HAL
open science

Tropical rainforests that persisted: inferences from the Quaternary demographic history of eight tree species in the Guiana shield

Stéphanie Barthe, Giorgio Binelli, Bruno Hérault, Caroline Scotti-Saintagne, Daniel Sabatier, Ivan Scotti

To cite this version:

Stéphanie Barthe, Giorgio Binelli, Bruno Hérault, Caroline Scotti-Saintagne, Daniel Sabatier, et al.. Tropical rainforests that persisted: inferences from the Quaternary demographic history of eight tree species in the Guiana shield. Molecular Ecology, 2016, 26 (4), 10.1111/mec.13949 . hal-01512150

HAL Id: hal-01512150

https://hal.science/hal-01512150v1

Submitted on 21 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MOLECULAR ECOLOGY

Molecular Ecology (2017) **26**, 1161–1174 doi: 10.1111/mec.13949

Tropical rainforests that persisted: inferences from the Quaternary demographic history of eight tree species in the Guiana shield

STÉPHANIE BARTHE,*† GIORGIO BINELLI,‡ BRUNO HÉRAULT,§ CAROLINE SCOTTI-SAINTAGNE,¶ DANIEL SABATIER** and IVAN SCOTTI¶

**UMR EcoFoG ("Ecologie des forêts de Guyane"), UAG, Campus Agronomique, BP 709, 97387 Kourou, French Guiana, †UMR EcoFoG ("Ecologie des forêts de Guyane"), INRA, Campus Agronomique, BP 709, 97387 Kourou, French Guiana, ‡DBSV ("Dipartimento di Biotecnologie e Scienze della Vita"), Università degli Studi dell'Insubria, via J.H. Dunant 3, 21100 Varese, Italy, §UMR EcoFoG ("Ecologie des forêts de Guyane"), CIRAD, Campus Agronomique, BP 709, 97387 Kourou, French Guiana, ¶URFM ("Écologie des Forêts Méditerranéennes"), INRA, Domaine Saint_Paul, Site Agroparc, 84914 Avignon, France, **UMR AMAP ("Botanique et modélisation de l'architecture des plantes et des végétations"), IRD, Boulevard de la Lironde, Parc Scientifique 2, 34398 Montpellier, France*

Abstract

How Quaternary climatic and geological disturbances influenced the composition of Neotropical forests is hotly debated. Rainfall and temperature changes during and/or immediately after the last glacial maximum (LGM) are thought to have strongly affected the geographical distribution and local abundance of tree species. The paucity of the fossil records in Neotropical forests prevents a direct reconstruction of such processes. To describe community-level historical trends in forest composition, we turned therefore to inferential methods based on the reconstruction of past demographic changes. In particular, we modelled the history of rainforests in the eastern Guiana Shield over a timescale of several thousand generations, through the application of approximate Bayesian computation and maximum-likelihood methods to diversity data at nuclear and chloroplast loci in eight species or subspecies of rainforest trees. Depending on the species and on the method applied, we detected population contraction, expansion or stability, with a general trend in favour of stability or expansion, with changes presumably having occurred during or after the LGM. These findings suggest that Guiana Shield rainforests have globally persisted, while expanding, through the Quaternary, but that different species have experienced different demographic events, with a trend towards the increase in frequency of light-demanding, disturbance-associated species.

Keywords: biogeography, forest ecology, historical demography, neotropics

Received 8 June 2016; revision received 14 November 2016; accepted 15 November 2016

Introduction

The long lifespan of forest trees often makes us forget that most terrestrial forests have also a long and eventful history. In temperate Europe and North America, sudden range expansions from refugia soon after the last glacial maximum (LGM, 18 000–20 000 years BP) (Taberlet *et al.* 1998; Hewitt 2000, 2004; Petit *et al.* 2002; Hu *et al.* 2009), bottlenecks (Heuertz *et al.* 2006) and founder events (Pyhäjärvi *et al.* 2007) are now well recognized. For Mesoamerica, a recent study (Jones *et al.* 2013) has also shown recent bottlenecks in four rainforest species, whose dates are compatible with the end of the LGM, whereas in Equatorial Africa tree populations appear to have gently contracted/expanded without dramatic range changes (Daïnou *et al.* 2010; Budde *et al.* 2013). In South America, several climatic changes during the Tertiary and Quaternary have likely shaped the current composition of Amazonian rainforests

Correspondence: Ivan Scotti, Fax: +33 (0)4 32 72 29 02; E-mail: ivan.scotti@inra.fr

© 2016 John Wiley & Sons Ltd

(Haberle 1997; Antonelli *et al.* 2009; Hoorn *et al.* 2010). However, the past distribution and the demographical history of South American rainforest communities is still relatively unknown (Hewitt 2004), although recent phylogeographic work (Motamayor *et al.* 2008; Hardesty *et al.* 2010; Scotti-Saintagne *et al.* 2013a, b) has revealed a complex history from genetic diversity patterns.

Fossil data are rare in the Amazon, due to rapid decay rates of organic matter in poor sedimentary soils, making it difficult to infer forest history. The oldest available biogeographical, geomorphological and palaeological data suggest that, at the LGM, Amazonian rainforests were fragmented into several isolated geographical areas (Haffer 1969; Prance 1982), probably due to a drier climate (Ledru *et al.* 1998). This is supported by several charcoal analyses suggesting Holocene drought-related palaeo-fires (Charles-Dominique *et al.* 1998; Tardy 1998; Mayle *et al.* 2000; Bush *et al.* 2008; Mayle & Power 2008; Power *et al.* 2010). Under the 'fragmentation' hypothesis, climate changes during Pleistocene glacials may have imposed demographic transitions on plant and animal populations and provoked large-scale fragmentation/recolonization processes (Aide & Rivera 1998; Caron *et al.* 2000; Dutech *et al.* 2004; Flanagan *et al.* 2004; Quijada-Mascarenas *et al.* 2007; Motamayor *et al.* 2008; Solomon *et al.* 2008; de Thoisy *et al.* 2010; Scotti-Saintagne *et al.* 2013a). The dry-climate-with-refugia hypothesis has nevertheless been challenged (Bush & De Oliveira 2006). One alternative hypothesis, suggested by Amazonian pollen records (Colinvaux *et al.* 1996), is that a 6 °C cooling during the LGM may have caused a floristic transition, rather than rainforest fragmentation and replacement of rainforest by open, drier tropical forests (Colinvaux *et al.* 2000; Bush *et al.* 2004; Freycon *et al.* 2010).

Past changes in effective population size can be detected using statistical approaches that rely on the comparison of observed patterns of genetic diversity at neutral molecular loci with the expectations provided by the neutral theory of molecular evolution (Kimura 1985). Highly polymorphic genetic markers such as microsatellites (or simple sequence repeats, SSRs), with a mutation rate of about 10^{-2} to 10^{-6} per site per generation (Thuillet *et al.* 2002), offer the opportunity to analyse demographic events that occurred in the recent evolutionary past. On the other hand, chloroplast sequence markers, with a lower mutation rate between 10^{-9} and 10^{-10} per site per generation (Wolfe *et al.* 1987; Richardson *et al.* 2001) and no or rare recombination events (Birky 1995), return more slowly to the equilibrium state after a demographic change. Genetic information carried by both diploid nuclear and haploid chloroplast markers allows describing and studying a wide range of past demographical events that may have affected current genetic diversity. Hypothesis testing, parameter estimation and selection of most likely demographic model can be performed by combining coalescent modelling (Kingman 1982), a powerful tool to model neutral molecular polymorphism (Nordborg 2004), with Bayesian or approximate Bayesian computation (ABC; Beaumont *et al.* 2002; Beaumont 2010; Budde *et al.* 2013; Jones *et al.* 2013).

The goal of the present study was to test whether past demographic transitions can be detected for eight tree species belonging to present-day forests of the eastern Guiana Shield. Focal tree species cover a gradient of ecological strategies from light-demanding to shade-tolerant. Common trends will be sought as evidence of Amazonian rainforest fragmentation and/or past reduction in global forest cover caused by dryer and/or cooler climate, while interspecific differences in demographic trends, when shown to correlate with ecological strategies, will be considered as evidence of changes in forest structure.

Material and methods

Choice of species

Eight tropical tree species, currently common in the eastern Guiana Shield forests, were studied: *Carapa surinamensis* Miq. (Meliaceae), *Dicorynia guianensis* Amshoff (Caesalpiniaceae), *Jacaranda copaia* (Aubl.) D. Don (Bignoniaceae), *Simarouba amara* Aubl. (Simaroubaceae), *Symphonia globulifera* L. f. (Clusiaceae) and the related subspecies *Symphonia* sp1 (Sabatier *et al.* 1997), *Virola michelii* Heckel (Myristicaceae) and *Vouacapoua americana* Aubl. (Caesalpiniaceae). These species show relatively ample variation in ecological strategies (Table S1, Supporting information), although they all are rainforest-dependent species. *D. guianensis*, *V. americana* and *S. sp1* are endemic to the Guiana Shield rainforests, while all others are widespread in the Neotropics, and *S. globulifera* has a bicontinental distribution (Africa, America).

Data sets

The data analysed here were drawn both from published and original data sets (Table S2, Supporting information). We report below the methods used for the production of original data. Whenever the new data were used to complete existing data sets, the methods used and reported here are the same as for previous studies.

Sampling and DNA extraction

Samples of leaf or cambium were collected from 37 to 356 individuals of each species during surveys

© 2016 John Wiley & Sons Ltd

conducted between 2000 and 2012 across French Guiana (Table S2 and Figure S1, Supporting information). Extractions of total genomic DNA were performed following Colpaert *et al.* (2005) or according to the Qiagen DNeasy® 2006 protocol with a DNeasy 96 plant kit (Qiagen, Hilden, Germany). DNA quality was checked by spectrophotometry and by agarose gel electrophoresis.

DNA amplification, genotyping and sequencing

Nuclear microsatellites. Nuclear genetic diversity was analysed at four to nine microsatellite (nSSR) markers depending on the species. Table S2 (Supporting information) provides the reference for each marker. Polymerase chain reactions (PCRs) for *S. amara* markers were performed in 15 μL volume containing 6 μL of 25 ng/μL DNA, 1× *Taq* buffer, 0.2 mM of dNTP, 0.04 U/μL *Taq* DNA polymerase (all products from Invitrogen, Carlsbad, CA, USA) and 0.75 μM of each primer (MWG Biotech, Ebersberg, Germany). An initial denaturation at 94 °C for 5 min was followed by 30 cycles of 94 °C for 30 s, annealing temperature of 53 °C for 30 s and 72 °C for 30 s, and a final extension at 72 °C for 5 min.

For *V. michelii*, PCRs were performed in a 15 μL volume containing 3 μL of 20 ng/μL DNA, 1× *Taq* buffer, 0.7 mM of $MgCl_2$, 0.1 mM of dNTP, 0.04 U/μL *Taq* DNA polymerase (all products from Invitrogen) and 0.2 μM of each primer (MWG Biotech, Ebersberg, Germany). PCR program was as follows: initial denaturation at 94 °C for 5 min, 35 cycles of 30 s at 94 °C, 40 s at 52 °C and 1 min at 72 °C, and a final extension at 72 °C for 10 min.

For the four *Symphonia globulifera* and *S. sp1* microsatellite markers, PCRs followed the M13-tailed primer PCR method (Schuelke 2000). Amplifications were carried out in a total volume of 15 μL containing 8.2 μL of 25 ng/μL diluted DNA, 1× *Taq* buffer, 0.125 mM of dNTP, 0.05 U/μL *Taq* DNA polymerase, 0.2 μM of forward M13-tailed and reverse primers, 0.05 μM of forward primer (MWG Biotech, Ebersberg, Germany). PCR conditions included a denaturation step at 94 °C for 4 min, followed by 30 cycles of 45 s at 94 °C, 45 s at annealing temperature (depending on the primer pair, see Table S2, Supporting information), and 45 s at 72 °C, followed by eight cycles of 45 s at 94 °C, 45 s at 53 °C, and 45 s at 72 °C, and a final extension step at 72 °C for 10 min.

The four *C. surinamensis* markers, *Cg*1L, *Cg*5, *Cg*7 and *Cg*17, were amplified according the protocol of the Type-it® multiplex-PCR kit (Qiagen, Hilden, Germany). Amplification conditions for *C. surinamensis* marker *Cg*01 are as reported in Vinson *et al.* (2005a), and nSSRs of *D. guianensis* were amplified as described by Latouche-Halle *et al.* (2002).

In all the protocols above, forward primers were fluorescently labelled (with PET, 6-FAM, NED or VIC) and fragments were separated on an ABI 3130XL capillary sequencer (Applied Biosystems, Foster City, CA, USA) using ABI POP4 and Applied Biosystems LIZ-500 as internal standard, following the manufacturer's instructions. Genotypes were scored using the GENEMAPPER 4.0 software package (Applied Biosystems).

Chloroplast sequences. Two universal chloroplast intergenic space regions, *trnH-psbA* and *trnC-ycf6*, were amplified, as described in Shaw *et al.* (2005), for all eight studied species. DNA sequencing was performed with the ABI BigDye Terminator® version 3.1 cycle sequencing kit (Applied Biosystems) according to the manufacturer's protocol. PCR products were purified with ExoSAP-IT (USB, Cleveland, OH, USA) and separated on an ABI 3130XL capillary sequencer (Applied Biosystems). DNA sequences were aligned with CODONCODE ALIGNER 1.6.3 (CodonCode Corporation, Dedham, MA, USA) and edited with BIOEDIT 7.0.9.0 (Hall 1999).

Data analyses

Microsatellite data analysis. To control for possible genetic substructuring within each species, we used Bayesian clustering analyses as implemented in the program STRUCTURE 2.3.1 (Pritchard *et al.* 2000) to probabilistically assign individuals to *K* clusters based on nuclear microsatellite genotypes. Parameter settings and choice of the most likely number of clusters, based on ΔK developed by Evanno *et al.* (2005), were as described in Duminil *et al.* (2006) and Scotti-Saintagne *et al.* (2013a). Individuals with membership coefficients >0.80 were assigned to the most likely cluster. Because the number of loci varied across species, and because the number of clusters that the algorithm detects may be sensitive to the amount of genetic data, we tested the sensitivity of the identification of the best *K* to the number of markers by randomly sampling (with replacement) a decreasing number of loci from the largest data set (*Jacaranda copaia*) and by rerunning STRUCTURE. The results of the analyses with a number of loci varying between nine (all loci) and three were compared.

Levels of polymorphism (mean number of alleles for all loci, A, and Nei's genetic diversity, H_e) were computed using the program ARLEQUIN 3.5.1.2 (Excoffier & Lischer 2010).

Chloroplast DNA data analysis. The best-fit nucleotide substitution model was inferred using the Akaike information criterion (AIC) implemented in the program

© 2016 John Wiley & Sons Ltd

MODELTEST 3.8 (Posada & Crandall 1998; Posada & Buckley 2004), starting from the neighbour-joining tree of chloroplast haplotypes obtained using the algorithm in PAUP 4.10b (Swofford 1991). The Hasegawa–Kishino–Yano model (HKY; Hasegawa *et al.* 1985) described eleven of twelve polymorphic chloroplast data sets and was therefore chosen for the coalescent modelling of both *trnH-psbA* and *trnC-ycf6* markers for all species.

Coalescent modelling (approximate Bayesian computation). Demographic history was modelled through coalescent-based approximate Bayesian computation (ABC; for reviews see Bertorelle *et al.* 2010; Csilléry *et al.* 2010), as implemented in the DIYABC version 2.1.0 algorithm (Cornuet *et al.* 2008). First, we modelled a 'free variation' in effective population size: both present and past population sizes were allowed to vary freely; this way, depending on the draws in each prior, all possible events (contraction, stability, expansion) could occur. This scenario was used for parameter estimation. To confirm the trends observed in the 'free variation' approach, we tested against each other the following scenarios under a standard 'scenario choice' approach: (1) a single event of population contraction and (2) a single event of population expansion. This allowed us to further check that a scenario with population size change was supported (when the two scenarios had close probabilities, we inferred that none was supported and concluded in favour of population stability). For comparison with the 'free variation' approach, parameter estimation was carried out under the most probable between (1) and (2) and used to check the results obtained with the first approach. Demographic parameters were drawn with the following distributions: log-normal (mean = 300; SD = 250) for demographic event times of all species and log-normal (mean = 1000; SD = 1000) for current or ancestral effective population sizes. We simulated nSSRs using a generalized stepwise-mutation model (GSM; Zhivotovsky *et al.* 1997; Estoup *et al.* 2002) with two parameters: the mean mutation rate (μ_{SSR}) and the mean of the geometric distribution of the probability of being a one-step mutation (P) were drawn, respectively, from Uniform [10^{-5}; 10^{-3}] and Uniform [0.1; 0.3] hyperprior distributions. Uniform prior distributions were preferred over a log-uniform for nSSR mutation rates because previous studies (Thuillet *et al.* 2002; Vigouroux *et al.* 2002; Ellegren 2004) showed values close or around the upper values chosen for this study. We therefore decided to sample more densely larger rather than smaller mutation rate values. Each locus was characterized by individual μ_i and P_i values, respectively, drawn from Gamma (mean = mean μ_{SSR}; shape = 2) and Gamma (mean = P; shape = 2) prior distributions. The two chloroplast sequence loci were assumed to follow the HKY evolutionary model in all species, as determined previously, with a fraction of constant sites fixed to 0% and the shape parameter of the Gamma distribution of mutations among sites equal to 2. The mean hyperprior mutation rate μ_{cpSeq} (i.e. the mean of the actual priors) was drawn in a Uniform [10^{-8}; 10^{-7}], and mutation rate of each locus was chosen from a Gamma prior (mean = μ_{cpSeq}; shape = 2). Coalescent modelling was applied to nuclear and chloroplast data, analysed together using DIYABC version 2.1.0 (Cornuet *et al.* 2008), on the ancestral genetic clusters identified by STRUCTURE. Because we held samples from a large number of sampling sites over an area of approximately 80 000 km^2, we could not consider our species-level samples as belonging to a single truly panmictic deme, even within clusters identified by STRUCTURE (see Results). We therefore subsampled each species' data set according to the 'scattered sampling' strategy (i.e. one or very few samples per deme, over the largest possible number of demes, here taken as corresponding to sampling sites). This sampling strategy permits to model a range-level data set as an effectively panmictic population and to correctly test demographic hypotheses (Stadler *et al.* 2009 and citations therein). Moreover, coalescent inferences are robust to sampling size (Nordborg 2004), suggesting that the results would be unaffected by subsampling; we nevertheless tested robustness empirically. Within each STRUCTURE cluster, a subset of individuals was sampled uniformly with respect to their geographical distribution. Thirty individuals were drawn from the complete data set for *C. surinamensis*, *J. copaia*, *S. amara*, *V. michelii*; fifteen individuals were drawn for *D. guianensis*, *S. globulifera*, *Symphonia* sp1 and for two ancestral clusters (see Results) for *V. americana*. A second draw (with replacement) was performed to evaluate the robustness of our analysis, except for *V. michelii* and one cluster of *V. americana* ('West') for which we could not sample different individuals. The summary statistics computed for nSSR loci were as follows: mean number of alleles, gene diversity (Nei 1987) and allele size variance; the latter was not estimated for *V. americana*, for which allele identity, but not allele size, was available (Dutech *et al.* 2004). For chloroplast DNA sequence loci, we chose number of distinct haplotypes, Tajima's D (Tajima 1989b) and mean of numbers of the rarest nucleotide at segregating sites. These summary statistics are influenced by demographic changes (Tajima 1989a; Reich *et al.* 1999; Cornuet *et al.* 2008). We simulated 1 000 000 data sets for each scenario for both the first draw and the second draw of empirical samples. In the 'scenario choice' approach, we estimated posterior probabilities P using a logistic regression method (Ghirotto *et al.* 2010) on 1–5% simulations

© 2016 John Wiley & Sons Ltd

closest to observed statistics. We computed type I (α) and II (β) error rates based on 1000 pseudo-observed data sets per scenario, selected within the 1% accepted set described above (Cornuet *et al.* 2008) using R (R Core Team, 2013) scripts. Posterior parameter distributions were estimated under the best scenario using a linear regression on the 1% closest simulations and applying a *logit* transformation to parameter values (Beaumont *et al.* 2002; Cornuet *et al.* 2008). Posterior distributions were obtained for three composite parameters: population diversity parameters $\theta_0 = N_0\mu_0$ (present), $\theta_1 = N_1\mu_1$ (past) and ratio $r_0 = \theta_0/\theta_1$, plus all single parameters. We focused interspecific comparisons on r_0 (the ratio of present-to-past θ's) because it correctly represents ratios of effective population sizes. Indeed, if we assume constant mutation rates over time ($\mu_0 = \mu_1 = \mu$), we obtain:

$$r_0 = \theta_0/\theta_1 = N_0\mu_0/N_1\mu_1 = N_0\mu/N_1\mu = N_1/N_1,$$

the ratio of present-to-past effective population sizes. This is the most interesting parameter of our analyses, because we focus on population size changes and because it can be compared across species.

To evaluate the reliability of chosen scenarios in the 'scenario choice' approach, we built an ad hoc 'confidence index' for each selected scenario, taking into account their posterior probability and associated errors, as follows. We first transformed each of the parameters $(1 - p)$, α and β by $X/(1 - X)$; the transformed values, which we call p', α' and β', vary between 0 and $+\infty$ and have the desirable property of increasing with increasing error rates and with decreasing probabilities, so that larger values indicate poorer scenarios. We next defined a three-dimensional space, with axes $W_p p'$, $W_\alpha \alpha'$ and $W_\beta \beta'$ obtained by multiplying the above variables by (arbitrary) weights (here, $W_p = 1$, $W_\alpha = 0.5$ and $W_\beta = 2$). Therefore, each scenario s can be assigned a point in this space, defined by coordinates $W_p p'_s$, $W_\alpha \alpha'_s$ and $W_\beta \beta'_s$. The arbitrary weights were chosen so that accepting a scenario when false is given a larger penalty than rejecting a scenario when true ($W_\beta \gg W_\alpha$). We then computed the Euclidean distance of a scenario s from the origin of the axes as: $d = \sqrt{[(W_p p'_s)^2 + (W_\alpha \alpha'_s)^2 + (W_\beta \beta'_s)^2]}$; such distance is larger for poorer scenarios. We finally defined the 'Probability-Alpha-Beta-based Confidence index' (PABC) as PABC = $(1/d)/[1 + (1/d)]$, which has the properties of increasing with increasing probability and decreasing error rates and being constrained between 0 and 1. Values of PABC close to 1 indicate reliable scenarios, while values close to 0 suggest that the chosen 'best' scenario cannot be really trusted. As an abstract example, a scenario with a high posterior probability ($P = 0.9$) and low type I and II errors ($\alpha = 0.1$ and $\beta = 0.1$) would have PABC = 0.80, while PABC = 0 for the worst scenario ($P = 0$, $\alpha = 1$ and $\beta = 1$). As a real-case example, from the results of Lander *et al.* (2011) in European beech, one of the few publications that reported type I and II errors for an ABC analysis, PABC = 0.61 was found. To represent the response surface of the confidence index, we studied variation of the PABC values around the weights (W) used. The sensitivity of PABC to weight choice was analysed by studying the relationship between PABC and each weight (Methods S1, Supporting information).

Coalescent modelling (maximum likelihood). At nSSRs, demographic history was also explored using the maximum-likelihood (ML) approach implemented in the MIGRAINE version 0.4 software (Leblois *et al.* 2014), applying the same 'free variation' scenario as in the ABC and a generalized stepwise-mutation model (GSM). Estimates and confidence intervals were computed for five parameters: scaled current population size $2\theta_0 = 2N_0\mu$, scaled time (in generations) of the demographic change $T = T/2N_0$, ancestral scaled population size $2\theta_1 = 2N_1\mu$, ratio of population sizes $r_0 = 2\theta_0/2\theta_1$ (see above for the interpretation of r_0), mutation parameter pGSM. Confidence intervals (CI) of parameter estimates were obtained from the two-dimensional profile likelihood ratio. Parameter values for the computation of likelihoods were drawn as follows (boundaries for the first iteration in parentheses): pGSM, uniform distribution (0.1–0.9); $2\theta_0$ and $2\theta_1$, log-uniform distribution (0.001–100); T, log-uniform distribution (0.01–10). The size of parameter points was 200, the number of genealogies explored by the IS algorithm was 1000 for the first ten iterations and 2000 from the eleventh iteration on (20 iterations in total). Maximum-likelihood and confidence intervals were obtained using a kriging scheme with a grid step of 40 in the last iteration (16 in the others). The ML approach could not be applied to the *V. americana* data sets because the original data, compiled from Dutech *et al.* (2004), did not report allele sizes.

Ecological characteristics. We tested correlations between values of θ ratios (r_0) and (i) functional traits describing the life history, leaf economic and stem economic strategy (Baraloto *et al.* 2010) as well as (ii) demographic parameters (turnover rates, growth potential and plasticity) derived from Hérault *et al.* (2011) and Aubry-Kientz *et al.* (2013).

Results

Genetic diversity and population structure at nSSRs

The average number of alleles per locus ranged from 8.20 in *Carapa surinamensis* to 25.29 in *V. michelii*;

© 2016 John Wiley & Sons Ltd

diversity (H_e) ranged from 0.32 in *C. surinamensis* to 0.94 in *Symphonia* sp1 (Table S2, Supporting information). Population structure was analysed to identify genetically homogeneous groups complying with the coalescent assumption of panmictic Wright–Fisher populations. *S. globulifera* and *S. sp1*, *C. surinamensis*, *Dicorynia guianensis*, *S. amara* and *V. michelii* showed weak ΔK values ($3 \leq \Delta K \leq 82$), and very few individuals assigned to a cluster with a membership coefficient greater than 0.80. We therefore considered these data sets as single clusters. A strong structure of two clusters was detected for *V. americana* ($\Delta K = 843.8$) and *Jacaranda copaia* ($\Delta K = 402.7$). The two *V. americana* clusters Q1 and Q2 were composed of 87 and 97 individuals, respectively, with an east–west distribution. We decided therefore to analyse separately the two genetic clusters, corresponding to Eastern and Western subpopulations. *J. copaia* structure comprised a large cluster Q1 with 80 individuals and a smaller cluster Q2 with 23 individuals only; the latter was removed from the data set used for further analyses. We thus had a total of nine data sets representing eight species. The identification of the most likely K was globally unaffected by the number of loci used in the analysis; Figure S3 (Supporting information) shows that for all data sets, except for the case with four loci, the most likely outcome is $K = 2$. For the four-loci case, the most likely K is between 3 and 4. Smaller numbers of loci do not miss hidden structure in any case.

cpDNA diversity and haplotype distribution

We identified between three (for *C. surinamensis*, *D. guianensis*, *S. amara* and *V. Americana*) and twenty-four haplotypes (for *S. sp1*) at the *trnH-psbA* locus. For *trnC-ycf6*, the number of haplotypes ranged from one in *D. guianensis*, *J. copaia* and *S. amara* to 15 in *S. sp1* (Table S2, Supporting information). For four data sets of nine (*D. guianensis*, *J. copaia*, *V. americana* (East and West), a major haplotype at high frequency was observed. *S. globulifera*, *S. sp1* and *V. michelii* displayed several haplotypes with intermediate frequencies. Haplotype frequency patterns in *S. amara* and *C. surinamensis* were intermediate between these two extremes (Figure S2, Supporting information).

Demographic history (ABC)

ABC detected a variation in N_e for at least eight of the genetic clusters considered in this study (i.e. the credible interval for r_0 does not include 1): a contraction episode was the most probable for eastern *V. americana*, *D. guianensis*, *S. globulifera* and *S. sp1* (Table 1; Fig. 1). For *S. amara*, *J. copaia*, *C. surinamensis* and western *V. americana* populations, the expansion scenario was the most likely (Table 1; Fig. 1), while we detected a possibly constant N_e for *V. michelii*, for which r_0 was larger than 1 for nuclear loci but not for chloroplast data (Table 1; Fig. 1). These results were robust to sampling, as shown by the very similar results obtained on the second drawing (Figure S4, Table S3, Supporting information). The size of demographic change roughly spanned between a 50-fold contraction (for *V. americana* East) and a 100-fold expansion (for *C. surinamensis*; Table 1). In these scenarios, the median age of demographic transitions of all species varied between 131 and 1280 generations (with confidence intervals up to 9350 generations; Table 1). With a generation time of 50 years, these demographic transitions fall within the Quaternary and were dated (median values) between 6550 and 64 000 years. To further check the ABC support for contraction and expansion, we tested, at the 1% acceptance threshold, the two alternative scenarios of contraction and expansion. We found virtually equal probabilities for *V. michelii* ($P = 0.52$ vs $P = 0.48$). For all other species, one of the two scenarios was preferred over the other: a contraction episode was again the most probable for eastern *V. americana* ($P = 0.97$), *D. guianensis* ($P = 0.74$), *S. sp1* ($P = 0.82$) and *S. globulifera* ($P = 0.66$) (Figure S4, Supporting information). For *S. amara*, *J. copaia*, *C. surinamensis* and western *V. americana* populations, the expansion scenario was the most likely, with probabilities of 0.86, 0.88, 0.99 and 0.86, respectively (Figure S5, Supporting information). The sizes of demographic transitions estimated by the 'scenario comparison' approach (Table S4, Figure S6, Supporting information) are similar to those estimated with the 'free variation' approach (Table 1; Fig. 1). PCAs based on the summary statistics computed on the data sets simulated under the most likely scenario and on empirical summary statistics show that the most likely 10 000 scenarios predict well the empirical data, with the possible exception of *V. americana* East (Figure S7, Supporting information). Scenario probabilities varied little when increasing the fraction of simulations used to compute the posterior distributions between 1% and 5% (Figure S4, Supporting information); moreover, probabilities were very similar in the analyses performed on the second data set (not shown). Type I error rate was between 0.04 in the two *Symphonia* species and 0.61 in *V. michelii*, while type II error rate was in the range 0.17–0.43 for seven species, with the two *Symphonia* species characterized by a very high value of 0.89 (Figure S8 and Table S5, Supporting information). Very high type II error rates led to low confidence indices for *S. sp1* and *S. globulifera* (PABC = 0.056 and PABC = 0.055, respectively); confidence indices were moderate to high for the other species, for which PABC

© 2016 John Wiley & Sons Ltd

varied between 0.39 for *V. michelii* and *J. copaia* and 0.68 for eastern *V. americana*. (Tables S4, S5, Supporting information). Results were very similar for the second data set (not shown).

To test for a link between demographic transition signals and ecological strategies, correlations between the medians of r_0 ratios, as obtained from the 'free variation' approach on nuclear markers (Table 1), and functional traits, as well as population dynamics parameters, were tested. Nineteen traits were taken into consideration, covering demography, life history, wood properties and leaf properties. A significant direct correlation was found for laminar toughness of the leaf ($r = 0.68$, $P = 0.042$; Table S6, Supporting information).

Demographic history (maximum likelihood)

Among the seven data sets analysed with the maximum-likelihood approach, three display a significant departure from population equilibrium (i.e. the 95% CI of their r_0 ratio does not encompass 1) with a r_0 ratio estimate higher than 1 suggesting a population expansion (Table 2, Fig. 2): *J. copaia*, *V. michelii* and *C. surinamensis*. For *J. copaia* and *V. michelii*, the expansion is close to fourfold with a rather narrow confidence interval. Estimate of T/2N was also close for the two species (0.0138 and 0.014 for *Jc* and *Vm*, respectively, with overlapping confidence intervals, Table 2). For *C. surinamensis*, the r_0 estimate was higher (335.6) with a larger confidence interval (2.68–4.3 × 10^6). For all other species, the confidence intervals estimated for r_0 are consistent with a stable effective population size.

Discussion

Demographic scenarios of population expansion and contraction were tested in the eight tree species by coalescent modelling coupled with ABC and maximum-likelihood (ML) approaches. Ratios of present-to-past diversity parameters, r_0, were reliably estimated at least for the ABC approach and for a subset of the species for the ML approach, as indicated by the narrow credible intervals and peaked posterior distributions (Tables 1 and 2, Figs 1 and 2), suggesting that demographic transitions were correctly captured by the model and the data, even in the absence of precise estimates for individual parameters (Tables 1 and 2). This allowed us to directly compare changes in effective population sizes over time across species. The ABC suggests that four genetic clusters (Eastern *V. americana*, *Dicorynia guianensis*, *S. globulifera* and *S. sp1*) underwent population contraction; another four (*S. amara*, *Jacaranda copaia*, *Carapa surinamensis* and Western *V. americana*) bear a signature of expansion, and *V. michelii* probably

Table 1 Parameter estimation for the ABC modelling with free population size variation. All parameter values are provided with their credible interval at 95%. *t*: Median of the time to demographic event; θ_0, θ_1: median of (respectively) past and present diversity index ($\theta = N_e\mu$, with N_e the effective population size and μ the mutation rate); r_0: median of the ratio θ_0/θ_1. Population diversity parameters and ratios are shown for both nuclear (nSSR) data and for chloroplast (sequence) data

			Nuclear data			Chloroplast data		
Species	Scenario	*t*	Θ_0	Θ_1	r_0	$\Theta_0 \times 10^{-4}$	$\Theta_1 \times 10^{-4}$	r_0
V. americana (East)	Contraction	189 [11.8–5160]	0.047 [0.002–1.73]	2.54 [0.056–40.2]	0.024 [0.022–0.025]	0.2 [0.01–7]	16 [0.35–80]	0.021 [0.0197–0.025]
D. guianensis	Contraction	208 [12–8330]	0.984 [0.003–29.5]	3.95 [0.51–26.2]	0.219 [0.21–0.23]	0.8 [0.037–27]	3.7 [0.32–30]	0.2 [0.19–0.21]
S. sp1	Contraction	131 [10.9–8110]	13.1 [0.3–88.3]	27 [3.12–98.9]	0.484 [0.47–0.50]	30 [0.74–90]	40 [2–87]	0.653 [0.63–0.68]
S. globulifera	Contraction	158 [11–8170]	15.6 [0.34–85.5]	17.2 [1.65–68.3]	0.879 [0.84–0.92]	30 [1–90]	40 [3–87]	0.82 [0.80–0.85]
V. michelii	Constant/expansion	148 [11.3–7850]	22.9 [0.73–89.2]	7.98 [1.86–35.4]	2.576 [2.48–2.69]	35 [1–90]	37 [8–77]	0.99 [0.96–1.02]
S. amara	Expansion	698 [15.2–9330]	11.8 [0.66–81.9]	1.63 [0.06–12.1]	7.521 [7.19–7.82]	18 [1–86]	1.6 [0.04–14]	10.85 [10.4–11.3]
J. copaia	Expansion	592 [12.9–9200]	52.1 [2.96–97.9]	1.4 [0.2–6.57]	32.37 [31.4–33.5]	30 [1.6–90]	8 [0.9–36]	4.05 [3.93–4.15]
C. surinamensis	Expansion	1280 [20–9350]	1.89 [0.11–40.1]	0.031 [0.0008–0.79]	80.81 [76.1–84.6]	4.6 [0.27–70]	0.1 [0.0029–4]	51.7 [49.4–54.5]
V. americana (West)	Expansion	629 [14.7–7870]	2.34 [0.19–52.1]	0.21 [0.0048–9.27]	19.97 [18.7–21.3]	2.5 [0.14–55]	0.22 [0.0046–10]	18.95 [17.9–20.1]

© 2016 John Wiley & Sons Ltd

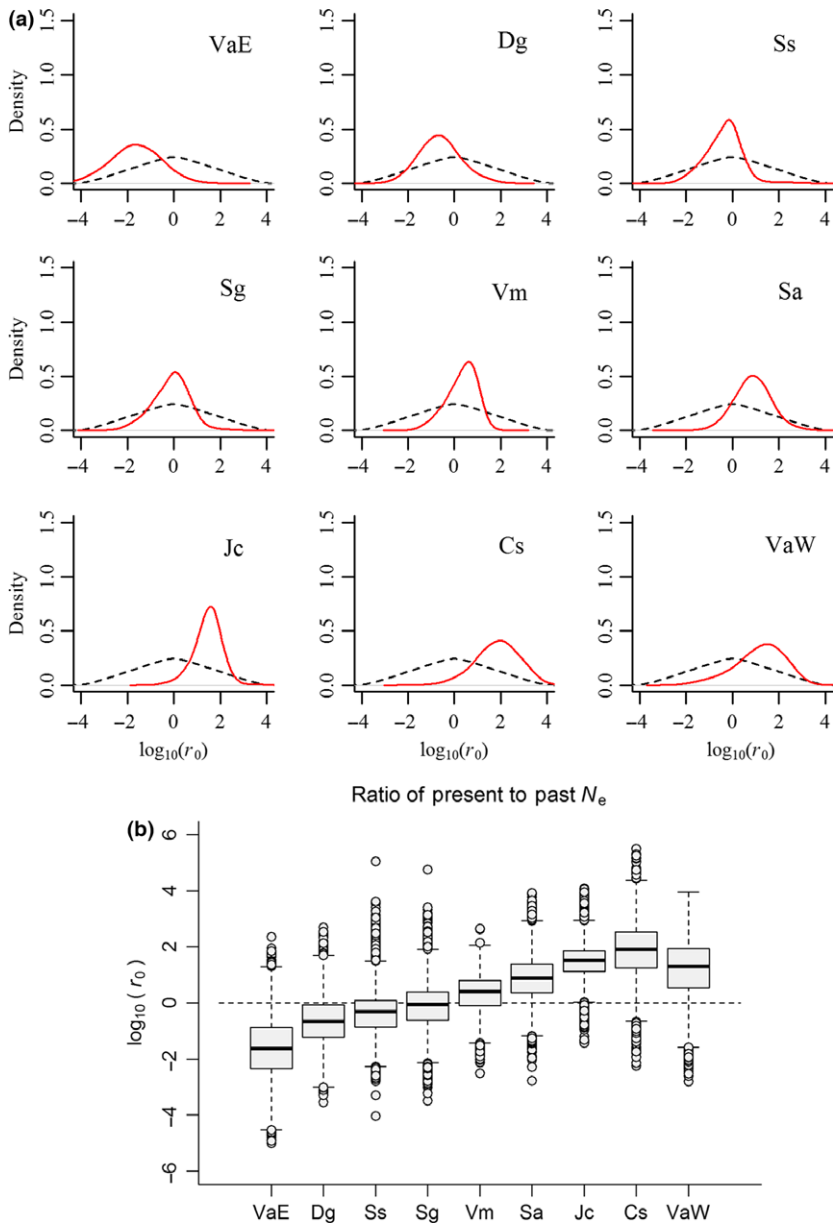

Fig. 1 Density distributions (a) and box plot (b) of posterior estimates of the ratio (r_0) of present (θ_0) and past (θ_1) diversity population index ($r_0 = \theta_0/\theta_1$) from the ABC 'free variation' approach. In figure 1a, dashed lines: prior distribution; full lines: posterior distribution. Species/group names: *Carapa surinamensis* (Cs), *Dicorynia guianensis* (Dg), *Jacaranda copaia* (Jc), *Simarouba amara* (Sa), *Symphonia globulifera* (Sg), *Symphonia sp1* (Ssp1), *Virola michelii* (Vm), eastern *Vouacapoua americana* (Va(E)) and western *Vouacapoua Americana* (Va(W)). [Colour figure can be viewed at wileyonlinelibrary.com].

maintained constant population size. The ML approach behaves somewhat more conservatively in the detection of population size changes, and only detected three instances of population expansion, of which two (*J. copaia*, *C. surinamensis*) in common with the ABC approach, while no contraction was detected. One explanation could be the low contraction detection rate (CDR) of MIGRAINE when the contraction is weak (CDR < 50% with $r_0 = 0.2$) (Leblois *et al.* 2014). According to the results of the ABC approach, the size of the contraction is indeed weak, ranging between 1 and 4 for *D. guianensis*, *S. sp1* and *S. globulifera*. Medians of r_0 estimates were nevertheless rather consistent between the two methods, with the only exception of *S. sp1*, for which the median is 0.484 (thus suggesting contraction) in the ABC and 260 541 for the ML (Tables 1 and 2). However, the latter value is definitely not realistic and is associated to a very large confidence interval, suggesting a very poor estimation. Among ecological and physiological properties, leaf laminar toughness shows marginally significant, positive correlation with present-to-past population size ratio r_0, possibly suggesting that tree species with tougher leaves have expanded since the Pleistocene, while species with less robust laminae have retreated. According to Kitajima & Poorter (2010) and Westbrook *et al.* (2011), higher leaf toughness is frequent in species having easier access to resources, and in particular light, at the juvenile stages; thus, leaf

© 2016 John Wiley & Sons Ltd

Table 2 Parameter estimates for each cluster from the maximum-likelihood approach. All parameter estimates are provided with their confidence interval at 95%; *t*: time to demographic event with $t = T_{\text{in generations}}/2N_{\text{present}}$; $2\theta = N_e\mu$, with N_e the past or current effective size of population and μ the mutation rate; r_0: the ratio $\theta_{(\text{present})}/\theta_{(\text{past})}$; pGSM: parameter of the geneneralized distribution of the generalized stepwise-mutation model

Species	Scenario	*t*	$2\Theta_{n(\text{present})}$	$2\Theta_{n(\text{past})}$	$r_{0\ n}$	pGSM
D. surinamensis	Constant	0.06 [1.21e−08–3560]	0.009 [8.76e−09–3.4–NA]	2.364 [0.000184–4164]	0.0039 [5.67e−09–9312]	0.495 [NA–0.739]
S. sp1	Constant	11.39 [7.81e−09–83.29]	17 [1.03e−06–716.5]	6.52e−05 [1.92e−05–NA]	260 541 [4.86e−09–1 327 030]	0.731 [0.225–0.881]
S. globulifera	Constant	11.7 [1.76e−05–NA]	1.26 [0.000571–88.93]	25.22 [6.34e−08–NA]	0.0499 [2.87e−05–786 629 957]	0.364 [NA–0.691]
V. michelii	Expansion	0.0144 [0.00342–0.0733]	28.18 [17.3–46.36]	7.172 [2.845–14.41]	3.929 [2.13–14.67]	0.656 [0.536–0.761]
S. amara	Constant	0.0171 [1.74e−09–NA]	4.85e−09 [1.2e−09–NA]	1.408 [0.00113–1366]	3.45e−09 [9.34e−10–1753]	0.368 [NA–0.627]
J. copaia	Expansion	0.0138 [0.0012–0.0594]	23.75 [13.1–85.671]	5.381 [2.661–9.667]	4.413 [2.155–10.68]	0.633 [0.538–0.719]
C. surinamensis	Expansion	0.117 [0.000285–0.408]	1.52 [0.72–43.95]	0.00453 [2.33e−07–0.808]	335.6 [2.681–4 349 801]	0.552 [0.369–0.735]

toughness is indirectly linked to forest gap dependence. Actually, the species showing positive r_0 values (either in the ABC or in the ML, or in both) are gap-dependent or favoured by gaps (Molino and Sabatier, 2001). Such a pattern suggests that stable, old-growth forests would have been partially replaced by disturbance-rich forests over the Pleistocene. The overall picture is, nevertheless, not as clear-cut as this sketchy description may suggest, because confidence in ABC scenario choice for two of the species having undergone contraction (*S. globulifera* and *S. sp1*) is quite poor, as indicated by very low PABC values (mostly due to very high type II error rates), and because the only very strong signature of contraction, in the Eastern population of *V. americana*, seems to be compensated by strong expansion for the other population of the same species. Therefore, we can be confident that observed expansions have occurred (except for *J. copaia*, which has a relatively low PABC, and for which cpDNA actually points to contraction, Table 1), while we cannot be sure of the interpretation of contraction signals (except for *D. guianensis*). On the other hand, the ML analysis strongly supports expansion in three cases out of seven, and never supports (strong) contraction, thus suggesting that some species persisted throughout geological changes in forest community composition while others took advantage of the Pleistocene–Holocene climate changes; in particular, the three species showing expansions all thrive in forest gaps, suggesting that more gaps may have become available since the last ice age,

Where does this leave us relative to currently available hypotheses on the biogeographic history of Amazonian forests? On one hand, the mixture of contraction, stability, and expansion signals shown by both methods agrees with the hypothesis of a continuous and permanent forest cover since the last glacial maximum in the north-western Amazon Basin (Colinvaux *et al.* (1996), Bush *et al.* (2002) and Mayle *et al.* (2004)), with different species reacting differently to climatic oscillations. On the other hand, the overall higher confidence in expansion signals, as well as the complete lack of support for contractions shown by the ML approach, agrees with refugial hypotheses (Haffer 1969; Prance 1982) and floristic transition assumptions (Colinvaux *et al.* 2000; Bush *et al.* 2004), possibly with fast-growing, light-demanding and pioneer species having responded more intensely than shade-tolerant ones to climatic change. Thus, Guiana Shield forests appear to have had a more complicated history than temperate and boreal forests (Taberlet *et al.* 1998; Hewitt 2000, 2004; Petit *et al.* 2002; Hu *et al.* 2009): a mixture of gradual compositional changes and expansion from refugia. *V. americana* may be a typical example of this situation, as it shows population contraction in its hypothesized

© 2016 John Wiley & Sons Ltd

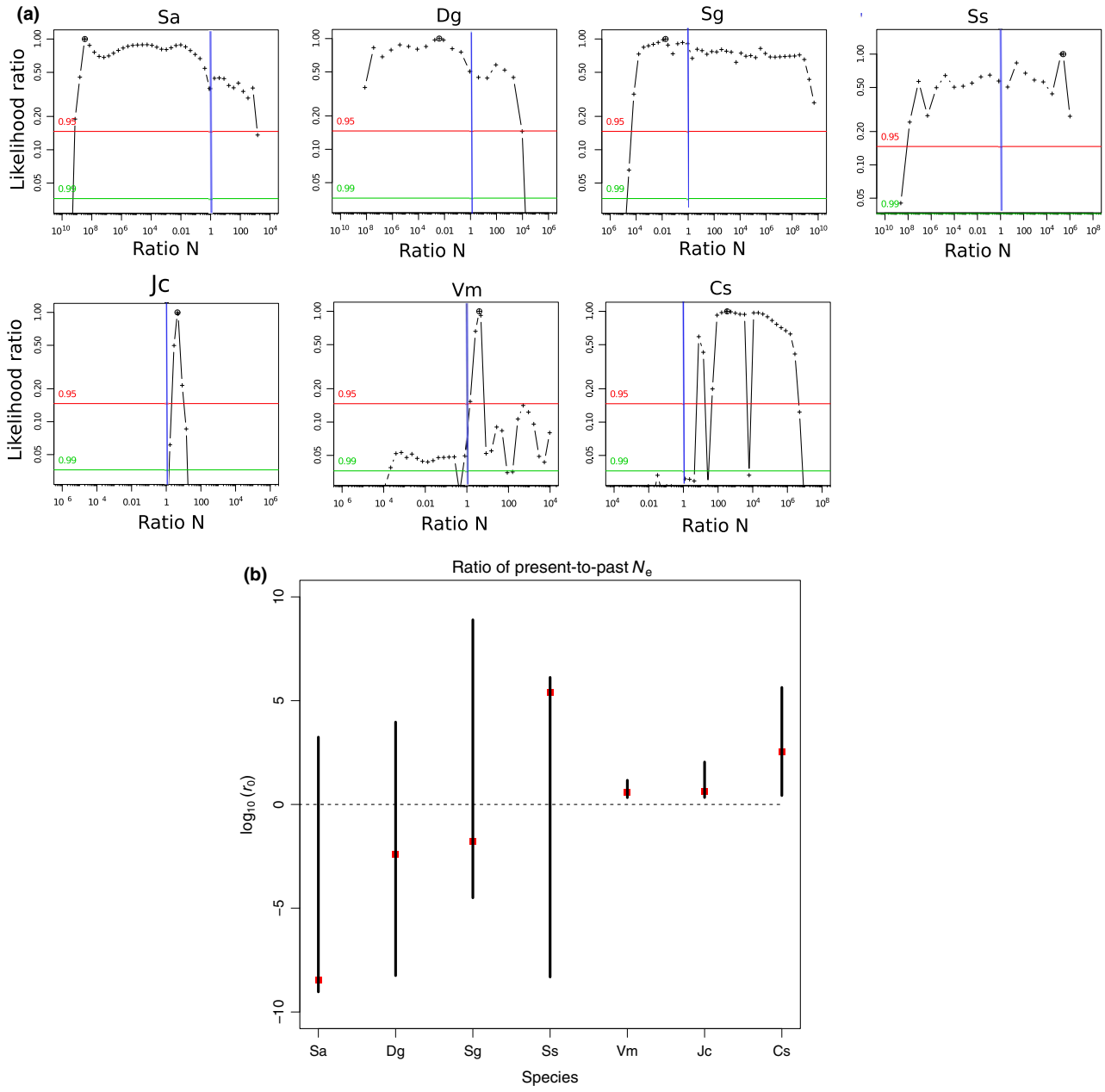

Fig. 2 Profile (a) and barplot (b) of likelihood of the ratio (r_0) of present (θ_0) and past (θ_1) diversity population index ($r_0 = \theta_0/\theta_1$) from the maximum likelihood approach. In pane (b), the bar represents the confidence interval and the square represents the maximum-likelihood point. Species/group names: *Carapa surinamensis* (Cs), *Dicorynia guianensis* (Dg), *Jacaranda copaia* (Jc), *Simarouba amara* (Sa), *Symphonia globulifera* (Sg), *Symphonia sp1* (Ssp1), *Virola michelii* (Vm). [Colour figure can be viewed at wileyonlinelibrary.com].

refuge (Dutech *et al.* 2003, 2004), and expansion in supposedly newly colonized areas. In addition, at least part of the demographic transitions may have been caused by purely neutral events, which would have caused random fluctuations in space and time of the composition of communities through demographic stochasticity (Hubbell 2001).

In Central America, Jones *et al.* (2013) detected historical bottlenecks that may coincide with the end of the LGM for *J. copaia* and *S. amara* but not for *S. globulifera*, for which the bottleneck is older. Our study agrees with Jones *et al.* (2013) for the behaviour of *S. globulifera*, which shows population contraction, while *J. copaia* and *S. amara* underwent population expansion. This suggests that the same species may have responded differently in different parts of their distribution range, as Mayle *et al.* (2004) had already suggested with their palaeo-vegetation model, perhaps because climatic

© 2016 John Wiley & Sons Ltd

conditions did not change in the same way, or to the same extent, in different regions (Colinvaux 1987). The fact that different species have responded differently to past climate changes, and that the same species may have responded differently in different regions, suggests that current climate change will likely have complex effects on tropical rainforests, and that predictions may have to consider responses at the species and population level, and to take genetic variation into account (Benito Garzón *et al.* 2011).

Conclusion

Our results are a direct, although inferential, evidence of complex changes undergone by the forest communities of the Guiana Shield in the distant climatic past, something which could not be precisely tested with standard historical methods due to the paucity of fossil evidence. We detect a trend towards both forest expansion and enrichment in light-dependent species, thus somehow merging two biogeographical hypotheses considered so far as alternative. The extensive application of this method to other regions of the Amazon basin may contribute, in the future, to the detailed reconstruction of the ecological history of Amazonian forests.

Acknowledgments

We wish to thank Cyril Dutech, Henri Caron, William Montaigne, Maxime Casalis, Christine Le Mounier and Luc Allard for providing a part of the molecular marker data; Valérie Troispoux and Saint-Omer Cazal for technical support; Jean-François Molino, the Hydreco Team and Pierre Sire for sampling; Katalin Csilléry for critically reading the manuscript; Raphaël Leblois for advice on Migraine runs; the Managing Editor and three anonymous reviewers for their very useful criticisms.

Funding

Financial support was provided by the EU-funded INCO 'SEEDSOURCE' project, by the EU-funded PO-FEDER 'ENERGIRAVI' program, by the ANR-funded 'BRIDGE' and 'FLAG' (ANR- 12-ADAP-0007-01) projects and by an 'Investissement d'Avenir' grant managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-0025). SB's Doctoral programme has been cofunded by the European Social Fund and by INRA. The funding agencies had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

Aide TM, 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, Johan JAA, Nylander AA, Persson C, Sanmartin I (2009) Tracing the impact of the Andean uplift on neotropical plant evolution. *Proceedings of the National Academy of Sciences of the USA*, **106**, 9749–9754.

Aubry-Kientz M, Hérault B, Ayotte-Trépanier C, Baraloto C, Rossi V (2013) Toward trait-based mortality models for tropical forests. *PLoS ONE*, **8**, e63678.

Baraloto C, Paine CET, Poorter L *et al.* (2010) Decoupled leaf and stem economics in rain forest trees. *Ecology Letters*, **13**, 1338–1347.

Beaumont MA (2010) Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution and Systematics*, **41**, 379–405.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Benito Garzón M, Alía R, Robson TM, Zavala MA (2011) Intraspecific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766–778.

Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Birky CW Jr (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the USA*, **92**, 11331–11338.

Budde KB, Gonzalez-Martinez SC, Hardy OJ, Heuertz M (2013) The ancient tropical rainforest tree *Symphonia globulifera* L. f. (Clusiaceae) was not restricted to postulated Pleistocene refugia in Atlantic Equatorial Africa. *Heredity*, **111**, 66–76.

Bush MB, De Oliveira PE (2006) The rise and fall of the Refugial hypothesis of Amazonian speciation: a paleo-ecological perspective. *Biota Neotropica*, **6**, bn00106012006.

Bush MB, Miller MC, De Oliveira PE, Colinvaux PA (2002) Orbital forcing signal in sediments of two Amazonian lakes. *Journal of Paleolimnology*, **27**, 341–352.

Bush MB, De Oliveira PE, Colinvaux PA, Miller MC, Moreno JE (2004) Amazonian paleoecological histories: one hill, three watersheds. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **214**, 359–393.

Bush MB, Silman MR, Mc Michael C, Saatchi S (2008) Fire, climate change and biodiversity in Amazonia: a Late-Holocene perspective. *Philosophical Transactions of the Royal Society B*, **363**, 1795–1802.

Caron H, Dumas S, Marque G *et al.* (2000) Spatial and temporal distribution of chloroplast DNA polymorphism in a tropical tree species. *Molecular Ecology*, **9**, 1089–1098.

Charles-Dominique P, Blanc P, Larpin D *et al.* (1998) Forest perturbations and biodiversity during the last ten thousand years in French Guiana. *Acta Oecologica*, **19**, 295–302.

Colinvaux PA (1987) Amazon diversity in light of the paleoecological record. *Quaternary Science Reviews*, **6**, 93–114.

Colinvaux PA, De Oliveira PE, Moreno JE, Miller MC, Bush MB (1996) A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science*, **274**, 85–88.

Colinvaux PA, De Oliveira PE, Bush MB (2000) Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. *Quaternary Science Reviews*, **19**, 141–169.

© 2016 John Wiley & Sons Ltd

Colpaert N, Cavers S, Bandou E, Caron H, Gheysen G, Lowe AJ (2005) Sampling tissue for DNA analysis of trees: trunk cambium as an alternative to canopy leaves. *Silvae Genetica*, **54**, 265–269.

Cornuet J-M, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, **25**, 410–418.

Daïnou K, Bizoux J-P, Doucet J-L, Mahy G, Hardy OJ, Heuertz M (2010) Forest refugia revisited: nSSRs and cpDNA sequences support historical isolation in a wide-spread African tree with high colonization capacity, *Milicia excelsa* (Moraceae). *Molecular Ecology*, **19**, 4462–4477.

Duminil J, Caron H, Scotti I, Cazal S-O, Petit RJ (2006) Blind population genetics survey of tropical rainforest trees. *Molecular Ecology*, **15**, 3505–3513.

Dutech C, Maggia L, Tardy C, Joly HI, Jarne P (2003) Tracking a genetic signal expansion of extinction-recolonization events in a neotropical tree species: *Vouacapoua americana* aublet in french Guiana. *Evolution*, **57**, 2753–2764.

Dutech C, Joly HI, Jarne P (2004) Gene flow, historical population dynamics and genetic diversity within French Guianan populations of a rainforest tree species, *Vouacapoua americana*. *Heredity*, **92**, 69–77.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Genetics*, **5**, 435–445.

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591–1604.

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, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Flanagan NS, Tobler A, Davison A *et al.* (2004) Historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. *Proceedings of the National Academy of Sciences*, **101**, 9704–9709.

Freycon V, Krencker M, Schwartz D, Nasi R, Bonal D (2010) The impact of climate changes during the Holocene on vegetation in northern French Guiana. *Quaternary Research*, **73**, 220–225.

Ghirotto S, Mona S, Benazzo A, Paparazzo F, Caramelli D, Barbujani G (2010) Inferring genealogical processes from patterns of bronze-age and modern DNA variation in Sardinia. *Molecular Biology and Evolution*, **27**, 875–886.

Haberle S (1997) Upper quaternary vegetation and climate history of the Amazon basin: correlating marine and terrestrial pollen records. In: *Proceedings of the Ocean Drilling Program, Scientific Results* (eds Flood R, Piper D, Klaus A, Peterson L), vol. 155, pp. 381–396. Ocean Drilling Program/Texas A&M University, College Station, TX.

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

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hardesty BD, Dick CW, Hamrick JL, Degen B, Hubbell SP, Bermingham E (2010) Geographic influence on genetic structure in the widespread neotropical tree *Simarouba amara* (Simaroubaceae). *Tropical Plant Biology*, **3**, 28–39.

Hasegawa M, Kishino H, Yano T-A (1985) Dating of the Human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hérault B, Bachelot B, Poorter L *et al.* (2011) Functional traits shape ontogenetic growth trajectories of rain forest tree species. *Journal of Ecology*, **99**, 1431–1440.

Heuertz M, De Paoli E, Kallman T *et al.* (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway Spruce [*Picea abies* (L.) Karst]. *Genetics*, **174**, 2095–2105.

Hewitt GM (2000) The genetic legacy of the quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2004) The structure of biodiversity – insights from molecular phylogeography. *Frontiers in Zoology*, **1**, doi: 10.1186/1742-9994-1-4.

Hoorn C, Wesselingh FP, ter Steege H *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Hu FS, Hampe A, Petit RJ (2009) Paleoecology meets genetics: deciphering past vegetational dynamics. *Frontiers in Ecology and the Environment*, **7**, 371–379.

Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

Jones FA, Ceron-Souza I, Hardesty BD, Dick CW (2013) Genetic evidence of quaternary demographic changes in four rain forest tree species sampled across the Isthmus of Panama. *Journal of Biogeography*, **40**, 720–731.

Kimura M (1985) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Kingman JFC (1982) On the genealogy of large populations. *Journal of Applied Probability*, **19**, 27–43.

Kitajima K, Poorter L (2010) Tissue-level leaf toughness, but not lamina thickness, predicts sapling leaf lifespan and shade tolerance of tropical tree species. *New Phytologist*, **186**, 708–721.

Lander TA, Oddou-Muratorio S, Prouillet-Leplat H, Klein EK (2011) Reconstruction of a beech population bottleneck using archival demographic information and Bayesian analysis of genetic data. *Molecular Ecology*, **20**, 5182–5196.

Latouche-Halle C, Ramboer A, Bandou E, Caron H, Decroocq S (2002) Isolation and characterization of microsatellite markers in the tropical tree species *Dicorynia guianensis* (Caesalpinaceae). *Molecular Ecology Notes*, **2**, 228–230.

Leblois R, Pudlo P, Bertaux F, Beeravolu C-R, Vitalis R, Rousset F (2014) Maximum-likelihood inference of population size contractions from microsatellite data. *Molecular Biology and Evolution*, **31**, 2805–2823.

Ledru M-P, Bertaux J, Sifeddine A (1998) Absence of last glacial maximum records in lowland tropical forests. *Quaternary Research*, **49**, 233–237.

Mayle FE, Power MJ (2008) Impact of a drier early–mid-Holocene climate upon Amazonian forests. *Philosophical Transactions of the Royal Society B*, **363**, 1829–1838.

© 2016 John Wiley & Sons Ltd

Mayle FE, Burbridge R, Killeen TJ (2000) Millennial-scale dynamics of southern Amazonian rain forests. *Science*, **290**, 2291–2294.

Mayle FE, Beerling DJ, Gosling WD, Bush MB (2004) Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philosophical Transactions of the Royal Society London B*, **359**, 499–514.

Molino JF, Sabatier D (2001) Tree diversity in tropical rain forests: a validation of the intermediate disturbance hypothesis. *Science*, **294**, 1702–1704.

Motamayor JC, Lachenaud P, da Silva e Mota JW *et al.* (2008) Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE*, **3**, e3311.

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

Nordborg M (2004) *Coalescent Theory. Handbook of STATISTICAL Genetics*. John Wiley & Sons, New York.

Petit RJ, Brewer S, Bordacs S *et al.* (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53**, 793–808.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Power MJ, Bush MB, Behling H, Horn SP, Mayle FE, Urrego DH (2010) Paleofire activity in tropical America during the last 21 ka: a regional synthesis based on sedimentary charcoal. *Pages News*, **18**, 73–75.

Prance GT (1982) A review of the phytogeographic evidences for Pleistocene climate changes in the neotropics. *Annals of the Missouri Botanical Garden*, **69**, 594–624.

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

Pyhäjärvi T, Garcia-Gil MR, Knurr T, Mikkonen M, Wachowiak W, Savolainen O (2007) Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics*, **177**, 1713–1724.

Quijada-Mascarenas JA, Ferguson JE, Pook CE, Da Graça Salomao M, Thorpe RS, Wüster W (2007) Phylogeographic patterns of trans-Amazonian vicariants and Amazonian biogeography: the Neotropical rattlesnake (*Crotalus durissus* complex) as an example. *Journal of Biogeography*, **34**, 1296–1312.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reich DE, Feldman MW, Goldstein DB (1999) Statistical properties of two tests that use multilocus data sets to detect population expansions. *Molecular Biology and Evolution*, **16**, 453–466.

Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM (2001) Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science*, **293**, 2242–2245.

Sabatier D, Grimaldi M, Prévost M-F *et al.* (1997) The influence of soil cover organization on the floristic and structural heterogeneity of a Guianan rain forest. *Plant Ecology*, **131**, 81–108.

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

Scotti-Saintagne C, Dick CW, Caron H *et al.* (2013a) Amazon diversification and cross-Andean dispersal of the widespread Neotropical tree species *Jacaranda copaia* (Bignoniaceae). *Journal of Biogeography*, **40**, 707–719.

Scotti-Saintagne C, Dick CW, Caron H *et al.* (2013b) Phylogeography of a species complex of lowland Neotropical rain forest trees (*Carapa*, Meliaceae). *Journal of Biogeography*, **40**, 676–692.

Shaw J, Lickey EB, Beck JT *et al.* (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Solomon SE, Bacci MJ, Martins JJ, Vinha GG, Mueller UG (2008) Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* spp.) provide new insight into the origins of Amazonian diversity. *PLoS ONE*, **3**, e2738.

Stadler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009) The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics*, **182**, 205–216.

Swofford DL (1991) PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima F (1989a) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.

Tajima F (1989b) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tardy C (1998) *Paléoincendies naturels, feux anthropiques et environnements forestiers de Guyane française du tardiglaciaire à l'holocène recent. Approches chronologique et anthracologique*. Thèse de Doctorat, Sciences et Techniques, Montpellier II, France, 343.

de Thoisy B, Da Silva AG, Ruiz-Garcia M *et al.* (2010) Population history, phylogeography, and conservation genetics of the last Neotropical mega-herbivore, the lowland tapir (*Tapirus terrestris*). *BMC Evolutionary Biology*, **10**, 278.

Thuillet A-C, Bru D, David J *et al.* (2002) Direct estimation of mutation rate for 10 microsatellite loci in Durum Wheat, *Triticum turgidum* (L.) Thell. ssp. *durum* desf. *Molecular Biology and Evolution*, **19**, 122–125.

Vigouroux Y, Jaqueth JS, Matsuoka Y *et al.* (2002) Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology and Evolution*, **19**, 1251–1260.

Vinson CC, Azevedo VCR, Sampaio I, Ciampi AY (2005a) Development of microsatellite markers for *Carapa guianensis* (Aublet), a tree species from the Amazon forest. *Molecular Ecology Notes*, **5**, 33–34.

Westbrook JW, Kitajima K, Burleigh JG, Kress WJ, Erickson DL, Wright SJ (2011) What makes a leaf tough? Patterns of correlated evolution between leaf toughness traits and demographic rates among 197 shade-tolerant woody species in a Neotropical forest. *The American Naturalist*, **177**, 800–811.

Wolfe K, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast,

© 2016 John Wiley & Sons Ltd

and nuclear DNAs. *Proceedings of the National Academy of Sciences of the USA*, **84**, 9054–9058.

Zhivotovsky LA, Feldman MW, Grishechkin SA (1997) Biased mutations and microsatellite variation. *Molecular Biology and Evolution*, **14**, 926–933.

S.B., B.H., C.S.S. and I.S. conceived the experiments. S.B., C.S.S. obtained nSSR genotypes and DNA sequences. D.S. contributed floristic abundance data. S.B., G.B., B.H., C.S.S. and I.S. analysed the data. All authors wrote the manuscript.

Data accessibility

Original data are accessible at Dryad (www.datadryad.org) at the following doi: http://dx.doi.org/10.5061/dryad.b58dv.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Some of life history traits related to reproductive strategies and ecological characteristics of studied species.

Table S2 (a) Sample size and genetic diversity per type of molecular markers and per data set for the eight studied species. (b) Characteristics of nuclear simple sequence repeats (nSSRs) of each species.

Table S3 Most strongly supported demographic scenario for each cluster and associated parameter values in the second sampling for the 'free variation' approach.

Table S4 Most strongly supported demographic scenario for each cluster, associated PABC index and parameter values.

Table S5 Scenario probabilities for the most likely scenario, error rates and PABC values for each data set.

Table S6 Biological criteria related to the ability of species to grow and colonize.

Figure S1 Location of sampling sites.

Figure S2 Pie charts of chloroplast haplotype frequencies (*trnH-psbA* and *trnC-ycf6* markers).

Figure S3 Likelihood of STRUCTURE *K* values for the *Jacaranda copaia* data and for varying numbers of loci.

Figure S4 Density distributions (1a) and box plot (1b) of posterior estimates of the ratio (r_0) of present (θ_0) and past (θ_1) diversity population index ($r_0 = \theta_0/\theta_1$) *for the second sample draw* and for the 'free variation' approach.

Figure S5 Probability of alternative scenarios as a function of the number of simulations used for computing posterior distributions for each genetic group in the 'scenario choice' approach.

Figure S6 Density distributions (1a) and box plot (1b) of posterior estimates of the ratio (r_0) of present (θ_0) and past (θ_1) diversity population index ($r_0 = \theta_0/\theta_1$) for the 'scenario choice' approach.

Figure S7 Principal Component Analysis plots for summary statistics for the most likely scenario for each species in the 'scenario choice' approach.

Figure S8 Type I and type II error rates for scenario choice for all species/groups.

Methods S1 Development of the PABC index.

© 2016 John Wiley & Sons Ltd