MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 1410–1422 doi: 10.1111/j.1365-294X.2010.04591.x

Molecular phylogeography of *Carapichea ipecacuanha*, an amphitropical shrub that occurs in the understory of both semideciduous and evergreen forests

LUIZ ORLANDO DE OLIVEIRA,* ANA APARECIDA BANDINI ROSSI,*† ERNANE RONIE MARTINS,‡ FLÁVIA REIS DE CARVALHO BATISTA* and ROBERTA SANTOS SILVA*

*Departamento de Bioquímica e Biologia Molecular, Universidade Federal de Viçosa, 36570-000, Viçosa (MG), Brazil, †Departamento de Ciências Biológicas, Universidade do Estado de Mato Grosso, 78580-000 Alta Floresta (MT), Brazil, ‡Núcleo de Ciências Agrárias; Universidade Federal de Minas Gerais; 39404-006 Montes Claros (MG), Brazil

Abstract

The medicinal shrub *Carapichea ipecacuanha* (ipecac) is an amphitropic species with three disjunct areas of distribution. In the Brazilian Atlantic and Amazonian ranges, the species was associated mostly with the understory of seasonal semideciduous forests, whereas in the Central American–Colombian range, the species occurred in the understory of moist evergreen forests. We examined the phylogeographic structure of ipecac using chloroplast *trn*T-*trn*L and nuclear internal transcribed spacer (ITS) sequences from 120 and 46 specimens, respectively. To complement existing data on root alkaloid profiles, we used high-performance liquid chromatography to assess the levels of emetine and cephaeline in 33 specimens from the two Brazilian ranges. The three ranges shared neither nuclear nor chloroplast haplotypes. The phylogeographic structures showed an uneven distribution of genetic diversity, sharp breaks and high levels of genetic differentiation among ranges. Our results suggest that the extant populations are descendents of at least four distinct ancestral lineages. The Atlantic ipecacs showed higher levels of genetic diversity than ipecacs from the other two ranges; it is likely that they derive from two ancestral lineages, with long-term persistence in that region. The Amazonian ipecacs were monomorphic with respect to the ITS and cpDNA sequences, which supports the view that there was a recent expansion from a single parental source after a strong genetic bottleneck. The existence of a fourth distinct lineage is apparent from the high levels of genetic and chemical differentiation that we identified in the Central American–Columbian ipecacs.

Keywords: ipecac, internal transcribed spacer, *Psychotria ipecacuanha*, root alkaloid, Rubiaceae, SDTF, *trnT-trnL*

Received 13 March 2009; revision received 18 January 2010; accepted 24 January 2010

Introduction

The neotropical seasonal forests comprise blocks of vegetation scattered throughout most of South and Central America. In Brazil, seasonal forests are one of the main types of vegetation of the Mata Atlântica Biome (Veloso *et al.* 1991; IBGE 2004). Another large disjunct block of seasonal forest in South America is the Chiquitano forest, which is located mostly in the eastern lowlands of Bolivia, but also extends into Central Brazil, specifically into areas of western Mato Grosso (Ribera *et al.* 1994; IBGE 2004). Some early biogeographical studies recognized the floristic relationships between these geographically distant fragments of forest (Bigarella *et al.* 1975; Ab'Saber 1977). More recent floristic inventories have listed many phylogenetically unrelated species that are found in neotropical seasonal forests but are absent

Correspondence: Luiz Orlando de Oliveira, Fax: +55 31 3899 2373; E-mail: lorlando@ufv.br

from both moist evergreen forests and savannas in neighbouring areas (Prado & Gibbs 1993; Prado 2000; Killeen *et al.* 2006; Pennington *et al.* 2006).

Characteristically, the seasonal forests occur in areas in which a severe dry season (up to 5 consecutive months) follows the rainy season. Nevertheless, seasonal forests also can occur in areas where the distribution of rainfall is more uniform, but where winter temperatures can fall below 15 °C, which can cause physiological drought (Veloso *et al.* 1991; Oliveira-Filho & Fontes 2000). In the Mata Atlântica Biome, seasonal semideciduous forests can be distinguished consistently from moist evergreen forests on the basis of physiognomy and floristic composition; the seasonality of rainfall makes an important contribution to this differentiation (Oliveira-Filho *et al.* 2006). Among seasonal forests, some authors use the degree of deciduousness to distinguish semideciduous forests (in which leaf fall occurs in 20–50% of the constituents) from deciduous forests (in which leaf fall occurs in more than 50% of the constituents) (Veloso *et al.* 1991). Other authors have noted that it is difficult to distinguish between the two types of forest on the basis of deciduousness alone, particularly in central Brazil, where the semideciduous and deciduous forests form a continuum whose characteristics are determined by local variations in soil moisture and fertility (Oliveira-Filho & Ratter 2002).

Several bodies of evidence suggest that neotropical seasonal forests experienced both contractions (Carnaval & Moritz 2008; Carnaval *et al.* 2009) and expansions (Auler *et al.* 2004; Mayle 2004) in response to climatic changes during the Quaternary period (Ledru 1993). These shifts in vegetation in the Neotropics were triggered when cycles of cooling/warming and dry/wet climates altered the ecological conditions at local and regional scales during the last 2 million years of the Pleistocene (Whitmore & Prance 1987), and possibly earlier (Mayle 2004; Pennington *et al.* 2004). According to the 'Pleistocenic Arc' hypothesis of Prado & Gibbs (1993), the present-day disjunct seasonal forests are relicts that were formed by the vicariance of a much larger single formation that covered South America during the Pleistocene. Rare long-distance dispersal of species associated with seasonal forests has been proposed as an alternative explanation for the disjunct distribution of seasonal forests (Gentry 1982; Mayle 2004).

Although phylogeographical studies have progressed substantially in many parts of the world, especially in Europe and North America, few studies have been reported on components of neotropical forests. Moreover, those studies that have been published have considered only woody tree species, such as *Caryocar brasiliense* (Collevatti *et al.* 2003) and *Hymenaea stigonocarpa* (Ramos *et al.* 2007), both of which occur in the Cerrado of Central Brazil; the conifer genus *Podocarpus* (Ledru *et al.* 2007) and *Caesalpinia echinata* (Cardoso *et al.* 1998), both of which occur in the Atlantic rainforest; and *Astronium urundeuva*, a tree species that is confined to deciduous forests (Caetano *et al.* 2008).

In the study described herein, we explored the evolutionary history of seasonal forests by investigating the molecular phylogeography of ipecac, *Carapichea ipecacuanha* (Brot.) L. Andersson [= *Psychotria ipecacuanha* (Brot.) Stokes; *Cephaelis ipecacuanha* (Brot.) A. Rich.] (Rubiaceae). Ipecac is a long-lived, perennial shrub that inhabits the shaded understory exclusively (Veloso 1947; Oliveira & Martins 1998). The species is sensitive to the long-term penetration of light to the canopy floor that is caused by clearing, selective cutting and incidental fires. In addition, populations of ipecac decline rapidly when exposed to forest-edge environments (Oliveira & Martins 2002). Remarkably, the species has an extensive distribution. Wild populations of ipecac occur in Brazil, Colombia and Central America and are used for medicinal purposes (Assis & Giulietti 1999; Rossi *et al.* 2009). Although ipecacs of Central America are similar morphologically to Brazilian ipecacs (Assis & Giulietti 1999), they can be distinguished readily on the basis of their root alkaloid profiles (Bruneton 1995). It is likely that the disjunction of ipecacs in Brazil did not result from human-mediated long-distance dispersals (Rossi *et al.* 2009). Several traits of ipecac, such as limited dispersal ability due to habitat specialization, a long life span, clonal propagation and a widely disjunct distribution, might have promoted the preservation of evolutionary signatures that are suitable for a phylogeographic study.

In this study, we characterized the geographic range of ipecac, carried out phylogeographic analyses with both chloroplast and nuclear DNA sequence data and compared root alkaloid profiles to address the following questions: (i) To what extent does ipecac occur in seasonal forests, at the expense of adjacent types of vegetation, such as Cerrado, Caatinga and moist evergreen forests? (ii) Does the phylogeographic structure of ipecac suggest the existence of recent colonized areas and refugia? (iii) Did the colonization and persistence of ipecac in distinct ranges follow a common pattern? (iv) Is the alkaloid profile related to phylogeographic structure? (v) What does the phylogeographic structure of ipecac reveal about the origin and persistence of disjunctions of seasonal forest over time? Our results shed light on the little-known evolutionary history of the neotropical flora, which so far has been characterized only in terms of woody tree species, with shrubs, lianas and herbs being overlooked.

Materials and methods

Characterization of geographic range

The geographic distributions of ipecac were estimated from field trips, herbarium records and databases of plant checklists. We have carried out extensive expeditions to collect Brazilian ipecacs since 1995. Wherever we found populations, we recorded the location using a Global Positioning System (GPS) receiver. Voucher specimens were prepared from specimens that contained reproductive structures at the time of sampling and were deposited in the herbarium VIC (Table S1, Supporting Information). We consulted 15 other herbaria and examined an additional 164 herbarium specimens of ipecac (Table S1, Supporting Information). The total collection period of samples from these sources spanned over 171 years (from 1837 to 2008). We checked specimens and their associated labels to confirm their identity. Duplicates, specimens with imprecise information about location, specimens that were probably collected from medicinal gardens and misidentified specimens were discarded. We also consulted databases of plant checklists for Ecuador (Jørgensen & León-Yánez 1999; Taylor 2006), Peru (Brako & Zarucchi 1993), the Guiana Shield (Funk *et al.* 2007) and the Brazilian states of Acre (Daly & Silveira 2008), São Paulo (Jung-Mendaçolli 2007), Goiás & Tocantins (Delprete PG, submitted) and Mato Grosso do Sul (Dubs 1998).

To define the vegetation formations in which ipecac occur, we consulted the Terrestrial Ecoregions of the World, which was published by the World Wildlife Fund (WWF) (Olson *et al.* 2001). The habitat classifications for Brazil from the Brazilian Institute of Geography and Statistics (IBGE 2004) provided further details about vegetation at a local level. We recorded our inventory of occurrences of ipecac on topographic maps (source: http://www.worldwildlife.org) using GPS software. After redundant plots that resulted from independent collections at the same locality were removed, the final map contained 97 unique plots (Table S1, Supporting Information).

Sampling strategy

For DNA analyses, we sampled 12 populations in the Mata Atlântica Biome, 18 populations of the Amazônia Biome and one population from Central America (Table 1). To minimize the chances of sampling clones, we avoided sampling plants that were close to each other by using 30 m as the minimum distance between individual samples. Sample sizes per population varied from 1 to 18, depending upon the number of clusters available for sampling within a given population.

DNA extraction and amplification

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 by 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 (Table S1, Supporting Information).

The *trn*T–*trn*L spacer of the chloroplast genome was amplified using standard PCR protocols. 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 A and B, which amplified the intergenic spacer between *trn*T (UGU) and the *trn*L (UAA) 5′ exon (Taberlet *et al.* 1991). The second was performed with primers C and D and amplified the *trn*L (UAA) intron (Taberlet *et al.* 1991). 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, with the universal primers ITS4 and ITS5 described by White *et al.* (1990). All amplifications were carried out in a GeneAmp PCR System 9700 (Applied Biosystems).

Sequencing of trnT–trnL spacer and ITS regions, sequence alignment, and Median Joining Network construction

DNA samples were sequenced using either a MegaBace DNA Analysis System 500 (Amersham Biosciences Corp.) or an ABI Prism 377 Genetic Analyzer (Applied Biosystems). For the ITS region, the electrophoretograms for direct sequencing of seven specimens from the Mata Atlântica Biome showed overlapping double peaks in both strands at particular sites (Fig. S1, Supporting Information), which suggested intragenomic variation. To resolve the ambiguities, we cloned these ITS sequences into the vector pGEM T-Easy (Promega) and resubmitted the clones for sequencing. The number of clones sequenced per specimen ranged from two (for TLM2) to six (for VRB5). The electrophoretograms for direct sequencing of the remaining 14 Atlantic specimens, 19 Amazonian specimens and five Central-American specimens showed no overlapping double peaks in both strands. For that reason, we did not clone these sequences. All sequences were imported into Sequencher version 4.8 (Gene Codes Corp.) for editing. The sequence alignments required manual corrections and

Table 1 Sampled locations of *Carapichea ipecacuanha* used in alkaloid and DNA analyses, with sample size (*N*), nucleotide diversity (*P*i), and haplotype diversity (Hd)

			DNA analyses					
			ITS			cpDNA		
Populations (acronym)	Location	Alkaloid analyses	*N*	*P*i	Hd	*N*	*P*i	Hd
Atlantic range			21	0.01033	0.801	73	0.00107	0.617
Carangola (CAR)	S20º44′28″W42º01′33″	x				1		
Caratinga (CAT)	S19º43′55″W41º49′00″	x						
Conceição do Macabu (CON)	S22º05′26″W41º52′12″	x	1			2	0.00000	0.000
Dores de Guanhães (DOG)	S19º02′17″W42º57′27″		1			2	0.00000	0.000
Eugenópolis (EUG)	S21º03′39″W42º25′01″	x						
Guaraciaba (GUA)	S20º34′14″W43º00′28″		2	0.0000	0.000	17	0.00106	0.632
Irupi (IRU)	S20º20′58″W41º38′25″	x				1		
Marilac (MAC)	S18º28′11″W42º10′04″	x						
Ponte Nova (PTN)	S20º25′26″W42º53′42″					1		
Raposo (RAP)	S21º06′40″W42º05′43″	x	5	0.01128	0.917	18	0.00087	0.399
Trilha Central (TCE)	S19º33′44″W42º37′57″		1			3	0.00188	1.000
Trilha da Lagoa do Meio (TLM)	S19º38′34″W42º30′56″		3	0.00658	0.833	5	0.00085	0.600
Trilha do Vinhático (TVI)	S19º45′44″W42º28′56″	x	4	0.00000	0.000	10	0.00066	0.467
Una (UNA)	S15º20′00″W39º22′00″	x	1			2	0.00000	0.000
Visconde do Rio Branco (VRB)	S21º00′37″W42º50′26″	x	3	0.01418	0.985	11	0.00123	0.709
Amazonian range			19	0.00000	0.000	41	0.00000	0.000
Abacatinho (ABA)	S11º14′14″W64º37′44″		1			1		
Abaitará (ABR)	S11º44′21″W61º31′36″	x	1			1		
Barra do Bugres (BBU)	S14º51′00″W57º48′00″					2	0.00000	0.000
Colorado (COL)	S11º14′49″W64º24′40″	x	2	0.00000	0.000	2	0.00000	0.000
Exu (EXU)	S15º40′41″W57º32′03″		1					
Fazenda Manilha (MAN)	S14º58′00″W57º57′00″					3	0.00000	0.000
Fazenda Raizama (FRA)	S14º53′00″W57º16′00″					3	0.00000	0.000
Figueirópolis D'oeste (FIG)	S15º26′37″W58º47′06″		3	0.00000	0.000	1		
Lambari D'Oeste (LAM)	S15º19′27″W58º00′10″		1			2	0.00000	0.000
Linhares (LIN)	S11º40′54″W61º32′25″	x	1					
Mozar (MOZ)	S15º04′57″W57º58′00″		1			3		
Prata (PRA)	S15º30′26″W58º01′50″		1			6	0.00000	0.000
Rio Vermelho (RVE)	S15º17′38″W57º51′43″		1			6	0.00000	0.000
Santa Maria (SMA)	S11º12′42″W64º46′45″	x	1			1		
Sítio Jesus (JES)	S11º31′29″W61º33′51″	x	2	0.00000	0.000	1		
Soroteca (SOR)	S15º31′34″W58º00′42″		2	0.00000	0.000	2	0.00000	0.000
Tangará da Serra (TAN)	S14º04′38″W57º03′45″		1			1		
Vila B. S. Trindade (VBS)	S15º00′29″W59º57′02″		1			6	0.00000	0.000
São Francisco (SAF)	S11º13′59″W64º39′30″	x						
São Pedro (SPE)	S11º36′43″W61º33′07″	x						
Central-American range			6	0.00000	0.000	6	0.00000	0.000
Barro Colorado Island (BCI)	N09º09′25″W79º51′16″		5	0.00000	0.000	5	0.00000	0.000
AF072020			1			1		

adjustments due to the presence of insertion/deletions (indels). After we had aligned the sequences, we pruned the ends of the sequences to eliminate fragments that could not be obtained for all specimens.

Gene genealogies were inferred using the Median Joining (MJ) network method (Bandelt *et al.* 1999) as implemented in NETWORK 4.5.0.2 (Fluxus Technology Ltd) using default parameters. The program was run such that indels were considered to be a fifth character state and coded such that each indel, regardless of its size, was considered to be a single mutation. One ITS sequence of *Psychotria ipecacuanha* (GenBank Accession AF072020) and its associated *trn*T-*trn*L spacer sequence (kindly provided by Molly Nepokroeff) were also included in the analyses. DnaSP v5 software (Librado & Rozas 2009) was used to estimate the gene diversity and nucleotide diversity (Nei 1987) indices at population and range levels.

Alkaloid analyses

Alkaloid profiles were obtained using the sample preparation and high-performance liquid chromatography (HPLC) techniques described by Garcia *et al.* (2005). We analysed one to six roots from each of 10 Atlantic populations and seven Amazonian populations (Table 1). Alkaloid concentrations were estimated by reference to standard calibration curves. Chromatographic analysis of each subsample was replicated and the average concentration recorded. All solvents used for these analyses were HPLC grade. Student's *t*-test (Neter & Wasserman 1974) was used to compare the contents of emetine and cephaeline between the two Brazilian ranges.

Results

Geographic range and associated vegetation formations

The spatial distribution of the sampled populations and herbarium specimens indicated that ipecac is an amphitropic species with three disjunct areas of distribution: (i) the Atlantic range; (ii) the Amazonian range; and (iii) the Central American-Colombian range (Fig. 1). The Amazonian range is at least 2500 km from the Central American–Colombian range and 1400 km from the Atlantic range. Plant checklists for intervening regions contained no mention of ipecac. Brazilian towns and villages with a strong oral history of gathering wild populations of ipecac for commercial purposes (which took place from the 18th century to the mid 1970s) were located within the limits of the Atlantic and Amazonian ranges.

In Brazil, ipecacs occurred mostly in seasonal semideciduous forests, whereas in Central America and Colombia, ipecacs were found in moist evergreen forests (Fig. 1). In the Atlantic range, ipecacs were present strongly in the seasonal semideciduous forests of the Bahia interior forests (ecoregion NT0104) of the Mata Atlântica Biome, a fringe forest formation that is located between the coastal moist forests and the drier cerrado towards Central Brazil. A few isolated plots indicated that the Atlantic range extended further into the adjacent seasonal semideciduous forests of the Paraná-Paraíba Interior Forests (ecoregion NT0150) and to the moist forests of the Atlantic Coast restingas, Bahia coastal forests and Serra do Mar coastal forests (ecoregions NT0102, NT0103 and NT0160, respectively). None of the plots fell within adjacent ecoregions of either cerrados or drier forest formations, such as the Caatinga and the isolated blocks of dry forest in Central Brazil. In the Amazonian range, ipecacs occurred in two ecoregions that were situated partially across transition zones: ipecacs of the Mato Grosso subrange occurred in seasonal semideciduous forests of the Chiquitano dry forests (ecoregion NT0212), whereas ipecacs of the Rondônia subrange occurred in the Madeira-Tapajós moist forests (ecoregion NT0135). An inspection of the local vegetation of the Madeira-Tapajós ecoregion, using both field visits and the IBGE vegetation system, revealed that the Rondônia subrange is located in a patchwork of seasonal semideciduous forests, submontane open ombrophilous forests and cerrado. In both the Madeira-Tapajos moist forests and the Chiquitano dry forests, the limits of the Amazonian range across the Bolivian border have not been defined exactly. Ipecacs of the

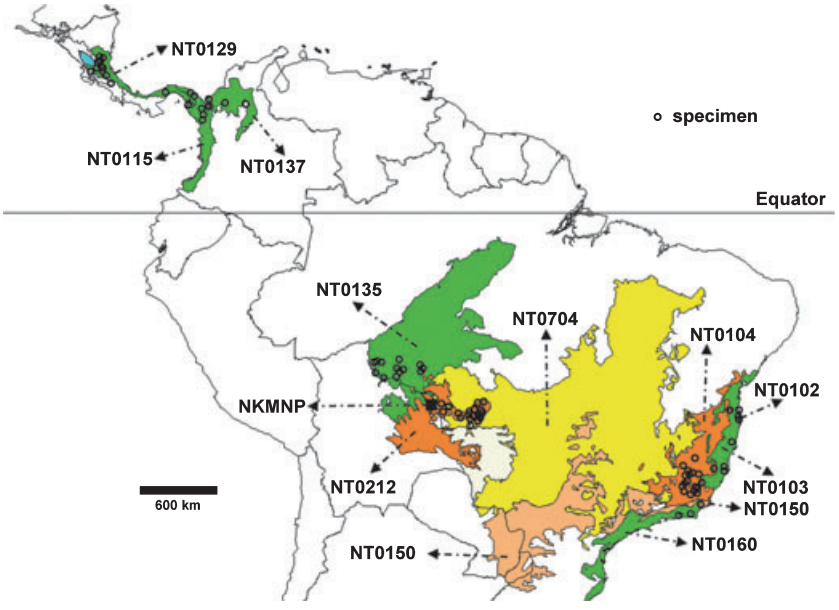

Fig. 1 The three geographic ranges of *Carapichea ipecacuanha* and associated vegetation formations. Terrestrial ecoregions are as follows: NT0102, Atlantic Coast restingas; NT0103, Bahia coastal forests; NT0104, Bahia interior forests; NT0115, Chocó-Darién moist forests; NT0129, Isthmian-Atlantic moist forests; NT0135, Madeira-Tapajós moist forests; NT0137, Magdalena-Urabá moist forests; NT0150, Paraná-Paraíba interior forests; NT0160, Serra do Mar coastal forests; NT0212, Chiquitano dry forests. Additional references: NKMNP, Noel Kempff Mercado National Park; ecoregion NT0704, Cerrado.

Central American–Colombian range were found within moist evergreen forests (Chocó-Darién moist forests—NT0115; Isthmian-Atlantic moist forests—NT0129; and Magdalena-Urabá moist forests—NT0137).

Nuclear ITS dataset

Our dataset for the nuclear ITS region consisted of aligned sequences of 532 bases. There were 23 ITS haplotypes among the 63 ribosomal sequences (Fig. 2). The ribosomal sequences comprised 24 clones plus 39 direct sequences, from a total of 46 specimens. Our pruning strategy eliminated the first 14 bases of the ITS1 spacer and the last eight bases of the ITS2 spacer. The sequences have been deposited in GenBank with the accession numbers GU361045–GU361109. Sequence alignment revealed the presence of 57 polymorphic sites that were characterized by either indels or base substitutions. The ITS-1 spacer showed 21 polymorphic sites, 19 of which were base substitutions and two of which were 1-bp indels. The ITS-2 spacer had a similar level of polymorphism: we identified 22 base substitutions and one 1-bp indel within this fragment. Sequencing of clones yielded a total of 20 unique ribosome sequences, among which we identified 16 intragenomic polymorphic sites in the ITS-1 spacer, 13 polymorphic sites in the 5.8S gene and 17 polymorphic sites in the ITS-2 spacer. Six out of 23 ITS haplotypes (7, 12, 16, 18, 19 and 20) featured 13 base substitutions within the 5.8S gene altogether, nine of which were methylation-related substitutions (C → T or G → A) that were found only in haplotypes 18, 19 and 20 (Fig. 2). In addition, three of these haplotypes (16, 19 and 20) showed at least three unique polymorphic sites within the ITS spacers.

cpDNA dataset

Sequencing of the target *trn*T-*trn*L spacer of the chloroplast genome yielded two segments of sequence. We aligned a total of 564 bases for fragment AB and 477 bases for fragment CD (GenBank accession numbers GU385028–GU385146). The assembly of these two sets of sequences resulted in a final chloroplast dataset that was 1041 bases long and contained sequences from 120 specimens. Sequence analysis revealed a total of 13 haplotypes and 22 polymorphic sites (Fig. 3). Fragment AB contained 14 polymorphic sites, nine of which were base substitutions and five of which were indels. The indels from segment AB were 14, 9, 7, 227 and 17 bases long, respectively. The 227-bp indel contained five base substitutions and four additional smaller indels and was present only in specimens of either Central-American or Amazonian origin. The absence of the entire 227-bp indel from all specimens of the Atlantic range precluded us from making any inferences about the ancestral states of these five base substitutions and four smaller indels; therefore, we discarded the information contained within 227-bp indel from our subsequent analyses. Segment CD showed lower levels of polymorphism; it contained five base substitutions and three indels that were 6, 10 and 71 bases in length, respectively.

```
                      1111111111  1122222222223  33333333334444444555555
           14466678890011223447  7812233457881  34566677890018999000222
          55034563178491903738O  8108967888276  55614705445798478568456
HAP #/@
 1   0/3 ACACCCAACAACCGACCCGTA  CGTGCTGCTCACC  CCCGGGACGCCGACACGGTAGAC
 2   0/1 G....................  .............  ..TT....................
 3   0/1 ............T........  .............  ........................
 4   0/1 ..G.................  .............  ........................
 5   0/1 ..................G..  .............  ..................G.....
 6   0/1 ..................G..  .............  ........A...............
 7   0/1 ....................  .....C.......  ........................
 8   0/1 ....................  .............  ..................G.....
 9   0/1 .....T..............  .............  A.....G..T.AG...........
10   0/1 ..............GT.T...  .............  A.....G..T..G...........
11 14/3 ..............GT.TA..  .............  A.....G..T..G...........
12   0/1 ..............GT.TA..  ........C....  A.....G..T..G...........
13   0/1 ..............GT.TA..  .............  A.....G..T..G.........T
14   0/1 ..............GT.TA..  .............  A.........T..G...........
15   0/1 .......G......GT.TA..  .............  A.........T..G...........
16   0/1 ....T.G...G...GT.TA..  ..........G..  A.........T..G...........
17   0/1 ..............GT.TA..  .............  A.....G..TT.G...........
18   0/1 ..............GT.TA..  ..C.T........  A.....G..T..G...........
19   0/1 .........T....GT.TA..  ...C..AT.T...  A.....GA.T..G...A.......
20   0/1 .T...........TTGT.TA..  TA.........TT  A-...AGA.T..G.T.........
21 19/0 ..............G..T.C.  .............  ......G..T..G......G....
22   5/0 ...-........A..G.A..C.  .............  T..TC.G..T...T.T.ACGC...
23* 1/0 ...-....-..A..G.A..C.  .............  T..TC.G..T...T.T.ACGC...
        └──────────────────┘  └─────────┘  └──────────────────┘
                 ITS-1            5.8S            ITS-2
```

Fig. 2 Sequence alignment of the variable sites in the nuclear ITS that define the 23 ITS haplotypes of *Carapichea ipecacuanha*. Each fragment spans 532 bases. Dots indicate similarity to haplotype 1 and hyphens indicate gaps. Numbers on top indicate the nucleotide position with haplotype 1, acting as the reference sequence during alignment. The numbers of occurrences of the haplotypes among direct sequences (#) and cloned sequences (@) were as indicated. *GenBank accession number AF072020.

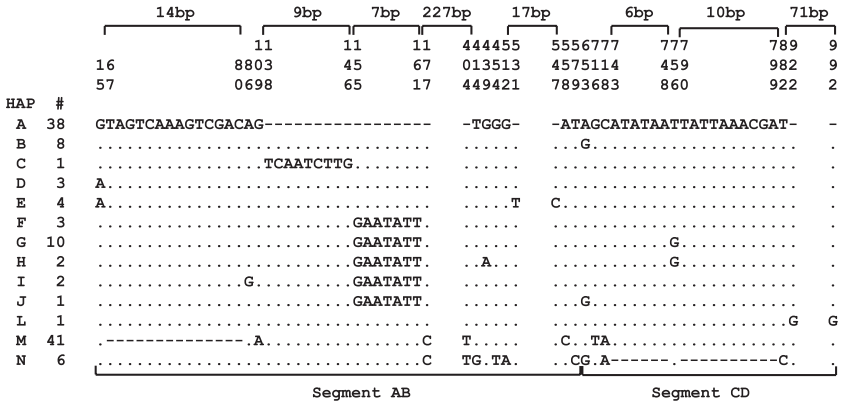

Fig. 3 Sequence alignment of the variable sites in the cpDNA *trn*T–*trn*L spacer that define the 13 haplotypes of *Carapichea ipecacuanha*. Each fragment spans 1041 bases. Dots indicate similarity to haplotype A and hyphens indicate gaps. Numbers on top indicate the nucleotide position with haplotype A acting as the reference sequence during alignment. The numbers of occurrences of each haplotype (#) were as indicated. Only the first base and the last base of indels larger than 15 bp are shown.

Genealogical and geographic relationships among lineages

Genetic structure was assessed with MJ networks, which were constructed for the cpDNA and nuclear ITS datasets separately. The results from both networks showed that genetic variation was partitioned into distinct lineages, which we referred to as ribogroups for the nuclear ITS dataset and haplogroups for the cpDNA dataset (Fig. 4). In both networks, the lineages were organized around the most frequent haplotypes, with the least common haplotypes being found at the periphery of the network. The genetic structure together with the geographic distributions of the ribogroups (Fig. 5) and haplogroups (Fig. 6) showed that there was a clear differentiation among lineages according to range. We found a complete reciprocal monophyly, such that ranges shared neither ITS haplotypes (Fig. 5) nor cpDNA haplotypes (Fig. 6). Both nuclear and chloroplast datasets showed that the genetic diversity of the Amazonian range was depauperate, whereas the Atlantic range showed high levels of genetic diversity. However, further samplings throughout the Central American–Colombian range are needed to confirm the trend of low genetic diversity that was inferred from population BCI.

The MJ network that was constructed with the nuclear dataset separated the Atlantic ITS haplotypes

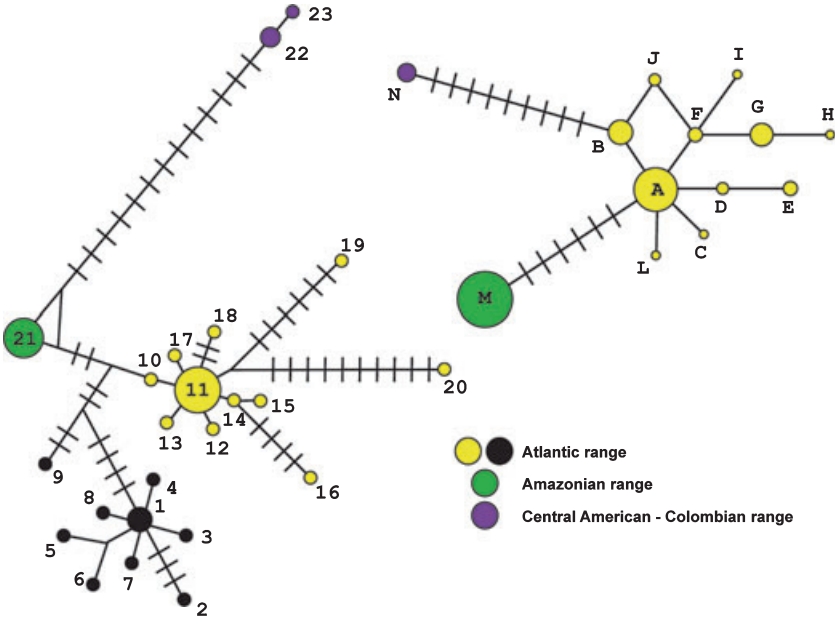

Fig. 4 Median-joining networks depicting the relationships among genealogical lineages of *Carapichea ipecacuanha* based on nuclear ITS haplotypes (left) and cpDNA *trn*T–*trn*L haplotypes (right). Circles represent ITS haplotypes (coded with numbers) or cpDNA haplotypes (coded with letters) and size is proportional to the relative frequencies. Numbers of substitutions are indicated with bars when more than one.

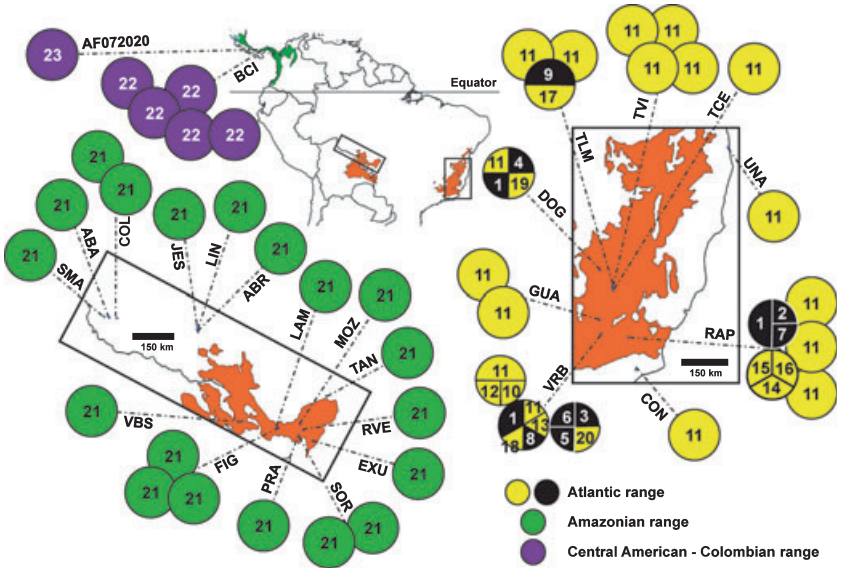

Fig. 5 Geographic distribution of the 23 ITS haplotypes of *Carapichea ipecacuanha*. Each circle represents a specimen; a number within a circle denotes the ITS haplotype; more than one number per circle denotes a specimen with intra-individual polymorphism for ITS. Population codes are as given in Table 1. AF072020 corresponds to a GenBank accession number. Ecoregions are as in Fig. 1.

into two highly divergent ribogroups, which were organized around the high frequency haplotypes 1 (referred to as Atlantic ribogroup I) and 11 (referred to as Atlantic ribogroup II), respectively (Fig. 4). ITS haplotypes 1 and 10 connected the two Atlantic ribogroups through seven intermediate missing haplotypes (Fig. 4). Members of Atlantic ribogroup I were less frequent and were identified in clones from five specimens only. Among the seven specimens that showed intra-individual variation for ITS sequences, we found one specimen in which all its ITS sequences belonged to Atlantic ribogroup I and two specimens in which all their ITS sequences belonged to Atlantic ribogroup II, whereas four specimens had ribosomal ITS sequences from both Atlantic ribogroups (Fig. 5). The MJ network showed that the ITS haplotypes of the Amazonian range (referred to as ribogroup III) had a central position and were connected with the two Central American–Colombian haplotypes (referred to as ribogroup IV) via 13 missing intermediate haplotypes and with the Atlantic ribogroups via four and nine missing intermediate haplotypes, respectively (Fig. 4).

In the MJ network that was based on the cpDNA dataset, the Atlantic haplogroup occupied the central position and was related more closely to the Amazonian haplogroup than to the Central American–Colombian haplogroup. The Atlantic haplogroup connected the Amazonian and Central American-Colombian haplogroups through six and nine missing intermediate haplotypes, respectively. The Amazonian and the

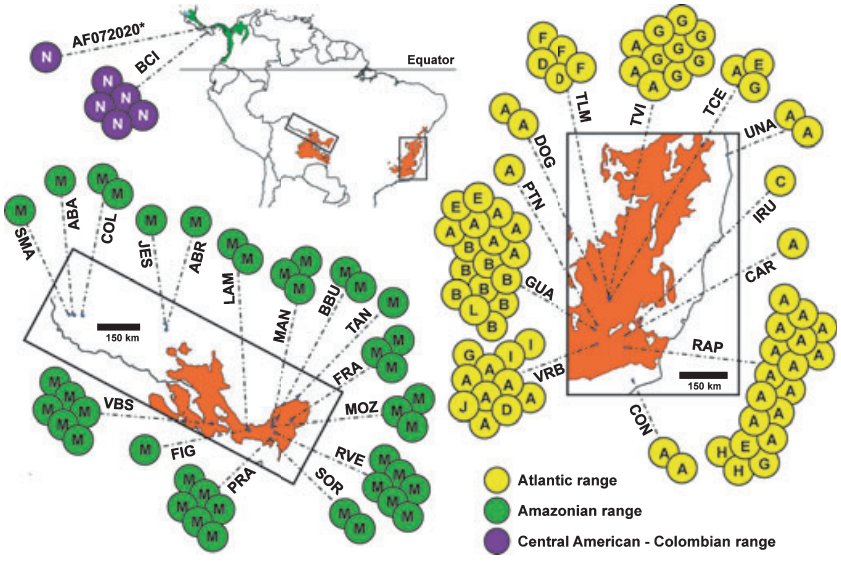

Fig. 6 Geographic distribution of the 13 cpDNA *trn*T–*trn*L haplotypes of *Carapichea ipecacuanha*. Each circle represents a specimen; a letter within a circle denotes the cpDNA haplotype. Population codes are as given in Table 1. AF072020* corresponds to the *trn*T–*trn*L sequence associated with the ITS sequence of GenBank accession AF072020. Ecoregions are as in Fig. 1.

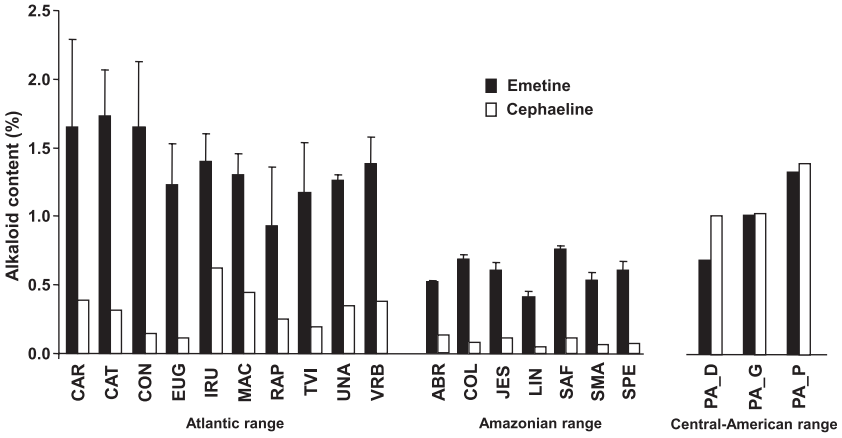

Fig. 7 Concentrations of emetine and cephaeline in dried roots of *Carapichea ipecacuanha*. Left, Brazilian populations evaluated for this study. Each bar on the graph represents the mean value of the alkaloid content of the population and vertical lines show standard deviations. Right, alkaloid content taken from data available in the literature: PA_G, Gupta *et al.* (1986); PA_D and PA_P, Hatfield *et al.* (1981). Population codes are as given in Table 1.

Central American-Colombian haplogroups possessed a single haplotype each. The Atlantic haplogroup contained 11 members, among which haplotype A was at the centre of a single clade. Haplotype A was identified in 38 out of 120 specimens examined and was the most widely distributed haplotype across the Atlantic range, being found in 10 out of 12 populations (Fig. 6).

The levels of emetine and cephaeline varied among the ranges (Fig. 7). Atlantic ipecacs accumulated significantly ($P < 0.001$) higher levels of emetine (1.37 ± 0.25; mean ± SD) than Amazonian ipecacs (0.59 ± 0.11; mean ± sd). Likewise, Atlantic ipecacs accumulated significantly ($P < 0.001$) more cephaeline (0.32 ± 0.15; mean ± SD) than Amazonian ipecacs (0.09 ± 0.03; mean ± sd). The content of emetine was always higher than that of cephaeline in all Brazilian populations examined. The ratio of emetine to cephaeline content (on average 4.28 in Atlantic ipecacs and 6.60 in Amazonian ipecacs) varied from 2.25 (in IRU) to 11.07 (in CON). Nevertheless, these results contrasted sharply with the ratio of emetine to cephaeline in Central-American ipecacs (0.88, on average), in which cephaeline was the most prevalent alkaloid.

Discussion

Presence of ipecac with respect to vegetation formation

Our inventory of occurrences of ipecac indicated that Brazilian ipecacs occurred mostly in seasonal semideciduous forests, with few occurrences in adjacent moist evergreen forests; whereas Central American–Colombian ipecacs occurred in moist evergreen forests. The three geographic ranges of ipecac were separated on a continent-wide scale and were each surrounded by a vast landscape of apparently unsuitable environments. For example, in the Atlantic range, the occurrence of ipecac corresponded remarkably well with the distribution of seasonal semideciduous forests in the Mata Atlântica Biome and was not found in conjunction with the adjacent vegetation, such as the cerrado of Central Brazil and drier forest formations. The Mato Grosso subrange appeared to be limited to a 'peninsula' of seasonal semideciduous forests that were immersed in a vast area of cerrado. These patterns of distribution suggest that the affinity of Brazilian ipecacs is for the understory of seasonal semideciduous forests.

In the Central American–Colombian range, ipecacs occur in the understory of moist evergreen forests. For example, they occur in the hyper-wet Chocó-Darién forests. This northern range may have experienced extended periods of genetic isolation, given that Central American–Colombian ipecacs can be distinguished from Brazilian ipecacs on the basis of their root alkaloid profiles, divergent ITS and cpDNA lineages and the vegetation formations in which they occur. The occurrence of ipecacs in the understory of both seasonal and moist forests is puzzling and deserves further investigation.

Evidence for four lineages

The high levels of genetic differentiation and the sharp breaks in the phylogeographic structures of cpDNA and ITS sequences among ipecacs constitute strong evidence that extant populations of ipecac are descendents of at least four distinct ancestral lineages. The Atlantic range harboured two lineages, whereas the Central American–Colombian range (represented in this study by the single population sampled in Barro Colorado Island, Panama) and the Amazonian range each harboured one lineage. The fact that the lineages were connected to each other through multiple missing intermediate haplotypes suggests long periods of genetic isolation. It is plausible that the phylogeographic signals in ipecac were preserved in a region-dependent manner, which could result in the ancestral–descendent relationships between these lineages appearing distorted in our analyses. For example, one cannot rule

out the possibility that the greater stability of vegetation in some refugial areas of southeastern Brazil (Carnaval & Moritz 2008; Carnaval *et al.* 2009) contributed to the long-term persistence of ipecac in those regions and therefore played a crucial role in preserving high levels of polymorphism (a characteristic of ancestral lineages). In contrast, recent drastic replacements of vegetation in southwestern Amazônia (Burbridge *et al.* 2004; Anhuf *et al.* 2006) might have caused the Amazonian lineage to become genetically depauperate (a characteristic of derived lineages). Palynological data gathered at Lagunas Bela Vista and Laguna Chaplin (which are sites located across the Bolivian border, ~200 km south of population ABR and ~130 km northwest of population VBS, respectively) suggest that the Madeira-Tapajos moist forests only became established in the region during the last few thousand years (Gosling *et al.* 2009). Here, we have refrained from making further inferences about the Central American–Columbian lineage because we only sampled a single population. The scenario described above provides a plausible explanation for the discrepancies in connectivity among ipecac lineages that were observed between the cpDNA and ITS networks. However, caution is required when interpreting the results. Therefore, we have avoided making inferences about ancestral–descendent relationships among ipecac lineages.

The Atlantic lineage

The Atlantic range has a complex phylogeographic structure, which indicates the long-term persistence of populations in this region. The existence of two ITS ribogroups and only one cpDNA haplogroup in this range is intriguing and suggests the existence of a 'missing' haplogroup. The effective population size of the chloroplast genome is one quarter to half that of the nuclear genome; therefore, the genetic diversity of chloroplast genes is more prone to decrease after changes in population size than that of nuclear genes (Petit *et al.* 1993). At this time, we cannot rule out the possibility that the 'missing' haplogroup might still be present in the Atlantic range, but at a frequency so low that it did not appear among our samples.

Some specimens of Atlantic ipecac harboured more than one copy of the ITS (up to five in one specimen of VRB), which indicates that these copies escaped homogenization via concerted evolution (Elder & Turner 1995). The presence of a zone of admixture around the Serra da Mantiqueira mountains that reconnected the two formerly isolated lineages (Atlantic ribogroups I and II) provides a plausible explanation for the intra-individual polymorphism in the ITS sequence among Atlantic ipecacs. According to this scenario, both lineages experienced genetic isolation in a different refuge and, over time, diverged independently through genetic drift, prior to hybridizing more recently. Intraregional variation is another possible scenario that could account for the observed pattern of intra-individual polymorphism in the ITS. However, this latter option does not explain how the two ITS ribogroups, in sympatry, could evolve the high level of genetic differentiation that we uncovered.

The Amazonian and Central American–Colombian lineages

The fact that the entire Amazonian range contained ipecacs that were monomorphic with respect to ITS and cpDNA sequences suggests that Amazonian ipecacs underwent a strong genetic bottleneck. Results obtained previously with inter-simple sequence repeat (ISSR) markers also pointed to low levels of genetic diversity and low differentiation among Amazonian ipecacs (Rossi *et al.* 2009). Such a pattern suggests a recent massive expansion of the parental source from a refuge to colonize a new territory. Had our geographic sampling been more extensive, we would be able to discuss the evolutionary history of ipecac in the Central American–Colombian range further. However, the existence of a distinct lineage from a fourth refuge is apparent from the high level of genetic differentiation that we detected in Central American ipecacs from Barro Colorado Island, Panama. Whether the distribution of genetic diversity of ipecac in the Central American–Colombian range follows the Amazonian pattern, which exhibited monomorphic populations throughout the entire range or the Atlantic pattern, which contained distinct polymorphic lineages, remains to be investigated.

Previous investigations reported that whereas emetine predominated in ipecac from Brazil, cephaeline was the major alkaloid in roots collected in Central America (Costa 1978; Hatfield *et al.* 1981; Gupta *et al.* 1986). Our results expanded these findings by showing that the content of emetine in Brazilian ipecacs was higher than that of cephaeline at the population level. Moreover, we found that the alkaloid profile was, to a certain extent, associated with phylogeographic structure. We postulated that an early event, either through vicariance or long-distance dispersion, split Central American-Colombian ipecacs from Brazilian ipecacs and the lack of gene flow thereafter led to genetic differentiation of the two gene pools at the level of alkaloid biosynthesis, which resulted in distinct alkaloid profiles. However, a more recent event, which split Amazonian ipecacs from Atlantic ipecacs, did not result in such profound changes in alkaloid profile.

Molecular biogeography

The phylogeographic structure of ipecac suggests that episodes of genetic isolation intermingled with extinction and dispersal events shaped the evolutionary history of this species. Currently, the ecological conditions of the open vegetation formations found in central Brazil, which are represented by the many variant forms of cerrado (Oliveira-Filho & Ratter 1995; Daly & Mitchell 2000), constitute a strong barrier to genetic exchange and connectivity among the populations of ipecac that are located in the Atlantic and Amazonian ranges. However, some palynological evidence and computer simulations suggest that in the past, temporary corridors connected the Atlantic coast to the Amazon and to central Brazil (Auler *et al.* 2004; Mayle 2004; Carnaval & Moritz 2008). The exact timing, duration and geographic extension of these corridors are not known, but it has been argued that the corridors emerged during warm humid periods when open vegetation, such as the cerrado, retreated and canopy-providing vegetation expanded (Oliveira-Filho & Ratter 1995; Costa 2003; Pennington *et al.* 2004). The existence of such corridors might have facilitated the dispersal of ipecac across otherwise inhospitable landscapes.

The pattern of distribution of genetic diversity that we have described herein is unlikely to be unique to ipecac. It is likely to be part of a larger phylogeographic signature that was retained as codistributed species of the associated flora experienced a common history of climatic changes. Understanding the comparative phylogeography of such groups will aid the development and evaluation of biogeographical hypotheses about the evolutionary history of the neotropical flora and have implications for the conservation of species.

Acknowledgements

We would like to express our sincere gratitude to the following people for providing assistance during field trips and sampling: Elibeu do Carmo e Silva (SEDAN-RO), Joseph Wright (STRI), Maria Antonia Carniello (UNEMAT), Paulo Fermiano da Silva (SEDAN-RO), Osvaldo S. Rossi, and Sebastião Marcelo de Oliveira (CEPLAC-RO). We thank Molly Nepokroeff for kindly providing cpDNA sequences of a Panamanian accession. João A. A. Meira Neto and Ana Carolina O. Q. Carnaval provided valuable comments on an earlier version of the manuscript. The authors are grateful to anonymous reviewers for their suggestions and comments, which improved the final version of this manuscript greatly. IEF-MG, IBAMA and ANAM provided collecting permits. This project was funded by grants from the International Foundation for Science (no. D/2815), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG—PPM-00148-09), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq—471127/2006-0). CAPES FAPEMIG, FENORTE, and UNEMAT also provided fellowships.

References

Ab'Saber AN (1977) Os domínios morfoclimáticos da América do Sul. *Primeira Aproximação. Geomorfologia*, **53**, 1–23.

Anhuf D, Ledru M-P, Behling H *et al.* (2006) Paleoenvironmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **239**, 510–527.

Assis MC, Giulietti AM (1999) Morphological and anatomical differentiation in populations of ''ipecacuanha''—*Psychotria ipecacuanha* (Brot.) Stokes (*Rubiaceae*). *Revista Brasileira de Botânica*, **22**, 205–216.

Auler AS, Wang X, Edwards RL *et al.* (2004) Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *Journal of Quaternary Science*, **19**, 693–701.

Bandelt H-J, Forster P, Rohl A (1999) Median-Joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Bigarella JJ, Andrade-Lima D, Riehs PJ (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.

Brako L, Zarucchi JL (1993) Catalogue of the flowering plants and gymnosperms of Peru. *Monographs in Systematic Botany from the Missouri Botanical Garden*, **45**, 1–1286.

Bruneton J (1995) *Pharmacognosy, Phytochemistry, Medicinal Plants*. Lavoisier, France, Paris.

Burbridge RE, Mayle FE, Killeen TJ (2004) Fifty-thousand-year vegetation and climate history of Noel Kempff Mercado National Park, Bolivian Amazon. *Quaternary Research*, **61**, 215–230.

Caetano S, Prado D, Pennington RT *et al.* (2008) The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Molecular Ecology*, **17**, 3147–3159.

Cardoso MA, Provan J, Powell W, Ferreira PCG, de Oliveira DE (1998) High genetic differentiation among remnant populations of the endangered *Caesalpinia echinata* Lam. (Leguminosae-Caesalpinioideae). *Molecular Ecology*, **7**, 601–608.

Carnaval AC, Moritz C (2008) Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.

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

Costa AF (1978) *Farmacognosia*. Calouste Gulbekian, Lisbon.

Costa LP (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.

Daly DC, Mitchell JD (2000) Lowland vegetation of tropical South America—an overview. In: *Imperfect Balance: Landscape Transformations in the Pre-Columbian America* (ed Lentz D), pp. 391–454. Columbia University Press, New York.

Daly DC, Silveira M (2008) *Primeiro catálogo da flora do Acre, Brasil/First catalogue of the flora of Acre, Brazil*. EDUFAC, Rio Branco.

Dubs B (1998) Prodromus florae matogrossensis—Rubiaceae. *The Botany of Mato Grosso*, **3**, 1–254.

Elder JF, Turner BJ (1995) Concerted evolution of repetitive DNA sequences in eukaryotes. *Quaternary Review of Biology*, **70**, 297–320.

Funk V, Hollowell T, Berry P, Kelloff C, Alexander SN (2007) Checklist of the plants of the Guiana shield (Venezuela: Amazonas, Bolivar, Delta Amacuro; Guyana, Surinam, French Guiana). *Smithsonian Institution Contributions from the United States National Herbarium*, **55**, 1–584.

Garcia RMA, Oliveira LO, Moreira MA, Barros WS (2005) Variation in emetine and cephaeline contents in roots of wild ipecac (*Psychotria ipecacuanha*). *Biochemical Systematics and Ecology*, **33**, 233–243.

Gentry AH (1982) Neotropical floristic diversity: phytogeographical 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.

Gosling WD, Mayle FE, Tate NJ, Killeen TJ (2009) Differentiation between Neotropical rainforest, dry forest, and savannah ecosystems by their modern pollen spectra and implications for the fossil pollen record. *Review of Palaeobotany and Palynology*, **153**, 70–85.

Gupta MP, Cedeno JE, Soto AA, Correa MDA (1986) Seasonal variation in the alkaloid content of Panamanian ipecac. *Fitoterapia*, **57**, 147–151.

Hatfield GM, Arteaga L, Dwyer JD, Arias TD, Gupta MP (1981) An investigation of Panamanian Ipecac: botanical source and alkaloid analysis. *Journal of Natural Products*, **44**, 452–456.

IBGE (2004) *Mapa de Vegetação do Brasil*. IBGE, Rio de Janeiro, Brazil.

Jørgensen PM, León-Yánez S (1999) Catalogue of the vascular plants of Ecuador. *Monographs in Systematic Botany from the Missouri Botanical Garden*, **75**, 1–1182.

Jung-Mendaçolli SL (2007) Rubiaceae. In: *Flora Fanerogâmica do Estado de São Paulo*, Vol. 5 (eds Wanderley MGL, Shepherd GJ, Melhem TS, Giulietti AM), pp. 259–460. Instituto de Botânica, São Paulo.

Killeen TJ, Chavez E, Pena-Claros M *et al.* (2006) The Chiquitano Dry Forest, the transition between humid and dry forest in eastern lowland Bolivia. In: *Neotropical Savannas and Seasonally Dry Forests: Diversity, Biogeography and Conservatio* (eds Pennington RT, Ratter JA, Lewis GP), pp. 213–233. CRC Press, Boca Raton.

Ledru MP (1993) Late quaternary environmental and climatic changes in central Brazil. *Quaternary Research*, **39**, 90–98.

Ledru M-P, Salatino MLF, 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.

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

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Neter J, Wasserman W (1974) *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*. Irwin, Homewood.

Oliveira LO, Martins ER (1998) *O Desafio das Plantas Medicinais Brasileiras: I-O Caso da Poaia (Cephaelis ipecacuanha)*. UENF, Campos dos Goytacazes.

Oliveira LO, Martins ER (2002) A quantitative assessment of genetic erosion in ipecac (*Psychotria ipecacuanha*). *Genetic Resources and Crop Evolution*, **49**, 607–617.

Oliveira-Filho AT, Fontes MAL (2000) Patterns of floristic differentiation among atlantic forests in southeastern Brazil and the influence of climate. *Biotropica*, **32**, 793–810.

Oliveira-Filho AT, Ratter JA (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 AT, Ratter JA (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 AT, Jarenkow JA, Rodal MJN (2006) Floristic relationships of seasonally dry forests of eastern South America based on tree species distribution patterns. In: *Neotropical Savannas and Seasonally Dry Forests: Diversity, Biogeography and Conservatio* (eds Pennington RT, Ratter JA, Lewis GP), pp. 159–192. CRC Press, Boca Raton.

Olson DM, Dinerstein E, Wikramanayake ED *et al.* (2001) Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**, 933–938.

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

Pennington RT, Lewis GP, Ratter JA (2006) An overview of the plant diversity, biogeography and conservation of neotropical savannas and seasonally dry forests. In: *Neotropical Savannas and Seasonally Dry Forests: Diversity, Biogeography and Conservatio* (eds Pennington RT, Ratter JA, Lewis GP), pp. 1–30. CRC Press, Boca Raton.

Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plants. *Heredity*, **71**, 630–641.

Prado DE (2000) Seasonally dry forests of tropical South America: from forgotten ecosystems to a new phytogeographic unit. *Edinburgh Journal of Botany*, **57**, 437–461.

Prado DE, Gibbs PE (1993) Patterns of species distribution in the dry seasonal forest of South America. *Annals of the Missouri Botanical Garden*, **80**, 902–927.

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

Ribera MO, Libermann M, Beck S, Moraes M (1994) *Mapa de la Vegetacion y Areas Protegidea de Bolivia. 1:1,500,000*. Centro de Investigaciones y Manejo de Recursos Naturales (CIMAR) and Universidad Autónoma Gabriel Rene Moreno (UAGRM), La Paz, Bolivia.

Rossi AAB, Oliveira LO, Venturini BA, Silva RS (2009) Genetic diversity and geographic differentiation of disjunct Atlantic and Amazonian populations of *Psychotria ipecacuanha* (Rubiaceae). *Genetica*, **136**, 57–67.

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.

Taylor CM (2006) Carapichea, Margaritopsis, Notopleura. In: *Flora of Ecuado* (eds Harling G, Persson C), pp. 1–74. University of Gotehnburg, Gotehnburg.

Veloso HP (1947) As condições ecológicas da *Cephaelis ipecacuanha* Rich. *Memórias do Instituto Oswaldo Cruz*, **45**, 361–372.

Veloso HP, Rangel Filho ALR, Lima JCA (1991) *Classificação da Vegetação Brasileira, Adaptada de um Sistema Universal*. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.

White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Application* (eds Innis MA, Gelfand DH, Sninsky JJ, White TJ), pp. 315–322. Academic Press, London.

Whitmore TC, Prance GT (1987) *Biogeography and Quaternary History in Tropical America*. Oxford Science Publications, Oxford.

An interest in the genetic diversity of native medicinal species was the common subject that brought this research team together. LOO is a molecular geneticist with broad interests in the geographic distribution of genetic diversity in native plant species. AABR is a plant geneticist and RSS is a botanist; both are interested in conservation genetics and the flora of Southwestern Amazonia. FRCB, an undergraduate student of biochemistry, seized this opportunity to advance her laboratory skills in molecular biology. ERM, a plant geneticist, is concerned about the genetic resources and cultivation of medicinal plant species.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Fig. S1 Partial electrophretograms from direct sequences of the ITS-1 spacer from two individuals of *Carapichea ipecacuanha*. Top: RAP1 sequence with two overlapping double peaks that resulted in the ambiguous readings R and Y, an indicative of ITS paralogues of same length. Bottom: TVI12 sequence with no double peaks at the corresponding sites. R, A or G; Y, C or T.

Table S1 Summary of herbarium and population data of *Carapichea ipecacuanha* used in this study

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.