma de cloroplasto

O alinhamento dos 1500 pares de bases das sequências revelou a presença de 22 sítios polimórficos que foram caracterizados por *indels* ou substituições de bases (Figura 3). Foram observadas dezessete substituições, 11 do tipo transição de bases e 6 do tipo transversão. Os cinco sítios polimórficos restantes apresentaram *indels* de 5, 7, 21, 22 e 1 par de bases. A análise dos 181 espécimes de *Cedrela fissilis* e dos 22 espécimes de *Cedrela odorata* revelou a existência de dezenove haplótipos (Figura 3). A presença de *indels* resultou na variação no tamanho da sequência dos haplótipos obtidos, que variaram de 1457 a 1495 pares de bases. A maior parte dos haplótipos (A, B, C, D, F, G, H, I, L, P, Q e R) apresentou sequência de 1474 pares de bases. O haplótipo M apresentou a sequência mais longa (1495 pb), com uma inserção de 21 pb entre as posições 317 e 337. Em contrapartida, o haplótipo S apresentou a sequência mais curta (1457 pb), com uma deleção de 22 pb entre as posições 421 e 442. Os haplótipos E, J e K apresentaram sequências de 1467 pares de bases. Os haplótipos N e O, apresentaram sequências de 1473 e 1478 pares de bases, respectivamente.

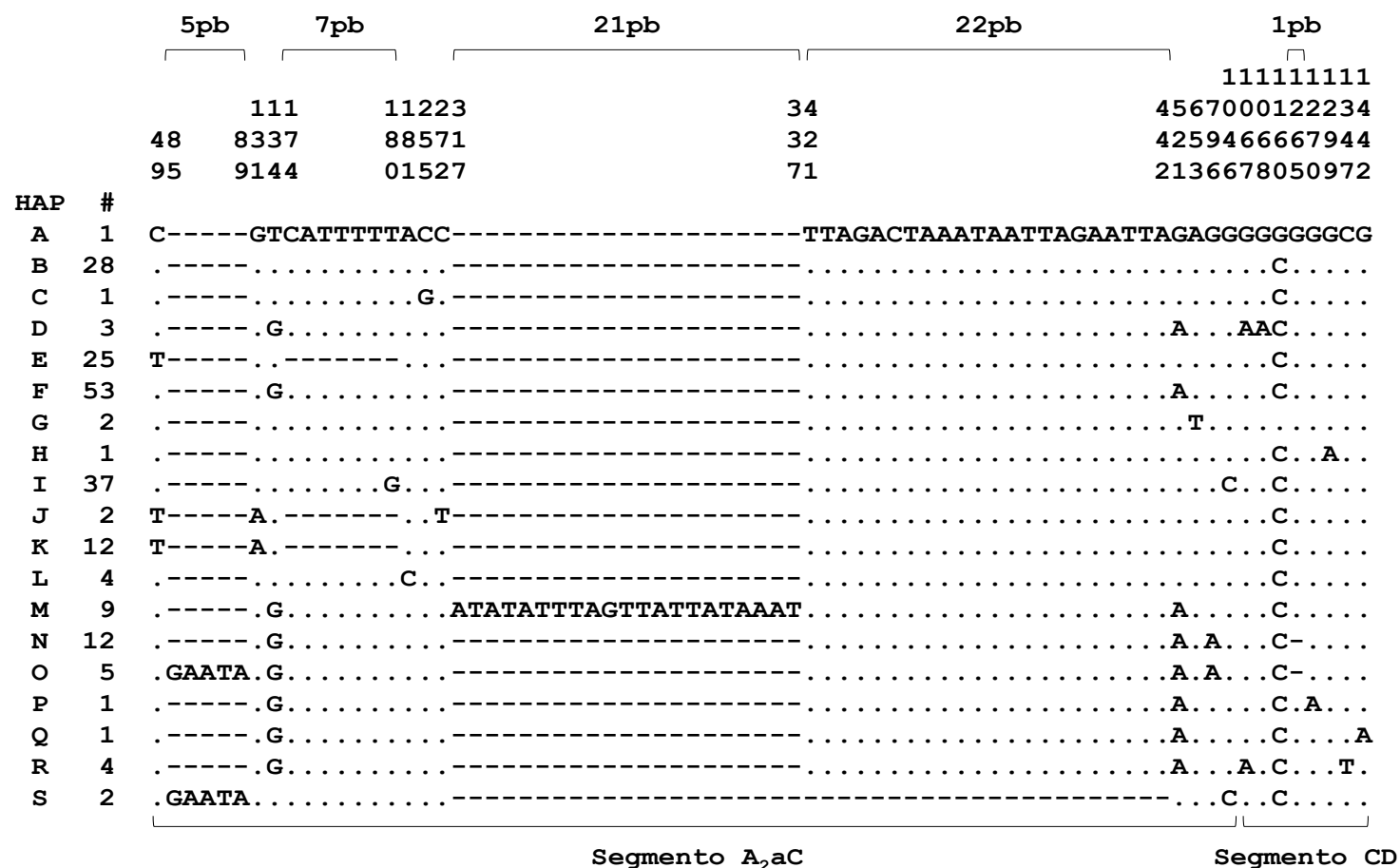


Figura 3. Alinhamento dos sítios polimórficos da região espaçadora *trnT-trnL* do genoma de cloroplasto que definem os dezenove haplótipos em *Cedrela fissilis* e *Cedrela odorata*. Cada fragmento combina 890 bp da região espaçadora intergênica entre o éxon 5' do *trnT* (UGU) e *trnL* (UAA) (segmento “A₂aC”), e 610 bp do intron do *trnL* (segmento “CD”). Os números acima do alinhamento representam a posição do sítio polimórfico no alinhamento. Os pontos indicam similaridade ao Haplótipo A; cada traço indica um *gap* e (#) indica o número de ocorrências de cada haplótipo entre as sequências analisadas.

Haplótipos da região de ITS do genoma nuclear

O alinhamento dos 636 pares de bases do conjunto de sequências revelou a presença de 38 sítios polimórficos caracterizados por substituição (Figura 4). Dezenove deles foram caracterizados como substituição do tipo transição de bases e as dezenove substituições restantes foram do tipo transversão. A distribuição dos sítios polimórficos entre os espaçadores indicou que 26 destes sítios pertenciam à região de ITS-1 e 12 à região de ITS-2, sendo que a região do rDNA 5.8S não apresentou nenhum sítio polimórfico.

A análise dos 47 espécimes de *Cedrela fissilis* e dos 10 espécimes de *Cedrela odorata* resultou na obtenção de 112 sequências (96 de *C. fissilis* e 16 de *C. odorata*), dentre as quais, 80 foram originadas por sequenciamento direto e 32 por clonagem. Foram encontrados 54 haplótipos entre as 112 sequências analisadas para a região de ITS (Figura 4). Não houve variação no tamanho das sequências entre os haplótipos encontrados para a região de ITS uma vez que as quatro sequências que apresentaram os *indels* de 36, 48 e 51 pares de bases foram excluídas da análise, conforme mencionado anteriormente. Cabe ressaltar que os três *indels* apresentados estavam inclusos na região de ITS-1 que também abrigou o maior número de sítios polimórficos.

		1111111111122222	444444455556
		23344667781112225789903345	234457867771
		72634077814681254063485730	948935345685
HAP	#/@		
1	4/0	CCGCACGCTCTTGCACATGCCGCC	CCTCCGCTTCCC
2	1/0		...T.....
3	3/0	..C.G.....G.....T.	
4	3/0	...G.....G.....T.	
5	2/0	...G...G.....G.....	
6	2/1	...G...G...T.CG.....	
7	2/0	G.....	
8	5/0	...G.....G.....	
9	1/0	...G...G...T.CG...TAAT.	
10	1/0	...G...G...T.CG...TA.T.	
11	6/4	...G...G...T.CGTG...TA.T.	
12	0/1	..C.G...G...T.CGTG...TA.T.	...A.....
13	0/1	...G...G...T.CGT.....	
14	1/1	..C.G.....G.....	
15	0/1	..C.G.....G.....	...A.....
16	1/0	.A..G...G...T.CGT...TA.T.	
17	1/0	...G...G...T.CGT...TA.T.	
18	1/0	...G...G...T.CGTG...TA.TT	
19	1/0	..C.G.....T.G.....T.	
20	1/0	...GT..G.....G.....	
21	1/0	...G.....CC...G.....	
22	1/0	..C.G...G.....CGT.C....T.	C....
23	1/0	..C.G...G.....GT.C....T.	C....
24	3/0	...G...G...T.CGTG...TA.T.	G
25	2/2	...G...G...T.CG.....	G.....
26	5/0	...G...G.....	
27	1/0	...G.....	
28	1/0	A.....	
29	2/0		GGA.
30	1/0	...G...G.....	G
31	1/0	T.....	
32	1/0	...G.....T.....	
33	0/1	...G...G...T.CGTG...TA.T.	.A...A....
34	0/1	...G.CGG.....T...G.A.T.	..C...A....
35	0/1	...G...G...T.CGTG...TA.T.	..C...A....
36	0/1	...G.CGG.....T...G.A.T.	
37	0/1	...G...G...T.CGT...A.T.	..C...A....
38	0/1	..C.G.....GT.....	A....
39	0/1	..C.G.....G.....	.A...A....
40	0/1	..AG.....G.....	
41	7/0	...G..GG.....T...G.A.T.	.A...A....
42	5/0	...G..GG.....T...G.A.T.	A....
43	1/3	...G..GG.....T...G.A.T.	..C...A....
44	4/2	...G..GG.....T..AG.A.T.	.A...A....
45	2/0	...G..GG.....T..AG.A.T.	A....
46	2/1	...G..GG.....T...A.T.	.A...A....
47	1/0	...G..GG.....T...G.A.T.	.A...AA....
48	2/0	...G..GG.....T...A.T.	A....
49	0/1	..AG...T...GT.....	A....
50	0/1	..AG...G.....T.....	GA...A....
51	0/1	..AG.....GT.....	.A...A....
52	0/1	...G..GG.....T..AG.A.T.	
53	0/2	...G..GG...T..GT.....	.A...A....
54	0/1	...G..GG.....T.....	A....

ITS - 1

ITS - 2

Figura 4. Alinhamento dos sítios polimórficos da região de ITS do genoma nuclear que definem os cinquenta e quatro haplótipos em *Cedrela fissilis* e *Cedrela odorata*. Cada fragmento inclui 636 pares de bases. Os números acima do alinhamento representam a posição do sítio polimórfico no alinhamento e os pontos indicam similaridade ao Haplótipo 1. O número de ocorrências de cada haplótipo entre as sequências obtidas por sequenciamento direto (#) ou clonagem (@) são indicadas.

Relação genealógica e geográfica entre as linhagens de cpDNA

A relação genealógica entre os haplótipos da região espaçadora *trnT-trnL* de *Cedrela fissilis* e *Cedrela odorata* foram apresentadas na rede de haplótipos apresentada na Figura 5-B. Na rede, a variação genética observada foi distribuída em linhagens distintas, organizadas em torno dos haplótipos mais frequentes, com os haplótipos de menor frequência encontrados na periferia da rede. Foram identificadas sete linhagens, denominadas haplogrupos, que estão representadas por cores distintas. A estrutura genética associada à distribuição geográfica dos haplogrupos evidenciou a diferenciação entre os haplogrupos presentes na região do Atlântico e na região de Chiquitano, com apenas dois haplótipos (F e I), em duas populações específicas (MAN e SRA), sendo compartilhados entre as regiões (Figura 5-A).

A rede separou os haplótipos da região do Atlântico em quatro haplogrupos pouco divergentes, separados entre si por um ou dois passos de mutação (Figura 5-B). No haplogrupo branco, o haplótipo B foi mais frequente, identificado em 28 dos 203 espécimes analisados. O haplótipo B ocupou posição central entre os demais haplogrupos da região do Atlântico e foi separado dos haplogrupos que integram a região de Chiquitano por apenas dois passos de mutação. Quanto à distribuição, o haplótipo B foi o mais amplamente distribuído, estando presente em seis (NT0101, NT0103, NT0104, NT0150, NT0160 e NT0704) das dez ecorregiões consideradas (Tabela 2). No haplogrupo amarelo o haplótipo mais frequente foi o I identificado em 37 espécimes, dentre os quais, 36 pertencem às florestas do interior do Paraná-Paraíba (ecorregião NT0150) localizada no limite sudoeste da região do Atlântico, e um único espécime (SRA_115) pertencente às florestas secas de Chiquitano (ecorregião NT0212). Além disso, o haplótipo S, também pertencente ao haplogrupo amarelo, foi identificado em apenas dois espécimes (SRA_113 e SRA_114) pertencentes à mesma população da ecorregião NT0212 da região de

Chiquitano. No haplogrupo laranja o haplótipo E identificado em 25 espécimes foi o mais frequente, seguido pelo haplótipo K, identificado em 12 espécimes e pelo haplótipo J, presente em 2 espécimes de uma única população (AMA) (Tabela 2). Quanto à distribuição, o haplogrupo laranja se concentrou nas florestas do interior do Paraná-Paraíba (ecorregião NT0150) e na ecorregião adjacente NT0101 caracterizada pelas florestas úmidas de Araucária. Um único espécime pertencente ao haplótipo E esteve presente na população FLE localizada no sudeste de Minas Gerais que integra as florestas do interior da Bahia (NT0104). O haplogrupo rosa formado unicamente pelo haplótipo G foi localizado no extremo Norte da região do Atlântico, em um fragmento de floresta estacional, as florestas litorâneas de Pernambuco (ecorregião NT0151) (Figura 5-A).

Na região de Chiquitano foram observados três haplogrupos pouco divergentes, separados entre si por dois passos de mutação (Figura 5-B). No haplogrupo preto o haplótipo F foi o mais frequente, identificado em 53 dos 203 espécimes analisados. Na rede, o haplótipo F apresentou posição central entre os demais haplogrupos da região de Chiquitano. Quanto à distribuição, o haplótipo F esteve presente na maioria das populações que integram as ecorregiões NT0135 (florestas úmidas do Madeira-Tapajós) e NT0212 (florestas secas de Chiquitano) (Tabela 2). Além disso, o haplótipo F esteve presente em um único espécime (MAN_284) amostrado no Norte de Minas Gerais, em uma zona de transição que inclui enclaves de floresta estacional (NT0202 – florestas secas do Atlântico), savana (Cerrado - NT0704) e savana estépica (NT1304 - Caatinga). O haplogrupo verde incluiu os haplótipos N (12 espécimes) e O (5 espécimes), e apresentou menor amplitude de distribuição, limitando sua ocorrência a seis populações da ecorregião NT0212 (florestas secas de Chiquitano) (Tabela 2 e Figura 5-A). Os membros do haplogrupo azul foram menos frequentes, limitando sua distribuição a duas populações NOX (haplótipo R, quatro espécimes) e PAL (haplótipo D, três

espécimes). A população NOX foi coletada em floresta de galeria de uma área de tensão ecológica que inclui elementos de floresta estacional e savana, inserida na ecorregião NT0704 (Cerrado), que se encontra geograficamente mais próxima às populações de Chiquitano. A população PAL está inserida em um fragmento isolado de floresta estacional localizado no Brasil Central, que está cercado por uma vegetação de savana (Cerrado - NT0704). Este fragmento disjunto faz parte das florestas do interior do Paraná-Paraíba (NT0150), que integra a região do Atlântico. Cabe ressaltar que essas populações foram geograficamente localizadas na região de limite entre as regiões do Atlântico e de Chiquitano (Figura 1).

Por fim, observamos que 11 dos 19 haplótipos encontrados para o cpDNA foram encontrados em uma única população (A, C, D, G, H, J, L, P, Q, R e S), sendo que os haplótipos A, C, H, P e Q, foram representados por um único espécime (Tabela 2). De maneira geral, a ecorregião NT0150 do Atlântico e a ecorregião NT0212 de Chiquitano, abrigaram o maior número de espécimes, 75 e 55, respectivamente. Além disso, essas ecorregiões abrigaram o maior número de haplótipos de cpDNA, 7 em cada uma das ecorregiões, dentre os quais cinco foram exclusivos para cada ecorregião e 2 deles (F e I) compartilhados entre essas ecorregiões.

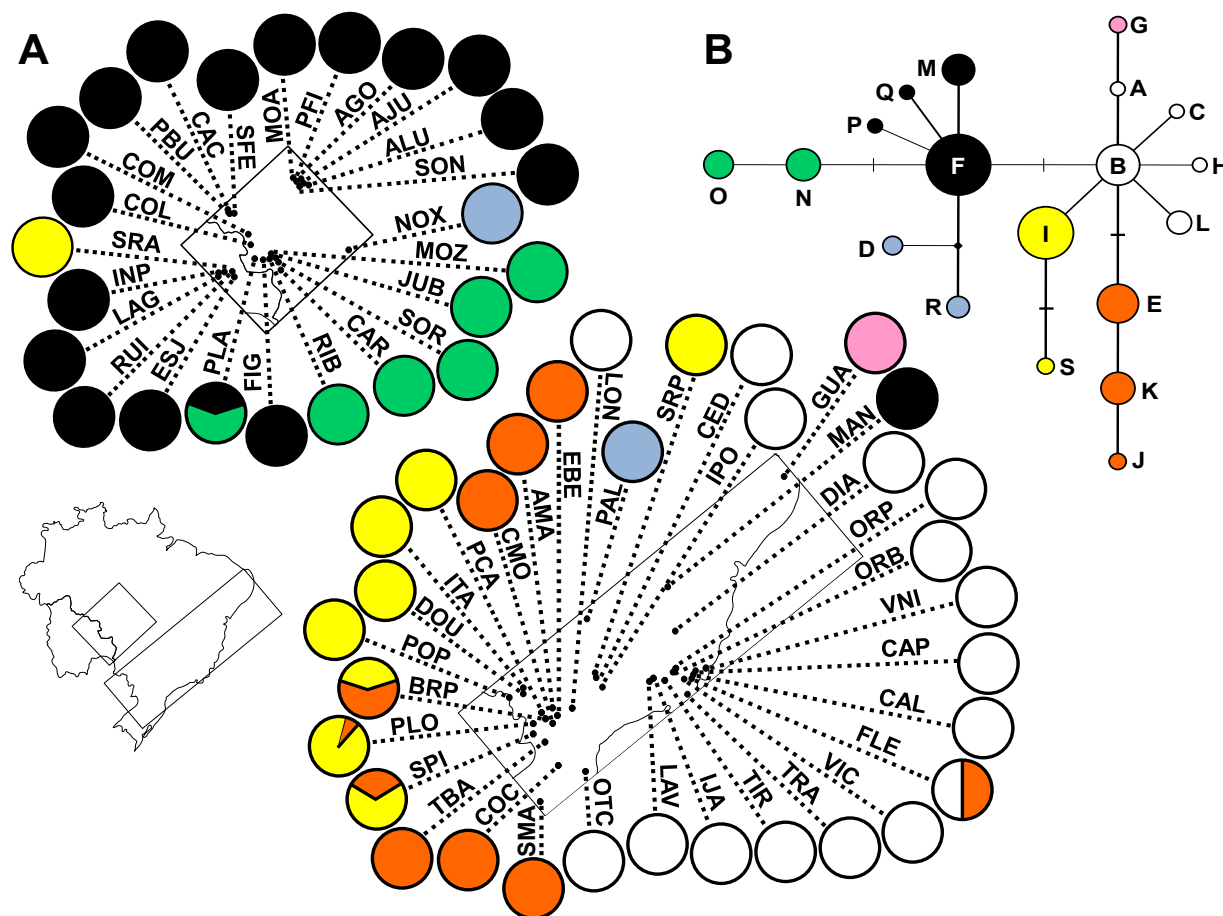


Figura 5. A: Distribuição geográfica dos 19 haplótipos de *trnT-trnL* de *Cedrela fissilis* Vell. e *Cedrela odorata* L. nas regiões do Atlântico e de Chiquitano. Cada círculo representa uma população; as cores dentro do círculo denotam os haplogrupos. Populações estão representadas por códigos de três letras. **B:** Rede de haplótipos *Median-joining* mostrando as relações entre as linhagens genealógicas em *C. fissilis* e *C. odorata*, com base nos haplótipos de *trnT-trnL*. Círculos representam haplótipos (codificados por letras), o tamanho do círculo é proporcional à frequência relativa do haplótipo e as cores dentro dos círculos denotam os haplogrupos. Números de substituições são indicados por barras quando mais do que uma substituição ocorreu.

Tabela 2. Distribuição dos 19 haplótipos referentes à região espaçadora *trnT-trnL* do genoma de cloroplasto obtidos a partir da análise de 203 sequências de *Cedrela fissilis* Vell. e *Cedrela odorata* L. O número de populações, de sequências analisadas (N) e o número total de haplótipos (NH) encontrados por ecorregião são indicados.

Ecorregiões	Populações	N	Distribuição dos haplótipos																			NH	
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	19	
NT0101	3	9		2			6						1									3	
NT0103	1	4	1	2	1																	3	
NT0104	8	16		15			1															2	
NT0135	9	28						19							9							2	
NT0150	15	75		1		3	18				36	2	11	4								7	
NT0151	1	2							2													1	
NT0160	1	2		2																		1	
NT0202	1	1						1														1	
NT0212	14	55						33			1					12	5	1	1		2	7	
NT0704	4	11		6						1										4		3	
TOTAL	57	203	1	28	1	3	25	53	2	1	37	2	12	4	9	12	5	1	1	4	2		

Relação genealógica e geográfica entre as linhagens de ITS

A segunda rede, construída a partir do conjunto de dados da região de ITS de *Cedrela fissilis* e *Cedrela odorata*, apresenta as relações genealógicas entre os 54 haplótipos encontrados (Figura 6-A). Na rede, a variação genética observada foi distribuída em três linhagens distintas, denominadas ribogrupos, que estão representadas por cores distintas. A estrutura genética associada à distribuição geográfica dos ribogrupos também evidenciou a diferenciação entre os ribogrupos presentes na região do Atlântico e na região de Chiquitano, com um único haplótipo sendo compartilhado entre as regiões (Figura 6-B e Tabela 3).

Os haplótipos da região do Atlântico foram divididos em dois ribogrupos, separados entre si por três passos de mutação. O ribogrupa amarelo foi organizado em torno do haplótipo 11, que apresentou a maior frequência, sendo identificado em 10 das 112 sequências analisadas. Quanto à distribuição, o Haplótipo 11 apresentou maior amplitude, estando presente em quatro ecorregiões da região do Atlântico (NT0101, NT0104, NT0150 e NT0704). Adicionalmente, o haplótipo 11 foi o único compartilhado entre as regiões do Atlântico e de Chiquitano, sendo também identificado em dois clones de um único espécime (SRA_113) pertencente a uma população das florestas secas de Chiquitano (ecorregião NT0212) (Tabela 3). O ribogrupa branco foi o mais numeroso apresentando 30 haplótipos distintos, entretanto, esses haplótipos foram pouco frequentes, sendo cada haplótipo identificado em no máximo 5 espécimes. Na rede, o ribogrupa branco assumiu posição central e foi conectado ao ribogrupa preto, característico da região de Chiquitano, por três passos de mutação (Figura 6-A). Quanto à distribuição, o haplogrupa branco foi encontrado em todas as ecorregiões pertencentes à região do Atlântico (NT0101, NT0103, NT0104, NT0150, NT0151, NT0160, NT0704), exceto nas florestas secas do Atlântico (ecorregião NT0202). Adicionalmente, dois clones do espécime PLA_75 e um clone do espécime SRA_113, pertencentes a

duas populações das florestas secas de Chiquitano (ecorregião NT0212) apresentaram os haplótipos 49, 50 e 51, respectivamente, pertencentes ao ribogruppo branco (Figura 6-B).

Os haplótipos da região de Chiquitano foram agrupados em um único ribogruppo, o ribogruppo preto, distribuído nas florestas úmidas do Madeira-Tapajós (ecorregião NT0135) e nas florestas secas de Chiquitano (ecorregião NT0212). Os haplótipos mais frequentes nesse ribogruppo foram o haplótipo 41, o 42 e o 44 que incluíram sete, cinco e seis sequências, respectivamente (Figura 6-A e Tabela 3). Adicionalmente, dois clones do espécime ITA_124 pertencente a uma população localizada no limite ocidental da região do Atlântico (ecorregião NT0150) apresentaram haplótipos 34 e 36, pertencentes ao ribogruppo preto, característico da região de Chiquitano (Figura 6-B). Entre os 54 haplótipos obtidos para a região de ITS, 42 haplótipos foram exclusivos, cada um, para uma única ecorregião, e destes, 33 três haplótipos foram representados por uma única sequência (Tabela 3).

Assim como observado na distribuição dos haplótipos de cpDNA, as florestas do interior do Paraná-Paraíba (ecorregião NT0150) localizadas na região do Atlântico e as florestas secas de Chiquitano (ecorregião NT0212) pertencentes à região de Chiquitano, incluíram o maior número de sequências, 35 e 30 respectivamente. Além disso, essas ecorregiões abrigaram o maior número de haplótipos, 20 para a ecorregião NT0150 e 13 para a ecorregião NT0212, sendo a maioria dos haplótipos exclusiva para sua respectiva ecorregião (Tabela 3). Além disso, cabe ressaltar que independente do ribogruppo analisado, a maior parte dos clones foram localizados nas extremidades da rede, ou nas regiões intermediárias entre os três ribogruppos.

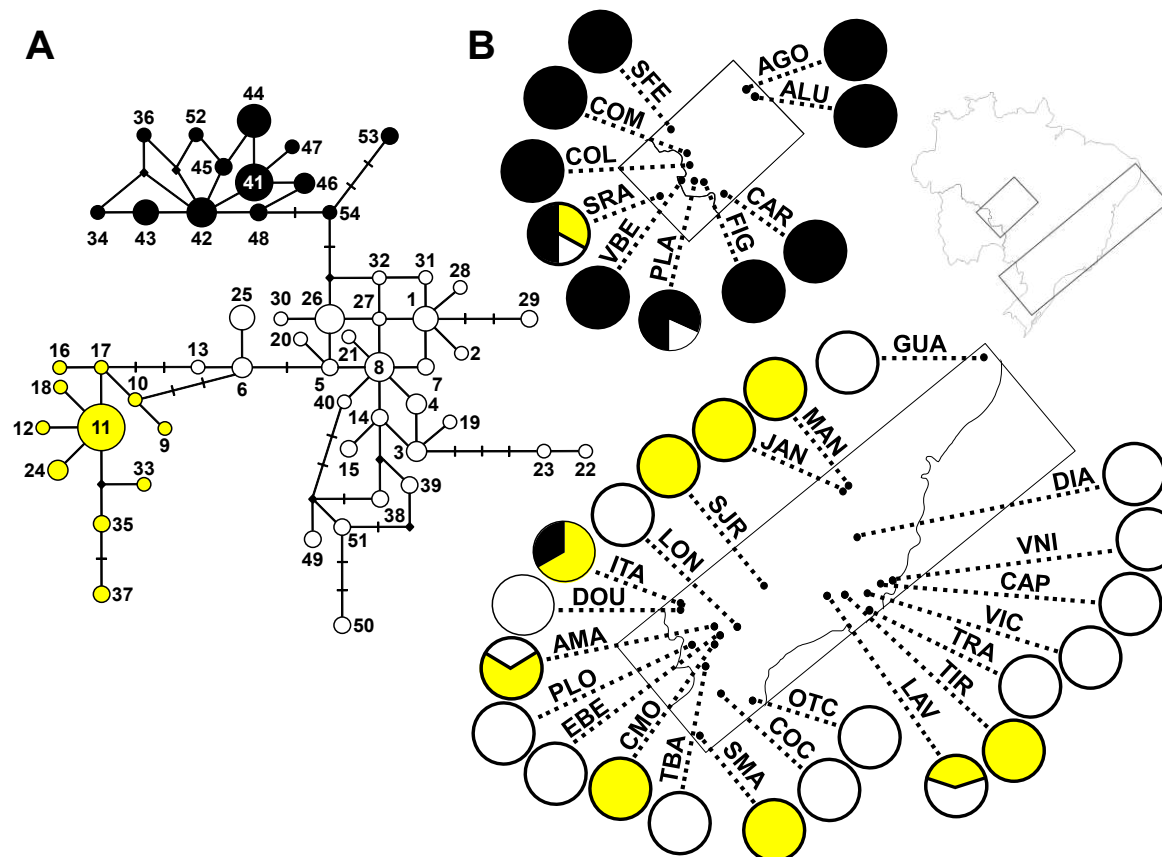


Figura 6. A: Rede de haplótipos *Median-joining* mostrando as relações entre as linhagens genealógicas em *Cedrela fissilis* Vell. e *Cedrela odorata* L., com base nos haplótipos de ITS. Círculos representam haplótipos (codificados por números), o tamanho do círculo é proporcional à frequência relativa do haplótipo, as cores dentro do círculo denotam os ribogrupos. Números de substituições são indicados por barras quando mais do que uma substituição ocorreu. **B:** Distribuição geográfica dos 54 haplótipos de ITS de *Cedrela fissilis* e *Cedrela odorata* nas regiões do Atlântico e de Chiquitano. Cada círculo representa uma população; as cores dentro do círculo denotam os ribogrupos; Populações estão representadas por códigos de três letras.

Tabela 3. Distribuição dos 54 haplótipos referentes à região de ITS do genoma nuclear, obtidos a partir da análise de 112 sequências de *Cedrela fissilis* Vell. e *Cedrela odorata* L. O número de populações, de espécimes e de sequências analisadas (N/S) e o número total de haplótipos (NH) encontrados por ecorregião são indicados.

Eco	NT0101	NT0103	NT0104	NT0135	NT0150	NT0151	NT0160	NT0202	NT0212	NT0704	Total
N/S	4/6	2/3	3/8	3/9	17/35	2/3	5/8	2/4	15/30	4/6	57/112
Pop	3	1	3	3	9	1	1	2	7	2	32
Distribuição dos haplótipos	1	2			1		1				4
	2	1									1
	3		1				2				3
	4		1				2				3
	5					1				1	2
	6				2					1	3
	7	1								1	2
	8	4								1	5
	9							1			1
	10							1			1
	11	1	1		5				2	1	10
	12		1								1
	13		1								1
	14		1				1				2
	15		1								1
	16							1			1
	17							1			1
	18									1	1
	19		1								1
	20					1					1
	21					1					1
	22						1				1
	23						1				1
	24				3						3
	25				4						4
	26				5						5
	27				1						1
	28				1						1
	29				2						2
	30				1						1
	31				1						1
	32				1						1
	33				1						1
	34				1						1
	35				1						1
	36				1						1
	37				1						1
	38				1						1
	39				1						1
	40				1						1
	41			2					5		7
	42			1					4		5
	43			3					1		4
	44								6		6
	45								2		2
	46								3		3
	47								1		1
	48								2		2
	49								1		1
	50								1		1
	51								1		1
	52								1		1
	53			2							2
	54			1							1
NH	3	2	8	5	20	3	6	4	13	6	

Compartilhamento de haplótipos entre *Cedrela fissilis* e *Cedrela odorata*

Foi observado que tanto para o conjuntos de dados de cpDNA quanto para ITS as sequências de *C. odorata* não foram restritas a um único haplótipo, nem a um grupo de haplótipos exclusivo para a espécie. Em função de sua identidade nucleotídica, sequências de *C. fissilis* e *C. odorata* compartilharam os haplótipos B, E, I, K e L de cpDNA e os haplótipos 1, 3, 4, 7, 8, 11, 14, 22, 23, 26, 28 e 30 de ITS.

Não foram obtidas redes de haplótipos espécie-específicas para *C. fissilis* e *C. odorata* para nenhum dos conjuntos de dados analisados, cpDNA e ITS. Em ambos os casos, as sequências das duas espécies foram agrupadas em uma única rede.

Índices de diversidade nucleotídica

Para o conjunto de dados da região cloroplastídica, os valores de diversidade haplotípica variaram consideravelmente entre as ecorregiões analisadas (variando de 0 a 1). A diversidade nucleotídica dentro das ecorregiões foi baixa, variando de 0 a 0.0015 (Tabela 4). Considerando as regiões de distribuição, os valores de diversidade haplotípica e diversidade nucleotídica foram mais elevados para a região do Atlântico (0.7924 e 0.0014) do que para a região de Chiquitano (0.6137 e 0.0010). Quanto aos testes de neutralidade seletiva realizados para o conjunto de dados da região cloroplastídica, o D de Tajima foi negativo e não significativo para a maior parte dos grupos analisados. Valores negativos do teste Fs de Fu foram obtidos para as regiões do Atlântico e de Chiquitano e ecorregiões NT0103 e NT0212, entretanto esses valores não foram significativos (Tabela 4).

Para o conjunto de dados da região de ITS, a diversidade haplotípica foi bastante semelhante e elevada para todos os grupos analisados, variando de 0.6000 a 1.0000. A diversidade nucleotídica dentro das ecorregiões variou de 0.0010 a 0.0100 (Tabela 5). Os valores de diversidade haplotípica foram elevados

em ambas as regiões do Atlântico (0.9711) e Chiquitano (0.9204). Entretanto, a diversidade nucleotídica variou bastante entre as regiões, apresentando valores mais elevados para a região do Atlântico (0.0090) em relação à de Chiquitano (0.0056). Para o conjunto de dados da região de ITS, o teste D de Tajima foi negativo e para as regiões do Atlântico e de Chiquitano e para as ecorregiões NT0160 e NT0212, entretanto valores significativos foram obtidos somente para a ecorregião NT0101 (-1.4228) (Tabela 5). Os valores do teste Fs de Fu foram negativos e significativos para a região do Atlântico (-24.9466) e ecorregiões NT0104 (-3.8109), NT0150 (-5.7173), NT0202 (-2.1811) e NT0704 (-2.1345). Valores negativos e não significativos foram obtidos para a região de Chiquitano e ecorregiões NT0151, NT0160 e NT0212 (Tabela 5).

Tabela 4. Medidas de diversidade nucleotídica e testes de neutralidade para *Cedrela fissilis* Vell. e *Cedrela odorata* L. por ecorregião amostrada, estimadas a partir da região espaçadora *trnT-trnL* do genoma de cloroplasto.

Grupos	Tamanho da amostra (N)	Número de haplótipos (H)	Diversidade haplotípica (H_d)	Diversidade nucleotídica (π)	Número de sítios polimórficos(S)	Tajima's D	Fu's Fs
Todos	203	19	0.8577	0.0019	22	-0.6406 (P=0.292)	-3.4136 (P=0.159)
Atlântico	116	12	0.7924	0.0014	14	-0.9617 (P=0.191)	-1.8977 (P=0.242)
NT0101	9	3	0.5556	0.0007	3	-0.5832 (P=0.309)	0.3512 (P=0.508)
NT0103	4	3	0.8333	0.0007	2	-0.7099 (P=0.282)	-0.8873 (P=0.091)
NT0104	16	2	0.1250	0.0002	2	-1.1622 (P=0.151)	0.1769 (P=0.309)
NT0150	75	7	0.6944	0.0015	10	-0.1233 (P=0.509)	1.3324 (P=0.774)
NT0151	2	1	0	0	0	0.0000 (P=1.000)	-
NT0160	2	1	0	0	0	0.0000 (P=1.000)	-
NT0202	1	1	1.0000	0	0	0.0000 (P=1.000)	-
NT0704	11	3	0.6182	0.0015	5	1.1471 (P=0.852)	2.4336 (P=0.899)
Chiquitano	87	9	0.6137	0.0010	12	-1.3181 (P=0.080)	-1.6003 (P=0.226)
NT0135	28	2	0.4524	0.0003	1	0.0000 (P=1.000)	1.4050 (P=0.672)
NT0212	55	7	0.5926	0.0011	9	-0.9305 (P=0.210)	-0.3834 (P=0.465)

Tabela 5. Medidas de diversidade nucleotídica e testes de neutralidade para *Cedrela fissilis* Vell. e *Cedrela odorata* L. por ecorregião amostrada, estimadas a partir da região de ITS do genoma nuclear.

Grupos	Tamanho da amostra (N)	Número de haplótipos (H)	Diversidade haplotípica (H_d)	Diversidade nucleotídica (π)	Número de sítios polimórficos(S)	Tajima's D	Fu's Fs
Todos	112	54	0.9757	0.0106	38	-0.1929 (P=0.494)	-24.9548** (P=0.000)
Atlântico	73	40	0.9711	0.0090	35	-0.6417 (P=0.292)	-24.9466** (P=0.000)
NT0101	6	3	0.6000	0.0047	9	-1.4228* (P=0.027)	1.8138 (P=0.810)
NT0103	3	2	0.6667	0.0010	1	0.0000 (P=0.988)	0.2007 (P=0.386)
NT0104	8	8	1.0000	0.0081	11	1.1020 (P=0.892)	-3.8109* (P=0.010)
NT0150	35	20	0.9479	0.0100	24	0.3328 (P=0.707)	-5.7173* (P=0.023)
NT0151	3	3	1.0000	0.0042	4	0.0000 (P=0.830)	-0.3409 (P=0.183)
NT0160	8	6	0.9286	0.0054	9	-0.0599 (P=0.492)	-1.3010 (P=0.150)
NT0202	4	4	1.0000	0.0026	3	0.1676 (P=0.733)	-2.1811* (P=0.013)
NT0704	6	6	1.0000	0.0079	10	0.9422 (P=0.828)	-2.1345* (P=0.049)
Chiquitano	39	15	0.9204	0.0056	18	-0.5186 (P=0.336)	-4.0054 (P=0.054)
NT0135	9	5	0.8611	0.0052	7	1.3124 (P=0.929)	0.2793 (P=0.530)
NT0212	30	13	0.9149	0.0056	18	-0.7274 (P=0.258)	-3.2643 (P=0.089)

DISCUSSÃO

Para ambos os conjuntos de dados, cpDNA e ITS não foram obtidas redes de haplótipos espécie-específicas para *C. fissilis* e *C. odorata*, sendo as sequências das duas espécies agrupadas em uma única rede. A obtenção de uma única rede de haplótipos para explicar a distribuição geográfica das linhagens genealógicas de *C. fissilis* de *C. odorata* demonstra que a similaridade das sequências nucleotídicas entre espécimes de *C. fissilis* e espécimes de *C. odorata* pertencentes à região do Atlântico foram maiores do que a similaridade existente entre as sequências nucleotídicas de espécimes de *C. fissilis* pertencentes às regiões do Atlântico e de Chiquitano. Adicionalmente, as sequências nucleotídicas de *C. odorata* não foram restritas a um único haplótipo, nem a um grupo de haplótipos exclusivo para a espécie, ao contrário, as sequências de *C. odorata* foram idênticas a algumas sequências nucleotídicas de *C. fissilis* e o agrupamento destas parece ter-se dado em função da distribuição geográfica. Entretanto, uma investigação mais detalhada a respeito da relação entre *C. fissilis* e *C. odorata* faz-se necessária, e para isso a amostragem em *C. odorata* deve ser ampliada e espécimes de *C. odorata* da região de Chiquitano devem ser incluídos.

Linhagens ancestrais

As análises filogeográficas fundamentadas nos genomas cloroplastídico e nuclear de *C. fissilis* e *C. odorata* evidenciaram a existência de sete haplogrupos para cpDNA (4 para região do Atlântico e 3 para Chiquitano) e apenas três ribogrupos para ITS (2 para região do Atlântico e apenas 1 para Chiquitano). Embora o número de grupos identificados pelos dois genomas seja distinto eles são

equivalentes, e apoiam a existência de níveis mais elevados de diferenciação genética na região do Atlântico, se comparado à região de Chiquitano.

Os altos níveis de diferenciação genética e a evidente divisão das linhagens entre as regiões do Atlântico e de Chiquitano permitiram inferir que as populações atuais de *C. fissilis* e *C. odorata* descendem de pelo menos duas linhagens ancestrais distintas. Segundo Avise (2000), a presença de linhagens espacialmente restritas e separadas por uma quantia considerável de passos mutacionais distingue as linhagens alopátricas ancestrais em uma árvore de genes. Uma separação evidente foi observada entre as linhagens do Atlântico e de Chiquitano, as quais compartilham apenas dois haplótipos de cpDNA (haplótipos I e F) e um haplótipo de ITS (haplótipo 11) em populações específicas localizadas em regiões ecotonais. Adicionalmente, as linhagens ancestrais das regiões do Atlântico e de Chiquitano estão separadas entre si por 2 e 3 (para cpDNA e ITS, respectivamente) haplótipos intermediários (hipotéticos), que podem ter sido extintos durante os eventos de expansão e contração das populações no processo de colonização e/ou estão em uma frequência tão baixa que não puderam ser amostrados. Dessa forma, a estrutura filogeográfica observada entre as linhagens ancestrais das regiões do Atlântico e de Chiquitano se deve, provavelmente, ao longo período de isolamento a que foram submetidas às populações dessas duas regiões.

Linhagens do Atlântico

A maior diversidade genética observada para *Cedrela fissilis* e *Cedrela odorata* encontra-se nas populações pertencentes à região do Atlântico. Uma região que apresenta altos índices de diversidade genética sugere: (i) que esta área pode ter sido um refúgio, ou seja, uma área que durante as flutuações climáticas manteve-se estável, favorecendo assim a manutenção da diversidade genética, ou

(ii) que esta região consiste em uma região intermediária que recebe imigrantes de refúgios adjacentes (Tzedakis et al., 2002; Carnaval e Moritz, 2008).

Estudos utilizando modelagem paleoambiental e registros palinológicos associados a padrões filogeográficos de espécies indicadoras, sugerem a existência de áreas historicamente estáveis na região do Atlântico que podem ter sido refúgios durante o Pleistoceno: (i) um corredor central ao longo da costa brasileira (que inclui o norte do Rio de Janeiro, Espírito Santo e Bahia), (ii) um pequeno refúgio no nordeste (Pernambuco), e (iii) um terceiro refúgio no sudeste (São Paulo) de tamanho intermediário e menos estável que os refúgios da porção central e norte do Atlântico (Carnaval e Moritz, 2008; Carnaval et al., 2009). Na porção central do Atlântico (NT0103, NT0104 e NT0160) o haplogrupo e o ribogrupo branco ocorrem em maior frequência e são amplamente distribuídos. Além disso, o haplogrupo e o ribogrupo branco ocupam a posição central da rede, sendo que o haplogrupo branco conecta os demais haplogrupos observados no Atlântico, características que segundo Avise et al. (1987) são inerentes a linhagens ancestrais. Na região Norte do Atlântico, também considerada um refúgio durante o Pleistoceno, a presença de dois haplótipos pertencentes ao ribogrupo branco foi observada. Para cpDNA um único haplótipo foi encontrado (haplótipo G) e os dados obtidos na genealogia de genes mostram que esse haplótipo é derivado do haplogrupo branco, sendo provável que a presença deste haplótipo na região nordeste seja fruto de um evento de expansão recente através da costa brasileira em direção ao norte do Atlântico.

Contrastando com a estabilidade presente nas regiões central e norte da costa do Atlântico, tem sido proposto que na região Sul as flutuações climáticas e o impacto das fases áridas do Pleistoceno tenham sido maiores, sendo a região Sul do Atlântico caracterizada como instável impossibilitando a existência de refúgios pleistocênicos nessas áreas (Carnaval e Moritz, 2008). Registros palinológicos

corroboram a instabilidade dessas áreas, sugerindo a substituição das florestas do sul do Atlântico por gramíneas durante o pleistoceno (Carnaval et al., 2009). Segundo Carnaval et al. (2009) a assinatura genética de expansão populacional em áreas instáveis reflete a colonização por múltiplas linhagens a partir de regiões de refúgio adjacentes, após o Último Máximo Glacial. Esse padrão se ajusta ao encontrado em cpDNA de *C. fissilis* e *C. odorata*, que apresenta os haplogrupos amarelo e laranja, derivados do haplogrupo ancestral branco, amplamente distribuídos na região Sul do Atlântico. Segundo Avise (2000) uma subestrutura espacial pode estar presente entre linhagens proximamente relacionadas dentro das regiões. Dentro da Região Sul do Atlântico foi possível observar que os haplogrupos amarelo e laranja se distribuíram preferencialmente em áreas distintas. O haplogrupo amarelo esteve presente em populações da região norte de SP, MS e noroeste do PR enquanto o haplogrupo laranja se concentrou nas populações do sudeste do PR, SC e RS. Entretanto uma região de sobreposição entre os dois haplogrupos, indicando o contato secundário entre os haplogrupos pode ser identificada nas populações do oeste do PR. De acordo com o padrão de distribuição espacial e temporal dos haplótipos na região Sul, é provável que após o último máximo glacial, as populações tenham se dispersado a partir dos refúgios central e sudeste, através de duas rotas: (i) uma rota central traçada pelos espécimes pertencentes ao haplogrupo amarelo que passa pelo sul do Mato Grosso do Sul e alcança o noroeste do Paraná, e (ii) uma rota litorânea traçada pelos espécimes do haplogrupo laranja, que contorna o litoral e alcança os estados do Paraná, Santa Catarina e Rio Grande do Sul. O afluxo significativo de imigrantes oriundos de populações localizadas em refúgios adjacentes pode explicar altos índices de diversidade genética encontrados nas florestas do interior do Paraná-Paraíba (ecorregião NT0150).

Para a região de ITS um único ribogruppo, denominado ribogruppo amarelo, foi derivado do ribogruppo branco ancestral. Os ribogruppos branco (ancestral) e amarelo (derivado) foram amplamente distribuídos entre as populações da região Sul do Atlântico, não sendo possível identificar nenhum padrão de subestrutura espacial entre esses ribogruppos. Segundo MacCauley (1995) os genes organelares frequentemente apresentam estrutura populacional consideravelmente mais distinta do que genes nucleares, simplesmente por causa de diferenças no tamanho efetivo da população. Em plantas, o fluxo gênico pode ser promovido pela dispersão de sementes e de pólen. Na maioria das angiospermas o cpDNA é maternalmente herdado e pode mover-se unicamente através da dispersão das sementes, dessa forma, as taxas de migração de cpDNA em angiospermas são muito menores do que as taxas de migração dos genes nucleares que podem mover-se através da dispersão de sementes e de pólen.

Segundo Briggs (1984) um determinado grupo tende a passar pela maior parte de seu desenvolvimento evolutivo em determinadas regiões que lhes são favoráveis, e a partir dessas regiões, eles podem se distribuir para outros locais, com seus representantes gradualmente invadindo regiões distantes. A região na qual a maior parte da história evolutiva de um determinado grupo ocorreu e onde, provavelmente, a maior diversidade pode ser encontrada (Centro de diversidade), pode ter sido o seu centro de origem, ou de qualquer forma, essa região provavelmente não está muito distante do centro de origem do grupo (Humphries e Parenti, 1986). Embora a Região do Atlântico não possa ser inferida como centro de origem para as populações de *C. fissilis* e *C. odorata* do Brasil ela constitui o Centro de diversidade para as espécies.

Linhagens de Chiquitano

Em geral, a diversidade genética encontrada na região de Chiquitano (ecorregiões NT0135 e NT0212) foi menor do que a diversidade presente na região do Atlântico. Esse fato sugere que as populações de Chiquitano são mais recentes ou sofreram gargalos populacionais em um passado recente, que eliminou grande parte da diversidade genética. Registros palinológicos evidenciaram que a composição florística das florestas secas da região de Chiquitano durante o Pleistoceno, difere muito da composição atual (Anhuf, 2006), sugerindo que as florestas secas hoje encontradas na região tenham origem recente, surgindo há aproximadamente 9.000 anos (Burbridge et al., 2004; Mayle, 2004).

Segundo Avise (2000) a comparação entre diversidade haplotípica e diversidade nucleotídica é uma forma simples de inferir a história demográfica de um grupo. Os valores de diversidade genética estimados para as ecorregiões de Chiquitano a partir de cpDNA e de ITS, apresentam valores moderados de diversidade haplotípica e baixa diversidade nucleotídica. Esse padrão permite inferir um rápido crescimento populacional a partir de uma população ancestral, com tempo suficiente para recuperação da variação por mutação, mas muito curto para acumular grandes diferenças nas sequências (Avise 2000). A análise filogeográfica fundamentada no genoma cloroplastídico demonstra que o haplogrupo ancestral preto foi identificado em 64 dos 88 indivíduos analisados e está amplamente distribuído na região de Chiquitano. Os haplogrupos verde e azul evoluíram a partir dele e são restritos a poucas populações respectivamente pertencentes às florestas secas de Chiquitano (ecorregião NT0212) e a uma população pertencente às florestas de galeria inserida na ecorregião no Cerrado (ecorregião NT0704). Adicionalmente, a análise filogeográfica fundamentada na região de ITS identificou um único ribogrupo distribuído pelas ecorregiões de Chiquitano.

Biogeografia molecular

A estrutura filogeográfica de *Cedrela fissilis* e *Cedrela odorata* sugere que episódios de isolamento genético associados a eventos de dispersão moldaram a história evolutiva das espécies. A existência de um evento vicariante que segregou as populações ancestrais de *C. fissilis* e *C. odorata* e promoveu um longo período de isolamento genético entre as regiões do Atlântico e de Chiquitano parece se ajustar a hipótese do “Arco Pleistocênico” de Prado e Gibbs (1993). De acordo com essa hipótese, a disjunção atual das florestas estacionais sazonalmente secas (Seasonally Dry Tropical Forests - SDTF) seriam relíquias formadas por vicariância de uma formação florestal pleistocênica de dimensão muito maior que cobria a América do Sul, e que foi submetida a perturbações ecológicas produzidas por ciclos de climas secos/úmidos com resfriamento/aquecimento que aconteceram nos últimos dois milhões de anos do Pleistoceno (Whitmore e Prance, 1987) ou possivelmente ainda em tempos anteriores (Pennington et al., 2004; Mayle 2004). A maior parte das populações de *C. fissilis* e *C. odorata* amostradas neste estudo estão localizadas em formações florestais inclusas no conceito de SDTF definido por Pennington et al. (2000) ou em regiões ecotonais entre áreas de SDTF e formações adjacentes, fato que corrobora a ideia da existência de refúgios. É plausível que o ambiente seco encontrado na faixa de vegetação mais aberta presente no Cerrado do Brasil central, que se interpõe entre as regiões do Atlântico e de Chiquitano, possa ter sido um fator de restrição importante na migração de *C. fissilis* e *C. odorata* entre as duas regiões durante períodos mais secos (Prado e Gibbs, 1993). O longo período de isolamento a que foram submetidas as populações do Atlântico e de Chiquitano permitiu a divergência entre elas, devido as diferentes forças de seleção natural ou deriva genica a que foram submetidas.

Entretanto, o isolamento por distancia não é o único modelo que explica a atual distribuição da diversidade em *C. fissilis* e *C. odorata*. Evidências de dispersão

recente entre as Regiões do Atlântico e de Chiquitano podem ser constatadas através da presença de espécimes do Atlântico que apresentaram haplótipos de cpDNA e ITS derivados de linhagens de Chiquitano, e de espécimes da região de Chiquitano que apresentaram haplótipos de cpDNA e ITS derivados de linhagens do Atlântico. Adicionalmente, os testes de neutralidade realizados para fazer inferências demográficas sobre a história das populações apresentaram valores negativos, sugerindo que as populações nessas ecorregiões sofreram expansão recente (Tajima, 1993). De um modo geral, a expansão das linhagens aconteceu através de duas rotas distintas: (i) através do Cerrado presente no Brasil Central e (ii) através da bacia do Rio Paraná ao Sul. A existência de corredores de vegetação atuando como possíveis rotas de migração para plantas entre as florestas Atlântica e Amazônica têm sido sugeridas. Bigarella et al. (1975) sugere a existência de uma ponte Sudeste-Noroeste que une as florestas do leste e do sudeste do Brasil a floresta Amazônica. Adicionalmente, Por (1992) citado por Costa (2003) relata a existência de uma rota principal de ligação através da bacia do Rio Paraná e uma rota secundária através das florestas de galeria dispostas ao longo dos rios que cortam o Cerrado do Brasil Central. Essas rotas são congruentes com possíveis rotas de migração observadas para as populações de *C. fissilis* e *C. odorata* e também e se sobrepõem as rotas observadas para a dispersão de pássaros (Silva, 1995) e pequenos mamíferos (Costa, 2003).

A expansão das linhagens de Chiquitano em direção ao Atlântico ocorreu através das duas rotas acima descritas; por meio do Cerrado as linhagens alcançaram as florestas do interior do Paraná-Paraíba (bloco disjunto da ecorregião NT0150) e as florestas secas do Atlântico (ecorregião NT0202); e por meio da Bacia do Rio Paraná alcançaram a ecorregião NT0150 (florestas do interior do Paraná-Paraíba) no sudeste do Brasil. A expansão das linhagens do Atlântico em direção a Chiquitano deu-se unicamente através da Bacia do Rio Paraná e permitiu

que as linhagens do Atlântico chegassem a Bolívia e ao Mato Grosso. Essas rotas de migração podem ter surgido em função da existência de florestas de galeria e de uma série de fragmentos de floresta estacional decídual e semi-decídual que constituem uma rede de florestas interligadas através da vegetação do Cerrado (Oliveira -Filho e Ratter, 1995) que pode ter favorecido o processo de dispersão.

A expansão dessas linhagens levou ao contato secundário entre as populações, reconectando-as. O elevado padrão de polimorfismo intraindividual e a evolução em concerto incompleta observada em ITS sugerem que a hibridação intraespecífica ocorreu quando linhagens alopátricas de *C. fissilis* e *C. odorata* estabeleceram zonas de contato secundárias. Os processos de hibridização interespecífica e intraespecífica reúnem os diferentes alelos originados das fontes parentais em um único genoma e, portanto, podem gerar polimorfismo intraindividual (Koch et al., 2003; Muir et al., 2001). Evidências de que o polimorfismo intraindividual foi muito difundido entre e dentro das regiões do Atlântico e de Chiquitano vem do fato de que o genoma da maior parte dos espécimes analisados combinou múltiplos parálogos de ITS advindos de linhagens distintas. A ausência de polimorfismo intraindividual para a região de ITS observada em alguns espécimes pode sugerir que estes espécimes sejam relíquias de populações ancestrais que escaparam da hibridação intraespecífica e, portanto, a presença desses espécimes pode sugerir a existência de refúgios nessas áreas ou em áreas próximas a ela. Além disso, a localização da maior parte dos clones obtidos para a região de ITS nas extremidades da rede ou em regiões intermediárias entre os três ribogrupos reafirma o caráter derivado desses parálogos, originados pelo processo de evolução em concerto incompleta.

Dispersão e vicariância são muitas vezes consideradas hipóteses concorrentes e um extenso debate em biogeografia tem abordado a questão da importância de cada um desses eventos para moldar padrões espaciais e temporais

da diversidade vegetal (Zink et al., 2000). O padrão filogeográfico obtido para *C. fissilis* e *C. odorata* reforça a idéia de que ambos os eventos podem ocorrer simultaneamente, atuando de maneira associada na determinação da distribuição atual dos táxons.

CONCLUSÕES

- 1) Para ambos os conjuntos de dados, cpDNA e ITS, não foram obtidas redes de haplótipos espécie-específicas para *C. fissilis* e *C. odorata*. A forma com que as sequências das duas espécies foram agrupadas na rede parece ter-se dado em função da distribuição geográfica.
- 2) Os níveis elevados de diferenciação genética e a evidente divisão das linhagens entre as regiões do Atlântico e de Chiquitano sugerem que as populações atuais de *C. fissilis* e *C. odorata* descendem de pelo menos duas linhagens ancestrais.
- 3) A região do Atlântico é o Centro de diversidade genética para *Cedrela fissilis* e *Cedrela odorata* no Brasil. A estrutura filogeográfica sugere a existência de refúgios nas regiões central e sudeste do Atlântico, a partir dos quais as linhagens migraram para o Nordeste, através de uma rota litorânea, e para o Sul através das rotas central e litorânea.
- 4). A estrutura filogeográfica sugere que as populações de Chiquitano são mais recentes ou sofreram gargalos populacionais em um passado recente. Os valores moderados de diversidade haplotípica e a baixa diversidade nucleotídica observados para essa região sugerem um rápido crescimento populacional a partir de uma população ancestral, com tempo suficiente para recuperação da variação por mutação, mas muito curto para acumular grandes diferenças nas sequências.

5) Episódios de isolamento genético associados a eventos de dispersão moldaram a história evolutiva de *C. fissilis* e *C. odorata*. A existência de um evento vicariante que segregou as populações ancestrais e promoveu um longo período de isolamento genético entre as regiões do Atlântico e de Chiquitano.

6) Eventos de expansão recente entre as Regiões do Atlântico e de Chiquitano reconectou essas regiões e tem favorecido o estabelecimento de zonas de contato secundário entre as populações nas florestas do interior do Paraná-Paraíba (ecorregião NT0150), nas florestas secas de Chiquitano (ecorregião NT0212) e nas florestas secas do Atlântico (ecorregião NT0202).

7) A expansão das linhagens entre as Regiões do Atlântico e de Chiquitano aconteceu através do Cerrado presente no Brasil Central e através da bacia do Rio Paraná ao Sul. Acredita-se que essas rotas de migração possam ter surgido em função da existência de florestas de galeria e de uma série de fragmentos de floresta estacional decidual e semi-decidual que constituem uma rede de florestas interligadas através da vegetação do Cerrado favorecendo o processo de dispersão.

REFERÊNCIAS BIBLIOGRÁFICAS

- Anhuf, D., Ledru, M.-P., Behling, H., Da Cruz Jr, F.W., Cordeiro, R.C., Van der Hammen, T., Karmann, I., Marengo, J.A., De Oliveira, P.E., Pessenda, L., Siffedine, A., Albuquerque, A.L., Da Silva Dias, P.L. 2006. Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology* 239, 510-527.
- Auler, A.S., Wang, X., Edwards, R.L., Cheng, H., Cristalli, P.S., Smart, P.L., Richards, D.A. 2004. Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *Journal of Quaternary Science* 19, 693-701.
- Avice, J.C. 2000. *Phylogeography: the history and formation of species*. Harvard Univ. Press, Cambridge. 447p.
- Avice, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18, 489-522.
- Bandelt, H-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16, 37-48.
- Baum, D.A., Small, R.L., Wendel, J.F. 1998. Biogeography and floral evolution of Baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Systematic Biology* 47, 181-207.
- Bigarella, J.J., Andrade-Lima, D., Riehs, P.J. 1975. Considerações a respeito das mudanças paleoambientais na distribuição de algumas espécies vegetais e animais no Brasil. *Anais da Academia Brasileira de Ciências* 47, 411-464.
- Briggs, J.C. 1984. *Centres of Origin in Biogeography*. Biogeographical Monographs vol. 1. University of Leeds. 1-79.
- Burbridge, R.E., Mayle, F.E., Killeen, T.J. 2004. Fifty-thousand-year vegetation and climate history of Noel Kempff Mercado National Park, Bolivian Amazon. *Quaternary Research* 61, 215-230.

- Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., Moritz, C. 2009. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. *Science* 323, 785-789.
- Carnaval, A.C., Moritz, C. 2008. Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography* 35, 1187-1201.
- Carvalho, P.E.R. 2003. *Espécies Arbóreas Brasileiras*, Vol. 1, Ed: Embrapa Florestas. 1040 p.
- Cavers, S.; Navarro, C. Lowe, A.J. 2003. Chloroplast DNA phylogeography reveals colonization history of a Neotropical tree, *Cedrela odorata* L., in Mesoamerica. *Molecular Ecology*. 12: 1451–1460.
- Costa, L.P. 2003. The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography* 30, 71-86.
- Cronn RC, Small RL, Haselkorn T, Wendel JF. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* 89, 707-725.
- Excoffier, L., Lawlor, S., Schhneider, S. 2005. Arlequin ver 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1, 47-50.
- Gardner, P.P., Daub, J., Tate, J.G., Nawrocki, E.P., Kolbe, D.L., Lindgreen, S., Wilkinson, A.C., Finn, R.D., Griffiths-Jones, S., Eddy, S.R., Bateman, A. 2008. Rfam: updates to the RNA families database. *Nucleic Acids Res.* 37, 136-140.
- Gentry, A.H. 1982. Neotropical floristic diversity: phylogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* 69, 557–593.
- Gillies, A.C.M., Cornelius, J.P., Newton, A.C., Navarro, C., Hernández, M., Wilson, J. 1997. Genetic variation in Costa Rican populations of the tropical timber species *Cedrela odorata* L., assessed using RAPDs. *Molecular Ecology* 6, 1133-1145.
- Humphries, C.J., Parenti, L.K. 1986. *Cladistic Biogeography*. Claredon Press. Oxford. Cap. 1.

- IBGE .2004. Mapa de vegetação do Brasil. IBGE, Rio de Janeiro, Brazil.
- IUCN. 2008. 2008 IUCN Red list of threatened species. Disponível em: <http://www.redlist.org/>.
- Janzen, D.H. 1988. Tropical dry forest: the most endangered major tropical ecosystem. In: Biodiversity (ed. Wilson EO) pp 130-137. National Academy Press, Washington.
- Kimura, M. 1969. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* 61, 893-903.
- Koch, M.A., Dobes, C., Mitchell-Olds, T. 2003. Multiple hybrid formation in natural populations: concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). *Molecular Biology and Evolution* 20, 338–350.
- Ledru, M.-P., Salatino, M.L.F., Ceccantini, G., Salatino, A., Pinheiro, F., Pintaud, J.-C. 2007. Regional assessment of the impact of climatic change on the distribution of a tropical conifer in the lowlands of South America. *Diversity and Distributions* 13, 761-771.
- Librado, P., Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451-1452.
- McCauley, D.E. 1995. The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution* 10, 198–202.
- Mast, A.R., KELSO, S., RICHARDS, A.J., LANG, D.J., FELLER, D.M.S., CONTI, E. 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. *International Journal of Plant Sciences* 162, 1381–1400.
- Mayle, F.E. 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.
- Muellner, A.N., Pennington, T.D., Chase, M.W. 2009. Molecular phylogenetics of Neotropical Cedreleae (mahogany family, Meliaceae) based on nuclear and plastid DNA sequences reveal multiple origins of "*Cedrela odorata*". *Molecular Phylogenetics and Evolution* 52, 461–469.

- Muellner, A.N., Pennington, T.D., Koecke, A.V., Renner, S.S., 2010. Biogeography of *Cedrela* (Meliaceae, Sapindales) in Central And South America. *American Journal of Botany* 97, 511-518.
- Muir, G., Fleming, C.C., Schlötterer, C., 2001. Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.. *Molecular Biology and Evolution* 18, 112–119.
- Oliveira-Filho, A.T., Fontes, M.A.L. 2000. Patterns of floristic differentiation among atlantic forests in southeastern Brazil and the influence of climate. *Biotropica* 32, 793–810.
- Oliveira-Filho, A.T., Ratter, J.A. 1995. A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. *Edinburgh Journal of Botany* 52, 141-194.
- Oliveira-Filho, A.T., Ratter, J.A. 2002. Vegetation physiognomies and woody flora of the Cerrado Biome. In: *The Cerrado of Brazil: Ecology and Natural History of a Neotropical Savann* (eds Oliveira PS, Marquis RJ), pp. 91–120. Columbia University Press, New York.
- Oliveira-Filho, A.T. Scolforo, J.R.S. Mello, J.M. 1994. Composição florística e estrutura de um remanescente de floresta semidecidual montana em Lavras, MG. *Revista Brasileira de Botânica* 17, 167-182.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Louks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R. 2001. Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* 51, 933-938.
- Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K., Butterworth, C.A. 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society* 359, 515-538.
- Pennington, R.T.; Prado, D.E.; Pendry, C.A. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography* 27, 261–273.
- Prado, D.E., Gibbs, P.E. 1993. Patterns of species distribution in the dry seasonal forest of South America. *Annals of the Missouri Botanical Garden* 80, 902–927.

- Rossi, A.A.B., Oliveira, L.O., Venturini, B.A., Silva, R.S. 2009. Genetic diversity and geographic differentiation of disjunct Atlantic and Amazonian populations of *Psychotria ipecacuanha* (Rubiaceae). *Genetica* 136, 57-67.
- Sanchez-Azofeifa, G.A., Quesada, M., Rodriguez, J.P., Nassar, J.M., Stoner, K.E., Castillo, A., Garvin, T., Zent, E.L., Calvo-Alvarado, J.C., Kalacska, M.E.R., Fajardo, L., Gamon, J.A., Cuevas-Reyes, P. 2005. Research priorities for Neotropical dry forests. *Biotropica* 37, 477-485.
- Sambrook, J., Fritsch, E. F., Maniatis, T. 1989. Molecular cloning, a Laboratory Manual. 2nd ed. USA: Cold Spring Harbor Laboratory.
- Selig, C., Wolf, M., Müller, T., Dandekar, T. and Schultz, J., 2008. The ITS2 Database II: homology modeling RNA structure for molecular systematic *Nucleic Acids Res.* 36, 377-380.
- Silva, J.M.C. 1996. Distribution of Amazonian and Atlantic birds in gallery forests of the cerrado region, South America. *Ornitologia Neotropical* 7, 1-18.
- 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.
- Tajima, F. 1993. Simple methods for testing molecular clock hypothesis. *Genetics* 135, 599-607.
- Tzedakis, P.C., Lawson, I.T., Forgley, M.R., Hewitt, G.M. 2002. Buffered tree population changes in a Quaternary refugium: evolutionary implications. *Science* 297, 2044–2047.
- Valera, F.P. 1997. Genetic Resources of *Swietenia* and *Cedrela* in the Neotropics: Proposals for Coordinated Action. Forest Resources Division, Forestry Department, Food and Agriculture Organisation of the United Nations, Rome.
- Veloso, H.P., Rangel Filho, A.L.R., Lima, J.C.A. 1991. Classificação da Vegetação Brasileira, Adaptada de um Sistema Universal. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.
- White, T.J., Bruns, T., Lee, S., Taylor, J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: a Guide to Methods and Applications. (Eds. Innis MA, Gelfand DH, Sninsky JJ and White TJ), pp. 315-322. Academic Press, London, UK.

- Whitmore, T.C., Prance, G.T. 1987. Biogeography and Quaternary History in Tropical America. Oxford Science Publications, Oxford, UK.
- Wuyts, J., Perriere, G., Van de Peer, Y., 2004. The European ribosomal RNA database. Nucleic Acids Res. 32, 101-103.
- ZINK, R.M.; BLACKWELL-RAGO, R.C.; RONQUIST, F. 2000. The shifting roles of dispersal and vicariance in biogeography. Proceedings of the Royal Society B: Biological Sciences 267, 497-503.