Colonization processes and the maintenance of genetic diversity: insights from a pioneer rainforest tree, *Aucoumea klaineana*

Céline Born^{1,2,*}, Finn Kjellberg², Marie-Hélène Chevallier^{2,3}, Hélène Vignes^{2,3}, Jean-Toussaint Dikangadissi⁴, Jodel Sanguié⁴, E. Jean Wickings¹ and Martine Hossaert-McKey²

¹CIRMF, UGENET, BP 769, Franceville, Gabon

²UMR 5175, CEFÉ, 1919 Route de Mende, 34293 Montpellier Cedex 5, France

³CIRAD—UR 37, TA C-37/D, 34398 Montpellier Cedex 5, France

⁴CIRMF, SEGC, BP 769, Franceville, Gabon

Despite recurrent episodes of range expansion and contraction, forest trees often harbour high genetic diversity. Studies of temperate forest trees suggest that prolonged juvenile phase and high pollen flow are the main factors limiting founder effects. Here, we studied the local colonization process of a pioneer rainforest tree in central Africa, *Aucoumea klaineana*. We identified 87% of parents among trees up to 20–25 years old and could thus compare direct parentage structure data with classical population genetics estimators.

In this species, genetic diversity was maintained during colonization. The absence of founder effects was explained by (i) local random mating and (ii) local recruitment, as we showed that 75% of the trees in the close neighbourhood participated in the recruitment of new saplings. Long-distance pollen flow contributed little to genetic diversity: pollen and seed dispersal was mainly within stand (128 and 118 m, respectively). Spatial genetic structure was explained by aggregated seed dispersal rather than by mother–offspring proximity as assumed in classical isolation-by-distance models.

Hence, *A. klaineana* presents a genetic diversity pattern typical of forest trees but does not follow the classical rules by which this diversity is generally achieved. We suggest that while high local genetic variability is of general importance to forest tree survival, the proximate mechanisms by which it is achieved may follow very different scenarios.

Keywords: Lopé National Park; dispersal; colonization dynamics; genetic diversity; founder events

1. INTRODUCTION

Because populations are mortal, and because climates change, the ecological and genetic processes associated with migration and colonization are crucial to species survival. Moreover, colonization processes can leave strong genetic signatures in terms of patterns of genetic diversity and genetic differentiation within and among populations. Two facets of colonization dynamics are often highlighted: species may expand either in a continuous contact process at the limit of their range or by discrete steps resulting in geographically isolated small new colonist populations. Both colonization dynamics may result in loss of genetic diversity and increased differentiation among populations (Wade & McCauley 1988; Austerlitz *et al.* 1997; Le Corre & Kremer 1998). Indeed, when only few individuals contribute to new populations, then strong founder effects are produced (Wright 1932; Mayr 1954). Although patterns of genetic diversity are often attributed *a posteriori* to founder effects, the actual dynamics of range expansion, including numbers of founders and their origins, remains largely unknown (see Anderson & Slatkin 2007; Leblois & Slatkin 2007).

The strong founder effects during colonization predicted by simple neutral models (Austerlitz *et al.* 1997; Le Corre & Kremer 1998) are not upheld by data on nuclear genes in temperate trees. Indeed, in most cases, empirical data evidence limited differentiation among populations and high diversity within populations (Hamrick *et al.* 1992). More realistic models taking into account the prolonged juvenile phase and high levels of pollen flow (i.e. among stand pollen flow) of many temperate tree species predict that founder effects should be much reduced in such species (Austerlitz *et al.* 2000). A long juvenile phase may allow multiple colonization events and the establishment of many seedlings in newly colonized locations before the first colonists start reproducing, thus limiting founder effects (Petit *et al.* 2001).

Furthermore, long-distance seed dispersal events should not be neglected. They account for the speed of recolonization of Europe by tree species during the last post-glacial period (Austerlitz & Garnier-Gere 2003; Bialozyt *et al.* 2006). At small spatial scales, only few studies have directly examined the connection between colonization events and plant population genetics (Ingvarsson & Giles 1999; Erickson *et al.* 2004; Litrico *et al.* 2005; Sezen *et al.* 2005; Jones *et al.* 2006; Raffl *et al.* 2006; Yang *et al.* 2008), and these have supported a causal relationship between the

* Author and address for correspondence: UMR 5175, CEFÉ, 1919 Route de Mende, 34293 Montpellier Cedex 5, France (celineborn@hotmail.com).

ecological circumstances of colonization and genetic diversity and population differentiation. Among them, most studies concerning large woody species (shrubs and trees) evidence (i) a major role of long dispersal mixing founders from multiple sources, (ii) combined with genetic homogeneity among sources due to gene flow, and (iii) the importance of the accumulation of new colonists during the juvenile phase of the first colonists (Erickson *et al.* 2004; Litrico *et al.* 2005; Yang *et al.* 2008). Nevertheless, even in the context of common long-distance pollen flow in plant species, spatial genetic structure (SGS) due to isolation-by-distance (IBD) processes is generally observed (Vekemans & Hardy 2004). This is generally assumed to be due to some short-distance (or, exceptionally, aggregated sibling dispersal; Jones *et al.* 2006) seed dispersal resulting in the presence of half-siblings close to their maternal parents.

Many tree species are long lived and hence founding individuals may persist for long periods of time. Hence, there are situations in which several cohorts corresponding to several discrete bouts of colonization, germination and seedling establishment can be monitored. In such situations, it may be possible to detect the origin of colonists and to compare parent–offspring genetics.

Here, we propose to analyse how the typical genetic diversity and structuring of trees are maintained in a species that does not seem to follow the basic rules underlying the models. Data on genetic variation in *Aucoumea klaineana* Pierre (Burseraceae) have revealed strong within-population genetic diversity and SGS due to IBD processes (Born *et al.* 2008). The species is an insect-pollinated, wind-dispersed pioneer forest tree colonizing savannah. Its advance is hindered by yearly savannah fires. However, if one year a zone escapes fire, the species advances in a discrete colonization step. As a result, discrete age cohorts are produced: there is no progressive accumulation of colonists as postulated in classical tree colonization models. Furthermore, *A. klaineana* does not regenerate within the forest, so that it is difficult to imagine offspring germinating (or surviving) very close to their maternal parent (Fuhr *et al.* 2001).

We chose a field situation constituted by a well-defined patch of old trees, presenting successively an adjacent band of younger trees and then a band of seedlings adjacent to the young trees. Such a set-up with well-defined colonization bouts and limited numbers of trees was used to increase the probability of locating female parents. We analysed the situation genetically in order to address the following questions: (i) is genetic diversity maintained from one generation to the next, from adults to seedlings? (ii) is long-distance pollen dispersal required to maintain genetic diversity? (iii) is progressive accumulation of colonists during a prolonged juvenile phase important for maintaining genetic diversity? and (iv) is SGS (IBD) due to parent–offspring proximity or offspring–offspring proximity (i.e. aggregated seed dispersal)?

2. MATERIAL AND METHODS

(a) Study species and sampling

Aucoumea klaineana Pierre (Burseraceae), forming a monotypic genus, is a long-lived pioneer tropical rainforest tree. It is virtually endemic to Gabon, its range including very small areas in Equatorial Guinea, Cameroon and Republic of Congo. The species is dioecious and is

pollinated by insects, including both social (Apidae: Apinae, Meliponinae) and solitary (Xylocopidae) bees, and flies (Diptera: Calliphoridae, Syrphidae). Its seeds (five seeds in each dehiscent fruit) are wind dispersed. Seeds of a fruit are generally disseminated independently but we have sometimes observed whole fruits dispersed by wind, allowing grouped dispersal of related seeds. In *A. klaineana*, flowering intensity is irregular. However, in a favourable year, flowers are abundant on all trees resulting in substantial seed production at the population level. The species does not regenerate within the forest, but requires larger open spaces (Fuhr *et al.* 2001).

We sampled the species in the forest–savannah mosaic of the Lopé National Park in Gabon. Most of the Park is covered by semi-evergreen lowland tropical rainforest with large areas of forest–savannah mosaic and gallery forest along its northern boundary.

(b) Insights into the colonization history of the population

Overall in Central Africa, savannahs are presently receding and, at the savannah–forest contact zone, only recurrent annual fire limits advance of the forest (White *et al.* 1996; Favier *et al.* 2004a,b). Our field site, the Lopé National Park in Gabon, is no exception to this rule and presents a shrinking savannah zone maintained by fire. If the fires were to cease, the whole area would be covered by forest. *Aucoumea klaineana* is one of the main initial forest trees colonizing the savannah. Seeds germinate in the savannah in January–February (just before one of the rainy seasons) and in the following fire period (during the long dry season) in July–August, seedlings are still too small to survive fires. However, if they escape fire the first year, the following year they have already become fire resistant. They form a zone of treelets, 1.5–3 m high, limiting the spread and intensity of fire, so that new seedlings may establish. Within very few years (1–3 years), the colonization front becomes saturated and as a result cohorts are very homogeneous in age (Favier *et al.* 2004a,b). Rapid local saturation by seedlings has been confirmed by analyses of growth rings performed in Atlantic coastal forests on *A. klaineana* (Nasi 1997).

We chose to analyse a forest edge isolated by more than 300 m of grassland from other forest fragments (latitude: S 0.20033, longitude: E 11.58738). Traces of three successive events of forest progression had been detected on the site and the pioneer trees were still present (figure 1). The first zone only presented saplings 5–15 cm in diameter at breast height (DBH) that must have established 3–5 years before. The second zone presented larger trees of homogeneous size 46 ± 9 cm in DBH (range 16–81 cm). This is a young–adult zone called AD2 in the following. These trees were approximately 20–25 years old (estimate based on DBH and known growth rate; Fuhr *et al.* 2001). The strong homogeneity in DBH in this cohort (table 1) suggests that they correspond to the founders in a very rapid episode of colonization of the forest margin. In this zone, competition for dominance begins to take place but no difference in DBH could yet be detected between dominant and suppressed trees (dominant: trees that reached the top of canopy or suppressed: overtopped trees; table 1). In the adjacent forest, a zone called AD1, we observed a strong heterogeneity in DBH (30–160 cm) associated with a continuous distribution of

Figure 1. Location of study site and samples. (a) Location of the Lopé National Park in Gabon; (b) location of the study area in the savannah–forest mosaic landscape and (c) distribution of reproductive trees in the forest and of collected saplings in the colonization front. DBH, diameter at breast height in cm (black circles, 5.00–14.58; grey circles, 14.59–80.60; white circles, 80.61–160.00); AD1, old adult tree cohort; AD2, young adult tree cohort; SAP, sapling cohort.

Table 1. Distribution of diameters measured at breast height (DBH; mean $\pm$ s.d.) among cohorts (AD1: youngest adult trees, AD2: oldest adult trees) and among social status. Here only trees with unambiguous genotypes are counted.

	dominant		suppressed		all	
	N	DBH (cm)	N	DBH (cm)	N	DBH (cm)
AD1	103	85.52 $\pm$ 28.85	18	56.13 $\pm$ 20.86	121	82.81 $\pm$ 29.93
AD2	18	46.48 $\pm$ 8.65	5	46.48 $\pm$ 12.05	26	46.48 $\pm$ 9.12

diameters (figure 1). This is typical of more mature forest with an important impact of social status on radial growth (table 1): AD1 is nowadays composed of a mosaic of trees of different social status (i.e. dominant versus suppressed). Therefore, age cohorts could not be defined within AD1. AD1 may be composed of one or several cohorts.

We collected 330 samples: 175 saplings (SAP; 3–5 years old), randomly sampled within the sapling zone and all 155 adult trees (AD) located within 300 m of the forest edge (figure 1). Coordinates and social status (dominant or suppressed) of all trees were recorded and DBH were measured for all adult trees. We collected leaves from saplings and bark fragments (cambium) from adult trees and immediately dried them in silica gel.

Depending on sample sizes required to meet statistical constraints, data analyses were performed either separating younger (AD2) and older (AD1) adult tree cohorts or pooling all adult trees (AD = AD1 + AD2) together.

(c) Genotyping

DNA was extracted using the DNeasy Plant minikit (Qiagen, Inc.). Ten specific microsatellite loci, which

previously tested on other populations, presented no heterozygote deficiency, no linkage disequilibrium and no null alleles (Born *et al.* 2006), were selected for genotyping: ak1, ak2, ak5, ak6, ak7, ak8, ak9, ak12, ak14, and ak16 (Born *et al.* 2006). Multiplex PCR was successful and primer pairs were associated as (i) ak1 and ak2, (ii) ak5, ak6, ak7 and ak8, (iii) ak9 and ak12, and (iv) ak14 and ak16 in primer mixes containing 2 μ M of each primer. Multiplex PCR was performed using the Multiplex PCR Kit (Qiagen, Inc.) following the recommended protocol using Q-solution in a final reaction volume of 10 μ l (5 μ l of 2 $\times$ Qiagen Multiplex Master Mix, 1 μ l of primer mix, 1 μ l of Q-solution, 1 μ l of H₂O and 2 μ l of template DNA). PCR conditions were: 15 min of denaturation at 95°C and 30 cycles of 30 s of initial denaturation at 94°C; 90 s of annealing at 59°C; 60 s of extension at 72°C; and 30 min of final elongation at 60°C. Amplifications were conducted in a GeneAmp 2700 Thermocycler (Applied Biosystems). One microlitre of diluted (1 : 20) PCR products of the four multiplexed PCR was pooled in 15 μ l deionized formamide and genotyping was performed on an ABI 3100, using 0.2 μ l

of GS500LIZ size standard (Applied Biosystems). All scoring was done using the GENEMAPPER software (Applied Biosystems).

To control genotyping quality, a strict screening protocol was applied to our data. First, all genotyping was carried out on the same sequencer with two previously typed samples in each plate as a reference. Scoring was automated using the GENEMAPPER software and allele sizing was subsequently hand checked. Finally, we retained only samples that produced non-ambiguous genotypes for at least 7 of the 10 markers. The final number of samples typed was 288, 147 originated from adult trees and 141 from saplings of which only 11 presented some missing data.

(d) Population genetics analysis

To characterize genetic processes associated with colonization, we calculated genetic diversity within each cohort in terms of expected heterozygosity (H_E) and allelic richness (A) using the FSTAT software (Goudet 1995).

Wright's inbreeding index, F_{IS} , was calculated and significance tested using the software SPAGeDi (Hardy & Vekemans 2002). We estimated differentiation between cohorts by calculating pairwise F_{ST} values following Weir & Cockerham (1984).

(e) Characterization of SGS

We assessed SGS using the software SPAGeDi (Hardy & Vekemans 2002). We chose Nason's estimator of the kinship coefficient (Loiselle *et al.* 1995) owing to its robust statistical properties (Vekemans & Hardy 2004). We first tested for within cohort SGS. For the AD and SAP cohorts (sample size too low to separate AD1 and AD2), kinship coefficient values (F_{ij}) were regressed on the logarithm of spatial distance between individuals $\ln(d_{ij})$ providing the regression slopes b_{Ld} . Standard errors were assessed by jackknifing data over each locus. 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 (Mantel test). We then tested for among-cohort SGS. In this case, we calculated F_{ij} considering only individuals i and j from different cohorts (i.e. among parents and offspring). To allow comparisons between these three analyses, all relatedness values were calculated considering allele frequencies calculated taking into account the whole dataset (i.e. allele frequencies obtained considering all individuals without consideration of cohort membership). To compare the extent of SGS among cohorts, we calculated the statistic Sp (Vekemans & Hardy 2004; Hardy *et al.* 2006), defined as $Sp = -b_{Ld}/(1 - F_N)$, where F_N is the mean F_{ij} between neighbouring individuals at most 20 m distant.

(f) Parentage analysis

To estimate pollen and seed dispersal distances within the study area, we used parentage analysis. We first checked the cumulative parentage exclusion capacity (EPC, i.e. the exclusion probability of a non-parental genotype) of the 10 microsatellite markers (Gerber *et al.* 2000). EPC was computed with the FAMoz software (Gerber *et al.* 2003) and was 99.998%, indicating that in the following genetic analysis, individuals recognized as potential parents were most likely to be the real parents.

We performed two parentage analyses. First, we searched within the older adult cohort (AD1) for the parents of the younger adult cohort (AD2). Second, we searched within the total adult cohort (AD1 + AD2) for the parents of the saplings (SAP). Maximum-likelihood estimates were obtained using the FAMoz software and the simulation procedure was used to define threshold values for parent/parent-pair assignment tests. When several potential parent/parent-pairs were identified (only five cases), we retained the parent/parent-pair presenting the highest likelihood ratio if it was much higher than the second highest parent/parent-pair likelihood ratio. We only accepted a parent-pair if individuals in the proposed parent-pair were also, separately, the most likely parents. Results led to three situations: no parents identified, only one parent identified or a parent-pair identified.

(g) A posteriori sex assignment

Aucoumea klaineana is a dioecious species but it is difficult to determine gender using floral morphology (Brunck *et al.* 1990). During the field session, in February 2007, it was impossible to determine the gender of collected individuals as the species was hardly fruiting. We therefore first assigned individuals to two mating groups (sex classes) and determined which of these mating groups was female the following year when fruits were more abundant.

Taking into account that parents in a parent-pair belong to different sex classes, we considered all parent-pairs and managed to assign most parents implicated in parent-pairs to a sex class ('sex1' or 'sex2') using the following line of reasoning: when a parent belongs to sex class 'sex1', then the other tree of the same parent-pair must belong to 'sex2'; if this latter tree is also implicated in another parent-pair, the other parent is assigned to sex1; and so forth to assign all parents to a sex class.

To determine the gender of trees assigned to each sex class, we went back in the field in February 2008, when fruits of *A. klaineana* were very abundant. We checked the presence of fruits on 12 trees: six trees of the class 'sex1' and other six of 'sex2' in order to know which sex class can be assigned as female trees.

(h) Estimating dispersal parameters using parentage data instead of population genetic parameters

Pollen dispersal within the study area corresponds to the distance calculated between trees of parent-pairs. Knowing the sex status of individuals, we also produced direct measurements of seed dispersal within the study area. Seed dispersal corresponds to the distance between mother trees and their offspring.

We used Mantel tests (9999 permutations) to test whether offspring from the same mother tree were aggregated or uniformly distributed within the colonization front, confronting the matrix of pairwise geographical distances separating all saplings and the matrix of pairwise distances separating half-siblings. In the same way, we tested whether trees of parent-pairs were located closer to each other than expected in the case of random crosses within the stand of reproductive trees, confronting the matrix of pairwise geographical distances between adult trees with the matrix of pairwise distances separating parents of the stand.

Table 2. Within-cohort genetic parameters. H_E , expected heterozygosity; A , allelic richness; F_{IS} , inbreeding index followed by s.d. The AD cohort includes all adult trees: oldest adult cohort (AD1) and youngest adult cohort (AD2). SAP designates the sapling cohort. (p -values: n.s. for $p > 0.05$; ** for $0.01 \leq p < 0.001$; *** for $p \leq 0.001$.)

cohorts	N	H_E	A	F_{IS} (s.d.); all loci	F_{IS} (s.d.); without ak4 and ak6
AD	147	0.614	4.89	0.109 *** (0.062)	0.033 n.s. (0.022)
AD1	121	0.617	4.82	0.103 *** (0.057)	0.033 n.s. (0.022)
AD2	26	0.606	5.04	0.107 ** (0.076)	0.037 n.s. (0.040)
SAP	141	0.624	4.58	0.137 *** (0.096)	0.008 n.s. (0.016)

To document whether female trees located close to the colonization front contributed more as parents of the saplings, we established the correlation between distance of female trees to the edge of the forest and number of observed offspring (REG procedure, SAS v. 9). Ideally, this test should include all female trees. However, trees that did not produce offspring could be male or female. We thus performed two regressions, one excluding all unsexed individuals and the other one including all unsexed individuals as though they were female.

As reproductive success may depend on gender, cohort and social status of adult trees, we used generalized linear models to analyse the relationship between total number of offspring produced by each adult tree and the two explanatory variables gender and social status. Because number of offspring is a categorical variable, we used a multinomial distribution and a cumulative logit function (GENMOD procedure, SAS v. 9).

Using a chi-squared test, we also tested whether the proportions of parents identified in the stand of SAP and of AD2 were similar.

3. RESULTS

(a) Cohort comparison

Details on sample sizes and DBH according to cohort are given in table 1. In the oldest adult tree cohort (AD1), suppressed trees had lower DBH than dominant trees. This was not the case in AD2.

Genetic diversity, as estimated both by expected heterozygosity (H_E) and by allelic richness (A), was very homogeneous among cohorts (table 2). H_E ranged from 0.606 in the youngest adult cohort (AD2) to 0.624 for the saplings (SAP). Taking into account all loci, Wright's inbreeding index was significant for each cohort and ranged between 0.103 and 0.137 (table 1). However, an analysis of results obtained for each locus showed that two loci (ak4 and ak6) displayed extremely high F_{IS} values (greater than 0.63) when compared with other loci (less than 0.05). This result suggests the presence of null alleles at these loci in the Lopé population. To compare inbreeding among cohorts, we used F_{IS} values calculated by omitting ak4 and ak6. No estimated value was any longer significantly different from zero and the estimators ranged from 0.008 to 0.037 (table 2).

Pairwise F_{ST} comparisons among cohorts (table 3) evidenced low though significant genetic differentiation between the sapling cohort and the adult cohorts. No significant differentiation was detected between AD1 and AD2.

(b) Characterization of SGS

Both cohorts (AD and SAP) displayed highly significant SGS, probably due to isolation by distance (figure 2).

Table 3. Genetic differentiation among cohorts estimated by pairwise F_{ST} . The sapling cohort (SAP) is weakly but significantly differentiated from adult cohorts (AD, AD1 and AD2). (p -values: n.s. for $p > 0.05$; *** for $p \leq 0.001$.)

	AD2	SAP
AD		0.0163 ***
AD1	0.0018 n.s.	0.0152 ***
AD2		0.0191 ***

Regression slopes were $b_{Ld} AD \pm s.d. = -0.0136 \pm 0.0033$ ($p < 0.001$) for adult trees and $b_{Ld} SAP \pm s.d. = -0.0118 \pm 0.0049$ ($p < 0.001$) for saplings. The extent of SGS, estimated by the Sp statistic, was very similar among cohorts, -0.014 for AD and -0.012 for SAP.

SGS was also observed among cohorts. The slope of the regression between kinship coefficient values (F_{ij}) and the logarithm of spatial distance between individuals $\ln(d_{ij})$ considering only individuals i and j from different cohorts (i.e. among parents and offspring) was b_{Ld} among $\pm s.d. = -0.0063 \pm 0.0020$ ($p < 0.001$). Compared with the within-cohort SGS, the extent of among-cohort SGS was low: $Sp = -0.006$. Furthermore, mean kinship coefficients reached much higher values within cohorts than among cohorts.

(c) Parentage analysis and sex assignment

Detailed results of parentage analyses for each cohort are given in table 4. Remarkably, 87% of parental trees were identified and belonged to the AD population (98% of parents of AD2 and 85% of parents of SAP). The mean number of offspring $\pm s.d.$ per tree was 1.878 ± 1.876 , the number of offspring per adult tree ranged between 0 and 9, and 75% of adult trees had at least one offspring.

Based on all parent-pairs (128), we assigned adult trees to a sex class ('sex1' or 'sex2') following the method described above. Fifty-five adult trees were assigned to sex1, 46 to sex2 and it was not possible to classify 40 trees either because they had no offspring (37) or because they were the single identified parent of only one offspring in our dataset.

Among the six trees assigned to the class 'sex1' and examined for fruits in February 2008, none carried fruits. On the contrary, abundant fruits were found on each of the six trees previously assigned to the class 'sex2'. Consequently, adult trees from class 'sex2' will be qualified, hereafter, as female or mother trees, and trees from class 'sex1' as male or father trees.

(d) Seed and pollen dispersal estimation

Estimates of dispersal are based on results from the parentage and sex assignment analyses considering all

Figure 2. Average kinship coefficient plotted against geographical distance between individuals for (a) the AD cohort, (b) the SAP cohort and (c) among individuals from different cohorts (AD and SAP). Standard errors were assessed by jackknifing data over each locus.

offspring: 87% of the parents were finally known. Seed dispersal ranged from 20 to 258 m and $d_s \pm \text{s.d.} = 118 \pm 62$ m. Pollen dispersal ranged between 16 and 364 m and $d_p \pm \text{s.d.} = 128 \pm 55$ m.

In the SAP cohort, offspring from the same mother tree were spatially aggregated (Mantel test $p = 0.008$). We did not detect any proximity effect on pollination: distances between trees from parent-pairs did not differ from distances between all adult trees ($p = 0.424$).

The number of offspring per mother tree was not correlated with its distance to the edge of the forest, irrespective of whether unsexed trees were included (as if they were female) or excluded ($R^2 = 0.0020$, $p = 0.6737$ and $R^2 = 0.0016$, $p = 0.7926$, respectively).

The number of undetermined parents for SAP was significantly higher than the number of undetermined parents for AD2 ($\chi^2_3 = 11.58$, $p = 0.009$). Surprisingly, we did not find any effect of gender and social status on the total number of offspring produced by adult trees ($\chi^2_1 = 0.31$, $p = 0.544$ and $\chi^2_1 = 0.33$, $p = 0.568$, respectively).

4. DISCUSSION

Aucoumea klaineana is one of the most important pioneer tree species in its distribution range. The species plays a major role in colonization of open and disturbed areas. As such, it is a species most sensitive to climate shift and has probably gone through major episodes of range expansion and regression. In tree species, founder effects associated with colonization have been postulated to be mainly counterbalanced by life-cycle characteristics such as the lengthy juvenile phase and high levels of gene flow through pollen dispersal (Austerlitz *et al.* 2000). However, direct data confirming these hypotheses is largely limited to

Table 4. Results of parentage analyses for each cohort.

	no. of parent	only one parent	parent-pair
individuals of AD2 with parents in AD1	0	1	25
individuals of SAP with parents in AD	5	33	103
total	5	34	128

temperate tree species, mostly wind-pollinated and with wind- and/or animal-dispersed seeds/fruits. Here, we investigated the consequences of colonization processes for genetic diversity in an insect-pollinated and wind-dispersed tree species of equatorial rainforest at a local scale. As we identified 87% of parents for trees up to 20–25 years old, we provide important access into demographic processes in a forest tree species.

(a) Impact of colonization process on genetic diversity

Genetic founder effect could result from cumulative foundation events. It is characterized by loss of genetic diversity and increased differentiation among populations. Here, genetic diversity ($H_E > 0.6$; table 2) was high in all cohorts. These values are within the range of genetic diversity typically observed in other rainforest species in the neotropics (Aldrich *et al.* 1998; Dutech *et al.* 2002; Latouche-Halle *et al.* 2003), in Malaysia (Fukue *et al.* 2007) and in Africa (Lourmas *et al.* 2007). Moreover, no reduction of genetic diversity was observed in association with the colonization process, as illustrated by the stability of expected heterozygosity and of allelic richness among cohorts (table 2). Furthermore, very limited differentiation (pairwise $F_{ST} < 2\%$) was observed between cohorts.

The absence of founder effect was largely due to the very high proportion of adult trees whose offspring participated in the establishment of the colonization front. Indeed, 75% of the adult trees genotyped by us were parent to at least one of the genotyped saplings, although the SAP cohort was installed within a very few years. Surprisingly, similar proportions of male (55) and female (46) trees, and of suppressed (48) and dominant (53) trees had offspring within the sapling cohort. As a limited number of saplings were collected, these proportions may be even higher. This high proportion of trees contributing to the next generation is possible owing to the among-tree homogeneity in the fruiting pattern of the species. Indeed, in the same study area, a monthly phenology survey of the species has been continued for 11 years on 43 trees (21 males and 22 females). While flowering and fruiting intensity varied from year to year, each individual tree flowered and females fruited almost every year (K. Abernethy & J.-T. Dikangadissi 2004, unpublished data). Moreover, as indicated above, AD is composed of at least two generations (perhaps more) and parents of saplings belong to both AD1 and AD2 cohorts. Hence, in *A. klaineana*, the stability of genetic diversity among cohorts and the absence of a strong founder effect are favoured by rather homogeneous fecundity of all trees and by the contribution of overlapping

generations to new cohorts. Nevertheless, the distribution of number of offspring per adult tree was somewhat aggregated, as variance was higher than the mean (i.e. mean number of offspring $\pm$ s.d. per tree was 1.878 ± 1.876): some trees contributed more than others to the saplings.

The classical image of forest tree functioning is based on temperate species. Some dominant trees contribute most of the seeds and a progressive accumulation of seedlings during a prolonged juvenile phase is the major factor counterbalancing founder effects (Mariette *et al.* 1997; Austerlitz *et al.* 2000). In this study, we managed to identify 87% of parents. This in turn allowed us to recognize 75% of adult trees as effective parents and to show that the contribution of trees to the next generation was very homogeneous comparatively to what is known from other forest trees. In this context, we may wonder whether *A. klaineana* is really a very peculiar species or whether (some) tropical forest trees behave very differently from temperate ones.

(b) Dispersal and colonization processes

Only 13% of parent trees remained unknown after parentage and sex assignment analyses. With such data, we were able to connect the values of classical population genetic parameters with actual data on effective pollen and seed dispersal. Offspring not assigned to parents may be offspring of (i) parent trees that were not well genotyped, (ii) undetected parents as parentage assignment failed, either due to genotyping ambiguities or because similarly plausible parents were present, (iii) adult trees outside the sampled stand, or (iv) parental trees within plot that died before our sampling, especially for AD2. The impact of the first two explanations on dispersal distance estimates is limited, while the presence of parental trees outside the stand would lead to underestimating dispersal distances. However, as only 300 m separate our plot from the next stands, long-distance dispersal events may be rare in the species. In our case, observed dispersal distances may give a good estimate of usual dispersal distances. The average pollen dispersal estimate, 128 m, is very low compared with other rainforest tree species pollinated by insects: in most studies, most pollen donors were located outside the study areas, which were generally much larger than our stand (Kenta *et al.* 2004; Latouche-Halle *et al.* 2004; Fukue *et al.* 2007; Lourmas *et al.* 2007). Such limited pollen dispersal may be one explanation of the species' pronounced genetic structuring observed at larger scales (Born *et al.* 2008).

Our average seed dispersal estimate, 118 m, is probably an underestimate as saplings only represent seedlings that escaped fire close to the forest edge. Nevertheless, based on direct counts of seedlings originating from isolated mother trees in open areas, it was shown that seedlings were abundant at a distance of 80 m from the mother tree and were rare at 200 m from the mother tree (Brunck *et al.* 1990). In our study, mother-sapling distances ranged from 20 to 258 m. Hence, our genetic estimates of seed dispersal through parentage assignment are very close to direct observations of dispersal distances.

In *A. klaineana*, effective seed and pollen dispersal distances were similar. While average dispersal distances were relatively short, they were not limiting. Indeed, within the stand, there was no proximity effect on which

male tree fertilized which female tree. Furthermore, location of female trees within the stand (relative to its edge) did not affect their reproductive success, although saplings only established outside the adult stand. Therefore, the lack of local proximity effects on dispersal may explain the observed lack of heterozygote deficit.

Hence, while limited dispersal of both pollen and seed is often associated with high inbreeding rates, even in dioecious species (Litrico *et al.* 2005), this is not the case in our species, at least when saplings are taken into account. However, biases earlier in development cannot be excluded. Indeed, selection against inbred individuals at pre-, post-zygotic or seedling level may occur but is difficult to document. At a later stage of development, selection against homozygotes as trees become older has been suggested as a mechanism for increasing heterozygosity for a number of rainforest tree species (e.g. Conte *et al.* 2003; Hufford & Hamrick 2003; Latouche-Halle *et al.* 2004). Such a process, which can always be invoked as an *ad hoc* explanation, is not required to explain our results for *A. klaineana*.

There was a significantly larger number of undetected parents in SAP compared with AD2. This was expected. Indeed, within the whole study area, forest is progressing into former savannah: forest patches slowly expand and thus get closer and closer to each other despite continued fire (Favier *et al.* 2004a,b). As a result, we expected the increase in the proportion of seeds originating from other, adjacent, forest edges in the younger cohort.

(c) SGS and colonization process

The IBD process was defined by Wright (1943) as the formation of local pedigree structures as a result of limited gene dispersal and local random genetic drift. Genetic statistical tools have been developed to measure SGS as the relationship between a similarity (or dissimilarity) coefficient with logarithmic distance (Rousset 2000). Significant SGS within a population, using several unlinked neutral markers, is generally interpreted as SGS due to IBD. In our case, SGS was significant for both adults and saplings. Interestingly, this was mainly due to aggregated dispersal of sibling seeds rather than to mother-offspring proximity (Hamrick & Nason 1996). Indeed, while there was a slight SGS between adults and offspring, mean values were very low.

The main patterns detected using pedigrees were (i) random within-stand gene flow, for both pollen and seeds and (ii) spatial proximity of siblings/half-siblings within the sapling cohort, probably due to joint dispersal of seeds. Strictly speaking, this is a case of SGS that is not explained by IBD as defined by Wright (1943). Had we not performed parentage analyses, then we would have concluded in favour of typical IBD, as have most studies. Though significant, the population genetics statistic measuring average relatedness values within neighbourhood remained quite low because siblings and non-siblings co-occurred within a neighbourhood. This shows the limits of such tools comparatively to direct pedigree analysis.

One of our results, co-dispersal of half-siblings has previously been speculated to explain the strong SGS found in seedlings of an expanding population of *Quercus rubra* (Jones *et al.* 2006). However, owing to the usual lack of formal genetic data on parentage, these results are

speculative. Furthermore, oaks are animal dispersed and aggregated sibling dispersal was suggested to result from dispersal and particularly caching by squirrels. In *A. klaineana*, we have formal genetic demonstration of grouped sibling dispersal. However, this species has wind-dispersed seeds and potential secondary grouped dispersal by animals cannot compensate for original mixing during the wind dispersal phase. Hence, we have evidence that wind dispersal processes may lead, as well as animal dispersal processes, to grouped seed dispersal, and this may result in SGS. The analysis of population genetic parameters can probably not substitute for direct analysis of parentage and demo-genetics for establishing the actual mechanisms involved.

In conclusion, *A. klaineana* shows genetic diversity patterns typical of forest trees but does not follow all the rules classically invoked to explain this diversity. We suggest that while high local genetic variability is of general importance to forest tree survival (Hamrick *et al.* 1992), the short-term biological determinants of how this diversity is achieved could be quite variable and could follow very different scenarios from those currently proposed.

This study is the product of collaboration between the CIRMF (Franceville, Gabon) and the CEFÉ (Montpellier, France). The CIRMF is funded by the Gabonese government, by Total-Gabon and by the French Ministry of Foreign Affairs. Genetic analyses were partly funded by a PICS (Programme International de Coopération Scientifique) grant (no. 864, to M. H.-M.) from the CNRS and by the FORINFO programme. The first author was funded by the French Ministry of Foreign Affairs and the Agence Universitaire de la Francophonie. We would like to thank all the management and technical staff of the Lopé National Park for permission to work and for their help in the park and more particularly the Great-Apes Study Station. We are especially grateful to Laure Benoît and Chantal Debain for help in genetic analyses and to Marie Charpentier for help in statistical analyses. We also thank Philippe Jarne and Doyle McKey for their comments on an earlier version of the manuscript. We would like to thank the editors and reviewers for their very kind and helpful comments. The constructive nature of the suggestions offered by the reviewers was particularly gratifying. Thank you all for your thoughtful attention to our manuscript. Finally, we are very obliged to some elephant families and *Colobus* groups for allowing us to work in their forest.

REFERENCES

- Aldrich, P. R., Hamrick, J. L., Chavarriaga, P. & Kochert, G. 1998 Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Mol. Ecol.* **7**, 933–944. (doi:10.1046/j.1365-294x.1998.00396.x)
- Anderson, E. C. & Slatkin, M. 2007 Estimation of the number of individuals founding colonized populations. *Evolution* **61**, 972–983. (doi:10.1111/j.1558-5646.2007.00080.x)
- Austerlitz, F. & Garnier-Gere, P. H. 2003 Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. *Heredity* **90**, 282–290. (doi:10.1038/sj.hdy.6800243)
- Austerlitz, F., Jung-Muller, B., Godelle, B. & Gouyon, P. H. 1997 Evolution of coalescence times, genetic diversity and structure during colonization. *Theor. Popul. Biol.* **51**, 148–164. (doi:10.1006/tpbi.1997.1302)
- Austerlitz, F., Mariette, S., Machon, N., Gouyon, P. H. & Godelle, B. 2000 Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* **154**, 1309–1321.
- Bialozyt, R., Ziegenhagen, B. & Petit, R. J. 2006 Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *J. Evol. Biol.* **19**, 12–20. (doi:10.1111/j.1420-9101.2005.00995.x)
- Born, C., Vignes, H., Muloko, N., Wickings, E. J., Hossaert-McKey, M. & Chevallier, M. H. 2006 Isolation and characterization of polymorphic microsatellite loci from *Aucoumea klaineana* Pierre (Burseraceae), a tropical rainforest tree of Central Africa. *Mol. Ecol. Notes* **6**, 1054–1056. (doi:10.1111/j.1471-8286.2006.01431.x)
- Born, C., Hardy, O. J., Chevallier, M.-H., Ossari, S., Attéké, C., Wickings, E. J. & Hossaert-McKey, M. 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. (doi:10.1111/j.1365-294X.2007.03685.x)
- Brunck, F., Grison, F. & Maitre, H.-F. 1990 *L'Okoumé*. Nogent-Sur-Marne, France: Centre Technique Forestier Tropical (CIRAD).
- Conte, R., Nodari, R. O., Vencovsky, R. & dos Reis, M. S. 2003 Genetic diversity and recruitment of the tropical palm, *Euterpe edulis* Mart., in a natural population from the Brazilian Atlantic Forest. *Heredity* **91**, 401–406. (doi:10.1038/sj.hdy.6800347)
- Dutech, C., Seiter, J., Petronelli, P., Joly, H. I. & Jarne, P. 2002 Evidence of low gene flow in a neotropical clustered tree species in two rainforest stands of French Guiana. *Mol. Ecol.* **11**, 725–738. (doi:10.1046/j.1365-294X.2002.01475.x)
- Erickson, D. L., Hamrick, J. L. & Kochert, G. D. 2004 Ecological determinants of genetic diversity in an expanding population of the shrub *Myrica cerifera*. *Mol. Ecol.* **13**, 1655–1664. (doi:10.1111/j.1365-294X.2004.02139.x)
- Favier, C., de Namur, C. & Dubois, M. A. 2004a Forest progression modes in littoral Congo, Central Atlantic Africa. *J. Biogeogr.* **31**, 1445–1461. (doi:10.1111/j.1365-2699.2004.01094.x)
- Favier, C., Chave, J., Fabing, A., Schwartz, D. & Dubois, M. A. 2004b Modelling forest-savanna mosaic dynamics in man-influenced environments: effects of fire, climate and soil heterogeneity. *Ecol. Model.* **171**, 85–102. (doi:10.1016/j.ecolmodel.2003.07.003)
- Fuhr, M., Nasi, R. & Deleue, M. A. 2001 Vegetation structure, floristic composition and growth characteristics of *Aucoumea klaineana* Pierre stands as influenced by stand age and thinning. *Forest Ecol. Manage.* **140**, 117–132. (doi:10.1016/S0378-1127(00)00320-0)
- Fukue, Y., Kado, T., Lee, S. L., Ng, K. K. S., Muhammad, N. & Tsumura, Y. 2007 Effects of flowering tree density on the mating system and gene flow in *Shorea leprosula* (Dipterocarpaceae) in Peninsular Malaysia. *J. Plant Res.* **120**, 413–420. (doi:10.1007/s10265-007-0078-z)
- Gerber, S., Mariette, S., Streiff, R., Bodenes, C. & Kremer, A. 2000 Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Mol. Ecol.* **9**, 1037–1048. (doi:10.1046/j.1365-294x.2000.00961.x)
- Gerber, S., Chabrier, P. & Kremer, A. 2003 FAMoz: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. *Mol. Ecol. Notes* **3**, 479–481. (doi:10.1046/j.1471-8286.2003.00439.x)
- Goudet, J. 1995 FSTAT (v. 1.2): a computer program to calculate *F*-statistics. *J. Hered.* **86**, 485–486.

- Hamrick, J. L. & Nason, J. D. 1996 Consequences of dispersal in plants. In *Population dynamics in ecological space and time* (eds O. E. Rhodess Jr, R. K. Chesser & M. H. Smith), pp. 203–236. Chicago, IL: University of Chicago Press.
- Hamrick, J. L., Godt, M. J. W. & Sherman-Broyles, S. L. 1992 Factors influencing levels of genetic diversity in woody plant species. *New Forests* **6**, 95–124. (doi:10.1007/BF00120641)
- Hardy, O. J. & Vekemans, X. 2002 SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **2**, 618–620. (doi:10.1046/j.1471-8286.2002.00305.x)
- Hardy, O. J. et al. 2006 Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Mol. Ecol.* **15**, 559–571. (doi:10.1111/j.1365-294X.2005.02785.x)
- Hufford, K. M. & Hamrick, J. L. 2003 Viability selection at three early life stages of the tropical tree, *Platypodium elegans* (Fabaceae, Papilionoideae). *Evolution* **57**, 518–526.
- Ingvarsson, P. K. & Giles, B. E. 1999 Kin-structured colonization and small-scale genetic differentiation in *Silene alba*. *Evolution* **53**, 605–611. (doi:10.2307/2640796)
- Jones, F. A., Hamrick, J. L., Peterson, C. J. & Squiers, E. R. 2006 Inferring colonization history from analyses of spatial genetic structure within populations of *Pinus strobus* and *Quercus rubra*. *Mol. Ecol.* **15**, 851–861. (doi:10.1111/j.1365-294X.2005.02830.x)
- Kenta, T., Isagi, Y., Nakagawa, M., Yamashita, M. & Nakashizuka, T. 2004 Variation in pollen dispersal between years with different pollination conditions in a tropical emergent tree. *Mol. Ecol.* **13**, 3575–3584. (doi:10.1111/j.1365-294X.2004.02345.x)
- Latouche-Halle, C., Ramboer, A., Bandou, E., Caron, H. & Kremer, A. 2003 Nuclear and chloroplast genetic structure indicate fine-scale spatial dynamics in a neotropical tree population. *Heredity* **91**, 181–190. (doi:10.1038/sj.hdy.6800296)
- Latouche-Halle, C., Ramboer, A., Bandou, E., Caron, H. & Kremer, A. 2004 Long-distance pollen flow and tolerance to selfing in a neotropical tree species. *Mol. Ecol.* **13**, 1055–1064. (doi:10.1111/j.1365-294X.2004.02127.x)
- Le Corre, V. & Kremer, A. 1998 Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *J. Evol. Biol.* **11**, 495–512. (doi:10.1007/s000360050102)
- Leblois, R. & Slatkin, M. 2007 Estimating the number of founder lineages from haplotypes of closely linked SNPs. *Mol. Ecol.* **16**, 2237–2245. (doi:10.1111/j.1365-294X.2007.03288.x)
- Litrico, I., Ronfort, J., Verlaque, R. & Thompson, J. D. 2005 Spatial structure of genetic variation and primary succession in the pioneer tree species *Antirhea borbonica* on La Reunion. *Mol. Ecol.* **14**, 1575–1584. (doi:10.1111/j.1365-294X.2005.02502.x)
- Loiselle, B. A., Sork, V. L., Nason, J. & Graham, C. 1995 Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425. (doi:10.2307/2445869)
- Lourmas, M., Kjellberg, F., Dessard, H., Joly, H. I. & Chevallier, M. H. 2007 Reduced density due to logging and its consequences on mating system and pollen flow in the African mahogany *Entandrophragma cylindricum*. *Heredity* **99**, 151–160. (doi:10.1038/sj.hdy.6800976)
- Mariette, S., Lefranc, M., Legrand, P., Taneyhill, D., Frascaria-Lacoste, N. & Machon, N. 1997 Genetic variability in wild cherry populations in France. Effects of colonizing processes. *Theor. Appl. Genet.* **94**, 904–908. (doi:10.1007/s001220050494)
- Mayr, E. 1954 Change of genetic environment and evolution. In *Evolution as a process* (eds H. A. C. Huxley & E. B. Ford), pp. 157–180. London, UK: Allen & Unwin.
- Nasi, R. 1997 Les peuplements d'okoumés au Gabon. Leur dynamique et croissance en zone côtière. *Bois et Forêts des Tropiques* **251**, 5–27.
- Petit, R. J., Bialozyt, R., Brewer, S., Cheddadi, R. & Comps, B. 2001 From spatial patterns of genetic diversity to postglacial migration processes in forest trees. In *Integrating ecology and evolution in a spatial context* (eds J. Silvertown & J. Antonovics), pp. 295–318. London, UK: Blackwell.
- Raffl, C., Schönswetter, P. & Erschbamer, B. 2006 'Sax-sess': genetics of primary succession in a pioneer species on two parallel glacier forelands. *Mol. Ecol.* **15**, 2433–2440. (doi:10.1111/j.1365-294X.2006.02964.x)
- Rousset, F. 2000 Genetic differentiation between individuals. *J. Evol. Biol.* **13**, 58–62. (doi:10.1046/j.1420-9101.2000.00137.x)
- Sezen, U. U., Chazdon, R. L. & Holsinger, K. E. 2005 Genetic consequences of tropical second-growth forest regeneration. *Science* **307**, 891. (doi:10.1126/science.1105034)
- Vekemans, X. & Hardy, O. J. 2004 New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol. Ecol.* **13**, 921–935. (doi:10.1046/j.1365-294X.2004.02076.x)
- Wade, M. J. & McCauley, D. E. 1988 Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution* **42**, 995–1005. (doi:10.2307/2408915)
- Weir, B. S. & Cockerham, C. C. 1984 Estimating *F*-statistics for the analysis of population-structure. *Evolution* **38**, 1358–1370. (doi:10.2307/2408641)
- White, L. J. T., Abernethy, K. A., Oslisly, R. & Maley, J. 1996 L'Okoumé (*Aucoumea klaineana*): expansion et déclin d'un arbre pionnier en Afrique Centrale au cours de l'Holocène. In *Dynamique à long terme des Ecosystèmes Forestiers Intertropicaux* (eds M. Servant & S. Servant-Vildary), pp. 195–198. Bondy, France: CNRS—ORSTOM.
- Wright, S. 1932 The roles of mutation, inbreeding, cross-breeding and selection in evolution. In *Proc. 6th Int. Congress of Genetics*, pp. 356–366.
- Wright, S. 1943 Isolation-by-distance. *Genetics* **28**, 114–138.
- Yang, S., Bishop, J. G. & Webster, M. S. 2008 Colonization genetics of an animal-dispersed plant (*Vaccinium membranaceum*) at Mount St Helens, Washington. *Mol. Ecol.* **17**, 731–740.