Journal of Biogeography (J. Biogeogr.) (2015) 42, 1443–1454

ORIGINAL ARTICLE

Late Pleistocene molecular dating of past population fragmentation and demographic changes in African rain forest tree species supports the forest refuge hypothesis

Jérôme Duminil[1,2]*, Stefano Mona[3], Patrick Mardulyn[1], Charles Doumenge[4], Frédéric Walmacq[1], Jean-Louis Doucet[5] and Olivier J. Hardy[1]

[1]Evolutionary Biology and Ecology, Université Libre de Bruxelles, 1050 Brussels, Belgium, [2]Bioversity International, Forest Genetic Resources Programme, Sub-Regional Office for Central Africa, Yaoundé, Cameroon, [3]Institut de Systématique, Évolution, Biodiversité ISYEB – UMR 7205 – CNRS, MNHN, UPMC, EPHE Ecole Pratique des Hautes Etudes, Sorbonne Universités, Paris, France, [4]CIRAD, UR B&SEF, Campus International de Baillarguet TA-C-105/D, F-34398 Montpellier cedex 5, France, [5]Laboratoire de Foresterie des Régions Tropicales et Subtropicales, Unité de Gestion des Ressources Forestières et des Milieux Naturels, Gembloux Agro-Bio Tech, Université de Liège, 5030 Gembloux, Belgium

*Correspondence: Jérôme Duminil, Evolutionary Biology and Ecology, CP160 / 12, Faculté des Sciences, Université Libre de Bruxelles, 50 Av. F. Roosevelt, 1050 Brussels, Belgium.
E-mail: jduminil@ulb.ac.be

ABSTRACT

Aim Phylogeographical signatures of past population fragmentation and demographic change have been reported in several African rain forest trees. These signatures have usually been interpreted in the light of the Pleistocene forest refuge hypothesis, although dating these events has remained impracticable because of inadequate genetic markers. We assess the timing of interspecific and intraspecific genetic differentiation and demographic changes within two rain forest *Erythrophleum* tree species (Fabaceae: Caesalpinioideae).

Location Tropical forests of Upper Guinea (West Africa) and Lower Guinea (Atlantic Central Africa).

Methods Six single-copy nuclear genes were used to characterize the phylogeographical patterns of the parapatric sister species *Erythrophleum suaveolens* (characteristic of semi-deciduous or gallery forests) and *Erythrophleum ivorense* (characteristic of evergreen forests). The number of gene pools within each species was determined and the timings of their divergence and past demographic changes were estimated using Bayesian-based coalescent approaches.

Results Three geographically separated gene pools were identified within *E. suaveolens*, and a single gene pool was inferred in *E. ivorense*. All gene pools show signatures of demographic bottlenecks concomitant with the last glacial period (*c.* 120–12 ka). Species-tree inferences show that the two species diverged *c.* 600 ka, whereas the divergence between *E. suaveolens* gene pools was dated to the late Pleistocene (first divergence *c.* 120 ka, second *c.* 60 ka).

Main conclusions (1) Molecular dating of demographic changes of two African tropical forest tree species is consistent with the Pleistocene forest refuge hypothesis. (2) Tree species from Guinean evergreen tropical forests might have been less affected by past climate change than semi-deciduous species. (3) Our phylogeographical data support a recent date (Holocene) of the last opening of the Dahomey Gap.

Keywords
Africa, coalescence methods, Guineo-Congolian forests, Last Glacial Maximum, nuclear gene, past climate change, phylogeography, refugia theory, tropical species, vanishing refuge model.

INTRODUCTION

The evolutionary mechanisms that have shaped the remarkable biodiversity of tropical rain forests are still poorly understood. West and Central African rain forests form the Guineo-Congolian phytochorion, which White (1983) subdivided into three subcentres of plant species endemism: Upper Guinea (UG; from Sierra Leone to Ghana), Lower

© 2015 John Wiley & Sons Ltd

Journal of Biogeography

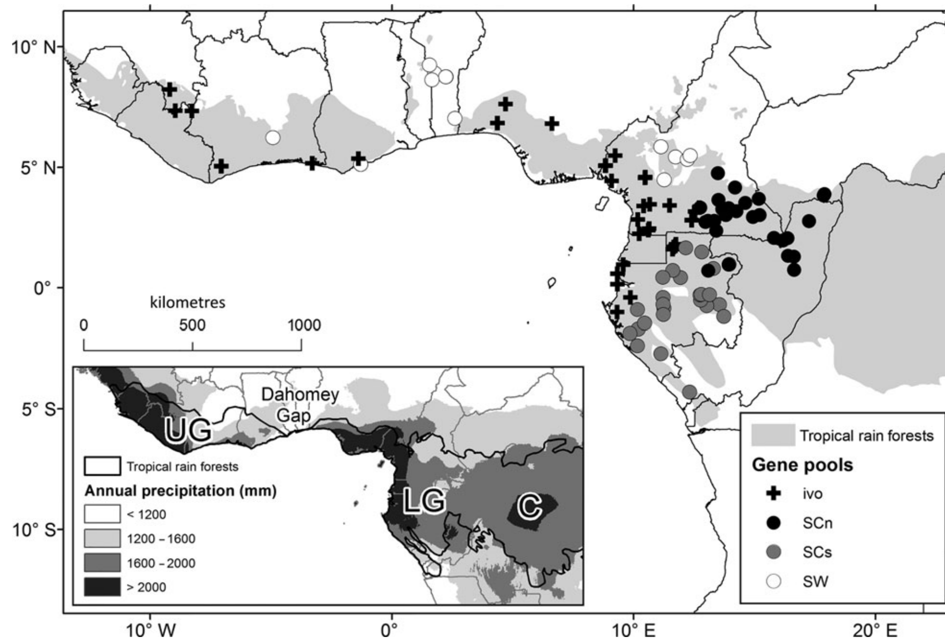

Figure 1 Gene pools in forest *Erythrophleum* species inferred by a Bayesian clustering method. Crosses, *E. ivorense*; circles, *E. suaveolens* (white, SW gene pool; black, SCn gene pool; grey, SCs gene pool). The inset shows the three subcentres of endemism as defined by White (1983): UG, Upper Guinea; LG, Lower Guinea; C, Congolia. White (1983) further defined three main types of rain forest in Guineo-Congolia on the basis of annual rainfall: evergreen (> 2000 mm), semi-deciduous (1600–2000 mm) and dry (1200–1600 mm).

Guinea (LG; from Nigeria to Gabon) and Congolia (Congo Basin; Fig. 1). Upper and Lower Guinea are separated by a dry savanna corridor known as the 'Dahomey Gap' (*c.* 0–3° E). White (1983) further recognized three main lowland forest vegetation types: evergreen forests (rainfall > 2000 mm), semi-deciduous forests (rainfall mostly 1600–2000 mm) and dry forests (rainfall 1200–1600 mm; Fig. 1). Although seminal, the work of White (1983) was mainly descriptive and did not discuss the evolutionary forces responsible for patterns of species endemism and diversity.

Palynological investigations indicate that African rain forests were affected by past climate changes (Dupont *et al.*, 2001; Bonnefille, 2007). In particular, Pleistocene climate oscillations led to the degradation and fragmentation of lowland tropical rain forests during the cool and dry glacial episodes, which favoured the expansion of savanna and mountain forests, whereas maximal rain forest expansion occurred during interglacials. Accordingly, the Pleistocene forest refuge hypothesis (also known as refugia theory) posits that the repeated isolation and re-expansion of rain forests has been a major driver of allopatric speciation (Lönnberg, 1929). Based on patterns of species distribution, different refugia have been proposed within Guineo-Congolia, mainly in mountainous regions (Maley, 1996; Sosef, 1996) or in river flood-plains (Robbrecht, 1996). The influence of Pleistocene climate changes on speciation patterns has been partly reconsidered in the light of molecular phylogenies showing that sister species often diverged before the Pleistocene (reviewed in Plana, 2004). Nevertheless, the refuge hypothesis might help to explain intraspecific patterns of genetic differentiation. Although there is little doubt that tropical Africa underwent substantial changes in vegetation during the Pleistocene, there is still much uncertainty regarding the intensity of forest degradation and the size and location of forest refugia (a few macrorefugia and/or many microrefugia) during glacial maxima, and the impacts these vegetation changes had on patterns of biodiversity.

Phylogeographical approaches, which study the spatial distribution of genetic lineages, can help decipher species evolution in relation to past climate change (Avise *et al.*, 1987). Under the forest refuge hypothesis, the following patterns of genetic diversity are expected in rain forest species: (1) genetic differentiation between populations from (previously) isolated refugia, generating distinct gene pools; (2) areas of genetic admixture between gene pools corresponding to their contact zones after post-glacial recolonization from refugia; (3) phylogeographical signals (i.e. closely related alleles tend to co-occur within populations) between populations that evolved in isolation for a long period; (4) signature of past demographic decline; (5) congruence of phylogeographical patterns across species that underwent similar demographic histories; and (6) timing of divergence between intraspecific gene pools and timing of demographic decline compatible with glacial periods.

In recent years, our knowledge of the organization of genetic diversity within African rain forest plant species has rapidly increased (Debout *et al.*, 2010; Duminil *et al.*, 2010, 2013; Budde *et al.*, 2013; Daïnou *et al.*, 2014; Ley & Hardy,

© 2015 John Wiley & Sons Ltd

2014) and has led to interspecific comparisons (Hardy *et al.*, 2013; Dauby *et al.*, 2014; Heuertz *et al.*, 2014). These studies repeatedly report genetic discontinuities, i.e. different gene pools occurring in parapatry, in regions where species display largely continuous distributions, suggesting past barriers to gene flow. A genetic discontinuity frequently found in LG occurs around the climatic hinge (latitude of N–S seasonal inversion, *c.* 2° N; Hardy *et al.*, 2013) and is congruent with floristic differentiation (Sosef, 1994; Gonmadje *et al.*, 2012). Together with genetic signatures of admixture zones (Debout *et al.*, 2010) and of historical demographic declines (Budde *et al.*, 2013), these findings are consistent with the Pleistocene forest refuge hypothesis. Further confirmation of this hypothesis, however, would require estimates of the timing of demographic changes and the divergence times between gene pools. Previous molecular dating attempts have been inconclusive because the genetic markers used either evolved too slowly, e.g. plastid DNA (pDNA), or were prone to substantial homoplasy, e.g. nuclear microsatellites (nSSR), causing difficulties in time calibration (Duminil *et al.*, 2013). To address this problem, nuclear DNA (nDNA) sequences offer unique advantages in characterizing genealogical relationships (Hare, 2001).

Erythrophleum suaveolens (Guill. & Perr.) Brenan (syn. *E. guineense* G. Don.) and *Erythrophleum ivorense* A. Chev. (syn. *E. micranthum* Harms) (Fabaceae–Caesalpinioideae) are two large sister tree species. *Erythrophleum ivorense* occurs in evergreen forests from Gambia to Gabon (Aubréville, 1950), and is characterized by a disjunct distribution, being present in both LG and UG, but absent from the Dahomey Gap (Akoègninou *et al.*, 2006). *Erythrophleum suaveolens* is more widespread, ranging from Senegal east to Sudan and Kenya, and south to Mozambique and Zimbabwe (Aubréville, 1950). It occurs in semi-deciduous Guineo-Congolian forests (at least in LG and Congolia) and adjacent gallery forests that form part of forest–savanna mosaic landscapes (Aubréville, 1950; Akoègninou *et al.*, 2006). Hence, it typically occupies drier climates than *E. ivorense*, and the two species are rarely found in sympatry (Duminil *et al.*, 2010). Unlike *E. ivorense*, *E. suaveolens* occurs also in the Dahomey Gap (Akoègninou *et al.*, 2006). The distinct ecogeographical characteristics of these species make them interesting models in phylogeographical studies to investigate the influence of past climate changes on the history of different lowland forest vegetation types. Their patterns of genetic diversity were previously analysed using pDNA and nSSR in the Guinean region (defined here as the region encompassing UG and LG; Duminil *et al.*, 2010, 2013). Within LG, nSSR data supported the presence of a north–south historical barrier to gene flow in both species at approximately 2° N. An east–west barrier to gene flow was also found within each species – at the transition between humid and dry tropical forests in Central Cameroon for *E. suaveolens*, and in Nigeria for *E. ivorense* (see Fig. 1 in Duminil *et al.*, 2013). We further detected a signal of historical demographic bottleneck within LG for both species (Duminil *et al.*, 2010, 2013), but our previous data did not allow us to date the inferred bottleneck events or the divergence between intraspecific gene pools.

In the present study, we use six single-copy intergenic nuclear DNA markers to investigate the genetic variation of each species in the Guinean region. We aim to assess whether the timing of divergence between intraspecific gene pools and signatures of past demographic changes provide support for a major role of Pleistocene climate changes and the forest refuge hypothesis. More specifically, we address the following questions: (1) Do nuclear genes confirm that different gene pools occur within *E. ivorense* and *E. suaveolens*? (2) Are the divergence times between these gene pools compatible with species distribution fragmentation driven by cooling events in the last glacial period? (3) Is there evidence of demographic bottlenecks concomitant with the last glacial period? Moreover, given that *Erythrophleum* species are found in different forest types, we tentatively interpret the available phylogeographical data in terms of biogeographical history of evergreen and semi-deciduous tropical forests in the Guinean region, including the Dahomey Gap.

MATERIALS AND METHODS

Model species

Erythrophleum (Fabaceae: Caesalpinioideae) is a pantropical woody genus containing three species in Southeast Asia and China, one species in Australia, one species in Madagascar and four species in mainland Africa. This study focuses on the two recognized African rain forest species: *E. ivorense* and *E. suaveolens*. These two sister species are difficult to distinguish using morphological characters, and so we must rely on molecular-based approaches (Duminil *et al.*, 2010, 2013).

Sampling and DNA extraction

Leaves or cambium of 139 individuals of *E. ivorense* or *E. suaveolens* were sampled (see Appendix S1 in Supporting Information). Samples were selected to maximize geographical spread and genetic diversity, according to prior knowledge of nSSR and pDNA diversity (Duminil *et al.*, 2010, 2013). For the purpose of molecular dating, we also used silica gel samples of four individuals of *E. africanum* from Benin, one individual of *E. chlorostachys* from Australia, three individuals of *E. fordii* from China, and one individual of the closely related *Calpocalyx dinklagei* from Cameroon (Bruneau *et al.*, 2008) (Appendix S1). Additional silica gel samples were provided by collaborators for other closely related species (according to Bruneau *et al.*, 2008) to be used as outgroups in phylogenetic analyses: *Tachigali paraensis*, *Tachigali bracteolata*, *Inga huberi*, *Inga alba*, *Enterolobium schomburgkii*, *Abarema jupunba* and *Zygia racemosa* (Appendix S1). Total DNA was isolated with a NucleoSpin plant kit (Macherey-Nagel, Düren, Germany).

© 2015 John Wiley & Sons Ltd

DNA sequences

We studied five single-copy nuclear conserved orthologue set (COS) genes using primers from Li *et al.* (2008), and one additional single-copy nuclear gene (phosphoenolpyruvate carboxylase, *PEPC*) using primers designed specifically for this study (Appendix S2). To improve amplification of the five COS genes of Li *et al.* (2008), internal primers were specifically designed for this study (Appendix S2).

For the *PEPC* gene, nine primer pairs were designed for the amplification of seven introns (Appendix S2), using Primer Analysis 3.3 (Rychlik & Rhoads, 1989). We tested the universal character of these primers on additional species (Appendix S2) using the PCR protocol described below for the *PEPC* E2/E3 fragment (hereafter, *E2E3*). Only *E2E3* was studied in more depth in this study.

PCR protocol

For the *Agt1*, *Apg1*, *At103*, *ChlP* and *Eif3E* regions (the COS genes), polymerase chain reactions (PCRs) were carried out in a TProfessional Thermocycler (Biometra, Göttingen, Germany). PCRs had a total volume of 25 μL, comprising 2 μL of template DNA (10–100 ng), 0.1 μL of Top *Taq* polymerase (Qiagen), 2 μL of PCR buffer, 1 μL of $MgCl_2$ (25 mm), 0.5 μL of dNTPs (10 μm), 0.25 μL of each primer (10 μm) and 18.9 μL of H_2O. The cycling profile for PCR of these regions included an initial step of 3 min at 94 °C followed by 35 cycles of 30 s at 94 °C, 30 s at 56 °C and 1 min at 72 °C, followed by a final incubation at 72 °C for 5 min. For *E2E3*, PCR reactions comprised 2 μL of template DNA (10–100 ng), 0.1 μL Phusion high-fidelity *Taq* DNA polymerase (Finnzymes, Espoo, Finland), 5 μL 5× Phusion HF buffer, 0.5 μL dNTPs (10 μm), 0.5 μL of each primer (10 μm), and 17.8 μL H_2O. The PCR cycling profile for *E2E3* included 30 s at 98 °C, followed by 35 cycles of 10 s at 98 °C, 30 s at 60 °C, and 1 min at 72 °C. Reactions were then incubated for 5 min at 72 °C. Sequences were resolved on a 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Haplotype definition

Because the six nuclear loci obtained by sequencing were from diploid individuals, multiple heterozygous positions were observed. Insertion–deletion polymorphisms were not considered. The phase 2.1 algorithm (Stephens & Donnelly, 2003), implemented in the software package DnaSP 5.10 (Librado & Rozas, 2009), was used to reconstruct two separate haplotypes per individual at each locus (posterior probability > 0.75) using the default settings. Based on these haplotypes, loci were then screened for evidence of recombination using rdp4.10 (Martin & Rybicki, 2000).

Bayesian clustering analyses

Only *E. suaveolens* and *E. ivorense* individuals that were characterized for at least four of the six nuclear genes were used in this analysis (108 individuals in total). structure 2.3.4 was run 10 times for each number of clusters, *K*, ranging from 1 to 10, in order to detect the presence of gene pools in our dataset (Pritchard *et al.*, 2000). The burn-in period and number of Markov chain Monte Carlo (MCMC) repetitions were set to 10,000 and 40,000, respectively, and no prior information on the population of origin of each sampled individual was included (Evanno *et al.*, 2005). An admixture model in which the fraction of ancestry from each cluster is estimated for each individual was used. The parameter of individual admixture, alpha, was chosen to be the same for all gene pools and was given a uniform prior. The allele frequencies were kept independent among gene pools in order to avoid overestimating the number of gene pools (Falush *et al.*, 2003). The optimum number of gene pools was determined using the delta-*K* method (Evanno *et al.*, 2005). Only individuals with an assignment probability above 0.8 were eventually assigned to a specific gene pool.

Genetic diversity analyses

For comparison with nSSR data, individuals were grouped in populations as in Duminil *et al.* (2013): 13 populations of *E. ivorense* and 20 of *E. suaveolens* were defined (Appendix S1). All nuclear fragments could be aligned unambiguously across all *E. suaveolens* and *E. ivorense* samples. The among-population genetic fixation index (G_{ST}) and the corresponding parameter (N_{ST}) that takes into account phylogenetic similarities between haplotypes (Pons & Petit, 1996), were calculated at the species level or at the level of structure-inferred gene pools (see 'Bayesian clustering analysis' results for the definition of gene pools) for all nuclear genes and for each nuclear gene separately, using spads 1.0 (Dellicour & Mardulyn, 2013). Only populations with at least three individuals were used for these analyses. The number of haplotypes, the allelic richness A_R (El Mousadik & Petit, 1996), the nucleotide diversity π (Nei & Li, 1979), and the nucleotide diversity ratio π_R (Mardulyn *et al.*, 2009) were estimated for each species and each defined gene pool.

Phylogeny and molecular dating analyses

Molecular dating requires calibration points, which were provided by fossil evidence at the scale of angiosperms (Bruneau *et al.*, 2008), but quickly evolving DNA sequences, which are informative at a shallow phylogenetic level, often cannot be aligned at the deep phylogenetic level relevant for fossil evidence. We therefore combined information from markers evolving at different rates, which could be aligned at the scale of: (1) angiosperms (pDNA gene *matK*), (2) *Erythrophleum* and closely related genera (nDNA genes *At103* and *Eif3E*), or (3) *Erythrophleum* species only (all six nDNA genes). We followed an established procedure (i.e. Quinzin & Mardulyn, 2014) to estimate (1) the evolutionary rate of *matK*, (2) the evolutionary rates of *At103* and *Eif3E* and the divergence time between sequences from different *Erythroph-*

© 2015 John Wiley & Sons Ltd

leum species, (3) the divergence time between *Erythrophleum* gene pools.

First, the rate of molecular evolution of *matK* within the *Erythrophleum* lineage was estimated using fossil calibration points on four nodes of the phylogeny (see Duminil *et al.*, 2013). The estimated number of substitutions per site per Myr in *matK* followed a normal distribution with a mean of 0.000276 and a standard deviation of 0.000027.

Second, this distribution was used in BEAST 1.7.4 (Drummond & Rambaut, 2007) to estimate the molecular evolutionary rate of *At103* and *Eif3E* and the timing of divergence between DNA sequences from *E. suaveolens*, *E. ivorense*, *E. africanum*, *E. chlorostachys*, *E. fordii*, *Calpocalyx dinklagei*, *Abarema jupunba*, *Enterolobium schomburgkii* and *Inga huberi*. DNA sequence alignment was conducted for each gene using the MUSCLE algorithm as implemented in CODONCODE ALIGNER 4.0.4. (CodonCode Corporation, MA). MrModeltest 2.3 (Nylander, 2004) was used to determine the best molecular substitution model for each gene according to the Akaike information criterion (AIC) (Table 1). The concatenated sequence alignment was analysed with an uncorrelated lognormal relaxed clock model and a Yule process of speciation. Two MCMC analyses were run for 100 million generations each with the parameters sampled every 5000 steps. Tracer 1.4 (Drummond & Rambaut, 2007) was used to assess convergence and measure effective sample sizes (ESS), ensuring that the MCMC chains had run long enough to estimate the parameters (burn-in 10%). The convergence between the two independent runs was also assessed by calculating the Gelman–Rubin statistic using the R package CODA (Plummer *et al.*, 2008). The mean and 95% highest posterior density intervals for divergence-time estimates were obtained in Tracer (Rambaut *et al.*, 2014). TreeAnnotator was used to obtain an estimate of a phylogenetic tree from the concatenation across the three genes (option 'Maximum clade credibility tree'). The probability distributions of the molecular evolution rates for *At103* and *Eif3E* were estimated as normal distributions with means and standard deviations of 0.00162 ± 0.00070 and 0.00138 ± 0.00065, respectively, which were used as priors in the following *BEAST analyses.

Third, *BEAST (Heled & Drummond, 2010) was used to infer the timing of divergence between *Erythrophleum* gene pools considering the six single-copy nuclear genes. This software estimates a species phylogenetic tree from multilocus and multiple-allele sequences under a Bayesian framework. Only individuals that were sequenced for at least four of the six nDNA genes were used. We did not use pDNA markers due to the presence of incomplete lineage sorting and/or plastid capture between gene pools (see Duminil *et al.*, 2013). The sequences were grouped into seven taxon sets, *E. chlorostachys* ($n = 2$), *E. fordii* ($n = 6$), *E. africanum* ($n = 8$), *E. ivorense* ($n = 64$) and three gene pools within *E. suaveolens*: SCn ($n = 50$), SCs ($n = 46$) and SW ($n = 36$) (see 'Bayesian clustering analyses' result section for the definition of the gene pools). The six genes were considered unlinked, allowing variation in the substitution models inferred by MrModeltest (Table 1). Previously obtained probability distributions of molecular evolution rates for *At103* and *Eif3E* were used as priors for these two genes and uninformative priors were selected for the four other genes. We used an uncorrelated lognormal relaxed clock model, the Yule process as species tree prior and the 'piecewise linear and constant' population size model. Two parallel MCMC analyses were run for 1 billion generations with sampling of the parameters every 50,000 steps. We used the same procedure as for the BEAST analysis to check the quality of the results.

Demographic changes analyses

The demographic history of each gene pool detected in *E. suaveolens* and *E. ivorense* was investigated by reconstructing the changes of the effective population size (N_e) through time using the multilocus extended Bayesian skyline plot (EBSP) implemented in BEAST 1.7.4 (Drummond & Rambaut, 2007). EBSP is a nonparametric Bayesian-based coalescent approach that makes no assumption of the demographic model of the population. Each coalescent interval has its own prior N_e distribution, which is sampled during the MCMC, together with the coalescent tree, the branch lengths and the evolutionary parameters. After the removal of the burn-in, N_e is evaluated at some specified time-point on the recorded iterations of the MCMC and then interpolated to obtain its variation through time. The most likely molecular evolutionary model was used for each of the six genes (Table 1). The coalescent units were converted to calendar years and the N_e was scaled based on: (1) an assumed

Table 1 Characteristics of the six nuclear gene fragments sequenced. Fragment sizes are based on the nucleotide alignment including only samples of *Erythrophleum ivorense* and *E. suaveolens*.

Gene	Size of the fragment	Variable (polymorphic) sites	Parsimony-informative sites	Number of haplotypes	Substitution model
Agt1	599	58	73	41	HKY+I+G
Apg1	393	8	6	11	HKY
At103	388	16	10	14	GTR
ChlP	648	14	12	15	HKY+I
E2E3	711	31	21	30	HKY+G
Eif3E	654	9	7	7	HKY

© 2015 John Wiley & Sons Ltd

generation time of 100 years, estimated by integrating data on annual diameter growth with minimal and maximal diameter of reproduction (Kouadio, 2009), and (2) a molecular clock calibration on *At103* and *Eif3E* as obtained below. A uniform prior between 10^{-12} and 10^{-6} substitutions per site per Myr was applied to the molecular evolutionary rates for the remaining four genes. The MCMC was run for 10^9 iterations with a 10% burn-in and a thinning interval of 10^5. Convergence was checked by running each analysis twice, by visual inspection of the trace of the likelihood, and by recording the ESS of each parameter (always higher than 500).

RESULTS

DNA sequencing

108 out of 139 individuals of *E. ivorense* and *E. suaveolens* used in this study amplified for at least four of the six genes. All eight *PEPC* primer pairs amplified in *Erythrophleum* species. Interestingly, these fragments represent valuable genomic resources to be used in phylogenetic or phylogeographical studies, because they amplified successfully in other tree species, even from other eudicot families (Table 1), and some fragments (*E8E9*, *E2E3* and *E7*) also amplified successfully in monocot taxa (e.g. Marantaceae; Alexandra Ley, Martin-Luther-Universität Halle–Wittenberg, pers. comm.).

Bayesian clustering analyses

Bayesian clustering analyses identified $K = 4$ as the best-supported number of gene pools in *E. ivorense* and *E. suaveolens* (Fig. 1, and see Appendix S3). A single gene pool was inferred across Guinea for *E. ivorense*, whereas three gene pools were identified within *E. suaveolens*: one spanning UG, the Dahomey Gap and the northern part of LG (named 'SW', for *E. suaveolens* in West Africa), and two others in LG. The latter two were approximately located on each side of the climatic hinge (north of the climatic hinge: 'SCn', south of the climatic hinge: 'SCs'; where 'SC' stands for *E. suaveolens* in Central Africa). Clustering solutions for $K = 2$ gene pools separated the two species, whereas for $K = 3$, *E. suaveolens* samples segregated into two gene pools corresponding to SW and the central African samples (i.e. SCn + SCs).

Genetic diversity analyses

At the genus level (considering both *E. suaveolens* and *E. ivorense*), *Agt1* was the most polymorphic locus, followed by *E2E3*, *ChlP*, *At103*, *Apg1* and *Eif3E* (Table 1). *Apg1* was monomorphic within gene pool SW and *Eif3E* was monomorphic within *E. suaveolens*. Every fragment except *Eif3E* was more polymorphic in *E. suaveolens* than in *E. ivorense* (π and π_R statistics; Table 2). Overall, *E. suaveolens* presents stronger population genetic structure than *E. ivorense* (G_{ST} statistics; Table 2). A significant phylogeographical signal, whereby related haplotypes are more likely to be found in the same population ($N_{ST} > G_{ST}$), was found in four out of five genes in *E. suaveolens* and in two out of six genes in *E. ivorense* (Table 2). At the gene-pool level, we found no clear differences regarding polymorphism or genetic structure.

Phylogeny and molecular dating analyses

The *BEAST analysis shows that *E. suaveolens* is monophyletic (its three gene pools form a clade; Fig. 2). The *BEAST and BEAST analyses show that *E. ivorense* and *E. suaveolens* are sister species. The sister group of *E. ivorense* and *E. suaveolens* appears to be either *E. africanum* (BEAST gene tree analysis) or *E. chlorostachys* and *E. fordii* (*BEAST species tree analysis), but with poor support in either case (Fig. 2).

Sequences from *E. suaveolens* and *E. ivorense* have an estimated mean divergence date of 3.11 Ma (BEAST, three genes), whereas the divergence time between these species is estimated at 0.628 Ma (*BEAST, six genes) (Table 3, Fig. 2). The *BEAST analysis supports a late Pleistocene divergence between the *E. suaveolens* gene pools (0.06 and 0.12 Ma; Table 3).

Demographic change analyses

For each *E. suaveolens* gene pool and for *E. ivorense*, the EBSP results suggest that there were historical demographic bottlenecks that started between 125,000 and 50,000 years ago (late Pleistocene; Fig. 3).

DISCUSSION

Gene differentiation patterns and congruence with nSSR results

In this study, six nuclear genes were sequenced to obtain phylogeographical inferences complementary to those already obtained from pDNA and nSSR in *E. suaveolens* and *E. ivorense* (Duminil *et al.*, 2010, 2013). Overall, *E. suaveolens* displayed higher levels of genetic diversity and genetic structure than *E. ivorense* (Table 3). We also detected a phylogeographical signal within *E. suaveolens* (four out of five genes) and *E. ivorense* (two out of six genes), but the multilocus test (based on six genes) was only significant within *E. suaveolens*. As this signal mostly disappears at the gene-pool level, the phylogeographical signal detected within *E. suaveolens* actually results from the substructure in three intraspecific gene pools. Overall, these phylogeographical patterns are congruent with those obtained previously using nSSR (Duminil *et al.*, 2013).

Within-species genetic discontinuities were only partly congruent with those inferred with nSSR. The same three gene pools were inferred within *E. suaveolens* but, within *E. ivorense*, a single gene pool was inferred instead of three as inferred with nSSR. This apparent discrepancy might result from the lower mutation rates of gene sequences than

© 2015 John Wiley & Sons Ltd

Table 2 Estimates of genetic diversity and differentiation, and tests of phylogeographical signals [n, total number of alleles; π, nucleotide diversity (Nei & Li, 1979); π_R, nucleotide diversity (Mardulyn *et al.*, 2009); A_R, allelic richness], at each nuclear gene within *Erythrophleum suaveolens* and *E. ivorense* and within each gene pool.

	Species/gene pool	n	No. of haplotypes	G_{ST}	N_{ST}	π	π_R	A_R
All six genes	*E. ivorense*	–	–	0.219	0.265	0.0016	0.314	–
	E. suaveolens	–	–	0.446	0.556*	0.0066	4.062	–
Agt1	*E. ivorense*	68	10	0.248	0.388*	0.0028	0.175	8
	E. suaveolens	134	32	0.481	0.580*	0.0160	5.722	32
	E. suaveolens SW	40	7	0.539	0.797*	0.0056	0.345	7
	E. suaveolens SCn	48	10	0.184	0.229	0.0088	0.446	10
	E. suaveolens SCs	46	19	0.152	0.229	0.0150	0.832	19
Apg1	*E. ivorense*	62	4	0.059	0.087	0.0006	0.183	4
	E. suaveolens	112	8	0.539	0.705*	0.0034	5.460	8
	E. suaveolens SW	34	1	–	–	–	–	1
	E. suaveolens SCn	36	6	−0.135	−0.063	0.0018	0.443	6
	E. suaveolens SCs	42	3	−0.007	0.000	0.0002	0.068	3
At103	*E. ivorense*	40	4	0.056	0.184*	0.0022	0.596	4
	E. suaveolens	134	11	0.665	0.773*	0.0037	1.678	11
	E. suaveolens SW	40	3	0.055	0.044	0.0006	0.075	3
	E. suaveolens SCn	52	6	0.025	0.041	0.0011	0.126	6
	E. suaveolens SCs	42	5	0.458	0.421	0.0011	0.115	5
ChlP	*E. ivorense*	64	5	0.345	0.327	0.0013	0.421	3
	E. suaveolens	144	10	0.382	0.592*	0.0032	2.371	10
	E. suaveolens SW	40	4	0.191	0.146	0.0013	0.263	4
	E. suaveolens SCn	54	4	0.070	0.087	0.0015	0.237	4
	E. suaveolens SCs	50	6	0.191	0.242	0.0017	0.293	6
E2E3	*E. ivorense*	66	7	0.031	0.025	0.0013	0.196	6
	E. suaveolens	112	23	0.227	0.303	0.0069	5.079	23
	E. suaveolens SW	36	8	0.195	0.023	0.0027	0.349	8
	E. suaveolens SCn	38	11	0.256	0.495*	0.0079	1.042	11
	E. suaveolens SCs	38	11	−0.081	−0.188	0.0064	0.805	11
Eif3E	*E. ivorense*	68	6	0.335	0.335	0.0015	–	6
	E. suaveolens	144	1	–	–	–	–	1
	E. suaveolens SW	40	1	–	–	–	–	1
	E. suaveolens SCn	54	1	–	–	–	–	1
	E. suaveolens SCs	50	1	–	–	–	–	1

*Significant ($P < 0.05$) phylogeographical signal ($N_{ST} > G_{ST}$).

SSRs. If the *E. ivorense* gene pools differentiated relatively recently, the higher mutation rates of SSRs would have allowed them to accumulate new alleles (contributing to allele frequency differentiation) faster than nuclear genes. It is also possible that the higher polymorphism of nSSR markers simply provided more power to detect a weak differentiation.

Molecular dating of demographic events supports the forest refuge hypothesis

Nuclear DNA markers allowed us to conduct molecular dating at the intraspecific level. The two species present signals of past demographic decline at the gene-pool level that occurred during the late Pleistocene, probably during the last glacial episode (*c.* 110–10 ka; Fig. 3). Furthermore, the divergence time between *E. suaveolens* gene pools also supports a late Pleistocene origin, possibly concomitant with the last (*c.* 115–10 ka) or penultimate (*c.* 200–130 ka) glacial periods (Fig. 2). To the best of our knowledge, this is the first successful attempt at molecular dating of the fragmentation of African forest tree populations. These results support the role of Quaternary glaciations in shaping biodiversity patterns of African rain forest trees through forest fragmentation, as suggested from fossil pollen evidence (Dupont *et al.*, 2001; Bonnefille, 2007).

© 2015 John Wiley & Sons Ltd

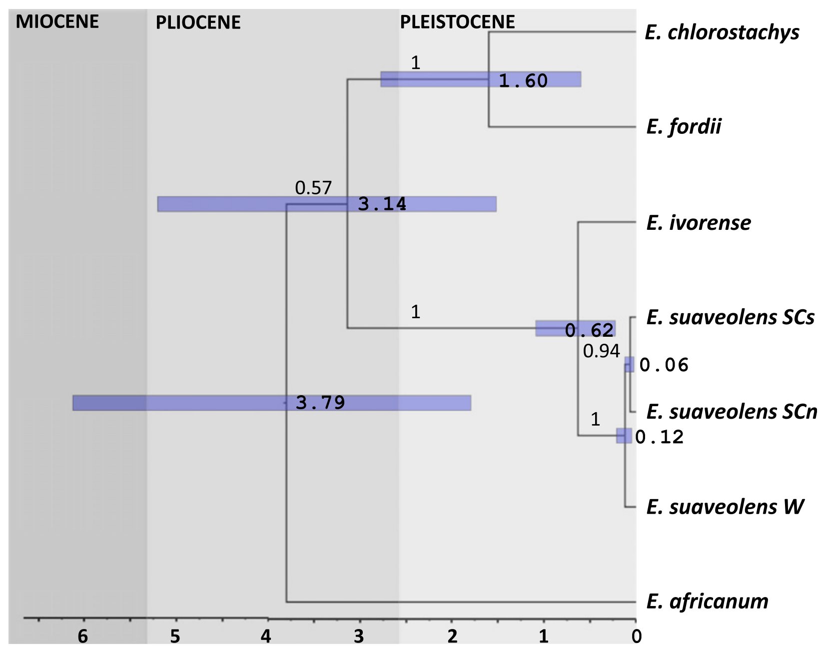

Figure 2 Phylogenetic relationships and *BEAST molecular dating of the divergence between *Erythrophleum* species and gene pools. Numbers on nodes indicate the mean age (in Ma) and numbers on branches the branch support (posterior probability). Bars indicate the 95% highest posterior density intervals around node ages. The different grey zones indicate, from left to right, the Miocene (23.03–5.33 Ma), the Pliocene (5.33–2.58 Ma) and the Pleistocene (2.58–0.01 Ma).

Table 3 Results of molecular dating analyses (mean, in Ma; 95% highest posterior density interval in parentheses) in the African species *Erythrophleum africanum*, *E. suaveolens* and *E. ivorense* and the Asian species *E. chlorostachys* and *E. fordii* at the six loci *Agt1*, *Apg1*, *At103*, *ChlP*, *E2E3* and *Eif3E*.

Estimate of divergence between	*Eif3E–At103–matK*, gene tree	Six genes, species tree
African *Erythrophleum* and Asian *Erythrophleum*	7.23 (1.25–15.48)	–
E. africanum and *E. suaveolens–E. ivorense*	6.42 (1.00–14.06)	–
E. chlorostachys and *E. fordii*	4.81 (0.74–10.59)	1.60 (0.59–2.77)
E. suaveolens and *E. ivorense*	3.11 (0.21–7.44)	0.62 (0.22–1.08)
E. suaveolens SC and *E. suaveolens* SW		0.117 (0.044–0.207)
E. suaveolens SCn and *E. suaveolens* SCs		0.064 (0.021–0.200)

It is worth noting that estimation of the divergence time between *E. ivorense* and *E. suaveolens* varies widely depending on the molecular markers and methods used. It is important here to emphasize the difference between phylogenetic gene trees, in which nodes correspond to the coalescence of gene lineages sampled in different species, and phylogenetic species trees, in which nodes indicate speciation events (Pamilo & Nei, 1988). Phylogenetic species trees are often equated to gene trees inferred from single samples per species, but this can lead to strongly overestimated divergence times in organisms characterized by long generation times and/or large ancestral effective population sizes (Hudson, 1990). This might explain the observed discrepancy between our gene-tree (e.g. DNA sequences sampled in *E. ivorense* and *E. suaveolens* coalesced *c.* 3.1 Ma) and species-tree (species diverging for *c.* 0.6 Myr according to *BEAST) estimates. More generally, divergence times between sister species previously estimated with a classic phylogenetic approach, i.e. without integrating the coalescence process into the model of speciation, may have been substantially overestimated in the case of tree species. This might cause us to reconsider the role of Pleistocene climate changes in promoting speciation (Moritz *et al.*, 2000; Plana, 2004).

A possible role of differentiation by ecological adaptation – the vanishing-refuge model

The oldest genetic discontinuity found within *E. suaveolens* separates a Central African gene pool inhabiting semi-deciduous forests from a West African gene pool inhabiting gallery forests in a relatively dry landscape. Thus, this genetic discontinuity corresponds to a deep ecological transition along a water-availability gradient, located between two main postulated glacial refugia (Mayr & O'Hara, 1986), one in Central Africa and one in West Africa. Hence, although our data support the influence of Quaternary glaciations through forest fragmentation, this does not exclude a potential synergetic role of ecological differentiation (niche divergence). The vanishing-refuge model of diversification (Damasceno *et al.*, 2014), integrates the impacts of habitat fragmentation and divergent selection under climate change in a way that may explain the origin of the West African *E. suaveolens*

© 2015 John Wiley & Sons Ltd

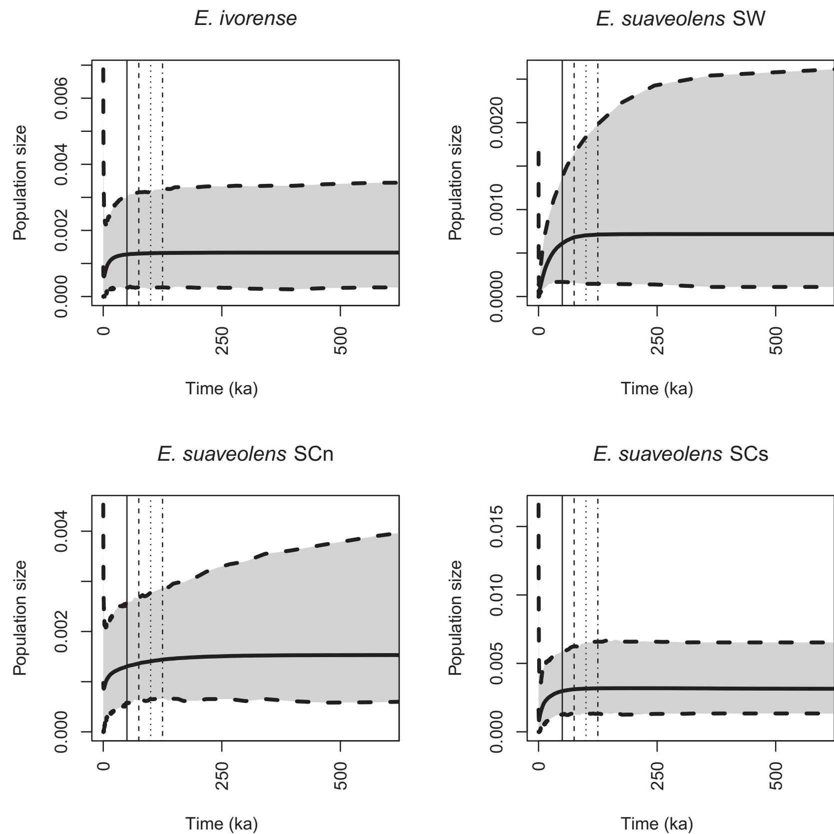

Figure 3 Historical demographic changes in each *Erythrophleum* gene pool according to Bayesian skyline analyses. A relative measure of effective population size ($N_e \times \mu$) is estimated according to the time before present. The grey zone (outlined by dashed lines) represents the 95% highest posterior density interval. From left to right, the vertical lines indicate 50, 75, 100 and 125 ka.

gene pool. If we assume that the ancestor of all three *E. suaveolens* gene pools was a forest species that was widespread during a previous interglacial period, then during the next glacial period, (1) the species might have retracted in Central Africa into two forest refuge regions, which led to the two Central African gene pools after generations of isolation, whereas (2) in West Africa, it might have adapted to the conditions reigning in gallery forests (riverine microrefugia), explaining the current ecological differentiation of the derived West African gene pool. These events might have occurred during different glacial periods, as suggested by our molecular dating. Additional data on the respective adaptations of the *E. suaveolens* gene pools would be needed to further test this scenario, but it already highlights that past climate changes might have influenced current phylogeographical patterns both through neutral differentiation driven by range fragmentation and through new adaptations to changing environmental conditions.

Forest refuge scenarios

Although our data support the forest refuge theory, the exact number and location of forest refugia remain speculative: the two *E. suaveolens* gene pools observed in LG suggest two main refugia, one located broadly in the north (in Cameroon, northern Republic of the Congo or Central African Republic) and one in the south (Gabon and southern Republic of the Congo). Given that similar north–south genetic differentiation has been observed in other tree species (Hardy *et al.*, 2013), there is substantial support for climate-driven rain forest fragmentation into two main sets of macrorefugia (relatively large areas where the climate remains favourable; refugia proposed for example by Maley, 1996), and/or two main groups of microrefugia (restricted favourable areas surrounded by inhospitable conditions, as along larger rivers; Leal, 2001). Nevertheless, additional data integrating samples from the Congolia (i.e. east of Lower Guinea) will be necessary to more accurately locate these refugia. In particular, it is possible that our current interpretation of a Northern Hemisphere refugium actually corresponds to a recent forest colonization front from refugia located in Congolia (Dauby *et al.*, 2014).

Climate-change effects may have differed between evergreen and semi-deciduous forests

The absence of intraspecific gene pools within *E. ivorense* using nDNA can be interpreted as a post-glacial colonization of the species from one only refugium. However, (1) we did not observe the spatial gradient of genetic diversity that would be expected under the refuge hypothesis, where refugia should harbour higher genetic diversity and endemism than recolonized areas (Hewitt, 1996), and (2) we previously reported genetic structure for pDNA (Duminil *et al.*, 2013) with one haplotype in LG and different haplotypes in UG, which seems incompatible with a single-refugium scenario. An alternative explanation is that *E. ivorense* was less severely affected by climate changes than *E. suaveolens*. Both species

© 2015 John Wiley & Sons Ltd

suffered a historical demographic decline, but the absence of clear genetic discontinuities within *E. ivorense* using nDNA (contrary to nSSR) suggests that the discontinuities within *E. suaveolens* are deeper than those in *E. ivorense*, suggesting that glacial refugia were more distant and/or separated for a longer period for *E. suaveolens* than for *E. ivorense*. *Erythrophleum ivorense* is mainly distributed in regions of higher rainfall, often in hilly landscapes, where most of the proposed Guinean macrorefugia were located (Maley, 1996; Sosef, 1996). Two of these refugia are already supported by molecular data: the Cameroonian volcanic line and the Ngovayang massif (Dauby *et al.*, 2014). Evergreen forests might thus have remained relatively well connected through gene flow, limiting population differentiation, whereas semi-deciduous forest would have persisted in fewer and/or more isolated refugia. Data from other characteristic evergreen and semi-deciduous forest species are needed to confirm this hypothesis.

A late Holocene opening of the Dahomey Gap

Dense humid tropical forests were probably recently replaced by more or less open, dry tropical forests in the Dahomey Gap. This pattern is supported by past vegetation reconstructions (Salzmann & Hoelzmann, 2005) and is confirmed by our molecular data. The split of the distribution of *E. ivorense* into two geographical areas on either side of the Dahomey Gap (LG and UG) might be recent, given the absence of genetic discontinuity using single-copy nuclear genes. Moreover, the presence of introgressed *E. ivorense* pDNA lineages within *E. suaveolens* in populations from the Dahomey Gap (Duminil *et al.*, 2013) supports the hypothesis that the species was present in this region in the past. This, in turn, supports the hypothesis of a recent occurrence of evergreen forest in the area as suggested by vegetation-reconstruction models (Hély *et al.*, 2009).

CONCLUSIONS

DNA sequences from six intergenic fragments add to the understanding of the evolutionary history of two widespread African tropical tree species, *E. ivorense* and *E. suaveolens*, in the Guinean region. Molecular dating of demographic changes supports the impact of late Pleistocene events on forest fragmentation. For the majority of tree species studied so far in the region, two main rain forest refuge areas have been identified in LG, roughly located on either side of the climatic hinge (*c.* 2° N). Nevertheless, the nature (many microrefugia or few macrorefugia) and location of these refugia remain highly speculative. The data suggest that trees from coastal evergreen forests may have been less affected by climate changes than trees from semi-deciduous forests further inland. Finally, our data support the hypothesis of a recent opening of the Dahomey Gap, as proposed by palynological studies. This work relied on the ability to amplify and sequence single-copy nuclear *PEPC* genes. To this end, we developed new primers for the *PEPC* gene, which were successfully amplified in a range of plant families, providing interesting new genetic resources for plant phylogeny and phylogeography.

ACKNOWLEDGEMENTS

Financial support was provided by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) (grants FRFC 2.4577.10, MIS F.4.519.10 and PDR T.0163.13) to O.J.H., by the Gembloux Agricultural University (FUSAGx, Belgium) via project PPR 10.000, and by the C3A and 'Îles Forestières Africaines' (IFORA) projects financed by the French ANR (Agence Nationale de la Recherche) under the ANR-BIODIV program. For their help with sampling, we also thank Guillaume Koffi, Eben-Ezer Ewédjè, Lambert Kouadio, Myriam Heuertz, Michel Arbonnier, Patrice Koumba Ipandi, Jean François Gillet, Nils Bourland, Kasso Dainou, Gilles Dauby, 'équipe Nature+ Pallisco', Gilbert Todou and Raymonde Mboma. We also thank the logging companies Pallisco, Wijma, Rougier, Precious Woods, ALPIcam–Grumcam, STBK, TRC, CUF, Rimbunan Ijau, Mont Pelé Bois, CBG and GWT for allowing access to their concessions, and ENEF (École Nationale des Eaux et Forêts), TEREA and Sylvafrica for facilitating sampling in Gabon. We thank Garry Cook for samples of *E. chlorostachys* from Australia, Chonglu Zhong and Zhigang Zhao for samples of *E. fordii* from China, and Jérôme Chave for samples of Caesalpinioideae species from French Guiana. We thank Arwen Bailey for her help in editing the manuscript as well as Cicely Marshall and several anonymous referees for helpful comments on earlier drafts of the manuscript.

REFERENCES

Akoègninou, A., van der Burg, W.J. & van der Maesen, L.J.G. (eds) (2006) *Flore analytique du Bénin*. Backhuys, Leiden.

Aubréville, A. (1950) *Flore forestière soudano-guinéenne. A.O.F. – Cameroun – A.O.F.* Société d'éditions géographiques, maritimes et coloniales, Paris.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. & Saunders, N.C. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Bonnefille, R. (2007) Rainforest responses to past climatic changes in tropical Africa. *Tropical rainforest responses to climate change* (ed. by M.B. Bush and J.R. Flenley), pp. 117–170. Springer, Berlin.

Bruneau, A., Mercure, M., Lewis, G.P. & Herendeen, P.S. (2008) Phylogenetic patterns and diversification in the caesalpinioid legumes. *Canadian Journal of Botany*, **86**, 697–718.

Budde, K., González-Martínez, S.C., Hardy, O.J. & 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.

© 2015 John Wiley & Sons Ltd

Daïnou, K., Mahy, G., Duminil, J., Dick, C.W., Doucet, J.-L., Donkpégan, A.S.L., Pluijgers, M., Sinsin, B., Lejeune, P. & Hardy, O.J. (2014) Speciation slowing down in widespread and long-living tree taxa: insights from the tropical timber tree genus *Milicia* (Moraceae). *Heredity*, **113**, 74–85.

Damasceno, R., Strangas, M.L., Carnaval, A.C., Rodrigues, M.T. & Moritz, C. (2014) Revisiting the vanishing refuge model of diversification. *Frontiers in Genetics*, **5**, 353.

Dauby, G., Duminil, J., Heuertz, M., Koffi, K.G., Stévart, T. & Hardy, O.J. (2014) Congruent phylogeographical patterns of eight tree species in Atlantic Central Africa provide insights into the past dynamics of forest cover. *Molecular Ecology*, **23**, 2299–2312.

Debout, G.D.G., Doucet, J.-L. & Hardy, O.J. (2010) Population history and gene dispersal inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae). *Heredity*, **106**, 88–99.

Dellicour, S. & Mardulyn, P. (2013) SPADS 1.0: a toolbox to perform spatial analyses on DNA sequence datasets. *Molecular Ecology Resources*, **14**, 647–651.

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

Duminil, J., Heuertz, M., Doucet, J.-L., Bourland, N., Cruaud, C., Gavory, F., Doumenge, C., Navascués, M. & Hardy, O.J. (2010) CpDNA-based species identification and phylogeography: application to African tropical tree species. *Molecular Ecology*, **19**, 5469–5483.

Duminil, J., Brown, R.P., Ewédjè, E.-E.B.K., Mardulyn, P., Doucet, J.-L. & Hardy, O.J. (2013) Large-scale pattern of genetic differentiation within African rainforest trees: insights on the roles of ecological gradients and past climate changes on the evolution of *Erythrophleum* spp (Fabaceae). *BMC Evolutionary Biology*, **13**, 195.

Dupont, L.M., Donner, B., Schneider, R. & Wefer, G. (2001) Mid-Pleistocene environmental change in tropical Africa began as early as 1.05 Ma. *Geology*, **29**, 195–198.

El Mousadik, A. & Petit, R.J. (1996) Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology*, **5**, 547–555.

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.

Falush, D., Stephens, M. & Pritchard, J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Gonmadje, C.F., Doumenge, C., Sunderland, T.C.H., Balinga, M.P.B. & Sonké, B. (2012) Analyse phytogéographique des forêts d'Afrique Centrale: le cas du massif de Ngovayang (Cameroun). *Plant Ecology and Evolution*, **145**, 152–164.

Hardy, O.J., Born, C., Budde, K., Daïnou, K., Dauby, G., Duminil, J., Ewédjè, E.-E.B.K., Gomez, C., Heuertz, M., Koffi, G.K., Lowe, A.J., Micheneau, C., Ndiade-Bourobou, D., Piñeiro, R. & Poncet, V. (2013) Comparative phylogeography of African rain forest trees: a review of genetic signatures of vegetation history in the Guineo-Congolian region. *Comptes Rendus Geoscience*, **345**, 284–296.

Hare, M.P. (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution*, **16**, 700–706.

Heled, J. & Drummond, A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.

Hély, C., Braconnot, P., Watrin, J. & Zheng, W. (2009) Climate and vegetation: simulating the African humid period. *Comptes Rendus Geoscience*, **341**, 671–688.

Heuertz, M., Duminil, J., Dauby, G., Savolainen, V. & Hardy, O.J. (2014) Comparative phylogeography in rainforest trees from Lower Guinea, Africa. *PLoS ONE*, **9**, e84307.

Hewitt, G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hudson, R.R. (1990) Gene genealogies and the coalescent process. *Oxford surveys in evolutionary biology* (ed. by D.G. Futuyma and J. Antonovics), pp. 1–44. Oxford University Press, Oxford.

Kouadio, L. (2009) *Mesures sylvicoles en vue d'améliorer la gestion des populations d'essences forestières commerciales de l'Est du Cameroun*. PhD Dissertation, Faculté Universitaire des Sciences Agronomiques de Gembloux, Gembloux, Belgium.

Leal, M.E. (2001) Microrefugia, small scale ice forest remnants. *Plant systematics and phytogeography for the understanding of African biodiversity* (ed. by E. Robbrecht, J. Degreef and I. Friis), pp. 1073–1077. National Botanic Garden, Meise.

Ley, A.C. & Hardy, O.J. (2014) Contrasting patterns of gene flow between sister plant species in the understorey of African moist forests – the case of sympatric and parapatric Marantaceae species. *Molecular Phylogenetics and Evolution*, **77**, 264–274.

Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H. & Varotto, C. (2008) Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics*, **24**, 727–745.

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

Lönnberg, E. (1929) The development and distribution of the African fauna in connection with and depending upon climatic changes. *Arkiv för Zoology*, **21**, 1–33.

Maley, J. (1996) Le cadre paléoenvironnemental des refuges forestiers africains: quelques données et hypothèses [The palaeoenvironmental structure of African forest reserves: some data and hypotheses]. *The biodiversity of African plants: Proceedings of the XIVth AETFAR Congress 22–27 August 1994, Wageningen, The Netherlands* (ed. by der Maesen L.J.G., der Burgt X.M. and van Medenbach de Rooy J.M.), pp. 519–535. Kluwer, Dordrecht (in French).

© 2015 John Wiley & Sons Ltd

Mardulyn, P., Mikhailov, Y.E. & Pasteels, J.M. (2009) Testing phylogeographic hypotheses in a Euro-Siberian cold-adapted leaf beetle with coalescent simulations. *Evolution*, **63**, 2717–2729.

Martin, D. & Rybicki, E. (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics*, **16**, 562–563.

Mayr, E. & O'Hara, R.J. (1986) The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution*, **40**, 55–67.

Moritz, C., Patton, J.L., Schneider, C.J. & Smith, T.B. (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533–563.

Nei, M. & Li, W.H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA*, **76**, 5269–5273.

Nylander, J.A.A. (2004) *MrModeltest 2.3*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala.

Pamilo, P. & Nei, M. (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568–583.

Plana, V. (2004) Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1585–1594.

Plummer, M., Best, N., Cowles, M.K. & Vines, K. (2008) *coda: output analysis and diagnostics for Markov chain Monte Carlo simulations*. R package version 0.13-3. Available at: http://cran.r-project.org/package=coda.

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.

Quinzin, M.C. & Mardulyn, P. (2014) Multi-locus DNA sequence variation in a complex of four leaf beetle species with parapatric distributions: mitochondrial and nuclear introgressions reveal recent hybridization. *Molecular Phylogenetics and Evolution*, **78**, 14–24.

Rambaut, A., Suchard, M.A., Xie, D. & Drummond, A.J. (2014) *Tracer v1.6*. University of Edinburgh, Edinburgh, UK. Available from http://beast.bio.ed.ac.uk/Tracer.

Robbrecht, E. (1996) Geography of African Rubiaceae with reference to glacial rain forest refuges. *The biodiversity of African plants: Proceedings of the XIVth AETFAR Congress 22–27 August 1994, Wageningen, The Netherlands* (ed. by L.J.G. Van Der Maesen, X.M. Van Der Burgt and J.M. Van Medenbach De Rooy), pp. 564–581. Kluwer, Dordrecht.

Rychlik, W. & Rhoads, R.E. (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA. *Nucleic Acids Research*, **17**, 8543–8551.

Salzmann, U. & Hoelzmann, P. (2005) The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *The Holocene*, **15**, 190–199.

Sosef, M.S.M. (1994) *Studies in Begoniaceae V. Refuge begonias. Taxonomy, phylogeny and historical biogeography of* Begonia *sect.* Loasibegonia *and sect.* Scutobegonia *in relation to glacial rain forest refuges in Africa*. Wageningen Agricultural University Papers, 94. Wageningen Agricultural University, Wageningen.

Sosef, M.S.M. (1996) Begonias and African rain forest refuges: general aspects and recent progress. *The biodiversity of African plants: Proceedings of the XIVth AETFAR Congress 22–27 August 1994, Wageningen, The Netherlands* (ed. by L.J.G. van der Maesen, X.M. van der Burgt And J.M. van Medenbach de Rooy), pp. 602–611. Kluwer, Dordrecht.

Stephens, M. & Donnelly, P. (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

White, F. (1983) *The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa*. Natural Resources Research no. 20. UNESCO, Paris, France.

SUPPORTING INFORMATION

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

Appendix S1 *Erythrophleum suaveolens* and *E. ivorense* sampling locations.
Appendix S2 Designation of single-copy nuclear genes and primers used.
Appendix S3 Optimal number of genetic pools chosen according to the results given by the delta-K method.

DATA ACCESSIBILITY

Sequences have been submitted to GenBank and accession numbers are indicated in Appendix S1.

BIOSKETCH

Jérôme Duminil is primarily interested in studying the evolution, reproductive biology and adaptation of tree species from tropical African rain forests.

Author contributions: J.D. and O.J.H. conceived the ideas; J.D., C.D., F.W. and J.-L.D. collected the data; J.D., S.M. and P.M. analysed the data; and J.D. and O.J.H. led the writing.

Editor: Peter Linder

© 2015 John Wiley & Sons Ltd