Journal of Biogeography (J. Biogeogr.) (2013) **40**, 707–719

SPECIAL ISSUE

Amazon diversification and cross-Andean dispersal of the widespread Neotropical tree species *Jacaranda copaia* (Bignoniaceae)

Caroline Scotti-Saintagne[1]*, Christopher W. Dick[6,7], Henri Caron[2], Giovanni G. Vendramin[3], Valérie Troispoux[1], Pierre Sire[1], Maxime Casalis[1], Anna Buonamici[3], Renato Valencia[4], Maristerra R. Lemes[5,8], Rogerio Gribel[5,8] and Ivan Scotti[1]

[1]*INRA–UMR ECOFOG campus agronomique, BP 709, 97387, Kourou Cedex, France,* [2]*INRA–UMR BIOGECO 69, 33312, Cestas Cedex, France,* [3]*Istituto di Genetica Vegetale, Sezione di Firenze, Consiglio Nazionale delle Ricerche, 50019, Sesto Fiorentino (Firenze), Italy,* [4]*Escuela de Ciencias Biológicas, Laboratorio de Ecología de Plantas, P. Universidad Católica del Ecuador, Apartado 17-01-2184, Quito, Ecuador,* [5]*Instituto Nacional de Pesquisas da Amazônia, Laboratório de Genética e Biologia Reprodutiva de Plantas, 69083-000, Manaus, AM, Brazil,* [6]*Department of Ecology and Evolutionary Biology and Herbarium, University of Michigan, Ann Arbor, MI, 48109, USA,* [7]*Smithsonian Tropical Research Institute, Gamboa, Republic of Panama,* [8]*Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rua Pacheco Leão 915, Jardim Botânico, 22460-030, Rio de Janeiro, RJ, Brazil*

*Correspondence: Caroline Scotti-Saintagne, INRA UMR 0745 EcoFoG ('Ecologie des forêts de Guyane'), Campus Agronomique, BP709, 97387 Kourou, French Guiana (France). E-mail: caroline.scotti@ecofog.gf

ABSTRACT

Aim The phylogeographical history of Neotropical species can be difficult to reconstruct because of superimposed Neogene and Quaternary histories, and because of taxonomic uncertainty. We analysed range-wide genetic diversity in a widespread pioneer tree species, *Jacaranda copaia* (Aubl.) D. Don, to characterize phylogeographical structure, date the evolutionary relationships among lineages, and evaluate the role of dispersal and vicariance in establishing the present geographical range.

Location Guiana Shield; central, southern and western Amazon Basin; Chocó region; Central America.

Methods We analysed nine nuclear simple sequence repeat loci (nuSSR), eight chloroplast SSRs (cpSSR), and two cpDNA intergenic sequences in 341 adult trees. Genetic differentiation at nuSSRs was inferred using Bayesian clustering. Dating of chloroplast lineage divergence was obtained using a range of published mutation rates and Bayesian coalescence analyses. Population divergence dating was performed using an isolation-with-migration model for eight loci (one cp sequence and seven nuSSRs).

Results Nuclear SSR variation identified three geographically overlapping clusters (*nu*-1, *nu*-2, *nu*-3). Twelve cpDNA haplotypes were clustered into two haplogroups (*cp*-1, *cp*-2) with the central Amazon harbouring the highest diversity. Molecular dating analysis suggests that cpDNA haplotype diversification started around the end of the Pliocene (2.61 Ma on average), whereas population divergence was more recent and occurred during the mid-Quaternary (point estimates between 357 and 436 ka).

Main conclusions The genetic variation of *J. copaia* in the Neotropics was shaped mainly by Pleistocene events. Chloroplast diversity did not display the expected *cis*/*trans* Andean disjunction, indicating recent dispersal. Nuclear variation revealed that separate regions share a recent history, with a centre of diversity in the central Amazon Basin. The geographical pattern of diversity is congruent with the distribution of the two subspecies, *J. copaia copaia* and *J. copaia spectabilis*, and evidence of nuSSR admixture between the two taxa supports their classification as subspecies.

Keywords

Amazonia, Bignoniaceae, centre of diversity, dispersal, historical biogeography, Neotropics, phylogeography, Pleistocene, tropical forest, vicariance.

Journal of Biogeography

INTRODUCTION

The biogeographical history of Neotropical regions remains uncertain despite the scientific importance of understanding mechanisms involved in the maintenance of its vast biodiversity. Several hypotheses have been proposed to explain tropical diversification but no general model has been accepted (Willig *et al.*, 2003; Mittelbach *et al.*, 2007). Diversification in Neotropical forest tree lineages has taken place throughout the Cenozoic (Pennington *et al.*, 2004; Rull, 2008). Major geological events, such as the uplift of the northern Andes (reaching modern elevations during the middle Pliocene; Gregory-Wodzicki, 2000), or the final closure of the Isthmus of Panama (3.5 Ma; Keigwin, 1982), may have shaped species distributions and divergence patterns during the Neogene. The driving forces behind Quaternary diversification are still under discussion. One hypothesis, Haffer's (1969) 'refuge' model, is that a drier climate favoured the expansion of savanna (van der Hammen & Hooghiemstra, 2000) and seasonally dry forests (Pennington *et al.*, 2004), resulting in the fragmentation of the rain forest and allopatric speciation (Mayr, 1954). This model of recent allopatric speciation has been strongly criticized and challenged by scenarios involving older speciation events and less dramatic vegetation changes during the Pleistocene (e.g. Bush, 1994; Aide & Rivera, 1998; Colinvaux *et al.*, 2000; Knapp & Mallet, 2003; Bush & Oliveira, 2006).

Phylogeography, which connects gene genealogy to geography (Avise, 2000), provides methods to infer intraspecific evolutionary history and to test hypotheses on the impact of Quaternary (and earlier) events on evolutionary divergence. Phylogeography also provides a conceptual framework for developing conservation strategies aiming to protect species and genetic diversity, as well as to ensure the continuation of evolutionary processes that maintain them (Moritz & Faith, 1998; Moritz, 2002). This is particularly important in a context of global climate changes where perspectives on the fate of tropical forests are rather pessimistic (Hubbell *et al.*, 2008).

In the Neotropics, broad phylogeographical studies of rain forest trees are limited to only a few species (e.g. Dick *et al.*, 2007; Dick & Heuertz, 2008; Rymer *et al.*, 2013; Scotti-Saintagne *et al.*, 2013). Most of these studies have revealed high levels of divergence between geographically isolated regions (Aide & Rivera, 1998; Cavers *et al.*, 2003; Dick *et al.*, 2003; Novick *et al.*, 2003; Hardesty *et al.*, 2010; Lemes *et al.*, 2010; Scotti-Saintagne *et al.*, 2013; but see Dick *et al.*, 2007 and Rymer *et al.*, 2013). Some of the deep phylogeographical patterns are temporally consistent with Neogene events, including the Pliocene uplifts of the Isthmus of Panama and the northern Andean cordilleras, and the late Miocene development of the Amazon drainage basin. Only a limited number of Neotropical studies have invoked Quaternary events to explain intraspecific genetic patterns (Dutech *et al.*, 2002; Poelchau & Hamrick, 2013; Scotti-Saintagne *et al.*, 2013).

We performed a phylogeographical analysis of the Neotropical pioneer tree species *Jacaranda copaia* (Aubl.) D. Don, which is a common canopy tree of Neotropical moist forests ranging from Belize to Brazil and Bolivia (Gentry, 1992), with two subspecies: *J copaia* subsp. *copaia* and *J. copaia* subsp. *spectabilis* (Mart. ex A. DC.) A.H. Gentry. *Jacaranda copaia* subsp. *copaia* is distributed in the Guiana Shield, whereas *J. copaia* subsp. *spectabilis* is widespread in lowland moist and wet forest from Belize to Bolivia (Gentry, 1992). Although the two subspecies can be distinguished by leaf and fruit characters, we made no *a priori* distinction between them due to the possibility of introgression.

To fully address the population history of *J. copaia*, we evaluated genetic diversity at both nuclear and chloroplast microsatellites (nuSSR and cpSSR) as well as chloroplast intergenic sequences (cpDNA). The use of organelle and nuclear markers allows us to capture imprints of historical demographic processes acting on different time-scales because of different modes of inheritance, effective population sizes and mutation rates. Our samples spanned central, southern and western Amazon, the Guiana Shield, Central America and the Andean slopes. Our main objectives were to: (1) characterize range-wide *J. copaia* phylogeographical patterns, (2) determine whether the divergence between lineages and populations could have been influenced by Neogene (geological) or Quaternary (climate change) events, (3) evaluate the role of dispersal and vicariance in establishing the geographical distribution of the genetic diversity, and (4) evaluate a hypothesis for species status of the *J. copaia* subspecies.

MATERIALS AND METHODS

Species description

The genus *Jacaranda* Juss. includes 49 species, which are broadly distributed in tropical and subtropical North and South America and the Caribbean. *Jacaranda copaia* is allogamous (James *et al.*, 1998). Its hermaphrodite flowers are pollinated by large bees (Maués *et al.*, 2008), which in other species have been shown to move pollen over large distances (reviewed in Dick *et al.*, 2008). Seeds weigh less than 2 g (Jones *et al.*, 2005), and are widely dispersed by wind (Dalling *et al.*, 2002; Jones *et al.*, 2005). Inference from fine-scale spatial genetic structure estimated gene dispersal distances between 327 and 1869 m (Hardy *et al.*, 2006), while direct measures indicate that the median seed dispersal from a maternal tree is 24 m (Jones *et al.*, 2005). *Jacaranda copaia* is considered a 'long-lived pioneer species' which experiences rapid growth in light-rich environments, yet it has a high wood density and persists in mature forest (Brokaw, 1985; Guariguata *et al.*, 1995). It is a valued timber tree and is recommended for use in agroforestry, reforestation and degraded land-recovery projects in South and Central America (Junior & Yared, 1991).

Sampling

This study is based on a range-wide sampling of natural populations from field ($n = 273$, see Appendix S1a in

Supporting Information) and herbarium specimens ($n = 68$) for a total of 341 *J. copaia* adult trees (Table 1). Herbarium specimens were obtained from Cayenne (IRD, French Guiana), Manaus (EMBRAPA and INPA, Brazil), Oxford (University Herbaria, UK) and Utrecht (National Herbarium, the Netherlands) (Appendix S1b). Two field-sampling strategies were conducted. For the analysis at nuclear SSRs, at least 30 samples were collected in each of three regions (hereafter named Western Amazon, Central Amazon and Guiana Shield) to obtain a reliable estimate of within- and among-region genetic diversity. The Western Amazon region was sampled at two sites (Napo and Orellana) in eastern Ecuador; Central Amazon was represented by three sites (Tefé, Manaus and Tapajos) in Brazil; and the Guiana Shield was represented by an exhaustive sampling over 21 sites in French Guiana (see Fig. 2). For the clustering analysis we also included five samples from the southern Amazon site of Boca do Acre in Brazil (Southern Amazon region), three samples from the Peruvian site of Jenarro Herera (Western Amazon) and eight samples from the Costa Rican site of Sarapiqui (Central America) (Fig. 2). For the cpDNA analysis, an average coverage of eight individuals per site (between one and 24 individuals) was obtained from 94 sites throughout the Neotropical range (341 samples in total). The DNA extraction of fresh material (leaf and/or cambium tissue) and herbarium samples followed the protocol of Scotti-Saintagne *et al.* (2013).

Nuclear microsatellites

Nuclear genetic diversity was analysed at nine microsatellite loci: JACC1.1, JACC2, JACC2.1, JAC22 from Jones & Hubbell (2003) and JC3A10, JC3C5, JC3F11, JC3F4, JC3H10 from Barthe *et al.* (2012) (see Appendix S1c). The polymerase chain reaction (PCR) protocol follows Barthe *et al.* (2012). Alleles were sized on an automated DNA sequencer (ABI 3130-XL; Applied Biosystems, Foster City, CA, USA) using GeneMapper 4.0 software (Applied Biosystems) and GeneScan LIZ-500 Size Standard (Applied Biosystems).

Chloroplast intergenic regions

Two chloroplast intergenic spacer regions, *trn*H–*psb*A and *trn*C–*ycf*6 (Shaw *et al.*, 2005), were tested for polymorphism within and between populations of *J. copaia* (see Appendix S1c). PCR and sequencing reactions were performed as described in Scotti-Saintagne *et al.* (2013). Sequences were aligned and edited using CodonCode Aligner 3.5.7 (Codoncode Corporation, Dedham, MA, USA). To overcome the high rate of failure in sequencing the degraded DNA of herbarium samples (success rate was 13%), internal primers were designed for a set of intergenic regions harbouring informative single nucleotide polymorphisms (SNPs, see Results). The sequences of these new primers are given in Appendix S1c; they gave a success rate of 78% for herbarium samples (68 samples out of 88 available).

Chloroplast microsatellites

Eight cpSSRs (see Appendix S1c) were genotyped on a subset of 94 individuals in order to estimate the genetic diversity at fast-evolving cp regions (Provan *et al.*, 1999). The amplification conditions are reported in Weising & Gardner (1999). Amplified products were sized on a 96 capillary MegaBACE 1000 automatic sequencer (GE Healthcare, Madison, WI, USA). Analysis of fragments was performed using the

Table 1 Number of samples of *Jacaranda copaia* per type of material collected in 11 countries in the Neotropics and sample size per type of molecular markers.

Region	Country	Fresh samples*	Herbarium samples*	Total sample size	cpDNA haplotype	cpDNA haplogroup	cpSSR	Nuclear SSR
Central Amazon	Brazil	33	13	46	27	46	14	40
Western Amazon	Bolivia		3	3		3		
	Ecuador	63		63	21	32	11	63
	Peru	4		4	6	4	3	3
Southern Amazon	Brazil	6	9	15	3	12	3	5
Guiana Shield	Brazil		3	3	1	3		
	French Guiana	175	18	193	145	193	23	98
	Guiana		8	8	4	8		
	Suriname		7	7	1	7		
	Venezuela		2	2		2		
Central America	Costa Rica	10		10	10	10	4	8
	Panama	13	1	14	7	14	1	
Andes	Columbia		4	4	2	4		
Total		304	68	372	227	338	59	217

*Fresh samples, leaves or pieces of cambium placed in silica gel; herbarium samples, small pieces of leaves of herbarium accessions (for a list of accessions see Appendix S1b).

cp, chloroplast; SSR, simple sequence repeat loci.

software Fragment Profiler 1.2 (GE Healthcare, Madison, WI, USA).

Statistical analyses

A summary of inferences made, sample set used, parameters estimated and software used is given in Fig. 1.

NuSSR diversity and genetic structure

Genetic diversity at the nine nuSSRs was analysed using Arlequin 3.5 (Excoffier *et al.*, 2005). Measures of nuclear diversity were allelic richness (A) and expected heterozygosity (H_E) with their standard deviation. To standardize the allelic richness estimates across populations, the rarefaction procedure implemented in rarefac was used (Petit *et al.*, 1998). The presence of null alleles was tested using micro-checker 2.2.3 (van Oosterhout *et al.*, 2004). Genetic differentiation among the three regions (Central Amazon, Guiana Shield, Western Amazon) was estimated by computing pairwise F_{ST} (genetic differentiation based on allele identity) and R_{ST} (genetic differentiation based on molecular distance between alleles assuming a stepwise mutation model; Slatkin, 1995). For loci with null alleles, a method taking null alleles into account (INA, Chapuis & Estoup, 2007) was used to provide genotype data corrected for null alleles and to re-estimate H_E. Finally, we refined the estimation of F_{ST} by a method excluding null alleles (ENA, Chapuis & Estoup, 2007). Both INA and ENA methods are implemented in FreeNA (Chapuis & Estoup, 2007).

To detect possible genetic substructuring within regions, we ran the genetic clustering algorithm implemented in structure 2.3.3 (Pritchard *et al.*, 2000) assuming a population admixture model. The most likely number of genetic clusters, K, was then inferred from the estimation of ΔK (Evanno *et al.*, 2005). To incorporate geographical coordinates to the results of structure, we also ran tess 2.3.1 (François *et al.*, 2006). More details on Bayesian clustering methods are given in Appendix S2a.

CpDNA diversity and genetic structure

Genetic diversity at chloroplast markers (cpDNA) was analysed using Arlequin 3.5 (Excoffier *et al.*, 2005). CpDNA diversity was measured as number of different haplotypes (N_H) standardized for comparison among populations with rarefac (Petit *et al.*, 1998), and haplotype diversity (H). Confidence intervals were computed as in Grundmann *et al.* (2001). Genetic differentiation among regions at cpDNA was computed based on haplotypic frequencies (F_{ST}) and on molecular distances between haplotypes (N_{ST}). For this purpose, a matrix of genetic distances was generated based on the number of mutational steps between haplotypes. The matrix of pairwise interhaplotypic distances was used to build a minimum spanning network among haplotypes using Arlequin 3.5.

Test for phylogeographical signals

To test whether the distribution of genetic diversity within *J. copaia* followed a phylogeographical pattern, we used the

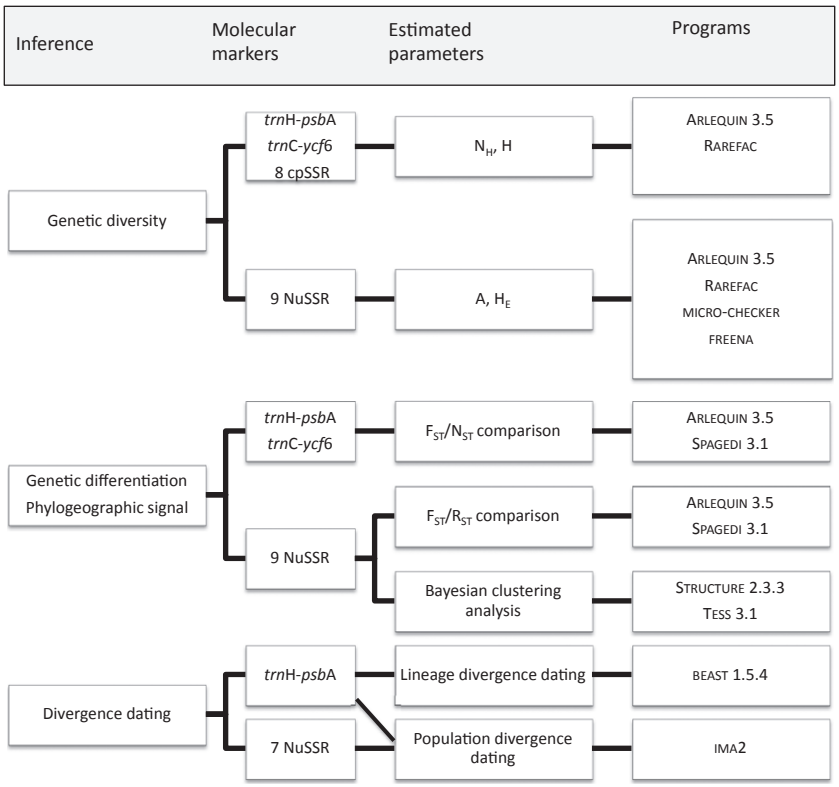

Figure 1 The data set used for the analyses, the estimated parameters and the software program employed for each type of inference.

procedure proposed by Pons & Petit (1996) for cpDNA and the global R statistics for nuSSRs (Hardy *et al.*, 2003), both implemented in SPAGEDI 1.3 (Hardy & Vekemans, 2002). At cpDNA we compared the coefficient of population differentiation G_{ST} (analogue to F_{ST}) with N_{ST}; at nuSSR we compared F_{ST} with $R_{ST.}$ These tests are based on the expectation that N_{ST} and R_{ST} should not be different respectively from G_{ST} and F_{ST} assuming random geographical distribution of evolutionarily related haplotypes. The significance test was based on 20,000 permutations.

Dating of divergence between cp lineages

Divergence time between cpDNA haplotypes was estimated using Bayesian inference on the cp sequence *trnH–psbA*. A publicly available sequence of *Tabebuia rosea* (accession number GQ982378.1) was used as the outgroup. *Tabebuia rosea* was chosen as outgroup because it is the closest relative of *J. copaia* appearing in the phylogeny published by Baraloto *et al.* (2012) for which a *trnH–psbA* sequence was available in public databases. We used BEAST 1.5.4 (Drummond & Rambaut, 2007) to reconstruct the haplotype phylogeny and estimate divergence times at key nodes. Phylogenetic reconstruction was performed using the 'GTR+I' substitution model, which follows a strict molecular clock (for details see Appendix S2b). The model was fitted with a uniform prior distribution for both tree root age and substitution rate, and a coalescent approach assuming constant population size as a tree prior model. The root age prior was constrained to between 20 and 40 Myr. These boundaries were chosen to encompass the divergence time between *Jacaranda* and the rest of Bignoniaceae estimated in a recent phylogeny (Baraloto *et al.*, 2012). In addition, we used a uniform distribution for the clock rate, ranging from 4×10^{-10} to 8.24×10^{-9} substitutions per site per year, which corresponds to the range of substitution rates estimated in the cp genome of plants (Richardson *et al.*, 2001). A Markov chain Monte Carlo (MCMC) was run for 500,000,000 steps, with a 50 million-step burn-in.

Dating of divergence between populations

The time of population divergence was estimated using the isolation-with-migration (IM) model implemented in IMa2 (Hey & Nielsen, 2004; Hey, 2010) based on eight molecular markers including *trn*H–*psb*A and seven nuSSR (two nuSSR with irregular allelic sizes were discarded). Given the limited number of markers available for analysis, we fitted simplified models and included only two populations with the assumption that ancestral population size was identical to that of the largest descendant population. To approximate, as closely as possible, panmictic Wright–Fisher populations, we chose to analyse two populations each belonging to only one STRUCTURE cluster, i.e. the Orellana population in Ecuador and the Tapajos population in Brazil. Analyses were performed on all 20 alleles from the Tapajos population and 20 alleles randomly drawn from the Orellana population. To check that diversity within the subsample was not different from the diversity of the 67 individuals samples, confidence intervals were calculated using the method developed by Grundmann *et al.* (2001). Since the confidence intervals overlapped for all loci, we considered the 20-individual sample as representative of the 67-individual samples.

Two models, with and without migration between populations, were fitted. Four independent runs were carried out for each model using different seed numbers, and using maximum values for the migration (m), the effective population size (q) and the divergence time (t) parameters identified through preliminary exploratory analyses ($q = 100$, $t = 200$, $m = 0.5$). Burn-in was at least 20 million steps for all runs. Thirty million steps were retained in each chain following burn-in and 300,000 genealogies were saved per locus. Mixing of the chains was considered satisfactory when effective sample size > 100 and when the posterior estimates overlapped between runs. Independent estimates of the mutation rates μ were used to convert parameter estimates into units of years. For the cp *trn*H–*psb*A sequence, we used the geometric mean estimated by BEAST multiplied by sequence length (3.51×10^{-9} multiplied by 274). For nuclear markers, we estimated mutation rates from an independent analysis in the population of Tapajos as follows. First, using the mutation rate estimated for the *trn*H–*psb*A cp sequence, we estimated the effective population size, N_e, from theta ($2N_e\mu$), using DNASP 5 (Librado & Rozas, 2009). Next, the mutation rate for each nuSSR locus was estimated from theta ($4N_e\mu$) using the default options of MISAT 1.0 (Nielsen, 1997). The geometric mean of the mutation rate over the seven microsatellite loci (5×10^{-5}) was used in IMa2 with a range of possible mutation rates among loci (2.47×10^{-4}–1.35×10^{-5}).

RESULTS

Nuclear markers

Genetic diversity was analysed at nine nuclear SSR in 201 samples belonging to three geographical regions: Guiana Shield, Central Amazon and Western Amazon (Fig. 2). No significant differences were observed between regional estimates of genetic diversity even when the data set was corrected for null alleles (Table 2). Estimates of global F_{ST} obtained with and without the ENA correction (Chapuis & Estoup, 2007) were 0.09 and 0.10, respectively, with overlapping confidence intervals meaning that the null alleles were evenly distributed among populations. R_{ST} of 0.55 was significantly higher than F_{ST}, suggesting the presence of a phylogeographical pattern. Pairwise F_{ST} (and R_{ST}) between regions revealed a low genetic divergence between Central Amazon and Guiana Shield ($F_{ST} = 0.035$; $R_{ST} = 0.053$) compared with the genetic divergence from Western Amazon ($F_{ST} = 0.10$ and $F_{ST} = 0.14$; $R_{ST} = 0.60$ and $R_{ST} = 0.66$ for central Amazon and Guiana Shield, respectively; Table 3). The STRUCTURE analysis revealed the presence of three nuSSR

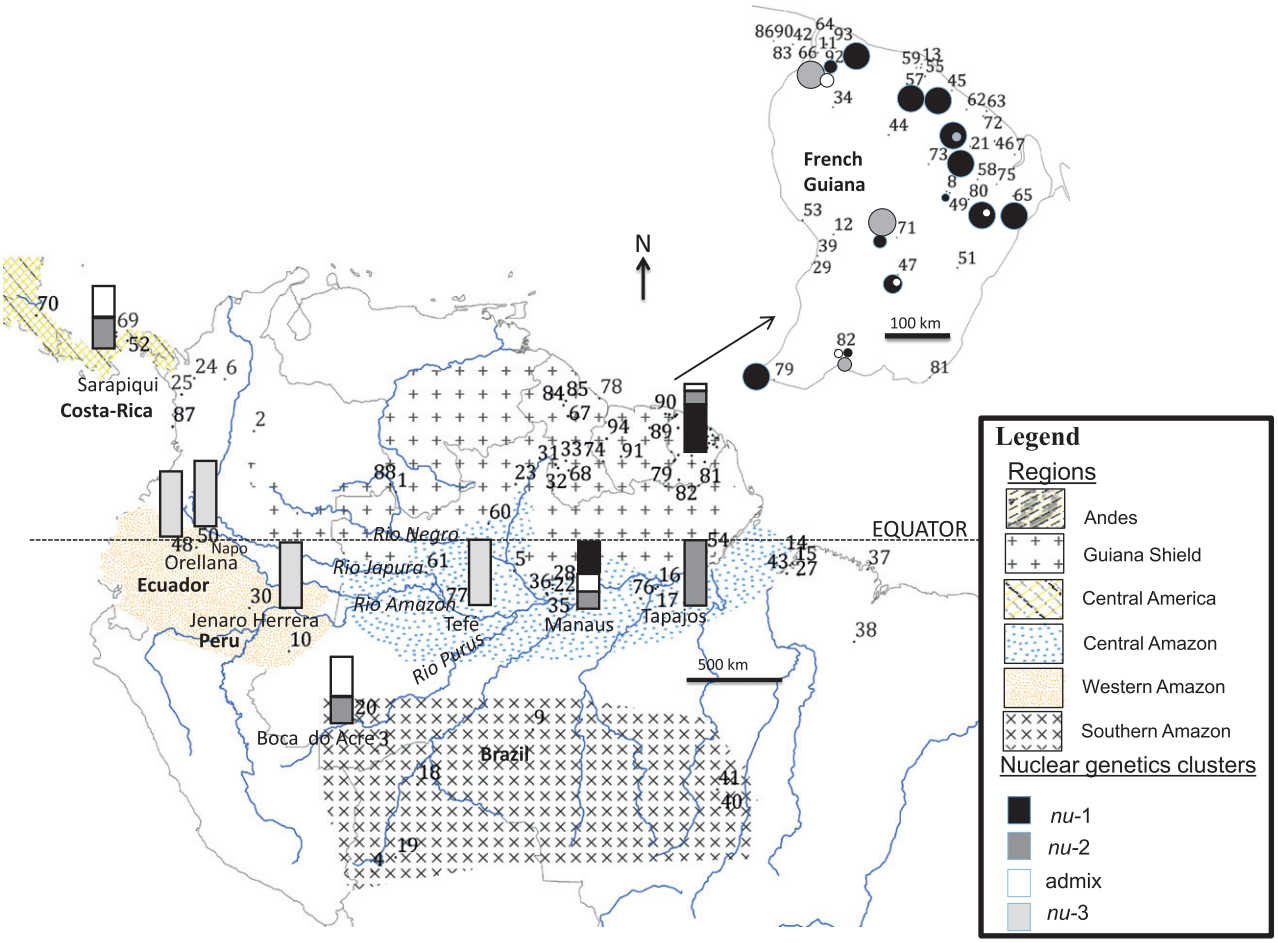

Figure 2 Geographical distribution of three nuclear clusters (*nu*-1, *nu*-2, *nu*-3) inferred by a Bayesian clustering analysis (STRUCTURE) after evaluating genetic diversity at nine nuSSRs over 217 individuals of *Jacaranda copaia* in five regions in the Neotropics (Western Amazon, Southern Amazon, Central Amazon, Guiana Shield, Central America). Label numbers correspond to the sampling site; details in Appendix S1.

Table 2 Summary of genetic diversity in *Jacaranda copaia* at nine microsatellites for each Neotropical region, averaged over loci (standard deviations in parentheses).

	n	N_A	N_{A_r}	H_E	$H_{E_}$cor
Central Amazon (BRA)	38.3*	20.1	20.1 (4.9)	0.90 (0.06)	0.90 (0.05)
Western Amazon (EC)	60	16.1	14.4 (7.4)	0.80 (0.19)	0.81(0.17)
Guiana shield (FG)	93.7	23.6	17.8 (10)	0.85 (0.13)	0.87 (0.09)

*Size used in the rarefaction procedure. BRA, Brazil; EC, Ecuador; FG, French Guiana. *n*, sample size; N_A, number of different alleles; N_{A_r}, number of different alleles after rarefaction; H_E, expected heterozygosity; $H_{E_}$cor, expected heterozygosity corrected for null alleles.

Table 3 Genetic differentiation at nuclear SSRs (F_{ST} and R_{ST}) estimated by pairs of regions in *Jacaranda copaia*.

	Central Amazon	Western Amazon	Guiana Shield
Central Amazon	–	**0.596**	**0.053**
Western Amazon	**0.10 (0.10)**	–	**0.660**
Guiana Shield	**0.035 (0.026)**	**0.14 (0.13)**	–

F_{ST} values are below the diagonal; R_{ST} values are above the diagonal. Bold type denotes significant values (*P* values based on 1000 permutations). F_{ST} values using ENA correction described by Chapuis & Estoup (2007) are shown in parentheses.

clusters (ΔK was maximal for $K = 3$, see Appendix S3a), which will be referred to as *nu*-1, *nu*-2 and *nu*-3. The nuclear clusters lacked any clear geographical structure, as shown in Fig. 2 and as confirmed by the TESS analysis (see Appendix S3b). The clusters *nu*-1 and *nu*-2 were present in the Guiana Shield, in Central Amazon, in Southern Amazon (site of Boca do Acre) and in Central America, and the group *nu*-3 was present in Western Amazon (Ecuador and Peru) and Central Amazon (site of Tefé) (Fig. 2, Appendix S3c).

Chloroplast markers

A higher level of polymorphism was observed in *trnH–psbA* due to the presence of a T/A-rich region. In total, 12 haplotypes were identified from the 226 samples, of which six were due to variability in the T/A rich region (see Appendix S3d). *trn*H–*psb*A and *ycf*6–*trn*C sequences have been deposited in GenBank under accession numbers JN661781–JN661792 and JN661796–JN661807. Two haplogroups emerged from the haplotype network (Fig. 3): *cp*-1, which contains eight haplotypes and *cp*-2, which includes four haplotypes; *cp*-1 and *cp*-2 differed by a minimum of three mutations including a 24 bp insertion/deletion.

Geographical distribution of chloroplast genetic diversity

Because it was not possible to obtain haplotype identity for all samples due to missing data at many SNPs in herbarium samples, we split the data into two data sets: the haplogroup and the haplotype data set. The haplogroup data set is based on the polymorphism at two sites: the 24 bp insertion/deletion in the *trn*H–*psb*A sequence and the SNP at position 138 in the *trn*C–*ycf*6 sequence (Appendix S3d). Both displayed the strongest genetic differentiation between the two cp lineages (F_{ST} = 1). *Cp*-1 is located in the Guiana Shield and Central Amazon whereas *cp*-2 is mainly present in Western and Southern Amazon and Central America (Fig. 3).

Haplotype richness (N_H) and haplotype diversity (H) were highest in the Central Amazon, with seven haplotypes out of 12 (N_H remained higher after the rarefaction correction) and H equal to 0.83 (confidence interval, CI = 0.72–0.88) (see Appendix S3e). In contrast, Central America displayed the lowest diversity (H = 0) with a single haplotype (H9).

Among the eight cpSSRs, two were monomorphic (ccmp1 and ccmp3) and the combination of the six remaining loci defined five haplotypes (see Appendix S3). Haplotypes A, B

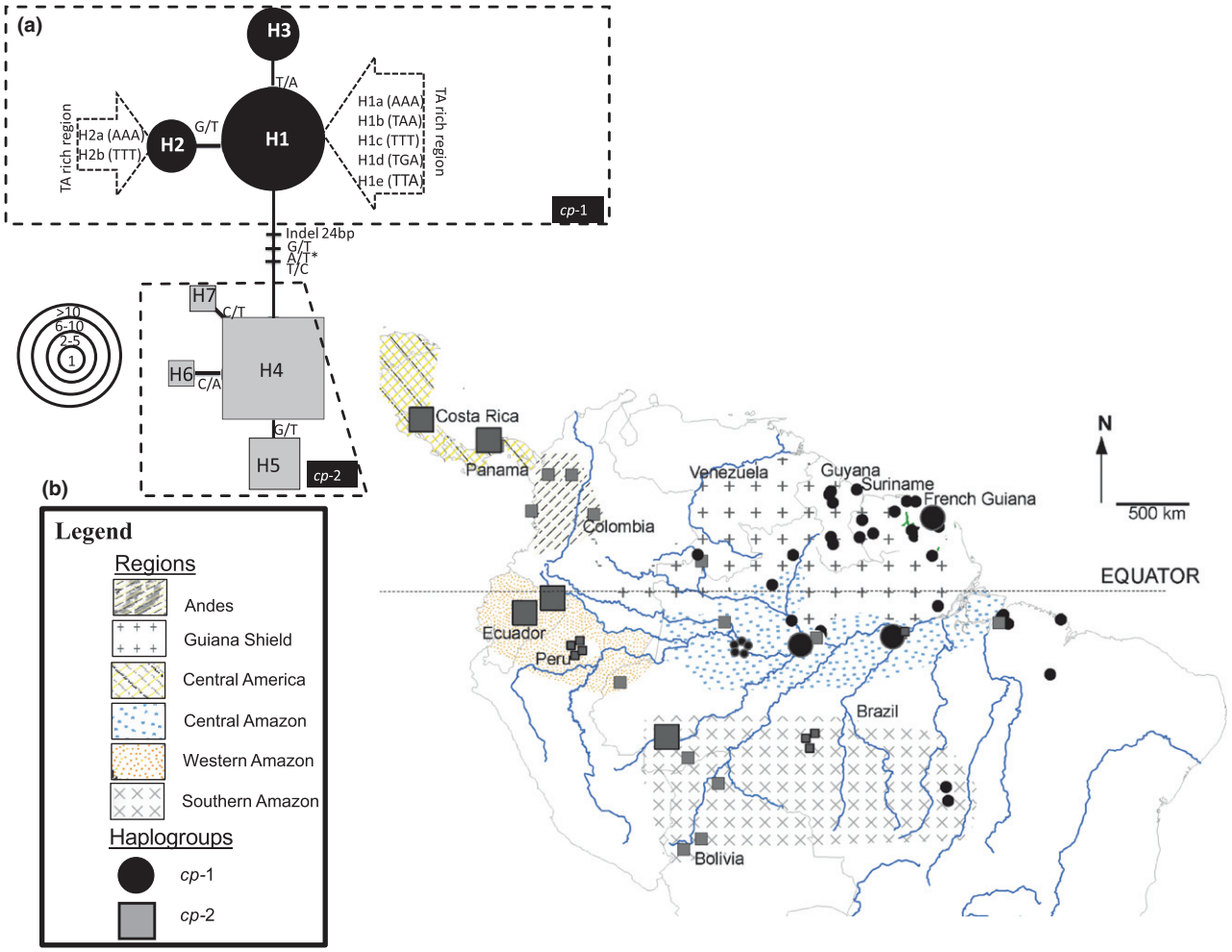

Figure 3 Chloroplast haplotypes of *Jacaranda copaia* in the Neotropics and geographical distribution of haplogroups in Central and South America. (a) Minimum spanning network computed from pairwise genetic distances (Arlequin) among 12 haplotypes observed in *J. copaia*. Dashes represent unobserved haplotypes, each one separated from its neighbours by one mutational step. When a mutation involves an insertion/deletion (indel), the size of the indel is given. Mutations within the T/A-rich region are indicated by a star. Symbol surface areas are proportional to haplotype frequency. (b) Geographical distribution of the two haplogroups (*cp*-1, *cp*-2) detected in 341 *J. copaia* samples across six regions of the Neotropics. Smaller symbols correspond to one sample; larger symbols correspond to more than five samples.

and C were associated with *cp*-1 and haplotypes D and E with *cp*-2: despite the higher evolutionary rate of cpSSRs, the phylogeographical pattern was maintained and haplotype E identified individuals from Central America within *cp*-2.

Population differentiation at cpDNA was significant based on both haplotype frequencies (F_{ST} = 0.84, P = 0.006) and distances between haplotypes (N_{ST} = 0.92, P = 0.006). N_{ST} was significantly higher than G_{ST} (0.71 vs. 0.58, P = 0.026) indicating the presence of phylogeographical structure.

Comparison of nuclear and chloroplast patterns

The chloroplast and nuclear data showed some degree of incongruence, consistent with incomplete lineage sorting or introgression, based on 140 individuals genotyped at both markers. The cp lineage *cp*-1 was associated with three nuclear groups *nu*-1, *nu*-2, *nu*-3 in Central Amazon and the cp lineage *cp*-2 with nuclear group *nu*-3 in Western Amazon and *nu*-2 in Southern Amazon (Boca do Acre) and Central America. The western population of the Central Amazon, Tefé, was genetically close to the populations of Western Amazon in Ecuador at nuclear markers, whereas its cp haplotypes belonged to *cp*-1 (present in Central Amazon); the population of Sarapiqui in Costa Rica and the population of Boca do Acre in Southern Amazon were genetically close to the populations of Central Amazon at nuclear markers, whereas their cp haplotypes belonged to *cp*-2 (present in Western Amazon).

Divergence dating

Figure 4 summarizes the Bayesian maximum clade credibility (MCC) trees for *trn*H–*psb*A. Diversification within *J. copaia* began in the late Pliocene with the split between *cp*-1 and *cp*-2, 2.61 Ma on average (95% CI = 0.83–4.77 Ma), while divergence within haplogroup *cp*-1 is estimated to have begun during the mid-Pleistocene (1.06 Ma, CI = 0.20–2.11 Ma). Divergence between the populations of Tapajos and Orellana was dated to the late Pleistocene for the model without migration, with the probability peak for t located between 357,309 and 435,982 years ago over four independent runs (Fig. 4) and the 95% credible interval from 114,732 to 1,458,738 years ago (see Appendix S3f). The model with migration converged poorly. The posterior distribution had a peak between 626,110 and 665,447 years ago (in the four independent runs) but had a flat right-end tail (see Appendix S3g), resulting in a very large, uninformative posterior credible interval.

DISCUSSION

Evidence for historical dispersal over or around the Andes

Both cpDNA and nuSSR cluster analyses showed a phylogeographical break in the Central Amazon region located approximately between Tefé (Rio Japura) and Manaus (Rio Negro). The cp haplotypes east of this region extended into the eastern Amazon and into the Guiana Shield (French Guiana), while the haplotype groups located to the west of this break extended from Central America through the Andean region and in Western and Southern Amazon as far south as Bolivia (Fig. 3). The location of the break is unusual among phylogeographical studies of Neotropical trees, which have either found major breaks separating *cis*- and *trans*-Andean clades (Dick *et al.*, 2003; Dick & Heuertz, 2008; Hardesty *et al.*, 2010; Lemes *et al.*, 2010) or weak to no phylogeographical structure (Dick *et al.*, 2007; Rymer *et al.*, 2013). The studies showing major divergence across the Andes are consistent with the expectations of the Andean vicariance hypothesis. Under Andean vicariance, lowland rain forest tree species were widespread in tropical South America prior to the uplift of the northern Andean cordilleras. The role of east, central and western cordilleras as biological barriers is not known with certainty (Antonelli *et al.*, 2009), and

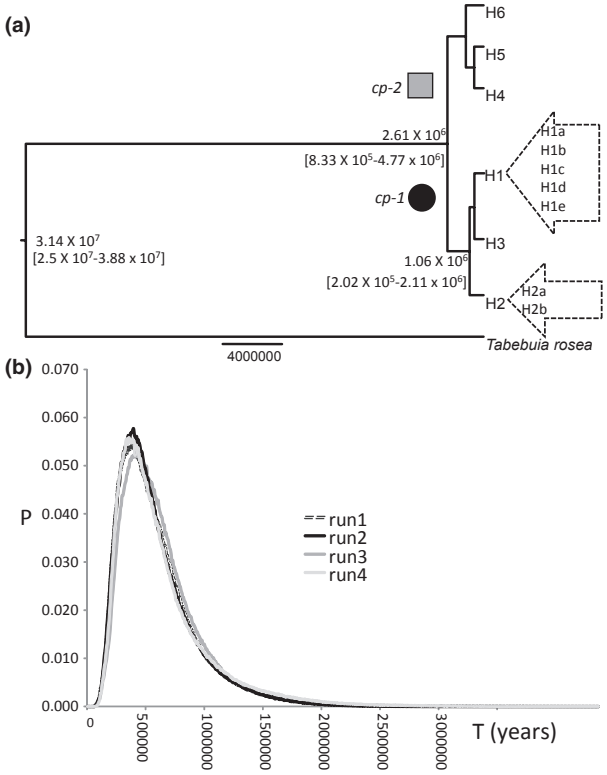

Figure 4 Lineage and population divergence dating in *Jacaranda copaia*. (a) Phylogenetic tree of *J. copaia* obtained from the cpDNA *trn*H–*psb*A sequence. The TA-rich region has been removed and the 24 bp insertion/deletion has been recoded as a single nucleotide polymorphism; as a consequence only six haplotypes out of 12 are displayed. Node labels display the mean and the lower and upper bounds of the 95% highest posterior density (HPD) interval of the node age in units of years. (b) Distribution of posterior estimates for 1000 values of divergence time between Tapajos and Orellana populations using a model without migration and four independent runs.

there may have been lowland passes available as dispersal corridors for lowland organisms even when average elevations exceeded the upper elevation limits of lowland forests. The most recent uplift events are associated with the eastern (Merida) cordilleras in Venezuela, which may have had modern elevations during the late Pliocene (*c.* 2.7 Ma).

The absence of nucleotide divergence between Western Amazon and Central American samples strongly suggest that the cross-Andean disjunction of *J. copaia* did not arise from vicariance. Actually, some nucleotide divergence is expected over a time-span of *c.* 3 Myr separation, as shown by Dick *et al.* (2007) for a species (*Ceiba pentandra*) with similar ecology, life cycle and population dynamics (suggesting similar effective population sizes). In *J. copaia*, the deepest phylogeographical divergence dates back to the late Pliocene, which suggests that the species became widespread after the uplift of the Northern Andes and other biogeographically important Neogene events (Pennington & Dick, 2010). The few widespread Neotropical tree species with weak phylogeographical breaks between Central America and the Amazon Basin share a number of ecological traits: (1) they are wind-dispersed species prone to long-distance seed dispersal, (2) they are long-lived pioneer species favoured by light gaps and disturbance yet persistent in mature forests, and (3) they tolerate drought and can therefore inhabit seasonally dry forests. The broad ecological tolerance of these species, and in particular their adaptations to xeric environments, could have favoured dispersal through a dry northern dispersal corridor along the Caribbean coast. Although this region may presently be too dry to permit contemporary migration of rain forest trees, it may have received more precipitation during warm periods of the Holocene or during previous interglacial periods.

As an alternative hypothesis to northern dispersal around the Andes, it is possible that widespread rain forest trees such as *J. copaia* penetrated the Andean barrier through lowland passes. Although lowland passes such as the portal to the Magdalene valley in Colombia (*c.* 2000 m a.s.l.) may be too high for contemporary *J. copaia* populations, the elevational range of *J. copaia* may have been higher at different times during the Pleistocene. As suggested by Dick *et al.* (2007), lowland tropical trees may have reproduced and dispersed at higher elevations during the mid-Holocene, when a warmer climate may have increased the elevation limit of lowland rain forest species (Bush *et al.*, 2004).

Divergence within the Amazon Basin

Divergence between *cp*-1 and *cp*-2 was estimated at 2.61 Ma (95% CI 0.83–4.77 Ma) and diversification within haplogroups was estimated at 1.06 Ma (95% CI 0.20–2.1 Ma). Divergence between populations of Central and Western Amazon was estimated by the model without migration at 0.36–0.44 Ma (95% CI 0.11–1.5 Ma), whereas the model with migration did not allow us to obtain a precise estimation of divergence time (this could be due to insufficient information in the data to correctly resolve migration rates; alternatively, the model may not account for gene flow from populations not included in the scenario, even though Bayesian clustering methods indicate that the populations used with IMa2 are genetically homogeneous; see Results). Diversification between and within haplotypes were deeper than population divergence time: the latter overlaps with many important Pleistocene climate changes. The diversification was larger in *cp*-1 (eight haplotypes versus four in *cp*-2) and the highest level of diversity was observed in Central Amazon, with private cpDNA haplotypes in Manaus and in Tapajos. This suggests that Central Amazon was a centre of diversification for the cp lineage *cp*-1, and thus would reinforce Gentry's (1982) hypothesis in which Manaus and Tapajos were putative refuge zones. In addition, the presence of few samples harbouring the cp lineage *cp*-2 along the Amazon River suggests that Central Amazon is a zone of secondary contact between the two cp lineages of *J. copaia*.

The distribution of *cp*-1 and *cp*-2 closely matches the geographical distribution of the two subspecies *J. copaia* subsp. *copaia* and *J. copaia* subsp. *spectabilis* (Gentry, 1992): *J. copaia* subsp. *spectabilis* is widespread in lowland moist and wet forest across Latin America, from Belize to Bolivia, up to 1200 m a.s.l.; *J. copaia* subsp. *copaia* ranges from the Guiana Shield to north-eastern Amazon and apparently does not exceed 600 m in elevation. No information is reported in the literature about ecological divergence between the two subspecies, but it is possible that they have undergone incipient speciation. This is supported by the partially overlapping geographical distributions (Gentry, 1992). However, the nuSSR clustering analysis suggests that there is genetic admixture of demes in support of the designation as subspecies rather than reproductively isolated species.

Gene flow between regions

Information provided by nuSSRs complemented the phylogeographical analysis run on cp markers by demonstrating the existence of three genetic clusters in *J. copaia*. Shallow genetic differentiation was detected between clusters, with the highest genetic differentiation ($F_{ST} = 0.10$) observed between the Western Amazon genetic group (*nu*-3) and the other two nuclear clusters (*nu*-1 or *nu*-2), in Central Amazon and Guiana Shield. Weak genetic differentiation between distant populations may be caused by long-distance gene dispersal or by relatively recent range expansions. The nuSSR groups showed a contrasting pattern of genetic differentiation at cp and nuclear markers, which suggests nuclear gene flow among distant populations characterized by divergent cp lineages: the Tefé population was genetically close to the populations of Western Amazon in Ecuador at nuclear markers, whereas its cp haplotypes belong to *cp*-1 (present in Central Amazon); the Sarapiqui population in Costa Rica and Boca do Acre in Southern Amazon were genetically close to the populations of Central Amazon and

Guiana Shield at nuclear markers, whereas their cp haplotypes belonged to *cp*-2 (present in Western Amazon). Lack of correspondence between genetic clusters obtained at cpDNA and nuSSRs suggests ongoing pollen-mediated gene flow rather than recent range expansion. SSR allele size homoplasy or incomplete lineage sorting may also explain SSR allele sharing between distant populations. However, we think that these scenarios are less likely: homoplasy should affect several loci simultaneously, which seems improbable; incomplete lineage sorting at nuSSRs seems unlikely to co-occur with simultaneous, complete cpDNA sorting, even though differences in effective population size between chloroplast and nuclear loci makes it possible that lineage sorting occurs more quickly for the former than for the latter. If strong gene flow between populations with divergent cp lineages occurs, it is correct to treat *J. copaia* subsp. *copaia* and *J. copaia* subsp. *spectabilis* as subspecies rather than as distinct species.

CONCLUSIONS

Our study suggests that patterns of genetic diversity within and among populations of Neotropical *J. copaia* were formed largely during the Pleistocene. A centre of diversity was observed in Central Amazon, which may be interpreted as the area of origin of *J. copaia* clades or as a zone of secondary contact between *J. copaia* lineages. The geographical distribution of the two main cp lineages overlaps with the ranges of two *J. copaia* subspecies. Nuclear variation among populations revealed gene flow between regional populations with divergent cp lineages, indicating that *J. copaia* subsp. *copaia* and *J. copaia* subsp. *spectabilis* belong to the same species under a biological species concept. These results highlight the importance of population-genetic investigations, based on both cp and nuclear genome-based markers, to retrace the past history of populations and clarify their taxonomic relationships. Finally, the absence of cross-Andean disjunction between populations of *J. copaia*, which was also observed in two other Neotropical tree species displaying similar ecological traits, suggests that species with pioneer traits may share similar patterns of intercontinental dispersal and weak phylogeographical structure.

ACKNOWLEDGEMENTS

This work was funded by the EU-funded INCO 'SEEDSOURCE' project, by the EU-funded PO-FEDER 'ENERGIRAVI' program and the Agence Nationale de la recherche (ANR) funded programme 'BRIDGE'. The authors thank Saint-Omer Cazal (UMR ECOFOG, INRA, Kourou, French-Guiana), Jean-Francois Molino (IRD, AMAP, Montpellier, France), Daniel Sabatier (IRD, AMAP, Montpellier, France), Heidy Villalobos (University of Costa Rica) and Carlos Navarro (CATIE, Costa Rica) for sampling; Sam Davies, Cayenne herbarium (IRD, French Guiana), Manaus herbarium (EMBRAPA and INPA, Brazil), Oxford herbarium (University Herbaria, UK) and Utrecht herbarium (National Herbarium, the Netherlands) for access to collections; and Eliane Louisanna (UMR ECOFOG, INRA, Kourou, French Guiana) for technical assistance with molecular biology.

REFERENCES

Aide, T.M. & Rivera, E. (1998) Geographic patterns of genetic diversity in *Poulsenia armata* (Moraceae): implications for the theory of Pleistocene refugia and the importance of riparian forest. *Journal of Biogeography*, **25**, 695–705.

Antonelli, A., Nylander, J.A.A., Persson, C. & Sanmartín, I. (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences USA*, **16**, 9749–9754.

Avise, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, MA.

Baraloto, C., Hardy, O.J., Paine, C.E.T., Dexter, K.G., Cruaud, C., Dunning, L.T., Gonzalez, M.-A., Molino, J.-F., Sabatier, D., Savolainen, V. & Chave, J. (2012) Using functional traits and phylogenetic trees to examine the assembly of tropical tree communities. *Journal of Ecology*, **100**, 690–701.

Barthe, S., Gugerli, F., Barkley, N.A., Maggia, L., Cardi, C. & Scotti, I. (2012) Always look on both sides: phylogenetic information conveyed by simple sequence repeat allele sequences. *PLoS ONE*, **7**, e40699.

Brokaw, N. (1985) Gap phase regeneration in a tropical forest. *Ecology*, **66**, 683–687.

Bush, M.B. (1994) Amazonian speciation: a necessarily complex model. *Journal of Biogeography*, **21**, 5–17.

Bush, M.B. & De Oliveira, P.E. (2006) The rise and fall of the refugial hypothesis of Amazonian speciation: a paleoecological perspective. *Biota Neotropica*, **6**, 1–17.

Bush, M.B., Silman, M.R. & Urrego, D.H. (2004) 48,000 years of climate and forest change in a biodiversity hot spot. *Science*, **303**, 827–829.

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.

Chapuis, M.-P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Colinvaux, P.A., De Oliveira, P.E. & Bush, M.B. (2000) Amazonian and Neotropical plant communities on glacial time-scale: the failure of the aridity and refuge hypothesis. *Quaternary Science Reviews*, **19**, 141–169.

Dalling, J.W., Muller-Landau, H.C., Wright, S.J. & Hubbell, S.P. (2002) Role of dispersal in the recruitment limitation of neotropical pioneer species. *Journal of Ecology*, **90**, 714–727.

Dick, C.W. & Heuertz, M. (2008) The complex biogeographic history of a widespread tropical tree species. *Evolution*, **62**, 2760–2774.

Dick, C.W., Abdul-Salim, K. & Bermingham, E. (2003) Molecular systematic analysis reveals cryptic Tertiary diversification of a widespread tropical rain forest tree. *The American Naturalist*, **162**, 691–703.

Dick, C.W., Bermingham, E., Lemes, M.R. & Gribel, R. (2007) Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* (Malvaceae) in Africa and the Neotropics. *Molecular Ecology*, **16**, 3039–3049.

Dick, C.W., Hardy, O.J., Jones, F.A. & Petit, R.J. (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology*, **1**, 20–33.

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

Dutech, C., Seiter, J., Petronelli, P., Joly, H.I. & Jarne, P. (2002) Evidence of low gene flow in a neotropical clustered tree species in two rainforest stands of French Guiana. *Molecular Ecology*, **11**, 725–738.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

François, O., Ancelet, S. & Guillot, G. (2006) Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics*, **174**, 805–816.

Gentry, A.H. (1982) Patterns of neotropical plant species diversity. *Evolution Biology*, **15**, 1–84.

Gentry, A.H. (1992) *Flora Neotropica. Monograph 25 (II) Bignoniaceae part II (tribe Tecomeaea)*. New York Botanical Garden, New York.

Gregory-Wodzicki, K.M. (2000) Uplift history of the Central and Northern Andes: a review. *Geological Society of America Bulletin*, **112**, 1091–1105.

Grundmann, H., Hori, S. & Tanner, G. (2001) Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. *Journal of Clinical Microbiology*, **39**, 4190–4192.

Guariguata, M.R., Rheingans, R. & Montagnini, F. (1995) Early woody invasion under tree plantations in Costa Rica: implications for forest restoration. *Restoration Ecology*, **3**, 252–260.

Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131–137.

van der Hammen, T. & Hooghiemstra, H. (2000) Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Reviews*, **19**, 725–742.

Hardesty, B.D., Dick, C.W., Hamrick, J.L., Degen, B., Hubbell, S.P. & Bermingham, E. (2010) Geographic influence on genetic structure in the widespread Neotropical tree *Simarouba amara* (Simaroubaceae). *Tropical Plant Biology*, **3**, 28–39.

Hardy, O.J. & Vekemans, X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hardy, O.J., Charbonnel, N., Freville, H. & Heuertz, M. (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Hardy, O.J., Maggia, L., Bandou, E., Breyne, P., Caron, H., Chevallier, M.-H., Doligez, A., Dutech, C., Kremer, A., Latouche-Hallé, C., Troispoux, V., Veron, V. & Degen, B. (2006) Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Molecular Ecology*, **15**, 559–571.

Hey, J. (2010) Isolation with migration models for more than two populations. *Molecular Biology Evolution*, **27**, 905–920.

Hey, J. & Nielsen, R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hubbell, S.P., He, F., Condit, R., Borda-de-Água, L., Kellner, J. & ter Steege, H. (2008) How many tree species are there in the Amazon and how many of them will go extinct? *Proceedings of the National Academy of Sciences USA*, **105**, 11498–11504.

James, T., Vege, S., Aldrich, P. & Hamrick, J.L. (1998) Mating systems of three tropical dry forest tree species. *Biotropica*, **30**, 587–594.

Jones, F.A. & Hubbell, S.P. (2003) Isolation and characterisation of microsatellites in the tropical tree *Jacaranda copaia* (Bignoniaceae). *Molecular Ecology Notes*, **3**, 403–405.

Jones, F.A., Chen, J., Weng, G.-J. & Hubbell, S.P. (2005) A genetic evaluation of seed dispersal in the Neotropical tree *Jacaranda copaia* (Bignoniaceae). *The American Naturalist*, **166**, 543–555.

Junior, S.B. & Yared, J.A.G. (1991) Agroforestry systems as an ecological approach in the Brazilian Amazon development. *Forest Ecology and Management*, **45**, 319–323.

Keigwin, L. (1982) Isotopic paleoceanography of the Caribbean and East Pacific: role of Panama uplift in late Neogene time. *Science*, **217**, 350–353.

Knapp, S. & Mallet, J. (2003) Refuting refugia? *Science*, **300**, 71–72.

Lemes, M.R., Dick, C.W., Navarro, C., Lowe, A.J., Cavers, S. & Gribel, R. (2010) Chloroplast DNA microsatellites reveal contrasting phylogeographic structure in mahogany (*Swietenia macrophylla* King, Meliaceae) from Amazonia and Central America. *Tropical Plant Biology*, **3**, 40–49.

Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Maués, M.M., De Oliveira, P.E. & Kanashiro, M. (2008) Pollination biology in *Jacaranda copaia* (Aubl.) D. Don. (Big-

noniaceae) at the 'Floresta nacional do Tapajós', Central Amazon. *Brazil. Revista Brazil Botanica*, **31**, 517–527.

Mayr, E. (1954) Change of genetic environment and evolution. *Evolution as a process* (ed. by J. Huxley, A.C. Hardy and E.B. Ford), pp. 157–180. Allen & Unwin, London.

Mittelbach, G.G., Schemske, D.W., Cornell, H.V. *et al.* (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, **10**, 315–331.

Moritz, C. (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Moritz, C. & Faith, D.P. (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular Ecology*, **7**, 419–429.

Nielsen, R. (1997) A maximum likelihood approach to population samples of microsatellite alleles. *Genetics*, **146**, 711–716.

Novick, R.R., Dick, C.W., Lemes, M.R., Navarro, C., Caccone, A. & Bermingham, E. (2003) Genetic structure of Mesoamerican populations of big-leaf mahogany (*Swietenia macrophylla*) inferred from microsatellite analysis. *Molecular Ecology*, **12**, 2885–2893.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Pennington, R.T. & Dick, C.W. (2010) Diversification of the Amazonian flora and its relation to key geological and environmental events: a molecular perspective. *Amazonia: landscape and species evolution* (ed. by C. Hoorn, H. Vonhof and F. Wesselingh), pp. 373–385. Blackwell, Oxford.

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 B: Biological Sciences*, **359**, 515–538.

Petit, R.J., El Mousadik, A. & Pons, O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Poelchau, M.L. & Hamrick, J.L. (2013) Comparative phylogeography of three common Neotropical tree species. *Journal of Biogeography*, **40**, 618–631.

Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Provan, J., Soranzo, N., Wilson, N.J., Goldstein, D.B. & Powell, W. (1999) A low mutation rate for chloroplast microsatellites. *Genetics*, **153**, 943–947.

Richardson, J.E., Pennington, R.T., Pennington, T.D. & Hollingsworth, P.M. (2001) Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science*, **293**, 2242–2245.

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

Rymer, P.D., Dick, C.W., Vendramin, G.G., Buonamici, A. & Boshier, D. (2013) Recent phylogeographic structure of a 'weedy' widespread Neotropical tree species, *Cordia alliodora* (Boraginaceae). *Journal of Biogeography*, **40**, 693–706.

Scotti-Saintagne, C., Dick, C.W., Caron, H., Vendramin, G.G., Guichoux, E., Buonamici, A., Duret, C., Sire, P., Valencia, R., Lemes, M.R., Gribel, R. & Scotti, I. (2013) Phylogeography of a species complex of lowland Neotropical rainforest trees (Carapa, Meliaceae). *Journal of Biogeography*, **40**, 676–692.

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

Slatkin, M. (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Weising, K. & Gardner, R. (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*, **42**, 9–19.

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology and Systematics*, **34**, 273–309.

SUPPORTING INFORMATION

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

Appendix S1 List of sampled sites for fresh and herbarium material of *Jacaranda copaia* and details of the genetic markers analysed.

Appendix S2 Description of the methods used to analyse population structure and to estimate divergence time.

Appendix S3 Genetic diversity and genetic differentiation *in Jacaranda copaia* in the Neotropics.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Caroline Scotti-Saintagne is a population geneticist interested in the demographic and environmental processes involved in the genetic divergence between closely related species of forest trees.

Author contributions: C.S.-S., I.S., C.D., H.C., G.G.V. designed the experiments; C.S.-S., H.C., I.S., R.V., M.L., R.G. collected the samples; C.S.-S., H.C., M.C., G.G.V., C.D., V.T., P.S., A.B. contributed to molecular analyses; C.S.-S., I.S. performed statistical analyses; C.S.-S., C.D., I.S. wrote the paper; all authors read and approved the manuscript.

Editor: Malte Ebach

This Special Issue was motivated by research undertaken by partners in the SEEDSOURCE consortium (www.seed-source.net), a collaborative project funded by the European Commission under FP6. The project aimed to establish the evolutionary basis and geographical distribution of genetic variation in a range of important Neotropical tree species, for better sourcing and control of germplasm and forest conservation.