

ORIGINAL ARTICLE

Population history and gene dispersal inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae)

GDG Debout¹, J-L Doucet² and OJ Hardy¹¹*Evolutionary Biology and Ecology, Faculté des Sciences, Université Libre de Bruxelles, Brussels, Belgium* and ²*Laboratory of tropical and subtropical forestry, Unit of forest and nature management, University of Liege, Gembloux, Belgium*

African rainforests have undergone major distribution range shifts during the Quaternary, but few studies have investigated their impact on the genetic diversity of plant species and we lack knowledge on the extent of gene flow to predict how plant species can cope with such environmental changes. Analysis of the spatial genetic structure (SGS) of a species is an effective method to determine major directions of the demographic history of its populations and to estimate the extent of gene dispersal. This study characterises the SGS of an African tropical timber tree species, *Distemonanthus benthamianus*, at various spatial scales in Cameroon and Gabon. Displaying a large continuous distribution in the Lower Guinea domain, this is a model species to detect signs of past population fragmentation and recolonization, and to estimate the extent

of gene dispersal. Ten microsatellite loci were used to genotype 295 adult trees sampled from eight populations. Three clearly differentiated gene pools were resolved at this regional scale and could be linked to the biogeographical history of the region, rather than to physical barriers to gene flow. A comparison with the distribution of gene pools observed for two other tree species living in the same region invalidates the basic assumption that all species share the same Quaternary refuges and recolonization pathways. In four populations, significant and similar patterns of SGS were detected. Indirect estimates of gene dispersal distances (σ) obtained for three populations ranged from 400 to 1200 m, whereas neighbourhood size estimates ranged from 50 to 110.

Heredity (2011) **106**, 88–99; doi:10.1038/hdy.2010.35; published online 14 April 2010

Keywords: spatial genetic structure; biogeography; dispersal; tropical tree; rainforest; Central Africa

Introduction

Geographical patterns of population genetic variation within plant species are shaped by many factors. Some are related to traits intrinsic to the species' biology (reproduction mechanisms, modes of seed and pollen dispersal, gregariousness), whereas others are linked to disturbances, mainly habitat fragmentation, induced by natural processes (for example, fire, climate change) or human impact (conversion of forested areas into agricultural land, logging). Quaternary climatic fluctuations caused important changes in the distribution ranges of many plant species and thereby strongly affected their large-scale genetic structure (Comes and Kadereit, 1998; Hewitt, 2000). The signature of the Quaternary ice ages in present-day patterns of genetic structure may be especially important for species with long life cycles such as trees, because probably fewer than 100–1000 generations

(depending on the species) have elapsed since the last glacial maximum 15 000 radiocarbon (¹⁴C) years ago (Walker, 1995). The most recent glacial maxima occurred from 160 000 to 130 000 and from 24 000 to 12 000 years before present (Maley, 1996). During these episodes, the climate in tropical Africa was drier and cooler, causing an expansion of savannah and a contraction of lowland tropical rainforests into regions still presenting sufficient humidity, as demonstrated by pollen records in lake sediments (Maley, 1996). Hence, fragmentation-expansion cycles of the lowland rain forest habitats likely occurred during the Pleistocene, causing population fragmentation of their constituent species (Flenley, 1998). More recently (around 2500 cal years BP), the rainforest in Africa has also been disturbed by a southwards shift of the intertropical convergence zone, as attested by a high-resolution pollen profile obtained from South West Cameroon (Ngomanda *et al.*, 2009).

Although the impact of the last ice age on the past distribution and the genetic structures of trees from Europe and North America is well documented, there is very little information for tropical tree species, particularly in African rainforests. The latter are mainly concentrated in the Guineo-Congolian phytogeographical region, which is composed of three domains (White,

Correspondence: Dr GDG Debout, Evolutionary Biology and Ecology, CP 160/12, Faculté des Sciences, Université Libre de Bruxelles, 50, av F Roosevelt, Brussels 1050, Belgium.

E-mails: debout_gabriel@yahoo.fr or gdebout@ulb.ac.be

Received 15 June 2009; revised 16 November 2009; accepted 18 February 2010; published online 14 April 2010

1979): the Upper-Guinean domain (comprising forests of West Africa), the Lower-Guinean domain (the Atlantic Central Africa, from Nigeria to Congo) and the Congolian domain (comprising the whole Congo basin until the Albertine Rift in the East of the Democratic Congo Republic). The Lower-Guinean domain harbours the richest flora, and several centres of high species diversity and/or endemism for many genera have been identified (for example, in West-Cameroon and in West and Central Gabon; see White, 1979; Sonké *et al.*, 2007). The richness of the Lower-Guinean domain has often been interpreted as the signature of past forest refuges, and the current distribution of plant species has been used to attempt to identify these refuges (Maley, 1996). However, given the paucity of direct evidences, such as pollen records, the intensity of forest fragmentation, the determination of historic localization of past forest fragments, and the degree of isolation between large fragments (gallery forests could constitute corridors or even refuges for various forest-dwelling species) are essentially matters of speculation. Moreover, centres of diversity or endemism are concordant with areas of complex topography and/or with high rainfall so that the effects of history and of environmental gradients on species diversity are generally confounded. As the last climate changes causing forest fragmentation occurred relatively recently, it might be more pertinent to investigate their effects on within-species patterns of diversity rather than on species endemism and richness.

Different forms of genetic signatures of past demographic changes may be exploited. The identification of differentiated and geographically coherent gene pools within species that nowadays show wide ranging and essentially continuous distributions could indicate past population fragmentation, potentially due to past changes in vegetation cover. Recent population expansion or decline may also leave a signature in the distribution of allele frequencies within locus due to mutation-drift disequilibrium (Cornuet and Luikart, 1996). Genetic markers keeping an information on the successive mutation events, such as DNA sequences or microsatellites mutating in a stepwise manner, may also be used to test the occurrence of a phylogeographic signal; for example, (Pons and Petit, 1996; Hardy *et al.*, 2003), giving clues about the relative time depth of observed genetic divergences between populations. Depending on the extent of gene dispersal, the genetic signature of past population fragmentation will progressively vanish once differentiated gene pools begin to exchange genes. Hence, one avenue to improve our understanding of past tropical forest cover changes is to analyse large-scale geographic patterns of genetic variation of forest-dwelling species as well as their gene dispersal distances. Evidence for converging patterns among species would give support for a strong determinism of past vegetation changes. However, nowadays, such data are very rare for central African species (Born *et al.*, 2008).

Gene dispersal in plants is possible only through the dispersal of pollen and seeds. Fairly reliable indirect estimates of gene dispersal distances can be obtained from the spatial genetic structure (SGS) (Fenster *et al.*, 2003; Hardy *et al.*, 2006). Such estimates can give insights into the number of generations during which a genetic signature of past fragmentation should remain. In addition, dispersal capacities may be constrained by

peculiar landscape features and demographic parameters. Comparing SGS among sites may provide insights into the range of variation of gene flow within species. For instance, forest fragmentation can increase SGS by breaking up gene flow and increasing local drift because of the reduction of the number of cohorts of many species (Dick *et al.*, 2008). In contrast, long-distance seed dispersal and landscape changes that increase dispersal capacities of species may reduce the stochastic action of drift and lead to a decrease in SGS.

Using the timber tree *Distemonanthus benthamianus* as a model species displaying, nowadays, a large continuous distribution in the Lower Guinea domain, the aim of this study is to characterize its genetic structure at various spatial scales in Cameroon and Gabon in order to detect possible signs of past population fragmentation and recolonization, and to estimate the extent of gene dispersal. More specifically, using microsatellite markers that were developed for this study, we address the following questions: (1) Are there large-scale geographic discontinuities in the pattern of genetic variation, indicating the occurrence of distinct gene pools? (2) Is there evidence for a contribution of stepwise mutations in the differentiation of gene pools, indicating ancient isolation? (3) Is there any genetic evidence of recent demographic changes? (4) What is the outcrossing rate and the extent of gene dispersal distances according to isolation by distance models? Our results are discussed in the light of possible causal processes and compared with those of a couple of other tree species from Central Africa to assess whether some general patterns emerge.

Understanding the genetic dynamics of tree populations and the origin of geographical patterns of genetic variation can provide key information about dispersal and regeneration for improving population management by logging companies (Lowe *et al.*, 2005). This is all the more important for a heavily exploited timber tree such as *D. benthamianus* that suffers very limited *in situ* regeneration in mature forest. Although this species did not receive any IUCN classification yet, it was ranked 7th by the FAO among the 10 species considered potentially most susceptible to loss of genetic variation due to logging in southern Cameroon (Namkoong *et al.*, 2002).

Materials and methods

Biological model and sampling

D. benthamianus Baillon (Caesalpinioideae) is a widely but sparsely distributed tree throughout the high forests of West and Central Africa, from Ivory Coast to Gabon and Congo. It is common in Cameroon and Gabon, where it is known as African satinwood, or more generally as Movingui. One of the big trees (40 m high and >1.20 m trunk diameter) of evergreen, semi-deciduous and secondary forests, *D. benthamianus* is traditionally used in African medicine to treat bacterial, fungal and viral infections (Nguelefack *et al.*, 2005). Its modern uses include cabinetwork, joinery, flooring and decorative veneers, and it is generally suggested as an oak alternative. Apart from its geographic range and data on its population density in a few forest concessions, little is known about its biology. Flowers are hermaphroditic; the fruit is a flat, thin-walled, indehiscent pod measuring 10 cm by 3.5 cm (Aubréville, 1968). Seeds are dispersed

by wind (Hecketsweiler, 1992) to distances of up to 50 m (Cordiez, 2000). At least three monkey species (*Lophocebus albigena*, *Cercopithecus nictitans*, *Cercopithecus pogonias*; Poulsen *et al.*, (2002)) prey on its seeds (Gautier-Hion *et al.*, 1993). Feeding only on immature fruits, they are quite unlikely to contribute to dispersal. Fruiting begins at a diameter at breast height (d.b.h.) of 20 cm, but only trees with a d.b.h. of 40 cm or larger fruit regularly. Trees grow on average between 3 and 4 mm in diameter annually. *D. benthamianus* is a light-demanding species and its natural regeneration in mature forest is very limited. Very few saplings are observed in closed-canopy forests (Doucet, 2003).

We collected samples (leaves or bits of cambium) from 295 adult *D. benthamianus* in eight different sites in Central Africa (three in Cameroon, five in Gabon), hereafter referred to as 'populations'. These are not in fact locally circumscribed groups of individuals isolated from each others, as the species is distributed rather continuously through the whole sampling area (Figure 1a). Samples were collected and immediately dried in silica gel. For spatial analyses, we only used 286 trees with precise GPS localizations obtained with a Garmin GPSmap 60CSx (Garmin, KS, USA) with a 10 m estimated error.

Development of microsatellite markers

Microsatellites were isolated using a fast and cheap protocol, modified from Hamilton *et al.* (1999) by TC Glenn (Savannah River Ecology Laboratory, University of Georgia, GA, USA). The detailed protocol of microsatellite isolation is available at request from us. Briefly, total genomic DNA was extracted from cambium (providing cleaner DNA than that from leaves) of two adult trees using our own CTAB protocol, modified from Doyle and Doyle (1987). A partial enriched genomic library was then constructed. A total of 192 white transformant clones were hybridized with (AC)₁₂. Around 50% of these clones gave a positive signal and the inserted DNA fragment was sequenced in all of these positive clones. Twenty-seven primer pairs, designed using the software OLIGO 3.3 (Rychlik and Rhoads, 1989), were tested and 17 gave satisfactory amplification patterns (that is, a PCR product of the predicted size, and no supernumerary bands). Protocol for the PCR amplifications is also available on request. We sized PCR products using an automatic ABI PRISM 3100, the 500 LIZ GeneScan size standard (Applied Biosystems, Foster City, CA, USA), and GENEMAPPER v3.0 software (Applied Biosystems). For genotyping we only used 10 out of the 17 microsatellite loci that were isolated from *D. benthamianus*, chosen so as to easily genotype individuals by multiplexing loci in one genotyping run. These 10 loci are named dis028, dis054, dis069, dis103, dis116, dis125, dis126, dis127, dis138 and dis140, and were given GenBank accession numbers FJ654552, FJ654554, FJ654555 and FJ654557–FJ654563 (Supplementary Table S1).

Genetic data analyses

We first controlled for the absence of heterozygote deficiency or linkage disequilibrium by genotyping 60 individuals, sampled from two populations (population 1 from Cameroon and population 5 from Gabon; Supplementary Table S1).

To infer the impacts of limited gene dispersal, population history and habitat fragmentation on isolation by distance, we followed a stepwise approach consisting of a Bayesian clustering method to detect differentiated gene pools followed by the analysis of kinship-distance curves.

Genetic diversity and testing demographic changes: We used either GENEPOP 3.4 (Raymond and Rousset, 1995) or FSTAT 2.9.3.2 (Goudet, 1995) to compute standard estimators of genetic polymorphism. We computed V , the variance of allele size, and calculated expected (H_E) and observed heterozygosity (H_O), averaged over loci. We tested for heterozygote deficiency using the Markov chain method. F_{IS} values were calculated for each population and their significance tested by permuting the distribution of alleles within populations. It provides information on the cumulative action of inbreeding, population substructure and potential null alleles. We used the rarefaction method (in FSTAT) to compute allelic richness (rA_{OL} , rarefied mean observed number of alleles per locus) because it gives more weight to rare alleles and thus better takes into account mutation accumulation that may occur in older populations.

To detect past demographic changes, the 'bottleneck' statistic T_2 (Cornuet and Luikart, 1996) was computed at the population level. It represents an average across loci of the deviation of the actual gene diversity H_E from the gene diversity expected from the number of alleles in the population H_A assuming mutation-drift equilibrium in a population of constant size. Positive values of T_2 reflect a gene diversity excess possibly caused by recent founder events, whereas negative values are consistent with recent population expansion without immigration. T_2 was computed with BOTTLENECK 1.2.02 (Piry *et al.*, 1999), according to the Stepwise Mutation Model (T_2 SMM). We used T_2 under the SMM model because alleles at our loci were separated by multiples of 2 bp with the exception of a single allele at locus dis054, suggesting that stepwise mutation is the norm. T_2 was also computed under the Infinite Allele Model to assess the robustness of the results with respect to the mutation model. Significant deviation from equilibrium gene diversity was determined using the Wilcoxon signed-rank test, which is the most appropriate test when only few polymorphic loci are analysed (Piry *et al.*, 1999).

Identification of spatial genetic discontinuities: Historical events may have produced spatial discontinuities in the genetic variation of *D. benthamianus* that could be detected by Bayesian cluster analyses. We used the model-based clustering method implemented in STRUCTURE 2.2.3 (Pritchard *et al.*, 2000). This method assigns individual multilocus genotypes (data consisting of unlinked markers) probabilistically to a user-defined number K of clusters (or gene pools), in order to approach linkage equilibrium within clusters. We ran STRUCTURE 10 times for each $K = 1-10$, for 1 00 000 iterations after a burn-in period of 40 000 on the data set with precise localizations ($n = 286$ individuals) with no earlier information on the population of origin of each sampled individual. We used the admixture model, in which the fraction of ancestry from each cluster is estimated for each individual. The parameter of individual admixture α was chosen to be the same for all clusters and it was given a uniform prior. The allele

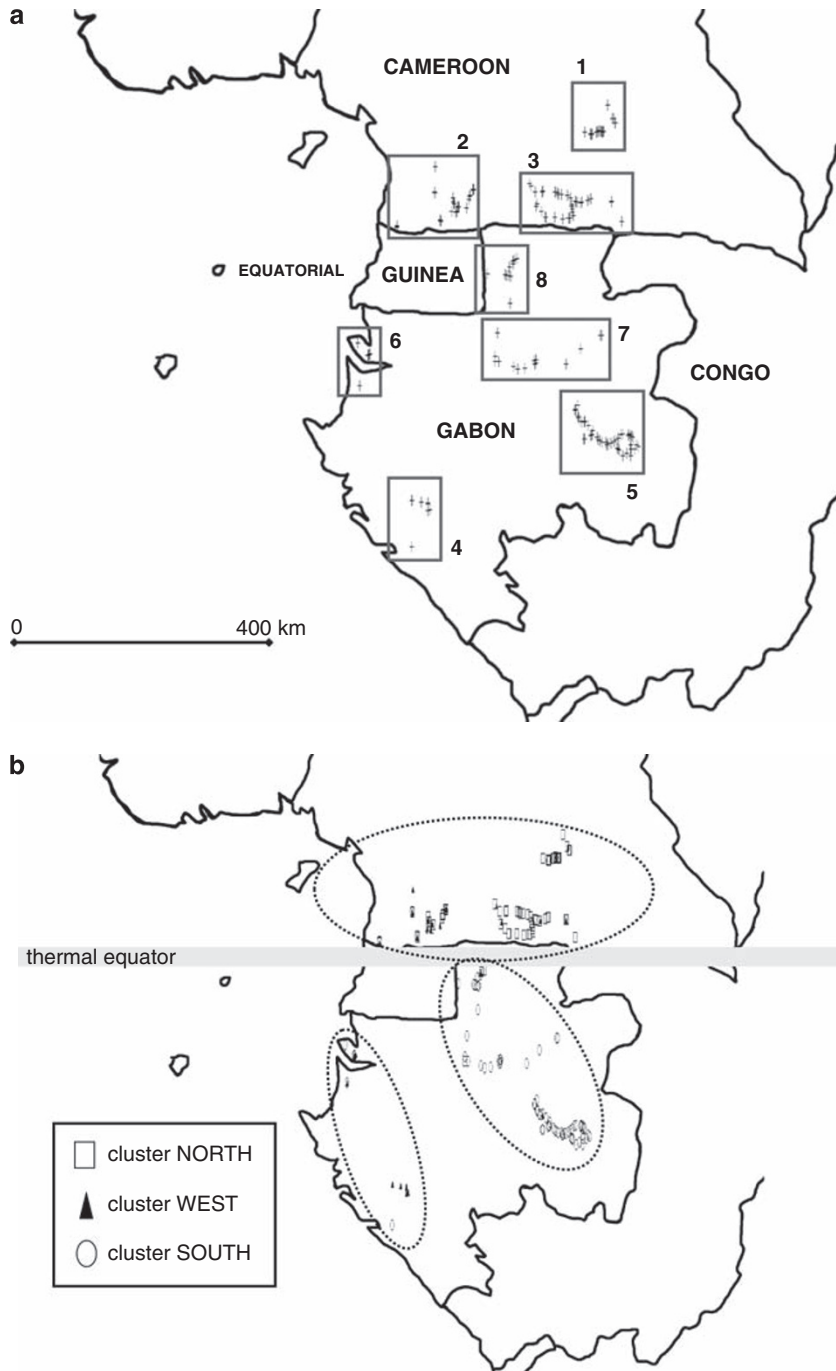


Figure 1 (a) The eight sampled populations of *D. benthamianus* in Central Africa. In Cameroon: 1-AbongMbang, 2-Ebolowa, 3-Sangmelima. In Gabon: 4-CBG, 5-CEB, 6-Libreville, 7-Mitzic, 8-Oyem. (b) Synthetic map of the genetic clustering of *D. benthamianus* in Central Africa, based on STRUCTURE results. Empty squares: individuals belonging to the cluster 'North'. Black triangles: cluster 'West'. Empty circles: cluster 'South'. The discontinuous lines delimit the three clusters. The thermal equator is figured, as it may be a current obstacle to gene flow by pollen due to the inversion of seasons.

frequencies were kept independent among clusters in order to avoid overestimating the number of clusters (Falush *et al.*, 2003).

To take into account the spatial coordinates, we used the spatial Bayesian clustering algorithm implemented in TESS 2.0 (Chen *et al.*, 2007), which assigns individuals to geographical clusters and estimates their probabilities of membership in clusters. As suggested by the authors of the programme, to include more realistic features of

spatial structure, we removed the very long edges in the neighbourhood network (those > 150 km). In our case, including these very long edges did not change the results, but greatly increased the computation time. We used the no-admixture model with an interaction parameter ψ ranging from 0.6 to 1. The interaction parameter is requested to implement TESS before distribution on cluster configuration and affects the relative weight given to spatial positions and genotypes when

assigning an individual to a cluster (authors suggest using ψ in the range (0.5, 1.0)). After preliminary computations up to $K=10$, the maximum number of clusters was fixed to $K=6$. We did 500 runs per ψ ($\psi=0.6, 0.8$ and 1), with a burn-in number of sweeps of 10 000 and 50 000 iterations.

Differences in allelic richness, observed and expected heterozygosities, and F_{IS} between clusters were tested using FSTAT with 5000 permutations.

Characterisation of SGS: Differentiation between populations was analysed based on allele identity with the statistic F_{ST} , or on allele size with R_{ST} . To characterize polymorphism and among-population differentiation for the species, multilocus estimates of these statistics were obtained. The contribution of stepwise-like mutations to population differentiation, that is, whether $R_{ST} > F_{ST}$, was investigated by testing whether the observed R_{ST} was larger than its value obtained after permuting allele sizes among allelic states (Hardy *et al.*, 2003).

We assessed SGS following the procedure described by Vekemans and Hardy (2004). We computed pairwise kinship coefficients between individuals using the software SPAGED1 v1.2g (Hardy and Vekemans, 2002) and J Nason's estimator of kinship coefficient (Loiselle *et al.*, 1995). Kinship coefficient values (F_{ij}) were regressed on the natural logarithm of the spatial distance between individuals, $\ln(d_{ij})$, providing the regression slopes b_{Ld} . Standard errors were assessed by jackknifing data over loci. To visualize SGS, kinship coefficient values were averaged over a set of distance intervals (d), giving $F(d)$, and plotted against geographical distance. Seven classes of pairwise distance intervals were chosen following a near-geometric distribution: the six boundaries were 1, 4, 8, 16, 32 and 64 km (a boundary at 2 km was removed because it caused an interval encompassing too few pairs of individuals). To test for SGS, spatial positions of individuals were permuted 9999 times in order to obtain the frequency distribution of b_{Ld} under the null hypothesis that F_{ij} and $\ln(d_{ij})$ were uncorrelated.

In four populations (1, 2, 3 and 5) in which at least 40 individuals were sampled, we examined intra-population genetic structure, estimating the statistic Sp (Vekemans and Hardy, 2004; Hardy *et al.*, 2006). To compare the extent of SGS among these populations over the same spatial scale, we calculated b_{Ld60} , the regression slopes of F_{ij} on $\ln(d_{ij})$ for $d_{ij} < 60$ km (the maximum distance that gave a good representation of pairs of individuals for each population, that is, at least 63% of pairwise comparisons were considered in each population). We then calculated Sp , defined as $Sp = -b_{Ld60}/(1-F_{(1)})$, where $F_{(1)}$ is the mean F_{ij} between neighbouring individuals, which was approximated by $F(d)$ for the first distance interval ($d_{ij} < 1$ km). Sp was also computed for $d_{ij} < 10$ km to allow comparisons with other tropical tree species at comparable spatial scales (Dick *et al.*, 2008), but a serious drop of precision was consequently achieved, as on average only 18% of all pairwise comparisons were considered.

Estimation of gene dispersal distances and outcrossing rate: Wright's neighbourhood size, $Nb \approx 4\pi D_e \sigma_g^2$, where D_e is the effective density and σ_g^2 is half the mean squared parent-offspring distance, can be estimated as $Nb = -(1-F_{(1)})/b_r$ when the regression slope, b_r , is

computed within the restricted distance range $\sigma_g > dij > 20\sigma_g$ (Hardy *et al.*, 2006). As σ_g is unknown, an iterative approach can be applied to estimate jointly Nb and σ_g knowing D_e , as implemented in the SPAGED1 software. The quality of the joint estimate is obviously highly sensitive to the quality of the value assumed for D_e (Hardy *et al.*, 2006). We estimated Nb and σ_g for populations 2, 3 and 5. For population 5 (Gabon), precise species-specific tree density (D) computed from forest inventories was obtained from the CEB, a logging company of the Precious Woods group. For both Cameroonian populations (2 and 3), we estimated D from a still-unpublished atlas map of densities for Cameroonian tree species (C Doumenge, personal communication) and from forest inventories from concessions in the vicinities of our sampling locations. In all cases, D was computed for trees with a minimum d.b.h. ≥ 40 cm, because in Cameroon these were the sole data available. Fortunately, this appears to be a biologically relevant threshold size, because *D. benthamianus* trees only flower regularly when their d.b.h. is more than 40 cm (Doucet, 2003). Moreover, at the study site for population 5, taking into account only trees with a d.b.h. more than 40 cm instead of over 30 cm caused a reduction of only 15% in estimated density (Supplementary Table S2). The slight underestimation of mature tree density will lead to a slight overestimation of σ_g^2 , and that will be minored when computing the square-root to obtain σ_g . D_e can be approximated as $D \times N_e/N$, where N_e/N is the ratio of the effective population size to census population size (Nunney, 1993; Vekemans and Hardy, 2004). Demographic studies show that this ratio typically ranges from 0.1 to 0.5 in populations of adult plants (Frankham, 1995). We thus used $D/2$, $D/4$ and $D/10$ as alternative estimates for D_e . For a fixed D_e , lower and upper bounds for the 95% confidence interval (CI) of Nb were computed as $(F_{(1)}-1)/(b_r + 2SE_b)$ and $(F_{(1)}-1)/(b_r - 2SE_b)$, respectively, SE_b being the standard error of the b_r estimates obtained by jackknifing data over each locus. The 95% CI of σ_g was obtained using an analogous procedure.

Intrapopulation multilocus selfing estimates were computed with the software RMES (David *et al.*, 2007). These estimations of selfing rates are based on the distribution of multilocus heterozygosity and, unlike those derived from F_{IS} , have been shown to be insensitive to technical artifacts such as null alleles (David *et al.*, 2007).

Results

Genetic diversity

In the two populations (1 and 5) used to control the quality of our markers, we observed no significant deviation from Hardy-Weinberg equilibrium nor from linkage equilibrium, using exact tests (GENEPOP 3.4 package; Raymond and Rousset, 1995) after correction for multiple tests with a sequential Bonferroni procedure (Supplementary Table S1).

All loci were polymorphic over the eight sampled populations, with the number of alleles ranging from 6 to 21, except for dis125, which exhibited only two alleles. We thus used a total of 115 different alleles. High levels of variation were found among loci, and observed

heterozygosity ranged from 0.00 to 0.93 per locus and per population. All loci were at Hardy–Weinberg and linkage equilibrium over the eight sampled populations.

Although the sampling areas of the eight populations varied greatly (from ~1000 to 5000 km²), no relation was found between the extent of the area sampled and the following genetic descriptors (Table 1): rA_{OL} (Spearman's $\rho = 0.262$; $P = 0.53$), V (Spearman's $\rho = 0.119$; $P = 0.78$), H_O (Spearman's $\rho = 0.405$; $P = 0.32$), H_E (Spearman's $\rho = -0.02$; $P = 0.96$) and F_{IS} (Spearman's $\rho = -0.786$; $P = 0.02$). We thus treated populations as equivalent units for further analysis. There was slight but significant genetic differentiation among these populations: overall value of F_{ST} was 0.0739.

Under a strict SMM model, all populations showed a negative value of T_2 (consistent with recent population expansion without immigration) and a significant heterozygosity deficit ($P < 0.01$). Under an Infinite Allele Model, no populations showed significant heterozygosity excess or deficit ($P > 0.05$), and T_2 was negative only in populations 2, 5 and 7.

Macrogeographic genetic discontinuities

STRUCTURE's posterior probabilities of K , log-likelihood of the microsatellite data on 286 *D. benthamianus* individuals given K clusters, and the corresponding ΔK statistic (Evanno *et al.*, 2005), all showed a peak at $K = 3$, indicating that three genetic clusters was the best solution for K given the data. Only 17 individuals out of the 286 (that is, 6%) could not be assigned to one of the three clusters ($P_{\text{assignment}} < 0.5$ to each cluster). These three clusters contained 20, 41 and 39% of all assigned

individuals (Table 2). These clusters corresponded quite closely to three geographical groups. The first (the least supported) contained 59% of its individuals from the Gabonese coast (populations 4 and 6), the others mostly coming from the western part of our Cameroonian sample (population 2); the second cluster contained 88% of its individuals coming from the inland of Gabon (populations 5, 7 and 8); and the third cluster contained 97% of its individuals coming from Cameroon (populations 1, 2 and 3), the few others coming from northern Gabon (population 8).

We then used the spatial Bayesian clustering algorithm TESS. For three different values of the interaction parameter (1, 0.8 and 0.6), we obtained a likely clustering of our individuals into three to five different clusters. When ψ equalled 1 (influence of spatial data greater than that of genetic data), we obtained three clusters (smallest DIC estimates) that were the Gabonese coast, inland Gabon and the Cameroonian populations. When the value of ψ was decreased to 0.8, a fourth cluster was obtained through the division of Cameroonian samples in western (population 2) and eastern (populations 1 and 3) parts. When $\psi = 0.6$, five clusters was the most likely solution for our data, but the fifth cluster was composed of only a single Gabonese coastal individual separated from the others.

From the STRUCTURE and TESS analyses, we thus concluded that the most likely clustering of our data was a division into three clusters: one called 'North', containing all three Cameroonian populations (1–3); 'West', containing the two coastal Gabonese populations (4 and 6) and 'South' that contained the three populations from inland Gabon (5, 7 and 8). There were no

Table 1 Description of the samples and parameters of genetic variation for the eight studied populations

Sampled population	N	Estimated area (km ²)	d_{ij} max (km)	ad_{ij} (km)	nd_{ij} (km)	rA_{OL} (SE)	V (SE)	H_E	F_{IS}	s (CI 95)
<i>Cameroon</i>										
1 AbongMbang	40	1102	60	18.1	0.8	5.26 (0.77)	20.67 (5.11)	0.602	0.094	0.018 (0, 0.133)
2 Ebolowa	51	2792	147	51.9	2.3	5.45 (0.79)	18.58 (4.75)	0.566	0.016	0.003 (0, 0.057)
3 Sangmelima	53	5043	171	49.7	3.8	5.42 (0.81)	23.20 (5.52)	0.616	0.012	0.033 (0, 0.111)
<i>Gabon</i>										
4 CBG	26	1980	77	24	4.1	5.28 (1.00)	22.40 (8.57)	0.567	0.053	0.004 (0, 0.195)
5 CEB	65	3298	125	46.8	2	4.25 (0.75)	14.34 (7.67)	0.481	0.055	0.004 (0, 0.070)
6 Libreville	17	1378	68	38.1	4.8	3.97 (0.68)	19.13 (7.87)	0.503	0.191	0.007 (0, 0.195)
7 Mitzic	28	2325	188	85.5	6	4.31 (0.80)	10.53 (4.78)	0.470	0.087	0.010 (0, 0.232)
8 Oyem	15	1164	75	26.1	8.8	4.36 (0.72)	13.89 (4.16)	0.526	0.173	0.004 (0, 0.110)

Abbreviations: N , number of sampled adult trees; d_{ij} max, maximal pairwise spatial distance (in km) between two individuals of the same population; ad_{ij} , average pairwise spatial distance within a population; nd_{ij} , average distance to nearest neighbour within the same population; rA_{OL} , rarefied mean observed number of alleles per locus, and standard error; V , variance of allele size, and standard error; H_E , expected heterozygosity; F_{IS} , fixation index; and s , selfing rate, and 95% confidence interval.

Table 2 Parameters of genetic variation for the three clusters (see Table 1 for signification of symbols) and mean pairwise differentiation between populations within and among clusters based on allele identity (mF_{ST}) or allele size (mR_{ST})

Cluster	N	rA_{OL}	H_O	H_E	F_{IS}	mF_{ST} / mR_{ST}
North	104	5.378	0.574	0.595	0.035	0.0267/0.0093
West	56	4.623	0.488	0.545	0.103	0.0532/0.0466
South	109	4.306	0.447	0.485	0.079	0.0184/0.0193
Two-sided P -value:		0.18	0.16	0.15	0.39	
Among clusters						0.0862/0.0921

P -values (two-tailed) for differences among clusters were obtained after 5000 permutations.

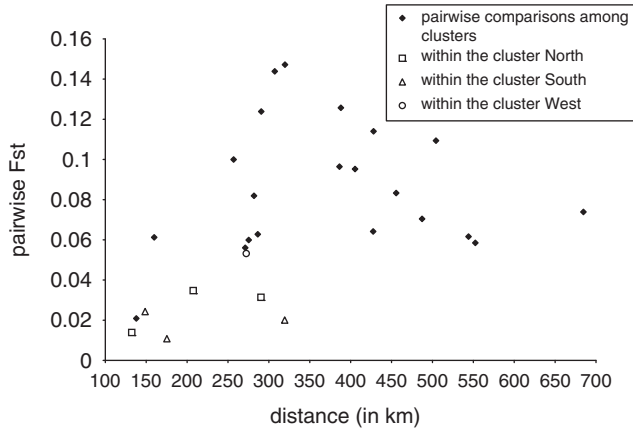


Figure 2 Among- and within-clusters pairwise F_{ST} between populations according to geographical distances (in km). Black diamonds: among-cluster comparisons; empty squares: comparisons within the 'North' cluster; empty triangles: comparisons within the 'South' cluster; empty circles: comparisons within the 'West' cluster.

overall significant differences among the three clusters in allelic richness, observed heterozygosity, gene diversity or F_{IS} (Table 2), and no pairwise significant differences (Wilcoxon tests; all $P > 0.05$), except for allelic richness, which showed a tendency to be higher in the 'North' cluster than in the 'South' cluster ($Z = -2.4$; $P = 0.014$).

Taking into account all eight populations, isolation by distance is marked ($P = 0.0133$). Isolation by distance may cause fictitious clusters when using clustering method such as the STRUCTURE algorithm. One way to check the reliability of the clusters proposed is to compare *a posteriori* pairwise F_{ST} between populations distinguishing within- and among-cluster pairs. For the same range of geographical distances (that is, < 350 km), average pairwise F_{ST} (mF_{ST}) increased threefold between within-cluster ($mF_{ST} = 0.027$) and among-cluster comparisons ($mF_{ST} = 0.086$; Figure 2). In addition, there was no F_{ST} -distance correlation for the among-cluster pairwise comparisons (slope $b \approx 0$, $R^2 = 0.005$, $P = 0.76$), suggesting that the genetic structure of each cluster was established independently. F_{ST} between the eight sampled populations revealed that the highest differentiations occurred between populations of the cluster "South" and the others. According to allele size randomization tests, R_{ST} did not exceed significantly F_{ST} (Table 2), suggesting that the number of generations since population isolation is lower than the reciprocal of the mutation rates of the locus considered (Hardy *et al.*, 2003).

Spatial genetic structure and estimation of gene dispersal and outcrossing rate

In the four populations (1, 2, 3 and 5) with at least 40 individuals, the obtained slope b_{Ld60} of the regression of the kinship coefficient values (F_{ij}) on the spatial distance was significant (all $P < 0.001$). Values of the statistic S_p were relatively similar, ranging from 0.008 to 0.013 (Table 3). Estimation of S_p at a smaller spatial scale (< 10 km) gave values ranging from 0.010 to 0.027 (Table 3).

We succeeded in estimating σ_{gr} , the square root of half the mean squared parent-offspring distance, in three

Table 3 Estimates of SGS and parameters of gene dispersal for four populations

Sampled population	N	D (Ind/km ²)	SGS parameters			Gene dispersal parameters				
			F_{IS}	F_{IT}	b_F	S_p (s.e.) for $d_{ij} \leq 60$ km	S_p (s.e.) for $d_{ij} \leq 10$ km	N_b	σ_g (m) ($D = D/2$)	σ_g (m) ($D = D/4$)
1 AbongMbang	40	10.7	0.094**	0.054***	-0.007*** (0.004)	0.008 (0.004)	0.017 (0.005)	—	—	—
2 Ebolowa	50	68.5	0.016 ^{NS}	0.052***	-0.011*** (0.003)	0.011 (0.003)	0.024 (0.007)	107 (28-∞)	993 [†] (391-∞)	1155 [†] (578-∞)
3 Sangmelima	52	55	0.012 ^{NS}	0.047***	-0.010*** (0.002)	0.010 (0.002)	0.010 (0.009)	56 (26-∞)	601 (441-1072)	746 (623-1783)
5 CEB	62	60	0.055 ^{NS}	0.105***	-0.012*** (0.004)	0.013 (0.004)	0.027 (0.006)	47 (18-∞)	426 [†] (310-∞)	—

Number of genotyped adult trees (n), estimated density of trees over 40 cm d.b.h. (D), SGS parameters: inbreeding coefficient (F_{IS}), average kinship coefficient between individuals separated by less than 1 km (F_{IT}), slope, and its s.e., of the regression of kinship coefficient values on the logarithm of spatial distance between individuals (b_F), and the intensity of SGS (S_p) and its s.e. Calculated for pairwise distances between individuals up to 60 km, and up to 10 km, in each population. Gene dispersal parameters within the spatial range $\sigma_{gr} < 20\sigma_{gr}$: neighbourhood size (N_b) and 95% confidence interval, gene dispersal distance (σ_g) and 95% confidence interval for three effective densities (D_e) estimated from the density of adults (D). Absence of values indicates that the estimation procedure did not succeed (see text). N_b was little affected by the assumed D_e , so that the value given is the mean estimate under the three assumed D_e (values in italics when the estimation procedure did not succeed for all assumed D_e), and the confidence interval considers the lowest lower and highest upper interval values. P -values: not significant (NS) for $P > 0.05$; ** for $0.01 \geq P > 0.001$; *** for $P \leq 0.001$.

[†] Averaged σ_g over the different values between which the iteration procedure cycled.

populations (2, 3 and 5); the algorithm did not converge for population 1, surely because of the low number of individuals ($n=40$). Over populations 2, 3 and 5, estimates of the neighbourhood size Nb ranged from 50 to 110 individuals (95% CI: 18– ∞) and estimates of σ_g ranged from ~ 400 to 1200 m (Table 3). CIs of these estimates usually had no upper bound limit. However, in population 3 in which the SGS was marked, we obtained more precise 95% CI for two effective densities (D_e) estimated from the density of adults (D): for $D_e = D/4$, $\sigma_g = 601$ m (95% CI: 441–1072) and for $D_e = D/10$, $\sigma_g = 746$ m (95% CI: 623–1783) (Table 3).

In all populations, both estimations of the multilocus selfing rate, $s(\hat{g}_2)$ and s , were not significantly different from zero, except for the estimate of $s(\hat{g}_2)$ in population 2 ($s(\hat{g}_2) = 0.12$; $P = 0.006$). Globally, there was no among-population significant difference in the selfing rates encountered, and upper bound estimates were always less than 0.23 (Table 1).

Discussion

Macrogeographic genetic structuring

Despite the continuous geographic distribution of *D. benthamianus* in tropical forests of Central Africa, we detected three distinct genetic clusters of this tree species in Gabon and Cameroon. The distribution of individuals in these three clusters demonstrated a clear geographical pattern (Figure 1b): a genetic discontinuity separated 'North' (104 individuals) and 'South' (109 individuals) clusters from inland regions, and a less clearly defined 'West' cluster (56 individuals) contained mostly individuals from coastal regions, but also encompassed a few individuals with further inland location. According to microsatellite allele size randomization tests (Hardy *et al.*, 2003), stepwise mutations have not contributed to the differentiation among these genetic clusters, which would be due to genetic drift only ($R_{ST} = F_{ST}$, no phylogeographic signal). Tests based on the distribution of allele frequencies suggest recent population expansion, at least when a model of stepwise mutations is assumed.

One potential explanation of the existence of several genetic clusters in the range of one species is the presence of barriers to gene flow. Along with historical factors (mainly, colonization history), environmental heterogeneity (for example, rainfall levels (Letouzey, 1968)) and physical barriers (for example, rivers (Anthony *et al.*, 2007)) are key factors that have already been demonstrated in Africa to influence species distribution, and thus their genetic structure, but mostly for animals species. Another potential ecological barrier is due to the annual movement of the intertropical convergence zone, a band of clouds that encircles the globe near the thermal equator and determines the seasonality of precipitations. South and north of the thermal equator, situated $\sim 2^\circ\text{N}$ in Africa, seasons are inverted, that is, they are shifted by 6 months, inducing a time-lag in plant flowering phenology that may slow down pollen-mediated gene flow, but should not affect seed-mediated gene flow. As for *D. benthamianus*, the limit between the 'North' and 'South' clusters matches the position of the thermal equator, while no physical barriers could explain that differentiation. Whether phenological isolation around

the thermal equator really occurs in this species should be investigated: we know that fruiting of *D. benthamianus* usually occurs at the end of the short dry season (Doucet, 2003), and it should be studied if, for instance, one gene pool is programmed to flower during one dry season, while another gene pool flowers 6 months later. However, this explanation seems unlikely given the high inter-annual variability of seasonality and the ability of plant species to respond phenologically to this variability. Moreover, temporal isolation would not prevent gene flow through seed dispersal and cannot explain the presence of the 'West' cluster. Presence of 'mountains' (zones of higher elevation) may act as barriers limiting contemporary gene flow in the transition zone between the 'West' genetic unit and the others, but the lower elevation of the plain of the Ogooué River (the main river of Gabon) significantly disrupts this potential physical barrier.

Hence, these clusters should rather be linked to the biogeographical history of the region. Forest fragmentation, for example, during cold and dry periods of the Last Maximum Glaciation at the end of the Pleistocene (Maley, 1996; Plana *et al.*, 2004), likely caused the decline of *D. benthamianus*, which would have survived only in isolated populations, each subject to drift. With the return of more favourable conditions, the species then recolonized the whole region from these relict populations. Being a light-demanding tree species, *D. benthamianus* may have undergone particularly fast population and range expansion during the early phase of forest recolonization. According to this hypothesis, genetic differentiation of several pools with distinct allelic frequencies could be the signature of expansion of differentiated relict populations following the Last Maximum Glaciation.

Dating past events from population genetics data is notoriously difficult because theoretical inferences are usually relative to a mutation rate or calibration points that are not available. We cannot rely on genetic data to obtain a narrow range of age estimates for past events given all the unknown parameters. We can only check whether the proposed demographic scenario occurred on a timescale compatible with genetic evidences, but this gives at best the order of magnitude of the time elapse since population isolation. In our case, as R_{ST} is close to F_{ST} , the number of generations of isolation should be less than the reciprocal of the mutation rate (Hardy *et al.*, 2003). Assuming that the latter is close to 10^{-3} , a reasonable value for microsatellite markers, we suggest that populations were isolated less than 1000 generations ago, otherwise R_{ST} would be expected to exceed F_{ST} . Assuming a mean generation time > 100 years, consistent with the facts that in *D. benthamianus* the minimal d.b.h. of fruiting trees is 20 cm, the mean d.b.h. of fruiting trees is 40–45 cm and the mean annual diameter increment is 3–4 mm (Doucet, 2003), our data seem compatible with fragmentation of populations during the last glacial period ($\sim 110\,000$ to 9700 years ago, just before the Holocene) of the current Quaternary period. With gene dispersal distances of the order of one kilometre per generation, many generations are needed to erase a genetic signature of past fragmentation throughout a region encompassing ~ 1000 km from one end to the other, such as the one studied. This is again compatible with recolonization beginning early in the Holocene. The limited ability of the clustering software

to assign some individuals to one of the three clusters, especially in western Gabon, may be a consequence of partial admixture. Forests in Central Africa have also been affected by major climatic events that occurred more recently, 4000 and 2500 years BP (Maley, 2002), and these might also have caused forest fragmentation. However, these disturbances lasted only for relatively short times, and population sizes in isolated fragments would have to have been very small to account for the observed F_{ST} (<150–250 individuals, assuming a generation time of 100 years and very recent population expansion). The intense drift would have resulted in relatively low levels of gene diversity, which is not observed. Hence, these recent events are unlikely to

account alone for the observed patterns. The fact that signs of recent population expansion under a strict stepwise mutation model were observed again supports a scenario of relatively recent recolonization after fragmentation occurring during the Last Maximum Glaciation, at the end of the Pleistocene. However, this signature of population growth is dependent on the mutation model assumed.

The observed distribution of clusters (Figure 3a) can be compared with that observed for two other tree species living in the same region and genotyped with nuclear microsatellites: *Milicia excelsa* (Moraceae; Figure 3b) exhibits one cluster in southern Cameroon and north-eastern Gabon and a second in the rest of Gabon



Figure 3 Maps of genetic clusters determined for three species of tropical African timber trees. Dashed lines stand for the known range limits of the species. (a) *D. benthamianus* (this study): the hatched zone in the upper light grey area shows the putative fourth cluster. (b) *Milicia excelsa* (Dainou K and Heuertz M, personal communication) with one cluster in southern Cameroon and north-eastern Gabon and one in the rest of Gabon. (c) *Aucoumea klaineana* (Born, 2007) with four clusters in Gabon and western Congo.

(Dainou K and Heuertz M, personal communication), whereas *Aucoumea klaineana* (Burseraceae; Figure 3c) exhibits four differentiated genetic clusters within Gabon (Born, 2007). Hence, within the areas where these three species co-occur, there is no clear geographical matching between their respective genetic clusters. The basic assumption that all species share the same Quaternary refuges and recolonization pathways is invalidated here. Likewise, in Europe, even closely related species living in the same habitats can display highly contrasted phylogeographic patterns (Michaux *et al.*, 2005). This suggests that the species-specific phylogeographic pattern may be a general feature and that the search for the influence of common historical factors to explain the phylogeographic structure of different species may turn out to be very complex but see reference Alvarez *et al.* (2009). Nevertheless, it is noteworthy that a north-south discontinuity of genetic variation occurring in the vicinity of the border between Cameroon and Gabon is observed both in *D. benthamianus* and in *M. excelsa* (Figure 3a and b), as well as in two other tree species using other genetic markers: *Irvingia gabonensis* according to RAPD (Lowe *et al.*, 2000) and *Greenwayodendron suaveolens* according to cpDNA sequences (Dauby G, personal communication). Moreover, the northern extension of *A. klaineana* also matches approximately this limit (Figure 3c). If this trend was confirmed in additional species, it might indicate that forests north and south of the thermal equator were isolated for enough time to cause population differentiation and some species might have survived in only one of them. In *D. benthamianus*, a (slight) relationship exists between belonging to a genetic cluster and some 'environmental' heterogeneity (presence of 'mountains' and localization of the thermal equator), contrary to *Aucoumea* (Born, 2007). The riverine barrier hypothesis, which has been verified with respect to the structure of genetic diversity mainly in primates (Eriksson *et al.*, 2004; Anthony *et al.*, 2007), remains largely untested in plants, notably in African rainforests. Here, it appears that the two main rivers of the sub-region (the Ogooué River in Gabon, and the Sanaga River in Cameroon), that both flow east to west, may not particularly act as physical barriers.

Similarity in SGS between populations

A slight SGS (estimated using the S_p statistic for distances <60 km; Table 3) was detected within four populations, with a maximum S_p statistic of 0.013. The extent of SGS was very similar among populations. Moreover, values of genetic diversity (H_E) and inbreeding index (F_{IS}) were relatively homogeneous among all populations. Hence, genetic dynamics appear similar even in regions distant from one another, suggesting that SGS reflects equilibrium isolation by distance patterns. The sole exception was the statistically significant inbreeding index (F_{IS}) for population 1, suggesting the occurrence of selfing and/or biparental inbreeding. This population is located at the margin of the distributional range of *D. benthamianus* (Vivien and Faure, 1985), but the low density (around one-fifth that observed in other populations) may be a more likely explanation for the higher level of inbreeding.

Considering only distances <10 km gave us doubled values of S_p (except for the population 3), with a

maximum S_p statistic of 0.027 (Table 3), a value in the last quartile of SGS estimates for tropical trees in the literature (Dick *et al.*, 2008). S_p values in *D. benthamianus* are typically of the same order of magnitude as those found in other tropical trees whose seeds are heavy gliders dispersed by the wind (for example, *Dicorynia* (Fabaceae), *Aucoumea* (Burseraceae); see Hardy *et al.*, 2006; Dick *et al.*, 2008). This is coherent with some field observations suggesting seed dispersal to distances up to 50 m (Hecketsweiler, 1992; Cordiez, 2000). The doubling of S_p values when considering distances <10 km instead of up to 60 km indicates that the curve of kinship coefficient values (F_{ij}) presents a concave trend, suggesting a 'fat-tailed' dispersal kernel of genes (Heuertz *et al.*, 2003). However, the precision obtained when considering distances <10 km is limited because only 7–40% of pairs of individuals were taken into account.

Gene dispersal distance in *D. benthamianus*

Information about gene dispersal is a prerequisite to understanding the distribution and evolutionary dynamics of species. Direct field measurements or parentage analyses, however, are not easy to perform for tropical species, and indirect approaches are often required (Heuertz *et al.*, 2003; Smouse and Sork, 2004; Hardy *et al.*, 2006), especially in species presenting long-distance pollen and/or seed dispersal. We were able to estimate the scale of gene dispersal based on the strength of SGS in three populations, and we found estimates of σ_g (~ 0.71 the quadratic mean parent-offspring distance) comprised between 400 and 1200 m—depending on the population and on the assumption made concerning effective density (D_e). Compared with previous estimates of σ_g in tropical trees, generally found to be below 500 m (Hardy *et al.*, 2006; Born *et al.*, 2008; but see Bizoux *et al.*, 2009), this dispersal distance is quite large. Nevertheless, it shows that at the geographic scale investigated (~ 1000 km diameter), past population fragmentation should keep a genetic signature in the form of differentiated gene pools for many generations after a secondary contact between fragments. Our indirect estimates of gene flow are thus compatible with a hypothesis that the large scale genetic discontinuities observed in *D. benthamianus* in the Lower-Guinean domain results from the fragmentation of central African lowland rainforests during the Pleistocene.

Our indirect estimates of gene dispersal distances cannot distinguish between seed and pollen dispersal, and according to our knowledge of seed dispersal vectors, we expect that pollen is the main contributor to long-distance gene dispersal. Nevertheless, only seed dispersal is involved in the dynamic of recolonization and should be quantified more accurately. To distinguish the relative roles of seed and pollen dispersal and elucidate the contributions of dispersal agents and distances, further insights should be obtained by genotyping progeny arrays that should provide contemporary estimates of pollen dispersal distances, by using, for example, a TwoGener approach (Smouse *et al.*, 2001; Robledo-Arnuncio *et al.*, 2006), or by conducting parentage analyses. In addition, sequencing of chloroplast and nuclear genes would complete the characterization of the large-scale genetic structure by a phylogeographic approach that might give additional

insights into the history of the species. General conclusions on the impact of past vegetation changes and the role of climate fluctuations on the patterns of genetic variation of forest species will only be reached when a sufficient number of species will be studied in Central Africa.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This study was supported by the Belgian Fund for Scientific Research (FNRS), in which OJH is a research associate, through Grant FRFC no. 2.4576.07, and by a Postdoctoral Research Fellowship attributed to GDG. It also benefited from the project IFORA, financed by the programme 'Biodiversité' of the 'Agence Nationale de la Recherche' (France), from the project PPR 10.000 financed by the Gembloux Agricultural University (Gembloux Agro-Bio Tech GxABT, Belgium), and from the industrial groups CEB (Precious Woods), Pallisco and Millet. Many thanks to Michel Rougeron, Loïc Douard and Richard Fétéké (Pallisco, Cameroon), Xavier Jaffret, Philippe Jeanmart and Jean-Marie Pasquier (CEB, Gabon) and Guillaume Fenart (CBG, Gabon) for welcoming us during our collecting. We are greatly indebted to Myriam Heuertz (ULB), Kasso Daïnou (Gembloux Agro-Bio Tech GxABT), Jérôme Laporte, Théophile Ayol and Armand Boubady (Nature +) and Patrice Koumba Ipani (Ecole Nationale des Eaux et Forêts, Libreville, Gabon) for their help with collecting samples. We thank Jean-Tristan Huillier for his help with some of the PCRs, Charles Doumenge for making unpublished data available to us, and Doyle McKey who revised the language.

References

- Alvarez N, Thiel-Egenter C, Tribisch A, Holderegger R, Manel S, Schönswetter P *et al.* (2009). History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecology Letters* **12**: 632–640.
- Anthony NM, Johnson-Bawe M, Jeffery K, Clifford SL, Abernethy KA, Tutin CE *et al.* (2007). The role of Pleistocene refugia and rivers in shaping gorilla genetic diversity in central Africa. *Proc Natl Acad Sci USA* **104**: 20432–20436.
- Aubréville A (1968). *Flore du Gabon* vol. 15 Caesalpinioïdées Muséum National d'Histoire Naturelle: Paris, 362 pp.
- Bizoux J-P, Daïnou K, Bourland N, Hardy OJ, Heuertz M, Mahy G *et al.* (2009). Spatial genetic structure in *Milicia excelsa* (Moraceae) indicates extensive gene dispersal in a low density wind pollinated tropical tree. *Mol Ecol* **18**: 4398–4408.
- Born C (2007). Diversité génétique et dynamique des forêts d'Afrique centrale—une étude multi-échelle de la structure de la diversité génétique d'un arbre pionnier, *Aucoumea klaineana* PhD. Thesis, Univ Montpellier 2 France, 124 pp.
- Born C, Hardy OJ, Chevallier MH, Ossari S, Attéké C, Wickings EJ *et al.* (2008). Small-scale spatial genetic structure in the Central African rainforest tree species *Aucoumea klaineana*: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. *Mol Ecol* **17**: 2041–2050.
- Chen C, Durand E, Forbes F, François O (2007). Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* **7**: 747–756.
- Comes HP, Kadereit JW (1998). The effect of Quaternary climatic changes on plant distribution and evolution. *Trends Plant Sci* **3**: 432–438.
- Cordiez F (2000). Etude des mécanismes de régénération naturelle de l'agba (*Gossweilerodendron blsamiferum* Harms), l'azobé (*Lophira alata* Banks Ex-Gaertn Fle movingui (*Distemonanthus benthamianus* Baill) et l'ozigo (*Dacryodes büttneri* (Engl) HJ Lam) au Gabon. Thesis Faculté Universitaire des Sciences agronomiques de Gembloux, 87 pp.
- Cornuet JM, Luikart G (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.
- David P, Pujol B, Viard F, Castella V, Goudet J (2007). Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol* **16**: 2474–2487.
- Dick CW, Hardy OJ, Jones FA, Petit RJ (2008). Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biol* **1**: 20–33.
- Doucet J-L (2003). L'alliance délicate de la gestion forestière et de la biodiversité dans les forêts du centre du Gabon PhD. Thesis, Faculté universitaire des Sciences agronomiques de Gembloux, 323 pp.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* **19**: 11–15.
- Eriksson J, Hohmann G, Boesch C, Vigilant L (2004). Rivers influence the population genetic structure of bonobos (*Pan paniscus*). *Mol Ecol* **13**: 3425–3435.
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.
- Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.
- Fenster CB, Vekemans X, Hardy OJ (2003). Quantifying gene flow from spatial genetic structure data in a metapopulation of *Chamaecrista fasciculata* (Leguminosae). *Evolution* **57**: 995–1007.
- Flenley JR (1998). Tropical forests under the climates of the last 30 000 years. *Climatic Change* **39**: 177–197.
- Frankham R (1995). Effective population size/adult population size ratios in wildlife—a review. *Gene Res* **66**: 95–107.
- Gautier-Hion A, Gautier JP, Maisels F (1993). Seed dispersal versus seed predation: an inter-site comparison of two related African monkeys. *Plant Ecology (formerly Vegetatio)* **107/108**: 237–244.
- Goudet J (1995). FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Heredity* **86**: 485–486.
- Hamilton MB, Pincus EL, Di Fiore A, Flesher RC (1999). Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques* **27**: 500–507.
- Hardy OJ, Vekemans X (2002). SPAGED1: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**: 618–620.
- Hardy OJ, Charbonnel N, Fréville H, Heuertz M (2003). Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics* **163**: 1467–1482.
- Hardy OJ, Maggia L, Bandou E, Caron H, Chevallier MH, Doligez A *et al.* (2006). Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Mol Ecol* **15**: 559–571.
- Hecketsweiler P (1992). Phénologie et saisonnalité en forêt gabonaise, l'exemple de quelques espèces ligneuses PhD. Thesis, Université de Montpellier II, 266 pp.
- Heuertz M, Vekemans X, Hausman J-F, Palada M, Hardy O (2003). Estimating seed vs pollen dispersal from spatial genetic structure in the common ash. *Mol Ecol* **12**: 2483–2495.
- Hewitt G (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

- Letouzey R (1968). *Etude phytogéographique du Cameroun*. Encyclopédie Biologique, LXIX: Paris. Edn Paul Lechevalier 513 pp.
- Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Botany* **82**: 1420–1425.
- Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for Neotropical trees. *Heredity* **95**: 255–273.
- Lowe AJ, Gillies ACM, Wilson J, Dawson IK (2000). Conservation genetics of bush mango from central/west Africa: implications from random amplified polymorphic DNA analysis. *Mol Ecol* **9**: 831–841.
- Maley J (1996). The African rain forest—main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. In: Alexander IJ, Swaine MD, Watling R (eds). *Proceedings of the Royal Society of Edinburgh—Essays on the Ecology of the Guinea-Congo Rain Forest*. The Royal Society of Edinburgh: Edinburgh, Scotland, pp 31–73.
- Maley J (2002). A catastrophic destruction of African forest about 2500 years ago still exerts a major influence on present vegetation formations. *IDS Bulletin* **33**: 14–30.
- Michaux JR, Libois RM, Filippucci MG (2005). So close and so different. Comparative phylogeography of two closely related rodent species (*Apodemus flavicollis* and *A. sylvaticus*) in the Western Palearctic region. *Heredity* **94**: 52–63.
- Namkoong G, Boyle T, El-Kassaby YA, Palmberg-Lerche C, Eriksson G, Gregorius HR et al. (2002). Criteria and indicators for sustainable forest management: assessment and monitoring of genetic variation. *Forest Genetic Resources Working Papers no FGR/37E* Forestry Department, Food and Agriculture Organization of the United Nations. Available online at <http://www.fao.org/docrep/005/AC649E/ac649e00.htm>.
- Nei M (1987). *Molecular Evolutionary Genetics* 2nd edn Columbia University Press: New York.
- Ngomanda A, Neumann K, Schweizer A, Maley J (2009). Seasonality change and the third millennium BP rainforest crisis in southern Cameroon (Central Africa). *Quaternary Res* **71**: 307–318.
- Nguelefack EMP, Ngu KB, Atchade A, Dimo T, Tsabang N, Mbafor JT (2005). Phytochemical composition and *in vitro* effects of the ethyl acetate bark extract of *Distemonanthus benthamianus* Baillon (Caesalpinaceae) on *Staphylococcus aureus* and *Streptococcus agalactiae*. *Cameroonian J Exp Biol* **1**: 50–53.
- Nunney L (1993). The influence of mating system and overlapping generations on effective population size. *Evolution* **47**: 1329–1341.
- Piry S, Luikart G, Cornuet JM (1999). Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Heredity* **86**: 502–503.
- Plana V, Gascoigne A, Forrest LL, Harris D, Pennington RT (2004). Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. *Mol Phylogenet Evol* **31**: 449–461.
- Pons O, Petit RJ (1996). Measuring and testing genetic differentiation with ordered vs. unordered alleles. *Genetics* **144**: 1237–1245.
- Poulsen JR, Clark CJ, Connor EF, Smith TB (2002). Differential resource use by primates and hornbills: implications for seed dispersal. *Ecology* **83**: 228–240.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Raymond M, Rousset F (1995). GENEPOP (Version 1.2): population genetics software for exact tests and oecumenism. *J Heredity* **86**: 248–249.
- Robledo-Arnuncio JJ, Austerlitz F, Smouse PE (2006). A new method of estimating the pollen dispersal curve independently of effective density. *Genetics* **173**: 1033–1045.
- Rychlik W, Rhoads RE (1989). A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA. *Nucleic Acids Res* **17**: 8543–8551.
- Smouse PE, Dyer RJ, Westfall RD, Sork VL (2001). Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. *Evolution* **55**: 260–271.
- Smouse PE, Sork VL (2004). Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management* **197**: 21–38.
- Sonké B, Neuba D, Kenfack D, De Block P (2007). An extraordinary new rheophyte in the genus *Leptactina* (Rubiaceae Pavetteae) from Rio Muni (Equatorial Guinea). *Botanical Journal of the Linnean Society* **153**: 109–113.
- Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* **13**: 921–935.
- Vivien J, Faure JJ (1985). *Arbres des forêts denses d'Afrique Centrale*. Ministère des Relations Extérieures, Coopération et Développement: Paris.
- Walker MJC (1995). Climatic changes in Europe during the Last Glacial/Interglacial transition. *Quaternary International* **28**: 63–76.
- White F (1979). The Guineo-Congolian Region and its relationships to other phytochoria. *Bulletin du Jardin Botanique National de Belgique* **49**: 11–55.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)